БИБЛИОТЕКА НИЖЕГОРОДСКОГО
ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
им. Н. И. ЛОБАЧЕВСКОГО

М. А. ФЕДОТКИН

МОДЕЛИ В ТЕОРИИ ВЕРОЯТНОСТЕЙ

*Допущено Учебно-методическим советом
по прикладной математике и информатике
УМО по классическому университетскому
образованию РФ в качестве учебника
для студентов высших учебных заведений,
обучающихся по специальностям «Прикладная
математика и информатика»
и «Информационные технологии»*

МОСКВА
ФИЗМАТЛИТ®
2012

НИЖЕГОРОДСКИЙ
ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ

УДК 519.2
ББК 22.17
Ф 34

Федоткин М. А. **Модели в теории вероятностей.** — М.: ФИЗМАТЛИТ, 2012. — 608 с. — ISBN 978-5-9221-1384-7.

Изложены фундаментальные и прикладные основы современной теории вероятностного моделирования реальных процессов и явлений. Основное внимание уделено проблеме математического задания и классификации реальных экспериментов, интуитивным понятиям и формализации допустимых, элементарных и наблюдаемых исходов, построению теоретико-множественной и вероятностной моделей, функциональным и числовым характеристикам измерителей исходов статистически устойчивых экспериментов. Характерной особенностью книги является наличие большого числа конкретных задач с подробными решениями и замечаниями с целью развития вероятностной интуиции.

Для широкого круга преподавателей, научных работников, инженеров, аспирантов, магистров и студентов, использующих вероятностно-статистические методы в прикладных и теоретических исследованиях реальных случайных экспериментов с применением компьютерных технологий.

Рецензенты: кафедра математической статистики МГУ им. М.В. Ломоносова (д.ф.-м.н., проф. В.Г. Ушаков); кафедра прикладной математики и компьютерного моделирования Российского государственного университета нефти и газа им. И.М. Губкина (д.ф.-м.н., проф. В.В. Рыков).

Учебное издание

ФЕДОТКИН Михаил Андреевич

МОДЕЛИ В ТЕОРИИ ВЕРОЯТНОСТЕЙ

Редактор *И.Л. Легостаева*
Оригинал-макет: *И.Г. Андреева*
Оформление переплета: *В.Ф. Киселев*

Подписано в печать 04.05.2012. Формат 60×90/16. Бумага офсетная. Печать офсетная. Усл. печ. л. 38. Уч.-изд. л. 41.8. Тираж 500 экз. Заказ №

Издательская фирма «Физико-математическая литература»
МАИК «Наука/Интерпериодика»
117997, Москва, ул. Профсоюзная, 90
E-mail: fizmat@maik.ru, fmlsale@maik.ru;
http://www.fml.ru

Отпечатано с электронных носителей издательства
в ГУП «ИПК Чувашия», 428019
г. Чебоксары, пр-т И.Яковлева, 13

ISBN 978-5-9221-1384-7

9 785922 113847

ISBN 978-5-9221-1384-7

ОГЛАВЛЕНИЕ

Предисловие

Учебник написан на основе многолетнего опыта в том числе применения вероятностно-статистических методов для решения реальных задач, и на основе преподавания общих лекционных курсов по вероятностному моделированию, которые на протяжении более 40 лет автор читал студентам различных факультетов Нижегородского государственного университета им. Н. И. Лобачевского. В зависимости от специальностей студентов (прикладная математика и информатика, прикладная информатика, информационные технологии, прикладная физика, экономика, социология) этот курс преподается в течение одного или двух семестров, сопровождается практикой и лабораторными работами. Учебник рассчитан на широкий круг преподавателей, научных работников, инженеров, аспирантов и студентов, которым приходится пользоваться методами теории вероятностей, математической статистики и теории случайных процессов в прикладных исследованиях. Это определило его структуру и изложение на основе теоретико-множественного подхода А. Н. Колмогорова с минимально возможным применением теории меры и функционального анализа, поэтому настоящий учебник значительно отличается от традиционных, так как содержит большое число прикладных задач с подробными решениями, с указаниями на типичные ошибки и возникающими в связи с этим парадоксы. Изложение материала является доступным и математически строгим. Однако в учебнике нет математических построений ради самих построений. Любая сложная математическая конструкция в книге имеет реальную интерпретацию и обоснование в ее необходимости.

Основными задачами этого учебника являются: 1) знакомство с методами построения и анализа адекватных стохастических моделей реальных экспериментов, которые дают различные исходы при их проведении или наблюдении; 2) изучение фундаментальных основ современной прикладной теории вероятностей; 3) развитие у слушателей курса интуиции вероятностно-статистического мировоззрения; 4) знакомство читателей с различными решениями конкретных задач с целью освоения основных понятий и идей вероятностного моделирования.

Для глубокого овладения этим курсом необходимы: 1) знания математики в объеме университетской или вузовской программы, в частности, начальные сведения по комбинаторному анализу, дискретной математике, функциональному анализу и теории множеств; 2) навыки математического моделирования, программирования и компьютерных технологий.

Во введении выявлен предмет теории вероятностей. Дан краткий исторический очерк теории вероятностей и на этой основе выделены условно пять этапов ее развития.

В первой главе даны основные неопределяемые понятия в теории построения вероятностных моделей. Предложен простейший способ задания реальных экспериментов и проведена их классификация. В частности, выделены так называемые статистически устойчивые эксперименты. Сформулированы и обоснованы на примерах аксиомы выбора элементарных исходов статистически устойчивого эксперимента и приведены способы построения его теоретико-множественной модели. В рамках этой модели над исходами статистически устойчивого эксперимента можно проследить логические связи и проводить теоретико-множественные операции. Установлено различие между допустимыми и наблюдаемыми случайными событиями статистически устойчивого эксперимента.

Вторая глава посвящена различным методам приближенного и точного вычисления вероятности наступления наблюдаемого исхода случайного эксперимента в зависимости от принадлежности его к тому или иному классу. Выяснена фундаментальная связь между эмпирическим и аксиоматическим подходами к определению вероятности случайных событий. Подробно изучены общие свойства вероятностной функции Колмогорова и на этой основе предложено построение вероятностной модели статистически устойчивого эксперимента.

В третьей главе дано описание условных экспериментов и решен важный вопрос обоснования целесообразности формулы, которая определяет условную вероятность. Для условных экспериментов предложены и исследованы унифицированная и локализованная вероятностные модели. Доказаны утверждения, которые устанавливают связь условных и априорных вероятностей. Проблема математического описания общеязыкового смысла причинной и статистической независимости событий и опытов решена с помощью простейших ограничений на вероятностную функцию.

В четвертой главе изучены свойства общей схемы конечной последовательности независимых испытаний. На примере схемы конечной последовательности независимых испытаний отработана общая методика построения вероятностной модели эволюционных экспериментов. Изложены методы получения формул для вычисления вероятностей событий в биномиальной схеме независимых испытаний Бернулли. Доказаны предельные теоремы Муавра–Лапласа и Пуассона с целью приближенного вычисления биномиальных вероятностей в схеме Бернулли. Выяснено фундаментальное значение предельных теорем Муавра–Лапласа в обосновании понятия статистической устойчивости для случайного эксперимента. Рассмотрена полиномиальная схема независимых испытаний.

Пятая глава посвящена составлению вероятностной модели простейшей формы зависимых испытаний — схемы Маркова с конечным числом состояний. Рассмотрено конструктивное задание такой схемы Маркова. На основе уравнений Колмогорова–Чепмена изучены алгебраические и предельные свойства распределений однородной схемы Маркова с конечным числом состояний. Предложена методика вычисления стационарных распределений схемы Маркова. Рассмотрена геометрическая теория схемы Маркова с точки зрения эволюции реальных систем.

В шестой главе изложены проблемы математического описания количественных показаний различных измерителей результатов статистически устойчивого эксперимента. Для выяснения теоретико-множественных и вероятностных свойств измерителей используются понятия борелевской σ-алгебры на действительной прямой и измеримости отображения. Таким образом, адекватными математическими моделями измерителей исходов статистически устойчивого эксперимента являются такие математические объекты, как одномерные случайные величины. Рассмотрены арифметические действия и операции предельного перехода над случайными величинами. Получен критерий функциональной зависимости случайных величин. Подробно исследованы свойства вероятностных законов распределения одномерных случайных величин и на этой основе проведена замкнутая классификация одномерных случайных величин.

В седьмой главе показана целесообразность рассмотрения упорядоченного семейства случайных величин с целью выяснения в дальнейшем функциональных и статистических связей между различными измерителями исходов случайного эксперимента. Подробно изучены свойства многомерных законов распределения. Предложена методика построения выборочного вероятностного пространства для многомерных случайных величин. Большое внимание уделено проблеме статистической зависимости и независимости семейства случайных величин. Проведена простейшая классификация случайных векторов. Исследованы свойства условных законов распределения случайных величин, получены формулы полной вероятности и Байеса для несчетного числа гипотез.

В восьмой главе на содержательном уровне показана важность рассмотрения в курсе вероятностного моделирования различных числовых характеристик одномерных и многомерных случайных величин. При этом числовые характеристики случайных величин являются математическими моделями интегральных свойств измерителей. Рассмотрены общие свойства характеристик положения, разброса, маловероятных и больших значений случайной величины и, наконец, свойства характеристик асимметрии и крутизны наклона законов распределения. Приведены формулы исчисления этих числовых характеристик. В связи с рассмотрением характеристики разброса случайного вектора введено понятие ковариационной матрицы и изучены ее простейшие свойства.

Девятая глава посвящена важной для практики проблеме функциональной и статистической зависимости между измерителями исходов случайных экспериментов. Прежде всего, решается трудная задача вычисления неизвестных законов распределения некоторого случайного вектора по информации о заданных законах распределения другого случайного вектора и о функциональной зависимости между ними. Для законов распределения неслучайных функций от случайных дискретных или непрерывных аргументов выведены эффективные для применения формулы. Проблема связи функциональной и линейной статистической зависимости между случайными величинами решена в рамках элементарной теории корреляции. Для решения задачи измерения степени статистической зависимости между произвольными случайными величинами рассмотрены общие свойства условного математического ожидания и изучены простейшие вопросы регрессии случайных величин.

В десятой главе рассмотрены наиболее распространенные законы распределения случайных величин. Среди законов исследованы: биномиальный, пуассоновский, нормальный, равномерный, показательный, распределение Бартлетта, геометрическое распределение, гипергеометрическое распределение, распределение Паскаля, распределение хи-квадрат, распределение Вейбулла–Гнеденко, распределение Парето.

Материалы, приведенные в одиннадцатой главе, посвященной предельным теоремам теории вероятностей, позволяют решать фундаментальную проблему об аппроксимации случайных величин и их законов распределения. Доказаны различные варианты законов больших чисел и приведена их интерпретация с точки зрения интуитивного понятия статистической устойчивости случайного эксперимента. На основе общих свойств характеристических функций доказаны простейшие варианты центральных предельных теорем.

По этому курсу в Нижегородском государственном университете им. Н. И. Лобачевского на 3-м курсе проводится лабораторный практикум, который включает следующие работы на персональном компьютере: свойства вероятности; моделирование случайных величин и проверка гипотезы о виде распределений; вычисление значений функций и определенного интеграла методом Монте-Карло; закон больших чисел, центральная предельная теорема и эмпирические функции распределения.

Автор понимает, что возможно улучшение учебника. Он будет искренне благодарен всем, кто укажет замечания, недостатки и пожелания: fma5@rambler.ru.

Автор

Введение

В.1. Цель настоящего учебника состоит в доступном изложении методов построения и анализа адекватных вероятностных моделей реальных процессов и явлений простейшего типа, основ современной прикладной теории вероятностей, а также способов решений конкретных задач математического моделирования статистически устойчивых экспериментов с целью развития интуиции вероятностно-статистического мировоззрения на мир.

Задача любой науки, в конечном счете, состоит в выявлении и исследовании закономерностей, которым подчиняются реальные процессы и явления. Найденные закономерности имеют не только теоретическую и познавательную ценность, но и широко применяются в естествознании, технике, планировании, управлении и прогнозировании. Рассматриваемые в теории вероятностей явления очень сложны. Они находятся под воздействием множества разнообразных и случайных факторов, поэтому каждое индивидуальное проявление процесса, как правило, будет отличаться от любого другого наблюдения. Лишь в массовой совокупности объектов наблюдений за процессом проявляются общие закономерности. Приведем простой пример, который иллюстрирует такую ситуацию.

Вы изучаете эксперимент, связанный с бросанием симметричной монеты. При однократном бросании такой монеты никаких существенных закономерностей выявить нельзя. Эксперимент может закончиться одним из двух возможных исходов, заранее неизвестных, например, выпадает герб или решетка. Но если вы бросите 1000 таких монет (или одну такую монету бросите 1000 раз), то здесь уже появляются общие закономерности следующего вида. Оказывается, что приблизительно в половине случаев выпадает герб. И чем больше монет вы бросите, тем ярче проявится указанная закономерность. Указанная закономерность, которую будем называть статистической, обладает некоторой устойчивостью. Эта устойчивость связана с неизменностью условий проведения эксперимента. Существенное изменение этих условий неизбежно приводит к изменению такого рода закономерности. Рассмотрим тот же пример с бросанием монеты. Допустим, что монета изготовлена из неоднородного материала. Та часть монеты, на которой изображен герб, является легкой, например, алюминиевой, а вторая часть монеты, на которой выбита решетка, является более тяжелой, например, стальной. Ясно, что при многократном бросании такой монеты герб будет выпадать чаще, чем решетка.

Со статистической устойчивостью сталкиваемся в повседневной жизни довольно часто и основываем на ней многие решения практического порядка. Например, характеризуя деятельность предприятия и принимая решения на дальнейший период, исходят из целого ряда средних характеристик. Характеристиками могут быть средняя выработка одного работающего, средний процент брака, средний расход сырья и материалов. Обычно оперируют средними показателями и руководствуются их устойчивостью, хотя в индивидуальном проявлении (в определенные дни или у одного рабочего) эти показатели будут колебаться в широких пределах.

Способы математического анализа данных, относящихся к массовым явлениям, с целью определения некоторых обобщающих эти данные характеристик и выявление статистических закономерностей составляют предмет теории вероятностей и математической статистики. Основным методом познания массовых явлений является построение и изучение абстрактных математических моделей, которые называются вероятностными моделями.

В.2. Всякий процесс, изучаемый в природе, ставит перед исследователями вопросы о соотношении случайного и необходимого, о законе, о причинности и т. п. Еще в древности философы знали, что реальные процессы вызываются бесчисленным множеством других явлений, или комплексом условий. Однако при изучении реального процесса приходится выбирать наиболее важные условия, определяющие основные закономерности данного процесса. В этом случае реальный комплекс условий исследователь заменяет модельным комплексом. Поэтому протекание процесса, за которым наблюдаем, во времени приводит к неопределенности и к отклонениям от ожидаемой закономерности. Процессы, которые протекают во времени, называются эволюционными. На содержательном и интуитивном уровне можно выделить следующие особенности или свойства исходов (результатов) реальных процессов и опытов:

— свойство однозначности и определенности допустимых исходов эволюционного процесса, когда весь его будущий ход и все его прошлое однозначно определяются исходом (состоянием) в некоторое фиксированное время. Так, например, небесная механика рассматривает движение космических тел, будущее и прошлое которых однозначно определяются их данными положениями и данными скоростями;
— свойство частичной однозначности и определенности допустимых исходов эволюционного процесса только по отношению к будущему. Другими словами, весь его будущий ход однозначно определяется состоянием в настоящее время, а прошлое — нет. Таким свойством обладает распространение тепла в нагретом стержне. Такие процессы рассматриваются подробно в электродинамике Максвелла, в теории колебаний и т. д.;

— свойство неоднозначности и неопределенности допустимых исходов эволюционного процесса не только по отношению к будущему, но и к прошлому. Такое свойство эксперимента на содержательном уровне называется случайным. В этом случае опыт дает различные исходы при его повторении в идентичных условиях. Следовательно, невозможно с полной уверенностью предсказать результаты такого рода экспериментов. На самом первом этапе развития человечества исследователи имели дело с такими экспериментами, например, при сборе и элементарном анализе различных статистических сведений. Император Древнего Рима Август (27 г. до н. э. — 14 г. н. э.) повелел подсчитывать население и количество ежегодно собираемого хлеба. Монголы после захвата Руси и других стран начали проводить переписи населения для сбора дани и податей. Религиозный трактат в Индии (1292 г.) предполагал относительно изучаемого явления следующие типы утверждений: а) есть; б) нет; в) есть и нет; г) неопределено; д) есть и неопределено; е) нет и неопределено; ж) есть и нет и неопределено. В XIV в. начинается подсчет шансов при страховании. В это же время по приходам появились бюллетени о числе похорон лиц, умерших от чумы. Существенным шагом в развитии представлений о неопределенных явлениях явились азартные игры в кости и карты. Шестая часть «Чистилища» из «Божественной комедии» А. Данте (1307–1321) начинается такими словами: «Когда кончается игра в три кости, то проигравший снова их берет и мечет их один в унылой злости...».

В.3. Азартные игры в кости появились в глубокой древности. Кости в Индии, Ираке были распространены уже в IV тыс. до н. э. Игральная кость представляет собой правильный куб с нанесенными на его гранях точками от одной до шести. В качестве примера рассмотрим широко распространенный во многих странах упрощенный вариант игры «Ям». Иллюстративным материалом для понимания правил этой игры будет служить табл. В.1.

В этой игре может принимать участие любое число n игроков, каждому из которых присваивается определенный номер от 1 до n. Игроки по очереди делают последовательные ходы, подбрасывая случайным образом (непреднамеренно) определенное число игральных костей, и заполняют, согласно правилам игры, информационную таблицу. Рассмотрим подробно правила этой игры и заполнения табл. В.1.

Игрок с номером i после каждого своего хода обязательно заполняет только одну свободную клетку i-го столбца, указывая число очков, которые он получает от набранной комбинации с помощью бросков игральных костей. В начале игры каждый игрок имеет одиннадцать

Таблица В.1

№ игрока	1	2	…	i	…	n
1	0	4	…	4	…	3
2	0	8	…	6	…	8
3	15	12	…	12	…	12
4	20	8	…	16	…	12
5	25	10	…	15	…	15
6	24	12	…	24	…	18
Σ^-	28	20	…	10	…	20
Σ^+	29	0	…	20	…	28
F	25	0	…	25	…	25
S	35	0	…	35	…	0
H	50	0	…	0	…	50
ПК	251	74	…	167	…	191
ОК	351	74	…	217	…	221

свободных комбинаций, которые приведены в первом столбце таблицы и обозначены символами 1, 2, 3, 4, 5, 6, Σ^-, Σ^+, F, S, H. Каждую комбинацию игрок объявляет в конце хода и только один раз.

Любой игрок должен выполнить ровно одиннадцать ходов и не более чем 33 броска. При первом ходе каждому игроку дается право совершить не более трех бросков с пятью игральными костями или отказаться от какого-то числа бросков и перенести эти броски в запас для последующего использования. В последнем случае он разумно экономит некоторое число бросков для будущих ходов. В частном случае, если игрок вообще отказывается от бросков, то он в запас переносит три броска и при этом обязан заполнить по своему желанию свободную клетку столбца нулем. Итак, при очередном ходе игрок имеет в своем распоряжении три новых броска и, возможно, еще некоторое число запасных бросков, с которыми он может поступить так же, как и при первом ходе. Игроку каждый раз разрешается бросать любые игральные кости и возможно меньшее их число, чем пять.

Если игрок объявляет комбинацию типа $\alpha = 1, 2, \ldots, 6$, то он получает число очков αm, где m — число игральных костей от одной до пяти, на каждой из которых после случайных бросков в конце хода имеется α точек. Легко видеть, что $0 \leqslant \alpha m \leqslant 30$, $0 \leqslant m \leqslant 5$. Комбинация Σ^- (нижняя сумма) дает число очков $\beta(\Sigma^-)$, равное сумме всех выпавших точек на пяти игральных костях. Точно так же комбинация Σ^+ (верхняя сумма) дает число очков $\beta(\Sigma^+)$, равное сумме всех выпавших точек на пяти игральных костях. При этом обязательно должно выполняться условие $\beta(\Sigma^-) < \beta(\Sigma^+)$. В противном случае одна из этих комбинаций, которая была объявлена позже, дает число очков, равное нулю. Следовательно, имеем неравенства $0 \leqslant \beta(\Sigma^-) \leqslant 30$, $0 \leqslant \beta(\Sigma^+) \leqslant 30$, $\beta(\Sigma^-) < \beta(\Sigma^+)$. Комбинация или фигура F называется фулом. Фул означает, что на трех любых игральных костях выпадает одинаковое число точек и на остальных двух игральных костях также выпадает одинаковое число точек. Появление объявленного фула дает 25 очков. Комбинация или фигура S называется стритом. Стрит означает, что на любых пяти игральных костях выпадают числа точек в последовательности 1, 2, 3, 4, 5 или в последовательности 2, 3, 4, 5, 6. Появление объявленного стрита приносит игроку 35 очков. Наконец, комбинация или фигура H называется покером. Покер означает, что на любых пяти игральных костях выпадает одинаковое число точек. Появление объявленного покера дает 50 очков. Игра заканчивается после осуществления одиннадцати ходов каждым игроком и заполнения всех клеток его столбца.

Для каждого игрока путем суммирования чисел всех клеток его столбца подсчитывается первоначальное количество очков (ПК), которое он получил в результате подбрасывания костей. Каждому игроку выдается одна из премий в размере 0, 30, 50, 100 очков, если суммарное число набранных им очков только за первые шесть комбинаций 1, 2, 3, 4, 5, 6 принадлежит соответственно одному из промежутков вида $[0, 59]$, $[60, 69]$, $[70, 79]$, $[80, 105]$. Из таблицы видно, что первый, второй, i-й и последний игроки получили соответственно премию в 100, 0, 50 и 30 очков. Окончательное количество очков (ОК) каждого игрока складывается из первоначального количества очков и премии. Каждое очко оценивается некоторой денежной единицей. При определении выигрыша и проигрыша каждым игроком предполагается, что данная игра является игрой с нулевой суммой.

В.4. Рассмотрим еще один пример азартной игры двух лиц в кости. Эта игра широко была распространена в XVII в. С помощью некоторого механизма на гладкую поверхность стола крупье подбрасывает симметричную игральную кость три раза и определяется сумма выпавших очков при этих бросках. У каждого из игроков заранее имеется возможность сделать одинаковые денежные ставки на сумму выпавших очков, например, равную только одиннадцати или только двенадцати. Предположим для определенности, что игрок номер один выбирает сумму очков, равную одиннадцати, а второй игрок выбирает сумму очков, равную двенадцати. При таком выборе выигрывает поставленные ставки игрок номер один, если сумма выпавших очков равна одиннадцати, и выигрывает эти же ставки второй игрок, если сумма выпавших очков равна двенадцати. В противном случае — ничья и ставки сохраняются, если игра повторяется, либо игра прекращается и сделанные соответствующие ставки возвращаются игрокам. Ясно, что у игрока номер один с самого начала имеется право выбора и суммы очков, равной двенадцати. Требуется определить разумное поведение каждого из игроков.

В.5. Первые задачи (подсчет числа возможных исходов при бросании нескольких игральных костей) были рассмотрены в работах Н. Тарталья (1499–1557), Дж. Кардано (1501–1576), Г. Галилея (1564–1642). Эти работы сыграли значительную роль в развитии комбинаторных основ явлений с неопределенностью. Основные понятия теории вероятностей были разработаны такими учеными XVI–XVII вв., как П. Ферма, Б. Паскаль, Х. Гюйгенс. Они на интуитивном уровне применяли теоремы сложения и умножения, владели такими понятиями, как зависимость и независимость событий, ввели математическое ожидание. Они решали известные задачи Де Мере (1607–1648) — философа и литератора, который был довольно значительной фигурой при дворе Людовика XIV. Де Мере полагал: «Если кто желает получить шесть очков при бросании одной кости, то он имеет преимущество, начиная с 4-х бросков. Если бросаются две кости, то получение двух шестерок ожидают с преимуществом, начиная с 24-х бросков». Позже Я. Бернулли (1654–1705) доказал предельную теорему для последовательности независимых испытаний.

Начало XVIII в. было ознаменовано в теории вероятностей выдающимися результатами П. Монмора, А. Муавра, Т. Симпсона, Т. Байеса. К середине XVIII в. теория вероятностей стала чаще применяться к страхованию, устройству лотерей и др. В то же время Л. Эйлер и Д. Бернулли начали использовать метод бесконечно малых для нахождения математических ожиданий. В своих работах Ж. Бюффон впервые среди естествоиспытателей XVIII в. начал применять вероятностные рассуждения в исследованиях (гипотеза о происхождении планет Меркурия, Венеры, Земли, Марса, Юпитера, Сатурна).

В дальнейшем П. Лаплас (1749–1827) продолжил исследования Я. Бернулли, связанные с простейшей схемой независимых испытаний (интегральная и локальная теоремы Лапласа). К. Ф. Гаусс (1777–1855) дал вывод нормального закона распределения случайных ошибок. С. Д. Пуассон (1781–1840) обобщил предельную теорему Я. Бернулли, получил распределение, которое носит теперь его имя.

В 1829–1830 гг. впервые в России С. Ревковский (Вильнюсский университет) стал читать курс теории вероятностей. В 1850 г. в МГУ начал читать аналогичный курс А. Ю. Давидов, который опубликовал несколько работ по применению теории вероятностей к статистике и медицине. Первыми серьезными работами в России по теории вероятностей были исследования Н. И. Лобачевского по определению распределения суммы независимых случайных величин. Доступные исследования В. Я. Буняковского (1804–1889) и М. В. Остроградского имели, прежде всего, значение в распространении идей теории вероятностей в России.

Петербургская математическая школа по теории вероятностей, которая была создана П. Л. Чебышевым (1821–1894), А. А. Марковым (1856–1922), А. М. Ляпуновым (1857–1918), математиками воспринимается как исходный пункт всего дальнейшего развития теории вероятностей. Законы больших чисел, марковские цепи, предельные теоремы — вот проблемы, которые они решали в первую очередь. Применение теории вероятностей к физике связано с открытием броуновского движения и было выполнено Л. Больцманом (1844–1906) и Д. В. Гиббсом (1833–1903). Наконец, С. Н. Бернштейн, А. Я. Хинчин и А. Н. Колмогоров завершили аксиоматику теории вероятностей. Теоретико-множественный подход способствовал тому, что теория вероятностей окончательно укрепилась как полноправная математическая дисциплина.

В.6. Хотелось бы отметить, что вся история теории вероятностей наполнена борьбой с неправильными и необоснованными ее применениями, борьбой за признание ее равноправной математической дисциплиной. Так, Ж. Даламбер, один из крупнейших математиков всех времен, при решении самых элементарных задач теории вероятностей допускал грубые ошибки. Рассмотрим пример такой задачи: «Монета бросается два раза. Какова вероятность, что хотя бы раз появится герб?» Ж. Даламбер решал и рассуждал так: «Герб появится либо при первом бросании, либо при втором, либо совсем не появится. Всех случаев три, из них благоприятствуют ожидаемому событию два, следовательно, искомая вероятность равна 2/3». В действительности искомая вероятность равна 3/4. Ж. Даламбер в этой задаче не различает равновозможные и неравновозможные случаи.

Другим типичным примером такого рода ошибок является решение известного парадокса Бертрана (рис. В.1). Подробное исследование парадокса Бертрана приведено во второй главе при решении задачи о гончарном круге.

Итак, требуется определить вероятность того, что взятая наудачу хорда окружности, радиус которой равен r, будет больше, чем сторона

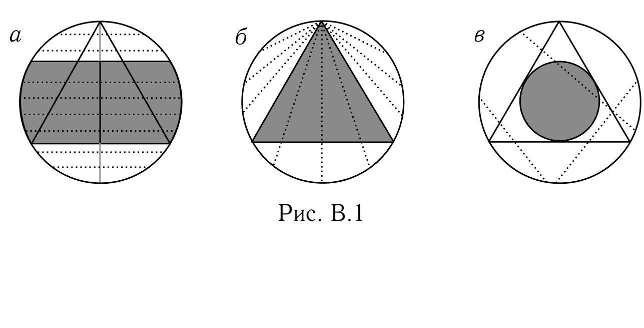


Рис. В.1

вписанного в окружность правильного треугольника. В данной задаче слова «взятая наудачу хорда» без дальнейших разъяснений в постановке задачи не дают никаких указаний о проведении конкретного опыта, а значит, вероятность не может быть определена однозначно в этих условиях. Так, по-разному дополнив условие задачи, можно получить отличные друг от друга результаты при ответе на поставленный вопрос.

В случае, показанном на рис. В.1, *а*, считаем, что хорды, взятые наудачу, всегда параллельны. В этом случае искомая вероятность p_a равна 1/2. В случае, изображенном на рис. В.1, *б*, все хорды имеют одну общую точку, лежащую на окружности; тогда искомая вероятность p_b будет равна 1/3. Если же точки пересечения каждой хорды с окружностью выбираются непреднамеренно, как в случае на рис. В.1, *в*, т. е. середина хорды будет произвольной точкой внутри окружности, то в этом случае искомая вероятность p_c будет вычисляться следующим образом: $p_c = \pi(r/2)^2/\pi r^2 = 1/4$.

В.7. В истории теории вероятностей можно условно выделить пять основных этапов.

1. Накопление материала и решение элементарных задач (до XVI в.). Здесь можно отметить таких выдающихся исследователей во многих областях науки и техники, как Дж. Кардано, Н. Тарталья, Г. Галилей.
2. Возникновение теории вероятностей как науки. Вырабатываются первые специфические понятия, такие, как математическое ожидание, устанавливаются теоремы сложения, умножения. Применяется теория вероятностей в страховом деле, в оценке ошибок наблюдения (П. Ферма, Б. Паскаль, Х. Гюйгенс).
3. Доказательство предельных теорем — простейшие случаи закона больших чисел, разработка аналитических методов в теории вероятностей (Я. Бернулли, А. Муавр, П. Лаплас, К. Ф. Гаусс, С. Пуассон, Л. Эйлер).
4. Дальнейшее доказательство предельных теорем, приложение идей теории вероятностей к физике (Н. И. Лобачевский). Возникновение Петербургской школы по теории вероятностей (П. Л. Чебышев, А. А. Марков, А. М. Ляпунов).
5. Современный этап развития теории вероятностей и ее приложений (С. Н. Бернштейн, Р. Мизес, Е. Борель, А. Н. Колмогоров, А. Я. Хинчин, Б. В. Гнеденко, Ю. В. Прохоров, А. А. Боровков). В настоящее время вероятностные законы открываются в самых различных сферах современного естествознания, и это определяется структурой реальных процессов, которые рассматриваются исследователями.

Глава 1

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ И ТЕОРЕТИКО-МНОЖЕСТВЕННАЯ МОДЕЛЬ СЛУЧАЙНЫХ ЭКСПЕРИМЕНТОВ

§ 1. Реальные эксперименты и их простейшее описание

1.1. Задание опытов на содержательном уровне. Рассмотрим следующие примеры реальных экспериментов и условия их проведения, при этом вместо термина эксперимент будем довольно часто использовать слова-синонимы: опыт, испытание, система, наблюдение и т. п.

Пример 1.1. На кончике указательного пальца левой руки находится достаточно маленький по размеру стальной шарик. В некоторый начальный момент времени $t = 0$ с помощью щелчка правой руки шарику сообщаем заданную скорость и наблюдаем при $t > 0$ его дальнейшее движение над поверхностью Земли под действием лишь силы тяжести. Если известно начальное положение кончика указательного пальца левой руки и если этот шарик с достаточной степенью точности можно принять за материальную точку, то траектория его полета однозначно определяется соответствующим дифференциальным уравнением. Отсюда следует, что для этого опыта с помощью решения такого рода уравнений однозначно вычисляется положение шарика в любой фиксированный момент времени t.

Пример 1.2. В чайнике нагревается один литр воды при естественных условиях до ста градусов по шкале Цельсия и наблюдается состояние воды. Однозначный исход при сохранении основных условий такого испытания состоит в том, что вода закипит.

Пример 1.3. Еще до Галилея многие исследователи проводили эксперимент с бросанием тел в поле тяжести Земли. Галилей с Пизанской башни наблюдал за свободным падением тел, одинаковых по форме, размеру и различных по массе (чугунные, деревянные и т. д.). Он нашел адекватную математическую модель, которая связывает высоту падения H и время падения t формулой $H(t) = gt^2/2$, где g — ускорение свободного падения.

Пример 1.4. Пусть в данном резервуаре поддерживается постоянный объем некоторого количества идеального газа. По определению идеальный газ удовлетворяет следующим двум условиям: 1) объем, приходящийся на молекулы газа, много меньше объема резервуара, и 2) радиус взаимодействия двух молекул значительно меньше среднего расстояния между ними. Тогда формула

$$U = a + bP \tag{1.1}$$

дает адекватную математическую модель, которая позволяет легко вычислить температуру U по заданному значению давления P, где a и b — некоторые постоянные, зависящие от вида газа.

Пример 1.5. Пусть в момент времени $t = 0$ замыкается ключ электрической цепи, изображенной на рис. 1.1, где C — емкость конденсатора, R — сопротивление резистора, $Q(t)$ — разряд электрической цепи в некоторый фиксированный момент времени t.

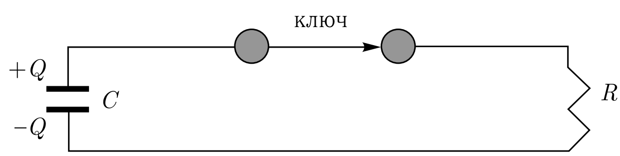


Рис. 1.1

Тогда процесс изменения $Q(t)$ с течением времени определяется функциональным соотношением $Q(t) = Q_0 \exp\{-\lambda t\}$, где $\lambda = 1/(RC)$, $Q(0) = Q_0$.

Пример 1.6. На рис. 1.2 изображена некоторая пружина в нерастянутом и растянутом состояниях.

Длина l пружины, растянутой грузом массы m (см. рис. 1.2, *б*), определяется равенством $l = l_0 + gm/k$. В этом равенстве l_0 — длина пружины в нерастянутом состоянии (см. рис. 1.2, *а*), g — ускорение

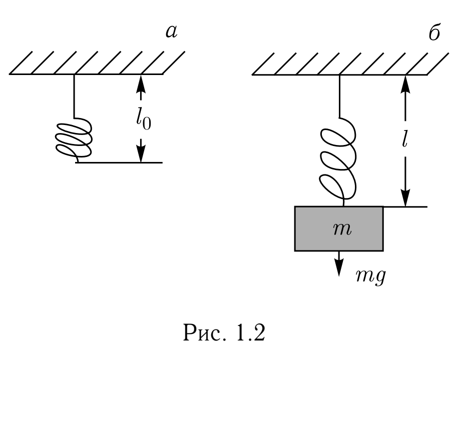


Рис. 1.2

свободного падения и k — коэффициент упругости пружины. При этом согласно закону Гука растянутая пружина создает гармоническую силу $F = (l - l_0)k$, если растяжение не слишком велико.

Пример 1.7. Температура $U(x, y, z, t)$ в точке (x, y, z) твердого тела в момент $t > 0$ удовлетворяет однородному дифференциальному уравнению в частных производных параболического вида:

$$\lambda^2 \left\{ \frac{\partial^2 U(x,y,z,t)}{\partial x^2} + \frac{\partial^2 U(x,y,z,t)}{\partial y^2} + \frac{\partial^2 U(x,y,z,t)}{\partial z^2} \right\} - \frac{\partial^2 U(x,y,z,t)}{\partial t^2} = 0,$$

где $U(x, y, z, 0) = \varphi(x, y, z)$ — известная функция, определяющая температуру такой системы в момент $t = 0$, и λ — коэффициент теплопроводности твердого тела (см. рис. 1.3).

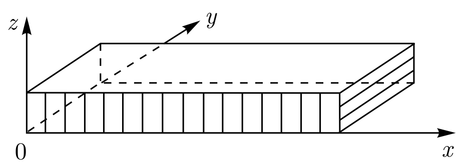


Рис. 1.3

Пример 1.8. Пусть симметричная монета бросается три раза с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Результатами этого опыта могут быть: три раза выпадает герб (орел), три раза выпадает решетка (цифра), один раз выпадает герб и два раза выпадает решетка, первый и третий раз выпадает решетка, второй раз выпадает герб и т. д.

Пример 1.9. Проводится денежно-вещевая лотерея, в которой имеются сто билетов. На билет с номером 50 выпадает выигрыш автомобиль, на билет с номером 60 выпадает денежный выигрыш в сумме 10 тыс. рублей, а на все остальные билеты выигрыша нет. Билет вытаскивается только один раз. Результатом этого эксперимента может быть вещевой выигрыш, денежный выигрыш или отсутствие выигрыша.

Пример 1.10. На поверхность стола с помощью некоторого механизма один раз подбрасываются две симметричные игральные кости, изготовленные из одного и того же материала, например, из слоновой кости. Грани каждой из костей занумерованы цифрами от 1 до 6. В этом опыте определяются числа выпавших очков на каждой из костей.

Пример 1.11. На отрезке $[0, 1]$ выбирается произвольно точка $K = \{x\}$ и определяется принадлежность ее абсциссы x некоторому подмножеству G множества $\{x\colon 0 \leqslant x \leqslant 1\} = [0,\, 1]$. Здесь подмножество G имеет длину.

Пример 1.12. Имеются три разных шара, которые произвольно распределяются по трем разным урнам. Наблюдается размещение шаров в урнах.

Пример 1.13. Пусть непреднамеренно выбирается группа людей из пятидесяти человек, которые проживают в Нижнем Новгороде. При этом интересуются только числом курящих людей среди выбранных.

1.2. Эксперимент и его интуитивное представление. Приведенные примеры говорят о большом разнообразии реальных экспериментов. Так, некоторые эксперименты проводятся непосредственно человеком, другие протекают без его участия, а человек выступает только в качестве наблюдателя и фиксатора происходящего. Это многообразие не позволяет дать точное или строго формализованное определение эксперимента. С другой стороны, в любой науке имеется ряд основных интуитивных понятий. Эти понятия не только не имеют точного определения, но и для каждого человека видоизменяются и усовершенствуются в течение всей его жизни. Так, например, в геометрии основными неопределяемыми (интуитивными) понятиями являются точка, прямая. В механике — сила, масса, путь и др. Основными неопределяемыми понятиями в теории вероятностей являются эксперимент, его результат (исход), статистически устойчивый эксперимент, элементарный результат, равновозможный (равновероятный) элементарный исход и др.

В дальнейшем будем считать, что любой реальный эксперимент E задается или вызывается некоторым множеством $\Sigma = \{u_1, u_2, \ldots\}$ условий u_1, $u_2,\ldots$ его проведения и множеством $\mathfrak{I} = \{A, B, C, A_1, A_2, \ldots\}$ всех его возможных исходов. Другими словами, любой эксперимент происходит только при выполнении соответствующих условий u_1, $u_2,\ldots$. Указанные исходы будем называть *допустимыми* и всегда будем обозначать латинскими прописными буквами без индексов или с индексами, например, A, B, C, A_1, A_2 и т. п. При этом среди допустимых исходов могут существовать такие, которые не наступают ни при каком проведении эксперимента E. Множества Σ и $\mathfrak{I}$ могут содержать конечное, счетное (бесконечное), несчетное число элементов. Как правило, последнее определяется содержанием реального эксперимента и целей его исследования. Приведенная запись $\{A, B, C, A_1, A_2, \ldots\}$ множества $\mathfrak{I}$ является не вполне корректной. Дело в том, что множество может быть несчетным. В этом случае в записи вида $\{A, B, C, A_1, A_2, \ldots\}$, ради простоты, указаны лишь элементы некоторого счетного подмножества множества $\mathfrak{I}$. Аналогичное замечание будем иметь в виду и при записи множества Σ. Множество Σ будем называть *комплексом условий проведения эксперимента E*. Ясно, что при проведении или наблюдении эксперимента часто невозможно определить и перечислить все его условия. Поэтому приходится выбирать и указывать конкретно наиболее важные из этих условий, и тем самым реальный комплекс условий Σ на самом деле заменять его приближением. Этот факт будем отображать следующим образом: $\Sigma = \{u_1, u_2, \ldots, u_s, \ldots\}$, т. е. здесь указано всего s основных условий, а остальные условия (случайные факторы), помеченные многоточием после символа u_s, нам точно неизвестны.

В приведенном примере 1.10 с подбрасыванием игральных костей имеем: u_1 — две симметричные игральные кости, u_2 — заданный механизм подбрасывания (например, с помощью щелчка пальцев правой руки), u_3 — количество бросков равно единице, u_4 — поверхность стола, u_5 — условие фиксации выпавших очков на каждой из костей (например, достаточная освещенность поверхности стола), $s = 5$. Если мы выполним эти пять основных условий, то произойдет известный эксперимент с непреднамеренным однократным подбрасыванием двух игральных костей. В этом эксперименте к случайным факторам можно отнести, например, температуру, давление, колебание воздуха и т. п.

Очень часто некоторый эксперимент E мы будем представлять как совокупность других экспериментов E_t, $t \in T$ (см. примеры 1.1, 1.3, 1.5, 1.7, 1.8). Здесь T — некоторое упорядоченное множество, содержащее конечное, счетное (бесконечное) или несчетное число элементов t. Будем индекс (метку) t интерпретировать как текущее время, а T — как заданный промежуток времени, например, конечный, бесконечный дискретный или бесконечный непрерывный.

Следует заметить, что каждый эксперимент E_t определяется комплексом условий Σ_t и множеством $\mathfrak{I}_t$ его допустимых исходов. Эксперимент E будем называть эволюционным или, другими словами, будем говорить, что он функционирует (протекает) во времени $t \in T$. При этом все допустимые исходы эксперимента E некоторым образом определяются допустимыми исходами каждого из составляющих его экспериментов E_t, $t \in T$.

Итак, чтобы провести эксперимент E нужно провести последовательно эксперименты E_t, $t \in T$. Этот факт формально мы будем обозначать следующим образом: $E = \{E_t, t \in T\}$, $\Sigma = \{\Sigma_t, t \in T\}$, $\mathfrak{I} = \bigsqcup_{t \in T} \mathfrak{I}_t$. Здесь возможен и более простой случай, когда множество T состоит из одного элемента (см. примеры 1.2, 1.4, 1.6, 1.9–13). В этом случае эксперимент будем называть *статическим*. Более того, при изучении некоторых свойств эволюционного эксперимента E, когда определены множество условий его проведения и множество его допустимых исходов, факт его проведения во времени с помощью составляющих экспериментов E_t, $t \in T$, уже может не играть существенной роли. В этом случае эволюционный эксперимент удобно представлять как единый эксперимент и считать его статическим. Можно сказать, что интуитивное понятие о статическом или эволюционном эксперименте является относительным. Другими словами, эксперимент E объявляется эволюционным по отношению к способу его задания с помощью однопараметрического семейства статических экспериментов E_t, $t \in T$. Приведем реальный пример такого эксперимента E.

Пример 1.14. Рассматривается работа простейшего автомата светофора по регулированию движения транспорта и пешеходов с 6-ти часов утра до 23-х часов вечера на пересечении улиц Родионова и Донецкой в Нижнем Новгороде. Предполагается, что в ночное время суток в силу незначительных интенсивностей поступления машин и пешеходов перекресток становится нерегулируемым. Светофор-автомат с помощью зеленого, желтого и красного сигналов в некоторые последовательные промежутки времени разрешает (запрещает) проезд через этот перекресток машинам и одновременно запрещает (разрешает) переход пешеходам. Таким способом обеспечивается безопасность движения машин и пешеходов, которые непредсказуемым образом прибывают на этот перекресток. Заметим, что для простейшего светофора-автомата длительности зеленого, желтого и красного сигналов являются постоянными величинами и периодически повторяются во времени. Эксперимент E_t для этой реальной задачи состоит в том, что инженером транспортником или некоторым автоматическим устройством регистрируются все автомобили и пешеходы, которые остановились на перекрестке в некоторый фиксированный момент времени $t \in T = [\,6; 23\,]$. Для определения длительностей зеленого, желтого и красного сигналов светофора с целью уменьшения задержек транспорта и пешеходов на этом перекрестке естественно рассматривать эксперимент $E = \{E_t, t \in T\}$. Результатом такого эксперимента E будет регистрация уже всех остановившихся машин и пешеходов за полный рабочий промежуток $T = [\,6; 23\,]$.

1.3. Классификация экспериментов. Перейдем теперь к классификации множества реальных экспериментов. Все реальные эксперименты условно можно классифицировать или разделить на группы по признаку однозначности наблюдения его результатов A, B, C, A_1, A_2,... и их предсказаний. В связи с этим исследователи, как правило, выделяют следующие три группы различных экспериментов.

К *первой группе* будем относить эволюционный эксперимент (процесс) E, называемый *детерминированным*, когда по любому $t_0 \in T$ и Σ_{t_0} однозначно определяется исход эксперимента E_t как при $t \leqslant t_0$, так и при $t > t_0$. Для такого эксперимента можно точно предсказать его исход, если имеется полная устойчивость исхода при выполнении комплекса условий Σ. Мы будем каждый раз повторять условия Σ такого эксперимента и, тем самым, можем наблюдать и фиксировать заранее известный его результат. Иначе, весь будущий ход, включая и настоящее, детерминированного эволюционного эксперимента и все его прошлое однозначно определяются комплексом условий Σ_{t_0} эксперимента E_{t_0} в настоящее время $t = t_0$. П. Лаплас, отец этой концепции, говорил: «Всеобъемлющий ум, который охватил бы все начальные условия, предсказал бы все будущее». Еще раз напомним, что классическая механика рассматривает движение материальных точек в пространстве, будущее и прошлое которых однозначно определяется начальными положениями и начальными скоростями всех точек. Подобная ситуация описана в примерах 1.1–1.6, при этом в примерах 1.2, 1.4, 1.6 промежуток T состоит из одной точки.

Ко *второй группе* будем относить так называемый полудетерминированный эволюционный эксперимент (процесс) E, если по любому $t_0 \in T$ и Σ_{t_0} однозначно определяется исход эксперимента E_t только при $t \geqslant t_0$. Иначе говоря, весь будущий ход полудетерминированного эволюционного эксперимента E полностью определяется его исходом или результатом в настоящее время $t = t_0$, а его прошлое не определяется. Распределение тепла в нагретом стержне (см. пример 1.7) является полудетерминированным процессом. Такого рода процессы протекают во времени и изучаются в электродинамике Максвелла, в теории колебаний, в квантовой механике и т. д. Математическое описание процессов из первых двух групп использует аппарат непрерывной и дискретной математики, например, аппарат дифференциальных уравнений.

Наконец, к *третьей группе* будем относить эволюционный эксперимент (процесс) E, если при его повторении практически в одних и тех же условиях он может давать различные, но вполне определенные результаты из множества $\mathfrak{I}$. Такой эксперимент будем называть случайным или стохастическим. Можно утверждать, что множества Σ_t, $t \in T$, и $\mathfrak{I}_t$, $t \in T$, исход эксперимента E определяют неоднозначно. Поэтому результат случайного эксперимента предсказать невозможно (см. примеры 1.8–1.13). Причина такого поведения случайного эксперимента, по-видимому, объясняется следующими обстоятельствами.

Во-первых, множество $\Sigma = \{u_1, u_2, \ldots, u_s, \ldots\}$ условий проведения случайного эксперимента (статического или эволюционного) содержит очень большое число разнообразных второстепенных условий (факторов), помеченных многоточием после символа u_s. Каждый из таких не основных (по отношению к решаемой задаче) факторов в отдельности очень мало влияет на исход эксперимента и может меняться неизвестным образом от опыта к опыту. Однако результат суммарного действия второстепенных факторов приводит к возможности появления заранее непредсказуемых исходов такого эксперимента.

Во-вторых, слишком малые и непреднамеренные изменения основных условий u_1, $u_2, \ldots, u_s$ не могут быть зафиксированы современными средствами (приборами). Поэтому очень трудно утверждать, что повторное проведение эксперимента осуществлялось совершенно точно в одних и тех же условиях. Вместе с тем появление того или иного исхода случайного эксперимента очень чувствительно к малым изменениям хотя бы одного из основных условий u_1, $u_2, \ldots, u_s$.

В-третьих, в природе действует известный принцип неопределенности, согласно которому невозможно абсолютно точное измерение некоторых физических величин. Например, в силу принципа неопределенности в квантовой механике нельзя точно и одновременно знать импульс частицы и ее место нахождения. Или еще пример: не существует способа, который определяет момент времени распада данного ядра урана.

Итак, первые два обстоятельства говорят о том, что при повторном проведении эксперимента E условия u_1, $u_2, \ldots, u_s, \ldots$, к нашему сожалению, можно сохранить только с определенной и доступной нам на сегодняшний момент степенью точности. Обстоятельство неопределенности отражает принципиальную суть вещей, а не только невозможность абсолютно точного измерения некоторых физических величин. Это приводит в случайных экспериментах к непредсказуемости его различных исходов. Например, нельзя заранее предсказать размножение популяции волков в лесах Нижегородской области, т. е. предвидеть в любой момент времени количество волков и в каком месте каждый волк будет находиться. Однако если результаты некоторого эксперимента мы не можем точно предсказать, то отсюда еще не следует, что эксперимент является случайным. Это может быть из-за отсутствия пока у нас некоторых знаний об эксперименте. Например, в настоящее время научная общественность не может утвердительно ответить на вопрос о том, существует или нет органическая жизнь в космосе, исключая, естественно, планету Земля.

1.4. Статистическая устойчивость. Среди всех случайных экспериментов рассмотрим класс экспериментов, которые обладают удивительным свойством статистической устойчивости. Статистически устойчивые эксперименты выделим с помощью следующих ограничений.

• Статистически устойчивый случайный эксперимент E можно проводить или наблюдать любое конечное число раз при одних и тех же Σ и $\mathfrak{I}$.

• Пусть $\mu(A, N)$ есть число наступлений результата $A \in \mathfrak{I}$ за N испытаний эксперимента E. Относительная частота $\mu(A, N)/N$ наблюдения исхода $A \in \mathfrak{I}$ за N испытаний эксперимента E обладает так называемой статистической устойчивостью. На содержательном уровне это означает, что относительная частота колеблется (группируется) около некоторого постоянного числа $\mathsf{P}(A)$ при неограниченном увеличении N и как угодно мало отличается от $\mathsf{P}(A)$. Как правило, число $\mathsf{P}(A)$, которое позднее будем называть вероятностью исхода A, исследователю не известно. Свойство статистической устойчивости должно выполняться для любого $A \in \mathfrak{J}$.

Однако последний факт, а именно близость относительной частоты $\mu(A, N)/N$ к $\mathsf{P}(A)$, не может быть гарантирован всегда, начиная с некоторого достаточно большого числа N. Строгое формальное определение статистической устойчивости частоты $\mu(A, N)/N$ наблюдения исхода A за N испытаний эксперимента E будет дано позднее в курсе теории вероятностей и математической статистики. Заметим, что для некоторых результатов $B \in \mathfrak{J}$ статистическая устойчивость может означать существование предела вида

$$\lim_{N\to\infty} \frac{\mu(B, N)}{N} = \mathsf{P}(B). \tag{1.2}$$

Статистически устойчивый эксперимент в теории вероятностей также является основным интуитивным и первоначальным понятием. Рассмотрим это понятие на следующем классическом примере.

Пусть эксперимент заключается в бросании симметричной монеты определенного достоинства с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Так, французский естествоиспытатель Ж. Бюффон бросал монету $N = 4\,040$ раз. Герб (исход A) появился $2\,048$ раз. Тогда отношение $\mu(A, N)/N = 2\,048/4\,040 \approx 0{,}507$. Английский математик К. Пирсон бросал монету $N = 24\,000$ раз и при этом герб выпал $12\,012$ раз. В этом случае отношение $\mu(A, N)/N = 0{,}5005$. В данном примере исход A означает выпадение герба, и мы имеем малое колебание около числа $\mathsf{P}(A) = 0{,}5$. При единичном проведении этого эксперимента невозможно предсказать его результат, например появление герба, и обнаружить какие-либо закономерности. Однако если рассмотреть последовательность большого числа проведения такого эксперимента, то можно обнаружить некоторую регулярность (статистическую устойчивость) в поведении относительной частоты $\mu(A, N)/N$ выпадения герба.

На рис. 1.4 по оси абсцисс отложено число N проведенных экспериментов, связанных с непреднамеренным подбрасыванием симметричной монеты достоинством в три копейки 1955 года выпуска, а по оси ординат отложено отношение $\mu(A, N)/N$.

График построен для наблюдаемой последовательности исходов следующего вида: B, A, A, A, A, B, B, A, B, B, A, B, A, B, B, A, B, A, B, A, B, A, A, B, A, B, B, A, A, B, B, A, A, B, B, A, A, B, B, A,…

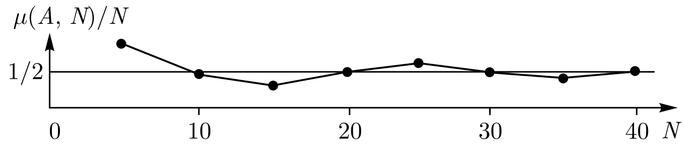


Рис. 1.4

Здесь исход A означает выпадение герба, а исход B — решетки. Из этого рисунка видно, что соединяющая черные точки ломаная линия все ближе и ближе прижимается к прямой $\mu(A, N)/N = 1/2$. Для одноразового подбрасывания этой симметричной монеты обозначим через C исход, который заключается в появлении герба или решетки. Тогда для результата C очевидно имеем выполнение равенства (1.2), ибо в этом случае величина $\mu(C, N) = N$ для всех $N = 1, 2, \ldots$ и, следовательно, $\mathsf{P}(C) = 1$.

Приведем теперь несколько искусственный пример статистически неустойчивого эксперимента. Пусть в некотором непрозрачном пакете достаточно большого объема находится достаточно большое число монет неодинакового достоинства, асимметричных, разной чеканки. Произвольно выбирается монета и подбрасывается с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Затем определяют выпавшую сторону этой монеты, наносят ей механическое повреждение и, наконец, возвращают в пакет. При этом частота выпадения орла и решки может измениться. На этом считаем, что эксперимент закончился. Если рассмотреть последовательность проведения большого числа такого эксперимента, то, возможно, мы и не обнаружим регулярности в поведении относительной частоты появления герба или относительной частоты появления решетки. Последний факт существенно зависит от закона механического повреждения выбранной монеты.

1.5. Фундаментальные проблемы теории вероятностей. Изложенная простейшая классификация реальных экспериментов позволяет точно выразить предмет теории вероятностей. Известно, что основополагающие знания об экспериментах могут быть получены следующими подходами:

- путем проведения значительного числа опытов над реальными экспериментами и обработки результатов наблюдений; это, как правило, требует значительных денежных и временны́х затрат;
- путем небольшого числа опытов над реальными экспериментами, составления адекватных математических моделей, получения фундаментальных знаний об этих моделях и, наконец, путем сопоставления этих знаний с реальными экспериментами.

Поясним более подробно второй подход с помощью рис. 1.5. Строится математическая модель реального процесса (эксперимента, объекта исследования). Затем средствами математики исследуется модель, и результаты исследования интерпретируются применительно к исходному явлению. Этот путь позволяет открыть закономерности реального мира, если математическая модель достаточно адекватна. На рис. 1.5 изображена наиболее полная связь между реальным объектом исследования и его математической моделью.

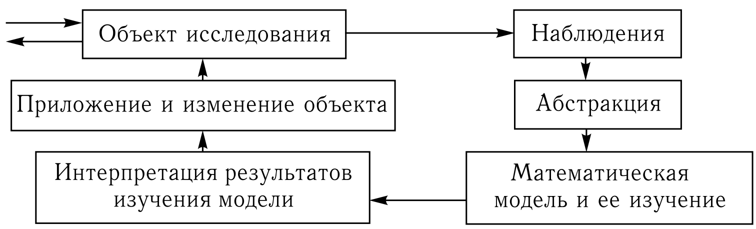


Рис. 1.5

Большая часть математических представлений о мире носит детерминированный и полудетерминированный характер, хотя природа в большей степени является стохастической. Так, в примере 1.4 функциональная зависимость (1.1) температуры U идеального газа от давления P есть суммарный результат соударений всех частиц о стенки резервуара. При этом количество таких частиц и их скорости, очевидно, носят стохастический характер. Это приводит к очень малым отклонениям функциональной зависимости (1.1). Эти отклонения пока не регистрируются современными приборами, которыми измеряют давление и температуру.

Детерминированные модели полезны, однако стохастические или вероятностные модели более адекватны. Интересно отметить, что детерминированные модели всегда будут слишком грубым приближением сложной действительности и, в то же время, они являются частным макроскопическим случаем вероятностных моделей. Однако составление вероятностных моделей и их анализ математическими средствами очень сложен и затруднителен. Широкий и всевозрастающий за последнее десятилетие интерес к теории вероятностей, математической статистике, теории случайных процессов и к применению вероятностно-статистических методов в самых разнообразных областях науки, техники, производства и экономики объясняется следующими двумя основными причинами:

1. Увеличением чувствительности современных измерительных, приемных и управляющих устройств. Вследствие этого случай-

ные отклонения количественных характеристик таких устройств от их средних значений играют все более существенную роль. Отказ от изучения роли стохастических отклонений и случайных механизмов часто приводит к техническим авариям, разрушению мостов, плотин, промышленных сооружений и гражданских зданий, поломкам самолетов и кораблей, транспортным катастрофам, выпуску некачественной и ненадежной продукции, экономическим и природным катаклизмам и т. п.

2. Развитием современных средств микропроцессорной техники, когда появилась реальная возможность хранения, поиска и обработки больших массивов вероятностно-статистической информации о реальных объектах. Используя мощные компьютеры и новые информационные технологии, теория вероятностей дает более широкую концепцию причинных связей между реальными процессами, позволяют найти закономерности природы там, где детерминированный подход оказывается бессильным. Например, в теории ошибок разного рода измерений, в молекулярной и статистической физике, в биологии, в рыночной экономике, в телефонии и процессах обслуживания, в процессах адаптивного управления и принятия решений, в управлении конфликтными транспортными потоками на магистралях крупных городов и т. д.

На основании изложенного материала в предыдущих разделах мы убедительно можем сформулировать основной предмет теории вероятностей. Теория вероятностей, во-первых, предлагает методы построения адекватных моделей реальных статистически устойчивых экспериментов, во-вторых, средствами математики изучает эти модели и тем самым открывает новые фундаментальные закономерности реального мира.

§ 2. Логические и функциональные связи между исходами реального эксперимента и их интерпретация

2.1. Аксиомы выбора элементарных исходов статистически устойчивого эксперимента. Случайные события и их классификация. Итак, в дальнейшем полагаем, что всякий эксперимент E определяется заданием комплекса $\Sigma = \{u_1, u_2, \ldots\}$ условий его проведения и некоторым множеством $\mathfrak{J} = \{A, B, C, A_1, A_2, \ldots\}$ допустимых исходов. В результате проведения эксперимента E мы можем фиксировать, наблюдать или регистрировать некоторые или все его допустимые исходы (результаты). Более того, иногда целесообразно считать, что множеству $\mathfrak{J}$ принадлежат исходы, которые никогда не наступают при проведении эксперимента E. Будем теперь рассматривать статистически устойчивые эксперименты E.

Среди всех допустимых исходов эксперимента E выделим некоторые: $A'_1 \in \mathfrak{J}$, $A'_2 \in \mathfrak{J}, \ldots$, которые будем называть *элементарными* (атомарными, неразложимыми, помеченными) *событиями*. Множество (пространство) элементарных исходов эксперимента E обозначим через $\mathfrak{J}' = \{A'_1, A'_2, \ldots\} \subset \mathfrak{J}$. Предполагается, что $\mathfrak{J}'$ не содержит одинаковых элементов. Заметим, что множество $\mathfrak{J}'$ может быть конечным, счетным (бесконечным) или несчетным. Это обстоятельство зависит от эксперимента E. Для каждого элементарного исхода A'_i из множества $\mathfrak{J}'$ выберем его описание с помощью некоторого языка. Такими языками могут быть русский, английский, немецкий, французский, итальянский, математический и т. п. Это описание (имя) будем обозначать через ω_i. Часто описание произвольного элементарного исхода из множества $\mathfrak{J}'$ будем обозначать просто греческой буквой ω без индекса. Таким образом, каждому элементарному исходу A' из класса $\mathfrak{J}'$ ставим во взаимно однозначное соответствие его описание ω и будем в дальнейшем писать равенство $A' = \{\omega\} \in \mathfrak{J}'$. Теперь одноточечное множество $A' = \{\omega\}$ является математическим объектом, и этот объект будем называть *элементарным событием*. Совокупность описаний всех отобранных элементарных исходов будем называть *множеством описаний элементарных исходов* и обозначать символом Ω.

При формальном построении модели реальных статистически устойчивых процессов E мы должны априори учитывать, что комплекс условий Σ может и не определять однозначно того, что называется элементарным исходом эксперимента. Поэтому при повторениях эксперимента с точным соблюдением комплекса условий Σ конкретное элементарное событие вида $A' = \{\omega\}$ может произойти, а может и не произойти. Такое элементарное событие естественно называть *случайным*. Заметим, что среди элементарных исходов иногда могут существовать такие, которые никогда не наступают при каждом проведении эксперимента E.

Итак, элементарное событие A' есть элементарный исход, представленный в виде одноточечного множества $\{\omega\}$ после выбора соответствующего описания. Следует заметить, что элементарное событие есть математический объект и его следует отличать от реального элементарного исхода эксперимента E.

Выделение элементарных результатов среди допустимых, как правило, является неоднозначным, и оно вызывает значительные трудности. Более того, не существует общего метода для решения этой проблемы. Каждый раз при решении этой проблемы исследователь чаще всего полагается на интуицию и опыт. Однако при выборе элементарных результатов (исходов) всегда будем накладывать следующие три аксиомы выбора:

- при проведении эксперимента E обязательно наступает один из элементарных исходов;
- если происходит некоторый элементарный исход, то все остальные элементарные исходы наступать в этом испытании не могут;
- по элементарному исходу, который происходит при проведении эксперимента E, можно определить наступление или ненаступление любого допустимого исхода A; в этом случае каждый допустимый исход A эксперимента E должен представляться через описания элементарных исходов.

Поясним эти ограничения более подробно. Прежде всего, рассмотрим широко распространенный на практике случай, когда любой из элементарных исходов когда-нибудь наступает при проведении эксперимента E. Соответствующее этому случаю каждое из множеств (пространств) $\mathfrak{I}'$ и Ω будем называть *регулярным*. Далее, для любого допустимого исхода A из множества $\mathfrak{I}$ с помощью многократного проведения эксперимента E согласно аксиомам выбора можно выделить все элементарные исходы A'_{a_1}, $A'_{a_2}, \ldots$, которые происходили одновременно с результатом A. Элементарные исходы вида A'_{a_1}, $A'_{a_2}, \ldots$ имеют описание на выбранном языке в виде соответствующих символов ω_{a_1}, $\omega_{a_2}, \ldots$. Тогда этот исход $A \in \mathfrak{I}$ естественно представлять в виде множества $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$, которое в зависимости от A и эксперимента E может быть конечным или бесконечным (счетным или несчетным). В дальнейшем по соглашению будем писать $A = \{\omega_{a_1}, \omega_{a_2}, \ldots\}$ и называть A после этого допустимым случайным событием эксперимента E. Если описание $\omega \in A$, например, $\omega = \omega_{a_1}$ и происходит элементарный исход $A'_{a_1} = \{\omega_{a_1}\}$, то также происходит и исход A. В силу этого будем говорить, что события A'_{a_1}, $A'_{a_2}, \ldots$ благоприятствуют появлению данного события A. Теперь можно сказать, что случайное событие A есть представленный в виде множества $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$ допустимый исход A реального эксперимента E.

Если нет элементарных исходов, которые благоприятствуют появлению данного исхода A, то этот исход никогда не происходит при проведении эксперимента E и, значит, не содержит ни одного описания $\omega \in \Omega$. Такой исход будем называть *невозможным*, и представлять его в виде пустого множества $\varnothing$.

Напротив, если каждый элементарный исход благоприятствует появлению данного исхода A, то этот исход наступает каждый раз при проведении эксперимента E. Такой исход будем называть *достоверным событием*, и представлять его в виде множества Ω всех описаний.

Итак, между реальным исходом A эксперимента E и случайным событием — подмножеством $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$ множества Ω установлено взаимно однозначное соответствие. Заметим, что как случайное элементарное событие, так и случайное событие есть математический объект.

В общем случае, когда множество элементарных исходов не является регулярным, каждому допустимому исходу A из $\mathfrak{J}$ можно поставить во взаимно однозначное соответствие некоторое подмножество множества Ω следующим способом.

Выберем произвольное непустое подмножество $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$ множества Ω такое, что при многократном проведении эксперимента E когда-нибудь наступает один из элементарных исходов $A'_{a_1}, A'_{a_2}, \ldots$. Тогда поставим во взаимно однозначное соответствие подмножеству $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$ некоторый допустимый исход A, который наступает тогда и только тогда, когда происходит один из элементарных исходов вида $A'_{a_1}, A'_{a_2}, \ldots$. При этом будем говорить, что подмножество $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$ представляет собой допустимый результат A эксперимента E. Ради простоты, будем обозначать подмножество вида $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$ тем же символом A и называть его случайным событием. Если оказалось, что такого допустимого исхода A из совокупности $\mathfrak{J}$ нет, то подмножество $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$ не является случайным событием и вместо подмножества $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$ выбираем с аналогичными свойствами другое подмножество $\{\omega_{b_1}, \omega_{b_2}, \ldots\}$.

Наконец, подмножеству $\{\omega_{c_1}, \omega_{c_2}, \ldots\}$, для которого каждый из элементарных исходов $A'_{c_1}, A'_{c_2}, \ldots$ никогда не происходит, ставим во взаимно однозначное соответствие некоторый допустимый исход $C \in \mathfrak{J}$, если он никогда не наступает. При этом математический объект $C = \{\omega_{c_1}, \omega_{c_2}, \ldots\}$ называется также *случайным событием*. Если такого допустимого исхода нет, то множество $\{\omega_{c_1}, \omega_{c_2}, \ldots\}$ не объявляем случайным событием. Описанную процедуру продолжаем до тех пор, пока не переберем все допустимые исходы из $\mathfrak{J}$. Невозможный исход $\varnothing$ никогда не наступает (необходимое условие) при проведении эксперимента E и обязательно не содержит ни одного описания $\omega \in \Omega$. Напротив, достоверный исход происходит каждый раз (необходимое условие) при проведении эксперимента E и отождествляется с пространством Ω.

Таким образом, все допустимые результаты из множества $\mathfrak{J}$ для эксперимента E должны представляться с помощью некоторой непустой совокупности подмножеств множества Ω. При этом различным подмножествам $\{\omega_{a_1}, \omega_{a_2}, \ldots\} = A$ и $\{\omega_{b_1}, \omega_{b_2}, \ldots\} = B$ множества Ω всегда соответствуют различные допустимые исходы A и B, хотя иногда исход A происходит тогда и только тогда, когда происходит исход B. В последнем случае каждый из подмножеств $\{\omega_{a_1}, \omega_{a_2}, \ldots\}$ и $\{\omega_{b_1}, \omega_{b_2}, \ldots\}$ включает одну и ту же совокупность описаний таких элементарных исходов, которые благоприятствуют появлению каждого из событий A и B. Теперь мы можем коротко зафиксировать предыдущие рассуждения.

Определение 1.1. Любую совокупность $A = \{\omega_{a_1}, \omega_{a_2}, \ldots\} \in \mathfrak{J}$ из описаний элементарных исходов будем называть *случайным событием*.

Случайные события, как соответствующие допустимым исходам эксперимента E, естественно обозначать латинскими прописными буквами без индексов или с индексами, например, A, B, C, A_1, A_2,.... О некотором случайном событии A имеет смысл говорить лишь в том случае, если при осуществлении комплекса условий Σ известно, что соответствующий ему допустимый исход произошел или не произошел. Если, например, случайное событие $A = \{\omega_{a_1}, \omega_{a_2}, \ldots\}$, то оно происходит всякий раз, когда наступает одно из элементарных событий $\{\omega_{a_1}\}$, $\{\omega_{a_2}\}$,.... Если $A \in \mathfrak{J}$ содержит только одно описание $\omega \in \Omega$, т. е. $A = \{\omega\}$, то $\{\omega\}$ есть случайное элементарное событие. Так как множество Ω содержит описания всех элементарных исходов эксперимента E и $A = \{\omega_{a_1}, \omega_{a_2}, \ldots\}$, то случайное событие A есть некоторое подмножество множества Ω и теперь можно писать $A \subset \Omega$ для всех $A \in \mathfrak{J}$. Известно, что $\Omega \subseteq \Omega$ и пустое множество $\varnothing \subset \Omega$. Поэтому целесообразно ввести следующие определения.

Определение 1.2. Множество Ω называется *достоверным событием*.

Элементами Ω являются описания всех элементарных исходов эксперимента E. Соответствующий достоверному событию Ω некоторый допустимый исход эксперимента E называется достоверным и естественно обозначается буквой Ω. На содержательном уровне достоверный исход в результате осуществления комплекса условий эксперимента неизбежно происходит и задается множеством $\mathfrak{J}'$.

Определение 1.3. Пустое множество $\varnothing$ по отношению к достоверному событию Ω эксперимента E будем называть *невозможным событием*.

Пустое множество $\varnothing$ не содержит никакой элемент $\omega \in \Omega$. Поэтому соответствующий невозможному событию $\varnothing$ некоторый исход эксперимента E заведомо не может произойти (необходимое условие) при осуществлении комплекса условий Σ и очевидно не может быть элементарным. Такой исход будем называть невозможным и обозначать символом $\varnothing$. Более того, невозможное событие может соответствовать всем таким исходам других экспериментов, которые никак не связаны с экспериментом E и никогда не наступают при его проведении. Целесообразность введения невозможного события будет ясна в дальнейшем при определении и изучении свойств теоретико-множественных операций над случайными событиями для любого статистически устойчивого эксперимента E. Ясно, что невозможное событие и достоверное событие можно рассматривать как частные случаи событий. Итак, для любого статистически устойчивого эксперимента E мы выделили элементарные события, случайные события, достоверное событие, невозможное событие.

В заключение этого раздела отметим важность соблюдения трех ограничений при выборе элементарных событий. На содержательном уровне три ограничения на выбор элементарных исходов означают, что множество $\mathfrak{J}'$ содержит почти всю информацию об эксперименте E. Например, при наступлении некоторого элементарного исхода мы в точности знаем, какие исходы из $\mathfrak{J}$ произошли, а какие — нет. С другой стороны, для некоторых экспериментов E могут быть выбраны элементарные исходы, которые никогда не наступают при его проведении. Такого рода элементарные исходы необходимо отличать от невозможного исхода эксперимента E. Такая ситуация будет подробно описана в следующей главе при построении математических моделей так называемых условных экспериментов. В следующем разделе этой главы обсудим значительное число примеров, которые будут играть определенную роль при изложении основ теории вероятностей.

2.2. Примеры выбора описаний исходов статистически устойчивого эксперимента $\boldsymbol{E}$. Прежде всего, рассмотрим подробно простые примеры различных способов выбора элементарных исходов и их описаний. При этом следует учитывать тот замечательный факт, что в примерах 1.15–1.19, 1.21–1.27 множество $\mathfrak{J}$ всех допустимых исходов включает регулярное множество $\mathfrak{J}'$ из элементарных исходов. Поэтому методически иногда удобно сначала выбрать все элементарные исходы, а затем, используя этот выбор, определить множество всех допустимых исходов.

Пример 1.15. Пусть эксперимент E заключается в случайном подбрасывании с помощью некоторого механизма на поверхность стола двух монет. Каждая из монет имеет две стороны. На одной стороне изображен герб, а на другой — решетка. Эти стороны целесообразно обозначать в виде знака Г (или цифры 0) и знака Р (или цифры 1). На этом опыте проиллюстрируем некоторое представление о возможностях того или иного способа выбора случайных элементарных исходов. Всего будет предложено семь способов. При этом различные способы выбора будем помечать верхними цифровыми индексами в круглых скобках.

Первый способ. В опыте с подбрасыванием двух различимых монет можно выбрать следующие элементарные исходы:

- результат $A_1^{(1)} = \{\omega_1^{(1)} = (\Gamma, \Gamma)\}$ выпадения на первой монете герба и на второй монете герба с описанием в виде вектора $\omega_1^{(1)} = (\Gamma, \Gamma) = (0, 0)$;
- результат $A_2^{(1)} = \{\omega_2^{(1)} = (\Gamma, \mathrm{P})\}$ выпадения на первой монете герба и на второй монете решетки с описанием в виде вектора $\omega_2^{(1)} = (\Gamma, \mathrm{P}) = (0,1)$;
- результат $A_3^{(1)} = \{\omega_3^{(1)} = (\mathrm{P}, \Gamma)\}$ выпадения на первой монете решетки и на второй монете герба с описанием в виде вектора $\omega_3^{(1)} = (\mathrm{P}, \Gamma) = (1, 0)$;
- результат $A_4^{(1)} = \{\omega_4^{(1)} = (\mathrm{P}, \mathrm{P})\}$ выпадения на первой монете решетки и на второй монете решетки с описанием в виде вектора $\omega_4^{(1)} = (\mathrm{P}, \mathrm{P}) = (1, 1)$.

Ниже, ради определенности, для описания исходов не будем использовать цифры 0 и 1. Первый способ выбора элементарных событий позволяет найти пространство $\Omega^{(1)} = \{\omega_1^{(1)}, \omega_2^{(1)}, \omega_3^{(1)}, \omega_4^{(1)}\}$ из описаний элементарных исходов и легко выписать все допустимые исходы этого эксперимента в виде множества

$$\begin{aligned}\mathfrak{I}^{(1)} = \{A\colon A \subset \Omega^{(1)}\} = \\ = \big\{&\{\omega_1^{(1)}\}, \{\omega_2^{(1)}\}, \{\omega_3^{(1)}\}, \{\omega_4^{(1)}\}, \{\omega_1^{(1)}, \omega_2^{(1)}\}, \{\omega_1^{(1)}, \omega_3^{(1)}\}, \\ &\{\omega_1^{(1)}, \omega_4^{(1)}\}, \{\omega_2^{(1)}, \omega_3^{(1)}\}, \{\omega_2^{(1)}, \omega_4^{(1)}\}, \{\omega_3^{(1)}, \omega_4^{(1)}\}, \\ &\{\omega_1^{(1)}, \omega_2^{(1)}, \omega_3^{(1)}\}, \{\omega_1^{(1)}, \omega_2^{(1)}, \omega_4^{(1)}\}, \\ &\{\omega_1^{(1)}, \omega_3^{(1)}, \omega_4^{(1)}\}, \{\omega_2^{(1)}, \omega_3^{(1)}, \omega_4^{(1)}\}, \Omega^{(1)}, \varnothing\big\}.\end{aligned}$$

Например, исход $\{\omega_2^{(1)}, \omega_3^{(1)}, \omega_4^{(1)}\} = \{(\Gamma, \mathrm{P}), (\mathrm{P}, \Gamma), (\mathrm{P}, \mathrm{P})\}$ состоит в том, что выпадает, по крайней мере, одна решетка.

Второй способ. Если при первом способе проведения эксперимента монеты считали различимыми, то при втором способе монеты являются физически неразличимыми. В этой задаче выберем теперь элементарные исходы несколько иначе:

- результат $A_1^{(2)} = \{\omega_1^{(2)} = \{\Gamma, \Gamma\}\}$ с описанием $\omega_1^{(2)}$ в виде множества $\{\Gamma, \Gamma\}$ означает, что обе монеты выпали гербами вверх;
- результат $A_2^{(2)} = \{\omega_2^{(2)} = \{\Gamma, \mathrm{P}\}\}$ с описанием $\omega_2^{(2)}$ в виде множества $\{\Gamma, \mathrm{P}\}$ означает, что монеты выпали разными сторонами;
- результат $A_3^{(2)} = \{\omega_3^{(2)} = \{\mathrm{P}, \mathrm{P}\}\}$ с описанием $\omega_3^{(2)}$ в виде множества $\{\mathrm{P}, \mathrm{P}\}$ означает, что обе монеты выпали решетками вверх.

В случае физически неразличимых монет пространство $\Omega^{(2)} = \{\omega_1^{(2)}, \omega_2^{(2)}, \omega_3^{(2)}\}$, а множество допустимых исходов равно

$$\begin{aligned}\mathfrak{I}^{(2)} = \big\{&\{\omega_1^{(2)}\}, \{\omega_2^{(2)}\}, \{\omega_3^{(2)}\}, \\ &\{\omega_1^{(2)}, \omega_2^{(2)}\}, \{\omega_1^{(2)}, \omega_3^{(2)}\}, \{\omega_2^{(2)}, \omega_3^{(2)}\}, \Omega^{(2)}, \varnothing\big\}.\end{aligned}$$

Тогда результат $\{\omega_1^{(2)}, \omega_2^{(2)}\} = \{\{\Gamma, \Gamma\}, \{\Gamma, \mathrm{P}\}\}$ состоит в том, что выпадает, по крайней мере, один герб, а событие $\{\omega_1^{(2)}, \omega_3^{(2)}\} = \{\{\Gamma, \Gamma\}, \{\mathrm{P}, \mathrm{P}\}\}$ заключается в том, что монеты выпали одинаковыми сторонами.

Заметим, что результат выпадения на первой монете герба и одновременно на второй монете решетки является случайным событием при первом способе, т. е. $A_2^{(1)} = \{\omega_2^{(1)} = (\Gamma, \mathrm{P})\}$, $A_2^{(1)} \in \mathfrak{J}^{(1)}$, и нельзя будет называть случайным событием при втором способе, т. е. $A_2^{(1)} \notin \mathfrak{J}^{(2)}$. Действительно, если произошло случайное элементарное событие $A_2^{(2)} = \{\omega_2^{(2)} = \{\Gamma, \mathrm{P}\}\}$, то при неразличимых монетах невозможно дать ни утвердительного, ни отрицательного ответа о выпадении на первой монете герба и на второй монете решетки. Аналогичное замечание можно сделать относительно следующих случайных событий:

$$\{\omega_3^{(1)}\},\ \{\omega_1^{(1)}, \omega_2^{(1)}\},\ \{\omega_1^{(1)}, \omega_3^{(1)}\},\ \{\omega_2^{(1)}, \omega_4^{(1)}\},\ \{\omega_3^{(1)}, \omega_4^{(1)}\},$$
$$\{\omega_1^{(1)}, \omega_2^{(1)}, \omega_4^{(1)}\},\ \{\omega_1^{(1)}, \omega_3^{(1)}, \omega_4^{(1)}\}$$

из $\mathfrak{J}^{(1)}$. Напротив, при появлении любого элементарного исхода, выбранного первым способом, мы в точности знаем, какие исходы из $\mathfrak{J}^{(2)}$ произошли, а какие — нет. Однако при этом следует помнить, что исходы из $\mathfrak{J}^{(2)}$ имеют несколько иное представление, чем исходы из $\mathfrak{J}^{(1)}$. Например, один и тот же исход, состоящий в том, что выпадает, по крайней мере, одна решетка, имеет следующие представления. В форме $\{\omega_2^{(1)}, \omega_3^{(1)}, \omega_4^{(1)}\} = \{(\Gamma, \mathrm{P}), (\mathrm{P}, \Gamma), (\mathrm{P}, \mathrm{P})\}$ — при первом способе выбора элементарных исходов и в виде $\{\omega_2^{(2)}, \omega_3^{(2)}\} = \{\{\Gamma, \mathrm{P}\}, \{\mathrm{P}, \mathrm{P}\}\}$ — при втором способе выбора элементарных исходов. Замечания, которые приведены в этом абзаце, можно высказать с небольшими изменениями и к остальным способам выбора элементарных исходов, например, к третьему способу.

Третий способ. В эксперименте с подбрасыванием двух различимых монет, когда невозможно фиксирование стороны второй монеты при выпадении на первой монете решетки, можно предложить следующий выбор элементарных исходов:

- результат $A_1^{(3)} = \{\omega_1^{(3)} = (\Gamma, \Gamma)\}$ выпадения на первой монете герба и на второй монете герба с описанием в виде вектора $\omega_1^{(3)} = (\Gamma, \Gamma)$;
- результат $A_2^{(3)} = \{\omega_2^{(3)} = (\Gamma, \mathrm{P})\}$ выпадения на первой монете герба и на второй монете решетки с описанием в виде вектора $\omega_2^{(3)} = (\Gamma, \mathrm{P})$;
- результат $A_3^{(3)} = \{\omega_3^{(3)} = (\mathrm{P}, \cdot)\}$ выпадения на первой монете решетки с описанием в виде вектора $\omega_3^{(3)} = (\mathrm{P}, \cdot)$.

Нетрудно найти пространство $\Omega^{(3)} = \{\omega_1^{(3)}, \omega_2^{(3)}, \omega_3^{(3)}\}$, а множество допустимых исходов этого эксперимента можно представить в виде:

$$\mathfrak{I}^{(3)} = \big\{\{\omega_1^{(3)}\}, \{\omega_2^{(3)}\}, \{\omega_3^{(3)}\},\\ \{\omega_1^{(3)}, \omega_2^{(3)}\}, \{\omega_1^{(3)}, \omega_3^{(3)}\}, \{\omega_2^{(3)}, \omega_3^{(3)}\}, \Omega^{(3)}, \varnothing\big\}.$$

Например, результат $\{\omega_2^{(3)}, \omega_3^{(3)}\} = \{(\Gamma, \mathrm{P}), (\mathrm{P}, \cdot)\}$ для такого выбора состоит в том, что выпадает, по крайней мере, одна решетка.

Приведем еще ряд способов выбора элементарных исходов в задаче о подбрасывании двух монет.

Четвертый способ. Этот способ используется при невозможности фиксирования стороны первой монеты, если на второй монете выпадает решетка. Выберем в качестве элементарных исходов:

- результат $A_1^{(4)} = \{\omega_1^{(4)} = (\Gamma, \Gamma)\}$ выпадения на первой монете герба и на второй монете герба с описанием в виде вектора $\omega_1^{(4)} = (\Gamma, \Gamma)$;
- результат $A_2^{(4)} = \{\omega_2^{(4)} = (\mathrm{P}, \Gamma)\}$ выпадения на первой монете решетки и на второй монете герба с описанием в виде вектора $\omega_2^{(4)} = (\mathrm{P}, \Gamma)$;
- результат $A_3^{(4)} = \{\omega_3^{(4)} = (\cdot, \mathrm{P})\}$ выпадения на второй монете решетки с описанием в виде вектора $\omega_3^{(4)} = (\cdot, \mathrm{P})$.

В этом случае пространство $\Omega^{(4)} = \{\omega_1^{(4)}, \omega_2^{(4)}, \omega_3^{(4)}\}$. Найдем допустимые исходы такого эксперимента в виде следующего множества:

$$\mathfrak{I}^{(4)} = \big\{\{\omega_1^{(4)}\}, \{\omega_2^{(4)}\}, \{\omega_3^{(4)}\},\\ \{\omega_1^{(4)}, \omega_2^{(4)}\}, \{\omega_1^{(4)}, \omega_3^{(4)}\}, \{\omega_2^{(4)}, \omega_3^{(4)}\}, \Omega^{(4)}, \varnothing\big\}.$$

Например, результат $\{\omega_1^{(4)}, \omega_2^{(4)}\} = \{(\Gamma, \Gamma), (\mathrm{P}, \Gamma)\}$ состоит в том, что на второй монете выпадает герб.

Пятый способ. Этот способ соответствует случаю невозможности фиксирования стороны второй монеты, если на первой монете выпадает герб. Выберем в качестве элементарных исходов:

- результат $A_1^{(5)} = \{\omega_1^{(5)} = (\mathrm{P}, \mathrm{P})\}$ выпадения на первой и второй монетах решетки с описанием в виде вектора $\omega_1^{(5)} = (\mathrm{P}, \mathrm{P})$;
- результат $A_2^{(5)} = \{\omega_2^{(5)} = (\mathrm{P}, \Gamma)\}$ выпадения на первой монете решетки и на второй монете герба с описанием в виде вектора $\omega_2^{(5)} = (\mathrm{P}, \Gamma)$;
- результат $A_3^{(5)} = \{\omega_3^{(5)} = (\Gamma, \cdot)\}$ выпадения на первой монете герба с описанием в виде вектора $\omega_3^{(5)} = (\Gamma, \cdot)$.

Получаем пространство $\Omega^{(5)} = \{\omega_1^{(5)}, \omega_2^{(5)}, \omega_3^{(5)}\}$. Найдем допустимые исходы эксперимента в виде следующего множества:

$$\mathfrak{J}^{(5)} = \big\{\{\omega_1^{(5)}\}, \{\omega_2^{(5)}\}, \{\omega_3^{(5)}\}, \\ \{\omega_1^{(5)}, \omega_2^{(5)}\}, \{\omega_1^{(5)}, \omega_3^{(5)}\}, \{\omega_2^{(5)}, \omega_3^{(5)}\}, \Omega^{(5)}, \varnothing\big\}.$$

Например, исход $\{\omega_1^{(5)}, \omega_2^{(5)}\} = \{(\text{Р}, \text{Р}), (\text{Р}, \Gamma)\}$ состоит в том, что на первой монете выпадает решетка.

Шестой способ. Рассмотрим теперь подбрасывание двух различимых монет, когда невозможно фиксирование стороны первой монеты, если на второй монете выпадает герб. Возьмем в качестве элементарных исходов:

- результат $A_1^{(6)} = \{\omega_1^{(6)} = (\text{Р}, \text{Р})\}$ выпадения на первой и второй монетах решетки с описанием в виде вектора $\omega_1^{(6)} = (\text{Р}, \text{Р})$;
- результат $A_2^{(6)} = \{\omega_2^{(6)} = (\Gamma, \text{Р})\}$ выпадения на первой монете герба и на второй монете решетки с описанием в виде вектора $\omega_2^{(6)} = (\Gamma, \text{Р})$;
- результат $A_3^{(6)} = \{\omega_3^{(6)} = (\cdot, \Gamma)\}$ выпадения на второй монете герба с описанием в виде вектора $\omega_3^{(6)} = (\cdot, \Gamma)$.

Здесь пространство $\Omega^{(6)} = \{\omega_1^{(6)}, \omega_2^{(6)}, \omega_3^{(6)}\}$. Все допустимые исходы при таком способе выбора записываются в виде следующего множества:

$$\mathfrak{J}^{(6)} = \big\{\{\omega_1^{(6)}\}, \{\omega_2^{(6)}\}, \{\omega_3^{(6)}\}, \\ \{\omega_1^{(6)}, \omega_2^{(6)}\}, \{\omega_1^{(6)}, \omega_3^{(6)}\}, \{\omega_2^{(6)}, \omega_3^{(6)}\}, \Omega^{(6)}, \varnothing\big\}.$$

Например, исход $\{\omega_2^{(6)}, \omega_3^{(6)}\} = \{(\Gamma, \text{Р}), (\cdot, \Gamma)\}$ состоит в том, что хотя бы один раз появится герб.

Седьмой способ. Последний способ проведения эксперимента целесообразно применять в случае невозможности определения герба и решетки, если монеты выпали одинаковыми сторонами. Выберем в качестве элементарных исходов:

- результат $A_1^{(7)} = \{\omega_1^{(7)} = (\Gamma, \text{Р})\}$ выпадения на первой монете герба и на второй монете решетки с описанием в виде вектора $\omega_1^{(7)} = (\Gamma, \text{Р})$;
- результат $A_2^{(7)} = \{\omega_2^{(7)} = (\text{Р}, \Gamma)\}$ выпадения на первой монете решетки и на второй монете герба с описанием в виде вектора $\omega_2^{(7)} = (\text{Р}, \Gamma)$;
- результат $A_3^{(7)} = \{\omega_3^{(7)} = \{(\Gamma, \Gamma), (\text{Р}, \text{Р})\}\}$ выпадения на монетах одинаковых сторон с описанием в виде множества $\omega_3^{(7)} = \{(\Gamma, \Gamma), (\text{Р}, \text{Р})\}$.

Пространство описаний элементарных исходов и множество допустимых исходов при таком способе выбора записываются в следующем виде: $\Omega^{(7)} = \{\omega_1^{(7)}, \omega_2^{(7)}, \omega_3^{(7)}\}$ и

$$\mathfrak{J}^{(7)} = \bigl\{\{\omega_1^{(7)}\}, \{\omega_2^{(7)}\}, \{\omega_3^{(7)}\},\\ \{\omega_1^{(7)}, \omega_2^{(7)}\}, \{\omega_1^{(7)}, \omega_3^{(7)}\}, \{\omega_2^{(7)}, \omega_3^{(7)}\}, \Omega^{(7)}, \varnothing\bigr\}.$$

Например, исход $\{\omega_1^{(7)}, \omega_2^{(7)}\} = \{(\Gamma, \mathrm{P}), (\mathrm{P}, \Gamma)\}$ состоит в том, что монеты выпали разными сторонами.

Подробно рассмотренный простой пример с подбрасыванием двух монет показывает, что не следует слишком полагаться на интуицию. При выборе элементарных исходов эксперимента E необходимо быть очень осторожным. Например, при втором и третьем способах на интуитивном уровне представлялось, что можно в качестве множества описаний всех элементарных исходов выбрать $\Omega^{(1)}$. При этом множество $\mathfrak{J}^{(1)}$ будет являться уже множеством допустимых исходов как для второго опыта, так и для третьего. Однако адекватный выбор для этих случаев оказался несколько другим и совершенно неожиданным по отношению к нашей интуиции. Каждый такой выбор должен основываться на глубоком знании основных условий проведения эксперимента E и его допустимых исходов. Так, в примере с подбрасыванием двух монет мы должны существенным образом учитывать условия, различаются или нет монеты физически, можно ли фиксировать стороны монет или нет.

Пример 1.16. Пусть монета бросается три раза. В комплекс условий входят фиксированная монета определенного достоинства, методика бросания монеты, поверхность и освещенность стола и т. п.

Пространство описаний элементарных событий для этого эксперимента будет равно

$$\Omega = \bigl\{\omega_1 = (\Gamma, \Gamma, \Gamma), \omega_2 = (\Gamma, \Gamma, \mathrm{P}), \omega_3 = (\Gamma, \mathrm{P}, \Gamma), \omega_4 = (\mathrm{P}, \Gamma, \Gamma),\\ \omega_5 = (\Gamma, \mathrm{P}, \mathrm{P}), \omega_6 = (\mathrm{P}, \Gamma, \mathrm{P}), \omega_7 = (\mathrm{P}, \mathrm{P}, \Gamma), \omega_8 = (\mathrm{P}, \mathrm{P}, \mathrm{P})\bigr\},$$

где символ «Г» условно обозначает выпадение герба, а символ «Р» условно обозначает выпадение решетки. Здесь, например, описание $\omega_3 = (\Gamma, \mathrm{P}, \Gamma)$ соответствует элементарному исходу выпадения решетки только при втором броске. Случайное событие A (выпало не менее двух гербов) можно описать совокупностью $A = \{(\Gamma, \Gamma, \Gamma), (\Gamma, \Gamma, \mathrm{P}), (\Gamma, \mathrm{P}, \Gamma), (\mathrm{P}, \Gamma, \Gamma)\} \subset \Omega$. Событие B (выпала в точности одна решетка) представляется совокупностью $B = \{(\Gamma, \Gamma, \mathrm{P}), (\Gamma, \mathrm{P}, \Gamma), (\mathrm{P}, \Gamma, \Gamma)\} \subset \Omega$. Появление четырех гербов — пример невозможного события $\varnothing$. Пример достоверного события Ω — выпадение герба или решетки.

Пример 1.17. Эксперимент заключается в проведении денежно-вещевой лотереи. В комплекс условий входят общее количество билетов, количество и наименование разыгрываемых предметов, правила проведения лотереи и т. п. Пусть в денежно-вещевую лотерею входят десять билетов, разыгрывается автомобиль и билет вытаскивается только один раз.

Пространство описаний элементарных событий состоит из номеров от 1 до 10, на каждый из которых может выпасть выигрыш с наименованием «автомобиль». Например, событие A заключается в том, что выигрыш выпадет на билет с четным номером. В этом случае $A = \{2, 4, 6, 8, 10\}$. Невозможное событие — выигрыша не будет.

Пример 1.18. Подбрасывается игральная кость и определяется выпавшее число очков. Для этого простого примера пространство $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ описаний элементарных исходов состоит из шести элементов. Здесь символ ω_k означает описание элементарного исхода о выпадении k очков, где $k = 1, 2, 3, 4, 5, 6$.

Пример 1.19. Подбрасываются две игральные кости и определяются числа выпавших очков на каждой из костей. Пространство описаний элементарных исходов имеет вид:

$$\begin{aligned}\Omega = \{&(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\ &(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\ &(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),\\ &(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\ &(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),\\ &(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}.\end{aligned}$$

Здесь при каждом $i = 1, 2, \ldots, 36$ символ $\omega_i = (x, y)$ есть описание такого i-го элементарного исхода, когда на первой кости выпадает x очков, а на второй — y очков.

Пример 1.20. Подбрасывается монета до первого выпадения герба (орла) с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Выпадение только решеток или появление герба здесь является достоверным событием Ω с описанием вида $\Omega = \{(\text{Р}, \text{Р}, \ldots), \text{Г}, (\text{Р}, \text{Г}), (\text{Р}, \text{Р}, \text{Г}), \ldots\}$; выпадение герба при втором броске будет случайным элементарным событием вида $(\text{Р}, \text{Г})$; наблюдение герба дважды — невозможное событие $\varnothing$; появление решетки (цифры) не более трех раз является случайным событием $A = \{\text{Г}, (\text{Р}, \text{Г}), (\text{Р}, \text{Р}, \text{Г}), (\text{Р}, \text{Р}, \text{Р}, \text{Г})\}$. Заметим, что для этого эксперимента E выпадение только решеток является выбранным элементарным исходом вида $\{(\mathrm{P}, \mathrm{P}, \ldots)\}$. Однако элементарный исход $\{(\mathrm{P}, \mathrm{P}, \ldots)\}$ никогда не наступает при проведении конечного числа испытаний, является одноточечным множеством и поэтому он не может быть объявлен невозможным событием. Поэтому множество Ω не является регулярным.

Пример 1.21. Колода в 52 игральные карты сдается четырем игрокам и определяется состав тринадцати карт, полученных первым игроком. Здесь $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, где $n = C_{52}^{13} = 635\,013\,559\,600$. При каждом $i = 1, 2, \ldots, n$ символ $\omega_i = \{x_1, x_2, \ldots, x_{13}\}$ есть описание такого элементарного исхода, когда первый игрок имеет множество $\{x_1, x_2, \ldots, x_{13}\}$ из конкретных 13 карт разного достоинства. Очевидно, что это множество является конкретным сочетанием из 52 карты по 13 карт разного достоинства.

Приведем еще три примера, которые часто рассматриваются в различных учебниках по теории вероятностей [21].

Пример 1.22. Имеются три разных шара с пометками a, b, c, и они произвольно распределяются по трем разным ящикам $(\cdot|\cdot|\cdot)$ с пометками Я$_1$, Я$_2$, Я$_3$. Определяется размещение шаров. В этой задаче имеем:

$$\begin{aligned}\Omega =\big\{&(a,b,c|-|-),(-|a,b,c|-),(-|-|a,b,c),(a,b|c|-),\\ &(a,c|b|-),(b,c|a|-),(a,b|-|c),(a,c|-|b),\\ &(b,c|-|a),(a|b,c|-),(b|a,c|-),(c|a,b|-),\\ &(a|-|b,c),(b|-|a,c),(c|-|a,b),(-|a,b|c),\\ &(-|a,c|b),(-|b,c|a),(-|a|b,c),(-|b|a,c),\\ &(-|c|a,b),(a|b|c),(a|c|b),(b|a|c),(b|c|a),(c|a|b),(c|b|a)\big\}.\end{aligned}$$

Здесь описание $\omega_1 = (a,b,c|-|-)$ соответствует такому размещению, когда в ящик Я$_1$ попали все три шара. Аналогичный смысл имеют и другие описания. Например, описание $\omega_{27} = (c|b|a)$ означает, что после опыта в ящиках Я$_1$, Я$_2$ и Я$_3$ соответственно оказались шары с пометками c, b и a. Следовательно, получаем 27 описаний элементарных исходов в этом эксперименте.

Пример 1.23. Имеются три одинаковых шара и они распределяются по трем разным ящикам Я$_1$, Я$_2$, Я$_3$. Для этого опыта легко получаем, что

$$\begin{aligned}\Omega =\big\{&(a,a,a|-|-),\ (-|a,a,a|-),\ (-|-|a,a,a),\ (a,a|a|-),\\ &(a,a|-|a),\ (a|a,a|-),\ (-|a,a|a),\ (a|-|a,a),\ (-|a|a,a),\ (a|a|a)\big\}.\end{aligned}$$

Итак, имеем десять описаний элементарных исходов в этом опыте.

Пример 1.24. Имеется три одинаковых шара и они распределяются по трем одинаковым ящикам $Я_1$, $Я_2$, $Я_3$. В этом испытании получаем три описания элементарных исходов и $\Omega = \{\{a, a, a|-|-\}, \{a, a|a|-\}, \{a|a|a\}\}$. Например, описание $\omega_1 = \{a, a, a|-|-\}$ состоит в том, что все три шара попали в один ящик.

Пример 1.25. Пусть выбирается группа из 50 человек. При этом мы интересуемся только числом курящих среди них людей. Пространство $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{50}, \omega_0\}$ описаний всех элементарных событий состоит из 51 элемента. Элементарное событие $\{\omega_i\}$ означает, что в отобранной группе будет ровно $i = 0, 1, \ldots, 50$ курящих людей. Например, событие A, которое означает, что большинство людей выборки будут курящими, является случайным. Событие B, которое заключается в том, что число курящих людей выборки будет не меньше 200, является невозможным.

Пример 1.26. Пусть опять выбирается группа из 50 человек. Но интересуемся теперь уже числом x_1 курящих мужчин, числом x_2 курящих женщин и числом x_3 некурящих мужчин, числом x_4 некурящих женщин.

В качестве пространства Ω описаний элементарных событий следует рассматривать множество всех упорядоченных четверок $\omega = (x_1, x_2, x_3, x_4)$ из целых чисел, заключенных между 0 и 50 и дающих в сумме 50. Например, описание $\omega_1 = (0, 0, 5, 45)$ соответствует такому элементарному событию, когда отобраны пять некурящих мужчин и 45 некурящих женщин. Событие A, которое состоит в том, что среди мужчин доля курящих больше, чем среди женщин, является случайным.

Пример 1.27. На отрезке $[0; 1]$ оси $0x$ выбирается произвольно точка. В качестве описания ω произвольно поставленной точки на этом отрезке можно взять ее абсциссу. Тогда $\Omega = \{\omega = x \colon 0 \leqslant x \leqslant 1\}$. При этом произвольный элементарный исход представляется в виде одноточечного множества $\{x\}$.

В отличие от примера 1.27, для которого достоверное случайное событие является несчетным множеством, в каждом из примеров 1.15–26 достоверное событие содержит счетное или конечное количество элементов. В связи с этим введем следующее определение.

Определение 1.4. Пространство $\mathfrak{J}'$ из элементарных событий с конечным или счетным множеством элементарных событий $\{\omega\}$ называется *дискретным*. Пространство любого другого типа будем называть *недискретным*.

2.3. Соотношения между случайными событиями. Пусть в результате проведения эксперимента наступило или не наступило некоторое случайное событие A. Совокупность $\mathfrak{J}$ всех случайных событий, связанных с данным экспериментом E, играет основную роль в нашем дальнейшем рассмотрении основ этого курса. Понятие случайного события уже имеет абстрактный характер, так как конкретная природа события не имеет значения. Существенно лишь то, что случайный исход A эксперимента E представляется в виде некоторой совокупности из описаний ω тех элементарных событий, которые могут одновременно наступать (не наступать) с исходом A, и что случайное событие $A \subset \Omega$ происходит или нет при осуществлении комплекса условий Σ.

Поэтому между событиями множества $\mathfrak{J}$ если и могут существовать соотношения, то только, в первую очередь, логического и теоретико-множественного характера. Если выбранное описание ω некоторого элементарного события $\{\omega\}$ принадлежит случайному событию A, то будем писать $\omega \in A$. Запись $A \in \mathfrak{J}$ означает, что случайное событие A принадлежит совокупности $\mathfrak{J}$. Противоположные утверждения, состоящие в том, что описание ω элементарного события $\{\omega\}$ не принадлежит случайному событию A и подмножество A не принадлежит семейству $\mathfrak{J}$, записываются в следующем виде: $\omega \notin A$ и $A \notin \mathfrak{J}$. Примем следующие определения.

Определение 1.5. Если A является случайным событием и каждое описание $\omega \in A$ принадлежит случайному событию B, то этот факт фиксируется символической записью $A \subset B$ или $B \supset A$. В этом случае будем говорить, что *случайное событие A есть часть B* или *B включает A*.

Из определения отношения включения вида $A \subset B$ непосредственно следует, что при каждом осуществлении комплекса условий Σ вместе с событием A обязательно наступает и событие B, т. е. событие A влечет за собой событие B или B следует за A. Однако обратное утверждение может и не иметь место, например, когда существуют элементарные исходы, которые никогда не наступают при проведении эксперимента E, и их описания являются элементами A и не являются элементами B. Можно сказать, что условие $A \subset B$ является более жестким, так как все элементы множества A входят в множество B. Однако если $\mathfrak{J}'$ является регулярным и при каждом осуществлении комплекса условий Σ вместе с событием A появляется и событие B, то $A \subset B$.

Определение 1.6. Если $A \subset B$ и $A \supset B$, то мы будем говорить, что *события A и B равносильны*, и обозначать это через $A = B$.

Из определений 1.5 и 1.6 следует, что при $A = B$ случайные события A и B состоят из одних и тех же элементов (описаний элементарных исходов). Следовательно, если $A = B$, то при каждом проведении эксперимента E событие A влечет за собой B и, в то же время, B влечет за собой A, т. е. события A и B оба наступают или оба не наступают.

Обратное утверждение справедливо не всегда. Проиллюстрируем определения 1.5 и 1.6 на следующих простых примерах.

Пример 1.28. Производится бросание игральной кости. Событие A — появление цифры 2, событие B — появление четной цифры. Легко видеть, что для этого опыта появление события A влечет за собой появление события B, т. е. $A \subset B$.

Пример 1.29. Производится выстрел по мишени, составленной из черного круга и ряда колец. Событие A — попадание в черный круг, а событие B — попадание в мишень. Для этого эксперимента при появлении события A происходит и событие B. Следовательно, $A \subset B$.

Пример 1.30. Производится шесть выстрелов по мишени. Событие A — попадание тремя пулями в мишень. Событие B — три промаха при шести выстрелах. При каждом проведении эксперимента E событие A влечет за собой B и, в то же время, B влечет за собой A. Здесь $A \subset B$ и одновременно $B \subset A$, следовательно, $A = B$.

В каждом из примеров 1.28–1.30 нетрудно выбрать множество $\mathfrak{I}'$, содержащее только такие элементарные исходы, которые когда-нибудь наступают при проведении соответствующего эксперимента. В этом случае $\mathfrak{I}'$ является регулярным и определения 1.5, 1.6 можно дать в терминах наступления событий A, B.

Определение 1.7. Два события A и B называются *несовместными* (*попарно несовместными*), если они не содержат общих описаний элементарных исходов. В противном случае события A и B называются совместными, т. е. они содержат хотя бы одно общее описание $\omega \in \Omega$.

Если A и B являются несовместными (несовместимыми), то появление одного из них исключает появление другого (необходимое условие). Обратное утверждение может не выполняться. Можно сказать, что событие, состоящее в одновременном наступлении двух несовместных событий A и B, является невозможным. Разным элементарным исходам соответствуют различные описания. Поэтому любые два элементарных события $\{\omega_i\}$ и $\{\omega_k\}$ являются несовместными при $i \neq k$. Значит, появление одного из элементарных исходов исключает появление другого. Этот факт естественно согласуется со второй аксиомой выбора элементарных исходов.

Определение 1.8. Событие B, содержащее все такие описания элементарных исходов из Ω, которые не принадлежат некоторому событию A, называется *противоположным* по отношению к событию A и обозначается через символ $\overline{A}$.

Например, при одноразовом и случайном подбрасывании симметричной игральной кости исход с выпадением четного числа очков противоположен результату выпадения нечетного числа очков.

Если события A и $\overline{A}$ являются противоположными, то событие A происходит тогда и только тогда, когда не происходит событие $\overline{A}$. Можно это утверждение сформулировать иначе. Если события A и $\overline{A}$ являются противоположными, то наступление одного из них исключает наступление другого и одно из них обязательно наступает при проведении эксперимента E. Обратное утверждение, как и, естественно, любое из обратных утверждений не всегда имеет место.

2.4. Теоретико-множественные операции над случайными событиями. Пусть дано произвольное множество $\{A, B, \ldots\}$ событий из $\mathfrak{J}$. Каждое такое случайное событие из $\mathfrak{J}$ есть некоторое множество из описаний вида ω. Тогда над случайными событиями A, $B, \ldots$ можно ввести известные теоретико-множественные операции, например, объединение $\cup$, пересечение $\cap$, симметрическую разность $\triangle$, разность $\setminus$ и т. д.

Определение 1.9. *Объединением случайных событий* A и B, обозначаемым через $A \cup B$, называется такое случайное событие $C = A \cup B$, которое содержит все описания события A и все те описания события B, которые не входят в A. Если событие $A_i \in \mathfrak{J}$ при каждом $i = 1, 2, \ldots$, то их объединение есть такое случайное событие $C = \bigcup_{i=1}^{\infty} A_i$, каждое описание которого принадлежит хотя бы одному из событий $A_1, A_2, \ldots$.

Следовательно, если $C = A_1 \cup A_2 \cup \ldots$, то случайное событие C происходит тогда и только тогда, когда происходит, по крайней мере, одно из случайных событий A_1, $A_2, \ldots$. Обратное утверждение имеет место, если множество Ω является регулярным.

Определение операции объединения в терминах случайных событий (исходов эксперимента E) можно проинтерпретировать в терминах теории множеств. Например, описание $\omega \in A \cup B$ означает или $\omega \in A$, или $\omega \in B$, другими словами, объединение $A \cup B$ реализуется известным логическим высказыванием «или». Приведем теперь примеры, в которых используется операция объединения случайных событий.

Пример 1.31. В Нижнем Новгороде имеются три транспортных моста через реку Ока. Рассматривается ежедневная возможность переезда наземного транспорта через реку. Обозначим через A_1 событие, которое заключается в исправности первого моста; через A_2 — событие, которое заключается в исправности второго моста, и, наконец, через A_3 — событие исправности третьего моста. Если A есть событие, означающее возможность переезда через реку, то $A = A_1 \cup A_2 \cup A_3 = \bigcup_{i=1}^{3} A_i$.

Пример 1.32. Производится испытание вычислительного узла, состоящего из трех параллельных ветвей, дублирующих друг друга (рис. 1.6).

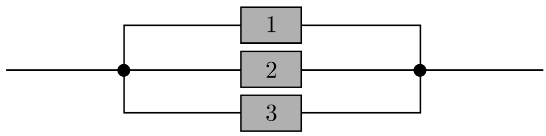


Рис. 1.6

Обозначим через A_1 событие исправности первой ветви; A_2 — событие исправности второй ветви; A_3 — событие исправности третьей ветви; A — событие исправности агрегата в целом. Нетрудно видеть, что событие $A = A_1 \cup A_2 \cup A_3$.

Пример 1.33. Производится три выстрела по мишени. Пусть событие A_1 означает попадание при первом выстреле, событие A_2 — попадание при втором выстреле, событие A_3 — попадание при третьем выстреле, наконец, событие A — попадание хотя бы одного из трех выстрелов. Отсюда легко находим, что $A = A_1 \cup A_2 \cup A_3$.

Определение 1.10. *Пересечением двух случайных событий A и B*, обозначаемым через $A \cap B$, называется такое случайное событие $D = A \cap B$, которое содержит только те описания, которые принадлежат как A, так и B. Если событие $A_i \in \mathfrak{J}$ при каждом $i = 1, 2, \ldots$, то их пересечение есть такое случайное событие $D = \bigcap_{i=1}^{\infty} A_i$, каждое описание которого принадлежит одновременно всем случайным событиям $A_1,\ A_2, \ldots$.

Пусть случайное событие $D = \bigcap_{i=1}^{\infty} A_i$. Отсюда следует, что событие D происходит тогда и только тогда, когда происходят все случайные события $A_1,\ A_2, \ldots$ одновременно. Обратное утверждение имеет место, если множество элементарных исходов является регулярным.

Определение операции пересечения в терминах случайных событий (исходов эксперимента E) можно естественным образом проинтерпретировать на языке теории множеств. Например, описание $\omega \in A \cap B$ означает $\omega \in A$ и $\omega \in B$, другими словами, пересечение $A \cap B$ реализуется логическим высказыванием «и».

Так как событие, состоящее в одновременном наступлении несовместных событий A и B, является невозможным, то $A \cap B = \varnothing$. Для различных элементарных событий $\{\omega_i\}$ и $\{\omega_k\}$ имеем равенство $\{\omega_i\} \cap \{\omega_k\} = \varnothing$. Наоборот, совместность событий A и B означает, что пересечение этих событий не является невозможным, т. е. $A \cap B \neq \varnothing$.

Приведем примеры, в которых используется операция пересечения случайных событий.

Пример 1.34. Узел ЭВМ состоит из трех последовательных элементов. Производится его испытание (рис. 1.7). Пусть A_1 — исправен первый элемент, A_2 — исправен второй элемент, A_3 — исправен третий элемент, A — исправен узел ЭВМ. Для этого опыта $A = A_1 \cap A_2 \cap A_3 = \bigcap_{i=1}^{3} A_i$.

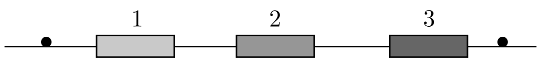


Рис. 1.7

Пример 1.35. Производится три выстрела по мишени. Пусть A_1 — попадание в мишень при первом выстреле, A_2 — попадание в мишень при втором выстреле, A_3 — попадание в мишень при третьем выстреле, A — ровно три попадания. В этом случае событие $A = A_1 \cap A_2 \cap A_3 = \bigcap_{i=1}^{3} A_i$. События A_1 и A_2 совместны, так как попадание в мишень при первом выстреле не исключает попадания в мишень при втором выстреле, т. е. $A_1 \cap A_2 \neq \varnothing$. Если B означает промах по мишени при трех выстрелах, то A и B являются несовместными событиями, следовательно, $A \cap B = \varnothing$.

Пример 1.36. Производится однократное бросание монеты. Появление герба и цифры при однократном бросании есть невозможное событие. Поэтому $A_1 \cap A_2 \neq \varnothing$, где A_1 — появление герба, A_2 — появление цифры.

С помощью операций объединения и пересечения можно формализовать определение 1.8 для противоположных событий. В самом деле, если случайное событие $\overline{A}$ есть совокупность всех тех описаний из Ω, которые не входят в совокупность A, то $A \cup \overline{A} = \Omega$, $A \cap \overline{A} = \varnothing$. Имеет место и обратное утверждение, т. е. если $A \cup B = \Omega$ и $A \cap B = \varnothing$, то события A и B будут противоположными.

Пример 1.37. Производится однократное бросание игральной кости. Пусть событие A — выпадение нечетного числа очков, т. е. $A = \{\omega_1, \omega_3, \omega_5\}$. Тогда $\overline{A}$ — выпадение четного числа очков и $\overline{A} = \{\omega_2, \omega_4, \omega_6\}$. Напомним, что для этого опыта множество $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ описаний элементарных исходов является регулярным и состоит из шести элементов, а элементарное событие $\{\omega_k\}$ означает выпадение на игральной кости k очков, где $k = 1, 2, 3, 4, 5, 6$.

Определение 1.11. *Разностью случайных событий* A и B называется такое случайное событие C, которое содержит описания из A

и не содержит описания из B. Разность событий A и B обозначается символом $A \backslash B$.

Определение операции разности $A \backslash B$ событий A и B в терминах случайных событий (исходов эксперимента E) можно проинтерпретировать в терминах теории множеств. Например, описание $\omega \in A \backslash B$ означает $\omega \in A$ и $\omega \notin B$. Так как разность двух множеств A и B содержит только те описания из A, которые не принадлежат B, то выполняются соотношения вида $A \backslash B = A \cap \overline{B}$. Если $C = A \backslash B$, то случайное событие C состоит в том, что A произошло, а B нет. Обратное утверждение имеет место, если множество Ω является регулярным. В примере 1.33 разность событий A_1 и A_2 представляет собой событие, которое заключается в попадании при первом выстреле и промахе при втором выстреле.

Определение 1.12. *Симметрической разностью случайных событий* A и B, которая обозначается символом $A \triangle B$, называется случайное событие вида $A \triangle B = (A \backslash B) \cup (B \backslash A)$.

Если $C = (A \backslash B) \cup (B \backslash A)$, то происходит только одно из событий: либо A, либо B. Обратное утверждение имеет место, если множество элементарных исходов является регулярным. Заметим, что с помощью основных операций объединения, пересечения и разности по аналогии с определением 1.12 можно определять новые теоретико-множественные операции над случайными событиями. Например, из определения разности случайных событий получаем, что случайное событие $\Omega \backslash A$ содержит описания из Ω и не содержит описания из A. Вспоминая определение противоположных или дополнительных событий, получаем равенство $\overline{A} = \Omega \backslash A$. Так, в эксперименте примера 1.35 при $C = A_1 \cup A_2 \cup A_3$ имеем $C \cap B = \varnothing$ и $C \cup B = \Omega$, т. е. $\overline{C} = B$. Здесь событие C означает хотя бы одно попадание в мишень, и оно является дополнением события B (промах). Одновременно с этим укажем, что событие $C \triangle A = A \triangle C$ происходит тогда и только тогда, когда имеется одно или два попадания при трех выстрелах в мишень.

Определение 1.13. События A_1, $A_2, \ldots$ *образуют полную группу несовместимых событий*, если их объединение $\bigcup_{i=1}^{\infty} A_i$ равно достоверному событию Ω и они попарно являются несовместными.

Например, множество случайных элементарных событий образует полную группу несовместимых событий. Противоположные события A и $\overline{A}$ образуют также полную группу несовместных событий. Если случайные события A_1, $A_2, \ldots, A_n$ образуют полную группу $\left(\bigcup_{i=1}^{n} A_i = \Omega\right)$, то в результате осуществления комплекса условий Σ обязательно произойдет хотя бы одно из них. Обратное утверждение не всегда имеет место.

2.5. Законы теоретико-множественных операций. Рассмотренные теоретико-множественные операции над всеми или некоторыми случайными событиями из $\mathfrak{J}$ удовлетворяют следующим основным законам:

- переместительному (коммутативному) закону для операций объединения и пересечения, т. е. $A \cup B = B \cup A$, $A \cap B = B \cap A$, где $A \in \mathfrak{J}$, $B \in \mathfrak{J}$;
- сочетательному (ассоциативному) закону для операций объединения и пересечения вида: $A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C)$, $A \cap B \cap C = (A \cap B) \cap C = A \cap (B \cap C)$, где $A \in \mathfrak{J}$, $B \in \mathfrak{J}$, $C \in \mathfrak{J}$;
- закону Де Моргана в двух формах: $\overline{A \cup B} = \overline{\overline{A} \cap \overline{B}}$, $A \cap B = \overline{\overline{A} \cup \overline{B}}$, где $A \in \mathfrak{J}$, $B \in \mathfrak{J}$;
- распределительному (дистрибутивному) закону для объединения и пересечения событий $A \in \mathfrak{J}$, $B \in \mathfrak{J}$, $C \in \mathfrak{J}$ в двух формах:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \quad (A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

Кроме того, из определения объединения и пересечения событий непосредственно вытекает: $A \cup A = A$, $A \cup \Omega = \Omega$, $A \cup \varnothing = A$, $A \cap A = A$, $A \cap \varnothing = \varnothing$, $A \cap \Omega = A$. Из определения противоположного события легко получаем, что $\overline{\overline{A}} = A$, $\overline{\Omega} = \varnothing$, $\overline{\varnothing} = \Omega$.

Все эти равенства доказываются одним и тем же способом. Докажем для примера равенство $A \cap B = \overline{\overline{A} \cup \overline{B}}$. При этом будет предложен метод, которым доказывается равенство двух произвольных случайных событий. Итак, пусть $\omega \in A \cap B$. Тогда последовательно имеем: $\omega \in A$ и $\omega \in B$. Отсюда следует, что $\omega \notin \overline{A}$ и $\omega \notin \overline{B}$. Значит, $\omega \notin \overline{A} \cup \overline{B}$. В итоге, описание $\omega \in \overline{\overline{A} \cup \overline{B}}$ и, следовательно, $A \cap B = \overline{\overline{A} \cup \overline{B}}$. Докажем обратное утверждение. Пусть $\omega \in \overline{\overline{A} \cup \overline{B}}$. Тогда $\omega \notin \overline{A} \cup \overline{B}$. Отсюда получаем, что $\omega \notin \overline{A}$ и $\omega \notin \overline{B}$. Теперь находим $\omega \in A$ и $\omega \in B$ и, значит, $\omega \in A \cap B$ или $\overline{\overline{A} \cup \overline{B}} \subset A \cap B$. И окончательно получаем, что $A \cap B = \overline{\overline{A} \cup \overline{B}}$. Точно так же можно показать справедливость ранее приведенного равенства $A \backslash B = A \cap \overline{B}$ и основных законов в случае, когда теоретико-множественные операции над событиями берутся в счетном числе. Например, имеем равенства

$$\left(\bigcup_{i=1}^{\infty} A_i \right) \cap B = \bigcup_{i=1}^{\infty} (A_i \cap B), \qquad \bigcup_{i=1}^{\infty} A_i = \overline{\bigcap_{i=1}^{\infty} \overline{A}_i}.$$

Легко видеть, что закон Де Моргана и соотношение $A \backslash B = A \cap \overline{B}$ позволяют получить все теоретико-множественные операции над случайными событиями с помощью только действий объединения и дополнения. Проиллюстрируем теперь введенные теоретико-множественные операции над случайными событиями на следующем простом эксперименте.

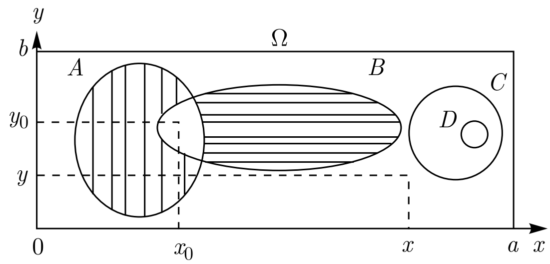


Рис. 1.8

Пример 1.38. На прямоугольный участок земли за некоторый фиксированный промежуток времени из космоса падает частица. В этом опыте будем интересоваться исходами, которые заключаются в попадании частицы в некоторые области прямоугольника, имеющие площадь (см. рис. 1.8).

Итак, в этом эксперименте достоверное событие, которое означает попадание частицы в прямоугольник, можно представить в виде множества $\Omega = \{\omega = (x, y)\colon 0 \leqslant x \leqslant a, 0 \leqslant y \leqslant b\}$. Например, элементарное событие $\{\omega = (x, y)\}$ означает прохождение частицы через точку с абсциссой x и ординатой y, а случайное элементарное событие $\{\omega_0 = (x_0, y_0)\}$ — прохождение частицы через точку с абсциссой x_0 и ординатой y_0 и т. д. Событие A — попадание частицы в эллипс, большая полуось которого расположена вертикально. Событие B — попадание частицы в эллипс, большая полуось которого расположена горизонтально. Событие C — попадание частицы в круг с большим диаметром. Событие D — попадание частицы в круг с маленьким диаметром.

Тогда событие $A \cup B$ соответствует попаданию частицы в область прямоугольника, которая получается наложением двух эллипсов без их перемещения. Событие $A \cap B$ соответствует попаданию частицы в светлую область любого из эллипсов. Событие $A \backslash B$ соответствует попаданию частицы в область, которая заштрихована вертикальными линиями. Случайное событие $B \backslash A$ соответствует попаданию частицы в область, которая заштрихована горизонтальными линиями. Событие $A \bigtriangleup B = B \bigtriangleup A$ соответствует попаданию частицы в область, которая заштрихована горизонтальными или вертикальными линиями. Случайное событие $\overline{C}$ соответствует попаданию частицы в прямоугольник и обязательно ее непопаданию в большой круг. Заметим, что при наблюдении элементарного события $\{\omega_0 = (x_0, y_0)\}$, например, произойдут события A, B, $A \cup B$, $A \cap B$, $\overline{C}$, $\overline{D}$ и не произойдут события $A \backslash B$, $B \backslash A$, $A \bigtriangleup B$. Наконец, события A и B являются совместными ($A \cap B \neq \varnothing$), события A и C — несовместными ($A \cap C = \varnothing$) и случайное событие $D \subset C$.

§ 3. Наблюдаемые события реального эксперимента

3.1. Алгебра и σ-алгебра событий. Некоторые события из множества $\mathfrak{J}$, которые когда-нибудь наступают при проведении эксперимента E, по различным причинам могут быть не доступны или не наблюдаемы для экспериментатора. Поэтому целесообразно выделить класс всех наблюдаемых исходов (событий) статистически устойчивого эксперимента.

Определение 1.14. Любое непустое подмножество $\mathcal{F}$ множества $\mathfrak{J}$ всех случайных допустимых событий, которые связаны с некоторым статистически устойчивым экспериментом E, будем называть *σ-алгеброй*, если справедливы следующие ограничения:

- введены теоретико-множественные операции (объединения $\cup$, пересечения $\cap$ и разность $\backslash$) над элементами из $\mathcal{F}$;
- из $A \in \mathcal{F}$ следует $\overline{A} \in \mathcal{F}$;
- из $A_1 \in \mathcal{F}$, $A_2 \in \mathcal{F}, \ldots$ следует $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$.

Если последнее ограничение имеет место для конечного числа случайных событий, то $\mathcal{F}$ называется *алгеброй*.

Так как множество $\mathcal{F}$ не является пустым, то существует $A \in \mathcal{F}$. Следовательно, $\overline{A} \in \mathcal{F}$. Теперь ясно, что $A \cup \overline{A} = \Omega \in \mathcal{F}$ и $\varnothing = \Omega \backslash \Omega = \overline{\Omega} \in \mathcal{F}$. Случайные события из $\mathcal{F}$ называются *наблюдаемыми*. Поэтому $\mathcal{F}$ будем называть множеством наблюдаемых исходов, хотя оно может содержать исходы, которые никогда не происходят при проведении эксперимента E. Более того, оно всегда содержит невозможный исход $\varnothing$. Теперь стало окончательно ясно, насколько важным было введение с самого начала в рассмотрение невозможного исхода $\varnothing$.

Легко показать, что σ-алгебра $\mathcal{F}$ замкнута относительно остальных теоретико-множественных операций. Прежде всего покажем, что $A \cap B \in \mathcal{F}$, $A \backslash B \in \mathcal{F}$, $A \vartriangle B \in \mathcal{F}$ при $A \in \mathcal{F}$, $B \in \mathcal{F}$. Действительно, все эти соотношения сразу следуют из представлений вида $A \cap B = \overline{\overline{A} \cup \overline{B}}$, $A \backslash B = A \cap \overline{B}$, $A \vartriangle B = (A \backslash B) \cup (B \backslash A)$ и последних двух ограничений на $\mathcal{F}$. Аналогичным способом доказывается, что σ-алгебра $\mathcal{F}$ замкнута относительно теоретико-множественных операций, взятых в счетном числе над случайными событиями.

Переход к множеству $\mathcal{F} \subset \mathfrak{J}$ вызван тем, что не над всеми допустимыми исходами могут быть введены теоретико-множественные операции и, более того, не все допустимые исходы эксперимента E могут происходить. Далее, могут существовать события из $\mathfrak{J}$, которые когда-нибудь наступают при проведении эксперимента E и не доступны наблюдению для исследователя. Поэтому очень часто такого рода события не включают в $\mathcal{F}$. Наконец, так как множество $\mathcal{F}$ всех наблюдаемых исходов статистически устойчивого эксперимента E является σ-алгеброй, то в результате теоретико-множественных операций над наблюдаемыми исходами мы получаем снова наблюдаемые исходы. Другими словами, мы будем иметь дело только с известными нам наблюдаемыми исходами (объектами).

Таким образом, каждому статистически устойчивому эксперименту E мы можем поставить в соответствие некоторую упорядоченную пару $(\Omega, \mathcal{F})$, которая называется его теоретико-множественной моделью. Приведенная ниже табл. 1.1 позволяет зафиксировать основные понятия и элементы теоретико-множественной модели эксперимента E.

Таблица 1.1.

Множество $\mathfrak{J}$ допустимых событий эксперимента E	Классификация случайных событий
Элементарное событие $\{\omega\}$	
Случайное событие A	
Достоверное событие Ω	
Невозможное событие $\varnothing$	
Из события A следует событие B: $A \subset B$	Соотношения между случайными событиями
Событие A равно событию B: $A = B \Leftrightarrow A \subset B,\ B \subset A$	
Совместность и несовместность событий A и B	
Противоположное к A событие $\overline{A}$	
Объединение событий A и B : $A \bigcup B$	Теоретико-множественные операции над случайными событиями
Пересечение событий A и B: $A \bigcap B$	
Разность событий A и B: $A \backslash B = A \bigcap \overline{B}$	
Симметрическая разность событий A и B: $A \vartriangle B = (A \backslash B) \bigcup (B \backslash A)$	
A и $\overline{A} = \Omega \backslash A$ противоположны $\Leftrightarrow A \bigcup \overline{A} = \Omega$, $A \bigcap \overline{A} = \varnothing$	
σ-алгебра $\mathcal{F}$ наблюдаемых событий	
$(\Omega, \mathcal{F})$ — теоретико-множественная модель	

Такого рода теоретико-множественные модели позволяют сделать следующее:

1) указать все допустимые исходы эксперимента E;
2) указать все элементарные исходы эксперимента E и определить пространство Ω описаний каждого элементарного исхода в некотором выбранном языке;

3) представить каждый допустимый исход A эксперимента E в виде множества из описаний $\omega \in \Omega$ только тех элементарных исходов, один из которых обязательно происходит одновременно с этим исходом, и тем самым A называть случайным событием;
4) определять равные события, соотношения между событиями и находить им интерпретацию через известные исходы эксперимента E;
5) выполнять теоретико-множественные операции над случайными событиями из множества $\mathfrak{J}$, порождать новые события и выяснять, как одни исходы с помощью теоретико-множественных операций выражаются через другие;
6) выделить множество $\mathcal{F}$ наблюдаемых исходов эксперимента E.

В заключение этого раздела приведем простейшие примеры σ-алгебр событий произвольного статистически устойчивого эксперимента E.

- $\mathcal{F} = \{\Omega, \varnothing\}$ — так называемая тривиальная σ-алгебра, состоящая из достоверного исхода Ω и невозможного события $\varnothing$. В этом простейшем случае в результате проведения эксперимента E можно фиксировать или наблюдать только достоверное событие. С практической точки зрения система $\mathcal{F} = \{\Omega, \varnothing\}$ оказывается бесполезной (самой «бедной»), ибо нельзя рассматривать (наблюдать) другие исходы реального эксперимента E кроме достоверного результата Ω.
- $\mathcal{F} = \{A, \overline{A}, \Omega, \varnothing\}$ — σ-алгебра, порожденная только исходом A. Для статистически устойчивого эксперимента E множество $\mathcal{F} = \{A, \overline{A}, \Omega, \varnothing\}$ содержит исходы A, $\overline{A}$, достоверный исход Ω и невозможное событие $\varnothing$. По отношению к эксперименту E система событий $\mathcal{F} = \{A, \overline{A}, \Omega, \varnothing\}$ является более информативной, чем $\mathcal{F} = \{\Omega, \varnothing\}$. В качестве простого примера такого эксперимента можно рассмотреть работу персонального компьютера за некоторое фиксированное время T, если пользователю интересна или наблюдаема только следующая информация: произошел отказ компьютера (случайное событие A) или нет (случайное событие $\overline{A}$).
- $\mathcal{F} = \{A\colon A \subseteq \Omega\}$ — совокупность всех (включая и пустое множество) подмножеств достоверного исхода Ω — множества описаний (имен) всех выбранных элементарных исходов. Если множество $\mathfrak{J}'$ всех выбранных элементарных исходов счетно (конечно или бесконечно), то множество $\mathcal{F} = \{A\colon A \subseteq \Omega\}$ — самая информативная («богатая») σ-алгебра и вполне обозримая, именно поэтому ее и рассматривают в качестве множества $\mathfrak{J}$ всех допустимых исходов статистически устойчивого эксперимента E.

3.2. Методы построения теоретико-множественных моделей случайных экспериментов. В простейшем варианте, когда выбран единственный элементарный исход $A'_1 = \Omega = \{\omega\}$, находим, что $\mathfrak{J} = \{A\colon A \subseteq \Omega\} = \{\Omega, \varnothing\} = \mathcal{F}$. При более сложном случае, когда выбраны два элементарных исхода $A = A'_1 = \{\omega_1\}$ и $\overline{A} = A'_2 = \{\omega_2\}$, получим, что $\mathfrak{J}' = \{A, \overline{A}\}$, $\Omega = \{\omega_1, \omega_2\}$ и $\mathfrak{J} = \{B\colon B \subseteq \Omega\} = \{A, \overline{A}, \Omega, \varnothing\} = \mathcal{F}$. В связи с этим заметим, что незначительное увеличение числа выбранных элементарных исходов или, что то же самое, числа элементов множества Ω связано с существенным увеличением числа всех возможных исходов статистически устойчивого эксперимента E. Например, если достоверное событие Ω состоит из n точек (описаний) $\omega_1, \omega_2, \ldots, \omega_n$, то $\mathfrak{J} = \{A\colon A \subseteq \Omega\} = \{A_1, A_2, \ldots, A_q\}$ уже включает $q = 2^n$ всех допустимых исходов эксперимента E.

Доказательство этого факта проведем методом индукции по n. Если $\Omega = \{\omega_1\}$, то $\mathfrak{J} = \{\{\omega_1\}, \varnothing\}$, $q = 2 = 2^1$ и, значит, равенство $q = 2^n$ установлено при $n = 1$. Предположим, что это равенство справедливо для некоторого числа n, и покажем, что оно выполняется для $n + 1$. Пусть теперь множество $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{n+1}\}$. Тогда имеем $\mathfrak{J} = \{A\colon A \subseteq \Omega\} = \{A\colon A \subseteq \Omega, \omega_{n+1} \in A\} \cup \{A\colon A \subseteq \Omega \backslash \{\omega_{n+1}\}\}$, где $\{A\colon A \subseteq \Omega, \omega_{n+1} \in A\}$ — множество всех подмножеств множества Ω, которые содержат описание ω_{n+1}, и $\{A\colon A \subseteq \Omega \backslash \{\omega_{n+1}\}\}$ — множество всех подмножеств множества Ω, которые уже не содержат элемент ω_{n+1}. Так как множество $\{A\colon A \subseteq \Omega, \omega_{n+1} \in A\} = \{A \cup \{\omega_{n+1}\}\colon A \subseteq \{\omega_1, \omega_2, \ldots, \omega_n\}\}$, то это множество состоит по предположению из 2^n элементов. Далее, ясно, что множество $\{A\colon A \subseteq \Omega \backslash \{\omega_{n+1}\}\} = \{A\colon A \subseteq \{\omega_1, \omega_2, \ldots, \omega_n\}\}$ также содержит 2^n элементов. Множества $\{A\colon A \subseteq \Omega, \omega_{n+1} \in A\}$, $\{A\colon A \subseteq \Omega \backslash \{\omega_{n+1}\}\}$ не содержат одинаковых элементов. Поэтому множество $\mathfrak{J} = \{A\colon A \subseteq \Omega\}$ включает $2^n + 2^n = 2^{n+1}$ элементов. Этим заканчивается доказательство рассматриваемого здесь утверждения.

Итак, появляется возможность представить большое число 2^n всех исходов статистически устойчивого эксперимента E через относительно малое число n всех выбранных элементарных исходов. Например, при $n = 23$ непосредственно находим $q = 2^n = 8\,388\,608$. Однако в случае несчетного множества Ω объект $\mathfrak{J} = \{A\colon A \subseteq \Omega\}$ оказывается слишком широким и необозримым. Поэтому в случае несчетного множества Ω в качестве множества всех наблюдаемых исходов статистически устойчивого эксперимента E, как правило, выбирают σ-алгебру $\mathcal{F}$ существенно более «узкую», чем σ-алгебра $\{A\colon A \subseteq \Omega\}$. Приведем пример такой ситуации.

Пример 1.39. Пусть опыт E заключается в непреднамеренном выборе точки на отрезке $[0; 1]$ действительной оси. Результат такого эксперимента может состоять в том, что точка взята из некоторого подмножества отрезка $[0;1]$. Простыми и хорошо известными еще из школьной математики являются тестовые или стандартные подмножества отрезка $[0;1]$ типа $[a,b] = \{\omega = x\colon a \leqslant x \leqslant b\}$ — отрезок, $[a,b) = \{\omega = x\colon a \leqslant x < b\}$ — замкнутый слева и открытый справа промежуток, $(a,b] = \{\omega = x\colon a < x \leqslant b\}$ — открытый слева и замкнутый справа промежуток, $(a,b) = \{\omega = x\colon a < x < b\}$ — интервал, $[a,a] = \{x\colon a \leqslant x \leqslant a\} = \{a\}$ — одноточечное подмножество.

В рассматриваемом эксперименте символ $\omega = x$ есть абсцисса (описание, имя) выбранной точки (элементарного исхода) вида $\{x\}$, а достоверный исход $\Omega = [0;1] = \{\omega = x\colon 0 \leqslant x \leqslant 1\}$ — несчетное множество. В большинстве случаев совокупность $\mathcal{F}$ всех наблюдаемых исходов этого эксперимента совпадает с самой «узкой» σ-алгеброй, содержащей все указанные выше тестовые подмножества отрезка $[0;1]$. Как известно, такая σ-алгебра называется борелевской, является достаточно широким классом исходов эксперимента E с точки зрения практических целей. О целесообразности использования борелевской σ-алгебры и ее точном определении будет еще сказано в следующих главах. Наряду с борелевской σ-алгеброй в качестве множества $\mathcal{F}$ всех наблюдаемых исходов эксперимента с выбором точки рассматривают также более «широкую» σ-алгебру подмножеств отрезка $[0;1]$, для которых определено понятие длины. Последнюю σ-алгебру называют системой лебеговских подмножеств отрезка $[0;1]$.

Этот несколько иллюстративный пример показывает, что одному и тому же множеству $\mathfrak{I}'$ всех элементарных исходов некоторого статистически устойчивого эксперимента E могут отвечать разные множества его наблюдаемых исходов. Поэтому задание статистически устойчивого эксперимента E определяется, в конечном счете, множеством $\mathfrak{I}$ всех его допустимых исходов. Это необходимо всегда принимать во внимание при построении адекватной вероятностной модели эксперимента E.

Итак, множество $\mathfrak{I}$ всех допустимых исходов и множество $\mathfrak{I}'$ элементарных исходов в общем случае следует выбирать относительно «большими», а множество $\mathcal{F}$ наблюдаемых исходов не обязательно должно быть наиболее «широким». Множество $\mathcal{F}$ существенно зависит от различных целей изучения эксперимента E и оно, как правило, содержит только те допустимые события, которые исследователь может наблюдать. Значит, если множества $\mathfrak{I}$ и Ω для эксперимента E определены заранее некоторым образом, то допускается рассмотрение различных σ-алгебр $\mathcal{F}$ наблюдаемых исходов. Например, конкретная σ-алгебра $\mathcal{F}$ может не содержать некоторые или все элементарные исходы. В простейших случаях, когда совокупность $\mathfrak{I}'$ всех элементарных исходов эксперимента E является счетным, например, конечным, множество Ω однозначно определяет наиболее широкое множество $\mathfrak{I} = \{A\colon A \subseteq \Omega\}$ всех допустимых исходов, и $\mathcal{F} \subset \mathfrak{I}$. В общем случае если $\mathfrak{I}$ есть σ-алгебра и является обозримым математическим объектом, то упорядоченную пару вида $(\Omega, \mathfrak{I})$ часто выбирают в качестве теоретико-множественной модели опыта E.

3.3. Теоретико-множественная модель эволюционных экспериментов. В первом параграфе на содержательном уровне было введено понятие эволюционного эксперимента $E = \{E_t; t \in T\}$. Рассмотрим теперь ради простоты эволюционный эксперимент E, для которого $T = \{1, 2, \ldots, m\}$. Таким образом, чтобы провести эксперимент E нужно провести последовательно эксперименты $E_1, E_2, \ldots, E_m$. При этом отмечалось, что исходы эксперимента $E = \{E_t; t \in T\}$ некоторым образом определяются исходами экспериментов $E_1, E_2, \ldots, E_m$. Другими словами, поведение всех m экспериментов $E_1, E_2, \ldots, E_m$ должно быть некоторым образом описано как единый эксперимент E. Рассмотрим один из наиболее часто встречающихся способов такого описания.

Пусть при каждом фиксированном $t = 1, 2, \ldots, m$ эксперимент E_t с комплексом условий Σ_t имеет теоретико-множественную модель наблюдаемых исходов следующего вида: $(\Omega_t, \mathcal{F}_t)$. При этом символ $\omega_t \in \Omega_t$ дает полное описание в некотором языке произвольного элементарного результата эксперимента E_t с допустимыми исходами из множества $\mathfrak{I}_t$. Такое описание позволяет произвольный элементарный исход эксперимента E_t представить в виде одноточечного множества $\{\omega_t\}$, а любой другой его результат $A_t \in \mathfrak{I}_t$ — в виде некоторого подмножества множества Ω_t, т. е. $A_t \subset \Omega_t$. Если мы последовательно наблюдали бы элементарные исходы $\{\omega_1\}$, $\{\omega_2\}, \ldots, \{\omega_m\}$ экспериментов $E_1, E_2, \ldots, E_m$, то естественно считать, что для единого эксперимента E осуществился элементарный исход $\{(\omega_1, \omega_2, \ldots, \omega_m)\}$ с описанием $\omega = (\omega_1, \omega_2, \ldots, \omega_m)$. В этом случае достоверное событие Ω единого эксперимента E представляет собой декартово произведение:

$$\Omega_1 \times \Omega_2 \times \ldots \times \Omega_m = \\ = \{\omega = (\omega_1, \omega_2, \ldots, \omega_m)\colon \omega_1 \in \Omega_1, \omega_2 \in \Omega_2, \ldots, \omega_m \in \Omega_m\}$$

достоверных исходов Ω_1, $\Omega_2, \ldots, \Omega_m$ экспериментов E_1, $E_2, \ldots, E_m$. Большой интерес с точки зрения практики представляют исходы эксперимента E вида

$$A_1 \times A_2 \times \ldots \times A_m = \\ = \{\omega = (\omega_1, \omega_2, \ldots, \omega_m)\colon \omega_1 \in A_1, \omega_2 \in A_2, \ldots, \omega_m \in A_m\},$$

где $A_1 \in \mathcal{F}_1$, $A_2 \in \mathcal{F}_1$,..., $A_m \in \mathcal{F}_m$.

Исход эксперимента E вида $A_1 \times A_2 \times \ldots \times A_m$ называют прямоугольником и такой исход заключается в том, что осуществились одновременно исходы A_1, $A_2, \ldots, A_m$ экспериментов E_1, $E_2, \ldots, E_m$. В дальнейшем произвольный исход эволюционного эксперимента E целесообразно обозначать полужирным латинским символом и курсивом, например, исход $\boldsymbol{A} = A_1 \times A_2 \times \ldots \times A_m$. К сожалению, все прямоугольники не образуют σ-алгебру подмножеств пространства $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_m$. Однако множество K всех прямоугольников $\boldsymbol{A} = A_1 \times A_2 \times \ldots \times A_m$ позволяет простым способом построить некоторую σ-алгебру подмножеств пространства $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_m$.

Определение 1.15. *Наименьшей σ-алгеброй* $\sigma(\mathsf{K})$, которая порождается множеством K всех прямоугольников вида $A_1 \times A_2 \times \ldots \times A_m$, называется такая σ-алгебра, для которой имеет место:

1) $\mathsf{K} \subset \sigma(\mathsf{K})$;
2) при любой σ-алгебре $\mathfrak{R} \supset \mathsf{K}$ выполняется соотношение $\sigma(\mathsf{K}) \subset \mathfrak{R}$.

Покажем, что наименьшая σ-алгебра $\sigma(\mathsf{K})$, содержащая множество K всех прямоугольников вида $A_1 \times A_2 \times \ldots \times A_m$, существует. В теории меры пространство $(\Omega, \sigma(\mathsf{K})) = (\Omega, \mathcal{F})$ называется прямым произведением пространств $(\Omega_t, \mathcal{F}_t)$, $t = 1, 2, \ldots, m$, и оно в нашем случае является теоретико-множественной моделью наблюдаемых исходов простейшего эволюционного эксперимента $E = \{E_t;\ t = 1, 2, \ldots, m\}$.

Теорема 1.1. *Наименьшая σ-алгебра $\sigma(\mathsf{K})$, содержащая множество K всех прямоугольников вида $A_1 \times A_2 \times \ldots \times A_m$, существует.*

Доказательство. Пусть $\Xi = \{\boldsymbol{B} \colon \boldsymbol{B} \subset \Omega\}$ есть множество всех подмножеств из множества $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_m$. Ясно, что $\Xi \supset \mathsf{K}$. Над элементами непустого множества Ξ, которые являются подмножествами Ω, вводим теоретико-множественные операции объединения ($\cup$), пересечения ($\cap$), разности ($\backslash$), симметрической разности ($\triangle$) и т. д. Если множества $\boldsymbol{B} \in \Xi$ и $\boldsymbol{B}_n \in \Xi$ для $n = 1, 2, \ldots$ или, что то же самое, $\boldsymbol{B} \subset \Omega$ и $\boldsymbol{B}_n \subset \Omega$ для $n \geqslant 1$, то $\overline{\boldsymbol{B}} = \Omega \backslash \boldsymbol{B} \subset \Omega$ и $\cup_{n=1}^{\infty} \boldsymbol{B}_n \subset \Omega$. Значит, $\overline{\boldsymbol{B}} \in \Xi$, $\cup_{n=1}^{\infty} \boldsymbol{B}_n \in \Xi$ и множество Ξ является σ-алгеброй.

Итак, существует σ-алгебра Ξ, содержащая множество K всех прямоугольников $A_1 \times A_2 \times \ldots \times A_m$. Каждую такую σ-алгебру будем обозначать $\mathfrak{R}_\gamma$, помечая ее индексом γ из некоторого непустого множества Γ. Например, при некотором индексе $\alpha \in \Gamma$ имеем $\Xi = \mathfrak{R}_\alpha$. Ясно, что $\mathfrak{R}_\gamma \supset \mathsf{K}$ и $\mathfrak{R}_\gamma \subset \Xi = \mathfrak{R}_\alpha$ для всех элементов γ из множества Γ. Далее покажем, что $\sigma(\mathsf{K}) = \bigcap\limits_{\gamma \in \Gamma} \mathfrak{R}_\gamma$. Так как при $\gamma \in \Gamma$ имеем $\mathfrak{R}_\gamma \subset \Xi$, то над элементами множества $\bigcap\limits_{\gamma \in \Gamma} \mathfrak{R}_\gamma$ введены все перечисленные теоретико-множественные операции. Пусть $\boldsymbol{C} \in \bigcap\limits_{\gamma \in \Gamma} \mathfrak{R}_\gamma$. Отсюда получаем, что $\boldsymbol{C} \in \mathfrak{R}_\gamma$ при $\gamma \in \Gamma$ в силу определения операции пересечения и, следовательно, $\overline{\boldsymbol{C}} \in \mathfrak{R}_\gamma$ при каждом $\gamma \in \Gamma$, так как всякое множество $\mathfrak{R}_\gamma$ является σ-алгеброй.

Следовательно, при $\boldsymbol{C} \in \cap_{\gamma\in\Gamma}\mathfrak{R}_\gamma$ имеем $\overline{\boldsymbol{C}} \in \cap_{\gamma\in\Gamma}\mathfrak{R}_\gamma$. Если теперь $\boldsymbol{C}_n \in \cap_{\gamma\in\Gamma}\mathfrak{R}_\gamma$ при $n \geqslant 1$, то при $\gamma \in \Gamma$ имеем $\boldsymbol{C}_n \in \mathfrak{R}_\gamma$ для $n \geqslant 1$. Тогда при $\gamma \in \Gamma$ имеем $\cup_{n=1}^{\infty}\boldsymbol{C}_n \in \mathfrak{R}_\gamma$, поскольку каждое из множеств $\mathfrak{R}_\gamma$ является σ-алгеброй. Значит, из определения операции пересечения находим, что $\cup_{n=1}^{\infty}\boldsymbol{C}_n \in \cap_{\gamma\in\Gamma}\mathfrak{R}_\gamma$ и, следовательно, множество $\cap_{\gamma\in\Gamma}\mathfrak{R}_\gamma$ будет σ-алгеброй.

Наконец, если некоторая σ-алгебра $\mathcal{R} \supset \mathsf{K}$, то существует такой элемент $\beta \in \Gamma$, что $\mathcal{R} = \mathfrak{R}_\beta$. Над σ-алгеброй $\cap_{\gamma\in\Gamma}\mathfrak{R}_\gamma$ проделаем следующие простые преобразования: $\cap_{\gamma\in\Gamma}\mathfrak{R}_\gamma = \mathfrak{R}_\beta \cap \big(\cap_{\gamma\in\Gamma}\mathfrak{R}_\gamma\big) \subset \mathfrak{R}_\beta$. Учитывая все предыдущее, доказали, что $\sigma(\mathsf{K}) = \cap_{\gamma\in\Gamma}\mathfrak{R}_\gamma$ есть наименьшая σ-алгебра, содержащая K. Эту σ-алгебру также называют произведением σ-алгебр $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_m$ и обозначают через $\mathcal{F}_1 \times \mathcal{F}_2 \times \ldots \times \mathcal{F}_m$ или через $\bigsqcup_{t=1}^{m} \mathcal{F}_t$.

Итак, в настоящем разделе предложен способ построения теоретико-множественной модели

$$(\Omega_1 \times \Omega_2 \times \ldots \times \Omega_m,\ \mathcal{F}_1 \times \mathcal{F}_2 \times \ldots \times \mathcal{F}_m)$$

статистически устойчивого эволюционного эксперимента $E = \{E_1, E_2, \ldots, E_m\}$ простейшего вида. Аналогичным способом поступают в случае, когда T произвольное упорядоченное множество, например, T есть некоторое несчетное подмножество числовой прямой. Если при каждом фиксированном $t \in T$ эксперимент E_t с комплексом условий Σ_t имеет теоретико-множественную модель наблюдаемых исходов вида $(\Omega_t, \mathcal{F}_t)$, то для эволюционного эксперимента $E = \{E_t; t \in T\}$ построенную таким способом его теоретико-множественную модель обозначают через $(\bigsqcup_{t\in T}\Omega_t, \bigsqcup_{t\in T}\mathcal{F}_t)$ и называют прямым произведением теоретико-множественных моделей или пространств $(\Omega_t, \mathcal{F}_t)$, $t \in T$. Здесь достоверный исход $\bigsqcup_{t\in T}\Omega_t$ эволюционного эксперимента E представляет собой множество всех функций $\omega = \omega(t)\colon T \to \Omega_t$, каждая из которых при фиксированном $t \in T$ принимает значение $\omega_t \in \Omega_t$. Множество $\bigsqcup_{t\in T}\mathcal{F}_t$ наблюдаемых исходов эволюционного эксперимента E совпадает с наименьшей σ-алгеброй, которая содержит все так называемые цилиндрические множества

$$\big\{\omega = \omega(t)\colon \omega(t_1) = \omega_{t_1} \in A_{t_1}, \omega(t_2) = \omega_{t_2} \in A_{t_2}, \ldots, \omega(t_n) = \omega_{t_n} \in A_{t_n}\big\},$$

где $t_1 \in T$, $t_2 \in T, \ldots, t_n \in T$, $A_{t_1} \in \mathcal{F}_{t_1}$, $A_{t_2} \in \mathcal{F}_{t_2}, \ldots, A_{t_n} \in \mathcal{F}_{t_n}$ и $n = 1, 2, \ldots$.

Рассмотрим теперь частный случай такого рода построений. Пусть $T = \{1, 2, \ldots\}$ и при любом фиксированном $t = 1, 2, \ldots$ эксперимент E_t совпадает с некоторым заданным экспериментом E_0. Если эксперимент E_0 имеет теоретико-множественную модель $(\Omega_0, \mathcal{F}_0)$, то для эволюционного эксперимента $E = \{E_1 = E_0, E_2 = E_0, \ldots\}$ теоретико-множественную модель естественно обозначить через $(\Omega_0^\infty, \mathcal{F}_0^\infty)$. Заметим, что достоверный исход Ω_0^∞ эволюционного эксперимента E совпадает с пространством упорядоченных последовательностей вида $\omega = (\omega_1, \omega_2, \ldots)$, где $\omega_t \in \Omega_0$ для всех $t = 1, 2, \ldots$. Множество $\mathcal{F}_0^\infty$ наблюдаемых исходов такого эволюционного эксперимента E равно наименьшей σ-алгебре, которая содержит все цилиндрические множества вида $\{\omega = (\omega_1, \omega_2, \ldots)\colon \omega_1 \in A_1, \omega_2 \in A_2, \ldots, \omega_n \in A_n\}$, где $A_1 \in \mathcal{F}_0$, $A_2 \in \mathcal{F}_0, \ldots, A_n \in \mathcal{F}_0$ и $n = 1, 2, \ldots$.

Краткий обзор

В главе 1 были введены основные интуитивные понятия теории вероятностей такие как эксперимент, условия проведения эксперимента, допустимый исход эксперимента, статистически устойчивый эксперимент, элементарный исход. Определен предмет теории вероятностей. Приводятся аксиомы выбора элементарных исходов. Дано определение случайного события и проведена классификация случайных событий. Установлены логические и теоретико-множественные связи между исходами статистически устойчивого эксперимента. Формализовано понятие семейства наблюдаемых исходов статистически устойчивого эксперимента. Приводятся способы построения теоретико-множественных моделей как для статических, так и для эволюционных случайных экспериментов. Такого рода модели позволяют установить теоретико-множественные и логические отношения между различными исходами статистически устойчивого эксперимента.

Контрольные вопросы и упражнения

1. Дать простейшую классификацию реальных экспериментов.
2. Привести примеры статистически устойчивых экспериментов, на которых проиллюстрировать различные способы выбора элементарных исходов.
3. Доказать основные законы, которым удовлетворяют теоретико-множественные операции над случайными событиями.
4. Используя законы для операций над событиями, доказать справедливость следующего равенства: $A \backslash (A \backslash B) = A \cap B$.
5. Привести примеры построения простейших σ-алгебр.
6. Опыт состоит в бросании трех монет. Пусть монеты занумерованы и события C_1, C_2 и C_3 означают выпадение герба на первой, второй и третьей монетах соответственно. Пусть событие A означает выпадение одного герба и двух цифр, а событие B

есть выпадение не более одного герба. Выразить через C_1, C_2, C_3 события A и B.

7. Доказать, что события A, $\overline{A} \cap B$, $\overline{A \cup B}$ образуют полную группу попарно несовместных событий.
8. Показать, что $A \cap B \neq \varnothing$ тогда и только тогда, когда имеет место соотношение вида $(A \cup B) \cap (\overline{A} \cup B) \cap (A \cup \overline{B}) \neq \varnothing$. Интерпретировать это утверждение на примерах 1.35, 1.38.
9. Пусть A и B — случайные события, которые были рассмотрены в эксперименте примера 1.28. Найти и интерпретировать на этом эксперименте событие C такое, что $\overline{\overline{(C \cup A)} \cup \overline{C \cup \overline{A}}} = B$.
10. Пусть A и B — случайные события, которые были рассмотрены в эксперименте примера 1.29. Найти и интерпретировать на этом эксперименте событие C такое, что $\overline{\overline{(C \cup A)} \cup \overline{C \cup \overline{A}}} = B$.
11. Пусть эксперимент E заключается в непреднамеренном подбрасывании с помощью некоторого механизма на поверхность стола трех монет. На одной стороне каждой монеты изображен герб, а на другой — решетка. Предположим, что невозможно зафиксировать сторону первой монеты, если на второй монете выпадает решетка. Построить теоретико-множественную модель для этого эксперимента $(\Omega, \mathcal{F})$.
12. Пусть эксперимент E заключается в непреднамеренном подбрасывании с помощью некоторого механизма на поверхность стола трех монет. На одной стороне каждой монеты изображен герб, а на другой — решетка. Предположим, что невозможно зафиксировать сторону второй монеты, если на первой монете выпадает герб. Построить теоретико-множественную модель $(\Omega, \mathcal{F})$ для этого эксперимента.
13. Пусть эксперимент E заключается в непреднамеренном подбрасывании с помощью некоторого механизма на поверхность стола трех монет. На одной стороне каждой монеты изображен герб, а на другой — решетка. В этом эксперименте нет возможности фиксировать герб и решетку, если на первой и второй монетах выпали одинаковые изображения. Построить теоретико-множественную модель $(\Omega, \mathcal{F})$ для такого эксперимента.
14. Построить теоретико-множественные модели для экспериментов, описание которых на содержательном уровне приведено в примерах 1.8–1.13.
15. Подбрасывается монета до первого выпадения герба (орла) с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на поверхность стола. Назвать для этого эксперимента достоверное событие, некоторое элементарное событие и невозможное собы-

тие. Показать, что появление решетки не более трех раз является случайным событием.

16. В Нижнем Новгороде имеются три транспортных моста через реку Ока. Рассматривается ежедневная возможность переезда наземного транспорта через реку. Обозначим через A_1 событие, которое заключается в исправности первого моста. Пусть событие A_2 состоит в исправности второго моста и, наконец, событие A_3 — в исправности третьего моста. Представить событие A через события A_1, A_2 и A_3, где событие A означает отсутствие возможности переезда через реку.
17. Производятся три выстрела по мишени. Пусть A_1 — попадание в мишень при первом выстреле, A_2 — при втором выстреле, A_3 — при третьем выстреле, A — ровно одно попадание. Выразить событие A через события A_1, A_2 и A_3.
18. Каждое из четырех изделий может быть либо бракованным, либо годным. Введем события: A — хотя бы одно изделие бракованное, B — бракованных не менее двух изделий. Что означают следующие события: $\overline{A}$, $\overline{B}$, $A \cup B$, $A \cap B$, $A \cup \overline{B}$, $A \cap \overline{B}$, $A \backslash B$?
19. В эксперименте примера 1.31 были введены события A_1, A_2 и A_3. Что означают события: a) $A_1 \vartriangle \overline{A}_2$; b) $\overline{A}_1 \vartriangle A_2$; c) $\overline{A}_1 \cap \overline{A}_2$; d) $\overline{A}_2 \backslash A_1$; e) $\overline{A}_1 \backslash A_3$; f) $\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3$; g) $\overline{A}_3 \backslash \overline{A}_2$?
20. Построить σ-алгебру, порожденную только событиями A и B.
21. Имеется классическая колода в 36 карт. Построить теоретико-множественную модель для эксперимента, состоящего из последовательного выбора без возвращения двух карт. При этом учитывается порядок следования отобранных карт.
22. На трассе нефтепровода между городами Γ_1 и Γ_2 в течение некоторого времени произошла авария. Расстояние между городами равно 25 км. Измеряется расстояние от города Γ_1 до места аварии. Построить теоретико-множественную модель для этого эксперимента.

Глава 2

ВЕРОЯТНОСТНЫЕ МОДЕЛИ АПРИОРНЫХ ЭКСПЕРИМЕНТОВ

§ 1. Субъективный способ измерения шанса наступления случайных событий

1.1. Понятие вероятности на интуитивном уровне. В главе 1 статистически устойчивому эксперименту E с комплексом априори фиксированных условий $\Sigma = \{u_1, u_2, \ldots, u_s, \ldots\}$ поставили в соответствие теоретико-множественную модель $(\Omega, \mathcal{F})$. Эта модель изучает свойства эксперимента E с качественной точки зрения, так как она, прежде всего, позволяет:

— указать допустимые, элементарные и наблюдаемые исходы для E;
— определить пространство Ω описаний всех элементарных исходов;
— провести теоретико-множественные операции над исходами;
— выразить одни исходы эксперимента E через другие.

В повседневной деятельности человек издавна имеет дело с реальными всевозможными экспериментами. Прежде всего, люди проводят или наблюдают самые различные эксперименты. В этом случае первый и самый трудный вопрос состоит в том, как измерить некоторое наше представление о конкретном результате эксперимента и как такому представлению поставить в соответствие число. Например, человек с давних пор проходил путь из одного пункта в другой и при этом чувствовал усталость. У людей разного возраста, преодолевших одно и то же расстояние, возникает неодинаковое ощущение усталости. Более того, одно и то же расстояние человек часто проходит как поднимаясь в гору, так и спускаясь с горы, и поэтому может устать больше или меньше.

Первоначально расстояние для человека ассоциировалось с тем чувством усталости, которое он приобретет при его прохождении. Интуитивное понятие физической усталости обогащается с течением всей жизни у каждого человека и становится более определенным. Затем, чтобы сделать это ощущение усталости понятием более ясным, люди стали приближенно измерять расстояние в некоторых единицах, например, в шагах, локтях и т. д. Прошло большое время, пока человечество не научилось объективно измерять это чувство усталости от преодоления расстояния, и это расстояние стали измерять в метрах. И тогда появилась возможность обратного сравнения: зная расстояние между пунктами, каждый человек для себя определяет, насколько он устанет, пройдя его.

Аналогичная ситуация возникала, когда человек приподнимал некоторый груз, например, ведро с водой из колодца. Человек чувствовал усталость от поднятия тяжести. Это побудило человека количественно измерять способность тяжести причинять ему усталость в некоторых единицах, например, в пудах, в фунтах, в граммах и т. д. Теперь человек, зная меру тяжести в некоторых единицах, может оценить свою усталость от поднятия этой тяжести. В настоящее время люди с некоторой степенью точности научились измерять расстояние между населенными пунктами, вес тела, его массу, температуру и т. д. При этом основной и доступный метод приближенного измерения заключается в следующем. Выбирают некоторый так называемый эталонный результат эксперимента и значение его измерения. Далее, тем или иным способом сравнивают представление об эталоне с представлением о некотором другом результате и, тем самым, получают приближенное измерение последнего. Перейдем теперь к аналогичной и в то же время необычной ситуации.

Человек давно и многократно наблюдал различные явления (события природы): выпадение осадков, изменение температуры воздуха, пожары, землетрясения, смену дня и ночи, времен года, падения метеоритов и т. д. Он мог точно предсказать некоторые из перечисленных явлений, например, смену дня и ночи, времен года. Какие-то он мог предвидеть с той или иной степенью уверенности (выпадение осадков, изменение температуры воздуха) и, наконец, явления типа пожаров, землетрясений, падения метеоритов он бессилен предугадать. Последние два типа явлений получили название случайных. Однако при достаточно большом числе N наблюдений за одним из такого рода явлений у человека возникает чувство уверенности в его предсказаниях. Количественное измерение возможности наступления некоторого события A выражается числом $0 \leqslant \mathsf{P}_{\text{с}}(A) \leqslant 1$, которое будем называть *субъективной вероятностью* или *степенью доверия* субъекта этому случайному явлению. При этом величина $\mathsf{P}_{\text{с}}(A) \cdot N$ приближенно измеряет число наступлений события A при N наблюдениях за экспериментом E. На содержательном уровне число $\mathsf{P}_{\text{с}}(A)$ позволяет субъекту предсказать результат A эксперимента E до его проведения или наблюдения. Если при этом $\mathsf{P}_{\text{с}}(A)$ близко к единице, то субъект может сказать, что событие A, по-видимому, наступит при единичном проведении эксперимента E. Напротив, при $\mathsf{P}_{\text{с}}(A)$ близком к нулю событие A скорее всего не наступит при единичном проведении эксперимента E. И, наконец, если $\mathsf{P}_{\text{с}}(A)$ близко к $1/2$, то факт наступления результата A вызывает у субъекта слишком неопределенные и даже непонятные чувства.

К сожалению, у разных людей чувства о шансе наступления некоторого случайного явления неодинаковы и количественному измерению поддаются с большим трудом и грубым приближением. Более того, в повседневной жизни слово «вероятность» часто употребляется в качестве степени уверенности в отношении наступления тех или иных единичных фактов (событий). Эта степень уверенности в большой мере также связана с нашими субъективными и психологическими пожеланиями. Таковы, например, утверждения следующего содержания: «вероятно, что теорема Ферма справедлива», «вероятно, что каждое натуральное четное число, большее двух, можно представить в виде суммы двух простых чисел (например, $4 = 2 + 2$, $6 = 3 + 3$ и т. д.)», «маловероятно, что 1 мая 2020 г. в Нижнем Новгороде выпадет снег».

Некоторые последователи интуитивно-субъективной концепции определяют субъективную вероятность как количественную меру степени психологической уверенности суждений каждого субъекта и для сугубо единичных явлений, и для исходов статистически устойчивых экспериментов. Из этого следует, что теория вероятностей оказывается чем-то вроде раздела психологии, а ее выводы носят чисто субъективный характер. Между тем, все «микроскопические» свойства газов были выведены с использованием результатов классической теории вероятностей. Это обстоятельство и другие факты применения результатов теории вероятностей говорят о том, что субъективная вероятность должна определяться не для единичных событий, а для массовых событий, которые порождаются любым статистически устойчивым экспериментом E при его многократном проведении. Поэтому такой подход является универсальным, хотя число $\mathsf{P}_c(A)$ различными субъектами может определяться неоднозначно. Заметим, что также неоднозначно различными субъектами на основании только своей интуиции и опыта получаются результаты измерений расстояния между некоторыми пунктами, массы тела, температуры воздуха и т. п.

1.2. Отношение предпочтения между случайными событиями эксперимента и вычисление субъективной вероятности. Рассмотрим теперь метод назначения субъективной вероятности или измерения шанса (чувства) возможного наступления наблюдаемого исхода A эксперимента E.

Пусть для эксперимента E с комплексом условий Σ его проведения построена теоретико-множественная модель $(\Omega, \mathcal{F})$. Выберем некоторые события $A_1, A_2 \in \mathcal{F}$. Каждый субъект, хорошо представляя эксперимент E, сначала только на основе своей интуиции сравнивает события A_1 и A_2, иначе говоря, устанавливает между ними отношение вида $A_1 \prec A_2$. Для субъекта это означает, что событие A_2 не менее предпочтительно при наступлении, или не реже наступает, чем A_1. Можно предложить следующую игру субъекта (некоторого человека) с экспериментом E. Если субъект считает $A_1 \prec A_2$, то он получает ставку в a денежных единиц при наступлении только события A_2 и теряет a денежных единиц при наступлении только события A_1. Во всех остальных случаях считается ничья, т. е. он не получает и не теряет ставки. Если субъект желает получить достойный выигрыш, то он должен на интуитивном уровне правильно установить *отношение предпочтения* $\prec$ между исходами случайного эксперимента. Далее, как правило, субъект назначает приближенные числа $\mathsf{P}_c(A_1) \leqslant \mathsf{P}_c(A_2)$ для наблюдаемых исходов A_1 и A_2, если только $A_1 \prec A_2$. Каждый субъект по-разному может установить отношение предпочтения. Однако желательно, чтобы числа $\mathsf{P}_c(A_1)$ и $\mathsf{P}_c(A_2)$, определяемые приближенно разными субъектами, несильно отличались друг от друга. С этой целью от любого отношения предпочтения $\prec$ событий из $\mathcal{F}$ будем требовать выполнение следующих естественных ограничений (аксиом):

- любая пара A и B из $\mathcal{F}$ сравнима, т. е. либо $A \prec B$, либо $B \prec A$. Если для некоторых событий C и D из $\mathcal{F}$ оказывается одновременно $C \prec D$ и $D \prec C$, то считают эти события C и D эквивалентными с точки зрения шансов их наступления. В последнем случае факт эквивалентности обозначают символом $C \cong D$;
- пусть $A_1 \cap A_2 = B_1 \cap B_2 = \varnothing$ и $A_1 \succ B_1$, $A_2 \succ B_2$, тогда событие $A_1 \cup A_2 \succ B_1 \cup B_2$. Если при этом события A_1 и B_1 не являются эквивалентными и/или события A_2 и B_2 не являются эквивалентными, то события $A_1 \cup A_2$ и $B_1 \cup B_2$ также не должны быть эквивалентными;
- считаем, что $\varnothing \prec A$ для любого $A \in \mathcal{F}$ и что невозможное событие и достоверное событие не являются эквивалентными, т. е. эксперимент имеет хотя бы один наблюдаемый исход;
- пусть $A_1 \supset A_2 \supset \ldots$ и $A_n \succ B$ при $n = 1, 2, \ldots$. Тогда можно записать, что событие $A = \cap_{n=1}^{\infty} A_n \succ B$, где $A_n \in \mathcal{F}$, $n \geqslant 1$, $B \in \mathcal{F}$. Отсюда следует, что событие A, которое наступает одновременно с наступлениями сразу всех событий $A_1, A_2, \ldots$, не менее предпочтительно, чем B;
- существует отображение $\xi(\omega)\colon \Omega \to [0,1]$, что $\{\omega\colon \xi(\omega) < x\} \in \mathcal{F}$ при всех $x \in R = \{x\colon -\infty < x < +\infty\}$, и такое, что условие $\{\omega\colon \xi(\omega) \in J_1\} \prec \{\omega\colon \xi(\omega) \in J_2\}$ равносильно $|J_1| \leqslant |J_2|$ для любых промежутков J_1, $J_2 \subset [0,1]$ соответственно с длинами $|J_1|$, $|J_2|$.

В этом случае в теории вероятностей доказывается (*Климов Г. П.* Теория вероятностей и математическая статистика. — М.: МГУ, 1983), что отношение предпочтения $\prec$ в $\mathcal{F}$ единственным образом определяет $\mathsf{P}_c(A)$, при этом если $A_1 \prec A_2$, то $\mathsf{P}_c(A_1) \leqslant \mathsf{P}_c(A_2)$. Пятое ограничение — наиболее трудное для субъективного восприятия и не является доступным и естественным для большинства исследователей. Поэтому это ограничение нельзя рекомендовать к применению при назначении субъективной вероятности, и, как правило, от него отказываются. Это естественно приводит к неоднозначности в определении числа $\mathsf{P}_c(A)$ различными субъектами. Следовательно, мы не можем точно определить шанс наступления некоторого результата эксперимента, но каждому исходу A можно приписать субъективную вероятность $\mathsf{P}_c(A)$. Если имеем $A_1 \prec A_2$ и $\mathsf{P}_c(A_2) \geqslant \mathsf{P}_c(A_1)$, то это означает, что шанс наступления события A_2 больше, чем шанс наступления события A_1. Эталонное событие здесь совпадает с достоверным исходом Ω. Любой другой результат A эксперимента будем сравнивать с результатом Ω. Ради удобства принимают $\mathsf{P}_c(\Omega) = 1$ и естественно всегда полагают, что $\mathsf{P}_c(\varnothing) = 0$. Поэтому субъективная вероятность $0 \leqslant \mathsf{P}_c(A) \leqslant 1$ для всех $A \in \mathcal{F}$, так как достоверное событие Ω всегда наступает и имеет место соотношение $\varnothing \prec A \prec \Omega$.

Итак, основная проблема — назначить субъективную вероятность $\mathsf{P}_c(A)$ для любого $A \in \mathcal{F}$, используя предпочтения. На содержательном уровне число $1/\mathsf{P}_c(A)$, которое соответствует субъективной вероятности события A, показывает, во сколько раз событие A наступает реже, чем Ω. Следует сказать, что субъективное назначение вероятности для случайных событий является ее первым и наиболее грубым приближенным вычислением. Однако субъективный подход к измерению шансов (чувства) наступления исхода A имеет большое значение для понимания природы и сути статистически устойчивого эксперимента E. Другими словами, субъективный подход развивает у исследователей навыки и интуицию вероятностно-статистического мировоззрения на мир. Для иллюстрации рассмотрим примеры назначения субъективных вероятностей.

Пример 2.1. Из хорошо перемешанной колоды игральных карт, состоящей из дамы, короля и туза, наудачу (непреднамеренно, случайным образом) вынимается одна карта. Назначить субъективные вероятности исходов данного опыта.

Решение. Обозначим через ω_1 (ω_2, ω_3) описание элементарного исхода данного эксперимента, который заключается в том, что появится дама (король, туз). В этом случае достоверный исход

$$\Omega = \{\omega_1, \omega_2, \omega_3\},$$
$$\mathfrak{I} = \mathcal{F} = \{\Omega, \varnothing, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}\}$$

и, следовательно, построена теоретико-множественная модель $(\Omega, \mathcal{F})$ этого опыта. На интуитивном уровне легко определить между случайными событиями из $\mathcal{F}$ следующее отношение предпочтения:

$$\{\omega_1\} \cong \{\omega_2\} \cong \{\omega_3\}, \quad \{\omega_1, \omega_2\} \cong \{\omega_1, \omega_3\} \cong \{\omega_2, \omega_3\},$$
$$\{\omega_1\} \prec \{\omega_1, \omega_2\} \prec \{\omega_1, \omega_2, \omega_3\} = \Omega \quad \text{и т. д.}$$

Субъекту, исходя из интуитивного представления этого эксперимента, нетрудно представить, что карта с изображением дамы — событие $\{\omega_1\}$ — выпадает, скорее всего, в три раза реже, чем появляется любая из трех карт — достоверное событие Ω. Поэтому при $\mathsf{P}_c(\Omega) = 1$ можно принять, что субъективная вероятность $\mathsf{P}_c(\{\omega_1\}) \approx 1/3$. Следовательно, в силу эквивалентности событий вида $\{\omega_1\}$, $\{\omega_2\}$, $\{\omega_3\}$ имеет место соотношение $\mathsf{P}_c(\{\omega_1\}) = \mathsf{P}_c(\{\omega_2\}) = \mathsf{P}_c(\{\omega_3\}) \approx 1/3$. Далее ясно, что туз выпадает в два раза реже, чем дама или король — событие $\{\omega_1, \omega_2\}$. В силу этого и эквивалентности событий $\{\omega_1, \omega_2\}$, $\{\omega_1, \omega_3\}$, $\{\omega_2, \omega_3\}$ можно с большой уверенностью положить, что субъективная вероятность $\mathsf{P}_c(\{\omega_1, \omega_2\}) = \mathsf{P}_c(\{\omega_1, \omega_3\}) = \mathsf{P}_c(\{\omega_2, \omega_3\}) \approx 2/3$. Назначение субъективных вероятностей $\mathsf{P}_c(A)$, $A \in \mathcal{F}$, для этого достаточно простого примера определено.

Пример 2.2. Тетраэдр $dabc$, грани $\triangle dab$, $\triangle dac$, $\triangle dcb$, $\triangle abc$ которого окрашены соответственно в белый, синий, красный и зеленый цвет, подбрасывается случайным образом (непреднамеренно) на поверхность стола с помощью некоторого механизма (см. рис. 2.1). При этом предполагается, что тетраэдр изготовлен из однородного материала. Фиксируется цвет грани, на которую падает тетраэдр, соприкасаясь с поверхностью стола. Обозначим через ω_1 (ω_2, ω_3, ω_4) описание в некотором языке элементарного исхода данного эксперимента, который заключается в том, что тетраэдр падает на грань $\triangle dab$ ($\triangle dac$, $\triangle dcb$, $\triangle abc$). В этом случае

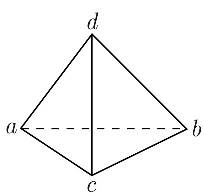


Рис. 2.1

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\},$$
$$\mathcal{F} = \{\Omega, \varnothing, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_1, \omega_4\}, \{\omega_2, \omega_3\},$$
$$\{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_4\}, \{\omega_1, \omega_3, \omega_4\}, \{\omega_2, \omega_3, \omega_4\}\}$$

и, значит, построена теоретико-множественная модель $(\Omega, \mathcal{F})$ этого эксперимента. Рассмотрим задачу о назначении субъективных вероятностей исходов данного опыта. На интуитивном уровне, почти с очевидностью, можно предложить между случайными событиями из $\mathcal{F}$ следующее отношение предпочтения:

$$\{\omega_1\} \cong \{\omega_2\} \cong \{\omega_3\} \cong \{\omega_4\},$$

$$\{\omega_1,\omega_2\} \cong \{\omega_1,\omega_3\} \cong \{\omega_1,\omega_4\} \cong \{\omega_2,\omega_3\} \cong \{\omega_2,\omega_4\} \cong \{\omega_3,\omega_4\},$$
$$\{\omega_1,\omega_2,\omega_3\} \cong \{\omega_1,\omega_2,\omega_4\} \cong \{\omega_1,\omega_3,\omega_4\} \cong \{\omega_2,\omega_3,\omega_4\},$$
$$\{\omega_1\} \prec \{\omega_1,\omega_2\} \prec \{\omega_1,\omega_2,\omega_3\} \prec \{\omega_1,\omega_2,\omega_3,\omega_4\} = \Omega \quad \text{и т. д.}$$

Исследователю, исходя из симметрии этого эксперимента, нетрудно допустить, что грань $\triangle dab$ с белым цветом — событие $\{\omega_1\}$ выпадает в четыре раза реже, чем появляется любая из четырех граней — достоверное событие Ω. Поэтому при $\mathsf{P}_c(\Omega) = 1$ можно принять $\mathsf{P}_c(\{\omega_1\}) = 1/4$. В силу эквивалентности событий $\{\omega_1\}$, $\{\omega_2\}$, $\{\omega_3\}$, $\{\omega_4\}$ имеем $\mathsf{P}_c(\{\omega_1\}) = \mathsf{P}_c(\{\omega_2\}) = \mathsf{P}_c(\{\omega_3\}) = \mathsf{P}_c(\{\omega_4\}) = 1/4$. На содержательном уровне очевидно, что белый цвет наступает в два раза реже, чем синий или красный — событие $\{\omega_2,\omega_3\}$. В силу этого и эквивалентности событий вида $\{\omega_1,\omega_2\}$, $\{\omega_1,\omega_3\}$, $\{\omega_1,\omega_4\}$, $\{\omega_2,\omega_3\}$, $\{\omega_2,\omega_4\}$, $\{\omega_3,\omega_4\}$ можно с большой уверенностью положить, что субъективная вероятность вида $\mathsf{P}_c(\{\omega_1,\omega_2\}) = \mathsf{P}_c(\{\omega_1,\omega_3\}) = \mathsf{P}_c(\{\omega_1,\omega_4\}) = \mathsf{P}_c(\{\omega_2,\omega_3\}) = \mathsf{P}_c(\{\omega_2,\omega_4\}) = \mathsf{P}_c(\{\omega_3,\omega_4\}) = 1/2$. Рассуждая аналогичным образом, получим, что $\mathsf{P}_c(\{\omega_1,\omega_2,\omega_3\}) = \mathsf{P}_c(\{\omega_1,\omega_2,\omega_4\}) = \mathsf{P}_c(\{\omega_1,\omega_3,\omega_4\}) = \mathsf{P}_c(\{\omega_2,\omega_3,\omega_4\}) = 3/4$. На этом заканчиваем рассмотрение назначения субъективной вероятности.

§ 2. Определение вероятности для опытов с конечным множеством равновозможных исходов

2.1. Классический подход. Субъективный подход, как уже отмечалось выше, дает различные числа $\mathsf{P}_c(A)$ при его использовании исследователями для измерения шанса наступления случайного события A. Поэтому естественно желание людей предложить другие методы и подходы, которые измеряли бы шанс наступления случайного события A в виде одного и того же числа $\mathsf{P}(A)$ и независимо от субъекта. Эти подходы в теории вероятностей выдвигают на первый план какую-либо одну сторону общей проблемы и, тем самым, позволяют предложить конкретные способы точного нахождения вероятностей. В силу этого такого рода способы сводят общие понятия в теории вероятностей к более простым и доступным для приложений понятиям и поэтому применяют не во всех случаях. Как правило, для этого рассматривают некоторые классы экспериментов, которые выделяются среди множества всех реальных статистически устойчивых экспериментов с помощью тех или иных ограничений. Первым таким множеством является так называемая совокупность классических экспериментов, для которых имеет место:

- конечное число элементарных исходов рассматриваемого эксперимента E, каждый из которых можно наблюдать;

- симметрия эксперимента E, которая позволяет на интуитивном уровне определенно утверждать о равных (равновозможных, равновероятных) шансах наступления каждого из элементарных исходов.

Примеры таких опытов приведены в первой главе под номерами 1.8–1.10, 1.12. Для такого рода экспериментов люди давно умели объективно измерять вероятности (шансы наступления) различных событий, производить вычисления с этими вероятностями, а также использовать результаты произведенных вычислений в практической деятельности и научных исследованиях. Это умение определять вероятности случайных событий, идущее от Ферма, Паскаля, Лапласа, Пуассона, Бернулли и Байеса, непосредственно связано с нашими интуитивными понятиями равновозможности, равновероятности и симметрии, а последние не подлежат формальному анализу.

Приведем примеры, иллюстрирующие понятие равновозможности и симметрии элементарных исходов эксперимента.

Пример 2.3. Пусть мы имеем опыт по извлечению белого или черного шара из урны, содержащей несколько одинаковых по размеру, весу и другим осязаемым признакам шаров, тщательно перемешанных перед извлечением. Если число белых шаров не равно числу черных, то можно присвоить каждому шару его единственный номер (имя, метка). Этим приемом добиваемся симметрии эксперимента. В данной ситуации говорят о равновозможности (равновероятности) появления любого имени шара или о том, что каждый шар имеет одинаковый шанс появления.

Пример 2.4. Появление герба при бросании симметричной монеты считается равновозможным с выпадением цифры.

Пример 2.5. Извлечение заранее указанной карты из хорошо перемешанной игральной колоды предполагается равновозможным по отношению к появлению любой другой карты.

Пусть теперь каждому классическому эксперименту E поставлена в соответствие теоретико-множественная модель $(\Omega, \mathcal{F})$. Тогда, согласно ограничениям для классического статистически устойчивого эксперимента E, будем иметь $\mathfrak{J}' = \{A'_1, A'_2, \ldots, A'_n\}$ и $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Здесь и далее используются обозначения п. 2.1 и п. 3.2 первой главы. Значит, при классическом построении основ теории вероятностей рассматривают множество $\mathfrak{J}'$, состоящее из конечного числа n попарно несовместимых равновозможных элементарных событий $\{\omega_k\}$, $k = 1, 2, \ldots, n$. Элементарные события образуют полную группу несовместимых событий. Составим теперь множество $\mathcal{F}$ событий, включающее невозможное событие $\varnothing$, все элементарные события и всевозможные события A, удовлетворяющие условию $A \subset \Omega$. В частности, достоверное событие Ω включается в $\mathcal{F}$.

Итак, множество $\mathcal{F} = \{A\colon A \subset \Omega\} = \{A_1, A_2, \ldots, A_q\}$, $q = 2^n$. Другими словами, множество $\mathcal{F}$ является множеством всех подмножеств множества Ω. Легко установить, что множество $\mathcal{F}$ образует σ-алгебру событий. Пусть при $m \leqslant n$ случайное событие $A = \{\omega_{a_1}, \omega_{a_2}, \ldots, \omega_{a_m}\}$. В этом случае будем говорить, что элементарные события $\{\omega_{a_1}\}$, $\{\omega_{a_2}\}$,…, $\{\omega_{a_m}\}$ благоприятствуют появлению события A.

Определение 2.1. Если $A = \{\omega_{a_1}, \omega_{a_2}, \ldots, \omega_{a_m}\}$, где $m \leqslant n$, и все элементарные события равновозможны, то *вероятностью события* A называется величина m/n, что обозначается $\mathsf{P}(A) = m/n$.

Другими словами, вероятностью события A называется отношение числа элементарных событий, благоприятствующих появлению данного события, к числу всех возможных элементарных событий. Отсюда следует, что $\mathsf{P}(A) = m/n = \mathsf{P}(\{\omega_{a_1}\}) + \mathsf{P}(\{\omega_{a_2}\}) + \ldots + \mathsf{P}(\{\omega_{a_m}\})$. Значит, вероятность события A определяется суммированием вероятностей тех элементарных исходов, которые могут наступать одновременно с исходом A.

Итак, классическое определение вероятностей основывается на так называемом принципе равных возможностей П. Лапласа, Я. Бернулли, Т. Байеса. Рассмотрим ряд примеров на классическое определение вероятностей.

Пример 2.6. В урне имеется шесть черных и три белых шара. Определить вероятность появления белого шара.

Решение. Так как урна содержит разное число белых и черных шаров, то данный эксперимент не является симметричным. Поэтому сначала пронумеруем все шары цифрами от единицы до девяти. Предположим, что все белые шары помечены цифрами от единицы до трех, а все черные шары помечены цифрами от четырех до девяти. После этого эксперимент стал симметричным, так как все шары перенумерованы, и каждый из них имеет только один номер. Пусть теперь ω_1, ω_2 и ω_3 суть описания таких элементарных исходов, когда появляются шары белого цвета с номерами 1, 2 и 3 соответственно. Аналогично, $\omega_4, \omega_5, \ldots, \omega_9$ суть описания таких элементарных исходов, когда появляются шары черного цвета с номерами 4, 5,…, 9 соответственно. Итак, для этой задачи событие A, которое заключается в появлении белого шара, равно $\{\omega_1, \omega_2, \omega_3\}$. Достоверное событие $\Omega = \{\omega_1, \omega_2, \ldots, \omega_9\}$. Тогда $n = 9$, $m = 3$ и $\mathsf{P}(A) = 3/9 = 1/3$.

Пример 2.7 (*задача Даламбера*). Пусть симметричная монета бросается два раза с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Найти вероятность того, что хотя бы раз появится герб (ожидаемое событие A).

Решение. По одной из легенд Даламбер и его последователи решали эту задачу и рассуждали следующим образом. Герб появится либо при первом бросании, и в этом случае второе бросание не нужно, либо только при втором, либо герб совсем не выпадает. Всех элементарных случаев три. Из них благоприятствуют ожидаемому событию A только два. Следовательно, искомая вероятность равна 2/3. Но практика показывала другой результат. А именно, из 100 проведенных такого рода опытов приблизительно в 75 случаях появлялось событие A. Только существенно в более позднее время стали подозревать, что такое решение является неверным.

Приведем правильное решение задачи Даламбера. Для этого построим теоретико-множественную модель $(\Omega, \mathcal{F})$ этого эксперимента. Пусть $\Sigma = \{u_1, u_2, u_3, u_4, u_5, \ldots\}$, где u_1 — выбрана симметричная монета определенного достоинства, u_2 — выбран механизм бросания монеты, u_3 — монета бросается два раза, u_4 — выбрана поверхность стола, u_5 — выбрана достаточная освещенность для определения выпавшей стороны монеты и т. д. Итак, перечислены основные условия u_1, u_2, u_3, u_4, u_5 проведения данного эксперимента.

Выберем теперь элементарные исходы и математический язык для их описания. Пусть $\omega_1 = (0,\ 0)$ означает описание такого элементарного исхода, когда выпадает герб при первом и втором бросках; $\omega_2 = (0,\ 1)$ означает описание такого элементарного исхода, когда выпадает герб при первом броске и решетка при втором; $\omega_3 = (1,\ 0)$ означает описание такого элементарного исхода, когда выпадает решетка при первом броске и герб при втором; $\omega_4 = (1,\ 1)$ означает описание такого элементарного исхода, когда выпадает решетка при первом и втором бросках. Тогда $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $\mathcal{F} = \{A_1, A_2, \ldots, A_{16}\}$, например, $A_i = \{\omega_i\}$ — элементарное случайное событие при $i = 1, 2, 3, 4$ и $A_{11} = \{\omega_1, \omega_2, \omega_3\}$ — случайное событие (ожидаемое по Даламберу событие), которое заключается в том, что хотя бы раз появится герб. В этом случае $\mathsf{P}(A_{11}) = 3/4$ согласно классическому определению вероятности случайного события.

Заметим, что при первом решении этой задачи были допущены следующие ошибки:

- нарушено в действительности не только условие проведения эксперимента, но и условие симметрии его проведения, так как, если герб появился при первом броске, то отказывались от проведения второго броска;
- выбранные по Даламберу исходы не равновозможны, т. е. имеют разные шансы наступления и, следовательно, классическое

определение вероятности применять нельзя. Действительно, если $A_5 = \{\omega_1, \omega_2\}$ — выпадение герба при первом бросании, $A_3 = \{\omega_3\}$ — выпадение герба только при втором бросании и $A_4 = \{\omega_4\}$ — герб совсем не появится, то $\mathsf{P}(A_5) = 1/2$ и $\mathsf{P}(A_3) = \mathsf{P}(A_4) = 1/4$.

Пример 2.8. Игральную кость подбрасывают один раз. Найти вероятность того, что выпавшее число очков не делится на три (событие A).

Решение. В этом простом эксперименте число n всех элементарных исходов равно шести. Число m элементарных исходов, для каждого из которых выпавшее число очков не делится на три, равно четырем. Поэтому вероятность $\mathsf{P}(A) = 4/6 = 2/3$.

2.2. Элементы комбинаторного анализа. Вычисление вероятности случайного события $A_i \in \mathcal{F}$ по классической схеме $\mathsf{P}(A_i) = m/n$ требует определенных навыков построения конечных множеств $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, $A_i = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_m}\}$, а также подсчета чисел n и m элементов этих множеств. В примерах 2.6, 2.8 и простой задаче Даламбера эти вопросы вполне успешно решаются методом полного перебора. Однако для большинства задач построение множеств $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, $A_i = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_m}\}$ и определение чисел n, m вызывают наибольшую трудность. Поэтому необходимо здесь рассмотреть элементарные задачи комбинаторики, которые играют важную роль при вычислении вероятностей различных событий так называемым классическим способом.

Пусть $G = \{a_1, a_2, \ldots, a_N\}$ — множество различных элементов произвольной природы. В комбинаторике рассматриваются комбинации (множества) из элементов множества G, составленные в соответствии с некоторыми правилами. Это могут быть упорядоченные и неупорядоченные множества, множества, содержащие все элементы из G или только часть этих элементов, множества с повторяющимися элементами или же множества, все элементы которых различны.

Комбинаторный анализ — раздел дискретной математики, который изучает способы построения некоторых новых множеств в соответствии с заданными правилами из элементов исходных множеств. Одна из задач комбинаторного анализа заключается в подсчете числа различных способов, которыми эти комбинации могут быть выполнены, и в определении числа элементов каждого из построенных множеств.

Рассмотрим простейшие правила построения множеств и приемы подсчета числа элементов этих множеств. К таким правилам, прежде всего, относятся:

- способ объединения множеств;
- способ прямого произведения множеств;
- способ перестановок;

- способ размещений;
- способ сочетаний.

Более сложные алгоритмы образования множеств и определения числа их элементов чаще всего основываются на указанных простейших правилах. Перейдем к изучению этих правил.

Способ объединения множеств. Пусть $A = \{a_1, a_2, \ldots, a_m\}$ и $B = \{b_1, b_2, \ldots, b_k\}$ — произвольные конечные множества, каждое из которых составлено из различных элементов. С помощью операций объединения и пересечения можно образовать новые множества: $A \cup B = \{c_1, c_2, \ldots, c_s\}$ и $A \cap B = \{d_1, d_2, \ldots, d_u\}$. В дальнейшем обозначим через символ $\aleph(A)$ число элементов произвольного множества A. Символ $\aleph(A)$ читается: «алеф от А». Тогда $\aleph(A) = m$, $\aleph(B) = k$, $\aleph(A \cup B) = s$ и $\aleph(A \cap B) = u$.

Число $\aleph(A) + \aleph(B)$ можно получить, перечисляя все элементы множества A, а потом — все элементы множества B. При этом общие элементы множеств A и B, число которых равно $\aleph(A \cap B)$, будут перечислены два раза. Отсюда $\aleph(A) + \aleph(B) = \aleph(A \cup B) + \aleph(A \cap B)$, и для числа элементов объединения двух множеств находим: $\aleph(A \cup B) = \aleph(A) + \aleph(B) - \aleph(A \cap B)$. Принимая во внимание равенства $A \cup B \cup C = A \cup (B \cup C)$, $\aleph(A) + \aleph(B) = \aleph(A \cup B) + \aleph(A \cap B)$, найдем формулу

$$\aleph(A \cup B \cup C) = \aleph(A) + \aleph(B) + \aleph(C) - \\ - \aleph(A \cap B) - \aleph(A \cap C) - \aleph(B \cap C) + \aleph(A \cap B \cap C) \quad (2.1)$$

для числа элементов объединения трех множеств. Далее, используя эти приемы и метод индукции, можем установить общую формулу для числа элементов объединения конечного числа множеств. В частности, если множества A_1, A_2,..., A_r попарно не пересекаются, то имеем формулу $\aleph(A_1 \cup A_2 \cup \ldots \cup A_r) = \aleph(A_1) + \aleph(A_2) + \ldots + \aleph(A_r)$.

Рассмотрим пример.

Пример 2.9. На столе стоят восемь кружек и несколько чашек. Любой из этих сосудов является либо кружкой, либо красным, либо пустым. Число всех красных кружек равно трем, из которых две пустые. Всего на столе пять красных сосудов, из которых три являются пустыми. Число пустых сосудов равно семи, из которых четыре кружки. Из этих условий можно определить число всех сосудов на данном столе. В самом деле, если через A, B и C обозначить множества кружек, красных сосудов и пустых сосудов, то $A \cap B$, $A \cap C$, $B \cap C$, $A \cap B \cap C$ являются множествами всех красных кружек, всех пустых кружек, всех красных и пустых сосудов, всех красных и пустых кружек. Из условий этой задачи последовательно получаем, что $\aleph(A) = 8$, число всех сосудов равно $\aleph(A \cup B \cup C)$, $\aleph(A \cap B) = 3$, $\aleph(A \cap C) = 3$,

$\aleph(A \cap B \cap C) = 2$, $\aleph(B) = 5$, $\aleph(B \cap C) = 3$, $\aleph(C) = 7$, $\aleph(A \cap C) = 4$. Теперь по формуле (2.1) легко находим:

$$\aleph(A \cup B \cup C) = \aleph(A) + \aleph(B) + \aleph(C) - \aleph(A \cap B) - \aleph(A \cap C) - \aleph(B \cap C) + \aleph(A \cap B \cap C) = 12.$$

Способ прямого произведения множеств. Пусть даны множества A_1, $A_2, \ldots, A_r$. Множество всех упорядоченных наборов (элементов) вида $(x_1, x_2, \ldots, x_r)$, где $x_1 \in A_1$, $x_2 \in A_2, \ldots, x_r \in A_r$, называется прямым произведением множеств A_1, $A_2, \ldots, A_r$ и обозначается через $A_1 \times A_2 \times \ldots \times A_r$. Итак, имеем $A_1 \times A_2 \times \ldots \times A_r = \{(x_1, x_2, \ldots, x_r)\colon x_1 \in A_1, x_2 \in A_2, \ldots, x_r \in A_r\}$. Покажем, что $\aleph(A_1 \times A_2) = \sum_{x_1 \in A_1} \aleph(\{x_1\} \times A_2) = \aleph(A_1) \times \aleph(A_2)$. Действительно, если x_1 — фиксированный элемент из A_1, то множество $\{x_1\} \times A_2 = \{(x_1, x_2)\colon x_2 \in A_2\}$ содержит $\aleph(A_2)$ элементов. Множества $\{x_1\} \times A_2$, $x_1 \in A_1$, попарно не имеют общих элементов и множество $A_1 \times A_2$ равно объединению всех этих множеств. На основании формулы для числа элементов объединения конечного числа попарно непересекающихся множеств получим, что $\aleph(A_1 \times A_2) = \sum_{x_1 \in A_1} \aleph(\{x_1\} \times A_2) = \aleph(A_1) \times \aleph(A_2)$. Используя метод индукции, обобщим теперь эту формулу для прямого произведения конечного числа множеств. Множества $A_1 \times A_2 \times \ldots \times A_r$ и $A_1 \times (A_2 \times \ldots \times A_r)$ эквивалентны, так как между их элементами можно установить взаимно однозначное соответствие. Например, элемент $(x_1, x_2, \ldots, x_r)$ соответствует элементу $(x_1, (x_2, \ldots, x_r))$. Поэтому

$$\aleph(A_1 \times A_2 \times \ldots \times A_r) = \aleph(A_1 \times (A_2 \times \ldots \times A_r)) = \aleph(A_1) \times \aleph(A_2 \times \ldots \times A_r) = \aleph(A_1) \times \aleph(A_2) \times \ldots \times \aleph(A_r).$$

Получили основное правило умножения комбинаторики или комбинаторного анализа. Рассмотрим задачу, когда можно применить это правило.

Пример 2.10. Лабиринт для мыши составлен из четырех клеток. Из первой клетки мышь может проникнуть только во вторую через три различные щели с метками a_1, a_2, a_3. Из второй клетки мышь может попасть только в третью через две различные норы с метками b_1, b_2. Наконец, из третьей клетки мышь может переползти только в четвертую через четыре различные отверстия с метками c_1, c_2, c_3, c_4. В четвертой клетке расположена кормушка для мыши. Сколькими способами мышь может достигнуть кормушки из первой клетки? Пусть $A_1 = \{a_1, a_2, a_3\}$, $A_2 = \{b_1, b_2\}$, $A_3 = \{c_1, c_2, c_3, c_4\}$. Каждому способу передвижения мыши из первой клетки в четвертую сопоставим упорядоченную тройку вида (x_1, x_2, x_3), где $x_1 \in A_1$, $x_2 \in A_2$, $x_3 \in A_3$. Множество $A_1 \times A_2 \times A_3 = \{(x_1, x_2, x_3)\colon x_1 \in A_1, x_2 \in A_2, x_3 \in A_3\}$

согласно правилу умножения комбинаторики содержит $3 \times 2 \times 4 = 24$ элемента. Поэтому у мыши имеется 24 возможных способа достижения кормушки.

Пример 2.11. Пусть некоторая операция D состоит в последовательном выполнении k действий D_1, $D_2, \ldots$, D_k. Действие D_1 можно выполнить n_1 различными способами, действие D_2 — n_2 способами, ..., действие D_k выполняется n_k способами. Если изменение результата любого из действий неизбежно ведет к изменению результата всей операции D, то операция D может быть выполнена $n_1 \times n_2 \times \ldots \times n_k$ различными способами.

Способ перестановок элементов неупорядоченного множества. Пусть дано неупорядоченное множество $G = \{a_1, a_2, \ldots, a_N\}$. Упорядоченный набор $(x_1,\ x_2, \ldots,\ x_N)$, где $x_1 \in G_1$, $x_2 \in G_2, \ldots$, $x_N \in G_N$ и $x_i \neq x_j$ при $i \neq j$, называется *перестановкой из N элементов множества G*. Отсюда следует, что различные перестановки из N элементов множества G отличаются только порядком следования элементов. Образуем теперь новое множество Ω_1, элементами которого являются всевозможные перестановки из N элементов множества G, т. е. имеем

$$\Omega_1 = \{(x_1, x_2, \ldots, x_N)\colon x_1 \in G, x_2 \in G, \ldots, x_N \in G;\ x_i \neq x_j \text{ при } i \neq j\}.$$

По правилу умножения комбинаторики непосредственно находим $\aleph(\Omega_1) = N \times (N-1) \times \ldots \times 1 = N!$ (читается: N-факториал). Значит, неупорядоченное множество $G = \{a_1, a_2, \ldots, a_N\}$ из N элементов можно упорядочить $N!$ способами. По определению полагают, что $0! = 1$. В качестве примера различных перестановок неупорядоченного множества $G = \{1, 2, \ldots, 7\}$ можно рассмотреть упорядоченные множества вида: $(1,2,3,4,5,6,7)$, $(3,1,2,4,5,6,7)$, $(7,6,5,4,3,2,1)$, $(4,5,6,1,2,3,7)$. Всего можно составить $7!$ различных перестановок из семи элементов. Используя эти стандартные комбинации, можно посчитать, например, каким числом способов можно назначить семь студентов на дежурство в семь различных пунктов. Ясно, что всего существует $7! = 5\,040$ различных таких назначений.

Рассмотрим еще пример.

Пример 2.12. Пять человек некоторым образом рассаживаются за круглым столом для игры в покер. Если $G = \{a_1, a_2, \ldots, a_5\}$ — неупорядоченное множество игроков в покер, то множество

$$\Omega_1 = \{(x_1, x_2, x_3, x_4, x_5)\colon$$
$$x_1 \in G, \quad x_2 \in G, \quad x_3 \in G, \quad x_4 \in G, \quad x_5 \in G; \quad x_i \neq x_j \text{ при } i \neq j\},$$

и всего будет $5! = 1 \times 2 \times 3 \times 4 \times 5 = 120$ способов расположений игроков за игральным столом.

Приведем теперь решение задачи на одновременное применение правила умножения комбинаторики и способа перестановок элементов данного неупорядоченного множества.

Пример 2.13. Сколькими способами можно расставить в один ряд m мужчин и m женщин так, чтобы каждая женщина имела четный номер? Пусть $G_1 = \{2, 4, \ldots, 2m\}$ — множество номеров для женщин и $G_2 = \{1, 3, \ldots, 2m-1\}$ — множество номеров для мужчин. Тогда множество вида $A_1 = \{y = (y_1, y_2, \ldots, y_m)\colon y_1 \in G_1,\, y_2 \in G_1, \ldots, y_m \in G_1;$ $y_i \neq y_j$ при $i \neq j\}$ содержит $m!$ всех расстановок номеров для женщин, а следующее множество $A_2 = \{z = (z_1, z_2, \ldots, z_m)\colon z_1 \in$ $\in G_2,\, z_2 \in G_2, \ldots, z_m \in G_2;$ $z_i \neq z_j$ при $i \neq j\}$ включает ровно $m!$ всех расстановок номеров для мужчин. Каждому фиксированному способу расстановки женщин с четными номерами соответствует $m!$ способов расстановки мужчин с нечетными номерами. Отсюда по основному правилу умножения комбинаторики следует, что $A_1 \times A_2 = \{(y, z)\colon y \in A_1,\, z \in A_2\}$ всех расстановок составлено из $m! \times m!$ элементов. Следовательно, общее число расстановок в один ряд m мужчин с нечетными номерами и m женщин с четными номерами равно $(m!)^2$.

Способ размещений элементов неупорядоченного множества. Пусть снова дано неупорядоченное множество $G = \{a_1, a_2, \ldots, a_N\}$. Упорядоченный набор $(x_1, x_2, \ldots, x_k)$, где $k \leqslant N$, $x_1 \in G$, $x_2 \in G, \ldots,$ $x_k \in G$ и $x_i \neq x_j$ при $i \neq j$ называется *размещением из N элементов по k*. Следовательно, различные размещения из N элементов по k отличаются либо своими элементами, либо их порядком следования. Образуем теперь новое множество Ω_2, элементами которого являются все размещения из N элементов по k. Тогда имеем

$$\Omega_2 = \{(x_1, x_2, \ldots, x_k)\colon x_1 \in G, x_2 \in G, \ldots, x_k \in G; x_i \neq x_j \text{ при } i \neq j\}.$$

По основному правилу умножения комбинаторики получим, что $\aleph(\Omega_2) = N \times (N-1) \times \ldots \times (N-k+1)$. Число $\aleph(\Omega_2)$ различных размещений из N по k обозначают специальным символом A_N^k и вычисляют по формуле $A_N^k = N!/(N-k)!$, где $k = \overline{0, N}$. Итак, можно образовать A_N^k упорядоченных k-элементных подмножеств множества G. При $k = N$ получаем в точности предыдущий способ построения множества Ω_2 и $N! = A_N^N$.

Примерами различных размещений из семи элементов по три могут служить упорядоченные множества вида: $(1, 2, 3)$, $(2, 1, 3)$, $(3, 1, 2)$, $(1, 5, 7)$, $(5, 1, 2)$. Всего можно составить $A_7^3 = 7!/4! = 210$ различных размещений из семи элементов по три. Посчитаем, сколько существует различных способов назначить трех студентов из семи на дежурство, если один должен дежурить в столовой, другой в библиотеке, а третий в университетском саду. Здесь можно использовать размещения из семи элементов по три, если условиться, что первый элемент размещения соответствует номеру студента, назначенному в столовую, второй элемент размещения — это номер студента, которому досталось дежурство в библиотеке, а третий элемент соответствует студенту, который пойдет работать в сад. Итак, число различных таких назначений равно 210.

Пусть теперь в упорядоченном наборе вида $(x_1, x_2, \ldots, x_k)$ могут быть повторяющиеся или одинаковые элементы из множества G. В этом случае упорядоченный набор $(x_1, x_2, \ldots, x_k)$ называется *размещением с повторениями из N элементов по k*. Если множество

$$\Omega_3 = \{(x_1, x_2, \ldots, x_k) \colon x_1 \in G,\ x_2 \in G, \ldots, x_k \in G\},$$

то число $\aleph(\Omega_3)$ всех различных размещений с повторениями из N элементов по $k = 0, 1, \ldots$ обозначают специальным символом $\overline{A}_N^k$, и по правилу умножения $\overline{A}_N^k = N^k$. Решим следующую задачу.

Пример 2.14. Дачник имеет два участка земли. Первый участок он разбил на четыре сектора, а второй — на три. На первом участке он решил посадить раздельно картофель, капусту и чеснок. На втором участке будет выращивать только лук и морковь. Каждая из выбранных овощных культур может быть посажена только на одном из указанных секторов, при этом лук и морковь можно сажать в одном секторе. Сколькими способами можно выбрать посадку данных сельскохозяйственных культур? Обозначим через $G_1 = \{1, 2, 3, 4\}$ множество меток для секторов первого земельного участка и через $G_2 = \{5, 6, 7\}$ множество меток для секторов второго земельного участка. Пусть x_1, x_2 и x_3 — номера тех секторов первого участка, на которых соответственно посажены картофель, капуста и чеснок. Тогда множество $A_1 = \{x = (x_1, x_2, x_3) \colon x_1 \in G_1,\ x_2 \in G_1,\ x_3 \in G_1;\ x_i \neq x_j$ при $i \neq j\}$ содержит $A_4^3 = 4 \times 3 \times 2 = 24$ элемента. Поэтому число всех различных посадок картофеля, капусты и чеснока на первом участке равно 24. Если теперь y_1 и y_2 — номера тех секторов, на которых соответственно посажены лук и морковь, то множество $A_2 = \{y = (y_1, y_2) \colon y_1 \in G_2,\ y_2 \in G_2\}$ содержит $3^2 = 9$ элементов. Значит, число всевозможных посадок лука и моркови на втором участке равно 9. Отметим, что множество A_1 составлено из размещений без повторений из четырех элементов по три, а множество A_2 — из размещений с повторениями из трех элементов по два. Любому заданному способу посадки овощей на первом участке соответствует 9 различных посадок овощей на втором участке. Следовательно, по основному правилу умножения комбинаторики находим, что множество $A_1 \times A_2 = \{(x, y) \colon x \in A_1,\ y \in A_2\}$ включает $24 \times 9 = 216$ элементов.

Число 216 совпадает с общим числом всевозможных посадок данных овощей на двух участках.

Способ сочетаний элементов неупорядоченного множества. Пусть опять задано неупорядоченное множество $G = \{a_1, a_2, \ldots, a_N\}$. Неупорядоченный набор $\{x_1, x_2, \ldots, x_k\}$, где $k \leqslant N$, $x_1 \in G$, $x_2 \in G, \ldots$, $x_k \in G$ и $x_i \neq x_j$ при $i \neq j$ называется *сочетанием из N элементов по k*. Значит, различные сочетания из N элементов по k отличаются лишь своими элементами. Образуем теперь новое множество Ω_4, элементами которого являются всевозможные сочетания из N элементов по k. Очевидно, что множество

$$\Omega_4 = \{\{x_1, x_2, \ldots, x_k\}\colon\ x_1 \in G,\ x_2 \in G, \ldots, x_k \in G;\ x_i \neq x_j \text{ при } i \neq j\}.$$

Число $\aleph(\Omega_4)$ элементов этого множества обозначают специальным символом C_N^k. Если теперь каждый элемент $\{x_1, x_2, \ldots, x_k\}$ множества Ω_4 упорядочить $k!$ способами, то получим всевозможные размещения или упорядоченные k-элементные подмножества множества G. При таком способе каждое размещение получается только один раз. Поэтому $C_N^k = N!/(k!\,(N-k)!) = N \times (N-1) \times \ldots \times (N-k+1)/k!$ и $k!\,C_N^k = A_N^k$, где $k = \overline{0, N}$. Итак, можно образовать C_N^k неупорядоченных k-элементных подмножеств множества G.

Рассмотрим пример составления различных сочетаний. Пусть множество G есть группа из семи студентов. Пронумеруем всех студентов, тогда $G = \{1, 2, \ldots, 7\}$. Различные неупорядоченные наборы по три студента будут являться примерами различных сочетаний из семи по три. Например, множества $\{1, 2, 3\}$, $\{1, 2, 4\}$, $\{1, 7, 8\}$, $\{3, 5, 6\}$, $\{4, 6, 7\}$ суть различные сочетания из семи по три. Всего можно составить ровно $C_7^3 = 7!/(3!\,(7-3)!) = (7 \times 6 \times 5)/3! = 35$ различных сочетаний из семи элементов по три. Если перед нами стоит задача вычисления числа различных способов, которыми можно выбрать трех студентов из семи для дежурства по столовой, то ответом будет число 35. Число различных сочетаний из семи по три получается в шесть раз меньше, чем число различных размещений из семи элементов по три. Это связано с тем, что размещения учитывают порядок следования элементов, а сочетания нет. Три различных элемента можно упорядочить ровно шестью различными способами.

При $k = N$ множество Ω_4 состоит из одного элемента $\{a_1, a_2, \ldots, a_N\}$. Напротив, множество $\{A\colon A \subseteq G\} = \{A_1, A_2, \ldots, A_q\}$ всех неупорядоченных подмножеств множества G, как уже было установлено в первой главе, содержит $q = 2^N$ элементов. Этот факт можно подтвердить еще и другим способом. Действительно, это множество состоит из пустого множества, например $A_1 = \varnothing$, из всех одноэлементных подмножеств множества G, из всех двухэлементных подмножеств множества G,..., и, наконец, только из одного N-элементного подмножества A_q множества G, которое естественно совпадает с самим множеством G. Отсюда, вспоминая из школьного курса математики разложение бинома Ньютона, непосредственно получаем, что $q = C_N^0 + C_N^1 + C_N^2 + \ldots + C_N^N = (1+1)^N = 2^N$.

Если неупорядоченный набор $\{x_1, x_2, \ldots, x_k\}$ может содержать одинаковые элементы из множества G, то аналогично размещениям с повторениями такой набор называется *сочетанием с повторениями из N элементов по k*. Множество всех сочетаний с повторениями из N элементов по k можно записать в виде

$$\Omega_5 = \big\{\{x_1, x_2, \ldots, x_k\} \colon x_1 \in G,\ x_2 \in G, \ldots, x_k \in G\big\}.$$

Число $\aleph(\Omega_5)$ различных сочетаний с повторениями обозначают символом $\overline{C}_N^k$, и имеет место равенство $\overline{C}_N^k = C_{N+k-1}^k$, где число $k = 0, 1, \ldots$. Действительно, поставим во взаимно однозначное соответствие каждому сочетанию с повторениями из N различных элементов по k вектор размерности $N+k-1$ из $N-1$ нулей и k единиц по следующему закону:

- первые компоненты вектора равны единице и их столько, сколько элементов a_1 содержит сочетание с повторениями;
- затем обязательно следует нулевая компонента, и ставим опять единичные компоненты и их столько, сколько элементов a_2 входит в это сочетание с повторениями, и т. д.

Отметим, что если сочетание с повторениями не содержит некоторый элемент множества G, то данный вектор не имеет единичной компоненты, соответствующей этому элементу. Поэтому любой рассматриваемый здесь вектор содержит k единичных и $N-1$ нулевых компонент и, значит, определяется номерами единичных или нулевых компонент. Если теперь $M = \{1, 2, \ldots, N+k-1\}$ и

$$\Omega_6 = \{\{y_1, y_2, \ldots, y_k\} \colon\ y_1 \in M,\ y_2 \in M, \ldots, y_k \in M;\ y_i \neq y_j \text{ при } i \neq j\},$$

то любой элемент (сочетание) из множества Ω_6 взаимно однозначно задает номера единичных компонент указанного выше вектора размерности $N+k-1$. Так как $\aleph(\Omega_6) = C_{N+k-1}^k$ и число всех векторов из k единичных и $N-1$ нулевых компонент совпадает с числом всех различных сочетаний с повторениями из N различных элементов по k, то $\aleph(\Omega_5) = \overline{C}_N^k = C_{N+k-1}^k$.

Изучим теперь следующий простой пример, в котором встречаются комбинации типа сочетаний и сочетаний с повторениями.

Пример 2.15. В детском уголке находятся семь различных по цвету кубиков, два одинаковых конструктора и два одинаковых мячика. Сколькими способами ребенок может отобрать шесть из указанных игрушек так, чтобы в уголке всегда оставалось три кубика? Обозначим через символ $G_1 = \{a_1, a_2, \ldots, a_7\}$ множество кубиков, каждый из которых окрашен в один из цветов радуги, и через $G_2 = \{b, b, c, c\}$ — множество, первые два элемента которого соответствуют наборам для конструирования, а вторые — мячикам. Множество G_2 содержит только два различных элемента: b и c. Пусть неупорядоченное множество $\{x_1, x_2, x_3, x_4\}$ означает выбор ребенком четырех различных кубиков. Заметим, что порядок отбора игрушек ребенком не играет роли. Тогда множество $A_1 = \{\{x_1, x_2, x_3, x_4\} \colon x_1 \in G_1,\ x_2 \in G_1,\ x_3 \in G_1,\ x_4 \in G_1;$ $x_i \neq x_j$ при $i \neq j\}$ сочетаний из семи различных элементов по четыре определяет все способы отбора ребенком кубиков. При этом число $\aleph(A_1)$ равно $C_7^4 = 35$. Если y_1 и y_2 — отобранные ребенком две игрушки, которые естественно не являются кубиками, то множество $A_2 = \{\{y_1, y_2\} \colon y_1 \in G_2,\ y_2 \in G_2\} = \{\{b, b\}, \{b, c\}, \{c, c\}\}$ сочетаний с повторениями из двух различных элементов по два содержит все способы отбора ребенком не кубиков. Число $\aleph(A_2)$ всех различных сочетаний с повторениями из двух различных элементов по два равно $C_{2+2-1}^2 = 3$. Если ребенок выбрал четыре кубика, то затем конструкторы или мячики могут быть им отобраны тремя способами. По правилу умножения комбинаторики находим, что у ребенка существует в точности $35 \times 3 = 105$ всевозможных вариантов выбора игрушек.

Рассмотрим теперь, как применяются правило прямого произведения и стандартные комбинации при решении ряда типичных задач комбинаторики и задач вычисления вероятностей случайных событий для простейших классических экспериментов.

Пример 2.16. Нужно составить комиссию из трех равноправных членов. Выбирать представителей комиссии можно из четырех семей (муж и жена), причем от каждой семьи не может избираться более одного человека. Приведем решение этой простой задачи первым способом. Пусть операция D означает составление комиссии. Тогда она может быть представлена как последовательное выполнение следующих двух действий:

1) действие D_1 есть выбор трех семей, члены которых будут представлены в комиссии;
2) действие D_2 означает выбор по одному человеку (муж или жена) в каждой из трех семей, представленных в комиссии.

Действие D_1 может быть выполнено $n_1 = C_4^3$ различными способами. Действие D_2 выполняется $n_2 = 2 \times 2 \times 2 = 8$ различными способами. Окончательно получаем $n = n_1 \times n_2$, и, значит, такую комиссию можно составить 32 различными способами. Следует заметить, что действия D_i, $i = 1, 2$, выбирались таким образом, что изменение результата любого из них ведет к новому составу комиссии.

Приведем теперь другое решение этой задачи. Операцию D представим в виде последовательного выполнения действий S_1 и S_2:

1) действие S_1 означает выбор в каждой из четырех семей по одному члену;
2) действие S_2 заключается в выборе трех членов комиссии из четырех, отобранных в результате выполнения S_1.

Действие S_1 может быть выполнено $n_1 = 2 \times 2 \times 2 \times 2$ различными способами, действие S_2 выполняется $n_2 = C_4^3$ различными способами. В результате имеем $n = 2^4 \times C_4^3 = 64$ различных вариантов комиссии. Ответы, естественно, не совпали.

Все дело здесь в том, что правило умножения во втором случае применено с ошибкой. Можно изменить результат действия S_1 и получить опять тот же состав комиссии. Поясним сказанное на конкретном примере. Пусть есть четыре семьи. Перечислим членов этих семей по именам и результат действий S_1 и S_2 изобразим на следующей схеме:

Семьи	1	2	3	4
	Валентина	Надежда	Нина	Елена
	Павел	Сергей	Александр	Андрей
Действие S_1	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
	Павел	Сергей	Александр	Андрей
Действие S_2		$\Downarrow$		
		Павел, Сергей, Александр		

В данном примере результат первого действия есть выбор Павла, Сергея, Александра и Андрея. Результат второго действия и есть комиссия в составе Павла, Сергея и Александра. Но можно изменить результат действия S_1, выбрав на первом этапе Павла, Сергея, Александра, Елену, и получить тот же состав комиссии, при этом на втором этапе опять же выбираем Павла, Сергея, Александра как показано на следующей схеме:

Семьи	1	2	3	4
	Валентина	Надежда	Нина	Елена
	Павел	Сергей	Александр	Андрей
Действие S_1	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
	Павел	Сергей	Александр	Елена
Действие S_2		$\Downarrow$		
		Павел, Сергей, Александр		

Таким образом, изменение результата действия S_1 не привело к изменению результата всей операции.

Пример 2.17. Сколько различных четырехзначных чисел n можно составить из цифр 0, 1, 2, 3, 4, 5, 6, если ни одна из цифр не повторяется? Можно решать данную задачу с использованием основного правила комбинаторики. Процесс составления четырехзначного числа можно разбить на четыре последовательных действия:

1) D_1 — выбор первой цифры;
2) D_2 — выбор второй цифры;
3) D_3 — выбор третьей цифры;
4) D_4 — выбор четвертой цифры.

На первом месте в числе может стоять одна из шести цифр 1, 2, 3, 4, 5, 6.

Если первая цифра выбрана, то на второе место может быть поставлена любая из шести оставшихся цифр (второй цифрой может быть ноль). После выбора первых двух цифр, третья выбирается пятью различными способами. Если первые три цифры в числе известны, то четвертую можно выбрать четырьмя различными способами. Тогда имеем: $n_1 = 6$, $n_2 = 6$, $n_3 = 5$, $n_4 = 4$ и $n = 6 \times 6 \times 5 \times 4 = 720$. Приведем решение этой задачи с использованием стандартных комбинаций. В данном случае можно применить размещения без повторений. Подсчитаем сначала, сколько различных векторов вида $\mathbf{b} = (a_1, a_2, a_3, a_4)$ можно составить, если все компоненты этого вектора различны и могут быть цифрами из множества $G = \{0, 1, 2, 3, 4, 5, 6\}$. Всего таких векторов равно числу A_7^4. Далее подсчитаем, сколько существует различных векторов $\mathbf{b}$, у которых $a_1 = 0$, а все другие компоненты различны и могут быть цифрами 1, 2, 3, 4, 5, 6. Таких векторов A_6^3. Четырехзначных чисел, удовлетворяющих условиям задачи, будет столько, сколько можно составить различных векторов $\mathbf{b}$, у которых все компоненты различны, могут принимать значения 0, 1, 2, 3, 4, 5, 6 и $a_1 \neq 0$. Число таких векторов равно $A_7^4 - A_6^3 = 720$.

Пример 2.18. Сколько различных пятизначных чисел n можно составить из цифр 1, 2, 3, 4, 5, 6 и 7, если есть только одна цифра, которая повторяется в числе ровно два раза, а все другие цифры разные? Процесс составления числа, удовлетворяющего условию задачи, представим в виде последовательного выполнения следующих трех независимых действий. Действие D_1 означает выбор цифры, которая будет повторяться два раза. Действие D_2 есть выбор двух мест в пятизначном числе для повторяющейся цифры. Наконец, действие D_3 означает выбор и расстановку трех разных цифр на оставшиеся три свободных места в пятизначном числе. Отсюда легко получаем, что $n_1 = 7$, $n_2 = C_5^2$, $n_3 = C_6^3 \times 3! = A_6^3$. Значит, $n = 7 \times 10 \times 120 = 8\,400$.

Пример 2.19. Сколькими различными способами можно разложить в n различных ящиках r белых и s черных шариков, если в каждом ящике может находиться любое количество шариков, а некоторые ящики могут оставаться пустыми?

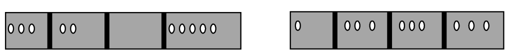
Решение. Операцию размещения шариков по ящикам разобьем на два последовательных действия:

1) D_1 есть распределение белых шариков;

2) D_2 означает распределение черных шариков по ящикам.

Следовательно, достаточно вычислить, каким числом способов можно выполнить каждое из действий, и использовать основной закон комбинаторики. Посчитаем количество различных разложений белых шариков по ящикам.

На рис. 2.2 изображены два различных расположения белых шариков по ящикам. Нетрудно видеть, что разложения шариков, изображенные здесь можно однозначно описать с помощью цепочки шариков и палочек (границы ящиков), как это сделано на рис. 2.3.



а *б*

Рис. 2.2

ooolooll ooooo (*а*) olooolooolooo (*б*)

Рис. 2.3

Если всего ящиков n, а белых шариков r, то в этой цепочке будет $n-1$ палочка и r шариков. Так что различных разложений n_1 белых шариков будет столько, сколько можно составить различных цепочек длины, равной величине $(n-1)+r$, из $n-1$ палочек и r шариков. А таких цепочек столько, сколькими различными способами можно выбрать r мест в цепочке под шарики, т. е. $n_1 = C^r_{n-1+r}$. Аналогично подсчитывается число n_2 разложений черных шариков по ящикам, т. е. имеем $n_2 = C^s_{n-1+s}$. Количество n распределений всех шариков по ящикам, согласно основному закону произведения комбинаторики, будет равно $C^r_{n-1+r} \times C^s_{n-1+s}$.

Эту задачу можно решить с использованием стандартных комбинаций. Процесс распределения белых шариков по ящикам может быть выполнен следующим образом. Берем белый шарик и выбираем для него любой из n ящиков, далее берем следующий белый шарик, опять выбираем для него любой из n ящиков и т. д., процедуру повторяем r раз, т. е. для каждого шарика выбираем ящик, причем каждый ящик может быть выбран несколько раз. Если некоторый ящик выбрали

q раз, то это означает, что в нем окажутся q шариков. Различных размещений белых шариков по ящикам равно числу способов, которыми можно выбрать число различных сочетаний с повторениями из n по r. Следовательно, имеет место равенство $n_1 = \overline{C}^r_n$. Число различных разложений белых и черных шариков по n ящикам будет равно $n = \overline{C}^r_n \times \overline{C}^s_n$. Решая эту задачу двумя различными способами, получаем общую формулу $\overline{C}^r_n = C^r_{n-1+r}$ для вычисления числа различных сочетаний с повторениями из n по r.

Пример 2.20. В партии из N изделий имеется Q бракованных. Из партии наугад выбирается S изделий. Определить вероятность $\mathsf{P}(A)$ того, что среди S изделий имеется ровно K бракованных.

Решение. Из N изделий S изделий можно выбрать наугад $n = C^S_N$ способами. Из N изделий, среди которых Q бракованных и $N - Q$ стандартных, можно выбрать K бракованных и $S - K$ стандартных $m = C^K_Q\, C^{S-K}_{N-Q}$ способами. Тогда, согласно классическому определению, искомая вероятность $\mathsf{P}(A) = m/n = C^K_Q\, C^{S-K}_{N-Q}/C^S_N$.

Пример 2.21. Бросают один раз три игральные кости. Найти вероятность того, что выпадут три различных числа очков на верхних гранях игральных костей.

Решение. Для этого опыта число n всех элементарных исходов равно $n = 6^3$, а число m элементарных исходов, когда выпадают три различных числа очков на верхних гранях игральных костей, равно $6 \times 5 \times 4$. Значит, $\mathsf{P}(A) = m/n = 120/216 = 5/9 \approx 0{,}556$.

Пример 2.22. Из колоды 52 карт вынимаются наудачу две карты. Какова вероятность того, что вынуты две фигуры? При этом фигурами объявляются валет, дама и король.

Решение. Из колоды 52 карт можно выбрать наугад две различные карты $n = C^2_{52}$ способами. При этом из 12 фигур две различные фигуры можно отобрать $m = C^2_{12}$ способами. Отсюда $\mathsf{P}(A) = m/n = C^2_{12}/C^2_{52} = 11/(13 \times 17) \approx 0{,}049$.

Пример 2.23. Колода в 36 карт делится пополам наудачу. Найти вероятность того, что в каждой части окажется по равному числу черных и красных карт.

Решение. Механизм деления пополам колоды карт можно выполнить следующим образом: из колоды в 36 карт наудачу без возвращения отбирают 18 карт. При этом порядок отобранных карт не играет роли. Поэтому в этом опыте $n = C^{18}_{36}$, $m = C^9_{18}\, C^{18-9}_{36-18} = C^9_{18}\, C^9_{18}$. Следовательно, вероятность $\mathsf{P}(A) = m/n = C^9_{18}\, C^9_{18}/C^{18}_{36} \approx 0{,}26$.

Пример 2.24. Две игральные кости брошены один раз. Какова вероятность того, что сумма выпавших очков есть простое число?

Решение. Пусть упорядоченная пара вида (x, y) означает описание такого элементарного исхода, когда на первой кости выпадает число очков $x = 1, 2, \ldots, 6$ и на второй выпадает $y = 1, 2, \ldots, 6$. Итак, здесь имеется всего $n = 36$ элементарных событий. В табл. 2.1 приведены всевозможные суммы $(x + y)$ очков, которые соответствуют всем элементарным исходам, которые могут появиться в этом опыте, $m = 1 + 2 + 4 + 6 + 2 = 15$.

Таблица 2.1

y\ x	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Из табл. 2.1 видно, что могут выпасть суммы вида: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 и 12. Из этих чисел простыми являются только числа 2, 3, 5, 7 и 11. С помощью табл. 2.2 легко получаем число m благоприятствующих рассматриваемому событию A элементарных исходов, т. е. $m = 1 + 2 + 4 + 6 + 2 = 15$. Итак, $\mathsf{P}(A) = m/n = 15/36 \approx 0{,}417$.

Таблица 2.2

Сумма	2	3	5	7	11
Число таких сумм	1	2	4	6	2

Пример 2.25. Пусть эксперимент заключается в бросании двух игральных костей, грани каждой из которых занумерованы от 1 до 6. Какова вероятность того, что на обеих костях выпадет одинаковое число очков?

Решение. Для этого опыта легко найти, что $n = 36$, $m = 6$ и, значит, $\mathsf{P}(A) = m/n = 6/36 = 1/6$. Решение этой простой задачи было известно еще в XVII столетии во Франции, во времена правления Людовика XIV.

2.3. Дискретные вероятностные модели симметричных экспериментов. Рассмотрим теперь ряд тестовых и поучительных примеров на непосредственный подсчет вероятностей в классической симметричной схеме и на объяснение возникающих при этом парадоксов. Решение большого числа конкретных и частных задач позволяет без особого труда строить дискретные вероятностные модели в общем случае.

Пример 2.26 (*задача Галилео Галилея – Де Мере*). С помощью некоторого механизма на некоторую поверхность стола подбрасывается симметричная игральная кость три раза и определяется сумма выпавших очков при этих бросках. Можно рассмотреть несколько другой вариант этого эксперимента. С помощью некоторого механизма на некоторую поверхность стола один раз подбрасываются три симметричные игральные кости, изготовленные из одного и того же материала, например, из слоновой кости, и определяется сумма выпавших очков на этих костях.

Согласно одной из легенд, в семнадцатом веке при дворе Людовика XIV была распространена приведенная в п. В. 4 игра для двух лиц. Мы приводим здесь более простой ее вариант. Крупье подбрасывает одну игральную кость три раза или три игральные кости один раз случайным образом с помощью некоторого механизма. У каждого из игроков заранее имеется возможность сделать одинаковые денежные ставки на сумму выпавших очков, например, равную одиннадцати или двенадцати. Предположим, что игрок номер один выбирает сумму очков, равную одиннадцати, а второй игрок выбирает сумму очков, равную двенадцати. Выигрывает поставленные ставки игрок номер один, если сумма выпавших очков равна одиннадцати, и выигрывает эти же ставки второй игрок, если сумма выпавших очков равна двенадцати. В противном случае объявляется ничья, и ставки сохраняются, если игра повторяется, либо игра прекращается, и сделанные соответствующие ставки возвращаются игрокам.

Как правило, игроки не придавали значения выбору суммы очков, равной одиннадцати или двенадцати, так как считали одинаковыми шансы их выпадения на следующем основании. Сумму очков одиннадцать можно получить следующими шестью различными способами:

1) на некоторой кости выпадает шесть очков, на другой кости четыре и на оставшейся выпадает одно очко;
2) на некоторой кости выпадает шесть очков, на другой кости три, а на оставшейся выпадает два очка;
3) на некоторой кости выпадает пять очков, на другой пять и на оставшейся кости выпадает одно очко;
4) на одной кости выпадает пять очков, на другой четыре, а на оставшейся кости выпадает два очка;
5) на одной кости выпадает пять очков, на другой три, а на оставшейся кости выпадает три очка;
6) на некоторой кости выпадает четыре очка, на другой четыре и на оставшейся кости выпадает три очка.

Сумма очков двенадцать может быть получена также шестью различными способами:

1) на одной кости выпадает шесть очков, на другой пять, а на оставшейся кости выпадает одно очко;
2) на некоторой кости выпадает шесть очков, на другой четыре, а на оставшейся кости выпадает два очка;
3) на одной кости выпадает шесть, на другой три и на оставшейся кости три очка;
4) на некоторой кости выпадает пять очков, на другой пять, а на оставшейся кости два очка;
5) на одной кости выпадает пять очков, на другой четыре и на оставшейся кости три очка;
6) на некоторой кости выпадает четыре очка, на другой четыре и, наконец, на оставшейся кости также выпадает четыре очка.

В более сжатом виде все это можно отобразить следующим образом. Одиннадцать очков можно представить в виде суммы трех чисел шестью способами: $6+4+1$, $6+3+2$, $5+5+1$, $5+4+2$, $5+3+3$, $4+4+3$. Двенадцать очков можно получить в виде суммы трех чисел тоже шестью способами: $6+5+1$, $6+4+2$, $6+3+3$, $5+5+2$, $5+4+3$, $4+4+4$. Равенство числа указанных способов, при которых выпадает сумма очков одиннадцать или сумма очков двенадцать, для многих игроков на интуитивном уровне означает, что шансы выигрыша игроков одинаковы. Однако в семнадцатом веке среди игроков была распространена легенда о том, что придворный французского королевского двора — шевалье Де Мере парадоксально часто выигрывал. Это средневековое предание можно объяснить следующим образом.

Попарно несовместимые исходы данной игры, взаимно однозначно соответствующие перечисленным способам получения суммы очков одиннадцать и суммы очков двенадцать, не являются равновозможными. Более того, сумму очков двенадцать можно получить, когда на каждой из трех костей выпадает одно и тоже число очков. Такая ситуация заведомо не может быть при получении суммы очков одиннадцать, т. е. мы имеем явную асимметрию при выборе игроками суммы очков одиннадцать или суммы очков двенадцать.

Приведем правильное решение этой задачи. Для этой игры или эксперимента комплекс априори заданных условий Σ содержит следующие основные элементы:

1) условие u_1 означает выбор симметричной игральной кости;
2) условие u_2 означает, что игральная кость подбрасывается три раза;
3) условие u_3 отвечает за непреднамеренный механизм подбрасывания игральной кости, например, с помощью крупье;

4) условие u_4 обеспечивает гладкую поверхность стола, на которую падает игральная кость;
5) условие u_5 определяет освещенность поверхности стола, которая позволяет зафиксировать число выпавших очков на игральной кости.

Обозначим через вектор (x, y, z) описание элементарного исхода этой игры в том случае, когда при первом броске выпадает x очков, при втором броске y и при третьем z. Каждая из компонент x, y, z может принимать целое значение от единицы до шести. Поэтому множество

$$\Omega = \{\omega = (x, y, z)\colon x, y, z \in \{1, 2, 3, 4, 5, 6\}\}.$$

По основному правилу умножения комбинаторики достоверное случайное событие Ω содержит $n = 216$ описаний всех элементарных исходов. Пусть $\mathcal{F}$ — множество всех подмножеств множества Ω, т. е. $\mathcal{F} = \{A_1, A_2, \ldots, A_{2^{216}}\}$.

Обозначим через $A \in \mathcal{F}$ и $B \in \mathcal{F}$ события, которые заключаются в том, что сумма выпавших очков равна одиннадцати (A) и, соответственно, двенадцати (B). Пусть произошло событие B и оно реализовалось суммой $x + y + z$, в которой числа x, y, z различны, например, $x = 6$, $y = 5$, $z = 1$. Тогда существуют шесть описаний $(6, 5, 1)$, $(6, 1, 5)$, $(5, 6, 1)$, $(5, 1, 6)$, $(1, 6, 5)$, $(1, 5, 6)$ различных элементарных исходов этой игры, благоприятствующих появлению события B. Если существует всего два одинаковых числа в сумме $x + y + z$, например, $x = 6$, $y = z = 3$, тогда имеем три описания: $(6, 3, 3)$, $(3, 6, 3)$, $(3, 3, 6)$ различных элементарных исходов, каждый из которых одновременно наступает с исходом B. Наконец, при $x = y = z$, будет одно описание $(4, 4, 4)$ элементарного исхода, одновременно наступающего с исходом B.

Таким образом, находим 25 элементарных исходов, благоприятствующих появлению события B. Рассуждая аналогичным образом, получаем 27 элементарных исходов, благоприятствующих появлению события A:

$$\begin{array}{ccc}
6+4+1 & 6+3+2 & 5+4+2 \\
6+1+4 & 6+2+3 & 5+2+4 \\
4+6+1 & 3+6+2 & 4+5+2 \\
4+1+6 & 3+2+6 & 4+2+5 \\
1+6+4 & 2+6+3 & 2+5+4 \\
1+4+6 & 2+3+6 & 2+4+5 \\
5+5+1 & 5+3+3 & 4+4+3 \\
5+1+5 & 3+5+3 & 4+3+4 \\
1+5+5 & 3+3+5 & 3+4+4
\end{array}$$

Выбор произвольного элементарного исхода и его описание в виде вектора (x, y, z) дают полное ощущение симметрии данного эксперимента, и, значит, шанс наступления каждого из элементарных исходов представляется различным субъектам одинаковым. По классическому определению вероятности событий получаем $\mathsf{P}(A) = 27/216 > \mathsf{P}(B) = 25/216$. Значит, на самом деле, более часто выпадает сумма в 11 очков, нежели сумма в 12 очков.

Отсюда возможно следующее разумное поведение в этой упрощенной игре, например, первого игрока. Если противник второй игрок делает ставку на исход A, то первый игрок требует ограничение на время игры и на денежные ставки, делая их минимально возможными, или отказывается от игры. В этом случае он, возможно, проигрывает незначительную сумму. В противном случае, когда второй игрок делает ставку на исход B, первый игрок, поставив на результат A, может играть достаточно долго с максимально возможными ставками и иметь большой шанс на значительный выигрыш.

В результате большого опыта игры в кости француз Де Мере (1607–1648) подметил, что при одновременном бросании трех игральных костей более часто выпадает комбинация, дающая в сумме 11 очков, чем комбинация, дающая в сумме 12 очков. Хотя с точки зрения интуиции его современников считалось, что эти комбинации имеют одинаковую вероятность. Такую ошибку интуиции игроков в XVII веке можно простить, так как $\mathsf{P}(A) - \mathsf{P}(B) = 2/216 \approx 0{,}01$. Возможно, игрок Де Мере — философ и литератор во времена правления Людовика XIV, используя свой опыт, и придерживался разумной стратегии в этой игре.

Пример 2.27 (*задача на размещение частиц в системе Максвелла–Больцмана*). Имеются k различных частиц и m различных ячеек, при этом $m \geqslant k$. В этой системе нет запрета Паули, т. е. в ячейке может не оказаться частиц, или быть одна частица, или две,..., и, наконец, все k частиц могут находиться в некоторой ячейке. Частицы могут быть, например, протоны, электроны. Ячейками могут быть энергетические уровни, объемы фазового пространства и т. д. Системе Максвелла–Больцмана, например, подчиняется обычный газ. Определить вероятность того, что:

1) в k определенных (зафиксированных) ячейках окажется по одной частице (событие A);
2) в каких-то k ячейках окажется по одной частице (событие B).

При решении этой задачи, которая играет важную роль в современной статистической физике, необходимо учитывать, что любые мыслимые здесь размещения частиц по ячейкам отличаются числом частиц в каждой из ячеек, индивидуальностью частиц и индивидуальностью ячеек. В силу этого присвоим частицам номера $1, 2, \ldots, k$ и ячейкам — номера $1, 2, \ldots, m$. После этого произвольный элементарный исход данного эксперимента целесообразно описывать вектором $\omega = (x_1, x_2, \ldots, x_k)$, в котором i-я компонента x_i определяет номер ячейки, куда попала i-я частица. Итак, $x_i \in \{1, 2, \ldots, m\}$, $i \in \{1, 2, \ldots, k\}$, а число n всех элементарных исходов очевидно равно m^k. Достоверное событие Ω для этого эксперимента имеет следующую структуру:

$$\Omega = \{\omega = (x_1, x_2, \ldots, x_k)\colon x_i = 1, 2, \ldots, m \text{ для всех } i = 1, 2, \ldots, k\} = \{\omega_1, \omega_2, \ldots, \omega_n\}.$$

При этом множество $\mathcal{F}$ всех наблюдаемых случайных событий можно представить в виде $\mathcal{F} = \{C\colon C \subset \Omega\}$. Из определения событий $A \in \mathcal{F}$ и $B \in \mathcal{F}$ легко понять, что $A = \{\omega = (x_1, x_2, \ldots, x_k)\colon x_i \neq x_j,$ $i \neq j;\ x_i, x_j \in \{c_1, c_2, \ldots, c_k\} \subset \{1, 2, \ldots, m\}\}$, где $c_1, c_2, \ldots, c_k$ суть номера заранее зафиксированных ячеек, и $B = \{\omega = (x_1, x_2, \ldots, x_k)\colon x_i \neq x_j,$ $i \neq j;\ \ x_i, x_j \in \{1, 2, \ldots, m\}\}$. Отсюда $A = \{\omega_{a_1}, \omega_{a_2}, \ldots, \omega_{a_r}\} \subset \Omega$, $B = \{\omega_{b_1}, \omega_{b_2}, \ldots, \omega_{b_s}\} \subset \Omega$, где $r = k!$ и $s = C_m^k\, k!$, и, значит, непосредственно имеем $\mathsf{P}(A) = k!/m^k$, $\mathsf{P}(B) = C_m^k\, k!/m^k$. Здесь интересно заметить, что вероятность $\mathsf{P}(A) = \mathsf{P}(B)$ при $k = m$ и событие $A = B = \varnothing$ при $k > m$.

Пример 2.28 (*задача на размещение частиц в системе Бозе–Эйнштейна*). Имеются k принципиально неразличимых частиц и m различных ячеек, при этом $m \geqslant k$. В этой системе также нет запрета Паули, т. е. в ячейке может оказаться любое число частиц от нуля до k. Частицы могут быть фотоны, пи-мезоны и т. д. Известно, например, что системе Бозе–Эйнштейна подчиняется фотонный газ. Заметим, что в отличие от системы Максвелла–Больцмана здесь рассматриваются одинаковые частицы. Определить вероятность того, что:

1) в k определенных (зафиксированных) ячейках окажется по одной частице (событие A);
2) в каких-то k ячейках окажется по одной частице (событие B).

В системе Бозе–Эйнштейна любые мыслимые здесь размещения частиц по ячейкам отличаются числом частиц в каждой из ячеек и индивидуальностью ячеек, но не индивидуальностью попавших частиц. Для наглядного представления этого эксперимента выстроим все ячейки с номерами от 1 до m в ряд вплотную друг к другу по некоторой прямой ab (см. рис. 2.4).

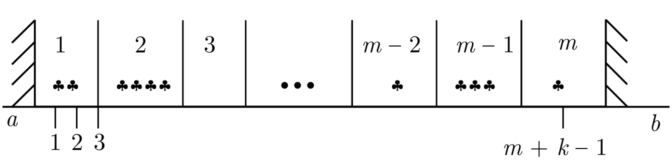


Рис. 2.4

Границы ячеек определим перегородками, среди которых две крайние жестко закреплены, а остальные $m-1$ внутренние и могут перемещаться. Расположим все частицы на этой прямой ab, перпендикулярной к указанным перегородкам. Частицы отмечены на рис. 2.4 символом «$\clubsuit$». Таким образом, на прямой ab имеются места с номерами 1, 2, ..., $m+k-1$ для всех k частиц и для всех $m-1$ внутренних перегородок. Если теперь при некотором данном размещении частиц по ячейкам поменять между собой местами любые две или более частиц, то в силу неразличимости частиц получим то же самое размещение. Точно так же перестановка между собой местами внутренних перегородок не дает нового размещения. Напротив, при перестановке местами частицы и перегородки получаем новое размещение. Более того, данное размещение частиц по ячейкам в системе Бозе–Эйнштейна однозначно определяет места на прямой ab для внутренних перегородок и, наоборот, выбранные места на прямой ab для внутренних перегородок однозначно задают размещение частиц по ячейкам. После этого произвольный элементарный исход данного эксперимента целесообразно описывать неупорядоченным множеством вида $\omega = \{y_1, y_2, \ldots, y_{m-1}\}$, в котором все его элементы различны и определяют номера мест на прямой ab для внутренних перегородок.

Таков первый способ описания элементарных исходов данного эксперимента. Достоверное событие Ω при этом имеет следующую структуру:

$$\Omega = \{\omega\colon \omega = \{y_1, y_2, \ldots, y_{m-1}\}, \quad y_i \neq y_j, \quad i \neq j;$$
$$y_i = 1, 2, \ldots, m+k-1 \text{ для всех } i = 1, 2, \ldots, m-1\}.$$

Число n всех элементарных исходов в рассматриваемом эксперименте равно числу всех $(m-1)$-элементных подмножеств множества из $(m+k-1)$ элементов: $n = C_{m+k-1}^{m-1}$. В качестве множества $\mathcal{F}$ всех наблюдаемых случайных событий для данного эксперимента, как это обычно делается в дискретных вероятностных моделях, выберем совокупность всех подмножеств множества Ω. Из определения события $A \in \mathcal{F}$ получаем, что $A = \{\omega_a = \{1, 2, \ldots, m+k-1\} \backslash \{a_1, a_2+1, \ldots, a_k+k-1\}\}$. Здесь $a_1 < a_2 < \ldots < a_k$ суть номера заранее зафиксированных ячеек и множество $\{a_1, a_2+1, \ldots, a_k+k-1\}$ задает места на прямой ab для всех k частиц. Немного труднее находим, что множество $B = \{\omega = \{1, 2, \ldots, m+k-1\} \backslash \{b_1, b_2+1, \ldots, b_k+k-1\}\colon b_i \neq b_j$ при $i \neq j$ и $i, j = 1, 2, \ldots, k$; $b_i = 1, 2, \ldots, m\} \in \mathcal{F}$. Следовательно, появлению события A благоприятствует только один элементарный исход, а наступлению события B благоприятствуют C_m^k элементарных исходов. Поэтому $\mathsf{P}(A) = 1/C_{m+k-1}^{m-1}$ и $\mathsf{P}(B) = C_m^k / C_{m+k-1}^{m-1}$. Здесь, как и в предыдущем примере, $\mathsf{P}(A) = \mathsf{P}(B)$ при $k = m$ и $A = B = \varnothing$ при $k > m$.

В заключение отметим, что очень часто в качестве описания ω произвольного элементарного исхода данного эксперимента выбирают вектор $(z_1, z_2, \ldots, z_m)$, компонента z_i которого определяет число частиц в i-й ячейке. В этом случае ясно, что $z_i \in \{0, 1, 2, \ldots, k\}$, $i \in \{1, 2, \ldots, m\}$ и $z_1 + z_2 + \ldots + z_m = k$. Достоверное событие Ω при таком выборе описания произвольного элементарного исхода имеет следующую структуру:

$$\Omega = \{\omega = (z_1, z_2, \ldots, z_m)\colon z_i = \\ = 0, 1, 2, \ldots, k \text{ для всех } i = 1, 2, \ldots, m; z_1 + z_2 + \ldots + z_m = k\}.$$

Если теперь некоторый элементарный исход имеет описание в виде вектора $(z_1, z_2, \ldots, z_m)$, то нетрудно отсюда найти единственное описание этого элементарного исхода в виде множества $\{z_1 + 1, z_1 + z_2 + 2, \ldots, z_1 + z_2 + \ldots + z_{m-1} + m - 1\} = \{y_1, y_2, \ldots, y_{m-1}\}$. При этом любым двум различным векторам $(z_1, z_2, \ldots, z_m)$ и $(z_1', z_2', \ldots, z_m')$ соответствуют различные множества вида $\{z_1 + 1, z_1 + z_2 + 2, \ldots, z_1 + z_2 + \ldots + z_{m-1} + m - 1\}$ и $\{z_1' + 1, z_1' + z_2' + 2, \ldots, z_1' + z_2' + \ldots + z_{m-1}' + m - 1\}$. Действительно, если это не имеет место, то $\{z_1 + 1, z_1 + z_2 + 2, \ldots, z_1 + z_2 + \ldots + z_{m-1} + m - 1\} = \{z_1' + 1, z_1' + z_2' + 2, \ldots, z_1' + z_2' + \ldots + z_{m-1}' + m - 1\}$. Значит, верны соотношения вида: $z_1 + 1 = z_1' + 1$, $z_1 + z_2 + 2 = z_1' + z_2' + 2, \ldots$, $z_1 + z_2 + \ldots + z_{m-1} + m - 1 = z_1' + z_2' + \ldots + z_{m-1}' + m - 1$. Отсюда и из равенств вида: $z_1 + z_2 + \ldots + z_m = k$, $z_1' + z_2' + \ldots + z_m' = k$ находим $z_1 = z_1'$, $z_2 = z_2', \ldots$, $z_m = z_m'$, а это противоречит неравенству $(z_1, z_2, \ldots, z_m) \neq (z_1', z_2', \ldots, z_m')$.

Наоборот, имея описание некоторого элементарного исхода в виде неупорядоченного множества $\{y_1, y_2, \ldots, y_{m-1}\}$, где ради простоты считаем $y_1 < y_2 < \ldots < y_{m-1}$, легко получим единственное описание этого элементарного исхода в виде вектора $(y_1 - y_0 - 1, y_2 - y_1 - 1, \ldots, y_m - y_{m-1} - 1) = (z_1, z_2, \ldots, z_m)$, где $y_0 = 0$ и $y_m = m + k$. Также показывается, что различным неупорядоченным множествам $\{y_1, y_2, \ldots, y_{m-1}\}$ и $\{y_1', y_2', \ldots, y_{m-1}'\}$ соответствуют различные векторы $(y_1 - y_0 - 1, y_2 - y_1 - 1, \ldots, y_m - y_{m-1} - 1)$ и $(y_1' - y_0' - 1, y_2' - y_1' - 1, \ldots, y_m' - y_{m-1}' - 1)$, где $y_1 < y_2 < \ldots < y_{m-1}$, $y_1' < y_2' < \ldots < y_{m-1}'$, $y_0 = y_0' = 0$ и, наконец, $y_m = y_m' = m + k$. Значит, между двумя различными описаниями элементарных исходов этого эксперимента установлено взаимно однозначное соответствие. Выбор того или иного способа описания элементарных исходов определяется кругом вопросов, на которые необходимо отвечать при рассмотрении данного эксперимента. Например, при подсчете числа всех элементарных исходов данного эксперимента целесообразно использовать первый способ описания.

Пример 2.29 (*задача на размещение частиц в системе Ферми–Дирака*). Имеются k принципиально неразличимых частиц и m различных ячеек, при этом $m \geqslant k$. В системе дополнительно действует принцип запрета Паули, т. е. в каждой ячейке может находиться не более одной частицы (см. рис. 2.5). Только этим свойством эта система отличается от системы Бозе–Эйнштейна.

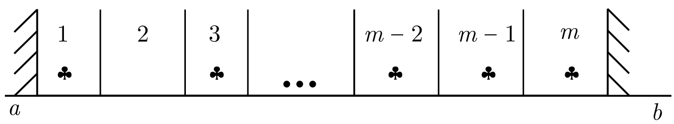


Рис. 2.5

Частицы могут быть электроны, протоны, нейтроны и т. д. Системе Ферми–Дирака подчиняется, например, электронный газ. Итак, имеем систему Бозе–Эйнштейна с ограничением Паули. Определить вероятность того, что:

1) в k определенных (зафиксированных) ячейках окажется по одной частице (событие A);
2) в каких-то k ячейках окажется по одной частице (событие B).

Как и в системе Бозе–Эйнштейна перестановка между собой местами только частиц или только внутренних перегородок не дает нового размещения. Однако при перестановке местами частицы и перегородки иногда в некоторой ячейке могут оказаться две частицы. Например, перестановка местами на рис. 2.5 первой внутренней перегородки и частицы из третьей ячейки приводит к размещению двух частиц в первой ячейке. Таким образом, для получения новых размещений в системе Ферми–Дирака требуется другой механизм.

Пусть мы имеем некоторое размещение частиц в системе Ферми–Дирака, тогда можно однозначно определить номера всех непустых ячеек. Обратно, если мы выберем любые k различных номеров ячеек из m и поместим в эти отобранные ячейки по одной частице, то в результате этого механизма однозначно получим некоторое размещение в системе Ферми–Дирака. Это размещение не зависит от порядка отобранных номеров ячеек. По этой причине произвольный элементарный исход данного эксперимента можно описать неупорядоченным множеством $\omega = \{u_1, u_2, \ldots, u_k\}$, в котором все его элементы различны и определяют номера непустых k ячеек из m возможных. При этом достоверное событие

$$\Omega = \{\omega = \{u_1, u_2, \ldots, u_k\} : u_i \neq u_j, \quad i \neq j; \\ u_i = 1, 2, \ldots, m \text{ для всех } i = 1, 2, \ldots, k\}.$$

Число n всех элементарных исходов в рассматриваемом эксперименте очевидно равно числу всех k-элементных подмножеств множества из m элементов, т. е. $n = C_m^k$. Теперь нетрудно видеть, что событие $A = \{\{c_1, c_2, \ldots, c_k\}\}$ совпадает с некоторым элементарным исходом $B = \Omega = \{\omega\colon \omega = \{u_1, u_2, \ldots, u_k\},\ u_i \neq u_j,\ i \neq j;$ $u_i = 1, 2, \ldots, m$ для всех $i = 1, 2, \ldots, k\}$ и поэтому для этой задачи имеем $\mathsf{P}(A) = 1/C_m^k = k!(m-k)!/m!$, $\mathsf{P}(B) = 1$. Здесь множество из k номеров заранее зафиксированных ячеек, в каждой из которых окажется по одной частице, обозначено через $\{c_1, c_2, \ldots, c_k\}$. Забавно отметить, что при $k = m$ система Ферми–Дирака становится детерминированной, т. е. $A = B = \Omega = \{\{1, 2, \ldots, m\}\}$, $\mathcal{F} = \{\Omega, \varnothing\}$.

Понятно, что в качестве описания ω произвольного элементарного исхода данного эксперимента можно выбрать вектор $(z_1, z_2, \ldots, z_m)$, компонента z_i которого определяет число частиц в i-й ячейке. Тогда имеем $z_i \in \{0, 1\}$, $i \in \{1, 2, \ldots, m\}$, $z_1 + z_2 + \ldots + z_m = k$ и

$$\Omega = \{\omega = (z_1, z_2, \ldots, z_m)\colon z_i = 0, 1, \quad \text{для всех}\ \ i = 1, 2, \ldots, m; \\ z_1 + z_2 + \ldots + z_m = k\}.$$

Рассуждая, как и в задаче на размещение частиц в рассмотренной системе Бозе–Эйнштейна, нетрудно установить взаимно однозначное соответствие между этими двумя различными описаниями элементарных исходов данного эксперимента.

Пример 2.30 (*задача на размещение частиц в системе Линден–Белла*). Имеются k различимых частиц и m различных ячеек, при этом $m \geqslant k$. В этой системе действует принцип запрета Паули и, значит, в каждой ячейке может находиться не более одной частицы. Итак, имеем систему Максвелла–Больцмана с ограничением Паули или систему Ферми–Дирака с принципиально различимыми частицами или систему Бозе–Эйнштейна с различными частицами и запретом Паули. Определить вероятность того, что:

1) в k зафиксированных ячейках окажется по одной частице (событие A);
2) в каких-то k ячейках окажется по одной частице (событие B).

Предполагается, что ограничениям Линден–Белла удовлетворяют элементарные объемы фазового пространства звездных систем. Так как мы имеем систему Максвелла–Больцмана с ограничением Паули, то произвольный элементарный исход данного эксперимента будем описывать вектором $\omega = (x_1, x_2, \ldots, x_k)$, в котором i-я компонента x_i определяет номер ячейки, куда попала i-я частица. Однако достоверное событие Ω для системы Линден–Белла имеет другую структуру, а именно:

$$\Omega = \{\omega = (x_1, x_2, \ldots, x_k)\colon x_i \neq x_j,\ i \neq j; \\ x_i = 1, 2, \ldots, m \text{ для всех } i = 1, 2, \ldots, k\}.$$

Из этого следует, что число n всех элементарных исходов для этого эксперимента равно числу размещений из m элементов по k, т. е. $n = A_m^k = k!\,C_m^k = m!/(m-k)! = m \times (m-1) \times \ldots \times (m-k+1)$. Из определения A и B для системы Линден–Белла легко видеть, что $A = \{\omega = (x_1, x_2, \ldots, x_k)\colon x_i \neq x_j,\ i \neq j;\ x_i \in \{c_1, c_2, \ldots, c_k\} \subset \subset \{1, 2, \ldots, m\}\}$, где $c_1, c_2, \ldots, c_k$ суть номера заранее зафиксированных ячеек, и $B = \{\omega = (x_1, x_2, \ldots, x_k)\colon x_i \neq x_j,\ i \neq j;\ x_i = 1, 2, \ldots, m$ для всех $i = 1, 2, \ldots, k\}$, т. е. $B = \Omega$. Следовательно, событию A благоприятствует появление $r = k!$ элементарных исходов данного эксперимента, а событию B благоприятствует появление $s = k!\,C_m^k$ элементарных исходов. Отсюда $\mathsf{P}(A) = 1/C_m^k$, $\mathsf{P}(B) = 1$. Найденные вероятности совпадают с аналогичными вероятностями в системе Ферми–Дирака. Отметим, что в задаче Линден–Белла $\mathsf{P}(A) = \mathsf{P}(B) = 1$ при $k = m$ и $A = B = \varnothing$ при $k > m$. Однако, в отличие от эксперимента Ферми–Дирака, система Линден–Белла при $k = m$ не является детерминированной.

Пример 2.31 (*задача о днях рождения*). Пусть происходит случайный выбор k людей в Нижнем Новгороде для прослушивания популярной лекции по теории вероятностей. Найти вероятность того, что, по крайней мере, у двух слушателей дни рождения совпадают (событие A).

Прежде всего, уточним этот эксперимент. Годы не одинаковы по продолжительности, а рождаемость в течение года не остается постоянной. Однако в достаточно хорошем приближении можно полагать, что в году 365 дней, и день рождения каждого слушателя равновозможен в один из этих дней. На самом деле это означает, что случайный выбор k людей и случайный выбор k дней рождения приводят к одним и тем же результатам. Для каждого $i = 1, 2, \ldots, k$ обозначим через x_i день рождения i-го слушателя, где $x_i \in \{1, 2, \ldots, 365\}$. Тогда произвольному элементарному исходу данного эксперимента можно поставить в соответствие описание ω в виде вектора $(x_1, x_2, \ldots, x_k)$ и, следовательно, достоверное событие

$$\Omega = \{\omega = (x_1, x_2, \ldots, x_k)\colon x_i = 1, 2, \ldots, 365 \ \text{ для всех } \ i = 1, 2, \ldots, k\}$$

содержит $n = (365)^k$ элементов. Если $\overline{A} = \Omega \backslash A$ есть событие, которое заключается в том, что ни один день рождения не встречается более одного раза, т. е. все k дней рождения различны, то

$\overline{A} = \{\omega = (x_1, x_2, \ldots, x_k)\colon x_i \neq x_j,\ i \neq j;\ x_i = 1, 2, \ldots, 365$ для $i = 1, 2, \ldots, k\}$, $A = \Omega \backslash \overline{A}$, $\aleph(A) = (365)^k - 365 \times 364 \times \ldots \times (365 - k + 1)$ и получаем $\mathsf{P}(\overline{A}) = 365 \times 364 \times \ldots \times (365 - k + 1)/(365)^k$, $\mathsf{P}(A) = ((365)^k - 365 \times 364 \times \ldots \times (365 - k + 1))/(365)^k = 1 - \mathsf{P}(\overline{A})$. Приведем теперь значения вероятности $\mathsf{P}(A)$ для некоторых k (см. табл. 2.3).

Таблица 2.3

k	4	10	16	20	22	23	30	40	50	60	64
$\mathsf{P}(A)$	0,016	0,117	0,284	0,411	0,476	0,507	0,706	0,891	0,970	0,994	0,997

Из табл. 2.3 видно, что при числе слушателей больше тридцати вероятность совпадения дней рождения довольно близка к единице. Такой вывод на интуитивном уровне, как правило, не ожидается, так как группа из 30 слушателей существенно меньше 366 человек. Заметим, что для числа слушателей, большего числа 365, по меньшей мере у двоих совпадают дни рождения, и в этом случае $\mathsf{P}(A) = 1$.

Пример 2.32 (*задача о выборе престижной квартиры*). Среди m квартир некоторого жилищно-строительного кооператива имеется k престижных квартир. В распределении всех квартир принимают участие $r < m$ членов жилищно-строительного кооператива. Оставшиеся $m - r$ квартир передаются в распоряжение городской администрации. Члены кооператива по очереди один за другим вытаскивают один билет, на котором указан номер его будущей квартиры. Вычислить вероятность того, что s-й из очереди член кооператива вытащит престижную квартиру.

Обозначим через x_1 номер квартиры, который вытащил первый из очереди член кооператива; через x_2 номер квартиры, который вытащил второй из очереди член кооператива;... и, наконец, через x_r номер квартиры, который вытащил последний из очереди член кооператива. Тогда произвольный элементарный исход этого эксперимента будем описывать вектором $\omega = (x_1, x_2, \ldots, x_r)$, а достоверное событие $\Omega = \{\omega\colon \omega = (x_1, x_2, \ldots, x_r),\ x_i \neq x_j,\ i \neq j;\ x_i = 1, 2, \ldots, m\}$. Отсюда находим, что число n всех элементарных исходов для эксперимента равно числу размещений из m элементов по r, т. е. $n = A_m^r = m \times (m-1) \times \ldots \times (m - r + 1)$. Пусть для определенности престижные квартиры имеют номера $1, 2, \ldots, k$. Если событие A заключается в том, что s-й из очереди член кооператива вытащит билет с номером престижной квартиры, то $A = \{\omega = (x_1, x_2, \ldots, x_{s-1}, x_s, x_{s+1}, \ldots, x_r)\colon x_i \neq x_j,\ i \neq j;\ x_i = 1, 2, \ldots, m;\ x_s = 1, 2, \ldots, k\}$. Событие A содержит $k \times (m-1) \times (m-2) \times \ldots \times (m - 1 - (r-1) + 1) = k(m-1)(m-2) \times \ldots \times (m - r + 1)$ описаний всех тех элементарных исходов, которые происходят с этим событием. Поэтому

$\mathsf{P}(A) = [k(m-1)(m-2) \times \ldots \times (m-r+1)] \times [m(m-1)(m-2) \times \times \ldots \times (m-r+1)]^{-1} = k/m$. Итак, вероятность вытащить билет с номером престижной квартиры не зависит от номера s очереди члена кооператива.

В заключение этого пункта важно заметить, что понятие равновероятности (равновозможности) элементарных исходов является первоначальным и интуитивным понятием. Это понятие основывается на симметрии опыта, без знания при этом численных значений вероятностей событий. Затем на основе этого предлагается простой способ вычисления вероятностей случайных событий.

Итак, при классическом определении вероятности в случае конечного числа возможных элементарных исходов E априори предполагается симметричность эксперимента E по отношению к таким исходам. Поэтому всегда остается открытым вопрос о симметричности, или однородности, или равновозможности. Наконец, возникает вопрос о существовании вероятностей в случае несимметричного эксперимента, например, в опыте с броском игральной кости при заведомо ее неоднородности.

§ 3. Вычисление вероятности для испытаний с несчетным числом равновозможных исходов

3.1. Геометрический подход. Классическое определение вероятности основано на рассмотрении симметричного эксперимента E, в котором выделена полная группа конечного числа n равновозможных (равновероятных) несовместных элементарных событий. При этом для любого случайного события $A = \{\omega_{a_1}, \omega_{a_2}, \ldots, \omega_{a_m}\} = \{\omega_{a_1}\} \cup \cup \{\omega_{a_2}\} \cup \ldots \cup \{\omega_{a_m}\} \in \mathcal{F}$, которому благоприятствуют $m \leqslant n$ элементарных событий, вероятность $\mathsf{P}(A) = m/n = \mathsf{P}(\{\omega_{a_1}\}) + \mathsf{P}(\{\omega_{a_2}\}) + \ldots + \mathsf{P}(\{\omega_{a_m}\})$. Однако имеется большое число экспериментов, в которых мыслимо бесконечное (несчетное, недискретное) множество элементарных исходов. В этом случае по-прежнему большую роль играет понятие симметрии эксперимента и понятие о равновозможности его выбранных элементарных исходов. Тем не менее, каждому элементарному исходу $\{\omega\}$ из $\mathfrak{J}'$ уже нельзя приписать одну и ту же положительную вероятность. Это можно объяснить следующими простыми рассуждениями. Произвольный случайный исход A, вообще говоря, может наблюдаться с бесконечным числом различных элементарных событий. Другими словами, случайное событие A может быть представлено в виде объединения бесконечного числа элементарных событий вида $\{\omega\}$. Поэтому вероятность $\mathsf{P}(A)$ такого случайного исхода нельзя вычислить суммированием вероятностей $\mathsf{P}(\{\omega\})$ тех элементарных исходов, которые могут наступать одновременно с исходом A. Для такого рода симметричных экспериментов положительная вероятность уже задается для некоторых случайных событий, которые не совпадают ни с одним из элементарных исходов. Переходим теперь непосредственно к понятию геометрической вероятности. Ради определенности и простоты рассмотрим последовательно следующие три случая.

Геометрическая вероятность на вещественной прямой. Пусть реальный эксперимент E таков, что каждый выбранный его элементарный исход можно описать абсциссой x некоторой точки из заданного промежутка G на действительной прямой R. При этом допускается, что длина промежутка G, обозначаемая через $\operatorname{mes}(G)$, существует и $\operatorname{mes}(G) \neq 0$. Наряду с экспериментом E будем рассматривать также ему адекватный модельный эксперимент E_m, связанный с непреднамеренным выбором точки из промежутка G действительной прямой R. Эта адекватность такова, что случайный выбор любого элементарного исхода эксперимента E с описанием $\omega = x$ взаимно однозначно соответствует непреднамеренному выбору точки из промежутка G с абсциссой x. Термин «непреднамеренный выбор» на содержательном уровне означает, что точка ставится случайным образом, причем ни одному положению внутри промежутка G не отдается предпочтение. Этим обеспечивается симметрия экспериментов E и E_m. Более того, наблюдаемый исход A эксперимента E и наблюдаемый исход A_m модельного эксперимента E_m взаимно однозначно соответствуют друг другу, если каждый из этих исходов представлен в виде одного и того же множества $g \subset G$. Другими словами, при наблюдении случайного события A мы наблюдаем случайное событие A_m, и наоборот. Предполагается, что длина $\operatorname{mes}(g)$ промежутка g существует. В этом случае по определению естественно принять, что для эксперимента E достоверный исход $\Omega = \{\omega = x\colon x \in G\} \subset R$, случайное событие $A = \{\omega = x\colon x \in g\}$, $\mathcal{F} = \{g\colon g \subset G,\ \operatorname{mes}(g)\ \text{существует}\}$ и вероятность $\mathsf{P}(A) = \operatorname{mes}(g)/\operatorname{mes}(G)$. Например, $\mathsf{P}(\Omega) = \operatorname{mes}(G)/\operatorname{mes}(G) = 1$.

Геометрическая вероятность на плоскости. Очень часто элементарный исход реального эксперимента E удобно описать в виде упорядоченной пары $(x,\, y)$, которая определяет некоторую точку из заданной области G на плоскости $x0y$. Предполагается, что область G имеет площадь, которая обозначается через $\operatorname{mes}(G)$, и $\operatorname{mes}(G) \neq 0$. По аналогии с первым случаем, во втором случае вместо реального эксперимента E мы рассматриваем соответствующий ему модельный эксперимент E_m, который заключается в случайном выборе точки из заданной области G. Пусть R^2 — плоскость, $G \subset R^2$, $g \subset G$ и площадь $\operatorname{mes}(g)$ области g существует. Тогда по определению принимаем, что для эксперимента E достоверное событие $\Omega = \{\omega = (x,\, y)\colon (x,\, y) \in G\} \subset R^2$, $\mathcal{F} = \{g\colon g \subset G, \operatorname{mes}(g)\ \text{существует}\}$, событие $A = \{\omega = (x,\, y)\colon (x,\, y) \in g\}$ и $\mathsf{P}(A) = \operatorname{mes}(g)/\operatorname{mes}(G)$.

Геометрическая вероятность и трехмерное пространство. Пусть теперь элементарный исход эксперимента E можно описать в виде упорядоченной тройки $(x,\, y,\, z)$, определяющей некоторую точку из трехмерной области G пространства R^3. Предполагается, что объем трехмерной области G, обозначаемый через $\mathrm{mes}\,(G)$, существует и $\mathrm{mes}\,(G) \neq 0$. По аналогии с первым и вторым случаем здесь вместо реального эксперимента E мы рассматриваем адекватный ему модельный эксперимент E_m, который заключается в произвольном выборе точки из трехмерной области G пространства R^3. Пусть теперь $g \subset G$ и объем $\mathrm{mes}\,(g)$ трехмерной области g существует. Не меняя рассуждения рассмотренных первых двух случаев, выводим, что для эксперимента E произвольный элементарный исход $A' = \{\omega = (x, y, z)\}$, достоверный исход $\Omega = \{\omega = (x, y, z)\colon (x, y, z) \in G\} \subset R^3$, произвольное случайное событие $A = \{\omega = (x, y, z)\colon (x, y, z) \in g\}$, $\mathcal{F} = \{g\colon g \subset G,$ $\mathrm{mes}\,(g)$ существует$\}$ и вероятность $\mathsf{P}(A) = \mathrm{mes}\,(g)/\mathrm{mes}\,(G)$.

Совершенно таким же образом можно рассмотреть общий случай геометрического подхода при определении вероятности случайных событий, когда элементарный исход эксперимента E описывается в виде упорядоченной системы $\omega = (x_1, x_2, \ldots, x_n)$ из n действительных чисел. Наподобие того, как мы пользовались понятием длины промежутка в R, площади плоской фигуры в R^2 и объема пространственного тела в R^3, в основе общего случая определения геометрической вероятности лежит понятие объема n-мерной области в евклидовом пространстве R^n. Следует заметить, что не существует методики, которая позволяла бы рационально использовать тот или иной случай геометрического подхода к вычислению вероятности случайных событий реальных экспериментов. Как правило, это определяется опытностью и научной зрелостью исследователя.

Итак, мы приходим к следующему определению геометрической вероятности. Пусть имеем некоторую область G описаний всех элементарных исходов, которая может быть представлена для каждого конкретного эксперимента E в виде:

1) некоторого промежутка на действительной оси, имеющего некоторую длину, например, отрезка $G = [a, b]$;
2) некоторой геометрической фигуры на плоскости, имеющей площадь, например, в виде замкнутого прямоугольника $G = \{(x, y)\colon a \leqslant x \leqslant b,\ c \leqslant y \leqslant d\}$;
3) некоторой трехмерной фигуры в пространстве, имеющей объем, например, замкнутого прямоугольного параллелепипеда $G = \{(x, y, z)\colon a \leqslant x \leqslant b,\ c \leqslant y \leqslant d,\ u \leqslant z \leqslant v\}$ и т. д. (см. рис. 2.6).

Здесь a, b, c, d, u, v — некоторые действительные числа.

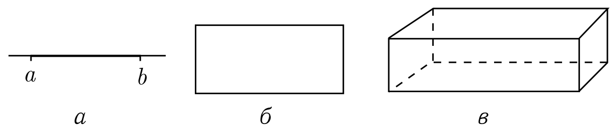


Рис. 2.6

Вместо реального эксперимента E будем теперь рассматривать адекватный ему модельный эксперимент E_m. Эксперимент E_m заключается в том, что в область G наудачу бросается точка. Термин «наудачу» означает, что точка ставится случайным образом, причем ни одному положению внутри области G не отдается предпочтение.

Определение 2.2. *Вероятность события* A реального эксперимента E с несчетным числом равновозможных исходов по определению равна вероятности попадания случайной точки внутрь некоторой области $g \subset G$ для модельного опыта E_m. Эта вероятность равна отношению меры области g к размеру всей области G, т. е. вероятность $\mathsf{P}(A) = \text{mes}(g)/\text{mes}\,(G)$.

Еще раз отметим, что здесь A есть такой исход, множество описаний которого совпадает с областью g. Появление события A реального эксперимента E адекватно выбору точки в области g для модельного эксперимента E_m. Рассмотрим простые примеры на геометрический подход вычисления вероятности.

Пример 2.33. В некоторой точке K электрического провода MN длиной L произошел разрыв (см. рис. 2.7). Определить вероятность того, что точка K удалена от точки M на расстояние не менее l.

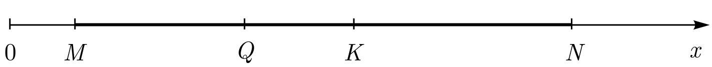


Рис. 2.7

Решение. Пусть $\Omega = \{\omega = x\colon a \leqslant x \leqslant b\}$, $G = \{x\colon a \leqslant x \leqslant b\}$, где x есть расстояние от некоторой начальной точки до точки K электрического провода (точки случайного обрыва). Можно также сказать, что x есть абсцисса точки K, выбранной случайным образом на отрезке $[a, b]$ оси $0x$. Здесь абсциссы точек M, Q и N соответственно равны a, c и b, где $b - a = L$ и $c - a = l$. Если событие $A = \{\omega = x\colon c < x \leqslant b\}$, то имеем промежуток вида $g = \{x\colon c < x \leqslant b\}$ и $\text{mes}\,(g) = L - l$, $\text{mes}\,(G) = L$. При геометрическом определении вероятности предполагается, что вероятность попадания точки K в какой угодно промежуток из отрезка $\{x\colon a \leqslant x \leqslant b\}$ не зависит от положения этого промежутка, а зависит лишь от его длины. Поэтому вероятность $\mathsf{P}(A) = \text{mes}\,(g)/\text{mes}\,(G) = (L - l)/L = 1 - l \cdot L^{-1}$.

Пример 2.34. Два парохода должны подойти к одному и тому же причалу, причем время прихода каждого из пароходов равновозможно в течение суток. Определить вероятность того, что ни одному из пароходов не придется ожидать освобождения причала, если время стоянки первого парохода составляет два часа, а второго — три часа.

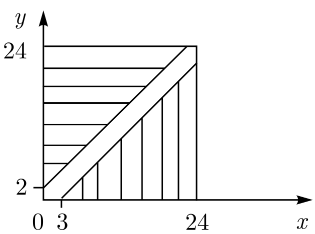


Рис. 2.8

Решение. Обозначим через $\omega = (x, y)$ описание такого элементарного исхода этого эксперимента, когда время прихода первого парохода равно x и время прихода второго парохода равно y. При таком выборе элементарных исходов и их описаний достоверный исход $\Omega = \{(x, y)\colon 0 \leqslant x, y \leqslant 24\}$. Можно, например, рассмотреть модельную и эквивалентную задачи бросания точки случайным образом в квадрат $G = \{(x, y)\colon 0 \leqslant x, y \leqslant 24\}$. Здесь уже x — абсцисса, а y — ордината точки (рис. 2.8). Далее, чтобы ни одному пароходу не пришлось ожидать для себя освобождения причала, необходимо и достаточно выполнение некоторых условий от чисел x и y. Во-первых, $y - x > 2$, если раньше пришел первый пароход, или $x - y > 3$, если раньше пришел второй пароход. Отсюда получаем ограничение вида $y > 2 + x$, $y > x$ или ограничение $y < x - 3$, $y < x$. Значит, область $g = \{(x, y)\colon 0 \leqslant x, y \leqslant 24,\ y > 2 + x,\ y > x\} \cup \{(x, y)\colon 0 \leqslant x,\ y \leqslant 24,$ $y < x - 3,\ y < x\}$ составлена из двух прямоугольных треугольников. Пусть событие A означает, что ни одному пароходу не пришлось ожидать освобождения причала. Отсюда получаем, что искомая вероятность случайного события A равна $\mathsf{P}(A) = \mathrm{mes}\,(g)/\mathrm{mes}\,(G) =$ $= (22 \times 22/2 + 21 \times 21/2)/(24 \times 24) = 462{,}5/576 \approx 0{,}8$.

Пример 2.35. Лодка перевозит груз с одного берега пролива на другой за один час. Какова вероятность того, что идущее вдоль пролива судно будет замечено (событие A), если с лодки обнаруживают судно в случае, когда пересекают его курс не ранее, чем за 20 мин до пересечения судном курса лодки, и не позднее, чем через 20 мин после пересечения судном курса лодки? Любой момент и любое место пересечения судном курса лодки равновозможны. Курс судна перпендикулярен курсу лодки.

Решение. Пусть t_1 есть время в минутах, которое затратит лодка от момента начала движения до момента пересечения судном курса лодки, а t_2 означает время в минутах, которое затратит лодка, чтобы доплыть в точку пересечения курса судна. Пространство описаний элементарных исходов (достоверное событие) изучаемого случайного эксперимента E можно представить в следующем виде: $\Omega = \{\omega = (t_1, t_2)\colon 0 \leqslant t_1 \leqslant 60,\ 0 \leqslant t_2 \leqslant 60\}$. Тогда событие $A = \{(t_1, t_2) \in \Omega\colon |t_1 - t_2| < 20\}$. На рис. 2.9 множество Ω помечено в виде квадрата, а случайное событие A представлено в виде заштрихованного шестиугольника. Следовательно, искомая вероятность $\mathsf{P}(A) = 3600/3600 - 40 \times \times 40/3600 = 5/9$.

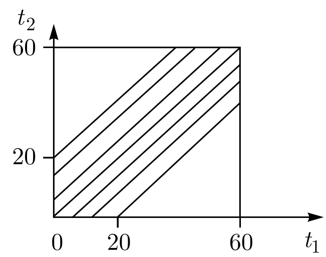


Рис. 2.9

Пример 2.36. Определить вероятность подрыва корабля (случайное событие A) при форсировании минного заграждения, если якорные контактные мины поставлены в один ряд через интервал $2l$, а курс корабля с линией мин составляет угол φ. Пересечение кораблем линии мин предполагается равновозможным в любой точке. Ширина корабля равна $2h$, $l > h$, диаметр мин равен $2r$ (см. рис. 2.10).

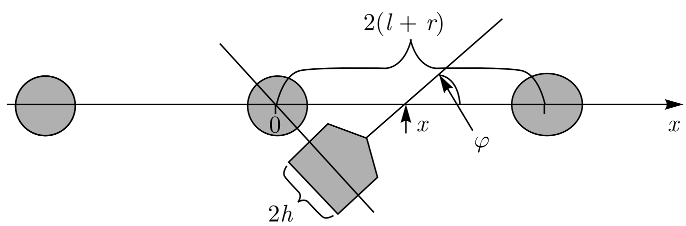


Рис. 2.10

Решение. В этом эксперименте случайной является точка пересечения курса корабля линии мин. Проведем через линию мин ось $0x$. Пусть $x = \omega$ есть координата точки пересечения курса корабля и линии мин. Тогда множество Ω описаний всех элементарных исходов данного эксперимента равно $\{\omega = x\colon 0 \leqslant x \leqslant 2(l+r)\}$. Так как событие $\overline{A} = \{\omega = x\colon (h+r)/\sin\varphi < x < 2l + 2r - (h+r)/\sin\varphi\}$, то событие $A = \{\omega = x\colon 0 \leqslant x \leqslant (h+r)/\sin\varphi\} \cup \{\omega = x\colon 2(l+r) - (h+r)/\sin\varphi \leqslant x \leqslant 2(l+r)\}$. Используя геометрический способ определения вероятности, непосредственно получим, что вероятность $\mathsf{P}(A) = [(h+r)/\sin\varphi + (h+r)/\sin\varphi]\cdot 2^{-1}(l+r)^{-1} = (h+r)/(l+r)\sin\varphi$ при $h + r \leqslant (l+r)\sin\varphi$, отсюда $\mathsf{P}(A) = 1$ при $h + r > (l+r)\sin\varphi$.

3.2. Примеры на построение адекватных вероятностных моделей для реальных экспериментов. Если проведение некоторого числа испытаний над реальным статистически устойчивым экспериментом E требует значительных временны́х и денежных затрат, то для изучения такого рода экспериментов, как правило, необходимо последовательное составление нескольких адекватных вероятностных моделей. Из этой последовательности построенных моделей выделяется наиболее простая, которая поддается исследованию методами теории вероятностей и математической статистики. Далее, результаты исследования такой модели интерпретируются и сопоставляются применительно к остальным моделям и, естественно, к исходному эксперименту E. Рассмотрим решение ряда такого сорта задач.

Пример 2.37 (*задача о гончарном круге*). Рассматривается кустарное производство глиняной посуды мастером с использованием цилиндрического гончарного круга радиуса r. В результате работы на верхней поверхности гончарного круга наудачу (случайным образом) образуются трещины типа хорды. Так как крепление гончарного круга при производстве посуды проходит через его центр, то трещины не могут через него проходить. Если длина трещины не превосходит $\sqrt{3}\,r$, то мастер может ее заделать и вновь использовать этот гончарный круг. При длине трещины больше $\sqrt{3}\,r$ мастер должен заменить этот гончарный круг новым (событие A). В противном случае он будет изготовлять посуду с браком. Чему равна вероятность того, что мастер воспользуется новым гончарным кругом? Так как верхняя поверхность гончарного круга представляет собой круг радиуса r, то вполне очевидно, что реальному эксперименту E можно поставить во взаимно однозначное соответствие первый модельный эксперимент E_{Γ} или хорошо известную задачу Бертрана (см. п. В.6). Наудачу выбирается хорда в круге радиуса r. Найти вероятность того, что ее длина превосходит $r\sqrt{3}$ — длину стороны вписанного в этот круг равностороннего треугольника abc (событие A_{Γ}).

Приведем несколько подходов решения задачи.

Решение. Первый подход. Предположим, что в плоскости данного круга выбрано некоторое направление $0y$, например, вертикальное (см. рис. 2.11). Точка b и точка O на рис. 2.11, *а* называются полюсами закрепления при изготовлении гончарного круга. Пусть хорды круга наудачу проводятся перпендикулярно к фиксированному направлению $0y$. Тогда произвольным элементарным исходом в эксперименте E_{Γ} является положение наудачу выбранной хорды.

Легко видеть, что положение хорды однозначно определяет ее расстояние до точки O, и наоборот. Поэтому в качестве описания элементарного исхода в эксперименте E_{Γ} естественно предложить расстояние x от точки O до выбранной хорды. Теперь вместо эксперимента E_{Γ} можно рассмотреть ему адекватный модельный эксперимент E_m, связанный с непреднамеренным выбором точки (см. рис. 2.11, *б*) из промежутка вида $G = \{x\colon 0 \leqslant x \leqslant 2r,\ x \neq r\}$, на оси $0x$. В этом случае

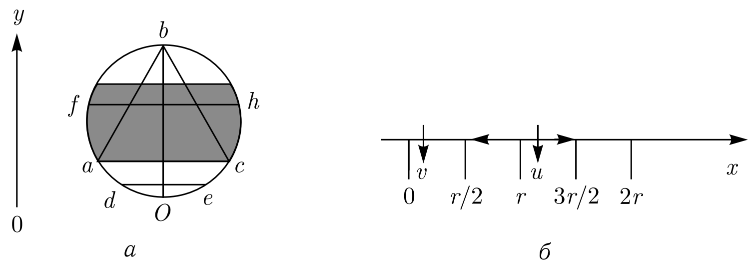
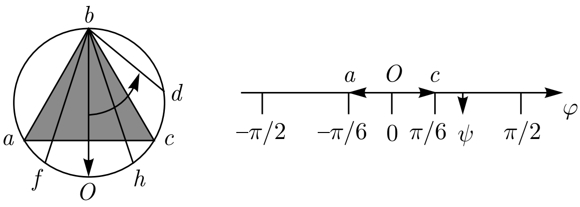


Рис. 2.11

для эксперимента E_Γ достоверное событие $\Omega_\Gamma = \{\omega = x\colon 0 \leqslant x \leqslant 2r,$ $x \neq r\}$, $\mathcal{F}_\Gamma = \{g\colon g \subset G,\ g$ имеет длину$\}$. Напомним, что длина стороны треугольника abc равна $\sqrt{3}\,r$. Например, хорда de с расстоянием v ($v < r/2$) до точки O имеет длину меньше, чем длина стороны треугольника abc. Напротив, хорда fh с расстоянием, равным u ($r < u < 3r/2$), имеет длину больше, чем длина стороны треугольника abc. Так как случайное событие $A_\Gamma = \{\omega = x\colon r/2 < x < 3r/2,$ $x \neq r\}$, то из определения геометрической вероятности на прямой легко получим, что $\mathsf{P}(A_\Gamma) = \operatorname{mes}((r/2, 3r/2))/\operatorname{mes}([0,2r]) = 1/2$. При наступлении случайного события A наступает случайное событие A_Γ, и наоборот. Поэтому для мастера вероятность замены гончарного круга новым равна 1/2. Заметим, что при технологии изготовлении гончарного круга, когда используются две закрепленные точки (см. рис. 2.11, a), направление возникшей трещины почти всегда перпендикулярно прямой Ob.

Второй подход. Закрепим один из концов хорды рис. 2.12, a, например, в точке b, а другой конец хорды непреднамеренно (наудачу) выбирается на окружности. Точка b называется полюсом гончарного круга. При этом в плоскости данного круга теперь удобнее выбрать направление, определяемое лучом bO. В этом случае в качестве описания ω положения наудачу выбранной хорды (элементарного исхода) в эксперименте E_Γ используем угол φ между лучом bO и хордой, измеряемый в радианах.

Угол φ будет положительным числом, если направление стрелки против часовой стрелки, и отрицательным в противном случае. Например, положение хорды bd, длина которой меньше, чем длина стороны треугольника abc, определяет угол ψ ($\pi/6 < \psi < \pi/2$). Напротив, каждая из хорд bh и bf имеет длину больше, чем длина стороны указанного треугольника. При этом отметим, что угол между лучом bO и хордой bf является отрицательным числом. Очевидно, что при втором подходе эксперименту E_Γ взаимно однозначно соответствует несколько другой по сравнению с первым случаем модельный



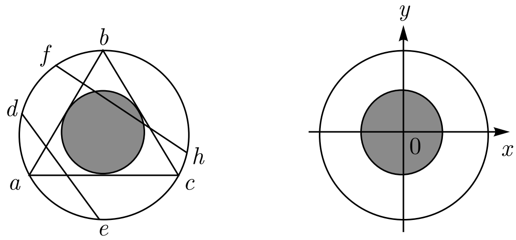
а *б*

Рис. 2.12

эксперимент E_m. Теперь эксперимент E_m заключается в том, что непреднамеренно выбирается точка (см. рис. 2.12, *б*) из промежутка $G = \{\omega = \varphi\colon -\pi/2 \leqslant \varphi \leqslant \pi/2,\ \varphi \neq 0\}$ действительной оси 0φ. Отсюда для эксперимента E_Γ легко получим, что достоверное событие $\Omega_\Gamma = \{\omega = \varphi\colon -\pi/2 \leqslant \varphi \leqslant \pi/2, \varphi \neq 0\}$, $\mathcal{F}_\Gamma = \{g\colon g \subset G,\ g$ имеет длину$\}$. Если описание $\omega = \varphi$ выбранной хорды удовлетворяет условию вида $-\pi/6 < \varphi < \pi/6$ и $\varphi \neq 0$, то ее длина будет больше величины $\sqrt{3}\,r$. Поэтому получим случайное событие $A_\Gamma = \{\omega = \varphi\colon -\pi/6 < \varphi < \pi/6, \varphi \neq 0\}$ и по определению геометрической вероятности на действительной прямой уже имеем другое значение вероятности $\mathsf{P}(A_\Gamma)$, а именно, $\mathsf{P}(A_\Gamma) = \mathrm{mes}\,((-\pi/6, \pi/6))/\mathrm{mes}\,([-\pi/2, \pi/2]) = 1/3$. Значит, вероятность замены гончарного круга равна 1/3. При технологии изготовления гончарного круга, когда используется одна точка b закрепления, трещины, как правило, появляются таким образом, как это показано на рис. 2.12, *а*.

Третий подход. Будем теперь наудачу проводить хорду, не закрепляя ни один из ее концов и не обращая внимания на выбранные направления оси $0y$ или луча bO при первых двух подходах. При этом хорда не может проходить через центр круга. Примеры случайного выбора положения хорд de и fh показаны на рис. 2.13, *а*. Это соответствует той ситуации, когда центр круга является единственной точкой закрепления при изготовлении гончарного круга. В этих ограничениях центр наудачу поставленной хорды однозначно определяется ее положением и наоборот. Поэтому в качестве описания положения наудачу выбранной хорды (элементарного исхода) в эксперименте E_Γ целесообразно использовать вектор $(x,\ y)$. Здесь x и y являются абсциссой и, соответственно, ординатой центра хорды в прямоугольной системе координат с началом в центре круга радиуса r. При этом плоскости прямоугольной системы координат $x0y$ и круга совпадают.

Ради удобства прямоугольная система координат показана отдельно на рис. 2.13, *б*. Тогда наряду с экспериментом E_Γ можно рассмотреть адекватный ему модельный эксперимент E_m, связанный со случайным



а *б*

Рис. 2.13

выбором точки из области $G = \{(x, y)\colon x^2 + y^2 \leqslant r^2,\ (x, y) \neq (0, 0)\}$ на плоскости $x0y$. Если в эксперименте E_m точка выбрана внутри заштрихованного круга радиуса $r/2$, то соответствующая этому выбору хорда в эксперименте E_Γ будет иметь длину больше $\sqrt{3}\, r$. В противном случае хорда в эксперименте E_Γ будет иметь длину не больше $\sqrt{3}\, r$. Например, из рис. 2.13, *а* видно, что длина хорды de меньше $\sqrt{3}\, r$, а длина хорды fh больше $\sqrt{3}\, r$. Принимая это во внимание, легко получим, что для эксперимента E_Γ достоверное событие $\Omega_\Gamma = \{(x, y)\colon x^2 + y^2 \leqslant r^2,$ $(x, y) \neq (0, 0)\}$, σ-алгебра $\mathcal{F}_\Gamma = \{g\colon g \subset G,\ g$ имеет площадь$\}$ и, наконец, случайное событие $A_\Gamma = \{(x, y)\colon x^2 + y^2 \leqslant r^2/4,\ (x, y) \in G\}$. Так как площадь области $\{(x, y)\colon x^2 + y^2 \leqslant r^2/4,\ (x, y) \neq (0, 0)\} = \pi r^2/4$ и площадь области $G = \pi r^2$, то, применяя определение геометрической вероятности на плоскости, получим, что вероятность $\mathsf{P}(A_\Gamma) = 1/4$. При наступлении случайного события A наступает случайное событие A_Γ, и наоборот. Поэтому для мастера вероятность замены гончарного круга новым в третьем случае равна $1/4$.

В задаче Бертрана на самом деле рассматриваются три различных эксперимента E_Γ, каждый из которых задается с помощью более детального описания случайного механизма выбора хорды. Поэтому вероятности получились различными. В реальной задаче с гончарным кругом (экспериментом E) легко понять появление разного рода трещин первого, второго или третьего типа в зависимости от первой, второй или третьей технологии изготовления круга и от материала, из которого он сделан. Забавно отметить, что древнегреческие купцы назначали цены C_1, C_2, C_3 для разного типа гончарных кругов в пропорции $C_1 : C_2 : C_3 = 2 : 3 : 4 = (1/2)^{-1} : (1/3)^{-1} : (1/4)^{-1}$. Так решается известный парадокс Бертрана.

Пример 2.38 (*задача Бюффона*). Рассмотрим теперь так называемую задачу Бюффона (рис. 2.14). На поверхности стола проведен ряд параллельных прямых, отстоящих друг от друга на расстоянии $2h$. На стол случайным образом бросают иглу, длина которой равна $2l$. Предполагается, что $l < h$. Эта ситуация условно показана на рис. 2.14, *а*.

Какова вероятность того, что игла пересекает одну из параллельных прямых?

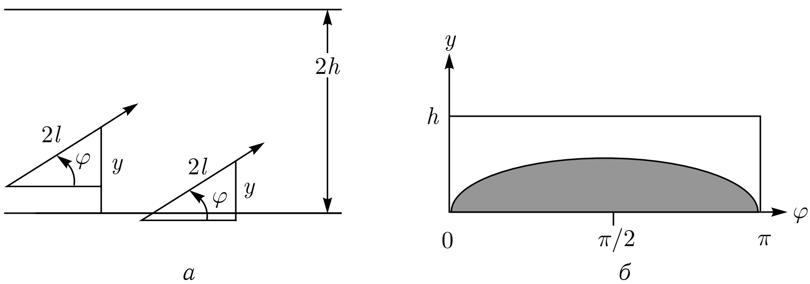


Рис. 2.14

Решение. На рис. 2.14, *а* приведены два случая, когда игла после броска пересекает какую-нибудь прямую и когда игла не пересекает ни одну из параллельных прямых. Нетрудно видеть, что положение иглы на поверхности стола после ее случайного броска полностью определяет расстояние y от центра иглы до ближайшей параллели и угол φ, составленный иглой с этой параллелью и отсчитываемый против хода часовой стрелки от параллели к игле, и, наоборот, с учетом параллельного сдвига. Тогда в качестве описания положения иглы (элементарного исхода) эксперимента Бюффона выберем упорядоченную пару чисел (φ, y), где $0 \leqslant \varphi \leqslant \pi$ и $0 \leqslant y \leqslant h$. Следовательно, в данной задаче реальный эксперимент E с бросанием иглы на поверхность стола, которая разграфлена параллельными прямыми, заменяется адекватным модельным экспериментом E_m с выбором наудачу точки в прямоугольнике вида $\{\omega = (\varphi, y)\colon 0 \leqslant \varphi \leqslant \pi,\ 0 \leqslant y \leqslant h\}$ на плоскости $\varphi 0 y$. При этом элементарный исход эксперимента Бюффона и соответствующий элементарный исход модельного эксперимента E_m имеют одно и то же описание (φ, y). Отсюда непосредственно получаем, что $\Omega = \{\omega = (\varphi, y)\colon 0 \leqslant \varphi \leqslant \pi,\ 0 \leqslant y \leqslant h\}$, область $G = \{(\varphi, y)\colon 0 \leqslant \varphi \leqslant \pi$ и $0 \leqslant y \leqslant h\}$ и $\mathcal{F} = \{g\colon g \subset G,\ g$ имеет площадь$\}$. Из рис. 2.14, *а* ясно, что игла пересечет одну из прямых или наступит событие A в опыте Бюффона только в случае выполнения неравенства вида $y \leqslant l \cdot \sin\varphi$. Тогда легко видеть, что $A = \{\omega = (\varphi, y)\colon 0 \leqslant \varphi \leqslant \pi,$ $0 \leqslant y \leqslant h,\ y \leqslant l \cdot \sin\varphi\}$. В модельном эксперименте E_m естественно рассматривается событие A_m, которое заключается в том, что точка будет наудачу выбрана в заштрихованной области рис. 2.14, *б*. При этом между событием A реального опыта Бюффона и событием A_m модельного эксперимента E_m установлено взаимно однозначное соответствие, так как эти события имеют одинаковое описание в виде множества (области) $g = \{(\varphi, y)\colon 0 \leqslant \varphi \leqslant \pi,\ 0 \leqslant y \leqslant h,\ y \leqslant l \cdot \sin\varphi\}$.

Отметим, что области G и g приведены на рис. 2.14, *б*. Площадь области $g = \int_0^\pi l \cdot \sin\varphi\, d\varphi = -l \cdot \cos\varphi\big|_0^\pi = 2l$, а площадь области G равна πh. Из определения геометрической вероятности на плоскости с учетом рис. 2.14, *б* получаем, что вероятность $\mathsf{P}(A)$ равна отношению заштрихованной площади к площади прямоугольника. Следовательно, вероятность $\mathsf{P}(A) = 2l\,(\pi h)^{-1}$.

§ 4. Эмпирический и аксиоматический подходы к определению вероятности случайных событий

4.1. Свойства относительных частот исходов эксперимента и статистическое определение вероятности. Одной из фундаментальных проблем теории вероятностей является построение математических моделей таких реальных экспериментов, которые могут давать различные результаты при одних и тех же условиях их проведения, и, следовательно, эти результаты не могут быть точно предсказаны заранее. Поэтому при использовании той или иной теоретико-вероятностной модели всегда возникает вопрос о том, насколько эта модель адекватна реальному эксперименту или опыту. На последний вопрос можно ответить только путем проведения наблюдений над исходами реальных экспериментов. К тому же на практике часто встречается случай, когда невозможно применить ни субъективный, ни классический, ни геометрический способы определения вероятности. Ведется стрельба из пистолета по плоской мишени. Фиксируется факт попадания пули в мишень. При этом стрелок старается попасть в центр мишени. Тогда вероятность попадания в некоторую область мишени зависит от площади области и ее положения.

В связи с этим многие естествоиспытатели определяют вероятность только эмпирическим путем или путем многократного проведения эксперимента, т. е. используют так называемый статистический подход в определении вероятности. В основу этой концепции положена стабильность относительной частоты события (отношения числа испытаний, в которых событие появилось, к числу всех проведенных испытаний) при достаточно большом числе опытов. Это, в свою очередь, является следствием понятия на содержательном уровне статистически устойчивого эксперимента. В дальнейшем относительную частоту события будем называть просто частотой события.

Пусть произведена серия из N опытов эксперимента E в неизменных условиях Σ, в каждом из которых может появиться или не появиться случайное событие $A \in \mathcal{F}$. Обозначим через $\mu(A, N)$ число опытов, в которых наблюдалось событие A при общем числе N проведенных испытаний. Тогда по определению *частотой события* A или *статистической вероятностью* называется отношение $\mu(A,N)/N$, которое обозначается $\mathsf{P}^*(A,N)$ или $\mathsf{P}^*(A)$, опуская ради простоты символ N, где это не вызывает путаницы.

При небольшом числе испытаний статистическая вероятность $\mathsf{P}^*(A)$ события A носит случайный нерегулярный характер и может заметно изменяться от опыта к опыту. Однако по мере увеличения числа испытаний частота события начинает терять случайный характер и стремится стабилизироваться около некоторого постоянного числа. Это число характеризует связь между комплексом условий Σ и наблюдаемым событием A. Из определения относительной частоты события легко показать простейшие ее свойства. Например, получаем, что $0 \leqslant \mathsf{P}^*(A) \leqslant 1$. Частота достоверного события Ω есть $\mathsf{P}^*(\Omega) = N/N = 1$ и частота невозможного события $\varnothing$ есть $\mathsf{P}^*(\varnothing) = 0/N = 0$. Более того, $\mathsf{P}^*(A \cup B) = \mathsf{P}^*(A) + \mathsf{P}^*(B)$, если только случайные события A и B являются несовместными.

Статистическое определение вероятности случайного события $A \in \mathcal{F}$ не есть формальное математическое определение его вероятности, а оно лишь постулирует при некоторых условиях ее существование и указывает метод ее приближенного вычисления. На содержательном уровне вероятность случайного события — это число $\mathsf{P}(A)$, около которого группируются частоты этого события по мере увеличения числа испытаний. В этом и заключается основной тезис при эмпирическом подходе к вычислению вероятности случайного события A. Основоположниками этого подхода являются Ричард фон Мизес и Ганс Рейхенбах. В качестве примера приведем опыт с бросанием английским математиком К. Пирсоном симметричной монеты $N = 12\,000$ раз. При этом герб выпал $6\,019$ раз. В этом случае отношение $\mu(A,N)/N = 0{,}5016$, где исход A означает выпадение герба. Значит, К. Пирсон получил эмпирическую или статистическую вероятность $\mathsf{P}^*(A) = 0{,}5016$. Заметим, что применение классического подхода к вычислению вероятности выпадение герба симметричной монеты дает число $\mathsf{P}(A) = 0{,}5$. Итак, для этого примера имеем $\mathsf{P}^*(A) \approx \mathsf{P}(A)$.

Статистическое понятие вероятности послужило созданию универсального вычислительного метода в современной компьютерной математике. Этот метод называется методом статистических испытаний или методом Монте-Карло. В свою очередь метод Монте-Карло послужил в дальнейшем одним из первоисточников при создании более универсального метода изучения свойств реальных явлений — метода имитационного моделирования реальных процессов. Проиллюстрируем этот метод на задаче Бюффона (пример 2.38). Покажем, что опыт Бюффона позволяет приближенно вычислить значение числа π. Действительно, проведем этот опыт достаточное число раз. Это число обозначим через N. Пусть при N бросках иглы на разграфленную поверхность в $\mu(A,N)$ опытах она пересекала прямые, иначе говоря, мы $\mu(A,N)$ раз наблюдали событие A. Эмпирически полученная статистическая вероятность события A будет равна $\mu(A,N)/N$. Из приближенного равенства $\mu(A,N)/N \approx 2l/(\pi h)$ находим приближенное значение числа π, а именно, $\pi \approx 2lN/(h \cdot \mu(A,N))$. В опытах, проведенных Вольфом в 1849 г., длина иглы составляла $2l = 36$ мм, а ширина полос была равна величине $2h = 45$ мм. При общем числе бросков $N = 5\,000$ игла пересекла полосы $\mu(A,N) = 2\,532$ раза, так что $\mathsf{P}(A) \approx 2532/5000 \approx$ $\approx 0{,}5064$. Отсюда значение $\pi = 2l/(h\,\mathsf{P}(A)) = 72/(45 \times 0{,}5064) \approx 3{,}159$. Отметим простоту и доступность такого способа определения числа π. Однако для более точного вычисления π необходимо выполнить большое число испытаний. Такого рода испытания можно проводить на современных компьютерах. Поэтому большое число испытаний, как правило, не является существенным ограничением. Большой вклад в развитие способа математических вычислений с помощью случайного эксперимента внесли Ричард фон Мизес и Ганс Рейхенбах.

4.2. Система аксиом Колмогорова и выбор адекватной вероятностной модели априорных экспериментов. В реальных экспериментах, например в военном деле, часто требуется определить вероятности таких исходов, для которых не применимы не только субъективный, классический и геометрический способы, но и статистический подход из-за чрезвычайной сложности и большой стоимости соответствующих экспериментов. В подобных случаях вероятности событий должны определяться исключительно из структуры самого априори заданного эксперимента E с помощью аксиом и теорем безотносительно к тому, проводится или не проводится так называемый априорный эксперимент. Таким образом, мы приходим к построению абстрактной теории вероятностей, которая должна быть адекватной моделью изучаемых реальных явлений случайного типа.

Итак, рассматривается статистически устойчивый случайный эксперимент E, который определяется комплексом условий $\Sigma = \{u_1, u_2, \ldots, u_s, \ldots\}$ и множеством $\mathfrak{J}$ его допустимых исходов. Пусть построена теоретико-множественная модель $(\Omega, \mathcal{F})$ эксперимента E, в которой Ω — множество описаний ω всех его элементарных исходов $\{\omega\}$ и $\mathcal{F}$ — множество наблюдаемых случайных событий $A \subset \Omega$. При этом $\mathcal{F} \neq \varnothing$ и является σ-алгеброй. Множество A содержит, по меньшей мере, все такие описания $\omega \in \Omega$ элементарных исходов эксперимента E, которые могут наблюдаться одновременно с A. Если множество $\mathcal{F}$ наблюдаемых событий образует σ-алгебру, то справедливы следующие ограничения:

1) из $A \in \mathcal{F}$ следует $\overline{A} \in \mathcal{F}$;

2) из $A_1 \in \mathcal{F}$, $A_2 \in \mathcal{F}$,... следует $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$. Тогда вероятность любого события $A \in \mathcal{F}$ вводится с помощью следующего определения.

Определение 2.3. *Вероятностью* любого события $A \in \mathcal{F}$ называется некоторое неотрицательное число $\mathsf{P}(A)$, которое для любой последовательности $A_1, A_2, \ldots$ попарно непересекающихся случайных событий из $\mathcal{F}$ удовлетворяет аксиоме счетной аддитивности вида $\mathsf{P}\big(\cup_{i=1}^{\infty} A_i\big) = \sum_{i=1}^{\infty} \mathsf{P}(A_i)$, а также аксиоме нормировки вида $\mathsf{P}(\Omega) = 1$.

Таким образом, вероятность $\mathsf{P}(A)$ есть неотрицательная, счетно-аддитивная и нормированная функция событий A из множества $\mathcal{F}$. Легко видеть, что три замечательные аксиомы Колмогорова для вероятности $\mathsf{P}(A)$ с областью определения $\mathcal{F} \subset \mathfrak{J}$ повторяют свойства относительных частот исходов реального эксперимента.

Определение 2.4. Упорядоченная тройка $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, состоящая из множества Ω описаний всех выбранных элементарных исходов эксперимента E, из выделенной σ-алгебры $\mathcal{F}$ случайных событий $A \subset \Omega$ и определенной на $\mathcal{F}$ неотрицательной, нормированной и счетно-аддитивной функции $\mathsf{P}(\cdot)$, называется *основным вероятностным пространством* или *вероятностной моделью априорного эксперимента E* (*Дуб Дж. Л.* Вероятностные процессы. — М.: ИЛ, 1956).

Следует заметить, что аксиомы Колмогорова и упорядоченная тройка $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ непротиворечивы и неполны. Эти аксиомы непротиворечивы, так как существуют реальные эксперименты, например классические эксперименты, удовлетворяющие всем этим аксиомам. Неполнота системы аксиом означает неоднозначность выбора теоретико-множественной модели $(\Omega, \mathcal{F})$ и вероятности $\mathsf{P}(\cdot)$ на множестве $\mathcal{F}$. Это вызвано тем, что для реальных экспериментов могут встретиться одинаковые теоретико-множественные модели $(\Omega, \mathcal{F})$ с различными вероятностными функциями $\mathsf{P}(\cdot)$, $A \in \mathcal{F}$.

Например, существуют игральные кости различной симметрии. Более того, для одного и того же эксперимента мы можем по-разному выбирать семейство элементарных исходов и их описания. То, что система аксиом не является полной, скорее есть положительный факт, чем отрицательный. В силу этого можно ставить вопрос о построении в некотором смысле оптимальной или более адекватной модели эксперимента E. В этом заключается принципиальная особенность и универсальность подхода Колмогорова.

Пример 2.39. *Проиллюстрируем сказанное на задаче Даламбера* (пример 2.7). В этом опыте достоверный исход $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, множество всех результатов $\mathfrak{J} = \mathcal{F} = \{A_1, A_2, \ldots, A_{16}\}$. Следовательно, в этом эксперименте можно наблюдать пятнадцать исходов. Напомним, что $\omega_1 = (0, 0)$ — описание выпадения герба при первом и втором бросках, $\omega_2 = (0, 1)$ — описание выпадения герба при первом броске и решетки при втором, $\omega_3 = (1, 0)$ — описание выпадения решетки при первом броске и герба при втором, $\omega_4 = (1, 1)$ — описание исхода решетки при первом и втором бросках. При этом случайные события из множества $\mathcal{F}$ представляются в следующем виде:

$$A_1 = \{\omega_1\}, \quad A_2 = \{\omega_2\}, \quad A_3 = \{\omega_3\}, \quad A_4 = \{\omega_4\}, \quad A_5 = \{\omega_1, \omega_2\},$$
$$A_6 = \{\omega_1, \omega_3\}, \quad A_7 = \{\omega_1, \omega_4\}, \quad A_8 = \{\omega_2, \omega_3\}, \quad A_9 = \{\omega_2, \omega_4\},$$
$$A_{10} = \{\omega_3, \omega_4\}, \quad A_{11} = \{\omega_1, \omega_2, \omega_3\}, \quad A_{12} = \{\omega_1, \omega_2, \omega_4\},$$
$$A_{13} = \{\omega_1, \omega_3, \omega_4\},\ A_{14} = \{\omega_2, \omega_3, \omega_4\},\ A_{15} = \{\omega_1, \omega_2, \omega_3, \omega_4\},\ A_{16} = \varnothing.$$

В силу симметрии эксперимента Даламбера вероятность $\mathsf{P}(A)$, $A \in \mathcal{F}$, можно вычислить, используя классический подход, например, $\mathsf{P}(A_1) = \mathsf{P}(A_2) = \mathsf{P}(A_3) = \mathsf{P}(A_4) = 1/4$, $\mathsf{P}(A_5) = 1/2$ и т. д. Таким образом, хотя бы для одной реальной задачи (задачи Даламбера) построили основное вероятностное пространство $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, для которого справедливы все аксиомы множества $\mathcal{F}$ и все аксиомы вероятности $\mathsf{P}(\cdot)$. Эти утверждения легко проверяются, если воспользоваться конкретным видом $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$.

Рассмотрим теперь другой способ выбора элементарных исходов, а именно пусть ω'_1 означает описание выпадения герба при первом броске, ω'_2 означает описание выпадения герба только при втором броске и ω'_3 означает описание невыпадения герба при двух бросаниях симметричной монеты. При таком выборе имеем

$$\Omega' = \{\omega'_1, \omega'_2, \omega'_3\}, \qquad \mathcal{F}' = \{B_1, B_2, \ldots, B_8\},$$

где

$$B_1 = \{\omega'_1\}, \quad B_2 = \{\omega'_2\}, \quad B_3 = \{\omega'_3\}, \quad B_4 = \{\omega'_1, \omega'_2\}, \quad B_5 = \{\omega'_1, \omega'_3\},$$
$$B_6 = \{\omega'_2, \omega'_3\}, \quad B_7 = \{\omega'_1, \omega'_2, \omega'_3\} = \Omega', \quad B_8 = \varnothing,$$

откуда

$$\mathsf{P}'(B_1) = 1/2, \quad \mathsf{P}'(B_2) = 1/4, \quad \mathsf{P}'(B_3) = 1/4,$$
$$\mathsf{P}'(B_4) = \mathsf{P}'(\{\omega'_1\}) + \mathsf{P}'(\{\omega'_2\}) = 3/4$$

и т. д., в результате чего получаем новую вероятностную модель $(\Omega', \mathcal{F}', \mathsf{P}'(\cdot))$ задачи Даламбера. В этой модели мы уже не можем наблюдать восемь исходов. Поэтому эта модель менее адекватна по сравнению с первой, в частности, мы не можем интересоваться вероятностью того, что орел выпадет два раза.

Наконец, выберем теперь несимметричную монету. Например, выберем монету, которая некоторое время была в употреблении в игре в орлянку. Тогда для первой построенной теоретико-множественной модели $(\Omega, \mathcal{F})$ может оказаться, что

$$\mathsf{P}''(A_1) = 1/9, \quad \mathsf{P}''(A_2) = \mathsf{P}''(A_3) = 2/9, \quad \mathsf{P}''(A_4) = 4/9,$$
$$\mathsf{P}''(A_5) = \mathsf{P}''(\{\omega_1\}) + \mathsf{P}''(\{\omega_2\}) = \mathsf{P}''(A_1) + \mathsf{P}''(A_2) = 3/9 = 1/3$$

и т. д. В этом несимметричном случае мы получаем третью вероятностную модель, которая отличается от первой только выбором другой вероятности $\mathsf{P}''(\cdot)$ для каждого события.

Заметим, что для любого классического эксперимента E, когда множество $\mathfrak{J}$ содержит конечное число случайных равновероятных элементарных событий $\{\omega_1\}$, $\{\omega_2\}, \ldots, \{\omega_n\}$, имеет место аксиоматика Колмогорова для функции $\mathsf{P}(\cdot)$. Если, например, $A = \{\omega_{a_1}, \omega_{a_2}, \ldots, \omega_{a_m}\} \in \mathcal{F}$, то вероятность $\mathsf{P}(A) = m/n \geqslant 0$ и $\mathsf{P}(\Omega) = n/n = 1$. Итак, свойства неотрицательности и нормировки для функции $\mathsf{P}(\cdot)$ выполняются. Пусть дополнительно $B = \{\omega_{b_1}, \omega_{b_2}, \ldots, \omega_{b_k}\} \in \mathcal{F}$, а события A и B несовместны. Поэтому случайное событие $A \cup B = \{\omega_{a_1}, \omega_{a_2}, \ldots, \omega_{a_m}, \omega_{b_1}, \omega_{b_2}, \ldots, \omega_{b_k}\}$ и $\mathsf{P}(A \cup B) = (m+k)/n = (m/n) + (k/n) = \mathsf{P}(A) + \mathsf{P}(B)$. Аналогичное равенство можно показать для любого конечного числа попарно непересекающихся событий, т. е. имеет место аксиома конечной аддитивности.

4.3. Простейшие свойства вероятностной функции Колмогорова. Построение вероятностного пространства $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ является основным этапом в создании математической модели того или иного статистически устойчивого случайного эксперимента E. Однако не меньшую роль играет выяснение основных свойств вероятности $\mathsf{P}(A)$, $A \in \mathcal{F}$, которые непосредственно следуют из системы аксиом Колмогорова. Перейдем к формулировке и доказательству основных свойств вероятности $\mathsf{P}(A)$, $A \in \mathcal{F}$.

Теорема 2.1. *Имеет место неравенство* $1 \geqslant \mathsf{P}(A) \geqslant 0$ *для всех* $A \in \mathcal{F}$, *другими словами, область значений вероятностной функции* $\mathsf{P}(A)$, $A \in \mathcal{F}$, *есть отрезок* $[0,\ 1]$.

Доказательство. Из определения противоположных событий A и $\overline{A}$ имеем равенства: $\overline{A} \cap A = \varnothing$, $\overline{A} \cup A = \Omega$. Отсюда, используя аксиомы нормировки и счетной аддитивности, сразу находим $\mathsf{P}(\overline{A}) + \mathsf{P}(A) = 1$ или $\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A)$, и, значит, на основании аксиомы неотрицательности будет $1 - \mathsf{P}(A) \geqslant 0$ или $1 \geqslant \mathsf{P}(A)$. Итак, окончательно имеем $1 \geqslant \mathsf{P}(A) \geqslant 0$. Теперь, если в качестве A взять достоверное событие Ω, как следствие этой теоремы получим $\mathsf{P}(\varnothing) = 0$. Здесь также отметим важность полученной формулы $\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A)$, которая впервые встречалась в этой главе и была установлена в конкретной задаче о днях рождения. Эта формула будет использоваться на протяжении всего курса.

Теорема 2.2. *Если* $A \subset B$, $A \in \mathcal{F}$, $B \in \mathcal{F}$, *то* $\mathsf{P}(B) \geqslant \mathsf{P}(A)$ (*свойство монотонности вероятностной функции* $\mathsf{P}(A)\colon \mathcal{F} \to [0,\ 1]$).

Доказательство. При $A \subset B$ выполняется равенство в событиях вида: $B = A \cup (B \backslash A)$. Более того, имеем $A \cap (B \backslash A) = \varnothing$. Равенство в событиях, как правило, доказывается по следующей схеме. Пусть $\omega \in B$, тогда либо $\omega \in A$ и, следовательно, $\omega \in A \cup (B \backslash A)$; либо $\omega \notin A$ и поэтому $\omega \in B \backslash A$. Отсюда непосредственно находим $\omega \in A \cup (B \backslash A)$ и, следовательно, событие $B \subset A \cup (B \backslash A)$. Наоборот, если $\omega \in A \cup (B \backslash A)$, то либо $\omega \in A$ и, так как $A \subset B$, имеем $\omega \in B$; либо $\omega \in B \backslash A$ и из определения разности событий имеем $\omega \in B$, т. е. $A \cup (B \backslash A) \subset B$. На основании определения равенства событий сразу получаем, что событие $B = A \cup (B \backslash A)$. То, что событие $A \cap (B \backslash A) = \varnothing$, т. е. указанные события A и $B \backslash A$ не пересекаются, будем доказывать методом от противного. Если $A \cap (B \backslash A) \neq \varnothing$, то существует хотя бы одно такое $\omega_0 \in \Omega$, что $\omega_0 \in A$ и одновременно $\omega_0 \in B \backslash A$. Из определения разности событий B, A находим $\omega_0 \in B$ и $\omega_0 \notin A$. Итак, получаем противоречивые утверждения $\omega_0 \in A$ и $\omega_0 \notin A$.

Теперь из доказанного равенства $B = A \cup (B \backslash A)$ получаем на основании аксиомы счетной аддитивности $\mathsf{P}(B) = \mathsf{P}(A) + \mathsf{P}(B \backslash A)$, откуда $\mathsf{P}(B) \geqslant \mathsf{P}(A)$, поскольку $\mathsf{P}(B \backslash A) \geqslant 0$ на основании аксиомы неотрицательности. Отсюда также можно получить важное равенство $\mathsf{P}(B \backslash A) = \mathsf{P}(B) - \mathsf{P}(A)$ при $A \subset B$.

Теорема 2.3. *Если* $A \in \mathcal{F}$, $B \in \mathcal{F}$, *то имеет место теорема сложения вероятностей вида*: $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$.

Доказательство. Имеем следующие равенства в событиях:

$$\begin{gathered} A \cup B = A \cup (B \backslash A), \quad A \cap (B \backslash A) = \varnothing, \\ B = (A \cap B) \cup (B \backslash A), \quad (A \cap B) \cap (B \backslash A) = \varnothing. \end{gathered} \tag{2.2}$$

Для установления этих равенств можно воспользоваться либо методом, приведенным в доказательстве теоремы 2.2, либо основными законами теоретико-множественных операций над случайными событиями. Например, первые два из этих равенств можно установить с помощью тождественных преобразований:

$$A \cup (B \backslash A) = A \cup (B \cap \overline{A}) = (A \cup B) \cap (A \cup \overline{A}) = (A \cup B) \cap \Omega = A \cup B,$$
$$A \cap (B \backslash A) = A \cap (B \cap \overline{A}) = \varnothing.$$

Из соотношений (2.2), используя аксиому счетной аддитивности, находим равенства в вероятностях:

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B \backslash A), \qquad \mathsf{P}(B) = \mathsf{P}(A \cap B) + \mathsf{P}(B \backslash A).$$

Вычитая эти равенства и перенося затем $\mathsf{P}(B)$ в правую часть, окончательно для любых двух случайных событий из $\mathcal{F}$ получим теорему сложения вероятностей вида: $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$.

Используя теорему 2.3 и законы теоретико-множественных операций над событиями, легко показать теорему сложения для трех событий A, B, C. Например, последовательно имеем:

$$\begin{aligned}
&\mathsf{P}(A \cup B \cup C) = \mathsf{P}(A \cup B) + \mathsf{P}(C) - \mathsf{P}((A \cup B) \cap C) = \\
&= \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(A \cap B) - \mathsf{P}((A \cap C) \cup (B \cap C)) = \\
&= \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(A \cap B) - \mathsf{P}(A \cap C) - \mathsf{P}(B \cap C) + \mathsf{P}(A \cap B \cap C).
\end{aligned}$$

Так как теорема сложения доказана для двух и трех событий, то методом математической индукции легко получить формулу для любого конечного числа событий. Итак, для любого конечного числа событий A_1, $A_2, \ldots, A_m$ имеет место общее равенство:

$$\begin{aligned}
&\mathsf{P}\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{m} \mathsf{P}(A_i) - \sum_{1 \leqslant i_1 < i_2 \leqslant m}^{m} \mathsf{P}(A_{i_1} \cap A_{i_2}) + \\
&+ \sum_{1 \leqslant i_1 < i_2 < i_3 \leqslant m}^{m} \mathsf{P}(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \ldots + (-1)^{m-1} \mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_m).
\end{aligned}$$

В целях применения этого общего равенства решим следующую задачу.

Пример 2.40 (*задача о туристах*). Туристам из m человек в гостинице выделили одноместные комнаты с номерами $1, 2, \ldots, m$. Любая комната открывается только одним ключом, номер которого совпадает с номером такой комнаты. Каждый из туристов случайным образом взял из ящика дежурного стола только один ключ. Найти вероятность того, что хотя бы один турист откроет комнату.

Решение. Пусть x_i — номер выбранного ключа туристом, которому выделили комнату с номером i. Случайное событие, которое состоит в том, что турист из первой комнаты отберет ключ с номером x_1, турист из второй комнаты отберет ключ с номером x_2, ..., наконец, турист из m-й комнаты отберет ключ с номером x_m, объявляем элементарным с описанием ω в виде вектора $(x_1, x_2, \ldots, x_m)$. В этом случае достоверное событие

$$\Omega = \{\omega = (x_1, x_2, \ldots, x_m) \colon x_1 \neq x_2 \neq \ldots \neq x_m,\ x_i = 1, 2, \ldots, m\},$$

$\aleph(\Omega) = m!$ и $\mathcal{F} = \{A\colon A \subset \Omega\}$. Для каждого фиксированного $i = 1, 2, \ldots, m$ обозначим через $A_i = \{\omega = (x_1, x_2, \ldots, x_m)\colon x_1 \neq x_2 \neq \ldots \neq x_m,\ x_i = i,\ x_j = 1, 2, \ldots, m,\ j \neq i\}$ случайное событие, которое заключается в том, что турист из i-й комнаты отобрал свой ключ. Множество A_i содержит все перестановки из $(m-1)$ элементов. Следовательно, число $\aleph(A_i)$ равно $(m-1)!$ и оно не зависит от i. Для любых различных чисел i_1, $i_2, \ldots$, i_r из $\{1, 2, \ldots, m\}$ рассмотрим событие $A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_r} = \{\omega = (x_1, x_2, \ldots, x_m)\colon x_1 \neq x_2 \neq \ldots \neq x_m;\ x_{i_1} = i_1,\ x_{i_2} = i_2, \ldots,\ x_{i_r} = i_r;\ x_j = 1, 2, \ldots, m;\ j \notin \{i_1, i_2, \ldots, i_r\}\}$, которое состоит в том, что туристы из комнат с номерами i_1, $i_2, \ldots$, i_r отобрали свои ключи. Множество $A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_r}$ содержит все перестановки из $(m-r)$ элементов. Поэтому число $\aleph(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_r})$ равно $(m-r)!$ и оно не зависит от набора чисел i_1, $i_2, \ldots$, i_r. Используя классическое определение вероятности, непосредственно найдем

$$\mathsf{P}(A_i) = \frac{(m-1)!}{m!},$$
$$\mathsf{P}(A_{i_1} \cap A_{j_2}) = \frac{(m-2)!}{m!},\ \ldots,\ \mathsf{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_r}) = \frac{(m-r)!}{m!}.$$

Если теперь событие A состоит в том, что хотя бы один турист откроет выделенную ему комнату, то $A = \cup_{i=1}^{m} A_i$ и по общей формуле легко находим

$$\mathsf{P}\left(\bigcup_{i=1}^{m} A_i\right) = \\ = C_m^1 \frac{(m-1)!}{m!} - C_m^2 \frac{(m-2)!}{m!} + C_m^3 \frac{(m-3)!}{m!} - \ldots + (-1)^{m-1} C_m^m \frac{1}{m!} = \\ = 1 - (1/2!) + (1/3!) - \ldots + (-1)^{m-1} (m!)^{-1}.$$

Заметим, что вероятность $\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A) = (1/2!) - (1/3!) + (1/4!) - \ldots + (-1)^m (m!)^{-1}$. Здесь $\overline{A}$ — событие, состоящее в том, что ни один турист не выбрал нужный ему ключ. Так как $e^{-1} = (1/2!) - (1/3!) + (1/4!) - (1/5!) + \ldots$, то при больших значениях m вероятность $\mathsf{P}(A) \approx 1 - e^{-1}$ и вероятность $\mathsf{P}(\overline{A}) \approx e^{-1}$. Например, для $m = 6$ с использованием точных формул получим, что значение вероятности вида $\mathsf{P}(A) = 91/144 \approx 0{,}631944$ и значение вероятности $\mathsf{P}(\overline{A}) = 53/144 \approx 0{,}368055$. Простые вычисления по приближенным формулам приводят к достаточно близкому результату, а именно, вероятность $\mathsf{P}(A) \approx 0{,}632121$ и вероятность $\mathsf{P}(\overline{A}) \approx 0{,}367879$.

Из теоремы сложения для двух и трех событий следуют неравенства: $\mathsf{P}(A \cup B) \leqslant \mathsf{P}(A) + \mathsf{P}(B)$ и $\mathsf{P}(A \cup B \cup C) \leqslant \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C)$, так как $\mathsf{P}(A \cap B) \geqslant \mathsf{P}(A \cap B \cap C) \geqslant 0$ и $\mathsf{P}(A \cap C) \geqslant 0$, $\mathsf{P}(B \cap C) \geqslant 0$.

Можно доказать общее неравенство

$$\mathsf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} \mathsf{P}(A_i), \tag{2.3}$$

если $A_i \in \mathcal{F}$ для всех $i = 1, 2, \ldots$. В самом деле, пусть $B_1 = A_1$ и $B_i = \overline{A}_1 \cap \overline{A}_2 \cap \ldots \cap \overline{A}_{i-1} \cap A_i$ при $i = 2, 3, \ldots$. Из определения B_1, $B_2, \ldots$ видно, что они попарно не пересекаются, удовлетворяют соотношению $B_i \subset A_i$ для каждого $i = 1, 2, \ldots$ и, наконец, $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} B_i$. Поэтому легко находим, что $\mathsf{P}\bigl(\cup_{i=1}^{\infty} A_i\bigr) = \mathsf{P}\bigl(\cup_{i=1}^{\infty} B_i\bigr) = \sum_{i=1}^{\infty} \mathsf{P}(B_i) \leqslant \leqslant \sum_{i=1}^{\infty} \mathsf{P}(A_i)$.

4.4. Предел последовательности случайных событий и аксиомы непрерывности. Пусть $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ — вероятностное пространство. Рассмотрим любую последовательность A_1, $A_2, \ldots$ случайных событий из $\mathcal{F}$. Эту последовательность будем также обозначать через $\{A_i; i \geqslant 1\}$.

Определение 2.5. *Верхним пределом* $\overline{\lim\limits_{n\to\infty}} A_n$, или $\limsup\limits_{n\to\infty} A_n$ *последовательности* $\{A_i; i \geqslant 1\}$ случайных событий из $\mathcal{F}$ называется такое событие $A^* \in \mathcal{F}$, что $A^* = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$, т. е. A^* — событие, которое состоит только из тех $\omega \in \Omega$, которые принадлежат бесконечному числу событий A_k.

На содержательном уровне событие A^* состоит в том, что произойдет бесконечно много событий из последовательности A_1, $A_2, \ldots$.

Определение 2.6. *Нижним пределом* $\underline{\lim\limits_{n\to\infty}} A_n$, или иначе $\liminf\limits_{n\to\infty} A_n$ *последовательности* $\{A_i; i \geqslant 1\}$ случайных событий из $\mathcal{F}$ называется такое событие $A_* \in \mathcal{F}$, что $A_* = \cup_{i=1}^{\infty} \cap_{k=n}^{\infty} A_k$, т. е. A_* — событие, которое состоит из тех $\omega \in \Omega$, которые принадлежат всем событиям A_1, $A_2, \ldots$, за исключением, может быть, конечного числа.

Другими словами, событие A_* наступит, если произойдут все события A_1, $A_2, \ldots$, за исключением, быть может, только конечного числа.

Лемма 2.1. *Имеет место соотношение вида $A_* \subset A^*$.*

Доказательство. Если $\omega \in A_*$, то существует такое натуральное число N, что элемент $\omega \in A_N \cap A_{N+1} \cap \ldots$. Отсюда получим, что $\omega \in A_N$, $\omega \in A_{N+1}, \ldots$. Тогда $\omega \in \cup_{k=1}^{\infty} A_k$, $\omega \in \cup_{k=2}^{\infty} A_k, \ldots$ и, следовательно, $\omega \in A^*$, т. е. $A_* \subset A^*$.

Обратное утверждение неверно.

Определение 2.7. Если верхний и нижний пределы последовательности $\{A_i; i \geqslant 1\}$ совпадают, т. е. $A_* = A^*$, то говорят, что эта *последовательность событий из $\mathcal{F}$ имеет предел*. В этом случае будем писать $A = \lim\limits_{n\to\infty} A_n$.

Теорема 2.4. *Вероятностная функция* $\mathsf{P}(\cdot)\colon \mathcal{F} \to [0,1]$ *непрерывна снизу, т. е.* $\lim\limits_{n\to\infty} \mathsf{P}(A_n) = \mathsf{P}\bigl(\lim\limits_{n\to\infty} A_n\bigr) = \mathsf{P}\bigl(\cap_{n=1}^{\infty} A_n\bigr)$ *для любой последовательности* $\{A_i; i \geqslant 1\}$ *случайных событий из* $\mathcal{F}$ *таких, что* $A_1 \supset A_2 \supset \ldots$.

Доказательство этого утверждения проведем в два этапа. Сначала покажем, что предел последовательности $\{A_i; i = 1, 2, \ldots\}$ существует и $\lim\limits_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n = A \in \mathcal{F}$. С этой целью вычислим A_* и A^*. Для событий A_* и A^*, непосредственно учитывая $A_1 \supset A_2 \supset \ldots$, находим

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = (A_1 \cap A_2 \cap \ldots) \cup (A_2 \cap A_3 \cap \ldots) \cup \ldots =$$
$$= (A_1 \cap A_2 \cap \ldots) \cup (A_1 \cap A_2 \cap \ldots) \cup \ldots = \bigcap_{n=1}^{\infty} A_n = A_*,$$

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = (A_1 \cup A_2 \cup \ldots) \cap (A_2 \cup A_3 \cup \ldots) \cap \ldots =$$
$$= (A_1 \cap A_2 \cap \ldots) = \bigcap_{n=1}^{\infty} A_n = A^*.$$

Отсюда $A_* = A^*$ и, следовательно, $\lim\limits_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n = A$.

На втором этапе покажем, что $\lim\limits_{n\to\infty} \mathsf{P}(A_n) = \mathsf{P}\bigl(\lim\limits_{n\to\infty} A_n\bigr)$. В самом деле, любое событие A_n при $A_1 \supset A_2 \supset \ldots$ можно представить в следующем виде: $A_n = \cup_{k=n}^{\infty}(A_k \cap \overline{A}_{k+1}) \cup \bigl(\cap_{k=1}^{\infty} A_k\bigr)$. Легко видеть, что события $\cap_{k=1}^{\infty} A_k$ и $\cup_{k=n}^{\infty}(A_k \cap \overline{A}_{k+1})$ не пересекаются. События $A_1 \cap \overline{A}_2$, $A_2 \cap \overline{A}_3, \ldots$ также являются попарно непересекающимися. Тогда справедливы равенства:

$$\mathsf{P}(A_1) = \sum_{k=1}^{\infty} \mathsf{P}(A_k \cap \overline{A}_{k+1}) + \mathsf{P}\biggl(\bigcap_{k=1}^{\infty} A_k\biggr),$$
$$\ldots$$
$$\mathsf{P}(A_n) = \sum_{k=n}^{\infty} \mathsf{P}(A_k \cap \overline{A}_{k+1}) + \mathsf{P}\biggl(\bigcap_{k=1}^{\infty} A_k\biggr).$$

Ряд $\sum_{k=1}^{\infty} \mathsf{P}(A_k \cap \overline{A}_{k+1})$ сходится, так как его сумма равна $\mathsf{P}(A_1) - \mathsf{P}\Bigl(\cap_{k=1}^{\infty} A_k\Bigr)$, а это конечное число. Поэтому при $n \to \infty$ остаток $\sum_{k=n}^{\infty} \mathsf{P}(A_k \cap \overline{A}_{k+1}) \to 0$. Переходя к пределу во втором равенстве для $\mathsf{P}(A_n)$, непосредственно получаем: $\lim\limits_{n\to\infty} \mathsf{P}(A_n) = \mathsf{P}\bigl(\cap_{k=1}^{\infty} A_k\bigr) = \mathsf{P}\bigl(\lim\limits_{n\to\infty} A_n\bigr)$. Следовательно, теорема 2.4 (первая аксиома непрерывности) доказана.

Теорема 2.5. *Вероятностная функция* $\mathsf{P}(\cdot)\colon \mathcal{F} \to [0,1]$ *непрерывна сверху, т.е. предел* $\lim\limits_{n\to\infty} \mathsf{P}(A_n) = \mathsf{P}\Big(\lim\limits_{n\to\infty} A_n\Big) = \mathsf{P}\big(\cup_{n=1}^{\infty} A_n\big)$ *для любой неубывающей последовательности событий* $A_1 \subset A_2 \subset \ldots$ *из* $\mathcal{F}$.

Доказательство. Во-первых, покажем, что $\lim\limits_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n = A \in \mathcal{F}$. При $A_1 \subset A_2 \subset \ldots$ получим

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = (A_1 \cup A_2 \cup \ldots) \cap (A_2 \cup A_3 \cup \ldots) \cap \ldots =$$
$$= (A_1 \cup A_2 \cup \ldots) \cap (A_1 \cup A_2 \cup \ldots) \cap \ldots = \bigcup_{n=1}^{\infty} A_n = A.$$

Аналогично рассуждая, легко найдем

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = (A_1 \cap A_2 \cap \ldots) \cup (A_2 \cap A_3 \cap \ldots) \cup \ldots =$$
$$= A_1 \cup A_2 \cup \ldots = \bigcup_{n=1}^{\infty} A_n = A.$$

Поэтому, $\overline{\lim\limits_{n\to\infty}}\, A_n = \underline{\lim\limits_{n\to\infty}}\, A_n = \lim\limits_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n = A$. На заключительном этапе введем новую последовательность событий: $B_1 = A \backslash A_1$, $B_2 = A \backslash A_2, \ldots,$ $B_n = A \backslash A_n, \ldots.$ Ясно, что $B_1 \supset B_2 \supset \ldots,$ так как $A_1 \subset A_2 \subset \ldots.$ Покажем, что $\cap_{n=1}^{\infty} B_n = \varnothing$. Доказательство проводится от противного. Предположим, что это утверждение неверно и, значит, $\cap_{n=1}^{\infty} B_n = \varnothing$. Тогда существует такое $\omega_0 \in \Omega$, что $\omega_0 \in \cap_{n=1}^{\infty} B_n$. Поэтому $\omega_0 \in B_1$, $\omega_0 \in B_2, \ldots.$ Так как событие $B_n = A \backslash A_n$, то $\omega_0 \in A$ и $\omega_0 \notin A_n$ для каждого $n \geqslant 1$. Отсюда одновременно $\omega_0 \in A$ и $\omega_0 \notin \cup_{n=1}^{\infty} A_n = A$. Это противоречие и доказывает, что $\cap_{n=1}^{\infty} B_n = \varnothing$.

Для событий B_n, $n \geqslant 1$, верна теорема 2.4 о непрерывности снизу. Поэтому имеем $\lim\limits_{n\to\infty} \mathsf{P}(B_n) = \mathsf{P}\Big(\lim\limits_{n\to\infty} B_n\Big) = \mathsf{P}\Big(\cap_{n=1}^{\infty} B_n\Big) = \mathsf{P}(\varnothing) = 0$. В то же время имеет место равенство $\lim\limits_{n\to\infty} \mathsf{P}(B_n) = \lim\limits_{n\to\infty} \mathsf{P}(A \backslash A_n)$. Так как $A = \cup_{n=1}^{\infty} A_n$, то $A \supset A_n$ для всех $n \geqslant 1$. Поэтому $\mathsf{P}(B_n) = \mathsf{P}(A) - \mathsf{P}(A_n)$ и, значит, $\lim\limits_{n\to\infty} \mathsf{P}(B_n) = \mathsf{P}(A) - \lim\limits_{n\to\infty} \mathsf{P}(A_n) = 0$. Таким образом, получаем, что $\lim\limits_{n\to\infty} \mathsf{P}(A_n) = \mathsf{P}(A) = \mathsf{P}\Big(\cup_{n=1}^{\infty} A_n\Big) = \mathsf{P}\Big(\lim\limits_{n\to\infty} A_n\Big)$; теорема 2.5 (вторая аксиома непрерывности) доказана.

При доказательстве теорем 2.4 и 2.5 использовалась аксиома счетной аддитивности. Наоборот, если для любого конечного набора попарно непересекающихся событий $A_i \in \mathcal{F}$ выполняется равенство $\mathsf{P}\big(\cup_{i=1}^{k} A_i\big) = \sum_{i=1}^{k} \mathsf{P}(A_i)$ и имеет место хотя бы одна из этих теорем, то аксиому счетной аддитивности можно доказать (*Гихман И. И.*, *Скороход А. В.*, *Ядренко М. И.* Теория вероятностей и математическая статистика. — Киев: Вища школа, 1979). Поэтому теоремы 2.4 и 2.5 называют аксиомами непрерывности.

Лемма 2.2 (первая лемма Бореля–Кантелли). *Пусть задана* $\{A_i; i = 1, 2, \ldots\}$ *— последовательность событий из* $\mathcal{F}$, $A^* \in \mathcal{F}$ и $\sum_{i=1}^{\infty} \mathsf{P}(A_i) < \infty$, *тогда* $\mathsf{P}(A^*) = 0$.

Доказательство. Так как $A^* = \cap_{i=1}^{\infty} \cup_{k=i}^{\infty} A_k$, то для любого $i \geqslant 1$ имеем $A^* \subset \cup_{k=i}^{\infty} A_k$. Используя теперь свойство монотонности неотрицательной вероятностной функции $\mathsf{P}(A)\colon \mathcal{F} \to [0, 1]$ и неравенство (2.3), последовательно находим, что $0 \leqslant \mathsf{P}(A^*) \leqslant \mathsf{P}\big(\cup_{k=i}^{\infty} A_k\big) \leqslant \sum_{k=i}^{\infty} \mathsf{P}(A_k)$ при любом $i \geqslant 1$. Отсюда и из условия сходимости ряда $\sum_{i=1}^{\infty} \mathsf{P}(A_i)$ следует, что $0 \leqslant \mathsf{P}(A^*) \leqslant \lim\limits_{i\to\infty} \sum_{k=i}^{\infty} \mathsf{P}(A_k) = 0$, т. е. $\mathsf{P}(A^*) = 0$. Значит, вероятность события A^*, состоящего в том, что наступает бесконечное число событий из последовательности $\{A_i; i = 1, 2, \ldots\}$, равна нулю. В этом случае с вероятностью единица произойдет только конечное число событий A_i, $i \geqslant 1$.

Таковы первые предварительные свойства вероятностной функции $\mathsf{P}(A)$, $A \in \mathcal{F}$, которые будут постоянно применяться в дальнейшем. Более того, свойства вероятностей и различные равенства в событиях позволяют находить вероятности одних событий через известные вероятности других событий. Например, если мы знаем вероятности событий A, B и $A \cap B$, то вероятность события $A \triangle B$ можно вычислить следующим способом:

$$\mathsf{P}(A \triangle B) = \mathsf{P}((A \backslash B) \cup (B \backslash A)) = \mathsf{P}(A \backslash B) + \mathsf{P}(B \backslash A) = \\ = \mathsf{P}(A) - \mathsf{P}(A \cap B) + \mathsf{P}(B) - \mathsf{P}(A \cap B) = \mathsf{P}(A) + \mathsf{P}(B) - 2\,\mathsf{P}(A \cap B).$$

Здесь были использованы равенства: $A = (A \backslash B) \cup (A \cap B)$, $B = (B \backslash A) \cup (A \cap B)$.

Краткий обзор

В главе 2 рассмотрен нетрадиционный подход приближенного вычисления вероятности исходов произвольного статистически устойчивого эксперимента. Такие вероятности называются субъективными. Субъективное назначение вероятности для случайных событий развивает у исследователей навыки и интуицию вероятностно-статистического мировоззрения на мир. Для случайных экспериментов с различными свойствами приводятся традиционные подходы (классический, геометрический, статистический и аксиоматический) к определению и вычислению вероятности случайных событий. Установлена связь между этими подходами и доказаны общие свойства вероятностной функции. В частности, установлено свойство непрерывности вероятностной функции и доказана важная лемма Бореля–Кантелли. Теоретические построения и результаты главы, как правило, иллюстрируются решениями большого числа задач, например, задачи Даламбера, задачи Галилея–Де Мере, задачи на размещение частиц в системе Максвелла–Больцмана, в системе Бозе–Эйнштейна, в системе Ферми–Дирака, в системе Линден–Белла, задачи о днях рождения, задачи о выборе престижной квартиры, задачи о гончарном круге, задачи Бюффона, задачи о туристах и т. п. Основной результат главы состоит в изложении методики построения и предварительного изучения вероятностной модели статистически устойчивого эксперимента E, для которого комплекс заданных условий его проведения априори остается неизменным.

Контрольные вопросы и упражнения

1. Привести пример назначения субъективных вероятностей и проинтерпретировать на этом примере аксиомы выбора отношения предпочтения событий.
2. Найти субъективные вероятности исходов для каждого эксперимента, описание которых приведено в примерах 1.8, 1.10, 1.24.
3. Сформулировать основной закон комбинаторики (правило умножения). Дать определения и доказать основные формулы для сочетаний, размещений, перестановок, сочетаний с повторениями, размещений с повторениями.
4. Найти все различные четырехзначные числа, если можно использовать только цифры 0, 1, 2, 3, 4, 5, 6 и цифры могут повторяться.
5. Из колоды, которая содержит 36 игральных карт, случайным образом отбираются сразу четыре карты. Вычислить вероятность того, что среди них окажется точно одна дама.
6. Из колоды, которая содержит 36 игральных карт, случайным образом отбираются сразу четыре карты. Определить вероятность того, что среди них окажется хотя бы одна дама.
7. На складе имеется 15 холодильников. Четыре из них являются бракованными. Определить вероятность того, что среди выбранных наудачу для проверки качества сразу трех холодильников ровно два будут бракованными.
8. Имеется две непрозрачные корзины. В первой корзине находятся два белых мячика и три черных, а во второй корзине — один белый и три черных мячика. Мячики отличаются только цветом. Наудачу без возвращения последовательно извлекаются все мячики следующим образом: сначала из первой корзины, потом из второй корзины, затем снова из первой корзины и т. д. Построить

вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ для этого эксперимента. Найти вероятность того, что первый белый мячик будет извлечен из второй корзины.

9. В урне m белых и n черных шаров ($m \geqslant 2$). Из урны вынимают сразу два шара. Найти вероятность того, что оба шара будут белыми.
10. На первом этаже девятиэтажного дома в лифт входят три человека. Каждый пассажир с одинаковой возможностью может выбрать любой этаж со второго до девятого. Найти вероятность того, что все они выйдут на разных этажах.
11. Игральный кубик случайным образом подбросили два раза. Найти вероятность того, что сумма выпавших очков равна пяти.
12. Слово «вероятность» разрезали на отдельные буквы и получили семь одинаковых по размеру карточек. Карточки с буквами тщательно перемешали и наудачу выложили в ряд. Найти вероятность того, что в результате будет сложено слово «вероятность».
13. Наудачу выбирается четырехзначное число, при этом цифра ноль не может стоять на первом месте. Построить вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ для этого эксперимента. Какова вероятность того, что выбранное число состоит из двух пар одинаковых цифр?
14. В сборочном цехе завода работают шесть мужчин и семь женщин. По табельным номерам наудачу отобрали пять человек. Найти вероятность того, что среди отобранных лиц окажутся хотя бы три женщины.
15. Стержень длины l разломали в двух наудачу выбранных точках на три части. Какова вероятность того, что из полученных частей можно составить треугольник?
16. Два судна должны подойти к одному и тому же причалу. Их появления суть независимые случайные события, равновозможные в течение суток. Найти вероятность того, что одному из судов придется ждать освобождения причала, если время стоянки первого судна равно одному часу, а второго равно двум часам.
17. Радиосигналы поступают от двух станций обнаружения объектов на видеоэкран в любой промежуток времени длительностью Θ. На видеоэкране появляется отметка о цели, если разность между моментами поступления радиосигналов будет меньше Θ. Найти вероятность обнаружения объекта, если каждая из станций с равной возможностью пошлет только по одному сигналу.
18. На окружности радиуса r наугад выбрано две точки. Какова вероятность того, что расстояние между ними не превышает r?

Глава 3

УНИФИЦИРОВАННАЯ И ЛОКАЛИЗОВАННАЯ ВЕРОЯТНОСТНЫЕ МОДЕЛИ УСЛОВНЫХ ЭКСПЕРИМЕНТОВ

§ 1. Условные вероятности и их свойства

1.1. Описание условных экспериментов и измерение зависимости между случайными событиями. Напомним, что статистически устойчивый эксперимент E задается множеством Σ условий его проведения и совокупностью $\mathfrak{J}$ его допустимых исходов. В основе построения вероятностной модели $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, где $\mathcal{F} \subset \mathfrak{J}$, лежит указанная совокупность Σ. При этом комплекс условий Σ остается неизменным и никаких дополнительных ограничений на эксперимент E не налагается. В силу этого эксперимент E называется *априорным*, а вероятности вида $\mathsf{P}(A)$, $A \in \mathcal{F}$, называются *доопытными* или *априорными*. Однако в ряде случаев при анализе того или иного случайного явления перед исследователем возникает трудный вопрос: как на возможность осуществления события A влияет дополнительное условие о том, что произошло некоторое событие $B \in \mathcal{F}$? Другими словами, очень важно установить связь между случайными наблюдаемыми исходами эксперимента E. В простейшем варианте между двумя событиями может быть причинно-следственная связь, когда наступление одного из событий ведет к обязательному наблюдению другого или же, наоборот, когда наступление одного события исключает осуществление другого. В более сложном случае, когда такая причинно-следственная зависимость отсутствует, некоторая связь между событиями все же имеется. Чтобы пояснить сказанное, приведем пример.

Пример 3.1. Наудачу выбирается одна из двух урн, а затем из выбранной урны вытаскивается случайным образом один шар. В первой урне содержится 100 черных и 2 белых шара. Во второй урне имеется 100 белых и 2 черных шара. Пусть событие A означает выбор второй урны, а событие B состоит в том, что вынут белый шар. В этом примере было бы неверно утверждать, что одно из этих событий влечет за собой другое или что, наоборот, одно из них исключает другое.

С другой стороны очевидно, что между событиями A и B имеется какая-то зависимость. В самом деле, при отсутствии предварительной информации о цвете выбранного шара шанс наступления события A оцениваются числом 1/2. Если же считать наступившим событие B, то вероятнее всего шар вынут из второй урны, т. е. шанс наступления события A при этом условии значительно повышается.

Итак, возникает проблема определения характеристики зависимости одних случайных событий от других. Более того, на практике часто встречаются случаи, когда нам недоступны по той или иной причине наблюдения элементарных исходов эксперимента E. В то же время мы можем утвердительно ответить на вопрос о том, произошло или не произошло некоторое случайное событие B эксперимента E. Например, в процессе случайного поступления автомобилей к стоп-линии перекрестка в течение суток мы не имеем технической возможности регистрировать их число. Однако нам доступна опосредованная информация от сотрудников дорожной инспекции о количестве нарушений за это время. Другими словами, мы наблюдаем событие $B = \{\text{количество нарушений}\}$. В этом и ему аналогичных экспериментах к комплексу $\Sigma = \{u_1, u_2, \ldots, u_s, \ldots\}$ условий проведения эксперимента E добавляется еще одно условие: «Произошло некоторое событие $B \in \mathcal{F}$». В силу этого мы фактически должны рассматривать новый статистически устойчивый эксперимент E_y с комплексом условий $\Sigma_y = \{u_1, u_2, \ldots, u_s, B, \ldots\}$.

Эксперимент E_y по отношению к эксперименту E будем называть условным. Этот термин появляется потому, что с самого начала имеем дело с экспериментом E, хотя для проведения эксперимента E требуется выполнение условий $\Sigma = \{u_1, u_2, \ldots, u_s, \ldots\}$. Однако необходимо иметь в виду следующее важное замечание. Так как событие $B \in \mathcal{F}$ является случайным, то реализовать на практике условия $\Sigma_y = \{u_1, u_2, \ldots, u_s, B, \ldots\}$ эксперимента E_y, как правило, не удается. В этом случае для проведения эксперимента E_y требуется проведение схемы серий эксперимента E, в каждой из которых опыт E повторяется независимым образом заданное число раз. Затем то испытание эксперимента E, при котором наступает случайное событие B, отождествляем с проведением условного эксперимента E_y. Значит, для проведения условного эксперимента E_y можно либо воспользоваться многократным проведением эксперимента E, либо непосредственно выполнить все условия из множества $\Sigma_y = \{u_1, u_2, \ldots, u_s, B, \ldots\}$. В силу этого можно построить различные теоретико-множественные и вероятностные модели условного эксперимента E_y, которые будем обозначать соответственно через символы $(\Omega_y, \mathcal{F}_y)$ и $(\Omega_y, \mathcal{F}_y, \mathsf{P}_y(\cdot))$. Перейдем к построению этих моделей.

Унифицированная вероятностная модель. Нетрудно видеть, что у эксперимента E_y не могут появиться новые исходы, отличные от допустимых исходов эксперимента E. По этой причине мы можем считать, что множество $\mathfrak{J}_y$ допустимых исходов эксперимента E_y равно множеству $\mathfrak{J}$ допустимых исходов эксперимента E. Аналогично рассуждая, мы примем, что множество Ω_y описаний всех элементарных исходов эксперимента E_y совпадает с множеством Ω, т.е. $\Omega_y = \Omega$. Так как $\mathcal{F}$ является σ-алгеброй, то можно в качестве множества $\mathcal{F}_y$ взять $\mathcal{F}$. Итак, мы непосредственно приходим к построению первой теоретико-множественной модели $(\Omega_y, \mathcal{F}_y) = (\Omega, \mathcal{F})$ эксперимента E_y. Следует заметить, что некоторые элементарные исходы и случайные события из $\mathcal{F}_y = \mathcal{F}$ не происходят при проведении эксперимента E_y, а это именно те события, которые не наблюдаются одновременно с B. С этой точки зрения построенная первая теоретико-множественная модель $(\Omega, \mathcal{F})$ условного эксперимента E_y является избыточной. Однако ее не следует отвергать, так как она, во-первых, совпадает с уже хорошо изученной теоретико-множественной моделью эксперимента E, и, во-вторых, остается неизменной, если мы будем рассматривать различные события B из $\mathcal{F}$. В этом смысле эту модель назовем *согласованной* или *унифицированной*. Для построения вероятностной модели эксперимента E_y в простейшем случае предполагается, что безусловная вероятность $\mathsf{P}(B) \neq 0$. Введем следующее определение.

Определение 3.1. *Условной* (апостериорной, послеопытной, байесовской) *вероятностью осуществления события* $A \in \mathcal{F}$ при наблюдении события B, для которого $\mathsf{P}(B) \neq 0$, называется число $\mathsf{P}(A \cap B)/\mathsf{P}(B)$, которое будем обозначать символом $\mathsf{P}(A|B)$.

Итак, по определению полагаем, что условная вероятность

$$\mathsf{P}_y(A) = \mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) \text{ при } \mathsf{P}(B) \neq 0. \tag{3.1}$$

Теорема 3.1. *Упорядоченная тройка* $(\Omega, \mathcal{F}, \mathsf{P}(\cdot \mid B))$ *является основным вероятностным пространством для условного эксперимента* E_y.

Доказательство. Так как $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ с самого начала является основным вероятностным пространством Колмогорова априорного эксперимента E, то упорядоченная пара $(\Omega, \mathcal{F})$ обладает всеми необходимыми свойствами в аксиоматике Колмогорова. Поэтому необходимо проверить только свойства нормировки, неотрицательности и счетной аддитивности для условной вероятности $\mathsf{P}(A|B)$, $A \in \mathcal{F}$. Свойство нормировки следует сразу из равенств $\mathsf{P}(\Omega|B) = \mathsf{P}(\Omega \cap B)/\mathsf{P}(B) = \mathsf{P}(B)/\mathsf{P}(B) = 1$, а свойство неотрицательности следует из (3.1) и аксиомы неотрицательности для функции $\mathsf{P}(\mathrm{A})$, $A \in \mathcal{F}$. Действительно, условная вероятность вида $\mathsf{P}_y(A) = \mathsf{P}(A|B) = \mathsf{P}(A\cap B)/\mathsf{P}(B) \geqslant 0$, так как $\mathsf{P}(A\cap B) \geqslant 0$ и $\mathsf{P}(B) > 0$.

Пусть теперь имеем любую счетную последовательность A_1, $A_2, \ldots$ попарно непересекающихся событий из $\mathcal{F}$. Тогда последовательно находим:

$$\mathsf{P}\Bigl(\bigcup_{i=1}^{\infty} A_i|B\Bigr) = \frac{\mathsf{P}\bigl((\cup_{i=1}^{\infty} A_i)\cap B\bigr)}{\mathsf{P}(B)} = \frac{\mathsf{P}\bigl(\cup_{i=1}^{\infty}(A_i\cap B)\bigr)}{\mathsf{P}(B)} = \\ = \sum_{i=1}^{\infty}\frac{\mathsf{P}(A_i\cap B)}{\mathsf{P}(B)} = \sum_{i=1}^{\infty}\mathsf{P}(A_i|B)$$

для случайных событий A_1, $A_2, \ldots$. Заметим, что $A_i \cap B \subset A_i$ для всех $i = 1, 2, \ldots$ и, значит, последовательность $A_1\cap B$, $A_2\cap B, \ldots$ составлена из попарно непересекающихся случайных событий. Итак, свойство счетной аддитивности имеет место для условной вероятности. Теорема 3.1 доказана, так как выполняется вся система аксиом Колмогорова для упорядоченной тройки $(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$. Можно сказать, что условная вероятность обладает всеми свойствами априорной вероятности, т. е. она является неотрицательной, нормированной и счетно-аддитивной функцией на множестве $\mathcal{F}$.

Условные частоты и условные вероятности. Здесь следовало бы обосновать целесообразность определения условной вероятности или вид формулы (3.1) с экспериментальной точки зрения. Для этого проведем N раз статистически устойчивый эксперимент E. Пусть при N опытах этого эксперимента событие A произошло $\mu(A,N)$ раз, событие B произошло $\mu(B,N)$ раз и событие $A\cap B$ произошло $\mu(A\cap B,N)$ раз. Теперь рассмотрим лишь те $\mu(B,N)$ опытов, в которых произошло событие B. Число наступлений события A только в этих $\mu(B,N)$ опытах, очевидно, равно числу $\mu(A\cap B,N)$ наступлений события $A\cap B$ во всех N опытах. Здесь важно отметить, что при достаточно большом N в силу колебаний статистической вероятности $\mu(B,N)/N$ около априорной вероятности $\mathsf{P}(B) > 0$ будет велико и число $\mu(B,N)$ тех испытаний, в которых осуществляется событие B. Тогда относительная частота события A при условии, что событие B произошло, равна

$$\mu(A\cap B,N)/\mu(B,N) = (\mu(A\cap B,N)/N) : (\mu(B,N)/N). \tag{3.2}$$

Отношение $\mu(A\cap B,N)/\mu(B,N)$ показывает, какую долю от числа опытов, в которых наступило событие B, составляет число таких опытов, в которых, наряду с B, наступило также и событие A. Естественно назвать отношение $\mu(A\cap B,N)/\mu(B,N)$ условной частотой события A при условии, что событие B произошло, и обозначить это отношение через символ $\mathsf{P}^*(A|B,N)$, или, еще проще, через $\mathsf{P}^*(A|B)$. Принимая во внимание это обозначение и ранее введенное обозначение для частоты произвольного события при определении статистической вероятности для априорного эксперимента, равенство (3.2) запишем в таком виде:

$$\mathsf{P}^*(A|B) = \mathsf{P}^*(A \cap B)/\mathsf{P}^*(B). \tag{3.3}$$

Выражение (3.3) связывает условную частоту $\mathsf{P}^*(A|B)$ события A с безусловными частотами $\mathsf{P}^*(A \cap B)$ и $\mathsf{P}^*(B)$ событий $A \cap B$ и B. Из соотношения (3.1) легко видеть, что равенство (3.3) и положено в основу определения условной вероятности. В самом деле, безусловная частота $\mathsf{P}^*(A \cap B)$ группируется около вероятности $\mathsf{P}(A \cap B)$, а безусловная частота $\mathsf{P}^*(B)$ колеблется около вероятности $\mathsf{P}(B)$. Поэтому правая часть вида $\mathsf{P}^*(A \cap B)/\mathsf{P}^*(B)$ равенства (3.3) группируется около дроби $\mathsf{P}(A \cap B)/\mathsf{P}(B)$. На этом основании можно сказать, что условная частота $\mathsf{P}^*(A|B)$ группируется около условной вероятности $\mathsf{P}(A|B)$. Естественно величину вида $|\mathsf{P}(A) - \mathsf{P}(A|B)|$ можно принять за меру статистической или вероятностной зависимости случайного события A от случайного события B.

Локализованная вероятностная модель условных экспериментов. Переходим теперь к построению второй вероятностной модели условного статистически устойчивого случайного эксперимента E_y. Пусть $\mathfrak{J}'$ есть множество элементарных исходов, которое было рассмотрено в п. 2.1 первой главы. Сначала отметим следующий важный факт. Любой элементарный исход из $\mathfrak{J}'$, который не пересекается с B, не происходит при проведении эксперимента E_y. Следовательно, такого рода элементарные исходы из $\mathfrak{J}'$ могут быть отнесены к невозможному событию $\varnothing$ по отношению к эксперименту E_y. С другой стороны, исход B всегда наступает при проведении эксперимента E_y. В силу этого, в качестве достоверного события Ω_y для эксперимента E_y можно взять случайное событие $B \in \mathcal{F}$. Далее, все случайные события из $\mathcal{F}$, которые не пересекаются с B, также не происходят при проведении эксперимента E_y, и поэтому могут быть отнесены к ненаблюдаемым событиям по отношению к эксперименту E_y. Эти несложные рассуждения позволяют выбрать в качестве множества $\mathcal{F}_y$ наблюдаемых исходов условного эксперимента E_y совокупность вида $\{A \cap B\colon A \in \mathcal{F}\}$.

Пусть теперь $\mathcal{F} \cap B = \{A \cap B\colon A \in \mathcal{F}\} = \mathcal{F}_y$ является совокупностью всех наблюдаемых событий условного эксперимента E_y. Покажем, что $\mathcal{F} \cap B$ есть σ-алгебра подмножеств из B. Так как множество $\mathcal{F} \cap B \subset \mathcal{F}$, то в $\mathcal{F} \cap B$ введены все те же теоретико-множественные операции, что и в $\mathcal{F}$. Равенство $B = B \cap B$ позволяет утверждать, что событие $B \in \mathcal{F} \cap B$ и, значит, $\mathcal{F} \cap B \neq \varnothing$. Далее, если $C \in \mathcal{F} \cap B$, то событие $C = C_0 \cap B$, где C_0 — некоторое случайное событие из $\mathcal{F}$. Отсюда нетрудно получить следующие соотношения:

$$\overline{C} = \Omega_y \backslash C = B \backslash C = B \backslash (C_0 \cap B) = B \cap \overline{C_0 \cap B} = B \cap (\overline{C}_0 \cup \overline{B}) =$$

$$= (B \cap \overline{C}_0) \cup (B \cap \overline{B}) = (\overline{C}_0 \cap B) \cup \varnothing = (\overline{C}_0 \cap B) \in \mathcal{F} \cap B,$$

ибо $\overline{C}_0 \in \mathcal{F}$. Наконец, если A_1, $A_2, \ldots$ суть последовательность множеств из $\mathcal{F} \cap B$, то имеем $A_1 = D_1 \cap B$, $A_2 = D_2 \cap B \ldots$, где D_1, $D_2, \ldots$ — некоторые случайные события из $\mathcal{F}$. Найдем теперь случайное событие вида $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty}(D_i \cap B) = \left(\cup_{i=1}^{\infty} D_i\right) \cap B$ и, значит, $\cup_{i=1}^{\infty} A_i \in \mathcal{F} \cap B$, если учесть, что $\cup_{i=1}^{\infty} D_i \in \mathcal{F}$. Итак, для $\mathcal{F} \cap B$ выполняются все аксиомы σ-алгебры.

Теперь необходимо проверить аксиомы Колмогорова для условной вероятности $\mathsf{P}(A|B)$, $A \in \mathcal{F} \cap B$. Свойство нормировки получим из равенств $\mathsf{P}(B|B) = \mathsf{P}(B \cap B)/\mathsf{P}(B) = \mathsf{P}(B)/\mathsf{P}(B) = 1$. Свойства неотрицательности и счетной аддитивности условной вероятности $\mathsf{P}(A|B)$, $A \in \mathcal{F} \cap B$, следует из $\mathcal{F} \cap B \subset \mathcal{F}$ и из того, что условная вероятность $\mathsf{P}(A|B)$ обладает указанными свойствами на $\mathcal{F}$. Как уже отмечалось ранее, при $A \in \mathcal{F} \cap B$ условную вероятность $\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) = \mathsf{P}(A)/\mathsf{P}(B)$ обозначают через $\mathsf{P}_y(A)$, подчеркивая тем самым, что мы рассматриваем эксперимент E_y и его теоретико-множественную модель $(\Omega_y, \mathcal{F}_y) = (B, \mathcal{F} \cap B)$, которая существенно зависит от B. При этом вероятность $\mathsf{P}_y(A)$, $A \in \mathcal{F} \cap B$, можно найти, рассматривая эксперимент E_y как априорный эксперимент и используя различные подходы к определению вероятности, изложенные во второй главе. Можно сказать, что при построении модели $(B, \mathcal{F} \cap B, \mathsf{P}_y(A))$ условного эксперимента E_y используются такие элементарные исходы априорного эксперимента E, которые полностью определяют локальные свойства исхода B. При этом исход B эксперимента E считаем одним из условий проведения эксперимента E_y. Поэтому такую модель будем называть локализованной по отношению к событию B.

На этом этапе рассуждений хотелось бы отметить, что для статистически устойчивых экспериментов простейшей структуры, например для классических экспериментов, формулу (3.1) даже можно доказать. Действительно, пусть для некоторого классического эксперимента E событию A благоприятствует m элементарных исходов, а случайному событию B благоприятствуют k элементарных исходов. Тогда $A = \{\omega_{a_1}, \omega_{a_2}, \ldots, \omega_{a_m}\}$ и $B = \{\omega_{b_1}, \omega_{b_2}, \ldots, \omega_{b_k}\}$. В этом случае событию $A \cap B$ благоприятствуют r элементарных исходов, т.е. событие $A \cap B = \{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_r}\}$, $A \cap B \subset A$, $A \cap B \subset B$ и $m \geqslant r$, $k \geqslant r$. Предположим теперь, что событие B произошло. Значит, наблюдался один из k равновозможных элементарных исходов. Из этих k элементарных исходов событию A благоприятствуют только r равновозможных исходов, которые наблюдаются одновременно с исходом $A \cap B$. На основании классического определения вероятности для эксперимента E_y с конечным числом k равновозможных исходов $\{\omega_{b_1}\}$, $\{\omega_{b_2}\},\ldots$, $\{\omega_{b_k}\}$ последовательно получаем, что $\mathsf{P}_y(A) = \mathsf{P}(A|B) = r/k = (r/n)/(k/n) = \mathsf{P}(A \cap B)/\mathsf{P}(B)$, и формула (3.1) доказана.

1.2. Способы выбора моделей условного эксперимента. Этапы построения двух предложенных вероятностных моделей условного эксперимента E_y наглядно проиллюстрируем на примерах.

Пример 3.2. В урне находятся два белых шара и один черный. Из урны последовательно наудачу вынимают два шара. Известно, что второй шар — белый. Какова вероятность, что первый шар — тоже белый?

Решение. Способ первый. Сначала построим вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ априорного эксперимента E. Комплекс условий его проведения $\Sigma = \{u_1$ — взята некоторая урна с тремя шарами, u_2 — два белых шара в урне, u_3 — один черный шар в урне, u_4 — последовательно наудачу вынимают два шара, $\ldots\}$. В комплекс условий проведения эксперимента E не входит утверждение о том, что вторым вынут белый шар. Пусть теперь $\omega_1 = (\text{б}, \text{б})$ — описание элементарного исхода, заключающегося в том, что оба вынутых шара — белых. Далее, $\omega_2 = (\text{б}, \text{ч})$ — описание элементарного исхода, заключающегося в том, что первым вытаскивают белый шар, вторым — черный. Наконец, $\omega_3 = (\text{ч}, \text{б})$ — описание элементарного исхода, заключающегося в том, что первым вытаскивают черный шар, а вторым — белый. Тогда для такого эксперимента достоверное событие

$$\Omega = \{(\text{б}, \text{б}), (\text{б}, \text{ч}), (\text{ч}, \text{б})\}$$

и σ-алгебра

$$\mathcal{F} = \{\varnothing, \{(\text{б}, \text{б}), (\text{б}, \text{ч}), (\text{ч}, \text{б})\}, \{(\text{б}, \text{б})\}, \{(\text{б}, \text{ч})\}, \{(\text{ч}, \text{б})\}, \\ \{(\text{б}, \text{б}), (\text{б}, \text{ч})\}, \{(\text{б}, \text{б}), (\text{ч}, \text{б})\}, \{(\text{ч}, \text{б}), (\text{б}, \text{ч})\}\}.$$

Используя классическое определение вероятности, найдем

$$\mathsf{P}(\varnothing) = 0, \quad \mathsf{P}(\Omega) = 1, \quad \mathsf{P}(\{(\text{б}, \text{б})\}) = \mathsf{P}(\{(\text{б}, \text{ч})\}) = \mathsf{P}(\{(\text{ч}, \text{б})\}) = 1/3, \\ \mathsf{P}(\{(\text{б}, \text{б}), (\text{б}, \text{ч})\}) = \mathsf{P}(\{(\text{б}, \text{б}), (\text{ч}, \text{б})\} = \mathsf{P}(\{(\text{ч}, \text{б}), (\text{б}, \text{ч})\}) = 2/3.$$

Следовательно, построена вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ априорного эксперимента E для этого примера.

Элементы Ω_y, $\mathcal{F}_y$ первой вероятностной модели рассматриваемого условного эксперимента E_y с комплексом условий $\Sigma_y = \{u_1$ — взята некоторая урна с тремя шарами, u_2 — два белых шара в урне, u_3 — один черный шар в урне, u_4 — последовательно вынимают случайным образом два шара таким образом, что второй шар — белый шар, $\ldots\}$ имеют следующий вид: $\Omega_y = \Omega$, $\mathcal{F}_y = \mathcal{F}$. Теперь если $A = \{(б,б),(б,ч)\}$ — событие, которое заключается в том, что первым вытаскивают белый шар и $B = \{(б,б),(ч,б)\}$ есть событие, которое заключается в том, что вторым вытаскивают белый шар, тогда $A \cap B = \{(б,б)\}$. Теперь по формуле (3.1) найдем, что послеопытная вероятность $\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) = \mathsf{P}(\{(б,б)\})/\mathsf{P}(\{(б,б),(ч,б)\}) = \mathsf{P}(\{\omega_1\})/\mathsf{P}(\{\omega_1,\omega_3\}) = 1/2$. Это и есть окончательный ответ на поставленный вопрос в данном примере. Заметим, что, используя формулу (3.1), нетрудно найти условные вероятности $\mathsf{P}(\cdot|B)$ для остальных событий из $\mathcal{F}_y = \mathcal{F}$.

Второй способ решения этой задачи. Элементы Ω_y, $\mathcal{F}_y$ локализованной вероятностной модели данного условного эксперимента E_y с комплексом условий Σ_y имеют следующий вид:

$$\Omega_y = B = \{(б,б),(ч,б)\},$$
$$\mathcal{F}_y = \mathcal{F} \cap B = \{\varnothing, \{(б,б),(ч,б)\}, \{(б,б)\}, \{(ч,б)\}\} \subset \mathcal{F}.$$

Теперь для события $\{(б,б)\} = A \cap B \in \mathcal{F}_y$, в котором уже учтено, что событие B произошло и $B = \Omega_y$, имеем: $\mathsf{P}(\{(б,б)\}|B) = \mathsf{P}(A \cap B|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) = 1/2$. Заметим, что эту же вероятность можно определить, используя классическое определение для эксперимента E_y всего с двумя элементарными исходами. Тогда $\mathsf{P}_y(\{(б,б)\}) = 1/2$, так как для E_y имеем $n = 2$, $m = 1$.

Пример 3.3. Пусть эксперимент заключается в однократном бросании игральной кости. Пусть A — событие, заключающееся в выпадении четного числа очков, и B — событие, заключающееся в выпадении числа очков, меньшего трех. Определить условную вероятность $\mathsf{P}(A|B)$.

Решение. Для априорного опыта достоверное событие Ω равно множеству $\{1,2,3,4,5,6\}$, а множество $\mathcal{F}$ всех его наблюдаемых исходов можно представить в таком виде:

$$\begin{gathered}\{\varnothing, \{1,2,3,4,5,6\}, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\},\\ \{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{1,6\}, \{1,2,3\}, \{1,2,4\}, \{1,2,5\}, \{1,2,6\},\\ \{1,3,4\}, \{1,3,5\}, \{1,3,6\}, \{1,4,5\}, \{1,4,6\}, \{1,5,6\}, \{2,3\}, \{2,4\},\\ \{2,5\}, \{2,6\}, \{2,3,4\}, \{2,3,5\}, \{2,3,6\}, \{2,4,5\}, \{2,4,6\}, \{2,5,6\},\\ \{3,4\}, \{3,5\}, \{3,6\}, \{3,4,5\}, \{3,4,6\}, \{3,5,6\}, \{4,5\}, \{4,6\},\\ \{5,6\}, \{4,5,6\}, \{1,2,3,4\}, \{1,2,3,5\}, \{1,2,3,6\}, \{1,2,4,5\},\\ \{1,2,4,6\}, \{1,2,5,6\}, \{1,3,4,5\}, \{1,3,4,6\}, \{1,3,5,6\}, \{1,4,5,6\},\\ \{2,3,4,5\}, \{2,3,4,6\}, \{2,3,5,6\}, \{2,4,5,6\}, \{3,4,5,6\}, \{1,2,3,4,5\},\\ \{1,2,3,4,6\}, \{1,2,3,5,6\}, \{1,2,4,5,6\}, \{1,3,4,5,6\}, \{2,3,4,5,6\}\}.\end{gathered}$$

Множество $\mathcal{F}$ содержит 64 элемента или 2^6 подмножеств множества Ω. Легко найти, что $A = \{2,4,6\}$, $B = \{1,2\}$ и $\mathsf{P}(\{1,2\}) = 2/6$. Из формулы (3.1) сразу получим, что условная вероятность

$$\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) = \mathsf{P}(\{2\})/\mathsf{P}(\{1,2\}) = (1/6) : (2/6) = 1/2.$$

Можно решать такого рода задачи, рассуждая следующим образом. Для условного эксперимента E_y достоверное событие Ω_y равно множеству $\{1,2\}$, а множество $\mathcal{F}_y$ всех его наблюдаемых исходов можно записать в виде $\{\varnothing, \{1,2\}, \{1\}, \{2\}\}$. Итак, условный эксперимент E_y имеет два элементарных события, три наблюдаемых события и невозможное событие $\varnothing$. Можно сказать, что условный эксперимент E_y существенно проще по сравнению с априорным экспериментом E. Теперь событию $A \cap B$ из $\mathcal{F}_y$ благоприятствует только один элементарный исход эксперимента E_y из двух возможных. Поэтому искомая вероятность равна 1/2 согласно классическому определению вероятности для эксперимента E_y.

Пример 3.4. Две игральные кости бросают один раз. Найти вероятность выпадения простой суммы, если сумма выпавших очков больше пяти.

Решение. Напомним информацию из таблиц 2.1 и 2.2 примера 2.24. Простая сумма выпавших очков означает, что эта сумма есть простое число. Обозначим через A событие, которое заключается в том, что сумма очков есть простое число. Пусть B есть событие, которое заключается в том, что сумма очков больше пяти. Тогда событие $A \cap B$ заключается в том, что сумма очков есть простое число, большее пяти.

Рассмотрим первый способ вычисления условной вероятности $\mathsf{P}(A|B)$. Из таблиц 2.1 и 2.2 примера 2.24 и определения условной вероятности получим, что $\mathsf{P}(A \cap B) = 8/36$, $\mathsf{P}(B) = 26/36$ и $\mathsf{P}(A|B) = 8/26 = 4/13 \approx 0{,}308$. Далее, рассмотрим второй способ решения, который основан на построении локализованной вероятностной модели $(B, \mathcal{F} \cap B, \mathsf{P}(A \cap B)/\mathsf{P}(B))$ для условного эксперимента E_y. Достоверному событию $\Omega_y = B$ условного эксперимента E_y благоприятствует 26 элементарных событий. И только 8 среди этих элементарных событий благоприятствуют появлению события A. Поэтому, используя классическое определение вероятности для эксперимента E_y, получим, что $\mathsf{P}(A|B) = 8/26 = 4/13 \approx 0{,}308$.

Приведем теперь иллюстративные примеры на различные способы вычисления условных вероятностей по формуле (3.1).

Пример 3.5. Один раз брошены две игральные кости. Чему равна вероятность того, что сумма выпавших на них очков равна восьми (событие A), если известно, что эта сумма есть четное число (событие B).

Решение. Используя методику решения примера 2.24, легко найдем, что событие $A = \{(x,y)\colon x + y = 8\} = \{(2,6), (6,2), (3,5), (5,3), (4,4)\}$

и событие $B = \{(x, y)\colon x + y = 2, 4, 6, 8, 10, 12\} = \{(1, 1),\ (1, 3),\ (3, 1),\ (1, 5),\ (5, 1),\ (2, 2),\ (2, 4),\ (4, 2),\ (2, 6),\ (6, 2),\ (3, 3),\ (3, 5),\ (5, 3),\ (4, 4),\ (4, 6),\ (6, 4),\ (5, 5),\ (6, 6)\}$. Согласно классическому определению вероятности, получим $\mathsf{P}(A) = 5/36$ и $\mathsf{P}(B) = 18/36$. Так как $A \subset B$, то $A = A \cap B$, поэтому $\mathsf{P}(A \cap B) = 5/36$ и, следовательно, искомая условная вероятность вида $\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) = 5/18$.

Пример 3.6. Из колоды в 36 карт, последовательно вынуты две карты. Определить вероятность того, что вторая карта туз. Определить вероятность того, что вторая карта — туз, если первоначально был вынут туз.

Решение. Перенумеруем все карты от 1 до 36. Пусть символ $\omega = (x, y)$ означает описание такого элементарного исхода, когда первая выбранная карта имеет номер x, а вторая y. Тогда достоверное событие Ω равно $\{\omega = (x, y)\colon x, y \in \{1, 2, \ldots, 36\},\ x \neq y\}$ и включает $n = 36 \times 35$ описаний. Поэтому $\mathsf{P}(\{\omega\}) = 1/(36 \times 35)$, $\omega \in \Omega$. Обозначим через A событие, состоящее в том, что вторая вынутая карта окажется тузом, и через B событие, состоящее в том, что первая вынутая карта окажется тузом. Число элементарных исходов, благоприятствующих событиям A, $B \cap A$ и B, соответственно равно 35×4, 4×3 и 4×35. Отсюда легко находим, что $\mathsf{P}(A) = (35 \times 4)/(36 \times 35) = 1/9$, $\mathsf{P}(A \cap B) = (4 \times 3)/(36 \times 35) = 1/105$, $\mathsf{P}(B) = (4 \times 35)/(36 \times 35) = 1/9$ и условная или апостериорная вероятность $\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) = (1/105) : (1/9) = (4 \times 3)/(4 \times 35) = 3/35$.

1.3. Типичные ошибки при вычислении условных вероятностей. При вычислении условной вероятности случайного события $A \in \mathcal{F}$ по формуле (3.1) необходимо, во-первых, очень четко на содержательном уровне представлять событие B, во-вторых, уметь выражать событие B через множество описаний ω всех тех элементарных исходов из $\mathfrak{J}'$, которые одновременно могут происходить с B. Далее, нельзя заменять событие B некоторым событием $C \supset B$, хотя в этом случае событие C также произошло. В противном случае могут возникать ошибки и парадоксы. Для подтверждения этого высказывания приведем ряд примеров.

Пример 3.7. Для социологических исследований случайным образом выбирается семья. В семье имеется двое детей. Трубку телефона, по которому позвонили в эту семью, взял случайно мальчик. Найти вероятность того, что в семье оба мальчика, т. е. и другой ребенок будет мальчик.

Очень часто (см., например, *Lipschuz S.* Theory and Problems of Probability. — New York: McGraw-Hill Book Corp., 1968) эту задачу по существу решают следующим образом. В качестве пространства Ω описаний всех элементарных исходов выбирают множество $\{(\text{м}, \text{м}), (\text{м}, \text{д}), (\text{д}, \text{м}), (\text{д}, \text{д})\}$. Здесь одноточечные множества $\{(\text{м}, \text{м})\}$, $\{(\text{м}, \text{д})\}$, $\{(\text{д}, \text{м})\}$ и, наконец, множество $\{(\text{д}, \text{д})\}$ соответственно означают, что в семье старший ребенок и младший — мальчики, старший мальчик и младшая девочка, старшая девочка и младший мальчик и, наконец, старшая и младшая девочки. Пусть событие B состоит в том, что трубку взял мальчик. Обозначим через A событие вида $\{(\text{м}, \text{м})\}$ и через C событие, которое заключается в том, что в семье есть мальчик, т. е. $C = \{(\text{м}, \text{м}), (\text{м}, \text{д}), (\text{д}, \text{м})\}$. Ясно, что $C \supset B$, более того, многие ошибочно считают эти события равными. Тогда, по их мнению, условная вероятность вида $\mathsf{P}(A|B) = \mathsf{P}(A|C) = \mathsf{P}(A \cap C)/\mathsf{P}(C) = \mathsf{P}(A)/\mathsf{P}(C) = (1/4) : (3/4) = 1/3$. Этот ответ противоречит нашей интуиции о том, что появление мальчика или девочки в качестве другого ребенка, который не брал трубку телефона, имеет одну и ту же вероятность, равную $1/2$.

Рассмотрим теперь правильное решение этого примера. Следует понимать, что в нашем эксперименте действуют два случайных механизма. Первый состоит в том, что случайным образом выбирается семья из двух детей, а второй — в том, что один из детей (старший или младший) случайно взял трубку телефона. В качестве пространства описаний всех элементарных исходов из $\mathfrak{J}'$ выберем множество $\Omega = \{(\text{м}, \text{м}, 1), (\text{м}, \text{м}, 2), (\text{м}, \text{д}, 1), (\text{м}, \text{д}, 2), (\text{д}, \text{м}, 1), (\text{д}, \text{м}, 2), (\text{д}, \text{д}, 1), (\text{д}, \text{д}, 2)\}$. Здесь первые два элемента в каждой упорядочной тройке представляют пол детей в последовательности их рождения, а третий элемент характеризует старшинство ребенка, взявшего телефонную трубку. Третий элемент равен единице, если телефонную трубку снял старший ребенок, и равен двойке в противном случае. При таком подходе $A = \{(\text{м}, \text{м}, 1), (\text{м}, \text{м}, 2)\}$, $B = \{(\text{м}, \text{м}, 1), (\text{м}, \text{м}, 2), (\text{м}, \text{д}, 1), (\text{д}, \text{м}, 2)\}$, $C = \{(\text{м}, \text{м}, 1), (\text{м}, \text{м}, 2), (\text{м}, \text{д}, 1), (\text{м}, \text{д}, 2), (\text{д}, \text{м}, 1), (\text{д}, \text{м}, 2)\}$ и, значит, $C \supset B$, $C \neq B$, $\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) = (2/8) : (4/8) = 1/2$, $\mathsf{P}(A|C) = \mathsf{P}(A \cap C)/\mathsf{P}(C) = (2/8) : (6/8) = 1/3$, что вполне соответствует нашей интуиции. Заметим, что при первом решении не удается представить событие B через элементы множества $\Omega = \{(\text{м}, \text{м}), (\text{м}, \text{д}), (\text{д}, \text{м}), (\text{д}, \text{д})\}$.

Пример 3.8. Случайным образом выбирается семья, имеющая двух детей. По телефону от одного из родителей узнаем, что один из их детей — мальчик. Найти вероятность того, что в семье оба мальчика.

В примере 3.8 имеется только один случайный механизм, который состоит в непреднамеренном выборе семьи, имеющей двух детей. Предполагается, что родители говорят правду. Поэтому в качестве пространства Ω описаний всех элементарных исходов из $\mathfrak{J}'$ для этого эксперимента можно предложить множество $\{(\text{м}, \text{м}, \text{да}), (\text{м}, \text{д}, \text{да}),$

(д, м, да), (д, д, нет)}. Для выбранного описания первые два элемента в каждой упорядочной тройке $\omega = (\cdot, \cdot, \cdot)$ представляют пол детей в последовательности их рождения, а третий элемент соответствует высказыванию одного из родителей о наличии или отсутствии мальчика в семье. Теперь легко находим равенства $A = \{(\text{м}, \text{м}, \text{да})\}$, $C = \{(\text{м}, \text{м}, \text{да}), (\text{м}, \text{д}, \text{да}), (\text{д}, \text{м}, \text{да})\}$. Поэтому условная вероятность $\mathsf{P}(A|C) = \mathsf{P}(A)/\mathsf{P}(C) = (1/4) : (3/4) = 1/3$ и напомним, что событие B не является исходом данного эксперимента.

Наконец, приведем еще более яркий и поучительный пример.

Пример 3.9 (*задача о трех фермерских хозяйствах*). Рассматривается работа трех фермерских хозяйств Φ_1, Φ_2, Φ_3 по производству мяса. Фермерское хозяйство Φ_1 имеет 100 белых бычков, хозяйство Φ_2 — 50 белых и 50 черных, наконец, фермерское хозяйство Φ_3 содержит 100 черных бычков. Наудачу выбирается одно из фермерских хозяйств, а затем из выбранного хозяйства непреднамеренно отбирают 50 бычков, которые все оказались белыми. Найти вероятность того, что выбрано фермерское хозяйство Φ_1.

Пусть при $i = 1, 2, 3$ событие A_i, $A_1 + A_2 + A_3 = \Omega$, означает выбор фермерского хозяйства Φ_i, событие B означает отбор 50 белых бычков и $C = A_1 \cup A_2$. Требуется найти вероятность того, что выбрано фермерское хозяйство Φ_1 при условии, что все отобранные 50 бычков оказались белыми, т. е. найти $\mathsf{P}(A_1|B)$.

Будем поступать теперь по схеме неверного решения примера 3.7. Так как $C \supset B$, то событие C также наступило, и при $i = 1, 2$ имеем $\mathsf{P}(A_i|C) = \mathsf{P}(A_i \cap C)/\mathsf{P}(C) = (1/3) : (2/3) = 1/2$, т. е. условные вероятности $\mathsf{P}(A_1|C)$ и $\mathsf{P}(A_2|C)$ совпали. При таком способе решения исключается из рассмотрения фермерское хозяйство, в котором нет белых бычков, т. е. нарушается случайный выбор фермы.

Для построения правильного решения возьмем в качестве описания ω произвольного элементарного исхода данного эксперимента вектор (x, y). В векторе $\omega = (x, y)$ первый элемент x определяет выбор некоторого фермерского хозяйства из трех, а второй элемент y задает некоторый отбор 50 различных бычков из 100 уже выбранного фермерского хозяйства. Например, при $x = \Phi_1$ и $y = \{1, 2, \ldots, 50\}$ случайное элементарное событие вида $\{(\Phi_1, \{1, 2, \ldots, 50\})\}$ имеет описание $\omega_1 = (\Phi_1, \{1, 2, \ldots, 50\})$ и заключается в том, что выбрано фермерское хозяйство Φ_1 и в нем отобрано 50 первых бычков. При этом, ради простоты, бычки в каждом хозяйстве пронумерованы от 1 до 100. Легко видеть, что для этого опыта достоверное событие $\Omega = \{\omega\colon \omega = \omega_1, \omega_2, \ldots, \omega_n\}$ и $n = 3\,C_{100}^{50}$. Далее, событие B содержит $C_{100}^{50} + 1$ описание ω всех таких элементарных исходов, каждый из которых может происходить с событием B. Значит, безусловная вероятность $\mathsf{P}(B) = \big(C_{100}^{50} + 1\big)/3\, C_{100}^{50}$, а условные вероятности вида $\mathsf{P}(A_1|B)$ и $\mathsf{P}(A_2|B)$ можно найти следующим образом:

$$\mathsf{P}(A_1|B) = \mathsf{P}(A_1 \cap B)/\mathsf{P}(B) = \mathsf{P}(B|A_1) \cdot \mathsf{P}(A_1)/\mathsf{P}(B) = \\ = C_{100}^{50}/(C_{100}^{50} + 1) = 1 - (C_{100}^{50} + 1)^{-1} \approx 1,$$

$$\mathsf{P}(A_2|B) = \mathsf{P}(A_2 \cap B)/\mathsf{P}(B) = \mathsf{P}(B|A_2) \cdot \mathsf{P}(A_2)/\mathsf{P}(B) = \\ = \big(1/3\, C_{100}^{50}\big) : \big(\big(C_{100}^{50} + 1\big)/3\, C_{100}^{50}\big) = \big(C_{100}^{50} + 1\big)^{-1} \approx 0.$$

Заметим, что здесь $A_2 \cap B = \{(\Phi_2, \{1, 2, \ldots, 50\})\}$ является случайным элементарным событием, которое заключается в том, что выбрано фермерское хозяйство Φ_2 и в нем случайным образом отобрано 50 бычков, которые все оказались белыми. Полученные числовые значения для условных вероятностей $\mathsf{P}(A_1|B)$ и $\mathsf{P}(A_2|B)$ хорошо соответствуют интуиции.

§ 2. Формулы и методика вычисления вероятностей

2.1. Связь условных и априорных вероятностей. Теорема умножения. При изучении общих свойств вероятностной функции $\mathsf{P}(A)\colon \mathcal{F} \to [0,1]$ в гл. 2 были получены простейшие формулы для вычисления вероятностей одних событий через вероятности других событий. Например, при $A \in \mathcal{F}$ имеем $\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A)$, при $A, B \in \mathcal{F}$ и $A \subset B$ верно равенство $\mathsf{P}(B \backslash A) = \mathsf{P}(B) - \mathsf{P}(A)$, наконец, при $A, B \in \mathcal{F}$ справедливо соотношение вида $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$. Далее, рассмотрим два случайных события $A \in \mathcal{F}$, $B \in \mathcal{F}$, и пусть $\mathsf{P}(A) \neq 0$ и $\mathsf{P}(B) \neq 0$. По определению условных вероятностей из (3.1), имеем $\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B)$, $\mathsf{P}(B|A) = \mathsf{P}(B \cap A)/\mathsf{P}(A)$. Каждое из этих равенств эквивалентно так называемой теореме умножения для двух событий, т. е. справедливы соотношения $\mathsf{P}(A \cap B) = \mathsf{P}(B \cap A) = \mathsf{P}(A) \cdot \mathsf{P}(B|A) = \mathsf{P}(B) \cdot \mathsf{P}(A|B)$ — вероятность пересечения двух событий равна вероятности одного события, умноженной на условную вероятность второго события при условии, что первое событие произошло. Повторное применение этой теоремы для трех событий A, B и C дает: $\mathsf{P}(A \cap B \cap C) = \mathsf{P}((A \cap B) \cap C) = \mathsf{P}(A \cap B)\, \mathsf{P}(C|(A \cap B)) = \mathsf{P}(A)/\mathsf{P}(B|A)\, \mathsf{P}(C|(A \cap B))$.

Докажем теперь часто используемое на практике следующее утверждение.

Теорема 3.2 (теорема умножения). *Пусть случайные события A_1, $A_2, \ldots$, A_{n-1}, A_n принадлежат σ-алгебре $\mathcal{F}$ и вероятность*

$\mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n) > 0$. *Тогда имеет место следующее равенство*:

$$\mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n) = \\ = \mathsf{P}(A_1)\,\mathsf{P}(A_2|A_1)\,\mathsf{P}(A_3|A_1 \cap A_2) \times \ldots \times \mathsf{P}(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

Доказательство. Утверждение этой теоремы было установлено для двух и трех событий. Завершим дальнейшее доказательство для n событий методом математической индукции. Пусть

$$\mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) = \\ = \mathsf{P}(A_1)\,\mathsf{P}(A_2|A_1)\,\mathsf{P}(A_3|A_1 \cap A_2) \times \ldots \times \mathsf{P}(A_{n-1}|A_1 \cap A_2 \cap \ldots \cap A_{n-2})$$

для $(n-1)$ событий. Докажем подобное равенство для n событий. Обозначим вероятность $\mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n)$ через $\mathsf{P}(A \cap B)$, где случайное событие $A = A_1 \cap A_2 \cap \ldots \cap A_{n-1}$ и событие $B = A_n$. Далее имеем

$$\mathsf{P}(A \cap B) = \\ = \mathsf{P}(A)\,\mathsf{P}(B|A) = \mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \cdot \mathsf{P}(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}) = \\ = \mathsf{P}(A_1)\,\mathsf{P}(A_2|A_1)\,\mathsf{P}(A_3|A_1 \cap A_2) \times \ldots \times \mathsf{P}(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

Заметим, что для доказательства этой теоремы достаточно потребовать более слабое ограничение $\mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$, из которого следует $\mathsf{P}(A_1) > 0$, $\mathsf{P}(A_1 \cap A_2) > 0, \ldots$, $\mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$. Однако в формулировке теоремы требуется более жесткое ограничение $\mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n) > 0$, которое позволяет доказать и использовать $n!$ ее различных форм. Например, форму вида

$$\mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n) = \\ = \mathsf{P}(A_n)\,\mathsf{P}(A_{n-1}|A_n)\,\mathsf{P}(A_{n-2}|A_{n-1} \cap A_n) \times \ldots \times \mathsf{P}(A_1|A_2 \cap A_3 \cap \ldots \cap A_n).$$

Наконец, еще отметим, что теорема умножения применима, если вероятность только одного из событий A или B равна нулю. Пусть для определенности $\mathsf{P}(A) = 0$ и $\mathsf{P}(B) \neq 0$. Отсюда получаем: $0 = \mathsf{P}(A) \geqslant \mathsf{P}(A \cap B) = 0$. Поэтому имеем $\mathsf{P}(A|B) = 0$ и, значит, $0 = \mathsf{P}(A \cap B) = \mathsf{P}(B)\,\mathsf{P}(A|B) = 0$.

Классические эксперименты и теорема умножения. Выясним связь между построением моделей классических экспериментов и теоремой умножения. Эту связь продемонстрируем на простых примерах.

Пример 3.10. Буквы слова «арба» написаны каждая на отдельной карточке. Одинаковые по форме карточки тщательно перемешаны, после чего последовательно и без возвращения извлекаются три.

Выбранные карточки кладутся слева направо. Определить вероятность того, что в результате получается слово «раб» или, другими словами, наступает событие A. Рассмотрим различные способы решения этой задачи.

1. *Классический способ.* Пусть для отличия одинаковых букв с целью симметрии данного эксперимента E буквы в слове «арба» занумерованы слева направо цифрами 1, 2, 3 и 4. Обозначим множество $\{1, 2, 3, 4\}$ символом M. Комплекс условий эксперимента E имеет вид $\Sigma = \{u_1$ — четыре карточки, на каждой из которых написаны одна буква из слова «арба» и ей соответствующая метка — цифра из множества M; u_2 — наудачу по одной отбираются три карточки без возвращения и последовательно располагаются слева направо; u_3 — условие для прочтения записей на отобранных карточках;...$\}$.

В опыте E предлагается под отдельным элементарным событием $\{\omega_i\}$ понимать упорядоченный список из трех последовательно отобранных карточек с описанием в виде вектора $\omega_i = (x_1, x_2, x_3)$, причем среди компонент x_1, x_2, x_3 нет одинаковых цифр из множества M. Здесь x_1 — цифра на первой отобранной карточке, x_2 — цифра на второй отобранной карточке и, наконец, x_3 — цифра на третьей отобранной карточке. Например, описание (3, 1, 2) соответствует такому последовательному отбору трех карточек, при котором получается слово «бар». Очевидно, что при таком подходе элементарные исходы данного эксперимента будут равновероятными.

Равенства $\Omega = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3;\, x_1, x_2, x_3 \in M\}$ и $A = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3; x_1 = 2; x_2 \in M; x_3 = 3\} = \{(2, 1, 3),\, (2, 4, 3)\}$ определяют достоверное событие и событие появления слова «раб». Если число элементов каждого из множеств Ω и A соответственно равно $\aleph(\Omega)$ и $\aleph(A)$, то по правилам комбинаторики находим $\aleph(\Omega) = 4 \times 3 \times 2 = 24$ и $\aleph(A) = 1 \times 2 \times 1 = 2$. Отсюда, применяя классическое определение вероятности, сразу получим $\mathsf{P}(A) = 2/24 = 1/12$. Отметим, что для этого классического эксперимента σ-алгебра $\mathcal{F} = \{B\colon B \subset \Omega\}$.

2. *Решение с использованием теоремы умножения.* Пусть событие $A_1 \in \mathcal{F}$ обозначает появление первой карточки с надписью буквы «р», событие $A_2 \in \mathcal{F}$ обозначает появление второй карточки с надписью буквы «а», и, наконец, событие $A_3 \in \mathcal{F}$ обозначает появление третьей карточки с надписью буквы «б». Из определения событий A_1, A_2 и A_3 получим:

$$A_1 = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3;\ x_1 = 2;\ x_2, x_3 \in M\},$$
$$A_2 = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3;\ x_2 = 1, 4;\ x_1, x_3 \in M\},$$
$$A_3 = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3;\ x_3 = 3;\ x_1, x_2 \in M\},$$

$$A_1 \cap A_2 = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3;\ x_1 = 2;\ x_2 = 1, 4,\ x_3 \in M\},$$
$$A_1 \cap A_2 \cap A_3 = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3;\ x_1 = 2;\ x_2 = 1, 4;\ x_3 = 3\} = \{(2, 1, 3), (2, 4, 3)\}.$$

Следовательно, событие $A \in \mathcal{F}$ наступает, если происходят одновременно случайные события A_1, A_2 и A_3. Тогда $A = A_1 \cap A_2 \cap A_3$ и $\mathsf{P}(A) = \mathsf{P}(A_1 \cap A_2 \cap A_3) = \mathsf{P}(A_1)\,\mathsf{P}(A_2|A_1)\,\mathsf{P}(A_3|A_1 \cap A_2)$. Для вычисления условной вероятности $\mathsf{P}(A_2|A_1)$ рассмотрим по отношению к первоначальному эксперименту E первый условный эксперимент E_{y1} с комплексом условий $\Sigma_{y1} = \{u_1$ — четыре карточки, на каждой из которых написаны одна буква из слова «арба» и ей соответствующая метка цифра из множества M; u_2 — наудачу по одной отбирается три карточки без возвращения и последовательно располагаются слева направо; u_3 — на первой карточке появляется буква «р»; u_4 — условие для прочтения записей на отобранных карточках;...}.

Для эксперимента E_{y1} строим вторую теоретико-множественную модель $(\Omega_{y1}, \mathcal{F}_{y1}) = (A_1, \mathcal{F} \cap A_1)$, а вероятности событий из $\mathcal{F} \cap A_1$ определяем классическим способом (см. примеры 3.2–3.4). Так как событие $A_1 = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3;\ x_1 = 2;\ x_2, x_3 \in M\} = \{(2, 1, 3),\ (2, 3, 1),\ (2, 1, 4),\ (2, 4, 1),\ (2, 3, 4),\ (2, 4, 3)\}$ и случайное событие $A_1 \cap A_2 = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3;\ x_1 = 2;\ x_2 = 1, 4;\ x_3 \in M\} = \{(2, 1, 3),\ (2, 1, 4),\ (2, 4, 1),\ (2, 4, 3)\}$, то $\aleph(A_1) = 6$, $\aleph(A_1 \cap A_2) = 4$. Отсюда, учитывая равенство $\aleph(\Omega) = 24$, находим, что вероятность $\mathsf{P}(A_1) = 6/24 = 1/4$ и условная вероятность вида $\mathsf{P}(A_2|A_1) = \mathsf{P}(A_1 \cap A_2|A_1) = 2/3$.

Совершенно аналогично определяется условная вероятность $\mathsf{P}(A_3|A_1 \cap A_2)$. Для этого введем по отношению к первоначальному эксперименту E еще один условный эксперимент E_{y2} с комплексом условий $\Sigma_{y2} = \{u_1$ — четыре карточки, на каждой из которых написаны одна буква из слова «арба» и ей соответствующая метка — цифра из множества M; u_2 — наудачу по одной отбирается три карточки без возвращения и последовательно располагаются слева направо; u_3 — на первой карточке появляется буква «р»; u_4 — на второй карточке появляется буква «а»; u_5 — условие для прочтения записей на отобранных карточках; ...}. Для условного эксперимента E_{y2} строим теоретико-множественную модель $(\Omega_{y2}, \mathcal{F}_{y2})$ вида $(A_1 \cap A_2, \mathcal{F} \cap (A_1 \cap A_2))$, а вероятности событий из $\mathcal{F} \cap (A_1 \cap A_2)$ определяем также классическим способом.

Так как события $A_1 \cap A_2 = \{(2, 1, 3),\ (2, 1, 4),\ (2, 4, 1),\ (2, 4, 3)\}$, $A_1 \cap A_2 \cap A_3 = \{\omega_i = (x_1, x_2, x_3)\colon x_1 \neq x_2 \neq x_3;\ x_1 = 2;\ x_2 = 1, 4;\ x_3 = 3\} = \{(2, 1, 3), (2, 4, 3)\}$, то $\aleph(A_1 \cap A_2) = 4$, $\aleph(A_1 \cap A_2 \cap A_3) = 2$ и условная вероятность $\mathsf{P}(A_3|A_1 \cap A_2) = \mathsf{P}(A_1 \cap A_2 \cap A_3|A_1 \cap A_2) = 2/4 = 1/2$.

В результате окончательно находим вероятность

$$\mathsf{P}(A) = \mathsf{P}(A_1 \cap A_2 \cap A_3) = \\ = \mathsf{P}(A_1)\,\mathsf{P}(A_2|A_1)\,\mathsf{P}(A_3|A_1 \cap A_2) = (1/4) \times (2/3) \times (1/2) = 1/12.$$

Пример 3.11. В урне находится a белых шаров и b черных. Извлекаются последовательно два шара без возвращения. Найти вероятности того, что первый шар черный, а второй — белый.

Решение. Пусть A — событие, состоящее в том, что первый шар черный, B — событие, состоящее в том, что второй шар белый. Тогда по теореме умножения для двух событий A и B, используя понятие условного эксперимента и методику построения локализованной вероятностной модели, сразу получим:

$$\mathsf{P}(A \cap B) = \mathsf{P}(A)\,\mathsf{P}(B|A) = \frac{b}{a+b} \cdot \frac{a}{a+b-1} = \frac{ab}{(a+b)(a+b-1)}\,.$$

Решая задачу классическим способом, находим, что в этом опыте число всех элементарных исходов $n = (a+b)(a+b-1)$, а число благоприятствующих $m = ab$, и получаем тот же ответ.

Пример 3.12. Буквы слова «панама» написаны каждая на отдельной карточке, тщательно перемешаны и последовательно извлечены три карточки без возвращения. Какова вероятность получить слово «нам»?

Решение. Пусть A_1, A_2 и A_3 обозначают появление последовательно букв *н*, *а*, *м*. Тогда по теореме умножения для трех событий A_1, A_2 и A_3 находим $\mathsf{P}(A_1 \cap A_2 \cap A_3) = \mathsf{P}(A_1)\,\mathsf{P}(A_2|A_1)\,\mathsf{P}(A_3|A_1 \cap A_2) = (1/6) \times \times (3/5) \times (1/4) = 1/40 = 0{,}025$. Условные вероятности были вычислены методом построения соответствующей локализованной вероятностной модели.

Итак, в этом разделе предложен метод, который позволяет находить вероятности пересечений событий априорного эксперимента E через вероятности событий специально построенных и более простых условных экспериментов, например, экспериментов E_{y1} и E_{y2} в примере 3.10.

2.2. Теоремы о полной вероятности и о гипотезах. Если последовательность $H_1, H_2, \ldots$ образует полную группу попарно несовместимых событий из $\mathcal{F}$, то события $H_1, H_2, \ldots$ принято называть гипотезами. С теоремой умножения и определением гипотез тесно связаны две формулы, с помощью которых можно вычислять априорные и апостериорные вероятности случайных событий. С этой целью перейдем к теоремам о полной вероятности и о гипотезах, которые справедливы, как для конечного, так и для бесконечного числа гипотез.

Теорема 3.3. *Пусть задано вероятностное пространство* $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ *и последовательность* H_1, $H_2, \ldots$ *образует полную группу попарно несовместных событий из* $\mathcal{F}$. *Если* $\mathsf{P}(H_i) > 0$ *для* $i \geqslant 1$ *и* $A \in \mathcal{F}$, *то имеет место формула полной вероятности*: $\mathsf{P}(A) = \sum_{i=1}^{\infty} \mathsf{P}(H_i)\,\mathsf{P}(A|H_i)$.

Доказательство. Прежде всего, заметим, что $\cup_{i=1}^{\infty} H_i = \Omega$. Отсюда имеем $A = \cup_{i=1}^{\infty}(A \cap H_i)$ и $\mathsf{P}(A) = \mathsf{P}\big(\cup_{i=1}^{\infty}(A \cap H_i)\big)$. Так как события H_1, $H_2, \ldots$ попарно несовместны, т.е. $H_i \cap H_j = \varnothing$ при $i \neq j$, то и их части $A \cap H_i$, $i = 1, 2, \ldots$, также будут попарно несовместны. Учитывая все это, а также аксиому счетной аддитивности, условие $\mathsf{P}(H_i) > 0$ для $i \geqslant 1$ и теорему умножения, получим, что $\mathsf{P}(A) = \mathsf{P}\big(\cup_{i=1}^{\infty}(A \cap H_i)\big) = \sum_{i=1}^{\infty} \mathsf{P}(A \cap H_i) = \sum_{i=1}^{\infty} \mathsf{P}(H_i)\,\mathsf{P}(A|H_i)$.

Рассмотрим теперь несколько другую формулировку теоремы о полной вероятности. Пусть последовательность H_1, $H_2, \ldots$ составлена из попарно несовместимых событий, событие $A \subset \cup_{i=1}^{\infty} H_i \subset \Omega$ и $\mathsf{P}(H_i) > 0$ для всех $i \geqslant 1$. Тогда случайное событие $A = \cup_{i=1}^{\infty}(A \cap H_i)$, т.е. событие A может осуществиться с одним и только с одним из событий H_1, $H_2, \ldots$. Далее доказываем формулу полной вероятности аналогичным образом. Таким образом, последовательность H_1, $H_2, \ldots$ может и не образовывать полную группу. Более того, если предварительно ввести следующее дополнительное событие $H_0 = \Omega \setminus \cup_{i=1}^{\infty} H_i$ и учесть, что $\mathsf{P}(A \cap H_0) = 0$, то события H_0, $H_1, \ldots$ уже будут гипотезами. Приведем простые примеры на применение формулы из теоремы 3.3.

Пример 3.13. Имеется пять урн. Первая урна состоит из двух белых и одного черного шара, вторая — состоит только из 10 черных шаров, третья урна содержит два белых шара и один черный, четвертая — три белых и один черный, и, наконец, пятая урна также содержит три белых и один черный шар. Наудачу выбирается урна, и из нее наудачу вынимается шар. Определить вероятность того, что вынутый шар будет белый.

Решение. Пусть H_i — событие, которое заключается в выборе урны с номером i, где $i = 1, 2, 3, 4, 5$. События H_1, H_2, H_3, H_4, H_5 являются несовместными и равновозможными. Пусть A есть событие, заключающееся в том, что вынут белый шар из какой-либо урны. Итак, событие A может произойти только с одним из событий H_1, H_2, H_3, H_3, H_5. Поэтому имеем $A = (A \cap H_1) \cup (A \cap H_2) \cup (A \cap H_3) \cup (A \cap H_4) \cup (A \cap H_5)$. Тогда по формуле полной вероятности легко вычислить, что

$$\mathsf{P}(A) = \sum_{i=1}^{5} \mathsf{P}(H_i)\,\mathsf{P}(A|H_i) = (1/5) \sum_{i=1}^{5} \mathsf{P}(A|H_i) =$$
$$= (1/5) \times (2/3 + 0 + 2/3 + 3/4 + 3/4) = 17/30.$$

Пример 3.14. В урне находится $n = 20$ шаров, из которых $m = 9$ белых, а остальные черные. Из урны последовательно извлекаются три шара без возвращения, и при каждом излечении определяется цвет шара. Какова вероятность появления белого шара при третьем извлечении?

Решение. Введем следующие четыре события:

1) гипотеза H_1 состоит в том, что первый вынутый шар белый и одновременно второй шар белый;
2) гипотеза H_2 состоит в том, что первый шар белый и одновременно второй шар черный;
3) гипотеза H_3 состоит в том, что первый шар черный и одновременно второй шар белый;
4) гипотеза H_4 состоит в том, что первый шар черный и одновременно второй шар черный.

Теперь с помощью теоремы умножения определим вероятности гипотез H_1, H_2, H_3, H_4. Так, $\mathsf{P}(H_1) = (m/n) \times [(m-1)/(n-1)]$. В самом деле, вероятность $\mathsf{P}(H_1) = \mathsf{P}(B_1)\,\mathsf{P}(B_2|B_1)$, где B_1 — событие, заключающееся в появлении белого шара при первом извлечении, и B_2 — появление белого шара при втором извлечении. Здесь вероятность $\mathsf{P}(B_1) = m/n$, вероятность $\mathsf{P}(B_2|B_1) = (m-1)/(n-1)$. Отсюда $\mathsf{P}(H_1) = (m/n) \times [(m-1)/(n-1)]$. Аналогично определяются вероятности остальных гипотез: $\mathsf{P}(H_2) = (m/n) \times [(n-m)/(n-1)]$, $\mathsf{P}(H_3) = [(n-m)/n] \times [m/(n-1)]$, а также $\mathsf{P}(H_4) = [(n-m)/n] \times \times [(n-m-1)/(n-1)]$. Если событие A есть появление белого шара при третьем извлечении, то условная вероятность $\mathsf{P}(A|H_i)$ при каждом $i = 1, 2, 3, 4$ определяется построением соответствующей модели для условных экспериментов. В результате применения этого способа получим, что $\mathsf{P}(A|H_1) = (m-2)/(n-2)$, $\mathsf{P}(A|H_2) = (m-1)/(n-2)$, $\mathsf{P}(A|H_3) = (m-1) \times (n-2)$ и, наконец, $\mathsf{P}(A|H_4) = m/(n-2)$. Теперь по формуле полной вероятности искомая вероятность $\mathsf{P}(A) = = \mathsf{P}(H_1)\,\mathsf{P}(A|H_1) + \mathsf{P}(H_2)\,\mathsf{P}(A|H_2) + \mathsf{P}(H_3)\,\mathsf{P}(A|H_3) + \mathsf{P}(H_4)\,\mathsf{P}(A|H_4) = = m/n = 0{,}45$.

Перейдем к рассмотрению теоремы о гипотезах или теоремы Байеса.

Теорема 3.4. *Пусть выполняются все условия теоремы* 3.3 *и еще* $\mathsf{P}(A) > 0$. *Тогда для любого* $k = 1, 2, \ldots$ *имеет место формула Байеса*:

$$\mathsf{P}(H_k|A) = \mathsf{P}(H_k)\,\mathsf{P}(A|H_k) \Big/ \sum_{i=1}^{\infty} \mathsf{P}(H_i)\,\mathsf{P}(A|H_i).$$

Доказательство. Учитывая условие $\mathsf{P}(A) > 0$, определение условной вероятности, теорему умножения и формулу полной вероятности, для каждого фиксированного $k = 1, 2, \ldots$ находим:

$$\mathsf{P}(H_k|A) = \mathsf{P}(H_k \cap A)/\mathsf{P}(A) = (\mathsf{P}(H_k)\,\mathsf{P}(A|H_k))/\left(\sum_{i=1}^{\infty} \mathsf{P}(H_i)\,\mathsf{P}(A|H_i)\right).$$

Это выражение называют формулой Байеса, и теорема о гипотезах доказана.

Полезно отметить, что функция $\mathsf{P}(H_i)$ одна и та же, изменяется здесь только аргумент — событие H_i из σ-алгебры $\mathcal{F}$. Но функции $\mathsf{P}(A|H_i)$, вообще говоря, различные. Здесь для каждого события H_i определяется своя функция $\mathsf{P}(A|H_i)\colon \mathcal{F} \to [0,1]$. При этом аргументы у этих функций рассматриваются одинаковые и равные событию A из σ-алгебры $\mathcal{F}$. Далее, смысл формулы Байеса таков, что если наблюдали событие A, то эту информацию можно использовать для уточнения вероятностей событий — гипотез H_1, $H_2, \ldots$. Рассмотрим пример такой ситуации.

Пример 3.15. Персональные компьютеры собираются из высококачественных элементов и из элементов обычного качества. Известно, что 80 % компьютеров собирается из высококачественных элементов и 20 % — из схем обычного качества. Если компьютер собран из высококачественных схем, то его надежность, или вероятность безотказной работы в течение некоторого времени t, равна 0,9. Если компьютер собран из элементов обычного качества, то его надежность равна 0,5. При испытании в течение времени t компьютер работал безотказно. Определить вероятность того, что компьютер собран из высококачественных элементов.

Решение. Пусть H_1 и H_2 суть события, которые означают, что компьютер собран из высококачественных элементов и, соответственно, из элементов обычного качества. Наконец, событие A состоит в том, что компьютер безотказно работал в течение времени t. Тогда из условия задачи сразу имеем: $\mathsf{P}(H_1) = 4/5$, $\mathsf{P}(H_2) = 1/5$, $\mathsf{P}(A|H_1) = 9/10$, $\mathsf{P}(A|H_2) = 1/2$.

Используя эти условия и формулу Байеса, пересчитаем вероятность H_1: $\mathsf{P}(H_1|A) = [\mathsf{P}(H_1)\,\mathsf{P}(A|H_1)]/[\mathsf{P}(H_1)\,\mathsf{P}(A|H_1) + \mathsf{P}(H_2)\,\mathsf{P}(A|H_2)] = 0{,}72/(0{,}8 \times 0{,}9 + 0{,}2 \times 0{,}5) = 0{,}878 \approx 0{,}88$. Полученный результат имеет следующий вероятностный смысл. При эксплуатации достаточно большой партии компьютеров в течение времени t в среднем 88% компьютеров, успешно прошедших испытания, были собраны из схем высокого качества. В то же время напомним, что общий процент компьютеров, собранных из высококачественных схем, составляет лишь 80 % из всей партии.

Приведем теперь показательные примеры решений типовых задач с использованием формулы полной вероятности и формулы Байеса.

Пример 3.16 (*задача о легком экзаменационном билете*). Преподаватель по теории вероятностей составил для r студентов $n > r$ экзаменационных билетов, среди которых m легких. Билеты вытаскиваются студентами один за другим по очереди в порядке их прибытия на экзамен случайным образом и без возвращения. Найти вероятность того, что второй извлеченный билет окажется легким.

Пусть гипотеза H_1 означает появление легкого билета для студента, который пришел на экзамен первым, а гипотеза H_2 появление у такого студента трудного билета. Вероятности событий H_1 и H_2 легко вычисляются с использованием классического подхода следующим образом: $\mathsf{P}(H_1) = m/n$, $\mathsf{P}(H_2) = (n-m)/n$. Если событие A есть появление легкого билета для студента, который был в очереди вторым, то $A = (A \cap H_1) \cup (A \cap H_2)$. Отсюда по формуле полной вероятности находим, что вероятность $\mathsf{P}(A) = \mathsf{P}(H_1)\,\mathsf{P}(A|H_1) + \mathsf{P}(H_2)\,\mathsf{P}(A|H_2)$. Используя способ построения вероятностых моделей $(H_1, \mathcal{F} \cap H_1, \mathsf{P}(\cdot|H_1))$ и $(H_2, \mathcal{F} \cap H_2, \mathsf{P}(\cdot|H_2))$ для условных экспериментов и классическое определение вероятности, без труда вычислим условные вероятности появления легкого билета у второго студента при различных гипотезах H_1 и H_2: $\mathsf{P}(A|H_1) = (m-1)/(n-1)$, $\mathsf{P}(A|H_2) = m/(n-1)$. Окончательно, вероятность $\mathsf{P}(A) = (m/n) \times [(m-1)/(n-1)] + [(n-m)/n] \times$ $\times\, [m/(n-1)] = m/n$.

Итак, вероятность вытащить легкий билет для первого студента или второго оказывается одна и та же. Аналогично, можно показать, что вероятность вытащить легкий билет для s-го $(2 < s \leqslant r)$ студента равна m/n. Следовательно, не играет роли в какое время идти студенту на экзамен с точки зрения появления у него легкого билета, и он может не нарушать свой привычный распорядок сна в день сдачи экзамена.

Заметим, что этот пример можно решить, используя классический способ вычисления вероятностей случайных событий (см. решенную ранее задачу о выборе престижной квартиры, пример 2.32). Однако формула полной вероятности иногда позволяет получить ответ значительно быстрее, минуя построение теоретико-множественной модели $(\Omega, \mathcal{F})$ для данного эксперимента.

Пример 3.17. Страховая компания разделяет застрахованных лиц по трем классам риска: первый класс — имеется малый риск, второй класс — имеется средний риск, третий класс — имеется большой риск. Среди всех клиентов компании 50 % составляют первый класс риска, 30 % — второй и 20 % — третий. Вероятность наступления страхового случая для первого класса риска равна 0,01; второго — 0,03 и третьего — 0,08. Какова вероятность того, что клиент, получивший денежное вознаграждение за период страхования, относится к группе малого риска?

Решение. Пусть событие A означает, что клиент компании получил вознаграждение. Событие A может наступить лишь совместно с одним из трех следующих попарно несовместных событий. Гипотеза H_1 означает, что клиент относится к первому классу риска. Гипотеза H_2 означает, что клиент относится ко второму классу риска. Наконец, гипотеза H_3 означает, что клиент относится к третьему классу риска. Необходимо определить условную вероятность $\mathsf{P}(H_1|A)$. Из условия задачи определяем вероятности гипотез: $\mathsf{P}(H_1) = 0{,}5$; $\mathsf{P}(H_2) = 0{,}3$; $\mathsf{P}(H_3) = 0{,}2$. Известны также условные вероятности: $\mathsf{P}(A|H_1) = 0{,}01$; $\mathsf{P}(A|H_2) = 0{,}03$ и $\mathsf{P}(A|H_3) = 0{,}08$. Искомую вероятность вычисляем по формуле Байеса, т. е.

$$\mathsf{P}(H_1|A) = \frac{\mathsf{P}(H_1)\,\mathsf{P}(A|H_1)}{\mathsf{P}(H_1)\,\mathsf{P}(A|H_1) + \mathsf{P}(H_2)\,\mathsf{P}(A|H_2) + \mathsf{P}(H_3)\,\mathsf{P}(A|H_3)} =$$

$$= 5/30 = 1/6 < \mathsf{P}(H_1) = 0{,}5.$$

§ 3. Статистически независимые события эксперимента

3.1. Причинная независимость исходов одного и того же эксперимента и различных экспериментов на содержательном уровне. Общеязыковой смысл понятия причинной независимости исходов A и B реального эксперимента E означает, что осуществление события B не влияет на возможность наступления события A или $\overline{A}$. Можно сказать, что свойства условий проведения эксперимента E, которые отвечают за наступление события A, совершенно безразличны к осуществлению исхода B или $\overline{B}$. В этом случае для независимых исходов A и B эксперимента E при достаточно большом числе N его испытаний с большой уверенностью можно ожидать, что отношение $\mu(A\cap B,N)/\mu(B,N)$ приближенно равно отношению $\mu(A\cap\overline{B},N)/\mu(\overline{B},N)$. Приближенное равенство

$$\mu(A\cap B,N)/\mu(B,N) \approx \mu(A\cap\overline{B},N)/\mu(\overline{B},N) \tag{3.4}$$

для независимых реальных событий установлено эмпирически многовековой практикой человека. Итак, для независимых исходов A и B относительная частота случаев появления результата $A\cap B$ по отношению к числу появления результата B и относительная частота случаев появления результата $A\cap\overline{B}$ по отношению к числу появления результата $\overline{B}$ приблизительно совпадают. Выпишем следующие очевидные равенства:

$$\mu(B,N) + \mu(\overline{B},N) = N, \quad \mu(A,N) = \mu(A\cap B,N) + \mu(A\cap\overline{B},N). \tag{3.5}$$

Из (3.4), используя (3.5), легко находим, что

$$\mu(A \cap B, N)/\mu(B, N) \approx [\mu(A, N) - \mu(A \cap B, N)]/[N - \mu(B, N)],$$
$$\mu(A \cap B, N) \times N - \mu(A \cap B, N) \times \mu(B, N) \approx$$
$$\approx \mu(B, N) \times \mu(A, N) - \mu(B, N) \times \mu(A \cap B, N),$$
$$\mu(A \cap B, N) \times N \approx \mu(B, N) \times \mu(A, N),$$
$$\mu(A \cap B, N)/N \approx (\mu(A, N)/N) \times (\mu(B, N)/N).$$

Учитывая теперь обозначение для частоты произвольного события при определении статистической вероятности, получим основополагающее соотношение

$$\mathsf{P}^*(A \cap B) \approx \mathsf{P}^*(A) \times \mathsf{P}^*(B), \tag{3.6}$$

если только N достаточно велико.

Очень трудно на интуитивном уровне выявить независимые исходы A и B некоторого статистически устойчивого эксперимента E. Например, при непреднамеренном одноразовом подбрасывании симметричной игральной кости исходы, состоящие в выпадении нечетного числа очков (событие A) и выпадении числа очков, кратного трем (событие B), не представляются на первый взгляд независимыми. Однако многочисленные опыты с этим экспериментом подтверждают справедливость соотношения (3.6) и, следовательно, эти события мы должны считать независимыми. Но настоящая экзотика в нашей интуиции наступает, если рассмотреть для этого эксперимента новые три пары исходов:

1) выпадение четного числа очков (результат A_1) и выпадение числа очков, кратного трем (результат B);
2) выпадение четного числа очков (результат A_1) и выпадение числа очков, некратного трем (событие B_1);
3) выпадение нечетного числа очков (результат A) и выпадение числа очков, некратного трем (результат B_1).

Опытным путем легко убедиться, что для этих пар при достаточно большом числе N также имеет место условие (3.6). Для этого примера, как и для многих других, понятие независимости событий может показаться несколько искусственным. На содержательном уровне не так просто себе вообразить, что одни и те же условия этого или другого эксперимента некоторые его исходы A и B делают безразличными по отношению друг к другу в смысле их наступления или ненаступления.

Настоящая сфера рассмотрения и применения интуитивного понятия независимости относится к случаю, когда приходится одновременно или последовательно проводить несколько причинно-независимых экспериментов E_1, $E_2, \ldots$, E_m. Например, при игре в спортивный бридж разными судьями происходит раздача двух колод карт в разных комнатах. В этой игре участвуют две команды, каждая из которых включает по четыре спортсмена. По два спортсмена из каждой команды находятся в разных комнатах. Легко представить здесь причинную независимость комбинаций розданных карт для игроков в одной комнате (эксперимент E_1) от комбинаций розданных карт для игроков в другой комнате (эксперимент E_2). Приведем еще более простой пример проведения двух причинно-независимых экспериментов E_1 и E_2.

Пусть в одной комнате бросается симметричная монета, а в другой комнате подбрасывается симметричная игральная кость. Совершенно ясно, что выпадение той или другой стороны монеты не оказывает никакого влияния на выпадение той или иной грани игральной кости. Значит, выпадение, например, герба на монете (событие A) и выпадение четного числа очков на игральной кости (событие B) очевидно всеми признается независимыми.

Рассмотрим теперь два причинно-независимых статистически устойчивых эксперимента E_1 и E_2 произвольной природы. Пусть эксперименту E_1 с описанием ω_1 его произвольного элементарного исхода соответствует теоретико-множественная модель $(\Omega_1, \mathcal{F}_1)$, а эксперименту E_2 с описанием ω_2 его произвольного элементарного исхода соответствует теоретико-множественная модель $(\Omega_2, \mathcal{F}_2)$. Заметим, что $\omega_1 \in \Omega_1$ и $\omega_2 \in \Omega_2$. Проведем эксперимент E_1 ровно N_1 раз и независимым образом эксперимент E_2 в точности N_2 раз. Введем в рассмотрение произвольные исходы $A \in \mathcal{F}_1$ и $B \in \mathcal{F}_2$ эксперимента E_1 и, соответственно, эксперимента E_2. Предположим, что для исхода A число всех его появлений при N_1 испытаниях первого эксперимента равно $\mu_1(A, N_1)$. Аналогично, для исхода B число его появлений при N_2 испытаниях второго эксперимента равно $\mu_2(B, N_2)$.

Фиксированное испытание с номером $i = 1, 2, \ldots, N_1$ эксперимента E_1 можно рассматривать с любым номером $j = 1, 2, \ldots, N_2$ испытания эксперимента E_2. Наоборот, фиксированное испытание с номером $j = 1, 2, \ldots, N_2$ эксперимента E_2 можно рассматривать с любым номером $i = 1, 2, \ldots, N_1$ испытания эксперимента E_1. Тогда испытание с номером i эксперимента E_1 и испытание с номером j эксперимента E_2 можно представить как испытание с векторным номером (i, j) совместного и более сложного эксперимента $E = \{E_1, E_2\}$. Если при этом испытание с номером i эксперимента E_1 дает некоторый его исход A и испытание с номером j эксперимента E_2 дает некоторый его исход B, то будем считать, что испытание с векторным номером (i, j) совместного эксперимента $E = \{E_1, E_2\}$ дает некоторый его результат, обозначаемый символом $A \times B$.

Например, исход $A \times \Omega_2$ в некотором испытании для совместного эксперимента $E = \{E_1, E_2\}$ означает появление результата A для первого эксперимента и любого события для второго эксперимента. Аналогично, результат вида $\Omega_1 \times B$ в некотором испытании для совместного эксперимента $E = \{E_1, E_2\}$ означает появление любого события для первого эксперимента и исхода B для второго эксперимента. Напомним, что исход Ω_1 является достоверным для эксперимента E_1, а исход Ω_2 является достоверным для эксперимента E_2. В силу этого $\Omega = \Omega_1 \times \Omega_2$ будет, очевидно, достоверным исходом для совместного эксперимента $E = \{E_1, E_2\}$. Эти рассуждения позволяют в качестве теоретико-множественной модели эксперимента $E = \{E_1, E_2\}$ принять наиболее простую модель вида $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$, которая подробно рассматривалась в первой главе.

Легко подсчитать, что общее число испытаний, которое соответствует совместному эксперименту $E = \{E_1, E_2\}$ здесь будет равно $N = N_1 \times N_2$. При проведении указанным выше способом ровно $N = N_1 \times N_2$ раз совместного эксперимента $E = \{E_1, E_2\}$ событие вида $A \times \Omega_2$ появлялось ровно $\mu(A \times \Omega_2, N) = \mu_1(A, N_1) \times N_2$ раз, событие $\Omega_1 \times B$ появлялось очевидно ровно $\mu(\Omega_1 \times B, N) = \mu_2(B, N_2) \times N_1$ раз и, наконец, событие $A \times B$ (одновременно и событие $A \in \mathcal{F}_1$ и событие $B \in \mathcal{F}_2$ наблюдалось $\mu(A \times B, N) = \mu_1(A, N_1) \times \mu_2(B, N_2)$ раз. Здесь следует предостеречь от тривиальной ошибки, когда событие $A \times B$ реального эксперимента $E = \{E_1, E_2\}$ выдают за несуществующий объект $A \cap B$. Напомним, что операция пересечения применяется только к событиям из одной и той же σ-алгебры, а мы имеем $A \in \mathcal{F}_1$ и $B \in \mathcal{F}_2$.

Итак, для относительных частот исходов $A \times B$, $A \times \Omega_2$, $\Omega_1 \times B$ эксперимента $E = \{E_1, E_2\}$ выполняются равенства вида:

$$\begin{aligned}\mu(A \times B, N)/N &= (\mu_1(A, N_1)\,\mu_2(B, N_2))/N = \\ &= (\mu_1(A, N_1)/N_1)\,(\mu_2(B, N_1)/N_2) = \\ &= (\mu_1(A, N_1)N_2/N_1N_2)\,(\mu_2(B, N_2)N_1/N_1N_2) = \\ &= (\mu(A \times \Omega_2, N/N))\,(\mu(\Omega_1 \times B, N)/N).\end{aligned}$$

Используя обозначение частоты произвольного события при определении статистической вероятности, получим для совместного эксперимента $E = \{E_1, E_2\}$ и отдельных экспериментов E_1, E_2 соотношение $\mathsf{P}^*(A \times B) = \mathsf{P}_1^*(A)\,\mathsf{P}_2^*(B) = \mathsf{P}^*(A \times \Omega_2)\,\mathsf{P}^*(\Omega_1 \times B)$. Напомним, что символы $\mathsf{P}^*(A \times B)$, $\mathsf{P}_1^*(A)$ и $\mathsf{P}_2^*(B)$ обозначают статистические вероятности, соответствующие экспериментам E, E_1 и E_2. Так как $A \times B = \{\omega = (\omega_1, \omega_2) \colon \omega_1 \in A, \omega_2 \in B\} = (A \times \Omega_2) \cap (\Omega_1 \times B)$, то окончательно для статистических вероятностей (относительных частот) получаем следующие равенства

$$\begin{aligned}\mathsf{P}^*((A \times \Omega_2) \cap (\Omega_1 \times B)) = \mathsf{P}^*(A \times \Omega_2) \cdot \mathsf{P}^*(\Omega_1 \times B),\\ \mathsf{P}^*((A \times \Omega_2) \cap (\Omega_1 \times B)) = \mathsf{P}_1^*(A) \cdot \mathsf{P}_2^*(B).\end{aligned} \tag{3.7}$$

3.2. Математическое описание независимости случайных событий и его фундаментальная роль в теории вероятностей. Итак, для причинно-независимых статистически устойчивых экспериментов E_1 и E_2 произвольной природы получены удивительные равенства (3.7). В конечном результате мы имеем фундаментальные соотношения (3.6) и (3.7) для статистических вероятностей причинно-независимых событий и общее свойство устойчивости относительных частот. Это позволяет в рамках вероятностной модели $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ произвольного статистически устойчивого эксперимента E математически формализовать интуитивное понятие независимости исходов. С этой целью введем определение.

Определение 3.2. События A, B из $\mathcal{F}$ называются *статистически независимыми* относительно вероятности $\mathsf{P}(\cdot)$ или *попарно независимыми*, если $\mathsf{P}(A \cap B) = \mathsf{P}(A)\,\mathsf{P}(B)$. В противном случае они называются *зависимыми*.

Используя это определение, докажем, что из независимости событий A и B следует независимость $\overline{A}$ и B, A и $\overline{B}$, $\overline{A}$ и $\overline{B}$. Действительно, находим, что $\overline{A} \cap B = B \backslash (A \cap B)$ и вероятность $\mathsf{P}(\overline{A} \cap B) = \mathsf{P}(B) - \mathsf{P}(A \cap B) = \mathsf{P}(B) - \mathsf{P}(A)\,\mathsf{P}(B) = \mathsf{P}(B)(1 - \mathsf{P}(A)) = \mathsf{P}(B)\,\mathsf{P}(\overline{A})$. Таким образом, события $\overline{A}$ и B являются статистически независимыми. Это предложение можно показать еще более простым способом. Так как $A \cup \overline{A} = \Omega$ и $B \cap (A \cup \overline{A}) = B$, то $\mathsf{P}(B) = \mathsf{P}(B \cap A) + \mathsf{P}(B \cap \overline{A})$. Поскольку $\mathsf{P}(B \cap A) = \mathsf{P}(A)\,\mathsf{P}(B)$, то $\mathsf{P}(B) = \mathsf{P}(A)\,\mathsf{P}(B) + \mathsf{P}(B \cap \overline{A})$ и $\mathsf{P}(B \cap \overline{A}) = \mathsf{P}(B) - \mathsf{P}(A)\,\mathsf{P}(B) = \mathsf{P}(B)(1 - \mathsf{P}(A)) = \mathsf{P}(B)\,\mathsf{P}(\overline{A})$. В результате имеем $\mathsf{P}(B \cap \overline{A}) = \mathsf{P}(B)\,\mathsf{P}(\overline{A})$. Для остальных пар вида A и $\overline{B}$, $\overline{A}$ и $\overline{B}$ доказательство проверяется аналогичным способом.

Приведем теперь важное определение статистической независимости событий в совокупности.

Определение 3.3. События $A_1, A_2, \ldots, A_n$ называются *независимыми в совокупности*, если для любого набора $A_{j_1}, A_{j_2}, \ldots, A_{j_s}$ событий из семейства $\{A_1, A_2, \ldots, A_n\}$ выполняется равенство

$$\mathsf{P}(A_{j_1} \cap A_{j_2} \cap \ldots \cap A_{j_s}) = \mathsf{P}(A_{j_1}) \times \mathsf{P}(A_{j_2}) \times \ldots \times \mathsf{P}(A_{j_s}),$$

где $j_1 < j_2 < \ldots < j_s$, $1 < s \leqslant n$.

Для $n = 2$ при $\mathsf{P}(A_1) > 0$, $\mathsf{P}(A_2) > 0$ и в случае статистической независимости A_1 и A_2 с использованием формулы (3.1) находим, что

$$\begin{aligned} \mathsf{P}(A_1|A_2) &= \frac{\mathsf{P}(A_1 \cap A_2)}{\mathsf{P}(A_2)} = \frac{\mathsf{P}(A_1) \times \mathsf{P}(A_2)}{\mathsf{P}(A_2)} = \mathsf{P}(A_1), \\ \mathsf{P}(A_2|A_1) &= \frac{\mathsf{P}(A_1 \cap A_2)}{\mathsf{P}(A_1)} = \frac{\mathsf{P}(A_1) \times \mathsf{P}(A_2)}{\mathsf{P}(A_1)} = \mathsf{P}(A_2). \end{aligned} \tag{3.8}$$

Таким образом, в случае независимости случайных событий A_1 и A_2 их условные и безусловные вероятности совпадают, т. е. осуществление события A_2 не изменяет вероятности появления события A_1. В этом и заключается содержательный смысл статистической независимости событий. Однако следует заметить, что равенства (3.8) имеет место только при $\mathsf{P}(A_1) \neq 0$, $\mathsf{P}(A_2) \neq 0$, и поэтому их неразумно использовать в качестве определения статистической независимости. Более того, пусть для проверки статистической независимости событий A_1 и A_2 мы используем определение независимости в виде равенства $\mathsf{P}(A_1|A_2) = \mathsf{P}(A_1)$, статистическое определение условной и безусловной вероятности, и приближенное равенство $\mathsf{P}^*(A_1|A_2) \approx \mathsf{P}^*(A_1)$. Тогда на практике необходимо было бы проводить условный эксперимент E_y. Как правило, последнее обстоятельство вызывает значительные трудности.

Рассмотрим теперь простую связь между совместностью и независимостью событий A и B. Пусть события A и B несовместны и $\mathsf{P}(B) \neq 0$, тогда имеем $A \cap B = \varnothing$, $\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) = 0$, и если $\mathsf{P}(A) \neq 0$, то события A и B зависимы, а если $\mathsf{P}(A) = 0$, то A и B независимы.

Статистическая независимость является одним из фундаментальных понятий, и это понятие играет центральную роль в различных приложениях теории вероятностей. Определения 3.2 и 3.3 фактически формализуют интуитивный смысл независимости исходов многих реальных явлений и процессов. Однако статистическая независимость не тождественна причинной независимости, которую мы часто употребляем и используем на интуитивном уровне в обыденной жизни.

Особенности статистической независимости случайных событий отражение некоторых основных свойств их причинной независимости. Прежде всего, заметим, что если пара $(\Omega, \mathcal{F})$ фиксирована, то некоторые события A и B из $\mathcal{F}$ могут быть статистически независимыми относительно одной вероятностной функции и в то же время оказаться статистически зависимыми относительно другой вероятностной функции. Следующий пример подтверждает это вполне очевидное замечание.

Пример 3.18. Пусть опыт заключается в случайном бросании симметричной игральной кости, грани которой помечены цифрами от 1 до 6. Определяется метка (цифра) выпавшей грани. В этом эксперименте комплекс условий $\Sigma = \{u_1$ — выбрана некоторая симметричная игральная кость, u_2 — игральная кость подбрасывается один раз, u_3 — непреднамеренный механизм подбрасывания игральной кости, u_4 — поверхность стола, на которую падает игральная кость, u_5 — освещенность поверхности стола, позволяющая различать цифры, $\ldots\}$.

При каждом $i = 1, 2, \ldots, 6$ символом ω_i обозначим описание такого элементарного исхода этого опыта, когда при броске выпадает грань с цифрой i. Поэтому множество Ω содержит 6 описаний всех элементарных исходов, т. е. $\Omega = \{\omega_1, \omega_2, \ldots, \omega_6\}$. Пусть $\mathcal{F}$ — множество всех подмножеств множества Ω, т. е. $\mathcal{F} = \{A_1, A_2, \ldots, A_{64}\}$. Рассмотрим исходы A и B этого эксперимента, которые заключаются соответственно в появлении четной грани и в появлении любой грани с цифрой от 3 до 6. Тогда $A = \{\omega_2, \omega_4, \omega_6\}$, $B = \{\omega_3, \omega_4, \omega_5, \omega_6\}$, $A \cap B = \{\omega_4, \omega_6\}$ и, следовательно, используя классическое определение вероятности, легко находим $\mathsf{P}(A) = 3/6 = 1/2$, $\mathsf{P}(B) = 4/6 = 2/3$, $\mathsf{P}(A \cap B) = 2/6 = 1/3 = \mathsf{P}(A)\,\mathsf{P}(B)$. Значит, в этом опыте события A и B независимы. Если же у этой игральной кости так нарушить симметрию, что $\mathsf{P}(\{\omega_1\}) = \mathsf{P}(\{\omega_3\}) = \mathsf{P}(\{\omega_4\}) = \mathsf{P}(\{\omega_5\}) = \mathsf{P}(\{\omega_6\}) = 4/25$, $\mathsf{P}(\{\omega_2\}) = 1/5$, то вероятность $\mathsf{P}(A) = \mathsf{P}(\{\omega_2, \omega_4, \omega_6\}) = \mathsf{P}(\{\omega_2\}) + \mathsf{P}(\{\omega_4\}) + \mathsf{P}(\{\omega_6\}) = 13/25$, $\mathsf{P}(B) = \mathsf{P}(\{\omega_3, \omega_4, \omega_5, \omega_6\}) = \mathsf{P}(\{\omega_3\}) + \mathsf{P}(\{\omega_4\}) + \mathsf{P}(\{\omega_5\}) + \mathsf{P}(\{\omega_6\}) = 16/25$, $\mathsf{P}(A \cap B) = \mathsf{P}(\{\omega_4, \omega_6\}) = 8/25$. Поскольку в этом случае $\mathsf{P}(A \cap B) \neq \mathsf{P}(A)\,\mathsf{P}(B)$, то события A и B для несимметричной игральной кости уже станут зависимыми. Легко видеть, что эксперимент с симметричной игральной костью и эксперимент с несимметричной игральной костью имеют одну и ту же теоретико-множественную модель $(\Omega, \mathcal{F})$. Более интересным является такой факт, когда совсем незначительное изменение условий эксперимента приводит к изменению $(\Omega, \mathcal{F})$ и самое главное к потере независимости событий. Проиллюстрируем это на примере.

Пример 3.19. Предположим, что имеется стандартная колода из 36 карт. Каждая карта различается одной из четырех мастей (пики, трефы, червы, бубны) и одним из девяти значений (6, 7, 8, 9, 10, валет, дама, король, туз). Эксперимент состоит в том, чтобы случайно выбрать одну карту. Рассмотрим исходы A и B этого опыта, которые заключаются в появлении короля и соответственно в появлении карты пиковой масти. Тогда событие $A \cap B$ означает, что вынут король пиковой масти. В этом простом эксперименте имеем $\mathsf{P}(A) = 4/36 = 1/9$, $\mathsf{P}(B) = 9/36 = 1/4$, $\mathsf{P}(A \cap B) = 1/36 = \mathsf{P}(A) \times \mathsf{P}(B)$ и, следовательно, события A и B независимы.

Изменим теперь незначительно этот опыт. Известно, что при игре в покер необходимо добавить в колоду карту джокер. Итак, в колоде уже содержится 37 карт. В последнем опыте события A и B станут зависимыми, так как $\mathsf{P}(A) = 4/37$, $\mathsf{P}(B) = 9/37$, $\mathsf{P}(A \cap B) = 1/37 \neq \mathsf{P}(A) \times \mathsf{P}(B)$.

Из определения 3.3 следует, что для доказательства независимости в совокупности событий A_1, $A_2, \ldots$, A_n достаточно проверить равенство в вероятностях $\mathsf{P}(A_{j_1} \cap A_{j_2} \cap \ldots \cap A_{j_s}) = \mathsf{P}(A_{j_1}) \times \mathsf{P}(A_{j_2}) \times \ldots \times \mathsf{P}(A_{j_s})$ для любых $1 \leqslant j_1 < j_2 < \ldots < j_s \leqslant n$, $1 < s \leqslant n$. Если равенство $\mathsf{P}(A_{j_1} \cap A_{j_2}) = \mathsf{P}(A_{j_1}) \times \mathsf{P}(A_{j_2})$ имеет место для $1 \leqslant j_1 < j_2 \leqslant n$, т. е. для каждого из C_n^2 наборов $\{A_{j_1}, A_{j_2}\}$ событий из последовательности

A_1, $A_2, \ldots$, A_n, то события A_1, $A_2, \ldots$, A_n называют *попарно независимыми*.

Лемма 3.1 (вторая лемма Бореля–Кантелли). *Если* $\{A_i;\ i = 1, 2, \ldots\}$ *последовательность событий из* $\mathcal{F}$ *составлена из независимых в совокупности случайных событий,* $A^* = \cap_{i=1}^{\infty} \cup_{k=i}^{\infty} A_k$ *и* $\sum_{i=1}^{\infty} \mathsf{P}(A_i) = \infty$, *то* $\mathsf{P}(A^*) = 1$.

Доказательство. Так как события A_1, $A_2, \ldots$ независимы, то независимыми будут события $\overline{A}_1$, $\overline{A}_2, \ldots$. Пусть случайное событие $B_i = \cup_{k=i}^{\infty} A_k$ и случайное событие $C_{in} = \cap_{k=i}^{i+n-1} \overline{A}_k$ при $i, n = 1, 2, \ldots$. Тогда $B_1 \supset B_2 \supset \ldots$, $C_{i1} \supset C_{i2} \supset \ldots$ и $\lim\limits_{i\to\infty} B_i = \cap_{i=1}^{\infty} B_i = \cap_{i=1}^{\infty} \cup_{k=i}^{\infty} A_k = A^*$, $\lim\limits_{i\to\infty} C_{in} = \cap_{n=1}^{\infty} C_{in} = \cap_{k=1}^{\infty} \overline{A}_k$. Отсюда в силу первой аксиомы непрерывности, закона де Моргана и независимости событий $\overline{A}_1$, $\overline{A}_2, \ldots$ получим следующие равенства:

$$
\begin{aligned}
\mathsf{P}(A^*) &= \mathsf{P}\Big(\lim_{i\to\infty} B_i\Big) = \lim_{i\to\infty} \mathsf{P}(B_i) = \lim_{i\to\infty} \mathsf{P}\Big(\bigcup_{k=i}^{\infty} A_k\Big) = \\
&= \lim_{i\to\infty} \mathsf{P}\Big(\Omega \setminus \bigcap_{k=i}^{\infty} \overline{A}_k\Big) = 1 - \lim_{i\to\infty} \mathsf{P}\Big(\bigcap_{k=i}^{\infty} \overline{A}_k\Big) = \\
&= 1 - \lim_{i\to\infty} \mathsf{P}\Big(\lim_{n\to\infty} C_{in}\Big) = 1 - \lim_{i\to\infty} \lim_{n\to\infty} \mathsf{P}(C_{in}) = \\
&= 1 - \lim_{i\to\infty} \lim_{n\to\infty} \mathsf{P}\Big(\bigcup_{k=i}^{i+n-1} \overline{A}_k\Big) = 1 - \lim_{i\to\infty} \lim_{n\to\infty} \prod_{k=i}^{i+n-1} \mathsf{P}(\overline{A}_k) = \\
&= 1 - \lim_{i\to\infty} \lim_{n\to\infty} \prod_{k=i}^{i+n-1} (1 - \mathsf{P}(A_k)) = 1 - \lim_{i\to\infty} \prod_{k=i}^{\infty} (1 - \mathsf{P}(A_k)) = 1.
\end{aligned}
$$

Здесь равенство $\lim\limits_{i\to\infty} \prod_{k=i}^{\infty} (1 - \mathsf{P}(A_k)) = 0$ вытекает из свойств бесконечного произведения при условии расходимости ряда $\sum_{i=1}^{\infty} \mathsf{P}(A_i)$, которые рассматриваются в курсе математического анализа.

В заключение доказательства этой леммы с учетом утверждения первой леммы Бореля–Кантелли сформулируем так называемый закон «нуля или единицы» для последовательности $\{A_i;\ i = 1, 2, \ldots\}$ независимых событий. Итак, в случае независимых событий A_1, $A_2, \ldots$ вероятность наступления бесконечного числа указанных событий равна нулю или единице в зависимости от того, сходится или расходится ряд $\sum_{i=1}^{\infty} \mathsf{P}(A_i)$.

3.3. Типы зависимостей случайных событий. Рассмотрим простые примеры различных зависимостей случайных событий A, B и C. Первый пример показывает, что из попарной независимости событий A_1, $A_2, \ldots$, A_n, вообще говоря, не следует их независимость в совокупности.

Пример 3.20 (*пример Бернштейна*). Тетраэдр $dabc$, отображенный на рис. 3.1, изготовлен из однородного материала. Грани dab, dac, dcb, abc окрашены соответственно в белый, синий, красный и, наконец, во все эти три цвета. Тетраэдр подбрасывается случайным образом (непреднамеренно) на поверхность стола с помощью некоторого механизма. Появление грани, которая соприкасается с поверхностью стола, обеспечивает устойчивое положение тетраэдра. Естественно предполагается, что тетраэдр не может устойчиво стоять на любой из его вершин или на любом из его ребер.

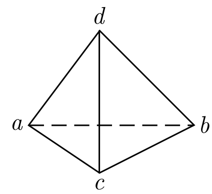


Рис. 3.1

Элементарным исходом эксперимента Бернштейна является выпадение той или иной грани тетраэдра. При этом фиксируется цвет грани, на которую устойчиво падает тетраэдр. Обозначим через ω_1 (ω_2, ω_3, ω_4) описание в некотором языке элементарного исхода данного эксперимента, который заключается в том, что тетраэдр падает на грань dab (dac, dcb, abc). В этом случае достоверный исход

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\},$$
$$\mathcal{F} = \{\Omega, \varnothing, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_1, \omega_4\},$$
$$\{\omega_2, \omega_3\}, \{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_4\},$$
$$\{\omega_1, \omega_3, \omega_4\}, \{\omega_2, \omega_3, \omega_4\}\},$$
$$\mathsf{P}(\{\omega_1\}) = \mathsf{P}(\{\omega_2\}) = \mathsf{P}(\{\omega_3\}) = \mathsf{P}(\{\omega_4\}) = 1/4,$$

и, значит, построена вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ этого классического эксперимента.

Пусть теперь $A = \{\omega_1, \omega_4\}$, $B = \{\omega_2, \omega_4\}$ и $C = \{\omega_3, \omega_4\}$ события из $\mathcal{F}$, которые заключаются в том, что тетраэдр падает на грань, имеющую, соответственно, белый, синий и красный цвет. С использованием классического подхода нетрудно найти вероятности следующих событий: A, B, C, $A \cap B$, $A \cap C$, $B \cap C$, $A \cap B \cap C$. Последовательно имеем:

$$\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = 2/4 = 1/2,$$
$$\mathsf{P}(A \cap B) = \mathsf{P}(\{\omega_4\}) = 1/4 = (1/2) \times (1/2) = \mathsf{P}(A) \times \mathsf{P}(B),$$
$$\mathsf{P}(A \cap C) = \mathsf{P}(\{\omega_4\}) = 1/4 = (1/2) \times (1/2) = \mathsf{P}(A) \times \mathsf{P}(C),$$
$$\mathsf{P}(B \cap C) = \mathsf{P}(\{\omega_4\}) = 1/4 = (1/2) \times (1/2) = \mathsf{P}(B) \times \mathsf{P}(C),$$
$$\mathsf{P}(A \cap B \cap C) = \mathsf{P}(\{\omega_4\}) = 1/4 \neq \mathsf{P}(A) \times \mathsf{P}(B) \times \mathsf{P}(C) = 1/8.$$

Итак, в этом примере имеем попарную независимость. Однако из последнего соотношения следует, что для событий A, B, C не выполняется независимость в совокупности. Следовательно, события A, B

и C являются зависимыми в совокупности, более того, из физических соображений имеем очевидное равенство $\mathsf{P}(A|B \cap C) = 1 \neq \mathsf{P}(A)$. Значит, условная и безусловная вероятности события A не совпадают.

Пример 3.21. Из урны, содержащей три белых и семь красных шаров, наудачу последовательно и без возвращения извлекаются два шара. Пусть события A, B и C соответственно означают, что первый шар извлечен белый, второй шар вынут белый и, наконец, по крайней мере, один из вынутых шаров белый. Вычислить вероятности $\mathsf{P}(A|B)$, $\mathsf{P}(B|A)$, $\mathsf{P}(A|C)$. Являются ли независимыми в совокупности события A, B и C ?

Решение. Занумеруем белые шары числами 1, 2, 3, а красные шары — числами 4, 5,..., 10. Тогда множество элементарных исходов есть $\Omega = \{\omega = (i,j)\colon i,j = 1,2,\ldots,10;\ i \neq j\}$. Отсюда следует, что число всех описаний равновозможных элементарных исходов равно $A_{10}^2 = 90$. Теперь события A, B и $A \cap B$ можно представить:

$$\begin{aligned} A &= \{\omega = (i,j) \in \Omega\colon i = 1,2,3;\ j = 1,2,\ldots,10;\ i \neq j\},\\ B &= \{\omega = (i,j) \in \Omega\colon i = 1,2,\ldots,10;\ j = 1,2,3;\ i \neq j\},\\ A \cap B &= \{\omega = (i,j) \in \Omega\colon i = 1,2,3;\ j = 1,2,3;\ i \neq j\}. \end{aligned}$$

Поэтому $\mathsf{P}(A) = 0{,}3$, $\mathsf{P}(B) = 0{,}3$ и $\mathsf{P}(A \cap B) = 1/15$. Для вычисления условных вероятностей воспользуемся формулой вида $\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B)$. Далее находим $\mathsf{P}(A|B) = \mathsf{P}(A \cap B)/\mathsf{P}(B) = 2/9$, $\mathsf{P}(B|A) = \mathsf{P}(A \cap B)/\mathsf{P}(A) = 2/9$. Так как $A \subset C$, то $\mathsf{P}(A \cap C) = \mathsf{P}(A) = 0{,}3$. Легко видеть, что вероятность $\mathsf{P}(C) = 1 - \mathsf{P}(\overline{C}) = 1 - 42/90 = 8/15$. Значит, $\mathsf{P}(A|C) = \mathsf{P}(A)/\mathsf{P}(C) = 9/16$. События A и B являются попарно зависимыми, так как имеет место соотношение: $\mathsf{P}(A \cap B) = 1/15 \neq \mathsf{P}(A) \times \mathsf{P}(B) = 0{,}09$. Отсюда делаем вывод, что события A, B и C тем более не являются независимыми в совокупности.

Заметим также, что для независимости в совокупности событий A_1, $A_2, \ldots, A_n$ недостаточно выполнения равенства $\mathsf{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathsf{P}(A_1) \times \mathsf{P}(A_2) \times \ldots \times \mathsf{P}(A_n)$. Для подтверждения этого рассмотрим эксперимент.

Пример 3.22. С помощью некоторого механизма на поверхность стола случайным образом подбрасывается симметричная игральная кость два раза и определяется число выпавших очков при каждом броске. Найти такие случайные события A, B и C, для которых независимость тройки не влечет их попарную независимость. В этом эксперименте комплекс условий $\Sigma = \{u_1$ — выбрана некоторая симметричная игральная кость, u_2 — игральная кость подбрасывается два раза, u_3 — случайный механизм подбрасывания игральной кости,

u_4 — поверхность стола, на которую падает игральная кость, u_5 — освещенность поверхности стола, $\ldots\}$.

Решение. Обозначим теперь через вектор (x, y) описание элементарного исхода этого опыта в том случае, когда при первом броске выпадает x очков, а при втором броске — y. Каждая из компонент x, y может принимать целое значение от единицы до шести. Итак, здесь имеется всего $n = 36$ равновозможных элементарных событий вида $\{(x, y)\}$. В табл. 3.1 приведены все описания элементарных исходов этого опыта.

Таблица 3.1

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Поэтому достоверное событие $\Omega = \{\omega = (x, y)\colon x, y = 1, 2, \ldots, 6\}$. Если $\mathfrak{J} = \mathcal{F}$ — множество всех подмножеств множества Ω, то $\mathcal{F} = \{A_1, A_2, \ldots, A_{2^n}\}$, где $n = 36$. Обозначим через $A \in \mathcal{F}$ событие, которое заключается в том, что сумма выпавших очков при втором броске не более трех. Обозначим также через $B \in \mathcal{F}$ событие, которое означает, что сумма выпавших очков при втором броске не менее трех и, одновременно, не более пяти. Пусть событие C состоит в том, что сумма выпавших очков при двух бросках равна девяти. Тогда

$$\begin{aligned}
A &= \{\omega = (x, y)\colon x = 1, 2, \ldots, 6,\ y = 1, 2, 3\} = \\
&= \{(1,1),\ (2,1),\ (3,1),\ (4,1),\ (5,1),\ (6,1),\ (1,2),\ (2,2),\ (3,2),\ (4,2),\\
&\qquad (5,2),\ (6,2),\ (1,3),\ (2,3),\ (3,3),\ (4,3),\ (5,3),\ (6,3)\},\\
B &= \{\omega = (x, y)\colon x = 1, 2, \ldots, 6,\ y = 3, 4, 5\} = \\
&= \{(1,3),\ (2,3),\ (3,3),\ (4,3),\ (5,3),\ (6,3),\ (1,4),\ (2,4),\ (3,4),\\
&\qquad (4,4),\ (5,4),\ (6,4),\ (1,5),\ (2,5),\ (3,5),\ (4,5),\ (5,5),\ (6,5)\},\\
C &= \{\omega = (x, y)\colon x, y = \overline{1,6},\ x + y = 9\} = \{(6,3),\ (5,4),\ (4,5),\ (3,6)\}.
\end{aligned}$$

Отсюда

$$\mathsf{P}(A) = \mathsf{P}(B) = 1/2, \quad \mathsf{P}(C) = 1/9,$$

$$\begin{aligned}
\mathsf{P}(A \cap B) = \mathsf{P}(\{(1,3),\ (2,3),\ (3,3),\ (4,3),\ (5,3),\ (6,3)\}) =\\
= 1/6 \neq \mathsf{P}(A)\,\mathsf{P}(B),
\end{aligned}$$

$$\mathsf{P}(A \cap C) = \mathsf{P}(\{(6,3)\}) = 1/36 \neq \mathsf{P}(A)\,\mathsf{P}(C),$$

$$\mathsf{P}(B \cap C) = \mathsf{P}(\{(6,3),\,(5,4),\,(4,5)\}) = 1/12 \neq \mathsf{P}(B)\,\mathsf{P}(C).$$

В то же время находим, что $\mathsf{P}(A \cap B \cap C) = \mathsf{P}(\{(6,3)\}) = 1/36 = (1/2) \times (1/2) \times (1/9) = \mathsf{P}(A)\,\mathsf{P}(B)\,\mathsf{P}(C)$.

Итак, имеет место независимость тройки событий A, B и C, но в то же время всевозможные пары этих событий являются зависимыми объектами. Другими словами, попарная независимость для этого эксперимента не следует из равенства $\mathsf{P}(A \cap B \cap C) = \mathsf{P}(A)\,\mathsf{P}(B)\,\mathsf{P}(C)$.

3.4. Типичные ошибки при использовании основных понятий и формул теории вероятностей. Понятие и определение независимости в теории вероятностей часто согласуются с нашим обычным представлением о независимости в реальности и помогают объяснить в связи с этим различные поведения субъектов в тех или иных жизненных ситуациях. Проиллюстрируем сказанное на результатах эксперимента по приему экзамена по теории вероятностей и математической статистики в Нижегородском государственном университете им. Н. И. Лобачевского (ННГУ), которые подробно приведены в работе: *Федоткин М.А.* Управляющие системы и решение парадокса независимости // Вестник ННГУ. 2011. №3(2). С. 152–161.

Пример 3.23. Студенты сдают экзамен по теории вероятностей, каждый из которых должен ответить только на три вопроса раздельно профессору либо ассистенту. Профессору студент отвечает положительно на любой вопрос с вероятностью p и ассистенту — с вероятностью q. Здесь рассмотрим случай, когда имеет место неравенство $p < q$. Это означает, что студенту легче ответить на вопрос ассистенту. Совершенно аналогично можно решить задачу при $p \geqslant q$, когда студенту не труднее ответить профессору. Студенту предлагается выбрать одну из двух стратегий последовательности ответа. Первая стратегия заключается в том, что сначала студент отвечает профессору, затем — ассистенту и, наконец, снова — профессору. При второй стратегии на первый вопрос студент отвечает ассистенту, на второй вопрос профессору и на последний вопрос — ассистенту. Студент сдает экзамен, если он два раза подряд положительно отвечает на вопросы. Как правило, студенты выбирают более легкую на их взгляд вторую стратегию, так как в этом случае приходится отвечать два раза ассистенту и один раз профессору. Рассмотрим решение этой удивительной и поучительной задачи.

Решение задачи о поведении на экзаменах хорошо подготовленных студентов. Обозначим через $(\Omega_1, \mathcal{F}_1, \mathsf{P}_1(\cdot))$ и через $(\Omega_2, \mathcal{F}_2, \mathsf{P}_2(\cdot))$ вероятностные модели данного эксперимента для первой и второй стратегии соответственно. Пусть события A_1, A_2 и A_3 из множества $\mathcal{F}_1$

означают, что на первый вопрос студент положительно ответил профессору, на второй вопрос — ассистенту и на третий вопрос — опять профессору. Обозначим через A событие из $\mathcal{F}_1$, которое означает, что студент сдаст экзамен при первой стратегии. Аналогичные события введем, если студент выбрал вторую стратегию. Пусть события B_1, B_2 и B_3 из множества $\mathcal{F}_2$ заключаются в том, что на первый вопрос студент положительно ответил ассистенту, на второй вопрос — профессору и на третий вопрос — снова ассистенту. Обозначим через B событие из $\mathcal{F}_2$, которое состоит в том, что студент сдаст экзамен при второй стратегии. Для событий $A \in \mathcal{F}_1$ и $B \in \mathcal{F}_2$ легко находим равенства

$$A = (A_1 \cap A_2) \cup (A_2 \cap A_3), \qquad B = (B_1 \cap B_2) \cup (B_2 \cap B_3). \tag{3.9}$$

Тогда при независимых ответах студента преподавателям имеем

$$\begin{aligned} \mathsf{P}_1(A) &= \mathsf{P}_1((A_1 \cap A_2) \cup (A_2 \cap A_3)) = \\ &= \mathsf{P}_1(A_1 \cap A_2) + \mathsf{P}_1(A_2 \cap A_3) - \mathsf{P}_1(A_1 \cap A_2 \cap A_3) = pq + qp - pqp, \end{aligned}$$

если он выбрал первую стратегию, и имеем

$$\begin{aligned} \mathsf{P}_2(B) &= \mathsf{P}_2((B_1 \cap B_2) \cup (B_2 \cap B_3)) = \\ &= \mathsf{P}_2(B_1 \cap B_2) + \mathsf{P}_2(B_2 \cap B_3) - \mathsf{P}_2(B_1 \cap B_2 \cap B_3) = qp + pq - qpq, \end{aligned}$$

если он выбрал вторую стратегию. Отсюда находим, что $\mathsf{P}_1(A) > \mathsf{P}_2(B)$ при $q > p$. Итак, для отлично подготовленных студентов, правильные ответы которых, скорее всего, являются независимыми случайными событиями, целесообразно пользоваться первой стратегией. Еще раз уточним, что полученное решение этой задачи и выводы существенно опирались на факт независимости в совокупности событий A_1, A_2 и A_3 относительно вероятности $\mathsf{P}_1(\cdot)$ и на факт независимости в совокупности событий B_1, B_2 и B_3 относительно вероятности $\mathsf{P}_2(\cdot)$.

Приведем теперь решение этой задачи в случае, когда нет указанной независимости. В этом случае естественно предположить, что слабо подготовленные студенты стараются улучшить свои знания по теории вероятностей после своего очередного ответа преподавателю. С другой стороны, очень часто на экзаменах преподаватели высокой квалификации используют последнюю возможность обучения плохо подготовленных студентов при их неудачных ответах.

Линейная модель обучения на экзаменах плохо подготовленных студентов. По сравнению с равенствами (3.9) запишем здесь несколько другие соотношения для исходов $A \in \mathcal{F}_1$ и $B \in \mathcal{F}_2$, а именно:

$$\begin{aligned} A &= (A_1 \cap A_2 \cap \overline{A}_3) \cup (\overline{A}_1 \cap A_2 \cap A_3) \cup (A_1 \cap A_2 \cap A_3), \\ B &= (B_1 \cap B_2 \cap \overline{B}_3) \cup (\overline{B}_1 \cap B_2 \cap B_3) \cup (B_1 \cap B_2 \cap B_3). \end{aligned} \tag{3.10}$$

Используя несовместимость событий в правых частях равенств (3.10) и теорему умножения при зависимых ответах студента преподавателям, получим, что

$$\begin{aligned}
&\mathsf{P}_1(A) = \mathsf{P}_1((A_1 \cap A_2 \cap \overline{A}_3) \cup (\overline{A}_1 \cap A_2 \cap A_3) \cup (A_1 \cap A_2 \cap A_3)) = \\
&= \mathsf{P}_1(A_1 \cap A_2 \cap \overline{A}_3) + \mathsf{P}_1(\overline{A}_1 \cap A_2 \cap A_3) + \mathsf{P}_1(A_1 \cap A_2 \cap A_3) = \\
&= \mathsf{P}_1(A_1)\,\mathsf{P}_1(A_2|A_1)\,\mathsf{P}_1(\overline{A}_3|A_1 \cap A_2) + \mathsf{P}_1(\overline{A}_1)\,\mathsf{P}_1(A_2|\overline{A}_1)\,\mathsf{P}_1(A_3|\overline{A}_1 \cap A_2) + \\
&\qquad\qquad + \mathsf{P}_1(A_1)\,\mathsf{P}_1(A_2|A_1)\,\mathsf{P}_1(A_3|A_1 \cap A_2) = \\
&= (\mathsf{P}_1(\overline{A}_3|A_1 \cap A_2) + \mathsf{P}_1(A_3|A_1 \cap A_2)) \times \mathsf{P}_1(A_1)\,\mathsf{P}_1(A_2|A_1) + \\
&\qquad\qquad + \mathsf{P}_1(\overline{A}_1)\,\mathsf{P}_1(A_2|\overline{A}_1)\,\mathsf{P}_1(A_3|\overline{A}_1 \cap A_2) = \\
&= \mathsf{P}_1(A_1)\,\mathsf{P}_1(A_2|A_1) + \mathsf{P}_1(\overline{A}_1)\,\mathsf{P}_1(A_2|\overline{A}_1)\,\mathsf{P}_1(A_3|\overline{A}_1 \cap A_2),
\end{aligned}$$

если он выбрал первую стратегию. Аналогичным способом имеем:

$$\mathsf{P}_2(B) = \mathsf{P}_2(B_1)\,\mathsf{P}_2(B_2|B_1) + \mathsf{P}_2(\overline{B}_1)\,\mathsf{P}_2(B_2|\overline{B}_1)\,\mathsf{P}_2(B_3|\overline{B}_1 \cap B_2),$$

если он выбрал вторую стратегию.

Пусть теперь условные вероятности $\mathsf{P}_1(A_2|\overline{A}_1)$ и $\mathsf{P}_2(B_3|\overline{B}_1 \cap B_2)$ равны $q + \varepsilon < 1$ $(\varepsilon < 1 - q)$. Это означает увеличение на $\varepsilon > 0$ вероятности следующего ответа плохо подготовленного студента ассистенту после неудачного его предыдущего ответа. Таким способом произошло обучение студента профессором. Используя равенство $q = \mathsf{P}_1(A_2) = \mathsf{P}_1(A_2 \cap A_1) + \mathsf{P}_1(A_2 \cap \overline{A}_1)$ и формулу (3.1), находим $\mathsf{P}_1(A_2 \mid A_1) = (pq + p\varepsilon - \varepsilon)p^{-1} = q - (1-p)\varepsilon p^{-1} < q$. Итак, произошло уменьшение вероятности следующего ответа плохо подготовленного студента ассистенту после удачного предыдущего ответа профессору. Это обстоятельство можно проинтерпретировать тем, что у студента появляется некоторое психологическое самомнение после удачного ответа профессору. Будем еще считать, что условные вероятности $\mathsf{P}_1(A_3|\overline{A}_1 \cap A_2)$ и $\mathsf{P}_2(B_2|B_1)$ равны безусловным, т. е. профессор оценивает студента более строго и, в то же время, более объективно по сравнению с ассистентом. Другими словами, оценка студента профессором не зависит от удачных или неудачных его предыдущих ответов. Таким способом здесь заложена простейшая линейная модель обучения плохо подготовленных студентов на экзаменах.

Так как условная вероятность $\mathsf{P}_1(A_2|A_1) = (pq + p\varepsilon - \varepsilon)p^{-1} > 0$, то отсюда имеем еще одно ограничение $\varepsilon < pq(1-p)^{-1}$ на параметр обучения ε. Подставляя теперь значения указанных условных и безусловных вероятностей в выражения для $\mathsf{P}_1(A)$ и $\mathsf{P}_2(B)$, непосредственно получим для $\mathsf{P}_1(A) = (2-p)qp - (1-p)^2\varepsilon$ и $\mathsf{P}_2(B) = (2-q)qp + p(1-q)\varepsilon$. Учитывая это, находим $\mathsf{P}_2(B) > \mathsf{P}_1(A)$, если только значение параметра обучения выбирать из следующего условия:

$$(q-p)pq(p(1-q) + (1-p)^2)^{-1} < \varepsilon < \min\{1-q, pq(1-p)^{-1}\}.$$

Поскольку при $0 < p < q < 1$ имеем очевидное неравенство

$$(q-p)pq(p(1-q)+(1-p)^2)^{-1} < pq(1-p)^{-1},$$

то в этом случае на параметры p и q получаем ограничение

$$(q-p)pq(p(1-q)+(1-p)^2)^{-1} < 1-q,$$

откуда $q < (p-0{,}5)^2 + 0{,}75$.

На рис. 3.2, с учетом неравенства $0<p<q<1$, приведена заштрихованная область $G = \{(p,q)\colon 0 < p < q < 1,\, q < (p-0{,}5)^2 + 0{,}75\}$ значений параметров p и q, составленная из двух областей: $G_1 = \{(p,q)\colon 0 < p < q < 1,\, p+q \geqslant 1,\, q < (p-0{,}5)^2 + 0{,}75\}$ и $G_2 = \{(p,q)\colon 0 < p < q < 1,\, p+q < 1\}$.

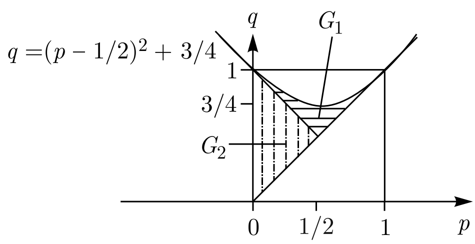


Рис. 3.2

Область G_1 заштрихована сплошными горизонтальными линиями, а область G_2 штрихпунктирными вертикальными линиями. Область G соответствует тому случаю, когда слабо подготовленным студентам при их ответах преподавателям целесообразно пользоваться первой или второй стратегией в зависимости от значения параметра обучения ε. Как следствие из предыдущих рассуждений, в следующем абзаце впервые приведем доказательство последнего утверждения и общее решение известного парадокса о независимости. Этому классическому парадоксу, например, посвящены следующие замечательные книги:

— *Мостеллер Ф.* Пятьдесят занимательных вероятностных задач с решениями.— М.: Наука, 1975.

— *Секей Г.* Парадоксы в теории вероятностей и математической статистике.— М.: Мир, 1990.

Пусть параметры p, q и ε, которые соответственно определяют степень подготовки студентов и возможность их обучения на экзаменах, удовлетворяют условиям $0 < \varepsilon \leqslant (q-p)pq(p(1-q)+(1-p)^2)^{-1}$, $(p,q) \in G$. Отсюда $\mathsf{P}_1(A) \geqslant \mathsf{P}_2(B)$ и, значит, студенты должны использовать первую стратегию. В противном слечае, при $(p,q) \in G_1$ и $(q-p)pq(p(1-q)+(1-p)^2)^{-1} < \varepsilon < 1-q$ получаем $\mathsf{P}_2(B) > \mathsf{P}_1(A)$, и студенты должны применять вторую стратегию.

Если $(q-p)pq(p(1-q)+(1-p)^2)^{-1} < \varepsilon < pq(1-p)^{-1}$ и $(p,q) \in G_2$, то находим $\mathsf{P}_2(B) > \mathsf{P}_1(A)$, и студенты опять должны использовать вторую стратегию.

Наконец, при $(p,q) \in \{(p,q)\colon 0 < p < q < 1,\ q \geqslant (p-0{,}5)^2 + 0{,}75\}$ имеем: $1-q \leqslant (q-p)pq(p(1-q)+(1-p)^2)^{-1} < pq(1-p)^{-1}$. В этом случае для всех возможных значений $0 < \varepsilon < 1-q$ следует придерживаться только первой стратегии, так как $\mathsf{P}_1(A) > \mathsf{P}_2(B)$.

Итак, совет использовать только первую стратегию, который дается в указанных книгах, не всегда является верным. В этих книгах ключевым фактом в рассуждениях является важность правильного ответа на второй вопрос. Действительно, только в этом случае студент будет иметь возможность подряд два правильных ответа и может сдать экзамен. Другими словами, большая вероятность ответить на второй вопрос ассистенту якобы играет более существенную роль, чем меньшее число ответов профессору. Однако многочисленные опросы студентов и сами экзамены показывают, что применение первой или второй стратегии существенно зависит от подготовки и обучаемости на экзаменах студентов.

Пример 3.24 (*парадокс Мизеса*). Обратим внимание на важность рассмотрения пространств $(\Omega_1, \mathcal{F}_1, \mathsf{P}_1(\cdot))$ и $(\Omega_2, \mathcal{F}_2, \mathsf{P}_2(\cdot))$ при решении задачи об экзаменах. Более того, часто теоретико-множественные модели $(\Omega_1, \mathcal{F}_1)$ и $(\Omega_2, \mathcal{F}_2)$ экспериментов совпадают, т. е. $\Omega_1 = \Omega_2$ и $\mathcal{F}_1 = \mathcal{F}_2$, а вероятностные функции разные, т. е. $\mathsf{P}_1(\cdot) \neq \mathsf{P}_2(\cdot)$. Если это не учитывать, то могут возникать всевозможные ошибки. Иллюстрацией этого может служить решение такого парадокса. Студенты сдают экзамен по теории вероятностей, каждый из которых должен отвечать по билету только одному из преподавателей: либо доценту, либо профессору. Студентов непреднамеренно вызывает отвечать по билету преподаватель. Профессору студент отвечает положительно по билету с вероятностью p, а доценту — с вероятностью q. Требуется найти вероятность того, что студент сдаст экзамен?

Ошибочное решение. По условию эксперимента событие, которое заключается в том, что студент сдаст экзамен профессору, и событие, которое состоит в том, что студент сдаст экзамен доценту, являются несовместимыми. Отсюда по теореме сложения искомая вероятность равна $p+q$. Абсурдность такого решения становится очевидной, если взять численные значения для $p = 0{,}75$ и $q = 0{,}85$. Тогда имеем $p+q = 1{,}6 > 1$. В этом и заключается парадокс Мизеса. Здесь нельзя применять теорему сложения, так как вероятности p и q относятся к разным вероятностным пространствам. Ниже приведем решение этого парадокса.

Решение с использованием соответствующих вероятностных пространств. Для этого эксперимента E разумно ввести четыре элементарных исхода:

1) студент будет отвечать профессору и сдаст экзамен, это элементарное событие будем обозначать через $\{\omega_1\}$;
2) студент будет отвечать профессору и не сдаст экзамен, это элементарное событие будем обозначать через $\{\omega_2\}$;
3) студент будет отвечать доценту и сдаст экзамен, это элементарное событие будем обозначать через $\{\omega_3\}$;
4) студент будет отвечать доценту и не сдаст экзамен, это элементарное событие будем обозначать через $\{\omega_4\}$.

В этих обозначениях символы ω_1, ω_2, ω_3 и ω_4 отвечают за описание в некотором языке введенных элементарных событий. Теперь легко найти элементы $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ и $\mathcal{F} = \{\Omega, \varnothing, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}, \{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_4\}, \{\omega_1, \omega_3, \omega_4\}, \{\omega_2, \omega_3, \omega_4\}\}$ теоретико-множественной модели данного опыта E. Пусть событие A из $\mathcal{F}$ означает, что студент сдаст экзамен какому-либо преподавателю, а событие B из $\mathcal{F}$ состоит в том, что студент будет отвечать по билету профессору. Отсюда получим, что $A = \{\omega_1, \omega_3\}$ и $B = \{\omega_1, \omega_2\}$, при этом событие $\overline{A} = \{\omega_2, \omega_4\}$ заключается в том, что студент получит отрицательную оценку на экзамене, а событие $\overline{B} = \{\omega_3, \omega_4\}$ означает, что студент будет сдавать экзамен доценту. Тогда имеем $A \cap B = \{\omega_1\}$, $\overline{A} \cap B = \{\omega_2\}$, $A \cap \overline{B} = \{\omega_3\}$, $\overline{A} \cap \overline{B} = \{\omega_4\}$. Используя теперь теорему умножения и значения p и q для заданных условных вероятностей $\mathsf{P}(A|B)$ и $\mathsf{P}(A|\overline{B})$ найдем:

$$\begin{aligned} &\mathsf{P}(\{\omega_1\}) = \mathsf{P}(A \cap B) = \mathsf{P}(B)\,\mathsf{P}(A|B) = \mathsf{P}(B)\,p, \\ &\mathsf{P}(\{\omega_2\}) = \mathsf{P}(\overline{A} \cap B) = \mathsf{P}(B)\,\mathsf{P}(\overline{A}|B) = \mathsf{P}(B)\,(1-p), \\ &\mathsf{P}(\{\omega_3\}) = \mathsf{P}(A \cap \overline{B}) = \mathsf{P}(\overline{B})\,\mathsf{P}(A|\overline{B}) = (1-\mathsf{P}(B))\,q, \\ &\mathsf{P}(\{\omega_4\}) = \mathsf{P}(\overline{A} \cap \overline{B}) = \mathsf{P}(\overline{B})\,\mathsf{P}(\overline{A}|\overline{B}) = (1-\mathsf{P}(B))\,(1-q). \end{aligned} \tag{3.11}$$

Здесь следует обратить особое внимание на следующую тонкость.

Правомерность использования в (3.11) теоремы умножения логически нельзя доказать, так как вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ еще полностью не построена. Однако если бы вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ была заранее известна, то все равенства соотношения (3.11) должны были бы выполняться. Такова мотивировка целесообразности использования теоремы умножения. Вероятности элементарных событий в соотношении (3.11) полностью определяются величинами $\mathsf{P}(B)$, p, q и, тем самым, позволяют находить вероятности любого события из $\mathcal{F}$. Например, искомая вероятность $\mathsf{P}(A) = \mathsf{P}(\{\omega_1, \omega_3\}) = \mathsf{P}(\{\omega_1\}) + \mathsf{P}(\{\omega_3\}) = \mathsf{P}(B)\,p + (1-\mathsf{P}(B))\,q$. Если теперь для определенности положить, что $\mathsf{P}(B) = 0{,}5$, $p = 0{,}75$

и $q = 0{,}85$, то $\mathsf{P}(A) = 0{,}8$. Таким образом, вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ данного опыта задается его условиями и гипотетическими числами $\mathsf{P}(B)$, p, q с применением теоремы умножения. Здесь еще раз обнаруживается полезность теоремы умножения, если мы уже знаем условные вероятности. Условные вероятности, как правило, задаются содержанием задачи или приближенно заменяются условными частотами. Для окончательного решения парадокса Мизеса переходим теперь к рассмотрению других вероятностных пространств опыта E.

Пусть студента вызвал отвечать по билету профессор, т. е. наступило событие B. В этом случае мы должны рассматривать условный эксперимент E_{y1} с вероятностной моделью $(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$. Значит, для такого условного эксперимента вероятность того, что студент сдаст экзамен профессору, равна $\mathsf{P}(A \cap B|B) = \mathsf{P}(A|B) = p$. Эту вероятность также можно получить из равенства $\mathsf{P}(A \cap B)/\mathsf{P}(B) = \mathsf{P}(B)\,p/\mathsf{P}(B) = p$, если заранее принять вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ априорного эксперимента E с вероятностями элементарных исходов вида (3.11). Тогда из определения условной вероятности получим, что $\mathsf{P}(A|B) = p$.

Совершенно так же рассуждаем, если студента вызвал отвечать по билету доцент, т. е. наступило событие $\overline{B}$. В этом случае мы естественно рассматриваем условный эксперимент E_{y2} с вероятностной моделью $(\Omega, \mathcal{F}, \mathsf{P}(\cdot|\overline{B}))$. Условная вероятность $\mathsf{P}(A|\overline{B}) = \mathsf{P}(A \cap \overline{B}|\overline{B}) = \mathsf{P}(A \cap \overline{B})/\mathsf{P}(\overline{B}) = q(1 - \mathsf{P}(B))/\mathsf{P}(\overline{B}) = q$. Окончательно получаем решение задачи Мизеса. Вероятности $p = \mathsf{P}(A|B)$ и $q = \mathsf{P}(A|\overline{B})$, естественно, относятся к разным вероятностным моделям $(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$ и $(\Omega, \mathcal{F}, \mathsf{P}(\cdot|\overline{B}))$. Значит, не имеет смысла применять теорему сложения при первом ошибочном решении.

В заключение этого раздела отметим, что теорема умножения 3.2 и теоремы 3.3 и 3.4 имеют простые доказательства. Однако эти теоремы очень часто используются в теории вероятностей и математической статистике. Здесь было бы полезно привести пример эксперимента, когда применение теоремы умножения, формул полной вероятности, Байеса и определения 3.3 происходит формально, т. е. без проверки соответствующих условий указанных утверждений. Это неизбежно может привести к серьезным недоразумениям и даже к ошибкам. С этой целью можно предложить довольно типичный пример 7.8 из «Сборника задач по теории вероятностей, математической статистике и теории случайных функций» под редакцией А. А. Свешникова (М.: Наука, 1970).

Пример 3.25 (*пример Свешникова*). Вероятности попадания при каждом выстреле для трех стрелков равны соответственно 4/5, 3/4, 2/3. Предполагается, что стрельба ведется независимым образом. Определить вероятность того, что промахнулся третий стрелок, если при одновременном выстреле всех трех стрелков имелось два попадания. Пусть событие A означает два попадания. В указанном выше сборнике задач в качестве гипотезы H_i рекомендуется взять событие, которое заключается в том, что i-й стрелок промахнулся, где $i = 1, 2, 3$. Тогда имеем:

$$\mathsf{P}(H_1) = 1/5, \qquad \mathsf{P}(H_2) = 1/4, \qquad \mathsf{P}(H_3) = 1/3,$$
$$A = (\overline{H}_1 \cap \overline{H}_2 \cap H_3) \cup (\overline{H}_1 \cap H_2 \cap \overline{H}_3) \cup (H_1 \cap \overline{H}_2 \cap \overline{H}_3),$$
$$\mathsf{P}(A|H_1) = \mathsf{P}(A \cap H_1)/\mathsf{P}(H_1) = \mathsf{P}(\overline{H}_2)\,\mathsf{P}(\overline{H}_3) = 1/2,$$
$$\mathsf{P}(A|H_2) = \mathsf{P}(A \cap H_2)/\mathsf{P}(H_2) = \mathsf{P}(\overline{H}_1)\,\mathsf{P}(\overline{H}_3) = 8/15,$$
$$\mathsf{P}(A|H_3) = \mathsf{P}(A \cap H_3)/\mathsf{P}(H_3) = \mathsf{P}(\overline{H}_1)\,\mathsf{P}(\overline{H}_2) = 3/5,$$
$$\mathsf{P}(A) = \sum_{i=1}^{3} \mathsf{P}(H_i)\,\mathsf{P}(A|H_i) = 13/30,$$
$$\mathsf{P}(H_3|A) = \mathsf{P}(H_3)\,\mathsf{P}(A|H_3)/\mathsf{P}(A) = (1/3) \times (3/5) : (13/30) = 6/13.$$

Решая пример Свешникова таким методом, студенты получают верный ответ. Более того, для этого примера имеем $\mathsf{P}(H_1) + \mathsf{P}(H_2) + \mathsf{P}(H_3) = 47/60 < 1$, а события H_i, H_j при $i \neq j$ несовместимы при наблюдении события A. Эти соображения придают уверенность студентам в правильности выбора метода решения этого примера.

Изменим числовые данные рассматриваемого примера. Пусть теперь вероятности попадания при каждом выстреле для трех стрелков равны соответственно 7/10, 3/5, 2/5. Тогда для этих числовых данных вероятности $\mathsf{P}(H_1)$, $\mathsf{P}(H_2)$, $\mathsf{P}(H_3)$ получаются равными 3/10, 2/5, 3/5 и $\mathsf{P}(H_1) + \mathsf{P}(H_2) + \mathsf{P}(H_3) = 13/10 > 1$. Последнее неравенство объясняется тем, что события H_1, H_2, H_3 являются попарно совместимыми и, следовательно, эти события не могут быть выбраны в качестве гипотез.

Приведем правильное решение примера с первоначальными данными. Для этого выберем в качестве гипотез события H_3, $\overline{H}_3$. Применяем формулу (3.1) для вычисления условной вероятности вида $\mathsf{P}(A|\overline{H}_3)$. Используя представление исхода A через события H_1, H_2, H_3, несовместность событий $(\overline{H}_1 \cap H_2 \cap \overline{H}_3)$, $(H_1 \cap \overline{H}_2 \cap \overline{H}_3)$ и независимость случайных событий H_1, H_2, H_3, получим, что условная вероятность

$$\mathsf{P}(A|\overline{H}_3) = \mathsf{P}(A \cap \overline{H}_3)/\mathsf{P}(\overline{H}_3) =$$
$$= \mathsf{P}((\overline{H}_1 \cap H_2 \cap \overline{H}_3) \cup (H_1 \cap \overline{H}_2 \cap \overline{H}_3))/\mathsf{P}(\overline{H}_3) =$$
$$= \mathsf{P}(\overline{H}_1)\,\mathsf{P}(H_2) + \mathsf{P}(H_1)\,\mathsf{P}(\overline{H}_2) = (4/5) \times (1/4) + (1/5) \times (3/4) = 7/20.$$

Следовательно,

$$\mathsf{P}(A) = \mathsf{P}(H_3)\,\mathsf{P}(A|H_3) + \mathsf{P}(\overline{H}_3)\,\mathsf{P}(A|\overline{H}_3) =$$
$$= (1/3) \times (3/5) + (2/3) \times (7/20) = 13/30,$$

$$\mathsf{P}(H_3|A) = \mathsf{P}(H_3)\,\mathsf{P}(A|H_3)/\mathsf{P}(A) = 6/13.$$

Теперь поясним причину получения правильного ответа при решении данного примера, если следовать указаниям, приведенным в указанном сборнике. С этой целью рассмотрим события $A \cap H_1$, $A \cap H_2$, $A \cap H_3$, $\Omega \backslash A$. Вспоминая представление исхода A через события H_1, H_2, H_3, легко показать, что $A \cap H_1$, $A \cap H_2$, $A \cap H_3$, $\Omega \backslash A$ образуют полную группу несовместимых событий, т. е. являются гипотезами. Тогда для случайного события A имеем

$$\mathsf{P}(A) = \sum_{i=1}^{3} \mathsf{P}(A \cap H_i)\,\mathsf{P}(A|A \cap H_i) + \mathsf{P}(\Omega\backslash A)\,\mathsf{P}(A|\Omega\backslash A) =$$

$$= \sum_{i=1}^{3} \mathsf{P}(A \cap H_i) = \sum_{i=1}^{3} \mathsf{P}(H_i \cap A) = \sum_{i=1}^{3} \mathsf{P}(H_i)\,\mathsf{P}(A|H_i) = 13/30,$$

$$\mathsf{P}(H_3|A) = \mathsf{P}(H_3)\,\mathsf{P}(A|H_3)/\mathsf{P}(A) = (1/3) \times (3/5) : (13/30) = 6/13.$$

В силу полученного для вероятности $\mathsf{P}(A)$ равенства $\mathsf{P}(A) = \sum_{i=1}^{3} \mathsf{P}(H_i)\,\mathsf{P}(A|H_i)$ и создается ошибочное мнение, что в качестве гипотез можно взять события H_1, H_2, H_3.

Краткий обзор

В главе 3 предложено описание условных экспериментов и измерение зависимости между случайными событиями. Рассмотрена методика построения унифицированной и локализованной вероятностных моделей условных экспериментов. С экспериментальной точки зрения дано обоснование целесообразности определения условной вероятности с помощью соотношения (3.1). Доказаны формулы для вычисления априорных и условных вероятностей. Решены примеры, в которых указаны типичные ошибки при вычислении условных вероятностей. На содержательном уровне выяснен физический смысл причинной независимости исходов реального эксперимента. На этом основании в теории вероятностей формализовано фундаментальное понятие статистической независимости случайных событий. Установлен так называемый закон «нуля или единицы» для последовательности $\{A_i;\ i = 1, 2, \ldots\}$ независимых событий, который играет важную роль в теории вероятностей и математической статистике. Предложена и изучена линейная модель обучения на экзаменах плохо подготовленных студентов. Приводится общее решение известного парадокса о независимости.

Контрольные вопросы и упражнения

1. Подбрасываются два игральных кубика. Известно, что сумма выпавших очков равна восьми. Найти вероятность того, что выпадет хотя бы один раз шесть очков.

2. Вероятность того, что два близнеца — однополые, равна 0,64. Вероятность того, что новорожденный младенец (первенец) мальчик равна 0,51. Вероятность рождения девочки первой и мальчика вторым равна 0,19. Найти вероятность того, что второй из близнецов мальчик, если известно, что первый из них является мальчиком.
3. При переписи населения Уэльса в 1891 г. оказалось, что темноглазые отцы и темноглазые сыновья составляют 5 % из обследованных пар, темноглазые отцы и светлоглазые сыновья — 7,9 %, светлоглазые отцы и светлоглазые сыновья — 78,2 %. Найти условную вероятность рождения темноглазого сына у светлоглазого отца.
4. В первой урне содержится десять шаров, из них восемь белых; во второй урне двенадцать шаров, из них четыре белых. Из каждой урны наудачу извлекли по одному шару, а затем из этих двух шаров наудачу взят один. Найти вероятность того, что взят белый шар.
5. Известно, что 10 % всей продукции является браком, а 45 % всей продукции удовлетворяет требованиям первого сорта. Определить вероятность того, что выбранное наудачу небракованное изделие является первосортным.
6. Батарея из трех орудий произвела залп. При этом два снаряда попали в цель. Вычислить вероятность того, что первое орудие дало попадание, если вероятности попадания в цель первым, вторым и третьим орудиями соответственно равны 2/5, 3/10 и 1/2.
7. В урну, содержащую n шаров, опущен белый шар. Затем наудачу извлекают один шар. Найти вероятность того, что извлеченный шар окажется белым, если все допущения о первоначальном числе белых шаров в урне равновозможны.
8. В урне находятся n шаров, из которых m белых, а остальные черные. Из урны последовательно и без возвращения извлекаются два шара, и при каждом извлечении определяется цвет шара. Какова вероятность появления белого шара при втором извлечении?
9. Найти вероятность того, что написанная наудачу простая дробь будет несократима (задача Чебышева).
10. По крейсеру выпустили независимо три торпеды, и он затонул. Вероятность попадания каждой из них равна 0,2. При попадании одной торпеды вероятность гибели крейсера равна 1/10, при двух — 3/10, а при трех — 7/10. Определить наиболее вероятное число попаданий торпед в крейсер.
11. В урне находятся черные и белые шарики. Если из урны последовательно и без возвращения наудачу извлекается два шарика, то

вероятность того, что они оба будут белыми, оказывается равна 1/2. Каково минимально возможное число шариков в такой урне?

12. События A_1, $A_2, \ldots$, A_n независимы в совокупности и вероятность $\mathsf{P}(A_k) = p_k$. Найти вероятности следующих событий: 1) не произойдет ни одного из событий A_1, $A_2, \ldots$, A_n; 2) произойдет хотя бы одно из событий A_1, $A_2, \ldots$, A_n; 3) произойдет одно и только одно из событий A_1, $A_2, \ldots$, A_n.
13. Две различимые игральные кости брошены один раз. Пусть событие A состоит в том, что на второй кости выпадает цифра или 1, или 2, или 5. Далее, событие B означает, что на второй кости появится цифра или 4, или 5, или 6. Наконец, событие C означает выпадение суммы очков на игральных костях, равной 9. Доказать, что в этом эксперименте независимость тройки событий A, B и C не влечет их попарную независимость.
14. Пусть достоверное событие $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ и $\mathsf{P}(\omega_1) = \mathsf{P}(\omega_2) = 3/8$, $\mathsf{P}(\omega_3) = \mathsf{P}(\omega_4) = 1/8$, $A = \{\omega_1, \omega_3\}$, $B = \{\omega_2, \omega_3\}$, $C = \{\omega_2, \omega_3\}$. Показать, что события A, B, C не являются независимыми в совокупности.
15. Родители назначили своим сыновьям премию в случае успешной сдачи экзаменов в школе. Младший брат сдает экзамен с вероятностью 7/10, а старший — с вероятностью 0,25. Каким образом родители должны поделить премию между сыновьями, если экзамен сдал только один из сыновей, и родителям не сообщается, кто из них сдал экзамен? При этом раздел премии между сыновьями пропорционален их условным вероятностям сдачи экзамена.
16. В городах Γ_1 и Γ_2 непреднамеренно подбрасывают игральную кость и соответственно симметричную монету. Пусть событие A означает выпадение герба, а событие B состоит в том, что число выпавших очков есть простое число. Докажите независимость событий A и B.
17. Преподаватель по теории вероятностей составил для r студентов n экзаменационных билетов $(n > r)$, среди которых m легких. Билеты вытаскиваются студентами один за другим по очереди в порядке их прибытия на экзамен случайным образом и без возвращения. Пусть событие A означает появление легкого билета для студента, который пришел на экзамен первым, а событие B появление у второго студента трудного билета. Являются ли эти события независимыми?

Глава 4

КОНЕЧНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ПРИЧИННО-НЕЗАВИСИМЫХ ИСПЫТАНИЙ

§ 1. Общая схема независимых испытаний

1.1. Метод составления вероятностной модели и его обоснование. Рассмотрим эволюционный эксперимент $E = \{E_t; t \in T\}$, для которого $T = \{1, 2, \ldots, n\}$. Будем предполагать, что каждый из опытов E_1, $E_2, \ldots$, E_n имеет конечное число элементарных исходов. Тогда при каждом $t = 1, 2, \ldots, n$ эксперимент E_t с комплексом условий Σ_t имеет конечное вероятностное пространство вида $(\Omega_t, \mathcal{F}_t, \mathsf{P}_t(\cdot))$. При этом символ $\omega_t \in \Omega_t$ дает в некотором языке полное описание произвольного элементарного результата $\{\omega_t\}$ эксперимента E_t и $\{\omega_t\} \in \mathcal{F}_t$. Следовательно, σ-алгебра $\mathcal{F}_t$ есть множество всех подмножеств множества Ω_t. Более того, будем считать, что эксперимент E составлен из n причинно-независимых испытаний E_1, $E_2, \ldots$, E_n. Приведем примеры конечного числа причинно-независимых испытаний, каждое из которых имеет конечное число исходов.

Пример 4.1. Имеются следующие предметы: симметричная монета определенного достоинства, симметричная игральная кость и тетраэдр. Указанные предметы находятся в трех различных комнатах. Пусть в этих комнатах совершенно непреднамеренно и независимо подбрасывают сначала монету, потом игральную кость и, наконец, тетраэдр. Здесь опыт E_1 состоит в подбрасывании монеты, опыт E_2 игральной кости и опыт E_3 тетраэдра. При этом интересуются сторонами, на которые падают указанные предметы.

Пример 4.2. В урне находятся 10 белых и 15 черных шаров, которые отличаются только цветом. Из урны непреднамеренно извлекают шар, фиксируют его цвет, после чего возвращают шар обратно в урну. Описанный опыт повторяют пять раз, и на этом эксперимент E считается завершенным. В этом примере E_t означает извлечение шара из урны, в которой лежат 10 белых и 15 черных шаров, в попытке с номером $t = 1, 2, 3, 4, 5$.

Пример 4.3. Самолет последовательно пролетает зоны обнаружения трех станций с номерами 1, 2 и 3. Каждая станция независимо от остальных регистрирует пролет самолета в своей зоне обнаружения. При этом первая станция обнаруживает самолет с вероятностью 9/10, вторая — с вероятностью 4/5 и третья — с вероятностью 3/5. Фиксируются номера станций, которым удалось обнаружить самолет.

В первой главе предложен простой способ построения для статистически устойчивого эволюционного эксперимента $E = \{E_1, E_2, \ldots, E_n\}$ теоретико-множественной модели $(\Omega, \mathcal{F})$, где $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$ и $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2 \times \ldots \times \mathcal{F}_n$. В этих обозначениях имеет место следующая теорема.

Теорема 4.1. *Множество всех подмножеств множества Ω совпадает с σ-алгеброй $\mathcal{F}$.*

Доказательство. Так как $\{\omega_t\} \in \mathcal{F}_t$ при любом $t = 1, 2, \ldots, n$, то каждый прямоугольник вида $\{\omega_1\} \times \{\omega_2\} \times \ldots \times \{\omega_n\}$ является элементарным событием из $\mathcal{F}$. Поэтому всевозможные конечные объединения прямоугольников такого вида и невозможное событие $\varnothing$ также принадлежат σ-алгебре $\mathcal{F}$. Известно, что совокупность всех (включая и пустое множество $\varnothing$) подмножеств конечного множества $\Omega = \{\omega = (\omega_1, \omega_2, \ldots, \omega_n)\colon \omega_1 \in \Omega_1, \omega_2 \in \Omega_2, \ldots, \omega_n \in \Omega_n\}$ можно получить путем включения в нее всевозможных конечных объединений прямоугольников вида $\{\omega_1\} \times \{\omega_2\} \times \ldots \times \{\omega_n\}$ и невозможного события $\varnothing$. Эта совокупность является σ-алгеброй и содержит самое большое число подмножеств из Ω. Отсюда и следует утверждение этой теоремы, ибо σ-алгебра $\mathcal{F}$ включает невозможное событие $\varnothing$ и всевозможные конечные объединения прямоугольников вида $\{\omega_1\} \times \{\omega_2\} \times \ldots \times \{\omega_n\}$.

Если теперь принять во внимание соотношение (3.7), которое имеет место для причинно-независимых испытаний, то вероятности элементарных событий $\{\omega\} = \{(\omega_1, \omega_2, \ldots, \omega_n)\} = \{\omega_1\} \times \{\omega_2\} \times \ldots \times \{\omega_n\}$ совместного эксперимента E целесообразно определить по формуле

$$\mathsf{P}(\{\omega\}) = \mathsf{P}_1(\{\omega_1\}) \times \mathsf{P}_2(\{\omega_2\}) \times \ldots \times \mathsf{P}_n(\{\omega_n\}). \tag{4.1}$$

Отсюда вероятности событий $\boldsymbol{A} \in \mathcal{F}$ можно находить по формуле

$$\mathsf{P}(\boldsymbol{A}) = \sum_{\omega \in \boldsymbol{A}} \mathsf{P}(\{\omega\}). \tag{4.2}$$

Напомним, что здесь и далее случайное событие из $\mathcal{F}$ обозначается полужирным латинским символом и курсивом.

Покажем, что функция $\mathsf{P}(\boldsymbol{A})$, $\boldsymbol{A} \in \mathcal{F}$, определенная формулами (4.1) и (4.2), удовлетворяет аксиомам Колмогорова.

Действительно, $\mathsf{P}(\boldsymbol{A}) \geqslant 0$ для любого $\boldsymbol{A} \in \mathcal{F}$, так как $\mathsf{P}(\{\omega\}) = \mathsf{P}_1(\{\omega_1\}) \times \mathsf{P}_2(\{\omega_2\}) \times \ldots \times \mathsf{P}_n(\{\omega_n\}) \geqslant 0$ при всех $\omega \in \Omega$. Далее непосредственно находим

$$\mathsf{P}(\Omega) = \sum_{\omega\in\Omega} \mathsf{P}(\{\omega\}) = \\ = \sum_{\omega_1\in\Omega_1} \mathsf{P}_1(\{\omega_1\}) \times \sum_{\omega_2\in\Omega_2} \mathsf{P}_2(\{\omega_2\}) \times \ldots \times \sum_{\omega_n\in\Omega_n} \mathsf{P}_n(\{\omega_n\}) = 1.$$

Наконец, если $\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_k$ — любая конечная последовательность попарно несовместимых событий из $\mathcal{F}$, то

$$\mathsf{P}\left(\bigcup_{i=1}^{k} \boldsymbol{B}_i\right) = \sum_{\omega\in\cup_{i=1}^{k}\boldsymbol{B}_i} \mathsf{P}(\{\omega\}) = \sum_{i=1}^{k} \sum_{\omega\in\boldsymbol{B}_i} \mathsf{P}(\{\omega\}) = \sum_{i=1}^{k} \mathsf{P}(\boldsymbol{B}_i).$$

Построенная таким методом вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ совместного эксперимента E называется прямым произведением вероятностных моделей $(\Omega_1, \mathcal{F}_1, \mathsf{P}_1(\cdot))$, $(\Omega_2, \mathcal{F}_2, \mathsf{P}_2(\cdot)), \ldots, (\Omega_n, \mathcal{F}_n, \mathsf{P}_n(\cdot))$ причинно-независимых экспериментов $E_1, E_2, \ldots, E_n$. Построения этого раздела существенно основывались на определении вероятности случайных элементарных исходов составного эксперимента $E = \{E_t;\, t = 1, 2, \ldots, n\}$ по формуле (4.1). Очень часто при любом фиксированном $t = 1, 2, \ldots, n$ эксперимент E_t содержит конечное число $\aleph(\Omega_t)$ равновозможных (равновероятных, симметричных) элементарных исходов. Такого рода случайные эксперименты называются классическими. Если к тому же эксперимент E составлен из n причинно-независимых испытаний $E_1, E_2, \ldots, E_n$, то эксперимент E будет классическим с числом элементарных исходов, равным $\aleph(\Omega) = \aleph(\Omega_1) \times \aleph(\Omega_2) \times \ldots \times \aleph(\Omega_n)$. В силу определения классической вероятности для элементарного события $\{(\omega_1, \omega_2, \ldots, \omega_n)\}$ эксперимента E и элементарных событий $\{\omega_1\}, \{\omega_2\}, \ldots, \{\omega_n\}$ опытов $E_1, E_2, \ldots, E_n$ получим, что $\mathsf{P}(\{(\omega_1, \omega_2, \ldots, \omega_n)\}) = \big(\aleph(\Omega_1) \times \aleph(\Omega_2) \times \ldots \times \aleph(\Omega_n)\big)^{-1} = (\aleph(\Omega_1))^{-1} \times (\aleph(\Omega_2))^{-1} \times \ldots \times (\aleph(\Omega_n))^{-1} = \mathsf{P}_1(\{\omega_1\}) \times \mathsf{P}_2(\{\omega_2\}) \times \ldots \times \mathsf{P}_n(\{\omega_n\})$, и, следовательно, формула (4.1) установлена. Этим фактом можно еще раз подкрепить целесообразность введения формулы (4.1).

1.2. Свойства общей схемы независимых испытаний. При каждом фиксированном $t = 1, 2, \ldots, n$ любому случайному событию $A_t \in \mathcal{F}_t$ поставим во взаимно однозначное соответствие случайное событие — прямоугольник вида $\boldsymbol{A}_t = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{t-1} \times A_{t+1} \times \ldots \times \Omega_{t+1} \times \Omega_n \in \mathcal{F}$. Покажем теперь следующее утверждение.

Теорема 4.2. *Для любых событий* $A_t \in \mathcal{F}_t$ и $\boldsymbol{A}_t \in \mathcal{F}_t$ *имеет место равенство* $\mathsf{P}_t(A_t) = \mathsf{P}(\{\boldsymbol{A}_t\})$ *и каждая последовательность событий* $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_n$ *независима в совокупности относительно вероятности* $\mathsf{P}(\cdot)$.

Доказательство. Из определения события $\boldsymbol{A}_t$, соотношений (4.1), (4.2) и равенств $\mathsf{P}_t(A_t) = \sum_{\omega_t \in A_t} \mathsf{P}_t(\{\omega_t\})$, $\sum_{\omega_t \in \Omega_t} \mathsf{P}_t(\{\omega_t\}) = 1$, $t = 1, 2, \ldots, n$, сразу получим, что

$$\begin{aligned}
\mathsf{P}(\boldsymbol{A}_t) &= \mathsf{P}(\Omega_1 \times \Omega_2 \times \ldots \times \Omega_{t-1} \times A_t \times \Omega_{t+1} \times \ldots \times \Omega_n) = \\
&= \sum_{\omega \in \boldsymbol{A}_t} \mathsf{P}(\{\omega\}) = \sum_{\omega_1 \in \Omega_1} \mathsf{P}_1(\{\omega_1\}) \times \sum_{\omega_2 \in \Omega_2} \mathsf{P}_2(\{\omega_2\}) \times \ldots \times \\
&\times \sum_{\omega_{t-1} \in \Omega_{t-1}} \mathsf{P}_{t-1}(\{\omega_{t-1}\}) \times \sum_{\omega_t \in A_t} \mathsf{P}_t(\{\omega_t\}) \times \sum_{\omega_{t+1} \in \Omega_{t+1}} \mathsf{P}_{t+1}(\{\omega_{t+1}\}) \times \ldots \times \\
&\times \sum_{\omega_n \in \Omega_n} \mathsf{P}_n(\{\omega_n\}) = \sum_{\omega_t \in A_t} \mathsf{P}_t(\{\omega_t\}) = \mathsf{P}_t(A_t).
\end{aligned}$$

Пусть теперь $\boldsymbol{A}_{t_1}$, $\boldsymbol{A}_{t_2}, \ldots$, $\boldsymbol{A}_{t_m}$ есть любая подпоследовательность последовательности событий $\boldsymbol{A}_1$, $\boldsymbol{A}_2, \ldots$, $\boldsymbol{A}_n$, где $1 \leqslant t_1 < t_2 < \ldots < t_m \leqslant n$, $2 \leqslant m \leqslant n$. Ради сокращения обозначим событие $\boldsymbol{A}_{t_1} \cap \boldsymbol{A}_{t_2}, \cap \ldots \cap \boldsymbol{A}_{t_m}$ через $\boldsymbol{A}_{1m}$. В этом случае легко видеть, что

$$\begin{aligned}
\boldsymbol{A}_{1m} = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{t_1-1} \times A_{t_1} \times \Omega_{t_1+1} \times \ldots \times \Omega_{t_2-1} \times A_{t_2} \times \Omega_{t_2+1} \times \\
\times \ldots \times \Omega_{t_m-1} \times A_{t_m} \times \Omega_{t_m+1} \times \ldots \times \boldsymbol{\Omega}_n,
\end{aligned}$$

$$\begin{aligned}
&\mathsf{P}(\boldsymbol{A}_{t_1} \cap \boldsymbol{A}_{t_2} \cap \ldots \cap \boldsymbol{A}_{t_m}) = \\
&= \mathsf{P}(\boldsymbol{A}_{1m}) = \sum_{\omega \in \boldsymbol{A}_{1m}} \mathsf{P}(\{\omega\}) = \sum_{\omega_1 \in \Omega_1} \mathsf{P}(\{\omega_1\}) \times \sum_{\omega_2 \in \Omega_2} \mathsf{P}(\{\omega_2\}) \times \ldots \times \\
&\times \sum_{\omega_{t_1-1} \in \Omega_{t_1-1}} \mathsf{P}_{t_1-1}(\{\omega_{t_1-1}\}) \times \sum_{\omega_{t_1} \in A_{t_1}} \mathsf{P}_{t_1}(\{\omega_{t_1}\}) \times \\
&\times \sum_{\omega_{t_1+1} \in \Omega_{t_1+1}} \mathsf{P}_{t_1+1}(\{\omega_{t_1+1}\}) \times \ldots \times \sum_{\omega_{t_2-1} \in \Omega_{t_2-1}} \mathsf{P}_{t_2-1}(\{\omega_{t_2-1}\}) \times \\
&\times \sum_{\omega_{t_2} \in A_{t_2}} \mathsf{P}_{t_2}(\{\omega_{t_2}\}) \times \sum_{\omega_{t_2+1} \in \Omega_{t_2+1}} \mathsf{P}_{t_2+1}(\{\omega_{t_2+1}\}) \times \\
&\times \ldots \times \sum_{\omega_{t_m-1} \in \Omega_{t_m-1}} \mathsf{P}_{t_m-1}(\{\omega_{t_m-1}\}) \times \sum_{\omega_{t_m} \in A_{t_m}} \mathsf{P}_{t_m}(\{\omega_{t_m}\}) \times \\
&\times \sum_{\omega_{t_m+1} \in \Omega_{t_m+1}} \mathsf{P}_{t_m+1}(\{\omega_{t_m+1}\}) \times \ldots \times \sum_{\omega_n \in \Omega_n} \mathsf{P}_n(\{\omega_n\}) = \\
&= \sum_{\omega_{t_1} \in A_{t_1}} \mathsf{P}_{t_1}(\{\omega_{t_1}\}) \times \sum_{\omega_{t_2} \in A_{t_2}} \mathsf{P}_{t_2}(\{\omega_{t_2}\}) \times \ldots \times \sum_{\omega_{t_m} \in A_{t_m}} \mathsf{P}_{t_m}(\{\omega_{t_m}\}) = \\
&= \mathsf{P}(\boldsymbol{A}_{t_1})\, \mathsf{P}(\boldsymbol{A}_{t_2}) \times \ldots \times \mathsf{P}(\boldsymbol{A}_{t_m}).
\end{aligned}$$

Итак, теорема 4.2 доказана. Поэтому вместо любого события A_t из вероятностной модели $(\Omega_t, \mathcal{F}_t, \mathsf{P}_t(\cdot))$ эксперимента E_t можно рассматривать изоморфные ему событие $\boldsymbol{A}_t$ из вероятностной модели $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ более сложного опыта E и наоборот. Другими словами, утверждать, что наступил исход A_t эксперимента E_t, все равно, что сказать, что наступил исход $\boldsymbol{A}_t$ опыта E, т. е. эксперимент E_t закончился исходом A_t, и при этом остальные из $n-1$ опытов закончились какими угодно исходами. Причинная независимость экспериментов E_1, $E_2, \ldots$, E_n совершенно естественно приводит к статистической независимости в совокупности случайных событий $\boldsymbol{A}_1$, $\boldsymbol{A}_2, \ldots$, $\boldsymbol{A}_n$ специального вида. Каждая из вероятностных моделей $(\Omega_t, \mathcal{F}_t, \mathsf{P}_t(\cdot))$ при заданном $t = 1, 2, \ldots, n$ естественно проще. Однако более сложная вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ позволяет изучать общие вероятностные свойства сразу всех опытов E_1, $E_2, \ldots$, E_n.

Вероятностную модель последовательности независимых испытаний рекомендуется применять в том случае, когда эксперимент E по физическому содержанию носит сложный характер, и найти соответствующие вероятности бывает очень громоздко и трудно. Тогда опыт E, если это возможно, представляют составным и разбивают на последовательность n независимых испытаний E_1, $E_2, \ldots$, E_n. Естественно этот прием следует использовать, если вероятностная модель каждого из испытаний E_t, где $t = 1, 2, \ldots, n$, определяется совсем просто либо с помощью классического способа, либо с помощью статистических вероятностей.

Рассмотрим иллюстративный пример построения составной вероятностной модели.

Пример 4.4. Буквы слова «математика» написаны каждая на отдельной карточке. Одинаковые по форме карточки тщательно перемешаны, после чего последовательно и без возвращения извлекаются четыре. Выбранные карточки кладутся слева направо. Определить вероятность получения слова «тема» или, другими словами, наступления событие $\boldsymbol{A}$.

Решение. Эту задачу можно решить способами, которые использованы при исследовании примера 3.10. Однако указанные способы приводят к громоздким вычислениям. Комплекс условий составного эксперимента E имеет вид $\Sigma = \{u_1$ — десять карточек, на каждой из которых написаны одна буква из слова «математика»; u_2 — наудачу по одной отбирается четыре карточки без возвращения, и последовательно располагаются слева направо; u_3 — условие для прочтения записи на отобранных карточках;...$\}$.

Введем в рассмотрение следующие эксперименты.

• Эксперимент E_1 с комплексом условий $\Sigma_1 = \{u_{11}$ — десять карточек, на каждой из которых написаны одна из букв «а», «а», «а», «е», «и», «к», «м», «м», «т», «т»; u_{12} — наудачу отбирается одна карточка без возвращения и располагается первой в записи будущего слова; u_{13} — условие для прочтения записи на отобранной карточке;...$\}$. Если событие A_1 означает появление буквы «т» при проведении этого эксперимента, то классическое определение вероятности дает

$\mathsf{P}_1(A_1) = 2/10 = 1/5$. Не представляет особого труда построить вероятностную модель $(\Omega_1, \mathcal{F}_1, \mathsf{P}_1(\cdot))$ для классического эксперимента E_1.

• Эксперимент E_2 с комплексом условий $\Sigma_2 = \{u_{21}$ — девять карточек, на каждой из которых написаны одна из букв «а», «а», «а», «е», «и», «к», «м», «м», «т»; u_{22} — наудачу отбирается одна карточка без возвращения и располагается второй в записи будущего слова; u_{23} — условие для прочтения записи на отобранной карточке;...$\}$. Если событие A_2 означает появление буквы «е» при проведении эксперимента E_2, то классическое определение вероятности дает $\mathsf{P}_2(A_2) = 1/9$. Здесь вероятностная модель классического эксперимента E_2 есть $(\Omega_2, \mathcal{F}_2, \mathsf{P}_2(\cdot))$.

• Эксперимент E_3 с комплексом условий $\Sigma_3 = \{u_{31}$ — восемь карточек, на каждой из которых написаны одна из букв «а», «а», «а», «и», «к», «м», «м», «т»; u_{32} — наудачу отбирается одна карточка без возвращения и располагается третьей в записи будущего слова; u_{33} — условие для прочтения записи на отобранной карточке;...$\}$. Если событие A_3 означает появление буквы «м» при проведении эксперимента E_3, то классическое определение вероятности дает $\mathsf{P}_3(A_3) = 2/8 = 1/4$. Вероятностная модель классического эксперимента E_3 обозначена через $(\Omega_3, \mathcal{F}_3, \mathsf{P}_3(\cdot))$.

• Эксперимент E_4 с комплексом условий $\Sigma_4 = \{u_{41}$ — семь карточек, на каждой из которых написаны одна из букв «а», «а», «а», «и», «к», «м», «т»; u_{42} — наудачу отбирается одна карточка без возвращения и располагается четвертой в записи будущего слова; u_{43} — условие для прочтения записи на отобранной карточке;...$\}$. Если случайное событие A_4 означает появление буквы «а» при проведении эксперимента E_4, то классическое определение вероятности дает $\mathsf{P}_4(A_4) = 3/7$. Здесь вероятностная модель классического эксперимента E_4 имеет вид $(\Omega_4, \mathcal{F}_4, \mathsf{P}_4(\cdot))$.

Представим теперь составной эксперимент E в виде последовательности независимых испытаний E_1, E_2, E_3 и E_4. Пусть событие $\boldsymbol{A}$ состоит в том, что в результате проведения эксперимента $E = \{E_t;\, t = 1, 2, 3, 4\}$ получается слово «тема». Тогда вероятность $\mathsf{P}(\boldsymbol{A}) = \mathsf{P}_1(A_1) \times \mathsf{P}_2(A_2) \times \mathsf{P}_3(A_3) \times \mathsf{P}_4(A_4) = (1/5) \times (1/9) \times (1/4) \times (3/7) = 1/420 \approx 0{,}002$. Здесь вероятностная модель классического эксперимента E есть $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$.

§ 2. Схема независимых испытаний Бернулли

2.1. Вычисление вероятностей событий в биномиальной схеме независимых испытаний. Рассмотрим теперь частный случай процедуры последовательности независимых испытаний. Для наглядности обратимся к примерам 4.1–4.3. В отличие от примеров 4.1 и 4.3 в примере 4.2 последовательность причинно-независимых испытаний такова, что один и тот же опыт E_0 повторяется пять раз. На самом деле в реальных задачах довольно часто встречается так называемая биномиальная схема или схема Бернулли. В такой схеме, во-первых, некоторый заданный статистически устойчивый эксперимент E_0 приходится n-кратно повторять причинно-независимым образом. Следовательно, при каждом фиксированном номере $t = 1, 2, \ldots, n$ опыт E_t состоит в проведении эксперимента E_0. Отсюда все вероятности $\mathsf{P}_1(\{\omega_1\})$, $\mathsf{P}_2(\{\omega_2\}), \ldots$, $\mathsf{P}_n(\{\omega_n\})$ будут равными, если символы ω_1, $\omega_2, \ldots$, ω_n задают описание одного и того же элементарного исхода эксперимента E_0. Во-вторых, в каждом испытании E_t, отвлекаясь от возможного разнообразия его исходов, прежде всего, интересуются тем, наступило или не наступило некоторое заранее фиксированное событие A эксперимента E_0. Обычно исход A, ради простоты, условно называют успехом эксперимента E_0, а исход $\overline{A}$ — неудачей эксперимента E_0, причем нас интересует не исход отдельного опыта, а результат всей серии независимых испытаний из n опытов над экспериментом E_0. Договоримся считать несколько фиксированных опытов независимыми, если вероятность появления или не появления каждого исхода в том или ином опыте не зависит от того, как исходы появлялись в остальных опытах. Рассмотрим три примера независимых испытаний:

1) несколько последовательных бросков монеты, например, четыре раза;
2) многократное бросание игральной кости, например, пятикратное;
3) несколько последовательных извлечений карты из игральной колоды представляет собой схему независимых испытаний при условии, что вынутая карта каждый раз возвращается в колоду, и карты тщательно перемешиваются.

Приведем два примера зависимых опытов:

1) многократное извлечение карты из колоды без ее возвращения;
2) проведение нескольких выстрелов из орудия с корректировкой, т. е. с учетом результатов предыдущих выстрелов.

Итак, в схеме Бернулли при каждом фиксированном $t = 1, 2, \ldots, n$ событие A_t теперь означает наступление исхода A в эксперименте E_t или, что то же самое, наступление события A в испытании E_0 за номером t. В связи с этим при любом $t = 1, 2, \ldots, n$ получим равенства:

$$\mathsf{P}_t(A_t) = \sum_{\omega_t \in A_t} \mathsf{P}_t(\{\omega_t\}) = \sum_{\omega_t \in A} \mathsf{P}_t(\{\omega_t\}) = \\ = \sum_{\omega_1 \in A} \mathsf{P}_1(\{\omega_1\}) = \sum_{\omega_1 \in A_1} \mathsf{P}_1(\{\omega_1\}) = \mathsf{P}_1(A_1).$$

Поэтому вероятности $\mathsf{P}_1(A_1)$, $\mathsf{P}_2(A_2), \ldots$, $\mathsf{P}_n(A_n)$ можно обозначить через p, а вероятности $\mathsf{P}_1(\overline{A}_1)$, $\mathsf{P}_2(\overline{A}_2), \ldots$, $\mathsf{P}_n(\overline{A}_n)$ через q, где $p + q = 1$. Для составного эксперимента $E = \{E_1, E_2, \ldots, E_n\}$ в схеме Бернулли с параметрами n и p наибольший практический интерес представляют следующие его исходы и соответствующие им вероятности.

• Исход $\boldsymbol{B}_m \in \mathcal{F}$, который заключается в том, что в результате n-кратного проведения опыта E_0 событие A будет наблюдаться ровно m раз, где $m = 1, 2, \ldots, n$. Для наступления события $\boldsymbol{B}_m$ безразлично, в каких именно опытах произойдет событие A, а в каких $\overline{A}$. Важно лишь, что бы общее число появлений события A в данной серии испытаний равнялось m. Зафиксируем, прежде всего, значение $m = 1, 2, \ldots, n$. Пусть t_1, $t_2, \ldots$, t_m суть номера испытаний опыта E_0, в каждом из которых появится успешный исход A. Для определенности будем полагать, что $1 \leqslant t_1 < t_2 < \ldots < t_m \leqslant n$. Тогда очевидно числа 1, $2, \ldots$, $t_1 - 1$, $t_1 + 1, \ldots$, $t_2 - 1$, $t_2 + 1, \ldots$, $t_m - 1$, $t_m + 1, \ldots$, n будут номерами испытаний опыта E_0, в каждом из которых наступит неудачное событие $\overline{A}$. При фиксированном m число всех неупорядочных наборов (сочетаний) из n элементов — номеров множества $\{1, 2, \ldots, n\}$ по m вида $\{t_1, t_2, \ldots, t_m\}$, как известно, равно C_n^m. Из определения случайного события $\boldsymbol{B}_m$ получим, что

$$\boldsymbol{B}_m = \bigcup_{\{t_1, t_2, \ldots, t_m\}} \overline{A}_1 \times \overline{A}_2 \times \ldots \times \overline{A}_{t_1 - 1} \times A_{t_1} \times \overline{A}_{t_1+1} \times \ldots \times$$
$$\times \overline{A}_{t_2 - 1} \times A_{t_2} \times \overline{A}_{t_2+1} \times \ldots \times \overline{A}_{t_m - 1} \times A_{t_m} \times \overline{A}_{t_m+1} \times \ldots \times \overline{A}_n. \quad (4.3)$$

Здесь объединение несовместимых событий из $\mathcal{F}$ ведется по всевозможным сочетаниям вида $\{t_1, t_2, \ldots, t_m\}$. Такое представление случайного события $\boldsymbol{B}_m$ при каждом $m = 1, 2, \ldots, n$ и очевидное представление события $\boldsymbol{B}_0 = \overline{A}_1 \times \overline{A}_2 \times \ldots \times \overline{A}_n$ позволяют вычислить вероятность заданного числа появлений события A при n причинно-независимых испытаниях опыта E_0. Действительно, из (4.3) с учетом аксиомы аддитивности, утверждений и доказательства теоремы 4.2 непосредственно находим:

$$\mathsf{P}(\boldsymbol{B}_m) = \sum_{\{t_1, t_2, \ldots, t_m\}} \mathsf{P}_1(\overline{A}_1) \times \mathsf{P}_2(\overline{A}_2) \times \ldots \times \mathsf{P}_{t_1-1}(\overline{A}_{t_1-1}) \times$$
$$\times \mathsf{P}_{t_1}(\overline{A}_{t_1}) \times \mathsf{P}_{t_1+1}(\overline{A}_{t_1+1}) \times \ldots \times \mathsf{P}_{t_2-1}(\overline{A}_{t_2-1}) \times \mathsf{P}_{t_2}(\overline{A}_{t_2}) \times \mathsf{P}_{t_2+1}(\overline{A}_{t_2+1}) \times$$
$$\times \ldots \times \mathsf{P}_{t_m-1}(\overline{A}_{t_m-1}) \times \mathsf{P}_{t_m}(\overline{A}_{t_m}) \times \mathsf{P}_{t_m+1}(\overline{A}_{t_m+1}) \times \ldots \times \mathsf{P}_n(\overline{A}_n) =$$
$$= \sum_{\{t_1, t_2, \ldots, t_m\}} p^m q^{n-m} = C_n^m p^m q^{n-m},$$

$$\mathsf{P}(\boldsymbol{B}_0) = \mathsf{P}(\overline{A}_1 \times \overline{A}_2 \times \ldots \times \overline{A}_n) = \prod_{t=1}^{n} \mathsf{P}_t(\overline{A}_t) = q^n.$$

Так как $\boldsymbol{B}_0 \cup \boldsymbol{B}_1 \cup \ldots \cup \boldsymbol{B}_n = \Omega$ и случайные события $\boldsymbol{B}_0$, $\boldsymbol{B}_1, \ldots$, $\boldsymbol{B}_n$ попарно несовместимы, то $\mathsf{P}(\boldsymbol{B}_0 \cup \boldsymbol{B}_1 \cup \ldots \cup \boldsymbol{B}_n) = \mathsf{P}(\boldsymbol{B}_0) + \mathsf{P}(\boldsymbol{B}_1) + \ldots + \mathsf{P}(\boldsymbol{B}_n) = 1$. То же самое показывается следующим образом: $1 = (p + q)^n = \sum_{m=0}^{n} C_n^m p^m q^{n-m} = \sum_{m=0}^{n} \mathsf{P}(\boldsymbol{B}_m)$, если принять во внимание известное разложение бинома и равенство $C_n^0 = 1$. Поэтому изучаемую схему из n причинно-независимых испытаний и называют *биномиальной схемой*. Итак, последовательность $\boldsymbol{B}_0$, $\boldsymbol{B}_1, \ldots$, $\boldsymbol{B}_n$ образует полную группу несовместимых событий из $\mathcal{F}$.

• Исход $\cup_{m=1}^n \boldsymbol{B}_m \in \mathcal{F}$, который состоит в том, что в результате n причинно-независимых испытаний опыта E_0 событие A появится по крайней мере один раз. Так как последовательность $\boldsymbol{B}_0$, $\boldsymbol{B}_1, \ldots$, $\boldsymbol{B}_n$ образует полную группу несовместимых событий из $\mathcal{F}$, то $\mathsf{P}\big(\cup_{m=1}^n \boldsymbol{B}_m\big) = 1 - q^n$.

• Исход $\cup_{m=2}^n \boldsymbol{B}_m \in \mathcal{F}$, который состоит в том, что в результате n причинно-независимых испытаний опыта E_0 событие A появится не менее двух раз. Для вероятности события $\cup_{m=2}^n \boldsymbol{B}_m$ точно так же находим $\mathsf{P}\big(\cup_{m=2}^n \boldsymbol{B}_m\big) = 1 - \mathsf{P}(\boldsymbol{B}_0) - \mathsf{P}(\boldsymbol{B}_1) = 1 - q^n - C_n^1 pq^{n-1} = 1 - q^n - npq^{n-1}$.

• Исход $\cup_{m=0}^k \boldsymbol{B}_m \in \mathcal{F}$, который означает, что в результате n причинно-независимых испытаний опыта E_0 событие A появится не более k раз. Для вероятности события $\cup_{m=0}^k \boldsymbol{B}_m$ имеем $\mathsf{P}\big(\cup_{m=0}^k \boldsymbol{B}_m\big) = \sum_{m=0}^k C_n^m p^m q^{n-m}$.

• Исход $\cup_{m=k}^n \boldsymbol{B}_m \in \mathcal{F}$ заключается в том, что в результате n причинно-независимых испытаний опыта E_0 событие A появится не менее k раз. Для вероятности события $\cup_{m=k}^n \boldsymbol{B}_m$ легко находим $\mathsf{P}\big(\cup_{m=k}^n \boldsymbol{B}_m\big) = \sum_{m=k}^n C_n^m p^m q^{n-m}$.

• Исход $\cup_{m=k}^s \boldsymbol{B}_m \in \mathcal{F}$ заключается в том, что в результате n испытаний опыта E_0 событие A появится не менее k раз и не более s раз. Для вероятности случайного события $\cup_{m=k}^s \boldsymbol{B}_m$ имеем $\mathsf{P}\big(\cup_{m=k}^s \boldsymbol{B}_m\big) = \sum_{m=k}^s C_n^m p^m q^{n-m}$, где $0 \leqslant k \leqslant s \leqslant n$.

В биномиальной схеме знаменитый математик Якоб Бернулли (1654–1705) рассматривал опыт E_0 лишь с двумя элементарными исходами A и $\overline{A}$. В этом простейшем варианте биномиальной схемы элементарные исходы A и $\overline{A}$ могут быть описаны цифрами 1 и 0, т. е. $A = \{1\}$ и $\overline{A} = \{0\}$. В соответствии с этим при каждом $t = 1, 2, \ldots, n$ будем иметь $\Omega_t = \{0, 1\}$, $\mathcal{F}_t = \{A, \overline{A}, \Omega_t, \varnothing\}$, $\mathsf{P}_t(A) = p$, $\mathsf{P}_t(\overline{A}) = q$. Поэтому получаем, что $\Omega = \{\omega = (\omega_1, \omega_2, \ldots, \omega_n)\colon \omega_1, \omega_2, \ldots, \omega_n \in \{0, 1\}\}$, $\mathcal{F} = \{\boldsymbol{A}\colon \boldsymbol{A} \subset \Omega\}$, $\mathsf{P}(\{\omega\}) = p^m q^{n-m}$. Здесь $m = \sum_{t=1}^n \omega_t$ — число успешных исходов при n независимых испытаниях данного опыта E_0 и при наблюдении элементарного события $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ составного эксперимента E. Из формулы $\mathsf{P}(\{\omega\}) = p^m q^{n-m}$ видно, что элементарные исходы эксперимента E не являются равновозможными при $p \neq q$. Число всех тех элементарных исходов $\{\omega = (\omega_1, \omega_2, \ldots, \omega_n)\}$, для каждого из которых имеет место равенство $\sum_{t=1}^n \omega_t = m$, равно C_n^m. Отсюда получим, что $\sum_{\omega \in \Omega} \mathsf{P}(\{\omega\}) = \sum_{m=0}^n C_n^m p^m q^{n-m}$, где

$\mathsf{P}(\boldsymbol{B}_m) = C_n^m p^m q^{n-m}$. Последнее равенство называется *формулой Бернулли*.

Таким образом, если в биномиальной схеме будем интересоваться только лишь тем, наступило или не наступило в результате опыта E_0 событие A, то даже для эксперимента E_0 с большим многообразием его исходов естественно для составного эксперимента E принять более простую вероятностную модель с двумя элементарными исходами A и $\overline{A}$.

2.2. Биномиальные вероятности и их свойства. Если $p = \mathrm{const}$, то вероятность $\mathsf{P}(\boldsymbol{B}_m) = C_n^m p^m q^{n-m}$ будет функцией от двух аргументов n и m. Очень часто эту вероятность обозначают через $P_n(m)$ и называют биномиальной. Зафиксируем теперь параметры n, p биномиальной схемы последовательности независимых испытаний. Тогда $P_n(m)$ будет функцией от аргумента m, принимающего значения $0, 1, \ldots, n$. Определим все такие экстремальные значения m_0, для которых имеет место соотношение вида: $P_n(m_0) = \max\{P_n(0), P_n(1), \ldots, P_n(n)\}$.

Каждая такая величина m_0 есть наивероятнейшее число появлений события A в схеме Бернулли. Здравый смысл и интуиция говорят о том, что значения m_0 должны быть близки к числу np при достаточно большом n. Действительно, статистически устойчивый эксперимент E_0 можно проводить, по крайней мере, теоретически любое конечное число раз причинно независимым образом. Тогда появление или непоявление некоторого исхода A эксперимента E_0 в каждом конкретном испытании не будет зависеть от исходов предыдущих испытаний. Статистическая вероятность $\mu(A,n)/n$ события A при n испытаниях опыта E_0 колеблется около теоретической вероятности p. Отсюда число $\mu(A,n)$ наступлений события A при n испытаниях опыта E_0, скорее всего, будет близќо к величине np. Перейдем теперь к точным выкладкам.

С этой целью рассмотрим последовательность из вероятностей $P_n(0)$, $P_n(1), \ldots$, $P_n(n)$ и отношение последующей вероятности к предыдущей вероятности. Используя формулу Бернулли и проделывая простые преобразования, последовательно при $m = 1, 2, \ldots, n$ находим:

$$P_n(m)/P_n(m-1) = C_n^m p^m q^{n-m} / C_n^{m-1} p^{m-1} q^{n-m+1} =$$
$$= \frac{n(n-1)\times\ldots\times(n-m+1)\,p\,(m-1)!}{n(n-1)\times\ldots\times(n-m+2)\,q\,m!} =$$
$$= \frac{(n-m+1)\,p}{mq} = 1 + \frac{(1+n)\,p - m}{mq}.$$

Отсюда, для всех значений m меньших, чем $(n+1)\,p$, справедливо неравенство $P_n(m-1) < P_n(m)$. Если $(n+1)\,p$ есть натуральное число, то имеет место равенство $P_n(m-1) = P_n(m)$ при $m = (n+1)\,p$. Наконец, при $m > (n+1)\,p$ выполняется неравенство $P_n(m-1) > P_n(m)$.

Учитывая все это, выводим, что при значениях m, меньших $(n+1)\,p$, функция $P_n(m)$ возрастает, а при значениях m, больших $(n+1)\,p$, функция $P_n(m)$ убывает. Тем самым, если число $(n+1)\,p$ не является целым, то функция $P_n(m)$ имеет единственный максимум при $m_0 = [(n+1)p]$. Здесь символ $[(n+1)\,p]$ обозначает целую часть числа $(n+1)\,p$ и он равен наибольшему целому числу, меньшему $(n+1)\,p$. Если же $(n+1)\,p$ есть натуральное число, то функция $P_n(m)$ достигает одного и того же максимума при двух различных значениях m_0, а именно, при $m_0 = (n+1)\,p - 1 = np - q$ и при $m_0 = (n+1)p$. В качестве иллюстрации применения приведенных формул и рассуждений в биномиальной схеме рассмотрим примеры.

Пример 4.5. Вероятность поражения цели при одиночном выстреле любого из n орудий равна $p = 0{,}01$. Сколько потребуется орудий, что бы вероятность поражения цели при одном залпе была больше половины?

Решение. Пусть стреляют одновременно из n орудий. Используя ранее приведенные формулы, найдем вероятность того, что заряд хотя бы одного из орудий попадет в цель: $\mathsf{P}\big(\cup_{m=1}^{n}\boldsymbol{B}_m\big) = 1 - q^n = 1 - (0{,}99)^n$. Затем из требования задачи шаг за шагом выводим $1 - (0{,}99)^n > 2^{-1}$, $(0{,}99)^n < 2^{-1}$, $n\lg(99/100) < \lg(1/2)$, $n(\lg 99 - \lg 100) < -\lg 2$, $\lg 2 < n(2 - \lg 99)$, $n > \lg 2/(2 - \lg 99) \approx 70$. Итак, число орудий должно быть более 70.

Пример 4.6. Ракетная установка может за некоторое фиксированное время в своей зоне выпустить только одну ракету по воздушной цели. Вероятность попадания в истребитель при одиночном пуске ракеты любой из n однотипных ракетных установок равна $p = 1/3$. Считается, что для поражения современного истребителя требуется хотя бы одно попадание. Определить наименьшее число n_0 ракетных установок, при котором вероятность поражения истребителя при одновременном пуске ракет из всех установок была больше $211/243 \approx 0{,}87$. Найти наивероятнейшее число попаданий ракет при одновременном их запуске по истребителю.

Решение. Вероятность поражения истребителя при одновременном пуске ракет из n установок равна $\mathsf{P}\big(\cup_{m=1}^{n}\boldsymbol{B}_m\big) = 1 - q^n = 1 - (2/3)^n$, где $q = 2/3$. По условию задачи имеем неравенство $1 - (2/3)^n > 211/243$, или, что то же самое, $(2/3)^5 > (2/3)^n$. Отсюда находим $n > 5$, следовательно, наименьшее число ракетных установок $n_0 = 6$, а наивероятнейшее число попаданий ракет при одновременном их запуске по истребителю определяется формулой $m_0 = [(n_0+1)\,p] = 2$. Вероятность такого числа попаданий найдем по формуле Бернулли: $P_6(2) = C_6^2(1/3)^2(2/3)^4 \approx 0{,}33$.

Пример 4.7. Два равносильных шахматиста играют ряд партий, и ничьи в счет не идут. Что вернее всего в счете: (1:1), или (2:2), или (3:3) и т. д.?

Решение. Пусть шахматисты сыграли n результативных партий. Вероятность выигрыша первого шахматиста, так же, как и вероятность выигрыша второго, одинакова для всех партий и равна $p = 1/2$. Нам нужно найти вероятность того, что в n результативных партиях счет будет равный. На серию $n = 2k$ результативных партий смотрим как на серию $n = 2k$ независимых испытаний Бернулли. Вероятность успеха (будем называть успехом выигрыш первого шахматиста) одинакова для всех испытаний и равна $p = 0{,}5$. Возьмем для удобства $n = 2k$, где k — целое число. Необходимо найти вероятность $P_n(k) = P_{2k}(k)$. По формуле Бернулли определяем

$$P_{2k}(k) = C_{2k}^{k}(0{,}5)^k(1-0{,}5)^{2k-k} = (2k)!\,(k!\,k!)^{-1}(0{,}5)^{2k}.$$

Значит, вероятность того, что в $n = 2k$ результативных партиях счет будет $(k:k)$ равна

$$P_{2k}(k) = (2k)!\,(k!\,k!)^{-1}(0{,}5)^{2k}.$$

Определим вероятность того, что в $n = 2(k+1)$ результативных партиях счет будет $((k+1):(k+1))$. В этом случае следует изучать серию из $n = 2(k+1)$ независимых испытаний с вероятностью успеха $p = 0{,}5$. Аналогичным способом получаем, что

$$P_{2(k+1)}(k+1) = (2(k+1))!\,((k+1)!\,(k+1)!)^{-1}(0{,}5)^{2(k+1)}.$$

Преобразуя теперь выражение для $P_{2k}(k)$ с целью определения связи между вероятностями $P_{2k}(k)$ и $P_{2(k+1)}(k+1)$, находим, что

$$P_{2k}(k) = \frac{(2k)!\,(2k+1)\,(2k+2)\,(k+1)^2 2^2}{(k!)^2\,(2k+1)\,(2k+2)\,(k+1)^2 2^{2k+2}} = \frac{2k+2}{2k+1}P_{2k+2}(k+1) =$$
$$= (1+(2k+1)^{-1})\,P_{2(k+1)}(k+1).$$

Из полученного соотношения

$$P_{2k}(k) = (1+(2k+1)^{-1})\,P_{2(k+1)}(k+1)$$

видно, что счет $(k:k)$ более вероятен, чем счет $((k+1):(k+1))$. Значит, наиболее вероятен счет вида $(1:1)$. Расчеты по формуле Бернулли показывают, что шахматисты могут закончить игру при счете $(1:1)$ с вероятностью 64/128, при счете $(2:2)$ — с вероятностью 48/128, при счете $(3:3)$ — с вероятностью 40/128 и т. д.

Пример 4.8. Вероятность p поражения самолета из ракеты равна 0,3. Выпущено $n = 15$ ракет. Построить график биномиальных вероятностей $P_n(m)$, $m = 0, 1, \ldots, n$, и найти наивероятнейшее число попаданий в цель.

Решение. График функции $P_n(m)$ представлен на рис. 4.1.

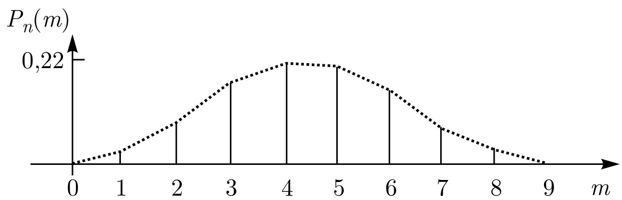


Рис. 4.1

Наивероятнейшее значение m_0 определяется следующим способом: $m_0 = [(1+n)p] = [(1+15) \times 0{,}3] = 4$. Поэтому для этого опыта $P_{15}(4) = C_{15}^4 (3/10)^4 (7/10)^{11} \approx 0{,}22$. Для сравнения укажем, что $P_{15}(3) \approx 0{,}17$, $P_{15}(2) \approx 0{,}09$, $P_{15}(1) \approx 0{,}03$, $P_{15}(5) \approx 0{,}21$, $P_{15}(6) \approx 0{,}15$, $P_{15}(7) \approx 0{,}08$.

§ 3. Приближенные формулы в биномиальной схеме

3.1. Предельные теоремы Муавра–Лапласа. В прикладных задачах при больших значениях n, m, k, s часто приходится вычислять вероятности

$$\mathsf{P}(\boldsymbol{B}_m) = C_n^m p^m q^{n-m}, \quad \mathsf{P}\Big(\bigcup_{m=k}^{s} \boldsymbol{B}_m\Big) = \sum_{m=k}^{s} C_n^m p^m q^{n-m}.$$

Непосредственный подсчет этих вероятностей по этим формулам представляет значительные трудности даже на современных персональных компьютерах, главным образом из-за больших чисел C_n^m и малых чисел $p^m q^{n-m}$. Пусть требуется решить такую задачу.

Пример 4.9. По каналу связи независимым образом передано 10 000 знаков. Вероятность искажения любого передаваемого знака равна 0,0008. Какова вероятность того, что получится: 1) ровно 14 искажений, 2) число искажений не менее 10 и не более 19?

Решение. В этом примере имеем биномиальную схему с параметрами $n = 10\,000$, $p = 0{,}0008$. Поэтому по формулам Бернулли получим:

1) $\mathsf{P}(\boldsymbol{B}_{14}) = C_{10\,000}^{14} (0{,}0008)^{14} (0{,}9992)^{9\,986}$, $m = 14$;
2) $\mathsf{P}\big(\cup_{m=10}^{19} \boldsymbol{B}_m\big) = C_{10\,000}^{m} (0{,}0008)^{m} (0{,}9992)^{10\,000-m}$, $k = 10$, $s = 19$.

Эти выражения достаточно просты, но получить их точные числовые значения слишком затруднительно. Поэтому возникает необходимость в получении приближенных формул для вероятностей $\mathsf{P}(\boldsymbol{B}_m) = C_n^m p^m q^{n-m}$, $\mathsf{P}\big(\cup_{m=k}^{s}\boldsymbol{B}_m\big) = \sum_{m=k}^{s} C_n^m p^m q^{n-m}$ при достаточно больших значениях n, m, k, s. Такие формулы, позволяющие с достаточной степенью точности вычислять указанные вероятности, впервые были получены Муавром, Лапласом и Пуассоном в результате доказательства простейших предельных теорем в схеме Бернулли. Доказательства этих теорем основываются на комбинаторных методах с применением хорошо известной формулы Стирлинга–Муавра вида

$$n! = (2n\pi)^{1/2} n^n \exp\{-n\} \exp\{\theta_n\}, \qquad 0 < \theta_n < 1/12\,n, \tag{4.4}$$

а также на теории пределов и на определении интегральных сумм из математического анализа. Другими словами, эти доказательства носят чисто технический характер и не связаны с вероятностными методами. Заметим, что в предельных теоремах величины n, m, k, s естественно становятся переменными и стремятся к символу $+\infty$, ибо ни одно в отдельности взятое число не может быть квалифицировано, как достаточно или очень большое. Перейдем сначала к теоремам Муавра–Лапласа.

Теорема 4.3 (локальная предельная теорема Муавра–Лапласа). *Пусть в схеме независимых испытаний* $0 < p, q < 1$ *и при* $n \to \infty$ *выполняется соотношение* $|(m - np)/(npq)^{1/2}| < c$, *где* c *— любая фиксированная постоянная. Тогда*

$$\lim_{n\to\infty} C_n^m p^m q^{n-m} (2\pi npq)^{1/2} (\exp\{-(m-np)^2/(2npq)\})^{-1} = 1.$$

Это предельное равенство выполняется равномерно для всех таких m, *для которых величина* $x = x_m = x_m(n) = (m - np)/(npq)^{1/2}$ *принадлежит отрезку* $[-c, c]$.

Доказательство. Из соотношения $x = (m - np)/(npq)^{1/2}$ получим, что $m = np + x(npq)^{1/2}$ и $n - m = nq - x(npq)^{1/2}$. Отсюда следует, что при любом фиксированном значении величины $x \in [-c, c]$ и при $n \to \infty$ числа m, $n - m$ стремятся к бесконечности. Теперь следующую последовательность $\{(2\pi npq)^{1/2}\mathsf{P}(\boldsymbol{B}_m)/\exp\{-x^2/2\};\ n = 1, 2, \ldots\}$ можно представить в виде

$$\begin{aligned}(2\pi npq)^{1/2}\mathsf{P}(\boldsymbol{B}_m)\exp\{-x^2/2\} &= (2\pi npq)^{1/2} C_n^m p^m q^{n-m} / \exp\{-x^2/2\} = \\ &= (2\pi npq)^{1/2} C_n^{np+x\sqrt{npq}}\, p^{np+x\sqrt{npq}}\, q^{nq-x\sqrt{npq}} / \exp\{-x^2/2\} = u_n(x), \\ &\qquad n = 1, 2, \ldots,\end{aligned}$$

и рассматривать ее как функциональную последовательность по переменной $x \in [-c, c]$. Поэтому в формулировке теоремы и говорится о равномерной сходимости последовательности $\{u_n(x);\ n = 1, 2, \ldots\}$ по аргументу $x \in [-c, c]$. Используя формулу Стирлинга–Муавра (4.4), непосредственно находим:

$$
\begin{aligned}
&(2\pi npq)^{1/2}\mathsf{P}(\boldsymbol{B}_m) = (2\pi npq)^{1/2}C_n^m p^m q^{n-m} = \\
&\qquad = p^m q^{n-m}(2\pi npq)^{1/2} n! \times (m!\,(n-m)!)^{-1} = \\
&\qquad = p^m q^{n-m}(2\pi npq)^{1/2}(2\pi n)^{1/2} n^n \exp\{-n\}\exp\{-n\}\exp\{\theta_n\}\times \\
&\qquad \times (2\pi m)^{-1/2} \times [m^m \exp\{-m\}\exp\{\theta_m\}(2\pi(n-m))^{1/2}(n-m)^{n-m}\times \\
&\qquad \times \exp\{-(n-m)\}\exp\{\theta_{n-m}\}]^{-1} = [pqn^2m^{-1}(n-m)^{-1}]^{1/2}\times \\
&\qquad\qquad \times (nm^{-1}p)^m[n(n-m)^{-1}q]^{n-m}\exp\{\theta_n - \theta_m - \theta_{n-m}\}.
\end{aligned}
$$

Простые выкладки показывают, что имеет место

$$
\begin{aligned}
&\lim_{n\to\infty}[pqn^2m^{-1}(n-m)^{-1}]^{1/2} = \\
&\qquad = \lim_{n\to\infty}\{pqn^2[np + x(npq)^{1/2}]^{-1}[nq - x(npq)^{1/2}]^{-1}\}^{1/2} = 1
\end{aligned}
$$

равномерно относительно $x \in [-c, c]$.

Действительно,

$$
\begin{aligned}
&\ln[pqn^2m^{-1}(n-m)^{-1}]^{1/2} = \ln n + 2^{-1}\ln p + 2^{-1}\ln q - 2^{-1}\ln m - \\
&\quad - 2^{-1}\ln(n-m) = \ln n + 2^{-1}\ln p + 2^{-1}\ln q - 2^{-1}\ln(np + x(npq)^{1/2}) - \\
&\quad - 2^{-1}\ln(nq - x(npq)^{1/2}) = \ln n + 2^{-1}\ln p + 2^{-1}\ln q - 2^{-1}\ln np - \\
&\quad - 2^{-1}\ln[1 + xq^{1/2}(np)^{-1/2}] - 2^{-1}\ln nq - 2^{-1}\ln[1 - xp^{1/2}(nq)^{-1/2}] = \\
&\quad = \ln n + 2^{-1}\ln p + 2^{-1}\ln q - 2^{-1}\ln np - 2^{-1}xq^{1/2}(np)^{-1/2} + \\
&\quad + 4^{-1}x^2q(np)^{-1} + O(n^{-3/2}) - 2^{-1}\ln nq + 2^{-1}xp^{1/2}(nq)^{-1/2} + \\
&\quad + 4^{-1}x^2p(nq)^{-1} + O(n^{-3/2}) = n^{-1/2}2^{-1}x(p^{1/2}q^{-1/2} - p^{-1/2}q^{1/2}) + \\
&\qquad + n^{-1}4^{-1}x^2(pq^{-1} + p^{-1}q) + O(n^{-3/2}) = O(n^{-1/2})
\end{aligned}
$$

или, что то же самое,

$$
[pqn^2m^{-1}(n-m)^{-1}]^{1/2} = \exp\{O(n^{-1/2})\} = 1 + O(n^{-1/2}). \tag{4.5}
$$

Здесь $O(u_n)$ и u_n суть бесконечно малые величины одного порядка при $n \to \infty$ и еще мы воспользовались разложением в степенной ряд логарифма $\ln(1+y) = y - 2^{-1}y^2 + O(y^3)$, когда величины $q^{1/2}(np)^{-1/2}$, $p^{1/2}(nq)^{-1/2}$ при достаточно больших n могут быть сделаны как угодно малыми. Так как

$$
\begin{aligned}
&|\theta_n - \theta_m - \theta_{n-m}| \leqslant |\theta_n| + |\theta_m| + |\theta_{n-m}| \leqslant \\
&\qquad \leqslant (12)^{-1}[(n^{-1} + (np + x(npq)^{1/2})^{-1} + (nq - x(npq)^{1/2})^{-1}] = \\
&\qquad\qquad = O(n^{-1/2}),
\end{aligned} \tag{4.6}
$$

то $\lim\limits_{n\to\infty}(\theta_n - \theta_m - \theta_{n-m}) = 0$ равномерно относительно $x \in [-c, c]$ и, следовательно, $\lim\limits_{n\to\infty}\exp\{\theta_n - \theta_m - \theta_{n-m}\} = 1$ также равномерно относительно x в указанном отрезке. Перейдем теперь к рассмотрению величины вида:

$$\begin{aligned}
&\ln\{(nm^{-1}p)^m[n(n-m)^{-1}q]^{n-m}\} = \\
&\quad = -(np + x(npq)^{1/2})\ln[1 + xq^{1/2}(np)^{-1/2}] - \\
&\qquad - (nq - x(npq)^{1/2})\ln[1 - xp^{1/2}(nq)^{-1/2}] = \\
&\quad = -(np + x(npq)^{1/2})[xq^{1/2}(np)^{-1/2} - 2^{-1}x^2q(np)^{-1} + O(n^{-3/2})] - \\
&\qquad - (nq - x(npq)^{1/2}[-xp^{1/2}(nq)^{-1/2} - 2^{-1}x^2p(nq)^{-1} + O(n^{-3/2})] = \\
&\quad = -2^{-1}x^2 + O(n^{-1/2}).
\end{aligned}$$

Теперь окончательно находим, что переменная величина

$$\begin{aligned}
(nm^{-1}p)^m[n(n-m)^{-1}q]^{n-m} &= \exp\{-2^{-1}x^2\}\exp\{O(n^{-1/2})\} = \\
&= (1 + O(n^{-1/2}))\exp\{-2^{-1}x^2\} \qquad (4.7)
\end{aligned}$$

равномерно относительно $x \in [-c, c]$ стремится к функции $\exp\{-2^{-1}x^2\}$ при $n \to \infty$. Приведенные рассуждения доказывают локальную предельную теорему, из которой получаем первую приближенную формулу вида

$$\mathsf{P}(\boldsymbol{B}_m) \approx \frac{1}{\sqrt{npq}}\,\frac{1}{\sqrt{2\pi}}\,\exp\Big\{-\frac{1}{2}\Big(\frac{m-np}{\sqrt{npq}}\Big)^2\Big\}. \qquad (4.8)$$

Если ввести в рассмотрение функцию $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}\exp\{-2^{-1}x^2\}$, для которой имеются хорошо известные таблицы (см., например, табл. П.1 из приложения или *Большев Л. Н.*, *Смирнов С. В.* Таблицы математической статистики. — М.: Наука, 1983), то $\mathsf{P}(\boldsymbol{B}_m) \approx \approx \dfrac{1}{\sqrt{npq}}\,\varphi(x_m)$, где $x_m = \dfrac{m-np}{\sqrt{npq}}$. Из соотношений (4.5)–(4.7) непосредственно получаем оценку для приближения (4.8), а именно, при $x_m \in [-c, c]$ для вероятности $\mathsf{P}(\boldsymbol{B}_m)$ имеем

$$\begin{aligned}
\mathsf{P}(\boldsymbol{B}_m) &= \frac{1}{\sqrt{npq}}\,\varphi(x_m) + \frac{1}{\sqrt{npq}}\,\varphi(x_m)\,O\Big(\frac{1}{\sqrt{n}}\Big) = \\
&= \frac{1}{\sqrt{npq}}\,\varphi(x_m)\,\Big(1 + O\Big(\frac{1}{\sqrt{n}}\Big)\Big). \qquad (4.9)
\end{aligned}$$

Отсюда видно, что при конечных значениях n и при больших отклонениях значений вероятностей p или q от $1/2$ погрешность может оказаться очень большой. С другой стороны, абсолютная погрешность приближения биномиальных вероятностей $\mathsf{P}(\boldsymbol{B}_m)$ не превышает величины d/n, а относительная погрешность равна $dn^{-1/2}$, где $d > 0$ — некоторая постоянная. В инженерной практике локальную приближенную формулу Муавра–Лапласа часто применяют при $n > 100$ и $npq > 20$. Для сравнения точных и приближенных значений вероятностей $\mathsf{P}(\boldsymbol{B}_m)$ приведем пример.

Пример 4.10. На поверхность пола, разграфленную параллельными прямыми на расстоянии $2h$ друг от друга, непреднамеренно и независимым образом 1156 раз бросается игла длиной $h\pi/2$. Найти вероятность того, что игла пересечет какие-нибудь прямые ровно 595 раз.

Решение. Вероятность того, что игла пересечет какую-нибудь прямую в каждом опыте равна $(2^{-1}h\pi)/(h\pi) = 1/2$ (см. задачу Бюффона, пример 2.38). Итак, в примере имеем схему Бернулли с параметрами $n = 1156$, $p = 1/2$. Поэтому при $m = 595$ для вероятности $\mathsf{P}(\boldsymbol{B}_{595})$ найдем точное значение по формуле Бернулли $C_{1156}^{595}(0{,}5)^{595}(0{,}5)^{561}$, равное 0,014236. При $x = (595 - 1156/2)/\sqrt{1156/4} = 1$, $\sqrt{npq} = \sqrt{1156/4} = 17$, $\varphi(1) = 0{,}241971$ получаем приближенное значение $\mathsf{P}(\boldsymbol{B}_{595}) \approx 0{,}014234$. Абсолютная погрешность такого приближения равна 0,000002, а относительная равна 0,00014. Из этого примера видно, что при достаточно большом значении $n = 1156$ почти все вероятности $\mathsf{P}(\boldsymbol{B}_m)$ близки к нулю. Действительно, даже для максимальной вероятности $\mathsf{P}(\boldsymbol{B}_{578})$ при $m_0 = [(n+1)p] = [(1156+1)/2] = 578$, $x = (578 - 1156/2)/17 = 0$, $\varphi(0) = 0{,}398942$ получаем приближенное значение, равное 0,023467. Для сравнения укажем, что приближенное значение вероятности $\mathsf{P}(\boldsymbol{B}_{510})$ при $x = (510 - 578)/17 = -4$, $\varphi(-4) = 0{,}000134$ равно 0,000008.

Рассмотрим еще два примера на применение локальной предельной теоремы Муавра–Лапласа.

Пример 4.11. По каналу связи передано независимым образом 100 знаков. Вероятность искажения любого передаваемого знака равна $p = 0{,}2$. Какова вероятность получения 16 искажений или, другими словами, какова вероятность события $\boldsymbol{B}_{16}$?

Решение. Здесь можно считать, что параметры $n = 100$ и $m = 16$ схемы Бернулли являются достаточно большими, а значения $p = 0{,}2$ и $q = 0{,}8$ отделены от нуля и от единицы. Поэтому возможно применение локальной теоремы Муавра–Лапласа. На основании этого

$$\mathsf{P}(\boldsymbol{B}_{16}) = P_{100}(16) = C_{100}^{16}p^{16}q^{100-16},$$

$$(m - np)/(npq)^{1/2} = (16 - 100 \times 0{,}2) \times (100 \times 0{,}2 \times 0{,}8)^{-1/2} = -1,$$

$$\mathsf{P}(\boldsymbol{B}_{16}) \approx (npq)^{-1/2}\varphi((m - np)/(npq)^{1/2}) = (1/4) \times \varphi(-1) \approx 0{,}060492.$$

Пример 4.12. При производстве деталей вероятность брака равна 0,1. Какова вероятность того, что из 400 наугад взятых деталей ровно 50 будут бракованными?

Решение. В данном примере при $n = 400$, $p = 0{,}1$ и $q = 0{,}9$ необходимо вычислить вероятность $\mathsf{P}(\boldsymbol{B}_{50}) = C_{400}^{50}(0{,}1)^{50}(0{,}9)^{350}$. Непосредственный подсчет этой вероятности связан с большими техническими трудностями. Используем локальную теорему Муавра–Лапласа и соответствующую приближенную формулу, сразу получим:

$$\mathsf{P}(\boldsymbol{B}_{50}) = (400 \times 0{,}1 \times 0{,}9)^{-1/2} \times \varphi((50 - 400 \times 0{,}1)\times \\ \times (400 \times 0{,}1 \times 0{,}9)^{-1/2}) = 6^{-1}\varphi(5/3) \approx 0{,}016542.$$

Для событий $\boldsymbol{A}$ и $\boldsymbol{B}_m$, введенных в п. 2.1, выполняется следующая теорема.

Теорема 4.4 (интегральная предельная теорема Муавра–Лапласа). *Пусть величина* $x_m = x_m(n) = (m - np)/(npq)^{1/2}$ *для* $m = 0, 1, \ldots, n$ *и* $-\infty \leqslant a < b \leqslant +\infty$. *Тогда равенство*

$$\lim_{n\to\infty} \mathsf{P}\Big(\bigcup_{\{m:\, a\leqslant x_m<b\}} \boldsymbol{B}_m\Big) = \frac{1}{\sqrt{2\pi}} \int\limits_a^b \exp\{-x^2/2\}\, dx$$

выполняется равномерно по a *и* b, *где объединение распространяется на все такие значения* m, *для которых* $a \leqslant x_m < b$, *и вероятность* p $(0 < p < 1)$ *успеха в каждом из* n *независимых испытаний опыта* E_0 *постоянна.*

Доказательство. Вначале докажем теорему для фиксированных a и b из какого-либо конечного отрезка $[-c, c]$. При $x_m \in [a, b)$ согласно локальной предельной теореме, а точнее, благодаря соотношению (4.9), имеем следующее соотношение:

$$\mathsf{P}\Big(\bigcup_{\{m:\, a\leqslant x_m<b\}} \boldsymbol{B}_m\Big) = \sum_{\{m:\, a\leqslant x_m<b\}} \mathsf{P}(\boldsymbol{B}_m) = \sum_{\{m:\, a\leqslant x_m<b\}} \frac{1}{\sqrt{npq}}\varphi(x_m)\times \\ \times \Big(1 + O\Big(\frac{1}{\sqrt{n}}\Big)\Big) = \sum_{\{m:\, a\leqslant x_m<b\}} \frac{1}{\sqrt{npq}}\varphi(x_m)+ \\ + \sum_{\{m:\, a\leqslant x_m<b\}} \frac{1}{\sqrt{npq}}\varphi(x_m)\, O\Big(\frac{1}{\sqrt{n}}\Big) = S_n(a,b) + R_n(a,b),$$

где

$$S_n(a,b) = \sum_{\{m:\, a\leqslant x_m<b\}} \frac{1}{\sqrt{npq}}\varphi(x_m),$$

$$R_n(a,b) = \sum_{\{m:\, a\leqslant x_m<b\}} \frac{1}{\sqrt{npq}}\varphi(x_m)\, O\Big(\frac{1}{\sqrt{n}}\Big).$$

Пусть теперь $n \geqslant \max\{c^2pq^{-1}, c^2qp^{-1}\}$, тогда $x_0 \leqslant a < b \leqslant x_n$ для любых чисел a и b из отрезка $[-c, c]$. Если

$$\{m\colon 0 \leqslant m \leqslant n,\ x_m \in [a, b)\} = \varnothing,$$

то

$$\mathsf{P}\Big(\bigcup_{\{m\colon a \leqslant x_m < b\}} \boldsymbol{B}_m\Big) = 0.$$

При этом

$$\frac{1}{\sqrt{2\pi}} \int\limits_a^b \exp\{-x^2/2\}\, dx < \frac{1}{\sqrt{npq}},$$

так как

$$0 < \varphi(x) = (2\pi)^{-1/2} \exp\{-x^2/2\} < 1$$

и $\Delta x_m = x_{m+1} - x_m = (m + 1 - np)(npq)^{-1/2} - (m - np)(npq)^{-1/2} = (npq)^{-1/2} > b - a$ для всех $m = 0, 1, \ldots, n$. Если множество $\{m\colon 0 \leqslant m \leqslant n,\ x_m \in [a, b)\} \neq \varnothing$, то $S_n(a, b)$ можно записать в виде суммы

$$S_n(a, b) = \sum_{\{m\colon a \leqslant x_m < b\}} \varphi(x_m)\Delta x_m,$$

которая отличается не более чем на величину $(npq)^{-1/2}$ от выбранной подходящим образом интегральной суммы $S'_n(a, b)$, соответствующей интегралу

$$\frac{1}{\sqrt{2\pi}} \int\limits_a^b \exp\{-x^2/2\}\, dx.$$

Это утверждение можно доступно разъяснить с помощью рис. 4.2.

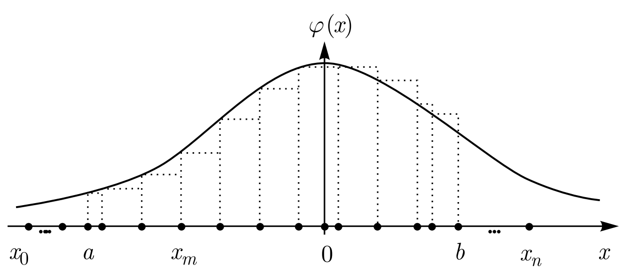


Рис. 4.2

Итак, на рис. 4.2 предполагается, что интегральная сумма $S'_n(a, b)$ однозначно выбирается абсциссами $x = a$, $x = b$, $x = x_m$, $x_m \in [a, b)$, осью $0x$, графиком функции $\varphi(x) = (2\pi)^{-1/2} \exp\{-x^2/2\}$ и, наконец, ее значениями $\varphi(a)$, $\varphi(x_m)$, $a \leqslant x_m < b$. Уточним, что при таком выборе интегральной суммы $S'_n(a, b)$ ее значение отличается не более чем двумя слагаемыми от суммы $S_n(a,b)$. Теперь $S_n(a,b)$ представим в таком виде

$$S_n(a,b) = S'_n(a,b) + O(1/\sqrt{npq}\,) = \\ = (2\pi)^{-1/2}\int\limits_a^b \exp\{-x^2/2\}\,dx + R'_n(a,b) + O(1/\sqrt{npq}\,),$$

где $R'_n(a,b)$ — остаточный член при замене интегральной суммы $S'_n(a,b)$ интегралом

$$(2\pi)^{-1/2}\int\limits_a^b \exp\{-x^2/2\}\,dx.$$

В силу известных свойств интегральных сумм для равномерно непрерывной функции $\varphi(x)$ на промежутке $[-c,c]$ легко находим, что при $n\to\infty$ член $|R'_n(a,b)|\to 0$ равномерно по a и b, где $-c\leqslant a<b\leqslant c$. Очевидно также, что

$$|R_n(a,b)| = \Big|\sum_{\{m:\ a\leqslant x_m<b\}} \frac{1}{\sqrt{npq}}\,\varphi(x_m)\,O\Big(\frac{1}{\sqrt{n}}\Big)\Big| \leqslant \\ \leqslant \sum_{\{m:\ a\leqslant x_m<b\}} \frac{1}{\sqrt{npq}}\,\varphi(x_m)\Big|\,O\Big(\frac{1}{\sqrt{n}}\Big)\Big| \leqslant \frac{d}{\sqrt{n}}\sum_{\{m:\ a\leqslant x_m<b\}}\varphi(x_m)\Delta x_m \leqslant \\ \leqslant \frac{d}{\sqrt{n}}\sum_{\{m:\ a\leqslant x_m<b\}}\Delta x_m \leqslant \frac{d}{\sqrt{n}}\Big(b-a+\frac{1}{\sqrt{n}}\Big) \leqslant \\ \leqslant \frac{d}{\sqrt{n}}\Big(2c+\frac{1}{\sqrt{npq}}\Big) \leqslant \frac{d}{\sqrt{n}}\Big(2c+\frac{1}{cp}\Big), \quad d = \mathrm{const} > 0,$$

так как $0<\varphi(x)<1$ и $n\geqslant \max\{c^2pq^{-1}, c^2qp^{-1}\}$. Значит, при $n\to\infty$ переменная величина $R_n(a,b)$ стремится к нулю равномерно по a и b из отрезка $[-c,c]$. Таким образом, утверждение теоремы доказано при любых фиксированных a и b, $-\infty<a<b<+\infty$.

Покажем теперь, что утверждение теоремы справедливо и для $-\infty\leqslant a<b\leqslant+\infty$. Прежде всего, напомним факт из математического анализа, что интеграл Эйлера–Пуассона

$$\int\limits_0^\infty \exp\{-x^2/2\}\,dx = \pi^{1/2}2^{-1/2}.$$

Отсюда легко получаем

$$(2\pi)^{-1/2}\int\limits_{-\infty}^{+\infty} \exp\{-x^2/2\}\,dx = 1.$$

В силу этого для любого $\varepsilon > 0$ можно найти столь большое строго положительное число $c = c(\varepsilon)$, что

$$\frac{1}{\sqrt{2\pi}} \int\limits_{-c}^{c} \exp\{-x^2/2\}\, dx > 1 - \varepsilon/4, \tag{4.10}$$

$$\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{-c} \exp\{-x^2/2\}\, dx = \frac{1}{\sqrt{2\pi}} \int\limits_{c}^{+\infty} \exp\{-x^2/2\}\, dx < \varepsilon/8. \tag{4.11}$$

Согласно доказанному утверждению для случая $-\infty < a < b < +\infty$ выберем столь большое число $N = N(\varepsilon)$, что при $-c \leqslant a < b \leqslant c$ и $n > N$:

$$\left| \sum_{\{m\colon\, a \leqslant x_m < b\}} \mathsf{P}(\boldsymbol{B}_m) - \frac{1}{\sqrt{2\pi}} \int\limits_{a}^{b} \exp\{-x^2/2\}\, dx \right| < \varepsilon/4. \tag{4.12}$$

По этой причине и из (4.10) получаем, что

$$\sum_{\{m\colon\, -c \leqslant x_m < c\}} \mathsf{P}(\boldsymbol{B}_m) > 1 - \varepsilon/2, \qquad \sum_{\{m\colon\, -\infty \leqslant x_m < -c\}} \mathsf{P}(\boldsymbol{B}_m) + \\ + \sum_{\{m\colon\, c \leqslant x_m < \infty\}} \mathsf{P}(\boldsymbol{B}_m) = 1 - \sum_{\{m\colon\, -c \leqslant x_m < c\}} \mathsf{P}(\boldsymbol{B}_m) < \varepsilon/2, \tag{4.13}$$

где

$$\sum_{\{m\colon\, -\infty \leqslant x_m < -c\}} \mathsf{P}(\boldsymbol{B}_m) = \lim_{s \downarrow -\infty} \sum_{\{m\colon\, s \leqslant x_m < -c\}} \mathsf{P}(\boldsymbol{B}_m),$$
$$\sum_{\{m\colon\, c \leqslant x_m < \infty\}} \mathsf{P}(\boldsymbol{B}_m) = \lim_{s \uparrow +\infty} \sum_{\{m\colon\, c \leqslant x_m < s\}} \mathsf{P}(\boldsymbol{B}_m).$$

Рассмотрим теперь один из возможных конечных случаев принадлежности чисел a и b подмножествам $(-\infty, -c)$, $[-c, c)$, $[c, \infty)$ действительной прямой, а именно, $-\infty \leqslant a \leqslant -c < c \leqslant b \leqslant +\infty$. Остальные случаи разбираются аналогичным образом. Очевидно, что для любых $-\infty \leqslant a \leqslant -c < c \leqslant b \leqslant +\infty$ легко находим:

$$\frac{1}{\sqrt{2\pi}} \int\limits_{a}^{b} \exp\{-x^2/2\}\, dx = \frac{1}{\sqrt{2\pi}} \int\limits_{a}^{-c} \exp\{-x^2/2\}\, dx + \\ + \frac{1}{\sqrt{2\pi}} \int\limits_{-c}^{c} \exp\{-x^2/2\}\, dx + \frac{1}{\sqrt{2\pi}} \int\limits_{c}^{b} \exp\{-x^2/2\}\, dx,$$

$$\sum_{\{m\colon\, a \leqslant x_m < b\}} \mathsf{P}(\boldsymbol{B}_m) = \sum_{\{m\colon\, c \leqslant x_m < -c\}} \mathsf{P}(\boldsymbol{B}_m) + \\ + \sum_{\{m\colon\, -c \leqslant x_m < c\}} \mathsf{P}(\boldsymbol{B}_m) + \sum_{\{m\colon\, c \leqslant x_m < b\}} \mathsf{P}(\boldsymbol{B}_m).$$

Учитывая это и соотношения (4.11)–(4.13), последовательно выводим, что

$$\Big|\sum_{\{m:\ a\leqslant x_m<b\}} \mathsf{P}(\boldsymbol{B}_m)-\frac{1}{\sqrt{2\pi}}\int\limits_a^b \exp\{-x^2/2\}\,dx\Big|\leqslant$$

$$\leqslant\Big|\sum_{\{m:\ a\leqslant x_m<-c\}} \mathsf{P}(\boldsymbol{B}_m)-\frac{1}{\sqrt{2\pi}}\int\limits_a^{-c} \exp\{-x^2/2\}\,dx\Big|+$$

$$+\Big|\sum_{\{m:\ -c\leqslant x_m<c\}} \mathsf{P}(\boldsymbol{B}_m)-\frac{1}{\sqrt{2\pi}}\int\limits_{-c}^{c} \exp\{-x^2/2\}\,dx\Big|+$$

$$+\Big|\sum_{\{m:\ c\leqslant x_m<b\}} \mathsf{P}(\boldsymbol{B}_m)-\frac{1}{\sqrt{2\pi}}\int\limits_c^b \exp\{-x^2/2\}\,dx\Big|\leqslant$$

$$\leqslant\sum_{\{m:\ -\infty\leqslant x_m<-c\}} \mathsf{P}(\boldsymbol{B}_m)+\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{-c} \exp\{-x^2/2\}\,dx+$$

$$+\Big|\sum_{\{m:\ -c\leqslant x_m<c\}} \mathsf{P}(\boldsymbol{B}_m)-\frac{1}{\sqrt{2\pi}}\int\limits_{-c}^{c} \exp\{-x^2/2\}\,dx\Big|+$$

$$+\sum_{\{m:\ c\leqslant x_m<\infty\}} \mathsf{P}(\boldsymbol{B}_m)+\frac{1}{\sqrt{2\pi}}\int\limits_c^{\infty} \exp\{-x^2/2\}\,dx\leqslant\frac{\varepsilon}{2}+\frac{\varepsilon}{4}+\frac{\varepsilon}{8}+\frac{\varepsilon}{8}=\varepsilon$$

для любых $-\infty\leqslant a\leqslant -c<c\leqslant b\leqslant +\infty$, и, значит, интегральная предельная теорема Муавра–Лапласа доказана.

Из интегральной теоремы Муавра–Лапласа непосредственно следует вторая приближенная формула вида

$$\mathsf{P}\Big(\bigcup_{m=k}^{s}\boldsymbol{B}_m\Big)\approx\frac{1}{\sqrt{2\pi}}\int\limits_{\frac{k-np}{\sqrt{npq}}}^{\frac{s-np}{\sqrt{npq}}}\exp\{-x^2/2\}\,dx=\frac{\sqrt{2}}{2\sqrt{\pi}}\int\limits_0^{\frac{s-np}{\sqrt{npq}}}\exp\{-x^2/2\}\,dx-$$

$$-\frac{\sqrt{2}}{2\sqrt{\pi}}\int\limits_0^{\frac{k-np}{\sqrt{npq}}}\exp\{-x^2/2\}\,dx=\frac{1}{2}\Big[\Phi\Big(\frac{s-np}{\sqrt{npq}}\Big)-\Phi\Big(\frac{k-np}{\sqrt{npq}}\Big)\Big],\quad (4.14)$$

где $\Phi(x)=\dfrac{\sqrt{2}}{\sqrt{\pi}}\int_0^x\exp\{-t^2/2\}\,dt$, $x\in(-\infty,+\infty)$, называется *функцией Лапласа*, или *интегралом ошибок*, или, наконец, *интегралом вероятностей*. График функции Лапласа представлен на рис. 4.3.

Отметим свойства функции Лапласа. Из ее определения непосредственно следует, что $\Phi(0)=0$ и $\Phi(-x)=-\Phi(x)$, т. е. функция Лапласа

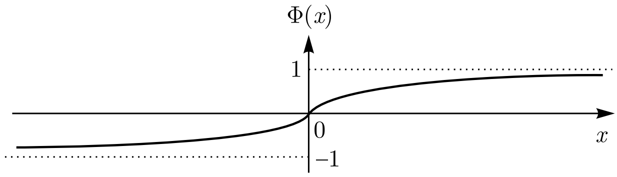


Рис. 4.3

является нечетной. Действительно, используя замену $t = -u$, $dt = -du$, легко найдем

$$\Phi(-x) = \frac{\sqrt{2}}{\sqrt{\pi}} \int\limits_0^{-x} \exp\{-t^2/2\}\,dt = -\frac{\sqrt{2}}{\sqrt{\pi}} \int\limits_0^{x} \exp\{-u^2/2\}\,du = -\Phi(x).$$

Далее, при увеличении аргумента x от нуля до $+\infty$ функция Лапласа $\Phi(x)$ возрастает от нуля до единицы, так как интеграл Эйлера–Пуассона

$$\int\limits_0^{\infty} \exp\{-t^2/2\}\,dt$$

равен $\pi^{1/2}2^{-1/2}$. В силу нечетности функции Лапласа имеем $\Phi(-x) = -\Phi(x)$, поэтому

$$\Phi(-\infty) = \lim_{x\to\infty} \Phi(-x) = -\lim_{x\to\infty} \Phi(x) = -\Phi(+\infty) = -1.$$

Отсюда выводим, что при уменьшении аргумента x от нуля до $-\infty$ функция Лапласа $\Phi(x)$ убывает от нуля до числа $\Phi(-\infty) = -\Phi(+\infty) = -1$. График функции $\Phi(x)$ имеет две асимптотические прямые $y = +1$ и $y = -1$. Для функции $\Phi(x)$, как и для функции $\varphi(x) = (2\pi)^{-1/2}\exp\{-x^2/2\}$, составлены подробные таблицы (см., например, табл. П.2 приложения или работы [2, 3, 6, 16, 22]). Для нечетной функции $\Phi(x)$ таблицы приведены только при $x > 0$. Так как

$$(2\pi)^{-1/2} \int\limits_{-\infty}^{x} \exp\{-t^2/2\}\,dt = (2\pi)^{-1/2} \int\limits_{-\infty}^{0} \exp\{-t^2/2\}\,dt + (2\pi)^{-1/2} \int\limits_0^{x} \exp\{-t^2/2\}\,dt = 2^{-1}(1 + \Phi(x)),$$

то иногда в статистических таблицах приводятся не значения функции Лапласа $\Phi(x)$, а значения одного из следующих интегралов

$$(2\pi)^{-1/2} \int\limits_{-\infty}^{x} \exp\{-t^2/2\}\, dt,$$

$$(2\pi)^{-1/2} \int\limits_{x}^{\infty} \exp\{-t^2/2\}\, dt = 2^{-1}(1 - \Phi(x))$$

или даже функции $2^{-1}\Phi(x)$. При использовании таблиц в формуле (4.14), возможно, придется произвести пересчет, если в конкретной таблице даны значения не функции Лапласа. Вычисления по формуле (4.14), как и по формуле (4.8), дают вполне хорошие приближения в случае, когда n достаточно велико, а p не очень близко́ к нулю или единице. Рассмотрим примеры использования интегральной предельной теоремы Муавра–Лапласа.

Пример 4.13. При передаче сообщения вероятность искажения каждого знака равна $p = 0{,}005$. Передано $n = 10000$ знаков. Найти вероятность того, что искажено не более 70 знаков.

Решение. По условию задачи имеем $p = 0{,}005$, $q = 0{,}995$, $n = 10\,000$. Параметры p и q «плохо» отделены от нуля и от единицы. Однако значение n, равное $10\,000$, достаточно велико. Поэтому можно воспользоваться приближенной интегральной формулой Муавра–Лапласа. В результате ее применения получим

$$\begin{aligned}\mathsf{P}\Big(\bigcup_{m=0}^{70} \boldsymbol{B}_m\Big) &= \sum_{m=0}^{70} \mathsf{P}(\boldsymbol{B}_m) \approx \\ &\approx 2^{-1}\Phi((70 - 10000 \times 0{,}005)/\sqrt{10000 \times 0{,}005 \times 0{,}995}\,) - \\ &\quad - 2^{-1}\Phi((0 - 10000 \times 0{,}005)/\sqrt{10000 \times 0{,}005 \times 0{,}995}\,) \approx \\ &\approx 2^{-1}\Phi(20/\sqrt{49{,}75}\,) + 2^{-1}\Phi(50/\sqrt{49{,}75}\,) \approx \\ &\approx 2^{-1}[\Phi(2{,}84) + \Phi(7)] \approx (1/2) \times [\,0{,}9949 + 1] \approx 0{,}997.\end{aligned}$$

Пример 4.14. Найти вероятность того, что при $n = 400$ бросаниях симметричной монеты герб появится от $m_1 = 180$ до $m_2 = 210$ раз.

Решение. Из условия эксперимента имеем $n = 400$, $m_1 = 180$, $m_2 = 210$, $p = q = 1/2$, $np = 400 \times 0{,}5 = 200$, $\sqrt{npq} = 10$. В этом эксперименте имеем идеальные условия для применения приближенной интегральной формулы, так $p = q = 0{,}5$. В результате несложных вычислений с помощью приближенной интегральной формулы легко найдем

$$\mathsf{P}\left(\bigcup_{m=180}^{210}\boldsymbol{B}_m\right) = \sum_{m=180}^{210}\mathsf{P}(\boldsymbol{B}_m) \approx$$
$$\approx 2^{-1}[\Phi((210-200)/(10) - \Phi((180-200)/10)] =$$
$$= 2^{-1}\Phi(1) + 2^{-1}\Phi(2) \approx 2^{-1}[0{,}6827 + 0{,}9545] \approx 0{,}8186.$$

Пример 4.15. Завод отправил на базу 100 доброкачественных изделий. Вероятность повреждения в пути для каждого изделия равна 0,2.

Найти вероятность того, что неповрежденными на базу прибудут: 1) не менее 75 и не более 90 изделий; 2) не менее 75 изделий.

Решение. По условию задачи вероятность того, что изделие не повредится в пути, равна $p = 0{,}8$. Всего было отправлено $n = 100$ доброкачественных изделий. Для решения задачи 1) необходимо вычислить вероятность $\mathsf{P}\big(\cup_{m=75}^{90}\boldsymbol{B}_m\big)$. Применяя интегральную теорему Муавра–Лапласа, получим,

$$\mathsf{P}\left(\bigcup_{m=75}^{90}\boldsymbol{B}_m\right) \approx 0{,}5\,[\Phi(x_2) - \Phi(x_1)],$$

где $x_1 = \dfrac{75 - 100\times 0{,}8}{\sqrt{100\times 0{,}8\times 0{,}2}} = -1{,}25$ и $x_2 = \dfrac{90 - 100\times 0{,}8}{\sqrt{100\times 0{,}8\times 0{,}2}} = 2{,}5$. Учитывая равенство $\Phi(-1{,}25) = -\Phi(1{,}25)$, найдем

$$\mathsf{P}\left(\bigcup_{m=75}^{90}\boldsymbol{B}_m\right) \approx 2^{-1}[\Phi(2{,}5) + \Phi(1,25)] \approx 0{,}888.$$

Аналогично получаем ответ для случая 2), а именно:

$$\mathsf{P}\left(\bigcup_{m=75}^{100}\boldsymbol{B}_m\right) \approx 2^{-1}\Phi\left(\frac{100 - 100\times 0{,}8}{\sqrt{100\times 0{,}8\times 0{,}2}}\right) - 2^{-1}\Phi\left(\frac{75 - 100\times 0{,}8}{\sqrt{100\times 0{,}8\times 0{,}2}}\right) =$$
$$= 2^{-1}[\Phi(5) + \Phi(1,25)] \approx 0{,}894.$$

Пример 4.16. На хоккейный матч продается 5000 билетов. Предполагается, что болельщики приходят на стадион парами и каждая пара независимо от других с равной вероятностью выбирает один из двух однотипных гардеробов дворца спорта. Сколько мест должен иметь каждый гардероб, чтобы с вероятностью 0,95 болельщики могли раздеться в выбранном ими гардеробе?

Решение. Из условия задачи непосредственно следует, что имеем дело со схемой независимых испытаний Бернулли, для которой число испытаний n равно числу пар, т. е. $n = 5000 : 2 = 2500$. Ради определенности допустим, что с вероятностью p каждая пара выбирает первый гардероб и с вероятностью q — второй гардероб. Согласно условию задачи, $p = q = 0{,}5$. Если обозначим через y число мест в каждом гардеробе и через m число пар, которые выбрали первый гардероб, то величина $(2\,500 - m)$ определяет число пар, которые выбрали второй гардероб. Очевидно, что при $m \leqslant 2^{-1}y$ и одновременно при $2\,500 - m \leqslant 2^{-1}y$ болельщики могут раздеться в выбранном ими гардеробе. Отсюда получим ограничение на число пар, которые выбрали первый гардероб в виде следующих неравенств: $2500 - 2^{-1}y \leqslant m \leqslant 2^{-1}y$. Используя интегральную теорему Муавра–Лапласа, условие $0{,}95 = \sum_{m=2\,500-y/2}^{y/2} P_n(m)$ и значение $np = 1250$, можно записать следующие равенства:

$$0{,}95 = \sum_{m=2\,500-y/2}^{y/2} P_n(m) \approx 2^{-1}\left(\Phi\left(\frac{y/2 - 1\,250}{25}\right) - \Phi\left(\frac{1\,250 - y/2}{25}\right)\right) = \\ = \Phi\left(\frac{y/2 - 1\,250}{25}\right),$$

или $(y/2 - 1\,250)/25 = 1{,}96$. Отсюда, каждый гардероб должен иметь $y = 2\,598$ мест, чтобы с вероятностью 0,95 болельщики могли раздеться в выбранном ими гардеробе.

3.2. Предельные теоремы Муавра–Лапласа и понятие статистической устойчивости эксперимента. Фундаментальное значение предельных теорем Муавра–Лапласа в обосновании понятия статистической устойчивости эксперимента поясним на решении ряда примеров.

Пример 4.17. При сборке микросхем на кристаллах применяется около 19 технологических операций. Статистические данные производства показывают, что процент выхода годных интегральных схем после каждой технологической операции достаточно высок. Однако после выполнения всех технологических операций число годных электронных изделий составляет около 50 %. Для контроля готовой продукции обследуются 22 500 интегральных схем серии КМ155 в корпусе ТО-116 из общего объема 5 876 тысяч штук, выпускаемого некоторым предприятием за год.

Найти вероятность того, что среди 22 500 наудачу отобранных для контроля микросхем число годных будет: 1) в точности 11250, 2) не менее $k = 11\,100$ и не более $s = 11\,400$.

Решение. Рассмотрим сначала решение первого пункта. В данном примере $n = 22\,500 > 100$, $p = 1/2$, $q = 1/2$, $npq = 22\,500/4 = 5\,625 > 20$, $(npq)^{1/2} = 75$, $m = 11\,250$. Воспользовавшись приближенной формулой (4.8), в которой полагаем $x_m = (m - np)/(npq)^{1/2} = (11\,250 - 11\,250)/75 = 0$, получим соотношение $\mathsf{P}(\boldsymbol{B}_{11250}) \approx \varphi(0)/75 = 0{,}398942/75 = 0{,}0053$. Эта вероятность близка к нулю. Этот результат для данного и многих аналогичных примеров приводит к следующим важным рассуждениям.

Рассмотрим, ради простоты, случай, когда число отобранных для контроля интегральных схем является четным. В этом предположении вероятность $\mathsf{P}(\boldsymbol{B}_{n/2})$ того, что число годных микросхем будет в точности равно числу бракованных, стремится к нулю при увеличении n, так как имеет место приближенное равенство $\mathsf{P}(\boldsymbol{B}_{n/2}) \approx \varphi(0)/(npq)^{1/2}$. В этом примере статистически устойчивый эксперимент E_0 состоит в том, что происходит обследование микросхемы на предмет ее годности (событие A) или брака (событие $\overline{A}$). Отсюда при n-кратном повторении причинно-независимым образом этого опыта вероятность $\mathsf{P}(\boldsymbol{B}_{n/2})$ совпадает с вероятностью того, что относительная частота $\mu(A,n)/n$ в точности равна $p = 1/2$. Следовательно, вероятность того, что относительная частота $\mu(A,n)/n$ в точности равна $p = 1/2$, также стремится к нулю с ростом n. Отсюда, на первый взгляд, получается некоторое противоречие с интуитивным понятием статистической устойчивости опыта E_0. Это недоразумение можно объяснить следующим образом. Пусть $m_0 = [(n+1)p]$ определяет наивероятнейшее число появлений события A при n-кратном повторении опыта E_0. Тогда при достаточно большом n число $n/2$ будет близко к числу $m_0 = [(n+1)/2]$. Далее, теоретический результат заключается в том, что вероятность $\mathsf{P}(\boldsymbol{B}_{m_0}) = \max\{\mathsf{P}(\boldsymbol{B}_0), \mathsf{P}(\boldsymbol{B}_1), \ldots, \mathsf{P}(\boldsymbol{B}_n)\}$. Эти несложные рассуждения позволяют написать соотношение вида $\mathsf{P}(\boldsymbol{B}_{n/2}) \approx \max\{\mathsf{P}(\boldsymbol{B}_0), \mathsf{P}(\boldsymbol{B}_1), \ldots, \mathsf{P}(\boldsymbol{B}_n)\}$.

Итак, вероятнее всего абсолютная частота $\mu(A,n)$ будет принимать значения около числа $n/2$ или относительная частота $\mu(A,n)/n$ будет колебаться вблизи вероятности $\mathsf{P}(A) = 1/2$. Получили первый фундаментальный факт, который соединяет наше интуитивное представление о статистически устойчивом эксперименте с теоретическими результатами изучения вероятностной модели. Переходим к решению второго пункта.

В пункте 2) нас интересует вероятность того, что среди 22 500 наудачу отобранных для контроля микросхем будем иметь число годных деталей от 11 100 до 11 400 включительно. Применяя приближенную интегральную формулу Муавра–Лапласа (4.14), в которой $k = 11\,100$, $s = 11\,400$, $(k-np)/(npq)^{1/2} = (11\,100 - 11\,250)/\sqrt{22\,500/4} = -2$, $(s-np)/(npq)^{1/2} = (11\,400 - 11\,250)/\sqrt{22\,500/4} = 2$, с использованием таблицы для функции $\Phi(x)$ находим

$$\mathsf{P}\left(\bigcup_{m=11\,100}^{11\,400} \boldsymbol{B}_m\right) \approx 2^{-1}[\Phi(2) - \Phi(-2)] = \Phi(2) \approx 0{,}954499.$$

Событие $\cup_{m=11\,100}^{11\,400}\boldsymbol{B}_m$ означает, что абсолютная частота $\mu(A,n)$ при $n = 22\,500$ примет значение из отрезка $[11\,100, 11\,400]$ или относительная частота $\mu(A,n)/n$ примет значение из отрезка

$$\left[\frac{11100-11250}{22500}+\frac{1}{2}, \frac{11400-11250}{22500}+\frac{1}{2}\right] = \left[-\frac{3}{450}+\frac{1}{2}, \frac{3}{450}+\frac{1}{2}\right].$$

Значит, относительная частота $\mu(A,n)/n$ получения годной микросхемы при $n = 22\,500$ отклонится от $\mathsf{P}(A) = 1/2$ не более чем на $\varepsilon = 3/450 \approx 0{,}0067$ с вероятностью 0,9545, близкой к единице.

При вычислении вероятности $\mathsf{P}\left(\cup_{m=11\,100}^{11\,400}\boldsymbol{B}_m\right)$ по приближенной формуле (4.14) мы естественно допустили некоторую ошибку. В настоящее время предлагаются различные более точные приближенные формулы, некоторые из них основаны на незначительном изменении аргумента функции (x) в соотношении (4.14). Например (см. [17]), можно предложить более точную приближенную формулу вида

$$\mathsf{P}\left(\bigcup_{m=k}^{s}\boldsymbol{B}_m\right) \approx \frac{1}{2}\left[\Phi\left(\frac{s+0{,}5-np}{\sqrt{npq}}\right) - \Phi\left(\frac{k-0{,}5-np}{\sqrt{npq}}\right)\right], \tag{4.15}$$

которая получается, если $S_n(a,b) = S_n\left((k-np)/\sqrt{npq}\,, (s-np)/\sqrt{npq}\,\right)$ при доказательстве интегральной теоремы Муавра–Лапласа заменить более точной интегральной суммой $S_n'(a_1,b_1)$, где число

$$a_1 = a - 0{,}5/\sqrt{npq} = (k-0{,}5-np)/\sqrt{npq}\,,$$
$$b_1 = b + 0{,}5/\sqrt{npq} = (s+0{,}5-np)/\sqrt{npq}\,.$$

Если формула (4.14) при $n \geqslant 1000$ вычисляет вероятность

$$\mathsf{P}\left(\bigcup_{m=11\,100}^{11\,400}\boldsymbol{B}_m\right),$$

как правило, с точностью до двух знаков после запятой, то формула (4.15) обеспечивает значения этой вероятности с точностью до четырех знаков. Для подтверждения этого факта рассмотрим пример 4.17 с несколько измененными данными, а именно, пусть для контроля отобрано $n = 1000$ микросхем и в пункте 2) предполагаем, что $k = 470$, $s = 530$. В этом случае точное значение для вероятности

$$\mathsf{P}\left(\bigcup_{m=470}^{530}\boldsymbol{B}_m\right) = \sum_{m=470}^{530} C_n^m p^m q^{n-m}$$

равно 0,9463, приближенная формула (4.14) дает значение 0,9422 и приближенная формула (4.15) — значение 0,9463. Другими словами, с достаточно большой вероятностью 0,9463 относительная частота

$\mu(A,n)/n$ появления события A при $n = 1000$ испытаниях опыта E_0 примет значения из отрезка

$$[2^{-1} + (470 - 500)/1000, 2^{-1} + (530 - 500)/1000] = \\ = [-0{,}03 + 1/2,\ 0{,}03 + 1/2].$$

Замечание 4.1. При ответах на вопрос 2) примера 4.17 было установлено, что относительная частота появления события A при $n = 1000$ и $n = 22\,500$ испытаниях опыта E_0 будет незначительно отклонятся от $\mathsf{P}(A) = 1/2$ с вероятностью, достаточно близкой к единице. Это утверждение справедливо в общем случае и непосредственно следует из интегральной предельной теоремы Муавра–Лапласа. Действительно, для любого постоянного $\varepsilon > 0$ последовательно имеем

$$1 = \lim_{n\to\infty} \frac{1}{\sqrt{2\pi}} \int\limits_{-\varepsilon\sqrt{n/pq}}^{\varepsilon\sqrt{n/pq}} \exp\{-x^2/2\}\,dx = \\ = \lim_{n\to\infty} \left(\mathsf{P}\left(\bigcup_{a_n \leqslant x_m \leqslant b_n} \boldsymbol{B}_m\right) + O(1/(npq)^{1/2})\right) = \lim_{n\to\infty} \mathsf{P}\left(\bigcup_{m=k}^{s} \boldsymbol{B}_m\right),$$

где $a_n = -\varepsilon\, n^{1/2}(pq)^{-1/2}$, $b_n = \varepsilon\, n^{1/2}(pq)^{-1/2}$. Здесь символ $k =]np - n\varepsilon[$ есть такое наименьшее целое число, что $np - n\varepsilon \leqslant]np - n\varepsilon[$, и символ $s = [np + n\varepsilon]$ есть такое наибольшее целое число, что $[np + n\varepsilon] \leqslant \leqslant np + n\varepsilon$. Событие $\cup_{m=k}^{s} \boldsymbol{B}_m$ означает, что абсолютная частота $\mu(A,n)$ примет значение из отрезка вида $[np - n\varepsilon, np + n\varepsilon]$ или относительная частота $\mu(A,n)/n$ появления события A в схеме независимых испытаний Бернулли отклонится от вероятности $\mathsf{P}(A) = p$ не более, чем на величину ε. Отсюда получаем, что вероятность колебания с точностью $\varepsilon > 0$ относительной частоты $\mu(A,n)/n$ около $\mathsf{P}(A) = p$ стремится к единице при увеличении n до бесконечности. Таким образом, из предельных теорем Муавра–Лапласа получили второй фундаментальный факт. Этот факт прокладывает мост между удивительным свойством статистической устойчивости реальных случайных экспериментов и аналогичным свойством вероятностных моделей.

3.3. Предельные теоремы Пуассона. Нередко встречаются реальные эксперименты (см. пример 4.9), в которых рассматривается большое число n независимых испытании. При этом вероятность p наступления успешного события A при каждом отдельном испытании опыта E_0 мала. В этом случае исход A опыта E_0 является редким событием, а приближенные формулы (4.8), (4.14) и (4.15) дают значительные ошибки при вычислении вероятностей $\mathsf{P}(\boldsymbol{B}_m)$, $\mathsf{P}\big(\cup_{m=k}^{s} \boldsymbol{B}_m\big)$ даже при больших значениях n. Последнее утверждение непосредственно следует из доказательств предельных теорем Муавра–Лапласа. Итак, возникает задача отыскания для вероятностей $\mathsf{P}(\boldsymbol{B}_m)$, $\mathsf{P}\big(\cup_{m=k}^{s}\boldsymbol{B}_m\big)$ лучшую аппроксимацию при малых значениях p или q. Если вероятность p достаточно мала, то математическая формализация этого факта означает, что величина p становится переменной, например, зависящей от n, т. е. $p = p(n)$, $\lim\limits_{n\to\infty} p(n) = 0$. Однако при получении точной формулы $\mathsf{P}(\boldsymbol{B}_m) = C_n^m p^m q^{n-m}$ вероятность $\mathsf{P}(A) = p$ наступления успешного исхода A должна быть постоянной. Ситуации, когда вероятность успешного исхода A стремится к нулю, добиться нельзя в рамках классической схемы независимых испытаний Бернулли. В силу этого приходится рассматривать несколько другую конструкцию причинно-независимых испытаний с использованием последовательности эволюционных экспериментов.

Итак, рассмотрим последовательность E_1, $E_2, \ldots$ из причинно-независимых эволюционных экспериментов $E_n = \{E_{n,1}, E_{n,2}, \ldots, E_{n,n}\}$, где $n = 1, 2, \ldots$. Будем предполагать, что каждый эволюционный эксперимент E_n в свою очередь составлен из n опытов $E_{n,1}, E_{n,2}, \ldots, E_{n,n}$ и, более того, представляет собой серию из n независимых испытаний Бернулли над одним и тем же опытом $E_{n,0}$. При этом вероятность $p = p(n)$ наступления успешного исхода для опыта $E_{n,0}$ зависит от номера n эволюционного эксперимента E_n и естественно остается постоянной в пределах каждой серии, как это требуется в схеме независимых испытаний Бернулли. В дальнейшем из физических соображений будем полагать, что вероятность $\mathsf{P}(\boldsymbol{B}_m) = P_n(m) = 0$ при $m > n$. Тогда справедлива предельная теорема.

Теорема 4.5 (локальная предельная теорема Пуассона). *Пусть при* $n \to +\infty$ *величина* $p(n) = \lambda n^{-1} + o(n^{-1})$, *где* $\lambda = \mathrm{const} > 0$ и $o(n^{-1})$ *есть бесконечно малая величина по сравнению с* n^{-1}. *Тогда для каждого фиксированного* $m = 0, 1, 2, \ldots$ *имеет место предельное равенство*:
$$\lim_{n\to\infty} P_n(m) = (m!)^{-1}\lambda^m \exp\{-\lambda\}.$$

Доказательство. По условию теоремы имеем $p(n) = \lambda n^{-1} + o(n^{-1})$. Отсюда для любого фиксированного $m = 0, 1, 2, \ldots$ находим
$$\lim_{n\to\infty} P_n(m) = \lim_{n\to\infty} C_n^m p^m q^{n-m} =$$
$$= \lim_{n\to\infty}\Big\{\frac{n\cdot(n-1)\times\ldots\times(n-m+1)}{m!}\,(\lambda n^{-1} + o(n^{-1}))^m \times$$
$$\times\,(1 - \lambda n^{-1} + o(n^{-1}))^{n-m}\Big\} =$$

$$= \lambda^m (m!)^{-1} \lim_{n\to\infty} \{n \cdot (n-1) \times$$
$$\times \ldots \times (n-m+1) n^{-m} (1 + n\cdot\lambda^{-1}\cdot o(n^{-1}))^m\} \times$$
$$\times \lim_{n\to\infty} \{(1-\lambda n^{-1} + o(n^{-1}))^n\} \times \lim_{n\to\infty} \{(1-\lambda n^{-1} + o(n^{-1}))^{-m}\} =$$
$$= \frac{\lambda^m}{m!} \exp\{-\lambda\}.$$

Эти равенства и доказывают локальную предельную теорему Пуассона.

Теорема 4.6 (интегральная предельная теорема Пуассона). *Пусть при* $n \to +\infty$ *величина* $p(n) \to 0$ *таким образом, что* $np(n) \to \to \lambda = \mathrm{const} > 0$. *Тогда для заданных* $k, s = 0, 1, \ldots$ *имеем*

$$\lim_{n\to\infty} \mathsf{P}\left(\bigcup_{m=k}^{s} \boldsymbol{B}_m\right) = \sum_{m=k}^{s} \frac{\lambda^m}{m!} \exp\{-\lambda\}.$$

Это равенство сразу следует из локальной предельной теоремы Пуассона.

Из предельных теорем Пуассона при больших n, малых p и при $\lambda = np$ непосредственно вытекают следующие приближенные формулы:

$$\mathsf{P}(\boldsymbol{B}_m) = C_n^m p^m q^{n-m} \approx (\lambda^m/m!) \exp\{-\lambda\}, \tag{4.16}$$

$$\mathsf{P}\left(\bigcup_{m=k}^{s} \boldsymbol{B}_m\right) = \sum_{m=k}^{s} C_n^m p^m q^{n-m} \approx \sum_{m=k}^{s} (\lambda^m/m!) \exp\{-\lambda\}. \tag{4.17}$$

Легко показать, что $\sum_{m=0}^{\infty} (\lambda^m/m!) \exp\{-\lambda\} = 1$. Формулы (4.16) и (4.17) применяют для подсчета биномиальных вероятностей, которые связаны с так называемыми редкими событиями. В качестве примеров редких явлений можно указать такие, как достижение столетнего возраста некоторым человеком, рождение близнецов у некоторой женщины среднего возраста, автомобильная авария на некотором перекрестке за сутки, сбой в работе персонального компьютера за некоторый фиксированный промежуток времени, падение метеорита крупного размера на некоторый участок поверхности Земли за фиксированный промежуток времении т. п.

В дальнейшем вероятность $\mathsf{P}(\boldsymbol{B}_m) = C_n^m p^m q^{n-m}$, рассматриваемую как функцию от трех переменных n, m и p, будем иногда обозначать через $Q(n, m, p)$, а пуассоновскую функцию $(\lambda^m/m!) \exp\{-\lambda\}$ от двух переменных λ и m — через $\pi(m; \lambda)$. Численные значения функций $\pi(m; \lambda)$ и $\sum_{m=s}^{\infty} \pi(m; \lambda)$ берутся или из таблиц П.3, П.4 приложения или из известных учебников по теории вероятностей [6, 13, 14, 19]. В настоящее время существуют простые программы получения значений функции $\pi(m; \lambda)$ на компьютере.

При использовании теорем Пуассона следует обратить внимание на удивительную особенность, которая состоит в том, что функция

$Q(n, m, p)$ зависит от трех переменных, а функция $\pi(m; \lambda)$ — только от двух переменных. Следовательно, для вычисления правых частей формул (4.16) и (4.17) вовсе не требуется знание n и p, если задан параметр λ. Это обстоятельство объясняется следующим образом. В реальной задаче, когда имеется в виду схема из n причинно-независимых испытаний Бернулли, мы наблюдаем не только число m успешных исходов эксперимента E_0, но и число $n-m$ неудачных его исходов. В предельной схеме Бернулли параметры n, p становятся неопределенными ($n \to \infty$, $p \to 0$). В этом случае вероятность наступления m успешных исходов может быть задана более простым выражением $\pi(m; \lambda)$, зависящим от двух параметров m и λ. При этом число неудачных исходов не определяется. Приведем пример, поясняющий эту ситуацию.

Рассмотрим дорожно-транспортные происшествия на улице крупного города в течение суток. Вероятность $p = \mathsf{P}(A)$ того, что с некоторым человеком произойдет дорожно-транспортное происшествие, очень мала. Предполагается, что люди на улицах крупного города подвергаются несчастным дорожно-транспортным случаям независимо один от другого. Представим себе, что в крупном городе в течение суток на улицах может побывать n пешеходов. Тогда вероятность того, что среди n пешеходов ровно m пострадает от дорожно-транспортных происшествий, равна $C_n^m p^m q^{n-m}$. Поскольку число n пешеходов достаточно велико, вероятность $p = p(n)$ очень мала и np устойчиво меняется около фиксированной величины λ, то по локальной теореме Пуассона имеем $C_n^m p^m q^{n-m} \approx \pi(m; \lambda)$. Здесь параметр λ есть ожидаемое число дорожно-транспортных происшествий с пешеходами за сутки в большом городе. В таком городе не только постоянно проживает большое число людей, но имеется также значительное и неопределенное число туристов, командировочных, гостей и т. п. Поэтому специальными службами регистрируется только число m пострадавших пешеходов и не фиксируется количество n всех пешеходов на улицах, которое становится слишком большим и одновременно неопределенным. Таким образом, в реальной задаче о несчастных случаях с пешеходами специальные службы имеют дело как бы уже с предельной схемой Бернулли.

Приведем решение некоторых простых задач с применением теорем Пуассона.

Пример 4.18. Какова вероятность того, что в обществе из 500 человек ровно 5 человек родились в день нового года?

Решение. Предполагается, что для произвольного человека все его дни рождения равновероятны, поэтому вероятность $p = 1/365 \approx 0{,}0027$ мала. Из условия задачи имеем: $m = 5$, $\lambda = np = 500/365 \approx 1{,}37$. Тогда $P_{500}(5) \approx \pi(5; 1{,}37) \approx 0{,}0102$. Отметим, что точный подсчет вероятности $P_{500}(5)$ с указанием только четырех десятичных знаков дает результат 0,0101.

Пример 4.19. Фирма отправила на сборку компьютеров 5 000 доброкачественных узлов. Вероятность того, что в пути узел повредится, очень мала и равна 0,0002. Найти вероятность того, что на сборку компьютеров поступит ровно три негодных узла.

Решение. По условию задачи количество узлов равно $n = 5\,000$, вероятность повреждения каждого узла при транспортировке равна $p = 0{,}0002$. При этом параметр λ равен $np = 5\,000 \times 0{,}0002 = 1$. Тогда, используя локальную предельную теорему Пуассона, можно сосчитать искомую вероятность $\mathsf{P}(\boldsymbol{B}_3)$: $\mathsf{P}(\boldsymbol{B}_3) \approx \exp\{-1\}/3! = \pi(3;1) \approx 0{,}0613$.

Пример 4.20. Телефонная станция обслуживает 400 абонентов. Вероятность позвонить на станцию для каждого абонента равна 0,01. Найти вероятность того, что: 1) в течение часа 5 абонентов позвонят на станцию; 2) в течение часа не более 4 абонентов позвонят на станцию.

Решение. Так как $p = 0{,}01$ мало и $n = 400$ велико, то для решения этой задачи будем использовать локальную и интегральную приближенные формулы Пуассона при $\lambda = np = 400 \times 0{,}01 = 4$. В результате простых вычислений получим

1) $\mathsf{P}(\boldsymbol{B}_5) = P_{400}(5) \approx \pi(5;4) \approx 0{,}1563$;

2) $$\mathsf{P}\Bigl(\bigcup_{m=0}^{4} \boldsymbol{B}_m\Bigr) = \sum_{m=0}^{4} C_{400}^{m}(0{,}01)^m(0{,}99)^{400-m} \approx$$
$$\approx \sum_{m=0}^{4} 4^m (m!)^{-1}\exp\{-4\} = 1 - \sum_{m=5}^{\infty} \pi(m;4) \approx 1 - 0{,}37116 \approx 0{,}6288.$$

3.4. Оценка погрешности при вычислениях биномиальных вероятностей. Важный вопрос об оценке совершаемой погрешности при замене вероятности $Q(n,m,p)$ на асимптотическую функцию $\pi(m;\lambda)$ решается с помощью следующей теоремы.

Теорема 4.7. *Пусть M является любым подмножеством числового множества $L = \{0, 1, 2, \ldots\}$. Тогда имеет место неравенство*

$$\Bigl|\sum_{m\in M} Q(n,m,p) - \sum_{m\in M} \pi(m;\lambda)\Bigr| \leqslant np^2,$$

в частности, если множество M состоит из одного числа m, то

$$\Bigl|C_n^m p^m q^{n-m} - \frac{\lambda^m}{m!}\exp\{-\lambda\}\Bigr| \leqslant np^2.$$

Доказательство. Из определений функций $Q(n,m,p)$, $\pi(m;\lambda)$ непосредственно следует, что $Q(n,m,p)=0$ при $m>n$, $\pi(m;\lambda)>0$ для $m\in L$ и $\sum_{m\in L} Q(n,m,p)=1$,

$$\sum_{m=0}^{\infty}\pi(m;\lambda)=\exp\{-\lambda\}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!}=\exp\{-\lambda\}\cdot\exp\{\lambda\}=1.$$

Заметим, что ряд $\sum_{m\in L} Q(n,m,p)$ лишь формально является бесконечным, так как при $m>n$ члены ряда равны нулю. Введем в рассмотрение множества следующего вида:

$$L_+=\{m\colon m\in L, Q(n,m,p)-\pi(m;\lambda)>0\},$$
$$L_-=\{m\colon m\in L, Q(n,m,p)-\pi(m;\lambda)\leqslant 0\}.$$

Отсюда сразу имеем $L_+\cap L_-=\varnothing$ и $L_+\cup L_-=L$, т.е. множество L всех неотрицательных целых чисел разбили на подмножества L_+, L_-. Используя равенства

$$\sum_{m\in L} Q(n,m,p)=\sum_{m\in L}\pi(m;\lambda)=1,$$

получим, что

$$\sum_{m\in L} Q(n,m,p)-\sum_{m\in L}\pi(m;\lambda)=\sum_{m\in L}(Q(n,m,p)-\pi(m;\lambda))=$$
$$=\sum_{m\in L_+}(Q(n,m,p)-\pi(m;\lambda))+\sum_{m\in L_-}(Q(n,m,p)-\pi(m;\lambda))=0,$$

$$\sum_{m\in L_+}(Q(n,m,p)-\pi(m;\lambda))=-\sum_{m\in L_-}(Q(n,m,p)-\pi(m;\lambda))=\Sigma_n, \tag{4.18}$$

откуда последовательно находим, что

$$\begin{aligned}&\sum_{m\in L}|Q(n,m,p)-\pi(m;\lambda)|=\\&=\sum_{m\in L_+}|Q(n,m,p)-\pi(m;\lambda)|+\sum_{m\in L_-}|Q(n,m,p)-\pi(m;\lambda)|=\\&=\sum_{m\in L_+}(Q(n,m,p)-\pi(m;\lambda))-\sum_{m\in L_-}(Q(n,m,p)-\pi(m;\lambda))=\\&=2\sum_{m\in L_+}(Q(n,m,p)-\pi(m;\lambda))=2\Sigma_n,\end{aligned}$$

или, в конце концов,

$$\Sigma_n=2^{-1}\sum_{m\in L}|Q(n,m,p)-\pi(m;\lambda)|. \tag{4.19}$$

Далее, проделывая очевидные преобразования и используя (4.18), получим соотношение

$$\Big|\sum_{m\in M} Q(n,m,p) - \sum_{m\in M} \pi(m;\lambda)\Big| = \Big|\sum_{m\in M} (Q(n,m,p) - \pi(m;\lambda))\Big| =$$
$$= \Big|\sum_{m\in M\cap L_+} (Q(n,m,p) - \pi(m;\lambda)) + \sum_{m\in M\cap L_-} (Q(n,m,p) - \pi(m;\lambda))\Big| \leqslant$$
$$\leqslant \max\Big\{\sum_{m\in M\cap L_+} (Q(n,m,p) - \pi(m;\lambda)),$$
$$\Big|\sum_{m\in M\cap L_-} (Q(n,m,p) - \pi(m;\lambda))\Big|\Big\} \leqslant \Sigma_n$$

для любого подмножества $M \subset L$. Теперь для доказательства теоремы достаточно, например, методом полной математической индукции установить, что $\Sigma_n \leqslant np^2$. Действительно, при $n = 1$ и $\lambda = p$, учитывая справедливое при $0 \leqslant p \leqslant 1$ неравенство $0 \leqslant 1 - \exp\{-p\} \leqslant p$, имеем

$$|Q(1,0,p) - \pi(0;p)| = \exp\{-p\} + p - 1,$$
$$|Q(1,1,p) - \pi(1;p)| = p - p\exp\{-p\},$$
$$|Q(1,m,p) - \pi(m;p)| = p^m (m!)^{-1} \exp\{-p\},\ m \geqslant 2.$$

Затем из (4.19) при $n = 1$ следует соотношение

$$\Sigma_1 = 2^{-1} \sum_{m\in L} |Q(1,m,p) - \pi(m;p)| =$$
$$= 2^{-1}(\exp\{-p\} + p - 1 + p - p\,\exp\{-p\} + \sum_{m=2}^{\infty} \frac{p^m}{m!} \exp\{-p\}) =$$
$$= 2^{-1}(\exp\{-p\} + p - 1 + p - p\,\exp\{-p\} + 1 - \exp\{-p\} - p\exp\{-p\}) =$$
$$= p(1 - \exp\{-p\}) \leqslant p^2, \tag{4.20}$$

так как $0 \leqslant 1 - \exp\{-p\} \leqslant p$ при $0 \leqslant p \leqslant 1$. Предположим теперь, что имеет место неравенство

$$\Sigma_{n-1} \leqslant (n-1)\,p^2. \tag{4.21}$$

Установим ряд важных соотношений для функций $Q(n,m,p)$ и $\pi(m;\lambda)$. Используя формулу для $Q(n,m,p) = C_n^m p^m q^{n-m}$, легко убедиться в справедливости рекуррентного соотношения вида

$$Q(n,m,p) = Q(n-1,m,p)\,Q(1,0,p) + Q(n-1,m-1,p)\,Q(1,1,p) =$$
$$= \sum_{k=0}^{m} Q(n-1,m-k,p)\,Q(1,k,p). \tag{4.22}$$

В самом деле,

$$\begin{aligned}
&Q(n-1,m,p)\,Q(1,0,p)+Q(n-1,m-1,p)\,Q(1,1,p)=\\
&\quad=C_{n-1}^{m}p^{m}q^{n-m-1}C_1^0p^0q^1+C_{n-1}^{m-1}p^{m-1}q^{n-m}C_1^1p^1q^0=\\
&\quad=p^mq^{n-m}\Big(\frac{(n-1)!}{m!(n-m-1)!}+\frac{(n-1)!}{(m-1)!(n-m)!}\Big)=\\
&\quad=\frac{n!}{m!(n-m)!}p^mq^{n-m}=Q(n,m,p).
\end{aligned}$$

Если еще учесть равенства $Q(1,k,p)=0$, $k=2,3,\ldots$, то окончательно получим соотношение (4.22). Из определения функции $\pi(m;\lambda)$ при $\lambda_1\geqslant 0$ и $\lambda_2\geqslant 0$ имеем

$$\begin{aligned}
\pi(m;\lambda_1+\lambda_2)&=\frac{(\lambda_1+\lambda_2)^m}{m!}\exp\{-(\lambda_1+\lambda_2)\}=\exp\{-\lambda_1\}\exp\{-\lambda_2\}\times\\
&\times\frac{1}{m!}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}\lambda_1^{m-k}\lambda_2^k=\sum_{k=0}^{m}\pi(m-k;\lambda_1)\pi(k;\lambda_2). \qquad (4.23)
\end{aligned}$$

Если принять во внимание соотношения (4.19)–(4.23) и равенства $\lambda=np$, $\lambda_1=(n-1)p$, $\lambda_2=p$ и $\lambda=\lambda_1+\lambda_2$, то доказательство этой теоремы завершает следующая цепочка равенств и неравенств:

$$\begin{aligned}
&Q(n,m,p)-\pi(m;\lambda)=\\
&\quad=\sum_{k=0}^{m}Q(n-1,m-k,p)\,Q(1,k,p)-\sum_{k=0}^{m}\pi(m-k;(n-1)p)\,\pi(k;p)=\\
&\quad=\sum_{k=0}^{m}Q(n-1,m-k,p)\,Q(1,k,p)-\sum_{k=0}^{m}\pi(m-k;(n-1)p)\,Q(1,k,p)+\\
&\qquad\qquad+\sum_{k=0}^{m}Q(1,k,p)-\pi(k;p))\pi(m-k;(n-1)p)
\end{aligned}$$

и

$$\begin{aligned}
\Sigma_n&=2^{-1}\sum_{m\in L}|Q(n,m,p)-\pi(m;\lambda)|\leqslant\\
&\leqslant 2^{-1}\sum_{m\in L}\sum_{k=0}^{m}|Q(n-1,m-k,p)-\pi(m-k;(n-1)p)|\,Q(1,k,p)+\\
&\qquad+2^{-1}\sum_{m\in L}\sum_{k=0}^{m}|Q(1,k,p)-\pi(k;p)|\,\pi(m-k;(n-1)p)=\\
&=2^{-1}\sum_{m\in L}|Q(n-1,m,p)-\pi(m;(n-1)p)|\,Q(1,0,p)+\\
&\qquad+2^{-1}\sum_{m\in L}|Q(n-1,m-1,p)-\pi(m-1;(n-1)p)|\,Q(1,1,p)+
\end{aligned}$$

$$
\begin{aligned}
&+ 2^{-1} \sum_{m\in L} \pi(m;(n-1)p)\,|Q(1,0,p) - \pi(0;p)| + \\
&+ 2^{-1} \sum_{m\in L} \pi(m-1;(n-1)p)\,|Q(1,1,p) - \pi(1;p)| + \\
&+ 2^{-1} \sum_{k=2}^{\infty} \pi(k;p) \sum_{m\in L} \pi(m-k;(n-1)p) = \\
= {} & 2^{-1} \sum_{m\in L} |Q(n-1,m,p) - \pi(m;(n-1)p)|\,(1-p) + \\
&+ 2^{-1} \sum_{m\in L} |Q(n-1,m-1,p) - \pi(m-1;(n-1)p)|p + \\
&+ 2^{-1} \sum_{m\in L} \pi(m;(n-1)p)\,(\exp\{-p\} + p - 1) + \\
&+ 2^{-1} \sum_{m\in L} \pi(m-1;(n-1)p)\,(p - p\exp\{-p\}) + \\
&+ 2^{-1} \sum_{k=2}^{\infty} \pi(k;p) \sum_{m\in L} \pi(m-k;(n-1)p) = \\
= {} & 2^{-1} \sum_{m\in L} |Q(n-1,m,p) - \pi(m;(n-1)p)| + \\
&+ 2^{-1}\Big(\exp\{-p\} + p - 1 + p - p\exp\{-p\} + \sum_{m=2}^{\infty} \frac{p^m}{m!}\exp\{-p\}\Big) \leqslant \\
\leqslant {} & \Sigma_{n-1} + p^2 \leqslant (n-1)p^2 + p^2 = np^2.
\end{aligned}
$$

Рассмотрим теперь некоторые типичные задачи, решаемые с помощью теорем Пуассона.

Пример 4.21. На птицеферме имеется 100 куриц. По многолетним наблюдениям фермер оценил, что каждая несушка в конце зимы ежедневно откладывает одно яйцо с вероятностью 0,02. Определить вероятность того, что фермер в течение дня обнаружит в кладке пять яиц.

Решение. Каждая из $n = 100$ куриц ежедневно в конце зимы независимо друг от друга откладывает одно яйцо с вероятностью $p = 0{,}02$ и не откладывает с вероятностью $q = 0{,}98$. Таким образом, эта реальная задача хорошо описывается схемой независимых испытаний Бернулли. Следовательно, искомая вероятность $\mathsf{P}(\boldsymbol{B}_5) = C_{100}^5(0{,}02)^5(0{,}98)^{95} \approx 0{,}0353$. Поскольку число куриц, равное 100, достаточно велико и вероятность успешного исхода отложить одно яйцо отдельной несушкой, равная 1/50, мала, то по приближенной формуле Пуассона при $m = 5$, $\lambda = np = 100 \times (0{,}02) = 2$ с использованием статистических таблиц для известной функции $\pi(m;\lambda)$ имеем $\mathsf{P}(\boldsymbol{B}_5) = P_{100}(5) = C_{100}^5(0{,}02)^5(0{,}98)^{95} \approx \pi(5;2) \approx 0{,}0361$.

Погрешность такого приближения согласно теореме 4.7 будет не более $np^2 = 100(0{,}02)^2 = 0{,}04$. Однако для этой конкретной задачи находим, что замена точного значения искомой вероятности $\mathsf{P}(\boldsymbol{B}_5)$ на приближенное значение $\pi(5;2)$ дает ошибку $0{,}0361 - 0{,}0353 = 0{,}0008$. Эта ошибка существенно меньше числа 0,04. Заметим, что вычисленное значение искомой вероятности $\mathsf{P}(\boldsymbol{B}_5) = C_{100}^5(0{,}02)^5(0{,}98)^{95}$ по приближенной формуле Муавра–Лапласа (4.8) при $np = 2$, $(npq)^{1/2} = (100 \times 0{,}02 \times 0{,}98)^{1/2} = 7/5$, $m = 5$, $x_5 = (m - np)(npq)^{-1/2} \approx 2{,}143$ и $\varphi(x_5) \approx 0{,}0402$ равно 0,0287. Погрешность приближенной формулы (4.8) для этого примера равна $0{,}0353 - 0{,}0287 = 0{,}0066$. Это намного больше погрешности 0,0008 приближенной формулы (4.16). Такой результат объясняется тем, что в этой задаче равная 0,02 вероятность успешного исхода мала и $npq = 1{,}96 < 20$.

Имея в наличии приближенные формулы (4.16) и (4.17), мы можем теперь получить численное решение примера 4.9. В этом примере нам нужно было определить численное значение вероятностей $\mathsf{P}(\boldsymbol{B}_{14})$ и $\mathsf{P}\big(\cup_{m=10}^{19}\boldsymbol{B}_m\big)$ при $n = 10\,000$, $p = 0{,}0008$ и $np = 8$. Отсюда с точностью до четвертого десятичного знака легко находим: $\pi(14;8) = 8^{14}(14!)^{-1}\exp\{-8\} \approx 0{,}0169$ и $\sum_{m=10}^{19} 8^m(m!)^{-1}\exp\{-8\} \approx 0{,}2831$. Следовательно, $\mathsf{P}(\boldsymbol{B}_{14}) \approx 0{,}0169$ и вероятность $\mathsf{P}\big(\cup_{m=10}^{19}\boldsymbol{B}_m\big) \approx 0{,}2831$ с абсолютной ошибкой меньше величины $np^2 = 10\,000 \times (0{,}0008)^2 = 0{,}0064$.

Как уже подчеркивали, приближенная формула (4.8) даже при очень больших значениях n может давать совсем непригодные результаты, если вероятность p успешного исхода слишком мала и $\lambda = np \ll 20$. Однако при $\lambda = np \geqslant 9$ аппроксимация (4.8) уже дает вполне приемлемый результат. Более того, при $\lambda = np \geqslant 16$ используют приближенные формулы (4.8) и (4.14), так как очень трудно найти таблицы функции $\pi(m;\lambda)$ для таких значений λ. Рассмотрим конкретный пример такой ситуации.

Пример 4.22. Имеются четыре образца радиоактивного вещества, каждый из которых отличается только количеством нестабильных ядер. Первый образец содержит $0{,}9 \times 10^{20}$ ядер, второй 9×10^{20} ядер, третий 30×10^{20} и, наконец, четвертый образец содержит 90×10^{-20} ядер. Каждое из ядер этих образцов за любую фиксированную минуту может распасться с вероятностью $p = 10^{-20}$. Вычислить вероятность наблюдения за одну минуту следующих событий: 1) произошло ноль распадов для первого образца; 2) произошло девять распадов для второго образца; 3) произошло тридцать распадов для третьего образца; 4) произошло 89 распадов для четвертого образца.

Решение. В каждом образце имеется n ядер, а вероятность того, что любое одно ядро независимо от других ядер распадется, равна p.

Легко понять, что мы имеем дело с четырьмя отдельными схемами независимых испытаний Бернулли. В каждом таком эксперименте число n ядер огромно ($n \sim 10^{20}$), а вероятность успешного исхода (распада отдельного ядра) ничтожна, мала ($p = 10^{-20}$). При этих условиях можно применять приближенную формулу (4.16). Тогда для первого, второго, третьего и четвертого эксперимента соответственно имеем:

1) $n = 0{,}9 \times 10^{20}$, $p = 10^{-20}$, $\lambda = np = 0{,}9$, $Q(0{,}9 \times 10^{20}, 0, 10^{-20}) \approx \approx \pi(0; 0{,}9) \approx 0{,}406570$;
2) $n = 9 \times 10^{20}$, $p = 10^{-20}$, $\lambda = np = 9$, $Q(9 \times 10^{20}, 9, 10^{-20}) \approx \approx \pi(9; 9) \approx 0{,}131756$;
3) $n = 30 \times 10^{20}$, $p = 10^{-20}$, $\lambda = np = 30$, $Q(30 \times 10^{20}, 30, 10^{-20}) \approx \approx \pi(30; 30) \approx 0{,}072635$;
4) $n = 90 \times 10^{20}$, $p = 10^{-20}$, $\lambda = np = 90$, $Q(90 \times 10^{20}, 89, 10^{-20}) \approx \approx \pi(89; 90)$.

Для четвертого образца радиоактивного вещества не удается приближенно вычислить вероятность $Q(90 \times 10^{20}, 89, 10^{-20})$ наблюдения 89 распадов за одну минуту, так как нет таблиц для функции $\pi(m; \lambda)$ при $\lambda = 90$. Значение $\pi(89; 90)$ можно вычислить с помощью простой программы на компьютере. Если использовать приближенную формулу (4.8), то для этих задач находим:

1) $n = 0{,}9 \times 10^{20}$, $p = 10^{-20}$, $npq = 0{,}9 \times (1 - 10^{-20}) \approx 9/10$, $(npq)^{1/2} \approx 0{,}948683$, $m = 0$, $(m - np) \times (npq)^{-1/2} \approx \approx -0{,}948683$, $\varphi(-0{,}948683) \approx 0{,}254542$, $Q(0{,}9 \times 10^{20}, 0, 10^{-20}) \approx \approx \varphi(-0{,}948683)/0{,}948683 \approx 0{,}268310$;
2) $n = 9 \times 10^{20}$, $p = 10^{-20}$, $npq = 9 \times (1 - 10^{-20}) \approx 9$, $(npq)^{1/2} \approx 3$, $m = 9$, $(m - np) \times (npq)^{-1/2} = 0$, $\varphi(0) \approx 0{,}398942$ и, наконец, $Q(9 \times 10^{20}, 9, 10^{-20}) \approx \varphi(0)/3 \approx 0{,}13298$;
3) $n = 30 \times 10^{20}$, $p = 10^{-20}$, $npq = 30 \times (1 - 10^{-20}) \approx 30$, $(npq)^{1/2} \approx \approx 5,477225$, $m = 30$, $(m - np) \times (npq)^{-1/2} = 0$, $\varphi(0) \approx 0{,}398942$ и вероятность $Q(30 \times 10^{20}, 30, 10^{-20}) \approx \varphi(0)/5,477225 \approx 0{,}072836$;
4) если $n = 90 \times 10^{20}$, $p = 10^{-20}$, $npq = 90 \times (1 - 10^{-20}) \approx 90$, $(npq)^{1/2} \approx 9,486832$, $m = 89$, то $(m - np)(npq)^{-1/2} \approx -0{,}105409$, $\varphi(-0{,}105409) \approx 0{,}396791$ и, наконец, $Q(90 \times 10^{20}, 89, 10^{-20}) \approx \approx \varphi(-0{,}105409)/9,486832 \approx 0{,}041825$.

Для этих четырех экспериментов, каждый из которых отличается только числом ядер в образце, приближенная формула Пуассона (4.16) согласно теореме 4.7 дает абсолютную ошибку меньше $np^2 \leqslant 90 \times 10^{-20} < 10^{-18}$. Поэтому приближенная локальная формула Муавра–Лапласа для задачи 1) дает совершенно непригодный результат $0{,}268831 \ll 0{,}406570$. Для задачи 2) локальная формула Муавра Лапласа вычисляет искомую вероятность с точностью до двух десятичных знаков, так как $0{,}132980 \approx 0{,}131756$. Для задачи 3) приближенная формула (4.8) вычисляет искомую вероятность с точностью до трех десятичных знаков, ибо 0,072836 ≈ 0,072635. Значит, для задачи 4), где $npq \approx 90 \gg 30$, локальная формула Муавра–Лапласа дает вполне хороший результат. Более того, компьютер дает приближенное значение для функции $\pi(m;\lambda)$ при $m = 89$ и $\lambda = 90$, равное 0,041825.

В заключение заметим, что если число испытаний n принимает значение от 10 до 20, то приближенные формулы используются для грубых и предварительных расчетов. При этом формула Пуассона применяется в том случае, когда величины np или nq изменяются в пределах от нуля до двух при $n = 10$ и от нуля до трех при $n = 20$. В противном случае необходимо пользоваться формулами Муавра–Лапласа. Если параметр n изменяется от двадцати до ста, то приближенные формулы уже можно использовать для предварительных инженерных расчетов. Формулу Пуассона рекомендуется использовать при $n = 100$, когда величины np или nq заключены в промежутке от нуля до пяти. Если величина n больше числа 100, то практически при любых инженерных расчетах можно обойтись приближенными формулами. При $n = 1000$ применяется формула Пуассона, когда величины np или nq заключены в промежутке от нуля до десяти. Наконец, при $n > 1000$ даже специальные таблицы рассчитываются с помощью приближенных формул, однако для увеличения точности используются специальные поправки.

3.5. Вывод формул для вероятностей событий в полиномиальной схеме независимых испытаний. Рассмотрим несколько усложненный вариант схемы последовательных независимых испытаний Бернулли. Уже в каждом испытании E_t будем интересоваться тем, наступило или не наступило некоторое событие C_s из конечного семейства $\{C_1, C_2, \ldots, C_r\}$, составляющего полную группу несовместимых событий эксперимента E_0. Такую схему последовательности независимых испытаний Бернулли будем называть *полиномиальной*. Следовательно, в полиномиальной схеме Бернулли при каждом фиксированном $t = 1, 2, \ldots, n$ наступление события $C_{s,t}$ в эксперименте E_t означает то же самое, что и наступление события C_s в t-том испытании опыта E_0. Легко видеть, что при $r = 2$ полиномиальная схема Бернулли будет биномиальной. Так как вероятности $\mathsf{P}_1(\{\omega_1\})$, $\mathsf{P}_2(\{\omega_2\}), \ldots$, $\mathsf{P}_n(\{\omega_n\})$ в любой схеме Бернулли равны, если символы ω_1, $\omega_2, \ldots$, ω_n задают описание одного и того же элементарного исхода эксперимента E_0, то при любом $t = 1, 2, \ldots, n$ имеем

$$\mathsf{P}_t(C_{s,t}) = \sum_{\omega_t \in C_{s,t}} \mathsf{P}_t(\{\omega_t\}) = \sum_{\omega_t \in C_s} \mathsf{P}_t(\{\omega_t\}) = \sum_{\omega_1 \in C_s} \mathsf{P}_1(\{\omega_1\}) = \\ = \sum_{\omega_1 \in C_{s,1}} \mathsf{P}_1(\{\omega_1\}) = \mathsf{P}_1(C_{s,1}). \quad (4.24)$$

Поэтому вероятности $\mathsf{P}_1(C_{s,1})$, $\mathsf{P}_2(C_{s,2}), \ldots$, $\mathsf{P}_n(C_{s,n})$ можно обозначить через p_s, где $p_1 + p_2 + \ldots + p_r = 1$. Итак, p_s здесь определяет вероятность наступления события C_s эксперимента E_0 при любом его испытании.

Для составного эксперимента $E = \{E_1, E_2, \ldots, E_n\}$ в полиномиальной схеме Бернулли с параметрами n и p_1, $p_2, \ldots$, p_r, как правило, рассматривают исход $\boldsymbol{C}(m_1, m_2, \ldots, m_r)$, который состоит в том, что в каких-то испытаниях событие C_1 наступит ровно m_1 раз, событие C_2 наступит ровно m_2 раз,..., событие C_r наступит ровно m_r раз. При этом имеем очевидное равенство $m_1 + m_2 + \ldots + m_r = n$. Найдем теперь вероятность события вида $\boldsymbol{C}(m_1, m_2, \ldots, m_r) \in \mathcal{F}$.

Пусть при каждом фиксированном $s = 1, 2, \ldots, r$ величины $t(s,1)$, $t(s,2), \ldots$, $t(s, m_s)$ суть различные номера испытаний опыта E_0, в каждом из которых появится его исход C_s. Теперь вместо события $C_{s,t}$ из вероятностной модели $(\Omega_t, \mathcal{F}_t, \mathsf{P}_t(\cdot))$ эксперимента E_t будем рассматривать изоморфное ему событие — прямоугольник вида

$$\boldsymbol{C}_{s,t} = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{t-1} \times C_{s,t} \times \Omega_{t+1} \times \ldots \times \Omega_n$$

из вероятностной модели $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ более сложного опыта E. Тогда событие

$$\bigcap_{s=1}^{r} \bigcap_{k=1}^{m_s} \boldsymbol{C}_{s,t(s,k)}$$

означает, что в испытаниях опыта E_0 с фиксированными номерами $t(1,1)$, $t(1,2), \ldots$, $t(1, m_1)$ событие C_1 наступит ровно m_1 раз, в испытаниях опыта E_0 с фиксированными номерами $t(2,1)$, $t(2,2), \ldots$, $t(2, m_2)$ событие C_2 наступит ровно m_2 раз,... и, наконец, в испытаниях опыта E_0 с фиксированными номерами $t(r,1)$, $t(r,2), \ldots$, $t(r, m_r)$ событие C_r наступит ровно m_r раз. По теореме 4.2, используя соотношение (4.24), легко находим

$$\mathsf{P}\Big(\bigcap_{s=1}^{r} \bigcap_{k=1}^{m_s} \boldsymbol{C}_{s,t(s,k)}\Big) = \prod_{s=1}^{r} \prod_{k=1}^{m_s} \mathsf{P}(\boldsymbol{C}_{s,t(s,k)}) = \prod_{s=1}^{r} \prod_{k=1}^{m_s} \mathsf{P}_{t(s,k)}(C_{s,t(s,k)}) =$$
$$= \prod_{s=1}^{r} p_s^{m_s} = p_1^{m_1} \times p_2^{m_2} \times \ldots \times p_r^{m_r}$$

при любых различных номерах $t(1,1)$, $t(1,2), \ldots$, $t(1, m_1)$, $t(2,1)$, $t(2,2), \ldots$, $t(2, m_2), \ldots$, $t(r,1)$, $t(r,2), \ldots$, $t(r, m_r)$.

Среди всех номеров $1, 2, \ldots, n$ различные номера $t(1,1)$, $t(1,2), \ldots$, $t(1, m_1)$ испытаний опыта E_0, в каждом из которых появится исход C_1, можно выбрать $C_n^{m_1}$ способами; из оставшихся $n - m_1$ номеров различные номера $t(2,1)$, $t(2,2), \ldots$, $t(2, m_2)$ испытаний опыта E_0, в каждом из которых появится исход C_2, можно выбрать $C_{n-m_1}^{m_2}$ способами и т. д. Наконец, из оставшихся $n - m_1 - \ldots - m_{r-1}$ номеров различные номера $t(r,1)$, $t(r,2)$, …, $t(r,m_r)$ испытаний опыта E_0, в каждом из которых появится исход C_r, можно выбрать только единственным способом, так как число $m_r = n - m_1 - \ldots - m_{r-1}$ и $C^{m_r}_{n-m_1-\ldots-m_{r-1}} = C^{m_r}_{m_r} = 1$. Итак, существует

$$C_n^{m_1} \times C_{n-m_1}^{m_2} \times \ldots \times C^{m_r}_{n-m_1-\ldots-m_{r-1}} = \frac{n!}{m_1! \times m_2! \times \ldots \times m_r!} \stackrel{\text{def}}{=} \\ \stackrel{\text{def}}{=} \begin{pmatrix} n \\ m_1, m_2, \ldots, m_r \end{pmatrix}$$

различных наборов множеств

$$\mathsf{t} = \{t(1,1), t(1,2), \ldots, t(1,m_1), t(2,1), t(2,2), \ldots, t(2,m_2), \ldots, \\ t(r,1), t(r,2), \ldots, t(r,m_r)\}.$$

Нетрудно также убедиться в том, что событие

$$\boldsymbol{C}(m_1, m_2, \ldots, m_r) = \bigcup_{\mathsf{t}} \bigcap_{s=1}^{r} \bigcap_{k=1}^{m_s} \boldsymbol{C}_{s,t(s,k)}.$$

В этом равенстве объединение ведется по всевозможным наборам — множествам вида

$$\mathsf{t} = \{t(1,1), t(1,2), \ldots, t(1,m_1), t(2,1), t(2,2), \ldots, t(2,m_2), \ldots, \\ t(r,1), t(r,2), \ldots, t(r,m_r)\}.$$

Принимая во внимание, что различные случайные события вида

$$\bigcap_{s=1}^{r} \bigcap_{k=1}^{m_s} \boldsymbol{C}_{s,t(s,k)}$$

являются несовместимыми и равновероятными, получим

$$\mathsf{P}(\boldsymbol{C}(m_1, m_2, \ldots, m_r)) = \mathsf{P}\Big(\bigcup_{\mathsf{t}} \bigcap_{s=1}^{r} \bigcap_{k=1}^{m_s} \boldsymbol{C}_{s,t(s,k)}\Big) = \sum_{t} \mathsf{P}\Big(\bigcap_{s=1}^{r} \bigcap_{k=1}^{m_s} \boldsymbol{C}_{s,t(s,k)}\Big) = \\ = \begin{pmatrix} n \\ m_1, m_2, \ldots, m_r \end{pmatrix} p_1^{m_1} \times p_2^{m_2} \times \ldots \times p_r^{m_r}. \quad (4.25)$$

Эта вероятность является коэффициентом при члене $x_1^{m_1} \cdot x_2^{m_2} \cdot \ldots \cdot x_r^{m_r}$ в следующем разложении полинома вида

$$(p_1x_1 + p_2x_2 + \ldots + p_rx_r)^n = \sum_{m_1+m_2+\ldots+m_r=n} \begin{pmatrix} n \\ m_1, m_2, \ldots, m_r \end{pmatrix} \times \\ \times p_1^{m_1} \times p_2^{m_2} \times \ldots \times p_r^{m_r} \times x_1^{m_1} \times x_2^{m_2} \times \ldots \times x_r^{m_r}.$$

При $x_1 = x_2 = \ldots = x_r = 1$ имеем равенство

$$1 = (p_1 + p_2 + \ldots + p_r)^n =$$
$$= \sum_{m_1+m_2+\ldots+m_r=n} \binom{n}{m_1, m_2, \ldots, m_r} p_1^{m_1} \times p_2^{m_2} \times \ldots \times p_r^{m_r}.$$

Поэтому описанный эксперимент E, составленный из n последовательных независимых испытаний опыта E_0 с несовместимыми и образующими полную группу исходами C_1, $C_2, \ldots$, C_r, называется *полиномиальной схемой Бернулли*. Еще раз напомним, что в частном случае при $r = 2$, $p_1 = p$, $p_2 = q$, $m_1 = m$, $m_2 = n - m$ полиномиальная схема будет биномиальной схемой Бернулли. Рассмотрим простой пример полиномиальной схемы Бернулли.

Пример 4.23. Грани сделанного из однородного материала кубика пронумерованы цифрами от единицы до шести. Причем грань, занумерованная цифрой единица, заштрихована белым цветом, грани с четными номерами окрашены синим цветом, и, наконец, грани с номерами три и пять помечены красным цветом. Опыт E_0 состоит в непреднамеренном подбрасывании этого кубика и фиксировании выпавшей цифры на верхней его грани. Следовательно, опыт E_0 порождает шесть элементарных событий, например, выпадение грани кубика с номером $i = 1, 2, \ldots, 6$. Составной эксперимент $E = \{E_1, E_2, \ldots, E_{12}\}$ здесь заключается в последовательном и независимом подбрасывании кубика ровно двенадцать раз.

В опыте E_0, например, возможно появление одного из трех несовместимых исходов C_1, C_2, C_3, где C_1, C_2 или C_3 соответственно означает наблюдение белого, синего или красного цвета. На основе классического определения вероятностей для эксперимента E_0 получим $p_1 = \mathsf{P}_t(C_{1,t}) = 1/6$, $p_2 = \mathsf{P}_t(C_{2,t}) = 1/2$, $p_3 = \mathsf{P}_t(C_{3,t}) = 1/3$ для всех $t = 1, 2, \ldots, 12$. Тогда вероятность того, что белый цвет появится ровно два раза, синий цвет — ровно шесть раз и красный цвет — ровно четыре раза, вычислим по формуле (4.25). Эта вероятность равна $\mathsf{P}(\boldsymbol{C}(2, 6, 4)) = 12! \times (1/2!) \times (1/6!) \times (1/4!) \times (1/6)^2 \times (1/2)^6 \times$ $\times\, (1/3)^4 \approx 0{,}0742$.

Краткий обзор

В главе 4 предложен метод составления вероятностной модели общей схемы причинно-независимых испытаний, при этом каждое испытание имеет конечное число исходов. Изучены простейшие свойства вероятностной модели общей схемы независимых испытаний.

Подробно рассмотрена так называемая схема независимых испытаний Бернулли. Для схемы Бернулли с параметрами n и p выводятся формулы для вероятностей случайных событий, которые представляют наибольший практический интерес. С целью получения приближенных формул для вычисления биномиальных вероятностей в схеме независимых испытаний Бернулли доказаны предельные теоремы Муавра–Лапласа и Пуассона. Рассмотрен важный вопрос об оценке совершаемой погрешности при вычислении биномиальных вероятностей с помощью приближенных формул. Выяснена фундаментальная связь между предельными теоремами в схеме независимых испытаний Бернулли и понятием статистической устойчивости случайного эксперимента. Приведена методика получения формул для вероятностей событий в полиномиальной схеме независимых испытаний Бернулли. Материал этой главы иллюстрируется большим числом решенных задач.

Контрольные вопросы и упражнения

1. Привести примеры конечного числа причинно-независимых испытаний, каждое из которых имеет конечное число исходов, и примеры биномиальной схемы Бернулли.
2. Построить простейшую вероятностную модель для схемы последовательных независимых опытов серии объема n. Вывести формулу наивероятнейшего числа появлений события A в схеме Бернулли.
3. Дать простейшее объяснение близости наивероятнейшего числа появлений события A в схеме Бернулли к величине np.
4. Буквы слова «панама» написаны каждая на отдельной карточке. Одинаковые по форме карточки тщательно перемешаны, после чего последовательно и без возвращения извлекаются три. Выбранные карточки кладутся слева направо. Определить вероятность того, что в результате получается слово «пан» или, другими словами, наступает событие A. Рассмотреть и сравнить различные способы решения этой задачи.
5. При одном выстреле вероятность перелета цели равна 2/3, а вероятность недолета — 1/3. Найти наивероятнейшую комбинацию недолетов и перелетов при четырех независимых выстрелах. Определить вероятность такой комбинации.
6. Вероятность поражения цели при одном выстреле любого из n одинаковых орудий равна 0,01. Найти минимальное число орудий, чтобы вероятность поражения цели при залпе была больше 3/4.
7. Вероятность поражения самолета одной ракетой равна $p = 1/3$. По самолету выпускается $n = 12$ ракет. Найти наивероятнейшее число попаданий в самолет.

8. Указать условия, когда можно пользоваться предельными теоремами Муавра–Лапласа, Пуассона.
9. Каково фундаментальное значение предельных теорем Муавра–Лапласа в связи с понятием статистической устойчивости случайного эксперимента?
10. Привести пример, на котором можно обосновать приближенное вычисление значения функции $Q(n, m, p)$ от трех переменных n, m, p с помощью значения функции $\pi(m; \lambda)$ только от двух переменных m, λ.
11. Вероятность рождения мальчика примем равной 0,5. Используя таблицы П.1 и П.2, найти вероятность того, что среди 100 новорожденных детей будет: 1) 40 мальчиков; 2) 50 мальчиков; 3) от 40 до 60 мальчиков.
12. По каналу связи передано независимым образом $n = 2100$ знаков. Вероятность искажения любого передаваемого знака равна $p = 0{,}3$. Какова вероятность получения ровно 651 искажений?
13. Найти вероятность того, что при $n = 500$ бросаниях симметричной монеты герб появится от 240 до 260 раз включительно.
14. Используя табл. П.3, найти вероятность того, что в обществе из 1095 человек ровно 10 человек родились в первый день нового года.
15. Вероятность выхода из строя одного компьютера за некоторое фиксированное время равна 0,2. Определить вероятность того, что из 100 такого типа компьютеров в течение этого времени выйдет из строя от 14 до 26.
16. По цели производится 1000 независимых выстрелов. Вероятность попадания в цель при одном выстреле равна 0,002. Используя табл. П.4, вычислить вероятность того, что в результате стрельбы будет зафиксировано не менее четырех попаданий.
17. Некто купил 300 карточек для игры в тираже «Спортлото 6 из 49». На каждой из этих карточек он случайным образом и независимо от остальных карточек отмечает только шесть из 49 различных видов спорта. Максимальный выигрыш на каждую карточку выпадает, если некто отметил на этой карточке шесть видов спорта, которые совпадают с непреднамеренным выбором шесть из 49 спортивных номеров лотерейной комиссией. Какова вероятность того, что некто получит ровно два максимальных выигрыша?
18. Лабиринт для мыши составлен из клеток с номерами от 1 до 100. Мышь помещается в первую клетку и за единицу времени с равной возможностью остается в той же клетке либо переползает в клетку, номер которой увеличивается на единицу. Найти

вероятность того, что за 99 единиц времени мышь окажется в клетке с номером 99.

19. В школе 600 учеников. Найти наивероятнейшее число учащихся, родившихся 1 мая.
20. Большая партия изделий содержит один процент брака. Каков должен быть объем n случайной выборки, чтобы наивероятнейшее число появлений в ней небракованных изделий равнялось 30.
21. Вероятность попадания в корабль каждой сбрасываемой с самолета независимо одна от другой торпеды равна 0,2. Для гибели корабля необходимо не менее двух попаданий. Какое количество торпед необходимо сбросить, чтобы вероятность гибели корабля была не менее 0,9?
22. В квартире, состоящей из четырех сообщающихся и равных по объему комнат, находятся $4n$ пылинок. Объем квартиры равен V. Определить вероятность того, что в первой комнате будет находиться ровно n пылинок.
23. Два равносильных шахматиста играют ряд партий и ничьи в счет не идут. Что вероятнее для первого игрока выиграть три партии из четырех или четыре партии из пяти?
24. Среди всех собранных компьютеров 98 % являются небракованными, 1 % имеют устранимый брак и 1 % компьютеров будут с неустранимым браком. Вычислить вероятность того, что в результате купленных наудачу трех компьютеров хотя бы один из них небракованный и хотя бы один из них — с устранимым браком.

Глава 5

ПРОСТЕЙШАЯ ФОРМА ПОСЛЕДОВАТЕЛЬНОСТИ ЗАВИСИМЫХ ИСПЫТАНИЙ

§ 1. Общая схема Маркова

1.1. Составление вероятностной модели схемы Маркова. Рассмотрим теперь эволюционный эксперимент $E = \{E_t;\, t{=}1, 2, \ldots, n\}$, для которого каждый из опытов E_1, $E_2, \ldots$, E_n имеет одну и ту же теоретико-множественную модель с конечным числом r элементарных исходов или состояний $\{s_1\}$, $\{s_2\}, \ldots$, $\{s_r\}$. В этом случае при $t = 1, 2, \ldots, n$ положим $(\Omega_1, \mathcal{F}_1) = (\Omega_t, \mathcal{F}_t)$ и $\Omega_1 = \{s_1, s_2, \ldots, s_r\}$, $\mathcal{F}_1 = \{A\colon A \subset \Omega_1\}$. Построенную в первой главе теоретико-множественную модель $(\Omega, \mathcal{F})$ для статистически устойчивого эволюционного эксперимента $E = \{E_1, E_2, \ldots, E_n\}$ удобно обозначить через упорядоченную пару $(\Omega_1^n, \mathcal{F}_1^n)$, где достоверное событие $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n = \Omega_1^n$ и σ-алгебра $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2 \times \ldots \times \mathcal{F}_n = \mathcal{F}_1^n$.

Для составного эксперимента E с конечным числом 2^{r^n} всевозможных исходов или для последовательности конечного числа n испытаний E_1, $E_2, \ldots$, E_n в общем случае вероятностную модель $(\Omega_1^n, \mathcal{F}_1^n, \mathsf{P}(\cdot))$ задают с помощью вероятностей $\mathsf{P}(\{\omega\})$ элементарных исходов с описанием следующего вида: $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$. Здесь $\omega_t \in \Omega_t$ при каждом $t = 1, 2, \ldots, n$. Если событие из $\mathcal{F}_1^n$ обозначить полужирным латинским символом и курсивом, то вероятность события $\boldsymbol{A} \in \mathcal{F}_1^n$ можно определить следующим образом: $\mathsf{P}(\boldsymbol{A}) = \sum_{\omega \in \boldsymbol{A}} \mathsf{P}(\{\omega\})$, $\boldsymbol{A} \in \mathcal{F}_1^n$. При этом вероятности $\mathsf{P}(\{\omega\})$ элементарных исходов уже не вычисляются по формуле (4.1) и, значит, эксперимент E составлен из n последовательных зависимых испытаний E_1, $E_2, \ldots$, E_n. Для вероятностей элементарных исходов составного эксперимента E заранее требуется выполнение естественных условий $\mathsf{P}(\{\omega\}) \geqslant 0$, $\sum_{\omega \in \Omega_1^n} \mathsf{P}(\{\omega\}) = 1$. Отсюда видно, что модель $(\Omega_1^n, \mathcal{F}_1^n, \mathsf{P}(\cdot))$ удовлетворяет аксиомам Колмогорова. Заметим, что проверка условия $\mathsf{P}\bigl(\bigcup\limits_{\gamma=1}^{\varkappa} \boldsymbol{B}_\gamma\bigr) = \sum_{\gamma=1}^{\varkappa} \mathsf{P}(\boldsymbol{B}_\gamma)$ для любой конечной последовательности $\boldsymbol{B}_1$, $\boldsymbol{B}_2, \ldots$, $\boldsymbol{B}_\varkappa$ попарно несовместимых событий из $\mathcal{F}_1^n$ дословно повторяет случай независимых испытаний из четвертой главы.

Введем в рассмотрение события — прямоугольники из $\mathcal{F}_1^n$:

$$\begin{aligned}
\boldsymbol{A}_1(\omega_1) &= \{\omega_1\} \times \Omega_2 \times \Omega_3 \times \ldots \times \Omega_n,\\
\boldsymbol{A}_2(\omega_2) &= \Omega_1 \times \{\omega_2\} \times \Omega_3 \times \ldots \times \Omega_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\boldsymbol{A}_n(\omega_n) &= \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{n-1} \times \{\omega_n\}.
\end{aligned}$$

В этих равенствах $\{\omega_t\}$ есть элементарное событие эксперимента E_t при $t = 1, 2, \ldots, n$. Тогда имеем равенство $\{\omega = (\omega_1, \omega_2, \ldots, \omega_n)\} = \boldsymbol{A}_1(\omega_1) \cap \boldsymbol{A}_2(\omega_2) \cap \ldots \cap \boldsymbol{A}_n(\omega_n)$. Используя для всех $\omega \in \Omega_1^n$ дополнительное ограничение $\mathsf{P}(\{\omega\}) > 0$ и теорему умножения, вероятности элементарных исходов эксперимента E представим в виде

$$\begin{aligned}
\mathsf{P}(\{\omega\}) &= \mathsf{P}(\boldsymbol{A}_1(\omega_1) \cap \boldsymbol{A}_2(\omega_2) \cap \ldots \cap \boldsymbol{A}_n(\omega_n)) = \\
&= \mathsf{P}(\boldsymbol{A}_1(\omega_1)) \times \mathsf{P}(\boldsymbol{A}_2(\omega_2)|\boldsymbol{A}_1(\omega_1)) \times \mathsf{P}(\boldsymbol{A}_3(\omega_3)|\boldsymbol{A}_1(\omega_1), \boldsymbol{A}_2(\omega_2)) \times \\
&\quad \times \ldots \times \mathsf{P}(\boldsymbol{A}_n(\omega_n)|\boldsymbol{A}_1(\omega_1), \boldsymbol{A}_2(\omega_2), \ldots, \boldsymbol{A}_{n-1}(\omega_{n-1})). \quad (5.1)
\end{aligned}$$

Отметим, что при дополнительном ограничении $\mathsf{P}(\{\omega\}) > 0$, $\omega \in \Omega_1^n$, все величины в правой части соотношения (5.1) определены. Пусть условные вероятности в правой части равенства (5.1) при $t = 2, 3, \ldots, n$ и при $\omega_1, \omega_2, \ldots, \omega_n \in \{s_1, s_2, \ldots, s_r\}$ удовлетворяют следующим условиям:

$$\mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_1(\omega_1), \boldsymbol{A}_2(\omega_2), \ldots, \boldsymbol{A}_{t-1}(\omega_{t-1})) = \mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1})). \quad (5.2)$$

Отсюда соотношение (5.1) примет более простую форму:

$$\begin{aligned}
\mathsf{P}(\{\omega\}) &= \mathsf{P}(\boldsymbol{A}_1(\omega_1) \cap \boldsymbol{A}_2(\omega_2) \cap \ldots \cap \boldsymbol{A}_n(\omega_n)) = \\
&= \mathsf{P}(\boldsymbol{A}_1(\omega_1)) \times \mathsf{P}(\boldsymbol{A}_2(\omega_2)|\boldsymbol{A}_1(\omega_1)) \times \mathsf{P}(\boldsymbol{A}_3(\omega_3)|\boldsymbol{A}_2(\omega_2)) \times \\
&\quad \times \ldots \times \mathsf{P}(\boldsymbol{A}_n(\omega_n)|\boldsymbol{A}_{n-1}(\omega_{n-1})). \quad (5.3)
\end{aligned}$$

Последовательность зависимых испытаний, для которой имеет место соотношение (5.2) и, следовательно, вероятности элементарных исходов определяются формулой (5.3), называется *схемой Маркова*. Если при этом условные вероятности $\mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1}))$ зависят от t, то схема Маркова называется *неоднородной*, а в противном случае называется *однородной*. Здесь следует отметить, что схема Маркова является простейшим и вместе с тем очень важным для практики обобщением схемы независимых испытаний на особый случай, когда $\mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1})) \neq \mathsf{P}(\boldsymbol{A}_t(\omega_t))$, и, значит, испытания E_1, $E_2, \ldots$, E_n будут зависимыми.

1.2. Свойства марковости. При $t = 2, 3, \ldots, n$ установим справедливость следующего обобщения равенства (5.2):

$$\mathsf{P}(\boldsymbol{A}_t | \boldsymbol{A}_1(\omega_1), \boldsymbol{A}_2(\omega_2), \ldots, \boldsymbol{A}_{t-1}(\omega_{t-1})) = \mathsf{P}(\boldsymbol{A}_t | \boldsymbol{A}_{t-1}(\omega_{t-1})), \tag{5.4}$$

где $\boldsymbol{A}_t = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{t-1} \times A_t \times \Omega_{t+1} \times \ldots \times \Omega_n$ является событием — прямоугольником из $\mathcal{F}_1^n = \mathcal{F}$ и A_t есть любое событие из $\mathcal{F}_t = \mathcal{F}_1$. Если $A_t = \{\omega_t\}$, то $\boldsymbol{A}_t = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{t-1} \times \{\omega_t\} \times \Omega_{t+1} \times \ldots \times \Omega_n = \boldsymbol{A}_t(\omega_t)$.

Доказательство (5.4) следует из представления $\boldsymbol{A}_t = \cup_{\omega_t \in A_t} \boldsymbol{A}_t(\omega_t)$, свойств условных вероятностей, равенства (5.2) и цепочки равенств:

$$\begin{aligned}
&\mathsf{P}(\boldsymbol{A}_t | \boldsymbol{A}_1(\omega_1), \boldsymbol{A}_2(\omega_2), \ldots, \boldsymbol{A}_{t-1}(\omega_{t-1})) = \\
&\qquad = \mathsf{P}\Bigg(\bigcup_{\omega_t \in A_t} \boldsymbol{A}_t(\omega_t) | \boldsymbol{A}_1(\omega_1), \boldsymbol{A}_2(\omega_2), \ldots, \boldsymbol{A}_{t-1}(\omega_{t-1}) \Bigg) = \\
&\qquad = \sum_{\omega_t \in A_t} \mathsf{P}(\boldsymbol{A}_t(\omega_t) | \boldsymbol{A}_1(\omega_1), \boldsymbol{A}_2(\omega_2), \ldots, \boldsymbol{A}_{t-1}(\omega_{t-1})) = \\
&\qquad = \sum_{\omega_t \in A_t} \mathsf{P}(\boldsymbol{A}_t(\omega_t) | \boldsymbol{A}_{t-1}(\omega_{t-1})) = \\
&\qquad = \mathsf{P}\Bigg(\bigcup_{\omega_t \in A_t} \boldsymbol{A}_t(\omega_t) | \boldsymbol{A}_{t-1}(\omega_{t-1}) \Bigg) = \mathsf{P}(\boldsymbol{A}_t | \boldsymbol{A}_{t-1}(\omega_{t-1})).
\end{aligned}$$

Несколько иначе устанавливается более сильное обобщение равенства (5.2), а именно, соотношение следующего вида:

$$\mathsf{P}(\boldsymbol{A}_t | \boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_{t-2}, \boldsymbol{A}_{t-1}(\omega_{t-1})) = \mathsf{P}(\boldsymbol{A}_t | \boldsymbol{A}_{t-1}(\omega_{t-1})). \tag{5.5}$$

Действительно, имеем очевидные равенства

$$\begin{aligned}
&\mathsf{P}(\boldsymbol{A}_t | \boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_{t-2}, \boldsymbol{A}_{t-1}(\omega_{t-1})) = \\
&\qquad = \mathsf{P}\Bigg(\bigcup_{\omega_t \in A_t} \boldsymbol{A}_t(\omega_t) | \boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_{t-2}, \boldsymbol{A}_{t-1}(\omega_{t-1}) \Bigg) = \\
&\qquad = \sum_{\omega_t \in A_t} \mathsf{P}(\boldsymbol{A}_t(\omega_t) | \boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_{t-2}, \boldsymbol{A}_{t-1}(\omega_{t-1})) = \\
&\qquad = \sum_{\omega_t \in A_t} \frac{\mathsf{P}(\boldsymbol{A}_1 \cap \boldsymbol{A}_2 \cap \ldots \cap \boldsymbol{A}_{t-2} \cap \boldsymbol{A}_{t-1}(\omega_{t-1}) \cap \boldsymbol{A}_t(\omega_t))}{\mathsf{P}(\boldsymbol{A}_1 \cap \boldsymbol{A}_2 \cap \ldots \cap \boldsymbol{A}_{t-2} \cap \boldsymbol{A}_{t-1}(\omega_{t-1}))}.
\end{aligned} \tag{5.6}$$

Пусть

$$\begin{aligned}
\boldsymbol{A} = \boldsymbol{A}_1 \cap \boldsymbol{A}_2 \cap \ldots \cap \boldsymbol{A}_{t-2} \cap \boldsymbol{A}_{t-1}(\omega_{t-1}) \cap \boldsymbol{A}_t(\omega_t) = \\
= A_1 \times A_2 \times \ldots \times A_{t-2} \times \{\omega_{t-1}\} \times \{\omega_t\} \times \Omega_{t+1} \times \ldots \times \Omega_n.
\end{aligned}$$

Так как при каждом $t = 1, 2, \ldots, n$ выполняются очевидные равенства

$$\bigcup_{\omega_t \in \Omega_t} \boldsymbol{A}_t(\omega_t) = \Omega_1^n, \quad \mathsf{P}\Big(\bigcup_{\omega_t \in \Omega_t} \boldsymbol{A}_t(\omega_t) \Big) = \sum_{\omega_t \in \Omega_t} \mathsf{P}(\boldsymbol{A}_t(\omega_t)) = 1,$$

то для условной вероятности $\mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1}))$ при $t = 2, 3, \ldots, n$ и $\omega_{t-1} \in \{s_1, s_2, \ldots, s_r\}$ имеем

$$\sum_{\omega_t \in \Omega_t} \mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1})) = \\ = \mathsf{P}\Bigg(\bigcup_{\omega_t \in \Omega_t} \boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1})\Bigg) = \mathsf{P}\big(\Omega_1^n|\boldsymbol{A}_{t-1}(\omega_{t-1})\big) = 1.$$

Отсюда следует соотношение

$$\mathsf{P}(\boldsymbol{A}_1 \cap \boldsymbol{A}_2 \cap \ldots \cap \boldsymbol{A}_{t-2} \cap \boldsymbol{A}_{t-1}(\omega_{t-1}) \cap \boldsymbol{A}_t(\omega_t)) = \mathsf{P}(\boldsymbol{A}) = \sum_{\omega \in \boldsymbol{A}} \mathsf{P}(\{\omega\}) = \\ = \sum_{\omega \in \boldsymbol{A}} \mathsf{P}(\boldsymbol{A}_1(\omega_1)) \times \mathsf{P}(\boldsymbol{A}_2(\omega_2)|\boldsymbol{A}_1(\omega_1)) \times \ldots \times \mathsf{P}(\boldsymbol{A}_{t-2}(\omega_{t-2})|\boldsymbol{A}_{t-3}(\omega_{t-3})) \times \\ \times \mathsf{P}(\boldsymbol{A}_{t-1}(\omega_{t-1})|\boldsymbol{A}_{t-2}(\omega_{t-2})) \times \mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1})) \times \\ \times \mathsf{P}(\boldsymbol{A}_{t+1}(\omega_{t+1})|\boldsymbol{A}_t(\omega_t)) \times \ldots \times \mathsf{P}(\boldsymbol{A}_n(\omega_n)\ |\boldsymbol{A}_{n-1}(\omega_{n-1})) = \\ = \sum_{\omega_1 \in A_1, \omega_2 \in A_2, \ldots, \omega_{t-2} \in A_{t-2}} \mathsf{P}(\boldsymbol{A}_1(\omega_1)) \times \mathsf{P}(\boldsymbol{A}_2(\omega_2)\ |\boldsymbol{A}_1(\omega_1)) \times \ldots \times \\ \times \mathsf{P}(\boldsymbol{A}_{t-2}(\omega_{t-2})|\boldsymbol{A}_{t-3}(\omega_{t-3})) \times \mathsf{P}(\boldsymbol{A}_{t-1}(\omega_{t-1})|\boldsymbol{A}_{t-2}(\omega_{t-2})) \times \\ \times \mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1})) = \mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1})) \times \\ \times \sum_{\omega_1 \in A_1, \omega_2 \in A_2, \ldots, \omega_{t-2} \in A_{t-2}} \mathsf{P}(\boldsymbol{A}_1(\omega_1)) \times \mathsf{P}(\boldsymbol{A}_2(\omega_2)|\boldsymbol{A}_1(\omega_1)) \times \ldots \times \\ \times \mathsf{P}(\boldsymbol{A}_{t-1}(\omega_{t-1})|\boldsymbol{A}_{t-2}(\omega_{t-2})).$$

Поступая аналогичным образом, получим равенство вида:

$$\mathsf{P}(\boldsymbol{A}_1 \cap \boldsymbol{A}_2 \cap \ldots \cap \boldsymbol{A}_{t-2} \cap \boldsymbol{A}_{t-1}(\omega_{t-1})) = \\ = \sum_{\omega_1 \in A_1, \omega_2 \in A_2, \ldots, \omega_{t-2} \in A_{t-2}} \mathsf{P}(\boldsymbol{A}_1(\omega_1)) \times \mathsf{P}(\boldsymbol{A}_2(\omega_2)|\boldsymbol{A}_1(\omega_1)) \times \ldots \times \\ \times \mathsf{P}(\boldsymbol{A}_{t-1}(\omega_{t-1})|\boldsymbol{A}_{t-2}(\omega_{t-2})).$$

Подставляя в соотношение (5.6) найденные значения для вероятностей

$$\mathsf{P}(\boldsymbol{A}_1 \cap \boldsymbol{A}_2 \cap \ldots \cap \boldsymbol{A}_{t-2} \cap \boldsymbol{A}_{t-1}(\omega_{t-1})) \cap \boldsymbol{A}_t(\omega_t), \\ \mathsf{P}(\boldsymbol{A}_1 \cap \boldsymbol{A}_2 \cap \ldots \cap \boldsymbol{A}_{t-2} \cap \boldsymbol{A}_{t-1}(\omega_{t-1})),$$

выводим

$$\mathsf{P}(\boldsymbol{A}_t|\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_{t-2}, \boldsymbol{A}_{t-1}(\omega_{t-1})) = \\ = \sum_{\omega_t \in A_t} \mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1})) = \mathsf{P}(\boldsymbol{A}_t|\boldsymbol{A}_{t-1}(\omega_{t-1})),$$

что и доказывает равенство (5.5). Однако если событие $\boldsymbol{A}_{t-1}(\omega_{t-1})$ в равенстве (5.5) заменить общим событием:

$$\boldsymbol{A}_{t-1} = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{t-2} \times A_{t-1} \times \Omega_t \times \ldots \times \Omega_n,$$

где $\boldsymbol{A}_{t-1} \in \mathcal{F}_{t-1}$ и $\boldsymbol{A}_{t-1} \neq \{\omega_{t-1}\}$, то в этом случае равенство (5.5) уже остается справедливым не всегда. Соотношения (5.2), (5.4) и (5.5) называются *условиями марковости эволюционного эксперимента* E.

1.3. Конструктивное задание однородной схемы Маркова. Для однородной схемы Маркова при $2 \leqslant t \leqslant n$, $\omega_{t-1} = s_i$ и $\omega_t = s_j$ условные вероятности $\mathsf{P}(\boldsymbol{A}_t(s_j)|\boldsymbol{A}_{t-1}(s_i))$, называются *вероятностями перехода* от события $\{s_i\}$ в $(t-1)$-м испытании к событию $\{s_j\}$ в t-м испытании и обозначаются через $p_{i,j}$, где $1 \leqslant i, j \leqslant r$. Вероятности $\mathsf{P}(\boldsymbol{A}_1(\omega_1))$, $\omega_1 = s_1, s_2, \ldots, s_r$, называются *начальным распределением схемы Маркова*. В дальнейшем при $\omega_1 = s_i$ так называемую начальную вероятность $\mathsf{P}(\boldsymbol{A}_1(s_i))$ будем обозначать через p_i, где $1 \leqslant i \leqslant r$. Рассуждения предыдущего абзаца показывают, что начальные и переходные вероятности схемы Маркова при любых $i, j = 1, 2, \ldots, r$ удовлетворяют следующим очевидным соотношениям:

$$p_i \geqslant 0, \quad p_{i,j} \geqslant 0, \quad \sum_{i=1}^{r} p_i = 1, \quad \sum_{j=1}^{r} p_{i,j} = 1. \tag{5.7}$$

Покажем, что множества $\{p_i : i = 1, 2, \ldots, r\}$, $\{p_{i,j} : i, j = 1, 2, \ldots, r\}$, элементы которых удовлетворяют условиям (5.7), могут быть использованы для превращения последовательности конечного числа n зависимых испытаний E_1, $E_2, \ldots$, E_n в схему Маркова. Для этого на теоретико-множественной модели $(\Omega_1^n, \mathcal{F}_1^n)$ эволюционного эксперимента $E = \{E_1, E_2, \ldots, E_n\}$ определим функцию $\mathsf{P}(\cdot)$ или его распределение:

$$\begin{gathered}\mathsf{P}(\{\omega\}) = p_{i_1} \times p_{i_1, i_2} \times p_{i_2, i_3} \times \ldots \times p_{i_{n-1}, i_n}, \\ \omega = (\omega_1, \omega_2, \ldots, \omega_n) = (s_{i_1}, s_{i_2}, \ldots s_{i_n}) \in \Omega_1^n, \\ \mathsf{P}(\boldsymbol{A}) = \sum_{\omega \in \boldsymbol{A}} \mathsf{P}(\{\omega\}), \quad \boldsymbol{A} \in \mathcal{F}_1^n.\end{gathered} \tag{5.8}$$

Из (5.7) легко находим

$$\mathsf{P}(\{\omega\}) = p_{i_1} \times p_{i_1, i_2} \times p_{i_2, i_3} \times \ldots \times p_{i_{n-1}, i_n} \geqslant 0,$$

$$\sum_{\omega \in \Omega_1^n} \mathsf{P}(\{\omega\}) = \sum_{i_1=1}^{r} \sum_{i_2=1}^{r} \ldots \sum_{i_n=1}^{r} p_{i_1} \times p_{i_1, i_2} \times p_{i_2, i_3} \times \ldots \times p_{i_{n-1}, i_n} = 1,$$

т. е. $\mathsf{P}(\boldsymbol{A})$, $\boldsymbol{A} \in \mathcal{F}_1^n$, является вероятностной функцией. Эта функция удовлетворяет условию марковости (5.2). На самом деле, из равенств (5.7) и (5.8) при $\mathsf{P}(\boldsymbol{A}_1(\omega_1) \cap \boldsymbol{A}_2(\omega_2)) \cap \ldots \cap \boldsymbol{A}_{t-1}(\omega_{t-1})) > 0$ имеем

$$\mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_1(\omega_1), \boldsymbol{A}_2(\omega_2), \ldots, \boldsymbol{A}_{t-1}(\omega_{t-1})) =$$

$$= \frac{\mathsf{P}(\boldsymbol{A}_1(\omega_1) \cap \boldsymbol{A}_2(\omega_2) \cap \ldots \boldsymbol{A}_{t-1}(\omega_{t-1}) \cap \boldsymbol{A}_t(\omega_t))}{\mathsf{P}(\boldsymbol{A}_1(\omega_1) \cap \boldsymbol{A}_2(\omega_2) \cap \ldots \boldsymbol{A}_{t-1}(\omega_{t-1}))} =$$

$$= \frac{\mathsf{P}(\{\omega_1\} \times \{\omega_2\} \times \ldots \times \{\omega_{t-1}\} \times \{\omega_t\} \times \Omega_{t+1} \times \ldots \times \Omega_n)}{\mathsf{P}(\{\omega_1\} \times \{\omega_2\} \times \ldots \times \{\omega_{t-1}\} \times \Omega_t \times \ldots \times \Omega_n)} =$$

$$= \frac{\sum\limits_{i_{t+1}=1}^{r} \ldots \sum\limits_{i_n=1}^{r} \mathsf{P}(\{(s_{i_1}, s_{i_2}, \ldots, s_{i_{t-1}}, s_{i_t}, s_{i_{t+1}}, \ldots, s_{i_n})\})}{\sum\limits_{i_t=1}^{r} \ldots \sum\limits_{i_n=1}^{r} \mathsf{P}(\{(s_{i_1}, s_{i_2}, \ldots, s_{i_{t-1}}, s_{i_t}, \ldots, s_{i_n})\})} =$$

$$= \frac{\sum\limits_{i_{t+1}=1}^{r} \ldots \sum\limits_{i_n=1}^{r} p_{i_1} \times p_{i_1,i_2} \times \ldots \times p_{i_{t-2},i_{t-1}} \times p_{i_{t-1},i_t} \times p_{i_t,i_{t+1}} \times \ldots \times p_{i_{n-1},i_n}}{\sum\limits_{i_t=1}^{r} \ldots \sum\limits_{i_n=1}^{r} p_{i_1} \times p_{i_1,i_2} \times \ldots \times p_{i_{t-2},i_{t-1}} \times p_{i_{t-1},i_t} \times p_{i_t,i_{t+1}} \times \ldots \times p_{i_{n-1},i_n}} =$$

$$= p_{i_{t-1},i_t} \times \frac{p_{i_1} \times p_{i_1,i_2} \ldots \times p_{i_{t-2},i_{t-1}}}{p_{i_1} \times p_{i_1,i_2} \ldots \times p_{i_{t-2},i_{t-1}}} = p_{i_{t-1},i_t},$$

$$\mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1})) = \frac{\mathsf{P}(\boldsymbol{A}_{t-1}(\omega_{t-1}) \cap \boldsymbol{A}_t(\omega_t))}{\mathsf{P}(\boldsymbol{A}_{t-1}(\omega_{t-1}))} =$$

$$= \frac{\mathsf{P}(\Omega_1 \times \Omega_2 \times \ldots \times \Omega_{t-2} \times \omega_{t-1} \times \omega_t \times \Omega_{t+1} \times \ldots \times \Omega_n)}{\mathsf{P}(\Omega_1 \times \Omega_2 \times \ldots \times \Omega_{t-2} \times \omega_{t-1} \times \Omega_t \times \Omega_{t+1} \times \ldots \times \Omega_n)} =$$

$$= p_{i_{t-1},i_t} \frac{\sum\limits_{i_1=1}^{r} \ldots \sum\limits_{i_{t-2}=1}^{r} p_{i_1} \times p_{i_1,i_2} \times \ldots \times p_{i_{t-2},i_{t-1}}}{\sum\limits_{i_1=1}^{r} \ldots \sum\limits_{i_{t-2}=1}^{r} p_{i_1} \times p_{i_1,i_2} \times \ldots \times p_{i_{t-2},i_{t-1}}} = p_{i_{t-1},i_t}.$$

Отсюда $\mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_1(\omega_1), \boldsymbol{A}_2(\omega_2), \ldots, \boldsymbol{A}_{t-1}(\omega_{t-1})) = \mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1}))$ и соотношение (5.2) выполняется. Таким образом, получаем удобный способ определения конкретной схемы Маркова с помощью множества $\Omega_1 = \{s_1, s_2, \ldots, s_r\}$, начального распределения $\{p_i \colon i = 1, 2, \ldots, r\}$ и матрицы из вероятностей перехода за одно испытание вида $\Pi = \|p_{i,j}\|_{i,j=1,2,\ldots,r}$. Матрица перехода Π, элементы которой удовлетворяют неравенству $p_{ij} \geqslant 0$ при каждом $i, j = 1, 2, \ldots, r$ и равенству $\sum_{j=1}^{r} p_{ij} = 1$ при любом $i = 1, 2, \ldots, r$, называется *стохастической*. Особо важно здесь подчеркнуть, что в теории вероятностей, как правило, рассматривают неконкретную схему Маркова, а целый класс. Этот класс схем Маркова определяется множеством $\Omega_1 = \{s_1, s_2, \ldots, s_r\}$, матрицей Π и некоторым семейством начальных распределений $\{p_i \colon i = 1, 2, \ldots, r\}$.

1.4. Уравнение Колмогорова–Чепмена. Вероятность любого события $\boldsymbol{A}$ в схеме Маркова можем найти с помощью формулы (5.8). Однако вероятности для некоторых событий $\boldsymbol{A} \in \mathcal{F}_1^n$, которыми наиболее часто интересуются на практике, легко вычислить с использованием формулы полной вероятности. Покажем это на примере событий $\boldsymbol{A}_t(s_j)$, $j = 1, 2, \ldots, r$, при любом $t = 2, 3, \ldots, n$. Вероятность $\mathsf{P}(\boldsymbol{A}_t(s_j))$ того, что в t-м испытании схемы Маркова наступит событие $\{s_j\}$, будем обозначать через $p_j^{(t)}$. В этом случае, естественно, полагаем $p_j^{(1)} = p_j$ для всех $j = 1, 2, \ldots, r$. В силу соотношений $\cup_{\omega_t \in \Omega_t} \boldsymbol{A}_t(\omega_t) = \Omega_1^n$, $\boldsymbol{A}_t(s_j) \cap \boldsymbol{A}_t(s_i) = \varnothing$, где $t = 1, 2, \ldots, n$ и $j \neq i$, по формуле полной вероятности имеем

$$p_j^{(2)} = \sum_{i=1}^{r} p_i^{(1)} p_{i,j}, \quad p_j^{(3)} = \sum_{i=1}^{r} p_i^{(2)} p_{i,j}, \ldots, p_j^{(n)} = \sum_{i=1}^{r} p_i^{(n-1)} p_{i,j}. \qquad (5.9)$$

В дальнейшем будем считать известными из курса алгебры основные операции над матрицами: сложение матриц, умножение матриц, умножение матрицы на число, транспонирование матрицы и взятие предела последовательности матриц. Пусть вектор

$$\mathbf{p}^{(t)} = (p_1^{(t)}, p_2^{(t)}, \ldots, p_r^{(t)}),$$

тогда равенства (5.9) можно записать в матричной форме:

$$\mathbf{p}^{(2)} = \mathbf{p}^{(1)} \times \Pi, \quad \mathbf{p}^{(3)} = \mathbf{p}^{(2)} \times \Pi, \ldots, \quad \mathbf{p}^{(n)} = \mathbf{p}^{(n-1)} \times \Pi.$$

Выражая эти равенства через так называемое начальное распределение $(p_1, p_2, \ldots, p_r)$, матрицу Π и учитывая обозначение

$$\mathbf{p}^{(1)} = (p_1, p_2, \ldots, p_r) = \mathbf{p},$$

окончательно находим вектор

$$\mathbf{p}^{(t)} = \mathbf{p}^{(1)} \times \Pi^{t-1} = \mathbf{p}\Pi^{t-1},$$

который называется *абсолютным распределением вероятностей схемы Маркова в момент* $t = 2, 3, \ldots, n$. Например, при $t = 3$ имеем

$$\mathbf{p}^{(3)} = \mathbf{p}^{(2)} \times \Pi = \mathbf{p}^{(1)} \times \Pi \times \Pi = \mathbf{p}\Pi^2.$$

Выясним вероятностный смысл элементов матрицы Π^t при $t = 2, 3, \ldots, n-1$. С этой целью при $2 \leqslant t = h + u \leqslant n$ рассмотрим условную вероятность $\mathsf{P}(\boldsymbol{A}_{h+u}(s_j)|\boldsymbol{A}_h(s_i))$. Эту вероятность будем называть переходной вероятностью схемы Маркова за u шагов от события $\{s_i\}$ в h-м испытании к событию $\{s_j\}$ в $(h+u)$-м испытании и обозначать через $p_{i,j}(h, u)$, где $i, j = 1, 2, \ldots, r$. Методом индукции по u докажем, что для однородной схемы Маркова $p_{i,j}(h, u)$ не зависит от h и является элементом в i-й строчке и j-м столбце u-й степени матрицы Π. Поэтому эту вероятность целесообразно в дальнейшем обозначать через $p_{i,j}(u)$, которая естественно равна $p_{i,j}$ при $u = 1$. Итак, при $u = 2$ по формуле полной вероятности, используя свойства марковости (5.2) и однородности схемы Маркова, для условной вероятности $p_{i,j}(h,2) = \mathsf{P}(\boldsymbol{A}_{h+2}(s_j)|\boldsymbol{A}_h(s_i))$ находим равенства:

$$\begin{aligned}
&\mathsf{P}(\boldsymbol{A}_{h+2}(s_j)|\boldsymbol{A}_h(s_i)) = \\
&\qquad = \sum_{k=1}^{r} \mathsf{P}(\boldsymbol{A}_{h+1}(s_k)|\boldsymbol{A}_h(s_i))\,\mathsf{P}(\boldsymbol{A}_{h+2}(s_j)|\boldsymbol{A}_h(s_i), \boldsymbol{A}_{h+1}(s_k)) = \\
&\qquad = \sum_{k=1}^{r} \mathbf{P}(\boldsymbol{A}_{h+1}(s_k)|\boldsymbol{A}_h(s_i))\,\mathsf{P}(\boldsymbol{A}_{h+2}(s_j)|\boldsymbol{A}_{h+1}(s_k)) = \sum_{k=1}^{r} p_{i,\,k}p_{k,j}.
\end{aligned}$$

Следовательно, условная вероятность $p_{i,j}(h,2)$ за два шага не зависит от h, т. е. $p_{i,j}(h,2) = p_{i,j}(2)$. Если учесть правило умножения матриц, то легко видеть, что $\|p_{i,j}(2)\|_{i,j=1,2,\ldots,r} = \Pi^2$. Пусть $\|p_{i,j}(u-1)\|_{i,j=1,2,\ldots,r} = \Pi^{u-1}$ — переходная матрица схемы Маркова за $u-1$ шагов. Отсюда с помощью тех же рассуждений для $p_{i,j}(h,u) = \mathsf{P}(\boldsymbol{A}_{h+u}(s_j)|\boldsymbol{A}_h(s_i))$ имеем

$$\begin{aligned}
&\mathsf{P}(\boldsymbol{A}_{h+u}(s_j)|\boldsymbol{A}_h(s_i)) = \\
&\qquad = \sum_{k=1}^{r} \mathsf{P}(\boldsymbol{A}_{h+1}(s_k)|\boldsymbol{A}_h(s_i))\,\mathsf{P}(\boldsymbol{A}_{h+u}(s_j)|\boldsymbol{A}_h(s_i), \boldsymbol{A}_{h+1}(s_k)) = \\
&\qquad = \sum_{k=1}^{r} \mathsf{P}(\boldsymbol{A}_{h+1}(s_k)|\boldsymbol{A}_h(s_i))\,\mathsf{P}(\boldsymbol{A}_{h+u}(s_j)|\boldsymbol{A}_{h+1}(s_k)) = \\
&\qquad = \sum_{k=1}^{r} p_{i,\,k}p_{k,j}(u-1) = p_{i,\ j}(u).
\end{aligned}$$

Снова получили величину, которая не зависит от h и равна элементу в i-й строчке и j-м столбце матрицы $\Pi \times \Pi^{u-1} = \Pi^u$. Итак, элементы матрицы Π^t совпадают с вероятностями перехода схемы Маркова за t испытаний. Отсюда при $t = h + u$, используя снова правило перемножения матриц, находим $\Pi^t = \Pi^{h+u} = \Pi^h \times \Pi^u$, или

$$\|p_{i,j}(h+u)\|_{i,j=1,2,\ldots,r} = \|p_{i,j}(h)\|_{i,j=1,2,\ldots,r} \times \|p_{i,j}(u)\|_{i,j=1,2,\ldots,r}.$$

Матричная форма последнего равенства позволяет написать для переходных вероятностей $p_{i,j}(t)$ схемы Маркова знаменитое уравнение Колмогорова–Чепмена:

$$p_{i,j}(h+u) = \sum_{k=1}^{r} p_{i,k}(h)\,p_{k,j}(u). \tag{5.10}$$

Так как $\mathbf{p}^{(t)} = \mathbf{p}\Pi^{t-1}$, то теперь для вероятности события $\boldsymbol{A}_t(s_j)$ или, другими словами, для вероятности $\mathsf{P}(\boldsymbol{A}_t(s_j))$ того, что в t-м испытании схемы Маркова наступит событие $\{s_j\}$, можно записать простую формулу

$$p_j^{(t)} = \sum_{i=1}^{r} p_i p_{i,j}(t-1), \quad t = 2, 3, \ldots, n; \quad j = 1, 2, \ldots, r. \tag{5.11}$$

§ 2. Эргодические и стационарные схемы Маркова

2.1. Свойства абсолютного распределения схемы Маркова. Вычисление по формуле (5.11) абсолютных вероятностей $p_j^{(t)}$ при любых $j = 1, 2, \ldots, r$ и больших значениях числа t вызывает значительные трудности. Поэтому важной задачей является исследование поведения вероятностей $p_j^{(t)}$, $j = 1, 2, \ldots, r$, при неограниченном увеличении t. Во многих случаях решение этой задачи может быть найдено на основе следующей эргодической теоремы.

Теорема 5.1 (*теорема Маркова*). *Если при некотором числе t_0 шагов все вероятности $p_{i,j}(t_0)$ строго положительны, то существуют так называемые предельные вероятности $\pi_j = \lim\limits_{t\to\infty} p_{i,j}(t) = \lim\limits_{t\to\infty} p_j^{(t)}$, $j = 1, 2, \ldots, r$. При этом $\sum_{j=1}^r \pi_j = 1$, $\pi_j > 0$ для всех возможных значений j и, наконец, система уравнений*

$$\sum_{i=1}^r x_j = 1, \quad x_j = \sum_{i=1}^r x_i p_{i,j}, \quad j = 1, 2, \ldots, r, \tag{5.12}$$

имеет единственное решение $x_j = \pi_j$, $j = 1, 2, \ldots, r$.

Доказательство. Определим числа $m_j(t)$, $M_j(t)$, α равенствами: $m_j(t) = \min\limits_i p_{i,j}(t)$, $M_j(t) = \max\limits_i p_{i,j}(t)$, $\alpha = \min\limits_{i,j} p_{i,j}(t_0)$. Принимая во внимание уравнение Колмогорова–Чепмена (5.10), при $h = 1$ и $u = t$ легко находим

$$m_j(t+1) = \min_i p_{i,j}(t+1) =$$
$$= \min_i \sum_{k=1}^r p_{i,k} p_{k,j}(t) \geqslant m_j(t) \times \min_i \sum_{k=1}^r p_{i,k} = m_j(t), M_j(t+1) =$$
$$= \max_i p_{i,j}(t+1) = \max_i \sum_{k=1}^r p_{i,k} p_{k,j}(t) \leqslant M_j(t) \max_i \sum_{k=1}^r p_{i,k} = M_j(t).$$

Значит, последовательность $\{m_j(t);\, t \geqslant 1\}$ возрастает и ограничена сверху, а последовательность $\{M_j(t);\, t \geqslant 1\}$ убывает и ограничена снизу, следовательно, пределы вида $\lim\limits_{t\to\infty} m_j(t)$, $\lim\limits_{t\to\infty} M_j(t)$, $\lim\limits_{t\to\infty} (M_j(t) - m_j(t))$ существуют. Из уравнения Колмогорова–Чепмена (5.10) при $h = t_0$ и $u = t$ следует, что

$$p_{i,j}(t_0 + t) = \sum_{k=1}^r p_{i,k}(t_0)\, p_{k,j}(t) = \sum_{k=1}^r (p_{i,k}(t_0) - \alpha p_{j,k}(t))\, p_{k,j}(t) +$$
$$+ \sum_{k=1}^r \alpha p_{j,k}(t)\, p_{k,j}(t) = \alpha p_{j,j}(2t) + \sum_{k=1}^r (p_{i,k}(t_0) - \alpha p_{j,k}(t))\, p_{k,j}(t).$$

Поскольку $\alpha = \min\limits_{i,j} p_{i,j}(t_0)$ и $0 \leqslant p_{j,\,k}(t) \leqslant 1$, то $p_{i,\,k}(t_0) - \alpha p_{j,\,k}(t) \geqslant 0$ и

$$p_{i,\,j}(t_0+t) \geqslant m_j(t) \times \sum_{k=1}^{r} (p_{i,\,k}(t_0) - \alpha p_{j,\,k}(t)) + \alpha p_{j,j}(2t) = \\ = m_j(t)(1-\alpha) + \alpha p_{j,j}(2t),$$

$$p_{i,j}(t_0+t) \leqslant M_j(t) \sum_{k=1}^{r} (p_{i,\,k}(t_0) - \alpha p_{j,\,k}(t)) + \alpha p_{j,j}(2t) = \\ = M_j(t)(1-\alpha) + \alpha p_{j,j}(2t).$$

Значит,

$$m_j(t_0+t) \geqslant m_j(t)(1-\alpha) + \alpha p_{j,j}(2t)$$

и

$$M_j(t_0+t) \leqslant M_j(t)(1-\alpha) + \alpha p_{j,j}(2t),$$

или

$$0 \leqslant M_j(t_0+t) - m_j(t_0+t) \leqslant (M_j(t) - m_j(t))(1-\alpha).$$

Применяя последнее неравенство последовательно $m = 1, 2, \ldots$ раз, получим, что

$$M_j(mt_0+t) - m_j(mt_0+t) \leqslant [M_j((m-1)\,t_0+t) - m_j((m-1)\,t_0+t)] \times \\ \times (1-\alpha) \leqslant \ldots \leqslant (M_j(t) - m_j(t))(1-\alpha)^m. \quad (5.13)$$

Из условий теоремы $\alpha > 0$. Поэтому

$$\lim_{m\to\infty} (M_j(mt_0+t) - m_j(mt_0+t)) = 0$$

по некоторой подпоследовательности $\{mt_0+t; m = 1, 2, \ldots\}$ последовательности $\{t;\ t = 1, 2, \ldots\}$, а значит,

$$\lim_{t\to\infty} (M_j(t) - m_j(t)) = 0, \quad \lim_{t\to\infty} m_j(t) = \lim_{t\to\infty} M_j(t).$$

Этот предел будем обозначать через π_j. Если $t \geqslant t_0$, то

$$m_j(t) \geqslant m_j(t_0) = \min_i p_{i,j}(t_0) \geqslant \alpha > 0$$

и, следовательно, $\pi_j > 0$. Из определения $m_j(t)$, $M_j(t)$ выводим, что

$$m_j(t) \leqslant p_{i,j}(t) \leqslant M_j(t),$$

откуда, учитывая равенство

$$\lim_{t\to\infty} m_j(t) = \lim_{t\to\infty} M_j(t),$$

получаем $\lim\limits_{t\to\infty} p_{i,j}(t) = \pi_j$ для любого $i = 1, 2, \ldots, r$.

Полагая в уравнении Колмогорова–Чепмена $h = t$, $u = 1$ и $t \to \infty$, мы поэтапно получаем

$$p_{i,j}(t+1) = \sum_{k=1}^{r} p_{i,k}(t)\, p_{k,j},$$

$$\lim_{t\to\infty} p_{i,j}(t+1) = \lim_{t\to\infty} \sum_{k=1}^{r} p_{i,k}(t)\, p_{k,j} = \sum_{k=1}^{r} \lim_{t\to\infty} p_{i,k}(t)\, p_{k,j} = \sum_{k=1}^{r} \pi_k p_{k,j},$$

$$\pi_j = \sum_{k=1}^{r} \pi_k p_{k,j}, \quad j = 1, 2, \ldots, r.$$

Из равенства $\sum_{k=1}^{r} p_{i,k}(t) = 1$ непосредственно следует, что

$$\lim_{t\to\infty} \sum_{k=1}^{r} p_{i,k}(t) = 1, \quad \sum_{k=1}^{r} \lim_{t\to\infty} p_{i,k}(t) = 1, \quad \sum_{k=1}^{r} \pi_k = 1.$$

Если в равенстве (5.11) принять $t \to \infty$, то находим

$$\lim_{t\to\infty} p_j^{(t)} = \lim_{t\to\infty} \sum_{i=1}^{r} p_i p_{i,j}(t-1) = \sum_{i=1}^{r} p_i \lim_{t\to\infty} p_{i,j}(t-1) = \\ = \sum_{i=1}^{r} p_i \pi_j = \pi_j \sum_{i=1}^{r} p_i = \pi_j.$$

Наконец, установим, что система (5.12) имеет единственное решение. Пусть $\{\lambda_j \colon j = 1, 2, \ldots, r\}$ есть какое-нибудь ее решение. Тогда при $t = 1$ имеем

$$\lambda_j = \sum_{i=1}^{r} \lambda_i p_{i,j} = \sum_{i=1}^{r} \lambda_i p_{i,j}(1), \quad j = 1, 2, \ldots, r,$$

Предположим, что эти равенства справедливы при некотором t. Применяя опять уравнение Колмогорова–Чепмена и изменяя порядок суммирования, для $t+1$ легко найдем, что

$$\sum_{i=1}^{r} \lambda_i p_{i,j}(t+1) = \sum_{i=1}^{r} \lambda_i \sum_{k=1}^{r} p_{i,k} p_{k,j}(t) = \\ = \sum_{k=1}^{r} \Big(\sum_{i=1}^{r} \lambda_i p_{i,k} \Big) p_{k,j}(t) = \sum_{k=1}^{r} \lambda_k p_{k,j}(t) = \lambda_j.$$

Итак, для $t \geqslant 1$ по индукции установлены равенства

$$\lambda_j = \sum_{i=1}^{r} \lambda_i p_{i,j}(t), \quad j = 1, 2, \ldots, r. \tag{5.14}$$

Переходя в этих равенствах к пределу по $t \to \infty$, сразу находим

$$\lambda_j = \lim_{t\to\infty} \sum_{i=1}^{r} \lambda_i p_{i,j}(t) = \sum_{i=1}^{r} \lambda_i \lim_{t\to\infty} p_{i,j}(t) = \sum_{i=1}^{r} \lambda_i \pi_j = \pi_j$$

при каждом $j = 1, 2, \ldots, r$. Теорема А. А. Маркова установлена.

Из теоремы А. А. Маркова выводим следующий замечательный факт. Предельные вероятности $\pi_1 > 0$, $\pi_2 > 0, \ldots$, $\pi_r > 0$ не зависят от начального распределения p_1, $p_2, \ldots$, p_r и могут быть найдены путем решения системы линейных уравнений (5.12). Такого рода схемы Маркова называются *эргодическими*. Если схема Маркова является эргодической, то при некотором натуральном числе t_0 имеет место неравенство $\min\limits_{i,j} p_{i,j}(t_0) > 0$, поскольку $\lim\limits_{t\to\infty} p_{i,j}(t) = \pi_j > 0$ для конечного набора величин $i, j = 1, 2, \ldots, r$.

Отметим еще одно важное свойство. Для эргодической схемы Маркова при $t \geqslant t_0$ и каждом $j = 1, 2, \ldots, r$ сходимость абсолютной вероятности $p_j^{(t)}$ или переходной вероятности $p_{i,j}(t)$ к предельному (эргодическому) значению π_j происходит с геометрической скоростью. Вот простое доказательство этого утверждения. Представим натуральное число $t \geqslant t_0$ в таком виде: $t = m\,t_0 + t_1$, где $m = 1, 2, \ldots$ и $t_1 = 0, 1, \ldots, t_0 - 1$. Тогда из соотношений (5.11), (5.13) и (5.14) при $\lambda_j = \pi_j$, $j = 1, 2, \ldots, r$, находим

$$\begin{aligned}
|p_j^{(t+1)} - \pi_j| &= \Big|\sum_{i=1}^{r} p_i p_{i,j}(t) - \sum_{i=1}^{r} \pi_i p_{i,j}(t)\Big| \leqslant \\
&\leqslant \Big|M_j(t)\sum_{i=1}^{r} p_i - m_j(t)\sum_{i=1}^{r} \pi_i\Big| = |M_j(t) - m_j(t)| = \\
&= |M_j((m-1)\,t_0 + t_0 + t_1) - m_j((m-1)\,t_0 + t_0 + t_1)| \leqslant \\
&\leqslant (M_j(t_0 + t_1) - m_j(t_0 + t_1))(1-\alpha)^{m-1} \leqslant (1-\alpha)^{m-1}.
\end{aligned}$$

Рассмотрим пример, который иллюстрирует теорему Маркова и результат скорости сходимости переходной вероятности $p_{i,j}(t)$.

Пример 5.1 (*задача о процессах протекания болезни и выздоровления*). Некоторая страховая компания еженедельно контролирует здоровье мужчин в возрасте от 60 до 70 лет, проживающих в большом городе с населением около одного миллиона человек. Представим себе, что произвольный мужчина в таком возрасте на момент контроля может находиться в одном из трех альтернативных состояний (элементарных исходов): состояние $\{s_1\}$ означает быть здоровым, состояние $\{s_2\}$ означает недомогание и, наконец, состояние $\{s_3\}$ означает быть больным. Есть основание считать, что эволюционные процессы протекания болезни и выздоровления человека обладают марковским свойством. Такое предположение объясняется тем, что границы места проживания, пола и возраста человека строго очерчены. Эмпирическим путем были получены оценки вероятностей переходов схемы Маркова, которые сведены в матрицу вида

$$\Pi = \begin{pmatrix} 0{,}5 & 0{,}25 & 0{,}25 \\ 0{,}5 & 0 & 0{,}5 \\ 0{,}25 & 0{,}25 & 0{,}5 \end{pmatrix}.$$

Матрица вероятностей переходов за две недели протекания процесса болезни или выздоровления человека после начальной стадии наблюдения равна

$$\Pi^2 = \begin{pmatrix} 0{,}4375 & 0{,}1875 & 0{,}3750 \\ 0{,}3750 & 0{,}2500 & 0{,}3750 \\ 0{,}3750 & 0{,}1875 & 0{,}4375 \end{pmatrix}.$$

Так как при $t_0 = 2$ все элементы матрицы Π^2 строго положительны, то по теореме Маркова эргодическое распределение (π_1, π_2, π_3) определяется из решения системы: $\pi_1 + \pi_2 + \pi_3 = 1$, $\pi_1/2 + \pi_2/2 + \pi_3/4 = \pi_1$, $\pi_1/4 + \pi_3/4 = \pi_2$, $\pi_1/4 + \pi_2/2 + \pi_3/2 = \pi_3$. Эта система имеет единственное решение вида $(\pi_1, \pi_2, \pi_3) = (2/5, 1/5, 2/5)$. Доказано, что скорость сходимости абсолютной вероятности $p_j^{(t)}$ или переходной вероятности $p_{i,j}(t)$ к предельному распределению является геометрической. Поэтому даже при малом значении $t = 5$ строчки матрицы вида

$$\Pi^5 = \begin{pmatrix} 0{,}4004 & 0{,}2002 & 0{,}3994 \\ 0{,}4004 & 0{,}1992 & 0{,}4004 \\ 0{,}3994 & 0{,}2002 & 0{,}4004 \end{pmatrix}$$

и эргодическое распределение $(2/5, 1/5, 2/5)$ оказываются близки. Это означает, что сегодняшнее состояние здоровья пожилого мужчины оказывает небольшое влияние на прогноз его здоровья уже через пять недель.

Приведем пример схемы Маркова, для которой существует единственное предельное распределение и отсутствует эргодическое.

Пример 5.2. Пусть $r = 3$, $\Omega_1 = \{s_1, s_2, s_3\}$ и матрица

$$\Pi = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Тогда при любом $t = 2, 3, \ldots$ имеем

$$\Pi^t = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix},$$

и, следовательно,

$$\begin{aligned} \lim_{t\to\infty} p_{1,1}(t) = \lim_{t\to\infty} p_{2,1}(t) = \lim_{t\to\infty} p_{3,1}(t) = \lim_{t\to\infty} p_{1,2}(t) = \\ = \lim_{t\to\infty} p_{2,2}(t) = \lim_{t\to\infty} p_{3,2}(t) = 1/2, \\ \lim_{t\to\infty} p_{1,3}(t) = \lim_{t\to\infty} p_{2,3}(t) = \lim_{t\to\infty} p_{3,3}(t) = 0. \end{aligned}$$

Предельные вероятности $\pi_1 = 1/2$, $\pi_2 = 1/2$, $\pi_3 = 0$ не зависят от начального распределения и удовлетворяют следующей системе уравнений: $\pi_1 + \pi_2 + \pi_3 = 1$, $\pi_1 = (\pi_1 + \pi_2)/2 + \pi_3$, $\pi_2 = (\pi_1 + \pi_2)/2$, $\pi_3 = 0$. Однако предельное распределение $(1/2, 1/2, 0)$ не будет эргодическим, так как $\pi_3 = 0$.

2.2. Стационарное распределение схемы Маркова. Рассмотрим теперь однородную схему Маркова произвольного типа, для которой система уравнений (5.12) имеет неотрицательное решение $\lambda_1,\ \lambda_2, \ldots,\ \lambda_r$. Выберем эти числа в качестве начального распределения схемы Маркова, т. е. $p_i = \lambda_i$ для всех $i = 1, 2, \ldots, r$. Тогда для $j = 1, 2, \ldots, r$ имеем соотношения

$$p_j^{(2)} = \sum_{i=1}^{r} p_i p_{i,j} = p_j,$$
$$p_j^{(3)} = \sum_{i=1}^{r} p_i^{(2)} p_{i,j} = \sum_{i=1}^{r} p_i p_{i,j} = p_j, \ldots,$$
$$p_j^{(n)} = \sum_{i=1}^{r} p_i^{(n-1)} p_{i,j} = \sum_{i=1}^{r} p_i p_{i,j} = p_j.$$

Эти соотношения показывают, что абсолютная вероятность $p_j^{(t)}$ не зависит от номера испытания. В этом случае схема Маркова называется стационарной. Всякое неотрицательное решение системы (5.12) определяет так называемое стационарное распределение вероятностей схемы Маркова. Так как предельное распределение $\pi_1,\ \pi_2, \ldots,\ \pi_r$ удовлетворяет системе (5.12), то оно естественно является стационарным. Однако обратное утверждение не всегда имеет место. Для доказательства сказанного рассмотрим простой пример схемы Маркова.

Пример 5.3. Допустим, что каждое испытание однородной схемы Маркова имеет три элементарных исхода, т. е. $r = 3$, $\Omega_1 = \{s_1, s_2, s_3\}$ и матрица вероятностей перехода за один шаг

$$\Pi = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

В этом случае при $t \geqslant 1$ находим, что

$$\Pi^{2t-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}, \qquad \Pi^{2t} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

Отсюда видно, что пределы $\lim\limits_{t\to\infty} p_{1,1}(t)$, $\lim\limits_{t\to\infty} p_{1,2}(t)$, $\lim\limits_{t\to\infty} p_{2,1}(t)$, $\lim\limits_{t\to\infty} p_{2,2}(t)$ не существуют и $\lim\limits_{t\to\infty} p_{1,3}(t) = \lim\limits_{t\to\infty} p_{2,3}(t) = \lim\limits_{t\to\infty} p_{3,3}(t) = 0$, $\lim\limits_{t\to\infty} p_{3,1}(t) = \lim\limits_{t\to\infty} p_{3,2}(t) = 1/2$. Следовательно, для этой схемы Маркова предельное распределение, а тем более эргодическое, не существует. Однако величины $\lambda_1 = \lambda_2 = 1/2$, $\lambda_3 = 0$ являются единственным решением системы вида: $\lambda_1 + \lambda_2 + \lambda_3 = 1$, $\lambda_1 = \lambda_2 + \lambda_3/2$, $\lambda_2 = \lambda_1 + \lambda_3/2$, $\lambda_3 = 0$, которая непосредственно получается из системы (5.12) при $r = 3$, $p_{1,1} = p_{1,3} = p_{2,2} = p_{2,3} = p_{3,3} = 0$, $p_{1,2} = p_{2,1} = 1$,

$p_{3,1} = p_{3,2} = 1/2$. Итак, эта схема Маркова имеет единственное стационарное распределение $(1/2, 1/2, 0)$. В общем случае, имеет место следующее удивительное свойство.

Теорема 5.2 (теорема существования стационарного распределения). *Однородная схема Маркова с конечным множеством элементарных исходов* $\{s_1\}$, $\{s_2\}, \ldots$, $\{s_r\}$ *всегда имеет хотя бы одно стационарное распределение.*

Доказательство. Из определения стохастической матрицы $\Pi = \|p_{i,j}\| = \|p_{i,j}\|_{i,j=1,2,\ldots,r}$ следует, что система линейных уравнений

$$y_i = \sum_{j=1}^{r} y_j p_{i,j}, \quad i = 1, 2, \ldots, r,$$

имеет решение $y_1 = y_2 = \ldots = y_r = 1$. Следовательно, вектор $(1, 1, \ldots, 1)$ размерности r будет собственным для матрицы Π с характеристическим числом $\chi = 1$. Известно, что модуль любого характеристического числа χ матрицы Π не превышает величины $\max\limits_{i} \sum_{j=1}^{r} p_{i,j} = 1$. Теперь можно сказать, что является максимальным характеристическим числом для матрицы Π. Пусть матрица Π' получается из матрицы Π транспонированием и $\mathbf{E}$ является единичной матрицей размерности r. Тогда определители $|\Pi - \chi\mathbf{E}|$ и $|\Pi' - \chi\mathbf{E}|$ будут равны. Отсюда получаем, что характеристические числа матриц Π и Π' одни и те же. Поэтому $\chi = 1$ также будет максимальным характеристическим числом для матрицы Π'. Из спектральных свойств произвольной матрицы с неотрицательными элементами (*Гантмахер Ф. Р.* Теория матриц. — М.: ФИЗМАТЛИТ, 2009) следует, что максимальному характеристическому числу $\chi = 1$ матрицы Π' соответствует неотрицательный собственный вектор $(y_1, y_2, \ldots, y_r)$, для которого имеем $\sum_{j=1}^{r} y_j > 0$, $y_j = \sum_{i=1}^{r} y_i p_{i,j}$, $j = 1, 2, \ldots, r$. Очевидно, что неотрицательные величины $\lambda_j = y_j \left(\sum_{i=1}^{r} y_i\right)^{-1}$, $j = 1, 2, \ldots, r$, являются решением системы (5.12), следовательно, эти величины можно взять в качестве стационарного распределения схемы Маркова. Теорема доказана.

2.3. Вычисление стационарных распределений схемы Маркова. Имеет место следующее утверждение

Теорема 5.3 (теорема о вычислении стационарных распределений). *Множество всех стационарных распределений любой однородной схемы Маркова с конечным числом элементарных исходов* $\{s_1\}$, $\{s_2\}, \ldots$, $\{s_r\}$ *задается строками квадратной матрицы* $\Lambda = \|\lambda_{i,j}\|$ *порядка* r, *где величина*

$$\lambda_{i,j} = \lim_{t \to \infty} \frac{1}{t} \sum_{m=1}^{t} p_{i,j}(m) \geqslant 0$$

для всех $i, j = 1, 2, \ldots, r$.

Доказательство. Прежде чем приступить к доказательству существования любого элемента $\lambda_{i,j}$ предельной матрицы Λ заметим, что обычный предел $\lim\limits_{t\to\infty} p_{i,j}(t)$ совпадает с пределом Чезаро:

$$\lim_{t\to\infty} \frac{1}{t} \sum_{m=1}^{t} p_{i,j}(m).$$

Следовательно, в условиях теоремы Маркова имеем

$$\lim_{t\to\infty} \frac{1}{t} \sum_{m=1}^{t} p_{i,j}(m) = \lim_{t\to\infty} p_{i,j}(t) = \pi_j > 0, \quad i,j = 1,2,\ldots,r.$$

В этом случае матрица Λ имеет одинаковые строчки вида $(\pi_1, \pi_2, \ldots, \pi_r)$, которые определяют единственное стационарное распределение. Пусть

$$\lambda_{i,j}(t) = \frac{1}{t} \sum_{m=1}^{t} p_{i,j}(m) \text{ при } t \geqslant 1.$$

Тогда для всех чисел $i,j = 1,2,\ldots,r$ и $t = 1,2,\ldots$ имеем $\lambda_{i,j}(t) \geqslant 0$ и

$$\sum_{j=1}^{r} \lambda_{i,j}(t) = \sum_{j=1}^{r} \frac{1}{t} \sum_{m=1}^{t} p_{i,j}(m) = \frac{1}{t} \sum_{m=1}^{t} \sum_{j=1}^{r} p_{i,j}(m) = 1.$$

При любых фиксированных $i,j = 1,2,\ldots,r$ последовательность вида $\{\lambda_{i,j}(t);\ t \geqslant 1\}$ ограничена единицей. По лемме Больцано–Вейерштрасса всегда существует подпоследовательность натуральных чисел вида $t_1 < t_2 < \ldots$, для которой соответствующая частичная подпоследовательность вида $\{\lambda_{i,j}(t_1), \lambda_{i,j}(t_2), \ldots\}$ сходится к конечному пределу, т. е. имеем $\lim\limits_{c\to\infty} \lambda_{i,j}(t_c) = q_{i,j}$. Более того, $q_{i,j} \geqslant 0$ и $\sum_{j=1}^{r} q_{i,j} = 1$, так как при любом $c = 1,2,\ldots$ имеем $\lambda_{i,j}(t_c) \geqslant 0$ и $\sum_{j=1}^{r} \lambda_{i,j}(t_c) = 1$. Следовательно, из чисел $q_{i,j}$, $i,j = 1,2,\ldots,r$, можно составить стохастическую матрицу $Q = \|q_{i,j}\|$ порядка r. Таким способом полученную матрицу Q будем называть предельной для последовательности натуральных чисел $t_1 < t_2 < \ldots$. Обозначим через $a_{i,j}$ и $b_{i,j}$ элементы матриц $Q\Pi$ и ΠQ, которые стоят на пересечении i-й строки и j-го столбца. Вспоминая правила умножения матриц, свойства пределов и уравнение Колмогорова–Чепмена (5.10), последовательно и легко найдем, что

$$a_{i,j} = \sum_{v=1}^{r} q_{i,v} p_{v,j} = \sum_{v=1}^{r} \lim_{c\to\infty} \lambda_{i,v}(t_c)\, p_{v,j} = \sum_{v=1}^{r} \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=1}^{t_c} p_{i,v}(m)\, p_{v,j} =$$
$$= \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=1}^{t_c} \sum_{v=1}^{r} p_{i,v}(m)\, p_{v,j} = \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=1}^{t_c} p_{i,j}(m+1) =$$

$$= \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=2}^{t_c+1} p_{i,j}(m), b_{i,j} = \sum_{v=1}^{r} p_{i,v} q_{v,j} = \sum_{v=1}^{r} p_{i,v} \lim_{c\to\infty} \lambda_{v,j}(t_c) =$$

$$= \sum_{v=1}^{r} p_{i,v} \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=1}^{t_c} p_{v,j}(m) = \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=1}^{t_c} \sum_{v=1}^{r} p_{i,v} p_{v,j}(m) =$$

$$= \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=1}^{t_c} p_{i,j}(m+1) = \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=2}^{t_c+1} p_{i,j}(m), \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=2}^{t_c+1} p_{i,j}(m) =$$

$$= \lim_{c\to\infty} \left(\frac{1}{t_c} \sum_{m=1}^{t_c} p_{i,j}(m) - t_c^{-1} p_{i,j} + t_c^{-1} p_{i,j}(t_c+1) \right) =$$

$$= \lim_{c\to\infty} \frac{1}{t_c} \sum_{m=1}^{t_c} p_{i,j}(m) = \lim_{c\to\infty} \lambda_{i,j}(t_c) = q_{i,j}.$$

Используя эти соотношения, находим: $Q = Q\Pi = \Pi Q$. Умножая равенства $Q = Q\Pi$ и $Q = \Pi Q$ последовательно на Π справа и слева, получим $Q\Pi = Q\Pi^2$, $\Pi Q = \Pi^2 Q$ или $Q = Q\Pi^2$, $Q = \Pi^2 Q$. Повторяя многократно этот процесс, выводим для любого $t = 1, 2, \ldots$ равенства $Q = Q\Pi^t$, $Q = \Pi^t Q$. Последние равенства позволяют для любого $t = 1, 2, \ldots$ получить следующие соотношения:

$$Q \times \left(\frac{1}{t} \sum_{m=1}^{t} \Pi^m \right) = \frac{1}{t} \sum_{m=1}^{t} (Q \times \Pi^m) = Q, \quad \sum_{v=1}^{r} q_{i,v} \frac{1}{t} \sum_{m=1}^{t} p_{v,j}(m) = q_{i,j},$$

$$\left(\frac{1}{t} \sum_{m=1}^{t} \Pi^m \right) \times Q = \frac{1}{t} \sum_{m=1}^{t} (\Pi^m \times Q) = Q, \quad \sum_{v=1}^{r} \frac{1}{t} \sum_{m=1}^{t} p_{i,\nu}(m) q_{v,j} = q_{i,j}. \tag{5.15}$$

Если теперь $W = \|w_{i,j}\|$ — некоторая предельная матрица порядка r для другой подпоследовательности натуральных чисел вида $t'_1 < t'_2 < \ldots$, полученной согласно лемме Больцано–Вейерштрасса, то из (5.15) имеем

$$\sum_{v=1}^{r} q_{i,v} w_{v,j} = \sum_{v=1}^{r} q_{i,v} \lim_{c\to\infty} \lambda_{v,j}(t'_c) = \sum_{v=1}^{r} q_{i,v} \lim_{c\to\infty} (t'_c)^{-1} \sum_{m=1}^{t'_c} p_{v,j}(m) =$$

$$= \lim_{c\to\infty} \sum_{v=1}^{r} q_{i,v} (t'_c)^{-1} \sum_{m=1}^{t'_c} p_{v,j}(m) = \lim_{c\to\infty} q_{i,j} = q_{i,j},$$

$$\sum_{v=1}^{r} w_{i,v} q_{v,j} = \sum_{v=1}^{r} \lim_{c\to\infty} \lambda_{i,v}(t'_c) q_{v,j} = \sum_{v=1}^{r} \lim_{c\to\infty} (t'_c)^{-1} \sum_{m=1}^{t'_c} p_{i,v}(m)\, q_{v,j} =$$

$$= \lim_{c\to\infty} \sum_{v=1}^{r} (t'_c)^{-1} \sum_{m=1}^{t'_c} p_{i,v}(m)\, q_{v,j} = \lim_{c\to\infty} q_{i,j} = q_{i,j},$$

или, что то же самое, в матричной форме $Q = QW = WQ$. В силу того, что W — также предельная матрица, то $W = WQ = QW$ и, значит, существует только одна предельная матрица

$$Q = W = \Lambda = \|\lambda_{i,j}\| = \lim_{t\to\infty} \frac{1}{t} \sum_{m=1}^{t} \Pi^m = \Lambda^2.$$

Теперь можно написать матричное равенство $\Lambda = \Lambda\Pi$ и поэтому при любом $i = 1, 2, \ldots, r$ строчка $(\lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,r})$ является стационарным распределением схемы Маркова. Более того, если $\lambda_1 \geqslant 0$, $\lambda_2 \geqslant 0, \ldots$, $\lambda_r \geqslant 0$ и $\lambda_1 + \lambda_2 + \ldots + \lambda_r = 1$, то получим соотношение

$$\left(\sum_{i=1}^{r} \lambda_i \times (\lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,r}) \times \Pi\right) = \\ = \sum_{i=1}^{r} \lambda_i \times \big((\lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,r}) \times \Pi\big) = \sum_{i=1}^{r} \lambda_i \times (\lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,r}).$$

Отсюда следует, что линейная комбинация

$$\sum_{i=1}^{r} \lambda_i \times (\lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,r})$$

строк матрицы Λ будет стационарным распределением схемы Маркова. Значит, в случае существования хотя бы двух различных строк этой матрицы схема Маркова имеет континуум стационарных распределений.

Наоборот, если $(\lambda_1, \lambda_2, \ldots, \lambda_r)$ — произвольное стационарное распределение схемы Маркова, то для всех $j = 1, 2, \ldots, r$ и $t = 1, 2, \ldots$ имеем $\lambda_j = \sum_{i=1}^{r} \lambda_i p_{i,j}(t)$. Отсюда получим, что $\lambda_j = \sum_{i=1}^{r} \lambda_i \frac{1}{t} \sum_{m=1}^{t} p_{i,j}(m)$. Переходя в последнем равенстве к пределу, найдем $\lambda_j = \sum_{i=1}^{r} \lambda_i \lambda_{i,j}$ при $t \to \infty$. Поэтому произвольное стационарное распределение $(\lambda_1, \lambda_2, \ldots, \lambda_r)$ представляется в виде линейной комбинации строк матрицы Λ. Теорема доказана.

При доказательстве теоремы 5.3 подпоследовательность Больцано–Вейерштрасса выбирается произвольным образом. Поэтому можно предложить другой способ определения стационарных распределений, который основан не на решении системы (5.12), а на специальном подборе подпоследовательности $t_1 < t_2 < \ldots$ Больцано–Вейерштрасса. Поясним этот способ на следующем примере.

Пример 5.4 (*задача о развитии экономики*). Экономика многих стран развивается циклически: за продолжительным периодом роста экономики наступает не менее продолжительный период застоя и даже кризиса. Рассмотрим эволюцию экономики некоторой страны через дискретные промежутки времени. Под действием рыночных и других факторов экономика страны на каждом периоде наблюдения может находиться только в одном из трех состояний: состояние $\{s_1\}$ соответствует кризису, состояние $\{s_2\}$ — переходному режиму и, наконец, состояние $\{s_3\}$ соответствует подъему экономики. Допустим теперь, что экономика, достигнув состояния кризиса или подъема, точно возвращается в состояние переходного режима. Если экономика на каком-то шаге была в переходном режиме, то на следующем периоде с одинаковой вероятностью экономика страны будет идти к упадку или будет успешно развиваться. Матрица вероятностей перехода за один шаг имеет в этом случае вид:

$$\Pi = \begin{pmatrix} 0 & 1 & 0 \\ 0{,}5 & 0 & 0{,}5 \\ 0 & 1 & 0 \end{pmatrix}.$$

Для этой схемы Маркова при $t = 1, 2, \ldots$ получим

$$\Pi^{2t-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0{,}5 & 0 & 0{,}5 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \Pi^{2t} = \begin{pmatrix} 0{,}5 & 0 & 0{,}5 \\ 0 & 1 & 0 \\ 0{,}5 & 0 & 0{,}5 \end{pmatrix}.$$

Значит, при любом $i = 1, 2, 3$ и каждом $j = 1, 2, 3$ пределы $\lim\limits_{t\to\infty} p_{i,j}(t)$ не существуют. Пусть матрица $\Lambda(t) = \|\lambda_{i,j}(t)\| = \frac{1}{t}\sum_{m=1}^{t} \Pi^m$. Тогда при специальном выборе подпоследовательности Больцано–Вейерштрасса $\{t_c = 2c\colon c = 1, 2, \ldots\}$ легко находим:

$$\|\lambda_{i,j}(2c)\| = \frac{c}{2c}\begin{pmatrix} 0 & 1 & 0 \\ 0{,}5 & 0 & 0{,}5 \\ 0 & 1 & 0 \end{pmatrix} + \frac{c}{2c}\begin{pmatrix} 0{,}5 & 0 & 0{,}5 \\ 0 & 1 & 0 \\ 0{,}5 & 0 & 0{,}5 \end{pmatrix} =$$

$$= \frac{1}{2}\begin{pmatrix} 0{,}5 & 1 & 0{,}5 \\ 0{,}5 & 1 & 0{,}5 \\ 0{,}5 & 1 & 0{,}5 \end{pmatrix} = \begin{pmatrix} 0{,}25 & 0{,}5 & 0{,}25 \\ 0{,}25 & 0{,}5 & 0{,}25 \\ 0{,}25 & 0{,}5 & 0{,}25 \end{pmatrix},$$

$$\Lambda = \begin{pmatrix} 0{,}25 & 0{,}5 & 0{,}25 \\ 0{,}25 & 0{,}5 & 0{,}25 \\ 0{,}25 & 0{,}5 & 0{,}25 \end{pmatrix}.$$

Отсюда видно, что предельная матрица Λ имеет одинаковые строчки вида $(1/4, 1/2, 1/4)$. Поэтому вектор $(1/4, 1/2, 1/4)$ определяет единственное стационарное распределение этой схемы Маркова.

В результате доказательства последней теоремы появляется простая возможность приближенного вычисления стационарных распределений схемы Маркова с конечным числом элементарных исходов без ограничений на матрицу Π. В качестве демонстрации такого вычисления воспользуемся несколько идеализированным примером 5.3. В этом примере было получено, что

$$\Pi^{2t-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}, \quad \Pi^{2t} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix} \quad \text{при } t = 1, 2, \ldots .$$

Если $\Lambda(t) = \|\lambda_{i,j}(t)\| = \frac{1}{t}\sum_{m=1}^{t} \Pi^m$, то при $t_c = 2c - 1$, $t'_c = 2c$, где $c \geqslant 1$, имеем

$$\|\lambda_{i,j}(2c)\| = \frac{1}{2c}\begin{pmatrix} c & c & 0 \\ c & c & 0 \\ c & c & 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix},$$

$$\|\lambda_{i,j}(2c-1)\| = \frac{1}{2c-1} \times \begin{pmatrix} c-1 & c & 0 \\ c & c-1 & 0 \\ (2c-1)/2 & (2c-1)/2 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 2^{-1} - 2^{-1}(2c-1)^{-1} & 2^{-1} + 2^{-1}(2c-1)^{-1} & 0 \\ 2^{-1} + 2^{-1}(2c-1)^{-1} & 2^{-1} - 2^{-1}(2c-1)^{-1} & 0 \\ 2^{-1} & 2^{-1} & 0 \end{pmatrix}.$$

Отсюда сразу получаем, что независимо от i будут справедливы равенства $\lambda_{i,3} = \lim\limits_{c\to\infty} \lambda_{i,3}(t_c) = \lim\limits_{c\to\infty} \lambda_{i,3}(t'_c) = 0$, $\lambda_{i,j} = \lim\limits_{c\to\infty} \lambda_{i,j}(t_c) = \lim\limits_{c\to\infty} \lambda_{i,j}(t'_c) = 1/2$ при $j = 1, 2$. Если теперь за приближенное значение стационарной вероятности $\lambda_{i,j}$ принять величину $\lambda_{i,j}(t)$, то модуль абсолютной погрешности при таком способе вычисления каждой из стационарных вероятностей не будет превышать величины $2^{-1}(2c-1)^{-1}$.

§ 3. Схемы Маркова с точки зрения эволюции реальных систем

3.1. Геометрическая интерпретация схемы Маркова. Схема Маркова является хорошей вероятностной моделью многих реальных эволюционных экспериментов $E = \{E_t;\ t = 1, 2, \ldots, n\}$, для которых условная вероятность $\mathsf{P}(\boldsymbol{A}_t(\omega_t)|\boldsymbol{A}_{t-1}(\omega_{t-1}))$ элементарного исхода $\{\omega_t\}$ в t-м испытании однозначно определяется элементарным исходом $\{\omega_{t-1}\}$ в $(t-1)$-м испытании и не изменяется от добавочной информации об элементарных исходах в более ранних испытаниях. В этом случае часто используют следующую удобную интерпретацию схемы Маркова. Мы будем говорить о некоторой физической, технической, биологической, экологической или экономической системе $E = \{E_t;\ t = 1, 2, \ldots, n\}$. Система E функционирует во времени и может находиться в каждый момент времени t в одном и только одном из состояний $\{s_1\}$, $\{s_2\}, \ldots,$ $\{s_r\}$. Поэтому здесь вместо интуитивного понятия об элементарном исходе вида $\{s_i\}$ эксперимента E_t используется более распространенный в теории систем термин состояние $\{s_i\}$ с номером i. Эту терминологию мы частично уже использовали в задаче о процессе протекания болезни и в задаче о развитии экономики.

Теперь последовательность зависимых испытаний в схеме Маркова можно наглядно представить на плоскости следующим образом. Каждому состоянию $\{s_i\}$ схемы Маркова (системы) поставим во взаимно однозначное соответствие точку на плоскости с обозначением i. Здесь i также означает номер состояния схемы Маркова. Любой упорядочной паре (i, j) таких точек на плоскости, для которых имеет место неравенство $p_{i,j} > 0$, сопоставляется направленная дуга. Эта дуга на плоскости изображается стрелкой, идущей в направлении от точки с меткой i к точке с меткой j. Итак, схема Маркова удобно представляется в виде ориентированного графа на плоскости. Вершинами такого графа являются помеченные точки на плоскости, а ребрами — направленные дуги. Этапы построения ориентированного графа на плоскости продемонстрируем на примере.

Пример 5.5 (*задача о сборе грибов*). Пусть имеется большой массив леса, который разбит на участки. Всего выделено пять участков (см. рис. 5.1). Первый и пятый участки имеют общие границы. Через первый и пятый участки проложена автомобильная дорога. Четвертый участок занимает незначительную по размерам площадь и имеет общие границы с пятым, вторым и третьим. Второй и третий участки составляют большую часть лесного массива, имеют очень непротяженные проходимые границы с четвертым, отделены от первого и пятого непроходимым болотом и расположены внутри этого массива. Представим человека, который был доставлен вертолетом в случайно выбранное место массива леса для сбора грибов. Каждые десять минут, например, по виду деревьев или по шуму автомобилей грибник может определить только номер участка, а не свое точное местонахождение. С помощью исследования некоторой схемы Маркова, адекватно описывающей хождение в лесу грибника, можно предсказать его положение в конце

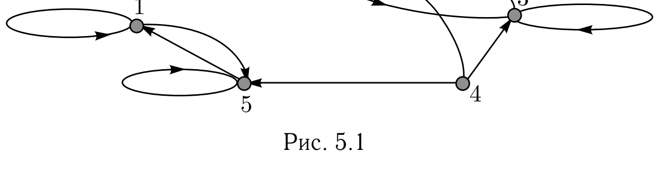


Рис. 5.1

сбора грибов и, тем самым, решить вопрос о его возвращении домой с помощью вертолета или рейсового автобуса.

Для этой системы в качестве вероятностной модели можно предложить однородную схему Маркова с пятью состояниями $\{s_1\}$, $\{s_2\}$, $\{s_3\}$, $\{s_4\}$, $\{s_5\}$. Пусть для этой задачи матрица вероятностей перехода за один шаг имеет вид

$$\Pi = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 5/6 & 1/6 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 & 0 \\ 0 & 1/18 & 1/18 & 0 & 8/9 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix}.$$

При этом состояние $\{s_i\}$ схемы Маркова означает, что в конце некоторого десятиминутного промежутка времени грибник проводил сборы на i-м участке. Сделаем несколько пояснений относительно нулевых и ненулевых значений вероятностей перехода за один шаг. Элементы первой и пятой строк матрицы Π показывают, что грибник, попав на первый или пятый участок, перемещается совершенно случайным образом только по этим участкам. Если грибник попал на четвертый участок, то он, скорее перейдет на пятый участок, чем на второй или третий. Такое естественное поведение можно объяснить тем, что грибник желает вернуться домой автобусом и не заблудиться, ориентируясь постоянно по шуму автомобилей на трассе. Первый участок и незначительный по размерам четвертый участок не имеют общей границы. Поэтому $p_{1,4} = p_{4,1} = p_{4,4} = 0$. Так как второй и третий участки занимают большую площадь, практически не имеют проходимых границ с четвертым и отделены от первого и пятого, то естественно принять $p_{1,2} = p_{1,3} = p_{2,1} = p_{2,4} = p_{2,5} = p_{3,1} = p_{3,4} = p_{3,5} = 0$. Числовые значения ненулевых элементов матрицы вероятностей перехода за один шаг естественно носят условно-иллюстративный характер и, в действительности, могут быть найдены только экспериментальным путем.

Перейдем теперь непосредственно к построению ориентированного графа, соответствующего рассматриваемой схеме Маркова (рис. 5.1).

Для этого совершенно произвольно выберем пять разных точек на плоскости и пометим их цифрами 1, 2, 3, 4, 5. Далее согласно всем ненулевым элементам матрицы Π проведем направленные дуги, которые соединяют некоторые из помеченных точек. В результате получим ориентированный граф, который изображен на рис. 5.1. На этом рисунке имеется особая дуга (петля), которая начинается и заканчивается в точке с меткой единица, так как $p_{1,1} = 1/2 > 0$. Каждая из точек с метками 2, 3, 5 также по этой причине имеют петлю. Точка с меткой 4 петли не имеет, так как $p_{4,4} = 0$. Естественно нет направленных дуг, которые соединяют точки с метками 1 и 4, так как $p_{1,4} = p_{4,1} = 0$. Аналогичная ситуация имеет место для остальных нулевых элементов матрицы Π. Число ненулевых элементов всегда совпадает с количеством направленных дуг. В нашем случае это число равно 11. Легко проверить выполнение следующих равенств:

$$(1/2, 0, 0, 0, 1/2) \times \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 5/6 & 1/6 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 & 0 \\ 0 & 1/18 & 1/18 & 0 & 8/9 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix} = (1/2, 0, 0, 0, 1/2),$$

$$(0, 4/5, 1/5, 0, 0) \times \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 5/6 & 1/6 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 & 0 \\ 0 & 1/18 & 1/18 & 0 & 8/9 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix} = (0, 4/5, 1/5, 0, 0),$$

т. е. распределения $(1/2, 0, 0, 0, 1/2)$, $(0, 4/5, 1/5, 0, 0)$ будут стационарными для этой схемы Маркова. Тогда для любого $0 \leqslant \lambda \leqslant 1$ имеем

$$\begin{aligned} &(\lambda/2, 4(1-\lambda)/5, (1-\lambda)/5, 0, \lambda/2) \times \Pi = \\ &\qquad = \{\lambda \times (1/2, 0, 0, 0, 1/2) + (1-\lambda) \times (0, 4/5, 1/5, 0, 0)\} \times \Pi = \\ &\qquad = \lambda \times (1/2, 0, 0, 0, 1/2) \times \Pi + (1-\lambda) \times (0, 4/5, 1/5, 0, 0) \times \Pi = \\ &\qquad = \lambda \times (1/2, 0, 0, 0, 1/2) + (1-\lambda) \times (0, 4/5, 1/5, 0, 0) = \\ &\qquad = (\lambda/2, 4(1-\lambda)/5, (1-\lambda)/5, 0, \lambda/2). \end{aligned}$$

Поэтому для этой схемы Маркова существует континуум стационарных распределений вида $(\lambda/2, 4(1-\lambda)/5, (1-\lambda)/5, 0, \lambda/2)$, $0 \leqslant \lambda \leqslant 1$.

3.2. Стандартные типы схем Маркова. Как уже отмечалось, схемы Маркова часто используются для адекватного описания эволюции самых разнообразных реальных систем. В силу этого поведения схем Маркова при $t \to \infty$ могут существенно отличаться друг от друга. Доказанные теоремы и рассмотренные примеры в этой главе позволяют выделить различные типы схем Маркова. Для решения этой задачи будем использовать особенности и свойства предельной матрицы $\Lambda = \|\lambda_{i,j}\|$, которая существует для любой схемы Маркова, и следующую терминологию.

Схема Маркова называется *апериодической*, если для всех i, j существуют пределы $\lim\limits_{t\to\infty} p_{i,j}(t) = \lambda_{i,j} \geqslant 0$. В противном случае схема Маркова называется *периодической*. Апериодическая схема Маркова называется *регулярной*, если дополнительно каждая величина $\lambda_{i,j}$ не зависит от i или, что то же самое, все строки предельной матрицы Λ одинаковы. Заметим, что предельная матрица Λ может иметь одинаковые строки и при этом не быть регулярной из-за отсутствия пределов $\lim\limits_{t\to\infty} p_{i,j}(t)$, как это было установлено в примере 5.3. Схема Маркова называется *разложимой*, если существует хотя бы один элемент предельной матрицы Λ, равный нулю. При $\lambda_{i,j} > 0$ для всех i, j схема Маркова называется *неразложимой*. Заметим, что предельная матрица Λ не может иметь разные строки, если схема Маркова является неразложимой. В самом деле, строго пололожительную и стохастическую матрицу Λ можно рассматривать как матрицу переходных вероятностей некоторой однородной схемы Маркова. Для такой схемы $\Lambda = \Lambda^2 = \lim\limits_{t\to\infty} \Lambda^t$ и имеет место теорема Маркова. Поэтому матрица Λ должна иметь одинаковые строки.

Понятия и замечания этого раздела позволяют разбить множество схем Маркова на следующие шесть типов:

1) *регулярные и неразложимые марковские* системы или *эргодические* марковские системы с распределением $\pi_1 > 0$, $\pi_2 > 0, \ldots$, $\pi_r > 0$ (см. теорему Маркова и пример 5.1 — задачу о процессе протекания болезни и выздоровления);
2) *регулярные и разложимые* марковские системы, т. е. существует предельное распределение $\pi_j = \lim\limits_{t\to\infty} p_j^{(t)} \geqslant 0$, $j = 1, 2, \ldots, r$, которое не зависит от начального распределения p_1, $p_2, \ldots$, p_r, совпадает с единственным стационарным распределением и, наконец, существует хотя бы одно такое состояние $\{s_k\}$, что вероятность $\pi_k = 0$ (см. пример 5.2);
3) *апериодические*, *нерегулярные и разложимые* марковские системы, т. е. для всех i, j существуют пределы $\lim\limits_{t\to\infty} p_{i,j}(t) = \lambda_{i,j} \geqslant 0$, континуум стационарных распределений и хотя бы один элемент предельной матрицы Λ равен нулю (см. пример 5.5 — задачу о сборе грибов);
4) *периодические*, *неразложимые и с одинаковыми строчками матрицы* Λ марковские системы, т. е. имеется единственное стационарное и неэргодическое распределение $\lambda_1 > 0$, $\lambda_2 > 0, \ldots$, $\lambda_r > 0$ (см. пример 5.4 — задачу о развитии экономики);
5) *периодические*, *разложимые и с одинаковыми строчками матрицы* Λ марковские системы, для которых имеется единственное стационарное распределение $\lambda_1, \lambda_2, \ldots, \lambda_r$ и $\lambda_1 \times \lambda_2 \times \ldots \times \lambda_r = 0$ (см. пример 5.3);
6) *периодические*, *разложимые и с разными строчками матрицы* Λ марковские системы, т. е. предельная матрица Λ имеет нулевые элементы и существует континуум стационарных распределений (см. пример 5.6).

Пример 5.6. Рассмотрим несколько другой вариант случайного блуждания человека при сборе грибов. Допустим теперь, что грибник с первого участка обязательно переходит на пятый и, наоборот, с пятого участка точно переходит на первый. Остальные условия передвижения человека по лесу остаются такими же, как и в первоначальной задаче. В этом случае соответствующая схема Маркова имеет матрицу вероятностей перехода вида

$$\Pi = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 5/6 & 1/6 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 & 0 \\ 0 & 1/18 & 1/18 & 0 & 8/9 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

а ориентированный граф переходов состояний для этой схемы Маркова изображен на рис. 5.2.

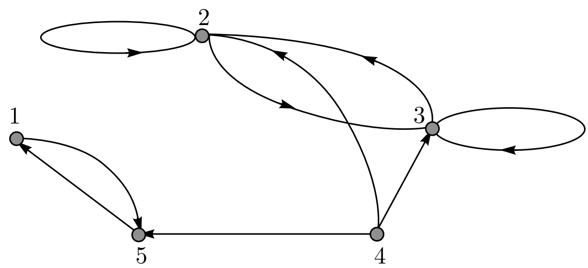


Рис. 5.2

На этом рисунке уже каждая из точек с метками 1, 5 петлю не имеет, так как в новой схеме Маркова $p_{1,1} = p_{5,5} = 0$. Матрица Π и ориентированный граф на рис. 5.2 позволяют без труда найти $p_{1,1}(2t-1) = p_{5,5}(2t-1) = 0$, $p_{1,1}(2t) = p_{5,5}(2t) = 1$ для всех $t = 1, 2, \ldots$. Следовательно, пределы вида $\lim\limits_{t\to\infty} p_{1,1}(t)$, $\lim\limits_{t\to\infty} p_{5,5}(t)$ не существуют. По меньшей мере, элементы четвертого столбца предельной матрицы Λ равны нулю, так как для всех $t = 1, 2, \ldots$ и независимо от i имеем $p_{i,4}(t) = 0$. Точно таким же способом, как мы это доказали в первоначальной задаче о сборе грибов, установим, что данная схема Маркова имеет континуум стационарных распределений следующего вида: $(\lambda/2, 4(1-\lambda)/5, (1-\lambda)/5, 0, \lambda/2)$, $0 \leqslant \lambda \leqslant 1$.

3.3. Разбиение пространства состояний схемы Маркова на замкнутые классы и их свойства. В предыдущем пункте выделили шесть различных типов схем Маркова. Для того чтобы глубже и точнее разобраться в асимптотическом поведении каждой такой схемы Маркова, введем классификацию ее состояний. В качестве признака классификации выберем возможность перехода системы из одного состояния в другое за некоторое конечное число шагов. Назовем любые состояния $\{s_i\}$, $\{s_j\}$ сообщающимися, если существуют такие целые числа $t_1 \geqslant 0$ и $t_2 \geqslant 0$, что одновременно $p_{i,j}(t_1) > 0$ и $p_{j,i}(t_2) > 0$. Этот факт обозначают символом $\{s_i\} \leftrightarrow \{s_j\}$. По определению полагаем, что $p_{i,j}(0) = 1$ при $i = j$ и $p_{i,j}(0) = 0$ при $i \neq j$, следовательно, всегда имеем $\{s_i\} \leftrightarrow \{s_i\}$. Если два разных состояния $\{s_i\}$ и $\{s_j\}$ не сообщаются, то или $p_{i,j}(t) = 0$ при всех $t = 0, 1, \ldots$, или $p_{j,i}(t) = 0$ при всех $t = 0, 1, \ldots$, или эти условия выполняются вместе. На содержательном уровне можно сказать, что при сообщаемости состояний $\{s_i\}$ и $\{s_j\}$ система за конечное число шагов попадает из состояния $\{s_i\}$ в состояние $\{s_j\}$ и одновременно состояние $\{s_i\}$ достижимо системой из состояния $\{s_j\}$ также за конечное число шагов. Свойство сообщаемости состояний позволяет сформулировать следующее утверждение.

Теорема 5.4. *Множество всех состояний схемы Маркова с помощью отношения $\{s_i\} \leftrightarrow \{s_j\}$ разбивается на непересекающиеся классы.*

Доказательство. Свойство сообщаемости состояний $\{s_i\}$ и $\{s_j\}$ представляет собой отношение эквивалентности, другими словами, является рефлексивным, симметричным и транзитивным. В самом деле, по определению это отношение является рефлексивным и симметричным, т. е. имеет место $\{s_i\} \leftrightarrow \{s_i\}$, и из условия $\{s_i\} \leftrightarrow \{s_j\}$ следует $\{s_j\} \leftrightarrow \{s_i\}$. Далее, если $\{s_i\} \leftrightarrow \{s_j\}$ и $\{s_j\} \leftrightarrow \{s_v\}$, то существуют такие неотрицательные целые числа t_1, t_2, m_1 и m_2, что $p_{i,j}(t_1) > 0$, $p_{j,v}(t_2) > 0$, $p_{v,j}(m_1) > 0$ и $p_{j,i}(m_2) > 0$. Отсюда, в силу уравнения Колмогорова–Чепмена (5.10) и неотрицательности всех $p_{i,k}(t_1)$, $p_{k,v}(t_2)$, $p_{v,k}(m_1)$ и $p_{k,i}(m_2)$ имеем

$$\begin{aligned} p_{i,v}(t_1 + t_2) &= \sum_{k=1}^{r} p_{i,\,k}(t_1)\, p_{k,v}(t_2) \geqslant p_{i,j}(t_1)\, p_{j,v}(t_2) > 0, \\ p_{v,i}(m_1 + m_2) &= \sum_{k=1}^{r} p_{v,\,k}(m_1)\, p_{k,\,i}(m_2) \geqslant p_{v,j}(m_1)\, p_{j,i}(m_2) > 0, \end{aligned} \tag{5.16}$$

следовательно, отношение сообщаемости является также транзитивным. Известно [11], что с помощью отношения эквивалентности множество всех состояний схемы Маркова можно разбить на непересекающиеся классы K_1, $\mathsf{K}_2, \ldots$, K_b. Так как каждый класс может состоять из двух или более сообщающихся между собой состояний, либо из одного состояния, то целое число b очевидно удовлетворяет условию $1 \leqslant b \leqslant r$. При этом состояния объединяются в один класс только тогда, когда они сообщаются друг с другом. Теорема доказана.

Рассмотрим теперь такие свойства любого состояния, которыми будут обладать все сообщающиеся с ним состояния. Тогда такие свойства можно объявить свойствами класса, а класс представляет собой совокупность состояний с едиными свойствами. Сформулируем теперь эти свойства.

Состояние $\{s_i\}$ назовем *несущественным*, если существуют такое состояние $\{s_j\}$ и натуральное число m, что $p_{i,j}(m) > 0$, однако $p_{j,i}(t) = 0$ для каждого $t \geqslant 1$. В этом случае возможен переход (не обязательно за один шаг) из состояния $\{s_i\}$ в некоторое состояние $\{s_j\}$. При этом не допускается возвращение в состояние $\{s_i\}$. В противном случае состояние назовем *существенным*. Состояние $\{s_i\}$ будет существенным тогда и только тогда, когда возможен переход из состояния $\{s_j\}$ в $\{s_i\}$ для всех таких состояний $\{s_j\}$, для которых из состояния $\{s_i\}$ можно достигнуть состояния $\{s_j\}$ схемы Маркова.

Периодом $d(i)$ состояния $\{s_i\}$ называется наибольший общий делитель всех натуральных чисел m, для которых $p_{i,i}(m) > 0$. Если $p_{j,j}(m) = 0$ при всех $m \geqslant 1$, то по определению полагаем $d(j) = 0$. Состояние $\{s_i\}$ схемы Маркова называется *периодическим* при $d(i) > 1$ и *апериодическим* при $d(i) \leqslant 1$. Покажем, что указанные свойства являются свойствами класса.

Теорема 5.5. *Все состояния из одного класса марковской цепи могут быть только существенными, либо только несущественными.*

Доказательство. Ради простоты покажем это утверждение для первого класса K_1. Если этот класс состоит из одного состояния, то утверждение теоремы носит очевидный характер. Пусть, например, состояние $\{s_i\}$ из класса K_1 является существенным, и предположим, что существует хотя бы одно несущественное состояние $\{s_j\}$ из этого класса. Так как состояния $\{s_i\}$ и $\{s_j\}$ сообщаются, то существуют такие натуральные числа t_1 и t_2, что $p_{i,j}(t_1) > 0$ и $p_{j,i}(t_2) > 0$. Поскольку $\{s_j\}$ по предположению является несущественным состоянием, то существуют такое состояние $\{s_v\}$ и натуральное число m_1, что $p_{j,v}(m_1) > 0$ и $p_{v,j}(t) = 0$ для всех $t = 1, 2, \ldots$. Из последних двух предложений с использованием способа получения неравенств (5.16) выводим, что $p_{i,v}(t_1 + m_1) > 0$. Если теперь для некоторого натурального числа m_2 будет $p_{v,i}(m_2) > 0$, то так же получим, что $p_{v,j}(m_2 + t_1) > 0$. Последнее неравенство противоречит условиям $p_{v,j}(t) = 0$ для всех $t = 1, 2, \ldots$. Если такого натурального числа m_2 не существует, т. е. для каждого $t = 1, 2, \ldots$ верно равенство $p_{v,i}(t) = 0$, тогда состояние $\{s_i\}$ оказывается несущественным. Аналогично доказывается вторая часть этой теоремы, когда предполагаем, что есть несущественное состояние $\{s_i\}$ в классе K_1.

Теперь термин существенный или несущественный можно применять к произвольному классу K_g, где $g = 1, 2, \ldots, b$.

Легко видеть, что существуют схемы Маркова, каждая из которых не имеет несущественных классов. В качестве простого примера можно взять схему Маркова, для которой все элементы матрицы Π из вероятностей переходов строго положительны. Из следующего утверждения следует, что аналогичный факт для существенных классов не имеет место.

Теорема 5.6. *Среди непересекающихся классов* K_1, $\mathsf{K}_2, \ldots, \mathsf{K}_b$ *любой схемы Маркова с конечным множеством состояний существует хотя бы один существенный класс.*

Доказательство. Допустим обратное. В этом случае все состояния схемы Маркова и классы K_1, $\mathsf{K}_2, \ldots, \mathsf{K}_b$ являются несущественными. Не ограничивая общности, выберем на первом шаге произвольное состояние $\{s_i\}$ из некоторого несущественного класса, например, из класса K_1. Удобно в дальнейшем несущественный класс K_1 обозначить через $\mathsf{K}(s_i)$. Напомним, что класс $\mathsf{K}(s_i)$ содержит только такие состояния, которые сообщаются с $\{s_i\}$. Так как состояние $\{s_i\}$ является несущественным, то существуют натуральное число t_1 и состояние $\{s_{i_1}\}$ из несущественного класса $\mathsf{K}(s_{i_1})$, что будет $p_{i,i_1}(t_1) > 0$ и $p_{i_1,i}(t) = 0$ для каждого натурального числа t. Итак, состояния $\{s_i\}$ и $\{s_{i_1}\}$ не сообщаются. Поэтому классы $\mathsf{K}(s_i)$ и $\mathsf{K}(s_{i_1})$ не пересекаются. Далее, на втором шаге для несущественного состояния $\{s_{i_1}\}$ найдутся такое натуральное число t_2 и состояние $\{s_{i_2}\}$ из несущественного класса $\mathsf{K}(s_{i_2})$, что $p_{i_1,i_2}(t_2) > 0$ и $p_{i_2,i_1}(t) = 0$ для каждого натурального числа t. При этом классы $\mathsf{K}(s_i)$, $\mathsf{K}(s_{i_1})$, $\mathsf{K}(s_{i_2})$ попарно не пересекаются. Продолжая аналогичным образом этот процесс, на b-м шаге имеем некоторое несущественное состояние $\{s_{i_{b-1}}\}$. К этому шагу мы перебрали все несущественные классы K_1, $\mathsf{K}_2, \ldots, \mathsf{K}_b$. Поэтому для несущественного состояния $\{s_{i_{b-1}}\}$ не найдутся такое натуральное число t_b и состояние $\{s_{i_b}\}$, что $p_{i_{b-1},i_b}(t_b) > 0$ и $p_{i_b,i_{b-1}}(t) = 0$ для каждого натурального числа t. Значит, состояние $\{s_{i_{b-1}}\}$ не является несущественным. Получили противоречие с тем, что все состояния схемы Маркова были несущественными. Отметим, что при доказательстве этой теоремы существенно использовалось то обстоятельство, что схема Маркова имела конечное число состояний. Теорема доказана.

Теорема 5.7. *Состояния из одного класса имеют один и тот же период.*

Доказательство. Утверждение теоремы становится тривиальным в том случае, когда класс содержит единственное состояние. Пусть теперь $\{s_i\}$ и $\{s_j\}$ — два произвольных состояния, например, из класса K_1, т. е. $\{s_i\} \leftrightarrow \{s_j\}$. Тогда найдутся такие натуральные числа t_1 и t_2, для которых будет $p_{i,j}(t_1) > 0$ и $p_{j,i}(t_2) > 0$. Выберем любое m, для которого $p_{i,i}(m) > 0$. Отсюда, как и при выводе соотношения (5.16), последовательно получим

$$p_{j,j}(t_2 + m + t_1) = \sum_{k=1}^{r} p_{j,k}(t_2)\, p_{k,j}(m + t_1) \geqslant$$
$$\geqslant p_{j,i}(t_2)\, p_{i,j}(m + t_1) \geqslant p_{j,i}(t_2)\, p_{i,i}(m)\, p_{i,j}(t_1) > 0,$$

$$p_{i,i}(2m) = p_{i,i}(m+m) \geqslant p_{i,i}(m)\,p_{i,i}(m) > 0,$$
$$p_{j,j}(t_2 + 2m + t_1) \geqslant p_{j,i}(t_2)\,p_{i,i}(2m)\,p_{i,j}(t_1) > 0.$$

Так как $p_{j,j}(t_2+m+t_1) > 0$ и $p_{j,j}(t_2+2m+t_1) > 0$ и $d(j)$ является периодом состояния $\{s_j\}$, то натуральные числа t_2+m+t_1, t_2+2m+t_1 делятся на $d(j)$. Значит, разность $(t_2+2m+t_1)-(t_2+m+t_1)=m$ также делится на $d(j)$. Это свойство имеет место для любого m, для которого $p_{i,i}(m) > 0$. Учитывая, что $\{s_i\}$ имеет период $d(i)$, получим неравенство $d(j) \leqslant d(i)$. Если в приведенных выше рассуждениях поменять местами состояния $\{s_i\}$ и $\{s_j\}$, то придем к неравенству $d(i) \leqslant d(j)$. Окончательно имеем $d(i) = d(j)$ для любых состояний $\{s_i\}$ и $\{s_j\}$ из данного класса K_1. Поэтому этот общий период целесообразно называть периодом класса и обозначать через символ $d(\mathsf{K}_1)$. Теорема доказана.

Значение теоремы 5.7 состоит в том, что при ее доказательстве фактически был установлен следующий замечательный факт. Схема Маркова с некоторой положительной вероятностью через конечное число шагов выйдет из любого несущественного класса так, что возврат ее в исходный класс уже невозможен. Более того, через некоторое конечное число шагов схема Маркова из всего множества несущественных классов также с положительной вероятностью попадает в один из существенных классов сообщающихся состояний. В силу этого общего свойства все несущественные состояния схемы Маркова можно собрать в один класс K_0. Ради простоты ниже существенные классы будем обозначать символами K_1, $\mathsf{K}_2, \ldots$, K_b. Здесь символ b уже обозначает число существенных классов. Таким образом, окончательно произведено разбиение множества всех состояний схемы Маркова на классы K_1, $\mathsf{K}_2, \ldots$, K_b, K_0.

Так как множество всех состояний является конечным, то при $b > 1$ и $\mathsf{K}_0 \neq \varnothing$ можно перенумеровать состояния так, чтобы последовательные номера состояний, кроме первого и последнего состояний из каждого класса, принадлежали одному и тому же классу. Если номера состояниям присваивать таким способом, то сначала идут номера состояний из существенного класса K_1, затем — из существенного класса $\mathsf{K}_2, \ldots$, потом — из существенного класса K_b и, наконец, последними идут номера из несущественного класса K_0.

Пусть теперь Π_1, $\Pi_2, \ldots$, Π_b, N суть квадратные блоки (матрицы), расположенные вдоль главной диагонали матрицы Π, а остальные блоки в общем случае — прямоугольные. Через символ O обозначен каждый из нулевых блоков, а через $N_1, N_2, \ldots, N_b$ обозначены блоки, которые соответствуют несущественным состояниям марковской цепи. В этом случае матрица вероятностей перехода за один шаг примет так называемый блочный (клеточный) вид:

$$\Pi = \begin{pmatrix} \Pi_1 & O & O & \bullet & \bullet & \bullet & O & O \\ O & \Pi_2 & O & O & \bullet & \bullet & \bullet & O \\ O & O & \Pi_3 & O & \bullet & \bullet & \bullet & O \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ O & O & \bullet & \bullet & \bullet & O & \Pi_b & O \\ N_1 & N_2 & \bullet & \bullet & \bullet & N_{b-1} & N_b & N \end{pmatrix}. \tag{5.17}$$

Напомним, что в курсе высшей алгебры матрица такого вида называется *разложимой*. В противном случае, когда с помощью указанного выше способа матрица Π не может быть приведена к такому виду, она называется *неразложимой*. Для неразложимой матрицы Π число $b = 1$ и соответствующая ей схема Маркова имеет только существенные состояния. При $g = 1, 2, \ldots, b$ размерность блока Π_g равна числу состояний существенного класса K_g, а размерность блока N совпадает с числом всех несущественных состояний в классе K_0. Число строк каждого прямоугольного блока равно числу строк некоторого квадратного блока, который расположен впереди или после этого прямоугольного блока. Число столбцов каждого прямоугольного блока равно числу столбцов некоторого квадратного блока, который расположен выше или ниже этого прямоугольного блока. Появление нулевых блоков вызвано наличием непересекающихся классов K_1, $\mathsf{K}_2, \ldots$, K_b, K_0. Каждое состояние из любого такого класса не сообщается ни с одним состоянием из любого другого класса.

Так как K_0 является множеством всех несущественных состояний, то существует хотя бы один ненулевой блок из семейства блоков вида N_1, $N_2, \ldots$, N_b. Отсюда вытекает, что у квадратной матрицы N хотя бы одна строчная сумма меньше единицы или, другими словами, квадратная матрица N не является стохастической. Напротив, квадратные матрицы Π_1, $\Pi_2, \ldots$, Π_b будут стохастическими, так как любой из непересекающихся классов K_1, $\mathsf{K}_2, \ldots$, K_b является существенным. При фиксированном значении $g = 1, 2, \ldots, b$ любые состояния из класса K_g сообщаются между собой и соответствующая этому классу матрица Π_g является стохастической, т. е. $p_{i,j} = 0$ для каждого $\{s_i\} \in \mathsf{K}_g$ и любого $\{s_j\} \notin \mathsf{K}_g$. Это означает, что схема Маркова после достижения некоторого состояния из класса K_g навсегда остается в этом классе. В этом смысле класс K_g является замкнутым множеством. Более того, класс K_g не содержит ни одного собственного подмножества, обладающего таким свойством. Поэтому иногда любой класс K_g называют минимальным множеством для схемы Маркова.

Пример 5.7 (*задача о выборах по партийным спискам*). На следующем примере рассмотрим различные задачи, которые обычно решают в связи с введенной классификацией схем Маркова. В странах с переходной экономикой очень часто проходят разного рода выборы по партийным спискам. Некоторый избиратель перед очередными выборами имеет право реализовать одну из большого числа следующих своих возможностей:

1) он пропускает выборы и это означает для него состояние $\{s_1\}$;
2) голосует за коммунистов — состояние $\{s_2\}$;
3) голосует за демократов левого толка — состояние $\{s_3\}$;
4) голосует против всех — состояние $\{s_4\}$;
5) голосует за демократов правого толка — состояние $\{s_5\}$;
6) голосует за патриотов — состояние $\{s_6\}$;
7) голосует за центристов — состояние $\{s_7\}$;
8) голосует за монархию — состояние $\{s_8\}$.

Опытным путем установлено, что для избирателя некоторой местности схема Маркова процесса голосования имеет матрицу вероятностей перехода

$$\Pi = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0{,}8 & 0 & 0 & 0 & 0{,}2 & 0 & 0 \\ 0 & 0 & 0{,}4 & 0 & 0 & 0 & 0{,}6 & 0 \\ 0{,}5 & 0 & 0{,}1 & 0 & 0 & 0{,}4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0{,}6 & 0 & 0{,}4 & 0 \\ 0 & 0{,}3 & 0 & 0 & 0 & 0{,}7 & 0 & 0 \\ 0 & 0 & 0{,}5 & 0 & 0{,}5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}2 & 0 & 0{,}3 & 0 & 0{,}5 \end{pmatrix}$$

и ее ориентированный граф переходов состояний имеет вполне конкретный вид (см. рис. 5.3). Видно, что имеется направленная дуга от состояния $\{s_8\}$ к состоянию $\{s_6\}$ и нет ни одной дуги, которая входит в состояние $\{s_8\}$ из состояний $\{s_1\}$, $\{s_2\}$, $\{s_3\}$, $\{s_4\}$, $\{s_5\}$, $\{s_6\}$, $\{s_7\}$, поэтому состояние $\{s_8\}$ является несущественным.

Далее, так как вероятность перехода $p_{8,8} = 0{,}5 > 0$, то период состояния $\{s_8\}$ равен единице, и это состояние является апериоди-

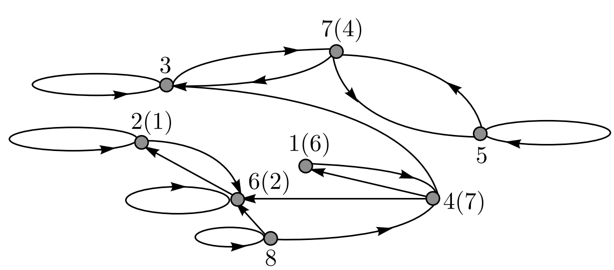


Рис. 5.3

ческим. Аналогичным образом устанавливается, что сообщающиеся только между собой состояния $\{s_1\}$ и $\{s_4\}$ будут несущественными. Более того, из состояния $\{s_1\}$ или $\{s_4\}$ можно вернуться снова в это же состояние только за четное число шагов. Следовательно, каждое из состояний $\{s_1\}$ и $\{s_4\}$ имеет период, равный двум. Состояния $\{s_2\}$, $\{s_6\}$ сообщаются только между собой. Схема Маркова не может попасть из состояний $\{s_2\}$, $\{s_6\}$ в любое другое состояние, отличное от этих состояний, и $p_{2,2} = 0{,}8 > 0$. Значит, состояния $\{s_2\}$, $\{s_6\}$ являются существенными и апериодическими. Точно по такой же причине состояния $\{s_3\}$, $\{s_5\}$, $\{s_7\}$ являются существенными, сообщающимися и апериодическими.

Итак, ориентированный граф схемы Маркова процесса голосования содержит два существенных, непериодических и минимально-замкнутых класса $\mathsf{K}_1 = \{\{s_2\}, \{s_6\}\}$, $\mathsf{K}_2 = \{\{s_3\}, \{s_5\}, \{s_7\}\}$ и два несущественных незамкнутых класса $\mathsf{K}_3 = \{\{s_1\}, \{s_4\}\}$, $\mathsf{K}_4 = \{\{s_8\}\}$. При этом период класса K_3 равен двум, а класс K_4 является апериодическим. Теперь номера состояний $\{s_3\}$, $\{s_5\}$, $\{s_8\}$ оставим без изменения, а остальным состояниям присвоим новые номера, указанные на рис. 5.3 в круглых скобках. После этого объединим два несущественных класса K_3 и K_4 в один несущественный класс K_0 и учтем при этом новую нумерацию состояний. Тогда без труда найдем: $b = 2$, $\mathsf{K}_1 = \{\{s_1\}, \{s_2\}\}$, $\mathsf{K}_2 = \{\{s_3\}, \{s_4\}, \{s_5\}\}$, $\mathsf{K}_0 = \{\{s_6\}, \{s_7\}, \{s_8\}\}$. В результате таких действий матрица вероятностей перехода рассматриваемой схемы Маркова приводится к блочному виду:

$$\Pi = \begin{pmatrix} 0{,}8 & 0{,}2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0{,}3 & 0{,}7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0{,}4 & 0{,}6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0{,}5 & 0 & 0{,}5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}4 & 0{,}6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0{,}4 & 0{,}1 & 0 & 0 & 0{,}5 & 0 & 0 \\ 0 & 0{,}3 & 0 & 0 & 0 & 0 & 0{,}2 & 0{,}5 \end{pmatrix}.$$

При этом матрицы Π_1, Π_2, N, N_1 и N_2 имеют вид:

$$\Pi_1 = \begin{pmatrix} 0{,}8 & 0{,}2 \\ 0{,}3 & 0{,}7 \end{pmatrix}, \quad \Pi_2 = \begin{pmatrix} 0{,}4 & 0{,}6 & 0 \\ 0{,}5 & 0 & 0{,}5 \\ 0 & 0{,}4 & 0{,}6 \end{pmatrix},$$

$$N = \begin{pmatrix} 0 & 1 & 0 \\ 0{,}5 & 0 & 0 \\ 0 & 0{,}2 & 0{,}5 \end{pmatrix}, \quad N_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0{,}4 \\ 0 & 0{,}3 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0{,}1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

На этом конкретном примере убеждаемся, что квадратные матрицы Π_1, Π_2 являются стохастическими, квадратная матрица N не является таковой и матрицы N_1, N_2 образуют ненулевые блоки.

3.4. Разбиение минимального множества всех состояний схемы Маркова на циклические подклассы. Пусть множество всех состояний схемы Маркова является минимальным. В этом случае $b = 1$ и все ее состояния будут существенными. Далее, рассмотрим два случая. Первый случай означает, что схема Маркова имеет период $d = 1$. Для такой схемы Маркова матрица Π вероятностей перехода за один шаг является не только неразложимой, но и примитивной, так как $d = 1$. Известно, что для примитивной матрицы Π всегда существует такое натуральное число m, для которого $\Pi^m > 0$. Поэтому такая система по теореме Маркова будет эргодической.

Второй случай означает, что схема Маркова имеет период $d > 1$. Выберем произвольное состояние, например, $\{s_1\}$ и образуем подклассы (подмножества) $\mathfrak{H}_1$, $\mathfrak{H}_2, \ldots$, $\mathfrak{H}_d$ из состояний схемы Маркова следующего вида: $\mathfrak{H}_1 = \{\{s_i\}\colon$ для некоторого целого числа $t \geqslant 0$ будет $p_{1,i}(td) > 0\}$, $\mathfrak{H}_2 = \{\{s_i\}\colon$ для некоторого целого числа $t \geqslant 0$ будет $p_{1,i}(td+1) > 0\}$, $\ldots$, $\mathfrak{H}_d = \{\{s_i\}\colon$ для некоторого целого числа $t \geqslant 0$ будет $p_{1,i}(td+d-1) > 0\}$. Итак, при $1 \leqslant l \leqslant d$ состояние $\{s_i\} \in \mathfrak{H}_l$, если существует такое целое и неотрицательное число t, что $p_{1,i}(td+l-1) > 0\}$. Имеет место следующее утверждение.

Теорема 5.8. *Если множество всех состояний схемы Маркова с периодом $d > 1$ является минимальным, то оно разбивается на непересекающиеся подклассы $\mathfrak{H}_1$, $\mathfrak{H}_2, \ldots$, $\mathfrak{H}_d$.*

Доказательство. Покажем от противного, что никакое состояние не может принадлежать одновременно двум подклассам. Пусть, например, при $a \neq l$, $1 \leqslant a, l \leqslant d$ существует такое состояние $\{s_i\}$, что $\{s_i\} \in \mathfrak{H}_a$ и $\{s_i\} \in \mathfrak{H}_l$. Тогда из определения подклассов $\mathfrak{H}_a$, $\mathfrak{H}_l$ выводим, что существуют такие натуральные числа t_1 и t_2, для которых будут выполняться неравенства $p_{1,i}(t_1 d + a - 1) > 0$ и $p_{1,i}(t_2 d + l - 1) > 0$. Так как состояния $\{s_i\}$ и $\{s_1\}$ сообщаются между собой, то найдется такое натуральное число t_3, для которого имеем $p_{i,1}(t_3) > 0$. Из полученных выше трех неравенств легко выводим $p_{1,1}(t_3 + t_1 d + a - 1) > 0$, $p_{1,1}(t_3 + t_2 d + l - 1) > 0$. Состояние $\{s_1\}$ имеет период d, то существуют натуральные числа t_4 и t_5, что $t_3 + t_1 d + a - 1 = t_4 d$, $t_3 + t_2 d + l - 1 = t_5 d$. Отсюда получаем противоречивое равенство $(t_1 - t_2 + t_5 - t_4)\, d = l - a$, так как $a \neq l$, $1 \leqslant a, l \leqslant d$. Следовательно, подклассы $\mathfrak{H}_1$, $\mathfrak{H}_2, \ldots$, $\mathfrak{H}_d$ попарно не пересекаются.

Установим, что $\{\{s_1\}, \{s_2\}, \ldots, \{s_r\}\} = \mathfrak{H}_1 \cup \mathfrak{H}_2 \cup \ldots \cup \mathfrak{H}_d$. В силу определения подмножеств $\mathfrak{H}_1$, $\mathfrak{H}_2, \ldots$, $\mathfrak{H}_d$ имеем $\mathfrak{H}_1 \cup \mathfrak{H}_2 \cup \ldots \cup \mathfrak{H}_d \subset$ $\subset \{\{s_1\}, \{s_2\}, \ldots, \{s_r\}\}$. Далее, любые два состояния рассматриваемой схемы Маркова являются сообщающимися. Поэтому существует натуральное число t_6, что для любого состояния $\{s_j\} \in \{\{s_1\}, \{s_2\}, \ldots, \{s_r\}\}$ имеем неравенство $p_{1,j}(t_6) > 0$. Представим натуральное число t_6 в виде $t_6 = t_7 d + l - 1$, где t_7, l — некоторые неотрицательные целые числа и $1 \leqslant l \leqslant d$. Отсюда получаем $p_{1,j}(t_6) = p_{1,j}(t_7 d + l - 1) > 0$ и $\{s_j\} \in \mathfrak{H}_l$. Следовательно, имеем $\{s_j\} \in \mathfrak{H}_1 \cup \mathfrak{H}_2 \cup \mathfrak{H}_3 \cup \ldots \cup \mathfrak{H}_d$ и $\{\{s_1\}, \{s_2\}, \ldots, \{s_r\}\} = \mathfrak{H}_1 \cup \mathfrak{H}_2 \cup \ldots \cup \mathfrak{H}_d$.

Дополнительно к утверждениям этой теоремы покажем, что движение схемы Маркова из подкласса в подкласс происходит циклически, т. е. система за один шаг при любом $l = 1, 2, \ldots, (d-1)$ из подкласса $\mathfrak{H}_l$ переходит в подкласс $\mathfrak{H}_{l+1}$, а из подкласса $\mathfrak{H}_d$ переходит в подкласс $\{s_i\} \in \mathfrak{H}_1$. Действительно, если $l = 1, 2, \ldots, (d-1)$ и $\{s_i\} \in \mathfrak{H}_l$, то для некоторого целого числа $t_8 \geqslant 0$ имеем $p_{1,i}(t_8 d + l - 1) > 0$. Пусть теперь состояние $\{s_j\}$ такое, что $p_{i,j} > 0$. Тогда будет выполняться неравенство $p_{1,j}(t_8 d + l) > 0$ и, следовательно, состояние $\{s_j\}$ принадлежит подклассу $\mathfrak{H}_{l+1}$. Наконец, если $\{s_i\} \in \mathfrak{H}_d$ и $p_{ij} > 0$, то для некоторого целого и неотрицательного числа t_9 выводим $p_{1,i}(t_9 d + d - 1) > 0$, $p_{1,j}(t_9 d + d) = p_{1,j}((t_9 + 1)d) > 0$. Значит, любое такое состояния $\{s_j\}$ является элементом подкласса $\mathfrak{H}_1$.

Выбирая вместо состояния $\{s_1\}$ любое другое и применяя аналогичные рассуждения, легко убедиться, что разбиение множества состояний схемы Маркова с периодом d на подклассы не зависит от указанного выбора. В силу этого перенумеруем состояния схемы Маркова так, чтобы сначала шли состояния из подкласса $\mathfrak{H}_1$, затем — из подкласса $\mathfrak{H}_2$, … и, наконец, в последнюю очередь шли состояния из подкласса $\mathfrak{H}_d$. Из теоремы 5.8 и приведенного дополнительного утверждения легко видеть, что матрица вероятностей перехода за один шаг для периодической схемы Маркова с учетом новой нумерации будет иметь следующую блочную структуру:

$$\Pi = \begin{pmatrix} O & \Pi_{1,2} & O & O & \bullet & \bullet & \bullet & O \\ O & O & \Pi_{2,3} & O & \bullet & \bullet & \bullet & O \\ O & O & O & \Pi_{3,4} & \bullet & \bullet & \bullet & O \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ O & O & \bullet & \bullet & \bullet & O & O & \Pi_{(d-1),d} \\ \Pi_{d,1} & O & O & \bullet & \bullet & \bullet & O & O \end{pmatrix}.$$

Здесь отличные от нуля блоки обозначены через символы $\Pi_{1,2}, \Pi_{2,3}, \ldots, \Pi_{(d-1),d}$, $\Pi_{d,1}$ и все нулевые блоки — одним символом O. Заметим, что вдоль главной диагонали стоят нулевые квадратные блоки. При $l = 1, 2, \ldots, d-1$ блок $\Pi_{l,(l+1)}$ содержит число строк, равное числу состояний подкласса $\mathfrak{H}_l$, и имеет число столбцов, равное числу состояний подкласса вида $\mathfrak{H}_{l+1}$. Наконец, число строк блока $\Pi_{d,1}$ равно числу состояний подкласса $\mathfrak{H}_d$, а число столбцов равно числу состояний подкласса $\mathfrak{H}_1$. После этого замечания можем утверждать, что матрица вероятностей перехода имеет циклический вид.

Пусть схема Маркова с периодом d в начальный момент $t = 1$ находится в некотором состоянии из подкласса $\mathfrak{H}_a$, тогда в моменты

$t = d, 2d, \ldots$ она снова будет находиться в этом подклассе. При этом периодическая схема Маркова циклически проходит подклассы, т.е. в таком порядке: $\mathfrak{H}_a$, $\mathfrak{H}_{a+1}, \ldots$, $\mathfrak{H}_d$, $\mathfrak{H}_1$, $\mathfrak{H}_2, \ldots$, $\mathfrak{H}_{a-1}$, $\mathfrak{H}_a, \ldots$. Поэтому матрица вероятностей переходов за d шагов такой схемы Маркова имеет вид

$$\|p_{i,j}(d)\| = \Pi^d = \begin{pmatrix} \Pi_{1,1}(d) & O & \bullet & \bullet & \bullet & O \\ O & \Pi_{2,2}(d) & O & \bullet & \bullet & O \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ O & O & \bullet & \bullet & \bullet & \Pi_{d,d}(d) \end{pmatrix}.$$

Вспоминая правило умножения блочных матриц, легко найдем $\Pi_{a,a}(d) = \Pi_{a,(a+1)} \times \Pi_{(a+1),(a+2)} \times \ldots \times \Pi_{(d-1),d} \times \Pi_{d,1}\Pi_{1,2} \times \Pi_{2,3} \times \ldots \times \Pi_{(a-1),a}$ при $1 \leqslant a \leqslant d$. В частности, имеем $\Pi_{1,1}(d) = \Pi_{1,2} \times \Pi_{2,3} \times \ldots \times \Pi_{(d-1),d} \times \Pi_{d,1}$ и $\Pi_{d,d}(d) = \Pi_{d,1} \times \Pi_{1,2} \times \Pi_{2,3} \times \ldots \times \Pi_{(d-1),d}$.

Итак, элементы матрицы Π^d равны нулю, за исключением элементов, которые входят в квадратные и стохастические блоки, расположенные на главной диагонали. Поэтому вместо исходной периодической схемы Маркова можно рассматривать d вспомогательных эргодических схем Маркова. Каждая такая схема Маркова с номером $a = 1, 2, \ldots, d$ будет иметь состояния из подкласса $\mathfrak{H}_a$ и матрицу вероятностей перехода за один шаг $\Pi_{a,a}(d)$. Следовательно, для каждого фиксированного $a = 1, 2, \ldots, d$ при $\{s_i\} \in \mathfrak{H}_a$, $\{s_j\} \in \mathfrak{H}_a$ имеем $\lim\limits_{t\to\infty} p_{i,j}(td) = \pi_{a,j} > 0$, $\sum_{j\in C_a} \pi_{a,j} = 1$, где $C_a = \{i\colon \{s_i\} \in \mathfrak{H}_a\}$.

В общем случае, если состояния $\{s_i\} \in \mathfrak{H}_a$ и $\{s_j\} \in \mathfrak{H}_l$, то сразу получаем

$$\lim_{t\to\infty} p_{i,j}(td + z) = \lim_{t\to\infty} \sum_{k\in C_l} p_{i,k}(z)\, p_{k,j}(td) = \\ = \sum_{k\in C_l} p_{i,k}(z) \lim_{t\to\infty} p_{k,j}(td) = \sum_{k\in C_l} p_{i,k}(z)\, \pi_{l,j}.$$

Этот предел, окончательно, равен $\pi_{l,j}$, если $l - a - z$ делится на d, и равен нулю во всех остальных случаях. По теореме о вычислении стационарных распределений предел $\lim\limits_{t\to\infty} t^{-1} \sum_{m=1}^{t} p_{i,j}(m)$ всегда существует и равен λ_{ij} при $i, j = 1, 2, \ldots, r$. Выбирая теперь в последнем пределе частичную подпоследовательность $\left\{(tc)^{-1} \sum_{m=1}^{t_c} p_{i,j}(m);\ c = 1, 2, \ldots\right\}$ с помощью последовательности $\{t_c = cd + z\colon l - a - z = -d, 0;\ 0 \leqslant z < d;\ c = 1, 2, \ldots\}$ из натуральных чисел, легко находим, что для периодической схемы Маркова $\lambda_{ij} = \pi_{l,j}/d$. Получили удивительный факт. Предельное среднее по времени значение вероятности того, что схема Маркова будет находиться в состоянии $\{s_j\}$, не зависит от начального состояния $\{s_i\}$. Например, если $a = 1$ и $l = 2$, то $\{s_i\} \in \mathfrak{H}_1$

и $\{s_j\} \in \mathfrak{H}_2$. В этом случае $z = 1$ и указанная частичная подпоследовательность задается последовательностью чисел $d+1,\ 2d+1, \ldots$, для которой получим, что предел

$$\lim_{c\to\infty}(cd+1)^{-1}\sum_{m=1}^{cd+1} p_{i,j}(m) = d^{-1}\lim_{c\to\infty} c^{-1}\sum_{m=1}^{cd+1} p_{i,j}(m) = \pi_{2,j}/d.$$

Если $a = 3$, $l = 2$, $\{s_i\} \in \mathfrak{H}_3$, $\{s_j\} \in \mathfrak{H}_2$ и $d > 2$, то $z = d - 1$ и получаем, что

$$\lim_{c\to\infty}(cd+d-1)^{-1}\sum_{m=1}^{cd+d-1} p_{i,j}(m) = d^{-1}\lim_{c\to\infty} c^{-1}\sum_{m=1}^{cd+d-1} p_{i,j}(m) = \pi_{2,j}/d.$$

Теоретические результаты для периодических схем Маркова поясним на следующей реальной задаче.

Пример 5.8 (*задача об управлении потоками транспорта и пешеходов*). При функционировании системы регулирования трех конфликтных транспортных потоков на изолированных перекрестках очень часто используется алгоритм с так называемой вызывной кнопкой для пешеходов. Система регулирования или светофор имеет шесть состояний: $\{s_1\}$, $\{s_2\}, \ldots$, $\{s_6\}$. В состоянии $\{s_1\}$ светофор разрешает проезд через перекресток машинам первого направления, переход улиц пешеходам и запрещает движение машин других направлений. В состоянии $\{s_3\}$ светофор разрешает проезд через перекресток только машинам второго направления. В состоянии $\{s_5\}$ светофор разрешает проезд через перекресток лишь машинам третьего направления. Наконец, в состояниях $\{s_2\}$, $\{s_4\}$, $\{s_6\}$ светофор запрещает начинать движение машинам всех направлений и пешеходам и позволяет завершить движение машинам и пешеходам.

Переключение сигналов (состояний) светофора происходит следующим образом. Пусть в некоторый начальный момент, ради определенности, светофор включает состояние $\{s_1\}$. Это означает, что в течение времени T_1 горит зеленый свет для машин первого направления, зеленый свет для пешеходов и одновременно красный свет для машин других направлений. По истечении времени T_1 светофор включает для всех желтый свет или состояние $\{s_2\}$ на время T_2. Затем светофор переходит в состояние $\{s_3\}$, т.е. на время T_3 включает зеленый свет для машин второго направления, красный свет для машин других направлений и красный свет для пешеходов. В течение времени T_3 некоторый пешеход нажимает вызывную кнопку светофора с вероятностью α, где $0 < \alpha < 1$. Если пешеход нажимает кнопку, то после состояния $\{s_3\}$ включается состояние $\{s_1\}$, в противном случае система регулирования переходит в режим $\{s_4\}$ и на время T_4 для всех загорается желтый свет. Из состояния $\{s_4\}$ светофор переходит в состояние $\{s_5\}$ и в течение времени T_5 горит зеленый свет для машин третьего направления, красный свет для машин других направлений и красный свет для пешеходов. После сигнала $\{s_5\}$ включается для всех направлений желтый свет — состояние $\{s_6\}$ на время T_6. Наконец, из режима $\{s_6\}$ светофор переходит в состояние $\{s_1\}$, и вышеописанный алгоритм изменения сигналов светофора повторяется во времени.

Этот алгоритм позволяет выписать элементы матрицы вероятностей перехода этой системы за один такт ее работы в таком виде: $p_{1,2} = p_{2,3} = p_{4,5} = p_{5,6} = p_{6,1} = 1$, $p_{3,1} = \alpha$, $p_{3,4} = 1 - \alpha$. Остальные элементы матрицы вероятностей перехода такой системы за один такт ее работы равны нулю. Этой матрице соответствует изображенный на рис. 5.4 граф переключений состояний светофора.

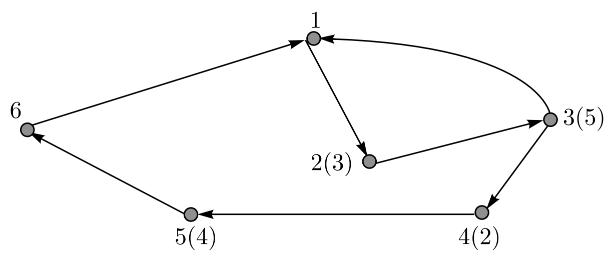


Рис. 5.4

Из рис. 5.4 непосредственно видно, что при любом $i = 1, 2, 3$ условная вероятность $p_{i,i}(t) > 0$ только для $t = 0, 3, 6, 9, \ldots$ и при каждом $i = 4, 5, 6$ условная вероятность $p_{i,i}(t) > 0$ лишь для $t = 0, 6, 9, 12 \ldots$. Отсюда следует, что период d данной схемы Маркова равен трем. Отметим здесь следующую особенность данного эволюционного эксперимента. Хотя период каждого из состояний $\{s_4\}$, $\{s_5\}$, $\{s_6\}$ равен трем, однако возвращение в каждое из указанных состояний невозможно за число шагов, равное трем. Подклассы $\mathfrak{H}_1$, $\mathfrak{H}_2$, $\mathfrak{H}_3$ определяются следующими равенствами:

$$\mathfrak{H}_1 = \{\{s_i\}: \text{для некоторого целого} \quad t \geqslant 0 \quad \text{будет} \\ p_{1,i}(td) > 0\} = \{\{s_1\}, \{s_4\}\},$$

$$\mathfrak{H}_2 = \{\{s_i\}: \text{для некоторого целого} \quad t \geqslant 0 \quad \text{будет} \\ p_{1,i}(td + 1) > 0\} = \{\{s_2\}, \{s_5\}\},$$

$$\mathfrak{H}_3 = \{\{s_i\}: \text{для некоторого целого} \quad t \geqslant 0 \quad \text{будет} \\ p_{1,i}(td + 2) > 0\} = \{\{s_3\}, \{s_6\}\}.$$

В силу такого разбиения на подклассы номера состояний $\{s_1\}$, $\{s_6\}$ оставим без изменения, а состояниям $\{s_2\}$, $\{s_3\}$, $\{s_4\}$ и $\{s_5\}$ присвоим новые номера 3, 5, 2 и 4 соответственно. Новые номера указаны в круглых скобках на рис. 5.4. После таких изменений нумераций состояний матрица вероятностей переключения сигналов светофора за один шаг и матрица вероятностей переключения сигналов светофора за три шага приводятся к виду

$$\Pi = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \alpha & 1-\alpha & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Pi^3 = \begin{pmatrix} \alpha & 1-\alpha & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha & 1-\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 1-\alpha \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

где $\Pi_{1,2} = \Pi_{2,3} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\Pi_{3,1} = \begin{pmatrix} \alpha & 1-\alpha \\ 1 & 0 \end{pmatrix}$ и $\begin{pmatrix} \alpha & 1-\alpha \\ 1 & 0 \end{pmatrix} = \Pi_{1,1}(3) = \Pi_{2,2}(3) = \Pi_{3,3}(3)$. В силу этого для каждого фиксированного $a = 1, 2, 3$ находим, что $\pi_{a,1} = \pi_{a,3} = \pi_{a,5} = 1/(2-\alpha)$ и $\pi_{a,2} = \pi_{a,4} = \pi_{a,6} = (1-\alpha)/(2-\alpha)$.

Теперь, используя формулу $\lambda_{i,j} = \pi_{l,j}/d$ при $l = 1, 2, 3$ и $1 \leqslant j \leqslant 6$, легко получим

$$\lambda_{i,1} = \lambda_{i,3} = \lambda_{i,5} = 1/3(2-\alpha),$$
$$\lambda_{i,2} = \lambda_{i,4} = \lambda_{i,6} = (1-\alpha)/3(2-\alpha)$$

для $1 \leqslant i \leqslant 6$. Отсюда следует замечательный факт. Предельное среднее значение вероятности того, что светофор будет пропускать через перекресток поток машин третьего направления, меньше аналогичных вероятностей для потока машин других направлений и потока пешеходов. Например, при $\alpha = 0{,}5$ имеем $\lambda_{i,1} = \lambda_{i,5} = 2/9$, $\lambda_{i,4} = 1/9$. Поэтому этот алгоритм управления транспортом и пешеходами целесообразно применять, когда транспортные потоки можно разделить на три группы. При этом первое и второе направления движения машин являются наиболее интенсивными, третье направление движения машин оказывается малоинтенсивным и, наконец, пешеходам необходимо создать приоритет в переходе перекрестка. Заметим, что при $\alpha = 0$ получаем известный алгоритм с фиксированным ритмом переключения сигналов светофора [20], а соответствующая этому случаю схема Маркова будет иметь период, равный шести.

3.5. Исследование асимптотического поведения общей схемы Маркова. Пусть теперь множество $\{\{s_1\}, \{s_2\}, \ldots, \{s_r\}\}$ всех возможных состояний схемы Маркова разбивается на минимальные замкнутые множества K_1, $\mathsf{K}_2, \ldots$, K_b из существенных состояний и на множество K_0 всех несущественных состояний. В этом общем случае матрица Π вероятностей перехода за один шаг для схемы Маркова имеет вид (5.17). Непосредственным возведением в степень t матрицы Π, или используя правило умножения блочных матриц, легко убедиться в справедливости равенства

$$\Pi^t = \begin{pmatrix} (\Pi_1)^t & O & O & \bullet & \bullet & \bullet & O & O \\ O & (\Pi_2)^t & O & O & \bullet & \bullet & \bullet & O \\ O & O & (\Pi_3)^t & O & \bullet & \bullet & \bullet & O \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ O & O & \bullet & \bullet & \bullet & O & (\Pi_b)^t & O \\ N_{1,t} & N_{2,t} & \bullet & \bullet & \bullet & N_{(b-1),t} & N_{b,t} & N^t \end{pmatrix}. \tag{5.18}$$

Здесь прямоугольные блоки $N_{1,t}$, $N_{2,t}, \ldots$, $N_{b,t}$ матрицы Π^t при $t = 1, 2, \ldots$ могут быть найдены более простым способом из следующих b равенств:

$$\begin{pmatrix} \Pi_g & O \\ N_g & N \end{pmatrix}^t = \begin{pmatrix} (\Pi_g)^t & O \\ N_{g,t} & N^t \end{pmatrix}, \qquad g = 1, 2, \ldots, b.$$

Обозначим через P_g блочную квадратную матрицу $\begin{pmatrix} \Pi_g & O \\ N_g & N \end{pmatrix}$. Так как $P_g(P_g)^t = (P_g)^{t+1}$, то для матриц $N_{g,t}$, $t = 1, 2, \ldots$, легко получить рекуррентные соотношения вида

$$N_{g,(t+1)} = N_g(\Pi_g)^t + NN_{g,t}, \quad N_{g,1} = N_g. \tag{5.19}$$

Блочная квадратная матрица P_g является нижней квазитреугольной с квадратными диагональными блоками. Тогда характеристические числа этой матрицы совпадают с характеристическими числами квадратных диагональных блоков Π_g и N. Пусть теперь стохастическая матрица Π_g является не только неразложимой, но и примитивной, т. е. ее период $d_g = 1$. Следовательно, матрица Π_g имеет единственное характеристическое число, равное единице, а все остальные характеристические числа этой матрицы по модулю меньше единицы. Далее, у квадратной матрицы N хотя бы одна строчная сумма меньше единицы, то все ее характеристические числа по модулю также меньше единицы. Итак, матрица P_g имеет единственное характеристическое число, равное единице, а все остальные характеристические числа этой матрицы по модулю меньше единицы. Воспользуемся известной формулой для степени t квадратных матриц P_g и N. В результате этого легко установить, что матричные пределы $\lim\limits_{t\to\infty}(P_g)^t$ и $\lim\limits_{t\to\infty}(N)^t$ существуют. Более того, матрица $\lim\limits_{t\to\infty}(N)^t$ является нулевой, так как все характеристические числа матрицы N по модулю меньше единицы.

Если матрица $\lim\limits_{t\to\infty}(N)^t$ является нулевой, то определитель матрицы вида $\mathbf{E} - N$ будет отличен от нуля. Здесь $\mathbf{E}$ — единичная матрица, размерность которой равна числу несущественных состояний. Действительно, из легко проверяемого тождества

$$(\mathbf{E} - N)(\mathbf{E} + N + N^2 + \ldots + N^{t-1}) = \mathbf{E} - N^t$$

по $t = 1, 2, \ldots$ получим

$$\lim_{t\to\infty} \big((\mathbf{E} - N)(\mathbf{E} + N + N^2 + \ldots + N^{t-1})\big) = \mathbf{E}.$$

Определитель матрицы $\mathbf{E}$ равен единице. Следовательно, при достаточно больших t определитель произведения матриц $\mathbf{E} - N$ и $\mathbf{E} + N + N^2 + \ldots + N^{t-1}$ отличен от нуля. Так как определитель произведения двух матриц равен произведению определителей, то определитель матрицы $(\mathbf{E} - N)$ также не равен нулю. На основании этого в дальнейшем мы можем применять матрицу $(\mathbf{E} - N)^{-1}$. Используя теперь тот факт, что матрица $\lim\limits_{t\to\infty} (N)^t$ всегда является нулевой, принимая во внимание блочный вид (5.18) матрицы Π^t и не предполагая здесь обязательного выполнения равенства $d_g = 1$, находим для каждого $g = 1, 2, \ldots, b$ равенство $\lim\limits_{t\to\infty} p_{i,j}(t) = 0$ при любых $\{s_i\} \in \mathsf{K}_g$, $\{s_j\} \notin \mathsf{K}_g$ и равенство $\lim\limits_{t\to\infty} p_{i,j}(t) = 0$ при всех $\{s_i\} \in \mathsf{K}_0$, $\{s_j\} \in \mathsf{K}_0$.

Если снова $d_g = 1$, то, переходя по t к пределу в матричном равенстве (5.19) и обозначая

$$\lim_{t\to\infty} N_{g,t} = L_g, \qquad \lim_{t\to\infty} (\Pi_g)^t = R_g,$$

находим

$$L_g = N_g R_g + N L_g, \quad L_g = (\mathbf{E} - N)^{-1} N_g R_g. \tag{5.20}$$

Из последнего равенства следует, что для определения матрицы L_g предварительно надо вычислить матрицу $(\mathbf{E} - N)^{-1}$ с неотрицательными элементами и найти эргодическое распределение некоторой вспомогательной схемы Маркова. Множество состояний этой схемы Маркова совпадает с множеством из существенных состояний минимального замкнутого класса K_g, а ее матрица вероятностей перехода за один шаг равна Π_g.

Рассмотрим теперь схему Маркова с матрицей вероятностей перехода за один шаг общего типа (5.17), для которой период каждого класса K_g удовлетворяет условию $d_g \geqslant 1$. Предельное поведение условных вероятностей $p_{i,j}(t)$ такой схемы Маркова можно выяснить двумя способами. При первом способе отдельно для каждого $g = 1, 2, \ldots, b$ достаточно изучить асимптотическое поведение матрицы $(P_g)^{d_g t}$ при $t \to \infty$. При втором способе сначала находим наименьшее общее кратное d чисел d_1, $d_2, \ldots$, d_b. Число d называется периодом схемы Маркова.

При фиксированном $g = 1, 2, \ldots, b$ каждый диагональный блок $(\Pi_g)^d$ матрицы

$$\Pi^d = \begin{pmatrix} (\Pi_1)^d & O & O & \bullet & \bullet & \bullet & O & O \\ O & (\Pi_2)^d & O & O & \bullet & \bullet & \bullet & O \\ O & O & (\Pi_3)^d & O & \bullet & \bullet & \bullet & O \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ O & O & \bullet & \bullet & \bullet & O & (\Pi_b)^d & O \\ N_{1,d} & N_{2,d} & \bullet & \bullet & \bullet & N_{(b-1),d} & N_{b,d} & N^d \end{pmatrix}$$

теперь является неразложимой примитивной и стохастической матрицей. Поэтому существует матричный предел $\lim\limits_{t\to\infty}((\Pi)^d)^t = P$. Для определения элементов матрицы P можно полностью воспользоваться методикой предыдущего абзаца. Итак, последовательность Π, Π^2, $\Pi^3, \ldots$ из стохастических матриц определяет асимптотическое поведение условных вероятностей $p_{i,j}(t)$, $i, j = 1, 2, \ldots, r$. Эта последовательность в самом сложном случае содержит d сходящихся подпоследовательностей вида: Π^d, $(\Pi^d)^2, \ldots \to P$; $\Pi\Pi^d$, $\Pi(\Pi^d)^2, \ldots \to \Pi P$; $\Pi^{d-1}\Pi^d$, $\Pi^{d-1}(\Pi^d)^2, \ldots \to \Pi^{d-1}P$. Отсюда для общей схемы Маркова с периодом d легко получить формулу

$$\Lambda = \|\lambda_{i,j}\| = \lim_{t\to\infty} \frac{1}{t} \sum_{m=1}^{t} \Pi^m = \frac{1}{d}(\mathbf{E} + \Pi + \Pi^2 + \ldots + \Pi^{d-1})\, P, \qquad (5.21)$$

где $\mathbf{E}$ — единичная матрица размерности r. Из соотношения (5.21), учитывая равенства $\lim\limits_{t\to\infty} p_{i,j}(t) = 0$, $\{s_i\} \in \mathsf{K}_0$, $\{s_j\} \in \mathsf{K}_0$, непосредственно следует, что матрица Λ имеет блочную структуру:

$$\Lambda = \begin{pmatrix} \Lambda_1 & O & O & \bullet & \bullet & \bullet & O & O \\ O & \Lambda_2 & O & O & \bullet & \bullet & \bullet & O \\ O & O & \Lambda_3 & O & \bullet & \bullet & \bullet & O \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ O & O & \bullet & \bullet & \bullet & O & \Lambda_b & O \\ M_1 & M_2 & \bullet & \bullet & \bullet & M_{b-1} & M_b & O \end{pmatrix},$$

которая аналогична блочному виду матрицы P. При доказательстве теоремы о вычислении стационарных распределений было получено матричное равенство $\Lambda = \Pi\Lambda$. Из этого равенства, вспоминая перемножение блочных матриц, без труда при каждом $g = 1, 2, \ldots, b$ находим равенства

$$M_g = N_g\Lambda_g + NM_g, \quad M_g = (\mathbf{E} - N)^{-1} N_g\Lambda_g, \qquad (5.22)$$

которые полностью аналогичны равенствам (5.20). Далее, пользуясь начальным распределением $\mathbf{p}^{(1)} = (p_1, p_2, \ldots, p_r) = \mathbf{p}$, обозначением

$\mathbf{p}^{(t)} = (p_1^{(t)}, p_2^{(t)}, \ldots, p_r^{(t)})$ и равенством $\mathbf{p}^{(t)} = \mathbf{p}\Pi^{t-1}$, можно выяснить при неограниченном увеличении t поведение вероятности $p_j^{(t)}$ того, что схема Маркова в момент $t = 2, 3, \ldots$ находится в состоянии $\{s_j\}$. Окончательный результат состоит в том, что последовательность $\mathbf{p}^{(1)}$, $\mathbf{p}^{(2)}, \ldots$ из абсолютных распределений разбивается на d сходящихся частичных подпоследовательностей: $\mathbf{p}$, $\mathbf{p}\Pi^d$, $\mathbf{p}\Pi^{2d}, \ldots \to \mathbf{p}P$; $\mathbf{p}\Pi$, $\mathbf{p}\Pi\Pi^d$, $\mathbf{p}\Pi\Pi^{2d}, \ldots \to \mathbf{p}\Pi P$; $\ldots$; $\mathbf{p}\Pi^{d-1}$, $\mathbf{p}\Pi^{d-1}\Pi^d$, $\mathbf{p}\Pi^{d-1}\Pi^{2d}, \ldots \to \mathbf{p}\Pi^{d-1}P$.

Пример 5.9 (*задача об аквариуме*). Применение схем Маркова с матрицей вероятностей перехода за один шаг общего типа (5.17) покажем на следующем простом примере. Неоновую рыбку запускают в аквариум, который разделен стеклянными перегородками на одиннадцать ячеек с метками s_1,s_2,..., s_{11}. Каждая ячейка имеет хотя бы один направленный выход через стеклянные перегородки в некоторую соседнюю ячейку. Неоновая рыбка плавает по ячейкам аквариума случайным образом. В последовательные моменты $t = 1, 2, \ldots$ фиксируется нахождение рыбки в одной из одиннадцати ячеек. Если в некоторый момент t рыбка плавает в ячейке с меткой s_i, то в следующий момент $t+1$ она может оказаться в другой ячейке с меткой s_j, либо остаться в той же ячейке. Под состоянием $\{s_i\}$ этой системы будем понимать тот факт, что неоновая рыбка плавает в ячейке с меткой s_i. Предполагается, что перемещение неоновой рыбки образует схему Маркова с вероятностями перехода за один шаг вида: $p_{1,4} = p_{3,1} = p_{5,1} = p_{9,8} = p_{10,8} = 1$, $p_{11,11} = 2/3$, $p_{2,5} = p_{2,6} = p_{4,3} = p_{4,5} = p_{6,2} = p_{6,11} = p_{7,7} = p_{7,11} = p_{8,9} = p_{8,10} = 1/2$, $p_{11,7} = 1/3$. В остальных случаях вероятности перехода за один шаг для этой схемы Маркова равны нулю.

Ориентированный граф возможных перемещений неоновой рыбки по ячейкам аквариума приведен на рис. 5.5.

Анализируя структуру этого графа, легко выводим, что множество состояний данной схемы Маркова разбивается на три минимальных замкнутых множества $\{\{s_7\}, \{s_{11}\}\}$, $\{\{s_8\}, \{s_9\}, \{s_{10}\}\}$, $\{\{s_1\}, \{s_3\}, \{s_4\}, \{s_5\}\}$ из существенных состояний и на множество $\{\{s_2\}, \{s_6\}\}$ всех несущественных состояний. При этом

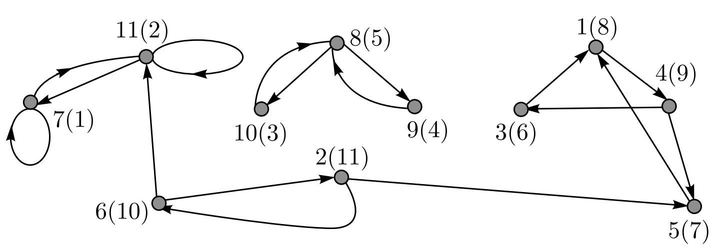


Рис. 5.5

класс $\{\{s_7\}, \{s_{11}\}\}$ является апериодическим, каждый из классов $\{\{s_2\}, \{s_6\}\}$, $\{\{s_8\}, \{s_9\}, \{s_{10}\}\}$ имеет период два и, наконец, класс $\{\{s_1\}, \{s_3\}, \{s_4\}, \{s_5\}\}$ имеет период три. Это разбиение множества состояний данной схемы и периодичность второго и третьего классов позволяет присвоить состояниям новые номера. Эти номера указаны на рис. 5.5 в круглых скобках. В дальнейшем для этого примера мы будем иметь в виду только эти номера. Принимая во внимание новую нумерацию состояний системы, окончательно получим вид замкнутых классов $\mathsf{K}_1 = \{\{s_1\}, \{s_2\}\}$, $\mathsf{K}_2 = \{\{s_3\}, \{s_4\}, \{s_5\}\}$, $\mathsf{K}_3 = \{\{s_6\}, \{s_7\}, \{s_8\}, \{s_9\}\}$ и несущественного класса $\mathsf{K}_0 = \{\{s_{10}\}, \{s_{11}\}\}$. При этом класс K_1 имеет период $d_1 = 1$, подклассы $\mathfrak{H}_{21} = \{\{s_3\}, \{s_4\}\}$, $\mathfrak{H}_{22} = \{\{s_5\}\}$ разбивают класс K_2 с периодом $d_2 = 2$, а подклассы $\mathfrak{H}_{31} = \{\{s_6\}, \{s_7\}\}$, $\mathfrak{H}_{32} = \{\{s_8\}\}$, $\mathfrak{H}_{33} = \{\{s_9\}\}$ составляют класс K_3 с периодом $d_3 = 3$. Если учесть новые номера состояний системы и предыдущие рассуждения, то матрица вероятностей перемещения рыбки за один шаг приводится к блочному виду:

$$\Pi = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 1/2 & 0 \end{pmatrix}.$$

Теперь для матрицы Π с периодом $d = d_1 \times d_2 \times d_3 = 1 \times 2 \times 3 = 6$ легко получить, что

$$\Pi_1 = \begin{pmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{pmatrix}, \quad \Pi_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}, \quad \Pi_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \end{pmatrix},$$

$$N_1 = N_{1,1} = \begin{pmatrix} 0 & 1/2 \\ 0 & 0 \end{pmatrix}, \quad N_2 = N_{2,1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$N_3 = N_{3,1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \end{pmatrix}, \quad N = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix},$$

$$N^3 = \begin{pmatrix} 0 & 1/8 \\ 1/8 & 0 \end{pmatrix}, \quad N^6 = \begin{pmatrix} 1/64 & 0 \\ 0 & 1/64 \end{pmatrix}, \quad N_{2,2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$N_{3,3} = \begin{pmatrix} 0 & 0 & 1/4 & 0 \\ 0 & 1/8 & 0 & 1/2 \end{pmatrix}, \quad (\mathbf{E} - N)^{-1} = \begin{pmatrix} 4/3 & 2/3 \\ 2/3 & 4/3 \end{pmatrix},$$

$$(\mathbf{E} - N^3)^{-1} = \begin{pmatrix} 64/63 & 8/63 \\ 8/63 & 64/63 \end{pmatrix}, \quad \lim_{t\to\infty} (\Pi_1)^t = \begin{pmatrix} 2/5 & 3/5 \\ 2/5 & 3/5 \end{pmatrix},$$

$$\lim_{t\to\infty} N_{1,\,t} = (\mathbf{E} - N)^{-1} N_1 \lim_{t\to\infty} (\Pi_1)^t = \begin{pmatrix} 4/15 & 2/5 \\ 2/15 & 1/5 \end{pmatrix},$$

$$\lim_{t\to\infty} ((\Pi_2)^2)^t = \lim_{t\to\infty} \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}^t = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\lim_{t\to\infty} N_{2,2t} = (\mathbf{E} - N^2)^{-1} \times N_{2,2} \lim_{t\to\infty} ((\Pi_2)^2)^t = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lim_{t\to\infty} ((\Pi_3)^3)^t = \lim_{t\to\infty} \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^t = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\lim_{t\to\infty} N_{3,3t} = (\mathbf{E} - N^3)^{-1} N_{3,3} \quad \times \quad \lim_{t\to\infty} ((\Pi_3)^3)^t.$$

Из последнего равенства выводим, что

$$\lim_{t\to\infty} N_{3,3t} = \begin{pmatrix} 1/126 & 1/126 & 16/63 & 4/63 \\ 4/63 & 4/63 & 2/63 & 32/63 \end{pmatrix}.$$

Так как имеет место соотношение $\lim\limits_{t\to\infty} (\Pi_1)^t = \lim\limits_{t\to\infty} ((\Pi_1)^6)^t$, $\lim\limits_{t\to\infty} N_{1,t} = \lim\limits_{t\to\infty} N_{1,6t}$, $\lim\limits_{t\to\infty} ((\Pi_2)^2)^t = \lim\limits_{t\to\infty} ((\Pi_2)^6)^t$, $\lim\limits_{t\to\infty} N_{2,2t} = \lim\limits_{t\to\infty} N_{2,6t}$, $\lim\limits_{t\to\infty} ((\Pi_3)^3)^t = \lim\limits_{t\to\infty} ((\Pi_3)^6)^t$, $\lim\limits_{t\to\infty} N_{3,3t} = \lim\limits_{t\to\infty} N_{3,6t}$ то для этой задачи можно найти предельную матрицу $P = \lim\limits_{t\to\infty} ((\Pi)^6)^t$ в таком виде:

$$P = \begin{pmatrix}
2/5 & 3/5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2/5 & 3/5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
4/15 & 2/5 & 0 & 0 & 0 & 1/126 & 1/126 & 16/63 & 4/63 & 0 & 0 \\
2/15 & 1/5 & 0 & 0 & 0 & 4/63 & 4/63 & 2/63 & 32/63 & 0 & 0
\end{pmatrix}.$$

Используя формулу (5.21), получим

$$\Lambda = \frac{1}{6}(\mathbf{E} + \Pi + \Pi^2 + \ldots + \Pi^5)P.$$

Отсюда

$$M_1 = \begin{pmatrix} 4/15 & 2/5 \\ 2/15 & 1/5 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$M_3 = \begin{pmatrix} 1/18 & 1/18 & 1/9 & 1/9 \\ 1/9 & 1/9 & 2/9 & 2/9 \end{pmatrix},$$

$$\Lambda_1 = \begin{pmatrix} 2/5 & 3/5 \\ 2/5 & 3/5 \end{pmatrix}, \quad \Lambda_2 = \begin{pmatrix} 1/4 & 1/4 & 1/2 \\ 1/4 & 1/4 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{pmatrix},$$

$$\Lambda_3 = \begin{pmatrix} 1/6 & 1/6 & 1/3 & 1/3 \\ 1/6 & 1/6 & 1/3 & 1/3 \\ 1/6 & 1/6 & 1/3 & 1/3 \\ 1/6 & 1/6 & 1/3 & 1/3 \end{pmatrix}.$$

Наконец, матрица

$$\Lambda = \begin{pmatrix} 2/5 & 3/5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2/5 & 3/5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 1/4 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 1/4 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 1/4 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/3 & 0 & 0 \\ 4/15 & 2/5 & 0 & 0 & 0 & 1/18 & 1/18 & 1/9 & 1/9 & 0 & 0 \\ 2/15 & 1/5 & 0 & 0 & 0 & 1/9 & 1/9 & 2/9 & 2/9 & 0 & 0 \end{pmatrix}.$$

Заметим, что для этой схемы Маркова можно проверить соотношение (5.22). Действительно, при $g = 3$ имеем следующее матричное равенство:

$$\begin{pmatrix} 1/18 & 1/18 & 1/9 & 1/9 \\ 1/9 & 1/9 & 2/9 & 2/9 \end{pmatrix} =$$

$$= \begin{pmatrix} 4/3 & 2/3 \\ 2/3 & 4/3 \end{pmatrix} \times \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 1/6 & 1/6 & 1/3 & 1/3 \\ 1/6 & 1/6 & 1/3 & 1/3 \\ 1/6 & 1/6 & 1/3 & 1/3 \\ 1/6 & 1/6 & 1/3 & 1/3 \end{pmatrix}.$$

Из анализа процесса перемещения неоновой рыбки по ячейкам аквариума можно сделать такие выводы:

1. Стационарные распределения

$$\mathbf{p}_1 = (2/5, 3/5, 0, 0, \ldots, 0), \quad \mathbf{p}_2 = (0, 0, 1/4, 1/4, 1/2, 0, 0, \ldots, 0),$$
$$\mathbf{p}_3 = (0, 0, \ldots, 0, 1/6, 1/6, 1/3, 1/3, 0, 0),$$

которые соответствуют трем минимальным замкнутым классам K_1, K_2 и K_3, порождают континуум стационарных распределений $\lambda_1\mathbf{p}_1 + \lambda_2\mathbf{p}_2 + \lambda_3\mathbf{p}_3$, $\lambda_1 \geqslant 0$, $\lambda_2 \geqslant 0$, $\lambda_3 \geqslant 0$, $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Например, последняя строка $(2/15, 1/5, 0, 0, 0, 1/9, 1/9, 2/9, 2/9, 0, 0)$ матрицы Λ является стационарным распределением и в этом случае $\lambda_1 = 1/3$, $\lambda_2 = 0$, $\lambda_3 = 2/3$.

2. Пусть рыбку помещают плавать в одну из ячеек $\{s_{10}\}$ и $\{s_{11}\}$ аквариума. Тогда рыбка, достигнув пространства, состоящего из секций $\{s_1\}$ и $\{s_2\}$ или из ячеек $\{s_6\}$, $\{s_7\}$, $\{s_8\}$ и $\{s_9\}$, остается

в этом пространстве навсегда. При этом она не может заплыть в ячейки $\{s_3\}$, $\{s_4\}$ и $\{s_5\}$.

3. Если рыбку запускают в часть аквариума, которая состоит из ячеек $\{s_1\}$ и $\{s_2\}$, то она будет перемещаться только по этим ячейкам. Аналогично можно написать два следующих утверждения, которые относятся к замкнутым классам $\mathsf{K}_2 = \{\{s_3\}, \{s_4\}, \{s_5\}\}$ и $\mathsf{K}_3 = \{\{s_6\}, \{s_7\}, \{s_8\}, \{s_9\}\}$.
4. Если рыбку запускают в часть аквариума, которая состоит из ячеек $\{s_3\}$, $\{s_4\}$ и $\{s_5\}$, то она будет перемещаться только по этим ячейкам.
5. Наконец, если неоновую рыбку запускают в часть аквариума, которая состоит из ячеек $\{s_6\}$, $\{s_7\}$, $\{s_8\}$ и $\{s_9\}$, то она будет перемещаться только по этим ячейкам.

Итак, на задаче об аквариуме показан метод построения вероятностной модели для такого простого эксперимента. В качестве вероятностной модели была выбрана однородная схема Маркова. Проведен подробный анализ такой схемы Маркова, который позволяет прогнозировать перемещение неоновой рыбки по аквариуму.

Краткий обзор

В главе 5 рассмотрен способ построения вероятностной модели эволюционного эксперимента, составленного из конечного числа марковских испытаний. При этом каждое испытание имеет конечное число одних и тех же элементарных исходов. Устанавлены условия марковости такого рода эволюционных экспериментов. Предложено конструктивное задание однородной схемы Маркова. Для переходных вероятностей схемы Маркова получено уравнение Колмогорова–Чепмена. На основе доказанной эргодической теоремы Маркова выяснено предельное поведение абсолютных вероятностей для схемы Маркова. Для эргодической схемы Маркова определена скорость сходимости абсолютных вероятностей к предельному распределению. Изучены алгебраические и предельные свойства абсолютного распределения однородной схемы Маркова с конечным числом состояний. Доказана теорема о существовании стационарного распределения. Приводится метод вычисления стационарных распределений схемы Маркова. Предложена геометрическая интерпретация схемы Маркова с точки зрения эволюции реальных систем. Это позволяет схемы Маркова использовать для адекватного описания эволюции самых разнообразных реальных систем. Выделены и изучены следующие стандартные схемы Маркова:

1) эргодические схемы Маркова;
2) регулярные и разложимые марковские системы;
3) апериодические, нерегулярные и разложимые марковские системы;

4) периодические, неразложимые и с одинаковыми строчками матрицы Λ марковские системы;
5) периодические, разложимые и с одинаковыми строчками матрицы Λ марковские системы;
6) периодические, разложимые и с разными строчками матрицы Λ марковские системы.

Приведена методика разбиения пространства состояний схемы Маркова на классы. Подробно изучены так называемые периодические схемы Маркова. Приведено асимптотическое исследование поведения общей схемы Маркова. Теоретические результаты интерпретируются на конкретных примерах.

Контрольные вопросы и упражнения

1. Построить вероятностную модель последовательности зависимых испытаний с двумя элементарными исходами $\{s_1\}$ и $\{s_2\}$, для которой имеет место марковское свойство.
2. Найти явное представление вероятностей перехода за t шагов однородной схемы Маркова с двумя элементарными исходами $\{s_1\}$ и $\{s_2\}$ через элементы $p_{1,2}$ и $p_{2,1}$ матрицы перехода за одно испытание.
3. Пусть некоторый водитель маршрутного такси в течение каждого рабочего дня непреднамеренно нарушает или не нарушает правила дорожного движения. Если в течение некоторого рабочего дня водитель случайно нарушил правила, то, естественно, на следующий день он будет более внимателен и, тем самым, уменьшит вероятность наступления нарушений. Напротив, если некоторый рабочий день водителя оказался удачным, то, как правило, на следующий день он ослабевает контроль и, скорее всего, увеличивает вероятность наступления нарушений. Обозначим через $\{s_1\}$ результат работы водителя, когда он нарушает правила дорожного движения, и через $\{s_2\}$ — результат работы водителя, когда он не нарушает правила дорожного движения. Построить и изучить три конкретные однородные схемы Маркова, которые адекватно описывают различные поведения трех типов водителей маршрутного такси. Найти приближенное значение вероятности нарушений в последний день недели. Какова вероятность работы водителя без нарушений в течение недели?
4. Элементы матрицы вероятностей перехода за одно испытание однородной схемы Маркова с тремя состояниями равны: $p_{1,1} = 0{,}2$; $p_{1,2} = 0{,}3$; $p_{2,1} = 0{,}1$; $p_{2,2} = 0{,}4$; $p_{3,1} = 0{,}5$; $p_{3,2} = 0{,}4$. Для этой схемы Маркова начальное распределение задается вектором $(1, 0, 0)$, т. е. $p_1 = 1$ и $p_2 = p_3 = 0$. Определить абсолютное распределение $\mathbf{p}^{(t)} = (p_1^{(t)}, p_2^{(t)}, p_3^{(t)})$ схемы Маркова в момент $t = 3$. Используя

теорему Маркова, вычислить эргодическое распределение этой последовательности зависимых испытаний.

5. Построить граф переходов состояний схемы Маркова для задачи о процессе протекания болезни и выздоровления. Определить тип этой схемы Маркова.
6. Построить граф переходов состояний схемы Маркова для задачи о развитии экономики. Определить тип такой схемы Маркова.
7. Пусть некоторая футбольная команда в игре на первенство своей страны ведет себя следующим образом. Эта команда выигрывает следующий матч с вероятностями 7/10, 3/5 или 1/10, если предыдущий матч ей был выигран, сыгран вничью или проигран соответственно. Команда делает ничью в следующем матче с вероятностями 1/5, 1/10 или 4/5 соответственно, если предыдущий матч ей был выигран, сыгран вничью или проигран. Наконец, команда проигрывает следующий матч с вероятностями 1/10, 3/10 или 1/10, если предыдущий матч она выиграла, сыграла вничью или проиграла соответственно. Построить вероятностную модель выступления такой футбольной команды в первенстве страны. Проинтерпретировать результаты анализа схемы Маркова, которая соответствует данной задаче.
8. Привести пример неразложимой и апериодической схемы Маркова, для которой матрица вероятностей перехода за один шаг имеет только нулевые диагональные элементы.
9. Схема Маркова имеет четыре состояния. Найти существенные, несущественные, периодические и апериодические состояния, если матрица вероятностей перехода за один шаг для схемы Маркова имеет только следующие ненулевые элементы: $p_{1,1} = p_{1,2} = p_{1,3} = p_{1,4} = 1/4$, $p_{2,3} = p_{3,2} = 1$, $p_{4,1} = 1$. Определить все пары сообщающихся состояний.
10. Доказать, что неразложимая схема Маркова является эргодической, если хотя бы один диагональный элемент ее матрицы вероятностей перехода за один шаг отличен от нуля.
11. При регулировании на 20 перекрестках автомобиль задерживается следующим светофором с вероятностью 1/2, если его остановил предыдущий светофор, и с вероятностью 2/5, когда такого рода остановки не было. Вероятность остановки перед первым светофором равна 3/10. Построить и изучить схему Маркова, которая соответствует такому движению машин на магистрали. Найти приближенное значение вероятности остановки автомашины последним светофором. Какова вероятность проезда без остановок первых шести светофоров?

Глава 6

ВЕРОЯТНОСТНЫЕ МОДЕЛИ ИЗМЕРИТЕЛЕЙ ИСХОДОВ СТАТИСТИЧЕСКИ УСТОЙЧИВОГО ЭКСПЕРИМЕНТА

§ 1. Одномерные случайные величины

1.1. Измерители элементарных исходов. Элементарный результат $\{\omega\}$ характеризует эксперимент E с качественной точки зрения. Вероятность $\mathsf{P}(\{\omega\})$ определяет шанс наступления элементарного исхода $\{\omega\}$. Однако каждому элементарному случайному событию опыта можно дать и количественную характеристику. В этом случае можно говорить, что эксперимент E характеризуется с количественной точки зрения. С этой целью, как правило, над элементарными исходами статистически устойчивых экспериментов с помощью приборов (инструментов) проводят различные измерения. Результатом этих измерений обычно являются некоторые числа. Приведем следующие примеры.

Пример 6.1. Опыт состоит в наблюдении процесса распада некоторого количества радиоактивного вещества. Например, химический элемент радий (Ra) превращается в химический элемент радон (Rn), и при этом излучается альфа-частица — ядро атома гелия (He). С помощью счетчика Гейгера–Мюллера в течение десятиминутного интервала фиксируется число распавшихся атомов. Это число — количественная характеристика наблюдаемого элементарного результата данного эксперимента. Возможными значениями числа распавшихся атомов за каждые десять минут могут быть все неотрицательные целые числа.

Пример 6.2. Эксперимент заключается в испытании некоторого узла космического корабля в течение двухсот часов непрерывной работы. Определяется время работы узла до возникновения отказа. В этом эксперименте количественная характеристика результата измерения его элементарных исходов может принимать значения в отрезке $[0, 200]$.

Пример 6.3. Опыт заключается в случайном бросании симметричной игральной кости с помощью некоторого случайного механизма на поверхность стола. Пусть за выпадение грани, на которой имеется один глазок, засчитывается двадцать очков. За появление грани, на которой имеется два глазка, засчитывается одно очко. При выпадении грани, на которой существует три глазка, засчитывается тридцать очков. Если выпадает одна из неназванных граней, то засчитывается сто очков. Таким образом, каждый эксперимент с бросанием игральной кости характеризуется одним из натуральных чисел множества $\{1, 20, 30, 100\}$.

Можно привести большое число разнообразных реальных экспериментов, в каждом из которых легко подметить следующую общую ситуацию. В результате проведения статистически устойчивого эксперимента E наблюдают некоторый его элементарный исход A_i' с выбранным описанием ω_i. Если над элементарными исходами экспериментов с помощью некоторых инструментов проводят различные измерения, то в результате этих измерений получают набор или множество некоторых чисел. Можно сказать, что с помощью конкретного измерителя получают количественную характеристику статистически устойчивого эксперимента E, которую будем обозначать через $\xi(\omega_i)$. Другими словами, с реальным экспериментом связывают числовую функцию $\xi(\omega)\colon \Omega \to X$ с областью определения Ω и с областью значения $X \subset R = \{x\colon -\infty < x < \infty\}$. При этом значения аргумента ω отображения $\xi(\omega)\colon \Omega \to X$, в отличие от функций, которые рассматриваются в математическом анализе, не находятся в нашем распоряжении, а определяются исходом статистически устойчивого эксперимента E. Итак, некоторые значения аргумента ω функции $\xi(\omega)\colon \Omega \to X$ появляются часто, а некоторые редко, и, следовательно, соответствующие им значения функции также появляются часто и редко. Разумеется, очень важно, какие значения принимает функция $\xi(\omega)\colon \Omega \to X$, но не менее важно уметь находить вероятности тех или иных значений функции $\xi(\omega)\colon \Omega \to X$. Например, уметь определить вероятность того, что функция $\xi(\omega)\colon \Omega \to X$ принимает значения из промежутка $\{x\colon a \leqslant x < b\} = [a, b)$, или из отрезка $\{x\colon a \leqslant x \leqslant b\} = [a, b]$, или из интервала $\{x\colon a < x < b\} = (a, b)$ и т. д. Для иллюстрации этого рассмотрим пример.

Пример 6.4. Разыгрывается денежная лотерея. Участник лотереи называет шесть чисел из множества $\{1, 2, \ldots, 49\}$. Он выигрывает 50 тысяч рублей, если он угадал все шесть чисел, которые заранее выбираются комиссией с помощью некоторого случайного механизма. В противном случае он проигрывает десять рублей. В этом эксперименте $\xi(\omega)$ — число, которое определяет выигрыш в рублях участника лотереи при выборе им шести чисел. Например, конкретный выбор может иметь описание $\omega = \{1, 2, 3, 4, 5, 10\}$. Легко видеть, что для функции $\xi(\omega)\colon \Omega \to X$ область ее значений $X = \{-10, 50\,000\}$. Используя классический способ вычисления вероятностей случайных событий, легко находим, что функция $\xi(\omega)$ принимает значение 50 000 с вероятностью $(C_{49}^6)^{-1} \approx 10^{-8} \times 7,2$ и значение -10 с вероятностью 0,999999928. Таким образом, участник лотереи практически всегда будет проигрывать в этой игре. Поэтому в настоящее время очень мало разумных людей играет в различные лотереи.

В приведенных выше примерах 6.1–6.4 этого параграфа мы имели всегда одну и ту же картину. А именно, с каждым статистически устойчивым экспериментом связывали определенную числовую функцию. При этом, коль скоро эксперимент проведен и зарегистрирован определенный его элементарный исход с описанием ω, то значение функции $\xi(\omega)$ оказывается известным. Далее при определении вероятностей того, что функция $\xi(\omega)\colon \Omega \to X$ будет принимать значения из промежутков вида $[a,b]$, (a,b), $[a,b)$, $(a,b]$, $\{a\}$ и $\{b\}$, необходимо потребовать, чтобы соответствующие им множества $\{\omega\colon a \leqslant \xi(\omega) \leqslant b\}$, $\{\omega\colon a < \xi(\omega) < b\}$, $\{\omega\colon a \leqslant \xi(\omega) < b\}$, $\{\omega\colon a < \xi(\omega) \leqslant b\}$, $\{\omega\colon \xi(\omega) = a\}$ и $\{\omega\colon \xi(\omega) = b\}$ принадлежали $\mathcal{F}$. Другими словами, эти подмножества множества Ω должны быть случайными событиями. Только в этом случае мы можем найти указанные вероятности.

1.2. Математические модели измерителей элементарных событий. Отмеченные ранее некоторые стандартные множества $[a,b]$, (a,b), $[a,b)$, $(a,b]$, $\{a\}$ и $\{b\}$ на действительной прямой R являются для нас привычными. Однако очень часто эти множества не удовлетворяют в полной мере практическим и теоретическим запросам, возникающим при детальном изучении вероятностных свойств измерителей (количественных характеристик) реальных экспериментов.

Переходим теперь к выделению и конструктивному построению других стандартных подмножеств на действительной прямой R, которые называются в теории вероятностей борелевскими. При этом известные теоретико-множественные операции над борелевскими множествами приводят снова к борелевским множествам. Этим фактом частично можно объяснить использование новых стандартных подмножеств на действительной прямой R, наряду с широко известными подмножествами $[a,b]$, (a,b), $[a,b)$, $(a,b]$, $\{a\}$ и $\{b\}$. Итак, пусть по-прежнему $R = \{x\colon -\infty < x < \infty\}$ есть числовая прямая и $\mathsf{K} = \{[a,b)\colon -\infty < a < b < \infty\}$ — класс (множество) всех интервалов вида $[\,a,b) \subset R$. Дополнение до промежутка $[a,\infty)$, естественно, равно $(-\infty,a)$. Поэтому условимся под промежутком $[-\infty,a)$ понимать интервал $(-\infty,a)$. Это соглашение позволяет иметь дело только с промежутками, замкнутыми слева и открытыми справа, при рассмотрении свойств класса K. Множество K позволяет выполнить конструктивное задание так называемой борелевской σ-алгебры на действительной прямой R. С этой целью дадим следующее определение.

Определение 6.1. *Борелевской* σ-алгеброй на действительной прямой называется такая σ-алгебра $\mathcal{B}$ подмножеств из R, для которой $\mathsf{K} \subset \mathcal{B}$ и для любой σ-алгебры $\mathcal{R}$, содержащей K, имеем $\mathcal{B} \subset \mathfrak{R}$.

Можно сказать, что борелевская σ-алгебра $\mathcal{B}$ является наименьшей σ-алгеброй на действительной прямой, содержащей класс K, т. е. $\mathcal{B} = \sigma(\mathsf{K})$ (см. определение 1.15). Аналогично теореме 1.1 имеет место утверждение для борелевской σ-алгебры $\mathcal{B} = \sigma(\mathsf{K})$. Более того, доказательство этого утверждения, если использовать по необходимости те же обозначения для аналогичных математических объектов, во многом повторяет красивый метод доказательства теоремы 1.1. Поэтому имеет смысл еще раз продемонстрировать его универсальность при доказательстве следующего утверждения.

Теорема 6.1. *Борелевская σ-алгебра $\mathcal{B} = \sigma(\mathsf{K})$ на прямой R существует.*

Доказательство. Пусть Ξ — множество всех подмножеств множества R, т. е. $\Xi = \{M : M \subset R\}$. Покажем, что Ξ есть σ-алгебра.

Заметим, что над элементами множества Ξ введены теоретико-множественные операции и что $R \subseteq R$, следовательно, $R \in \Xi$, т. е. Ξ — непустое множество. Далее, если элемент $M \in \Xi$, тогда и $\overline{M} \in \Xi$, так как множество $\overline{M} \in R \backslash M \subset R$. Пусть теперь $M_i \in \Xi$, $i \geqslant 1$, тогда $\cup_{i=1}^{\infty} M_i \subset R$ и множество $\cup_{i=1}^{\infty} M_i \in \Xi$. Так как любой промежуток $[a, b) \subset R$, то и класс $\mathsf{K} \subset \Xi$. Итак, существует хотя бы одна σ-алгебра, содержащая класс K.

Пусть теперь $\mathfrak{R}$ — множество всех σ-алгебр на прямой, которые содержат класс K. Запишем это в виде $\mathfrak{R} = \{\mathfrak{R}_\gamma : \gamma \in \Gamma, \mathfrak{R}_\gamma$ — σ-алгебра на прямой, $\mathsf{K} \subset \mathfrak{R}_\gamma\}$, где Γ — непустое множество индексов. В частности, при некотором $\alpha \in \Gamma$ имеем $\mathfrak{R}_\alpha = \Xi$. Обозначим через $\mathcal{B} = \cap_{\gamma \in \Gamma} \mathfrak{R}_\gamma$. Покажем, что $\mathcal{B}$ — σ-алгебра. Действительно, в каждой σ-алгебре $\mathfrak{R}_\gamma$ введены теоретико-множественные операции, а так как $\mathcal{B} \subset \mathfrak{R}_\gamma$ при любом $\gamma \in \Gamma$, то и на множестве $\mathcal{B}$ эти операции введены. Так как $R \in \mathfrak{R}_\gamma$ при любом $\gamma \in \Gamma$, то $R \in \mathcal{B}$ и множество $\mathcal{B}$ не пустое. Пусть $M \in \cap_{\gamma \in \Gamma} \mathfrak{R}_\gamma$, тогда множество $M \in \mathfrak{R}_\gamma$ при каждом $\gamma \in \Gamma$ и, значит, $\overline{M} \in \mathfrak{R}_\gamma$ при каждом $\gamma \in \Gamma$, следовательно, $\overline{M} \in \cap_{\gamma \in \Gamma} \mathfrak{R}_\gamma = \mathcal{B}$. Пусть $M_i \in \cap_{\gamma \in \Gamma} \mathfrak{R}_\gamma$, $i \geqslant 1$. Покажем, что элемент $M' = \cup_{i=1}^{\infty} M_i \in \cap_{\gamma \in \Gamma} \mathfrak{R}_\gamma$. В самом деле, $M_i \in \mathfrak{R}_\gamma$ при $i \geqslant 1$ и, значит, $M' \in \mathfrak{R}_\gamma$, так как множество $\mathfrak{R}_\gamma$ есть σ-алгебра. Здесь γ — любой фиксированный номер из Γ. Тогда элемент $M' = \cup_{i=1}^{\infty} M_i \in \cap_{\gamma \in \Gamma} \mathfrak{R}_\gamma$.

Итак, $\mathcal{B}$ — σ-алгебра. Так как $\mathsf{K} \subset \mathfrak{R}_\gamma$ при любом $\gamma \in \Gamma$, то и $\mathcal{B} = \cap_{\gamma \in \Gamma} \mathfrak{R}_\gamma \supset \mathsf{K}$. Пусть имеется некоторая σ-алгебра $\mathcal{R} \supset \mathsf{K}$, тогда $\mathcal{R} \in \mathfrak{R}$ и, значит, найдется такой индекс $\beta \in \Gamma$, что $\mathcal{R} = \mathfrak{R}_\beta$. Отсюда следует, что $\cap_{\gamma \in \Gamma} \mathfrak{R}_\gamma = \left(\cap_{\gamma \in \Gamma \backslash \{\beta\}} \mathfrak{R}_\gamma\right) \cap \mathfrak{R}_\beta$. Поэтому $\mathcal{R} = \mathfrak{R}_\beta \supset \left(\cap_{\gamma \in \Gamma} \mathfrak{R}_\gamma\right) = \mathcal{B}$. Значит, $\mathcal{B} = \sigma(\mathsf{K})$ — наименьшая σ-алгебра, включающая K. Другими словами, $\mathcal{B}$ является борелевской σ-алгеброй на действительной прямой.

Рассмотрим примеры стандартных подмножеств на R, которые содержит σ-алгебра $\mathcal{B}$. Так как $\mathcal{B} = \sigma(\mathsf{K}) \supset \mathsf{K}$, то $[a, b) \in \mathcal{B}$. Аналогично рассуждая, получим, что одноточечное множество $\{a\} = \cap_{n=1}^{\infty}[\,a, a+1/n)$ принадлежит $\mathcal{B}$. Так как $(a, b) = [a, b)\backslash\{a\} \in \mathcal{B}$, то интервал $(a, b) \in \mathcal{B}$. Наконец, так как промежуток $(a, b] = (a, b) \cup \{b\}$ и отрезок $[a, b] = [a, b) \cup \{b\}$, то $(a, b] \in \mathcal{B}$ и $[a, b] \in \mathcal{B}$. Таким образом, хорошо известные стандартные подмножества $[a, b]$, (a, b), $[a, b)$, $(a, b]$, $\{a\}$ и $\{b\}$ являются борелевскими. Далее, обозначим через $\mathcal{L}$ множество всех подмножеств действительной прямой R, каждое из которых имеет длину (меру Лебега). В теории функций показывается, что $\mathcal{L}$ является σ-алгеброй. Более того, $\mathcal{L} \supset \mathsf{K}$, так как произвольный элемент $[a, b)$ из K имеет длину $b - a$. Следовательно, $\mathcal{L} \supset \mathcal{B}$. В теории вероятностей, как правило, рассматривается борелевская σ-алгебра $\mathcal{B}$, которая, с одной стороны, содержит такие известные подмножества действительной прямой, как отрезки, интервалы, полузамкнутые интервалы, одноточечные множества, объединения и пересечения в конечном или счетном числе указанных подмножеств, а с другой стороны, борелевская σ-алгебра не содержит некоторые специальные подмножества действительной прямой, которые имеют длину, так как $\mathcal{L} \supset \mathcal{B}$. Борелевская σ-алгебра содержит практически все необходимые для естествоиспытателей подмножества на действительной прямой R. Поэтому борелевская σ-алгебра на действительной прямой играет фундаментальную роль при изучении вероятностных свойств количественных характеристик элементарных событий эксперимента E. Итак, теперь мы имеем весь необходимый материал для строгого определения нового математического объекта — одномерной случайной величины.

Определение 6.2. *Одномерной случайной величиной* $\xi(\omega)$ на $(\Omega, \mathcal{F})$ называется всякое отображение Ω на R, для которого $\{\omega\colon \xi(\omega) \in B\} \in \mathcal{F}$ при любом борелевском множестве B из σ-алгебры $\mathcal{B}$.

Отображение $\xi(\omega)\colon \Omega \to R$, для которого $\{\omega\colon \xi(\omega) \in B\} \in \mathcal{F}$ при любом $B \in \mathcal{B}$, называется *$\mathcal{F}/\mathcal{B}$-измеримым*. Поэтому одномерная случайная величина есть всякое $\mathcal{F}/\mathcal{B}$-измеримое отображение Ω на R. Случайные величины (функциональные зависимости от описания ω) будем обозначать буквами греческого алфавита $\xi(\cdot)$, $\eta(\cdot)$, $\xi_i(\cdot)$, $\eta_i(\cdot)$,… без индексов или с индексами, а их возможные значения — соответственно строчными (малыми) буквами латинского алфавита x, y, x_i, y_i,….

Рассмотрим теперь некоторый частный случай, когда борелевское множество имеет вид $B = (-\infty, x)$, где $\{x\}$ — фиксированная точка на действительной оси. В этом случае множество $\{\omega\colon \xi(\omega) \in (-\infty, x)\} = \{\omega\colon \xi(\omega) < x\}$ будет событием из $\mathcal{F}$ при любом фиксированном $x \in R$, если $\xi(\omega)$ является случайной величиной. Самое удивительное — имеет место обратное утверждение.

Теорема 6.2. *Если $\{\omega\colon \xi(\omega) < x\} \in \mathcal{F}$ для любого действительного числа $x \in R$, то отображение $\xi(\omega)\colon \Omega \to R$ будет случайной величиной.*

Доказательство. Рассмотрим множество $\mathcal{L}_0$ подмножеств M действительной прямой R вида $\mathcal{L}_0 = \{M\colon M \subset R, \{\omega\colon \xi(\omega) \in M\} \in \mathcal{F}\}$. Покажем, что $\mathcal{L}_0$ является σ-алгеброй. Так как $\{\omega\colon \xi(\omega) \in R\} = \Omega \in \mathcal{F}$ и $R \subset R$, то $R \in \mathcal{L}_0$ и $\mathcal{L}_0 \neq \varnothing$. Установим, что для любых $M \subset R$, $M_1 \subset R$, $M_2 \subset R, \ldots$ справедливы следующие два равенства:

$$\begin{gathered}\{\omega\colon \xi(\omega) \in \overline{M}\} = \Omega \backslash \{\omega\colon \xi(\omega) \in M\},\\ \{\omega\colon \xi(\omega) \in \cup_{i=1}^{\infty} M_i\} = \cup_{i=1}^{\infty}\{\omega\colon \xi(\omega) \in M_i\}.\end{gathered} \tag{6.1}$$

В самом деле, пусть $\omega_0 \in \{\omega\colon \xi(\omega) \in \overline{M}\}$. Это означает, что $\xi(\omega_0) \in \overline{M}$, т. е. $\xi(\omega_0) \notin M$ или $\omega_0 \notin \{\omega\colon \xi(\omega) \in M\}$. Но тогда $\omega_0 \in \Omega \backslash \{\omega\colon \xi(\omega) \in M\}$. Обратно, если $\omega_0 \in \Omega \backslash \{\omega\colon \xi(\omega) \in M\}$, то $\omega_0 \notin \{\omega\colon \xi(\omega) \in M\}$. Отсюда $\xi(\omega_0) \notin M$ или $\xi(\omega_0) \in \overline{M}$, т. е. $\omega_0 \in \{\omega\colon \xi(\omega) \in \overline{M}\}$. Следовательно, первое равенство установлено.

Переходим к рассмотрению второго равенства. Предположим, что $\omega_0 \in \{\omega\colon \xi(\omega) \in \cup_{i=1}^{\infty} M_i\}$. Значит, $\xi(\omega_0) \in \cup_{i=1}^{\infty} M_i$, т. е. существует такое натуральное число k, что $\xi(\omega_0) \in M_k$. Поэтому имеем соотношение $\omega_0 \in \{\omega\colon \xi(\omega) \in M_k\}$ и $\omega_0 \in \cup_{i=1}^{\infty}\{\omega\colon \xi(\omega) \in M_i\}$. Наоборот, выберем любой элемент ω_0 из множества $\cup_{i=1}^{\infty}\{\omega\colon \xi(\omega) \in M_i\}$. Тогда существует такое натуральное число m, что $\omega_0 \in \{\omega\colon \xi(\omega) \in M_m\}$, и, следовательно, $\xi(\omega_0) \in M_m$, $\xi(\omega_0) \in \cup_{i=1}^{\infty} M_i$. Поэтому $\omega_0 \in \{\omega\colon \xi(\omega) \in \cup_{i=1}^{\infty} M_i\}$. Это доказывает второе равенство соотношения (6.1).

Пусть теперь $M \in \mathcal{L}_0$. Тогда, во-первых, $M \subset R$, $\overline{M} = R \backslash M \subset R$ и, во-вторых, $\{\omega\colon \xi(\omega) \in \overline{M}\} = \Omega \backslash \{\omega\colon \xi(\omega) \in M\} \in \mathcal{F}$, т. е. $\overline{M} \in \mathcal{L}_0$. Далее, если множества $M_1 \in \mathcal{L}_0$, $M_2 \in \mathcal{L}_0, \ldots$, то $\{\omega\colon \xi(\omega) \in M_i\} \in \mathcal{F}$ и $\cup_{i=1}^{\infty} M_i \subset R$ для каждого натурального числа i. Так как $\mathcal{F}$ является σ-алгеброй, то имеем $\{\omega\colon \xi(\omega) \in \cup_{i=1}^{\infty} M_i\} = \cup_{i=1}^{\infty}\{\omega\colon \xi(\omega) \in M_i\} \in \mathcal{F}$ и, значит, $\cup_{i=1}^{\infty} M_i \in \mathcal{L}_0$. Суммируя все факты, заключаем, что $\mathcal{L}_0$ является σ-алгеброй.

Переходим к завершению доказательства этой теоремы. Легко видеть, что $\{\omega\colon a \leqslant \xi(\omega) < b\} = \{\omega\colon \xi(\omega) < b\} \backslash \{\omega\colon \xi(\omega) < a\}$ при любых a, b. Отсюда и из условий теоремы получаем, что $\{\omega\colon a \leqslant \xi(\omega) < b\} \in \mathcal{F}$. Таким образом, σ-алгебра $\mathcal{L}_0$ содержит все подмножества действительной прямой R вида $[a, b)$, т. е. $\mathcal{L}_0 \supset \mathsf{K}$. На основании определения борелевской σ-алгебры получаем, что $\mathcal{B} \subset \mathcal{L}_0$. Теперь из определения σ-алгебры $\mathcal{L}_0$ можно написать соотношение $\{\omega\colon \xi(\omega) \in B\} \in \mathcal{F}$ при любом множестве B из σ-алгебры $\mathcal{B}$, т. е. $\xi(\omega)$ будет случайной величиной при условиях теоремы.

Замечание 6.1. Здесь следует сделать совершенно необходимое замечание. Утверждение теоремы 6.2 и легко устанавливаемые равенства

$$\begin{aligned}
\{\omega\colon \xi(\omega) \geqslant a\} &= \Omega \backslash \{\omega\colon \xi(\omega) < a\},\\
\{\omega\colon \xi(\omega) \leqslant a\} &= \bigcap_{i=1}^{\infty} \{\omega\colon \xi(\omega) < a + 1/i\},\\
\{\omega\colon \xi(\omega) > a\} &= \Omega \backslash \left(\bigcap_{i=1}^{\infty} \{\omega\colon \xi(\omega) < a + 1/i\} \right),\\
\{\omega\colon a \leqslant \xi(\omega) < b\} &= \{\omega\colon \xi(\omega) < b\} \backslash \{\omega\colon \xi(\omega) < a\},\\
\{\omega\colon a < \xi(\omega) < b\} &= \{\omega\colon \xi(\omega) < b\} \backslash \bigcap_{i=1}^{\infty} \{\omega\colon \xi(\omega) < a + 1/i\},\\
\{\omega\colon \xi(\omega) = a\} &= \bigcap_{i=1}^{\infty} \{\omega\colon \xi(\omega) < a + 1/i\} \backslash \{\omega\colon \xi(\omega) < a\}
\end{aligned} \tag{6.2}$$

позволяют в дальнейшем более подробно рассматривать только свойства вероятностей случайных событий вида $\{\omega\colon \xi(\omega) < x\}$, $x \in R$.

Если теперь принять во внимание результаты этого раздела, то функцию $\xi(\omega)\colon \Omega \to R$ естественно рассматривать в качестве математической модели измерителя случайных элементарных событий эксперимента E.

1.3. Фундаментальное значение требования измеримости для математических моделей количественных характеристик. Рассмотрим теперь простые примеры случайных величин и примеры некоторых функций $\xi(\omega)$, которые по различным причинам не являются случайными величинами. На этих примерах продемонстрируем принципиальное значение ограничения измеримости на отображение вида $\xi(\omega)\colon \Omega \to R$.

Пример 6.5. Пусть симметричная монета бросается два раза с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Для задачи Даламбера некто подсчитывает общее число $\xi(\cdot)$ появлений гербов в этом опыте. Во второй главе была построена теоретико-множественная модель $(\Omega, \mathcal{F})$ для такого эксперимента, где $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $\mathcal{F} = \{A_1, A_2, \ldots, A_{16}\}$. Из определения функции $\xi(\cdot)$ и выбора элементарных исходов задачи Даламбера легко найти ее значения при каждом $\omega \in \Omega$: $\xi(\omega_1) = 2$, $\xi(\omega_2) = \xi(\omega_3) = 1$, $\xi(\omega_4) = 0$. Отсюда непосредственно получаем, что $\{\omega\colon \xi(\omega) < x\} = \varnothing$ при $x \leqslant 0$, $\{\omega\colon \xi(\omega) < x\} = \{\omega_4\}$ при $0 < x \leqslant 1$, $\{\omega\colon \xi(\omega) < x\} = \{\omega_2, \omega_3, \omega_4\}$ при $1 < x \leqslant 2$ и $\{\omega\colon \xi(\omega) < x\}$ равно Ω при $x > 2$. На основании теоремы 6.2 число появлений гербов будет случайной величиной вида $\xi(\omega)\colon \Omega \to \{0, 1, 2\} = X$, так как множества

$\varnothing$, $\{\omega_4\}$, $\{\omega_2, \omega_3, \omega_4\}$ и Ω являются элементами из $\mathcal{F}$ или случайными событиями.

Пример 6.6. Пусть $(\Omega, \mathcal{F})$ есть теоретико-множественная модель произвольного статистически устойчивого эксперимента E и A — некоторое случайное событие из $\mathcal{F}$. Рассмотрим отображение $\xi(\omega)$: $\Omega \to \{0, 1\} = X$, которое принимает значение 1 при $\omega \in A$ и значение 0 при $\omega \in \overline{A}$. Эта функция является случайной величиной, так как $\{\omega\colon \xi(\omega) < x\} = \varnothing \in \mathcal{F}$ при $x \leqslant 0$, $\{\omega\colon \xi(\omega) < x\} = \overline{A} \in \mathcal{F}$ при $0 < x \leqslant 1$ и $\{\omega\colon \xi(\omega) < x\} = \Omega \in \mathcal{F}$ при $x > 1$, т. е. при $x \in R$ множество $\{\omega\colon \xi(\omega) < x\} \in \mathcal{F}$. Такую случайную величину называют индикатором события A и обозначают через $I_A(\omega)$.

Пример 6.7. Подбрасывается монета до первого выпадения герба с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на поверхность стола. В этом случае ω_i при фиксированном $i \geqslant 1$ есть описание такого элементарного исхода, когда $(i-1)$ раз выпадает решетка, а последний раз выпадает герб. Пусть ω_∞ есть описание такого элементарного исхода, когда герб не выпадает. Для этого эксперимента достоверное событие $\Omega = \{\omega_\infty, \omega_1, \omega_2, \ldots\}$, $\mathcal{F} = \{A\colon A \subset \Omega\}$ — множество всех подмножеств множества Ω. Если теперь $\xi(\omega)$ определяет число произведенных бросков в каждом таком опыте, то естественно положить $\xi(\omega_i) = i$ при $i \geqslant 1$ и $\xi(\omega_\infty) = +\infty$. При $x \in R$ получаем, что $\{\omega\colon \xi(\omega) < x\} \subset \Omega$ и, значит, $\{\omega\colon \xi(\omega) < x\} \in \mathcal{F}$. Однако такая функция не будет случайной величиной, так как $\xi(\omega_\infty) \notin R$ и, более того, это отображение имеет вид $\xi(\omega)\colon \Omega \to \{0, 1, \ldots\} \cup \{+\infty\}$.

Пример 6.8. Имеются две симметричные монеты. Одна из этих монет принадлежит отцу семейства, а другая — младшему сыну. Рассмотрим теперь следующую игру отца и младшего сына. Каждый из них в комнате для отдыха подбрасывает один раз свою монету. Сын подбрасывает монету первым и случайным образом. Механизм подбрасывания монеты отцом неизвестен. Отец выигрывает один рубль у младшего сына, если на монете отца выпадает герб и на монете сына решетка, во всех остальных случаях отец проигрывает ему такую же сумму. Старшему сыну после многократного повторения этой игры требуется сделать заключение о механизме подбрасывания монеты отцом в двух различных случаях:

1) старший сын находится в игровой комнате и полностью наблюдает за игрой;
2) старший сын находится в своем кабинете, и ему каждый раз сообщают результаты бросков отца и сына, если на монете брата выпадает герб, в противном случае для него нет дополнительной информации.

В случае 1) для старшего сына достоверное событие

$$\Omega = \{(\Gamma, \mathrm{P}),\ (\Gamma, \Gamma),\ (\mathrm{P}, \mathrm{P}),\ (\mathrm{P}, \Gamma)\},$$

а множество всех наблюдаемых исходов равно множеству всех подмножеств множества Ω, т. е.

$$\begin{gathered}\mathcal{F}_1 = \{\{(\Gamma,\mathrm{P})\},\{(\Gamma,\Gamma)\},\{(\mathrm{P},\mathrm{P})\},\{(\mathrm{P},\Gamma)\},\{(\Gamma,\mathrm{P}),(\Gamma,\Gamma)\},\{(\Gamma,\mathrm{P}),(\mathrm{P},\mathrm{P})\},\\ \{(\Gamma,\mathrm{P}),(\mathrm{P},\Gamma)\},\{(\Gamma,\Gamma),(\mathrm{P},\mathrm{P})\},\{(\Gamma,\Gamma),(\mathrm{P},\Gamma)\},\{(\mathrm{P},\mathrm{P}),(\mathrm{P},\Gamma)\},\\ \{(\Gamma,\mathrm{P}),(\Gamma,\Gamma),(\mathrm{P},\mathrm{P})\},\{(\Gamma,\mathrm{P}),(\Gamma,\Gamma),(\mathrm{P},\Gamma)\},\{(\Gamma,\mathrm{P}),(\mathrm{P},\mathrm{P}),(\mathrm{P},\Gamma)\},\\ \{(\Gamma,\Gamma),(\mathrm{P},\mathrm{P}),(\mathrm{P},\Gamma)\},\{(\Gamma,\mathrm{P}),(\Gamma,\Gamma),(\mathrm{P},\mathrm{P}),(\mathrm{P},\Gamma)\},\varnothing\}.\end{gathered}$$

Заметим, что здесь (Γ, P) есть описание такого элементарного исхода, когда младший сын выбрасывает герб, а отец — решетку. Аналогичный смысл имеют описания вида (Γ, Γ), (P, P), (P, Γ) для остальных элементарных исходов этого эксперимента. Например, вектор (P, Γ) описывает выпадение решетки на монете младшего сына и появление герба на монете отца. Пусть теперь функция $\xi(\omega)\colon \Omega \to \{-1, 1\}$ определяет выигрыш отца в каждой такой игре. Из условий игры находим, что $\xi(\omega) = -1$ при $\omega \in \{(\Gamma, \Gamma), (\mathrm{P}, \mathrm{P}), (\Gamma, \mathrm{P})\}$ и $\xi(\omega) = 1$ при $\omega = (\mathrm{P}, \Gamma)$. Отсюда следует, что $\{\omega\colon \xi(\omega) < x\} = \varnothing \in \mathcal{F}_1$ при $x \leqslant -1$, $\{\omega\colon \xi(\omega) < x\} = \{(\Gamma, \Gamma), (\mathrm{P}, \mathrm{P}), (\Gamma, \mathrm{P})\} \in \mathcal{F}_1$ при $-1 < x \leqslant 1$ и, наконец, $\{\omega\colon \xi(\omega) < x\} = \Omega \in \mathcal{F}_1$ при $x > 1$, т. е. отображение вида $\xi(\omega)\colon \Omega \to \{-1, 1\}$ является случайной величиной.

Предположим, что отец подбрасывает свою монету непреднамеренно (случайно). Тогда старший сын, используя классическое определение вероятности, легко найдет, что отец проигрывает с вероятностью $\mathsf{P}(\{\omega\colon \xi(\omega) = -1\}) = \mathsf{P}(\{(\Gamma, \Gamma), (\mathrm{P}, \mathrm{P}), (\Gamma, \mathrm{P})\}) = 3/4$ и выигрывает с вероятностью $\mathsf{P}(\{\omega\colon \xi(\omega) = 1\}) = \mathsf{P}(\{(\mathrm{P}, \Gamma)\}) = 1/4$. При многократном наблюдении за этой игрой старший сын может найти статистическую вероятность наблюдаемого события $\{(\mathrm{P}, \Gamma)\}$ того, что отец выиграет. Если эта статистическая вероятность близка к числу $1/4$, то отец непреднамеренно бросает монету, и он действительно дает фору младшему сыну или ведет игру в его пользу. Если же эта статистическая вероятность существенно больше числа $1/4$, то отец с помощью некоторого механизма преднамеренно чаще выбрасывает герб, и он на самом деле ведет игру в свою пользу. Например, старший сын по наблюдениям вычислил статистическую вероятность выигрыша отца и она равна 0,45. Тогда механизм подбрасывания монеты отцом таков, что он обеспечивает вероятность появления герба, приблизительно равную 0,9.

Итак, старший сын в задаче 1) может достаточно точно судить о механизме подбрасывания монеты отцом.

В случае 2) для старшего сына по-прежнему достоверное событие $\Omega = \{(\text{Г}, \text{Р}), (\text{Г}, \text{Г}), (\text{Р}, \text{Р}), (\text{Р}, \text{Г})\}$. Однако он не может определить какая из сторон выпала на монете отца, если брат выбросил решетку. Поэтому множество всех наблюдаемых для него исходов будет равно

$$\mathcal{F}_2 = \{\{(\text{Г}, \text{Г})\}, \{(\text{Г}, \text{Р})\}, \{(\text{Р}, \text{Г}), (\text{Р}, \text{Р})\}, \{(\text{Г}, \text{Г}), (\text{Г}, \text{Р})\}, \{(\text{Г}, \text{Г}), (\text{Р}, \text{Г}), (\text{Р}, \text{Р})\}, \{(\text{Г}, \text{Р}), (\text{Р}, \text{Г}), (\text{Р}, \text{Р})\}, \{(\text{Г}, \text{Г}), (\text{Р}, \text{Р}), (\text{Р}, \text{Г}), (\text{Г}, \text{Р})\}, \varnothing\} \subset \mathcal{F}_1.$$

Итак, при $-1 < x \leqslant 1$ имеем $\{\omega\colon \xi(\omega) < x\} = \{(\text{Г}, \text{Г}), (\text{Р}, \text{Р}), (\text{Г}, \text{Р})\} \notin \mathcal{F}_2$. Следовательно, отображение вида $\xi(\omega)\colon \Omega \to \{-1, 1\}$ не является случайной величиной относительно σ-алгебры $\mathcal{F}_2$. В частности, при многократном повторении игры старший сын уже не может вычислить статистическую вероятность выигрыша отца, так как он не в состоянии наблюдать исход $\{(\text{Р}, \text{Г})\}$. Более того, если младший сын выбросил решетку, т. е. старший сын наблюдал событие вида $\{(\text{Р}, \text{Г}), (\text{Р}, \text{Р})\}$ из $\mathcal{F}_2$, то старший сын не может утвердительно ответить на вопрос о том, выиграл ли отец.

Пусть отец с помощью некоторого механизма выбрасывает герб с вероятностью 0,99, если у младшего сына появилась решетка. Естественно при этом считать, что отец выбрасывает решетку с вероятностью 0,01, если у младшего сына появилась решетка. Наконец, отец с равной вероятностью выбрасывает герб или решетку, если у младшего сына появился герб. Используя теоремы умножения и сложения, сразу находим:

$$\mathsf{P}(\{(\text{Г}, \text{Г})\}) = \mathsf{P}(\{(\text{Г}, \text{Р})\}) = 0{,}25;$$
$$\mathsf{P}(\{(\text{Р}, \text{Г}), (\text{Р}, \text{Р})\}) = 0{,}5 \times 0{,}99 + 0{,}5 \times 0{,}01 = 0{,}5;$$
$$\mathsf{P}(\{(\text{Г}, \text{Г}), (\text{Г}, \text{Р})\}) = 0{,}25 + 0{,}25 = 0{,}5;$$
$$\mathsf{P}(\{(\text{Р}, \text{Г}), (\text{Г}, \text{Р}), (\text{Р}, \text{Р})\}) = \mathsf{P}(\{(\text{Г}, \text{Г}), (\text{Р}, \text{Г}), (\text{Р}, \text{Р})\}) = 1 - 0{,}25 = 0{,}75;$$
$$\mathsf{P}(\{(\text{Г}, \text{Г}), (\text{Р}, \text{Р}), (\text{Г}, \text{Р}), (\text{Р}, \text{Г})\}) = 1.$$

Полученные вероятности наблюдаемых событий из $\mathcal{F}_2$, которые соответствуют броскам отца с помощью специального механизма, полностью совпадают с вероятностями этих событий для симметричной задачи или для совершенно случайных и независимых бросков отца и сына. Поэтому у старшего сына в случае 2) нет никакой возможности сделать заключение о механизме подбрасывания монеты отцом.

Заметим, что для случайных и независимых бросков игроков отец выигрывает с вероятностью 0,25, в то же время специальный механизм подбрасывания монеты отцом реализует ему вероятность выигрыша, равную $0{,}495 \approx 0{,}5$, т. е. отец в действительности почти на равных играет с сыном.

Этот пример позволяет сделать важный для практики вывод: если функция $\xi(\omega)\colon \Omega \to R$ не является случайной величиной, то ее вероятностные свойства в общем случае не могут быть определены однозначно.

Рассмотренные простые примеры с одной стороны показывают значение требования измеримости относительно σ-алгебры $\mathcal{F}$ для изучения вероятностных свойств некоторой количественной характеристики вида $\xi(\omega)\colon \Omega \to R$ эксперимента E. С другой стороны, свойство измеримости случайных величин позволяет выделить из множества $\mathcal{F}$ всех наблюдаемых событий некоторые его подмножества. Каждое такое подмножество событий по сравнению с множеством $\mathcal{F}$, как правило, оказывается более простым и доступным для исследования. Это дает нам возможность изучить вероятностные свойства эксперимента E, используя некоторый набор случайных величин. Рассмотрим подробно этот вопрос.

Если ξ является одномерной случайной величиной, то она порождает семейство $\{\{\omega\colon \xi(\omega) \in B\}\colon B \in \mathcal{B}\}$ событий из $\mathcal{F}$. Так как имеет место соотношение (6.1) и $\mathcal{B}$ является σ-алгеброй, то это семейство также образует σ-алгебру. В дальнейшем семейство $\{\{\omega\colon \xi(\omega) \in B\}\colon B \in \mathcal{B}\}$ событий из $\mathcal{F}$ по аналогии с определением 1.15 будем обозначать символом $\sigma(\xi)$, при этом имеем $\sigma(\xi) \subset \mathcal{F}$. В качестве иллюстрации применения понятия σ-алгебры, порожденной случайной величиной ξ, рассмотрим опыт Даламбера (см. примеры 2.7 и 2.39 или пример 6.5).

Пример 6.9. Напомним, что для опыта Даламбера достоверное событие $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ и множество наблюдаемых исходов

$$\begin{gathered}\mathcal{F} = \{A_1, A_2, \ldots, A_{16}\} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_1, \omega_2\},\\ \{\omega_1, \omega_3\}, \{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}, \{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3\},\\ \{\omega_1, \omega_2, \omega_4\}, \{\omega_1, \omega_3, \omega_4\}, \{\omega_2, \omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3, \omega_4\}, \varnothing\}.\end{gathered}$$

Выберем сначала случайную величину $\zeta(\omega_i) = i$, $i = 1, 2, 3, 4$. Случайная величина $\zeta(\cdot)$ определяет номер индекса описания элементарного исхода. Тогда $\{\omega\colon \zeta(\omega) = i\} = \{\omega_i\}$ и, следовательно, $\sigma(\zeta) = \mathcal{F}$. Значит, наблюдая в этом опыте только значение случайной величины ζ, мы в точности знаем, какие исходы из множества $\mathcal{F}$ произошли, а какие нет. Другими словами, случайная величина ζ содержит всю информацию о вероятностных свойствах этого эксперимента.

В примере 6.5 этого раздела была рассмотрена другая случайная величина, $\xi(\cdot)$, которая означает число появлений гербов при двух бросках симметричной монеты. Для случайной величины $\xi(\omega)$ легко найдем: $\sigma(\xi) = \{\{\omega_1\}, \{\omega_4\}, \{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_3\}, \{\omega_2, \omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3, \omega_4\}, \varnothing\} \subset \mathcal{F}$. Случайная величина $\xi(\omega)$ выделяет только восемь событий из множества $\mathcal{F}$. Если случайная величина $\xi(\omega)$ в результате проведения опыта Даламбера приняла значение, равное единице, то относительно событий $\{\omega_2\}$, $\{\omega_3\}$, $\{\omega_1,\omega_2\}$, $\{\omega_1,\omega_3\}$, $\{\omega_2,\omega_4\}$, $\{\omega_3,\omega_4\}$, $\{\omega_1,\omega_2,\omega_4\}$, $\{\omega_1,\omega_3,\omega_4\}$ мы не можем утверждать, произошли они или нет. Таким образом, при изучении только случайной величины $\xi(\omega)$ уже будет потеряна некоторая информация о вероятностных свойствах этого опыта.

Для этого опыта введем дополнительно случайные величины $\xi_1(\omega)$, $\xi_2(\omega)$ и $\xi_3(\omega)$. Здесь $\xi_1(\omega)$ принимает значение, равное двум, если монета падает одной и той же стороной, иначе $\xi_1(\omega)$ подсчитывает число появлений гербов при первом броске. Случайные величины $\xi_2(\omega)$ и $\xi_3(\omega)$ вычисляют число появлений гербов при первом броске и втором соответственно. Из определения этих случайных величин находим:

$$\sigma(\xi_1) = \{\{\omega_2\},\{\omega_3\},\{\omega_1,\omega_4\},\{\omega_2,\omega_3\},\{\omega_1,\omega_2,\omega_4\},$$
$$\{\omega_1,\omega_3,\omega_4\},\{\omega_1,\omega_2,\omega_3,\omega_4\},\varnothing\},$$
$$\sigma(\xi_2) = \{\{\omega_1,\omega_2\},\{\omega_3,\omega_4\},\{\omega_1,\omega_2,\omega_3,\omega_4\},\varnothing\},$$
$$\sigma(\xi_3) = \{\{\omega_1,\omega_3\},\{\omega_2,\omega_4\},\{\omega_1,\omega_2,\omega_3,\omega_4\},\varnothing\}.$$

Отсюда имеем равенство $\sigma(\xi)\cup\sigma(\xi_1)\cup\sigma(\xi_2)\cup\sigma(\xi_3)=\mathcal{F}$. Значит, вероятностные свойства семейства случайных величин $\xi(\omega)$, $\xi_1(\omega)$, $\xi_2(\omega)$, $\xi_3(\omega)$ определяют вероятностные свойства этого эксперимента.

Пусть теперь $\eta(\omega)$ является индикатором ожидаемого по Даламберу события $A_{11}=\{\omega_1,\omega_2,\omega_3\}$, т. е. случайная величина $\eta(\omega)$ равна единице, если хотя бы раз появится герб, и равна нулю, если решетка появится два раза. Нетрудно получить, что случайная величина $\eta(\omega)$ в эксперименте Даламбера порождает σ-алгебру $\sigma(\eta)=\{\Omega,A_{11},\{\omega_4\},\varnothing\}$, которая включает всего четыре события из $\mathcal{F}$. Можно сказать, что количественный признак $\eta(\omega)$ содержит бедную информацию о свойствах эксперимента Даламбера. Наконец, если случайная величина $\chi(\omega)$ есть число бросков в опыте Даламбера, то $\chi(\omega)\equiv 2$ и $\sigma(\chi)=\{\Omega,\varnothing\}$. На содержательном уровне количественный признак $\chi(\omega)$ с точки зрения выяснения вероятностных свойств опыта Даламбера является бесполезным. Сравнивая все рассмотренные случаи примера 6.9, получаем соотношение вида $\sigma(\chi)\subset\sigma(\eta)\subset\sigma(\xi)\subset\sigma(\zeta)$, $\sigma(\zeta)=\sigma(\xi)\cup\sigma(\xi_1)\cup\sigma(\xi_2)\cup\sigma(\xi_3)=\mathcal{F}$. При этом чем больше мощность σ-алгебры количественного признака, тем детальнее можно изучить свойства исходов случайного эксперимента E.

1.4. Арифметические действия и операции предельного перехода над случайными величинами. В результате рассмотрения примера 6.9 на самом деле было показано, что над σ-алгебрами количественных признаков произвольного эксперимента E можно проводить хорошо известные теоретико-множественные операции и операции сравнения с целью получения его вероятностных свойств. Более того, над измеримыми количественными характеристиками $\xi(\omega)$, $\eta(\omega)$, $\xi_1(\omega)$, $\xi_2(\omega), \ldots$ произвольного эксперимента E можно проводить различные арифметические действия и операции предельного перехода, если только все эти действия и операции при любом $\omega \in \Omega$ определены в R. Например, арифметическая операция $\xi(\omega)/\eta(\omega)$ рассматривается, если $\eta(\omega) \neq 0$ для всех $\omega \in \Omega$. Как правило, при $c \in R$ на практике еще применяются операции типа: $c \times \xi(\omega)$, $\xi(\omega) + c$, $\xi(\omega) \pm \eta(\omega)$, $\xi(\omega) \times \eta(\omega)$, $|\xi(\omega)|$, $\sup\limits_n \xi_n = \sup\{\xi_n(\omega);\, n = 1, 2, \ldots\}$, $\inf \xi_n = \inf\{\xi_n(\omega);\, n = 1, 2, \ldots\}$, $\overline{\lim}\, \xi_n(\omega)$, $\underline{\lim}\, \xi_n(\omega)$. В результате этих операций мы получаем снова случайные величины.

В качестве примера покажем это утверждение для конкретных операций вида $c\xi(\omega)$, $\xi(\omega) - \eta(\omega)$. Так как $c \in R$, $\xi(\omega) \in R$, $\eta(\omega) \in R$ при любом $\omega \in \Omega$, то $c\xi(\omega) \colon \Omega \to R$ и $\xi(\omega) - \eta(\omega) \colon \Omega \to R$. Множество $\{\omega \colon c\xi(\omega) < x\}$ равно $\varnothing \in \mathcal{F}$ при $c = 0$, $x \leqslant 0$ и равно $\Omega \in \mathcal{F}$ при $c = 0$, $x > 0$. Пусть теперь $c \neq 0$, тогда множество $\{\omega \colon c\xi(\omega) < x\}$ равно $\{\omega \colon \xi(\omega) < x/c\} \in \mathcal{F}$ при $c > 0$ и равно $\{\omega \colon \xi(\omega) > x/c\} \in \mathcal{F}$ при $c < 0$. Итак, $c\xi(\omega)$ является случайной величиной. Далее, для множества R_0 всех рациональных чисел и любого $x \in R$ имеем:

$$\begin{aligned}\{\omega \colon \xi(\omega) - \eta(\omega) < x\} &= \{\omega \colon \xi(\omega) < \eta(\omega) + x\} = \\ &= \bigcup_{r \in R_0} \{\omega \colon \xi(\omega) < r < \eta(\omega) + x\} = \\ &= \bigcup_{r \in R_0} \big(\{\omega \colon \xi(\omega) < r\} \cap \{\omega \colon \eta(\omega) > r - x\}\big) \in \mathcal{F}.\end{aligned}$$

Значит, $\xi(\omega) - \eta(\omega)$ является случайной величиной или измеримой количественной характеристикой эксперимента E.

Для остальных операций доказательство измеримости можно провести аналогичным образом, используя следующие равенства:

$$\{\omega \colon \inf\{\xi_n(\omega);\, n = 1, 2, \ldots\} < x\} = \bigcup_{n=1}^{\infty} \{\omega \colon \xi_n(\omega) < x\},$$

$$\sup\{\xi_n(\omega);\, n = 1, 2, \ldots\} = -\inf\{-\xi_n(\omega);\, n = 1, 2, \ldots\},$$

$$\overline{\lim}\, \xi_n(\omega) = \inf_n \sup_{m \geqslant n} \xi_m, \qquad \underline{\lim}\, \xi_n(\omega) = \sup_n \inf_{m \geqslant n} \xi_m.$$

Наконец, рассмотрим еще один тип операции, а именно, так называемую поточечную сходимость последовательности случайных величин $\xi_1(\omega)$, $\xi_2(\omega), \ldots$ к функции $\xi(\omega)$. Имеет место следующее простое утверждение.

Лемма 6.1. *Если последовательность ξ_1, $\xi_2, \ldots$ поточечно сходится к функции $\xi \colon \Omega \to R$, то $\xi(\omega) = \lim\limits_{n \to \infty} \xi_n(\omega)$ является случайной величиной.*

Доказательство. Это утверждение сразу следует из соотношения вида $\{\omega\colon \overline{\lim}\,\xi_n(\omega) < x\} \in \mathcal{F}$ и следующей цепочки равенств:

$$
\begin{aligned}
&\{\omega\colon \xi(\omega) < x\} = \\
&\quad = \{\omega\colon \lim_{n\to\infty} \xi_n(\omega) < x\} = \{\omega\colon \overline{\lim}\,\xi_n(\omega) = \underline{\lim}\,\xi_n(\omega),\ \overline{\lim}\,\xi_n(\omega) < x\} = \\
&\qquad = \{\omega\colon \overline{\lim}\,\xi_n(\omega) = \underline{\lim}\,\xi_n(\omega)\} \cap \{\omega\colon \overline{\lim}\,\xi_n(\omega) < x\} = \\
&\qquad\qquad = \Omega \cap \{\omega\colon \overline{\lim}\,\xi_n(\omega) < x\} = \{\omega\colon \overline{\lim}\,\xi_n(\omega) < x\} \in \mathcal{F}.
\end{aligned}
$$

Лемма 6.2. *Пусть последовательность* $\xi_1(\omega), \xi_2(\omega), \ldots$ *задана на* $(\Omega, \mathcal{F})$. *Тогда множество* $A = \{\omega\colon \lim\limits_{n\to\infty} \xi_n(\omega)$ существует$\}$ *является событием из* $\mathcal{F}$.

Доказательство. Если $\omega \in A$, то для любого $m = 1, 2, \ldots$ существует такое натуральное число $n(m)$, зависящее от m, что для всех $k > n$ и $i > n$ будет выполняться $|\xi_k(\omega) - \xi_i(\omega)| < 1/m$. Отсюда легко получим, что

$$\omega \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k>n, i>n} \{\omega\colon |\xi_k(\omega) - \xi_i(\omega)| < 1/m\}.$$

Обратно, если

$$\omega \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k>n, i>n} \{\omega\colon |\xi_k(\omega) - \xi_i(\omega)| < 1/m\},$$

то $\lim\limits_{n\to\infty} \xi_n(\omega)$ существует. Используя ранее доказанное соотношение

$$\{\omega\colon |\xi_k(\omega) - \xi_i(\omega)| < 1/m\} \in \mathcal{F},$$

можно написать

$$A = \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k>n, i>n} \{\omega\colon |\xi_k(\omega) - \xi_i(\omega)| < 1/m\} \in \mathcal{F}.$$

Лемма 6.2 доказана.

§ 2. Способы задания одномерных случайных величин

2.1. Функции от случайных величин и критерий функциональной зависимости между случайными величинами. В предыдущем разделе рассмотрены простейшие функции от случайных величин. Остановимся теперь на общей задаче, а именно, на возможности построения или задания новых случайных величин с помощью более сложных функций от заданных случайных величин.

Будем говорить, что функция $g(x)\colon R \to R$ является борелевской или $\mathcal{B}/\mathcal{B}$-измеримой, если для любого борелевского множества B имеем соотношение $\{x\colon g(x) \in B\} = g^{-1}(B) \in \mathcal{B}$. Иначе, полный прообраз $g^{-1}(B)$ любого борелевского множества B есть опять борелевское множество. Класс таких функций достаточно богат. Например, он включает кусочно-непрерывные функции. Покажем следующее свойство.

Лемма 6.3. *Если $\xi(\omega)$ — случайная величина, $g(x)\colon R \to R$ является борелевской функцией и $\eta(\omega) = g(\xi(\omega))$, то $\sigma(\eta) \subset \sigma(\xi) \subset \mathcal{F}$ и, следовательно, сложная функция $\eta(\omega) = g(\xi(\omega))$ будет также случайной величиной.*

Доказательство. Прежде всего, необходимо установить, что $\eta(\omega)$ однозначно отображает множество Ω описаний элементарных исходов в некоторую область действительной прямой или на всю действительную прямую. Действительно, для любого $\omega \in \Omega$ случайная величина $\xi(\omega)$ принимает одно и только одно действительное значение x, т.е. имеем $\xi(\omega) = x \in R$. Функция $g(x)$ также является однозначной и принимает некоторое действительное значение y при любом x. Итак, для любого $\omega \in \Omega$ существует единственное число $y \in R$ такое, что $\eta(\omega) = g(\xi(\omega)) = g(x) = y$, т.е. имеем отображение $\eta(\omega)\colon \Omega \to R$. Далее, так как $g(x)$ является борелевской функцией, то $g^{-1}(B) \in \mathcal{B}$ для любого борелевского множества B. Отсюда получаем, что $\{\omega\colon \eta(\omega) \in B\} = \{\omega\colon g(\xi(\omega)) \in B\} = \{\omega\colon \xi(\omega) \in g^{-1}(B)\} \in \sigma(\xi)$ для любого борелевского множества B. Поэтому имеем $\sigma(\eta) \subset \sigma(\xi) \subset \mathcal{F}$ и, значит, $\eta(\omega)$ является случайной величиной. Лемма 6.3 доказана.

Следующее утверждение является ключевым при решении очень важной проблемы построения такого простого критерия, который дает возможность обнаружить факт существования функциональной зависимости между случайными величинами.

Лемма 6.4. *Пусть случайные величины $\xi(\omega)$ и $\eta(\omega)$, заданные на основном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, таковы, что $\sigma(\eta) \subset \sigma(\xi)$.*

Тогда найдется такая борелевская функция $g(x)\colon R \to R$, что $\eta(\omega) = g(\xi(\omega))$ при каждом $\omega \in \Omega$, т.е. случайные величины $\xi(\omega)$ и $\eta(\omega)$ функционально связаны.

Доказательство. При заданных $m = \pm 0, \pm 1, \ldots$ и $n = 0, 1, \ldots$ рассмотрим случайное событие

$$A_{m,n} = \{\omega\colon m/2^n \leqslant \eta(\omega) < (m+1)/2^n\} \in \sigma(\eta).$$

При любом фиксированном n семейство $\{A_{m,n}\colon m = \pm 0, \pm 1, \ldots\}$ образует полную группу из счетного (конечного или бесконечного) числа попарно несовместимых событий. Так как $\sigma(\eta) \subset \sigma(\xi)$, то

$A_{m,n} \in \sigma(\xi)$ и, следовательно, существует такое борелевское множество $B_{m,n}$ на действительной прямой, что $A_{m,n} = \{\omega \colon \xi(\omega) \in B_{m,n}\}$. Пусть при фиксированных m и n множество $C_{m,n} = B_{m,n} \setminus \cup_{k \neq m} B_{k,n}$. Определим теперь при любом фиксированном n новое семейство $\{C_{m,n} \colon m = \pm 0, \pm 1, \ldots\}$ из попарно непересекающихся борелевских множеств. Из равенств $\{\omega \colon \xi(\omega) \in B_{m,n} \cap B_{i,n}\} = \{\omega \colon \xi(\omega) \in B_{m,n}\} \cap \cap \{\omega \colon \xi(\omega) \in B_{i,n}\} = A_{m,n} \cap A_{i,n} = \varnothing$ при $m \neq i$ следует, что событие $A_{m,n} = \{\omega \colon \xi(\omega) \in C_{m,n}\}$.

Введем $C_n = (C_{0,n} \cup C_{1,n} \cup \ldots) \cup (C_{-1,n} \cup C_{-2,n} \cup \ldots)$ и $\overline{C}_n = R \setminus C_n$ — борелевские множества. Итак, при фиксированном $n = 0, 1, \ldots$ семейство $\{C_{m,n} \colon m = \pm 0, \pm 1, \ldots\} \cup \{\overline{C}_n\}$, которое составляет разбиение действительной прямой R, поставим в соответствие семейству $\{A_{m,n} \colon m = \pm 0, \pm 1, \ldots\}$, которое образует разбиение пространства Ω. Определим функцию вида $r_n(x) \colon R \to R$, положив $r_n(x) = m/2^n$ при $x \in C_{m,n}$ и $r_n(x) = 0$ при $x \in \overline{C}_n$. Заметим, что множество $\{x \colon r_n(x) = m/2^n\} = C_{m,n}$ при $m \neq 0$ и множество $\{x \colon r_n(x) = 0\} = C_{0,n} \cup \overline{C}_n$. Функция $r_n(x) \colon R \to R$ является борелевской, так как при любом значении $y \in R$ имеем $\{x \colon r_n(x) < y\} = \cup_{m, m < 2^n y} \{x \colon r_n(x) = m/2^n\} \in \mathcal{B}$. Теперь следует воспользоваться теоремой 6.2 при $(\Omega, \mathcal{F}) = (R, \mathcal{B})$.

Рассмотрим последовательность $\{\eta_n(\omega); n = 0, 1, \ldots\}$ случайных величин, где $\eta_n(\omega) = m/2^n$ при $\omega \in A_{m,n}$. Из определения события $A_{m,n}$ следует, что $m = [2^n \eta(\omega)]$. Здесь символ $[2^n \eta(\omega)]$ означает целую часть числа $2^n \eta(\omega)$. Значит, $\eta_n(\omega) = [2^n \eta(\omega)]/2^n$ и поэтому $\lim\limits_{n \to \infty} \eta_n(\omega) = \eta(\omega)$ при всех $\omega \in \Omega$. Если $\omega \in A_{m,n}$, то $\xi(\omega) = x \in C_{m,n}$, $r_n(\xi(\omega)) = m/2^n = \eta_n(\omega)$ и $\lim\limits_{n \to \infty} r_n(\xi(\omega)) = \lim\limits_{n \to \infty} \eta_n(\omega) = \eta(\omega)$. Так как семейство $\{A_{m,n} \colon m = \pm 0, \pm 1, \ldots\}$ образует разбиение пространства Ω, то при всех $\omega \in \Omega$ имеем $\lim\limits_{n \to \infty} r_n(\xi(\omega)) = \lim\limits_{n \to \infty} \eta_n(\omega) = \eta(\omega)$ или предел $\lim\limits_{n \to \infty} r_n(x)$ существует при каждом $x \in \{\xi(\omega) = x \colon \omega \in \Omega\} = \xi(\Omega)$. Здесь подмножество $\xi(\Omega)$ действительной прямой R в общем случае может не быть борелевским. Поэтому необходимо продолжить предельную функцию $\lim\limits_{n \to \infty} r_n(x)$ как борелевскую на действительную прямую R. Это можно сделать следующим способом.

Обозначим через B_0 множество $\{x \colon \lim\limits_{n \to \infty} r_n(x)$ существует$\}$, которое по лемме 6.2 при $(\Omega, \mathcal{F}) = (R, \mathcal{B})$ является борелевским. Обозначим этот предел через $r(x)$. В силу предыдущего абзаца очевидно имеем $B_0 \supset \xi(\Omega)$. Определим теперь новую функцию $g(x)$, которая равна $r(x)$ на B_0 и равна нулю на $R \setminus B_0$. Так как $B_0 \supset \xi(\Omega)$, то построенная таким способом функция $g(x)$ удовлетворяет условию $g(\xi(\omega)) = r(\xi(\omega)) = \eta(\omega)$ для любого $\omega \in \Omega$. Более того, функция $g(x) \colon R \to R$ является борелевской. Поскольку, при любых $n = 0, 1, \ldots$ и $B \in \mathcal{B}$ имеем $\{x \colon x \in B_0,\ r_n(x) \in B\} = \{x \colon r_n(x) \in B\} \cap B_0 \in \mathcal{B} \cap B_0$.

Отсюда непосредственно получаем, что функция $r_n(x)\colon B_0 \to R$ является $\mathcal{B} \cap B_0/\mathcal{B}$-измеримым отображением B_0 на R. Используя лемму 6.1 при $(\Omega, \mathcal{F}) = (B_0, \mathcal{B} \cap B_0)$ и поточечную сходимость последовательности $r_0(x)$, $r_1(x), \ldots$ при $x \in B_0$, получаем, что предельная функция $r(x)\colon B_0 \to R$ является $\mathcal{B} \cap B_0/\mathcal{B}$-измеримым отображением B_0 на R. Для любого $B \in \mathcal{B}$ можно написать следующие соотношения: $\{x\colon g(x) \in B\} = \{x\colon x \in B_0,\, g(x) \in B\} \cup \{x\colon x \in R\backslash B_0,\, g(x) \in B\} = \{x\colon x \in B_0,\, r(x) \in B\} \cup \{x\colon x \in R\backslash B_0, g(x) \in B\} \in \mathcal{B}$. Множество $\{x\colon x \in R\backslash B_0, g(x) \in B\}$ равно $R\backslash B_0$, если число ноль является элементом борелевского множества B, и равно пустому множеству в противном случае. Лемма 6.4 доказана.

Пример 6.10. Для опыта Даламбера в примере 6.9 было получено, что $\sigma(\eta) \subset \sigma(\xi)$. Напомним, что случайная величина $\xi(\omega)$ считает число появлений гербов при двух бросках симметричной монеты, а случайная величина $\eta(\omega)$ равна единице, если хотя бы раз появится герб, и равна нулю, если решетка появится два раза. Пусть теперь борелевская функция $g(x)$ равна единице при $x \in \{1, 2\}$ и равна нулю при $x \notin \{1, 2\}$. Непосредственной проверкой убеждаемся, что $\eta(\omega) = g(\xi(\omega))$ для $\omega \in \Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Еще раз заметим, что теоретико-множественная модель $(\Omega, \mathcal{F})$ эксперимента Даламбера была подробно рассмотрена во второй главе.

2.2. Поточечное задание и задание с помощью распределений случайных величин. Поточечное задание случайных величин заключается в том, что для любого $\omega \in \Omega$ указывается численное значение каждой рассматриваемой для этого эксперимента случайной величины, например, $\xi(\omega) = x \in R$, $\eta(\omega) = y \in R$ и т. д. Так как элементарные результаты появляются или не появляются согласно вероятностной функции $\mathsf{P}(\cdot)$, то наблюдаемые значения случайной величины меняются от опыта к опыту случайным образом. Поэтому для полной характеристики свойств случайной величины недостаточно задать множество $\{\xi(\omega)\colon \omega \in \Omega\}$ всех элементов вида $\xi(\omega)$. Необходимо дополнительно указать, с какой вероятностью следует ожидать в каждом отдельном опыте появление того или иного значения случайной величины. В силу этого естественным образом возникает третий способ задания одномерной случайной величины ξ — задание случайной величины с помощью ее распределения. В силу измеримости отображения $\xi(\omega)\colon \Omega \to R$ можно определить вероятности вида $\mathsf{P}(\{\omega\colon \xi(\omega) \in B\})$, $B \in \mathcal{B}$, и изучать их свойства. Семейство или набор $\{\mathsf{P}(\{\omega\colon \xi(\omega) \in B\})\colon B \in \mathcal{B}\}$ из вероятностей называется распределением случайной величины $\xi(\cdot)$. Однако после сделанного замечания 6.1 мы можем рассматривать более простой случай, когда $B = (-\infty, x)$, где $\{x\}$ — любая точка на числовой оси R. В этом частном случае можно определить функцию $F(x)$ с помощью равенства $F(x) = \mathsf{P}(\{\omega\colon \xi(\omega) \in (-\infty, x)\}) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\})$. Функция $F(x)$ называется интегральной функцией распределения случайной величины $\xi(\cdot)$. Интегральная функция устанавливает соответствие между возможными значениями случайной величины и вероятностями некоторых событий из $\mathcal{F}$, связанных определенным образом с этими возможными значениями. Интегральную функцию распределения еще называют интегральным законом распределения случайной величины. Теперь естественно предложить следующее определение.

Определение 6.3. Значение при $x \in R$ интегральной функции распределения $F(x)$ для величины $\xi(\omega)$, определенной на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, равно вероятности события $A_x = \{\omega\colon \xi(\omega) < x\} \in \mathcal{F}$.

Таким образом, по определению полагаем $F(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\})$ или, для краткости записи, иногда будем отображать это в более простой форме $F(x) = \mathsf{P}(\xi(\omega) < x) = \mathsf{P}(\xi < x)$ или $F(x) = \mathsf{P}(\xi < x)$.

Если мы рассматриваем в контексте несколько случайных величин, например, $\xi(\cdot)$, $\eta(\cdot), \ldots$, то естественно обозначать их интегральные функции распределения соответственно через $F_\xi(x)$, $F_\eta(x), \ldots$. Ниже будет показано, что вероятностные свойства случайной величины $\xi(\cdot)$ полностью определяются ее интегральной функцией распределения $F(x)$. В частности, с помощью $F(x)$ довольно просто вычисляется вероятность того, что значения случайной величины $\xi(\cdot)$ принадлежат тому или иному подмножеству вида $[a, b]$, (a, b), $[a, b)$, $(a, b]$, $\{a\}$, $\{b\}$. Нетрудно видеть, что поточечное задание случайной величины $\xi(\cdot)$, т. е. задание значения $\xi(\omega)$ для любого $\omega \in \Omega$, определяет полностью $F(x)$, $x \in R$, но наоборот — не всегда. Другими словами, восстановить случайную величину $\xi(\omega)\colon \Omega \to R$ по ее интегральной функции распределения в общем случае нельзя, так как различные случайные величины могут определять одну и ту же функцию $F(x)$. Таким образом, задание случайной величины с помощью интегральной функции распределения является более бедным в информативном смысле, чем ее поточечное задание. В качестве пояснения последних утверждений приведем пример следующего эксперимента.

Пример 6.11. На отрезок $[0, 1]$ действительной прямой непреднамеренно бросают точку. Это означает, что ни одно место на отрезке $[0, 1]$, куда попадает брошенная точка, не имеет преимущества перед другими. Результат этого эксперимента может состоять в том, что точка взята из некоторого подмножества отрезка $[0, 1]$, например, из промежутка (0, 1/2).

В данном эксперименте элементарным исходом $\{\omega\}$ является выбранная точка на отрезке $[0, 1]$ с описанием в виде абсциссы ω. Здесь достоверный исход $\Omega = [0, 1] = \{\omega\colon 0 \leqslant \omega \leqslant 1\}$, множество наблюдаемых исходов $\mathcal{F}$ есть σ-алгебра борелевских подмножеств на $[0,1]$, наконец, вероятность $\mathsf{P}(A)$ любого борелевского подмножества (случайного события) $A \subset [0,1]$ равна его длине. Например, вероятность случайного события $\{\omega\colon 0 \leqslant \omega < 1/2\}$ равна $1/2$. Определим поточечно две функции $\xi(\omega)$ и $\eta(\omega)$ на Ω следующим образом. Функция $\xi(\omega)$ равна единице при $\omega \in [0,1/2)$ и равна нулю при $\omega \in [1/2,1]$, а функция $\eta(\omega)$ равна единице при $\omega \in [0,1/4) \cup [1/2,3/4)$ и равна нулю при $\omega \in [1/4,1/2) \cup [3/4,1]$. Графики этих функции приведены на рис. 6.1.

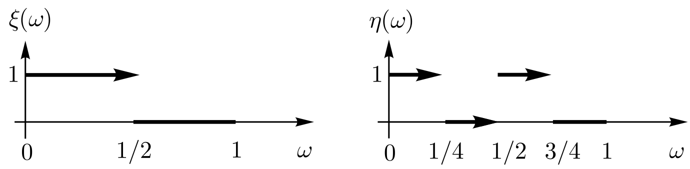


Рис. 6.1

Отметим, что множество $\{\omega\colon \xi(\omega) < x\}$ равно $\varnothing \in \mathcal{F}$ при $x \leqslant 0$, равно отрезку вида $[1/2,1] \in \mathcal{F}$ при $0 < x \leqslant 1$ и равно $\Omega = [0,1] \in \mathcal{F}$ при $x > 1$. Аналогичным образом находим, что множество $\{\omega\colon \eta(\omega) < x\}$ равно $\varnothing \in \mathcal{F}$ при $x \leqslant 0$, равно промежутку $[1/4,1/2) \cup [3/4,1] \in \mathcal{F}$ при $0 < x \leqslant 1$ и равно $\Omega = [0,1] \in \mathcal{F}$ при $x > 1$. Следовательно, различные функции $\xi(\omega)$ и $\eta(\omega)$ на Ω являются случайными величинами. Теперь легко получить, что интегральная функция распределения $F_\xi(x)$ для величины $\xi(\omega)$ равна нулю при $x \leqslant 0$, равна $1/2$ при $0 < x \leqslant 1$ и равна единице при $x > 1$. Интегральная функция $F_\eta(x)$ для случайной величины $\eta(\omega)$ совпадает с $F_\xi(x)$. На рис. 6.2 приведены графики интегральных функций $F_\xi(x)$ и $F_\eta(x)$, которые соответствуют случайным величинам $\xi(\omega)$ и $\eta(\omega)$.

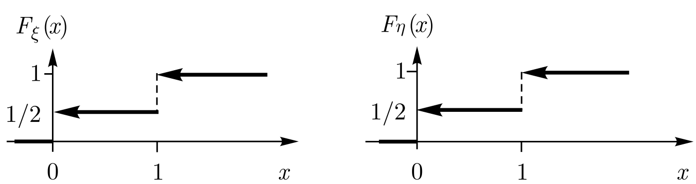


Рис. 6.2

В заключение рассмотрения этого простого примера все же отметим следующее очень важное достоинство интегральной функции распределения, а именно, функция $F_\xi(x)$ содержит всю информацию о вероятностных свойствах случайной величины $\xi(\omega)\colon \Omega \to R$. Приведем еще один показательный пример на построение интегральной функции.

Пример 6.12. Рассмотрим снова задачу о гончарном круге (пример 2.37). Пусть $\xi(\omega)$ определяет длину трещины, которая случайным образом возникает на верхней поверхности гончарного круга. Как уже было подмечено ранее, в зависимости от материала и технологии изготовления гончарного круга мы имеем дело с тремя различными экспериментами.

Первый эксперимент. Пусть технология изготовления гончарного круга такова, что при его использовании трещина в виде хорды появляется наудачу перпендикулярно к некоторому фиксированному направлению $0y$ (см. рис. 2.11). Если ω — абсцисса точки пересечения некоторой трещины (хорды de) с выбранным направлением $0y$, то пространство описаний элементарных исходов опыта имеет вид $\Omega = \{\omega\colon 0 \leqslant \omega \leqslant 2r\}$, а σ-алгебра наблюдаемых исходов $\mathcal{F} = \{A\colon A \subset \Omega,\ A \text{ имеет длину}\}$. Из рис. 2.11 получаем, что длина $\xi(\omega)$ трещины равна $2(r^2 - (r-\omega)^2)^{1/2}$. Из вида функции $\xi(\omega)\colon \Omega \to [0, 2r]$ находим, что множество

$$\{\omega\colon \xi(\omega) < x\} = \{\omega\colon 2(r^2 - (r-\omega)^2)^{1/2} < x\} = \begin{cases} \varnothing \in \mathcal{F} & \text{при } x \leqslant 0, \\ \Omega \in \mathcal{F} & \text{при } x > 2r. \end{cases}$$

Затем при $0 < x \leqslant 2r$ имеем:

$$\begin{gathered} \{\omega\colon \xi(\omega) < x\} = \{\omega\colon 2(r^2 - (r-\omega)^2)^{1/2} < x\} = \\ = \{\omega\colon \omega^2 - 2r\omega + x^2/4 > 0\} = \\ = \{\omega\colon 0 \leqslant \omega < r - (r^2 - x^2/4)^{1/2}\} \cup \{\omega\colon r + (r^2 - x^2/4)^{1/2} < \omega \leqslant 2r\} \in \mathcal{F}. \end{gathered}$$

В частности, если $x = 2r$, то множество вида $\{\omega\colon \xi(\omega) < 2r\} = \Omega \setminus \{r\} \in \mathcal{F}$. Следовательно, показали, что длина $\xi(\omega) = 2(r^2 - (r-\omega)^2)^{1/2}$ хорды в первом эксперименте является случайной величиной. Вспоминая вид множества $\{\omega\colon \xi(\omega) < x\}$ при $-\infty < x < +\infty$ и определение геометрической вероятности на прямой, найдем интегральную функцию распределения $F_\xi(x)$ для случайной длины хорды. А именно, интегральная функция распределения $F_\xi(x)$ равна нулю при $x \leqslant 0$, равна $2(r - (r^2 - x^2/4)^{1/2})/2r = 1 - (1 - x^2/4r^2)^{1/2}$ при $0 < x \leqslant 2r$ и, наконец, равна единице при $x > 2r$. На рис. 6.3 приведен график *а* для $F_\xi(x)$.

Второй эксперимент. Пусть теперь технология изготовления гончарного круга порождает трещину гончарного круга в виде хорды, для которой один из концов всегда закреплен в точке — полюсе b, а другой конец непреднамеренно выбирается на окружности (см. рис. 2.12). Напомним, что в этом опыте описание ω элементарного исхода совпадает с углом φ между лучом bO и выбранной наудачу хордой, достоверный исход имеет вид $\Omega = \{\omega\colon -\pi/2 \leqslant \omega \leqslant \pi/2\}$ и, наконец, σ-алгебра наблюдаемых исходов $\mathcal{F} = \{A\colon A \subset \Omega,\ A \text{ имеет длину}\}$. Непосредственно из рис. 2.12 находим, что длина трещины определяется по формуле $\xi(\omega) = 2r\cos\omega$. Из вида функции $\xi(\omega)\colon \Omega \to [0, 2r]$ легко получим, что

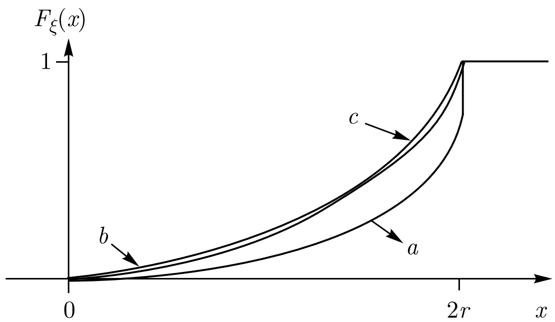


Рис. 6.3

множество $\{\omega\colon \xi(\omega) < x\} = \{\omega\colon 2r\,\cos\omega < x\}$ равно $\varnothing \in \mathcal{F}$ при $x \leqslant 0$, равно $\Omega \in \mathcal{F}$ при $x > 2r$. Наконец, при $0 < x \leqslant 2r$ сразу находим, что

$$\{\omega\colon \xi(\omega) < x\} = \{\omega\colon 2r\,\cos\omega < x\} = \{\omega\colon\ \cos\omega < x/2r\} =$$
$$= \{\omega\colon -\pi/2 \leqslant \omega < -\arccos(x/2r)\} \cup \{\omega\colon\ \arccos(x/r) < \omega \leqslant \pi/2\} \in \mathcal{F}.$$

В частном случае, при $x = 2r$ множество $\{\omega\colon \xi(\omega) < 2r\} = \Omega\backslash\{0\} \in \mathcal{F}$. Следовательно, длина $\xi(\omega) = 2r\,\cos\omega$ наудачу выбранной хорды при втором опыте также является случайной величиной. Используя вид множества $\{\omega\colon \xi(\omega) < x\}$ при $-\infty < x < +\infty$ и определение геометрической вероятности на прямой, найдем интегральную функцию распределения $F_\xi(x)$ для случайной длины хорды. Итак, интегральная функция распределения $F_\xi(x)$ равна нулю при $x \leqslant 0$, равна $1 - (2/\pi)\,\arccos(x/2r)$ при $0 < x \leqslant 2r$ и, наконец, равна единице при $x > 2r$. На рис. 6.3 изображен график b для этой функции.

Третий эксперимент. В этом опыте процесс изготовления гончарного круга (см. рис. 2.13) таков, что его центр является единственной точкой закрепления. Поэтому в результате производства глиняной посуды появление произвольной трещины не связано с выбором направления оси $0y$ или луча bO, как это было в первых двух экспериментах. Заметим, что в этом испытании описание ω элементарного исхода в виде трещины есть упорядоченная пара (u, v). Здесь символы u и v соответственно определяют абсциссу и ординату центра трещины на верхней поверхности гончарного круга радиуса r. Следовательно, в последнем варианте достоверное событие имеет вид $\Omega = \{\omega = (u, v)\colon u^2 + v^2 \leqslant r^2,\ (u, v) \neq (0, 0)\}$, а множество всех наблюдаемых исходов $\mathcal{F}$ равно $\{A\colon A \subset \Omega, A$ имеет площадь$\}$.

Из рис. 2.13 выводим, что длина трещины определяется по формуле $\xi(\omega) = 2(r^2 - u^2 - v^2)^{1/2}\colon \Omega \to [0, 2r)$. Отсюда сразу находим, что множество $\{\omega = (u, v)\colon \xi(\omega) < x\} = \{\omega = (u, v)\colon 2(r^2 - u^2 - v^2)^{1/2} < x\}$ равно $\varnothing \in \mathcal{F}$ при $x \leqslant 0$ и равно $\Omega \in \mathcal{F}$ при $x > 2r$. Если $0 < x \leqslant 2r$, то получим $\{\omega\colon \xi(\omega) < x\} = \{\omega = (u,v)\colon 2(r^2 - u^2 - v^2)^{1/2} < x\} = \{\omega = (u,v)\colon u^2 + v^2 > r^2 - x^2/4\} \in \mathcal{F}$. Например, при $x = 2r$ множество $\{\omega\colon \xi(\omega) < 2r\} = \Omega \in \mathcal{F}$. Таким образом, длина $\xi(\omega) = 2(r^2 - u^2 - v^2)^{1/2}$ наудачу образованной в гончарном круге трещины при третьем эксперименте тоже является случайной величиной. Используя вид множества $\{\omega\colon \xi(\omega) < x\}$ при $-\infty < x < +\infty$ и определение геометрической вероятности на плоскости, найдем функцию $F_\xi(x)$ для случайной длины хорды. Интегральная функция $F_\xi(x)$ равна нулю при $x \leqslant 0$, равна $x^2/4r^2$ при $0 < x \leqslant 2r$ и, наконец, равна единице при $x > 2r$. На рис. 6.3 приведен график c для этой функции.

2.3. Свойства интегральной функции распределения одномерной случайной величины. Перейдем теперь к основным свойствам интегральной функции распределения. Некоторые из этих свойств доказываются очень просто, но мы для единообразия будем формулировать их в виде теорем.

Теорема 6.3. *Интегральная функция распределения $F(x)$ определена при $-\infty \leqslant x \leqslant +\infty$ и принимает значение из отрезка* $[0,1]$.

Доказательство. В самом деле, $\xi(\omega) \neq -\infty$ и $\xi(\omega) \neq +\infty$, так как $-\infty$ и $+\infty$ это не числа, а специальные математические объекты. Суть этих математических объектов состоит в том, что символ «$-\infty$» меньше любого действительного числа, а символ «$+\infty$» больше любого действительного числа. В таком случае множество $\{\omega\colon \xi(\omega) < +\infty\} = \Omega$, так как неравенство $\xi(\omega) < +\infty$ выполняется при любом $\omega \in \Omega$, а $\{\omega\colon \xi(\omega) < -\infty\} = \varnothing$, ибо неравенство $\xi(\omega) < -\infty$ не имеет решения. Отсюда, сразу получаем, что $F(+\infty) = \mathsf{P}(\{\omega\colon \xi(\omega) < +\infty\}) = 1$, $F(-\infty) = \mathsf{P}(\{\omega\colon \xi(\omega) < -\infty\}) = 0$. Для остальных значений аргумента $x \in (-\infty, +\infty)$ функции $F(x)$ имеем: $F(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\})$. Вспоминая свойства вероятностной функции, теперь можно написать, что функция $F(x)\colon \{-\infty\} \cup R \cup \{+\infty\} \to [0,1]$.

Следовательно, интегральная функция распределения $F(x)$ произвольной случайной величины $\xi(\omega)$ определена на расширенной действительной прямой $R \cup \{-\infty\} \cup \{+\infty\}$ с областью значений из отрезка $[0,1]$.

Теорема 6.4. *Интегральная функция распределения $F(x)$ любой случайной величины есть неубывающая функция своего аргумента.*

Доказательство. Рассмотрим любые действительные числа a_1 и a_2 такие, что выполняется неравенство $a_1 < a_2$ (рис. 6.4). Покажем, что имеет место неравенство $F(a_1) \leqslant F(a_2)$.

Введем события вида $A_{a_1} = \{\omega\colon \xi(\omega) < a_1\}$, $A_{a_2} = \{\omega\colon \xi(\omega) < a_2\}$ и $A_{a_1,a_2} = \{\omega\colon a_1 \leqslant \xi(\omega) < a_2\}$, которые порождаются случайной величиной $\xi(\omega)$. Покажем, что $A_{a_2} = A_{a_1} \cup A_{a_1,a_2}$. Для этого необходимо

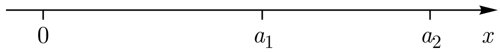


Рис. 6.4

показать, что имеет место: 1) если $\omega \in A_{a_2}$, то $\omega \in A_{a_1} \cup A_{a_1,a_2}$; и 2) если $\omega \in A_{a_1} \cup A_{a_1,a_2}$, то описание $\omega \in A_{a_2}$. Итак, пусть $\omega \in A_{a_2}$, тогда $\xi(\omega) < a_2$. Отсюда, либо выполняется $\xi(\omega) \in (-\infty, a_1)$, либо имеет место соотношение $a_1 \leqslant \xi(\omega) < a_2$. Значит, $\omega \in A_{a_1}$ или $\omega \in A_{a_1,a_2}$, т. е. $\omega \in A_{a_1} \cup A_{a_1,a_2}$. Обратно, пусть элемент $\omega \in A_{a_1} \cup A_{a_1,a_2}$, тогда $\xi(\omega) < a_1$ или $a_1 \leqslant \xi(\omega) < a_2$ и, значит, имеет место $\xi(\omega) < a_2$, т. е. $\omega \in A_{a_2}$. Далее, покажем, что $A_{a_1} \cap A_{a_1,a_2} = \varnothing$. Доказательство проводится методом от противного. Предположим, что существует $\omega^* \in \Omega$ такое, что $\omega^* \in A_{a_1} \cap A_{a_1,a_2}$. Следовательно, одновременно выполняются неравенства $\xi(\omega^*) < a_1$ и $a_1 \leqslant \xi(\omega^*) < a_2$, а такого не может быть. Приходим к противоречию, а это значит, что первоначальное предположение о том, что множество $A_{a_1} \cap A_{a_1,a_2} \neq \varnothing$ неверно. Итак, имеем $A_{a_1} \cap A_{a_1,a_2} = \varnothing$. На основании аксиомы счетной аддитивности для $\mathsf{P}(\cdot)$ теперь можно написать равенство $\mathsf{P}(A_{a_2}) = \mathsf{P}(A_{a_1}) + \mathsf{P}(A_{a_1,a_2})$. Так как $\mathsf{P}(A_{a_1,a_2}) \geqslant 0$, то $\mathsf{P}(A_{a_2}) \geqslant \mathsf{P}(A_{a_1})$, и, значит, неравенство $F(a_2) \geqslant F(a_1)$ или свойство монотонности для интегральной функции $F(x)$ доказано.

Замечание 6.2. Из этой теоремы получаем важную формулу $\mathsf{P}(A_{a_1,a_2}) = \mathsf{P}(\{\omega\colon a_1 \leqslant \xi(\omega) < a_2\}) = F(a_2) - F(a_1)$. Содержательно эта формула означает хорошо известный факт. А именно, вероятность того, что случайная величина примет значение из полузамкнутого промежутка $[a_1, a_2)$, равна разности значений ее интегральной функции в точках a_2 и a_1.

Используя результаты теорем 6.3 и 6.4, докажем следующее утверждение.

Лемма 6.5. *Множество значений x, в каждом из которых интегральная функция $F(x)$ терпит скачок, или является пустым, или счетным.*

Доказательство. Напомним из курса математического анализа, что функция $F(x)$ имеет при $x = a$ скачок, если $F(a+0) - F(a-0) > 0$. Пусть символ $H_i = \{a\colon F(a+0) - F(a-0) = h_i,\ 2^{-i} < h_i \leqslant 2^{-i+1}\}$ при каждом фиксированном $i = 1, 2, \ldots$ обозначает множество таких действительных чисел a, для которых величина h_i i-го скачка интегральной функции распределения $F(x)$ принимает значение из промежутка $(2^{-i}, 2^{-i+1}]$. Из теорем 6.3 и 6.4 непосредственно следует, что множество H_1 может содержать не более одного элемента, множество H_2 — не более трех. В общем случае множество H_i может иметь не более $(2^i - 1)$ элементов. Иначе говоря, число скачков интегральной функции $F(x)$ размером $h_i \in (2^{-i}, 2^{-i+1}]$ не превышает $(2^i - 1)$. Каждое из множеств H_i может оказаться пустым, и, следовательно, в этом случае интегральная функция распределения $F(x)$ будет непрерывной (см. пример 6.12). В противном случае, все элементы множества $\cup_{i=1}^{\infty} H_i$ можно пронумеровать, откуда и следует его счетность. Итак, интегральная функция распределения $F(x)$ непрерывна всюду, за исключением, быть может, счетного множества значений аргумента.

Теорема 6.5. *Функция распределения $F(x)$ непрерывна слева для любого действительного числа x.*

Доказательство. Рассмотрим $F(a)$, где a — любое число на действительной оси. Для доказательства утверждения достаточно для любой строго возрастающей числовой последовательности $\{a_n; n = 1, 2, \ldots\}$, которая сходится к числу a (см. рис. 6.5), показать выполнение равенства $F(a-0) = \lim\limits_{n\to\infty} F(a_n) = F(a)$.

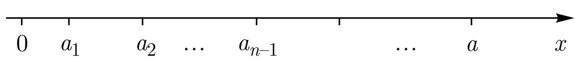


Рис. 6.5

Выберем некоторую числовую последовательность $\{a_n; n \geqslant 1\}$ такую, что $a_1 < a_2 < \ldots < a_n < a_{n+1} < \ldots < a$ и $\lim\limits_{n\to\infty} a_n = a$. Для доказательства равенства $\lim\limits_{n\to\infty} F(a_n) = F(a)$ введем последовательность $A_1 = \{\omega\colon \xi(\omega) < a_1\}$, $A_2 = \{\omega\colon \xi(\omega) < a_2\}, \ldots$, $A_n = \{\omega\colon \xi(\omega) < a_n\}, \ldots$ из множеств. В силу измеримости функции $\xi(\omega)$ все эти множества являются случайными событиями и принадлежат некоторой σ-алгебре $\mathcal{F}$. Заметим, что верно равенство $A = \{\omega\colon \xi(\omega) < a\} = \cup_{n=1}^{\infty}\{\omega\colon \xi(\omega) < a_n\}$. Из определения последовательностей $\{a_n; n \geqslant 1\}$, $\{A_n; n \geqslant 1\}$ следует, что $A_1 \subset A_2 \subset \ldots$. Поэтому $\lim\limits_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty}\{\omega\colon \xi(\omega) < a_n\}$. Далее, используя определение $F(a)$ и свойство непрерывности сверху вероятностной функции (вторую аксиому непрерывности), легко находим равенства

$$F(a) = \mathsf{P}(\{\omega\colon \xi(\omega) < a\}) = \mathsf{P}\Big(\bigcup_{n=1}^{\infty}\{\omega\colon \xi(\omega) < a_n\}\Big) =$$
$$= \mathsf{P}\Big(\lim_{n\to\infty}\{\omega\colon \xi(\omega) < a_n\}\Big) = \lim_{n\to\infty}\mathsf{P}(\{\omega\colon \xi(\omega) < a_n\}) =$$
$$= \lim_{n\to\infty} F(a_n) = F(a-0).$$

Теорема доказана.

Теорема 6.6. *Вероятность* $\mathsf{P}(\{\omega\colon \xi(\omega) \leqslant a\})$ *равна* $F(a+0)$, $a \in R$.

Доказательство. Рассмотрим $F(a)$, где a — любое число, и последовательность $\{a_n; n = 1, 2, \ldots\}$ такую, что $a_1 > a_2 > \ldots > a_n > a_{n+1} > \ldots > a$ и $\lim\limits_{n\to\infty} a_n = a$ (рис. 6.6). Покажем, что $\lim\limits_{n\to\infty} F(a_n) = \mathsf{P}(\{\omega\colon \xi(\omega) \leqslant a\})$.

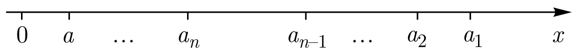


Рис. 6.6

Числовая последовательность $\{a_n; n = 1, 2, \ldots\}$ и случайная величина $\xi(\omega)$ порождают невозрастающую последовательность событий $A_1 = \{\omega\colon \xi(\omega) < a_1\}$, $A_2 = \{\omega\colon \xi(\omega) < a_2\}, \ldots$, $A_n = \{\omega\colon \xi(\omega) < a_n\}, \ldots$, которые принадлежат некоторой σ-алгебре $\mathcal{F}$.

Итак, имеем $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$. Поэтому по свойству непрерывности вероятностной функции снизу и в силу равенства $\{\omega\colon \xi(\omega) \leqslant a\} = \cap_{n=1}^{\infty}\{\omega\colon \xi(\omega) < a_n\}$ получаем соотношение вида: $\mathsf{P}(\{\omega\colon \xi(\omega) \leqslant a\}) = \mathsf{P}(\cap_{n=1}^{\infty}\{\omega\colon \xi(\omega) < a_n\}) = \mathsf{P}(\lim\limits_{n\to\infty}\{\omega\colon \xi(\omega) < a_n\}) = \lim\limits_{n\to\infty}\mathsf{P}(\{\omega\colon \xi(\omega) < a_n\}) = \lim\limits_{n\to\infty}F(a_n) = F(a+0)$. Теорема доказана.

Заметим, что в теоремах 6.5 и 6.6 предполагали a конечным числом. Рассмотрим теперь оставшиеся случаи, когда $a = -\infty$ или $a = +\infty$.

Теорема 6.7. *Для любых последовательностей* $\{a_n; n \geqslant 0\}$ и $\{a_{-n}; n \geqslant 1\}$ *таких, что* $a_0 < a_1 < \ldots < a_n < a_{n+1} < \ldots$, $\lim\limits_{n\to\infty} a_n = +\infty$ *и* $a_0 > a_{-1} > \ldots > a_{-n} > a_{-n-1} > \ldots$, $\lim\limits_{n\to\infty} a_{-n} = -\infty$, *интегральная функция распределения* $F(x)$ *случайной величины удовлетворяет следующим предельным свойствам*: $\lim\limits_{n\to\infty} F(a_n) = F(+\infty) = 1$ *и* $\lim\limits_{n\to\infty} F(a_{-n}) = F(-\infty) = 0$.

Доказательство. Используя предварительные результаты при доказательстве теоремы 6.3 и несовместимость событий $\{\omega\colon a_n \leqslant \xi(\omega) < a_{n+1}\}$, $n = 0, \pm 1, \ldots$, последовательно найдем:

$$1 = F(+\infty) = \mathsf{P}(\{\omega\colon \xi(\omega) < +\infty\}) = \\ = \mathsf{P}\left(\bigcup_{n=-\infty}^{+\infty}\{\omega\colon a_n \leqslant \xi(\omega) < a_{n+1}\}\right) = \sum_{n=-\infty}^{+\infty}\mathsf{P}(\{\omega\colon a_n \leqslant \xi(\omega) < a_{n+1}\}).$$

Так как вероятность $\mathsf{P}(\{\omega\colon a_n \leqslant \xi(\omega) < a_{n+1}\}) \geqslant 0$ для любого $n = 0, \pm 1, \ldots$ и ряд вида

$$\sum_{n=-\infty}^{+\infty}\mathsf{P}(\{\omega\colon a_n \leqslant \xi(\omega) < a_{n+1}\})$$

сходится, то его частичная сумма

$$S_n = \sum_{k=1-n}^{n}\mathsf{P}(\{\omega\colon a_{k-1} \leqslant \xi(\omega) < a_k\}) = \\ = \sum_{k=1-n}^{n}(F(a_k) - F(a_{k-1})) = F(a_n) - F(a_{-n}) \to 1$$

при $n \to +\infty$. Итак, $S_n = F(a_n) - F(a_{-n}) \to 1$ при $n \to +\infty$. Неубывающая последовательность $\{F(a_n); n = 1, 2, \ldots\}$ ограничена сверху единицей, поэтому она имеет предел $\lim\limits_{n\to\infty} F(a_n) = a_+ \leqslant 1$. Невозрастающая последовательность $\{F(a_{-n}); n = 1, 2, \ldots\}$ ограничена снизу нулем, поэтому она имеет предел $\lim\limits_{n\to\infty} F(a_{-n}) = a_- \geqslant 0$. Тогда будет верно предельное равенство $a_+ - a_- = 1$, которое для $0 \leqslant a_-$, $b_+ \leqslant 1$ может выполняться только при $a_+ = 1$ и $a_- = 0$, т. е. $\lim\limits_{n\to\infty} F(a_n) = 1 = F(+\infty)$ и $\lim\limits_{n\to\infty} F(a_{-n}) = 0 = F(+\infty)$.

Замечание 6.3. Утверждения теоремы 6.7 позволяют в дальнейшем для функции распределения $F(x)$ писать соотношения вида: $\lim\limits_{x\to\infty} F(x) = 1$ и $\lim\limits_{x\to-\infty} F(x) = 0$. Далее, из соотношений

$$\{\omega\colon \xi(\omega) \leqslant x\} = \{\omega\colon \xi(\omega) = x\} \cup \{\omega\colon \xi(\omega) < x\},$$

$$\{\omega\colon \xi(\omega) = x\} \cap \{\omega\colon \xi(\omega) < x\} = \varnothing,$$

теоремы 6.6 и равенств (6.2) замечания 6.1 легко находим важные формулы:

$$\begin{aligned}
\mathsf{P}(\{\omega\colon \xi(\omega) = x\}) &= \mathsf{P}(\{\omega\colon \xi(\omega) \leqslant x\}) - \mathsf{P}(\{\omega\colon \xi(\omega) < x\}) = \\
&= F(x+0) - F(x), \\
\mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) \leqslant b\}) &= \mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) < b\}) + \mathsf{P}(\{\omega\colon \xi(\omega) = b\}) = \\
&= F(b) - F(a) + F(b+0) - F(b) = F(b+0) - F(a), \\
\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant x\}) &= 1 - \mathsf{P}(\{\omega\colon \xi(\omega) < x\}) = 1 - F(x), \\
\mathsf{P}(\{\omega\colon \xi(\omega) > x\}) &= 1 - \mathsf{P}(\{\omega\colon \xi(\omega) \leqslant x\}) = 1 - F(x+0), \\
\mathsf{P}(\{\omega\colon a < \xi(\omega) < b\}) &= \mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) < b\}) - \mathsf{P}(\{\omega\colon \xi(\omega) = a\}) = \\
&= F(b) - F(a) - F(a+0) + F(a) = F(b) - F(a+0).
\end{aligned}$$

Таким образом, с учетом замечаний 6.2 и 6.3 вероятность того, что случайная величина примет значение из некоторого стандартного промежутка, определяется ее интегральной функцией.

Обобщим результаты доказанных теорем, сформулировав следующее окончательное

Утверждение. *Если $F(x)$ есть интегральная функция некоторой случайной величины, то необходимо выполнение условий теорем* 6.3–6.5 *и* 6.7, *а утверждение леммы* 6.5 *есть их следствие.*

Современное математическое мировоззрение на реальный мир, в котором закономерности определяются случаем, формулируется следующим образом. Каждому статистически устойчивому эксперименту E с комплексом условий Σ его проведения ставится в соответствие вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. В качестве вероятностных моделей количественных характеристик эксперимента E выступают так называемые случайные величины — измеримые отображения на основном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. При этом все вероятностные свойства количественной характеристики $\xi(\omega)$ определяются ее интегральной функцией распределения $F_\xi(x)$. Интегральная функция распределения $F_\xi(x)$ однозначно определяется, если известны основное вероятностное пространство $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ для эксперимента E и поточечное задание отображения $\xi(\omega)\colon \Omega \to R$.

К сожалению, чаще всего построить $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ или найти поточечное задание отображения $\xi(\omega)\colon \Omega \to R$ не удается в силу сложности конкретного статистически устойчивого эксперимента E. Например, в простейшем случае может быть неизвестна область определения Ω для функции $\xi(\omega)$. Поэтому говорить в этом и других аналогичных вариантах о случайной величине $\xi(\omega)$ и о ее вероятностных свойствах, по меньшей мере, математически некорректно. Однако в результате многократного и достаточно большого числа проведения эксперимента E удается вычислить статистические вероятности $\mathsf{P}^*(\{\omega\colon \xi(\omega) < x\})$, $x \in R$. Эти вероятности на основании интегральной предельной теоремы Муавра–Лапласа восстанавливают функцию $F_\xi(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\})$ с заданной степенью точности (см. замечание 4.1 из четвертой главы). Поэтому на практике, не имея возможности построить математические объекты $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ и $\xi(\omega)\colon \Omega \to R$, исследователи свойств некоторой количественной характеристики эксперимента E все же используют полученную экспериментальным образом интегральную функцию $F_\xi(x)$. Способ применения интегральной функции $F_\xi(x)$ при изучении свойств некоторого количественного признака сложного эксперимента E и обоснование такого подхода дается фактически следующей важной для практики теоремой.

Теорема 6.8 (теорема о выборочном вероятностном пространстве). *Пусть функция $G(x)\colon R \cup \{-\infty\} \cup \{+\infty\} \to [0,1]$ является неубывающей, непрерывной слева и $\lim\limits_{x\to+\infty} G(x) = 1$, $\lim\limits_{x\to-\infty} G(x) = 0$. Тогда существует некоторый искусственный эксперимент $E_\text{и}$, для которого можно построить вероятностную модель $(\Omega_\text{и}, \mathcal{F}_\text{и}, \mathsf{P}_\text{и}(\cdot))$ и случайную величину $\xi_\text{и}(\omega_\text{и})\colon \Omega_\text{и} \to R$ такую, что ее интегральная функция распределения равна $G(x)$.*

Доказательство. Рассмотрим искусственный (модельный) эксперимент $E_\text{и}$, который заключается в бросании точки на действительную прямую R с помощью некоторого случайного механизма. При этом случайный механизм выбора места положения точки на прямой R будет разъяснен в дальнейшем, и он существенно будет зависеть от заданной функции $G(x)$. Исход этого модельного эксперимента может заключаться в том, что точка взята из некоторого стандартного промежутка на действительной прямой, например, из интервала (a, b), или из отрезка $[a, b]$ и т. п. Обозначим через $\omega_{и}$ абсциссу случайным образом поставленной точки на R. В эксперименте $E_{и}$ достоверный исход можно представить в виде $\Omega_{и} = \{\omega_{и}\colon -\infty < \omega_{и} < +\infty\}$, а множество всех наблюдаемых исходов $\mathcal{F}_{и} = \mathcal{B}$ есть σ-алгебра борелевских подмножеств на R. Значит, построена теоретико-множественная модель $(\Omega_{и}, \mathcal{F}_{и})$ эксперимента $E_{и}$.

Трудным шагом в доказательстве будет определение вероятностной функции $\mathsf{P}_{и}(A)$ на $\mathcal{F}_{и}$. Пусть $\mathcal{L}_0$ есть класс подмножеств действительной прямой, каждое из которых является конечным объединением непересекающихся промежутков вида $[a, b) \subset R$. При этом допускается, в частном случае, $a = -\infty$, $b = +\infty$. Легко проверить, что класс $\mathcal{L}_0$ является алгеброй. Если случайное событие $A = \cup_{i=1}^{n}[a_i, b_i) \in \mathcal{L}_0$ и $[a_i, b_i) \cap [a_j, b_j) = \varnothing$ при $i \neq j$, то, по определению, полагаем, $\mathsf{P}_{и}(\cup_{i=1}^{n}[a_i, b_i)) = \sum_{i=1}^{n}(G(b_i) - G(a_i))$. Используя свойства заданной функции $G(x)\colon R \cup \{-\infty\} \cup \{+\infty\} \to [0, 1]$, легко доказывается [1], что $\mathsf{P}_{и}(A)$ является неотрицательной, нормированной и счетно-аддитивной функцией на алгебре $\mathcal{L}_0$. Заметим, что в определении 6.1 борелевской σ-алгебры $\mathcal{B}$ и при доказательстве теоремы 6.1 вместо множества K всех интервалов вида $[a, b)$ можно взять класс $\mathcal{L}_0$. Значит, борелевская σ-алгебра $\mathcal{B}$ является наименьшей σ-алгеброй на действительной прямой, содержащей алгебру $\mathcal{L}_0$. Отсюда по теореме Каратеодори [1] вероятность $\mathsf{P}_{и}(\cdot)$ на алгебре $\mathcal{L}_0$ можно единственным образом продолжить на борелевскую σ-алгебру $\mathcal{B}$.

На этом шаге заканчивается построение вероятностной модели $(\Omega_{и}, \mathcal{F}_{и}, \mathsf{P}_{и}(\cdot))$ для модельного эксперимента $E_{и}$. Наконец пришло время пояснить механизм выбора точки на действительной прямой. Случайный механизм должен выбирать точку из произвольного промежутка $[a, b)$ на действительной прямой с вероятностью $\mathsf{P}_{и}([a, b))$, равной величине $G(b) - G(a)$.

На последнем этапе в качестве количественной характеристики $\xi_{и}(\omega_{и})\colon \Omega_{и} \to R$ модельного эксперимента $E_{и}$ выберем абсциссу места положения выбранной случайным образом точки на действительной прямой. Значит, для величины $\xi_{и}(\omega_{и})$ на $(\Omega_{и}, \mathcal{F}_{и}, \mathsf{P}_{и}(\cdot))$ имеем тождество $\xi_{и}(\omega_{и}) \equiv \omega_{и}$, а для ее интегральной функции сразу находим

$$
\begin{aligned}
F_{\xi_{и}}(x) = \mathsf{P}_{и}(\{\omega_{и}\colon \xi_{и}(\omega_{и}) < x\}) = \mathsf{P}_{и}(\{\omega_{и}\colon \omega_{и} < x\}) = \\
= \mathsf{P}_{и}(\{\omega_{и}\colon -\infty < \omega_{и} < x\}) = G(x) - G(-\infty) = G(x),
\end{aligned}
$$

и теорема о выборочном вероятностном пространстве доказана.

Замечание 6.4. Пусть теперь для случайной величины $\xi(\omega)\colon \Omega \to R$ имеет место равенство $\mathsf{P}(\{\omega\colon \xi(\omega) \in B\}) = \mathsf{P}_{и}(B)$ для любого борелевского множества B на действительной прямой. Тогда будет полностью определено вероятностное пространство $(\Omega_{\text{и}}, \mathcal{F}_{\text{и}}, \mathsf{P}_{\text{и}}(\cdot))$, где $\Omega_{\text{и}} = R$ и $\mathcal{F}_{\text{и}} = \mathcal{B}$. Отсюда для любого $x \in R$ имеем для интегральной функции распределения

$$F_\xi(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\}) = \mathsf{P}_{\text{и}}((-\infty, x)) = \mathsf{P}_{\text{и}}(\{\omega_{\text{и}}\colon \xi_{\text{и}}(\omega_{\text{и}}) < x\}) = F_{\xi_{\text{и}}}(x),$$

если только $\xi_{\text{и}}(\omega_{\text{и}}) \equiv \omega_{\text{и}}$. Теперь остается в качестве функции $G(x)$ взять $F_\xi(x)$ и воспользоваться теоремой о выборочном вероятностном пространстве. Еще раз напомним, что согласно этой теореме функция $F_\xi(x)$ однозначно определяет вероятность $\mathsf{P}_{\text{и}}(B)$ для любого борелевского множества B на действительной прямой.

Итак, на практике при изучении конкретной количественной характеристики $\xi(\omega)$ сложного эксперимента E можно вместо неизвестной его вероятностной модели $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ рассматривать вероятностную модель $(\Omega_{\text{и}}, \mathcal{F}_{\text{и}}, \mathsf{P}_{\text{и}}(\cdot))$ более простого эксперимента $E_{\text{и}}$. Далее, возможно незаданную поточечно на $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ случайную величину $\xi(\omega)\colon \Omega \to R$ необходимо заменить случайной величиной вида $\xi_{\text{и}}(\omega_{\text{и}})\colon \Omega_{\text{и}} = R \to R$. Так как интегральные функции распределения $F_\xi(x)$ и $F_{\xi_{\text{и}}}(x)$ совпадают, то вероятностные свойства случайной величины $\xi(\omega)$ будут определяться вероятностными свойствами $\xi_{\text{и}}(\omega_{\text{и}})$, т. е. вероятность $\mathsf{P}(\{\omega\colon \xi(\omega) \in B\}) = \mathsf{P}_{\text{и}}(\{\omega_{\text{и}}\colon \xi_{\text{и}}(\omega_{\text{и}}) \in B\})$ для любого $B \in \mathcal{B}$. В силу этого и тождества $\xi_{\text{и}}(\omega_{\text{и}}) \equiv \omega_{\text{и}}$ в дальнейшем вероятность $\mathsf{P}_{\text{и}}(\{\omega_{\text{и}}\colon \xi_{\text{и}}(\omega_{\text{и}}) \in B\})$ будем обозначать через функцию $\mathsf{P}_\xi(B)$. Можно сказать, что с точки зрения количественных характеристик $\xi(\omega)$, $\xi_{\text{и}}(\omega_{\text{и}})$ эксперимент $E_{\text{и}}$ является моделью эксперимента E. При этом основное вероятностное пространство $(\Omega_{\text{и}}, \mathcal{F}_{\text{и}}, \mathsf{P}_{\text{и}}(\cdot)) = (R, \mathcal{B}, \mathsf{P}_\xi(\cdot))$ для эксперимента $E_{\text{и}}$ называется *выборочным* для случайной величины $\xi(\omega)\colon \Omega \to R$. Таково практическое значение теоремы о выборочном вероятностном пространстве.

Подводя итог рассмотренным свойствам функции распределения, можно сказать, что функция распределения любой случайной величины является неубывающей и непрерывной слева функцией, удовлетворяющей неравенствам $0 \leqslant F(x) \leqslant 1$ и предельным равенствам $\lim\limits_{x\to\infty} F(x) = F(+\infty) = 1$ и $\lim\limits_{x\to-\infty} F(x) = F(-\infty) = 0$. Однако верно и в некотором смысле обратное утверждение. Если для функции $G(x)$, выполнены утверждения теорем 6.3–6.5 и 6.7, то существует такой искусственный эксперимент $E_{\text{и}}$ с комплексом условий $\Sigma_{\text{и}}$ и с исходами $\mathcal{F}_{\text{и}}$, для которого можно построить основное вероятностное пространство $(\Omega_{\text{и}}, \mathcal{F}_{\text{и}}, \mathsf{P}_{\text{и}}(\cdot))$ и ввести искусственную случайную величину $\xi_{\text{и}}(\omega_{\text{и}})$ такую, что ее функция распределения $F_{\xi_{\text{и}}}(x)$ будет тождественно совпадать с функцией $G(x)$. Следовательно, можно сказать, что утверждения теорем 6.3–6.5 и 6.7 являются необходимыми и достаточными для существования интегральной функции распределения.

§ 3. Классификация одномерных случайных величин и их законы распределения

Разнородность реальных экспериментов и большое число приборов, которые измеряют количественные характеристики каждого из них, порождают великое многообразие случайных величин. Например, множество принимаемых случайными величинами значений может быть конечным, счетным или несчетным. Значения некоторой случайной величины могут быть расположены дискретно и любые два из них разделены интервалом. Множество значений другой случайной величины может быть счетным и всюду плотным на действительной прямой. Поэтому возникает проблема классификации случайных величин, столь различных по своей физической природе и математической структуре. Классификация математических, как и любых других объектов, существенно зависит от удачно выбранного для этой цели признака. В замечании 6.4 для произвольной случайной величины $\xi(\omega)\colon \Omega \to R$ было введено так называемое выборочное пространство вида $(R, \mathcal{B}, \mathsf{P}_\xi(\cdot))$. В этом пространстве только вероятность $\mathsf{P}_\xi(\cdot)$ опосредованно через интегральную функцию $F_\xi(x)$ зависит от $\xi(\omega)$. Поэтому в качестве одного из основных признаков классификации выберем такие локальные свойства интегральной функции распределения, как ее разрывы, непрерывность и абсолютная непрерывность. Если через $X \subset R$, как и ранее, обозначить множество значений произвольной случайной величины, то вторым дополнительным признаком классификации является мощностные свойства этого числового множества. Очевидно, что множество X может быть конечным, счетным или несчетным. Из множества всех случайных величин, прежде всего, выделим так называемые дискретные случайные величины, которые очень часто используются на практике.

3.1. Функциональные характеристики измерителей элементарных исходов с дискретным распределением. Сначала рассмотрим случай, когда в качестве основного признака классификации одномерной случайной величины ξ используются некоторые особые свойства структуры множества $X = \{\xi(\omega)\colon \omega \in \Omega\}$. В связи с этим введем следующее определение.

Определение 6.4. Случайная величина $\xi(\omega)$ называется *дискретной*, если из множества X можно выделить такое счетное подмножество $X_1 = \{x_1', x_2', \ldots\}$ так называемых возможных ее значений, что $\mathsf{P}(\{\omega\colon \xi(\omega) = x_1'\}) = p_1 > 0$, $\mathsf{P}(\{\omega\colon \xi(\omega) = x_2'\}) = p_2 > 0, \ldots$ и $p_1 + p_2 + \ldots = 1$.

В простейшем варианте подмножество X_1 может содержать лишь конечное число m элементов, т.е. $X_1 = \{x_1', x_2', \ldots, x_m'\}$. Если случайная величина ξ принимает только счетное множество возможных значений, то достаточно использовать второй признак классификации, так как в этом варианте требования определения 6.4 очевидно выполняются и случайная величина будет дискретной. Заметим, что принятое определение 6.4 не исключает такую необычную для восприятия ситуацию, когда дискретная случайная величина ξ может принимать несчетное число различных значений из некоторого множества $X_2 \subset X$. В этом случае множество $X = X_1 \cup X_2$ и $X_1 \cap X_2 = \varnothing$. Однако всегда дискретная случайная величина $\xi(\omega)$ принимает значения из некоторого счетного множества, другими словами, для $\xi(\omega)$ выполняется так называемое условие нормировки.

Для иллюстрации сказанного рассмотрим следующий простой эксперимент.

Пример 6.13 (*задача о консультациях*). Преподаватель и студент условились для консультаций встречаться на кафедре по понедельникам. При этом преподаватель и студент из-за транспортных задержек приходят на кафедру независимо и наудачу в помежутке с восьми до девяти часов. Аналогичный эксперимент рассмотрен в [6]. Пусть t_1 и t_2 суть моменты прихода на кафедру преподавателя и студента соответственно. Тогда множество Ω описаний ω элементарных исходов для этого эксперимента является квадрат $\{\omega = (t_1, t_2)\colon 8 \leqslant t_1 \leqslant 9,$ $8 \leqslant t_2 \leqslant 9\}$ на плоскости $t_1 0 t_2$ с единицей масштаба в один час. Выберем в качестве σ-алгебры $\mathcal{F}$ наблюдаемых исходов эксперимента множество всех подмножеств квадрата Ω, имеющих площадь. Так как площадь квадрата Ω равна единице, то каждому наблюдаемому событию A из $\mathcal{F}$ можно приписать геометрическую вероятность $\mathsf{P}(A)$, которая равна площади множества A [6]. Таким образом, для этого эксперимента посроена вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. Рассмотрим теперь случайную величину $\xi(\omega)$ на $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, которая следующим образом определяет моменты начала консультации. Если преподаватель появился раньше студента, т. е. произошло событие $B = \{\omega = (t_1, t_2)\colon t_1 < t_2,\ (t_1, t_2) \in \Omega\}$, то консультация начинается в девять часов 3 минуты и, следовательно $\xi(\omega) = 9{,}05$ при $\omega \in B$. Если студент появился раньше преподавателя, т. е. произошло событие $C = \{\omega = (t_1, t_2)\colon t_2 < t_1,\ (t_1, t_2) \in \Omega\}$, то консультация начинается в девять часов 6 минут и, значит, $\xi(\omega) = 9{,}1$ при $\omega \in C$. Наконец, если они приходят на кафедру одновременно в моменты t, т. е. произошло событие $D = \{\omega = (t_1, t_2)\colon t_1 = t_2 = t,\ 8 \leqslant t \leqslant 9\}$, то консультация начинается немедленно и поэтому $\xi(\omega) = t$ при $\omega = (t, t) \in D$. Так как $B \cup C \cup D = \Omega$, $B \cap C \cap D = \varnothing$, то случайная величина $\xi(\omega)$ определена на $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ и принимает значение x из множества $X = \{9{,}05, 9{,}1\} \cup \{t\colon 8 \leqslant t \leqslant 9\}$. Найдем закон распределения вероятностей для $\xi(\omega)$: $\mathsf{P}(\{\omega\colon \xi(\omega) = 9{,}05\}) = \mathsf{P}(B) = 1/2,$

$\mathsf{P}(\{\omega\colon \xi(\omega) = 9{,}1\}) = \mathsf{P}(C) = 1/2$, $\mathsf{P}(\{\omega\colon 8 \leqslant \xi(\omega) \leqslant 9\}) = \mathsf{P}(D) = 0$. Отсюда множество $X_1 = \{9{,}05;\ 9{,}1\}$, а множество $X_- = \{x\colon 8 \leqslant x \leqslant 9\}$, которое является несчетным. Итак, приведен пример дискретной случайной величины $\xi(\omega)$ на $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, которая принимает значения x из несчетного множества X.

Так как $\xi(\omega)$ — измеримая функция, то множество $\{\omega\colon \xi(\omega) = x_i'\}$ принадлежит σ-алгебре $\mathcal{F}$. Поэтому всегда можно определить вероятность вида $\mathsf{P}(\{\omega\colon \xi(\omega) = x_i'\})$, где $i = 1, 2, \ldots$ или $i = 1, 2, \ldots, m$. Совокупность вида $\{p_i\colon i = 1, 2, \ldots\}$ называется *распределением дискретной случайной величины*. В частном случае, когда $X_1 = \{x_1', x_2', \ldots, x_m'\}$, т. е. случайная величина принимает конечное число различных значений с вероятностью единица, ее распределение естественно примет вид $\{p_1, p_2, \ldots, p_m\}$. Для дискретной случайной величины ξ в общем случае $X = X_1 \cup X_2$, где $X_2 = X \backslash X_1$ и, следовательно,

$$\Bigl(\bigcup_{i\colon x_i' \in X_1} \{\omega\colon \xi(\omega) = x_i'\}\Bigr) \cup \{\omega\colon \xi(\omega) \in X_2\} = \Omega,$$

$$\Bigl(\bigcup_{i\colon x_i' \in X_1} \{\omega\colon \xi(\omega) = x_i'\}\Bigr) \cap \{\omega\colon \xi(\omega) \in X_2\} = \varnothing,$$

$$\{\omega\colon \xi(\omega) = x_i'\} \cap \{\omega\colon \xi(\omega) = x_j'\} = \varnothing \text{ при } i \neq j,$$

$$\{\omega\colon \xi(\omega) < x\} = \Bigl(\bigcup_{i\colon x_i' < x} \{\omega\colon \xi(\omega) = x_i'\}\Bigr) \cup \{\omega\colon \xi(\omega) \in X_2, \xi(\omega) < x\}.$$

Из этих равенств в событиях сразу следуют равенства в вероятностях:

$$\mathsf{P}\Bigl(\bigcup_{i\colon x_i' \in X_1} \{\omega\colon \xi(\omega) = x_i'\}\Bigr) = \sum_{i\colon x_i' \in X_1} \mathsf{P}\Bigl(\{\omega\colon \xi(\omega) = x_i'\}\Bigr) = \sum_{i\colon x_i' \in X_1} p_i = 1,$$

$$\mathsf{P}(\{\omega\colon \xi(\omega) \in X_2\}) = 0,$$

$$\begin{aligned} F(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\}) = \mathsf{P}\Bigl(\bigcup_{i\colon x_i' < x} \{\omega\colon \xi(\omega) = x_i'\}\Bigr) + \\ + \mathsf{P}(\{\omega\colon \xi(\omega) \in X_2, \xi(\omega) < x\}) = \mathsf{P}\Bigl(\bigcup_{i: x_i' < x} \{\omega\colon \xi(\omega) = x_i'\}\Bigr) = \\ = \sum_{i: x_i' < x} \mathsf{P}(\xi = x_i') = \sum_{i\colon x_i' < x} p_i. \end{aligned}$$

Заметим, что в формуле для функции $F(x) = \sum_{i\colon x_i' < x} p_i$ суммирование распространяется на все те индексы, для которых $x_i' < x$.

Рассмотрим теперь важный для практики случай, когда значения дискретной случайной величины из множества X_1 можно занумеровать и расположить на действительной прямой R в виде строго возрастающей последовательности $x'_1 < x'_2 < \ldots$. При этом значения дискретной случайной величины $\xi(\omega)$ из множества X_1 отделены интервалами. Например, если X_1 содержит конечное число элементов, то это всегда можно сделать. График такой интегральной функции распределения $F(x)$ имеет следующий качественный вид (рис. 6.7).

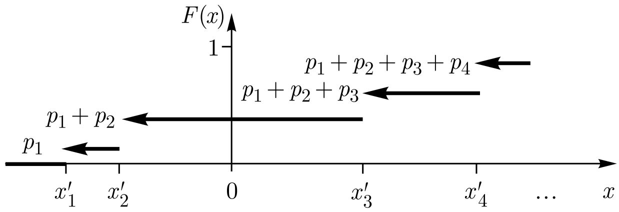


Рис. 6.7

Отсюда легко видеть, что интегральная функция распределения $F(x)$ является ступенчатой функцией или кусочно-постоянной. Более того, производная интегральной функции распределения дискретной случайной величины всюду равна нулю, исключая точки из X_1, в которых $F(x)$ терпит разрывы и $F(x'_i + 0) - F(x'_i) = p_i$, $i = 1, 2, \ldots$. В самом деле, так как $F(x) = \sum_{i:\, x'_i < x} p_i$, то эта функция на промежутке $(-\infty, x'_1]$ принимает значение ноль. На каждом промежутке вида $(x'_i, x'_{i+1}]$ значение $F(x)$ равно $p_1 + p_2 + \ldots + p_i$, где $i = 1, 2, \ldots$. Если случайная величина $\xi(\omega)$ принимает конечное число различных значений с вероятностью единица, то функция $F(x) = p_1 + p_2 + \ldots + p_m = 1$ при $x \in (x'_m, +\infty)$.

Для дискретной случайной величины $\xi(\omega)$, которая рассмотрена в предыдущем абзаце, очень часто используют другие законы распределения. Эти законы полностью эквивалентны интегральной функции $F(x)$ и распределению $\{p_i\colon i = 1, 2, \ldots, \sum_{i=1}^{\infty} p_i = 1\}$. В качестве такого закона распределения удобно применять так называемый ряд распределения. Рядом распределения дискретной случайной величины называется таблица

Таблица 6.1

$\xi(\omega)$	x'_1	x'_2	$\ldots$	x'_i	$\ldots$
$\mathsf{P}(\cdot)$	p_1	p_2	$\ldots$	p_i	$\ldots$

при $X_1 = \{x'_1, x'_2, \ldots\}$ и, естественно, таблица

Таблица 6.2

$\xi(\omega)$	x'_1	x'_2	$\ldots$	x'_m
$\mathsf{P}(\cdot)$	p_1	p_2	$\ldots$	p_m

при $X_1 = \{x'_1, x'_2, \ldots, x'_m\}$. В верхней строке каждой такой таблицы указываются все последовательные значения случайной величины, имеющие строго положительную вероятность, а в нижней — вероятности того, что случайная величина примет такие значения. Отметим следующие простейшие свойства ряда распределения:

1) $x'_1 < x'_2 < \ldots < x'_i < x'_{i+1} < \ldots$;

2) $p_i > 0$, $i = 1, 2, \ldots$;

3) $\sum_{i=1}^{\infty} p_i = 1$.

Иногда ряд распределения дискретной случайной величины изображают графически. Для этого вероятностные характеристики случайной величины дискретного типа и ее численные значения удобно представить наглядно в прямоугольной системе координат на плоскости. В этой системе координат на плоскости для каждого фиксированного $i = 1, 2, \ldots$ поставим точку с абсциссой x'_i и ординатой p_i. Затем каждые две соседние точки соединим отрезком. Полученный график в виде ломанной кривой определяет так называемый многоугольник или полигон распределения вероятностей. На рис. 6.8 ради иллюстрации отображен пример, когда дискретная случайная величина $\xi(\omega)$ принимает только значения x'_1, x'_2, x'_3 и x'_4 с вероятностями p_1, p_2, p_3, и p_4.

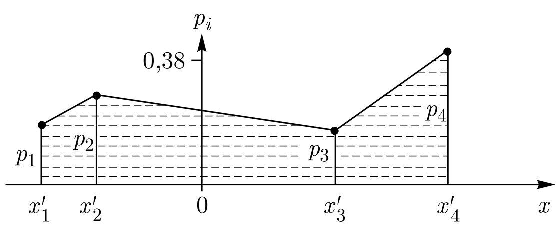


Рис. 6.8

Многоугольник распределения вероятностей для этого частного случая помечен штриховкой и является еще одной формой закона распределения. Этот график дает для исследователя наглядное представление о вероятностных свойствах дискретной случайной величины. Например, вероятность того, что дискретная случайная величина примет значения строго меньше нуля, равна сумме $p_1 + p_2$ длин вертикальных отрезков, расположенных левее прямой $x = 0$.

Замечание 6.5. В заключение отметим, что интегральная функция $F(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\})$ дискретной случайной величины ξ, как правило, представляет собой кусочно-постоянную или ступенчатую кривую с наличием интервалов постоянства. Например, на рис. 6.7 имеется не менее пяти таких интервалов. Однако в общем случае интервалов постоянства у интегральной функции распределения дискретной случайной величины может и не быть. Действительно, пусть значениями дискретной случайной величины $\xi(\omega)$ будут все рациональные числа. Множество рациональных чисел счетно. Поэтому все рациональные числа можно занумеровать натуральными числами. Тогда множество всех рациональных чисел можно записать в виде: $\{r_i\colon i = 1, 2, \ldots\}$. Далее примем, что эта дискретная случайная величина принимает некоторое значение r_i с вероятностью $p_i = 1/2^i$, где r_i — рациональное число с номером $i = 1, 2, \ldots$. Легко видеть, что $p_i = 1/2^i > 0$ и $\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} 1/2^i = 0{,}5/(1 - 0{,}5) = 1$, т. е. выполняется условие нормировки. Множество рациональных чисел является всюду плотным на прямой. Поэтому нельзя определить никакого интервала между двумя соседними рациональными числами, в котором нет других рациональных значений случайной величины. Следовательно, интервалов постоянства у интегральной функции распределения для такой дискретной случайной величины быть не может. Более того, не существует минимального рационального числа $r \in \{r_i\colon i = 1, 2, \ldots\}$. Отсюда следует, что для такой случайной величины нельзя написать ряд распределения, построить многоугольник распределения и, наконец, графически отобразить функцию распределения.

3.2. Примеры распределений дискретных случайных величин. Рассмотрим теперь типичные примеры на определение и представление различных законов распределения дискретной случайной величины.

Пример 6.14. Стрелок ведет стрельбу по мишени до первого попадания, имея боезапас из четырех патронов. Вероятность попадания при каждом выстреле равна 0,6. Построить функцию распределения боезапаса, оставшегося неизрасходованным.

Решение. Запишем пространство описаний элементарных исходов или достоверное событие этого опыта в виде: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$. Здесь при каждом $i = 1, 2, 3, 4$ описание ω_i соответствует такому элементарному исходу, когда первое попадание происходит на i-м выстреле. Описание ω_5 соответствует такому элементарному исходу, когда стрелок четыре раза промахнулся. Введем случайную величину, которая определяет число неизрасходованных патронов после испытания. Эта случайная величина такова, что $\xi(\omega_1) = 3$, $\xi(\omega_2) = 2$, $\xi(\omega_3) = 1$, $\xi(\omega_4) = \xi(\omega_5) = 0$. Итак, имеем поточечное задание рассматриваемой случайной величины. В этой простой задаче $X = X_1$, $X_2 = \varnothing$. Отсюда получим распределение случайной величины в виде конечной последовательности вида:

$$\begin{aligned} p_1 &= \mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) = \mathsf{P}(\{\omega_4, \omega_5\}) = \\ &= 0{,}4 \times 0{,}4 \times 0{,}4 \times 0{,}6 + 0{,}4 \times 0{,}4 \times 0{,}4 \times 0{,}4 = 0{,}064, \\ p_2 &= \mathsf{P}(\{\omega\colon \xi(\omega) = 1\}) = \mathsf{P}(\{\omega_3\}) = 0{,}4 \times 0{,}4 \times 0{,}6 = 0{,}096, \end{aligned}$$

$$p_3 = \mathsf{P}(\{\omega\colon \xi(\omega) = 2\}) = \mathsf{P}(\{\omega_2\}) = 0{,}4 \times 0{,}6 = 0{,}24,$$
$$p_4 = \mathsf{P}(\{\omega\colon \xi(\omega) = 3\}) = \mathsf{P}(\{\omega_1\}) = 0{,}6.$$

Удобно для этой случайной величины информацию о вероятностях тех или иных значениях величины ξ свести в табл. 6.3, состоящую из двух строк и пяти столбцов.

Таблица 6.3

$\xi(\omega)$	$x_1' = 0$	$x_2' = 1$	$x_3' = 2$	$x_4' = 3$
p_i	$p_1 = 0{,}064$	$p_2 = 0{,}096$	$p_3 = 0{,}24$	$p_4 = 0{,}6$

Интегральную функцию распределения легко вычислить с помощью формулы $F(x) = \sum_{i\colon x_i' < x} p_i$ и представить графически (рис. 6.9).

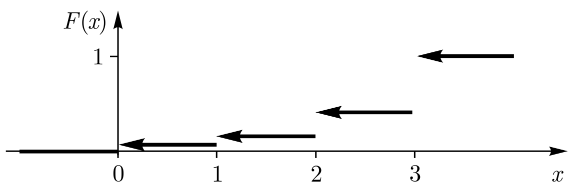


Рис. 6.9

Например, $F(2) = \mathsf{P}(\{\omega\colon \xi(\omega) < 2\}) = \mathsf{P}(\{\omega\colon \xi(\omega) = 0\} \cup \{\omega\colon \xi(\omega) = 1\}) = \mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) + \mathsf{P}(\{\omega\colon \xi(\omega) = 1\} = p_1 + p_2 = 0{,}16.$

Пример 6.15. На зачете студент получил пять задач. Вероятность того, что студент решит задачу, одинакова для всех задач и равна 0,7. Описать закон распределения для случайного количества решенных задач. Определить наиболее вероятное количество решенных задач. Найти вероятность того, что студент решит не менее трех задач.

Решение. Опишем пространство элементарных исходов рассматриваемого случайного эксперимента. В данном эксперименте случаен факт решения каждой из пяти предложенных задач. Элементарный исход эксперимента можно описать, например, как упорядоченную последовательность пяти элементов. Элементами могут быть нули и единицы. Если i-й элемент равен единице, то это означает, что i-я задача решена; если этот элемент нулевой, то соответствующая задача не решена. Получаем $\Omega = \{\omega = (i_1, i_2, i_3, i_4, i_5)\colon i_j \in \{0, 1\},\ j = 1, 2, \ldots, 5\}$. В качестве алгебры $\mathcal{F}$ наблюдаемых случайных событий рассмотрим всевозможные подмножества множества Ω. Вероятность для каждого элементарного случайного события $\{\omega\}$ естественно определить по следующей формуле:

$$\mathsf{P}(\{\omega\}) = (0{,}7)^{\sum_{j=1}^{5} i_j} (0{,}3)^{5 - \sum_{j=1}^{5} i_j}.$$

Значит, вероятность элементарного случайного события $\{\omega\}$ определяется количеством единиц в записи вектора ω.

Для вычисления вероятности случайного события $A \in \mathcal{F}$ суммируем вероятности всех элементарных исходов, благоприятствующих появлению события A. Далее, случайную величину $\xi(\omega)$ определим следующим образом: $\xi(\omega) = \xi((i_1, i_2, i_3, i_4, i_5)) = \sum_{j=1}^{5} i_j$. Например, $\xi((0,1,1,0,0)) = 2$, $\xi((1,1,1,0,1)) = 4$, $\xi((0,0,0,0,0)) = 0$ и т. д. Случайная величина $\xi(\omega)$ принимает значения $0, 1, 2, 3, 4, 5$. Для описания распределения случайной величины $\xi(\omega)$ достаточно указать вероятности принятия этих значений. Например, вероятности $\mathsf{P}(\xi = 0)$ и $\mathsf{P}(\xi = 1)$ вычисляются следующим образом:

$$\begin{aligned}
&\mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) = \mathsf{P}(\{(0,0,0,0,0)\}) = (0{,}3)^5 = 0{,}00243,\\
&\mathsf{P}(\xi = 1) = \mathsf{P}(\{\omega\colon \xi(\omega) = 1\}) =\\
&\qquad = \mathsf{P}(\{(1,0,0,0,0)\},\ \{(0,1,0,0,0)\},\ \{(0,0,1,0,0)\}) +\\
&\qquad + \mathsf{P}(\{(0,0,0,1,0)\},\ \{(0,0,0,0,1)\}) = 5 \times (0{,}7) \times (0{,}3)^4 = 0{,}02835.
\end{aligned}$$

Используя этот способ, легко получим, что

$$\mathsf{P}(\xi = 2) = C_5^2 (0{,}7)^2 (0{,}3)^3 = 0{,}1323,$$

$$\mathsf{P}(\xi = 3) = C_5^3 (0{,}7)^3 (0{,}3)^2 = 0{,}3087,$$

$$\mathsf{P}(\xi = 4) = C_5^4 (0{,}7)^4 (0{,}3) = 0{,}36015$$

и, наконец,

$$\mathsf{P}(\xi = 5) = (0{,}7)^5 = 0{,}16807.$$

Для определения наиболее вероятного количества решенных задач необходимо найти такое значение этой случайной величины, которое имеет наибольшую вероятность. Зная закон распределения, на этот вопрос можно ответить сразу. Вероятность появления значения $x_5' = 4$ больше вероятностей других значений. Значит, студент, скорее всего, решит четыре задачи. Вероятность того, что студент решит не менее трех задач равна

$$\mathsf{P}(\xi \geqslant 3) = \mathsf{P}(\xi = 3) + \mathsf{P}(\xi = 4) + \mathsf{P}(\xi = 5) = 0{,}83692.$$

Ряд распределения $\xi(\omega)$ отображен в виде табл. 6.4:

Таблица 6.4

$\xi(\omega)$	$x_1' = 0$	$x_2' = 1$	$x_3' = 2$	$x_4' = 3$	$x_5' = 4$	$x_6' = 5$
$\mathsf{P}(\cdot)$	0,00243	0,02835	0,1323	0,3087	0,36015	0,16807

Пример 6.16. В студенческой группе организована лотерея. Разыгрываются две вещи: одна стоит 200 руб., другая — 300 руб. Найти все законы распределения суммы выигрыша для студента, который бесплатно приобрел один билет, а всего было распродано 50 билетов.

Решение. Перенумеруем все билеты, используя для этого числа от 1 до 50. Пусть билетам с выигрышами 200 и 300 рублей присвоены соответственно номера 2 и 50. В этом опыте множество описаний элементарных исходов $\Omega = \{1, 2, \ldots, 50\}$ содержит 50 элементов, $\mathcal{F} = \{A\colon A \subset \Omega\}$ и $\mathsf{P}(\{\omega\}) = 1/50$ для $\omega = 1, 2, \ldots, 50$. Здесь ω есть номер приобретенного наудачу студентом билета. Обозначим через $\xi(\omega)$ сумму выигрыша для студента, который приобрел билет. Из условий эксперимента непосредственно находим, что случайная величина $\xi(\omega)$ при различных $\omega \in \Omega$ принимает значение из множества $X = \{0, 200, 300\}$ и $\{\omega\colon \xi(\omega) = 0\} = \{1, 3, 4, \ldots, 49\}$, $\{\omega\colon \xi(\omega) = 200\} = \{2\}$, $\{\omega\colon \xi(\omega) = 300\} = \{50\}$. Событие $\{\omega\colon \xi(\omega) = 0\}$ означает, что студент не выиграет, событие $\{\omega\colon \xi(\omega) = 200\}$ означает, что студент выиграет 200 руб., и, наконец, событие $\{\omega\colon \xi(\omega) = 300\}$ означает, что студент выиграет 300 руб. Вероятности наступления этих событий таковы, что

$$p_1 = \mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) = \mathsf{P}(\{1, 3, 4, \ldots, 49\}) = 48/50,$$

$$p_2 = \mathsf{P}(\{\omega\colon \xi(\omega) = 200\}) = \mathsf{P}(\{2\}) = 1/50,$$

$$p_3 = \mathsf{P}(\{\omega\colon \xi(\omega) = 300\}) = \mathsf{P}(\{50\}) = 1/50.$$

Теперь для суммы выигрыша ξ можно записать ряд распределения (табл. 6.5) и построить многоугольник распределения (рис. 6.10).

Таблица 6.5

$\xi(\omega)$	0	200	300
$\mathsf{P}(\cdot)$	48/50	1/50	1/50

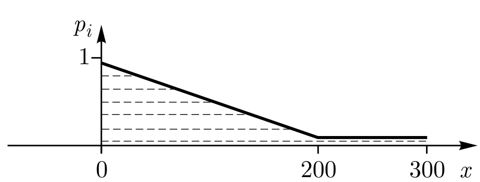


Рис. 6.10

График интегральной функции $F(x)$ представлен на рис. 6.11.

Рис. 6.11

Пример 6.17. Среди десяти лотерейных билетов имеется четыре выигрышных. Наудачу покупают два билета. Описать законы распределения вероятностей для числа выигрышных билетов среди двух купленных.

Решение. Пронумеруем все билеты. Пусть билеты с номерами 1–4 являются выигрышными, а билеты с номерами 5–10 выигрышными не являются. Случайный эксперимент заключается в отборе двух различных билетов из десяти. Порядок отбора билетов здесь не играет роли. Пространство описаний элементарных исходов данного случайного эксперимента имеет вид: $\Omega = \{\omega = \{i, j\}\colon i, j \in \{1, 2, \ldots, 10\},\ i \neq j\}$. В качестве алгебры $\mathcal{F}$ случайных событий будем рассматривать всевозможные подмножества множества Ω. Для каждого случайного события $A \in \mathcal{F}$ вероятность определим по классической формуле $\mathsf{P}(A) = \aleph(A)/\aleph(\Omega)$, где $\aleph(\Omega)$ подсчитывает общее число равновозможных элементарных исходов, а $\aleph(A)$ есть количество элементарных исходов, благоприятствующих наступлению события A. Если $\xi(\omega)$ определяет случайное число выигрышных билетов среди двух купленных, то такая случайная величина может принимать значения $0, 1, 2$. Вычислим вероятности этих значений:

$$p_1 = \mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) = \mathsf{P}(\xi = 0) = C_4^0 C_6^2 / C_{10}^2 = 1/3,$$

$$p_2 = \mathsf{P}(\{\omega\colon \xi(\omega) = 1\}) = \mathsf{P}(\xi = 1) = C_4^1 C_6^1 / C_{10}^2 = 8/15,$$

$$p_3 = \mathsf{P}(\{\omega\colon \xi(\omega) = 2\}) = \mathsf{P}(\xi = 2) = C_4^2 C_6^0 / C_{10}^2 = 2/15.$$

Ряд распределения ξ представлен в виде табл. 6.6.

Таблица 6.6

$\xi(\omega)$	$x'_1 = 0$	$x'_2 = 1$	$x'_3 = 2$
$\mathsf{P}(\cdot)$	$p_1 = 1/3$	$p_2 = 8/15$	$p_3 = 2/15$

Для проверки вычислений проверим условие нормировки: $p_1 + p_2 + p_3 = 1/3 + 8/15 + 2/15 = 1$. Описать распределение вероятностей случайной величины ξ можно было также с помощью интегральной функции распределения $F_\xi(x) = \mathsf{P}(\xi < x)$, которая равна нулю при $x \leqslant 0$, равна $1/3$ при $0 < x \leqslant 1$, равна $1/3 + 8/15 = 13/15$ при $1 < x \leqslant 2$ и, наконец, равна $1/3 + 8/15 + 2/15 = 1$ при $x > 2$.

Пример 6.18. По мишени производится один выстрел. Вероятность попадания в мишень равна 0,7. Пусть за удачный выстрел (попадание в мишень) дают вознаграждение 10 руб. Найти ряд, многоугольник и функцию распределения суммы вознаграждения за один выстрел.

Решение. Имеем множество $\Omega = \{\omega_1, \omega_2\}$ из описаний элементарных событий, где ω_1 есть описание попадания по мишени и ω_2 есть описание промаха. Строим множество всех событий: $\mathcal{F} = \{\varnothing, \{\omega_1, \omega_2\}, \{\omega_1\}, \{\omega_2\}\}$. В задаче предполагается следующее распределение вероятностей: $\mathsf{P}(\varnothing) = 0$, $\mathsf{P}(\{\omega_1\}) = 0{,}7$, $\mathsf{P}(\{\omega_2\}) = 0{,}3$, $\mathsf{P}(\{\omega_1, \omega_2\}) = 1$. Итак, для этого эксперимента построена простая вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. Случайная величина $\xi(\omega)$, которая определяет вознаграждение за удачный выстрел, имеет поточечное задание: $\xi(\omega_1) = 10$, $\xi(\omega_2) = 0$. Вероятности этих значений равны: $p_1 = \mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) = 0{,}3$ и $p_2 = \mathsf{P}(\{\omega\colon \xi(\omega) = 10\}) = 0{,}7$. Ряд распределения для ξ можно представить в виде табл. 6.7.

Таблица 6.7

$\xi(\omega)$	0	10
$\mathsf{P}(\cdot)$	0,3	0,7

Многоугольник распределения имеет вид (рис. 6.12), а график интегральной функции $F(x)$ для случайной величины представлен на рис. 6.13.

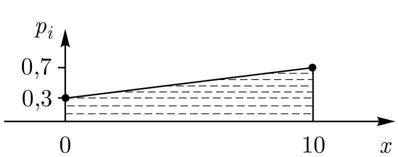


Рис. 6.12

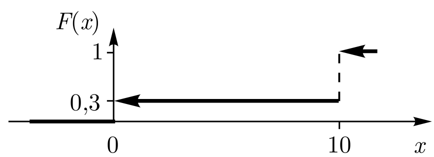


Рис. 6.13

3.3. Количественные характеристики элементарных исходов с несчетным множеством значений и абсолютно непрерывным распределением. Наряду с дискретными случайными величинами рассмотрим важный для практики класс непрерывных случайных величин. Для выделения этого класса будем использовать первый признак классификации. Каждая случайная величина из этого класса имеет абсолютно непрерывную интегральную функцию распределения $F(x)\colon R \to [0, 1]$ и, следовательно, принимает несчетное множество значений. Напомним, что функция $G(x)\colon [a, b] \to R$ называется *абсолютно непрерывной*, если для любого $\varepsilon > 0$ найдется такое число $\delta > 0$, что для любой конечной системы попарно непересекающихся интервалов $(a_k, b_k) \subset (a, b)$, $k = \overline{1, n}$, с суммой длин строго меньше δ, будет иметь место $\sum_{i=1}^{n} |G(b_k) - G(a_k)| < \varepsilon$.

Определение 6.5. Случайная величина $\xi(\omega)\colon \Omega \to X$ называется *непрерывной*, если существует такая неотрицательная функция $f(x)$ с областью определения R, что для любого действительного числа x имеет место равенство $F(x) = \int_{-\infty}^{x} f(u)\, du$.

Функция $f(x)$ называется *плотностью распределения случайной величины* $\xi(\omega)$. Если одновременно в контексте рассматривают несколько случайных величин, то плотность распределения обозначают $f_\xi(x)$ вместо $f(x)$. Непосредственно из определения 6.5 можно сделать следующие выводы. Функция $F(x)$ равна интегралу с переменным верхним пределом. Из курса математического анализа известно, что интеграл с переменным верхним пределом является абсолютно непрерывной функцией и, более того, эта функция имеет производную $dF(x)/dx = f(x)$ почти всюду по мере Лебега. Поэтому плотность распределения вероятностей для случайной величины $\xi(\omega)$ часто называют *дифференциальной функцией распределения*. Очевидно, что график функции распределения $F(x)$ не имеет скачков и, тем самым, имеет следующий качественный вид (рис. 6.14):

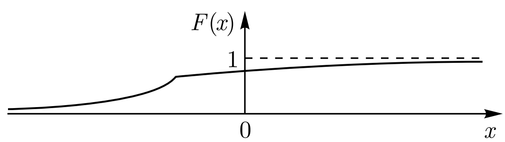


Рис. 6.14

Так как $f(x)$ — интегрируемая функция, то она является ограниченной. Значит, по теореме Лебега $f(x)$ будет не только почти всюду непрерывной по мере Лебега, но и измеримой в смысле Лебега. Для всякой измеримой по Лебегу функции существует такая борелевская функция, что указанные функции почти всюду совпадают по мере Лебега. Так как функция $F(x) = \int_{-\infty}^{x} f(u)\,du$, то подынтегральную функцию $f(x)$ можно изменить на любом множестве лебеговой меры ноль, не нарушив при этом величины интегральной функции распределения. Значит, функция $f(x)$ определяется неоднозначно. Поэтому функцию $f(x)$ можно считать в дальнейшем борелевской. Как правило, предполагают, что $f(x)$ имеет не более счетного числа точек разрыва на R, причем на конечном промежутке таких точек разрыва может быть лишь конечное число. Более того, на практике плотности распределения часто являются непрерывными функциями или становятся таковыми в результате соответствующего их изменения. Если плотность распределения $f(x)$ является непрерывной, то $dF(x)/dx = f(x)$ и $f(x) \geqslant 0$, так как $F(x)$ есть неубывающая функция. Поэтому в определении непрерывной случайной величины требуется естественное ограничение $f(x) \geqslant 0$.

Рассмотрим общие свойства произвольной непрерывной случайной величины и ее плотности распределения вероятностей. На рис. 6.15 приведен пример графика функции $f(x)$, который называется кривой распределения вероятностей случайной величины $\xi(\omega)$.

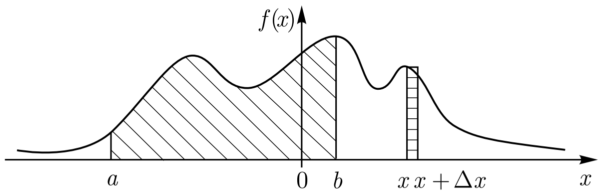


Рис. 6.15

Так как $F(x)$ является непрерывной функцией, то величина $\xi(\omega)$ принимает несчетное множество значений. Поскольку $\lim\limits_{x\to\infty} F(x) = F(+\infty) = 1$, то получаем известное условие нормировки: $\int_{-\infty}^{+\infty} f(u)\,du = 1$ для плотности распределения $f(x)$. Из равенства $\mathsf{P}(\{\omega\colon \xi(\omega) = x\}) = F(x+0) - F(x)$, непрерывности функции $F(x)$ и соотношения $\mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) < b\}) = F(b) - F(a)$ находим:

$$\mathsf{P}(\{\omega\colon \xi(\omega) = x\}) = 0,$$

$$\mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) < b\}) = \int\limits_{-\infty}^{b} f(u)\,du - \int\limits_{-\infty}^{a} f(u)\,du = \int\limits_{a}^{b} f(u)\,du.$$

Итак, если дифференциальная функция распределения $f(x)$ непрерывна, то $\mathsf{P}(\{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\}) = f(x)\Delta x + o(\Delta x) \approx f(x)\Delta x$.

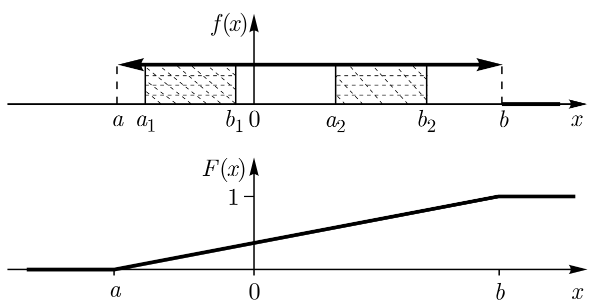


Рис. 6.16

На рис. 6.16 приведены более простой график плотности распределения некоторой случайной величины $\xi(\omega)$ и соответствующий ей график интегральной функции $F(x)$. В этом случае функция $f(x)$ равна $1/(b-a)$ при $x \in (a,b)$ и равна нулю при $x \notin (a,b)$. Легко видеть, что при всех $x \in R$ функция $f(x) \geqslant 0$ и что имеет место условие нормировки

$$\int\limits_{-\infty}^{\infty} f(u)\,du = \int\limits_{a}^{b} (b-a)^{-1} dx = 1.$$

Более того, при $(a_1, b_1) \subset (a, b)$, $(a_2, b_2) \subset (a, b)$ вероятности случайных событий вида $\{\omega\colon \xi(\omega) \in (a_1, b_1)\}$, $\{\omega\colon \xi(\omega) \in (a_2, b_2)\}$ вычисляются следующим образом:

$$\mathsf{P}(\{\omega\colon \xi(\omega) \in (a_1, b_1)\}) = \int\limits_{a_1}^{b_1} f(x)\,dx = (b_1 - a_1)(b-a)^{-1},$$

$$\mathsf{P}(\{\omega\colon \xi(\omega) \in (a_2, b_2)\}) = \int\limits_{a_2 1}^{b_2} f(x)\,dx = (b_2 - a_2)(b-a)^{-1}.$$

Каждая из этих вероятностей, которая зависит только от длины интервала и не зависит от его расположения на числовой оси, равна площади заштрихованной на рис. 6.16 одной из областей. Значит, при любых интервалах $(a_1, b_1) \subset (a, b)$ и $(a_2, b_2) \subset (a, b)$, для которых $b_1 - a_1 = b_2 - a_2$, получаем равенство: $\mathsf{P}(\{\omega\colon \xi(\omega) \in (a_1, b_1)\}) = \mathsf{P}(\{\omega\colon \xi(\omega) \in (a_2, b_2)\})$. Поэтому такая случайная величина $\xi(\omega)$ называется равномерно распределенной на конечном интервале (a, b).

Подытоживая все эти свойства и учитывая рис. 6.15, отметим следующие факты. Площадь между кривой распределения и осью абсцисс равна единице. Вероятность того, что случайная величина $\xi(\omega)$ примет заданное значение x, равна нулю. Вероятность попадания случайной величины в промежуток $[\,a, b)$ равна площади криволинейной трапеции, заштрихованной на рис. 6.15 наклонными линиями. Вероятность попадания случайной величины в промежуток $[x, x + \Delta x)$ при малом значении Δx приближенно равна площади заштрихованного на рис. 6.15 горизонтальными линиями прямоугольника со сторонами Δx и $f(x)$. В силу верного при $x \in R$ равенства $\mathsf{P}(\{\omega\colon \xi(\omega) = x\})$ имеем

$$\mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) < b\}) = \mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) \leqslant b\}) = \mathsf{P}(\{\omega\colon a < \xi(\omega) < b\}) = \\ = \mathsf{P}(\{\omega\colon a < \xi(\omega) \leqslant b\}) = \int\limits_a^b f(u)\,du.$$

Рассмотрим простой пример на построение непрерывной случайной величины и на определение ее законов распределения.

Пример 6.19. Между населенными пунктами O_1 и O_2 проведен газопровод вдоль некоторой прямой. Длина газопровода равна l. Ремонтная бригада базируется в пункте O_1. В течение некоторого промежутка времени случайным образом произошло повреждение газопровода в двух местах. Найти закон распределения расстояния, которое должна преодолеть бригада для устранения этих повреждений и возвращения на базу.

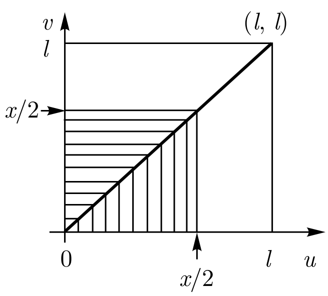


Рис. 6.17

Решение. Пусть u и v суть расстояния от места расположения базы ремонтников до первой и соответственно второй точек повреждений. Тогда вектор $\omega = (u, v)$ является описанием двух мест повреждения газопровода или некоторого элементарного исхода для этого эксперимента. Пространством описаний элементарных исходов данного опыта будет множество $\Omega = \{\omega = (u, v)\colon 0 \leqslant u \leqslant l, 0 \leqslant v \leqslant l\}$, представляющее собой квадрат на плоскости $u0v$ с вершинами $(0, 0)$, $(0, l)$, (l, l) и $(l, 0)$. Все это отображено на рис. 6.17. Множество $\mathcal{F}$ всех наблюдаемых исходов этого эксперимента совпадает с множеством всех подмножеств из Ω, которые имеют площадь или меру Лебега. Пусть $\xi(\omega)$ обозначает суммарное расстояние, которое проделает ремонтная бригада после возвращения ее на базу. Тогда расстояние $\xi(\omega)$ равно $2\,(v + u - v) = 2\,u$, если место первого повреждения расположено не ближе второго, и равно $2\,(u + v - u) = 2\,v$ в противном случае.

Итак, имеем отображение $\xi(\omega) = 2\max\{u, v\}\colon \Omega \to [0, 2l]$. Отсюда находим, что множество $\{\omega = (u, v)\colon 2\max\{u, v\} < x\} = \varnothing \in \mathcal{F}$ при $x \leqslant 0$ и равно $\Omega \in \mathcal{F}$ при $x > 2l$. Если $0 < x \leqslant 2l$, то множество $\{\omega\colon \xi(\omega) < x\} = \{(u, v)\colon 2\max\{u, v\} < x\} = \{(u, v)\colon u \geqslant v, u < x/2\} \cup$ $\cup\,\{(u, v)\colon u < v, v < x/2\} \in \mathcal{F}$. Первое из этих множеств заштриховано на рис. 6.17 вертикальными линиями, а второе — горизонтальными. Значит, проделанное ремонтной бригадой расстояние $\xi(\omega)$ является случайной величиной.

Используя теперь вид множества $\{\omega\colon \xi(\omega) < x\}$ при $-\infty < x < +\infty$ и определение геометрической вероятности на плоскости, найдем интегральную функцию распределения $F(x)$ для расстояния $\xi(\omega) =$ $= 2\max\{u, v\}$. Интегральная функция распределения $F(x)$ равна нулю при $x \leqslant 0$, равна $x^2/4l^2$ при $0 < x \leqslant 2l$ и, наконец, равна единице при $x > 2l$. На рис. 6.18 изображен график этой функции или первого закона распределения случайной величины $\xi(\omega)$.

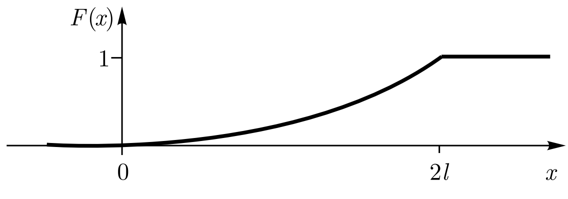


Рис. 6.18

Теперь легко найти второй закон распределения случайной величины $\xi(\omega)$ или плотность вероятности $dF(x)/dx = f(x)$. На рис. 6.19 приведен график дифференциальной функции $f(x)$, которая равна нулю при $x \leqslant 0$ и при $x > 2l$ и равна $x/2l^2$ при $0 < x \leqslant 2l$.

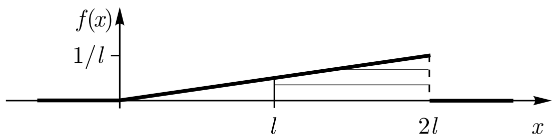


Рис. 6.19

Зная плотность распределения $f(x)$, можно, например, определить вероятность того, что ремонтная бригада преодолеет расстояние не менее l. Эта вероятность равна

$$\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant l\}) = \int\limits_l^\infty f(x)\,dx = \int\limits_l^{2l} f(x)\,dx = \int\limits_l^{2l} x/2l^2 dx = 3/4.$$

Можно вычислить эту вероятность с использованием интегральной функции $F(x)$:

$$\mathsf{P}(\{\omega\colon l \leqslant \xi(\omega) < \infty\}) = F(\infty) - F(l) = 1 - l^2/4l^2 = 3/4.$$

Заметим, что площадь заштрихованной области на рис. 6.19 равна 3/4.

3.4. Сингулярные случайные величины и нестандартные законы распределения. Кроме класса дискретных случайных величин, которые принимают конечное или счетное множество значений с вероятностью единица, и класса непрерывных случайных величин, интегральная функция распределения которых является не только непрерывной, но и абсолютно непрерывной, существует класс так называемых сингулярных случайных величин. Сингулярные случайные величины в реальных экспериментах практически не встречаются. Однако класс сингулярных случайных величин имеет большое теоретическое значение, так как позволяет понять все многообразие случайных величин и провести их полную классификацию. Для выделения этого класса предварительно рассмотрим понятие точки роста для интегральной функции распределения. Точка $\{a_0\}$ на действительной прямой называется точкой роста для интегральной функции распределения $F(x)$, если выполняется неравенство $F(a_0 + \varepsilon) - F(a_0 - \varepsilon) > 0$ для любого $\varepsilon > 0$. Например, изображенная на рис. 6.11 интегральная функция распределения имеет следующие точки роста: $\{0\}$, $\{200\}$ и $\{300\}$, а множество точек роста для приведенной на рис. 6.18 интегральной функции распределения совпадает с отрезком $[0, 2l]$.

Определение 6.6. Случайная величина $\xi(\omega)\colon \Omega \to X$ называется *сингулярной*, если ее интегральная функция распределения $F(x)$ непрерывна и точки роста $F(x)$ образуют множество нулевой меры Лебега.

Из этого определения непосредственно следует, что сингулярная случайная величина принимает несчетное множество значений и ее интегральная функция распределения $F(x)$ почти всюду по мере Лебега постоянна или $dF(x)/dx = 0$. Поэтому любая сингулярная случайная величина в отличие от непрерывной случайной величины не имеет плотности распределения. Заметим, что интегральная функция распределения $F(x)$ дискретной случайной величины также почти всюду по мере Лебега постоянна (или $dF(x)/dx = 0$), но при этом $F(x)$ не является непрерывной. Таковы в общих чертах отличия выделенных трех типов случайных величин. Приведем теперь пример сингулярной случайной величины.

Пример 6.20. Рассмотрим несколько экзотический эксперимент E, который заключается в случайном бросании точки на отрезок $[0,1]$ с помощью следующего механизма. На первом шаге делим отрезок $[0,1]$ на три равные по длине части. При этом внутренняя часть обязательно является интервалом. Точка может находиться в одном из полученных отрезков $[0,1/3]$, $[2/3,1]$ с вероятностью 2^{-1} или в интервале $(1/3,2/3)$ с нулевой вероятностью. На втором шаге этот процесс деления применяем к вновь полученнным отрезкам. Тогда точка может оказаться в одном из отрезков $[0,1/9]$, $[2/9,3/9]$, $[6/9,7/9]$, $[8/9,1]$ с вероятностью 2^{-2} или в одном из интервалов $(1/9,2/9)$, $(7/9,8/9)$ с нулевой вероятностью. На k-м шаге точка может оказаться в одном из 2^k отрезков с вероятностью 2^{-k} или в одном из 2^{k-1} интервалов с нулевой вероятностью. Это правило выбора новых отрезков и интевалов, как и механизм уточнения случайного положения точки в этих промежутках, применяется счетное число раз. В результате точка попадает в один из интервалов $(1/3,2/3)$, $1/9,2/9)$, $(7/9,8/9)$, $(1/27,2/27)$, $(7/27,8/27)$, $(19/27,20/27)$, $(25/27,26/27),\ldots$ с нулевой вероятностью. Исход этого эксперимента может заключаться в том, что точка появится в некотором стандартном промежутке отрезка $[0,1]$, например, в интервале $(1/3,8/9)$, или в отрезке $[2/27,25/27]$ и т. п.

Обозначим через символ ω абсциссу случайным образом поставленной точки на $[0,1]$. В опыте E достоверный исход $\Omega = \{\omega\colon 0 \leqslant \omega \leqslant 1\}$, а множество всех наблюдаемых исходов $\mathcal{F}$ есть σ-алгебра борелевских подмножеств на отрезке $[0,1]$. Этим завершается построение теоретико-множественной модели $(\Omega,\mathcal{F})$ эксперимента E и уточнение его случайного механизма.

В качестве числовой характеристики $\xi(\omega)$ эксперимента E выберем абсциссу места положения точки на отрезке $[0,1]$, выбранной с помощью описанного выше механизма. Следовательно, для функции $\xi(\omega)$ на $(\Omega, \mathcal{F})$ имеем тождество $\xi(\omega) \equiv \omega$. Отсюда легко находим, что $\{\omega\colon \xi(\omega) < x\} = \varnothing \in \mathcal{F}$ при $x \leqslant 0$, множество $\{\omega\colon \xi(\omega) < x\} = [0, x) \in \in \mathcal{F}$ при $0 < x \leqslant 1$ и $\{\omega\colon \xi(\omega) < x\} = [0, 1] = \Omega \in \mathcal{F}$ при $x > 1$. Итак, $\xi(\omega)\colon \Omega \to [0, 1]$ является случайной величиной на $(\Omega, \mathcal{F})$. Перейдем теперь к определению и непосредственному построению интегральной функции распределения $F(x)$ для такого рода случайной величины ξ (см. рис. 6.20).

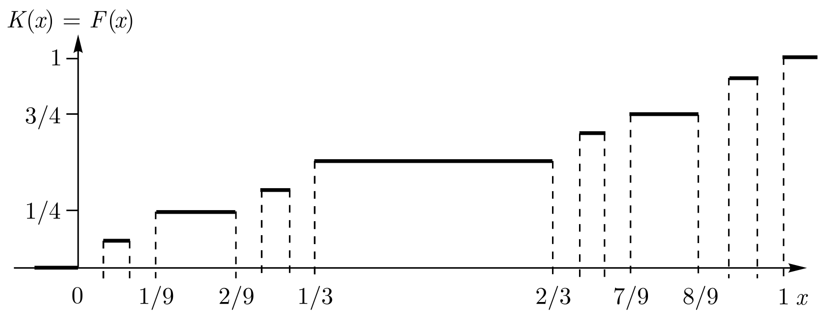


Рис. 6.20

Так как $\{\omega\colon \xi(\omega) < x\} = \varnothing$ при $x \leqslant 0$ и $\{\omega\colon \xi(\omega) < x\} = \Omega$ при $x > 1$, то $F(x) = 0$ при $x \leqslant 0$ и $F(x) = 1$ при $x > 1$. В этом эксперименте точка выбирается в интервале $(1/3, 2/3)$ с нулевой вероятностью или на отрезке $[2/3, 1]$ с вероятностью $1/2$. Поэтому функция $F(x) = 1 - 1/2 = 1/2$ при $1/3 < x < 2/3$. Так как точка выбирается в интервале $(1/9, 2/9)$ с нулевой вероятностью или на отрезке $[2/9, 1]$ с вероятностью $1/4 + 1/2$, то $F(x) = 1 - 1/4 - 1/2 = 1/4$ при $1/9 < x < 2/9$. Далее, точка выбирается в интервале $(7/9, 8/9)$ с нулевой вероятностью или на отрезке $[8/9, 1]$ с вероятностью $1/4$, то $F(x) = 1 - 1/4 = 3/4$ при $7/9 < x < 8/9$. Продолжая аналогичным образом этот процесс, определим интегральную функцию распределения $F(x)$ на всей последовательности интервалов: $(1/3, 2/3)$, $(1/9, 2/9)$, $(7/9, 8/9)$, $(1/27, 2/27)$, $(7/27, 8/27)$, $(19/27, 20/27)$, $(25/27, 26/27), \ldots$. Сумма длин всех этих интервалов постоянства функции $F(x)$, очевидно, равна $1/3 + 2 \times (1/9) + 4 \times (1/27) + \ldots = 1$. Значит, интегральная функция распределения $F(x)$ определена на R почти всюду по мере Лебега. В остальных точках отрезка $[0, 1]$ функцию $F(x)$ определим по непрерывности [11, 23]. Построенная таким способом функция распределения $F(x)$ является непрерывной и называется *кривой Кантора* $K(x)$ или *канторовой лестницей* [11].

Используя методику построения выборочного вероятностного пространства $(R, \mathcal{B}, \mathsf{P}_\xi(\cdot))$ с помощью интегральной функции $F(x) = K(x)$

и выводы замечания 6.4, нетрудно закончить построение вероятностной модели $(R, \mathcal{F}, \mathsf{P}(\cdot))$ для рассматриваемого здесь эксперимента.

Используя методику предыдущего абзаца, легко построить кривую Кантора $K(x)$, все изменения которой происходят на любом конечном отрезке $[a, b]$. Нетрудно видеть, что для $K(x)$ выполняются все свойства интегральной функции распределения, т. е. функция $K(x)$ удовлетворяет условию $0 \leqslant K(x) \leqslant 1$, является неубывающей, непрерывна на R и, наконец, удовлетворяет предельным соотношениям $\lim\limits_{x\to+\infty} K(x) = K(+\infty) = 1$ и $\lim\limits_{x\to-\infty} K(x) = K(-\infty) = 0$. Функция $K(x)$ на множестве лебеговской меры нуль возрастает без скачков от нуля до числа, равного единице. При этом множество точек роста функции $K(x)$ несчетно и образует так называемое борелевское множество Кантора [11, 19]. Этот несколько удивительный факт можно установить следующими простыми рассуждениями. Во-первых, величина изменения непрерывной функции в каждой отдельной точке равна нулю. Во-вторых, если бы это множество точек роста было счетным, то непрерывная функция $K(x)$ не могла бы измениться, т. е. была бы постоянной. Наконец отметим, что случайная величина с интегральной функцией $K(x)$ не является дискретной, так как функция $K(x)$ является непрерывной, а функция распределения дискретной величины является разрывной. Одновременно такая случайная величина не является непрерывной, так как кривая Кантора $K(x)$ почти всюду по мере Лебега имеет нулевую производную. Поэтому функция $K(x)$ не может быть представлена интегралом от своей производной. С другой стороны, случайная величина с интегральной функцией $K(x)$ напоминает непрерывную величину тем, что она принимает несчетное число значений, но для нее не существует плотности распределения. Далее, такого рода случайная величина напоминает дискретную случайную величину тем, что у ее функции распределения есть точки, которые образуют интервалы постоянства и, значит, не являются точками роста. Подводя итоги этого абзаца, можно дать более простое определение сингулярной случайной величины. А именно, случайная величина $\xi(\omega)$ называется *сингулярной*, если ее интегральная функция распределения является функцией типа канторовой.

Итак, были выделены три типа случайных величин: дискретные, непрерывные и сингулярные случайные величины. Следующая теорема Лебега о разложении показывает, что в действительности совокупность всех случайных величин покрывается указанными тремя типами.

Теорема 6.9 (теорема Лебега). *Любая интегральная функция распределения $F(x)$ может быть представлена в виде суммы $q_1 F_{\text{д}}(x) + q_2 F_{\text{ан}}(x) + q_3 F_{\text{с}}(x)$, где число $q_i \geqslant 0$, $i = 1, 2, 3$, и $q_1 + q_2 + q_3 = 1$, а функции $F_{\text{д}}(x)$, $F_{\text{ан}}(x)$ и $F_{\text{с}}(x)$ являются инте-*

гральными функциями распределения соответственно дискретной, непрерывной и сингулярной случайной величины.

Доказательство этой теоремы [1] требует владения специальным аппаратом из теории меры и поэтому здесь не приводится. Очевидно, что если случайная величина $\xi(\omega)$ является дискретной, то имеем $q_1 = 1$, $q_2 = q_3 = 0$. Аналогичным образом для непрерывной случайной величины $\xi(\omega)$ находим $q_2 = 1$, $q_1 = q_3 = 0$ и для сингулярной случайной величины $\xi(\omega)$ получим $q_3 = 1$, $q_1 = q_2 = 0$. Если, по крайней мере, два из чисел q_1, q_2 и q_3 в разложении Лебега интегральной функции отличны от нуля, то соответствующая ей величина $\xi(\omega)$ называется смешанной или произвольной. Теорема Лебега позволяет утверждать, что существует замкнутая классификация множества всех одномерных случайных величин. Иначе, среди множества всех одномерных случайных величин можно выделить только дискретные, непрерывные, сингулярные и смешанные.

Итак, кроме дискретных, непрерывных и сингулярных случайных величин, на практике встречаются случайные величины произвольного вида. Как правило, значения смешанной случайной величины сосредоточены как в конечном или счетном множестве изолированных точек (признак дискретной случайной величины), так и на участках между этими точками сосредоточены непрерывно (признак непрерывной случайной величины). Поэтому интегральная функция смешанных случайных величин имеет обыкновенные разрывы (скачки) в тех точках, вероятности попадания в которые отличны от нуля, а на участках между указанными точками изменяется непрерывно. Интегральная функция распределения смешанных случайных величин качественно имеет вид (рис. 6.21).

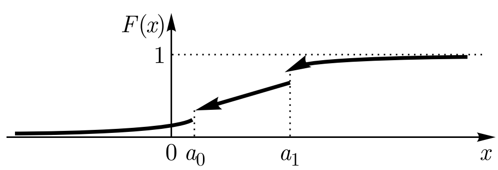


Рис. 6.21

Пример 6.21. Функция распределения $F(x)$ случайной величины ξ равна нулю при $x \leqslant 0$, равна $0,1 + x$ при $0 < x \leqslant 0{,}5$ и, наконец, равна единице при $x > 0{,}5$. Проверить необходимые и достаточные свойства $F(x)$, которые справедливы для любой интегральной функции распределения.

Решение. Для наглядности и простоты рассуждений изобразим график этой функции на рис. 6.22.

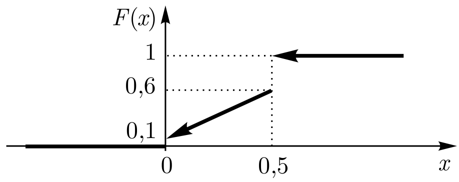


Рис. 6.22

Непосредственно из формулы для $F(x)$ или из ее графика следует, что $0 \leqslant F(x) \leqslant 1$, $\lim\limits_{x\to+\infty} F(x) = F(+\infty) = 1$, $\lim\limits_{x\to-\infty} F(x) = F(-\infty) = 0$. Далее, легко проверить, что функция $F(x)$ является неубывающей и непрерывной слева в точках с координатами $x = 0$ и $x = 0{,}5$. Действительно, для $\delta > 0$ имеем равенства $\lim\limits_{\delta\to 0}(F(0) - F(0-\delta)) = 0$, $\lim\limits_{\delta\to 0}(F(0{,}5) - F(0{,}5-\delta)) = 0$. Очевидно, что функция $F(x)$ имеет конечное число скачков в точках с абсциссами $x = 0$ и $x = 0{,}5$. Это видно из соотношений: $\mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) = \lim\limits_{\delta\to 0}(F(0+\delta) - F(0-\delta)) = 0{,}1 - 0 = 0{,}1$ и $\mathsf{P}(\{\omega\colon \xi(\omega) = 0{,}5\}) = \lim\limits_{\delta\to 0}(F(0{,}5+\delta) - F(0{,}5-\delta)) = 1 - 0{,}6 = 0{,}4$.

Рассмотрим еще пример смешанной случайной величины $\xi(\omega)$, когда ее интегральная функция распределения на некоторых промежутках прямой является непрерывной, а в некоторой точке имеет скачок (разрыв первого рода).

Пример 6.22. Компьютер выходит из строя в результате транспортировки с вероятностью $p_0 > 0$. Если компьютер не вышел из строя при транспортировке, то вероятность того, что он проработает в течение времени $t > 0$, равна $\exp\{-\lambda t\}$ при $\lambda = \mathrm{const} > 0$. Определить тип случайной величины ξ, которая считает время безотказной работы компьютера.

Решение. Многие детали и условия эксплуатации компьютера нам неизвестны. Поэтому мы находимся в таких обстоятельствах, когда нет никакой возможности построить для этого эксперимента вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ и поточечно задать его количественный признак (время безотказной работы) $\xi(\omega)\colon \Omega \to R$. Мы можем только постулировать факт существования основного вероятностного пространства вида $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ и отображения $\xi(\omega)\colon \Omega \to R$. Следуя теперь содержанию замечания 6.4, воспользуемся в дальнейшем для изучения вероятностных свойств времени безотказной работы компьютера выборочным пространством $(R, \mathcal{B}, \mathsf{P}_\xi(\cdot))$, в котором функция $\mathsf{P}_\xi(\cdot)$ пока не определена.

Из физических соображений имеем $\xi(\omega) \geqslant 0$. Отсюда вероятность вида $\mathsf{P}(\{\omega\colon \xi(\omega) < t\}) = \mathsf{P}_\xi((-\infty, t)) = 0$ для $t \leqslant 0$. По условию задачи вероятность того, что компьютер откажет мгновенно в момент включения, равна $\mathsf{P}(\{\omega\colon \xi(\omega)=0\}) = \mathsf{P}_\xi(\{0\}) = p_0$. Здесь для опыта E событие $\{\omega\colon \xi(\omega)=0\}$ выхода из строя компьютера при транспортировке взаимно однозначно соответствует событию $\{0\}$ выбора точки в начале действительной прямой модельного эксперимента $E_и$. Подобным же образом из условия задачи находим, что условная вероятность $\mathsf{P}(\{\omega\colon \xi(\omega)\geqslant t\} \mid \{\omega\colon \xi(\omega)>0\}) = \mathsf{P}_\xi([t,+\infty) \mid (0,+\infty)) = \exp\{-\lambda t\}$ при $t>0$. Здесь опять случайное событие $\{\omega\colon \xi(\omega)\geqslant t\}$ безотказной работы компьютера в течение времени $t>0$ взаимно однозначно соответствует событию выбора точки на промежутке $[t,+\infty)$ действительной прямой. По теореме умножения для событий из пространства $(R,\mathcal{B},\mathsf{P}_\xi(\cdot))$ имеем соотношение вида

$$\mathsf{P}_\xi((0,+\infty)\cap[t,+\infty)) = \mathsf{P}_\xi((0,+\infty))\,\mathsf{P}_\xi([t,+\infty) \mid (0,+\infty)) = \\ = \mathsf{P}_\xi([t,+\infty))\,\mathsf{P}_\xi((0,+\infty) \mid [t,+\infty))$$

или равенство

$$(1-p_0)\exp\{-\lambda t\} = \mathsf{P}_\xi([t,+\infty))\times 1.$$

Отсюда при $t>0$ имеем, что

$$\mathsf{P}_\xi((-\infty,t)) = 1-(1-p_0)\exp\{-\lambda t\}.$$

Поэтому функция распределения $F(t)$ времени ξ безотказной работы компьютера с учетом транспортировки и эксплуатации равна

$$\mathsf{P}(\{\omega\colon \xi(\omega)<t\}) = \mathsf{P}(\varnothing) = 0$$

при $t\leqslant 0$ и равна вероятности

$$\mathsf{P}(\{\omega\colon \xi(\omega)<t\}) = 1-\mathsf{P}(\{\omega\colon \xi(\omega)\geqslant t\}) = 1-(1-p_0)\exp\{-\lambda t\}$$

при $t>0$. Значит, вероятность

$$\mathsf{P}(\{\omega\colon \xi(\omega)=0\}) = F(+0)-F(0) = p_0.$$

Итак, в этом примере величина ξ не является дискретной, так как она с положительной вероятностью может принимать несчетное число значений из промежутка $(0,\infty)$. Она не является непрерывной, так как функция $F(t)$ имеет скачок в точке $t=0$. Эта случайная величина не имеет сингулярной компоненты, так как

$$dF(t)/dt = \lambda\,(1-p_0)\exp\{-\lambda t\} > 0$$

при $t>0$. График интегральной функции распределения $F(t)$ представлен на рис. 6.23.

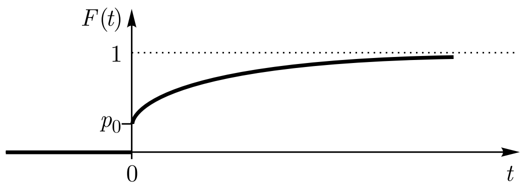


Рис. 6.23

Из этого графика видно, что $F(t)$ равна нулю при всех $t \leqslant 0$ и равна

$$1 - (1 - p_0)\exp\{-\lambda t\} = p_0 + (1 - p_0)(1 - \exp\{-\lambda t\}) = \\ = p_0 F_{\text{д}}(t) + (1 - p_0) F_{\text{ан}}(t)$$

при $t > 0$. Следовательно, каждая из функций $F_{\text{д}}(t)$ — дискретная компонента, $F_{\text{ан}}(t)$ — абсолютно непрерывная компонента, равная нулю при $t \leqslant 0$, и $F_{\text{д}}(t) = 1$, $F_{\text{ан}}(t) = 1 - \exp\{-\lambda t\}$ при $t > 0$. Так как интегральная функция распределения $F(t)$ на промежутке $(0, +\infty)$ действительной прямой является абсолютно непрерывной, а в начале координат имеет разрыв, величина которого равна p_0, то время $\xi(\omega)$ безотказной работы компьютера является смешанной случайной величиной, и представляет собой смесь дискретной и непрерывной случайных величин.

3.5. Интеграл Римана–Стилтьеса и произвольные случайные величины. Большое число вопросов и проблем в теории вероятностей решается с привлечением понятия интеграла Римана–Стилтьеса (или просто интеграла Стилтьеса). Целесообразно здесь и в дальнейшем воспользоваться общим определением интеграла Стилтьеса в пределах заданного конечного промежутка интегрирования $[a, b)$ от ограниченной функции $g(x)$ с интегрирующей функцией $F(x)$. Для этого разобьем $[a, b)$ на i промежутков Π_1, $\Pi_2, \ldots, \Pi_i$ точками с абсциссами $a = a_0 \leqslant a_1 \leqslant \ldots \leqslant a_{i-1} < a_i = b$. Для каждого из таких промежутков с номером $k = 1, 2, \ldots, i$ составим так называемое собственное приращение $\Delta F(\Pi_k)$ для интегрирующей функции $F(x)$ по следующей формуле:

$$\Delta F(\Pi_k) = \begin{cases} F(a_k) - F(a_{k-1}) & \text{при } \Pi_k = [a_{k-1}, a_k); \\ F(a_k + 0) - F(a_{k-1}) & \text{при } \Pi_k = [a_{k-1}, a_k]; \\ F(a_k) - F(a_{k-1} + 0) & \text{при } \Pi_k = (a_{k-1}, a_k); \\ F(a_k + 0) - F(a_{k-1} + 0) & \text{при } \Pi_k = (a_{k-1}, a_k]; \\ F(a_k + 0) - F(a_k) & \text{при } \Pi_k = [a_{k-1}, a_k], a_{k-1} = a_k. \end{cases} \tag{6.3}$$

Среди этих промежутков Π_1, $\Pi_2, \ldots, \Pi_i$ могут быть: 1) замкнутые слева и открытые справа; 2) отрезки; 3) интервалы; 4) открытые слева и замкнутые справа; 5) отдельные точки. Для разбиения $\Pi = \{\Pi_1, \Pi_2, \ldots, \Pi_i\}$ промежутка $[a, b)$ найдем две такие суммы: $\sum_{k=1}^{i} \inf\{g(x)\colon x \in \Pi_k\} \, \Delta F(\Pi_k)$ и $\sum_{k=1}^{i} \sup\{g(x)\colon x \in \Pi_k\} \, \Delta F(\Pi_k)$. Наконец, вычислим по всем возможным разбиениям Π промежутка $[a, b)$ точную верхнюю грань s^* первых сумм и точную нижнюю грань S^* вторых сумм. Если при этом $s^* = S^*$, то будем говорить, что существует интеграл Стилтьеса по промежутку $[a, b)$ от функции $g(x)$ относительно функции $F(x)$ и будем обозначать одинаковое числовое значение для s^* и S^* через $\int_a^b g(x) \, dF(x)$. Пусть теперь функция $g(x)$ определена в промежутке вида $(-\infty, +\infty)$ и интеграл $\int_a^b g(x) \, dF(x)$ существует при любых значенях a и b, для которых $-\infty < a < b < +\infty$. Далее, если существует предел вида $\lim\limits_{\substack{a \to -\infty \\ b \to +\infty}} \int_a^b g(x) \, dF(x)$ при стремлении произвольным способом величины a к $-\infty$ и величины b к $+\infty$, то по определению этот предел называется *несобственным интегралом Стилтьеса* в промежутке $(-\infty, +\infty)$ от функции $g(x)$ по интегрирующей функции $F(x)$ и обозначается через $\int_{-\infty}^{+\infty} g(x) \, dF(x)$. Подробно общие свойства интеграла Стилтьеса изучаются в различных учебниках по курсу математического анализа (*Гливенко В. И.* Интеграл Стилтьеса. — М.: Изд-во ЛКИ, 2007). Приведем простые задачи применения и вычисления интеграла Стилтьеса.

При изучении свойств произвольных случайных величин удобно представить вероятности $\mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) < b\})$, $\mathsf{P}(\{\omega\colon \xi(\omega) < x\})$ через интеграл Стилтьеса. Например, имеет место равенство вида

$$\mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) < b\}) = \int\limits_a^b dF(x).$$

В самом деле, для интеграла $\int_a^b dF(x)$ выполняется тождественное равенство $g(x) \equiv 1$. Используя это обстоятельство и непрерывность слева функции $F(x)$, для любого фиксированного разбиения $\Pi = \{\Pi_1, \Pi_2, \ldots, \Pi_i\}$ промежутка $[a, b)$ легко получим:

$$\sum_{k=1}^{i} \inf\{g(x)\colon x \in \Pi_k\} \, \Delta F(\Pi_k) = \sum_{k=1}^{i} \Delta F(\Pi_k) = F(b) - F(a),$$

$$\sum_{k=1}^{i} \sup\{g(x)\colon x \in \Pi_k\} \, \Delta F(\Pi_k) = \sum_{k=1}^{i} \Delta F(\Pi_k) = F(b) - F(a).$$

Отсюда находим, что

$$s^* = S^* = F(b) - F(a) = \int\limits_a^b dF(x).$$

Из равенства
$$\mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) < b\}) = F(b) - F(a)$$
имеем
$$\mathsf{P}(\{\omega\colon a \leqslant \xi(\omega) < b\}) = \int\limits_a^b dF(x).$$
Далее, из определения несобственного интеграла Стилтьеса получаем, что
$$\int\limits_{-\infty}^{x} dF(y) = \lim_{a\to-\infty}\int\limits_a^x dF(y) = \lim_{a\to-\infty}(F(x) - F(a)) = \\ = F(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\}).$$
Поэтому
$$\int\limits_{-\infty}^{x} dF(y) = \int\limits_{-\infty}^{x} f(y)\,dy = F(x)$$
для непрерывной случайной величины и
$$\int\limits_{-\infty}^{x} dF(y) = \sum_{k\colon x'_k<x} \mathsf{P}(\{\omega\colon \xi(\omega) = x'_k\}) = \sum_{k\colon x'_k<x} p_k = F(x)$$
для дискретной случайной величины. Используя теперь свойства интеграла Стилтьеса, легко вычислить $\int_{-\infty}^{x} dF(y)$, где функция $F(x) = p_0F_{\text{д}}(x) + (1-p_0)F_{\text{ан}}(x)$ является интегральной функцией распределения смешанной случайной величины из примера 6.21. Действительно, последовательно имеем
$$\int\limits_{-\infty}^{x} dF(y) = \int\limits_{-\infty}^{x} d(p_0F_{\text{д}}(y) + (1-p_0)F_{\text{ан}}(y)) = \int\limits_{-\infty}^{x} d(p_0F_{\text{д}}(y)) + \\ + \int\limits_{-\infty}^{x} d((1-p_0)F_{\text{ан}}(y)) = p_0\int\limits_{-\infty}^{x} dF_{\text{д}}(y) + (1-p_0)\int\limits_{-\infty}^{x} dF_{\text{ан}}(y).$$
Так как $F_{\text{д}}(x) = F_{\text{ан}}(x) = 0$ при $x \leqslant 0$ и $F_{\text{д}}(x) = 1$, $F_{\text{ан}}(x) = 1 - \exp\{-\lambda x\}$ при $x > 0$, то $\int_{-\infty}^{x} dF(y) = 0$ при $x \leqslant 0$ и
$$\int\limits_{-\infty}^{x} dF(y) = \int\limits_{-\infty}^{0} dF(y) + \int\limits_0^x dF(y) = p_0 + (1-p_0)\times\int\limits_0^x d(F_{\text{ан}}(y)) = \\ = p_0 + (1-p_0)(1-\exp\{-\lambda x\}) = 1 - (1-p_0)\exp\{-\lambda x\}$$
при $x > 0$. Эти простые вычисления интеграла Стилтьеса полностью повторяют значения интегральной функции распределения случайной величины ξ из примера 6.21. Результаты этого абзаца указывают на то, что для изучения вероятностных свойств произвольных случай-

ных величин целесообразно пользоваться интегралом Стилтьеса. При этом интеграл Стилтьеса сводится к обычному интегралу Римана для непрерывных случайных величин и к ряду для дискретных случайных величин.

Краткий обзор

В главе 6 на конкретных примерах дано понятие измерителей элементарных исходов статистически устойчивого эксперимента E. Приводится конструктивное задание борелевской σ-алгебры на действительной прямой, которая играет ключевую роль при определении одномерной случайной величины и изучении ее вероятностных свойств. Доказано утверждение о существовании борелевской σ-алгебры на действительной прямой. Введено определение одномерной случайной величины ξ, которая рассматривается в качестве адекватной математической модели данного измерителя элементарных событий случайного эксперимента E. На примерах показано фундаментальное значение требования измеримости для математических моделей количественных характеристик эксперимента E и понятия σ-алгебры, порожденной случайной величиной ξ.

Подробно излагаются способы задания одномерных случайных величин. Рассматриваются различные арифметические действия и операции предельного перехода над случайными величинами. Решена проблема критерия функциональной зависимости случайных величин. Для описания вероятностных свойств любой случайной величины определяется интегральная функции распределения, устанавливаются основные свойства этой функции. Отмечается практическое значение теоремы о выборочном вероятностном пространстве. На основе теоремы Лебега о разложении интегральной функции проведена замкнутая классификация одномерных случайных величин и их законов распределения. Среди множества всех одномерных случайных величин выделены дискретные, непрерывные, сингулярные и смешанные. Для изучения вероятностных свойств произвольных случайных величин дано определение интеграла Римана–Стилтьеса и приводятся его простейшие свойства.

Контрольные вопросы и упражнения

1. Дать обоснование применения наименьшей борелевской σ-алгебры $\mathcal{B} = \sigma(\mathsf{K})$ на действительной прямой при определении одномерной случайной величины.
2. Сформулировать основное отличие случайных величин от функций, которые рассматриваются в курсе математического анализа.

3. Дать физическую интерпретацию свойства измеримости случайных величин.
4. Дать обоснование использования выборочного вероятностного пространства.
5. Объяснить значение понятия σ-алгебры, порожденной случайной величиной ξ.
6. Перечислить основные законы распределения дискретных, непрерывных, сингулярных и смешанных случайных величин.
7. Пусть $(\Omega, \mathcal{F})$ есть теоретико-множественная модель произвольного статистически устойчивого эксперимента E и A — некоторое случайное событие из $\mathcal{F}$. Рассмотрим отображение $\xi(\omega)\colon \Omega \to R$, которое принимает значение единица при $\omega \in A$ и значение ноль при $\omega \notin A$. Доказать, что функция $\xi(\omega)\colon \Omega \to R$, которую называют *индикатором события* A и обозначают символом $I_A(\omega)$, является случайной величиной. Найти все вероятностные законы распределения случайной величины $I_A(\omega)$.
8. На отрезок $[0, 1]$ действительной прямой R непреднамеренно бросают точку, абсциссу которой обозначим через символ ω. Здесь множество $\mathcal{F}$ из наблюдаемых исходов есть σ-алгебра борелевских подмножеств на $[0, 1]$, а вероятность $\mathsf{P}(A)$ любого случайного события $A \subset [0, 1]$ равна его длине. Определим поточечно две функции $\xi(\omega)$ и $\eta(\omega)$ на $\Omega = [0, 1]$ следующим образом. Функция $\xi(\omega) = 1$ при $\omega \in [0, 1/2)$ и $\xi(\omega) = 0$ при $\omega \in [1/2, 1]$, а вторая функция $\eta(\omega) = 1$ при $\omega \in [0, 1/4) \cup [1/2, 3/4)$ и $\eta(\omega) = 0$ при $\omega \in [1/4, 1/2) \cup [3/4, 1]$. Доказать, что различные функции $2\,\xi(\omega)$ и $2\,\eta(\omega)$ на Ω являются случайными величинами, которые имеют одинаковые вероятностные законы распределения.
9. При работе ЭВМ время от времени возникают сбои. В этом случае количество сбоев за сутки есть случайная величина ξ с распределением вероятностей вида $\mathsf{P}(\xi = i) = \lambda^i (i!)^{-1} \exp\{-\lambda\}$, $\lambda > 0$, $i = 0, 1, \ldots$. Параметр λ определяет так называемую интенсивность числа сбоев за сутки. Найти интегральную функцию распределения случайной величины ξ и вероятность того, что в течение суток произойдет хотя бы один сбой.
10. При передаче сообщения по радиоканалу из-за наличия помех с вероятностью p сообщение не удается декодировать. Сообщение передается по радиоканалу до тех пор, пока оно не будет декодировано. Определить законы распределения случайного числа попыток для передачи сообщения и вероятность того, что на передачу сообщения понадобится больше трех попыток.
11. Автобусы подходят к остановке с интервалом в 5 мин. Случайная величина ξ определяет время ожидания автобуса пассажиром. Плотность распределения вероятностей $f(x) = 0$ при

$x \leqslant 0$, $f(x) = 0{,}2$ при $0 < x < 5$ и, наконец, $f(x) = 0$ при $x \geqslant 5$. Определить интегральную функцию распределения случайного времени ожидания автобуса и вероятность $\mathsf{P}(1 < \xi < 2{,}5)$.

12. При работе ЭВМ в случайные моменты времени возникают неисправности. Плотность распределения $f(x)$ времени работы ЭВМ до первой неисправности равна $\lambda \exp\{-\lambda t\}$ при $t > 0$ и равна нулю при $t \leqslant 0$, где параметр $\lambda > 0$. При возникновении неисправности она мгновенно обнаруживается и ЭВМ сразу поступает в ремонт. Ремонт продолжается время h, после чего ЭВМ снова включается в работу. Найти плотность распределения и функцию распределения промежутка времени ξ между двумя соседними неисправностями. Вычислить вероятность $\mathsf{P}(\xi > 2h)$.
13. Случайная величина ξ, сосредоточенная на отрезке $[1, 4]$, задана интегральной функцией распределения $F_\xi(x) = ax^2 + bx + c$, имеющей максимальное значение при $x = 4$. Вычислить вероятность попадания ξ в отрезок $[2, 3]$.
14. Построить интегральную функцию распределения случайной величины ξ и найти вероятности:

    1) $\mathsf{P}(-1 \leqslant \xi < 2)$;
    2) $\mathsf{P}(-1{,}1 < \xi \leqslant 2{,}1)$;
    3) $\mathsf{P}(0 < \xi < 2{,}1)$;
    4) $\mathsf{P}(0 \leqslant \xi \leqslant 4)$.

    Известно, что распределение вероятностей для случайной величины ξ задано рядом распределения

ξ	$-1{,}2$	0	$0{,}5$	1	$1{,}3$	2	3	4
$\mathsf{P}(\cdot)$	$0{,}05$	$0{,}02$	$0{,}04$	c	$2c$	$0{,}1$	$0{,}2$	$0{,}05$

Глава 7

СЕМЕЙСТВО ИЗМЕРИТЕЛЕЙ ИСХОДОВ СЛУЧАЙНОГО ЭКСПЕРИМЕНТА

§ 1. Многомерные случайные величины

1.1. Понятие о случайных векторах. При исследовании свойств статистически устойчивого эксперимента E часто приходится отдельно рассматривать различные количественные характеристики его элементарных исходов или случайные величины $\xi(\omega)$, $\eta(\omega), \ldots$. При этом основное вероятностное пространство $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ и свойства каждой из случайных величин $\xi(\omega)$, $\eta(\omega), \ldots$ позволяют подробно изучить эксперимент E как с качественной, так и с количественной точки зрения. Следующим шагом создания более информативной и исчерпывающей вероятностной модели статистически устойчивого эксперимента E является рассмотрение совместных свойств целого семейства (набора) $\Lambda = \{\xi(\omega), \eta(\omega), \ldots\}$ из случайных величин. К таким свойствам, например, можно отнести функциональные и статистические зависимости между случайными величинами семейства Λ. Более того, очень часто некоторые количественные признаки эксперимент E необходимо объединять и рассматривать как единую систему из двух или нескольких случайных величин. Учитывая все сказанное, мы естественным образом приходим к новой и более информативной вероятностной модели статистически устойчивого эксперимента E вида $(\Omega, \mathcal{F}, \mathsf{P}(\cdot), \Lambda)$. Приведем теперь примеры, которые поясняют целесообразность рассмотрения систем случайных величин на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$.

Пример 7.1. Опыт состоит в обнаружении станцией слежения в некотором пространстве летательного космического аппарата и пуске ракеты с целью его уничтожения. В качестве одной из количественных характеристик этого эксперимента можно взять точку разрыва снаряда ракеты вблизи летательного космического аппарата. Очевидно, что эта точка в трехмерном пространстве определяется координатами ξ_1, ξ_2 и ξ_3 или единой системой из трех упорядочных случайных величин ξ_1, ξ_2 и ξ_3.

Пример 7.2. Эксперимент заключается в наблюдении за поступлением некоторого пациента в областную больницу за фиксированный промежуток времени. Этот промежуток выбирается так, что обязательно будет только одно поступление. Элементарным исходом этого опыта является поступление конкретного больного, а описанием могут служить, например, отпечатки пальцев его правой руки. В качестве совершенно отдельных количественных характеристик этого опыта можно взять показания η_1, η_2, η_3 и η_4 приборов, которые соответственно измеряют вес, температуру, давление и простейший анализ крови больного. При этом давление η_3 само является совокупностью двух количественных характеристик, а именно артериального (верхнего) давления и венозного (нижнего) давления. Для определения предварительного диагноза больного важно знать различные типы зависимостей между случайными величинами η_1, η_2, η_3 и η_4. Поэтому состояние больного можно охарактеризовать, например, совокупностью $\Lambda = \{\eta_1, \eta_2, \eta_3, \eta_4\}$.

Определение 7.1. Упорядоченная система из n одномерных случайных величин ξ_1, $\xi_2, \ldots$, ξ_n, заданных на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, называется *случайным вектором* или *многомерной случайной величиной размерности* n.

Итак, каждому $\omega \in \Omega$ ставим в соответствие n-мерный вектор $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$, другими словами, имеем однозначное отображение $\xi(\omega)\colon \Omega \to R^n$. Здесь множество

$$R^n = \{\mathbf{x} = (x_1, x_2, \ldots, x_n)\colon -\infty < x_i < +\infty,\ i = 1, 2, \ldots, n\}$$

есть декартово (прямое) произведение n экземпляров действительной прямой $R = \{x\colon -\infty < x < +\infty\}$. Рассмотрим простой иллюстративный пример такой ситуации при $n = 2$, $x_1 = x$, $x_2 = y$.

Пример 7.3. Пусть эксперимент заключается в запуске ракеты для поражения цели на земле. Пусть положение объекта имеет координаты $(0, 0)$. Тогда положение случайной точки Q (рис. 7.1) падения ракеты на плоскости $x0y$ можно определить двумя случайными величинами $\xi_1(\omega)$ и $\xi_2(\omega)$. При этом важно знать не только, какие точки некоторой области на плоскости наблюдаемы, но и с какой вероятностью они появляются.

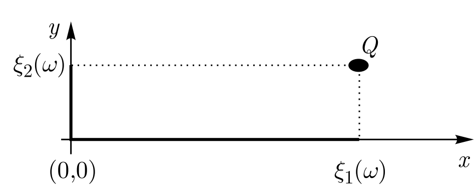


Рис. 7.1

Из определения 7.1 и рис. 7.1 следует, что каждой точке ω пространства Ω описаний элементарных событий статистически устойчивого эксперимента E соответствует упорядоченный набор чисел $(\xi_1(\omega), \xi_2(\omega))$. В этом случае говорят, что поточечно задана двумерная случайная величина, или задан двумерный случайный вектор. Как и в одномерном случае ($n = 1$) следует рассматривать измеримое отображение вида $\xi(\omega)\colon \Omega \to R^n$, для которого множество $\{\omega\colon \xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) \in B^n\} \in \mathcal{F}$ при любом борелевском множестве B^n из σ-алгебры $\mathcal{B}^n$. Напомним, что $\mathcal{B}^n$ есть наименьшая σ-алгебра, порожденная в пространстве R^n системой прямоугольников вида $\{(x_1, x_2, \ldots, x_n)\colon a_i \leqslant x_i < b_i,\ i = 1, 2, \ldots, n\}$. Семейство вероятностей

$$\{\mathsf{P}(\{\omega\colon (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) \in B^n\}) = \mathsf{P}_\xi(B^n)\colon B^n \in \mathcal{B}^n\}$$

называется *распределением многомерной случайной величины*. Поскольку множество $\{\omega\colon \xi_1(\omega) < x_1,\ \xi_2(\omega) < x_2, \ldots, \xi_n(\omega) < x_n\}$ равно $\bigcap_{i=1}^n \{\omega\colon \xi_i(\omega) < x_i\}$ и $\xi_1(\omega)$, $\xi_2(\omega), \ldots$, $\xi_n(\omega)$ являются одномерными случайными величинами, то случайное событие $\{\omega\colon \xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \ldots, \xi_n(\omega) < x_n\} \in \mathcal{F}$. Подобно тому, как и в одномерном случае, можно установить, что выполнение условия $\{\omega\colon \xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \ldots, \xi_n(\omega) < x_n\} \in \mathcal{F}$ для $(x_1, x_2, \ldots, x_n)$ из пространства R^n является достаточным для измеримости отображения $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))\colon \Omega \to R^n$. Поэтому часто можно встретить следующее определение многомерной случайной величины.

Определение 7.2. Всякое измеримое отображение $\xi(\omega)\colon \Omega \to R^n$ или, другими словами, $\{\omega\colon (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) \in B^n\} \in \mathcal{F}$, где $B^n \in \mathcal{B}^n$, называется *n-мерной случайной величиной*.

Однако определение 7.1 многомерной случайной величины следует признать естественным, а с практической точки зрения наиболее простым. Поскольку при фиксированном наборе чисел x_1, $x_2, \ldots$, x_n множество $\{\omega\colon \xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \ldots, \xi_n(\omega) < x_n\}$ является событием, то можно определить *многомерную интегральную функцию распределения* $F(x_1, x_2, \ldots, x_n) = \mathsf{P}(\{\omega\colon (\xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \ldots, \xi_n(\omega) < x_n\})$ для n-мерного вектора $\xi(\omega) = (\xi_1, \xi_2, \ldots, \xi_n)$. Эта функция определяет вероятность совместного наступления или пересечения событий $\{\omega\colon \xi_1(\omega) < x_1\}$, $\{\omega\colon \xi_2(\omega) < x_2\}, \ldots$, $\{\omega\colon \xi_n(\omega) < x_n\}$, которые порождаются случайными величинами ξ_1, $\xi_2, \ldots$, ξ_n. Ради упрощения записей и геометрической интерпретации изложения материала в дальнейшем будем подробно рассматривать только двумерные случайные величины ($n = 2$, $x_1 = x$, $x_2 = y$). Изучение свойств n-мерных случайных величин полностью аналогично. Более того, свойства n-мерных и двумерных функций распределения совпадают. Предварительно заметим, что при переходе от одномерных случайных величин к многомерным возникают дополнительные свойства интегральной функции распределения $F(x_1, x_2, \ldots, x_n)$.

1.2. Двумерная случайная величина и свойства ее интегральной функции распределения. Из определения 7.1 при $n = 2$ получаем, что *двумерной случайной величиной* $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega))\colon \Omega \to R^2$ называется упорядоченная система из одномерных случайных величин $\xi_1(\omega)$ и $\xi_2(\omega)$, заданных на $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. В дальнейшем для наглядности упорядоченный набор (ξ_1, ξ_2) из двух случайных величин можно отобразить случайной точкой на плоскости с абсциссой ξ_1 и ординатой ξ_2 (см. рис. 7.2).

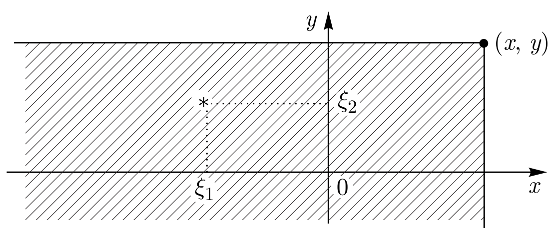


Рис. 7.2

Далее, при $(x, y) \in R^2$ интегральной функцией $F(x, y)$ двумерной случайной величины (ξ_1, ξ_2) называется вероятность $\mathsf{P}(\{\omega\colon \xi_1(\omega) < x, \xi_2(\omega) < y\})$. Можно сказать, что интегральная функция распределения $F(x, y)$ двумерной случайной величины $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega))$ есть вероятность того, что первая случайная величина $\xi_1(\omega)$ примет значение меньше числа x и одновременно вторая случайная величина $\xi_2(\omega)$ примет значение меньше y. Тогда двумерная функция распределения $F(x, y)$ есть вероятность того, что случайная точка (ξ_1, ξ_2) попадет в заштрихованную на рис. 7.2 область. При этом в эту область стороны прямого угла с вершиной в точке (x, y) не включаются. Такая геометрическая интерпретация и терминология во многом позволяет понять на содержательном уровне и наглядно проиллюстрировать основные свойства функции $F(x, y)$.

Сформулируем теперь такие свойства многомерной интегральной функции распределения, которые совершенно аналогичны свойствам интегральной функции распределения одномерной случайной величины. Более того, эти свойства $F(x, y)$ доказываются почти так же, как и в одномерном случае. Итак, справедливы следующие утверждения.

Утверждение 7.1. *Функция $F(x, y)$ принимает значение на отрезке $[0, 1]$, поскольку по определению она равна вероятности* $\mathsf{P}(\{\omega\colon \xi_1(\omega) < x, \xi_2(\omega) < y\})$.

Утверждение 7.2. *Функция $F(x, y)$ является неубывающей функцией по каждому из своих аргументов x и y.*

Для доказательства обратимся к рис. 7.3. Действительно, если $a_1 < a_2$ и события $A_{a_1} = \{\omega\colon \xi_1(\omega) < a_1,\ \xi_2(\omega) < y\}$, $A_{a_2} = \{\omega\colon \xi_1(\omega) < a_2,\ \xi_2(\omega) < y\}$, $A_{a_1,a_2} = \{\omega\colon a_1 \leqslant \xi_1(\omega) < a_2,\ \xi_2(\omega) < y\}$ порождены вектором $(\xi_1(\omega), \xi_2(\omega))$, то $A_{a_2} = A_{a_1} \cup A_{a_1,a_2}$, $A_{a_1} \cap A_{a_1,a_2} = \varnothing$ и вероятность $\mathsf{P}(A_{a_2}) = \mathsf{P}(A_{a_1}) + \mathsf{P}(A_{a_1,a_2})$, или интегральная функция $F(a_2, y) = F(a_1, y) + \mathsf{P}(\{\omega\colon a_1 \leqslant \xi_1(\omega) < a_2, \xi_2(\omega) < y\}) \geqslant F(a_1, y)$. Отсюда, используя для наглядности еще раз рис. 7.3, легко получаем важную формулу:

$$\mathsf{P}(\{\omega\colon a_1 \leqslant \xi_1(\omega) < a_2,\ \xi_2(\omega) < y\}) = F(a_2, y) - F(a_1, y). \qquad (7.1)$$

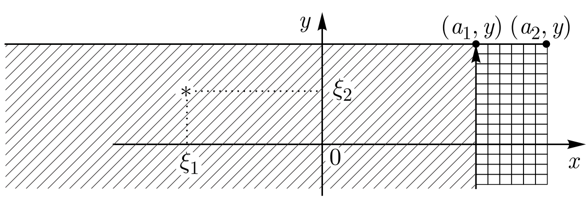


Рис. 7.3

Можно сказать, что формула (7.1) вычисляет вероятность попадания случайной точки (ξ_1, ξ_2) в область с двойной штриховкой (рис. 7.3). При этом граница, включаемая в эту область, помечена жирной линией со стрелкой. На интуитивном уровне из рис. 7.3 наглядно видно, что при увеличении первого аргумента от значения a_1 до a_2 заштрихованная косыми линиями область увеличивается на область с двойной штриховкой. Очевидно, вероятность попадания случайной точки (ξ_1, ξ_2) в большую область не может уменьшиться. По такой же схеме получим, что при $b_2 > b_1$ имеем неравенство $F(x, b_2) \geqslant F(x, b_1)$ и дополнительно простую формулу

$$\mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\ b_1 \leqslant \xi_2(\omega) < b_2\}) = F(x, b_2) - F(x, b_1). \qquad (7.2)$$

Утверждение 7.3. *Интегральная функция распределения $F(x, y)$ при $i \to \infty$ удовлетворяет предельным равенствам вида*:

$$\lim_{a_{-i} \to -\infty} F(a_{-i}, y) = \lim_{b_{-i} \to -\infty} F(x, b_{-i}) = 0,$$
$$\lim_{b_i \to +\infty} F(x, b_i) = F(x, +\infty) = F_{\xi_1}(x),$$
$$\lim_{a_i \to +\infty} F(a_i, y) = F(+\infty, y) = F_{\xi_2}(y),$$

$$\lim_{\substack{a_i \to +\infty \\ b_i \to +\infty}} F(a_i, b_i) = F(+\infty, +\infty) = 1,$$

$$\lim_{\substack{a_{-i} \to -\infty \\ b_{-i} \to -\infty}} F(a_{-i}, b_{-i}) = F(-\infty, -\infty) = 0,$$

где $F_{\xi_1}(x) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\})$ *и* $F_{\xi_2}(y) = \mathsf{P}(\{\omega\colon \xi_2(\omega) < y\})$ — *интегральные функции распределения одномерных случайных величин* $\xi_1(\omega)$ *и* $\xi_2(\omega)$.

Доказательство. Дополнительно можно определить, что функция вида

$$F(+\infty, y) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < +\infty,\ \xi_2(\omega) < y\}).$$

Поэтому имеем

$$\begin{aligned} F(+\infty, y) &= \mathsf{P}(\{\omega\colon \xi_1(\omega) < +\infty,\ \xi_2(\omega) < y\}) = \\ &= \mathsf{P}(\{\omega\colon \xi_1(\omega) < +\infty\} \cap \{\omega\colon \xi_2(\omega) < y\}) = \\ &= \mathsf{P}(\Omega \cap \{\omega\colon \xi_2(\omega) < y\}) = \mathsf{P}(\{\omega\colon \xi_2(\omega) < y\}) = F_{\xi_2}(y). \end{aligned}$$

В дальнейшем, ради простоты записи, вместо последовательностей $\{a_i;\ i = 0, 1, \ldots\}$ и $\{a_{-i};\ i = 1, 2, \ldots\}$ таких, что $a_0 < a_1 < \ldots < a_i < a_{i+1} < \ldots$, $\lim\limits_{i\to\infty} a_i = +\infty$ и $a_0 > a_{-1} > \ldots > a_{-i} > a_{-i-1} > \ldots$, $\lim\limits_{i\to\infty} a_{-i} = -\infty$, будем использовать последовательности $\{a_i = i;\ i = 0, 1, \ldots\}$, $\{a_{-i} = -i;\ i = 1, 2, \ldots\}$. Методика работы с более сложными последовательностями $\{a_i;\ i = 0, 1, \ldots\}$, $\{a_{-i};\ i = 1, 2, \ldots\}$ приведена при доказательстве теоремы 6.7. Далее, вероятность

$$\mathsf{P}(\{\omega\colon \xi_1(\omega) < +\infty\} \cap \{\omega\colon \xi_2(\omega) < y\})$$

можно представить в таком виде:

$$\begin{aligned} &\mathsf{P}(\{\omega\colon \xi_1(\omega) < +\infty\} \cap \{\omega\colon \xi_2(\omega) < y\}) = \\ &\quad = \mathsf{P}\left(\left(\bigcup_{k=-\infty}^{+\infty} \{\omega\colon k-1 \leqslant \xi_1(\omega) < k\}\right) \bigcap \{\omega\colon \xi_2(\omega) < y\}\right) = \\ &\quad = \mathsf{P}\left(\bigcup_{k=-\infty}^{+\infty} (\{\omega\colon k-1 \leqslant \xi_1(\omega) < k\} \cap \{\omega\colon \xi_2(\omega) < y\})\right) = \\ &\quad = \mathsf{P}\left(\bigcup_{k=-\infty}^{+\infty} \{\omega\colon k-1 \leqslant \xi_1(\omega) < k,\ \xi_2(\omega) < y\}\right) = \\ &\quad = \sum_{k=-\infty}^{+\infty} \mathsf{P}(\{\omega\colon k-1 \leqslant \xi_1(\omega) < k,\ \xi_2(\omega) < y\}). \end{aligned}$$

Итак,

$$F_{\xi_2}(y) = \sum_{k=-\infty}^{+\infty} \mathsf{P}(\{\omega\colon k-1 \leqslant \xi_1(\omega) < k,\, \xi_2(\omega) < y\}),$$

поэтому имеем неравенство

$$0 \leqslant \sum_{k=-\infty}^{+\infty} \mathsf{P}(\{\omega\colon k-1 \leqslant \xi_1(\omega) < k,\, \xi_2(\omega) < y\}) \leqslant 1,$$

т. е. указанный ряд сходится. В таком случае его частичные суммы вида

$$S_i = \sum_{k=1-i}^{i} \mathsf{P}(\{\omega\colon k-1 \leqslant \xi_1(\omega) < k,\, \xi_2(\omega) < y\}), \quad i = 0, 1, \ldots,$$

будут сходиться к функции $F_{\xi_2}(y)$. Используя формулу (7.1), сосчитаем частичную сумму

$$S_i = \sum_{k=1-i}^{i} \mathsf{P}(\{\omega\colon k-1 \leqslant \xi_1(\omega) < k,\, \xi_2(\omega) < y\}) = \\ = \sum_{k=1-i}^{i} (F(k,y) - F(k-1,y)) = F(i,y) - F(-i,y).$$

Последовательность $\{F(i,y); i \geqslant 0\}$ не убывает по i и ограничена сверху единицей, а последовательность $\{F(-i,y); i \geqslant 0\}$ не возрастает по i и ограничена снизу нулем. Поэтому существуют пределы $\lim\limits_{i\to+\infty} F(i,y)$, $\lim\limits_{i\to+\infty} F(-i,y)$. Теперь можно записать, что

$$\lim_{i\to+\infty} S_i = \lim_{i\to+\infty} F(i,y) - \lim_{i\to+\infty} F(-i,y) = F_{\xi_2}(y).$$

Так как при $i = 0, 1, \ldots$ имеем

$$F(i,y) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < i,\, \xi_2(\omega) < y\}) \leqslant \\ \leqslant \mathsf{P}(\{\omega\colon \xi_1(\omega) < +\infty,\, \xi_2(\omega) < y\}) = F_{\xi_2}(y),$$

то $\lim\limits_{i\to+\infty} F(i,y) \leqslant F_{\xi_2}(y)$. С другой стороны, $\lim\limits_{i\to+\infty} F(-i,y) \geqslant 0$, $\lim\limits_{i\to+\infty} F(i,y) = F_{\xi_2}(y) + \lim\limits_{i\to+\infty} F(-i,y)$, поэтому $\lim\limits_{i\to+\infty} F(i,y)$ не может быть меньше, чем $F_{\xi_2}(y)$. Отсюда сразу получаем, что $\lim\limits_{i\to+\infty} F(i,y) = F_{\xi_2}(y)$ и $\lim\limits_{i\to+\infty} F(-i,y) = 0$. Утверждения о том, что пределы $\lim\limits_{b_{-i}\to+\infty} F(x, b_{-i}) = 0$, $\lim\limits_{b_i\to+\infty} F(x, b_i) = F(x, +\infty) = F_{\xi_1}(x)$ и $\lim\limits_{\substack{a_i\to+\infty\\ b_i\to+\infty}} F(a_i, b_i) = 1$, $\lim\limits_{\substack{a_{-i}\to-\infty\\ b_{-i}\to-\infty}} F(a_{-i}, b_{-i}) = 0$ доказываются аналогичным образом. Утверждение 7.3 установлено.

Замечание 7.1. Полученные формулы $\lim\limits_{b_i \to +\infty} F(x, b_i) = F(x, +\infty) = F_{\xi_1}(x)$, $\lim\limits_{a_i \to +\infty} F(a_i, y) = F(+\infty, y) = F_{\xi_2}(y)$ позволяют найти одномерные (маргинальные) законы распределения каждой случайной величины ξ_1 и ξ_2 отдельно, зная закон распределения совокупности (ξ_1, ξ_2) случайных величин. В общем случае обратную задачу решить не удается. Действительно, если знаем одномерные интегральные функции $F_{\xi_1}(x)$ и $F_{\xi_2}(y)$, то двумерная функция распределения $F(x, y)$ не всегда может быть определена. В самом деле, из определения $F(x, y)$, используя теорему умножения, имеем

$$\begin{aligned} F(x, y) &= \mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\ \xi_2(\omega) < y\}) = \\ &= \mathsf{P}(\{\omega\colon \xi_2(\omega) < y\}) \mid \{\omega\colon \xi_1(\omega) < x\}) \times \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\}) = \\ &= F_{\xi_1}(x)\, \mathsf{P}(\{\omega\colon \xi_2(\omega) < y\}) \mid \{\omega\colon \xi_1(\omega) < x\}). \end{aligned}$$

Аналогичным способом выводим равенство

$$\begin{aligned} F(x, y) &= \mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\ \xi_2(\omega) < y\}) = \\ &= F_{\xi_2}(y)\, \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\} \mid \{\omega\colon \xi_2(\omega) < y\}). \end{aligned}$$

Однако условные вероятности $\mathsf{P}(\{\omega\colon \xi_2(\omega) < y\} \mid \{\omega\colon \xi_1(\omega) < x\})$, $\mathsf{P}(\{\omega\colon \xi_1(\omega) < x\} \mid \{\omega\colon \xi_2(\omega) < y\})$ неизвестны и невозможно их вычислить. Значит, одномерные распределения $F_{\xi_1}(x)$ и $F_{\xi_2}(y)$ не восстанавливают $F(x, y)$. Можно сказать, что совместное распределение $F(x, y)$ содержит больше информации, чем одномерные распределения $F_{\xi_1}(x)$ и $F_{\xi_2}(y)$. Это еще одно обстоятельство важности рассмотрения многомерных случайных величин. В случае, когда события $\{\omega\colon \xi_1(\omega) < x\}$ и $\{\omega\colon \xi_2(\omega) < y\}$ при $(x, y) \in R^2$ независимы, двумерную функцию распределения можно вычислить, зная одномерные распределения. При этом получаем

$$\begin{aligned} F(x, y) &= \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\} \cap \{\omega\colon \xi_2(\omega) < y\}) = \\ &= \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\})\, \mathsf{P}(\{\omega\colon \xi_2(\omega) < y\}) = F_{\xi_1}(x) \times F_{\xi_2}(y). \end{aligned}$$

Утверждение 7.4. *Интегральная функция распределения $F(x, y)$ непрерывна слева по каждому аргументу.*

Для этого достаточно показать, что для любых строго возрастающих числовых последовательностей $\{a_i; i \geqslant 1\}$ и $\{b_i; i \geqslant 1\}$, которые сходятся при $i \to \infty$ соответственно к числам a и b, выполняются равенства: $F(a-0, y) = \lim\limits_{a_i \to a,\, a_i < a} F(a_i, y) = F(a, y)$, $F(x, b-0) = \lim\limits_{b_i \to b,\, b_i < b} F(x, b_i) = F(x, b)$. Используя опыт доказательств свойств 7.1–7.3 для интегральной функции $F(x, y)$ и метод доказательства теоремы 6.5, нетрудно установить сформулированное здесь утверждение.

Рассмотрим теперь следующую игровую задачу, на которой проиллюстрируем указанные свойства интегральной функции распределения $F(x, y)$. Некто ежегодно наудачу выбирает точку на отрезке $[-1, 1]$ с абсциссой ξ_1 и вносит в благотворительный фонд развития культуры денежную сумму в размере $\xi_2 = (\xi_1)^2$ от некоторого единичного капитала. Прежде всего, найдем интегральную функцию распределения $F_{\xi_1}(x)$ случайной величины ξ_1. В данной задаче элементарным исходом $\{\omega\}$ является выбранная точка на отрезке $[-1, 1]$ с описанием в виде абсциссы $\omega = x$, достоверный исход $\Omega = [-1, 1] = \{\omega\colon -1 \leqslant \omega \leqslant 1\}$, а множество наблюдаемых исходов $\mathcal{F}$ есть σ-алгебра борелевских подмножеств на отрезке $[-1, 1]$. Вероятность $\mathsf{P}(A)$ любого борелевского подмножества (случайного события) $A \subset [-1, 1]$ равна половины его длины. Так как $\xi_1(\omega) = \omega$, то при $x \leqslant -1$ интегральная функция распределения $F_{\xi_1}(x) = \mathsf{P}(\{\omega \in \Omega\colon \xi_1(\omega) = \omega < x\}) = \mathsf{P}(\varnothing) = 0$. При $-1 < x \leqslant 1$ функция $F_{\xi_1}(x) = \mathsf{P}(\{\omega \in \Omega\colon \xi_1(\omega) = \omega < x\}) = \mathsf{P}(\{\omega\colon -1 \leqslant \omega < x\}) = (x+1)/2$. Наконец, при $x > 1$ функция $F_{\xi_1}(x) = \mathsf{P}(\{\omega \in \Omega\colon \xi_1(\omega) = \omega < x\}) = \mathsf{P}(\{\omega\colon -1 \leqslant \omega \leqslant 1\}) = 2/2 = 1$.

Найдем теперь интегральную функцию распределения $F(x, y)$ случайного вектора (ξ_1, ξ_2). С этой целью на плоскости $x0y$ введем в рассмотрение области S_1, S_2, S_3, S_4 и S_5 (см. рис. 7.4).

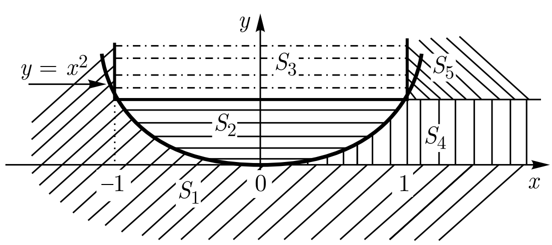


Рис. 7.4

При этом

$$S_1 = \{(x, y)\colon y \leqslant 0\} \cup \{(x, y)\colon x \leqslant -1\} \cup \{(x, y)\colon x \leqslant -y^{1/2},\ y > 0\},$$
$$S_2 = \{(x, y)\colon -y^{1/2} < x \leqslant y^{1/2},\ 0 < y \leqslant 1\},$$
$$S_3 = \{(x, y)\colon -1 < x \leqslant 1,\ y > 1\},$$
$$S_4 = \{(x, y)\colon x > y^{1/2},\ 0 < y \leqslant 1\},$$
$$S_5 = \{(x, y)\colon x > 1,\ y > 1\}.$$

В силу равенства $\xi_2 = (\xi_1)^2$ получаем, что при $(x, y) \in S_1$ интегральная функция распределения $F(x, y) = \mathsf{P}(\{\omega \in \Omega\colon \xi_1(\omega) < x, (\xi_1(\omega))^2 < y\}) = \mathsf{P}(\varnothing) = 0$.

Если $(x, y) \in S_2$, то $F(x, y) = \mathsf{P}(\{\omega \in \Omega\colon \xi_1(\omega) < x, (\xi_1(\omega))^2 < y\}) = \mathsf{P}(\{\omega \in \Omega\colon -y^{1/2} < \xi_1(\omega) < x\}) = F_{\xi_1}(x) - F_{\xi_1}(-y^{1/2}) = (x + y^{1/2})/2$.

При $(x,y)\in S_3$ интегральная функция $F(x,y)=\mathsf{P}(\{\omega\in\Omega\colon \xi_1(\omega)<x, (\xi_1(\omega))^2<y\})=\mathsf{P}(\{\omega\colon \xi_1(\omega)<x\})=F_{\xi_1}(x)=(x+1)/2$.

Для $(x,y)\in S_4$ легко получаем, что двумерная функция распределения $F(x,y)$ равна $\mathsf{P}(\{\omega\in\Omega\colon \xi_1(\omega)<x,\ (\xi_1(\omega))^2<y\})=$ $=\mathsf{P}(\{\omega\in\Omega\colon -y^{1/2}<\xi_1(\omega)<y^{1/2}\})=F_{\xi_1}(y^{1/2})-F_{\xi_1}(-y^{1/2})=y^{1/2}$.

Наконец, при $(x,y)\in S_5$ интегральная двумерная функция $F(x,y)=\mathsf{P}(\{\omega\in\Omega\colon \xi_1(\omega)<x,(\xi_1(\omega))^2<y\})=\mathsf{P}(\Omega)=1$.

Первые четыре свойства функции $F(x,y)$, очевидно, выполняются для областей S_1 и S_5, а для остальных областей эти свойства следуют из свойств одномерной функции распределения $F_{\xi_1}(x)$.

В качестве примера проверим выполнение предельного свойства $\lim\limits_{i\to\infty}F(i,y)=F_{\xi_2}(y)$. Сначала вычислим функцию $F_{\xi_2}(y)$ непосредственно из определения случайной величины $\xi_2(\omega)$. Так как $\xi_2=(\xi_1)^2$, то функция вида $F_{\xi_2}(y)=\mathsf{P}(\{\omega\colon \xi_2(\omega)<y\})=\mathsf{P}(\{\omega\colon (\xi_1(\omega))^2<y\})=0$ при $y\leqslant 0$ и $F_{\xi_2}(y)=\mathsf{P}(\{\omega\colon -y^{1/2}<\xi_1(\omega)<y^{1/2}\})=F_{\xi_1}(y^{1/2})-$ $-F_{\xi_1}(-y^{1/2})$ при $y>0$. Отсюда, используя значения функции $F_{\xi_1}(x)$, вычислим, что $F_{\xi_2}(y)=y^{1/2}$ при $0<y\leqslant 1$ и $F_{\xi_2}(y)=1$ при $y>1$. Теперь, принимая во внимание значения функции $F(x,y)$ в областях S_1, S_4 и S_5, легко найдем, что предел $\lim\limits_{i\to\infty}F(i,y)$ равен нулю при $y\leqslant 0$, равен $y^{1/2}$ при $0<y\leqslant 1$ и, наконец, равен единице при $y>1$. Итак, действительно, имеет место равенство $\lim\limits_{i\to\infty}F(i,y)=F_{\xi_2}(y)$.

Изучим теперь следующий вопрос. Зная функцию распределения $F(x,y)$ упорядоченной системы величин $\xi_1(\omega)$ и $\xi_2(\omega)$, найти вероятность попадания (ξ_1,ξ_2) в заштрихованную прямоугольную область на плоскости со сторонами, параллельными осям координат (см. рис. 7.5). При этом границы, включаемые в прямоугольник, помечены жирными линиями. Для решения этой задачи установим следующее утверждение.

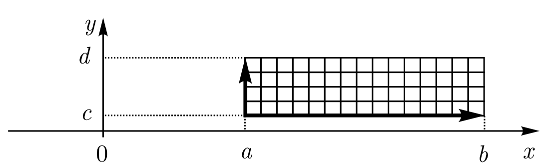


Рис. 7.5

Лемма 7.1. *Если* (ξ_1,ξ_2) *— двумерная случайная величина, то для любых* $a<b$ и $c<d$ *имеет место соотношение*

$$\mathsf{P}(\{\omega\colon a\leqslant\xi_1(\omega)<b,\ c\leqslant\xi_2(\omega)<d\})= \\ =F(b,d)-F(b,c)-F(a,d)+F(a,c)\geqslant 0.$$

Доказательство. Прежде всего, докажем равенство. С этой целью введем следующие случайные события:

$$A = \{\omega\colon \xi_1(\omega) < a,\ \xi_2(\omega) < c\},$$
$$B = \{\omega\colon \xi_1(\omega) < a,\ \xi_2(\omega) < d\},$$
$$C = \{\omega\colon \xi_1(\omega) < b,\ \xi_2(\omega) < d\},$$
$$D = \{\omega\colon \xi_1(\omega) < b,\ \xi_2(\omega) < c\},$$
$$G = \{\omega\colon a \leqslant \xi_1(\omega) < b,\ c \leqslant \xi_2(\omega) < d\}.$$

Эти события порождены случайным вектором $(\xi_1(\omega), \xi_2(\omega))$. Нетрудно проверить, что событие $C = G \cup B \cup D$, причем события G и $(B \cup D)$ несовместны, а случайное событие $(B \cap D) = A$. Тогда по теореме сложения вероятность $\mathsf{P}(C) = \mathsf{P}(G \cup B \cup D) = \mathsf{P}(G) + \mathsf{P}(B \cup D) = \mathsf{P}(B) + \mathsf{P}(D) - \mathsf{P}(B \cap D) + \mathsf{P}(G) = \mathsf{P}(B) + \mathsf{P}(D) - \mathsf{P}(A) + \mathsf{P}(G)$. Следовательно, $\mathsf{P}(C) = \mathsf{P}(B) + \mathsf{P}(D) - \mathsf{P}(A) + \mathsf{P}(G)$. Вспоминая теперь определения событий A, B, C, D, G и функции $F(x, y)$, находим равенство $F(b, d) = F(a, d) + F(b, c) - F(a, c) + \mathsf{P}(\{\omega\colon a \leqslant \xi_1(\omega) < b,\ c \leqslant \xi_2(\omega) < d\})$. Отсюда получаем соотношение

$$\mathsf{P}(\{\omega\colon a \leqslant \xi_1(\omega) < b,\ c \leqslant \xi_2(\omega) < d\}) = F(b, d) - F(b, c) - F(a, d) + F(a, c), \quad (7.3)$$

которое вычисляет вероятность попадания двумерной случайной точки в прямоугольник $\{(x, y)\colon a \leqslant x < b,\ c \leqslant y < d\}$ на плоскости (рис. 7.5). Заметим, что установленное равенство (7.3) позволяет получить еще раз соотношение (7.1) при $a = a_1$, $b = a_2$, $c = -\infty$, $d = y$ и формулу (7.2) при $a = -\infty$, $b = x$, $c = b_1$, $d = b_2$. Так как вероятностная функция неотрицательна, то необходимо выполнение неравенства $\mathsf{P}(\{\omega\colon a \leqslant \xi_1(\omega) < b,\ c \leqslant \xi_2(\omega) < d\}) \geqslant 0$. Значит, для функции распределения двумерной случайной величины должно выполняться следующее неравенство: $F(b, d) - F(b, c) - F(a, d) + F(a, c) \geqslant 0$, что и требовалось доказать.

1.3. Достаточные условия существования двумерной интегральной функции. Из теорем 6.3–6.5 и 6.7 непосредственно следует, что перечисленные утверждения 7.1–7.4 для многомерных случайных величин верны и для одномерных случайных величин. Эти свойства с учетом утверждения теоремы 6.8 о выборочном вероятностном пространстве являются необходимыми и достаточными для того, чтобы любую функцию $G(x)$ с такими свойствами объявить интегральной функцией распределения некоторой одномерной случайной величины. Однако в многомерном случае утверждения 7.1–7.4 не будут достаточными, чтобы любая функция с такими свойствами была интегральной. Для доказательства этого факта приведем простой и хорошо известный в теории вероятностей пример [6].

Пример 7.4. Пусть функция $G(u, v)$ равна нулю при $(u, v) \in \in \{(x, y)\colon x + y \leqslant 1\} \cup \{(x, y)\colon x \leqslant 0\} \cup \{(x, y)\colon y \leqslant 0\}$ и равна единице в остальной части плоскости $x0y$ (см. рис. 7.6, *а*). Показать, что функция $G(x, y)$ не является интегральным законом распределения некоторой двумерной случайной величины.

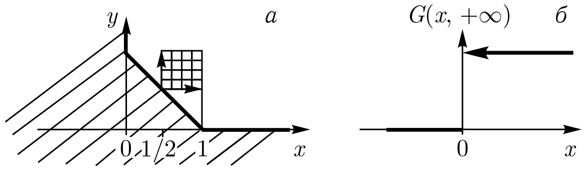


Рис. 7.6

Решение. По условию задачи функция $G(x, y)$ равна нулю в заштрихованной косыми линиями области рис. 7.6, *а* и равна единице в остальной части плоскости. Используя определение функции $G(x, y)$ и этот рисунок, легко проверить, что эта функция удовлетворяет утверждениям 7.1–7.4:

1) $0 \leqslant G(x, y) \leqslant 1$;
2) $G(x, y)$ — неубывающая функция по каждому из аргументов x и y;
3) $\lim\limits_{\substack{a_i \to +\infty \\ b_i \to +\infty}} G(a_i, b_i) = 1, \quad \lim\limits_{a_i \to -\infty} G(a_i, y) = \lim\limits_{b_i \to -\infty} G(x, b_i) = 0$ и каждая из функций $G(+\infty, y) = \lim\limits_{a_i \to +\infty} G(a_i, y)$, $G(x, +\infty) = = \lim\limits_{b_i \to +\infty} G(x, b_i)$ является так называемым вырожденным распределением, сосредоточенным в точке ноль (см. рис. 7.6, *б*);
4) функция $F(x, y)$ непрерывна слева по каждому из своих аргументов.

Однако, не смотря на все это, функцию $G(x, y)$ нельзя считать интегральной функцией распределения какой-либо двумерной случайной величины. В самом деле, предположим обратное. Итак, пусть существует некоторый эксперимент E, для которого можно построить вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ и некоторую двумерную случайную величину вида $(\xi_1(\omega), \xi_2(\omega))\colon \Omega \to R^2$ такую, что ее интегральная функция распределения $F(x, y) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\ \xi_2(\omega) < y\})$ равна $G(x, y)$. Из (7.3) найдем $\mathsf{P}(\{\omega\colon 1/2 \leqslant \xi_1(\omega) < 1,\ 1/2 \leqslant \xi_2(\omega) < 1\})$ — вероятность попадания двумерной случайной точки (ξ_1, ξ_2) в квадрат, который изображен на рис. 7.6, *а* в виде двойной заштрихованной области.

Итак, легко получаем: $\mathsf{P}(\{\omega\colon 1/2 \leqslant \xi_1(\omega) < 1,\ 1/2 \leqslant \xi_2(\omega) < 1\}) = = G(1, 1) - G(1/2, 1) - G(1, 1/2) + G(1/2, 1/2) = 1 - 1 - 1 + 0 = -1 < 0.$

Следовательно, приходим к абсурдному неравенству и функция $G(x,y)$ не может быть интегральной функцией распределения. На этом простом примере было продемонстрировано, что многомерный случай по существу отличается от одномерного. Более того, этот пример также доказывает, что пятое свойство двумерной интегральной функции распределения (утверждение леммы 7.1) не следует из первых четырех.

Заметим, что выполнение утверждений 7.1–7.4 и неравенства $F(b,d)-F(b,c)-F(a,d)+F(a,c) \geqslant 0$ при любых $a<b$, $c<d$ являются не только необходимыми, но и достаточными для существования интегральной функции распределения $F(x,y)$ некоторой двумерной случайной величины $\xi=(\xi_1(\omega),\xi_2(\omega))$. Доказательство этого факта полностью повторяет одномерный случай о построении выборочного вероятностного пространства $(R,\mathcal{B},\mathsf{P}_\xi(\cdot))$.

Приведем теперь упрощенную схему доказательства этого утверждения. В двумерном случае модельный эксперимент $E_и$ заключается в бросании точки на плоскость R^2 с помощью некоторого случайного механизма. Исход эксперимента $E_и$ может, например, состоять в том, что точка $\{\omega_и\}$ с описанием $\omega_и=(x,y)$ выбрана из прямоугольника вида $\{(x,y)\colon a\leqslant x<b,\ c\leqslant y<d\}$. Здесь x и y обозначают соответственно абсциссу и ординату случайным образом поставленной точки $\{\omega_и\}=\{(x,y)\}$ на плоскости R^2. В искусственном эксперименте $E_и$ достоверный исход $\Omega_и=\{\omega_и=(x,y)\colon -\infty<x<+\infty,\ -\infty<y<+\infty\}=R^2$, а множество всех наблюдаемых исходов $\mathcal{F}_и=\mathcal{B}^2$ есть наименьшая σ-алгебра, порожденная в пространстве R^2 системой прямоугольников вида $\{(x,y)\colon a\leqslant x<b,\ c\leqslant y<d\}$. Иначе $\mathcal{B}^2$ является борелевской σ-алгеброй на плоскости. Как и в одномерном случае, используя формулу $\mathsf{P}(\{\omega\colon a\leqslant\xi_1(\omega)<b,\ c\leqslant\xi_2(\omega)<d\}) = F(b,d)-F(b,c)-F(a,d)+F(a,c)$, найдем вероятность $\mathsf{P}_и(A)$, где A есть прямоугольник $\{(x,y)\colon a\leqslant x<b,\ c\leqslant y<d\}$ или любое конечное объединение непересекающихся прямоугольников такого типа.

Следовательно, случайный механизм эксперимента $E_и$ должен выбирать точку на плоскости из произвольного прямоугольника $\{(x,y)\colon a\leqslant x<b,\ c\leqslant y<d\}$ с вероятностью, которая равна $F(b,d)-F(b,c)-F(a,d)+F(a,c)$. По теореме Каратеодори [1] однозначно продолжаем эту вероятность на борелевскую σ-алгебру $\mathcal{B}^2$.

В качестве количественных характеристик эксперимента $E_и$ примем случайные величины $\eta_1(\omega_и)=\eta_1((x,y))=x$ и $\eta_2(\omega_и)=\eta_2((x,y))=y$, т. е. имеем $\eta(\omega_и)=(\eta_1(\omega_и),\ \eta_2(\omega_и))=\omega_и=(x,y)$. Теперь для двумерной случайной величины $\eta(\omega_и)$ на $(\Omega_и,\mathcal{F}_и,\mathsf{P}_и(\cdot))=(R^2,\mathcal{B}^2,\mathsf{P}_и(\cdot))$ можем сразу написать $F_\eta(x,y)=\mathsf{P}_и(\{\omega_и\colon \eta_1(\omega_и)<x,\eta_1(\omega_и)<y\})=F(x,y)-F(-\infty,y)-F(x,-\infty)+F(-\infty,-\infty)=F(x,y)$ для любой упорядоченной пары $(x,y)\in R^2$. Так как $F_\eta(x,y)=F(x,y)$ для всех $(x,y)\in R^2$ и функция $F(x,y)$ однозначно определяет вероятность $\mathsf{P}_и(B^2)$ для

любого борелевского множества B^2 из $\mathcal{B}^2$, то естественно вероятность $\mathsf{P}_{\text{и}}(B^2)$ обозначить через $\mathsf{P}_\xi(B^2)$. Как и в одномерном случае, теперь остается объявить пространство $(R^2, \mathcal{B}^2, \mathsf{P}_\xi(\cdot))$ выборочным для всех таких двумерных случайных величин ξ, для которых их интегральные функции равны $F(x,y)$.

Замечание 7.2. Итак, для того чтобы $F(x,y)$ была функцией распределения, достаточно выполнения утверждений 7.1–7.4 и неравенства вида $F(b,d) - F(b,c) - F(a,d) + F(a,c) \geqslant 0$ при любых $a < b$, $c < d$. Рассмотрим обобщение этого результата на случай $n > 2$. Пусть конечно-разностный оператор $\Delta^i_{h_i}$, который действует на интегральную функцию распределения $F(x_1, x_2, \ldots, x_n)$ случайного вектора $(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$, определяется соотношением $\Delta^i_{h_i} F(a_1, a_2, \ldots, a_n) = F(a_1, a_2, \ldots, a_{i-1}, a_i{+}h_i, a_{i+1}, \ldots, a_n) - F(a_1, a_2, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n)$, $h_i > 0$, $a_i \in R$, $i = 1, 2, \ldots, n$. Используя метод индукции (*Гихман И.И.*, *Скороход А.В.*, *Ядренко М.И.* Теория вероятностей и математическая статистика. — Киев: Вища школа, 1979), доказывается равенство вида $\mathsf{P}(\{\omega\colon a_i \leqslant \xi_i(\omega) < b_i,\ i = 1, 2, \ldots, n\}) = \Delta^1_{h_1} \Delta^2_{h_2} \ldots \Delta^n_{h_n} F(a_1, a_2, \ldots, a_n)$, где $h_i = b_i - a_i > 0$ для всех $i = \overline{1,n}$. В частности, при $n = 2$, $a_1 = a$, $b_1 = b$, $a_2 = c$, $b_2 = d$ получаем: $h_1 = b - a$, $h_2 = d - c$ и, значит, $\mathsf{P}(\{\omega\colon a_1 \leqslant \xi_1(\omega) < b_1,\ a_2 \leqslant \xi_2(\omega) < b_2\}) = \Delta^1_{h_1} \Delta^2_{h_2} F(a_1, a_2) = \Delta^1_{h_1}(F(a_1, b_2) - F(a_1, a_2)) = \Delta^1_{h_1} F(a_1, b_2) - \Delta^1_{h_1} F(a_1, a_2)) = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) = F(b,d) - F(a,d) - F(b,c) + F(a,c) = \mathsf{P}(\{\omega\colon a \leqslant \xi_1(\omega) < b,\ c \leqslant \xi_2(\omega) < d\})$.

В общем случае, когда рассматривается n-мерная случайная величина $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$, необходимыми и достаточными условиями существования функции распределения $F(x_1, x_2, \ldots, x_n)$ являются выполнение утверждений 7.1–7.4 и соотношения $\Delta^1_{h_1} \Delta^2_{h_2} \ldots \Delta^n_{h_n} F(a_1, a_2, \ldots, a_n) \geqslant 0$ при $h_i = b_i - a_i$ и $a_i < b_i$ для всех $i = \overline{1,n}$.

§ 2. Зависимость и независимость количественных характеристик эксперимента

2.1. Независимость семейства случайных величин. При изучении семейства двух случайных величин (ξ_1, ξ_2) было показано, что вероятностные свойства этой совокупности, как правило, не исчерпываются вероятностными свойствами отдельных случайных величин ξ_1 и ξ_2. Такая ситуация объясняется тем, что между случайными величинами ξ_1, ξ_2 может существовать разная степень зависимости. Эта зависимость может оказаться более или менее жесткой. Для некоторых конкретных экспериментов, зная значение одной количественной характеристики, можно совершенно точно определить значение другой.

Иначе можно сказать, что мы имеем дело с привычной для математики функциональной зависимостью. Например, если мы наблюдаем за освещенностью в подъезде некоторого жилого дома, то число ξ_1 всех вышедших из строя однотипных лампочек в течение года и их общая стоимость ξ_2 связаны очень простой функциональной зависимостью вида $\xi_2 = z\,\xi_1$, где стоимость z электрической лампочки предполагается постоянной величиной.

Напомним, что утверждение леммы 6.4 содержит простой критерий вида $\sigma(\eta) \subset \sigma(\xi)$, который позволяет обнаружить факт существования функциональной зависимости $\eta(\omega) = g(\xi(\omega))$ между случайными величинами $\xi(\omega)$ и $\eta(\omega)$, где $g(x)$ — некоторая борелевская функция.

В общем случае, зная значение одной случайной величины ξ_1, мы можем указать только закон распределения другой случайной величины ξ_2. При этом указанный закон распределения зависит от того, какое значение приняла случайная величина ξ_1. Такая зависимость между случайными величинами называется статистической (вероятностной, стохастической). Например, высота ξ_1 найденного наудачу гриба в лесу и его вес ξ_2 зависимы, но эта связь не является жесткой или функциональной. Из опыта грибников хорошо известно следующее утверждение: чем выше гриб, тем вероятнее, что он больше весит. С примерами функциональной или статистической зависимости мы постоянно встречаемся в технике, медицине, экономике, экологии и т. д. В то же время можно указать примеры совершенно другой ситуации, когда количественные признаки эксперимента на интуитивном уровне считаются причинно несвязанными. Например, нет оснований сомневаться в том, что рост ξ_1 выбранного наудачу взрослого человека в Нижнем Новгороде практически не связан с имеющейся у него суммой ξ_2 денег в рублях. На содержательном уровне это можно объяснить тем, что множество факторов, которые определяют значение роста взрослого человека, и другое множество факторов, которые определяют значение его суммы денег, почти различны.

Итак, в теории вероятностей возникают важные вопросы создания эффективных критериев, которые позволяют установить статистическую зависимость или независимость случайных величин.

Математическая формализация независимости естественно должна быть такой, что любые количественные характеристики, не связанные причинно (по происхождению), должны объявляться независимыми случайными величинами. Поскольку случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ единственным образом порождают σ-алгебры событий из $\mathcal{F}$ вида $\sigma(\xi_1)$, $\sigma(\xi_2), \ldots, \sigma(\xi_n)$ соответственно, то интуитивное понятие независимости случайных событий и формальное определение независимости событий целесообразно перенести на случайные величины. Перейдем теперь к определению независимости семейства случайных величин.

Определение 7.3. Случайные величины ξ_1, $\xi_2, \ldots$, ξ_n называются *статистически независимыми*, если для любых борелевских множеств B_1, $B_2, \ldots$, B_n на действительной прямой имеет место следующее равенство:

$$\mathsf{P}(\{\omega\colon \xi_1(\omega) \in B_1, \xi_2(\omega) \in B_2, \ldots, \xi_n(\omega) \in B_n\}) = \mathsf{P}(\{\omega\colon \xi_1(\omega) \in B_1\}) \times \\ \times \mathsf{P}(\{\omega\colon \xi_2(\omega) \in B_2\}) \times \ldots \times \mathsf{P}(\{\omega\colon \xi_n(\omega) \in B_n\}). \quad (7.4)$$

В противном случае случайные величины называются *зависимыми*.

Если для любого конечного $n \geqslant 2$ выполняется соотношение (7.4), то заданные на основном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ случайные величины последовательности $\{\xi_i(\omega);\ i = 1, 2, \ldots\}$ называются *статистически независимыми*.

Равенство (7.4) при $n = 2$ и $B_1 \in \mathcal{B}$, $B_2 \in \mathcal{B}$ можно переписать в эквивалентном виде: $\mathsf{P}(\{\omega\colon \xi_1(\omega) \in B_1\} \cap \{\omega\colon \xi_2(\omega) \in B_2\}) = \mathsf{P}(\{\omega\colon \xi_1(\omega) \in B_1\}) \times \mathsf{P}(\{\omega\colon \xi_2(\omega) \in B_2\})$. Так как для любых $B_1 \in \mathcal{B}$ и $B_2 \in \mathcal{B}$ множества $A_1 = \{\omega\colon \xi_1(\omega) \in B_1\}$ и $A_2 = \{\omega\colon \xi_2(\omega) \in B_2\}$ являются событиями из σ-алгебр $\sigma(\xi_1)$ и $\sigma(\xi_2)$, то для доказательства независимости случайных величин ξ_1 и ξ_2 достаточно показать выполнение равенства

$$\mathsf{P}(A_1 \cap A_2) = \mathsf{P}(A_1) \times \mathsf{P}(A_2), \qquad A_1 \in \sigma(\xi_1), \quad A_2 \in \sigma(\xi_2). \qquad (7.5)$$

Здесь важно заметить, что математическое определение 7.3 статистической независимости случайных величин является общим утверждением по сравнению с понятием их причинной независимости или принадлежности к причинно не связанным реальным явлениям. Другими словами, из причинной независимости случайных величин должна следовать их статистическая независимость, а наоборот не всегда. Этот факт можно пояснить следующими двумя простыми примерами.

Пример 7.5. Эксперимент E заключается в том, что два студента в тире независимым образом стреляют каждый по своей мишени. Стрелок с равной возможностью попадает в мишень и получает 10 рублей от организаторов аттракциона или промахивается и платит за стоимость выстрела 10 рублей. Если стрелок промахнулся, то мы применяем метку в виде цифры 0, в противном случае используем обозначение в виде цифры 1. Например, промах первого студента и попадание второго описываем вектором $(0, 1)$. При таком выборе описаний элементарных исходов данного симметричного опыта с конечным числом исходов имеем: достоверное событие $\Omega = \{\omega_1 = (0, 0), \omega_2 = (0, 1), \omega_3 = (1, 0), \omega_4 = (1, 1)\}$, вероятность $\mathsf{P}(\{(0, 0)\}) = \mathsf{P}(\{(0, 1)\}) = \mathsf{P}(\{(1, 0)\}) = \mathsf{P}(\{(1, 1)\}) = 1/4$. Пусть теперь ξ_1 и ξ_2 определяют выигрыш первого и второго студентов соответственно. Отсюда $\xi_1(\omega_1) = \xi_1(\omega_2) = -10$, $\xi_1(\omega_3) = \xi_1(\omega_4) = 10$, $\xi_2(\omega_1) = \xi_2(\omega_3) = -10$, $\xi_2(\omega_2) = \xi_2(\omega_4) = 10$ и, значит, $\sigma(\xi_1) = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega, \varnothing\}$ и $\sigma(\xi_2) = \{\{\omega_1, \omega_3\}, \{\omega_2, \omega_4\}, \Omega, \varnothing\}$. Для этого эксперимента нет никаких оснований считать наличие причинной связи между выигрышем ξ_1 первого студента и выигрышем ξ_2 второго. В самом деле, если предположить, что между случайными величинами $\xi_1(\omega)$ и $\xi_2(\omega)$ существует функциональная зависимость $\xi_2(\omega) = g(\xi_1(\omega))$, тогда при $\omega = \omega_3, \omega_4$ вычисляем $\xi_1(\omega_3) = \xi_1(\omega_4) = 10$, $\xi_2(\omega_3) = -10 = g(\xi_1(\omega_3)) = g(10)$, $\xi_2(\omega_4) = 10 = g(\xi_1(\omega_4)) = g(10)$. Отсюда получаем, что функция $g(x)$ принимает различные значения при $x = 10$, т. е. является неоднозначной.

Легко проверить выполнение равенств вида (7.5) для $\sigma(\xi_1) = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega, \varnothing\}$ и $\sigma(\xi_2) = \{\{\omega_1, \omega_3\}, \{\omega_2, \omega_4\}, \Omega, \varnothing\}$. Например, при $A_1 = \{\omega_1, \omega_2\} \in \sigma(\xi_1)$, $A_2 = \{\omega_1, \omega_3\} \in \sigma(\xi_2)$ имеем, что $\mathsf{P}(\{\omega_1, \omega_2\} \cap \{\omega_1, \omega_3\}) = \mathsf{P}(\{\omega_1\}) = 1/4 = (1/2) \times (1/2) = \mathsf{P}(\{\omega_1, \omega_2\}) \times \mathsf{P}(\{\omega_1, \omega_3\})$. Итак, причинно независимые случайные величины ξ_1 и ξ_2 являются статистически независимыми.

Пример 7.6. Грани Δsab, Δsac, Δscb, Δabc правильного тетраэдра $sabc$ окрашены соответственно в белый, синий, красный и, наконец, во все эти три цвета. Тетраэдр, изготовленный из однородного материала, подбрасывается случайным образом на поверхность стола с помощью некоторого механизма. Фиксируется цвет грани, на которую падает тетраэдр, соприкасаясь с поверхностью стола (см. пример 3.20 Бернштейна). Напомним, что выпадение грани Δsab, Δsac, Δscb или Δabc обозначают через символы ω_1, ω_2, ω_3 соответственно, или через ω_4. Поэтому для этого эксперимента достоверное событие $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ и $\mathsf{P}(\{\omega_i\}) = 1/4$, $i = 1, 2, 3, 4$. Здесь $\mathsf{P}(\{\omega_i\})$ есть вероятность того, что тетраэдр падает на грань с меткой или описанием ω_i.

Рассмотрим теперь связанную с бросанием раскрашенного тетраэдра игру отца, старшего и младшего сына. Если тетраэдр падает на грань, которая содержит белый цвет, то младший сын выигрывает у отца одно очко, иначе он получает ноль очков. Если тетраэдр падает на грань, которая имеет синий цвет, то старший сын получает от отца одно очко, в противном случае он проигрывает отцу одно очко. Пусть случайные величины ξ, η и ζ определяют выигрыши младшего сына, старшего и отца соответственно. Из определения этих случайных величин легко находим: $\xi(\omega_1) = \xi(\omega_4) = 1$, $\xi(\omega_2) = \xi(\omega_3) = 0$, $\eta(\omega_2) = \eta(\omega_4) = 1$, $\eta(\omega_1) = \eta(\omega_3) = -1$, $\zeta(\omega_1) = 0$, $\zeta(\omega_2) = -1$, $\zeta(\omega_3) = 1$, $\zeta(\omega_4) = -2$. Нетрудно установить, что между случайными величинами ξ, η и ζ существует простая функциональная связь $\zeta = -(\xi + \eta)$. Так как $\sigma(\xi) = \{\{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}, \Omega, \varnothing\}$ и $\sigma(\eta) = \{\{\omega_2, \omega_4\}, \{\omega_1, \omega_3\}, \Omega, \varnothing\}$, то легко проверяется равенство (7.5) для любых $A_1 \in \sigma(\xi)$ и $A_2 \in \sigma(\eta)$.

Итак, ξ и η являются статистически независимыми, хотя между ними существует функциональная зависимость.

2.2. О критериях независимости случайных величин. Рассмотрим теперь несколько простых и доступных для проверки критериев независимости случайных величин. Первый такой критерий формулируется в виде следующей теоремы.

Теорема 7.1. *Случайные величины ξ_1, $\xi_2, \ldots$, ξ_n статистически независимы тогда и только тогда, когда имеет место равенство*

$$F(x_1, x_2, \ldots, x_n) = F_{\xi_1}(x_1) \times F_{\xi_2}(x_2) \times \ldots \times F_{\xi_n}(x_n) \tag{7.6}$$

для каждой точки $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in R^n$.

Доказательство. Доказательство необходимости следует из равенства (7.4) при $B_i = (-\infty, x_i)$, $i = 1, 2, \ldots, n$. Для доказательства достаточности, ради простоты, положим $n = 2$. В силу равенств (7.3) и (7.6) вероятность $\mathsf{P}(\{\omega\colon \xi_1(\omega) \in [\,a, b), \xi_2(\omega) \in [\,c, d)\})$ равна величине $F(b, d) - F(a, d) - F(b, c) + F(a, c) = F_{\xi_1}(b) \times F_{\xi_2}(d) - F_{\xi_1}(a) \times \times F_{\xi_2}(d) - F_{\xi_1}(b) \times F_{\xi_2}(c) + F_{\xi_1}(a) \times F_{\xi_2}(c) = (F_{\xi_1}(b) - F_{\xi_1}(a)) \times \times (F_{\xi_2}(d) - F_{\xi_2}(c)) = \mathsf{P}(\{\omega\colon \xi_1(\omega) \in [\,a, b)\}) \times \mathsf{P}(\{\omega\colon \xi_2(\omega) \in [\,c, d)\})$. Зафиксируем промежуток $[\,c, d)$ и выберем любое семейство $\{[\,a_j, b_j)\colon j \in \{1, 2, \ldots, m\}\}$ из $m \in \{1, 2, \ldots\}$ непересекающихся промежутков на прямой R. Тогда из последнего равенства получим

$$\begin{aligned}
&\mathsf{P}\Bigg(\Bigg\{\omega\colon \xi_1(\omega) \in \bigcup_{j=1}^{m} [a_j, b_j),\ \xi_2(\omega) \in [c, d)\Bigg\}\Bigg) = \\
&\qquad = \mathsf{P}\Bigg(\Bigg\{\omega\colon \xi_1(\omega) \in \bigcup_{j=1}^{m} [a_j, b_j)\Bigg\} \bigcap \big\{\omega\colon \xi_2(\omega) \in [c, d)\big\}\Bigg) = \\
&\qquad = \mathsf{P}\Bigg(\Bigg(\bigcup_{j=1}^{m} \{\omega\colon \xi_1(\omega) \in [a_j, b_j)\}\Bigg) \bigcap \big\{\omega\colon \xi_2(\omega) \in [c, d)\big\}\Bigg) = \\
&\qquad = \mathsf{P}\Bigg(\bigcup_{j=1}^{m} \Big(\{\omega\colon \xi_1(\omega) \in [a_j, b_j)\} \bigcap \{\omega\colon \xi_2(\omega) \in [c, d)\}\Big)\Bigg) = \\
&\qquad = \sum_{j=1}^{m} \mathsf{P}\big(\{\omega\colon \xi_1(\omega) \in [a_j, b_j),\ \xi_2(\omega) \in [c, d)\}\big) = \\
&\qquad = \sum_{j=1}^{m} \mathsf{P}\big(\{\omega\colon \xi_1(\omega) \in [a_j, b_j)\}\big) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c, d)\}\big) = \\
&\qquad = \mathsf{P}\Bigg(\bigcup_{j=1}^{m} \{\omega\colon \xi_1(\omega) \in [a_j, b_j)\}\Bigg) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c, d)\}\big) = \\
&\qquad = \mathsf{P}\Bigg(\Bigg\{\omega\colon \xi_1(\omega) \in \bigcup_{j=1}^{m} [a_j, b_j)\Bigg\}\Bigg) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c, d)\}\big).
\end{aligned} \tag{7.7}$$

Пусть теперь $\mathcal{B}_1$ есть семейство таких борелевских множеств B на действительной прямой, для которых имеет место равенство

$$\mathsf{P}\big(\{\omega\colon \xi_1(\omega) \in B,\, \xi_2(\omega) \in [c,d)\}\big) = \\ = \mathsf{P}\big(\{\omega\colon \xi_1(\omega) \in B\}\big) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c,d)\}\big). \quad (7.8)$$

Следовательно, $\mathcal{B}_1 = \{B\colon B \in \mathcal{B},\, \mathsf{P}(\{\omega\colon \xi_1(\omega) \in B, c \leqslant \xi_2(\omega) < d\}) = \mathsf{P}(\{\omega\colon \xi_1(\omega) \in B\}) \times \mathsf{P}(\{\omega\colon c \leqslant \xi_2(\omega) < d\})\}$ и $\mathcal{B}_1 \subset \mathcal{B}$. Если $\mathcal{L}_0$ — алгебра подмножеств действительной прямой, составленных из конечных объединений непересекающихся промежутков вида $[a,b)$, и множество $B \in \mathcal{L}_0$, то в силу равенства (7.7) и определения $\mathcal{B}_1$ получаем $\mathcal{L}_0 \subset \mathcal{B}_1 \subset \mathcal{B}$.

Покажем, что семейство $\mathcal{B}_1$ замкнуто относительно предельного перехода монотонных последовательностей множеств из $\mathcal{B}_1$. Действительно, если теперь $B_k \in \mathcal{B}_1$ и $B_k \subset B_{k+1}$ для всех $k = 1, 2, \ldots$, то с учетом соотношений $\{\omega\colon \xi_1(\omega) \in B_k\} \subset \{\omega\colon \xi_1(\omega) \in B_{k+1}\}$, $\{\omega\colon \xi_1(\omega) \in B_k,$ $\xi_2(\omega) \in [c,d)\} \subset \{\omega\colon \xi_1(\omega) \in B_{k+1},\, \xi_2(\omega) \in [c,d)\}$ при $k = 1, 2, \ldots$ и известных равенств

$$\lim_{k\to\infty} B_k = \bigcup_{k=1}^{\infty} B_k, \quad \lim_{k\to\infty} \{\omega\colon \xi_1(\omega) \in B_k\} = \bigcup_{k=1}^{\infty} \{\omega\colon \xi_1(\omega) \in B_k\},$$

$$\lim_{k\to\infty} \{\omega\colon \xi_1(\omega) \in B_k,\, \xi_2(\omega) \in [c,d)\} = \bigcup_{k=1}^{\infty} \{\omega\colon \xi_1(\omega) \in B_k,\, \xi_2(\omega) \in [c,d)\}$$

последовательно находим следующую цепочку равенств:

$$\mathsf{P}\big(\{\omega\colon \xi_1(\omega) \in B_k, \xi_2(\omega) \in [c,d)\}\big) = \\ = \mathsf{P}\big(\{\omega\colon \xi_1(\omega) \in B_k\}\big) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c,d)\}\big),$$

$$\lim_{k\to\infty} \mathsf{P}\big(\{\omega\colon \xi_1(\omega) \in B_k, \xi_2(\omega) \in [c,d)\}\big) = \\ = \lim_{k\to\infty} \mathsf{P}\big(\{\omega\colon \xi_1(\omega) \in B_k\}\big) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c,d)\}\big),$$

$$\mathsf{P}\Big(\lim_{k\to\infty} \{\omega\colon \xi_1(\omega) \in B_k, \xi_2(\omega) \in [c,d)\}\Big) = \\ = \mathsf{P}\Big(\lim_{k\to\infty} \{\omega\colon \xi_1(\omega) \in B_k\}\Big) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c,d)\}\big),$$

$$\mathsf{P}\Big(\bigcup_{k=1}^{\infty} \{\omega\colon \xi_1(\omega) \in B_k, \xi_2(\omega) \in [c,d)\}\Big) = \\ = \mathsf{P}\Big(\bigcup_{k=1}^{\infty} \{\omega\colon \xi_1(\omega) \in B_k\}\Big) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c,d)\}\big),$$

$$\mathsf{P}\Big(\bigcup_{k=1}^{\infty} (\{\omega\colon \xi_1(\omega) \in B_k\} \cap \{\omega\colon \xi_2(\omega) \in [c,d)\})\Big) = \\ = \mathsf{P}\Big(\Big\{\omega\colon \xi_1(\omega) \in \bigcup_{k=1}^{\infty} B_k\Big\}\Big) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c,d)\}\big),$$

$$\mathsf{P}\Big(\Big\{\omega\colon \xi_1(\omega) \in \bigcup_{k=1}^{\infty} B_k\Big\} \cap \{\omega\colon \xi_2(\omega) \in [c,d)\}\Big) = \\ = \mathsf{P}\Big(\Big\{\omega\colon \xi_1(\omega) \in \bigcup_{k=1}^{\infty} B_k\Big\}\Big) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c,d)\}\big),$$

$$\mathsf{P}\Big(\Big\{\omega\colon \xi_1(\omega) \in \bigcup_{k=1}^{\infty} B_k, \xi_2(\omega) \in [c,d)\Big\}\Big) = \\ = \mathsf{P}\Big(\Big\{\omega\colon \xi_1(\omega) \in \bigcup_{k=1}^{\infty} B_k\Big\}\Big) \times \mathsf{P}\big(\{\omega\colon \xi_2(\omega) \in [c,d)\}\big).$$

Из последнего равенства следует, что множество $\cup_{k=1}^{\infty} B_k$ удовлетворяет (7.8), и, значит, $\cup_{k=1}^{\infty} B_k \in \mathcal{B}_1$ в силу определения $\mathcal{B}_1$. Отсюда $\lim\limits_{k\to\infty} B_k \in \mathcal{B}_1$, так как множество $\cup_{k=1}^{\infty} B_k$ равно $\lim\limits_{k\to\infty} B_k$. Итак, семейство или класс $\mathcal{B}_1$ является замкнутым относительно предельного перехода любой возрастающей последовательности множеств из $\mathcal{B}_1$. Аналогичным способом можно показать, что для любой убывающей последовательности $B_1 \supset B_2 \supset \ldots$ множеств из $\mathcal{B}_1$ будет верно соотношение $\lim\limits_{k\to\infty} B_k \in \mathcal{B}_1$.

В дальнейшем по определению любое семейство множеств будем называть монотонным, если предельный переход монотонных последовательностей его множеств не выводит из этого семейства. Например, семейство $\mathcal{B}_1$ подмножеств действительной прямой является монотонным классом. Обозначим через символ $M(\mathcal{L}_0)$ наименьший монотонный класс подмножеств из R, который содержит все множества из алгебры $\mathcal{L}_0$ и $M(\mathcal{L}_0) \subset \mathcal{B}_2$, если только $\mathcal{B}_2$ — произвольный монотонный класс подмножеств действительной прямой и $\mathcal{B}_2 \supset \mathcal{L}_0$. Способ доказательства существования наименьшего монотонного класса $M(\mathcal{L}_0)$ подмножеств из R полностью совпадает с методом доказательства теоремы 6.1 о существовании борелевской σ-алгебры $\mathcal{B}$ на прямой R. Так как $\mathcal{B}_1$ является монотонным классом и $\mathcal{L}_0 \subset \mathcal{B}_1 \subset \mathcal{B}$, то $\mathcal{L}_0 \subset M(\mathcal{L}_0) \subset \mathcal{B}_1 \subset \mathcal{B}$.

Установим теперь, что $M(\mathcal{L}_0)$ является алгеброй. Пусть $C \in M(\mathcal{L}_0)$, покажем, что $\overline{C} \in M(\mathcal{L}_0)$. Для этого введем вспомогательное множество $M_1 = \{B\colon B \in M(\mathcal{L}_0), \overline{B} \in M(\mathcal{L}_0)\}$. Ясно, что $M_1 \subset M(\mathcal{L}_0)$. Так как множество $\mathcal{L}_0$ является алгеброй, то при $B \in \mathcal{L}_0$ имеем $\overline{B} \in \mathcal{L}_0$. Поэтому из $B \in \mathcal{L}_0$, $\overline{B} \in \mathcal{L}_0$, $\mathcal{L}_0 \subset M(\mathcal{L}_0)$ следует, что $B \in M(\mathcal{L}_0)$ и $\overline{B} \in M(\mathcal{L}_0)$, т.е. $B \in M_1$. Окончательно получаем, что $\mathcal{L}_0 \subset M_1 \subset M(\mathcal{L}_0)$. Докажем, что M_1 является монотонным классом. Для любой возрастающей последовательности $B_1 \subset B_2 \subset \ldots$ множеств из M_1 будут справедливы следующие соотношения: $B_k \in M(\mathcal{L}_0)$, $\overline{B}_k \in M(\mathcal{L}_0)$, $k = 1, 2, \ldots$, $\lim\limits_{k\to\infty} B_k = \cup_{k=1}^{\infty} B_k \in M(\mathcal{L}_0)$, $\overline{B}_1 \supset \overline{B}_2 \supset \ldots$, $\lim\limits_{k\to\infty} \overline{B}_k = \cap_{k=1}^{\infty} \overline{B}_k \in M(\mathcal{L}_0)$, $\overline{\lim\limits_{k\to\infty} B_k} = \overline{\cup_{k=1}^{\infty} B_k} = \cap_{k=1}^{\infty} \overline{B}_k \in M(\mathcal{L}_0)$.

Отсюда непосредственно находим, что $\underline{\lim\limits_{k\to\infty}} B_k \in M(\mathcal{L}_0)$ и $\overline{\lim\limits_{k\to\infty} B_k} \in \in M(\mathcal{L}_0)$, т. е. $\lim\limits_{k\to\infty} B_k \in M_1$.

Аналогично доказывается, что $\lim\limits_{k\to\infty} B_k \in M_1$ для любой убывающей последовательности $B_1 \supset B_2 \supset \ldots$ множеств из M_1. Значит, M_1 является монотонным классом. Так как $M_1 \supset \mathcal{L}_0$ и $M(\mathcal{L}_0)$ есть наименьший монотонный класс, то $M_1 \supset M(\mathcal{L}_0)$. Отсюда, учитывая соотношение $M_1 \subset M(\mathcal{L}_0)$, получим $M_1 = M(\mathcal{L}_0)$. Если теперь $C \in M(\mathcal{L}_0)$, то $C \in M_1$, а из определения множества M_1 следует, что $\overline{C} \in M(\mathcal{L}_0)$. Значит, наименьший монотонный класс $M(\mathcal{L}_0)$ замкнут относительно взятия операции дополнения. В следующих абзацах установим, что класс $M(\mathcal{L}_0)$ замкнут относительно взятия операции объединения.

Рассмотрим множество $M_2(C) = \{B\colon B \in M(\mathcal{L}_0),\, C \cup B \in M(\mathcal{L}_0)\}$ для любого $C \in M(\mathcal{L}_0)$. Ясно, что $M_2(C) \subset M(\mathcal{L}_0)$. Для возрастающей последовательности $B_1 \subset B_2 \subset \ldots$ множеств из $M_2(C)$ имеем следующую цепочку очевидных соотношений: $C \in M(\mathcal{L}_0)$, $B_k \in M(\mathcal{L}_0)$, $C \cup B_k \in M(\mathcal{L}_0)$, $k \geqslant 1$;

$$\lim_{k\to\infty} B_k = \bigcup_{k=1}^{\infty} B_k \in M(\mathcal{L}_0), \quad (C \cup B_1) \subset (C \cup B_2) \subset \ldots,$$

$$\lim_{k\to\infty} (C \cup B_k) = \bigcup_{k=1}^{\infty} (C \cup B_k) = C \cup \left(\bigcup_{k=1}^{\infty} B_k\right) \in M(\mathcal{L}_0).$$

Отсюда $\lim\limits_{k\to\infty} B_k \in M_2(C)$, так как $\lim\limits_{k\to\infty} B_k \in M(\mathcal{L}_0)$ и $C \cup (\lim\limits_{k\to\infty} B_k) \in \in M(\mathcal{L}_0)$. Аналогично устанавливается, что $\lim\limits_{k\to\infty} B_k \in M_2(C)$ для любой убывающей последовательности $B_1 \supset B_2 \supset \ldots$ множеств из $M_2(C)$. Следовательно, $M_2(C)$ является монотонным классом для любого $C \in M(\mathcal{L}_0)$.

Легко проверить, что при $C \in \mathcal{L}_0$ имеем $\mathcal{L}_0 \subset M_2(C) \subset M(\mathcal{L}_0)$. В самом деле, если теперь множество C является элементом алгебры $\mathcal{L}_0$, то для любого множества $D \in \mathcal{L}_0$ имеем $C \cup D \in \mathcal{L}_0 \subset M(\mathcal{L}_0)$, т. е. $\mathcal{L}_0 \subset M_2(C) \subset M(\mathcal{L}_0)$. Поэтому $M_2(C) = M(\mathcal{L}_0)$ для каждого $C \in \mathcal{L}_0$, так как $M_2(C)$ — монотонный класс, а $M(\mathcal{L}_0)$ — наименьший монотонный класс. Пусть теперь D — произвольное множество из $M(\mathcal{L}_0)$ и C — любое множество из алгебры $\mathcal{L}_0$. Из равенства $M_2(C) = M(\mathcal{L}_0)$, где $C \in \mathcal{L}_0$, вытекает, что $D \in M_2(C)$. Отсюда $C \in M(\mathcal{L}_0)$, $D \in M(\mathcal{L}_0)$, $C \cup D \in M(\mathcal{L}_0)$ или $D \in M(\mathcal{L}_0)$, $C \in M(\mathcal{L}_0)$, $D \cup C \in M(\mathcal{L}_0)$ и, значит, $C \in M_2(D)$ для любого множества $C \in \mathcal{L}_0$ и каждого $D \in M(\mathcal{L}_0)$. Поэтому $\mathcal{L}_0 \subset M_2(D) \subset M(\mathcal{L}_0)$. Учитывая снова, что $M(\mathcal{L}_0)$ — наименьший монотонный класс, а $M_2(D)$ — монотонный класс, получим $M_2(D) = M(\mathcal{L}_0)$ для каждого $D \in M(\mathcal{L}_0)$. Итак, если $B \in M(\mathcal{L}_0)$ и $D \in M(\mathcal{L}_0)$, то $B \in M_2(D) = M(\mathcal{L}_0)$ или $D \cup B \in M(\mathcal{L}_0)$. Теперь можно утверждать, что класс $M(\mathcal{L}_0)$ замкнут относительно взятия операций дополнения и объединения. Другими словами, семейство $M(\mathcal{L}_0)$ является алгеброй. Более того, покажем, что $M(\mathcal{L}_0)$ является σ-алгеброй.

Действительно, если $B_k \in M(\mathcal{L}_0)$ для всех $k = 1, 2, \ldots$, то $C_m = \cup_{j=1}^{m} B_j \in M(\mathcal{L}_0)$ и $C_m \subset C_{m+1}$ при $m = 1, 2, \ldots$. Так как $M(\mathcal{L}_0)$ — монотонный класс, то $\lim\limits_{m\to\infty} C_m = \cup_{m=1}^{\infty} C_m = \cup_{m=1}^{\infty} B_m \in M(\mathcal{L}_0)$. Таким образом, $M(\mathcal{L}_0)$ является σ-алгеброй и $\mathcal{L}_0 \subset M(\mathcal{L}_0) \subset \mathcal{B}_1 \subset \mathcal{B}$, а $\mathcal{B}$ суть наименьшая σ-алгебра. Поэтому $M(\mathcal{L}_0) = \mathcal{B} = \mathcal{B}_1$ и, значит, имеет место равенство (7.8) для любого $B_1 \in \mathcal{B}$. Далее, зафиксируем борелевское множество B_1 и аналогичным способом устанавливаем, что верно соотношение (7.8), в котором только промежуток $[c, d)$ уже будет заменен на произвольное множество B_2 из $\mathcal{B}$. Затем этот метод применяем последовательно для остальных величин ξ_3, $\xi_4, \ldots$, ξ_n и доказываем равенство (7.4). Утверждение теоремы 7.1 доказано.

Теперь подчеркнем, что подробное рассмотрение теоремы 7.1 было вызвано вовсе не демонстрацией сложности, красоты и тонкостей этапов ее доказательства, а исключительно важной ролью этой теоремы в различных приложениях. Во-первых, статистическую независимость случайных величин можно определять более просто, а именно, в терминах многомерных интегральных функций распределения с помощью соотношения (7.6). Во-вторых, теорема 7.1 показывает, что в случае независимости случайных величин с помощью равенства (7.6) всегда можно восстановить интегральный закон распределения многомерной случайной величины по интегральным законам распределения одномерных случайных величин, входящих в упорядоченную систему.

Далее, восстановленная многомерная интегральная функция распределения $F(x_1, x_2, \ldots, x_n)$, $(x_1, x_2, \ldots, x_n) \in R^n$, однозначно определяет выборочное вероятностное пространство $(R^n, \mathcal{B}^n, \mathsf{P}_\xi(\cdot))$, где $(\xi_1, \xi_2, \ldots, \xi_n) = \xi$. При этом мультипликативная структура (7.6) интегральной функции распределения $F(x_1, x_2, \ldots, x_n)$, $(x_1, x_2, \ldots, x_n) \in R^n$, порождает мультипликативную структуру (7.4) для вероятностей $\mathsf{P}_\xi(B_1 \times B_2 \times \ldots \times B_n) = \mathsf{P}(\{\omega\colon \xi_1(\omega) \in B_1, \xi_2(\omega) \in B_2, \ldots, \xi_n(\omega) \in B_n\})$, $B_1 \in \mathcal{B}$, $B_2 \in \mathcal{B}, \ldots$, $B_n \in \mathcal{B}$. Наконец, для каждой случайной величины ξ_i с интегральной функцией распределения $F_{\xi_i}(x_i)$ можно построить выборочное вероятностное пространство $(R, \mathcal{B}, \mathsf{P}_{\xi_i}(\cdot))$, где $i = 1, 2, \ldots, n$. В силу определения вероятностей $\mathsf{P}_\xi(B_1 \times B_2 \times \ldots \times B_n)$, $\mathsf{P}_{\xi_i}(B_i)$, $i = 1, 2, \ldots, n$, и соотношения (7.4) находим

$$\begin{aligned}\mathsf{P}_\xi(B_1 \times B_2 \times \ldots \times B_n) &= \mathsf{P}(\{\omega\colon \xi_1(\omega) \in B_1,\, \xi_2(\omega) \in B_2, \ldots, \xi_n(\omega) \in B_n\}) = \\ &= \mathsf{P}(\{\omega\colon \xi_1(\omega) \in B_1\}) \times \mathsf{P}(\{\omega\colon \xi_2(\omega) \in B_2\}) \times \ldots \times \mathsf{P}(\{\omega\colon \xi_n(\omega) \in B_n\}) = \\ &= \mathsf{P}_{\xi_1}(B_1) \times \mathsf{P}_{\xi_2}(B_2) \times \ldots \times \mathsf{P}_{\xi_n}(B_n).\end{aligned}$$

В этом случае говорят, что выборочное вероятностное пространство $(R^n, \mathcal{B}^n, \mathsf{P}_\xi(\cdot))$ для случайного вектора $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ является прямым произведением выборочных вероятностных пространств $(R, \mathcal{B}, \mathsf{P}_{\xi_1}(\cdot))$, $(R, \mathcal{B}, \mathsf{P}_{\xi_2}(\cdot)), \ldots$, $(R, \mathcal{B}, \mathsf{P}_{\xi_n}(\cdot))$. Здесь уместно напомнить читателю, что с аналогичной ситуацией мы уже встречались в четвертой главе при построении вероятностной модели общей схемы независимых испытаний. Теперь можно сказать, что выборочные вероятностные пространства вида $(R, \mathcal{B}, \mathsf{P}_{\xi_1}(\cdot))$, $(R, \mathcal{B}, \mathsf{P}_{\xi_2}(\cdot)), \ldots$, $(R, \mathcal{B}, \mathsf{P}_{\xi_n}(\cdot))$ являются моделями схемы из n независимых опытов. При этом i-й опыт заключается в бросании точки на i-ю действительную прямую с помощью некоторого случайного механизма, определяемого интегральной функцией распределения $F_{\xi_i}(x_i)$. В перечисленных фактах и состоит исключительное значение теоремы 7.1.

Теорема 7.2. *Если случайные величины ξ_1, $\xi_2, \ldots$, ξ_n статистически независимы и функции $g_1(x)$, $g_2(x), \ldots$, $g_n(x)$ являются борелевскими, то независимы и случайные величины $\eta_1(\omega) = g_1(\xi_1(\omega))$, $\eta_2(\omega) = g_2(\xi_2(\omega)), \ldots$, $\eta_n(\omega) = g_n(\xi_n(\omega))$.*

Доказательство. По лемме 6.3 функции $\eta_1(\omega)$, $\eta_2(\omega), \ldots$, $\eta_n(\omega)$ будут случайными величинами и для любых борелевских множеств B_1, $B_2 \ldots$, B_n выполняются соотношения $\{\omega\colon \eta_1(\omega) \in B_1\} = \{\omega\colon g_1(\xi_1(\omega)) \in B_1\} = \{\omega\colon \xi_1(\omega) \in g_1^{-1}(B_1)\} \in \mathcal{F}$, $\{\omega\colon \eta_2(\omega) \in B_2\} = \{\omega\colon g_2(\xi_2(\omega)) \in B_2\} = \{\omega\colon \xi_2(\omega) \in g_2^{-1}(B_2)\} \in \mathcal{F}$, $\ldots$, $\{\omega\colon \eta_n(\omega) \in B_n\} = \{\omega\colon g_n(\xi_n(\omega)) \in B_n\} = \{\omega\colon \xi_n(\omega) \in g_n^{-1}(B_n)\} \in \mathcal{F}$. Отсюда последовательно находим:

$$\begin{aligned}
&\mathsf{P}(\{\omega\colon \eta_1(\omega) \in B_1, \eta_2(\omega) \in B_2, \ldots, \eta_n(\omega) \in B_n\}) = \\
&\quad = \mathsf{P}(\{\omega\colon \xi_1(\omega) \in g_1^{-1}(B_1), \xi_2(\omega) \in g_2^{-1}(B_2), \ldots, \xi_n(\omega) \in g_n^{-1}(B_n)\}) = \\
&\qquad = \mathsf{P}(\{\omega\colon \xi_1(\omega) \in g_1^{-1}(B_1)\}) \times \mathsf{P}(\{\omega\colon \xi_2(\omega) \in g_2^{-1}(B_2)\}) \times \ldots \times \\
&\qquad\quad \times \mathsf{P}(\{\omega\colon \xi_n(\omega) \in g_n^{-1}(B_n)\}) = \mathsf{P}(\{\omega\colon \eta_1(\omega) \in B_1\}) \times \\
&\qquad\qquad \times \mathsf{P}(\{\omega\colon \eta_2(\omega) \in B_2\}) \times \ldots \times \mathsf{P}(\{\omega\colon \eta_n(\omega) \in B_n\}),
\end{aligned}$$

так как ξ_1, $\xi_2, \ldots$, ξ_n суть независимые случайные величины и множества $g_1^{-1}(B_1)$, $g_2^{-1}(B_2), \ldots$, $g_n^{-1}(B_n)$ являются борелевскими. Теорема 7.2 доказана.

Теорема 7.3. *Если случайные величины ξ_1, $\xi_2, \ldots$, ξ_n статистически независимы, то будут статистически независимы случайные величины, совокупность которых является любым набором из величин ξ_1, $\xi_2, \ldots$, ξ_n.*

Доказательство. Ради определенности и простоты записи покажем утверждение этой теоремы, например, для случайных величин ξ_1 и ξ_2. В самом деле, для любых борелевских множеств B_1, B_2 из определения 7.3 и очевидного свойства многомерных распределений получаем

$$\mathsf{P}(\{\omega\colon \xi_1(\omega)\in B_1,\, \xi_2(\omega)\in B_2, \xi_3(\omega)\in \mathrm{R},\ldots,\xi_n(\omega)\in \mathrm{R}\}) = \\ = \mathsf{P}(\{\omega\colon \xi_1(\omega)\in B_1\})\times \mathsf{P}(\{\omega\colon \xi_2(\omega)\in B_2\})\times \mathsf{P}(\{\omega\colon \xi_3(\omega)\in \mathrm{R}\})\times \\ \times \mathsf{P}(\{\omega\colon \xi_4(\omega)\in \mathrm{R}\})\times\ldots\times \mathsf{P}(\{\omega\colon \xi_n(\omega)\in \mathrm{R}\}),$$

$$\mathsf{P}(\{\omega\colon \xi_1(\omega)\in B_1,\, \xi_2(\omega)\in B_2\}) = \\ = \mathsf{P}(\{\omega\colon \xi_1(\omega)\in B_1\})\times \mathsf{P}(\{\omega\colon \xi_2(\omega)\in B_2\}).$$

Из последнего равенства следует, что случайные величины ξ_1 и ξ_2 статистически независимы, так как для любых борелевских множеств B_1, B_2 имеет место независимость событий $\{\omega\colon \xi_1\in B_1\}$ и $\{\omega\colon \xi_2\in B_2\}$. Независимость событий $\{\omega\colon \xi_1\in B_1\}$ и $\{\omega\colon \xi_2\in B_2\}$ приводит к следующему замечательному утверждению. Если наблюдать некоторый эксперимент E и построить для него вероятностное пространство $(\Omega,\mathcal{F},\mathsf{P}(\cdot))$, затем изучить определенные на нем одномерные случайные величины ξ_1 и ξ_2, которые независимы, то при составлении вектора из двух независимых величин, вероятностные свойства этого вектора легко получить из свойств соответствующих одномерных распределений. В частном случае, пусть множества B_1, B_2 имеют вид $B_1=(-\infty,x)$, $B_2=(-\infty,y)$, $x,y\in R$. Тогда для независимых случайных величин справедливо следующее равенство:

$$\mathsf{P}(\{\omega\colon \xi_1(\omega)\in(-\infty,x),\, \xi_2(\omega)\in(-\infty,y)\}) = \\ = \mathsf{P}(\{\omega\colon \xi_1(\omega)\in(-\infty,x)\})\times \mathsf{P}(\{\omega\colon \xi_2(\omega)\in(-\infty,y)\}),$$

откуда сразу получаем $F(x,y)=F_{\xi_1}(x)\,F_{\xi_2}(y)$.

§ 3. Некоторые типы двумерных случайных величин

3.1. Дискретные двумерные случайные величины. Теорема Лебега о разложении интегральной функции распределения позволяет решить важную проблему классификации множества одномерных случайных величин. Было установлено, что это множество включает дискретные, непрерывные, сингулярные и смешанные одномерные случайные величины. Однако многомерные случайные величины по сравнению с одномерными случайными величинами более разнообразны по своей физической природе и математической структуре. Это разнообразие, прежде всего, порождается наличием функциональной или статистической зависимости между случайными величинами, которые составляют многомерную случайную величину. Поэтому не удается провести замкнутую и достаточно простую классификацию совокупности всех случайных векторов, как это было сделано в одномерном случае. Так же, как и для одномерных случайных величин, мы подробно рассмотрим наиболее часто встречающиеся в конкретных задачах дискретные и непрерывные многомерные случайные величины. Ради простоты пусть $n = 2$.

Определение 7.4. Двумерная случайная величина $\xi = (\xi_1, \xi_2)$ называется *дискретной*, если случайные величины ξ_1 и ξ_2 являются дискретными.

Далее, ради простоты, рассмотрим случай, когда каждая из величин ξ_1 и ξ_2 принимает не более чем счетное число возможных значений. Если отображение $(\xi_1(\omega), \xi_2(\omega))\colon \Omega \to R^2$ задано поточечно, то для дискретной величины (ξ_1, ξ_2) при всех $i, j = 1, 2, \ldots$ известны как ее значения (x_i, y_j), так и вероятности $\mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i, \xi_2(\omega) = y'_i\})$, которые связаны с этими значениями. Будем обозначать вероятность $\mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i, \xi_2(\omega) = y'_i\})$ через $r_{i,j}$. Множество $\{r_{i,j}\colon i, j = 1, 2, \ldots\}$ называется *распределением двумерной дискретной случайной величины* (ξ_1, ξ_2). В частном случае распределение может иметь вид $\{r_{i,j}\colon i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, s\}$. Найдем выражение для интегральной функции $F(x, y)$ в случае дискретной двумерной случайной величины. Для такой случайной величины имеем

$$\begin{aligned} F(x,y) &= \mathsf{P}(\{\omega\colon \xi_1(\omega) < x'_i,\ \xi_2(\omega) < y'_j\}) = \\ &= \mathsf{P}\Bigg(\bigcup_{(i,j)\colon x'_i<x, y'_j<y} \{\omega\colon \xi_1(\omega) = x'_i, \xi_2(\omega) = y'_j\}\Bigg) = \\ &= \sum_{i\colon x'_i<x}\ \sum_{j\colon y'_j<y} \mathsf{P}(\{\omega\colon \xi_1(\omega) = \xi'_i, \xi_2(\omega) = y'_j\}) = \sum_{i\colon x'_i<x}\ \sum_{j\colon y'_j<y} r_{i,j}. \end{aligned} \tag{7.9}$$

Здесь суммирование распространяется на все такие индексы i и j, для которых $x'_i < x$ и $y'_j < y$.

Далее, зная распределение $\{r_{i,j}\colon i, j = 1, 2, \ldots\}$ двумерной дискретной случайной величины (ξ_1, ξ_2), можно определить и интегральную функцию распределения каждой из одномерных случайных величин ξ_1, ξ_2. Действительно, из (7.9) имеем

$$F_{\xi_1}(x) = F(x, +\infty) = \sum_{i\colon x'_i<x}\ \sum_{j} r_{i,j}, \quad F_{\xi_2}(y) = F(+\infty, y) = \sum_{j\colon y'_i<y}\ \sum_{i} r_{i,j}.$$

Если двумерная дискретная случайная величина $\xi = (\xi_1, \xi_2)$ принимает не более чем счетное число значений, то каждая из ее компонент ξ_1, ξ_2 также принимает не более чем счетное число значений. Естественно, что справедливо и обратное утверждение.

Пусть элементы множеств $\{x'_i\colon i = 1, 2, \ldots\}$ и $\{y'_j\colon j = 1, 2, \ldots\}$ возможных значений дискретных случайных величин ξ_1 и ξ_2 можно занумеровать и упорядочить таким образом, что $x'_1 < x'_2 < \ldots < x'_i < \ldots$ и соответственно $y'_1 < y'_2 < \ldots < y'_i < \ldots$. Тогда наряду с распределением

$\{r_{i,j} : i, j = 1, 2, \ldots\}$ двумерной дискретной случайной величины (ξ_1, ξ_2) можно ввести понятие ее таблицы распределения (табл. 7.1).

Таблица 7.1

$\xi_2 \backslash \xi_1$	x'_1	x'_2	$\ldots$	x'_i	$\ldots$
y'_1	$r_{1,1}$	$r_{2,1}$	$\ldots$	$r_{i,1}$	$\ldots$
y'_2	$r_{1,2}$	$r_{2,2}$	$\ldots$	$r_{i,2}$	$\ldots$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
y'_j	$r_{1,j}$	$r_{2,j}$	$\ldots$	$r_{i,j}$	$\ldots$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$

Перечислим основные свойства табл. 7.1. Очевидно, что $r_{i,j} \geqslant 0$, а из (7.9) имеем $F(+\infty, +\infty) = \sum_{i,j} r_{i,j} = 1$.

Если событие $\{\omega\colon \xi_1(\omega) = x'_i\}$ означает, что случайная величина ξ_1 примет возможное значение x'_i, то события $\{\omega\colon \xi_1(\omega) = x'_i\}$, $i = 1, 2, \ldots$, образуют полную группу попарно несовместных событий. Тогда при $j = 1, 2, \ldots$ событие $\{\omega\colon \xi_2(\omega) = y'_j\} = \{\omega\colon \xi_2(\omega) = y'_j\} \cap \cap \left(\cup_i \{\omega\colon \xi_1(\omega) = x'_i\}\right) = \cup_i \{\omega\colon \xi_1(\omega) = x'_i,\ \xi_2(\omega) = y'_j\}$. Здесь событие $\{\omega\colon \xi_2(\omega) = y'_j\}$ означает, что случайная величина ξ_2 примет возможное значение y'_j. Легко видеть, что $\cup_j \{\omega\colon \xi_2(\omega) = y'_j\} = \Omega$, где события $\{\omega\colon \xi_2(\omega) = y'_j\}$, $j \geqslant 1$, попарно несовместны. Совершенно так же при любом $i = 1, 2, \ldots$ для события $\{\omega\colon \xi_1(\omega) = x'_i\}$ справедлива цепочка равенств $\{\omega\colon \xi_1(\omega) = x'_i\} = \{\omega\colon \xi_1(\omega) = x'_i\} \cap \left(\cup_j \{\omega\colon \xi_2(\omega) = y'_j\}\right) = = \cup_j \{\omega\colon \xi_1(\omega) = x'_i,\ \xi_2(\omega) = y'_j\}$. Так как события $\{\omega\colon \xi_1(\omega) = x'_i\}$, $i \geqslant 1$, попарно несовместны, то события $\{\omega\colon \xi_1(\omega) = x'_i,\ \xi_2(\omega) = y'_j\}$, $i \geqslant 1$, также попарно несовместны. Аналогичным образом, так как события вида $\{\omega\colon \xi_2(\omega) = y'_j\}$, $j \geqslant 1$, попарно несовместны, то случайные события $\{\omega\colon \xi_1(\omega) = x'_i,\ \xi_2(\omega) = y'_j\}$, $j \geqslant 1$, также попарно несовместны. Используя это и выражения вида

$$\{\omega\colon \xi_1(\omega) = x'_i\} = \cup_j \{\omega\colon \xi_1(\omega) = x'_i,\ \xi_2(\omega) = y'_j\},$$
$$\{\omega\colon \xi_2(\omega) = y'_j\} = \cup_i \{\omega\colon \xi_1(\omega) = x'_i,\ \xi_2(\omega) = y'_j\},$$

нетрудно найти формулы

$$\begin{aligned} p_i &= \mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i\}) = \sum_j r_{i,j}, \\ q_j &= \mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\}) = \sum_i r_{i,j} \end{aligned} \tag{7.10}$$

для подсчета распределений одномерных случайных величин ξ_1 и ξ_2.

Итак, если знаем таблицу распределения двумерной случайной величины, то легко находится распределение для каждой из одномерных случайных величин. Обратно неверно: если известны одномерные распределения $\{p_i\colon i \geqslant 1\}$ и $\{q_j\colon j \geqslant 1\}$, то двумерное распределение $\mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i, \xi_2(\omega) = y'_j\}) = \mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i\}) \times \mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\} \mid \{\omega\colon \xi_1(\omega) = x'_i\})$ не всегда можно вычислить, так как условное распределение $\mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\} \mid \{\omega\colon \xi_1(\omega) = x'_i\})$ в общем случае может быть неизвестным.

Рассмотрим случай, когда одномерные распределения случайных величин ξ_1 и ξ_2 полностью определяют двумерное распределение $\{r_{i,j}\colon i,j = 1,2,\ldots\}$. Пусть случайные величины ξ_1 и ξ_2 одновременно являются дискретными и статистически независимыми. Тогда из (7.4) при $n = 2$, $B_1 = B_{1,i} = \{x'_i\} \in \mathcal{B}$, $B_2 = B_{2,j} = \{y'_i\} \in \mathcal{B}$ получаем для любых $i,j = 1,2,\ldots$ равенство вида $r_{i,j} = \mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i, \xi_2(\omega) = y'_j\}) = \mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i\}) \times \mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\}) = p_i\, q_j$.

Обратно, пусть (ξ_1, ξ_2) является двумерной дискретной случайной величиной и $r_{i,j} = p_i\, q_j$ для всех $i,j = 1,2,\ldots$. Отсюда, используя (7.9) и формулы $F_{\xi_1}(x) = \sum_{i\colon x'_i<x} p_i$, $F_{\xi_2}(y) = \sum_{j\colon y'_j<y} q_j$, получим, что

$$F(x,y) = \sum_{i\colon x'_i<x}\ \sum_{j\colon y'_j<y} r_{i,j} = \sum_{i\colon x'_i<x} p_i \times \sum_{j\colon y'_j<y} q_j = F_{\xi_1}(x)\, F_{\xi_2}(y)$$

для каждой точки $(x,y) \in R^2$. Значит, случайные величины ξ_1 и ξ_2 являются статистически независимыми.

Итак, *для независимости дискретных случайных величин* ξ_1 и ξ_2 *необходимо и достаточно*, *чтобы для всех* $i,j = 1,2,\ldots$ *имело место равенство* $r_{i,j} = p_i\, q_j$. Приведем пример на определение законов распределения системы случайных величин.

Пример 7.7. Два игрока независимо друг от друга случайным образом бросают по два раза каждый свою несимметричную монету. Вероятность выпадения орла для первого игрока равна 0,7, а для второго игрока равна 0,4. Найти законы распределения системы случайных величин ξ_1 и ξ_2, которые определяют число выпадений орлов для первого и второго игрока.

Решение. В качестве описания $\omega = \omega_k$ произвольного элементарного исхода данной игры выберем вектор вида (u_1, u_2, u_3, u_4), где $u_c = 0, 1$ для всех $c = 1,2,3,4$. Символы u_1 и u_2 определяют число выпадений орлов для первого игрока при первом и соответственно при втором броске. Аналогично, символы u_3 и u_4 определяют число выпадений орлов для второго игрока при первом и соответственно при втором броске. Так, например, описание $\omega_0 = (0,0,0,0)$ означает, что игроки выбросили только решетки, а описание $\omega_{15} = (1,1,1,1)$ соответствует такому элементарному исходу, когда игроки выбросили только орлы. Ради определенности считаем, что при $\omega_k = (u_1, u_2, u_3, u_4)$ имеем $k = u_1\, 2^3 + u_2\, 2^2 + u_3\, 2^1 + u_4\, 2^0$.

Итак, для этого опыта имеем 16 элементарных исходов и 2^{16} всех наблюдаемых исходов. При этом достоверное событие $\Omega = \{\omega_0, \omega_1, \ldots, \omega_{15}\}$ и множество $\mathcal{F} = \{A\colon A \subset \Omega\}$. Используя независимость каждого из бросков, легко вычислить вероятности элементарных событий. Например, вероятность $\mathsf{P}(\{\omega_0\}) = 0{,}0324$ и вероятность $\mathsf{P}(\{\omega_{15}\}) = 0{,}0784$. Так как монеты не являются симметричными, то в этом эксперименте элементарные события не будут равновозможными. Пользуясь теоремой сложения, можно вычислить вероятности любого события $A \in \mathcal{F}$. Значит, построили вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ для этой простой игры. Легко найти поточечное задание случайных величин ξ_1 и ξ_2. Действительно, имеем $\{\omega\colon \xi_1(\omega) = 0\} = \{\omega_0, \omega_1, \omega_2, \omega_3\}$, $\{\omega\colon \xi_1(\omega) = 1\} = \{\omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9, \omega_{10}, \omega_{11}\}$, $\{\omega\colon \xi_1(\omega) = 2\} = \{\omega_{12}, \omega_{13}, \omega_{14}, \omega_{15}\}$, $\{\omega\colon \xi_2(\omega) = 0\} = \{\omega_0, \omega_4, \omega_8, \omega_{12}\}$, $\{\omega\colon \xi_2(\omega) = 1\} = \{\omega_1, \omega_2, \omega_5, \omega_6, \omega_9, \omega_{10}, \omega_{13}, \omega_{14}\}$, $\{\omega\colon \xi_2(\omega) = 2\} = \{\omega_3, \omega_7, \omega_{11}, \omega_{15}\}$. Таким образом, возможными значениями случайных величин ξ_1 и ξ_2 являются: $x_1' = 0$, $x_2' = 1$, $x_3' = 2$ и $y_1' = 0$, $y_2' = 1$, $y_3' = 2$. Поэтому двумерная случайная величина (ξ_1, ξ_2) принимает значения вида $(0,0)$, $(0,1)$, $(0,2)$, $(1,0)$, $(1,1)$, $(1,2)$, $(2,0)$, $(2,1)$, $(2,2)$. Теперь имеем всю необходимую информацию, чтобы выписать таблицу распределения для случайного вектора (ξ_1, ξ_2) и ряды распределения для случайных величин ξ_1 и ξ_2. Итак, распределение для (ξ_1, ξ_2) примет вид табл. 7.2.

Таблица 7.2

$\xi_2 \backslash \xi_1$	0	1	2
0	0,0324	0,1512	0,1764
1	0,0432	0,2016	0,2352
2	0,0144	0,0672	0,0784

Например, вероятность $r_{2,3}$, которая приведена на пересечении третьего столбца и четвертой строчки табл. 7.2, можно вычислить с использованием поточечного задания величин ξ_1 и ξ_2 следующим образом:

$$\begin{aligned} r_{2,3} &= \mathsf{P}(\{\omega\colon \xi_1(\omega) = 1, \xi_2(\omega) = 2\}) = \\ &= \mathsf{P}(\{\omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9, \omega_{10}, \omega_{11}\} \cap \{\omega_3, \omega_7, \omega_{11}, \omega_{15}\}) = \\ &= \mathsf{P}(\{\omega_7, \omega_{11}\}) = \mathsf{P}(\{\omega_7\}) + \mathsf{P}(\{\omega_{11}\}) = 0{,}0336 + 0{,}0336 = 0{,}0672. \end{aligned}$$

Для контроля правильности вычислений просуммируем все такие элементы этой таблицы, которые соответствуют вероятностям. Легко получить, что эта сумма равна единице.

Используя табл. 7.2, легко вычислить по формуле (7.9) совместную функцию распределения $F(x, y)$ случайных величин ξ_1 и ξ_2. Например, функция $F(x,y)=\sum_{i\colon x'_i<x}\sum_{j\colon y'_j<y} r_{i,j}=0$ при $x\leqslant 0$ и $y\leqslant 0$, так как $x'_i, y'_j\in\{0,1,2\}$ для всех $i,j=1,2,3$. Далее, если $0<x\leqslant 1$ и $0<y\leqslant 1$, то функция $F(x,y)=\sum_{i\colon x'_i<x}\sum_{j\colon y'_j<y} r_{i,j}=r_{1,1}=0{,}0324$. Если же $1<x\leqslant 2$ и $1<y\leqslant 2$, то интегральная функция $F(x,y)=$ $=\sum_{i\colon x'_i<x}\sum_{j\colon y'_j<y} r_{i,j}=r_{1,1}+r_{2,1}+r_{1,2}+r_{2,2}=0{,}0324+0{,}1512+$ $+0{,}0432+0{,}2016=0{,}4284$. Аналогично поступая, можно вычислить и все остальные значения функции $F(x,y)$.

Из табл. 7.2, используя приведенные ранее формулы (7.10), получим ряд распределения (табл. 7.3)

Таблица 7.3

ξ_1	0	1	2
$\mathsf{P}(\cdot)$	0,09	0,42	0,49

для одномерной случайной величины ξ_1 и ряд распределения (табл. 7.4)

Таблица 7.4

ξ_2	0	1	2
$\mathsf{P}(\cdot)$	0,36	0,48	0,16

для одномерной случайной величины ξ_2. Вероятности в этих рядах мы могли получить непосредственно из поточечного задания этих случайных величин. В самом деле, например, $\mathsf{P}(\{\omega\colon \xi_1(\omega)=0\})=$ $=\mathsf{P}(\{\omega_0,\omega_1,\omega_2,\omega_3\})=\mathsf{P}(\{\omega_0\})+\mathsf{P}(\{\omega_1\})+\mathsf{P}(\{\omega_2\})+\mathsf{P}(\{\omega_3\})=0{,}09$ или $\mathsf{P}(\{\omega\colon \xi_2(\omega)=2\})=\mathsf{P}(\{\omega_3,\omega_7,\omega_{11},\omega_{15}\})=\mathsf{P}(\{\omega_3\})+\mathsf{P}(\{\omega_7\})+$ $+\mathsf{P}(\{\omega_{11}\})+\mathsf{P}(\{\omega_{15}\})=0{,}16$.

Наконец, используя полученные таблицы при решении примера, легко проверить очевидное утверждение о том, что случайные величины ξ_1 и ξ_2, которые определяют число выпадений орлов для первого и соответственно второго игрока, являются статистически независимыми. Например, $r_{1,1}=0{,}0324=0{,}09\times 0{,}036=p_1q_1$, $r_{1,2}=0{,}0432=0{,}09\times 0{,}48=p_1q_2$ и т. д.

3.2. Непрерывные двумерные случайные величины. В качестве другого важного и наиболее удачного класса двумерных случайных величин рассмотрим так называемое множество непрерывных случайных величин.

Определение 7.5. Двумерная случайная величина $(\xi_1(\omega),\xi_2(\omega))$: $\Omega\to R^2$ называется *непрерывной*, если существует такая неотрицательная функция $f(x,y)$ с областью определения R^2, что для любой области H двумерной плоскости, для которой существует площадь, имеет место равенство

$$\mathsf{P}(\{\omega\colon (\xi_1, \xi_2) \in H\}) = \iint\limits_H f(x, y)\, dx\, dy. \tag{7.11}$$

Здесь через символ $\iint_H f(x, y)\, dx\, dy$ обозначается двойной интеграл от функции $f(x, y)$ по области H. Отсюда следует, что непрерывная двумерная случайная величина принимает несчетное число различных значений, сплошь заполняя либо всю двумерную плоскость R^2, либо некоторую область D на этой плоскости, причем площадь D должна быть строго положительна. Вспоминая специальное обозначение двойного интеграла на случай прямоугольной области, получим, что вероятность попадания вектора (ξ_1, ξ_2) в прямоугольник $H = \{(u, v)\colon u_1 \leqslant u < u_2,\ v_1 \leqslant v < v_2\}$ равна

$$\mathsf{P}(\{\omega\colon u_1 \leqslant \xi_1(\omega) < u_2,\ v_1 \leqslant \xi_2(\omega) < v_2\}) = \int\limits_{u_1}^{u_2} \int\limits_{v_1}^{v_2} f(x, y)\, dx\, dy. \tag{7.12}$$

Неотрицательная функция $f(x, y)$ называется *плотностью распределения вероятностей случайной величины* (ξ_1, ξ_2). Так как $\{\omega\colon (\xi_1, \xi_2) \in R^2\} = \Omega$, то из (7.11) имеем равенство

$$\mathsf{P}(\{\omega\colon (\xi_1, \xi_2) \in R^2\}) = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} f(x, y)\, dx\, dy = 1.$$

Это равенство называется *условием нормировки для плотности распределения* $f(x, y)$. При $H = \{(u, v)\colon -\infty < u < x,\ -\infty < v < y\}$ из равенства (7.12) найдем выражение для интегрального закона распределения $F(x, y)$ двумерной непрерывной случайной величины в следующем виде:

$$F(x, y) = \mathsf{P}(\{\omega\colon (\xi_1, \xi_2) \in H\}) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(u, v)\, du\, dv. \tag{7.13}$$

Используя формулы (7.3), (7.13) и свойство аддитивности двойного интеграла по области, легко вычислить вероятность попадания двумерной случайной величины в любой прямоугольник и, тем самым, еще раз подтвердить равенство (7.12). Следовательно, будем иметь формулу (7.11) для любой области H двумерной плоскости, для которой существует площадь. Поэтому в определении непрерывной случайной величины равенство (7.11) можно заменить на более простое равенство (7.13). Для эффективного вычисления функции $F(x, y)$ с помощью формулы (7.13) будем в дальнейшем предполагать, что при всех $(x, y) \in R^2$ имеет место следующее соотношение:

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv = \int_{-\infty}^{x}\Big(\int_{-\infty}^{y} f(u,v)\,dv\Big)\,du = \\ = \int_{-\infty}^{y}\Big(\int_{-\infty}^{x} f(u,v)\,du\Big)\,dv. \quad (7.14)$$

Соотношение (7.14), которое означает равенство двойного и повторных интегралов, имеет место, например, в случае, когда функция $f(x, y)$ непрерывна в каждой точке $(x, y) \in R^2$. Дифференцируя правую часть равенства (7.14) сначала по x, а затем по y, легко найдем $f(x, y) = \partial^2 F(x, y)/\partial x\,\partial y$. Если не предполагаем непрерывность функции $f(x, y)$ в каждой точке плоскости, то для плотности распределения вероятностей $f(x, y)$ почти всюду относительно лебеговской меры в R^2 верно такое представление. По этой причине функцию $f(x, y)$ еще называют *дифференциальной функцией распределения*.

Рассмотрим теперь важный для практики частный случай, когда величины ξ_1 и ξ_2 одновременно являются непрерывными и статистически независимыми. Покажем, что (ξ_1, ξ_2) будет двумерной непрерывной случайной величиной. Действительно, используя равенство (7.6) и связь между двойным и повторными интегралами, последовательно находим, что

$$F(x,y) = F_{\xi_1}(x) \times F_{\xi_2}(y) = \int_{-\infty}^{x} f_{\xi_1}(u)\,du \times \int_{-\infty}^{y} f_{\xi_2}(v)\,dv = \\ = \int_{-\infty}^{x}\Big(\int_{-\infty}^{y} f_{\xi_1}(u)\,f_{\xi_2}(v)\,dv\Big)\,du = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{\xi_1}(u)\,f_{\xi_2}(v)\,du\,dv.$$

Поэтому $f(x, y) = f_{\xi_1}(x) \times f_{\xi_2}(y)$ для каждой точки $(x, y) \in R^2$ и, значит, математический объект (ξ_1, ξ_2) является непрерывным случайным вектором. Обратно, пусть (ξ_1, ξ_2) является двумерной непрерывной случайной величиной и дифференциальная функция распределения $f(x, y) = f_{\xi_1}(x) \times f_{\xi_2}(y)$ для каждой точки $(x, y) \in R^2$. Отсюда для $(x, y) \in R^2$ получим, что

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{\xi_1}(u)\,f_{\xi_2}(v)\,du\,dv = \int_{-\infty}^{x}\Big(\int_{-\infty}^{y} f_{\xi_1}(u)\,f_{\xi_2}(v)\,dv\Big)\,du = \\ = \int_{-\infty}^{x} f_{\xi_1}(u)\,du \times \int_{-\infty}^{y} f_{\xi_2}(v)\,dv = F_{\xi_1}(x) \times F_{\xi_2}(y).$$

Следовательно, ξ_1 и ξ_2 являются статистически независимыми.

Заметим, что дифференциальная функция распределения $f(x, y)$ двумерной случайной величины (ξ_1, ξ_2) определяет дифференциальную функцию распределения каждой из одномерных случайных величин ξ_1 и ξ_2. Следовательно, если вектор (ξ_1, ξ_2) является непрерывным, то его составляющие ξ_1 и ξ_2 будут также непрерывными. В самом деле, из свойств функции $F(x, y)$ имеем равенства $F_{\xi_1}(x) = F(x, +\infty) = \int_{-\infty}^{x}\left(\int_{-\infty}^{+\infty} f(u, v)\,dv\right)du$, $F_{\xi_2}(y) = F(+\infty, y) = \int_{-\infty}^{y}\left(\int_{-\infty}^{+\infty} f(u, v)\,du\right)dv$ для одномерных случайных величин. Такое представление интегральных функций распределения $F_{\xi_1}(x)$ и $F_{\xi_2}(y)$ для случайных величин ξ_1 и ξ_2 в случае непрерывной двумерной случайной величины позволяет выписать соответствующие одномерные плотности вероятностей $f_{\xi_1}(x)$ и $f_{\xi_2}(y)$ в виде следующих формул:

$$f_{\xi_1}(x) = \int\limits_{-\infty}^{+\infty} f(x, v)\,dv, \qquad f_{\xi_2}(y) = \int\limits_{-\infty}^{+\infty} f(u, y)\,du. \tag{7.15}$$

Пример 7.8. Плотность распределения упорядоченной пары (ξ_1, ξ_2) имеет вид $f(x, y) = \pi^{-2}(x^2 + y^2 + x^2y^2 + 1)^{-1}$. Вычислите $f_{\xi_1}(x)$, $f_{\xi_2}(y)$ и определите, зависимы или независимы случайные величины ξ_1 и ξ_2.

Решение. Используя (7.15), найдем соответствующие одномерные плотности распределения вероятностей для случайных величин ξ_1 и ξ_2:

$$f_{\xi_1}(x) = \int\limits_{-\infty}^{+\infty} f(x, y)\,dy = \int\limits_{-\infty}^{+\infty} \pi^{-2}(x^2 + y^2 + x^2y^2 + 1)^{-1}\,dy = (\pi(1 + x^2))^{-1},$$

$$f_{\xi_2}(y) = \int\limits_{-\infty}^{+\infty} f(x, y)\,dx = \int\limits_{-\infty}^{+\infty} \pi^{-2}(x^2 + y^2 + x^2y^2 + 1)^{-1}\,dx = (\pi(1 + y^2))^{-1}.$$

Тогда выводим, что при любых $x, y \in R$ выполняется соотношение следующего вида: $f_{\xi_1}(x) \cdot f_{\xi_2}(y) = (\pi(1 + x^2))^{-1} \times (\pi(1 + y^2))^{-1} = \pi^{-2}(x^2 + y^2 + x^2y^2 + 1)^{-1} = f(x, y)$, т. е. случайные величины ξ_1 и ξ_2 независимы.

Выясним теперь геометрический смысл плотности распределения вероятностей или функции $f(x, y)$. Пусть (ξ_1, ξ_2) — непрерывный вектор с плотностью $f(x, y)$. Выберем на плоскости $x0y$, которая представляет собой область определения функции $z = f(x, y)$, так называемый элементарный прямоугольник $\{(u, v)\colon x \leqslant u < x + \Delta x, y \leqslant v < y + \Delta y\}$. Напомним, что вероятность $\mathsf{P}(\{\omega\colon x \leqslant \xi_1 < x + \Delta x, y \leqslant \xi_2 < y + \Delta y\})$ попадания случайной точки (ξ_1, ξ_2) в прямоугольник будет равна

$F(x+\Delta x, y+\Delta y) + F(x, y) - F(x+\Delta x, y) - F(x, y+\Delta y)$. Отсюда последовательно находим, что предел

$$\begin{aligned}\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} \mathsf{P}(\{\omega\colon x \leqslant \xi_1 < x+\Delta x, y \leqslant \xi_2 < y+\Delta y\})/\Delta x\,\Delta y =\\ = \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} \bigl(F(x+\Delta x, y+\Delta y) + F(x, y) -\\ - F(x+\Delta x, y) - F(x, y+\Delta y)\bigr)/\Delta x\,\Delta y =\\ = \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} (F(x+\Delta x, y+\Delta y) - F(x+\Delta x, y))/\Delta x\,\Delta y +\\ + \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} (F(x, y) - F(x, y+\Delta y))/\Delta x\,\Delta y.\end{aligned}$$

На рис. 7.7 изображена поверхность $z = f(x, y)$ и выбранный элементарный прямоугольник $\{(u, v)\colon x \leqslant u < x+\Delta x, y \leqslant v < y+\Delta y\}$.

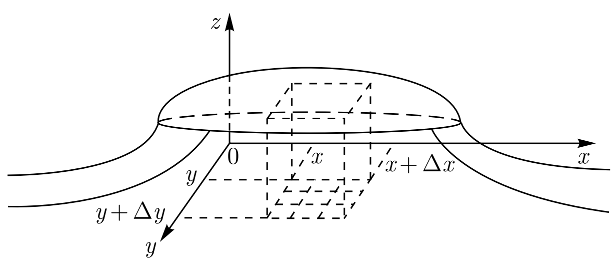


Рис. 7.7

Будем теперь предполагать, что все пределы здесь существуют. В этом случае можно продолжить преобразования следующим образом:

$$\begin{aligned}\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} (F(x+\Delta x, y+\Delta y) - F(x+\Delta x, y))/\Delta x\,\Delta y +\\ + \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} (F(x, y) - F(x, y+\Delta y))/\Delta x\,\Delta y =\\ = \lim_{\Delta x\to 0} \Bigl(\lim_{\Delta y\to 0} (F(x+\Delta x, y+\Delta y) - F(x+\Delta x, y))/\Delta y\Bigr)/\Delta x +\\ + \lim_{\Delta x\to 0} \Bigl(\lim_{\Delta y\to 0} (F(x, y) - F(x, y+\Delta y))/\Delta y\Bigr)/\Delta x =\\ = \lim_{\Delta x\to 0} (\partial F(x+\Delta x, y)/\partial y)/\Delta x - \lim_{\Delta x\to 0} (\partial F(x, y)/\partial y)/\Delta x =\\ = \lim_{\Delta x\to 0} (\partial F(x+\Delta x, y)/\partial y - \partial F(x, y)/\partial y)/\Delta x = \partial^2 F(x, y)/\partial x\,\partial y.\end{aligned}$$

Если снова предположить, что дифференциальная функция $f(x, y)$ непрерывна в точке (x, y), то $\partial^2 F(x, y)/\partial x\,\partial y = f(x, y)$. Итак, получили, что

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} \mathsf{P}(\{\omega\colon x \leqslant \xi_1 < x+\Delta x, y \leqslant \xi_2 < y+\Delta y\})/\Delta x\,\Delta y = f(x, y) \quad (7.16)$$

или $\mathsf{P}(\{\omega\colon x \leqslant \xi_1(\omega) < x + \Delta x, y \leqslant \xi_2(\omega) < y + \Delta y\}) = f(x, y)\,\Delta x\,\Delta y + o(\Delta x\,\Delta y) \approx f(x, y)\,\Delta x\,\Delta y$. Величина $f(x, y)\,\Delta x\,\Delta y$ равна объему параллелепипеда с основанием $\{(u, v)\colon x \leqslant u < x + \Delta x, y \leqslant v < y + \Delta y\}$, высотой $f(x, y)$ и вычисляет с точностью до бесконечно малых высших порядков вероятность попадания случайной точки (ξ_1, ξ_2) в элементарный прямоугольник. При этом величину $z = f(x, y)$ можно приближено считать вероятностью единицы площади элементарного прямоугольника. Итак, на плотность распределения можно смотреть как на некоторую функцию $z = f(x, y)$, определяющую некоторую поверхность плотности распределения. С ее помощью можно находить вероятность попадания двумерной случайной величины в элементарные прямоугольники. Рассмотрим пример непрерывного случайного вектора.

Пример 7.9. Снова обратимся к задаче о гончарном круге или к задаче Бертрана из второй главы. При этом здесь рассматриваем третий вариант этой задачи, когда случайное появление трещины на верхней поверхности гончарного круга не связано ни с закреплением одного из ее концов, ни с фиксированием ее направления. В этом случае положение трещины на верхней поверхности гончарного круга или положение хорды в круге радиуса r определяется ее серединой. Средняя точка $\{\omega\}$ хорды выбирается с помощью такого механизма, что вероятность попадания этой точки в любую область круга зависит только от площади области и не зависит от расположения области внутри круга и от ее формы. Обозначим через $|\xi_1|$ и $|\xi_2|$ расстояния от середины хорды до оси ординат и до оси абсцисс, если центр круга и центр прямоугольной системы координат совпадают. Найти законы распределения случайного вектора (ξ_1, ξ_2).

Решение. Во второй главе для этого эксперимента была построена вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, где достоверное событие имеет вид $\Omega = \{\omega = (x, y)\colon x^2 + y^2 \leqslant r^2,\ (x, y) \neq (0, 0)\}$, множество $\mathcal{F}$ всех наблюдаемых исходов равно $\{A\colon A \subset \Omega,\ A$ имеет площадь $\mathrm{mes}\,(A)\}$ и вероятность $\mathsf{P}(A) = \mathrm{mes}\,(A)/\pi r^2$. Отметим, что в этой задаче $\xi_1(\omega) = x$ и $\xi_2(\omega) = y$. Пусть числа u, v, Δu и Δv таковы, что событие $A_{uv} = \{\omega = (x, y)\colon u \leqslant \xi_1(\omega) < u + \Delta u,\ v \leqslant \xi_2(\omega) < v + \Delta v\}$ является прямоугольником. В этом случае, учитывая равенство $\mathrm{mes}\,(A_{uv}) = \Delta u\,\Delta v$, находим

$$\mathsf{P}(A_{uv}) = \mathrm{mes}\,(A_{uv})/\pi r^2 = \Delta u\,\Delta v/\pi r^2, \qquad \lim_{\substack{\Delta u \to 0\\ \Delta v \to 0}} \mathsf{P}(A_{uv})/\Delta u\,\Delta v = 1/\pi r^2.$$

Используя это и равенство (7.16), получим, что плотность $f(u, v) = 1/\pi r^2$ для всех $(u, v) \in \{(u, v)\colon u^2 + v^2 < r^2\}$. Так как случайная точка $\{\omega\}$ не может попасть за пределы области вида $\{(u, v)\colon u^2 + v^2 \leqslant r^2\}$, то для элемента $(u, v) \in \{(u, v)\colon u^2 + v^2 > r^2\}$ найдется такое $\varepsilon > 0$, что множество $A_{uv} = \varnothing$ при $|\Delta u| < \varepsilon$, $|\Delta v| < \varepsilon$. Отсюда получим $f(u, v) = 0$ для $(u, v) \in \{(u, v)\colon u^2 + v^2 > r^2\}$.

Осталось рассмотреть случай, когда $(u,v) \in \{(u,v)\colon u^2+v^2=r^2\}$. Для такого рода чисел u, v величина $\operatorname{mes}(A_{uv})$ существенно зависит от знака Δu и знака Δv. Например, для фиксированных и положительных u, v имеем $A_{uv} = \{\omega\colon u \leqslant \xi_1(\omega) < u+\Delta u,\ v \leqslant \xi_2(\omega) < v+\Delta v\}$,

$$\lim_{\substack{\Delta u\to 0\\ \Delta v\to 0}} \mathsf{P}(A_{uv})/\Delta u\,\Delta v = 0 \quad \text{при} \ \ \Delta u>0,\ \ \Delta v>0.$$

Если же $u>0$, $v>0$, $\Delta u<0$ и $\Delta v<0$, то, выбирая теперь в качестве события A_{uv} множество $\{\omega\colon u+\Delta u < \xi_1(\omega) \leqslant u,\ v+\Delta v < \xi_2(\omega) \leqslant v\}$, легко найдем, что

$$\lim_{\substack{\Delta u\to 0\\ \Delta v\to 0}} \mathsf{P}(A_{uv})/\Delta u\,\Delta v = 1/\pi r^2.$$

В силу этого при $(u,v) \in \{(u,v)\colon u^2+v^2=r^2\}$ можно положить $f(u,v)=0$. С целью контроля полученного выражения для дифференциальной функции распределения $f(u,v)$ проверим так называемое условие нормировки: $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(u,v)\,du\,dv = (1/\pi r^2)\iint_{H_0} du\,dv = 1$, где $H_0 = \{(u,v)\colon u^2+v^2<r^2\}$. Так как плотность $f(u,v)$ равна постоянной величине $1/\pi r^2$ при $(u,v)\in\{(u,v)\colon u^2+v^2<r^2\}$ и равна нулю при $(u,v)\in\{(u,v)\colon u^2+v^2\geqslant r^2\}$, то случайный вектор (ξ_1,ξ_2) для этой задачи называется равномерно распределенным внутри круга радиуса r.

Интегральный закон распределения $F(u,v)$ двумерной случайной величины (ξ_1,ξ_2) легко вычислим по формуле (7.13) следующим способом: $F(u,v) = \int_{-\infty}^{u}\int_{-\infty}^{v} f(x_1,x_2)\,dx_1\,dx_2 = (1/\pi r^2)\iint_H dx_1\,dx_2$, $H = \{(x_1,x_2)\colon x_1^2+x_2^2<r^2\} \cap \{(x_1,x_2)\colon -\infty<x_1<u,\ -\infty<x_2<v\}$. Например, $F(u,v)=0$ при $u\leqslant -r$ и $-\infty<v<+\infty$, или $F(u,v)=1$ при $u,v>r$. Используя равенства (7.15), можно вычислить плотности распределения одномерных случайных величин ξ_1 и ξ_2. На рис. 7.8 приведены графики функций $f_{\xi_1}(u)$ и $f_{\xi_2}(v)$.

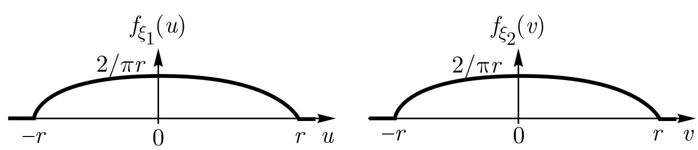


Рис. 7.8

Здесь

$$f_{\xi_1}(u) = \int_{-\infty}^{+\infty} f(u,v)\,dv = (1/\pi r^2)\int_{-\sqrt{r^2-u^2}}^{\sqrt{r^2-u^2}} dv = 2\sqrt{r^2-u^2}\,/\pi r^2 \quad \text{при } |u|<r$$

и

$$f_{\xi_2}(v) = \int_{-\infty}^{+\infty} f(u,v)\,du = (1/\pi r^2)\int_{-\sqrt{r^2-v^2}}^{\sqrt{r^2-v^2}} du = 2\sqrt{r^2-v^2}\,/\pi r^2 \quad \text{при } |v|<r.$$

Ясно, что $f_{\xi_1}(u) = 0$ при $|u| \geqslant r$ и $f_{\xi_2}(v) = 0$ при $|v| \geqslant r$. Из рис. 7.8 видно, что, несмотря на равномерное распределение двумерной случайной величины (ξ_1, ξ_2) внутри круга $\{(u, v)\colon u^2 + v^2 \leqslant r^2\}$, каждая из одномерных случайных величин ξ_1 и ξ_2 не распределена равномерно. Заметим, что для этого примера случайные величины ξ_1 и ξ_2 являются статистически зависимыми, так как $f(u, v) \neq f_{\xi_1}(u) \times f_{\xi_2}(v)$.

Итак, были рассмотрены дискретные и непрерывные двумерные случайные величины. Такого рода двумерные случайные величины наиболее часто используются на практике. Если ξ_1 и ξ_2 являются дискретными случайными величинами, то (ξ_1, ξ_2) всегда будет двумерным дискретным случайным вектором. Однако для непрерывных случайных величин это свойство может не всегда выполняться. В связи с этим рассмотрим пример, который показывает, что в многомерном случае существует более сложное разнообразие векторных случайных величин.

Пусть ξ_1 является одномерной непрерывной случайной величиной и случайная величина $\xi_2 = a\,\xi_1 + b$, где $a > 0$. Тогда ξ_2 будет одномерной непрерывной случайной величиной. Поэтому случайный вектор (ξ_1, ξ_2) не является дискретным. Легко видеть, что двумерная случайная величина (ξ_1, ξ_2) принимает заданное значение с нулевой вероятностью из несчетного множества, все элементы которого расположены только на прямой $y = ax + b$ двумерной плоскости. С другой стороны, при $y > ax + b$ интегральная функция вида $F(x, y) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x, \xi_2(\omega) < y\}) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x, a\xi_1 + b < y\}) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\}) = F_{\xi_1}(x)$, а при $y < ax + b$ интегральная функция $F(x, y) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x, a\xi_1 + b < y\}) = \mathsf{P}(\{\omega\colon a\xi_1 + b < y\}) = \mathsf{P}(\{\omega\colon \xi_1 < (y - b)\,a^{-1}\}) = F_{\xi_1}((y - b)\,a^{-1})$. Отсюда получаем, что при $y \neq ax + b$ имеет место равенство $\partial^2 F(x, y)/\partial x\,\partial y = 0$. Значит, если бы случайный вектор (ξ_1, ξ_2) был непрерывным, то $f(x, y) = \partial^2 F(x, y)/\partial x\,\partial y = 0$ для всех $y \neq ax + b$. Далее, площадь множества значений двумерной случайной величины (ξ_1, ξ_2) равна нулю. Поэтому неотрицательную функцию $f(x, y)$, для которой имеет место условие нормировки, здесь определить не удается. Значит, случайный вектор (ξ_1, ξ_2) нельзя назвать непрерывным. Этот пример показывает, что нельзя выполнить достаточно простую классификацию случайных векторов.

§ 4. Законы распределения количественных характеристик условного эксперимента

4.1. Унифицированная вероятностная модель и условные законы распределения случайных величин. Рассмотрим снова вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ статистически устойчивого эксперимента E, его некоторые количественные характеристики или случайные величины ξ_1, ξ_2 и произвольное событие $B \in \mathcal{F}$. При $\mathsf{P}(B) > 0$ в третьей главе для условного эксперимента $E_{\text{у}}$ была построена унифицированная вероятностная модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$, где условная вероятность $\mathsf{P}(A \mid B) = \mathsf{P}(A \cap B)/\mathsf{P}(B)$. Очевидно, что функции $\xi_1(\omega)$ и $\xi_2(\omega)$ будут случайными величинами и на условном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$. Естественно различные законы распределения случайных величин ξ_1 и ξ_2, которые рассматриваются на пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$, называются *условными*. Например, если $F_{\xi_1}(x) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\})$ суть интегральная функция распределения некоторой случайной величины $\xi_1(\omega)$, то ее условная интегральная функция распределения относительно случайного события $B \in \mathcal{F}$, обозначаемая через $F_{\xi_1}(x|B)$, вычисляется по формуле

$$F_{\xi_1}(x|B) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\}|B) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\} \cap B)/\mathsf{P}(B) = \\ = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\ B\})/\mathsf{P}(B). \tag{7.17}$$

Можно сказать, что условной функцией распределения случайной величины $\xi_1(\omega)$ относительно события $B \in \mathcal{F}$ называется вероятность события $\{\omega\colon \xi_1(\omega) < x\}$ при условии, что событие B произошло.

Для дискретной случайной величины $\xi_1(\omega)$ при $\mathsf{P}(B) > 0$ естественно выполняется равенство

$$\mathsf{P}(\{\omega\colon \xi_1(\omega) = x_i'\}|B) = \mathsf{P}(\{\omega\colon \xi_1(\omega) = x_i',\ B\})/\mathsf{P}(B) \tag{7.18}$$

и, значит, $F_{\xi_1}(x|B) = (\mathsf{P}(B))^{-1} \sum_{i\colon x_i' < x} \mathsf{P}(\{\omega\colon \xi_1(\omega) = x_i',\ B\})$.

Семейство $\{\mathsf{P}(\{\omega\colon \xi_1(\omega) = x_i',\ B\})/\mathsf{P}(B)\colon i = 1, 2, \ldots\}$ называют *условным распределением дискретной случайной величины* $\xi_1(\omega)$ *относительно события* B. Приведем простой пример вычисления условного распределения дискретной случайной величины.

Пример 7.10. Подбрасывается несимметричная монета до первого выпадения герба (орла) с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Пусть асимметрия монеты и механизм ее подбрасывания таковы, что в каждом броске герб выпадает с вероятностью p, а решетка — с вероятностью $q = 1 - p$, и, наконец, появление бесконечного числа решеток является невозможным событием. Найти ряд распределения числа выпадений решеток, если было выполнено в опыте более k бросков.

Решение. Обозначим через ω_i при каждом фиксированном $i = 1, 2, \ldots$ описание такого элементарного исхода, когда $i - 1$ раз выпадает решетка, а последний раз выпадает герб. Очевидно, что для этого эксперимента достоверное событие $\Omega = \{\omega_1, \omega_2, \ldots\}$, $\mathcal{F} = \{A\colon A \subset \Omega\}$ — множество всех подмножеств множества Ω и, наконец, вероятность $\mathsf{P}(\{\omega_i\})$ появления произвольного элементарного исхода ω_i равна pq^{i-1}.

Если теперь ξ_1 определяет случайное число выпадений решеток в каждом таком опыте, то случайная величина $\xi_1(\omega_i) = i - 1$ при $i = 1, 2, \ldots$ и, следовательно, априорное распределение $\mathsf{P}(\{\omega\colon \xi_1(\omega) = i - 1\}) = \mathsf{P}(\{\omega_i\}) = pq^{i-1}$, $i = 1, 2, \ldots$. Такое распределение в теории вероятностей называется *геометрическим*. Обозначим через B событие, которое заключается в том, что в опыте было выполнено более k бросков. Тогда событие $B = \{\omega_{k+1}, \omega_{k+2}, \ldots\} = \{\omega\colon \xi_1(\omega) \geqslant k\}$ и $\mathsf{P}(B) = \sum_{i=k}^{\infty} pq^i = q^k$. Отсюда, используя соотношение (7.18), при $i \geqslant k+1$ получим условное распределение для ξ_1 в виде

$$\begin{aligned}\mathsf{P}(\{\omega\colon \xi_1(\omega) = i-1\} \mid \{\omega\colon \xi_1(\omega) \geqslant k\}) &= \\ = \mathsf{P}(\{\omega\colon \xi_1(\omega) = i-1,\ B\})/\mathsf{P}(B) &= \\ = \mathsf{P}(\{\omega_i\})/q^k = pq^{i-1}/q^k = pq^{i-1-k} &= \\ = \mathsf{P}(\{\omega\colon \xi_1(\omega) = i-1-k\}), \qquad i \geqslant k+1.&\end{aligned}$$

Следовательно, условная вероятность того, что число решеток окажется ровно $i-1$, если известно, что их было не менее k, совпадает с априорной вероятностью того, что число выпадений решеток будет в точности равно $i - 1 - k = 0, 1, \ldots$. Эта вероятность при заданном числе $i - 1 - k$ не зависит от k. Такое свойство распределения называется *отсутствием последействия*. На содержательном уровне можно сказать, что механизм подбрасывания монеты не запоминает число уже выпавших решеток.

Пусть теперь $\xi_1(\omega)$ — непрерывная случайная величина с плотностью распределения $f_{\xi_1}(x)$ и при $\mathsf{P}(B) > 0$ существует такая неотрицательная функция $g(x, B)\colon R \to R$, что $F_{\xi_1}(x \mid B) = \int_{-\infty}^{x} g(u, B)\, du$ для всех $x \in R$. Тогда почти всюду $d(F_{\xi_1}(x \mid B))/dx = g(x, B)$ по мере Лебега. Естественно, функцию $g(x, B)$ называют *условной плотностью распределения случайной величины* $\xi_1(\omega)$ относительно события B и обозначают через $f_{\xi_1}(x \mid B)$.

При $\mathsf{P}(B) > 0$ условная вероятность $\mathsf{P}(\cdot \mid B)$ на $\mathcal{F}$ обладает всеми свойствами априорной вероятности $\mathsf{P}(\cdot)$ на $\mathcal{F}$. Поэтому условные и безусловные законы распределения любой случайной величины имеют одинаковые свойства. При использовании условных законов распределения часто приходится решать следующие два конкретных вопроса.

Во-первых, в реальных задачах далеко не всегда известна совместная интегральная функция распределения $F(x, y)$ двумерного случайного вектора (ξ_1, ξ_2). Поэтому по результатам наблюдений возникает потребность непосредственного вычисления условной интегральной функции распределения $F_{\xi_1}(x \mid B_y)$, $B_y = \{\omega\colon \xi_2(\omega) < y\}$, которая позволяет определить $F(x, y)$ следующим образом:

$$F(x, y) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x, \xi_2(\omega) < y\}) =$$
$$= \mathsf{P}(\{\omega\colon \xi_2(\omega) < y\}) \times \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\} \mid \{\omega\colon \xi_2(\omega) < y\}) = F_{\xi_2}(y) \times F_{\xi_1}(x \mid B_y).$$

Во-вторых, определение условных законов распределения одной случайной величины при известных значениях другой случайной величины. Например, пусть $\xi_2(\omega) = \xi_1(\omega) + \eta(\omega)$, где ξ_2 — наблюдаемая случайная величина и η — ошибка измерения непрерывной случайной величины ξ_1. Итак, необходимо найти условные законы распределения измеряемой случайной величины ξ_1 при условии, что с помощью некоторого прибора наблюдалось значение y случайной величины ξ_2. В этом примере случайное событие $B_y = \{\omega\colon \xi_2(\omega) = y\}$. Перейдем к решению этой проблемы.

4.2. Условные законы распределения случайной величины относительно значений другой случайной величины. Основная трудность при определении и непосредственном вычислении условных законов распределения случайной величины ξ_1 относительно события B возникает в том случае, когда $\mathsf{P}(B) = 0$. Например, событие B может порождаться некоторой непрерывной случайной величиной ξ_2 с плотностью $f_{\xi_2}(y)$, т. е. случайное событие $B_y = \{\omega\colon \xi_2(\omega) = y\}$ и, значит, $\mathsf{P}(B_y) = 0$. Покажем, как можно преодолеть такого рода трудности.

Сначала для ξ_1 рассмотрим некоторую функцию на плоскости R^2 от переменных x и y, которую удобно обозначить через $F_{\xi_1}(x \mid y)$. Если ξ_2 содержит дискретную компоненту, то абсциссы точек разрыва интегральной функции $F_{\xi_2}(y)$ в промежутке $[a, y) \subset R$, занумерованные каким-нибудь способом, обозначим через $y_1, y_2, \ldots$. Для $i = 1, 2, \ldots$ рассмотрим разбиение $\Pi_i = \{\Pi_{i,1}, \Pi_{i,2}, \ldots, \Pi_{i,r(i)}\}$ промежутка $[a, y)$ с помощью чисел $a = v_{i,0} \leqslant v_{i,1} \leqslant \ldots \leqslant v_{i,r(i)-1} < v_{i,r(i)} = y$, среди которых всегда используются все числа $y_1, y_2, \ldots, y_i$. В частном случае, если промежуток $[a, y)$ содержит конечное число L точек разрыва функции $F_{\xi_2}(y)$, то при каждом $i \geqslant L$ разбиение Π_i включает все эти точки. Заметим, что при $i = 1, 2, \ldots$ тип каждого из промежутков $\Pi_{i,k}$, $k = 1, 2, \ldots, r(i)$, представлен в формуле (6.3). Далее будем предполагать, что последовательность разбиений вида Π_1, $\Pi_2, \ldots$ удовлетворяет условию $\lim\limits_{i\to\infty} \max\limits_{1\leqslant k\leqslant r(i)} (v_{i,k} - v_{i,k-1}) = 0$. Здесь и в дальнейшем такой класс $\Pi = \{\Pi_i\colon i = 1, 2, \ldots\}$ разбиений промежутка $[a, y) \subset R$ с использованием свойств интегральной функции $F_{\xi_2}(y)$ будем называть st-допустимым. Затем для такого класса разбиений примем следующее определение.

Определение 7.6. Функцию $F_{\xi_1}(x \mid y)$ будем называть *условной интегральной функцией распределения случайной величины* $\xi_1(\omega)$ *относительно значений другой случайной величины* ξ_2, если:

1) для каждого фиксированного $x \in R$ функция $F_{\xi_1}(x \mid y)$ от переменной y является борелевской и удовлетворяет соотношению

$$F(x,y) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x, \xi_2(\omega) < y\}) = \\ = \int\limits_{-\infty}^{y} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v), \qquad (x,y) \in R^2, \tag{7.19}$$

в котором интеграл Римана–Стилтьеса $\int_{-\infty}^{y} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v)$ вычисляется следующим способом:

$$\int\limits_{-\infty}^{y} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v) = \lim_{a \to -\infty} \int\limits_{a}^{y} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v) = \\ = \lim_{a \to -\infty} \lim_{i \to +\infty} \sum_{k=1}^{r(i)} F_{\xi_1}(x \mid v'_{i,k})\, \Delta F_{\xi_2}(\Pi_{i,k}),$$

где число $v'_{i,k} \in \Pi_{i,k}$ и $\Delta F_{\xi_2}(\Pi_{i,k})$ есть собственное приращение функции $F_{\xi_2}(y)$ на промежутке $\Pi_{i,k} \subset [\, a, y)$, определяемое по формуле (6.3);

2) при каждом фиксированном $y \in R$ функция $F_{\xi_1}(x \mid y)$ от переменной x удовлетворяет условиям теоремы 6.8 о выборочном вероятностном пространстве, т. е. является интегральной функцией распределения для некоторой случайной величины.

При фиксированном $y \in R$ и наблюдении события $B_y = \{\omega\colon \xi_2(\omega) = y\}$ функцию $F_{\xi_1}(x \mid y)$ от переменной x называют *условной интегральной функцией распределения случайной величины* ξ_1 при условии, что случайная величина $\xi_2(\omega)$ приняла значение $y \in R$. Если при $y_0 \in R$ событие $\{\omega\colon \xi_2(\omega) = y_0\}$ является невозможным, то эту терминологию естественно не применяют. Доказательство существования условной интегральной функции распределения $F_{\xi_1}(x \mid y)$ для любой случайной величины ξ_1 относительно значений любой другой случайной величины ξ_2 здесь проводиться не будет, так как оно основано на рассмотрении интеграла Лебега–Стилтьеса вместо интеграла Римана–Стилтьеса в соотношении (7.19) и на применении аппарата абстрактной теории меры. Однако если случайная величина ξ_2 дискретная или случайный вектор (ξ_1, ξ_2) обладает совместной плотностью распределения вероятностей, то легко найти некоторый вариант функции $F_{\xi_1}(x \mid y)$, которая удовлетворяет определению 7.6. Рассмотрим сначала случай, когда ξ_2 является дискретной случайной величиной.

Итак, пусть $Y_1 = \{y'_j\colon j = 1, 2, \ldots\}$ есть множество возможных значений дискретной случайной величины ξ_2, занумерованных в каком угодно порядке. Значит, вероятность $\mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\}) = q_j > 0$ для всех $j = 1, 2, \ldots$ и $q_1 + q_2 + \ldots = 1$. Вероятность $\mathsf{P}(\{\omega\colon \xi_2(\omega) = y\}) > 0$ при $y \in Y_1$. В этом случае условную интегральную функцию $F_{\xi_1}(x \mid y)$ вычисляем по формуле $F_{\xi_1}(x \mid y) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x\} \mid \{\omega\colon \xi_2(\omega) = y\}) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x, \xi_2(\omega) = y\}) \times (\mathsf{P}(\{\omega\colon \xi_2(\omega) = y\}))^{-1}$, $y \in Y_1$. Если $y \notin Y_1$, то либо случайная величина ξ_2 не принимает значения из множества $R \backslash Y_1$, либо $\mathsf{P}(\{\omega\colon \xi_2(\omega) \in R \backslash Y_1\}) = \mathsf{P}_{\xi_2}(R \backslash Y_1) = 0$. Поэтому при $y \in R \backslash Y_1$ условную интегральную функцию распределения

$F_{\xi_1}(x \mid y)$ можно задать с помощью произвольной интегральной функции распределения $F(x)$ на R. Например, можно принять, что функция $F_{\xi_1}(x \mid y) = F_{\xi_1}(x)$ при $y \in R \backslash Y_1$. Для определенности в дальнейшем будем рассматривать этот вариант условной интегральной функции распределения $F_{\xi_1}(x \mid y)$. Значит, функция $F_{\xi_1}(x \mid y)$ определяется неоднозначно. Ясно, что для каждого фиксированного $x \in R$ функция вида $F_{\xi_1}(x \mid y)$ от переменной y является борелевской и при каждом фиксированном $y \in R$ функция $F_{\xi_1}(x \mid y)$ является интегральной функцией распределения для случайной величины ξ_1. Наконец, докажем, что функция $F_{\xi_1}(x \mid y)$ от переменных x и y удовлетворяет соотношению (7.19). Так как $0 \leqslant F_{\xi_1}(x \mid y) \leqslant 1$ и случайная величина ξ_2 является дискретной с распределением $\{\mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\}) = q_j : y_j \in Y_1\}$, то интеграл Римана–Стилтьеса $\int_{-\infty}^{y} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v)$ в равенстве (7.19) легко преобразовать следующим образом:

$$\begin{aligned}\int\limits_{-\infty}^{y} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v) &= \sum_{j\,:\, y'_j < y} F_{\xi_1}(x \mid y'_j)\, \mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\}) = \\ &= \sum_{j\,:\, y'_j < y} (\mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\}))^{-1} \mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\, \xi_2(\omega) = y'_j\}) \times \\ &\qquad \times \mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\}) = \\ &= \sum_{j\,:\, y'_j < y} \mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\, \xi_2(\omega) = y'_j\}) = \\ &= \mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\, \xi_2(\omega) = y\}) = F(x, y).\end{aligned}$$

Пусть двумерная случайная величина (ξ_1, ξ_2) обладает совместной плотностью распределения вероятностей $f(x, y)$. Если $F(x, y)$ является двумерной интегральной функцией распределения вероятностей для непрерывного случайного вектора (ξ_1, ξ_2) и имеем событие $B_{y,\Delta y} = \{\omega\colon y \leqslant \xi_2(\omega) < y + \Delta y\}$, то $\mathsf{P}(B_{y,\Delta y}) = F_{\xi_2}(y + \Delta y) - F_{\xi_2}(y)$, $\mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\, y \leqslant \xi_2(\omega) < y + \Delta y\}) = F(x, y + \Delta y) - F(x, y)$. Предположим, что для $\Delta y > 0$ будет $\mathsf{P}(\{\omega\colon y \leqslant \xi_2(\omega) < y + \Delta y\}) > 0$. Отсюда, используя (7.17), получим, что при любом $\Delta y > 0$ функция

$$\begin{aligned}F_{\xi_1}(x \mid B_{y,\Delta y}) &= F_{\xi_1}(x \mid \{\omega\colon y \leqslant \xi_2(\omega) < y + \Delta y\}) = \\ &= \mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\, y \leqslant \xi_2(\omega) < y + \Delta y\}) \times \\ &\qquad \times (\mathsf{P}(\{\omega\colon y \leqslant \xi_2(\omega) < y + \Delta y\}))^{-1} = \\ &= (F(x, y + \Delta y) - F(x, y))/(F_{\xi_2}(y + \Delta y) - F_{\xi_2}(y)).\end{aligned}$$

Далее, рассматриваем такие значения y, для которых имеет место неравенство $f_{\xi_2}(y) \neq 0$. Значит, $f_{\xi_2}(y) \neq 0$ при $y \in R \backslash B_0$, где множества $B_0 = \{y\colon f_{\xi_2}(y) = 0\}$ и $R \backslash B_0$ являются борелевскими. При этих ограничениях легко найти предел $\lim\limits_{\Delta y \to 0} F_{\xi_1}(x \mid \{\omega\colon y \leqslant \xi_2(\omega) < y + \Delta y\})$, который существует почти для всех y по мере Лебега на $R \backslash B_0$. В самом деле, если этот предел обозначить через $F_{\xi_1}(x \mid y)$, то с учетом соотношения (7.14) и (7.15) имеем

$$\begin{aligned} F_{\xi_1}(x \mid y) &= \lim_{\Delta y \to 0}((F(x, y+\Delta y) - F(x,y))/(F_{\xi_2}(y+\Delta y) - F_{\xi_2}(y))) = \\ &= \lim_{\Delta y \to 0}(\{(F(x, y+\Delta y) - F(x,y))/\Delta y\} \times \\ &\qquad \times \{(F_{\xi_2}(y+\Delta y) - F_{\xi_2}(y))/\Delta y\}^{-1}) = \\ &= \frac{\partial}{\partial y}(F(x,y)) \times \Big(\frac{d}{dy}(F_{\xi_2}(y))\Big)^{-1} = \Big(\int\limits_{-\infty}^{x} f(u,y)\,du\Big)/f_{\xi_2}(y). \end{aligned}$$

Обозначим через B_1 множество тех точек из $R \backslash B_0$, для которых это соотношение нарушается, например, для таких точек предел может вовсе не существовать или не выполняется последнее равенство. Так как мера Лебега множества B_1 равна нулю, то можно положить, что

$$F_{\xi_1}(x \mid y) = \Big(\int\limits_{-\infty}^{x} f(u,y)\,du\Big)\Big/ f_{\xi_2}(y) = \Big(\int\limits_{-\infty}^{x} f(u,y)\,du\Big)\Big/ \int\limits_{-\infty}^{+\infty} f(u,y)\,du$$

при $y \in B_1$. С другой стороны, вероятность того, что случайная величина ξ_2 примет значения из любого промежутка $[c,d) \subset B_0$, равна $\mathsf{P}(\{\omega\colon c \leqslant \xi_2(\omega) < d\}) = \int_c^d f_{\xi_2}(y)\,dy = 0$. Это обстоятельство позволяет в дальнейшем задавать функцию $F_{\xi_1}(x \mid y)$ при $y \in B_0$ произвольным образом. Например, можно считать, что функция $F_{\xi_1}(x \mid y) = F_{\xi_1}(x)$ при $y \in B_0$. Ради простоты будем рассматривать именно такой вариант условной интегральной функции распределения $F_{\xi_1}(x \mid y)$. Следовательно, в общем случае функция $F_{\xi_1}(x \mid y)$ определяется неоднозначно. Имеет место следующее важное утверждение.

Лемма 7.2. *Функция* $F_{\xi_1}(x \mid y)$, *которая равна* $F_{\xi_1}(x)$ при $y \in B_0$ *и равна* $\Big(\int_{-\infty}^{+\infty} f(u,y)\,du\Big)^{-1} \int_{-\infty}^{x} f(u,y)\,du$ *при* $y \in R \backslash B_0$, *является условной интегральной функцией распределения случайной величины* $\xi_1(\omega)$ *относительно значений другой случайной величины* ξ_2.

Доказательство. Сначала рассмотрим свойства функции $F_{\xi_1}(x \mid y)$ от переменной y для каждого фиксированного $x \in R$. В силу существования повторных интегралов в соотношении (7.14) и неравенства $f_{\xi_2}(y) \neq 0$ при любом $y \in R \backslash B_0$ получаем, что функция $F_{\xi_1}(x \mid y) = \Big(\int_{-\infty}^{x} f(u,y)\,du\Big)/f_{\xi_2}(y)$ по переменной y, заданная на множестве $R \backslash B_0$, является почти всюду непрерывной относительно лебеговской меры в R. Очевидно, что при фиксированном значении $x \in R$ функция $F_{\xi_1}(x \mid y) = F_{\xi_1}(x)$ по y, определенная на множестве

B_0, является непрерывной. Тогда $F_{\xi_1}(x \mid y)$ как функция от y будет почти всюду непрерывной и, значит, борелевской на R.

Установим теперь, что функция $F_{\xi_1}(x \mid y)$ удовлетворяет соотношению (7.19). Случайная величина ξ_2 является непрерывной с плотностью распределения $f_{\xi_2}(y)$. Используя это, неравенство $f_{\xi_2}(v) \geqslant 0$ и существование при любых конечных значениях $a < y$ объектов $\int_a^y F_{\xi_1}(x \mid v)\,dv$, $\int_a^y f_{\xi_2}(v)\,dv$, интеграл Стилтьеса $\int_a^y F_{\xi_1}(x \mid v)\,dF_{\xi_2}(v)$ можно преобразовать в следующем виде:

$$\int_a^y F_{\xi_1}(x \mid v)\,dF_{\xi_2}(v) = \int_a^y F_{\xi_1}(x \mid v)\,f_{\xi_2}(v)\,dv, \quad -\infty < a < y.$$

Вычислим интеграл Римана $\int_a^y F_{\xi_1}(x \mid v)\,f_{\xi_2}(v)\,dv$ в правой части последнего равенства с помощью предела интегральных сумм, которые будем строить специальным образом. Для этого разобьем промежуток интегрирования $[a, y)$ на s равных частей. При $h = (y - a)/s$ обозначим через k номер частичного промежутка $[a + (k-1)h, a + kh)$, где $k = 1, 2, \ldots, s$. Распределим все частичные промежутки на два класса. К первому классу отнесем частичные промежутки, в каждом из которых функция $f_{\xi_2}(v)$ тождественно равна нулю. Ко второму классу отнесем все остальные частичные промежутки. Пусть I есть множество номеров частичных промежутков из первого класса и J — из второго класса. Следовательно, при $J \neq \varnothing$ в каждом частичном промежутке с номером $j \in J$ существует хотя бы одна точка с абсциссой v_j, для которой $f_{\xi_2}(v_j) \neq 0$. Любую такую и только такую точку будем использовать при составлении интегральных сумм. В соответствии с этим при любом $s \geqslant 1$ интегральную сумму $\sigma_s = h\sum_{k=1}^s F_{\xi_1}(x \mid v_k)\,f_{\xi_2}(v_k)$ интеграла Римана $\int_a^y F_{\xi_1}(x \mid v)\,f_{\xi_2}(v)\,dv$ можно разложить на две суммы и вычислить каждую следующим способом:

$$\begin{aligned}
\sigma_s &= h\sum_{k=1}^{s} F_{\xi_1}(x \mid v_k)\,f_{\xi_2}(v_k) = \\
&= h\sum_{i\in I} F_{\xi_1}(x \mid v_i)\,f_{\xi_2}(v_i) + h\sum_{j\in J} F_{\xi_1}(x \mid v_j)\,f_{\xi_2}(v_j) = \\
&= h\sum_{j\in J} F_{\xi_1}(x \mid v_j)\,f_{\xi_2}(v_j) = h\sum_{j\in J}\Big(f_{\xi_2}(v_j)\int_{-\infty}^{x} f(u, v_j)\,du\Big)\Big/f_{\xi_2}(v_j) = \\
&= h\sum_{j\in J}\Big(\int_{-\infty}^{x} f(u, v_j)\,du\Big).
\end{aligned}$$

Заметим, что первая сумма

$$h\sum_{k\in I} F_{\xi_1}(x \mid v_k)\,f_{\xi_2}(v_k) = h\sum_{i\in I} F_{\xi_1}(x)\,f_{\xi_2}(v_i) = 0,$$

так как $f_{\xi_2}(v_i) = 0$ для каждого $i \in I$. Преобразуем интеграл Римана $\int_a^y \Big(\int_{-\infty}^x f(u,v)\,du\Big)\,dv$ таким образом:

$$\int\limits_a^y \Big(\int\limits_{-\infty}^x f(u,v)\,du\Big)\,dv =$$
$$= \sum_{i\in I}\Big(\int\limits_{a+(i-1)h}^{a+ih}\Big(\int\limits_{-\infty}^x f(u,v)\,du\Big)\,dv\Big) + \sum_{j\in J}\Big(\int\limits_{a+(j-1)h}^{a+jh}\Big(\int\limits_{-\infty}^x f(u,v)\,du\Big)\,dv\Big).$$

Пусть теперь $g_k = \inf\Big\{\int_{-\infty}^x f(u,v)\,du\colon a+(k-1)h \leqslant v < a+kh)\Big\}$ и $G_k = \sup\Big\{\int_{-\infty}^x f(u,v)\,du\colon a+(k-1)h \leqslant v < a+kh)\Big\}$ при $k = 1, 2, \ldots, s$. Здесь g_k и G_k определяют точные нижнюю и верхнюю границы значений функции $\int_{-\infty}^x f(u,v)\,du$ от переменной v на промежутке $[\,a+(k-1)h, a+kh)$. Напомним, что величина $(G_k - g_k)$ вычисляет колебание функции $\int_{-\infty}^x f(u,v)\,du$ от переменной v на k-м частичном промежутке $[\,a+(k-1)h, a+kh)$. Так как

$$\sum_{i\in I}\Big(\int\limits_{a+(i-1)h}^{a+ih}\Big(\int\limits_{-\infty}^x f(u,v)\,du\Big)\,dv\Big) =$$
$$= \mathsf{P}\Big(\Big\{\omega\colon \xi_1(\omega) < x, \xi_2(\omega) \in \bigcup_{i\in I}\,[a+(i-1)\,h, a+ih)\Big\}\Big) \leqslant$$
$$\leqslant \mathsf{P}\Big(\Big\{\omega\colon \xi_2(\omega) \in \bigcup_{i\in I}\,[a+(i-1)\,h, a+ih)\Big\}\Big) =$$
$$= \sum_{i\in I}\Big(\int\limits_{a+(i-1)h}^{a+ih} f_{\xi_2}(v)\Big)\,dv = 0,$$
$$\sum_{j\in J}\Big(\int\limits_{a+(j-1)h}^{a+jh}\Big(\int\limits_{-\infty}^x f(u,v)\,du\Big)\,dv\Big) = \sum_{j\in J}\mu_j h,$$

где число μ_j выбирается из промежутка $[g_j, G_j]$ согласно теореме о среднем значении для интеграла $\int_{a+(j-1)h}^{a+jh}\Big(\int_{-\infty}^x f(u,v)\,du\Big)\,dv$, то имеет место равенство $\int_a^y\Big(\int_{-\infty}^x f(u,v)\,du\Big)\,dv = \sum_{j\in J}\mu_j h$. Теперь для разности интегральной суммы σ_s и повторного интеграла $\int_a^y\Big(\int_{-\infty}^x f(u,v)\,du\Big)\,dv$ получим

$$\sigma_s - \int\limits_a^y\Big(\int\limits_{-\infty}^x f(u,v)\,du\Big)\,dv = h\sum_{j\in J}\Big(\int\limits_{-\infty}^x f(u,v_j)\,du\Big) - h\sum_{j\in J}\mu_j =$$
$$= h\sum_{j\in J}\Big(\int\limits_{-\infty}^x f(u,v_j)\,du - \mu_j\Big).$$

Ясно, что для $j \in J$ будет $\left|\int_{-\infty}^{x} f(u, v_j)\, du - \mu_j\right| \leqslant G_j - g_j$. Отсюда для разности $\sigma_s - \int_a^y \Bigl(\int_{-\infty}^{x} f(u,v)\, du\Bigr)\, dv$ следует оценка

$$\left|\sigma_s - \int\limits_a^y \Bigl(\int\limits_{-\infty}^{x} f(u,v)\, du\Bigr)\, dv\right| = \left|h \sum_{j\in J} \Bigl(\int\limits_{-\infty}^{x} f(u, v_j)\, du - \mu_j\Bigr)\right| \leqslant$$
$$\leqslant \sum_{j\in J} (G_j - g_j)\, h = \sum_{k=1}^{s} (G_k - g_k)\, h.$$

Последняя сумма при $h \to 0$ или, что то же самое, при $s \to \infty$, стремится к нулю. Поэтому интегральная сумма σ_s при $s \to \infty$ стремится к интегралу $\int_a^y \Bigl(\int_{-\infty}^{x} f(u,v)\, du\Bigr)\, dv$, следовательно, $\int_a^y F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v) = \int_a^y \Bigl(\int_{-\infty}^{x} f(u,v)\, du\Bigr)\, dv$ при любых конечных значениях $a < y$. Отсюда, учитывая равенство (7.14), получаем соотношение (7.19). Итак, функция $F_{\xi_1}(x \mid y)$ удовлетворяет ограничению 1) определения 7.6.

Наконец, легко проверить, что при каждом фиксированном $y \in R$ функция $F_{\xi_1}(x \mid y)\colon R \cup \{-\infty\} \cup \{+\infty\} \to [0, 1]$, которая при $y \in R\backslash B_0$ равна $\Bigl(\int_{-\infty}^{+\infty} f(u,v)\, du\Bigr)^{-1} \int_{-\infty}^{x} f(u,v)\, du$ и равна $F_{\xi_1}(x)$ при $y \in B_0$, является неубывающей, непрерывной слева и $\lim\limits_{x\to+\infty} F_{\xi_1}(x \mid y) = 1$, $\lim\limits_{x\to-\infty} F_{\xi_1}(x \mid y) = 0$. Поэтому по теореме о выборочном вероятностном пространстве функция $F_{\xi_1}(x \mid y)$ будет интегральной функцией распределения для некоторой случайной величины. Лемма 7.2 доказана.

Итак, свойства функции $F_{\xi_1}(x \mid y)$ в частном случае, когда ξ_2 является дискретной или вектор (ξ_1, ξ_2) обладает совместной плотностью вероятности, показывают целесообразность определения 7.6 для произвольных случайных величин. Однако следует отметить, что в соответствии с определением 7.6 условная интегральная функция распределения $F_{\xi_1}(x \mid y)$ вычисляется неоднозначно. Другими словами, интегральная функция распределения $F_{\xi_1}(x \mid y)$ определяется однозначно лишь с точностью до некоторого подмножества B_0 действительной прямой R, для которого вероятность $\mathsf{P}(\{\omega\colon \xi_2(\omega) \in B_0\}) = \mathsf{P}_{\xi_2}(B_0)$ равна нулю.

Так как для непрерывной случайной величины (ξ_1, ξ_2) имеем

$$F_{\xi_1}(x \mid y) = \begin{cases} \Bigl(\int\limits_{-\infty}^{\infty} f(u,y)\, du\Bigr)^{-1} \int\limits_{-\infty}^{x} f(u,y)\, du & \text{при } \ y \in R\backslash B_0; \\ F_{\xi_1}(x) = \int\limits_{-\infty}^{x} f_{\xi_1}(u)\, du & \text{при } \ y \in B_0, \end{cases}$$

то легко определить условную плотность распределения случайной величины ξ_1 относительно другой случайной величины ξ_2:

$$f_{\xi_1}(x \mid y) = \begin{cases} f(x,y)/f_{\xi_2}(y) & \text{при } \ y \in R\backslash B_0; \\ f_{\xi_1}(x) & \text{при } \ y \in B_0, \end{cases} \tag{7.20}$$

Рассуждая аналогичным образом, можно получить следующие формулы для условных законов распределения случайной величины ξ_2 относительно другой случайной величины ξ_1:

$$F_{\xi_2}(y \mid x) = \begin{cases} \Bigl(\int\limits_{-\infty}^{\infty} f(x,v)\,dv\Bigr)^{-1} \int\limits_{-\infty}^{y} f(x,v)\,dv & \text{при } \ x \in R\backslash A_0; \\ F_{\xi_2}(y) = \int\limits_{-\infty}^{y} f_{\xi_2}(v)\,dv & \text{при } \ x \in A_0, \end{cases}$$

$$f_{\xi_2}(y \mid x) = \begin{cases} f(x,y)/f_{\xi_1}(x) & \text{при } \ x \in R\backslash A_0; \\ f_{\xi_2}(y) & \text{при } \ x \in A_0. \end{cases} \tag{7.21}$$

В этих равенствах множество $A_0 = \{x\colon f_{\xi_1}(x) = 0\}$ и вероятность того, что случайная величина ξ_1 примет значения из любого промежутка $[a,b) \subset A_0$, равна $\mathsf{P}(\{\omega\colon a \leqslant \xi_1(\omega) < b\}) = \int_a^b f_{\xi_1}(x)\,dx = 0$. Теперь непосредственно из (7.20) и (7.21) для всех значений переменных величин x и y получаем равенство $f(x,y) = f_{\xi_2}(y)\,f_{\xi_1}(x \mid y) = f_{\xi_1}(x)\,f_{\xi_2}(y \mid x)$, т. е. справедлива известная теорема умножения для плотностей. Здесь следует отметить, что если случайные величины независимы, то для $x, y \in R$ имеем $f(x,y) = f_{\xi_1}(x)\,f_{\xi_2}(y) = f_{\xi_1}(x)\,f_{\xi_2}(y \mid x) = f_{\xi_2}(y)\,f_{\xi_1}(x \mid y)$, $f_{\xi_1}(x \mid y) = f_{\xi_1}(x)$, $f_{\xi_2}(y \mid x) = f_{\xi_2}(y)$.

Пример 7.11. В качестве применения полученных формул для условных законов распределения продолжим изучение свойств двумерной случайной величины (ξ_1, ξ_2) из примера 7.9. Напомним, что случайная величина (ξ_1, ξ_2) определяет среднюю точку трещины в задаче о гончарном круге радиуса r или центр хорды в задаче Бертрана. Используя вид функций $f(u,v)$, $f_{\xi_1}(u)$, $f_{\xi_2}(v)$ из примера 7.9, получим, что при $|v| < r$ условная плотность $f_{\xi_1}(u \mid v) = f(u,v)/f_{\xi_2}(v) = (1/\pi r^2)\bigl(2\sqrt{r^2 - v^2}\,/\pi r^2\bigr)^{-1} = 1/(2\sqrt{r^2 - v^2}\,)$ для всех $|u| < \sqrt{r^2 - v^2}$ и $f_{\xi_1}(u \mid v) = 0$ для всех $|u| \geqslant \sqrt{r^2 - v^2}$. Следовательно, условное распределение случайной величины ξ_1 является равномерным в интервале $(-\sqrt{r^2 - v^2}, \sqrt{r^2 - v^2}\,)$. Например при $r > 1$ и $v = \sqrt{r^2 - 1}$ график условной плотности $f_{\xi_1}(u \mid v)$ представлен на рис. 7.9.

Если $|v| > r$, то в соответствии с предыдущим соглашением условная плотность $f_{\xi_1}(u \mid v) = f_{\xi_1}(u)$, так как $f_{\xi_2}(v) = 0$ при $|v| > r$. Наконец, при $|v| = r$ значение случайной величины ξ_1 полностью опреде-

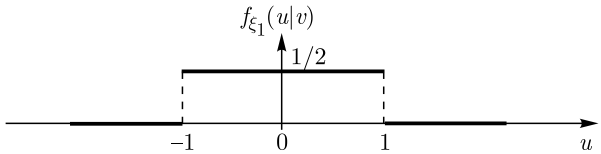


Рис. 7.9

лено и равно нулю. Более того, при $|v| = r$ плотность распределения $f_{\xi_2}(v) = 0$. Поэтому естественно считать, что условное распределение для случайной величины ξ_1 становится вырожденным при $u = 0$.

Аналогичным способом находим, что при $|u| < r$ условная плотность $f_{\xi_2}(v \mid u) = f(u,v)/f_{\xi_1}(u) = (1/\pi r^2)\big(2\sqrt{r^2-u^2}/\pi r^2\big)^{-1} = 1/(2\sqrt{r^2-u^2})$ для всех $|v| < \sqrt{r^2-u^2}$ и $f_{\xi_2}(v \mid u) = 0$ для всех $|v| \geqslant \sqrt{r^2-u^2}$. Так как $f_{\xi_1}(u) = 0$ при $|u| > r$, то $f_{\xi_2}(v \mid u) = f_{\xi_2}(v)$ для всех $|u| > r$. Далее, при $|u| = r$ значение случайной величины ξ_2 равно нулю и $f_{\xi_1}(u) = 0$. Следовательно, можно положить, что условное распределение для случайной величины ξ_2 является вырожденным при $v = 0$. Отметим важное обстоятельство. Поскольку $\{\omega\colon |\xi_1(\omega)| > r\} = \{\omega\colon |\xi_2(\omega)| > r\} = \varnothing$, то можно не определять условную плотность $f_{\xi_1}(u \mid v)$ для случайной величины ξ_1 при $|v| > r$ и условную плотность $f_{\xi_2}(v \mid u)$ для случайной величины ξ_2 при $|u| > r$.

4.3. Формулы полной вероятности и Байеса для несчетного числа гипотез. Напомним, что формулы полной вероятности и Байеса имеют вид $\mathsf{P}(A) = \sum_{k=1}^{\infty} \mathsf{P}(H_k)\,\mathsf{P}(A \mid H_k)$, $\mathsf{P}(H_i \mid A) = \mathsf{P}(H_i)\,\mathsf{P}(A \mid H_i)/\sum_{k=1}^{\infty} \mathsf{P}(H_k)\,\mathsf{P}(A \mid H_k)$ в случае, когда множество гипотез счетно или конечно. Однако понятие условной вероятности относительно событий, которые порождаются непрерывной случайной величиной, позволяет обобщить формулу полной вероятности и формулу Байеса на случай, когда множество гипотез имеет континуальную мощность. Пусть гипотезами являются возможные значения некоторой случайной величины ξ непрерывного типа с заданной плотностью $f(x)$, т. е. $H_x = \{\omega\colon \xi(\omega) = x\}$, $x \in R$, и $\mathsf{P}(H_x) = F(x+0) - F(x) = 0$.

Лемма 7.3. *Множества H_x, $x \in R$, образуют полную группу попарно несовместимых событий.*

Доказательство. Событие, которое заключается в том, что случайная величина ξ примет какое-либо действительное значение x, является достоверным. Поэтому имеем очевидное равенство: $\cup_{-\infty<x<+\infty} H_x = \Omega$. Второе утверждение леммы докажем от противного. Пусть оно неверно, и, следовательно, $H_x \cap H_u \neq \varnothing$ при $x \neq u$. Тогда найдется такое описание $\omega_0 \in \Omega$, что $\omega_0 \in H_x$ и $\omega_0 \in H_u$. Из определения множеств H_x и H_u следует, что $\xi(\omega_0) = x$ и $\xi(\omega_0) = u$.

Так как случайная величина $\xi(\omega)$ является однозначной функцией, то необходимо следует, что $x = u$. Получаем противоречие, так как по условию $x \neq u$. Лемма 7.3 доказана.

Итак, мы рассматриваем несчетное число попарно несовместных событий H_x, $x \in R$, которые образуют полную группу и для каждого из которых имеет место $\mathsf{P}(H_x) = 0$. При этом нельзя находить условную вероятность $\mathsf{P}(A \mid H_x)$ по известной формуле $\mathsf{P}(A \mid H_x) = \mathsf{P}(A \cap H_x)/\mathsf{P}(H_x)$, так как $\mathsf{P}(A \cap H_x) = 0$ и $\mathsf{P}(H_x) = 0$. В этом заключается существенное отличие данной ситуации от классического случая. По определению полагаем, что $\mathsf{P}(A \mid H_x) = \mathsf{P}(A \mid x) = \lim\limits_{\Delta x \to 0} \mathsf{P}(A \mid \{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\})$. Далее, вычислим этот предел следующим способом:

$$\begin{aligned}
&\lim_{\Delta x \to 0} \mathsf{P}(A \mid \{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\}) = \\
&= \lim_{\Delta x \to 0} (\mathsf{P}(A \cap \{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\})/\mathsf{P}(\{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\})) = \\
&= (\lim_{\Delta x \to 0} \mathsf{P}(A \cap \{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\})/\Delta x) \times \\
&\qquad \times (\lim_{\Delta x \to 0} \mathsf{P}(\{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\})/\Delta x)^{-1} = \\
&= (\lim_{\Delta x \to 0} \mathsf{P}(A \cap \{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\})/\Delta x)/f(x).
\end{aligned}$$

Итак, в случае существования пределов имеем

$$\mathsf{P}(A \mid x) = (\lim_{\Delta x \to 0} \mathsf{P}(A \cap \{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\})/\Delta x)/f(x).$$

Отсюда непосредственно находим, что

$$\mathsf{P}(A \cap \{\omega\colon x \leqslant \xi(\omega) < x + \Delta x\}) = \mathsf{P}(A \mid x')f(x')\,\Delta x + o(\Delta x), \qquad (7.22)$$

где $x' \in [\,x, x + \Delta x)$. В частности, можем положить $x' = x$.

Теорема 7.4 (о полной вероятности для несчетного числа гипотез). *Если* $\xi(\omega)$ *—непрерывная случайная величина с плотностью распределения* $f(x)$, *событие* $A \in \mathcal{F}$ *и* $\mathsf{P}(A \mid x)$ *— условная вероятность события* A *относительно значения* x *случайной величины* ξ, *тогда справедлива следующая формула полной вероятности*: $\mathsf{P}(A) = \int_{-\infty}^{+\infty} \mathsf{P}(A \mid x)\, f(x)\, dx$.

Доказательство. Ради упрощений записи предположим, что $\xi(\omega)$ — непрерывная случайная величина, принимающая значения из интервала (a, b) при любом $\omega \in \Omega$, где $-\infty \leqslant a < b \leqslant +\infty$. При этом $f(x) \geqslant 0$ для всех $x \in (a, b)$ и $f(x) = 0$ для $x \notin (a, b)$. Разобьем промежуток вида (a, b) на s частей следующим образом: $\Delta a_1 = a_1 - a_0$, $\Delta a_2 = a_2 - a_1, \ldots,$ $\Delta a_s = a_s - a_{s-1}$, где $a_0 = a$, $a_s = b$, $a_i \in (a, b)$,

$i = 1, 2, \ldots, s-1$. Введем для каждого значения $i = 1, 2, \ldots, s$ гипотезы $H_i = \{\omega\colon a_{i-1} \leqslant \xi(\omega) < a_i\}$. Так как $\cup_{i=1}^{s} H_i = \Omega$ и $H_i \cap H_j = \varnothing$ при $i \neq j$, то множество $\{H_i\colon i = 1, 2, \ldots, s\}$ образует полную группу попарно несовместимых событий. Значит, для $A \in \mathcal{F}$ выполняются равенства

$$A = A \cap \Omega = A \bigcap \Bigl(\bigcup_{i=1}^{s} H_i\Bigr) = \bigcup_{i=1}^{s} (A \cap H_i), \quad \mathsf{P}(A) = \sum_{i=1}^{s} \mathsf{P}(H_i \cap A).$$

Используя теперь при $i = 1, 2, \ldots, s$ и $a_i' \in [a_{i-1}, a_{i-1} + \Delta a_i)$ выражение (7.22), последовательно находим:

$$\mathsf{P}(A) = \sum_{i=1}^{s} (\mathsf{P}(A \mid a_i')\, f(a_i')\, \Delta a_i + o(\Delta a_i)),$$

$$\mathsf{P}(A) = \lim_{\substack{\Delta a_i \to 0\\ i=1,2,\ldots,s}} \Bigl(\sum_{i=1}^{s} (\mathsf{P}(A \mid a_i')\, f(a_i')\, \Delta a_i + o(\Delta a_i))\Bigr) = \int\limits_a^b \mathsf{P}(A \mid x)\, f(x)\, dx.$$

При $a = -\infty$ и $b = +\infty$, рассуждая аналогичным способом, получим, что $\mathsf{P}(A) = \int_{-\infty}^{+\infty} \mathsf{P}(A \mid x)\, f(x)\, dx$, и теорема доказана.

Теорема 7.5 (Байеса для условной плотности распределения). *Пусть выполняются все условия предыдущей теоремы и* $\mathsf{P}(A) > 0$. *Тогда для условной плотности* $f(x \mid A)$ *непрерывной случайной величины* ξ *имеет место равенство* $f(x \mid A) = (\mathsf{P}(A \mid x)\, f(x)) / \int_{-\infty}^{+\infty} \mathsf{P}(A \mid x)\, f(x)\, dx$.

Доказательство. Из определений условной плотности $f(x \mid A)$ и условной интегральной функции $F(x \mid A)$ последовательно имеем

$$\begin{aligned} f(x \mid A) &= d(F(x \mid A))/dx = \lim_{\Delta x \to 0} (F(x + \Delta x \mid A) - F(x \mid A))/\Delta x = \\ &= \lim_{\Delta x \to 0} (\mathsf{P}(\{\omega\colon \xi(\omega) < x + \Delta x\} \mid A) - \mathsf{P}(\{\omega\colon \xi(\omega) < x\} \mid A))/\Delta x = \\ &= \lim_{\Delta x \to 0} (\mathsf{P}(\{\omega\colon \xi(\omega) < x + \Delta x\} \cap A) - \\ &\qquad - \mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap A))/(\mathsf{P}(A)\Delta x). \end{aligned} \tag{7.23}$$

Нетрудно видеть, что справедливо соотношение вида

$$\{\omega\colon \xi < x + \Delta x\} = \{\omega\colon \xi < x\} \cup \{\omega\colon x \leqslant \xi < x + \Delta x\},$$

причем события $\{\omega\colon \xi < x\}$ и $\{\omega\colon x \leqslant \xi < x + \Delta x\}$ несовместны. Тогда

$$\{\omega\colon \xi < x + \Delta x\} \cap A = (\{\omega\colon \xi < x\} \cap A) \cup (\{\omega\colon x \leqslant \xi < x + \Delta x\} \cap A),$$

$$(\{\omega\colon \xi < x\} \cap A) \cap (\{\omega\colon x \leqslant \xi < x + \Delta x\} \cap A) = \varnothing,$$

$$\mathsf{P}(\{\omega\colon \xi < x + \Delta x\} \cap A) - \mathsf{P}(\{\omega\colon \xi < x\} \cap A) = \\ = \mathsf{P}(\{\omega\colon x \leqslant \xi < x + \Delta x\} \cap A).$$

Отсюда, используя (7.23) и (7.22), при $x' \in [\,x, x + \Delta x)$ имеем

$$f(x \mid A) = \lim_{\Delta x \to 0} (\mathsf{P}(\{\omega\colon x \leqslant \xi < x + \Delta x\} \cap A)/(\mathsf{P}(A)\Delta x)) = \\ = \lim_{\Delta x \to 0} (\mathsf{P}(A \mid x')\, f(x')\Delta x + o(\Delta x))/(\mathsf{P}(A)\Delta x) = \mathsf{P}(A \mid x)\, f(x)/\mathsf{P}(A).$$

Применяя теперь формулу полной вероятности для вероятности $\mathsf{P}(A)$ в случае несчетного числа гипотез, находим

$$f(x \mid A) = \mathsf{P}(A \mid x)\, f(x)/\mathsf{P}(A) = (\mathsf{P}(A \mid x)\, f(x)) \Big/ \int\limits_{-\infty}^{+\infty} \mathsf{P}(A \mid x)\, f(x)\, dx,$$

что и требовалось доказать.

Пример 7.12. Определить вероятность обнаружения космического тела, если расстояние до него, измеряемое в километрах, есть равномерная случайная величина, принимающая значения в интервале $(100, 200)$. При этом известно, что если расстояние до космического тела равно x, то вероятность его обнаружения равна $3000/x^2$.

Решение. Пусть $\xi(\omega)$ — расстояние от станции слежения до объекта. По условию задачи это непрерывная случайная величина с плотностью распределения вида

$$f(x) = \begin{cases} 0, & x \notin (100, 200), \\ 1/100, & x \in (100, 200). \end{cases}$$

Обозначим через A случайное событие, заключающееся в том, что станция слежения обнаружит объект. Из условия задачи $\mathsf{P}(A \mid \{\omega\colon \xi(\omega) = x\}) = \mathsf{P}(A \mid x) = 3000/x^2$. Используя теперь формулу полной вероятности для несчетного числа гипотез, легко подсчитаем вероятность события A, а именно:

$$\mathsf{P}(A) = \int\limits_{100}^{200} \mathsf{P}(A \mid x)\, f(x)\, dx = \int\limits_{100}^{200} (3000/x^2) \times (1/100)\, dx = 0{,}15.$$

Краткий обзор

В главе 7 на примерах поясняется целесообразность рассмотрения совокупности измерителей случайного эксперимента. Семейство измерителей количественно характеризуют статистически устойчивый эксперимент E с различных точек зрения. Рассмотрены различные случайные величины на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$.

Математическими моделями семейства измерителей являются многомерные случайные величины или случайные векторы. Дана геометрическая интерпретация многомерной случайной величины. Вводится понятие многомерной интегральной функции распределения и изучаются ее свойства. Показано, что многомерная интегральная функция распределения определяет вероятностные закономерности случайных векторов. Определены необходимые и достаточные условия существования многомерной интегральной функции распределения. Установлена связь между одномерными и многомерными законами распределения. Приводится методика построения выборочного вероятностного пространства для многомерной случайной величины.

Обсуждается проблема независимости семейства случайных величин. Доказаны теоремы о независимости случайных величин и, тем самым, предложены несколько простых и доступных критериев для проверки независимости случайных величин. Подробно изучены наиболее часто встречающиеся в конкретных задачах дискретные и непрерывные многомерные случайные величины. На основе унифицированной вероятностной модели условного эксперимента дано обоснование рассмотрения условных законов распределения случайных величин. Особое внимание уделено получению формул для условных законов распределения некоторой случайной величины относительно значений другой случайной величины. Доказана теорема, которая обобщает формулу полной вероятности для несчетного числа гипотез. Установлена формула для условной плотности распределения непрерывной случайной величины.

Контрольные вопросы и упражнения

1. Установить связь между различными определениями случайного вектора.
2. Доказать предельные свойства многомерной интегральной функции распределения.
3. Построить выборочное вероятностное пространство для случайных двумерных величин, интегральная функция которых равна $F(x, y)$.
4. Перечислить всевозможные законы распределения дискретных и непрерывных двумерных случайных величин.
5. Доказать теорему умножения для двумерной плотности вероятности.
6. Привести пример, когда из причинной независимости случайных величин следует их статистическая независимость.

7. Привести пример, когда случайные величины ξ и η являются статистически независимыми, хотя между ними существует функциональная зависимость.
8. Два игрока, независимо друг от друга, случайным образом бросают по одному разу игральную кость. Найти законы распределения системы двух случайных величин ξ_1 и ξ_2, которые определяют число выпадений очков для первого и второго игроков соответственно.
9. Одномерные случайные величины ξ_1 и ξ_2 независимы и равномерно распределены на промежутке $(0, 1]$ и на отрезке $[0, 1]$ соответственно. Пусть случайные величины ξ_1 и ξ_2 принимают значения на оси абсцисс $0x$ и на оси ординат $0y$ плоскости $x0y$ с перпендикулярными осями. Найти закон распределения случайного вектора $\xi = (\xi_1, \xi_2)$.
10. Упорядоченная система $\xi = (\xi_1, \xi_2)$ из одномерных случайных величин ξ_1 и ξ_2 равномерно распределена в некоторой области C^2 на плоскости $x_1 0 x_2$ с перпендикулярными осями и с правой ориентацией. При этом замкнутая область C^2 есть квадрат, стороны которого составляют углы $\pi/4$ с осями координат и равны по длине единице. Показать, что величины ξ_1 и ξ_2 зависимы.
11. Пусть плотность распределения $f(x, y)$ непрерывной двумерной случайной величины $\xi = (\xi_1, \xi_2)$ равна $\lambda^2 \exp\{-\lambda(x+y)\}$ при $x > 0$, $y > 0$ и равна нулю в противном случае, где $\lambda = \mathrm{const} > 0$. Найти законы распределения случайных величин ξ_1 и ξ_2.
12. Подбрасывается симметричная монета до первого выпадения герба (орла) с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Пусть появление бесконечного числа решеток является невозможным событием. Найти ряд распределения числа выпадений решеток, если было выполнено в опыте более k бросков.

Глава 8

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ИЗМЕРИТЕЛЕЙ ИСХОДОВ СЛУЧАЙНЫХ ЭКСПЕРИМЕНТОВ

Постановка задачи на содержательном уровне. Выполняя с помощью различных приборов измерения над элементарными исходами некоторого статистически устойчивого эксперимента, исследователь получает набор чисел. Математическим описанием этого набора чисел в простейшем случае является одномерная случайная величина, или в более сложном варианте — упорядоченное семейство из одномерных случайных величин. Случайная величина, как и ее различные законы распределения (ряд, функция, плотность и т. д.) являются функциональными количественными характеристиками статистически устойчивого эксперимента. Напомним, что случайная величина содержит более богатую информацию о свойствах статистически устойчивого эксперимента, чем ее законы распределения. С другой стороны, законы распределения полностью определяют случайную величину с вероятностной точки зрения. Например, знание одного из указанных законов распределения позволяет находить вероятность того, что показание прибора будет принадлежать некоторому отрезку на действительной прямой. К сожалению, на практике нам чаще всего случайная величина $\xi(\omega)\colon \Omega \to R$ и ее функция распределения $F(x)\colon R \cup \{-\infty\} \cup \{+\infty\} \to [0, 1]$ неизвестны. Если пространство Ω из описаний элементарных исходов содержит большое или бесконечное число элементов, то воссоздать случайную величину и ее интегральную функцию $F(x)$ непосредственно из экспериментальных наблюдений удается с большими трудностями и далеко не всегда. Более того, эксперимент E может быть настолько сложен, что не ставится даже вопрос о поточечном восстановлении случайной величины и ее функции распределения.

Далее, случайная величина $\xi(\omega)$ и ее закон распределения $F(x)$ в общем случае представляют довольно сложные функциональные характеристики измерителей исходов статистически устойчивого эксперимента. Наконец, в ряде случаев требуется знать только общее интегральное представление о результатах измерений. Суммируя все это, можно утверждать, что при решении многих практических задач очень часто нет возможности или в более простом случае нет необходимости исчерпывающим образом характеризовать количественные свойства элементарных исходов статистически устойчивого эксперимента с помощью случайных величин и их законов распределения. Поэтому уже давно стали применять простые и удобные, но вместе с тем и достаточно информативные параметрические характеристики. Эти числовые параметры некоторым существенным образом характеризуют специальные и суммарные свойства статистически устойчивого эксперимента E. В информационном отношении числовые характеристики беднее, чем случайная величина и ее интегральная функция распределения, но зато значительно проще. Числовые характеристики случайной величины отражают наиболее важные особенности ее законов распределения. К числовым характеристикам случайной величины относятся, прежде всего, такие важные константы, как математическое ожидание (среднее значение), квантили, медиана, мода, дисперсия, среднее квадратическое отклонение (стандарт), среднее отклонение, моменты, коэффициент асимметрии, эксцесс и т. д.

§ 1. Характеристики положения одномерных случайных величин

1.1. Математическое ожидание простой случайной величины. Сначала рассмотрим пример, который поясняет на интуитивном уровне понятие о математическом ожидании или о среднем значении случайных величин, в особенности случайных величин с конечным числом значений. Случайная величина с конечным множеством $\{x'_1, x'_2, \ldots, x'_m\}$ ее значений называется *простой*. Пусть простая случайная величина $\xi(\omega)$ определена на вероятностной модели $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ эксперимента E. Проведем теперь достаточно большое число N независимых испытаний эксперимента E, в каждом из которых случайная величина ξ может принять одно из значений x'_1, $x'_2, \ldots$, x'_m с вероятностями p_1, $p_2, \ldots$, p_m соответственно. В результате этих испытаний, согласно частотному методу приближенного вычисления вероятности каждого случайного события $A_k = \{\omega\colon \xi(\omega) = x_k\}$, получим, что $p_k = \mathsf{P}(\{\omega\colon \xi(\omega) = x'_k\}) \approx \approx \mu(A_k, N)/N$, $k = 1, 2, \ldots, m$. Здесь число $\mu(A_k, N)$ показывает, сколько раз наблюдали при N испытаниях случайное событие A_k, или сколько раз случайная величина ξ приняла значение x_k. Ясно, что величина $[\,\mu(A_1, N)\, x'_1 + \mu(A_2, N)\, x'_2 + \ldots + \mu(A_m, N)\, x'_m\,]/N$ определяет среднее арифметическое значений, принятых случайной величиной ξ в данной серии из N опытов. Отсюда непосредственно имеем $\sum_{k=1}^{m} \mu(A_k, N)\, x'_k/N \approx \sum_{k=1}^{m} N p_k x'_k/N = \sum_{k=1}^{m} x'_k p_k$.

Так как при увеличении числа N независимых испытаний эксперимента относительная частота $\mu(A_k, N)/N$ стремится стабилизироваться около вероятности p_k, то значение среднего арифметического будут группироваться около числа $\sum_{k=1}^{m} x'_k p_k$. На интуитивном уровне можно сказать, что между средним арифметическим наблюденных значений случайной величины ξ при увеличении числа N независимых испытаний эксперимента E и постоянным числом $\sum_{k=1}^{m} x'_k p_k$ существует такого же типа связь, как между относительной частотой и вероятностью случайного события. Математическое описание характера этой связи будет дано позднее в разделе о сходимости последовательности случайных величин.

Рассмотрим теперь несколько другую интерпретацию постоянного числа $\sum_{k=1}^{m} x'_k p_k$. Представим себе, что в m точках на оси абсцисс с координатами $(x'_1, 0)$, $(x'_2, 0), \ldots$, $(x'_m, 0)$ сосредоточены массы некоторого вещества в количестве p_1, $p_2, \ldots$, p_m соответственно. В механике такой способ распределения количества $\sum_{k=1}^{m} p_k$ некоторого вещества называется *дискретным*. В классической механике координата $x_{\text{ц}}$ центра масс таких материальных точек с учетом равенства $\sum_{k=1}^{m} p_k = 1$ вычисляется с помощью формулы $x_{\text{ц}} = \sum_{k=1}^{m} x'_k p_k / \sum_{k=1}^{m} p_k = \sum_{k=1}^{m} x'_k p_k$. Как известно, координата $x_{\text{ц}}$ центра масс является суммарной или интегральной характеристикой положения всей системы дискретного размещения материальных точек на прямой.

Из последних двух абзацев следует, что особое число $\sum_{k=1}^{m} x'_k p_k$ может служить грубой или ориентировочной оценкой положения значений случайной величины ξ на действительной оси. Все это оправдывает и делает понятным следующее определение.

Определение 8.1. *Математическим ожиданием* $\mathsf{M}(\xi)$ или *средним значением* простой случайной величины называется число

$$\mathsf{M}(\xi) = \sum_{k=1}^{m} x'_k p_k. \tag{8.1}$$

Пример 8.1. Бросаем несимметричную монету один раз. Герб появляется с вероятностью p, $0 < p < 1$. Случайная величина $\xi(\omega)$ показывает, сколько раз при одном испытании выпадает герб. Случайная величина $\xi(\omega)$ принимает значения из множества $\{0, 1\}$. Вероятности тех или иных значений для случайной величины $\xi(\omega)$, очевидно, вычисляются следующим образом: $\mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) = q$, $\mathsf{P}(\{\omega\colon \xi(\omega) = 1\}) = p$, $p + q = 1$. Тогда, по определению математического ожидания для такой простой случайной величины, получаем $\mathsf{M}(\xi) = 0 \times q + 1 \times p = p$.

Часто в обозначении $\mathsf{M}(\xi)$ опускают круглые скобки и пишут $\mathsf{M}\xi$, если это не приводит к очевидным недоразумениям и ошибкам. Непосредственно из формулы (8.1) следует, что значения x'_1, $x'_2, \ldots$, x'_m случайной величины ξ располагаются на действительной числовой оси как справа от точки $\mathsf{M}\xi$, так и слева. Причем случайная величина ξ, как правило, принимает реже те значения, которые располагаются дальше от точки $\mathsf{M}\xi$, и принимает чаще те значения, которые располагаются ближе к точке $\mathsf{M}\xi$.

Итак, математическое ожидание можно рассматривать как некоторое характерное для данной случайной величины постоянное число, как центр, относительно которого происходит рассеивание случайной величины. Поэтому математическое ожидание иногда называют центром распределения случайной величины. Таков содержательный смысл математического ожидания. Следует также отметить, что случайная величина может и не принимать значения, равного математическому ожиданию. Например, простая случайная величина $I_A(\omega)$, которая является индикатором некоторого события A, принимает значение ноль с вероятностью $1-p$ и будет равна единице с вероятностью $p = \mathsf{P}(A)$, где $0 < p < 1$. Математическое ожидание или среднее значение $\mathsf{M}(I_A(\omega))$ для такой случайной величины вычисляется просто и равно $0 \times (1-p) + 1 \times p = p = \mathsf{P}(A)$, однако не существует такого элемента $\omega \in \Omega$, что $I_A(\omega) = p$.

Для любой простой случайной величины $\xi(\omega)$ имеет место очевидное равенство: $\xi(\omega) = \sum_{k=1}^{m} x'_k I_{A_k}(\omega)$, где все числа x'_1, $x'_2, \ldots$, x'_m различны и событие $A_k = \{\omega\colon \xi(\omega) = x'_k\}$ при $k = 1, 2, \ldots, m$. Пусть теперь имеем любое другое представление: $\xi(\omega) = \sum_{i=1}^{s} x_{i1} I_{B_i}(\omega)$ случайной величины $\xi(\omega)$, в котором B_1, $B_2, \ldots$, B_s образуют полную группу попарно несовместимых событий, а среди x_{11}, $x_{21}, \ldots$, x_{s1} могут быть равные числа. Ясно, что при $\omega \in B_i$ случайная величина $\xi(\omega) = x_{i1}$. Из свойств событий A_1, $A_2, \ldots$, A_m и B_1, $B_2, \ldots$, B_s получаем, что $A_k = \sum_{i=1}^{s}(A_k \cap B_i)$, $B_i = \sum_{k=1}^{m}(B_i \cap A_k)$ и $\xi(\omega) = x'_k = x_{i1}$ при $\omega \in A_k \cap B_i$. Отсюда, используя равенство (8.1), имеем:

$$\begin{aligned}\mathsf{M}(\xi) &= \sum_{k=1}^{m} x'_k p_k = \sum_{k=1}^{m} x'_k \sum_{i=1}^{s} \mathsf{P}(A_k \cap B_i) = \sum_{k=1}^{m}\sum_{i=1}^{s} x'_k \mathsf{P}(A_k \cap B_i) = \\ &= \sum_{k=1}^{m}\sum_{i=1}^{s} x_{i1}\mathsf{P}(A_k \cap B_i) = \sum_{i=1}^{s} x_{i1} \sum_{k=1}^{m} \mathsf{P}(A_k \cap B_i) = \sum_{i=1}^{s} x_{i1}\mathsf{P}(B_i).\end{aligned}$$

Окончательно получим, что

$$\mathsf{M}\xi = \sum_{i=1}^{s} x_{i1}\mathsf{P}(B_i). \tag{8.2}$$

Итак, получили новую формулу для вычисления математического ожидания простой случайной величины ξ. Формула (8.2) часто используется для доказательства следующих свойств математического ожидания.

М.1. *Математическое ожидание неслучайной величины* $\xi(\omega) \equiv c$, $\omega \in \Omega$, *равно постоянной* c, так как в этом случае $c = cI_\Omega(\omega)$ и $\mathsf{M}c = c$ $\mathsf{P}(\Omega) = c$.

М.2. *Для любого $A \in \mathcal{F}$ математическое ожидание $I_A(\omega)$ равно вероятности* $\mathsf{P}(A)$, так как $I_A(\omega) = 0 \times I_{\overline{A}}(\omega) + 1 \times I_A(\omega)$ и $\mathsf{M}(I_A(\omega)) = \mathsf{P}(A)$.

М.3. *Постоянный множитель c можно выносить за символ математического ожидания*, так как $c\xi = \sum_{k=1}^{m} c\,x'_k I_{A_k}(\omega)$ и $\mathsf{M}(c\xi) = \sum_{k=1}^{m} c\,x'_k \mathsf{P}(A_k) = c\,\mathsf{M}\xi$.

М.4. *Математическое ожидание от неотрицательной случайной величины является неотрицательным числом*. Действительно, так как $\xi(\omega) \geqslant 0$ при $\omega \in \Omega$ и, значит, $x'_1 \geqslant 0$, $x'_2 \geqslant 0, \ldots, x'_m \geqslant 0$, то $\mathsf{M}\xi = \sum_{k=1}^{m} x'_k \mathsf{P}(A_k) \geqslant 0$.

М.5. *Если c и d суть постоянные, то* $\mathsf{M}(c\xi + d\eta) = c\,\mathsf{M}\xi + d\,\mathsf{M}\eta$. В самом деле, представим случайные величины ξ и η в виде: $\xi(\omega) = \sum_{i=1}^{s} x_{i1} I_{B_i}(\omega)$, $\eta(\omega) = \sum_{j=1}^{r} y_{j1} I_{C_j}(\omega)$. Заметим, что $C_1, C_2, \ldots, C_r$ составляют полную группу попарно несовместимых событий, а среди $y_{11}, y_{21}, \ldots, y_{r1}$ могут быть равные числа. Множество $\{B_i \cap C_j : i = \overline{1, s};\ j = \overline{1, r}\}$ образует полную группу попарно несовместимых случайных событий и $c\,\xi(\omega) + d\eta(\omega) = c\,x_{i1} + d\,y_{j1}$ для всех $\omega \in B_i \cap C_j$. Применяя теперь представление

$$c\,\xi(\omega) + d\eta(\omega) = \sum_{i=1}^{s} \sum_{j=1}^{r} (c\,x_{i1} + d\,y_{j1})\, I_{B_i \cap C_j}(\omega)$$

и формулу (8.2), легко найдем

$$\begin{aligned}\mathsf{M}(c\xi(\omega) + d\eta(\omega)) &= \sum_{i=1}^{s} \sum_{j=1}^{r} (c\,x_{i1} + d\,y_{j1})\, \mathsf{P}(B_i \cap C_j) = \\ &= c \sum_{i=1}^{s} x_{i1} \sum_{j=1}^{r} \mathsf{P}(B_i \cap C_j) + d \sum_{j=1}^{r} y_{j1} \sum_{i=1}^{s} \mathsf{P}(B_i \cap C_j) = \\ &= c \sum_{i=1}^{s} x_{i1} \mathsf{P}(B_i) + d \sum_{j=1}^{r} y_{j1} \mathsf{P}(C_j) = c\,\mathsf{M}\xi + d\,\mathsf{M}\eta.\end{aligned}$$

М.6. Пусть $\xi \geqslant \eta$ при $\omega \in \Omega$, тогда $\mathsf{M}\xi \geqslant \mathsf{M}\eta$. Действительно, при $\xi - \eta \geqslant 0$ и $\xi = \eta + (\xi - \eta)$, используя свойства М.4 и М.5, получим, что $\mathsf{M}\xi = \mathsf{M}\eta + \mathsf{M}(\xi - \eta) \geqslant \mathsf{M}\eta$.

М.7. *Математическое ожидание от произведения независимых случайных величин равно произведению их математических ожиданий, т.е. имеет место равенство*: $\mathsf{M}(\xi \times \eta) = \mathsf{M}\xi \times \mathsf{M}\eta$. Для доказательства этого свойства введем события $A_k = \{\omega\colon \xi(\omega) = x'_k\}$ и $D_j = \{\omega\colon \eta(\omega) = y'_j\}$, где $x'_1, x'_2, \ldots, x'_m$ — все различные значения простой случайной величины ξ и $y'_1, y'_2, \ldots, y'_s$ — все различные значения простой случайной величины η. Поэтому $\xi(\omega) = \sum_{k=1}^{m} x'_k I_{A_k}(\omega)$ и $\eta(\omega) = \sum_{j=1}^{s} y'_j I_{D_j}(\omega)$. Так как величины ξ и η независимы, то при любых $k = 1, 2, \ldots, m$ и $j = 1, 2, \ldots, s$ случайные события $A_k = \{\omega\colon \xi(\omega) = x'_k\}$ и $D_j = \{\omega\colon \eta(\omega) = y'_j\}$ являются независимыми, т. е. $\mathsf{P}(A_k \cap D_j) = \mathsf{P}(A_k)\,\mathsf{P}(D_j)$. Принимая во внимание все это, очевидное при всех $\omega \in A_k \cap D_j$ равенство $\xi(\omega) \times \eta(\omega) = x'_k \times y'_j$ и (8.2), получим, что $\xi(\omega)\,\eta(\omega) = \sum_{k=1}^{m}\sum_{j=1}^{s} x'_k y'_j I_{A_k \cap D_j}(\omega)$,

$$\mathsf{M}(\xi(\omega)\,\eta(\omega)) = \sum_{k=1}^{m}\sum_{j=1}^{s} x'_k y'_j \mathsf{P}(A_k \cap D_j) = \\ = \sum_{k=1}^{m} x'_k \mathsf{P}(A_k) \times \sum_{j=1}^{s} y'_j \mathsf{P}(D_j) = \mathsf{M}\xi \times \mathsf{M}\eta,$$

и свойство М.7 установлено.

М.8. *Для случайной величины* ξ *имеет место неравенство* $\mathsf{M}|\xi| \geqslant |\mathsf{M}\xi|$, так как $|\mathsf{M}\xi| = \left|\sum_{k=1}^{m} x'_k \mathsf{P}(A_k)\right| \leqslant \sum_{k=1}^{m} |x'_k|\,\mathsf{P}(A_k) = \mathsf{M}|\xi|$ с учетом равенства (8.2).

1.2. Математическое ожидание произвольной случайной величины. Рассмотрим сначала определение математического ожидания произвольной неотрицательной случайной величины $\xi(\omega)$.

Лемма 8.1. *Пусть случайная величина* $\xi(\omega) \geqslant 0$. *Тогда найдется такая последовательность* $\{\xi_i(\omega); i = 1, 2, \ldots\}$ *из простых случайных величин, что* $0 \leqslant \xi_1 \leqslant \xi_2 \ldots \leqslant \xi$ *и* $\lim\limits_{i\to\infty} \xi_i = \xi$ *для всех* $\omega \in \Omega$.

Доказательство. При фиксированных $i = 1, 2, \ldots$ и $k = 1, 2, \ldots, i2^i$ рассмотрим события $A_{k,i} = \{\omega\colon (k-1)/2^i \leqslant \xi(\omega) < k/2^i\}$, $A_i = \{\omega\colon \xi(\omega) \geqslant i\}$. Для каждого заданного i семейство $\{A_i\} \cup \{A_{k,i}\colon k = 1, 2, \ldots, i2^i\}$ образует полную группу из конечного числа $(i2^i + 1)$ попарно несовместимых событий. Это семейство порождает простую случайную величину

$$\xi_i(\omega) = \sum_{k=1}^{i\,2^i} (k-1)\,2^{-i} I_{A_{k,i}}(\omega) + i\, I_{A_i}(\omega), \qquad \omega \in \Omega. \tag{8.3}$$

Все значения $0, 1/2^i, \ldots, i$ случайной величины ξ_i совпадают с абсциссами точек деления отрезка $[0, i]$ на $i2^i$ равных частей. Из (8.3) и соотношений $\{\omega\colon (k-1)/2^i \leqslant \xi(\omega) < k/2^i\} = \{\omega\colon 2(k-1)/2^{i+1} \leqslant \xi(\omega) < (2k-1)/2^{i+1}\} \cup \{\omega\colon (2k-1)/2^{i+1} \leqslant \xi(\omega) < 2k/2^{i+1}\}$, $A_i = \{\omega\colon i\,2^{i+1}/2^{i+1} \leqslant \xi(\omega) < (i+1)\,2^{i+1}/2^{i+1}\} \cup A_{i+1}$ сразу следует, что $0 \leqslant \xi_i \leqslant \xi_{i+1} \leqslant \xi$ при всех $i = 1, 2, \ldots$. Поскольку для каждого фиксированного $\omega \in \Omega$ значение $\xi(\omega)$ есть действительное число, то существует такое натуральное число i, что $i \geqslant \xi(\omega)$. Из определения по формуле (8.3) случайной величины ξ_i вытекает, что при всех $i \geqslant \xi(\omega)$ имеет место неравенство $0 \leqslant \xi - \xi_i \leqslant 1/2^i$. Следовательно, для всех $\omega \in \Omega$ имеется неубывающая последовательность $\{\xi_i; i = 1, 2, \ldots\}$ из неотрицательных простых случайных величин, сходящаяся к $\xi(\omega)$. Последовательность $\{\eta_j; j = 1, 2, \ldots\}$, которая обладает указанными в лемме 8.1 свойствами, будем называть *аппроксимирующей для* $\xi(\omega)$. Теперь с каждой неотрицательной случайной величиной $\xi(\omega)$ можно связать некоторую аппроксимирующую последовательность $\{\eta_j; j = 1, 2, \ldots\}$, в частности, аппроксимирующую последовательность $\{\xi_i; i = 1, 2, \ldots\}$, элементы которой определяются формулой (8.3).

Лемма 8.2. *Если* $\{\xi_i; i = 1, 2, \ldots\}$ и $\{\eta_j; j = 1, 2, \ldots\}$ *две любые аппроксимирующие неотрицательную случайную величину* ξ *последовательности, то имеет место предельное равенство* $\lim\limits_{i\to\infty} \mathsf{M}\xi_i = \lim\limits_{j\to\infty} \mathsf{M}\eta_j$.

Доказательство. Из определения $\{\eta_j; j = 1, 2, \ldots\}$ следует, что $\eta_j \leqslant \xi$ для любого j. Выберем любое $\varepsilon > 0$ и определим событие $B_i = \{\omega \colon \xi_i < \eta_j(\omega) - \varepsilon\}$. Так как $\xi_1 \leqslant \xi_2 \leqslant \ldots$, то $B_1 \supset B_2 \supset \ldots$ и $\lim\limits_{i\to\infty} B_i = \cap_{i=1}^{\infty} B_i$. Ясно, что $\cap_{i=1}^{\infty} B_i = \varnothing$. Если допустить обратное, то существует такое описание $\omega_0 \in \Omega$, что $\xi_i(\omega_0) < \eta_j(\omega_0) - \varepsilon$ для $i = 1, 2, \ldots$. Отсюда получаем противоречие вида: $\xi(\omega_0) = \lim\limits_{i\to\infty} \xi_i(\omega_0) \leqslant \eta_j(\omega_0) - \varepsilon$. Из непрерывности снизу вероятностной функции находим, что $\lim\limits_{i\to\infty} \mathsf{P}(B_i) = 0$. Поскольку $\xi_i \geqslant 0$ и $\xi_i = \xi_i I_{\overline{B_i}} + \xi_i I_{B_i}$, то $\xi_i \geqslant \xi_i I_{\overline{B_i}} \geqslant (\eta_j - \varepsilon)\, I_{\overline{B_i}} = \eta_j - \eta_j I_{B_i} - \varepsilon I_{\overline{B_i}}$. Применяя свойства 6°, 5° и 2° математических ожиданий, легко получим, что $\mathsf{M}\xi_i \geqslant \mathsf{M}(\eta_j - \eta_j I_{B_i} - \varepsilon I_{\overline{B_i}}) \geqslant \mathsf{M}\eta_j - b_j\, \mathsf{P}(B_i)\varepsilon\, \mathsf{P}(\overline{B_i}) \geqslant \mathsf{M}\eta_j - \varepsilon - b_j\, \mathsf{P}(B_i)$, где $b_j = \max\{\eta_j(\omega) \colon \omega \in \Omega\}$. Итак, при любых фиксированных $\varepsilon > 0$ и $j \geqslant 1$ сразу для всех $i \geqslant 1$ имеет место неравенство $\mathsf{M}\xi_i \geqslant \mathsf{M}\eta_j - \varepsilon - b_j\, \mathsf{P}(B_i)$. Для любой аппроксимирующей последовательности $\{\xi_i; i = 1, 2, \ldots\}$ справедливы неравенства $0 \leqslant \mathsf{M}\xi_i \leqslant \mathsf{M}\xi_{i+1}$. Тогда существует предел вида $\lim\limits_{i\to\infty} \mathsf{M}\xi_i \geqslant 0$, который может, иногда, принимать значение $+\infty$. Отсюда, используя соотношение $\lim\limits_{i\to\infty} \mathsf{P}(B_i) = 0$, выводим, что $\lim\limits_{i\to\infty} \mathsf{M}\xi_i \geqslant \mathsf{M}\eta_j$ для всех $j \geqslant 1$ и, следовательно, имеет место $\lim\limits_{i\to\infty} \mathsf{M}\xi_i \geqslant \lim\limits_{j\to\infty} \mathsf{M}\eta_j$. Поскольку последовательности $\{\xi_i; i \geqslant 1\}$ и $\{\eta_j; j \geqslant 1\}$ в лемму 8.2 входят симметрично, то их можно поменять местами. Это обстоятельство позволяет получить неравенство $\lim\limits_{j\to\infty} \mathsf{M}\eta_j \geqslant \lim\limits_{i\to\infty} \mathsf{M}\xi_i$, и, значит, лемма 8.2 доказана.

Согласно лемме 8.2 предел $\lim\limits_{i\to\infty} \mathsf{M}\xi_i$ не зависит от выбора аппроксимирующей последовательности $\{\xi_i; i = 1, 2, \ldots\}$ для ξ. Поэтому определение математического ожидания для любой неотрицательной случайной величины ξ с помощью равенства $\mathsf{M}\xi = \lim\limits_{i\to\infty} \mathsf{M}\xi_i$ является корректным. Если теперь ξ — произвольная случайная величина, а случайные величины ξ^+ и ξ^- соответственно равны $\max\{0, \xi\}$

и $\max\{0, -\xi\}$, то верны очевидные соотношения $\xi = \xi^+ - \xi^-$, $\xi^+ \geqslant 0$, $\xi^- \geqslant 0$, и математическое ожидание в общем случае можно определить следующим образом.

Определение 8.2. Пусть $\min\{\mathsf{M}(\xi^+), \mathsf{M}(\xi^-)\} < \infty$, другими словами хоть одна из величин $\mathsf{M}(\xi^+)$ или $\mathsf{M}(\xi^-)$ конечна, тогда для любой случайной величины ξ полагаем $\mathsf{M}\xi = \mathsf{M}(\xi^+) - \mathsf{M}(\xi^-)$. В противном случае говорят, что математическое ожидание случайной величины ξ не существует.

Из определения 8.2 следует, что $\mathsf{M}\xi = \infty$, если $\mathsf{M}(\xi^+) = \infty$ и $\mathsf{M}(\xi^-) < \infty$, и $\mathsf{M}\xi = -\infty$, если $\mathsf{M}(\xi^+) < \infty$ и $\mathsf{M}(\xi^-) = \infty$. Пусть $\max\{\mathsf{M}(\xi^+), \mathsf{M}(\xi^-)\} < \infty$, то $|\mathsf{M}\xi| < \mathsf{M}(\xi^+) + \mathsf{M}(\xi^-) < \infty$, или, другими словами, $\mathsf{M}\xi$ будет конечным.

Итак, математическое ожидание $\mathsf{M}\xi$ определяется в три этапа:

1) сначала математическое ожидание определяется для простых случайных величин ξ по формуле (8.1): $\mathsf{M}(\xi) = \sum_{k=1}^{m} x'_k p_k$;

2) потом математическое ожидание определяется для неотрицательных случайных величин ξ с помощью предельного перехода $\mathsf{M}\xi = \lim\limits_{i\to\infty} \mathsf{M}\xi_i$, где $\{\xi_i; i \geqslant 1\}$ является аппроксимирующей последовательностью для $\xi \geqslant 0$;

3) и, наконец, математическое ожидание определяется для произвольной случайной величины ξ с использованием равенства $\mathsf{M}\xi = \mathsf{M}(\xi^+) - \mathsf{M}(\xi^-)$.

Такой способ позволяет распространить все свойства $\mathsf{M}\xi$, которые были доказаны для простых случайных величин, и на произвольные случайные величины.

1.3. Свойства математического ожидания. Установим теперь основные свойства $\mathsf{M}\xi$ для произвольных случайных величин.

Теорема 8.1. *Если математическое ожидание* $\mathsf{M}\xi$ *существует и* c *есть некоторая постоянная, то* $\mathsf{M}(c\xi) = c\mathsf{M}\xi$.

Доказательство. Предположим сначала, что $\xi \geqslant 0$, $c \geqslant 0$ и $\{\xi_i;\ i = 1, 2, \ldots\}$ является аппроксимирующей последовательностью для ξ. Из определения аппроксимирующей последовательности $\{\xi_i;$ $i = 1, 2, \ldots\}$ следует, что $0 \leqslant c\xi_1 \leqslant c\xi_2 \leqslant \ldots \leqslant c\xi$ и $\lim\limits_{i\to\infty}(c\xi_i) = c\lim\limits_{i\to\infty}\xi_i = c\xi$ для всех $\omega \in \Omega$. Так как для каждого $i = 1, 2, \ldots$ случайная величина $c\xi_i$ является простой, то последовательность $\{c\xi_i;$ $i = 1, 2, \ldots\}$ будет аппроксимирующей для $c\xi \geqslant 0$. Используя доказанное равенство $\mathsf{M}(c\xi_i) = c\mathsf{M}\xi_i$ для простых случайных величин, определение математического ожидания для неотрицательной случайной величины $c\xi$ и свойства пределов, последовательно найдем: $\mathsf{M}(c\xi) = \lim\limits_{i\to\infty}\mathsf{M}(c\xi_i) = c\lim\limits_{i\to\infty}\mathsf{M}\xi_i = c\mathsf{M}\xi$. Итак, для неотрицательной случайной величины ξ и для $c \geqslant 0$ утверждение теоремы доказано. Если теперь ξ является произвольной случайной величиной и по-прежнему $c \geqslant 0$, то для случайной величины типа $c\xi$ имеем: $(c\xi)^+ = c\xi^+$, $(c\xi)^- = c\xi^-$, $\mathsf{M}(c\xi) = \mathsf{M}((c\xi)^+) - \mathsf{M}((c\xi)^-) = \mathsf{M}(c\xi^+) - \mathsf{M}(c\xi^-) = c\,\mathsf{M}(\xi^+) - c\,\mathsf{M}(\xi^-) = c\,(\mathsf{M}(\xi^+) - \mathsf{M}(\xi^-)) = c\,\mathsf{M}(\xi)$. Наконец, пусть ξ является произвольной случайной величиной и $c < 0$, тогда для случайной величины $c\xi$ получим: $(c\xi)^+ = -c\xi^-$, $(c\xi)^- = -c\xi^+$. Отсюда $\mathsf{M}(c\xi) = \mathsf{M}((c\xi)^+) - \mathsf{M}((c\xi)^-) = \mathsf{M}(-c\xi^-) - \mathsf{M}(-c\xi^+) = -c\,\mathsf{M}(\xi^-) + c\,\mathsf{M}(\xi^+) = c\,(\mathsf{M}(\xi^+) - \mathsf{M}(\xi^-)) = c\,\mathsf{M}(\xi)$.

Лемма 8.3. *Пусть ξ и η неотрицательные случайные величины, тогда имеет место равенство $\mathsf{M}(\xi + \eta) = \mathsf{M}\xi + \mathsf{M}\eta$.*

Доказательство. Обозначим через $\{\xi_i; i = 1, 2, \ldots\}$ и $\{\eta_i; i = 1, 2, \ldots\}$ две аппроксимирующие последовательности для случайных величин ξ и η соответственно. Используя определение аппроксимирующей последовательности, свойства простых случайных величин, свойство аддитивности операции предельного перехода, определение математического ожидания для неотрицательных случайных величин и свойство аддитивности математического ожидания от простых случайных величин, получим, что $0 \leqslant \xi_i + \eta_i \leqslant \xi_{i+1} + \eta_{i+1} \leqslant \xi + \eta$, $\lim\limits_{i\to\infty} (\xi_i + \eta_i) = \lim\limits_{i\to\infty} \xi_i + \lim\limits_{i\to\infty} \eta_i = \xi + \eta$, $\mathsf{M}(\xi + \eta) = \lim\limits_{i\to\infty} \mathsf{M}(\xi_i + \eta_i) = \lim\limits_{i\to\infty} \mathsf{M}\xi_i + \lim\limits_{i\to\infty} \mathsf{M}\eta_i = \mathsf{M}\xi + \mathsf{M}\eta$.

Лемма 8.4. *Для того чтобы математическое ожидание $\mathsf{M}\xi$ случайной величины ξ было конечным, необходимо и достаточно, чтобы $\mathsf{M}|\xi| < \infty$.*

Доказательство. Если математическое ожидание $\mathsf{M}\xi$ конечно, т. е. $|\mathsf{M}\xi| < \infty$, то $0 \leqslant \mathsf{M}(\xi^+) < \infty$ и $0 \leqslant \mathsf{M}(\xi^-) < \infty$. Отсюда, используя представление $|\xi| = \xi^+ + \xi^-$, неравенства $\xi^+ \geqslant 0$, $\xi^- \geqslant 0$ и лемму 8.3, находим, что величина $\mathsf{M}|\xi| = \mathsf{M}(\xi^+) + \mathsf{M}(\xi^-) < \infty$. Пусть теперь $\mathsf{M}\xi < \infty$. Поскольку имеет место соотношение $\mathsf{M}|\xi| = \mathsf{M}(\xi^+) + \mathsf{M}(\xi^-) < \infty$, то $\mathsf{M}(\xi^+) < \infty$, $\mathsf{M}(\xi^-) < \infty$, и, значит, $|\mathsf{M}\xi| = |\mathsf{M}(\xi^+) - \mathsf{M}(\xi^-)| \leqslant \mathsf{M}(\xi^+) + \mathsf{M}(\xi^-) < \infty$. Лемма 8.4 доказана.

Из доказательства первой части леммы 8.4 следует, что при $\mathsf{M}|\xi| < \infty$ имеем $|\mathsf{M}\xi| = |\mathsf{M}(\xi^+) - \mathsf{M}(\xi^-)| \leqslant \mathsf{M}(\xi^+) + \mathsf{M}(\xi^-) = \mathsf{M}|\xi|$. Если $\mathsf{M}|\xi| = \infty$, то либо $\mathsf{M}(\xi^+) = \infty$ и $\mathsf{M}(\xi^-) < \infty$, либо $\mathsf{M}(\xi^+) < \infty$ и $\mathsf{M}(\xi^-) = \infty$. Поэтому $\mathsf{M}|\xi| = \mathsf{M}(\xi^+) + \mathsf{M}(\xi^-) = \infty$. Таким образом, если $\mathsf{M}\xi$ существует, то $|\mathsf{M}\xi| \leqslant \mathsf{M}|\xi|$.

Лемма 8.5. *Если $\mathsf{M}\xi$, $\mathsf{M}\eta$ существуют и $\xi \geqslant \eta$* при $\omega \in \Omega$, *то $\mathsf{M}\xi \geqslant \mathsf{M}\eta$.*

Доказательство. Проведем доказательство этого утверждения в четыре этапа.

Во-первых, рассмотрим случай, когда $\xi \geqslant 0$ и $\eta \geqslant 0$, и, следовательно, $\mathsf{M}\xi$ и $\mathsf{M}\eta$ всегда определены. Поэтому при $\xi \geqslant 0$ и $\eta \geqslant 0$ можно оставить только второе условие этой леммы. Используя эти ограничения, разложение $\xi = \eta + (\xi - \eta)$, неравенство $\xi - \eta \geqslant 0$ и лемму 8.3, получим, что $\mathsf{M}\xi = \mathsf{M}\eta + \mathsf{M}(\xi - \eta) \geqslant \mathsf{M}\eta$.

Во-вторых, если $\mathsf{M}\xi$, $\mathsf{M}\eta$ существуют и $\mathsf{M}\xi = \infty$ или $\mathsf{M}\eta = -\infty$, то с учетом только этого ограничения леммы и свойств символов ∞ и $-\infty$ непосредственно получаем $\mathsf{M}\xi \geqslant \mathsf{M}\eta$.

Далее, если потребовать только существование $\mathsf{M}\eta$ и выполнение дополнительного условия $\mathsf{M}\eta > -\infty$, то $\mathsf{M}\eta = \mathsf{M}(\eta^+) - \mathsf{M}(\eta^-) > -\infty$ и $\mathsf{M}\eta^- < \infty$. Так как $\xi \geqslant \eta$, то $\xi^+ \geqslant \eta^+ \geqslant 0$, $\eta^- \geqslant \xi^- \geqslant 0$ и получаем $\mathsf{M}(\eta^+) \geqslant 0$, $\mathsf{M}(\eta^-) \geqslant \mathsf{M}(\xi^-) \geqslant 0$, $\mathsf{M}(\xi^-) < \infty$. В силу этого имеем следующую цепочку соотношений: $-\infty < \mathsf{M}\eta = \mathsf{M}(\eta^+) - \mathsf{M}(\eta^-) \leqslant \leqslant \mathsf{M}(\xi^+) - \mathsf{M}(\xi^-) = \mathsf{M}(\xi)$, т. е. математическое ожидание $\mathsf{M}(\xi)$ будет определено, $\mathsf{M}(\xi) > -\infty$ и опять имеем $\mathsf{M}\xi \geqslant \mathsf{M}\eta$. Ясно, что при рассмотрении этого случая достаточно предположить только факт существования $\mathsf{M}\eta$ и выполнение второго условия $\xi \geqslant \eta$.

Наконец, пусть $\mathsf{M}\xi$ существует и дополнительно $\mathsf{M}\xi < \infty$. Тогда $\mathsf{M}(\xi) = \mathsf{M}(\xi^+) - \mathsf{M}(\xi^-) < \infty$ и, следовательно, $\mathsf{M}(\xi^+) < \infty$. Так как $\xi \geqslant \eta$, то, как и в предыдущем случае, последовательно имеем: $\xi^+ \geqslant \geqslant \eta^+ \geqslant 0$, $\eta^- \geqslant \xi^- \geqslant 0$, $\infty > \mathsf{M}(\xi^+) \geqslant \mathsf{M}(\eta^+) \geqslant 0$, $\mathsf{M}(\eta^-) \geqslant \mathsf{M}(\xi^-) \geqslant 0$, $\infty > \mathsf{M}\xi = \mathsf{M}(\xi^+) - \mathsf{M}(\xi^-) \geqslant \mathsf{M}(\eta^+) - \mathsf{M}(\eta^-) = \mathsf{M}\eta$, или, окончательно, математическое ожидание $\mathsf{M}\eta$ существует, $\mathsf{M}\eta < \infty$ и $\mathsf{M}\xi \geqslant \mathsf{M}\eta$. Отметим, что в последнем варианте мы не предполагали факт существования $\mathsf{M}\eta$. Лемма 8.5 доказана.

Лемма 8.6. *Пусть* ξ и η *являются неотрицательными случайными величинами, для которых* $\min\{\mathsf{M}(\xi), \mathsf{M}(\xi)\} < \infty$, *т. е. выражение* $\mathsf{M}\xi - \mathsf{M}\eta$ *определено. Тогда* $\mathsf{M}(\xi - \eta) = \mathsf{M}\xi - \mathsf{M}\eta$.

Доказательство. Обозначим через θ случайную величину вида $\xi - \eta$. Так как $\theta = \xi - \eta$, $\xi \geqslant 0$ и $\eta \geqslant 0$, то $\theta^+ \leqslant \xi$ и $\theta^- \leqslant \eta$. Используя этот факт и предыдущую лемму, находим, что $\mathsf{M}\theta^+ \leqslant \mathsf{M}\xi$ и $\mathsf{M}\theta^- \leqslant \mathsf{M}\eta$. Следовательно, $\min\{\mathsf{M}\theta^+, \mathsf{M}\theta^-\} \leqslant \min\{\mathsf{M}\xi, \mathsf{M}\eta\} < +\infty$ и $\mathsf{M}\theta = \mathsf{M}(\xi - \eta)$ существует. Поскольку $\theta = \xi - \eta$ и $\theta = \theta^+ - \theta^-$, то $\theta^+ + \eta = \theta^- + \xi$. Применяя теперь лемму 8.3, имеем $\mathsf{M}\theta^+ + \mathsf{M}\eta = \mathsf{M}\theta^- + \mathsf{M}\xi$, откуда $\mathsf{M}\theta^+ - \mathsf{M}\theta^- = \mathsf{M}\xi - \mathsf{M}\eta$. Отсюда получаем $\mathsf{M}\theta = \mathsf{M}(\xi - \eta) = \mathsf{M}\xi - \mathsf{M}\eta$. Лемма 8.6 доказана.

Лемма 8.7. *Если* ξ *— неотрицательная случайная величина и* $\mathsf{M}\xi = 0$, *то* $\mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) = 1$ *или почти всюду* ξ *равна нулю.*

Доказательство. Для каждого фиксированного $i = 1, 2, \ldots$ введем случайную величину $\eta_i(\omega)$, которая равна i^{-1} при $\omega \in \{\omega\colon \xi(\omega) \geqslant \geqslant i^{-1}\}$ и равна нулю при $\omega \in \{\omega\colon \xi(\omega) < i^{-1}\}$. Очевидно, что $\xi \geqslant \eta_i$

и случайная величина η_i принимает два значения: ноль и i^{-1}. Используя определение математического ожидания для простой случайной величины η_i и лемму 8.5, найдем: $\mathsf{M}\eta_i = i^{-1}\mathsf{P}(\{\omega\colon \eta_i(\omega) = i^{-1}\}) = i^{-1}\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant i^{-1}\}) \leqslant \mathsf{M}\xi = 0$ для всех $i = 1, 2, \ldots$. Отсюда получаем, что вероятности $\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant i^{-1}\}) = 0$ и $\mathsf{P}(\{\omega\colon \xi(\omega) < i^{-1}\}) = 1$ для каждого $i = 1, 2, \ldots$. Так как случайное событие $\{\omega\colon \xi(\omega) < (i+1)^{-1}\} \subset \{\omega\colon \xi(\omega) < i^{-1}\}$ при любом $i = 1, 2, \ldots$ и событие вида $\{\omega\colon \xi(\omega) \leqslant 0\} = \cap_{i=1}^{\infty}\{\omega\colon \xi(\omega) < i^{-1}\}$, то вероятность $\mathsf{P}(\{\omega\colon \xi(\omega) \leqslant 0\}) = \mathsf{P}(\cap_{i=1}^{\infty}\{\omega\colon \xi(\omega) < i^{-1}\}) = \mathsf{P}\big(\lim\limits_{i\to\infty}\{\omega\colon \xi(\omega) < i^{-1}\}\big) = \lim\limits_{i\to\infty}\mathsf{P}(\{\omega\colon \xi(\omega) < i^{-1}\})$. Из очевидных равенств $\{\omega\colon \xi(\omega) \leqslant 0\} = \{\omega\colon \xi(\omega) < 0\} \cup \{\omega\colon \xi(\omega) = 0\}$, $\{\omega\colon \xi(\omega) < 0\} = \varnothing$ непосредственно получим утверждение леммы 8.7.

Теорема 8.2. *Если ξ и η являются случайными величинами, для которых выражение $\mathsf{M}\xi + \mathsf{M}\eta$ определено, то $\mathsf{M}(\xi + \eta) = \mathsf{M}\xi + \mathsf{M}\eta$.*

Доказательство. Так как $\xi = \xi^+ - \xi^-$, $\eta = \eta^+ - \eta^-$, то имеет место равенство $\xi + \eta = (\xi^+ + \eta^+) - (\xi^- + \eta^-)$. Далее, поскольку выражение $\mathsf{M}\xi + \mathsf{M}\eta$ определено, то $\mathsf{M}\xi$ и $\mathsf{M}\eta$ существуют. Используя это и условие теоремы, методом от противного покажем, что $\min\{\mathsf{M}(\xi^+ + \eta^+), \mathsf{M}(\xi^- + \eta^-)\} < \infty$. В самом деле, предположим, что $\min\{\mathsf{M}(\xi^+ + \eta^+), \mathsf{M}(\xi^- + \eta^-)\} = \infty$. Тогда, учитывая факт существования $\mathsf{M}\xi$ и $\mathsf{M}\eta$, например, может оказаться верным следующий вариант: $\mathsf{M}\xi^+ = \infty$, $\mathsf{M}\xi^- < \infty$, $\mathsf{M}\eta^+ < \infty$, $\mathsf{M}\eta^- = \infty$. В этом случае $\mathsf{M}\xi = \infty$ и $\mathsf{M}\eta = -\infty$ и, значит, выражение $\mathsf{M}\xi + \mathsf{M}\eta = \infty - \infty$ не имеет смысла. Если имеет место $\mathsf{M}\xi^+ < \infty$, $\mathsf{M}\xi^- = \infty$, $\mathsf{M}\eta^+ = \infty$, $\mathsf{M}\eta^- < \infty$, то $\mathsf{M}\xi = -\infty$, $\mathsf{M}\eta = \infty$ и, значит, выражение $\mathsf{M}\xi + \mathsf{M}\eta = -\infty + \infty$ снова не определено. Поэтому имеет место следующее неравенство: $\min\{\mathsf{M}(\xi^+ + \eta^+), \mathsf{M}(\xi^- + \eta^-)\} < \infty$. Применяя леммы 8.6 и 8.3 к разложению $\xi + \eta = (\xi^+ + \eta^+) - (\xi^- + \eta^-)$, непосредственно получим: $\mathsf{M}(\xi + \eta) = \mathsf{M}((\xi^+ + \eta^+) - (\xi^- + \eta^-)) = \mathsf{M}(\xi^+ + \eta^+) - \mathsf{M}(\xi^- + \eta^-) = \mathsf{M}(\xi^+) - \mathsf{M}(\xi^-) + \mathsf{M}(\eta^+) - \mathsf{M}(\eta^-) = \mathsf{M}\xi + \mathsf{M}\eta$. Теорема доказана, и, значит, имеет место аддитивное свойство математического ожидания от суммы двух случайных величин.

Теорема 8.3. *Пусть ξ и η являются независимыми случайными величинами, для которых $\mathsf{M}|\xi| < \infty$, $\mathsf{M}|\eta| < \infty$. Тогда $\mathsf{M}(\xi \times \eta) = \mathsf{M}\xi \times \mathsf{M}\eta$.*

Доказательство. Прежде всего, напомним, что для простых случайных величин это утверждение было установлено. Рассмотрим теперь случай, когда ξ и η являются неотрицательными случайными величинами. Случайные величины ξ и η порождают аппроксимирующие последовательности $\{\xi_i = g_i(\xi); i = 1, 2, \ldots\}$ и $\{\eta_i = g_i(\eta); i = 1, 2, \ldots\}$, элементы $g_i(\xi)$ и $g_i(\eta)$ которых определяются функциональной зависимостью (8.3). Так как ξ и η независимы, то по теореме 7.2 простые случайные величины $\xi_i = g_i(\xi)$ и $\eta_i = g_i(\eta)$ также будут независимы и, следовательно, $\mathsf{M}(\xi_i \times \eta_i) = \mathsf{M}\xi_i \times \mathsf{M}\eta_i$. Поскольку $0 \leqslant \xi_i \leqslant \xi_{i+1} \leqslant \xi$, $0 \leqslant \eta_i \leqslant \eta_{i+1} \leqslant \eta$, $\lim\limits_{i\to\infty} \xi_i = \xi$ и $\lim\limits_{i\to\infty} \eta_i = \eta$, $\mathsf{M}(\xi_i \times \eta_i) = \mathsf{M}\xi_i \times \mathsf{M}\eta_i$, то, очевидно, имеем: $0 \leqslant \xi_i\eta_i \leqslant \xi_{i+1} + \eta_{i+1} \leqslant \xi\eta$, $\lim\limits_{i\to\infty}(\xi_i\eta_i) = (\lim\limits_{i\to\infty}\xi_i) \times \times (\lim\limits_{i\to\infty}\eta_i) = \xi\,\eta$, $\mathsf{M}(\xi\eta) = \lim\limits_{i\to\infty}\mathsf{M}(\xi_i\eta_i) = \lim\limits_{i\to\infty}(\mathsf{M}\xi_i \times \mathsf{M}\eta_i) = (\lim\limits_{i\to\infty}\mathsf{M}\xi_i) \times \times (\lim\limits_{i\to\infty}\mathsf{M}\eta_i) = \mathsf{M}\xi \times \mathsf{M}\eta$. Для неотрицательных случайных величин ξ и η теорема 8.3 справедлива.

Наконец, рассмотрим общий случай, когда $\xi = \xi^+ - \xi^-$ и $\eta = = \eta^+ - \eta^-$. Случайные величины $\xi^+ = \max\{0, \xi\}$ и $\xi^- = \max\{0, -\xi\}$ являются функциями от ξ, а случайные величины $\eta^+ = \max\{0, \eta\}$ и $\eta^- = \max\{0, -\eta\}$ суть функции от η. Отсюда, учитывая независимость ξ и η, выводим, что каждая из четырех пар ξ^+ и η^+, ξ^+ и η^-, ξ^- и η^+, ξ^- и η^- составлена из независимых величин. Используя соотношение вида $\xi\eta = (\xi^+ - \xi^-) \times (\eta^+ - \eta^-) = \xi^+\eta^+ - \xi^+\eta^- - \xi^-\eta^+ + \xi^-\eta^- = = (\xi^+\eta^+ - \xi^+\eta^-) + (\xi^-\eta^- - \xi^-\eta^+)$, ограничения $\mathsf{M}|\xi| < \infty$, $\mathsf{M}|\eta| < \infty$, леммы 8.4, 8.6 и теорему 8.2, легко вычислим:

$$
\begin{aligned}
\mathsf{M}(\xi \times \eta) &= \mathsf{M}((\xi^+\eta^+ - \xi^+\eta^-) + (\xi^-\eta^- - \xi^-\eta^+)) = \\
&= \mathsf{M}(\xi^+\eta^+ - \xi^+\eta^-) + \mathsf{M}(\xi^-\eta^- - \xi^-\eta^+) = \\
&= \mathsf{M}(\xi^+\eta^+) - \mathsf{M}(\xi^+\eta^-) + \mathsf{M}(\xi^-\eta^-) - \mathsf{M}(\xi^-\eta^+) = \\
&= \mathsf{M}(\xi^+)\mathsf{M}(\eta^+) - \mathsf{M}(\xi^+)\mathsf{M}(\eta^-) + \mathsf{M}(\xi^-)\mathsf{M}(\eta^-) - \mathsf{M}(\xi^-)\mathsf{M}(\eta^+) = \\
&= \mathsf{M}\xi \times \mathsf{M}\eta,
\end{aligned}
$$

причем условия $\mathsf{M}|\xi| < \infty$, $\mathsf{M}|\eta| < \infty$ обеспечивают применение лемм 8.4, 8.6 и теоремы 8.2.

Теорема 8.3 доказана, и, следовательно, имеет место так называемое мультипликативное свойство для математического ожидания.

Утверждения последних двух теорем или, другими словами, свойства аддитивности и мультипликативности математического ожидания с помощью метода математической индукции естественным образом распространяется на случай конечного числа случайных величин. Например, математическое ожидание суммы конечного числа n случайных величин равно сумме их математических ожиданий, т. е. $\mathsf{M}\big(\sum_{k=1}^n \xi_k\big) = \sum_{k=1}^n \mathsf{M}\xi_k$, если только правая часть этого равенства определена.

Свойства математического ожидания позволяют решить следующий важный вопрос о наилучшем приближении случайной величины ξ некоторой случайной величиной $\theta(\omega)$, где $\theta(\omega) = c$ для всех $\omega \in \Omega$ и $-\infty < c < +\infty$. Если $H_0 = \{\theta(\omega) \equiv c\colon -\infty < c < +\infty\}$, то вырожденную случайную величину $\theta(\omega) \equiv c$ называют *оценкой* или *прогнозом* случайной величины ξ в классе H_0. В этом случае ошибку такого приближения естественно измерять неотрицательным числом $\mathsf{M}(\xi - c)^2$. Пусть математическое ожидание $\mathsf{M}\xi$ конечно. Тогда при любом $-\infty < c < +\infty$ справедливы преобразования следующего вида: $\mathsf{M}(\xi - c)^2 = \mathsf{M}((\xi - \mathsf{M}\xi) + (\mathsf{M}\xi - c))^2 = \mathsf{M}((\xi - \mathsf{M}\xi)^2 + 2\,(\xi - \mathsf{M}\xi)(\mathsf{M}\xi - c) + (\mathsf{M}\xi - c)^2) = \mathsf{M}(\xi - \mathsf{M}\xi)^2 + 2\,(\mathsf{M}\xi - c)\,\mathsf{M}(\xi - \mathsf{M}\xi) + (\mathsf{M}\xi - c)^2 = \mathsf{M}(\xi - \mathsf{M}\xi)^2 + (\mathsf{M}\xi - c)^2 \geqslant \mathsf{M}(\xi - \mathsf{M}\xi)^2$. Если величина вида $\mathsf{M}(\xi - \mathsf{M}\xi)^2$ конечна, то ошибка принимает наименьшее значение при $c = \mathsf{M}\xi$. На практике при проведении эксперимента мы не всегда можем определить значения случайной величины ξ или не хотим иметь дело с поточечным ее заданием. В этом случае минимальное свойство математического ожидания позволяет случайную величину ξ заменить более простым математическим объектом вида: $\theta(\omega) \equiv \mathsf{M}\xi$. При этом точность такого приближения равна $\mathsf{M}(\xi - \mathsf{M}\xi)^2$. Этим фактом постоянно пользуются в прикладных задачах.

1.4. Формулы вычисления математических ожиданий. Следует отметить, что принятый здесь способ определения математического ожидания произвольной случайной величины ξ с помощью соответствующих ей последовательности простых случайных величин дает возможность строго вывести основные его свойства. К сожалению, такой способ не позволяет в общем случае получить доступные формулы исчисления математических ожиданий в прикладных задачах. Это обстоятельство связано с трудностями вычисления предела $\lim\limits_{i\to\infty} \mathsf{M}\xi_i$ для аппроксимирующей последовательности $\{\xi_i; i = 1, 2, \ldots\}$ в общих вероятностных пространствах $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. Поэтому возникает проблема получения простых формул для вычисления математических ожиданий. Решение этой проблемы дается с помощью следующих теорем.

Теорема 8.4. *Пусть на* $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ *заданы случайный вектор* $(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) = \xi(\omega)\colon \Omega \to R^n$ *и при* $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in R^n$ *борелевская функция* $g(\mathbf{x})\colon R^n \to R$. *Тогда математические ожидания* $\mathsf{M}g(\xi(\omega))$, $\mathsf{M}g(\mathbf{y}(\mathbf{x}))$ *существуют одновременно и* $\mathsf{M}g(\xi(\omega)) = \mathsf{M}g(\mathbf{y}(\mathbf{x}))$, *где* $\mathbf{y} = \mathbf{y}(\mathbf{x}) = (y_1(\mathbf{x}), y_2(\mathbf{x}), \ldots, y_n(\mathbf{x}))$ *есть многомерная случайная величина со значениями в* R^n, *определенная на выборочном вероятностном пространстве* $(R^n, \mathcal{B}^n, \mathsf{P}_\xi(\cdot))$ *вектора* $\xi = \xi(\omega)$, *и* $y_i(\mathbf{x}) = x_i$ *для всех* $i = 1, 2, \ldots, n$.

Доказательство. Так как $g(\mathbf{x})$ есть борелевская функция на n-мерном пространстве R^n со значениями в R, то $g(\xi(\omega))$ и $g(\mathbf{y}(\mathbf{x}))$ являются одномерными случайными величинами соответственно на основном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ и на выборочном вероятностном пространстве $(R^n, \mathcal{B}^n, \mathsf{P}_\xi(\cdot))$. Допустим сначала, что функция $g(\mathbf{x})$ неотрицательна. На основании определения математического ожидания для неотрицательных случайных величин $g(\xi)$ и $g(\mathbf{y})$

с учетом того факта, что эти случайные величины заданы на разных пространствах, имеем:

$$
\begin{aligned}
\mathsf{M}g(\xi) = \lim_{i\to\infty} \Bigg[\sum_{k=1}^{i2^i} (k-1)\,2^{-i}\mathsf{P}(\{\omega\colon (k-1)/2^i \leqslant g(\xi) < k/2^i\}) + {} \\
+ i\,\mathsf{P}(\{\omega\colon g(\xi) \geqslant i\}) \Bigg] = {} \\
= \lim_{i\to\infty} \Bigg[\sum_{k=1}^{i2^i} (k-1)\,2^{-i}\mathsf{P}(\{\omega\colon \xi(\omega) \in g^{-1}([(k-1)\,2^{-i}, k\,2^{-i}))\}) + {} \\
+ i\,\mathsf{P}(\{\omega\colon \xi(\omega) \in g^{-1}([i,\infty))\}) \Bigg] = {} \\
= \lim_{i\to\infty} \Bigg[\sum_{k=1}^{i2^i} (k-1)\,2^{-i}\mathsf{P}_\xi(\{\mathbf{x}\colon \mathbf{x} \in g^{1}([(k-1)\,2^{-i}, k\,2^{-i}))\}) + {} \\
+ i\,\mathsf{P}_\xi(\{\mathbf{x}\colon \mathbf{x} \in g^{1}([i,\infty))\}) \Bigg] = {} \\
= \lim_{i\to\infty} \Bigg[\sum_{k=1}^{i2^i} (k-1)\,2^{-i}\mathsf{P}_\xi(\{\mathbf{x}\colon (k-1)\,2^{-i} \leqslant g(\mathbf{y}(\mathbf{x})) < k\,2^{-i}\}) \Bigg] + {} \\
+ \lim_{i\to\infty} [i\,\mathsf{P}_\xi(\{\mathbf{x}\colon i \leqslant g(\mathbf{y}(\mathbf{x})) < \infty\})] = \mathsf{M}g(\mathbf{y}),
\end{aligned}
\tag{8.4}
$$

где $g^{-1}([(k-1)\,2^{-i}, k\,2^{-i})) = \{\mathbf{x}\colon (k-1)\,2^{-i} \leqslant g(\mathbf{y}(\mathbf{x})) < k\,2^{-i}\}$ и множество $g^{-1}([i,\infty)) = \{\mathbf{x}\colon i \leqslant g(\mathbf{y}(\mathbf{x})) < \infty\}$. В общем случае, когда функция $g(\mathbf{x})$ принимает значения разных знаков, представим ее в таком виде: $g(\mathbf{x}) = g^+(\mathbf{x}) - g^-(\mathbf{x})$, где функция $g^+(\mathbf{x}) = \max\{0, g(\mathbf{x}) \geqslant 0$ и функция $g^-(\mathbf{x}) = \max\{0, -g(\mathbf{x})\} \geqslant 0$. Если теперь воспользоваться представлениями $g(\xi(\omega)) = g^+(\xi(\omega)) - g^-(\xi(\omega))$, $\mathsf{M}g^+(\xi) = \mathsf{M}g^+(\mathbf{y})$, $\mathsf{M}g^-(\xi) = \mathsf{M}g^-(\mathbf{y})$, то $\mathsf{M}g(\xi) = \mathsf{M}(g^+(\xi) - g^-(\xi)) = \mathsf{M}g^+(\xi) - \mathsf{M}g^-(\xi) = \mathsf{M}g^+(\mathbf{y}) - \mathsf{M}g^-(\mathbf{y}) = \mathsf{M}g(\mathbf{y})$. При этом из существования математического ожидания $\mathsf{M}g(\mathbf{y})$ следует существование математического ожидания $\mathsf{M}g(\xi)$ и наоборот.

Отметим значение теоремы 8.4. Для вычисления математического ожидания $\mathsf{M}g(\xi)$, если пространство $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ гораздо сложнее выборочного пространства $(R^n, \mathcal{B}^n, \mathsf{P}_\xi(\cdot))$, можно воспользоваться вычислением математического ожидания $\mathsf{M}g(\mathbf{y})$. Напомним, что вероятностная функция $\mathsf{P}_\xi(\cdot)\colon \mathcal{B}^n \to [0,1]$ однозначно определяется интегральной функцией распределения $F_\xi(x_1, x_2, \ldots, x_n)$ многомерной случайной величины $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$. Следовательно, мы можем не знать полностью отображение $\xi(\omega)\colon \Omega \to R^n$ и вероятностную функцию $\mathsf{P}(\cdot)\colon \mathcal{F} \to [0,1]$. Подсчет математического ожидания $\mathsf{M}g(\mathbf{y})$ с использованием выборочного вероятностного пространства $(R^n, \mathcal{B}^n, \mathsf{P}_\xi(\cdot))$ в общем случае по-прежнему вызывает значительные трудности. Однако в некоторых случаях для $\mathsf{M}g(\mathbf{y})$ можно получить достаточно простые формулы вычисления. Перейдем теперь к выводу некоторых из такого рода формул при $n = 1$.

Лемма 8.8. *Пусть $\xi(\omega)$ — дискретная случайная величина и отображение $g(x)\colon R \to R$ является однозначным. Тогда имеет место формула*

$$\mathsf{M}g(\xi) = \sum_{k=1}^{m} g(x'_k)\, p_k, \tag{8.5}$$

если число m значений x'_1, $x'_2, \ldots$, x'_m для $\xi(\omega)$ конечно, и в противном случае

$$\mathsf{M}g(\xi) = \sum_{k=1}^{\infty} g(x'_k)\, p_k, \tag{8.6}$$

если ряд в правой части равенства (8.6) *сходится абсолютно.*

Доказательство. Если ξ является простой и вероятность $p_k = \mathsf{P}(\{\omega\colon \xi(\omega) = x'_k\}) = \mathsf{P}(A_k)$ для $k \in \{1, 2, \ldots, m\}$, то $g(\xi(\omega)) = \sum_{k=1}^{m} g(x'_k)\, I_{A_k}(\omega)$ — также простая. Отсюда по формуле (8.2) получаем (8.5). Этот результат получим иначе. Обозначим различные значения для случайной величины $g(\xi(\omega))$ через c_1, $c_2, \ldots$, c_r, где $r \leqslant m$. Используя методику получения формулы (8.2), распределение для $g(\xi(\omega))$ и обозначение p_k для вероятности $\mathsf{P}(\{\omega\colon \xi(\omega) = x'_k\})$, найдем

$$\mathsf{M}g(\xi(\omega)) = \sum_{i=1}^{r} c_i\, \mathsf{P}(\{\omega\colon g(\xi(\omega)) = c_i\}) =$$

$$= \sum_{i=1}^{r} c_i\, \mathsf{P}\Bigg(\bigcup_{k\colon g(x'_k) = c_i} \{\omega\colon \xi(\omega) = x'_k\} \Bigg) = \sum_{i=1}^{r} c_i \sum_{k\colon g(x'_k) = c_i} p_k = \sum_{k=1}^{m} g(x'_k)\, p_k.$$

Значит, формула (8.5) установлена.

Пусть теперь число различных значений x'_1, $x'_2, \ldots$ дискретной случайной величины $\xi(\omega)$ бесконечно. Рассмотрим сначала неотрицательную функцию $g(x)$. При любых фиксированных значениях $i \in \{1, 2, \ldots\}$, $k \in \{1, 2, \ldots, i\, 2^i\}$ и u, $u_0 \in \{x\colon (k-1)\, 2^{-i} \leqslant g(y(x)) < k\, 2^{-i}\}$, можно написать соотношения:

$$\mathsf{P}_\xi(\{x\colon (k-1)\, 2^{-i} \leqslant g(y(x)) < k\, 2^{-i}\}) = \sum_{j\colon (k-1)\, 2^{-i} \leqslant g(x'_j)) < k\, 2^{-i}} p_j,$$

$$\mathsf{P}_\xi(\{x\colon i \leqslant g(y(x)) < \infty\}) = \sum_{j\colon i \leqslant g(x'_j)} p_j, \quad |g(u) - g(u_0)| \leqslant 1/2^i.$$

Отсюда, используя неравенство $g(y(x)) \geqslant 0$, легко установим, что

$$\sum_{j\colon i \geqslant g(x'_j)} g(x'_j)\, p_j - 2^{-i} \leqslant$$

$$\leqslant \sum_{k=1}^{i\, 2^i} (k-1) 2^{-i} \mathsf{P}_\xi(\{x\colon (k-1)\, 2^{-i} \leqslant g(y(x)) < k\, 2^{-i}\}) +$$

$$+ i\, \mathsf{P}_\xi(\{x\colon i \leqslant g(y(x)) < \infty\}) \leqslant \sum_{j=1}^{\infty} g(x'_j)\, p_j.$$

Переходя в этих неравенствах к пределу по $i \to \infty$ и учитывая последнее равенство в соотношении (8.4), найдем:

$$\sum_{j=1}^{\infty} g(x_j')\, p_j \leqslant \lim_{i\to\infty}\left[\sum_{k=1}^{i\,2^i}(k-1)2^{-i}\mathsf{P}_\xi(\{x\colon (k-1)\,2^{-i} \leqslant g(y(x)) < k\,2^{-i}\})+\right.$$
$$\left. + i\,\mathsf{P}_\xi(\{x\colon i \leqslant g(y(x)) < \infty\})\right] = \mathsf{M}g(y(x)) \leqslant \sum_{j=1}^{\infty} g(x_j')\, p_j.$$

Значит, имеем равенство $\mathsf{M}g(y(x)) \leqslant \sum_{j=1}^{\infty} g(x_j')\, p_j$, которое доказывает утверждение этой леммы для неотрицательной функции $g(x)$.

Для произвольной однозначной функции $g(x)\colon R \to R$ имеем $g(x) = g^+(x) - g^-(x)$, где $g^+(x) = \max\{0, g(x)\}$ и $g^-(x) = \max\{0, -g(x)\}$. Из определений функций $g^+(x) \geqslant 0$ и $g^-(x) \geqslant 0$ и абсолютной сходимости ряда $\sum_{j=1}^{\infty} g(x_j)\, p_j$ следует, что $\mathsf{M}g^+(y(x)) = \sum_{j=1}^{\infty} g^+(x_j)\, p_j < \infty$ и $\mathsf{M}g^-(y(x)) = \sum_{j=1}^{\infty} g^-(x_j)\, p_j < \infty$. Используя лемму 8.6 и разложение $g(y(x)) = g^+(y(x)) - g^-(y(x))$, легко получим, что математическое ожидание вида

$$\mathsf{M}g(y) = \mathsf{M}(g^+(y) - g^-(y)) = \mathsf{M}g^+(y) - \mathsf{M}g^-(y) =$$
$$= \sum_{j=1}^{\infty} g^+(x_j')\, p_j - \sum_{j=1}^{\infty} g^-(x_j')\, p_j = \sum_{j=1}^{\infty} (g^+(x_j') - g^-(x_j'))\, p_j = \sum_{j=1}^{\infty} g(x_j')\, p_j.$$

Лемма 8.8 доказана.

При $g(x) = x$, естественно, формула (8.6) вычисляет математическое ожидание случайной величины $\xi(\omega)$, т. е.

$$\mathsf{M}\xi = \sum_{j=1}^{\infty} x_j'\, p_j \tag{8.7}$$

при $\sum_{j=1}^{\infty} |x_j'|\, p_j < \infty$. В формуле (8.7) предполагается, что ряд сходится абсолютно. В противном случае значение математического ожидания, как правило, не может быть вычислено однозначно и зависит от порядка перечисления значений случайной величины ξ. В последнем случае считается, что математическое ожидание для случайной величины не существует. Рассмотрим типичный пример такой ситуации.

Пример 8.2. Пусть в некотором эксперименте измеритель $\xi(\omega)$ принимает значения из счетного множества $\{(-1)^{i-1}2^i/i\colon i \geqslant 1\}$. Заметим, что для такой случайной величины нельзя указать ее минимальное значение. Определим закон распределения этой дискретной случайной величины равенствами $p_i = \mathsf{P}(\{\omega\colon \xi(\omega) = (-1)^{i-1}2^i/i\}) = 1/2^i$, $i \geqslant 1$. Так как $p_i > 0$ и $\sum_{i=1}^{\infty} p_i = 1$, то семейство чисел $p_i = 1/2^i$, $i \geqslant 1$, можно считать вероятностным распределением некоторой случайной величины. Легко видеть, что ряд $\sum_{i=1}^{\infty} |(-1)^{i-1}2^i i^{-1}|/2^i = \sum_{i=1}^{\infty} i^{-1}$ является расходящимся. Математическое ожидание для этой случайной величины не существует, так как его значение зависит от порядка перечисления значений такой случайной величины ξ. Действительно, легко находим (*Фихтенгольц Г.М.* Курс дифференциального и интегрального исчисления. Т. 2. — М.: ФИЗМАТЛИТ, 2008): $\sum_{i=1}^{\infty}(-1)^{i-1}2^i i^{-1}/2^i = 1 - 1/2 + 1/3 - 1/4 + 1/5 - 1/6 + 1/7 - 1/8 + \ldots = (1 - 1/2) + (1/3 - 1/4) + (1/5 - 1/6) + \ldots = \ln 2$, $\sum_{i=1}^{\infty}(-1)^{i-1}2^i i^{-1}/2^i = (1 - 1/2 - 1/4) + (1/3 - 1/6 - 1/8) + \ldots = 2^{-1}\ln 2$.

Лемма 8.9. *Пусть $\xi(\omega)\colon \Omega \to R$ есть непрерывная случайная величина с плотностью распределения $f(x)$ и заданное отображение $g(x)\colon R \to R$ является такой борелевской функцией, для которой $\int_{-\infty}^{+\infty} |g(x)|\, f(x)\, dx < \infty$. Тогда математическое ожидание случайной величины $g(\xi(\omega))$ равно*

$$\mathsf{M}g(\xi) = \int\limits_{-\infty}^{+\infty} g(x)\, f(x)\, dx. \tag{8.8}$$

Вывод равенства (8.8) и последующих формул, по которым можно вычислять математические ожидания $\mathsf{M}\xi$ и $\mathsf{M}g(\xi)$, осуществляется аналогично тому, как это было проделано при доказательстве леммы 8.8. Это утверждение проще всего продемонстрировать на доказательстве формулы (8.8) в частном случае, когда функция $g(x) = x$. При этом условии формула (8.8) вычисляет математическое ожидание $\mathsf{M}\xi$ случайной величины $\xi(\omega)$, т. е. при $g(x) = x$ и $\int_{-\infty}^{+\infty} x\, f(x)\, dx < \infty$ выполняется равенство

$$\mathsf{M}\xi = \int\limits_{-\infty}^{+\infty} x\, f(x)\, dx. \tag{8.9}$$

Действительно, рассмотрим теперь неотрицательные случайные величины $\xi^+ = \max\{0, \xi\}$ и $\xi^- = \max\{0, -\xi\}$. Так как ξ является непрерывной случайной величиной, то интегральные функции распределения $F_{\xi^+}(x)$ и $F_{\xi^-}(x)$ для случайных величин ξ^+ и ξ^- вычисляются по формулам:

$$\begin{gathered}
F_{\xi^+}(x) = \begin{cases} 0 & \text{при } x \leqslant 0; \\ F_\xi(x) = \int\limits_{-\infty}^{x} f(y)\, dy & \text{при } x > 0, \end{cases} \\
F_{\xi^-}(x) = \begin{cases} 0 & \text{при } x \leqslant 0; \\ 1 - F_\xi(-x) = \int\limits_{-x}^{\infty} f(y)\, dy & \text{при } x > 0. \end{cases}
\end{gathered} \tag{8.10}$$

Заметим, что из этих формул сразу следует, что при $0 < F_\xi(0) < 1$ случайные величины ξ^+ и ξ^- не являются непрерывными, так как $F_{\xi^+}(+0) \neq 0$ и $F_{\xi^-}(+0) \neq 0$. Далее, если $F_\xi(0) = 0$, то случайная величина ξ^+ будет непрерывной, а случайная величина ξ^- принимает значение ноль с вероятность единица, т. е. является вырожденной.

Наоборот, при $F_\xi(0) = 1$ случайная величина ξ^- будет непрерывной, а случайная величина ξ^+ принимает значение ноль с вероятность единица.

На основании определения математического ожидания для неотрицательных случайных величин ξ^+, ξ^- и равенства (8.4), которое было получено при доказательстве теоремы 8.4, с учетом $g(x) = x$ имеем

$$\begin{aligned}\mathsf{M}\xi^+ = \lim_{i\to\infty}\Bigg[\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\mathsf{P}(\{\omega\colon (k-1)/2^i \leqslant \xi^+ < k/2^i\})+\\ + i\,\mathsf{P}(\omega\colon \xi^+ \geqslant i)\Bigg] =\\ = \lim_{i\to\infty}\Bigg[\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\mathsf{P}_{\xi^+}(\{x\colon (k-1)\,2^{-i} \leqslant x < k\,2^{-i}\})+\\ + i\,\mathsf{P}_{\xi^+}(\{x\colon i \leqslant x < \infty\})\Bigg],\\ \mathsf{M}\xi^- = \lim_{i\to\infty}\Bigg[\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\mathsf{P}(\{\omega\colon (k-1)/2^i \leqslant \xi^- < k/2^i\})+\\ + i\,\mathsf{P}(\omega\colon \xi^- \geqslant i)\Bigg] =\\ = \lim_{i\to\infty}\Bigg[\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\mathsf{P}_{\xi^-}(\{x\colon (k-1)\,2^{-i} \leqslant x < k\,2^{-i}\})\Bigg]+\\ + \lim_{i\to\infty} i\,\mathsf{P}_{\xi^-}(\{x\colon i \leqslant x < \infty\}).\end{aligned} \tag{8.11}$$

Используя формулы (8.10), можно вычислить вероятности

$$\mathsf{P}_{\xi^+}(\{x\colon (k-1)\,2^{-i} \leqslant x < k\,2^{-i}\}) = \int\limits_{(k-1)\,2^{-i}}^{k2^{-i}} f(u)\,du, \quad k = 2, 3, \ldots, i\,2^i,$$

$$\mathsf{P}_{\xi^+}(\{x\colon 0 \leqslant x < 2^{-i}\}) = \int\limits_{-\infty}^{2^{-i}} f(u)\,du, \quad \mathsf{P}_{\xi^+}(\{x\colon i \leqslant x < \infty\}) = \int\limits_{i}^{\infty} f(u)\,du,$$

$$\mathsf{P}_{\xi^-}(\{x\colon (k-1)\,2^{-i} \leqslant x < k\,2^{-i}\}) = \int\limits_{-k2^{-i}}^{-(k-1)2^{-i}} f(u)\,du, \quad k = 2, 3, \ldots, i\,2^i,$$

$$\mathsf{P}_{\xi^-}(\{x\colon 0 \leqslant x < 2^{-i}\}) = \int\limits_{-2^{-i}}^{\infty} f(u)\,du, \quad \mathsf{P}_{\xi^-}(\{x\colon i \leqslant x < \infty\}) = \int\limits_{-\infty}^{-i} f(u)\,du,$$

которые указаны в соотношении (8.11). Далее легко установим для случайных величин ξ^+ и ξ^-, что

$$\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\mathsf{P}_{\xi^+}(\{x\colon (k-1)\,2^{-i}\leqslant x<k\,2^{-i}\})+ \\ +\,i\,\mathsf{P}_{\xi^+}(\{x\colon i\leqslant x<\infty\})= \\ =\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\int\limits_{(k-1)\,2^{-i}}^{k2^{-i}}f(u)\,du+i\int\limits_{i}^{\infty}f(u)\,du,$$

$$\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\mathsf{P}_{\xi^-}(\{x\colon (k-1)\,2^{-i}\leqslant x<k\,2^{-i}\})+ \\ +\,i\,\mathsf{P}_{\xi^-}(\{x\colon i\leqslant x<\infty\})= \\ =\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\int\limits_{-k2^{-i}}^{-(k-1)\,2^{-i}}f(u)\,du+i\int\limits_{-\infty}^{-i}f(u)\,du= \\ =\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\int\limits_{(k-1)2^{-i}}^{k2^{-i}}f(-u)\,du+i\int\limits_{i}^{-\infty}f(-u)\,du.$$

При любых фиксированных значениях $i\in\{1,2,\ldots\}$, $k\in\{1,2,\ldots,i\,2^i\}$ имеем $|u-u_0|\leqslant 1/2^i$, если только $u,u_0\in\{x\colon (k-1)/2^i\leqslant x<k/2^i\}$. Следовательно, будут справедливы следующие неравенства:

$$\int\limits_0^i uf(u)\,du-2^{-i}\leqslant\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\int\limits_{(k-1)\,2^{-i}}^{k2^{-i}}f(u)\,du+i\int\limits_i^{\infty}f(u)\,du\leqslant \\ +\leqslant\int\limits_0^{\infty}uf(u)\,du,$$

$$\int\limits_0^i uf(-u)\,du-2^{-i}\leqslant\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\int\limits_{(k-1)\,2^{-i}}^{k2^{-i}}f(-u)\,du+ \\ +\,i\int\limits_i^{\infty}f(-u)\,du\leqslant\int\limits_0^{\infty}uf(-u)\,du$$

для всех значений $i\in\{1,2,\ldots\}$. Переходя теперь в этих неравенствах к пределу по $i\to\infty$ и учитывая равенство (8.11), найдем

$$\int\limits_0^{\infty}uf(u)\,du\leqslant\lim_{i\to\infty}\left[\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\int\limits_{(k-1)\,2^{-i}}^{k2^{-i}}f(u)\,du+i\int\limits_i^{\infty}f(u)\,du\right]= \\ =\mathsf{M}\xi^+\leqslant\int\limits_0^{\infty}uf(u)\,du,\quad \mathsf{M}\xi^+=\int\limits_0^{\infty}uf(u)\,du,$$

$$\int_0^\infty uf(-u)\,du \leqslant \lim_{i\to\infty}\left[\sum_{k=1}^{i2^i}(k-1)\,2^{-i}\int\limits_{(k-1)\,2^{-i}}^{k2^{-i}} f(-u)\,du + i\int_i^\infty f(-u)\,du\right] =$$

$$= \mathsf{M}\xi^- \leqslant \int_0^\infty uf(-u)\,du, \quad \mathsf{M}\xi^- = \int_0^\infty uf(-u)\,du = -\int_{-\infty}^0 uf(u)\,du.$$

Для произвольной случайной величины $\xi(\omega)$ имеем $\xi(\omega) = \xi^+(\omega) - \xi^-(\omega)$. Из определений случайных величин $\xi^+(\omega) \geqslant 0$ и $\xi^-(\omega) \geqslant 0$ и неравенства $\int_{-\infty}^{+\infty} |x|\,f(x)\,dx < \infty$ следует, что

$$\mathsf{M}\xi^+ = \int_0^{+\infty} x\,f(x)\,dx < \infty \quad \text{и} \quad \mathsf{M}\xi^- = \int_0^{+\infty} x\,f(-x)\,dx < \infty.$$

Применяя теперь лемму 8.6 и равенство $\xi(\omega) = \xi^+(\omega) - \xi^-(\omega)$, получим, что

$$\mathsf{M}\xi = \mathsf{M}(\xi^+ - \xi^-) = \mathsf{M}\xi^+ - \mathsf{M}\xi^- = \\ = \int_0^{+\infty} x\,f(x)\,dx + \int_{-\infty}^0 x\,f(x)\,dx = \int_{-\infty}^{+\infty} x\,f(x)\,dx.$$

Значит, имеет место формула (8.9).

Приведем простой пример на вычисление математического ожидания $\mathsf{M}\xi$ непрерывной случайной величины ξ, если ее плотность распределения $f(x)$ является четной функцией и $\int_{-\infty}^{+\infty} |x|\,f(x)\,dx < \infty$. Используя (8.9), легко находим:

$$\int_{-\infty}^{+\infty} x\,f(x)\,dx = \int_{-\infty}^0 x\,f(x)\,dx + \int_0^{+\infty} x\,f(x)\,dx = \\ = -\int_{+\infty}^0 (-u)\,f(-u)\,du + \int_0^{+\infty} x\,f(x)\,dx = \\ = \int_0^{+\infty} (-u)\,f(-u)\,du + \int_0^{+\infty} x\,f(x)\,dx = -\int_0^{+\infty} u\,f(u)\,du + \int_0^{+\infty} x\,f(x)\,dx = 0.$$

Лемма 8.10. *Пусть $\xi(\omega)\colon \Omega \to R$ есть случайная величина с функцией распределения $F(x)$ и отображение $g(x)\colon R \to R$ является такой борелевской функцией, для которой интеграл Стилтьеса $\int_{-\infty}^{+\infty} |g(x)|\,dF(x) < \infty$. Тогда математическое ожидание случайной величины $g(\xi(\omega))$ равно*

$$\mathsf{M}g(\xi) = \int_{-\infty}^{+\infty} g(x)\,dF(x). \tag{8.12}$$

Здесь следует отметить, что интеграл Стилтьеса $\int_{-\infty}^{+\infty} |g(x)|\,dF(x)$ можно вычислять следующим образом:

$$\int\limits_{-\infty}^{+\infty} |g(x)|\,dF(x) = \lim_{\substack{a\to-\infty\\ b\to+\infty}} \int\limits_{a}^{b} |g(x)|\,dF(x) =$$

$$= \lim_{\substack{a\to-\infty\\ b\to+\infty}} \lim_{i\to+\infty} \sum_{k=1}^{i} |g(u'_k)|\,(F(a_k) - F(a_{k-1})) < \infty,$$

где $a = a_0 < a_1 < \ldots < a_i = b$, $u'_k \in [a_{k-1}, a_k)$ и $\lim\limits_{i\to+\infty} \max\limits_{1\leqslant k\leqslant i} (a_k - a_{k-1}) = 0$. Далее, в частном случае при $g(x) = x$ формула (8.12) вычисляет математическое ожидание случайной величины $\xi(\omega)$, т. е. при $\int_{-\infty}^{+\infty} |x|\,dF(x) < \infty$ имеем

$$\mathsf{M}\xi = \int\limits_{-\infty}^{+\infty} x\,dF(x). \tag{8.13}$$

Формула (8.13) позволяет предложить следующую геометрическую интерпретацию математического ожидания. С этой целью запишем равенство (8.13) в виде

$$\mathsf{M}\xi = \int\limits_{-\infty}^{+\infty} x\,dF(x) = \int\limits_{-\infty}^{0} x\,dF(x) - \int\limits_{0}^{+\infty} x\,d(1 - F(x)),$$

а затем проинтегрируем его по частям. После простых вычислений получим соотношение вида

$$\mathsf{M}\xi = -\int\limits_{-\infty}^{0} x\,dF(x) - \int\limits_{0}^{+\infty} (1 - F(x))\,dx = -S^- + S^+.$$

Правая часть в виде площадей S^- и S^+ отображена на рис. 8.1.

Площадь S^- заключена между интегральной функцией $F(x)$ в промежутке $(-\infty, 0]$, осью ординат и осью абсцисс в промежутке $[0, +\infty)$.

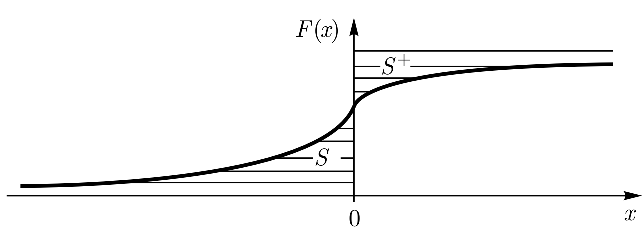


Рис. 8.1

Площадь S^+ заключена между осью ординат, интегральной функцией $F(x)$ в промежутке $[0,+\infty)$ и прямой, которая параллельна оси абсцисс и проходит через точку $(0,1)$. Итак, математическое ожидание равно разности указанных площадей. Используя эту интерпретацию, легко найдем математическое ожидание $\mathsf{M}\xi$ сингулярной случайной величины из примера 6.19. Из рис. 6.20 непосредственно получаем, что площадь S^- равна нулю, а площадь $S^+ = (1/3)\times(1/2) + (2/9)\times(1/2) + (4/27)\times(1/2) + \ldots = 1/2$.

Лемма 8.11. *Пусть $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))\colon \Omega \to R^n$ есть многомерная случайная величина с функцией распределения $F(x_1, x_2, \ldots, x_n)$ и отображение $g(x_1, x_2, \ldots, x_n)\colon R^n \to R$ является такой борелевской функцией, для которой многомерный интеграл Римана–Стилтьеса*

$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\ldots\int\limits_{-\infty}^{+\infty} |g(x_1, x_2, \ldots, x_n)|\, dF(x_1, x_2, \ldots, x_n) < +\infty.$$

Тогда математическое ожидание одномерной случайной величины вида $g(\xi_1(\omega)), g(\xi_2(\omega)), \ldots, g(\xi_n(\omega))$ можно вычислять по формуле

$$\mathsf{M}g(\xi_1(\omega)), \xi_2(\omega), \ldots, \xi_n(\omega)) = \\ = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\ldots\int\limits_{-\infty}^{+\infty} g(x_1, x_2, \ldots, x_n)\, dF(x_1, x_2, \ldots, x_n). \quad (8.14)$$

Если для каждого $k = 1, 2, \ldots, n$ случайная величина $\xi_k(\omega)$ принимает счетное число различных значений z_k из множества Z_k, набор вероятностей $\mathsf{P}(\{\omega\colon \xi_k(\omega) = z_k,\, k = 1, 2, \ldots, n\}) = p_{z_1, z_2, \ldots, z_n}$, $z_1 \in Z_1$, $z_2 \in Z_2, \ldots$, $z_n \in Z_n$, определяет распределение вектора $\xi(\omega) = (\xi_1(\omega)), \xi_2(\omega), \ldots, \xi_n(\omega))$ и, наконец, $\sum_{z_1\in Z_1}\sum_{z_2\in Z_2}\ldots\sum_{z_n\in Z_n} |g(z_1, z_2, \ldots, z_n)|\, p_{z_1, z_2, \ldots, z_n} < +\infty$, то формула (8.14) примет вид

$$\mathsf{M}g(\xi_1(\omega)), \xi_2(\omega), \ldots, \xi_n(\omega)) = \\ = \sum_{z_1\in Z_1}\sum_{z_2\in Z_2}\ldots\sum_{z_n\in Z_n} g(z_1, z_2, \ldots, z_n)\, p_{z_1, z_2, \ldots, z_n}. \quad (8.15)$$

Для непрерывной случайной величины $\xi(\omega) = (\xi_1(\omega)), \xi_2(\omega), \ldots, \xi_n(\omega))$ с плотностью распределения $f(x_1, x_2, \ldots, x_n)$ вычислительная формула (8.14) запишется в следующем виде:

$$\mathsf{M}g(\xi_1(\omega)), \xi_2(\omega), \ldots, \xi_n(\omega)) = \\ = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\ldots\int\limits_{-\infty}^{+\infty} g(x_1, x_2, \ldots, x_n)\, f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n, \quad (8.16)$$

если n-мерный интеграл этого равенства сходится абсолютно.

1.5. Прямой метод вычисления математического ожидания. Если отображение $g(x_1, x_2, \ldots, x_n)\colon R^n \to R$ есть борелевская функция, то среднее значение случайной величины $\eta(\omega) = g(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$ можно вычислить с использованием так называемого прямого метода. В этом случае сначала отыскивают значения $\eta(\omega)$, $\omega \in \Omega$, случайной величины или ее распределение. Затем, по этой информации с использованием соответствующих вычислительных формул (8.1), (8.7), (8.9), (8.13) окончательно получают $\mathsf{M}\xi(\eta(\omega))$. В силу этого большинство учебников по теории вероятностей, ориентированных на приложения, математическое ожидание непосредственно определяют с помощью формул (8.1), (8.5)–(8.9), (8.12)–(8.16). Такой способ на первый взгляд кажется наиболее доступным и простым. Однако здесь приходится вводить и запоминать значительное число определений. Более того, при таком подходе строгое доказательство некоторых свойств математического ожидания вызывает значительные трудности. Например, если $\xi(\omega)$ непрерывная случайная величина с плотностью распределения $f_\xi(x) > 0$ при $-\infty < x < +\infty$ и $\eta(\omega) = |\xi(\omega)|$ — непрерывная случайная величина с плотностью $f_{|\xi|}(x)$, то из определения математического ожидания с помощью формулы (8.8) получаем $\mathsf{M}\eta = \int_{-\infty}^{+\infty} |x|\, f_\xi(x)\, dx$, а из определения математического ожидания по формуле (8.9) имеем $\mathsf{M}\eta = \int_{-\infty}^{+\infty} x\, f_{|\xi|}(x)\, dx$. Это означает, что на основании свойств интеграла Римана в курсе анализа должно быть доказано удивительное равенство: $\int_{-\infty}^{+\infty} x\, f_{|\xi|}(x)\, dx = \int_{-\infty}^{+\infty} |x|\, f_\xi(x)\, dx$, в котором знак модуля можно убрать из индекса у плотности и перенести его на переменную интегрирования. Поэтому свойства математического ожидания при таком подходе, как правило, доказываются отдельно и разными методами для дискретных, непрерывных и произвольных случайных величин. Продемонстрируем это в случае, когда доказываются свойства аддитивности и мультипликативности математического ожидания для $n = 2$.

Предположим, что двумерная дискретная случайная величина (ξ_1, ξ_2) принимает значение (z_1, z_2) с вероятностью p_{z_1,z_2}, где $z_1 \in Z_1$, $z_2 \in Z_2$. Тогда $\theta = g(\xi_1, \xi_2) = \xi_1 + \xi_2$ есть дискретная случайная величина, которая принимает значения $z_1 + z_2$, $z_1 \in Z_1$, $z_2 \in Z_2$, часть из которых может совпадать. Значит, в общем случае $\mathsf{P}(\{\omega\colon \theta(\omega) = z_1 + z_2\}) \neq p_{z_1,z_2}$. Поэтому для вычисления $\mathsf{M}\theta = \mathsf{M}(\xi_1 + \xi_2)$ следует использовать формулу (8.15), а не формулу (8.7). Применяя сначала формулу (8.15), а на последнем этапе (8.7), найдем, что математическое ожидание величины θ равно

$$\mathsf{M}\theta = \sum_{z_1 \in Z_1} \sum_{z_2 \in Z_2} (z_1 + z_2)\, p_{z_1,z_2} = \sum_{z_1 \in Z_1} z_1 \sum_{z_2 \in Z_2} p_{z_1,z_2} + \sum_{z_2 \in Z_2} z_2 \sum_{z_1 \in Z_1} p_{z_1,z_2} =$$

$$= \sum_{z_1 \in Z_1} z_1 \, \mathsf{P}(\{\omega\colon \xi_1(\omega) = z_1\}) + \sum_{z_2 \in Z_2} z_2 \, \mathsf{P}(\{\omega\colon \xi_2(\omega) = z_2\}) = \mathsf{M}\xi_1 + \mathsf{M}\xi_2.$$

Рассмотрим теперь непрерывный случайный вектор (ξ_1, ξ_2) с плотностью $f(x, y)$. Тогда для случайной величины $\theta = g(\xi_1, \xi_2) = \xi_1 + \xi_2$, используя формулу (8.16), а затем (8.9), найдем математическое ожидание:

$$\mathsf{M}\theta = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x+y)\, f(x,y)\, dx\, dy = \int\limits_{-\infty}^{+\infty} x\Big(\int\limits_{-\infty}^{+\infty} f(x,y)\, dy\Big)\, dx +$$
$$+ \int\limits_{-\infty}^{+\infty} y\Big(\int\limits_{-\infty}^{+\infty} f(x,y)\, dx\Big)\, dy = \int\limits_{-\infty}^{+\infty} x\, f_{\xi_1}(x)\, dx + \int\limits_{-\infty}^{+\infty} y\, f_{\xi_2}(y)\, dy = \mathsf{M}\xi_1 + \mathsf{M}\xi_2.$$

Доказательство мультипликативного свойства математического ожидания приведем на примере произвольных случайных величин. Итак, если компоненты двумерной случайной величины (ξ_1, ξ_2) с функцией распределения $F(x, y)$ независимы, то в результате рассмотрения одномерной случайной величины $\eta(\omega) = g(\xi_1(\omega), \xi_2(\omega)) = \xi_1(\omega)\, \xi_2(\omega)$ и применения формулы (8.14), равенства $F(x, y) = F_{\xi_1}(x) \times F_{\xi_2}(y)$, свойства интеграла Римана–Стилтьеса и формулы (8.13) получим, что

$$\mathsf{M}\eta(\omega) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} g(x,y)\, dF(x,y) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xy\, dF(x,y) =$$
$$= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xy\, dF_{\xi_1}(x)\, dF_{\xi_2}(y) = \int\limits_{-\infty}^{+\infty} x\, dF_{\xi_1}(x) \int\limits_{-\infty}^{+\infty} y\, dF_{\xi_2}(y) = \mathsf{M}\xi_1 \cdot \mathsf{M}\xi_2.$$

Для вычисления математического ожидания ранее было отмечено, что можно использовать прямой метод. При таком методе необходимо задать случайную величину ξ или поточечно, или с помощью ее законов распределения. Далее, используя соответствующие вычислительные формулы, непосредственно находят математическое ожидание. Рассмотрим простые примеры, в которых будет использован этот метод.

Пример 8.3. Два игрока поочередно подбрасывают несимметричную монету до первого выпадения герба (орла). Каждый бросок осуществляется независимым образом с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Вероятность появления герба при одном броске равна p. Вычислить среднее число всех бросков, которое совершают игроки.

Решение. В этой игре ω_j при каждом фиксированном $j \geqslant 1$ есть описание такого элементарного исхода, когда $j-1$ раз выпадает решетка, а последний раз выпадает герб. Будем полагать, что в этом эксперименте выпадение только решеток есть невозможное событие. Для этого опыта достоверное событие $\Omega = \{\omega_1, \omega_2, \ldots\}$, а σ-алгебра $\mathcal{F} = \{A\colon A \subset \Omega\}$ есть множество всех подмножеств множества Ω. Пусть $\xi(\omega)$ определяет число выпавших в игре решеток. Легко видеть, что случайная величина $\xi(\omega_j) = j-1$, а вероятность $p_j = \mathsf{P}(\{\omega\colon \xi(\omega) = j-1\}) = pq^{j-1} > 0$ при $j \geqslant 1$ и $q = 1-p$. Так как $p_j > 0$ и $\sum_{j=1}^{\infty} p_j = 1$, то $\xi(\omega)$ является дискретной случайной величиной с так называемым геометрическим распределением. По формуле (8.7) с учетом равенства $p = 1-q$ имеем

$$\mathsf{M}\xi = \sum_{j=1}^{\infty}(j-1)\,pq^{j-1} = \sum_{j=1}^{\infty} jpq^{j-1} - \sum_{j=1}^{\infty} pq^{j-1} = p\,\frac{d}{dq}\Bigl(\sum_{j=1}^{\infty} q^j\Bigr) - 1 =$$
$$= p\,\frac{d}{dq}\Bigl(\frac{q}{1-q}\Bigr) - 1 = p(1-q)^{-1} + pq(1-q)^{-2} - 1 = q/p.$$

Если $\eta(\omega)$ определяет число всех произведенных бросков игроками, то $\eta(\omega) = \xi(\omega)+1$, и, следовательно, $\mathsf{M}\eta = \mathsf{M}(\xi+1) = qp^{-1} + 1 = p^{-1}$. Заметим, что если монета симметричная ($p = q = 1/2$), то среднее число бросков равно двум.

С помощью этого способа приведем теперь доказательство равенства $\mathsf{M}(c\,\xi) = c\,\mathsf{M}\xi$ в частном случае, когда выполняются условия теоремы 8.1, постоянная $c \neq 0$ и $\xi(\omega)$ является одномерной дискретной случайной величиной с распределением $\{\mathsf{P}(\{\omega\colon \xi(\omega) = z_1\}) = p_{z_1}\colon z_1 \in Z_1\}$. Очевидно, что случайная величина $\xi_1(\omega) = c\,\xi(\omega)$ будет дискретной, и иметь распределение $\mathsf{P}(\{\omega\colon \xi_1(\omega) = cz_1\}) = p_{z_1}$, $z_1 \in Z_1$. Теперь непосредственно по формуле (8.7) находим значение математического ожидания случайной величины ξ_1:

$$\mathsf{M}\xi_1 = \sum_{z_1 \in Z_1} (cz_1)\,\mathsf{P}(\{\omega\colon \xi_1(\omega) = cz_1\}) = c\sum_{z_1 \in Z_1} z_1 p_{z_1} = c\,\mathsf{M}\xi.$$

Проиллюстрируем прямой метод вычисления $\mathsf{M}\xi$ для случайных величин, рассмотренных в задаче о гончарном круге из шестой главы.

Пример 8.4. Найти среднюю длину $\xi(\omega)$ трещины, которая возникает на верхней плоской поверхности гончарного круга радиуса r.

Решение. Вариант 1. При рассмотрении первого эксперимента примера 6.12 было получено, что $\xi(\omega) = 2(r^2 - (r-\omega)^2)^{1/2}$ при $0 \leqslant \omega \leqslant 2r$. Интегральная функция распределения $F_\xi(x)$ равна нулю при $x \leqslant 0$, равна $1 - (1 - x^2/4r^2)^{1/2}$ при $0 < x \leqslant 2r$ и, наконец, равна единице при $x > 2r$. Напомним, что трещины появляются наудачу перпендикулярно к некоторому фиксированному направлению на плоской поверхности гончарного круга. Следовательно, плотность распределения $f_\xi(x)$ равна нулю при $x \leqslant 0$, равна $x(2r)^{-1}(4r^2 - x^2)^{-1/2}$ при $0 < x \leqslant 2r$ и, наконец, равна нулю при $x > 2r$. Отсюда средняя длина трещины равна

$$\mathsf{M}\xi = (2r)^{-1} \int\limits_0^{2r} x^2(4r^2 - x^2)^{-1/2} dx =$$
$$= \left[r \arcsin(x/2r) - (4r)^{-1} x(4r^2 - x^2)^{1/2}\right]\Big|_0^{2r} = \pi r/2 \approx 1{,}57\, r.$$

Вариант 2. Пусть теперь трещины на плоской поверхности гончарного круга образуют хорды, для которых один из их концов всегда закреплен в некоторой фиксированной точке его окружности, а другие концы возникают непреднамеренно. В этом случае (см. изучение второго эксперимента в примере 6.12) длина $\xi(\omega) = 2r\cos\omega$, $-\pi/2 \leqslant \omega \leqslant \pi/2$. Интегральный и дифференциальный законы распределения определяются следующими формулами: $F_\xi(x) = 0$ при $x \leqslant 0$, $F_\xi(x) = 1 - 2\pi^{-1}\arccos(x/2r)$ при $0 < x \leqslant 2r$ и $F_\xi(x) = 1$ при $x > 2r$; $f_\xi(x) = 0$ при $x \leqslant 0$, $f_\xi(x) = 2\pi^{-1}(4r^2 - x^2)^{-1/2}$ при $0 < x \leqslant 2r$ и $f_\xi(x) = 0$ при $x > 2r$. Получаем, что средняя длина трещины $\mathsf{M}\xi = 2\pi^{-1}\int_0^{2r} x\,(4r^2 - x^2)^{-1/2}dx = \left[-2\pi^{-1}(4r^2 - x^2)^{1/2}\right]\Big|_0^{2r} = 4r\pi^{-1} \approx$ $\approx 1{,}27\, r$.

Вариант 3. Наконец, если трещина пересекает плоскую поверхность гончарного круга по хорде, и если середина этой хорды выбирается наудачу, то можно воспользоваться результатами исследований третьего эксперимента в примере 6.12. Длина $\xi(\omega)$ такой хорды и ее интегральная функция $F_\xi(x)$ определяются с помощью формул вида: $\xi(\omega) = 2(r^2 - u^2 - v^2)^{1/2}$, $\omega = (u,v) \in \{(u,v)\colon u^2 + v^2 \leqslant r^2,\ (u,v) \neq$ $\neq (0,0)\}$; $F_\xi(x) = 0$ при $x \leqslant 0$, $F_\xi(x) = x^2/4r^2$ при $0 < x \leqslant 2r$ и $F_\xi(x) = 1$ при $x > 2r$. Отсюда сразу получаем, что $f_\xi(x) = 0$ при $x \leqslant 0$, $f_\xi(x) = (2r^2)^{-1}x$ при $0 < x \leqslant 2r$ и $f_\xi(x) = 0$ при $x > 2r$. Поэтому математическое ожидание $\mathsf{M}\xi = (2r^2)^{-1}\int_0^{2r} x^2 dx = 4r/3 \approx 1{,}33\, r$.

Итак, математическое ожидание длины трещины на поверхности круга принимает наибольшее значение при первой технологии его изготовления. Математические ожидания длины трещины практически совпадают при второй и третьей технологиях. Однако согласно решению задачи о гончарном круге (пример 2.37) следует выбирать круг, который изготовлен с использованием третьей технологии. В этом случае с наименьшей вероятностью возникает трещина, длина которой больше $\sqrt{3}\, r$, и мастер по производству посуды существенно реже будет заменять круг.

Пример 8.5. Государство выплачивает в течение года каждому пенсионеру по инвалидности компенсацию η за лекарство. Стоимость ξ всех лекарств для произвольного пенсионера есть случайная величина, плотность распределения $f_\xi(x)$ которой равна $(b-a)^{-1}$ при $x \in [a, b]$ и равна нулю при $x \notin [a, b]$. Предполагается, что значение $\eta(\omega)$ компенсации функционально связано со значением $\xi(\omega)$ стоимости лекарств. В простейшем случае компенсация η вычисляется по следующей простой функциональной зависимости $\eta = g(\xi)$ вида: $\eta = c$ при $\xi \leqslant c$, $\eta = \xi$ при $c < \xi \leqslant d$ и, наконец, $\eta = d$ при $\xi > d$. Здесь постоянные величины a, b, c и d удовлетворяют естественному условию $0 < a < c < d < b$. Вычислить математическое ожидание компенсации η.

Решение. Непосредственно из определения компенсации $\eta(\omega)$ найдем $\mathsf{P}(\{\omega\colon \eta(\omega) < c\}) = 0$, $\mathsf{P}(\{\omega\colon \eta(\omega) = c\}) = \mathsf{P}(\{\omega\colon \xi(\omega) \leqslant c\}) = (c-a)(b-a)^{-1} > 0$, $\mathsf{P}(\{\omega\colon \eta(\omega) > d\}) = 0$, $\mathsf{P}(\{\omega\colon \eta(\omega) = d\}) = \mathsf{P}(\{\omega\colon \xi(\omega) > d\}) = (b-d)(b-a)^{-1} > 0$, $\mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \mathsf{P}(\{\omega\colon \xi(\omega) < y\}) = (y-a)(b-a)^{-1}$ при $c < y \leqslant d$. Отсюда легко получим функцию распределения для случайной величины η, а именно: $F_\eta(y) = 0$ при $y \leqslant c$, $F_\eta(y) = (y-a)(b-a)^{-1}$ при $c < y \leqslant d$ и $F_\eta(y) = 1$ при $y > d$. График ее интегральной функции распределения $F_\eta(y)$ представлен на рис. 8.2.

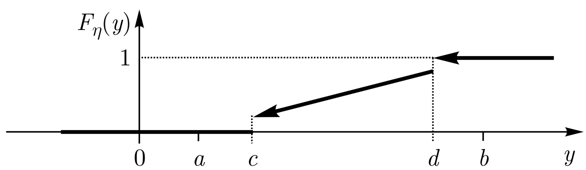


Рис. 8.2

Из рис. 8.2 легко видеть, что случайная величина η, которая содержит дискретную и непрерывную компоненту, является смешанной. Поэтому для вычисления математического ожидания случайной величины η будем использовать формулу (8.13), график функции $F_\eta(y)$ и свойства интеграла Римана–Стилтьеса. Учитывая все это, легко найдем: $\mathsf{M}\eta = \int_{-\infty}^{+\infty} y\,dF_\eta(y) = c\,\mathsf{P}(\{\omega\colon \eta(\omega) = c\}) + d\,\mathsf{P}(\{\omega\colon \eta(\omega) = d\}) + \int_c^d y\,dF_\eta(y) = c(c-a)(b-a)^{-1} + d(b-d)(b-a)^{-1} + 2^{-1}(d^2 - c^2) \times (b-a)^{-1} = 2^{-1}(c^2 + 2bd - 2ac - d^2)(b-a)^{-1}$. Используя для вычисления $\mathsf{M}\eta$ формулу (8.8), получим тот же результат. В самом деле, математическое ожидание $\mathsf{M}\eta = \mathsf{M}g(\xi) = \int_{-\infty}^{+\infty} g(x)\,f_\xi(x)\,dx = (b-a)^{-1}\left(\int_a^c c\,dx + \int_c^d x\,dx + \int_d^b d\,dx\right) = 2^{-1}(c^2 + 2\,bd - 2\,ac - d^2) \times (b-a)^{-1}$. Если компенсация полностью совпадает с затратами пенсионера на лекарства, т. е. $c = a$, $d = b$, то $\mathsf{M}\eta = 2^{-1}(a+b)$. При $c - a > b - d$ неожиданно получаем, что среднее значение компенсации больше среднего значения стоимости лекарств.

Отметим, что прямой способ вычисления математического ожидания часто бывает неэффективным. Например, когда законы распределения случайных величин чересчур громоздки или не выписываются в явном виде. В таких задачах удобнее использовать различные искусственные приемы, основанные, например, на общих свойствах математического ожидания, на представлении случайной величины в виде неслучайных функций от случайных аргументов и т. д. Рассмотрим простые примеры на применение такого рода приемов.

Пример 8.6. Каждый из десяти спортсменов независимым образом выполняет один выстрел по заданной мишени. Вероятность попадания в мишень i-м спортсменом равна $p_i = 0,8 + i/100$, где $i = 1, 2, \ldots, 10$. Найти математическое ожидание общего числа попадания в мишень.

Решение. Пусть ξ_i считает число попаданий i-тым спортсменом. Обозначим через ξ число всех попаданий при десяти выстрелах, тогда $\xi = \sum_{i=1}^{10} \xi_i$. По теореме о математическом ожидании для суммы случайных величин имеем:

$$\mathsf{M}\xi = \mathsf{M}\Big(\sum_{i=1}^{10} \xi_i\Big) = \sum_{i=1}^{10} \mathsf{M}(\xi_i) = \sum_{i=1}^{10} (0 \times (1 - p_i) + 1 \times p_i) = \\ = 0{,}81 + 0{,}82 + \ldots + 0{,}9 = 8{,}55.$$

Если решать эту задачу прямым методом, то построение вероятностной модели $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ для этого эксперимента, поточечное задание случайной величины $\xi(\omega)$ и вычисление на этой основе ее ряда распределения, хотя и не представляет особого труда, однако занимает значительное время. Еще более сложной становится эта задача по объему вычислений, когда число спортсменов велико.

Пример 8.7. На плоскости с прямоугольной системой координат $x0y$ через точку $(1, 0)$ наудачу проводится прямая b под углом ξ к оси абсцисс (см. рис. 8.3). Доказать, что математическое ожидание ординаты η точки $(0, \eta)$ пересечения этой прямой с осью $0y$ не существует.

Решение. По условию задачи предполагается, что все значения угла ξ, однозначно определяющего положение прямой b, равновозможны в интервале $(-\pi/2, \pi/2)$. Это означает, что для случайной величины ξ ее плотность $f_\xi(x) = 1/\pi$ при $x \in (-\pi/2, \pi/2)$ и равна нулю в противном случае. Из рис. 8.3 легко находим, что $\eta = g(\xi) = \operatorname{tg} \xi$, т. е. случайная величина η является неслучайной функцией от случайного аргумента ξ и принимает значения из промежутка $(-\infty, +\infty)$. Заметим, что функция $y = g(x) = \operatorname{tg} x$ в интервале $(-\pi/2, \pi/2)$ строго возрастает и имеет обратную функцию $x = \operatorname{arctg} y$. Производная от функции $x = \operatorname{arctg} y$ равна $1/(1 + y^2)$. Интегральная функция $F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) =$

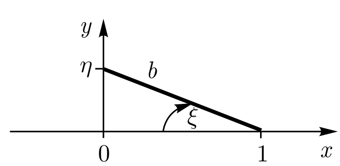


Рис. 8.3

$= \mathsf{P}(\{\omega\colon\ \operatorname{tg}(\xi(\omega)) < y\}) = \mathsf{P}(\{\omega\colon\ \xi(\omega) < \operatorname{arctg} y\}) = F_\xi(\operatorname{arctg} y)$ для любого вещественного значения y. Поэтому плотность распределения для случайной величины $\eta = \operatorname{tg}\xi$ равна $f_\eta(y) = = \dfrac{d}{dy}(F_\eta(y)) = \dfrac{d}{dy}(F_\xi(\operatorname{arctg} y)) = f_\xi(\operatorname{arctg} y)\,\dfrac{d}{dy}(\operatorname{arctg} y) = 1/\pi(1+y^2)$ при $y \in (-\infty, +\infty)$. В этом случае говорят, что случайная величина η распределена по закону Коши. Теперь для вычисления математического ожидания случайной величины η по формуле (8.8) необходимо найти значение несобственного интеграла вида:

$$\int\limits_{-\pi/2}^{\pi/2} -\pi^{-1}|\operatorname{tg} x|dx = -\pi^{-1}\int\limits_{-\pi/2}^{0} \operatorname{tg} x\,dx + \pi^{-1}\int\limits_{0}^{\pi/2} \operatorname{tg} x\,dx =$$
$$= \pi^{-1}\int\limits_{0}^{-\pi/2} \operatorname{tg} x\,dx + \pi^{-1}\int\limits_{0}^{\pi/2} \operatorname{tg} x\,dx = 2\pi^{-1}\int\limits_{0}^{\pi/2} \operatorname{tg} x\,dx =$$
$$= -2\pi^{-1}\ln|\cos x|\Big|_0^{\pi/2} = +\infty.$$

Следовательно, математическое ожидание случайной величины η, которая распределена по закону Коши, не существует.

Пример 8.8. На плоскости с прямоугольной системой координат $x0y$ случайным образом на оси $0y$ выбирается точка T, ордината ξ которой распределена по закону Коши с плотностью $1/\pi(1+y^2)$ при $y \in (-\infty, +\infty)$. Точка T и точка на оси абсцисс с координатами $(1, 0)$ соединяются отрезком b. Найти математическое ожидание расстояния η от отрезка b до начала координат (см. рис. 8.4).

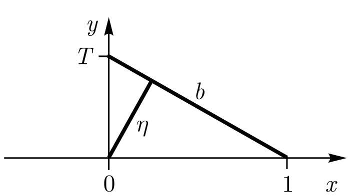


Рис. 8.4

Решение. Непосредственно из рис. 8.4 получаем, что случайная величина $\eta = g(\xi) = |\xi|\,(1+\xi^2)^{-1/2}$. Значит, случайная величина η является неслучайной функцией от случайного аргумента ξ и принимает значения из промежутка $[0, 1)$. Так как случайная величина η принимает неотрицательное значение, то ее математическое ожидание можно вычислить по формуле (8.8), т. е.

$$\mathsf{M}\eta = \int\limits_{-\infty}^{+\infty} \pi^{-1}|y|(1+y^2)^{-3/2}dy =$$
$$= -\pi^{-1}\int\limits_{-\infty}^{0} y(1+y^2)^{-3/2}dy + \pi^{-1}\int\limits_{0}^{+\infty} y(1+y^2)^{-3/2}dy =$$

$$= \pi^{-1} \int\limits_{0}^{+\infty} y(1+y^2)^{-3/2} dy + \pi^{-1} \int\limits_{0}^{+\infty} y(1+u^2)^{-3/2} dy =$$

$$= 2\pi^{-1} \int\limits_{0}^{+\infty} y(1+y^2)^{-3/2} dy = 2\pi^{-1}(-(1+y^2)^{-1/2})\big|_0^{+\infty} = 2\pi^{-1}.$$

1.6. Квантиль порядка p ($0 < p < 1$), медиана и мода. Так как вероятностная функция $\mathsf{P}(A)$, $A \in \mathcal{F}$, является безразмерной величиной, то из формулы (8.1) непосредственно следует, что случайная величина ξ и ее математическое ожидание $\mathsf{M}\xi$ имеют одинаковую размерность. Например, случайная величина и математическое ожидание измеряются в сантиметрах. Подчеркнем, что наименования для случайной величины ξ и для математического ожидания $\mathsf{M}\xi$ почти всегда опускаются в записях. Далее, значение математического ожидания $\mathsf{M}\xi$ и значения случайной величины ξ отмечаются точками на одной и той же действительной оси. При этом точки, которые соответствуют различным значениям случайной величины ξ, расположены, как слева, так и справа от точки с абсциссой $\mathsf{M}\xi$. Можно сказать, что математическое ожидание является типичным представителем множества значений случайной величины. Поэтому очень часто математическое ожидание называют характеристикой положения случайной величины. Если математическое ожидание не существует или $|\mathsf{M}\xi| = \infty$, то рассматривают другие характеристики положения случайной величины ξ. Такими характеристиками, например, являются квантиль порядка p ($0 < p < 1$), медиана и мода.

Определение 8.3. *Квантилью порядка* p ($0 < p < 1$) или *p-квантилью* случайной величины ξ называется такое число $\mathsf{K}_p\xi$, для которого имеет место соотношение $F(\mathsf{K}_p\xi) \leqslant p \leqslant F(\mathsf{K}_p\xi{+}0)$.

Так как интегральная функция $F(x)$ является неубывающей и принимает все значения от нуля до единицы, то существует по крайней мере одно такое конечное число $\mathsf{K}_p\xi$, для которого $F(\mathsf{K}_p\xi) \leqslant p$ и $F(\mathsf{K}_p\xi{+}0) \geqslant p$. На содержательном уровне выполнение этих неравенств означает, что интегральная функция распределения $F(x)$ в точке с абсциссой $\mathsf{K}_p\xi$ скачкообразно или непрерывно переходит от значений, не бо́льших p, к значениям, не меньших p. Если неравенства $F(\mathsf{K}_p\xi) \leqslant \leqslant p$, $F(\mathsf{K}_p\xi{+}0) \geqslant p$ записать в таком виде: $\mathsf{P}(\{\omega\colon \xi(\omega) < \mathsf{K}_p\xi\}) \leqslant p$, $\mathsf{P}(\{\omega\colon \xi(\omega) > \mathsf{K}_p\xi\}) \leqslant 1-p$, то число $\mathsf{K}_p\xi$ можно назвать характеристикой положения случайной величины ξ. Действительно, значения случайной величины ξ располагаются левее точки с абсциссой $\mathsf{K}_p\xi$ с вероятностью не больше p и правее этой же точки с вероятностью не больше $1-p$.

На рис. 8.5 представлен график интегральной функции распределения некоторой случайной величины. В точках разрыва $x = 0$ и

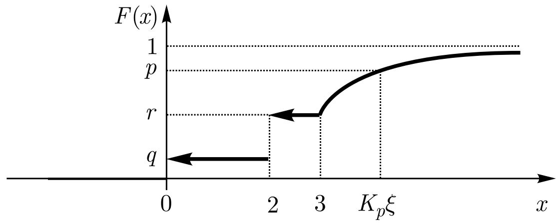


Рис. 8.5

$x = 2$ функция $F(x)$ принимает значения ноль и q соответственно, а $F(2+0) = r$. Решая неравенства $F(\mathsf{K}_p\xi) \leqslant p$, $F(\mathsf{K}_p\xi{+}0) \geqslant p$ и используя при этом рис. 8.5, легко получим, что p-квантиль случайной величины ξ: 1) равен нулю при $0 < p < q$; 2) равен любому числу $\mathsf{K}_p\xi$, где $0 \leqslant \mathsf{K}_p\xi \leqslant 2$, при $p = q$; 3) равен двум при $q < p < r$; 4) равен любому числу $\mathsf{K}_p\xi$, где $2 \leqslant \mathsf{K}_p\xi \leqslant 3$, при $p = r$. Наконец, p-квантиль случайной величины ξ при $r < p < 1$ равен $\mathsf{K}_p\xi$, где число $\mathsf{K}_p\xi$ есть единственное решение уравнения $F(\mathsf{K}_p\xi) = p$.

Итак, квантили существуют для любой случайной величины. При этом некоторые квантили могут совпадать для различных значений p, а некоторые определяются неоднозначно. Если интегральная функция $F(x)$ является непрерывной и строго монотонной, то $\mathsf{K}_p\xi$ есть однозначная функция p, обратная функции $F(x)$. Очевидно, что набор квантилей для некоторого целесообразно выбранного семейства $\{0{,}01;\ 0{,}05;\ 0{,}1;\ 0{,}25;\ 0{,}5;\ 0{,}75;\ 0{,}9;\ 0{,}95;\ 0{,}99\}$ значений p дает наглядное представление об интегральной функции распределения с достаточной степенью точности. Чаще всего используется значение p, равное 0,5. В связи с этим разумно дать следующее определение.

Определение 8.4. *Медианой* случайной величины ξ, которая обозначается через $\mathsf{Me}(\xi)$, называется квантиль порядка $p = 0{,}5$.

В дальнейшем, ради простоты, будем опускать в обозначении $\mathsf{Me}(\xi)$ круглые скобки. Из определения 8.3 следует, что медиана $\mathsf{Me}\,\xi = \mathsf{K}_{1/2}\xi$ случайной величины ξ всегда существует, иногда может быть определена неоднозначно и удовлетворяет следующим неравенствам:

$$\mathsf{P}(\{\omega\colon \xi(\omega) < \mathsf{Me}\,\xi\}) \leqslant 1/2, \qquad \mathsf{P}(\{\omega\colon \xi(\omega) > \mathsf{Me}\,\xi\}) \leqslant 1/2. \tag{8.17}$$

Это означает, что для случайной точки ξ вероятность попасть левее точки с абсциссой $\mathsf{Me}\,\xi$ не больше 0,5 и вероятность попасть правее этой точки также не больше 0,5. Отсюда можно условно сказать, что медиана характеризует некоторое центральное положение случайной величины ξ.

Более того, если вероятность $\mathsf{P}(\{\omega\colon \xi(\omega) = \mathsf{Me}\,\xi\}) = 0$, например, для непрерывной случайной величины, то из соотношения (8.17) и очевидного равенства $\mathsf{P}(\{\omega\colon \xi(\omega) < \mathsf{Me}\,\xi\}) + \mathsf{P}(\{\omega\colon \xi(\omega) > \mathsf{Me}\,\xi\}) = 1$ получаем, что $F(\mathsf{Me}\,\xi) = \mathsf{P}(\{\omega\colon \xi(\omega) < \mathsf{Me}\,\xi\}) = \mathsf{P}(\{\omega\colon \xi(\omega) > \mathsf{Me}\,\xi\}) = 1/2$. Следовательно, для таких случайных величин медиана есть абсцисса такой точки, в которой площадь под кривой плотности распределения вероятности случайной величины ξ делится пополам прямой, проходящей через эту точку и параллельной оси ординат. Для произвольной случайной величины ξ имеет место следующее утверждение.

Лемма 8.12. *Если $m = \mathsf{Me}\,\xi$ есть медиана случайной величины ξ, то*

$$\mathsf{M}(|\xi - c|) = \mathsf{M}(|\xi - m|) + 2\int\limits_c^m (x - c)\,dF_\xi(x) + \\ + (m - c)[\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant m\}) - \mathsf{P}(\{\omega\colon \xi(\omega) < m\})] \quad (8.18)$$

при $-\infty < c < m$ и

$$\mathsf{M}(|\xi - c|) = \mathsf{M}(|\xi - m|) + 2\int\limits_c^m (c - x)\,dF_\xi(x) + \\ + (c - m)[\mathsf{P}(\{\omega\colon \xi(\omega) < m\}) - \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant m\})] \quad (8.19)$$

при $m < c < +\infty$.

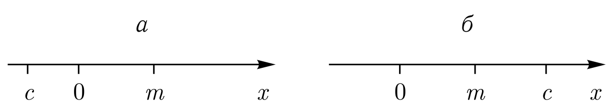


Рис. 8.6

Доказательство. Пусть сначала $c < m$. Используя формулу (8.12) для вычисления математического ожидания и рис. 8.6, *а*), легко находим:

$$\mathsf{M}(|\xi - c|) = \int\limits_{-\infty}^{+\infty} |x - c|\,dF_\xi(x) =$$

$$= \int\limits_{-\infty}^{c} |x - c|\,dF_\xi(x) + \int\limits_c^m |x - c|\,dF_\xi(x) + \int\limits_m^{+\infty} |x - c|\,dF_\xi(x) =$$

$$= \int\limits_{-\infty}^{c} (c - x)\,dF_\xi(x) + \int\limits_c^m (x - c)\,dF_\xi(x) + \int\limits_m^{+\infty} (x - c)\,dF_\xi(x) =$$

$$= \int\limits_{-\infty}^{c} (c - x)\,dF_\xi(x) + \Bigl(\int\limits_c^m x\,dF_\xi(x) - \int\limits_c^m x\,dF_\xi(x)\Bigr) +$$

$$+ \int\limits_c^m (x - c)\,dF_\xi(x) + \Bigl(\int\limits_{-\infty}^{m} m\,dF_\xi(x) - \int\limits_{-\infty}^{m} m\,dF_\xi(x)\Bigr) +$$

$$+\int\limits_{m}^{+\infty}(x-c)\,dF_\xi(x)+\int\limits_{m}^{+\infty}m\,dF_\xi(x)-\int\limits_{m}^{+\infty}m\,dF_\xi(x)=$$

$$=\int\limits_{-\infty}^{m}(m-x)\,dF_\xi(x)+\int\limits_{m}^{+\infty}(x-m)\,dF_\xi(x)+2\int\limits_{c}^{m}(x-c)\,dF_\xi(x)+$$

$$+\int\limits_{-\infty}^{m}c\,dF_\xi(x)-\int\limits_{m}^{+\infty}c\,dF_\xi(x)+\int\limits_{m}^{+\infty}m\,dF_\xi(x)-\int\limits_{-\infty}^{m}m\,dF_\xi(x)=$$

$$=\mathsf{M}(|\xi-m|)+2\int\limits_{c}^{m}(x-c)\,dF_\xi(x)+$$

$$+(m-c)\,[\mathsf{P}(\{\omega\colon\xi(\omega)\geqslant m\})-\mathsf{P}(\{\omega\colon\xi(\omega)<m\})].$$

При $m<c<+\infty$, используя рис. 8.6, *б*) и аналогичные выкладки, получим:

$$\mathsf{M}(|\xi-c|)=\int\limits_{-\infty}^{+\infty}|x-c|\,dF_\xi(x)=$$

$$=\int\limits_{-\infty}^{m}|x-c|\,dF_\xi(x)+\int\limits_{m}^{c}|x-c|\,dF_\xi(x)+\int\limits_{c}^{+\infty}|x-c|\,dF_\xi(x)=$$

$$=\int\limits_{-\infty}^{m}(c-x)\,dF_\xi(x)+\int\limits_{m}^{c}(c-x)\,dF_\xi(x)+\int\limits_{c}^{+\infty}(x-c)\,dF_\xi(x)=$$

$$=\int\limits_{-\infty}^{m}(c-x)\,dF_\xi(x)+\Big(\int\limits_{-\infty}^{m}m\,dF_\xi(x)-\int\limits_{-\infty}^{m}m\,dF_\xi(x)\Big)+$$

$$+\int\limits_{m}^{c}(c-x)\,dF_\xi(x)+\Big(\int\limits_{m}^{c}(x-m)\,dF_\xi(x)-\int\limits_{m}^{c}(x-m)\,dF_\xi(x)\Big)+$$

$$+\int\limits_{c}^{+\infty}(x-c)\,dF_\xi(x)+\int\limits_{c}^{+\infty}m\,dF_\xi(x)-\int\limits_{c}^{+\infty}m\,dF_\xi(x)=$$

$$=\mathsf{M}(|\xi-m|)+2\int\limits_{m}^{c}(c-x)\,dF_\xi(x)+\int\limits_{-\infty}^{m}c\,dF_\xi(x)-\int\limits_{m}^{+\infty}c\,dF_\xi(x)-$$

$$-\int\limits_{-\infty}^{m}m\,dF_\xi(x)+\int\limits_{m}^{+\infty}m\,dF_\xi(x)=$$

$$=\mathsf{M}(|\xi-m|)+2\int\limits_{m}^{c}(c-x)\,dF_\xi(x)+$$

$$+(c-m)\,[\mathsf{P}(\{\omega\colon\xi(\omega)<m\})-\mathsf{P}(\{\omega\colon\xi(\omega)\geqslant m\})].$$

Утверждение леммы доказано.

Лемма 8.13. *Величина* $\mathsf{M}(|\xi - c|)$ *принимает наименьшее значение, когда постоянное число* c *есть медиана.*

Доказательство. Пусть $m = \mathsf{Me}\,\xi$ одно из возможных значений медианы. Используя определение и свойства интеграла Стилтьеса [6], получим

$$\int\limits_{m}^{c} (c-x)\,dF_\xi(x) = \int\limits_{m+0}^{c} (c-x)\,dF_\xi(x) + (c-m)\,(F_\xi(m+0) - F_\xi(m)) = \\ = \int\limits_{m+0}^{c} (c-x)\,dF_\xi(x) + (c-m)\,\mathsf{P}(\{\omega\colon \xi(\omega) = m\}).$$

Подставляя это в равенство (8.19) и выполняя несложные преобразования, легко вычислим, что

$$\begin{aligned} \mathsf{M}(|\xi - c|) = \mathsf{M}(|\xi - m|) + 2 \int\limits_{m+0}^{c} (c-x)\,dF_\xi(x) + \\ + 2(c-m)\,\mathsf{P}(\{\omega\colon \xi(\omega) = m\}) + \\ + (c-m)\,[\mathsf{P}(\{\omega\colon \xi(\omega) < m\}) - \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant m\})] = \\ = \mathsf{M}(|\xi - m|) + 2 \int\limits_{m+0}^{c} (c-x)\,dF_\xi(x) + \\ + (c-m)\,[\mathsf{P}(\{\omega\colon \xi(\omega) \leqslant m\}) - \mathsf{P}(\{\omega\colon \xi(\omega) > m\})]. \end{aligned} \tag{8.20}$$

Вторые и третьи слагаемые в правых частях равенства (8.18) при $-\infty < c < m$ и соотношения (8.20) при $m < c < +\infty$, очевидно, неотрицательны. При этом постоянное число, обозначаемое через c, может совпадать с каким-нибудь другим значением медианы, если таковое значение существует. В этом случае указанные слагаемые обращаются в ноль. Отсюда непосредственно следует, что $\mathsf{M}(|\xi - c|) \geqslant \mathsf{M}(|\xi - m|)$ при $-\infty < c < +\infty$.

Доказанная лемма 8.13 используется для решения следующей интересной задачи. Случайную величину ξ, поточечное задание которой неизвестно, в простейшем случае заменяют некоторой постоянной c. Такую постоянную c называют *оценкой* (*прогнозом*) для случайной величины ξ. Абсолютная ошибка при такой замене равна случайной величине $|\xi - c|$. В качестве критерия оценивания случайной величины ξ естественно выбрать число, равное $\mathsf{M}(|\xi - c|)$. Будем говорить, что оценка c^* оптимальна в смысле выбранного критерия, если имеет место условие оптимальности: $\mathsf{M}(|\xi - c^*|) = \inf\{\mathsf{M}(|\xi - c|)\colon -\infty < c < +\infty\}$. Из леммы 8.13 следует, что $c^* = \mathsf{Me}\,\xi$. Другими словами, для случайной величины ξ любое значение медианы является оптимальной оценкой в классе действительных чисел.

Следующая числовая характеристика положения случайной величины ξ называется *модой* и обозначается через $\mathsf{Mo}(\xi)$ или, ради простоты, через $\mathsf{Mo}\,\xi$. Эта характеристика определяется только для непрерывных и дискретных случайных величин.

Определение 8.5. *Модой* $\mathsf{Mo}\,\xi$ непрерывной случайной величины ξ называется такое ее значение $x_{\max}$, которое доставляет плотности распределения $f_\xi(x)$ локальный максимум.

Пусть возможные значения дискретной случайной величины ξ занумерованы таким способом, что большему ее значению соответствует больший номер. Значит, если x'_k и x'_{k+1} возможные значения дискретной случайной величины ξ, то вероятности $p_k = \mathsf{P}(\{\omega\colon \xi(\omega) = x'_k\}) > 0$, $p_{k+1} = \mathsf{P}(\{\omega\colon \xi(\omega) = x'_{k+1}\}) > 0$ и $x'_k < x'_{k+1}$. Обозначим через N_ξ множество номеров k, которые таким способом перечисляют все различные значения дискретной случайной величины ξ.

Определение 8.6. *Модой* $\mathsf{Mo}\,\xi$ дискретной случайной величины называется такое ее возможное значение x'_i, что $p_{i-1} \leqslant p_i$ и $p_i \geqslant p_{i+1}$.

Если мода единственна, то распределение случайной величины называется *унимодальным*. Если случайная величина имеет несколько мод, то ее распределение называется *мультимодальным*. Наконец, целесообразно рассмотреть новую числовую характеристику положения, которая называется *наивероятнейшим значением случайной величины* ξ и обозначается через $\mathsf{Mo}^*\xi$. Для непрерывной случайной величины ξ наивероятнейшее значение $\mathsf{Mo}^*\xi$ по определению удовлетворяет условию $f_\xi(\mathsf{Mo}^*\xi) \geqslant f_\xi(x)$ для всех $x \in R$, т. е. $f_\xi(\mathsf{Mo}^*\xi)$ является наибольшим значением плотности $f_\xi(x)$. Для дискретной величины ξ наивероятнейшее значение $\mathsf{Mo}^*\xi$ совпадает с ее возможным значением x'_r, для которого $p_r = \mathsf{P}(\{\omega\colon \xi(\omega) = x'_r\}) = \max\{p_k\colon k \in N_\xi\}$. На содержательном уровне характеристика $\mathsf{Mo}^*\xi$ означает, что вероятность попадания ξ в некоторую окрестность точки с абсциссой $\mathsf{Mo}^*\xi$, как правило, будет не меньше, чем вероятность ее попадания в любую окрестность той же длины, которая не содержит эту точку.

Пример 8.9. Интегральная функция распределения случайной величины ξ равна $F(x) = a + b\,\mathrm{arctg}\,x$ при $-\infty < x < +\infty$. Найти числовые характеристики положения этой случайной величины.

Решение. Используя свойства $F(-\infty) = 0$, $F(+\infty) = 1$ интегральной функции распределения, получим равенства: $a - 2^{-1}b\pi = 0$, $a + 2^{-1}b\pi = 1$. Отсюда легко вычислим, что $a = 1/2$ и $b = 1/\pi$. Значит, для $x \in (-\infty, +\infty)$ интегральная функция распределения $F(x) = 2^{-1} + \pi^{-1}\,\mathrm{arctg}\,x$, а плотность вероятности $f(x) = 1/\pi(1 + x^2)$ и представлена на рис. 8.7.

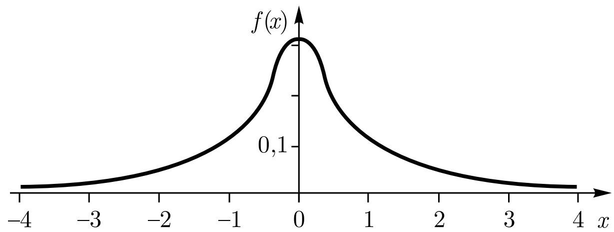


Рис. 8.7

Следовательно, случайная величина ξ распределена по закону Коши. При решении примера 8.7 было установлено, что математическое ожидание такой случайной величины не существует. Медиану $\mathsf{Me}\,\xi$ непрерывной случайной величины найдем из условия: $F(\mathsf{Me}\,\xi) = 2^{-1} + \pi^{-1}\operatorname{arctg}(\mathsf{Me}\,\xi) = 1/2$, откуда $\mathsf{Me}\,\xi = 0$. Для определения моды $\mathsf{Mo}\,\xi$ вычислим подозрительные точки на локальный максимум плотности $f(x)$. Производная от функции $f(x)$ равна $(-2x)\,\pi^{-1}(1 + x^2)^{-2}$ при всех $x \in (-\infty, +\infty)$. Следовательно, уравнение $(-2x)\,\pi^{-1}(1 + x^2)^{-2} = 0$ имеет единственное решение, равное нулю. Очевидно, что при $x = 0$ функция достигает локального максимума. Окончательно, находим, что $\mathsf{Mo}\,\xi = 0$. Более того, для этой случайной величины $\mathsf{Mo}^*\xi = \mathsf{Mo}\,\xi = 0$.

Пример 8.10. Вычислить числовые характеристики положения дискретной случайной величины ξ, для которой ряд распределения имеет вид:

Таблица 8.1

ξ	$x'_1 = -4$	$x'_2 = -1$	$x'_3 = 0$	$x'_4 = 5$
$\mathsf{P}(\cdot)$	$p_1 = 4\,p$	$p_2 = 2{,}5\,p$	$p_3 = 8\,p^2$	$p_4 = 0{,}5\,p$

Решение. Неизвестное значение для параметра p найдем из условия нормировки $4\,p + 2{,}5\,p + 8\,p^2 + 0{,}5\,p = 1$. Решая уравнение $8\,p^2 + 7\,p = 1$ и неравенства $0 < p < 1$, получим единственное значение $p = 1/8$. Следовательно, ряд распределения примет вид:

Таблица 8.2

ξ	$x'_1 = -4$	$x'_2 = -1$	$x'_3 = 0$	$x'_4 = 5$
$\mathsf{P}(\cdot)$	$p_1 = 1/2$	$p_2 = 5/16$	$p_3 = 1/8$	$p_4 = 1/16$

и $\mathsf{M}\xi = (-4) \times (1/2) + (-1) \times (5/16) + 0 \times (1/8) + 5 \times (1/16) = -2$. На рис. 8.8 представлен график интегральной функции распределения. Непосредственно из этого рисунка получаем, что $\mathsf{Me}\,\xi$ есть любое число $x_{1/2}$, удовлетворяющее неравенствам $-4 \leqslant x_{1/2} \leqslant -1$.

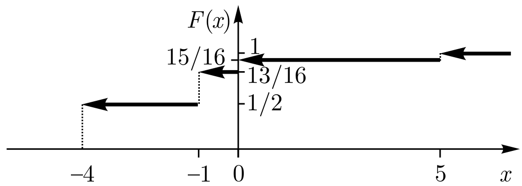


Рис. 8.8

Заметим, что для этой задачи $N_\xi = \{1, 2, 3, 4\}$. Так как неравенства $p_{k-1} \leqslant p_k$ и $p_k \geqslant p_{k+1}$ не имеют решения при $k = 2$ и при $k = 3$, то мода для этой случайной величины не существует. Легко проверить, что вероятность $p_1 = \mathsf{P}(\{\omega\colon \xi(\omega) = x_1'\}) = \max\{p_k\colon k \in N_\xi\} = 1/2$. Поэтому наивероятнейшее значение случайной величины ξ равно $\mathsf{Mo}^*\xi = x_1' = -4$.

§ 2. Характеристики степени разброса одномерных случайных величин

2.1. Дисперсия, среднее квадратичное отклонение и среднее отклонение случайной величины. Математическое ожидание, квантили, медиана и мода являются важными числовыми характеристиками положения случайной величины. Эти характеристики отвечают за некоторые типичные значения случайной величины. Каждая из этих числовых характеристик есть абсцисса некоторой точки на действительной оси, относительно которой происходит рассеивание или разброс случайной величины. При этом значение каждой из этих числовых характеристик ничего не говорит о величине этого рассеивания. Может оказаться, что случайные величины имеют, например, одинаковое математическое ожидание, но разные законы распределения. Рассмотрим тривиальный пример такой ситуации.

Пример 8.11. Некто имеет возможность купить только одну из двух типов акций. Первый тип акции стоит 100 рублей, а второй 100 000 рублей. По истечении некоторого времени некто с равной вероятностью получает доход в 10 % от стоимости акции или стоимость его акции уменьшается на 10 %. Необходимо на содержательном уровне выяснить, какой из двух типов акций, возможно, следует предпочесть в случае покупки таковой. Обозначим через ξ_1 и через ξ_2 случайный доход от покупки акции первого и соответственно второго типов. Из условий этой задачи найдем законы распределения вероятностей для дискретных случайных величин ξ_1 и ξ_2:

$$\mathsf{P}_{\xi_1}(\xi_1 = -10) = \mathsf{P}_{\xi_1}(\xi_1 = 10) = 1/2,$$
$$\mathsf{P}_{\xi_2}(\xi_2 = -10\,000) = \mathsf{P}_{\xi_2}(\xi_2 = 10\,000) = 1/2.$$

Итак, получили разные законы распределения вероятностей. При этом $\mathsf{M}\xi_1 = \mathsf{M}\xi_2 = 0$. Следовательно, математическое ожидание доходности от покупки каждого типа акций одинаковое. Однако не каждый рискнет с вероятностью 0,5 потерять 10 000 рублей из своего актива в результате покупки акции второго типа.

Такие случайные величины иногда можно различать по тому, как далеко и с какой вероятностью каждая из них отклоняется от математического ожидания. С этой целью перейдем к рассмотрению важнейших характеристик рассеивания или разброса случайной величины. Для характеристики рассеивания естественно рассмотреть разность $\xi - \mathsf{M}\xi$ между случайной величиной ξ и ее математическим ожиданием $\mathsf{M}\xi$. Разумеется, в этом случае всегда предполагается, что математическое ожидание $\mathsf{M}\xi$ не только существует, но и конечно. Тогда новая случайная величина $\overset{\circ}{\xi} = \xi - \mathsf{M}\xi$ называется *центрированной случайной величиной*, а значение $\xi - \mathsf{M}\xi$ — *отклонением случайной величины от ее математического ожидания*. Среднее значение случайной величины $\overset{\circ}{\xi}$ равно нулю для случайной величины ξ. Действительно, $\mathsf{M}\overset{\circ}{\xi} = \mathsf{M}(\xi - \mathsf{M}\xi) = \mathsf{M}\xi - \mathsf{M}\xi = 0$, поэтому $\mathsf{M}\overset{\circ}{\xi}$ не может служить характеристикой степени рассеивания возможных значений случайной величины ξ около ее математического ожидания.

Определение 8.7. *Дисперсией* случайной величины ξ или *характеристикой степени разброса возможных значений случайной величины около ее математического ожидания* называется среднее значение квадрата отклонения $(\xi - \mathsf{M}\xi)$ случайной величины от ее математического ожидания.

Обозначают эту константу символом $\mathsf{D}(\xi)$ или, ради простоты, через $\mathsf{D}\xi$. Итак, из определения дисперсии можно написать, что дисперсия $\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2$. Вспоминая вопрос о наилучшем приближении случайной величины ξ математическим ожиданием $\mathsf{M}\xi$ и определение дисперсии, можно утверждать, что дисперсия $\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2$ является как мерой ошибки такого приближения, так и мерой разброса случайной величины. Если дисперсия мала, то случайная величина имеет тенденцию принимать значения, близкие к ее математическому ожиданию. Поэтому, чем меньше дисперсия (рассеивание случайной величины), тем точнее и легче можно предсказать значения случайной величины ξ. Рассмотрим теперь интерпретацию дисперсии с точки зрения классической механики. С этой целью распределим единичную массу некоторого вещества вдоль действительной оси в соответствии со значением случайной величины ξ. Это означает, что количество массы вещества на любом фиксированном промежутке $[a, b)$ действительной оси равно вероятности попадания случайной величины ξ в этот промежуток. Тогда дисперсия случайной величины ξ является моментом инерции данного распределения масс относительно центра масс. Таков содержательный смысл числовой характеристики $\mathsf{D}\xi$ с точки зрения классической механики.

Интересно отметить, что дисперсия случайной величины ξ и дисперсия соответствующей ей центрированной случайной величины $(\xi - \mathsf{M}\xi)$ совпадают. В самом деле, так как математическое ожидание $\mathsf{M}\overset{\circ}{\xi} = 0$, то $\mathsf{D}\overset{\circ}{\xi} = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \mathsf{D}\xi$. Дисперсия имеет размерность квадрата случайной величины. Поэтому часто для удобства в качестве меры разброса пользуются величиной, размерность которой совпадает с размерностью случайной величины. В этом случае рассматривают среднее квадратичное отклонение $\boldsymbol{\sigma}(\xi) = \boldsymbol{\sigma}\xi$ случайной величины, которое определяется формулой: $\boldsymbol{\sigma}(\xi) = (\mathsf{D}\xi)^{1/2}$. Среднее квадратичное отклонение случайной величины называют *стандартом*.

Если математическое ожидание случайной величины не существует или принимает достаточно большое значение, то в качестве меры разброса рассматривают величину $\mathsf{E}\xi = \mathsf{M}(|\xi - \mathsf{Me}\,\xi|)$. Мера рассеивания $\mathsf{E}\xi$ называется *средним отклонением*. Медиана $\mathsf{Me}\,\xi$ характеризуется тем, что при замене случайной величины ξ постоянной c совершается ошибка $\mathsf{M}(|\xi - c|)$, которая принимает наименьшее значение при $c = \mathsf{Me}\,\xi$. Если рассеивание $\mathsf{E}\xi = \mathsf{M}(|\xi - \mathsf{Me}\,\xi|)$ мало́, то случайная величина в основном группируется около значения медианы. Поэтому целесообразно разброс измерять величиной $\mathsf{M}(|\xi - \mathsf{Me}\,\xi|)$. К сожалению, медиана $\mathsf{Me}\,\xi$ может определяться неоднозначно. Более того, все приемы вычисления, которые связаны с величиной рассеивания вида $\mathsf{M}(|\xi - \mathsf{Me}\,\xi|)$, как правило, являются далеко непростыми из вхождения модуля.

Пример 8.12. Найти дисперсию, среднеквадратическое отклонение и среднее отклонение числа появлений решетки несимметричной монеты в одном испытании, если вероятность появления решетки равна p.

Решение. Обозначим через ξ число появлений решетки в одном испытании, тогда случайная величина ξ имеет ряд распределения (табл. 8.3):

Таблица 8.3

ξ	0	1
$\mathsf{P}(\cdot)$	$1-p$	p

Математическое ожидание этой величины равно $\mathsf{M}\xi = p$. Дисперсию случайной величины ξ можно вычислить следующим способом: $\mathsf{D}\xi = \mathsf{M}(\xi - p)^2 = \mathsf{M}(\xi^2 - 2\,p\,\xi + p^2) = \mathsf{M}(\xi^2) - 2\,p\,\mathsf{M}\xi + p^2 = \mathsf{M}(\xi^2) - p^2$, $\mathsf{M}(\xi^2) = 0^2(1-p) + 1^2 p = p$, $\mathsf{D}\xi = \mathsf{M}(\xi^2) - p^2 = p - p^2 = p(1-p) = pq$, где $q = 1 - p$. Рассмотрим несколько другой способ подсчета дисперсии случайной величины ξ. Случайную величину $(\xi - p)^2$ можно рассматривать как неслучайную функцию $g(x) = (x - p)^2$ от случайной переменной ξ. Поэтому дисперсию, которая является математическим ожиданием от случайной величины $g(x) = (x - p)^2$, можно вычислить с использованием формулы (8.5) следующим образом: $\mathsf{D}\xi = \mathsf{M}g(\xi) = (0 - p)^2(1 - p) + (1 - p)^2 p = pq$. Среднее квадратическое отклонение для рассматриваемой случайной величины ξ равно $\sigma(\xi) = (\mathsf{D}\xi)^{1/2} = (pq)^{1/2}$.

На рисунках 8.9, *а* и 8.9, *б* представлен график интегральной функции распределения $F(x)$ числа появлений решетки несимметричной монеты при одном броске для $0{,}5 < p < 1$ и $0 < p < 0{,}5$ соответственно.

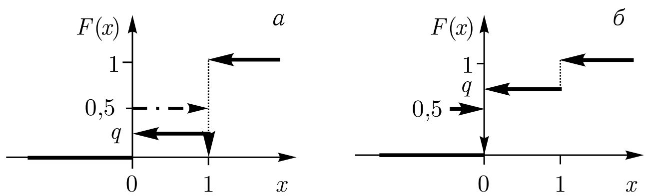


Рис. 8.9

Используя соотношение (8.17) и рис. 8.9 получим, что медиана $\mathsf{Me}\,\xi$ числа появлений решетки несимметричной монеты в одном испытании равна единицы при $0{,}5 < p < 1$ (рис. 8.9, *а*) и равна нулю при $0 < p < 0{,}5$ (рис. 8.9, *б*). Отсюда получаем, что среднее отклонение $\mathsf{E}\xi$ равно $|0-1|\,q + |1-1|\,p = q$ при $0{,}5 < p < 1$ и равно $|0-0|\,q + |1-0|\,p = p$ при $0 < p < 0,5$. Итак, $\mathsf{E}\xi = \min\{p, q\}$. Отметим, что в этой задаче случайная величина ξ принимает всего два значения. Поэтому мода для этой случайной величины не существует. Однако наивероятнейшее значение $\mathsf{Mo}^*\xi$ случайной величины ξ равно единице при $0{,}5 < p < 1$ и равно нулю при $0 < p < 0{,}5$. Значит, для этого примера $\mathsf{Mo}^*\xi = \mathsf{Me}\,\xi$. Так как $(pq)^{1/2} > \min\{p, q\}$, то для этого простого эксперимента стандарт $\sigma(\xi)$ больше среднего отклонения $\mathsf{E}\xi$. Следовательно, в игре в орлянку с использованием несимметричной монеты точнее значения случайной величины ξ будет предсказывать медиана. Это соответствует практике, когда игрок, зная асимметрию монеты, делает ставку на решетку при $p > q$ и на герб в противном случае.

2.2. Свойства дисперсии. Рассмотрим простейшие свойства дисперсии.

Д.1. *Дисперсия постоянной величины равна нулю*, т. е. $\mathsf{D}(c) = 0$. В самом деле, имеем: $\mathsf{D}(c) = \mathsf{M}(c - \mathsf{M}c)^2 = \mathsf{M}(c - c)^2 = \mathsf{M}0 = 0$.

Д.2. *Постоянную величину можно выносить за знак дисперсии, возводя эту величину в квадрат*, т. е. $\mathsf{D}(c\xi) = c^2\mathsf{D}(\xi)$. Действительно, последовательно найдем: $\mathsf{D}(c\xi) = \mathsf{M}(c\xi - \mathsf{M}(c\xi))^2 = \mathsf{M}(c\xi - c\mathsf{M}\xi)^2 = \mathsf{M}(c^2(\xi - \mathsf{M}\xi)^2) = c^2\mathsf{M}(\xi - \mathsf{M}\xi)^2 = c^2\mathsf{D}(\xi)$. Так как $\boldsymbol{\sigma}(\xi) = (\mathsf{D}\xi)^{1/2}$ и $\mathsf{D}(c\xi) = c^2\mathsf{D}\xi$, то, извлекая корень из обеих частей последнего равенства, получим $\boldsymbol{\sigma}(c\xi) = |c|\,\boldsymbol{\sigma}(\xi)$.

Д.3. *Если случайные величины ξ_1 и ξ_2 независимы и математические ожидания $\mathsf{M}\xi_1$, $\mathsf{M}\xi_2$ конечны, то дисперсия суммы этих случайных величин будет равна сумме дисперсий*, т. е. $\mathsf{D}(\xi_1 + \xi_2) = \mathsf{D}\xi_1 + \mathsf{D}\xi_2$.

Доказательство. Вспоминая определение дисперсии и свойства математического ожидания, найдем:

$$\begin{aligned}\mathsf{D}(\xi_1 + \xi_2) &= \mathsf{M}\{(\xi_1 + \xi_2) - \mathsf{M}(\xi_1 + \xi_2)\}^2 = \mathsf{M}\{(\xi_1 - \mathsf{M}\xi_1) + (\xi_2 - \mathsf{M}\xi_2)\}^2 = \\ &= \mathsf{M}(\xi_1 - \mathsf{M}\xi_1)^2 + 2\,\mathsf{M}(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2) + \mathsf{M}(\xi_2 - \mathsf{M}\xi_2)^2 = \\ &= \mathsf{D}\xi_1 + \mathsf{D}\xi_2 + 2\,\mathsf{M}(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2).\end{aligned}$$

Поскольку случайные величины ξ_1 и ξ_2 независимы, то центрированные случайные величины $\overset{\circ}{\xi}_1 = \xi_1 - \mathsf{M}\xi_1$ и $\overset{\circ}{\xi}_2 = \xi_2 - \mathsf{M}\xi_2$ тоже независимы. В самом деле, при любых действительных x и y, учитывая независимость ξ_1 и ξ_2, получим, что вероятность

$$\begin{aligned}&\mathsf{P}(\overset{\circ}{\xi}_1 < x, \overset{\circ}{\xi}_2 < y) = \mathsf{P}(\xi_1 - \mathsf{M}\xi_1 < x, \xi_2 - \mathsf{M}\xi_2 < y) = \\ &= \mathsf{P}(\xi_1 < x + \mathsf{M}\xi_1, \xi_2 < y + \mathsf{M}\xi_2) = \mathsf{P}(\xi_1 < x + \mathsf{M}\xi_1) \times \mathsf{P}(\xi_2 < y + \mathsf{M}\xi_2) = \\ &\qquad = \mathsf{P}(\xi_1 - \mathsf{M}\xi_1 < x) \times \mathsf{P}(\xi_2 - \mathsf{M}\xi_2 < y) = \mathsf{P}(\overset{\circ}{\xi}_1 < x) \times \mathsf{P}(\overset{\circ}{\xi}_2 < y).\end{aligned}$$

Итак, $\overset{\circ}{\xi}_1$ и $\overset{\circ}{\xi}_2$ независимы. Отсюда имеем $\mathsf{M}((\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)) = \mathsf{M}(\xi_1 - \mathsf{M}\xi_1) \times \mathsf{M}(\xi_2 - \mathsf{M}\xi_2) = 0$ и $\mathsf{D}(\xi_1 + \xi_2) = \mathsf{D}\xi_1 + \mathsf{D}\xi_2$. Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ попарно независимы и $|\mathsf{M}\xi_i| < \infty$ для всех $i \in \{1, 2, \ldots, n\}$, то аналогичным образом доказывается равенство $\mathsf{D}(\xi_1 + \xi_2 + \ldots + \xi_n) = \mathsf{D}\xi_1 + \mathsf{D}\xi_2 + \ldots + \mathsf{D}\xi_n$.

Д.4. *Дисперсия неотрицательна*, т. е. $\mathsf{D}(\xi) \geqslant 0$. Легко видеть, что $\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \int_{-\infty}^{+\infty}(\xi - \mathsf{M}\xi)^2 dF(x)$. Так как $(\xi - \mathsf{M}\xi)^2 \geqslant 0$ и $F(x)$ — неубывающая функция, то всякое приращение $dF(x) \geqslant 0$, и поэтому $\mathsf{D}(\xi) \geqslant 0$.

Случайная величина $\overline{\xi} = (\xi - \mathsf{M}\xi)/\boldsymbol{\sigma}(\xi)$ называется *центрированной и нормированной случайной величиной*, соответствующей случайной величине ξ. Для центрированной и нормированной случайной величины справедливы соотношения: $\mathsf{D}(\overline{\xi}) = 0$, $\mathsf{D}(\overline{\xi}) = 1$. В самом деле, легко находим: $\mathsf{M}(\overline{\xi}) = \mathsf{M}((\xi - \mathsf{M}\xi)/\boldsymbol{\sigma}(\xi)) = (1/\boldsymbol{\sigma}(\xi))\mathsf{M}(\xi - \mathsf{M}\xi) = (1/\boldsymbol{\sigma}(\xi))\mathsf{M}\overset{\circ}{\xi} = 0$, $\mathsf{D}(\overline{\xi}) = \mathsf{D}((\xi - \mathsf{M}\xi)/\boldsymbol{\sigma}(\xi)) = (1/\boldsymbol{\sigma}^2(\xi))\,\mathsf{D}(\xi - \mathsf{M}\xi) = (1/\mathsf{D}\xi)\,\mathsf{D}\xi = 1$.

Д.5. Приведем теперь формулы для вычисления дисперсии:

- поскольку дисперсия $\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \mathsf{M}\eta$, где $\eta = g(\xi) = (\xi - \mathsf{M}\xi)^2$, то из формулы (8.5) для дискретной случайной величины ξ получим: $\mathsf{D}\xi = \sum_{i=1}^{m} g(x_i')\, p_i = \sum_{i=1}^{m} (x_i' - \mathsf{M}\xi)^2 p_i$, если число возможных значений x_1', x_2', ..., x_m' для случайной величины $\xi(\omega)$ конечно, и по формуле (8.6) найдем: $\mathsf{D}\xi = \sum_{i=1}^{\infty} g(x_i')\, p_i = \sum_{i=1}^{\infty} (x_i' - \mathsf{M}\xi)^2 p_i$, если число таких значений для $\xi(\omega)$ бесконечно;
- если ξ — непрерывная случайная величина, то аналогично (18.8) из (8.8) имеем $\mathsf{D}\xi = \int_{-\infty}^{+\infty} g(x)\, f(x)\, dx = \int_{-\infty}^{+\infty} (x - \mathsf{M}\xi)^2 f(x)\, dx$. В общем случае дисперсия вычисляется при помощи формулы (8.12) с использованием интеграла Стилтьеса, т. е. $\mathsf{D}\xi = \int_{-\infty}^{+\infty} (x - \mathsf{M}\xi)^2\, dF(x)$. При этом, как правило, приведенные здесь ряд и интегралы сходятся, в противном случае говорят, что случайная величина не имеет конечной дисперсии;
- наконец, имеем $\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \mathsf{M}(\xi^2 - 2\,\xi\,\mathsf{M}\xi + (\mathsf{M}\xi)^2) = \mathsf{M}(\xi^2) - (\mathsf{M}\xi)^2$. Итак, получили получили новую формулу $\mathsf{D}\xi = \mathsf{M}(\xi^2) - (\mathsf{M}\xi)^2$ для вычисления дисперсии. Так как $\mathsf{D}\xi \geqslant 0$ и $\mathsf{D}\xi = \mathsf{M}(\xi^2) - (\mathsf{M}\xi)^2$, то $\mathsf{M}(\xi^2) \geqslant (\mathsf{M}\xi)^2$.

Перейдем к рассмотрению важных неравенств Чебышева, которые оценивают значения вероятностей случайных событий с использованием дисперсии.

Теорема 8.5. *Для любой случайной величины ξ и любого $\varepsilon > 0$ справедливо первое неравенство Чебышева*: $\mathsf{P}(\{\omega\colon |\xi(\omega)| \geqslant \varepsilon\}) \leqslant \mathsf{M}(\xi^2)/\varepsilon^2$.

Доказательство. Введем случайную величину $\eta(\omega)$ с помощью следующего равенства:

$$\eta(\omega) = g(\xi(\omega)) = \begin{cases} 0 & \text{при} \ \ \omega \in \{ : |\xi(\omega)| < \varepsilon\}, \\ \varepsilon & \text{при} \ \ \omega \in \{ : |\xi(\omega)| \geqslant \varepsilon\}. \end{cases}$$

Простая случайная величина $\eta(\omega)$ принимает два значения 0 и ε. Ясно, что $|\xi(\omega)| \geqslant |\eta(\omega)|$. Поэтому $\xi^2 \geqslant \eta^2$, и используя свойство Д.5, имеем $\mathsf{M}(\xi^2) \geqslant \mathsf{M}(\eta^2)$. Этот результат можно получить иначе. Так как $\xi^2 - \eta^2 \geqslant 0$, то $\mathsf{M}(\xi^2 - \eta^2) \geqslant 0$. Применяя теперь свойство Д.3, выводим: $\mathsf{M}(\xi^2) - \mathsf{M}(\eta^2) \geqslant 0$ и снова $\mathsf{M}(\xi^2) \geqslant \mathsf{M}(\eta^2)$. Рассмотрим новую случайную величину $\theta = g_1(\eta) = \eta^2$ и найдем $\mathsf{M}(\theta)$: $\mathsf{M}(\theta) = 0^2 \times \mathsf{P}(\{\omega\colon |\xi(\omega)| < \varepsilon\}) + \varepsilon^2 \mathsf{P}(\{\omega\colon |\xi(\omega)| \geqslant \varepsilon\}) = \varepsilon^2 \mathsf{P}(|\xi(\omega)| \geqslant \varepsilon)$. Используя неравенство $\mathsf{M}(\xi^2) \geqslant \mathsf{M}(\theta)$ и равенство $\mathsf{M}(\theta) = \varepsilon^2 \mathsf{P}(|\xi(\omega)| \geqslant \varepsilon)$, получаем искомое утверждение, а именно $\mathsf{P}(|\xi(\omega)| \geqslant \varepsilon) \leqslant \mathsf{M}(\xi^2)/\varepsilon^2$.

Теорема 8.6. *Для любой случайной величины ξ и любого $\varepsilon > 0$ справедливо второе неравенство Чебышева* $\mathsf{P}(\{\omega\colon |\xi - \mathsf{M}\xi| \geqslant \varepsilon\}) \leqslant \mathsf{D}\xi/\varepsilon^2$, *если только существует конечное математическое ожидание* $\mathsf{M}\xi$.

Доказательство. В самом деле, если рассматривать в качестве ξ центрированную случайную величину $\overset{\circ}{\xi} = \xi - \mathsf{M}\xi$, то первое неравенство Чебышева в этом случае примет вид: $\mathsf{P}(|\overset{\circ}{\xi}| \geqslant \varepsilon) \leqslant \leqslant \mathsf{M}((\overset{\circ}{\xi})^2)/\varepsilon^2$, или, иначе, $\mathsf{P}(|\overset{\circ}{\xi}| \geqslant \varepsilon) \leqslant \mathsf{D}(\overset{\circ}{\xi})/\varepsilon^2$. Так как $\mathsf{D}(\overset{\circ}{\xi}) = \mathsf{D}(\xi)$, то $\mathsf{P}(|\xi - \mathsf{M}\xi| \geqslant \varepsilon) \leqslant \varepsilon^{-2}\mathsf{D}(\xi)$, что и требовалось доказать.

Заметим, что при $\varepsilon \leqslant \sigma$, $\sigma^2 = \mathsf{D}\xi$, получаем $\varepsilon^{-2}\mathsf{D}(\xi) \geqslant 1$, и второе неравенство Чебышева дает тривиальную оценку $\mathsf{P}(|\xi - \mathsf{M}\xi| \geqslant \varepsilon) \leqslant 1$. Поэтому второе неравенство Чебышева применяют при $\varepsilon > \sigma$. Рассмотрим простой пример на применение второго неравенства Чебышева.

Пример 8.13. Случайная величина ξ имеет математическое ожидание $\mathsf{M}\xi = a$ и дисперсию $\mathsf{D}\xi = \sigma^2$. Оценить сверху вероятность того, что ξ отклонится от своего математического ожидания не меньше, чем на $\varepsilon = 3\,\sigma$.

Решение. Используя второе неравенство Чебышева, оценим вероятность $\mathsf{P}(\{\omega\colon |\xi - a| \geqslant 3\,\sigma\}) \leqslant \sigma^2/9\,\sigma^2 = 1/9$.

Такова оценка сверху для произвольного закона распределения случайной величины ξ, имеющей конечные математическое ожидание и дисперсию. Для многих случайных величин эта оценка является грубой. Например, рассмотрим случайную величину ξ, которая принимает значение -2 с вероятностью 0,5 и значение $+2$ с вероятностью 0,5. Так как $\mathsf{M}\xi = 0$, $\sigma\xi = 2$ и $\{\omega\colon |\xi| = 2\} = \Omega$, то получаем $\mathsf{P}(\{\omega\colon |\xi| \geqslant 3\,\sigma\}) = = \mathsf{P}(\{\omega\colon |\xi| \geqslant 6\}) = 0$. Этот факт показывает грубость оценки, которая получена с помощью неравенства Чебышева. С другой стороны, при $\varepsilon = 2$ вероятность вида $\mathsf{P}(\{\omega\colon |\xi| \geqslant 2\}) \leqslant \mathsf{D}\xi/\varepsilon^2 = 4/4 = 1$, и эту оценку уже улучшить нельзя, так как вероятность $\mathsf{P}(\{\omega\colon |\xi| \geqslant 2\}) = = \mathsf{P}(\{\omega\colon |\xi| = 2\}) = 1$. Следует отметить, что в теории вероятностей широко используется, как первое, так и второе неравенства Чебышева.

2.3. Начальные и центральные моменты k-го порядка, коэффициент асимметрии, эксцесс. На практике очень часто вводят так называемые вспомогательные числовые характеристики случайных величин. К вспомогательным числовым характеристикам, прежде всего, относятся начальные и центральные моменты случайной величины. Такого рода числовые характеристики учитывают влияние маловероятных, но больших по абсолютной величине значений случайной величины. Кроме того, с помощью этих характеристик определяют некоторые специфические свойства законов распределения, например, асимметрию плотности распределения для непрерывной случайной величины. Перейдем к рассмотрению начальных и центральных моментов случайной величины.

Определение 8.8. *Начальным моментом k-го порядка случайной величины* ξ называется математическое ожидание k-й степени случайной величины, и он обозначается через $\alpha_k(\xi)$ или, ради простоты, через $\alpha_k\xi$, т. е. $\alpha_k\xi = \mathsf{M}(\xi^k) = \mathsf{M}(g(\xi))$, где $g(\xi) = \xi^k$ и $k = 0, 1, \ldots$.

Нетрудно видеть, что $\alpha_0(\xi) = \mathsf{M}(\xi^0) = \mathsf{M}(1) = 1$ и $\alpha_1\xi = \mathsf{M}(\xi^1) = \mathsf{M}(\xi)$. При $k \geqslant 2$ получим новые числовые характеристики случайной величины. Для дискретных и непрерывных случайных величин с использованием формул (8.6) и (8.8) имеем $\alpha_k(\xi) = \mathsf{M}(\xi^k) = \sum_i (x_i')^k p_i$ и $\alpha_k(\xi) = \int_{-\infty}^{+\infty} x^k f(x)\,dx$, $k = 0, 1, \ldots$. В общем случае начальный момент k-го порядка вычисляется при помощи формулы (8.12) с использованием интеграла Стилтьеса, т. е. $\alpha_k\xi = \int_{-\infty}^{+\infty} x^k dF(x)$. При этом приведенные здесь ряд и интегралы должны абсолютно сходиться. В противном случае начальные моменты, или не существуют, или не имеют конечного значения.

На примере дискретной случайной величины легко проиллюстрировать причину введения таких числовых характеристик, как начальные моменты. При подсчете собственно математического ожидания случайной величины ξ фактически существенно учитываем первые n значений случайной величины. В самом деле, так как $\sum_i p_i = 1$, то $\lim\limits_{i\to\infty} p_i = 0$. Поэтому при подсчете математического ожидания $\mathsf{M}\xi = \sum_i x_i' p_i = \sum_{i\leqslant n} x_i' p_i + \sum_{i>n} x_i' p_i$ остаток ряда $\sum_{i>n} x_i' p_i$ для некоторых случайных величин может принимать маленькое значение за счет небольших значений вероятностей p_i при $i > n$. Для того чтобы учесть большие по абсолютной величине значения величины ξ и вводятся начальные моменты, при вычислении которых значения случайной величины возводятся в k-ю степень. При этом число k может быть выбрано достаточно большим. Часто говорят, что начальные моменты учитывают «хвосты» случайных величин.

Если математическое ожидание случайной величины ξ принимает конечное значение, то наряду с начальными моментами рассматривают семейство новых вспомогательных числовых характеристик, которые называются центральными моментами k-го порядка.

Определение 8.9. *Центральным моментом k-го порядка случайной величины* ξ при $|\mathsf{M}\xi| < \infty$ называется математическое ожидание k-й степени отклонения случайной величины от ее математического ожидания, и он обозначается через $\beta_k(\xi)$ или, ради простоты, через $\beta_k\xi$, где $k = 0, 1, \ldots$.

Другими словами, центральный момент k-го порядка есть математическое ожидание k-й степени центрированной случайной величины $\overset{\circ}{\xi}$. Тогда $\beta_k(\xi) = \mathsf{M}(\xi - \mathsf{M}\xi)^k$. Легко найти, что $\beta_0(\xi) = \mathsf{M}(\xi - \mathsf{M}\xi)^0 = 1$, $\beta_1(\xi) = \mathsf{M}(\xi - \mathsf{M}\xi)^1 = 0$ для любой случайной величины и центральный момент второго порядка $\beta_2(\xi) = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \mathsf{D}\xi$. При $k \geqslant 3$ имеем новые числовые характеристики случайной величины. Эти числовые характеристики ввели для учета так называемых «хвостов» отклонения случайной величины от своего математического ожидания. По аналогии с начальным моментом центральный момент k-го порядка для дискретных, непрерывных и произвольных случайных величин вычисляются по формулам: $\beta_k(\xi) = \sum_i (x_i' - \mathsf{M}\xi)^k p_i$, $\beta_k(\xi) = \int_{-\infty}^{+\infty} (x - \mathsf{M}\xi)^k f(x)\,dx$ и $\beta_k(\xi) = \int_{-\infty}^{+\infty} (x - \mathsf{M}\xi)^k dF(x)$, если ряд и интегралы абсолютно сходятся. Иначе: центральные моменты, либо не существуют, либо не имеют конечного значения. Между центральными и начальными моментами существует простая связь. Начальные моменты выражаются через центральные моменты и наоборот:

$$\alpha_k(\xi) = \mathsf{M}(\xi^k) = \mathsf{M}[((\xi - \mathsf{M}\xi) + \mathsf{M}\xi)^k] = \mathsf{M}\Big[\sum_{r=0}^{k} C_k^r (\xi - \mathsf{M}\xi)^r (\mathsf{M}\xi)^{k-r}\Big] =$$

$$= \sum_{r=0}^{k} C_k^r (\mathsf{M}\xi)^{k-r} \mathsf{M}[(\xi - \mathsf{M}\xi)^r] = \sum_{r=0}^{k} C_k^r (\alpha_1\xi)^{k-r} \beta_r \xi,$$

$$\beta_k(\xi) = \mathsf{M}[(\xi - \mathsf{M}\xi)^k] = \mathsf{M}[\{\xi + (-\mathsf{M}\xi)\}^k] = \mathsf{M}\Big(\sum_{r=0}^{k} C_k^r \xi^r (-\mathsf{M}\xi)^{k-r}\Big) =$$

$$= \sum_{r=0}^{k} C_k^r (-1)^{k-r} (\mathsf{M}\xi)^{k-r} \mathsf{M}(\xi^r) = \sum_{r=0}^{k} C_k^r (-1)^{k-r} (\alpha_1\xi)^{k-r} \alpha_r \xi.$$

Итак, если центральные моменты принимают конечные значения, то начальные моменты конечны, и наоборот. Центральные моменты третьего и четвертого порядка позволяют ввести еще такие вспомогательные характеристики, как коэффициент асимметрии и эксцесс случайных величин. Когда распределение случайной величины симметрично относительно ее математического ожидания, то нетрудно видеть, что центральные моменты нечетного порядка (если они конечны) равны нулю. В качестве примера покажем это для непрерывной случайной величины с плотностью $f(x)$, учитывая равенство $f(\mathsf{M}\xi - x) = f(\mathsf{M}\xi + x)$ при всех $-\infty < x < +\infty$ и выполняя преобразования:

$$\beta_{2k+1}(\xi) = \int_{-\infty}^{+\infty} (x - \mathsf{M}\xi)^{2k+1} f(x)\,dx =$$

$$= \int_{-\infty}^{\mathsf{M}\xi} (x - \mathsf{M}\xi)^{2k+1} f(x)\,dx + \int_{\mathsf{M}\xi}^{+\infty} (x - \mathsf{M}\xi)^{2k+1} f(x)\,dx =$$

$$= \int_{-\infty}^{0} t^{2k+1} f(\mathsf{M}\xi + t)\,dt + \int_{0}^{+\infty} t^{2k+1} f(\mathsf{M}\xi + t)\,dt =$$

$$= -\int_0^{-\infty} t^{2k+1} f(\mathsf{M}\xi + t)\,dt + \int_0^{+\infty} t^{2k+1} f(\mathsf{M}\xi + t)\,dt =$$
$$= -\int_0^{+\infty} u^{2k+1} f(\mathsf{M}\xi - u)\,du + \int_0^{+\infty} t^{2k+1} f(\mathsf{M}\xi + t)\,dt =$$
$$= -\int_0^{+\infty} u^{2k+1} f(\mathsf{M}\xi + u)\,du + \int_0^{+\infty} t^{2k+1} f(\mathsf{M}\xi + t)\,dt = 0.$$

В силу этого на практике вводят характеристику, которая отвечает за симметрию распределения случайной величины ξ.

Определение 8.10. Величина $\mathsf{Ka}(\xi) = \beta_3\xi/(\sigma\xi)^3$ называется *коэффициентом асимметрии случайной величины* ξ.

Выясним смысл коэффициента асимметрии $\mathsf{Ka}\,\xi$ на содержательном уровне. Если $\mathsf{Ka}\,\xi = 0$, то распределение случайной величины, как правило, симметрично относительно ее математического ожидания. При $\mathsf{Ka}\,\xi < 0$ случайная величина чаще принимает далекие и расположенные слева от $\mathsf{M}\xi$ значения, нежели далекие и расположенные справа от $\mathsf{M}\xi$ значения. Однако ξ чаще принимает близкие справа от $\mathsf{M}\xi$ значения, нежели близкие слева от $\mathsf{M}\xi$ значения. Такая асимметрия для непрерывной случайной величины отображена на рис.8. 10, *а* и называется *отрицательной*. Если $\mathsf{Ka}\,\xi > 0$, то это означает, что случайная величина далекие и расположенные слева от $\mathsf{M}\xi$ значения принимает реже, чем далекие и расположенные справа от $\mathsf{M}\xi$ значения. С другой стороны, случайная величина ξ чаще принимает близкие слева от $\mathsf{M}\xi$ значения, нежели близкие справа от $\mathsf{M}\xi$ значения. Такого рода асимметрия для непрерывной случайной величины представлена на рис. 8.10, *б* и называется *положительной*.

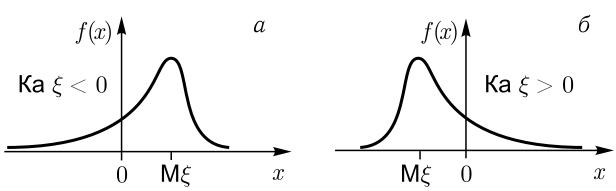


Рис. 8.10

Определение 8.11. *Эксцессом случайной величины* ξ называют числовую характеристику, равную $(\boldsymbol{\sigma}\xi)^{-4}\beta_4\xi - 3$ и обозначаемую через $\mathsf{Э}(\xi)$.

На содержательном уровне эксцесс характеризует крутизну наклона плотности вероятности или многоугольника распределения в окрестности наивероятнейшего значения случайной величины. Чем больше значение $Э\xi$, тем больше наклон соответствующего графика. Другими словами, эксцесс характеризует «островершинность» многоугольника распределения для дискретной случайной величины ξ и кривой распределения $f(x)$ для непрерывной случайной величины ξ.

Рассмотрим пример на вычисление вспомогательных характеристик.

Пример 8.14. Для непрерывной случайной величины ξ плотность вероятности $f(x) = 2^{-1}\exp\{-|x-a|\}$ для всех $-\infty < x < +\infty$, где a есть постоянная величина и $-\infty < a < +\infty$. На рис. 8.11 изображена функция $f(x)$.

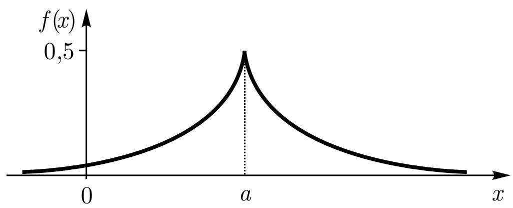


Рис. 8.11

Найти начальные и центральные моменты, коэффициент асимметрии и эксцесс случайной величины ξ.

Решение. Легко проверяется условие нормировки для плотности $f(x)$:

$$\int\limits_{-\infty}^{+\infty} 2^{-1}\exp\{-|x-a|\}\,dx = 2^{-1}\Big(\int\limits_{-\infty}^{a}\exp\{x-a\}\,dx + \int\limits_{a}^{+\infty}\exp\{a-x\}\,dx\Big) =$$

$$= \int\limits_{a}^{+\infty}\exp\{a-x\}\,dx = -\exp\{a-x)\Big|_{a}^{+\infty} = 1.$$

Будем говорить, что данная случайная величина ξ распределена по закону Лапласа. Математическое ожидание вычисляем следующим образом:

$$\mathsf{M}\xi = \int\limits_{-\infty}^{+\infty} x\,2^{-1}\exp\{-|x-a|\}\,dx =$$

$$= 2^{-1}\int\limits_{-\infty}^{a} x\exp\{x-a\}\,dx + 2^{-1}\int\limits_{a}^{+\infty} x\exp\{a-x\}\,dx =$$

$$= 2^{-1}\int\limits_{-\infty}^{0}(t+a)\,\mathrm{e}^t dt + 2^{-1}\int\limits_{-\infty}^{0}(a-t)\,\mathrm{e}^t dt = \int\limits_{-\infty}^{0} a\,\mathrm{e}^t dt = a.$$

Если $k = 2, 3, \ldots$, то для начальных моментов k-го порядка имеем:

$$\alpha_k \xi = \mathsf{M}(\xi^k) = \int\limits_{-\infty}^{+\infty} x^k 2^{-1} \exp\{-|x-a|\}\, dx =$$

$$= 2^{-1} \int\limits_{-\infty}^{a} x^k \exp\{x-a\}\, dx + 2^{-1} \int\limits_{a}^{+\infty} x^k \exp\{a-x\}\, dx =$$

$$= 2^{-1} \int\limits_{-\infty}^{0} (t+a)^k\, \mathrm{e}^t dt + 2^{-1} \int\limits_{-\infty}^{0} (a-t)^k\, \mathrm{e}^t dt =$$

$$= 2^{-1}\mathrm{e}^t[(t+a)^k - k(t+a)^{k-1} + k(k-1)\,(t+a)^{k-2} - \ldots + (-1)^k k!]\Big|_{-\infty}^{0} +$$

$$+ 2^{-1}\mathrm{e}^t[(a-t)^k + k(a-t)^{k-1} + k(k-1)\,(a-t)^{k-2} + \ldots + k!]\Big|_{-\infty}^{0} =$$

$$= 2^{-1}[a^k - ka^{k-1} + k(k-1)\,a^{k-2} - \ldots + (-1)^k k!] +$$

$$+ 2^{-1}[a^k + ka^{k-1} + k(k-1)\,a^{k-2} + \ldots + k!].$$

Отсюда получаем, что при $k = 2, 3, \ldots$ числовая характеристика $\alpha_k \xi$ равна $a^k + k(k-1)\,a^{k-2} + \ldots + k!\,a^0$ для четных начальных моментов и равна $a^k + k(k-1)\,a^{k-2} + \ldots + k!\,a$ для нечетных начальных моментов. Например, величина $\alpha_2 \xi = a^2 + 2$, а третий начальный момент равен $\alpha_3 \xi = a^3 + 6a$. Так как начальные моменты конечны и распределение случайной величины Лапласа симметрично относительно ее математического ожидания, то центральные моменты нечетного порядка равны нулю. Значит, коэффициент асимметрии $\mathsf{Ka}\,\xi = \beta_3\xi/(\sigma\xi)^3 = 0$. Для центральных моментов четного порядка получим:

$$\beta_{2k}\xi = 2^{-1} \int\limits_{-\infty}^{+\infty} (x - \mathsf{M}\xi)^{2k} \mathrm{e}^{-|x-a|} dx =$$

$$= 2^{-1} \int\limits_{-\infty}^{a} (x-a)^{2k} \mathrm{e}^{x-a} dx + 2^{-1} \int\limits_{a}^{+\infty} (x-a)^{2k} \mathrm{e}^{a-x} dx =$$

$$= 2^{-1} \int\limits_{-\infty}^{0} t^{2k} \mathrm{e}^t dt + 2^{-1} \int\limits_{-\infty}^{0} t^{2k} \mathrm{e}^t dt = \int\limits_{-\infty}^{0} t^{2k} \mathrm{e}^t dt =$$

$$= \mathrm{e}^t[t^{2k} - 2kt^{2k-1} + 2k(2k-1)\,t^{2k-2} - \ldots + (2k)!]\Big|_{-\infty}^{0} = (2k)!$$

Отсюда, дисперсия случайной величины Лапласа $\mathsf{D}\xi = \beta_2\xi = 2$, стандарт $\boldsymbol{\sigma}\xi = 2^{1/2}$ и, наконец, эксцесс $\text{Э}\xi = (\boldsymbol{\sigma}\xi)^{-4}\beta_4\xi - 3 = 24/4 - 3 = 3$. Легко вычислить, что для этой задачи $\mathsf{Mo}\,\xi = \mathsf{Mo}^*\xi = \mathsf{Me}\,\xi = \mathsf{M}\xi = a$. Теперь найдем среднее отклонение случайной величины Лапласа:

$$\mathsf{E}\xi = \mathsf{M}(|\xi - \mathsf{Me}\,\xi|) = 2^{-1}\int\limits_{-\infty}^{+\infty} |x-a|\,\mathrm{e}^{-|x-a|}dx =$$
$$= -2^{-1}\int\limits_{-\infty}^{a} (x-a)\,\mathrm{e}^{x-a}dx + 2^{-1}\int\limits_{a}^{+\infty} (x-a)^{2k}\mathrm{e}^{a-x}dx =$$
$$= -2^{-1}\int\limits_{-\infty}^{0} t\,\mathrm{e}^{t}dt - 2^{-1}\int\limits_{-\infty}^{0} t\,\mathrm{e}^{t}dt = -\int\limits_{-\infty}^{0} t\,\mathrm{e}^{t}dt = 1.$$

При решении этой задачи получили следующий замечательный факт. Случайную величину Лапласа можно приближено заменять математическим ожиданием или медианой. Однако ошибку такого приближения можно измерять величиной $(\mathsf{M}(\xi - \mathsf{M}\xi)^2)^{1/2} = \sigma\xi = 2^{1/2}$ или несколько другой величиной $\mathsf{M}(|\xi - \mathsf{Me}\,\xi|) = \mathsf{E}\xi = 1$. Так как $2^{1/2} > 1$, то имеем вполне естественный результат о том, что величина наименьшей ошибки существенно зависит от выбранного критерия точности приближения. При этом разные критерии дают для случайной величины Лапласа один и тот же оптимальный прогноз, который равен a.

§ 3. Характеристики положения и разброса семейства измерителей исходов эксперимента

3.1. Математическое ожидание многомерной случайной величины и ее смешанный центральный момент второго порядка. Напомним, что многомерная случайная величина $(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$ поточечно определяется на пространстве Ω описаний элементарных событий и является упорядоченной системой из n одномерных случайных величин $\xi_1(\omega)$, $\xi_2(\omega), \ldots$, $\xi_n(\omega)$, заданных на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. Отображение $(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))\colon \Omega \to R^n$ и совместная функция распределения вероятностей $F(x_1, x_2, \ldots, x_n) = \mathsf{P}(\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_n < x_n)$ n-мерной случайной величины являются полными и исчерпывающими характеристиками этой величины с функциональной и вероятностной точек зрения. Эти характеристики, как правило, при больших значениях размерности n и статистической зависимости случайных величин $\xi_1(\omega)$, $\xi_2(\omega), \ldots$, $\xi_n(\omega)$ являются сложными и громоздкими математическими объектами. Поэтому для приближенного вычисления этих функциональных характеристик необходимо многократно проводить или наблюдать эксперимент E. Это в свою очередь требует обработки большого объема экспериментальных данных. Следовательно, мы вынуждены решать проблему изучения свойств совокупности измерителей элементарных исходов эксперимента более простыми и доступными средствами. Более того, в некоторых инженерных приложениях можно обойтись без исчерпывающих характеристик, описав некоторые специальные свойства совокупности измерителей приближенно или в среднем. С этой целью вводят такие числовые характеристики многомерной случайной величины, как, например, математическое ожидание, дисперсию и корреляционные моменты.

Определение 8.12. *Математическим ожиданием* n-мерной случайной величины $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ называется точка n-мерного пространства R^n вида $\mathsf{M}\xi = (\mathsf{M}\xi_1, \mathsf{M}\xi_2, \ldots, \mathsf{M}\xi_n)$, где величина $\mathsf{M}\xi_k$ равна математическому ожиданию одномерной случайной величины ξ_k для каждого $k = 1, 2, \ldots, n$.

Таким образом, вычисление математического ожидания $\mathsf{M}\xi$ случайного вектора ξ сводится к вычислению математических ожиданий одномерных случайных величин ξ_1, $\xi_2, \ldots$, ξ_n. К сожалению, для вычисления дисперсии случайного вектора или n-мерной случайной величины ξ не достаточно знать дисперсии одномерных случайных величин. Для пояснения этого факта, ради простоты, рассмотрим две первые случайные величины ξ_1 и ξ_2. В этой главе было показано, что $\mathsf{D}(\xi_1 + \xi_2) = \mathsf{D}(\xi_1) + \mathsf{D}(\xi_2)$, если ξ_1 и ξ_2 независимые случайные величины. Если же ξ_1 и ξ_2 зависимые случайные величины, то указанное соотношения в общем случае не имеет место, так как $\mathsf{D}(\xi_1 + \xi_2) = \mathsf{D}(\xi_1) + \mathsf{D}(\xi_2) + 2\,\mathsf{M}((\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2))$ и величина $\mathsf{M}((\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2))$ может оказаться отличной от нуля. Значит, для вычисления $\mathsf{D}(\xi_1 + \xi_2)$ в общем случае не достаточно знать дисперсии $\mathsf{D}(\xi_1)$ и $\mathsf{D}(\xi_2)$, а необходимо еще знать $\mathsf{M}((\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2))$. Математический объект (ξ_1, ξ_2) существенно сложнее, чем $(\xi_1 + \xi_2)$, так как вектор (ξ_1, ξ_2) однозначно определяет сумму $(\xi_1 + \xi_2)$, а сумма вида $(\xi_1 + \xi_2)$ не восстанавливает вектор (ξ_1, ξ_2). Более того, даже более информативные математические объекты ξ_1 и ξ_2 по сравнению с объектом $(\xi_1 + \xi_2)$ определяют вектор (ξ_1, ξ_2) только при статистической независимости его компонент. Поэтому возникает естественный вопрос о характеристике зависимости величин ξ_1, ξ_2 и вопрос о свойствах величины $\mathsf{M}((\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2))$. В дальнейшем покажем, что при $|\mathsf{M}\xi_k| < \infty$ и $|\mathsf{M}\xi_s| < \infty$ величина вида $\mathsf{M}((\xi_k - \mathsf{M}\xi_k)(\xi_s - \mathsf{M}\xi_s))$ для любых двух случайных величин ξ_k и ξ_s из семейства $\{\xi_1, \xi_2, \ldots, \xi_n\}$ существенно влияет на величину рассеивания или на разброс многомерной случайной величины $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ около ее математического ожидания.

Определение 8.13. *Пирсоновским линейным корреляционным моментом*, или *ковариацией*, или *смешанным центральным моментом второго порядка случайных величин* ξ_k и ξ_s при $|\mathsf{M}\xi_k| < \infty$ и $|\mathsf{M}\xi_s| < \infty$ называется математическое ожидание произведения отклонений этих случайных величин от их математических ожиданий и обозначается через $\mathsf{cov}(\xi_k, \xi_s)$, т. е. $\mathsf{cov}(\xi_k, \xi_s) = \mathsf{M}[(\xi_k - \mathsf{M}\xi_k)(\xi_s - \mathsf{M}\xi_s)]$.

В дальнейшем будем говорить «корреляционным моментом».

Введем величину $\eta_{k,s} = (\xi_k - \mathsf{M}\xi_k)(\xi_s - \mathsf{M}\xi_s) = g(\xi_k, \xi_s)$. Если (ξ_k, ξ_s) — дискретный случайный вектор и следующие вероятности: $\mathsf{P}(\{\omega\colon \xi_k(\omega) = u_{ki}, \xi_s(\omega) = u_{sj}\}) = r_{i,j}$, $i, j = 1, 2, \ldots$, задают его распределение, тогда корреляционный момент равен

$$\mathsf{M}\eta_{k,s} = \mathsf{cov}(\xi_k, \xi_s) = \sum_{i,j} (u_{ki} - \mathsf{M}\xi_k)(u_{sj} - \mathsf{M}\xi_s)\, r_{i,j}. \tag{8.21}$$

Если (ξ_k, ξ_s) — непрерывный вектор и $f(x_k, x_s)$ — его плотность, то

$$\mathsf{M}\eta_{k,s} = \mathsf{cov}(\xi_k, \xi_s) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x_k - \mathsf{M}\xi_k)(x_s - \mathsf{M}\xi_s) f(x_k, x_s)\, dx_k\, dx_s. \tag{8.22}$$

Наконец, если (ξ_k, ξ_s) является произвольной двумерной случайной величиной и $F(x_k, x_s) = \mathsf{P}(\{\omega\colon \xi_k(\omega) < x_k, \xi_s(\omega) < x_s\})$, $x_k \in R$, $x_s \in R$, тогда

$$\mathsf{M}\eta_{k,s} = \mathsf{cov}(\xi_k, \xi_s) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x_k - \mathsf{M}\xi_k)(x_s - \mathsf{M}\xi_s) dF(x_k, x_s). \tag{8.23}$$

Если $\xi_k = \xi_s$, то $\mathsf{cov}(\xi_k, \xi_k) = \mathsf{M}((\xi_k - \mathsf{M}\xi_k)(\xi_k - \mathsf{M}\xi_k)) = \mathsf{M}[(\xi_k - \mathsf{M}\xi_k)^2] = \mathsf{D}\xi_k$. Из формул (8.21), (8.22) и (8.23) непосредственно следует, что корреляционный момент $\mathsf{cov}(\xi_k, \xi_s)$ двух случайных величин характеризует так называемый *смешанный разброс* этих *случайных величин*, или то, как разброс одной случайной величины влияет на разброс другой случайной величины. Если случайная величина ξ_k или ξ_s постоянна, то $\mathsf{cov}(\xi_k, \xi_s) = 0$. Из определения $\mathsf{cov}(\xi_k, \xi_s)$ получаем, что $\mathsf{cov}(\xi_k, \xi_s) = \mathsf{cov}(\xi_s, \xi_k)$. Размерность корреляционного момента равна произведению размерностей случайных величин. Этот факт является некоторым неудобством, так как при вычислениях корреляционного момента необходимо проверять его размерность. Поэтому вместо ковариации, если она принимает конечное значение, часто рассматривают следующую безразмерную числовую характеристику взаимного рассеивания двух случайных величин.

Определение 8.14. *Пирсоновским линейным коэффициентом корреляции* $\mathsf{corr}(\xi_k, \xi_s)$ *случайных величин* ξ_k и ξ_s при $0 < \boldsymbol{\sigma}\xi_k < \infty$, $0 < \boldsymbol{\sigma}\xi_s < \infty$ (и, значит, $|\mathsf{cov}(\xi_k, \xi_s)| < \infty$) называется отношение корреляционного момента к произведению их средних квадратических отклонений, т. е. $\mathsf{corr}(\xi_k, \xi_s) = \mathsf{cov}(\xi_k, \xi_s)/(\boldsymbol{\sigma}\xi_k \cdot \boldsymbol{\sigma}\xi_s)$.

В дальнейшем будем говорить «коэффициентом корреляции».

Если рассмотреть центрированные и нормированные случайные величины вида $\overline{\xi}_k = (\xi_k - \mathsf{M}\xi_k)/\boldsymbol{\sigma}\xi_k$, $\overline{\xi}_s = (\xi_s - \mathsf{M}\xi_s)/\boldsymbol{\sigma}\xi_s$, то $\mathsf{corr}(\xi_k, \xi_s) = \mathsf{M}(\overline{\xi}_k \times \overline{\xi}_s)$. Теперь нетрудно видеть, что $\mathsf{corr}(\overline{\xi}_k, \overline{\xi}_s) = \mathsf{M}(\overline{\xi}_k \times \overline{\xi}_s) = \mathsf{corr}(\xi_k, \xi_s)$. Это соотношение говорит о том, что коэффициент корреляции центрированных и нормированных случайных величин $\overline{\xi}_k$ и $\overline{\xi}_s$ совпадает с коэффициентом корреляции случайных величин ξ_k и ξ_s.

Лемма 8.14. *Если для произвольных случайных величин ξ_k и ξ_s существует величина $\mathsf{cov}(\xi_k, \xi_s)$, то $\mathsf{M}(\xi_k \times \xi_s) = \mathsf{M}\xi_k \times \mathsf{M}\xi_s + \mathsf{cov}(\xi_k, \xi_s)$.*

Доказательство. Так как

$$\mathsf{M}(\xi_k\xi_s) = \mathsf{M}[(\xi_k - \mathsf{M}\xi_k + \mathsf{M}\xi_k) \times (\xi_s - \mathsf{M}\xi_s + \mathsf{M}\xi_s)] = \\ = \mathsf{M}[(\xi_k - \mathsf{M}\xi_k)(\xi_s - \mathsf{M}\xi_s) + (\xi_k - \mathsf{M}\xi_k)\,\mathsf{M}\xi_s + (\xi_s - \mathsf{M}\xi_s)\,\mathsf{M}\xi_k + \mathsf{M}\xi_k\,\mathsf{M}\xi_s]$$

и величина $\mathsf{cov}(\xi_k, \xi_s)$ существует, то отсюда по теореме 8.2 следует необходимое равенство $\mathsf{M}(\xi_k\xi_s) = \mathsf{cov}(\xi_k, \xi_s) + \mathsf{M}\xi_k\,\mathsf{M}\xi_s$.

Лемма 8.15. *Если для произвольных случайных величин ξ_k и ξ_s существует величина $\mathsf{cov}(\xi_k, \xi_s) \neq -\infty$, то $\mathsf{D}(\xi_k + \xi_s) = \mathsf{D}\xi_k + \mathsf{D}\xi_s + 2\,\mathsf{cov}(\xi_k, \xi_s)$.*

Доказательство. Из определения дисперсии с учетом теоремы 8.2 находим:

$$\mathsf{D}(\xi_k + \xi_s) = \mathsf{M}[\xi_k + \xi_s - \mathsf{M}(\xi_k + \xi_s)]^2 = \mathsf{M}[(\xi_k - \mathsf{M}\xi_k) + (\xi_s - \mathsf{M}\xi_s)]^2 = \\ = \mathsf{M}(\xi_k - \mathsf{M}\xi_k)^2 + \mathsf{M}((\xi_s - \mathsf{M}\xi_s))^2 + 2\,\mathsf{M}[(\xi_k - \mathsf{M}\xi_k)(\xi_s - \mathsf{M}\xi_s)] = \\ = \mathsf{D}\xi_k + \mathsf{D}\xi_s + 2\,\mathsf{cov}(\xi_k, \xi_s),$$

что и требовалось доказать.

3.2. Дисперсия многомерной случайной величины. Перейдем теперь непосредственно к определению дисперсии n-мерной случайной величины. Пусть дана n-мерная случайная величина $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$. Ее рассеивание нельзя охарактеризовать вектором $(\mathsf{D}\xi_1, \mathsf{D}\xi_2, \ldots, \mathsf{D}\xi_n)$ из дисперсий случайных величин ξ_1, $\xi_2, \ldots$, ξ_n, так как по смыслу разброс должен измеряться неотрицательным числом. Отсюда можно утверждать, что в отличие от математического ожидания определение дисперсии многомерной случайной величины не является такой уж простой задачей.

Рассмотрим сначала тривиальный случай, когда $n = 2$ и $\xi_2 = \xi_1$. Тогда $\xi = (\xi_1, \xi_1)$ и все точки, координаты которых совпадают со значениями двумерной случайной величины $\xi = (\xi_1, \xi_1)$, лежат на прямой $\{(x, y)\colon y = x\}$ в R^2. Выберем прямую $\{(x, y)\colon y = x\}$ за новую ось абсцисс $0x_0$. Так как все двумерные случайные точки $\xi = (\xi_1, \xi_1)$ лежат на оси $0x_0$, то, очевидно, их можно считать одномерными случайными точками на оси $0x_0$. Легко вычислить, что указанные одномерные точки по отношению к оси $0x_0$ имеют координаты $\eta_0 = \sqrt{2}\,\xi_1$. Так как двумерная случайная величина $\xi = (\xi_1, \xi_1)$ и одномерная случайная величина $\eta_0 = \sqrt{2}\,\xi_1$ геометрически отображаются одними и теми же объектами на плоскости $x0y$, то разброс двумерной случайной величины $\xi = (\xi_1, \xi_1)$ в направлении оси $0x_0$ не может отличаться от дисперсии одномерной случайной величины $\eta_0 = \sqrt{2}\,\xi_1$. Отсюда получаем, что дисперсия двумерной случайной величины $\xi = (\xi_1, \xi_1)$ в направлении оси $0x_0$ равна $\mathsf{D}\eta_0 = \mathsf{D}(\sqrt{2}\,\xi_1) = 2\,\mathsf{D}\xi_1$.

Далее, двумерную случайную величину $\xi = (\xi_1, \xi_1)$ можно представить в виде другой одномерной случайной величины η. С этой целью выберем на плоскости $x0y$ произвольный единичный вектор вида $\mathbf{b} = (b_1, b_2)$, для которого $(b_1)^2 + (b_2)^2 = 1$. Ортогональная проекция случайных точек $\xi = (\xi_1(\omega), \xi_2(\omega))$, $\omega \in \Omega$, на направление, определяемое $\mathbf{b} = (b_1, b_2)$, и есть искомая одномерная случайная величина η. Ясно, что между двумерной случайной величиной $\xi = (\xi_1, \xi_1)$ и одномерной случайной величиной η установлено взаимно однозначное соответствие, если только направление, определяемое вектором $\mathbf{b} = (b_1, b_2)$, и прямая $\{(x, y)\colon y = x\}$ не являются ортогональными. Известно, что ортогональная проекция η равна скалярному произведению векторов $\mathbf{b} = (b_1, b_2)$ и $\xi = (\xi_1, \xi_1)$, т. е. в этом случае $\eta = b_1\xi_1 + b_2\xi_1 = (b_1 + b_2)\,\xi_1$. Таким способом двумерная случайная величина $\xi = (\xi_1, \xi_1)$ представляется одномерной случайной величиной $\eta = (b_1 + b_2)\,\xi_1$. Поэтому естественно считать, что дисперсии математических объектов $\xi = (\xi_1, \xi_1)$ и η должны совпадать. В силу этого дисперсия случайного вектора $\xi = (\xi_1, \xi_1)$ равна $\mathsf{D}\eta = (b_1 + b_2)^2\mathsf{D}\xi_1 = (1 + 2\,b_1b_2)\,\mathsf{D}\xi_1$. Так как эта дисперсия зависит от выбранного направления, то естественно обозначить ее символом $\mathsf{D}_\mathbf{b}\xi$. Итак, имеем $\mathsf{D}_\mathbf{b}\xi = (1 + 2\,b_1b_2)\,\mathsf{D}\xi_1$. Пусть теперь направление вектора $\mathbf{b} = (b_1, b_2)$ совпадает с прямой вида $\{(x, y)\colon y = x\}$ или с осью $0x_0$, т. е. $b_1 = b_2 = 1/\sqrt{2}$. В этом частном случае $\mathsf{D}_\mathbf{b}\xi = (1 + 2(1/\sqrt{2}\,)^2)\,\mathsf{D}\xi_1 = 2\,\mathsf{D}\xi_1$. Следовательно, при $\mathbf{b} = (1/\sqrt{2}\,, 1/\sqrt{2}\,)$ получили, что $\mathsf{D}_\mathbf{b}\xi = \mathsf{D}\eta_0$. Этот замечательный результат обосновывает целесообразность применения предложенного подхода к решению этой задачи в общем случае или к определению дисперсии многомерной случайной величины $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$.

Поступая аналогичным образом, можно определить дисперсию многомерной случайной величины в направлении произвольного единичного вектора $\mathbf{b} = (b_1, b_2, \ldots, b_n)$, при этом $(b_1)^2 + (b_2)^2 + \ldots + (b_n)^2 = 1$. Координаты единичного вектора $\mathbf{b}$ обычно называют *направляющими косинусами*. Найдем теперь проекцию $\varphi(\xi, \mathbf{b})$ многомерной случайной точки ξ на направление, определяемое вектором $\mathbf{b}$. Эта проекция равна скалярному произведению векторов ξ и $\mathbf{b}$, т. е.

$\varphi(\xi, \mathbf{b}) = (\xi, \mathbf{b}) = \sum_{k=1}^{n} b_k\xi_k$, и поэтому является одномерной случайной величиной. Отметим, что значения одномерной случайной величины $\varphi(\xi, \mathbf{b})$ зависят от вектора $\mathbf{b}$. Так как математическое ожидание $\mathsf{M}[\varphi(\xi, \mathbf{b})] = \mathsf{M}(\sum_{k=1}^{n} b_k\xi_k) = \sum_{k=1}^{n} b_k\mathsf{M}\xi_k$, то для дисперсии случайной величины $\varphi(\xi, \mathbf{b})$ последовательно получаем:

$$\mathsf{D}\varphi(\xi, \mathbf{b}) = \mathsf{M}\Big(\sum_{k=1}^{n} b_k\xi_k - \sum_{k=1}^{n} b_k\mathsf{M}\xi_k\Big)^2 = \mathsf{M}\Big(\sum_{k=1}^{n} b_k(\xi_k - \mathsf{M}\xi_k)\Big)^2 =$$
$$= \mathsf{M}\Big(\sum_{k=1}^{n} b_k(\xi_k - \mathsf{M}\xi_k) \sum_{s=1}^{n} b_s(\xi_s - \mathsf{M}\xi_s)\Big) = \sum_{k=1}^{n}\sum_{s=1}^{n} b_k b_s \mathsf{cov}(\xi_k, \xi_s).$$

Итак, при выбранном направлении единичного вектора $\mathbf{b}$ дисперсия $\mathsf{D}_{\mathbf{b}}\xi$ многомерной случайной величины $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ равна, по определению, дисперсии $\mathsf{D}\varphi(\xi, \mathbf{b})$ одномерной случайной величины $\varphi(\xi, \mathbf{b})$. В частности, если направление единичного вектора $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ совпадает с некоторой координатной осью, например, с осью $0x_i$, то $b_i = 1$, $b_k = 0$ для всех $k \neq i$. Тогда дисперсия вектора $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ в направлении координатной оси $0x_i$ равна $(b_i)^2\mathsf{cov}(\xi_i, \xi_i) = \mathsf{D}\xi_i$. Ортогональная проекция точек $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega)), \ldots, \xi_n(\omega))$, $\omega \in \Omega$, в пространстве R^n на координатную ось $0x_i$ совпадает с одномерной случайной величиной $\xi_i(\omega)$. Поэтому величину $\mathsf{D}\xi_i$ целесообразно интерпретировать как рассеивание n-мерной случайной величины $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ в направлении оси $0x_i$. Таков смысл $\mathsf{D}\xi_i$ по отношению к вектору $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$. Далее дисперсия $\mathsf{D}_{\mathbf{b}}\xi$ вычисляется через элементы коварицонной матрицы

$$\mathsf{K}(\xi) = \begin{pmatrix} \mathsf{cov}(\xi_1, \xi_1) & \mathsf{cov}(\xi_1, \xi_2) & \ldots & \mathsf{cov}(\xi_1, \xi_n) \\ \mathsf{cov}(\xi_2, \xi_1) & \mathsf{cov}(\xi_2, \xi_2) & \ldots & \mathsf{cov}(\xi_2, \xi_n) \\ \ldots & \ldots & \ldots & \ldots \\ \mathsf{cov}(\xi_n, \xi_1) & \mathsf{cov}(\xi_n, \xi_2) & \ldots & \mathsf{cov}(\xi_n, \xi_n) \end{pmatrix}$$

Определение 8.15. Две случайные величины называются *некоррелированными*, если их корреляционный момент равен нулю. Иначе эти случайные величины называются *коррелированными*.

Будем говорить, что последовательность ξ_1, $\xi_2, \ldots$ составлена из попарно некоррелированных случайных величин, если $\mathsf{cov}(\xi_k, \xi_s) = 0$ для всех $k, s \in \{1, 2, \ldots\}$ и $k \neq s$. Пусть теперь каждая компонента заданного направления $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ равна единице. Тогда, учитывая равенства $\varphi(\xi, \mathbf{c}) = \sum_{k=1}^{n} \xi_k$, $\mathsf{cov}(\xi_k, \xi_k) = \mathsf{D}\xi_k$, $\mathsf{cov}(\xi_k, \xi_s) = \mathsf{cov}(\xi_s, \xi_k)$, где $k, s \in \{1, 2, \ldots, n\}$, получим следующее обобщение леммы 8.15:

$$\mathsf{D}\,\varphi(\xi, \mathbf{c}) = \mathsf{D}\Big(\sum_{k=1}^{n} \xi_k\Big) = \sum_{k=1}^{n}\sum_{s=1}^{n} \mathsf{cov}(\xi_k, \xi_s) = \sum_{k=1}^{n} \mathsf{D}\xi_k + 2\sum_{k<s}^{n} \mathsf{cov}(\xi_k, \xi_s).$$

Отсюда
$$\mathsf{D}\Big(\sum_{k=1}^{n}\xi_k\Big)=\sum_{k=1}^{n}\mathsf{D}\xi_k+2\sum_{k<s}^{n}\mathsf{cov}(\xi_k,\xi_s).$$
Если случайные величины ξ_1, $\xi_2,\ldots$, ξ_n являются попарно некоррелированными, то $\mathsf{D}\Big(\sum_{k=1}^{n}\xi_k\Big)=\sum_{k=1}^{n}\mathsf{D}\xi_k$. Матрица $\mathsf{K}(\xi)$, которая в силу равенств $\mathsf{cov}(\xi_k,\xi_s)=\mathsf{cov}(\xi_s,\xi_k)$, k, $s\geqslant 1$ является симметричной, называется ковариационной или корреляционной. Далее, матрица $\mathsf{K}(\xi)$ является симметричной и для любых вещественных чисел ν_1, $\nu_2,\ldots$, ν_n выполняются соотношения вида:
$$\mathsf{M}\Big[\sum_{k=1}^{n}\nu_k(\xi_k-\mathsf{M}\xi_k)\Big]^2\geqslant 0,$$
$$\mathsf{M}\Big[\sum_{k=1}^{n}\nu_k(\xi_k-\mathsf{M}\xi_k)\Big]^2=\mathsf{M}\Big[\sum_{k=1}^{n}\nu_k(\xi_k-\mathsf{M}\xi_k)\sum_{s=1}^{n}\nu_s(\xi_s-\mathsf{M}\xi_s)\Big]=$$
$$=\sum_{k=1}^{n}\sum_{s=1}^{n}\nu_k\nu_s\mathsf{M}[(\xi_k-\mathsf{M}\xi_k)(\xi_s-\mathsf{M}\xi_s)]=\sum_{k=1}^{n}\sum_{s=1}^{n}\nu_k\nu_s\mathsf{cov}(\xi_s,\xi_k)\geqslant 0.$$
Тогда из теории квадратичных форм, получаем, что определитель
$$\begin{vmatrix}\mathsf{cov}(\xi_1,\xi_1) & \mathsf{cov}(\xi_1,\xi_2) & \ldots & \mathsf{cov}(\xi_1,\xi_s)\\ \mathsf{cov}(\xi_2,\xi_1) & \mathsf{cov}(\xi_2,\xi_2) & \ldots & \mathsf{cov}(\xi_2,\xi_s)\\ \ldots & \ldots & \ldots & \ldots\\ \mathsf{cov}(\xi_s,\xi_1) & \mathsf{cov}(\xi_s,\xi_2) & \ldots & \mathsf{cov}(\xi_s,\xi_s)\end{vmatrix}$$
при каждом $s=1,2,\ldots,n$. В случае, когда компоненты n-мерной случайной величины попарно некоррелированы, ковариационная матрица примет так называемый диагональный вид:
$$\mathsf{K}(\xi)=\begin{pmatrix}\mathsf{cov}(\xi_1,\xi_1) & 0 & \ldots & 0\\ 0 & \mathsf{cov}(\xi_2,\xi_2) & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & \mathsf{cov}(\xi_n,\xi_n)\end{pmatrix}.$$
В этом случае имеем: $\mathsf{D}_{\mathbf{b}}\xi=(b_1)^2\mathsf{cov}(\xi_1,\xi_1)+(b_2)^2\mathsf{cov}(\xi_2,\xi_2)+\ldots+$ $+(b_n)^2\mathsf{cov}(\xi_n,\xi_n)=(b_1)^2\mathsf{D}\xi_1+(b_2)^2\mathsf{D}\xi_2+\ldots+(b_n)^2\mathsf{D}\xi_n$.

Краткий обзор

В главе 8 изложена общая концепция о числовых параметрах, которые некоторым существенным образом характеризуют случайные количественные признаки статистически устойчивого эксперимента. Отмечается, что в информационном отношении для измерителей исходов эксперимента E числовые характеристики беднее, чем случайная величина и ее интегральная функция распределения, но зато значительно проще. Определяются основные числовые характеристики положения (математическое ожидание, медиана, квантили, мода) и основные числовые характеристики разброса (дисперсия, среднее квадратическое отклонение или стандарт, среднее отклонение) одномерных случайных величин. При этом основное внимание уделяется определению, свойствам и получению доступных формул вычисления математического ожидания. Математическое ожидание $\mathsf{M}\xi$ определяется сначала для так называемых простых случайных величин, затем для неотрицательных случайных величин и, наконец, для произвольных случайных величин. Такой подход позволяет подробно изучить общие свойства математического ожидания. Доказываются важные неравенства Чебышева, которые оценивают значения вероятностей событий, порожденных произвольными случайными величинами. Вводятся понятия вспомогательных числовых характеристик (начальных и центральных моментов, коэффициента асимметрии, эксцесса) случайных величин. Приводятся формулы вычисления как основных, так и вспомогательных числовых характеристик. Определяются и изучаются интегральные характеристики (среднее значение, корреляционный момент или ковариация, коэффициент корреляции, дисперсия) семейства измерителей исходов статистически устойчивого эксперимента. Приведено обоснование целесообразности применения предложенного подхода к определению дисперсии многомерной случайной величины. Устанавливается связь между дисперсией случайного вектора и дисперсиями одномерных случайных величин, которые его определяют. Получены общие формулы для математического ожидания от произведения двух случайных величин и для дисперсии от суммы конечного числа случайных величин. Дано важное понятие попарно некоррелированных случайных величин. Доказывается формула для дисперсии суммы попарно некоррелированных случайных величин. Теоретические результаты, как правило, иллюстрируются примерами.

Контрольные вопросы и упражнения

1. Привести случаи, когда математическое ожидание и начальные моменты случайных величин не существуют.
2. Дать интерпретацию числовых характеристик случайных величин.
3. В чем заключается основное отличие числовых характеристик одномерных и многомерных случайных величин?

4. Привести вид ковариационной матрицы для независимых случайных величин.
5. Два игрока поочередно подбрасывают несимметричную монету до первого выпадения герба (орла). Каждый бросок осуществляется независимым образом с помощью некоторого механизма, обеспечивающего произвольное вращение монеты в воздухе и ее непреднамеренное падение на некоторую поверхность стола. Вероятность появления герба при одном броске равна p. Вычислить среднее число бросков, которое совершает в этом опыте каждый из игроков.
6. На зачете студент получил шесть задач. Вероятность того, что студент решит задачу правильно, одинакова для всех задач и равна 0,8. Определить среднее количество и дисперсию решенных задач. Вычислить наиболее вероятное количество решенных задач.
7. При передаче сообщения по радиоканалу, вследствие наличия помех, с вероятностью p сообщение не удается декодировать. Сообщение передается до тех пор, пока оно не будет декодировано. Продолжительность передачи сообщения равна 2 мин. Найти математическое ожидание и дисперсию времени, которое уйдет на передачу сообщения.
8. Маршрутные такси подходят к остановке с интервалом в 5 мин. Предполагается, что случайное время ожидания маршрутного такси распределено равномерно. Определить среднее значение и среднее квадратическое отклонение времени ожидания такси.
9. Пусть плотность распределения вероятностей $f(x)$ непрерывной случайной величины ξ равна $0{,}5\,x$ при $0 \leqslant x \leqslant h$ и равна нулю при $x \notin [0, h]$. Вычислить математическое ожидание, моду, дисперсию, коэффициент асимметрии, эксцесс и квантиль порядка $p = 0{,}5$.
10. Пусть ξ_1 и ξ_2 являются некоррелированными случайными величинами. Вычислить математическое ожидание случайного вектора (η_1, η_2), если $\eta_1 = 3\,\xi_1^2 + 5\,\xi_2^2 + \xi_1\xi_2 + \xi_1 - 6\,\xi_2$, $\eta_2 = \xi_1^2 + 4\,\xi_2^2 + 7\,\xi_1\xi_2 + \xi_1$, $\mathsf{M}\xi_1 = 3$, $\mathsf{M}\xi_2 = 2$, $\mathsf{D}\xi_1 = 9$, $\mathsf{D}\xi_2 = 16$.
11. Пусть ξ_1, ξ_2, ξ_3, ξ_4 являются попарно некоррелированными случайными величинами. Найти коэффициент корреляции $\mathsf{corr}(\eta_1, \eta_2)$ двух случайных величин η_1 и η_2, где случайная величина $\eta_1 = \xi_1 + \xi_2$ и случайная величина $\eta_2 = \xi_3 + \xi_4$.
12. Наудачу подбрасываются один раз две монеты разной симметрии. Герб появляется с вероятностью p_1 на первой монете и с вероятностью p_2 на второй. Пусть ξ_1 определяет число выпавших гербов на первой монете и ξ_2 на второй. Суммарное число выпавших на

обеих монетах гербов обозначим через η_1, а решеток через η_2. Найти коэффициент корреляции случайных величин η_1 и η_2.

13. Пусть ξ и η являются некоррелированными случайными величинами. Плотность распределения $f_\xi(x)$ случайной величины ξ равна $2x$ при $0 \leqslant x \leqslant h_1$ и равна нулю при $x \notin [0, h_1]$. Плотность распределения $f_\eta(y)$ случайной величины η равна $2 - 2y$ при $0 \leqslant y \leqslant h_2$ и равна нулю при $y \notin [0, h_2]$. Вычислить дисперсию двумерной случайной величины (ξ, η) в направлении вектора $(1, 1)$.

Глава 9

ФУНКЦИОНАЛЬНАЯ И СТАТИСТИЧЕСКАЯ ЗАВИСИМОСТИ МЕЖДУ ИЗМЕРИТЕЛЯМИ СЛУЧАЙНЫХ ЭКСПЕРИМЕНТОВ

§ 1. Неслучайные функции от случайных аргументов

1.1. Понятие о функциональной зависимости между случайными величинами. В приложениях часто приходится встречаться с различными случайными величинами, которые связаны функциональной зависимостью с другими случайными величинами, т.е. являются функциями случайных аргументов. Например, на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ могут быть заданы два семейства случайных величин: $\{\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)\}$ и $\{\eta_1(\omega), \eta_2(\omega), \ldots, \eta_m(\omega)\}$, для которых справедливы следующие функциональные соотношения:

$$\begin{gathered} \eta_1 = g_1(\xi_1, \xi_2, \ldots, \xi_n), \\ \eta_2 = g_2(\xi_1, \xi_2, \ldots, \xi_n), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \eta_m = g_m(\xi_1, \xi_2, \ldots, \xi_n). \end{gathered} \tag{9.1}$$

Здесь $g_i(x_1, x_2, \ldots, x_n)$ — функция на n-мерном пространстве R^n со значениями в R при каждом $i = 1, 2, \ldots, m$. При $i = 1, 2, \ldots, m$ случайную величину $\eta_i = g_i(\xi_1, \xi_2, \ldots, \xi_n)$ называют функцией от n-мерной случайной величины $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ или функцией от n случайных аргументов.

Пусть теперь на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ заданы случайные величины $\xi_1(\omega)$, $\xi_2(\omega)$, …, $\xi_n(\omega)$ и на n-мерном пространстве R^n заданы борелевские функции $g_1(x_1, x_2, \ldots, x_n)$, $g_2(x_1, x_2, \ldots, x_n)$, …, $g_m(x_1, x_2, \ldots, x_n)$. В этом случае докажем утверждение, обобщающее лемму 6.3, а именно: *равенства* (9.1) *определяют случайные величины* $\eta_1(\omega)$, $\eta_2(\omega)$, …, $\eta_m(\omega)$. Действительно, для каждого $\omega \in \Omega$ имеем $\xi_1(\omega) = x_1 \in R$, $\xi_2(\omega) = x_2 \in R$, …, $\xi_n(\omega) = x_n \in R$ и $\eta_i(\omega) = g_i(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) \in R$. Следовательно, построено однозначное отображение $\eta_i(\omega)\colon \Omega \to R$. Из определения борелевской функции $g_i(x_1, x_2, \ldots, x_n)\colon R^n \to R$ следует, что для борелевского множества B на R имеем: $\{(x_1, x_2, \ldots, x_n)\colon g_i(x_1, x_2, \ldots, x_n) \in B\} = g_i^{-1}(B) \in \mathcal{B}^n$, где множество $g_i^{-1}(B)$ есть полный прообраз множества B и $\mathcal{B}^n$ — борелевская σ-алгебра в R^n. Учитывая это и измеримость отображения $(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))\colon \Omega \to R^n$, находим, что множество

$$\{\omega\colon \eta_i(\omega) \in B\} = \{\omega\colon g_i(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) \in B\} = \\ = \{\omega\colon (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) \in g_i^{-1}(B)\} \in \mathcal{F}.$$

Поэтому отображение $\eta_i(\omega)$ является случайной величиной. Так как $\eta_1(\omega)$, $\eta_2(\omega), \ldots, \eta_m(\omega)$ являются одномерными случайными величинами на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, то естественно рассмотреть и случайный вектор вида $\eta = (\eta_1, \eta_2, \ldots, \eta_m)$.

Будем теперь предполагать, что законы распределения n-мерного случайного вектора $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ и их свойства нам хорошо известны. Напротив, распределение m-мерной случайной величины $(\eta_1, \eta_2, \ldots, \eta_m)$ необходимо определить. В силу этого случайные величины ξ_1, $\xi_2, \ldots, \xi_n$ будем называть тестовыми. Очевидно, что для практики наибольший интерес представляют следующие две последовательно решаемые задачи:

1) во-первых, задача нахождения функциональной зависимости между каждой случайной величиной с неизвестным законом распределения и тестовыми случайными величинами;

2) во-вторых, задача вычисления неизвестных законов распределения случайного вектора η по информации о функциональных зависимостях (9.1) и о законах распределения многомерной случайной величины ξ.

В дальнейшем будем считать, что первая задача решена теми или иными средствами и, значит, функциональные зависимости (9.1) априори заданы. Перейдем теперь к подробному решению второй задачи в наиболее общем виде. Очевидно, что интегральную функцию распределения каждой случайной величины $\eta_i(\omega)$ можно найти следующим образом:

$$\begin{aligned} F_i(y_i) &= \mathsf{P}(\{\omega\colon \eta_i(\omega) < y_i\}) = \\ &= \mathsf{P}(\{\omega\colon g_i(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) \in (-\infty, y_i)\}) = \\ &= \mathsf{P}(\{\omega\colon (\xi_1, \xi_2, \ldots, \xi_n) \in g_i^{-1}((-\infty, y_i))\}), \quad -\infty < y_i < +\infty, \end{aligned} \tag{9.2}$$

где $g_i^{-1}((-\infty, y_i)) = \{(x_1, x_2, \ldots, x_n)\colon g_i(x_1, x_2, \ldots, x_n) < y_i\}$. Так как

$$\{\omega\colon \eta_i(\omega) < y_i,\ i = 1, 2, \ldots, m\} = \\ = \bigcap_{i=1}^{m} \{\omega\colon g_i(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) \in (-\infty, y_i)\} = \\ = \bigcap_{i=1}^{m} \{\omega\colon (\xi_1, \xi_2, \ldots, \xi_n) \in g_i^{-1}(-\infty, y_i))\} = \Big\{\omega\colon \xi \in \bigcap_{i=1}^{m} g_i^{-1}((-\infty, y_i))\Big\},$$

где $-\infty < y_1 < +\infty$, $-\infty < y_2 < +\infty, \ldots$, $-\infty < y_m < +\infty$. Поэтому функция распределения $F_\eta(y_1, y_2, \ldots, y_m)$ случайного вектора $\eta = (\eta_1, \eta_2, \ldots, \eta_m)$ равна

$$F_\eta(y_1, y_2, \ldots, y_m) = \mathsf{P}(\{\omega\colon \eta_i(\omega) < y_1,\ \eta_2(\omega) < y_2, \ldots \eta_m(\omega) < y_m\}) = \\ = \mathsf{P}\Big(\Big\{\omega\colon (\xi_1, \xi_2, \ldots, \xi_n) \in \bigcap_{i=1}^{m} g_i^{-1}((-\infty, y_i))\Big\}\Big). \quad (9.3)$$

Хорошо известно, что вероятности вида $\mathsf{P}(\{\omega\colon \xi \in g_i^{-1}((-\infty, y_i))\})$ и $\mathsf{P}(\{\omega\colon \xi \in \cap_{i=1}^{m} g_i^{-1}((-\infty, y_i))\})$ всегда могут быть вычислены через заданный закон распределения тестовой многомерной случайной величины $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$. В общем случае вычисления законов распределения $F_i(y_i)$ и $F_\eta(y_1, y_2, \ldots, y_m)$ с помощью формул (9.2) и (9.3) очень громоздки. В силу этого рассмотрим решение этой задачи для ряда конкретных случаев, а именно, для дискретных и непрерывных случайных величин, для $n = m = 1$, для $n \geqslant 2$ и $m = 1$, наконец, для $n \geqslant 2$ и $m \geqslant 2$. Перечисленные случай очень важны для практики. Сначала ограничимся простейшим случаем дискретных или непрерывных случайных величин и $n = m = 1$.

1.2. Дискретные одномерные случайные величины и их функциональная зависимость. Пусть $\xi(\omega)\colon \Omega \to R$ есть дискретная случайная величина и отображение $g(x)\colon R \to R$ является однозначным. Функция $\eta(\omega) = g(\xi(\omega))\colon \Omega \to R$, которая является суперпозицией функций $\xi(\omega)$ и $g(x)$, не может принимать больше значений, чем дискретная случайная величина $\xi(\omega)$. В самом деле, если случайная величина ξ принимает значения x_1', $x_2', \ldots$, то числа $g(x_1')$, $g(x_2'), \ldots$ будут возможными значениями функции η. Так как некоторые из чисел $g(x_1')$, $g(x_2'), \ldots$ могут быть равными, то мощность множества различных значений функции η не более мощности множества $\{x_1', x_2', \ldots\}$. Обозначим теперь через y_1', $y_2', \ldots$, различные значения функции η. Далее, так как при любом y_j' и любом $y \in R$ справедливы соотношения вида:

$$\begin{aligned} \{\omega\colon \eta(\omega) = y_j'\} &= \bigcup_{i\colon g(x_i') = y_j'} \{\omega\colon \xi(\omega) = x_i'\} \in \mathcal{F}, \\ \{\omega\colon \eta(\omega) < y\} &= \bigcup_{j\colon y_j' < y} \{\omega\colon \eta(\omega) = y_j'\} = \\ &= \bigcup_{j\colon y_j' < y}\ \bigcup_{i\colon g(x_i') = y_j'} \{\omega\colon \xi(\omega) = x_i'\} \in \mathcal{F}, \end{aligned} \quad (9.4)$$

то η будет дискретной случайной величиной. Итак, в случае дискретной случайной величины ξ на функцию $g(x)$ никаких дополнительных ограничений не накладывается.

Из соотношения (9.4) легко вычислить законы распределения случайной величины η через законы распределения случайной ξ: для дискретной случайной величины $\xi(\omega)$ с известной интегральной функцией $F_\xi(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\})$, $-\infty < x < +\infty$, всегда можно задать распределение в виде: $\mathsf{P}(\{\omega\colon \xi(\omega) = x'_i\}) = F_\xi(x'_i + 0) - F_\xi(x'_i) = p_i$, $i = 1, 2, \ldots$. Обозначая распределение одномерной дискретной случайной величины η через $q_j = \mathsf{P}(\{\omega\colon \eta(\omega) = y'_j\})$, $j = 1, 2, \ldots$, и используя соотношение (9.4), последовательно найдем ее законы распределения:

$$q_j = \mathsf{P}(\{\omega\colon \eta(\omega) = y'_j\}) = \sum_{i\colon g(x'_i)=y'_j} !\mathsf{P}(\{\omega\colon \xi(\omega) = x'_i\}) = \sum_{i\colon g(x'_i)=y'_j} p_i,$$

$$F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \\ = \sum_{j\,:\,y'_j<y} q_i = \sum_{j\,:\,y'_j<y}\ \sum_{i\colon g(x'_i)=y'_j} p_i, \quad -\infty < y < +\infty.$$

Приведем простой пример на использование полученных формул.

Пример 9.1. Пусть закон распределения дискретной случайной величины ξ задан в виде табл. 9.1.

Таблица 9.1

ξ	$x'_1 = -3$	$x'_2 = -1$	$x'_3 = 0$	$x'_4 = 1$	$x'_5 = 3$	$x'_6 = 10$
$\mathsf{P}(\cdot)$	$1/12$	$1/12$	$1/6$	$1/6$	$1/6$	$1/3$

Предположим, что задано следующее отображение: $y = g(x) = x^2$. Найти закон распределения случайной величины $\eta(\omega) = g(\xi(\omega)) = (\xi(\omega))^2$.

Решение. Случайная величина $\eta(\omega)$ будет принимать значения 9, 1, 0, 1, 9 и 100, которые соответствуют различным значениям случайной величины ξ. Различных значений для случайной величины $\eta(\omega)$ всего четыре, а именно, имеем $y'_1 = 0$, $y'_2 = 1$, $y'_3 = 9$, $y'_4 = 100$. Соответствующие вероятности q_j для значений случайной величины η вычисляются следующим образом:

$$q_1 = \mathsf{P}(\{\omega\colon \eta(\omega) = y'_1\}) = \mathsf{P}(\{\omega\colon \eta(\omega) = 0\}) = \sum_{i\colon g(x'_i)=0} p_i = p_3 = 1/6,$$

$$q_2 = \mathsf{P}(\{\omega\colon \eta(\omega) = y'_2\}) = \mathsf{P}(\{\omega\colon \eta(\omega) = 1\}) = \sum_{i\colon g(x'_i)=1} p_i = p_2 + p_4 = 1/4,$$

$$q_3 = \mathsf{P}(\{\omega\colon \eta(\omega) = y'_3\}) = \mathsf{P}(\{\omega\colon \eta(\omega) = 9\}) = \sum_{i\colon g(x'_i)=9} p_i = p_1 + p_5 = 1/4,$$

$$q_4 = \mathsf{P}(\{\omega\colon \eta(\omega) = y'_4\}) = \mathsf{P}(\{\omega\colon \eta(\omega) = 100\}) = \sum_{i\colon g(x'_i)=100} p_i = p_6 = 1/3.$$

Отсюда легко изобразить ряд распределения для дискретной случайной величины $\eta(\omega)$ в виде табл. 9.2.

Таблица 9.2

$\eta(\omega)$	$y'_1 = 0$	$y'_2 = 1$	$y'_3 = 9$	$y'_4 = 100$
$\mathsf{P}(\cdot)$	1/6	1/4	1/4	1/3

1.3. Непрерывные одномерные случайные величины и их функциональная зависимость. Для случайной величины ξ, которая не является дискретной, в силу леммы 6.3 будем требовать, чтобы функция $g(x)$ была борелевской. Если $g(x)$ — борелевская функция, то по лемме 6.3 функция $\eta(\omega) = g(\xi(\omega))\colon \Omega \to R$ является случайной величиной и $\sigma(\eta) \subset \sigma(\xi)$. Более того, при доказательстве этой леммы было установлено, что $\{\omega\colon \eta(\omega) \in B\} = \{\omega\colon \xi(\omega) \in g^{-1}(B)\}$. Отсюда следует, что распределение $\mathsf{P}(\{\omega\colon \eta(\omega) \in B\})$, $B \in \mathcal{B}$, случайной величины $\eta(\omega)$ можно вычислить через вероятности $\mathsf{P}(\{\omega\colon \xi(\omega) \in g^{-1}(B)\})$, $B \in \mathcal{B}$, которые определяются распределением случайной величины ξ. В частности, при $B = (-\infty, y)$ можем найти функцию $F_\eta(y)$ для случайной величины $\eta(\omega)$. Способы и сложность вычисления интегральной функции $F_\eta(y)$ существенно зависят от типа случайной величины ξ и от свойств функции $g(x)$.

Пусть теперь ξ — непрерывная случайная величина с функцией распределения $F_\xi(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\})$ и плотностью распределения $f_\xi(x) = dF_\xi(x)/dx$. Найдем интегральную функции $F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\})$ и плотность $f_\eta(y) = dF_\eta(y)/dy$ для случайной величины η, если считаем распределение для случайной величины ξ известным. Предположим сначала, что функция $g(x)$ строго возрастает и имеет первую производную, т. е. имеем $dg(x)/dx > 0$. В этом случае из курса математического анализа известно, что для функции $y = g(x)$ существуют обратная функция $x = w(y)$ и ее производная $w'(y) = dw(y)/dy > 0$. Пусть ради простоты случайная величина $\xi(\omega)$ принимает значения из отрезка $[a, b]$ для любого $\omega \in \Omega$. Значит, плотность вероятности $f_\xi(x)$ равна нулю при $x \notin [a, b]$. Если теперь при $x = a$ функция $g(a) = a_1$ и при $x = b$ функция $g(b) = b_1$, то случайная величина $\eta(\omega)$ принимает значения из отрезка $[a_1, b_1]$. На рис. 9.1 отображена именно такая ситуация.

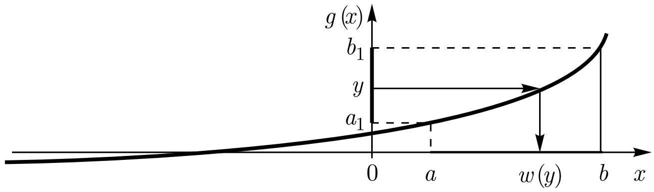


Рис. 9.1

Учитывая график функции $y = g(x)$ и свойства функции $x = w(y)$, легко найдем значение интегральной функции распределения $F_\eta(y)$ при любом $y \in R$. Ясно, что при $y \leqslant a_1$ интегральная функция распределения $F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \mathsf{P}(\varnothing) = 0$. Далее, при $a_1 < y \leqslant b_1$ функция

$$F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \mathsf{P}(\{\omega\colon \xi(\omega) < x = w(y)\}) = \\ = F_\xi(w(y)) = \int\limits_{-\infty}^{w(y)} f_\xi(u)\,du = \int\limits_{a}^{w(y)} f_\xi(u)\,du.$$

При $y > b_1$ функция $F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \mathsf{P}(\Omega) = 1$. Отсюда легко видеть, что $f_\eta(y) = 0$ при $y < a_1$, $f_\eta(y) = f_\xi(w(y))\,w'(y)$ при $a_1 \leqslant y \leqslant b_1$ и, наконец, $f_\eta(y) = 0$ при $y > b_1$.

Пример 9.2 (*задача имитации*). Изложенный в предыдущем абзаце прием часто используется при имитационном моделировании реальных случайных явлений на компьютере. В этом случае решается задача получения значений некоторой случайной величины η с заданной интегральной функцией распределения $F_\eta(x) = g(x)$. При этом допускается возможность вычисления реализаций равномерно распределенной на отрезке $[0,1]$ случайной величины ξ на компьютере с помощью стандартных программ и датчиков «псевдослучайных чисел». Предположим, что функция $y = F_\eta(x)$ является непрерывной и строго возрастающей. Тогда существует однозначная обратная функция $x = w(y)$, также непрерывная и строго возрастающая. Выберем теперь в качестве функциональной зависимости между случайными величинами ξ и η равенство $\eta = w(\xi)$. Из этих предположений, принимая во внимание, что плотность распределения $f_\xi(y)$ случайной величины ξ равна единице при $y \in [0,1]$ и равна нулю при $y \notin [0,1]$, имеем:

$$\mathsf{P}(\{\omega\colon \eta(\omega) < x\}) = \mathsf{P}(\{\omega\colon w(\xi(\omega)) < x\}) = \\ = \mathsf{P}(\{\omega\colon \xi(\omega) < F_\eta(x)\}) = \int\limits_{-\infty}^{F_\eta(x)} f_\xi(y)\,dy = \int\limits_{0}^{F_\eta(x)} dy = F_\eta(x).$$

Отсюда следует, что значение x случайной величины η с заданной интегральной функцией распределения $F_\eta(x)$ можно вычислить с помощью формулы $x = w(y)$, где y — значение случайной величины ξ. Наоборот, пусть мы умеем вычислять значение x случайной величины η с заданной интегральной функцией распределения $F_\eta(x)$. Тогда с помощью формулы $y = F_\eta(x)$ мы можем найти соответствующую реализацию случайной величины ξ, равномерно распределенной на отрезке $[0,1]$. Действительно, если $\xi = F_\eta(\eta)$, то

$$\mathsf{P}(\{\omega\colon \xi(\omega) < y\}) = \mathsf{P}(\{\omega\colon F_\eta(\eta(\omega)) < y\}) = \mathsf{P}(\Omega) = 1$$

при $y > 1$ и

$$\mathsf{P}(\{\omega\colon \xi(\omega) < y\}) = \mathsf{P}(\{\omega\colon F_\eta(\eta(\omega)) < y\}) = \mathsf{P}(\varnothing) = 0$$

при $y \leqslant 0$. Наконец, при $0 \leqslant y \leqslant 1$ имеем

$$\mathsf{P}(\{\omega\colon \xi(\omega) < y\}) = \mathsf{P}(\{\omega\colon \eta(\omega) < w(y)\}) = F_\eta(w(y)) = F_\eta(x) = y.$$

Значит, случайная величина ξ равномерно распределена на отрезке $[0,1]$.

Рассмотрим теперь несколько другой вариант такого рода задач. Пусть функция $g(x)$ строго убывает и имеет первую производную, т. е. $g'(x) < 0$. График функции $g(x)$ изображен на рис. 9.2.

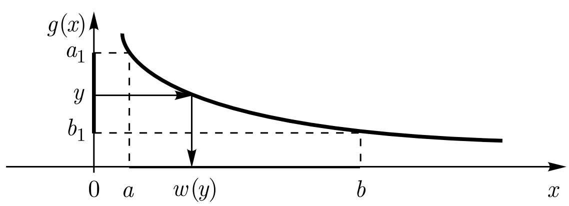


Рис. 9.2

И в этом случае известно, что для функции $y = g(x)$ существует обратная функция вида $x = w(y)$, но ее производная уже отрицательна, т. е. $w'(y) < 0$. Пусть опять для простоты случайная величина $\xi(\omega)$ принимает значения из отрезка $[a,b]$ для любого $\omega \in \Omega$. Очевидно, что при $x = a$ функция $g(a) = a_1$ и при $x = b$ функция $g(b) = b_1$. Следовательно, случайная величина $\eta(\omega)$ принимает значения из отрезка $[b_1, a_1]$. Эти свойства функций $g(x)$, $w(y)$, $\xi(\omega)$ и $\eta(\omega)$ легко можно пронаблюдать на рис. 9.2. Следуя методике предыдущего абзаца, вычислим при любом значении переменной $y \in R$ значение функции $F_\eta(y)$. При $y \leqslant b_1$ интегральная функция распределения $F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \mathsf{P}(\varnothing) = 0$. Затем, при $a_1 < y \leqslant b_1$ функция

$$\begin{aligned} F_\eta(y) &= \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \mathsf{P}(\{\omega\colon \xi(\omega) > w(y)\}) = \\ &= 1 - \mathsf{P}(\{\omega\colon \xi(\omega) \leqslant w(y)\}) = 1 - \mathsf{P}(\{\omega\colon \xi(\omega) < w(y)\}) = \\ &= 1 - F_\xi(w(y)) = 1 - \int\limits_{-\infty}^{w(y)} f_\xi(u)\,du = 1 - \int\limits_{a}^{w(y)} f_\xi(u)\,du. \end{aligned}$$

Наконец, при $y > a_1$ функция $F_\eta(y)$ равна $\mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \mathsf{P}(\Omega) = 1$. Поэтому плотность вероятностей для случайной величины η определяется следующим образом: $f_\eta(y) = 0$ при $y < b_1$, $f_\eta(y) = -f_\xi(w(y))\,w'(y)$ при $b_1 \leqslant y \leqslant a_1$ и $f_\eta(y) = 0$ при $y > a_1$.

Суммируя результаты предыдущих абзацев этого раздела и учитывая теперь знак производной $w'(y)$, можно написать общую формулу $f_\eta(y) = f_\xi(w(y))\,|w'(y)|$ для плотности вероятности. Здесь y принимает значения из промежутка значений случайной величины η. В остальной части действительной прямой плотность вероятности $f_\eta(y)$, очевидно, равна нулю.

Аналогичным образом можно рассмотреть несколько другой случай. Пусть для области определения дифференцируемой функции $g(x)$ можно выделить как участки монотонного возрастания, так и участки монотонного убывания. Значит, функция $g(x)$ является кусочно-монотонной и дифференцируемой. Выделяя участки монотонного возрастания и участки монотонного убывания, и повторяя предыдущие рассуждения с небольшими изменениями, можно найти интегральную функцию распределения и плотность вероятности для случайной величины η. В качестве иллюстрации сказанного рассмотрим простой пример.

Пример 9.3. Найти закон распределения случайной величины $\eta(\omega) = (\xi(\omega))^2$, если случайная величина ξ имеет плотность распределения вероятностей $f_\xi(x) = \exp\{-x_2/(2\,\sigma^2)\}/(\sigma\sqrt{2\pi}\,)$, $\sigma > 0$, $-\infty < x < +\infty$.

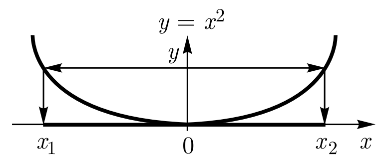


Рис. 9.3

Решение. Для этой функции $y = x^2$ (рис. 9.3) можно выделить два промежутка монотонности, а именно, полуинтервалы $(-\infty, 0\,]$ и $[\,0, +\infty)$.

Соответственно на этих полуинтервалах легко вычисляются обратные функции. В самом деле, из график функции $y = g(x) = x^2$ на рис. 9.3 видно, что для первого промежутка $(-\infty, 0\,]$ обратная функция $x_1 = w_1(y) = -(y)^{1/2}$ и для второго промежутка обратная функция $x_2 = w_2(y) = +(y)^{1/2}$.

Интегральная функция распределения $F_\eta(y)$ для случайной величины η при $y > 0$ вычисляется следующим образом:

$$\begin{aligned}
F_\eta(y) &= \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \mathsf{P}(\{\omega\colon w_1(y) < \xi < 0\} \cup \{\omega\colon 0 \leqslant \xi < w_2(y)\}) = \\
&= \mathsf{P}(\{\omega\colon w_1(y) < \xi(\omega) < 0\}) + \mathsf{P}(\{\omega\colon 0 \leqslant \xi(\omega) < w_2(y)\}) = \\
&= \int\limits_{w_1(y)}^{0} f_\xi(x)\,dx + \int\limits_{0}^{w_2(y)} f_\xi(x)\,dx = (1/(\sigma\sqrt{2\pi}\,))\int\limits_{-\sqrt{y}}^{\sqrt{y}} \exp\{-x^2/(2\sigma^2)\}\,dx = \\
&= 2^{1/2}\sigma^{-1}\pi^{-1/2}\int\limits_{0}^{\sqrt{y}} \exp\{-x^2/(2\sigma^2)\}\,dx.
\end{aligned}$$

Так как $\eta(\omega) = (\xi(\omega))^2 \geqslant 0$ для всех $\omega \in \Omega$, то интегральная функция распределения $F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) = \mathsf{P}(\varnothing) = 0$ при $y \leqslant 0$. Отсюда находим плотность распределения $f_\eta(y)$ случайной величины η: $f_\eta(y) = 0$ при $y \leqslant 0$ и $f_\eta(y) = (\sigma\sqrt{2\pi y}\,)^{-1}\exp\{-y/(2\,\sigma^2)\}$ при $y > 0$.

1.4. Многомерные случайные величины и их функциональная зависимость. Переходим к рассмотрению такого варианта, при котором $n \geqslant 2$, $m = 1$ и $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ является многомерной дискретной случайной величиной. В этом случае роль закона распределения случайного вектора ξ может играть набор вероятностей $\mathsf{P}(\{\omega\colon \xi_k(\omega) = z_k,\ k = 1, 2, \ldots, n\}) = p_{z_1, z_2, \ldots, z_n}$, $z_1 \in Z_1$, $z_2 \in Z_2, \ldots$, $z_n \in Z_n$, где z_k принимает счетное число различных значений из множества Z_k значений одномерной случайной величины $\xi_k(\omega)\colon \Omega \to Z_k$ для каждого $k = 1, 2, \ldots, n$. Пусть теперь случайная величина $\eta(\omega) = g(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$. В этом случае, принимая во внимание соотношение (9.2), функция $F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\})$ для $\eta(\omega)$ равна $\mathsf{P}(\{\omega\colon (\xi_1, \xi_2, \ldots, \xi_n) \in g^{-1}((-\infty, y))\}) = \sum_{g(z_1, z_2, \ldots, z_n) < y} p_{z_1, z_2, \ldots, z_n}$, $-\infty < y < +\infty$. Здесь суммирование ведется по таким наборам $(z_1, z_2, \ldots, z_n)$, для которых $z_1 \in Z_1$, $z_2 \in Z_2, \ldots$, $z_n \in Z_n$ и $g(z_1, z_2, \ldots, z_n) < y$.

Если теперь $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ есть непрерывная многомерная случайная величина с плотностью распределения $f_\xi(x_1, x_2, \ldots, x_n)$, то согласно (9.2) закон распределения $F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\})$ случайной величины $\eta(\omega) = g(\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega))$ при всех $-\infty < y < +\infty$ примет вид:

$$\begin{aligned} F_\eta(y) = \mathsf{P}(\{\omega\colon \eta(\omega) < y\}) &= \mathsf{P}(\{\omega\colon (\xi_1, \xi_2, \ldots, \xi_n) \in g^{-1}((-\infty, y))\}) = \\ &= \mathsf{P}(\{\omega\colon (\xi_1, \xi_2, \ldots, \xi_n) \in B^n\}) = \\ &= \int\limits_{B^n}\!\!\int \ldots \int f_\xi(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n. \end{aligned} \tag{9.5}$$

Здесь множество

$$B^n = g^{-1}((-\infty, y)) = \{(x_1, x_2, \ldots, x_n)\colon g(x_1, x_2, \ldots, x_n) < y\},$$

которое является областью интегрирования, зависит от y. Если $F_\eta(y)$ является абсолютно непрерывной функцией, то можно найти плотность распределения $f_\eta(y)$ для случайной величины $\eta(\omega)$ путем дифференцирования правой части равенства (9.5) по y. Рассмотрим пример такого случая, который часто встречается на практике.

Пример 9.4. Определить закон распределения суммы двух случайных величин ξ_1 и ξ_2, если плотность распределения двумерной непрерывной случайной величины $\xi = (\xi_1, \xi_2)$ равна $f_\xi(x_1, x_2)$.

Решение. Обозначим через η случайную величину, равную $g(\xi_1, \xi_2) = \xi_1 + \xi_2$, и через $F_\eta(y)$ функцию распределения случайной величины η. Функция $F_\eta(y)$ определяется из (9.5) при $n = 2$, $g(x_1, x_2) = x_1 + x_2$:

$$F_\eta(y) = \iint\limits_{B^2=\{(x_1,x_2)\colon\, x_1+x_2<y\}} f_\xi(x_1, x_2)\, dx_1\, dx_2 =$$

$$= \int\limits_{-\infty}^{+\infty} \Bigl(\int\limits_{-\infty}^{y-x_1} f_\xi(x_1, x_2)\, dx_2 \Bigr) dx_1 = \int\limits_{-\infty}^{+\infty} \Bigl(\int\limits_{-\infty}^{y-x_2} f_\xi(x_1, x_2)\, dx_1 \Bigr) dx_2,$$

$$-\infty < y < +\infty.$$

Здесь область $B^2 = \{(x_1, x_2)\colon x_1 + x_2 < y\}$ плоскости $x_1 0 x_2$ заштрихована на рис. 9.4.

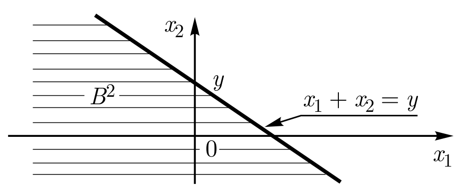


Рис. 9.4

Плотность распределения $f_\eta(y)$ случайной величины η примет вид: $f_\eta(y) = F'_\eta(y) = \int_{-\infty}^{+\infty} f_\xi(x_1, y - x_1)\, dx_1 = \int_{-\infty}^{+\infty} f_\xi(y - x_2, x_2)\, dx_2$, $-\infty < y < +\infty$. В том случае, когда случайные величины ξ_1 и ξ_2 независимы, т. е. плотность распределения $f_\xi(x_1, x_2) = f_{\xi_1}(x_1) \times f_{\xi_2}(x_2)$, получим $f_\eta(y) = \int_{-\infty}^{+\infty} f_{\xi_1}(x_1)\, f_{\xi_2}(y - x_1)\, dx_1 = \int_{-\infty}^{+\infty} f_{\xi_1}(y - x_2)\, f_{\xi_2}(x_2)\, dx_2$. Каждая из этих формул для $f_\eta(y)$ называется *сверткой функций* $f_{\xi_1}(x_1)$ и $f_{\xi_2}(x_2)$. Составление закона распределения суммы независимых случайных величин по законам распределения слагаемых называется *композицией законов распределения*.

Наконец, рассмотрим теперь случай, когда заданы n-мерная дискретная или непрерывная случайная величина $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ и m борелевских функций $g_1(x_1, x_2, \ldots, x_n)$, $g_2(x_1, x_2, \ldots, x_n), \ldots$, $g_m(x_1, x_2, \ldots, x_n)$. Пусть сначала $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ есть многомерная дискретная случайная величина и $\xi_k(\omega)\colon \Omega \to Z_k \subset R$ для каждого $k = \overline{1, n}$. Используя формулу (9.3), вычислим совместную функцию распределения $F_\eta(y_1, y_2, \ldots, y_m)$ случайного вектора $\eta = (\eta_1, \eta_2, \ldots, \eta_m)$ следующим образом:

$$F_\eta(y_1, y_2, \ldots, y_m) = \mathsf{P}\Bigl(\Bigl\{\omega\colon (\xi_1, \xi_2, \ldots, \xi_n) \in \bigcap_{i=1}^{m} g_i^{-1}((-\infty, y_i))\Bigr\}\Bigr) =$$

$$= \sum_{B^n} p_{z_1,z_2,\ldots,z_n}, \quad -\infty < y_1 < +\infty, -\infty < y_2 < +\infty, \ldots, -\infty < y_m < +\infty,$$

где суммирование ведется по всевозможным наборам $(z_1, z_2, \ldots, z_n)$ из множества $B^n = \{(z_1, z_2, \ldots, z_n)\colon z_k \in Z_k,\ g_k(z_1, z_2, \ldots, z_n) < y_k,\ k = 1, 2, \ldots, n\}$. Если теперь $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ есть непрерывный случайный вектор с плотностью распределения вероятностей $f_\xi(x_1, x_2, \ldots, x_n)$, то m-мерный интегральный закон $F_\eta(y_1, y_2, \ldots, y_m)$ вектора $\eta = (\eta_1, \eta_2, \ldots, \eta_m)$ при $\mathbf{y} = (y_1, y_2, \ldots, y_m)$, $-\infty < y_1 < +\infty$, $-\infty < y_2 < +\infty, \ldots$, $-\infty < y_m < +\infty$, можно определить с помощью формулы (9.3), а именно:

$$F_\eta(y_1, y_2, \ldots, y_m) = \mathsf{P}\Big(\Big\{\omega\colon (\xi_1, \xi_2, \ldots, \xi_n) \in \bigcap_{i=1}^{m} g_i^{-1}((-\infty, y_i))\Big\}\Big) = \\ = \int \underset{B^n(\mathbf{y})}{\int} \ldots \int f_\xi(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n, \tag{9.6}$$

где множество $B^n(\mathbf{y}) = \{(x_1, x_2, \ldots, x_n)\colon g_k(x_1, x_2, \ldots, x_n) < y_k,\ k = 1, 2, \ldots, m\}$. Рассмотрим важный для приложений частный случай формулы (9.6), когда $n = m$. Предположим теперь, что две системы функций:

$$v_1 = g_1(x_1, x_2, \ldots, x_n), v_2 = g_2(x_1, x_2, \ldots, x_n), \ldots, v_n = g_n(x_1, x_2, \ldots, x_n), \tag{9.7}$$

$$x_1 = t_1(v_1, v_2, \ldots, v_n), x_2 = t_2(v_1, v_2, \ldots, v_n), \ldots, x_n = t_n(v_1, v_2, \ldots, v_n) \tag{9.8}$$

устанавливают взаимно однозначное соответствие между n-мерной областью $C^n \subset R^n$ в пространстве значений действительных переменных x_1, $x_2, \ldots$, x_n и n-мерной областью $D^n \subset R^n$ в пространстве значений действительных переменных v_1, $v_2, \ldots$, v_n. При этом область C^n является множеством всех возможных значений случайного вектора $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ с непрерывной плотностью распределения $f_\xi(x_1, x_2, \ldots, x_n)$, а область D^n является множеством всех возможных значений многомерной случайной величины $\eta = (\eta_1, \eta_2, \ldots, \eta_n)$. Если n-мерные области C^n и D^n не совпадают с R^n, то их границы по предположению должны быть кусочно-гладкими поверхностями в n-мерном пространстве. При $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ и $-\infty < y_1 < +\infty$, $-\infty < y_2 < +\infty, \ldots$, $-\infty < y_n < +\infty$ обозначим через $D^n(\mathbf{y}) \subset D^n$ образ множества $B^n(\mathbf{y}) \cap C^n$ при отображении (9.7). Другими словами, формулы вида (9.7) осуществляют преобразование области $B^n(\mathbf{y}) \cap C^n$ в область $D^n(\mathbf{y})$, а соотношения (9.8), естественно, дают обратное преобразование области $D^n(\mathbf{y})$ в область $B^n(\mathbf{y}) \cap C^n$. Далее, считаем, что функции в соотношениях (9.7) и (9.8) непрерывны и имеют непрерывные частные производные первого порядка. При сделанных дополнительных и очень жестких ограничениях имеет место известная из курса математического анализа теорема о замене переменных в n-кратном интеграле. В силу этого формула (9.6) при $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, $-\infty < y_1 < +\infty$, $-\infty < y_2 < +\infty, \ldots$, $-\infty < y_n < +\infty$ и $n = m$ может быть преобразована следующим образом:

$$\begin{aligned} F_\eta(y_1, y_2, \ldots, y_n) &= \int\limits_{B^n(\mathbf{y})}\!\!\int \ldots \int f_\xi(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n = \\ &= \int\limits_{B^n(\mathbf{y}) \cap C^n}\!\!\int \ldots \int f_\xi(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n = \\ &= \int\limits_{D^n(\mathbf{y})}\!\!\int \ldots \int f_\xi(t_1(\cdot), t_2(\cdot), \ldots, t_n(\cdot))|J|\, dv_1\, dv_2 \ldots dv_n = \\ &= \int\limits_{D^n(\mathbf{y})}\!\!\int \ldots \int f_\eta(v_1, v_2, \ldots, v_n)\, dv_1\, dv_2 \ldots dv_n = \\ &= \int\limits_{-\infty}^{y_1} \int\limits_{-\infty}^{y_2} \ldots \int\limits_{-\infty}^{y_n} f_\eta(v_1, v_2, \ldots, v_n)\, dv_1\, dv_2 \ldots dv_n, \end{aligned} \tag{9.9}$$

где равенства $x_1 = t_1(v_1, v_2, \ldots, v_n) = t_1(\cdot)$, $x_2 = t_2(v_1, v_2, \ldots, v_n) = t_2(\cdot), \ldots$, $x_n = t_n(v_1, v_2, \ldots, v_n) = t_n(\cdot)$ являются обратными преобразованиями по отношению к системе функций $v_1 = g_1(x_1, x_2, \ldots, x_n)$, $v_2 = g_2(x_1, x_2, \ldots, x_n), \ldots$, $v_n = g_n(x_1, x_2, \ldots, x_n)$ и якобиан

$$J = J(v_1, v_2, \ldots, v_n) = \frac{D(t_1(\cdot), t_2(\cdot), \ldots, t_1(n\cdot))}{D(v_1, v_2, \ldots, v_n)} \neq 0$$

в n-мерной области D^n. Отсюда следует, что n-мерный случайный вектор $\eta = (\eta_1, \eta_2, \ldots, \eta_n)$ будет непрерывным, а его плотность распределения $f_\eta(v_1, v_2, \ldots, v_n)$, очевидно, равна $f_\xi(t_1(\cdot), t_2(\cdot), \ldots, t_n(\cdot))\,|J|$ в n-мерной области D^n и равна нулю в $R^n \backslash D^n$. Заметим, что этот факт на практике может иметь место и при более слабых ограничениях относительно плотности вероятности $f_\xi(x_1, x_2, \ldots, x_n)$ и преобразований (9.7) и (9.8). Например, требования, наложенные на функцию $f_\xi(x_1, x_2, \ldots, x_n)$ и на преобразования (9.7) и (9.8), могут нарушаться в отдельных точках или вдоль конечного числа кусочно-гладких линий и поверхностей в n-мерном пространстве R^n. Поясним сказанное следующим примером.

Пример 9.5. В области $C^2 = \{(x_1, x_2)\colon 0 < x_1 \leqslant 1,\ 0 < x_2 \leqslant 1\}$ на плоскости $x_1 0 x_2$ с перпендикулярными осями и с правой ориентацией (см. рис. 9.5), дана система функций

$$
\begin{aligned}
v_1 &= g_1(x_1, x_2) = (-\ln x_1)^{1/2} \cos(2\pi x_2),\\
v_2 &= g_2(x_1, x_2) = (-\ln x_1)^{1/2} \sin(2\pi x_2).
\end{aligned} \tag{9.10}
$$

Одномерные случайные величины ξ_1 и ξ_2 независимы и равномерно распределены на промежутке $(0, 1]$ и на отрезке $[0, 1]$ соответственно. Пусть случайные величины ξ_1 и ξ_2 принимают значения на оси абсцисс $0x_1$ и на оси ординат $0x_2$. Отсюда следует, что плотность распределения $f_\xi(x_1, x_2)$ случайного вектора $\xi = (\xi_1, \xi_2)$ равна единице при (x_1, x_2) из области C^2 и равна нулю при $(x_1, x_2) \notin C^2$. Найти закон распределения случайного вектора $\eta = (\eta_1, \eta_2)$, если величины $\eta_1 = g_1(\xi_1, \xi_2) = (-\ln \xi_1)^{1/2} \cos(2\pi\xi_2)$, $\eta_2 = g_2(\xi_1, \xi_2) = (-\ln \xi_1)^{1/2} \sin(2\pi\xi_2)$. Являются ли случайные величины η_1 и η_2 независимыми?

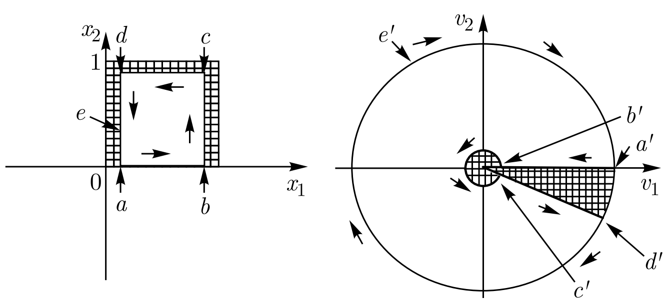


Рис. 9.5

Решение. Легко видеть (см. рис. 9.5), что система функций (9.10) каждой точке (x_1, x_2) из области $C^2 = \{(x_1, x_2)\colon 0 < x_1 \leqslant 1,\ 0 < x_2 \leqslant 1\}$ относит только одну точку (v_1, v_2) на плоскости $v_1 0 v_2$ с перпендикулярными осями и с правой ориентацией. Область $D^2 = \{(v_1, v_2)\colon -\infty < < v_1, v_2 < +\infty\}$ совпадает с R^2. Так как множество $\{\omega\colon \xi_1(\omega) = 0\} = \varnothing$ и $x_1 \neq 0$, то здесь мы имеем дело только с конечными вещественными значениями функций $g_1(x_1, x_2)$, $g_2(x_1, x_2)$, $\eta_1(\omega)$ и $\eta_2(\omega)$. Предположим, что случайные величины η_1 и η_2 принимают значения на оси абсцисс $0v_1$ и на оси ординат $0v_2$.

К сожалению, формулы преобразования (9.10) не будут однозначно разрешимы относительно x_1 и x_2. Действительно, при $v_1 = v_2 = 0$ из (9.10) получаем, что $x_1 = 1$ и $x_2 \in [0, 1]$. Следовательно, точке $(0, 0)$ на плоскости $v_1 0 v_2$ отвечает отрезок $\{(1, x_2)\colon 0 \leqslant x_2 \leqslant 1\}$ на плоскости $x_1 0 x_2$. Далее, любой фиксированной точке (v_1, v_2) из множества $\{(v_1, v_2)\colon v_1 > 0, v_2 = 0\}$ плоскости $v_1 0 v_2$ отвечают две точки $(\exp\{-v_1^2\}, 0)$ и $(\exp\{-v_1^2\}, 1)$ на плоскости $x_1 0 x_2$. Отсюда следует, что положительная ветвь оси абсцисс плоскости $v_1 0 v_2$ переходит либо в интервал $\{(x_1, 0)\colon 0 < x_1 < 1\}$ оси $0x_1$, либо в открытый промежуток $\{(x_1, 1)\colon 0 < x_1 < 1\}$ на плоскости $x_1 0 x_2$. Напротив, каждой точке (v_1, v_2) плоскости $v_1 0 v_2$, отличной от рассмотренных точек неотрицательной части оси $0v_1$, отвечает только одно значение x_1 и x_2 в следующих пределах: $0 < x_1 < 1$, $0 < x_2 < 1$. Эти значения для x_1 и x_2 определяются из (9.10) и вычисляются с помощью формул:

$$
\begin{gathered}
x_1 = t_1(v_1, v_2) = \\
= \exp\{-(v_1^2 + v_2^2)\}, (v_1, v_2) \in D^2 \backslash \{(v_1, v_2)\colon v_1 \geqslant 0, v_2 = 0\}, \\
x_2 = t_2(v_1, v_2) = \\
= \begin{cases} (2\pi)^{-1} \operatorname{arctg}(v_2 \times v_1^{-1}) & \text{при } v_1 > 0, v_2 > 0, \\ 1/4 & \text{при } v_1 = 0, v_2 > 0, \\ (2\pi)^{-1} \operatorname{arctg}(v_2 \times v_1^{-1}) + 1/2 & \text{при } v_1 < 0, -\infty < v_2 < \infty, \\ 3/4 & \text{при } v_1 = 0, v_2 < 0, \\ (2\pi)^{-1} \operatorname{arctg}(v_2 \times v_1^{-1}) + 1 & \text{при } v_1 > 0, v_2 < 0. \end{cases}
\end{gathered} \tag{9.11}
$$

Итак, в рассматриваемом примере для преобразования (9.10) и для его обратного преобразования (9.11) нарушаются условия непрерывности и взаимно однозначного соответствия в точках границы области C^2 или в точках неотрицательной части оси $0v_1$ плоскости $v_1 0 v_2$. Поэтому непосредственно нельзя применить теорему о замене переменных в двойном интеграле. Более того, область $D^2 = \{(v_1, v_2)\colon -\infty < v_1, v_2 < +\infty\}$ является неограниченной и охватывает всю плоскость. Эти трудности можно преодолеть следующим образом. Прежде всего, удалим из области C^2 на плоскости $x_1 0 x_2$ область Δ, которая на рис. 9.5 помечена двойной штриховкой. Тогда с помощью преобразования (9.10) область Δ сколь угодно малой площади перейдет в неограниченную область δ на плоскости $v_1 0 v_2$, а оставшаяся замкнутая область $C^2 \backslash \Delta$ перейдет в замкнутую и ограниченную область $D^2 \backslash \delta$. При этом точкам a, b, c, d, e границы области $C^2 \backslash \Delta$ отвечают точки a', b', c', d', e' границы области $D^2 \backslash \delta$. Из рис. 9.5 легко видеть, что положительному обходу контура a, b, c, d, e (область $C^2 \backslash \Delta$ остается слева) отвечает отрицательный обход контура a', b', c', d', e' (область $D^2 \backslash \delta$ остается справа). Заметим попутно, что область $D^2 \backslash \delta$ можно получить из большого круга на плоскости $v_1 0 v_2$ удалением области δ', помеченной на рис. 9.5 двойной штриховкой. К областям $C^2 \backslash \Delta$ и $D^2 \backslash \delta$ можно применить теорему о замене переменных в двойном интеграле. Воспользовавшись этим фактом, равенством $f_\xi(t_1(v_1, v_2), t_2(v_1, v_2)) = 1$ при $(v_1, v_2) \in D^2 \backslash \delta$ и соотношением (9.11), найдем:

$$\iint\limits_{C^2\setminus\Delta} f_\xi(x_1,x_2)\,dx_1dx_2 = \iint\limits_{D^2\setminus\delta} f_\xi(t_1(v_1,v_2),t_2(v_1,v_2))|J(v_1,v_2)|\,dv_1dv_2 = \\ = \iint\limits_{D^2\setminus\delta} \pi^{-1}\exp\{-(v_1^2+v_2^2)\}\,dv_1dv_2, \quad (9.12)$$

так как якобиан

$$J(v_1,v_2) = \frac{D(t_1(v_1,v_2),t_2(v_1,v_2))}{D(v_1,v_2)} = \begin{vmatrix} \dfrac{\partial t_1(v_1,v_2)}{\partial v_1} & \dfrac{\partial t_1(v_1,v_2)}{\partial v_2} \\ \dfrac{\partial t_2(v_1,v_2)}{\partial v_1} & \dfrac{\partial t_2(v_1,v_2)}{\partial v_2} \end{vmatrix} = \\ = \begin{vmatrix} -2v_1\exp\{-v_1^2-v_2^2\} & -2v_2\exp\{-v_1^2-v_2^2\} \\ (2\pi)^{-1}v_2(-v_1^2-v_2^2)^{-1} & (2\pi)^{-1}v_1(v_1^2+v_2^2)^{-1} \end{vmatrix} = \\ = -\pi^{-1}\exp\{-(v_1^2+v_2^2)\} \neq 0.$$

Здесь следует подчеркнуть, что отрицательное значение якобиана соответствует положительному обходу контура области $D^2\setminus\delta$ и отрицательному обходу контура области $C^2\setminus\Delta$. Переходим теперь в равенстве (9.12) к пределу при $\mathrm{mes}\,\Delta \to 0$ таким способом, чтобы контур a, b, c, d, e области $C^2\setminus\Delta$ всеми его точками приближался к границе области C^2. При этом контур области $D^2\setminus\delta$ всеми его точками будет удаляться в бесконечность, а область $D^2\setminus\delta$ будет стремиться к D^2. Из всего сказанного с учетом равенства $\iint_{C^2} f_\xi(x_1,x_2)\,dx_1\,dx_2 = \iint_{C^2} dx_1\,dx_2 = 1$ вытекают равенства

$$\iint\limits_{C^2} f_\xi(x_1,x_2)\,dx_1dx_2 = \iint\limits_{D^2} f_\xi(t_1(v_1,v_2),t_2(v_1,v_2))|J(v_1,v_2)|\,dv_1dv_2 = \\ = \iint\limits_{D^2} \pi^{-1}\exp\{-(v_1^2+v_2^2)\}\,dv_1dv_2 = 1,$$

Так как в данном примере области определения функций $g_1(x_1,x_2)$, $g_2(x_1,x_2)$ и множество значений вектора $\xi = (\xi_1,\xi_2)$ совпадают, то при любом $\mathbf{y} = (y_1,y_2) \in R^2$ имеет место: $B^2(\mathbf{y}) = \{(x_1,x_2)\colon g_1(x_1,x_2) < y_1; g_2(x_1,x_2) < y_2\} \subset C^2$, $B^2(\mathbf{y}) \cap C^2 = B^2(\mathbf{y})$. Применяя теперь к областям $B^2(\mathbf{y}) \subset C^2$, $D^2(\mathbf{y}) = \{(v_1,v_2)\colon v_1 < y_1;\ v_2 < y_2\} \subset D^2$ методику предыдущего абзаца и используя равенства (9.9), найдем для вектора $\boldsymbol{\eta} = (\eta_1,\eta_2)$ интегральную функцию:

$$F_{\boldsymbol{\eta}}(y_1,y_2) = \iint\limits_{B^2(\mathbf{y})} f_\xi(x_1,x_2)\,dx_1\,dx_2 = \\ = \iint\limits_{D^2(\mathbf{y})} f_\xi(t_1(v_1,v_2),t_2(v_1,v_2))\,|J(v_1,v_2)|\,dv_1\,dv_2 =$$

$$= \int_{-\infty}^{y_1}\int_{-\infty}^{y_2} \pi^{-1}\exp\{-(v_1^2+v_2^2)\}\,dv_1\,dv_2 = \int_{-\infty}^{y_1}\int_{-\infty}^{y_2} f_{\boldsymbol{\eta}}(v_1,v_2)\,dv_1\,dv_2.$$

Отсюда получаем решение задачи: $f_\eta(y_1,y_2) = \pi^{-1}\exp\{-(y_1^2+y_2^2)\}$, $(y_1,y_2)\in R^2$. Плотность $f_{\eta_1}(y_1)$ случайной величины η_1 и плотность распределения $f_{\eta_2}(y_2)$ величины η_2 вычисляются с помощью формулы (7.15):

$$f_{\eta_1}(y_1) = \int_{-\infty}^{+\infty} \pi^{-1}\exp\{-(y_1^2+y_2^2)\}\,dy_2 = \\ = \pi^{-1}\exp\{-y_1^2\}\pi^{1/2} = \pi^{-1/2}\exp\{-y_1^2\},$$

$$f_{\eta_2}(y_2) = \int_{-\infty}^{+\infty} \pi^{-1}\exp\{-(y_1^2+y_2^2)\}\,dy_1 = \\ = \pi^{-1}\exp\{-y_2^2\}\pi^{1/2} = \pi^{-1/2}\exp\{-y_2^2\}.$$

Так как $f_{\boldsymbol{\eta}}(y_1,y_2) = f_{\eta_1}(y_1)\times f_{\eta_2}(y_2)$ для всех $(y_1,y_2)\in R^2$, то случайные величины η_1 и η_2 являются независимыми.

Пример 9.6. Упорядоченная система $\boldsymbol{\xi} = (\xi_1,\xi_2)$ из одномерных случайных величин ξ_1 и ξ_2 равномерно распределена в области C^2 на плоскости $x_1 0 x_2$ с перпендикулярными осями и с правой ориентацией. Отображенная на рис. 9.6, *а* замкнутая область C^2 есть квадрат, стороны которого составляют углы $\pi/4$ с осями координат и равны по длине единице. Показать, что случайные величины ξ_1 и ξ_2 зависимы. Найти закон распределения случайного вектора $\boldsymbol{\eta} = (\eta_1,\eta_2)$, если $0 < \alpha \leqslant \pi/4$ и

$$\begin{aligned} \eta_1 &= g_1(\xi_1,\xi_2) = \xi_1\cos\alpha + \xi_2\sin\alpha, \\ \eta_2 &= g_2(\xi_1,\xi_2) = -\xi_1\sin\alpha + \xi_2\cos\alpha. \end{aligned} \tag{9.13}$$

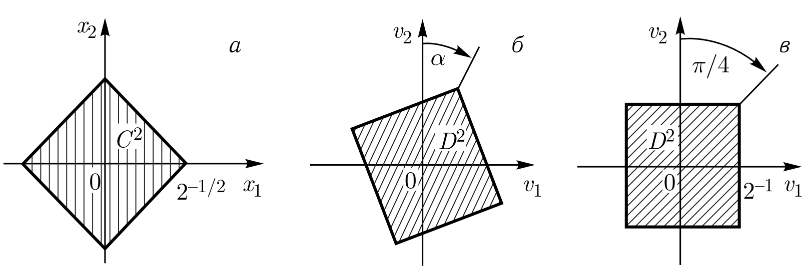


Рис. 9.6

Решение. Из условия задачи следует, что плотность $f_{\xi}(x_1, x_2)$ случайного вектора $\boldsymbol{\xi} = (\xi_1, \xi_2)$ равна единице при $(x_1, x_2) \in C^2$ и равна нулю при $(x_1, x_2) \notin C^2$. Применяя теперь равенства (7.15) и учитывая уравнения $x_1 + x_2 = 2^{-1/2}$, $-x_1 + x_2 = 2^{-1/2}$, $-x_1 - x_2 = 2^{-1/2}$, $x_1 - x_2 = 2^{-1/2}$ границы области C^2, легко вычислим плотности распределения случайных величин ξ_1 и ξ_2, а именно:

$$f_{\xi_1}(x_1) = \int\limits_{-\infty}^{+\infty} f_{\boldsymbol{\xi}}(x_1, x_2)\, dx_2 = \int\limits_{x_1 - 2^{-1/2}}^{2^{-1/2} - x_1} dx_2 = 2^{1/2} - 2x_1 \text{ при } 0 \leqslant x_1 \leqslant 2^{-1/2}$$

и

$$f_{\xi_1}(x_1) = \int\limits_{-x_1 - 2^{-1/2}}^{2^{-1/2} + x_1} dx_2 = 2^{1/2} + 2x_1 \quad \text{при} \quad -2^{-1/2} \leqslant x_1 \leqslant 0;$$

$$f_{\xi_2}(x_2) = \int\limits_{-\infty}^{+\infty} f_{\boldsymbol{\xi}}(x_1, x_2)\, dx_1 = \int\limits_{x_2 - 2^{-1/2}}^{2^{-1/2} - x_2} dx_1 = 2^{1/2} - 2x_2 \text{ при } 0 \leqslant x_2 \leqslant 2^{-1/2}$$

и

$$f_{\xi_2}(x_2) = \int\limits_{-x_2 - 2^{-1/2}}^{2^{-1/2} + x_2} dx_1 = 2^{1/2} + 2x_2 \quad \text{при} \quad -2^{-1/2} \leqslant x_2 \leqslant 0.$$

Ясно, что $f_{\xi_1}(x_1) = 0$ и $f_{\xi_2}(x_2) = 0$ при $|x_1|, |x_2| \geqslant 2^{-1/2}$. На рис. 9.7 приведены графики $f_{\xi_1}(x_1)$ и $f_{\xi_2}(x_2)$.

Итак, несмотря на равномерное распределение случайного вектора (ξ_1, ξ_2) в замкнутой области C^2, каждая из одномерных случайных величин ξ_1 и ξ_2 не распределена равномерно. Плотность распределения $f_{\xi_1}(x_1)$ или $f_{\xi_2}(x_2)$ носит название закона распределения Симпсона. Из неравенства $f_{\xi}(x_1, x_2) \neq f_{\xi_1}(x_1)\, f_{\xi_2}(x_2)$ при $(x_1, x_2) \in C^2$ следует, что случайные величины ξ_1 и ξ_2 являются статистически зависимыми.

Система функций (9.13) каждой точке $(\xi_1(\omega), \xi_2(\omega))$ из области C^2 на плоскости $x_1 0 x_2$ сопоставляет только одну точку $(\eta_1(\omega), \eta_2(\omega))$ из области D^2 на плоскости $v_1 0 v_2$ с перпендикулярными осями $0v_1$, $0v_2$ и с правой ориентацией. Отображенная на рис. 9.6, *б* область D^2 при $0 < \alpha < \pi/4$ является квадратом с центром в начале координат,

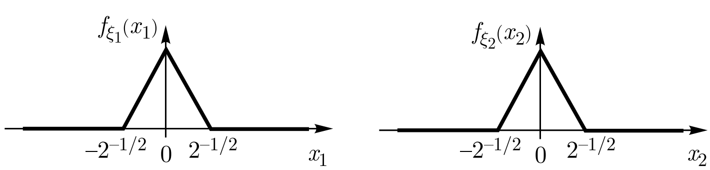


Рис. 9.7

одна из диагоналей которого составляет угол α с осью ординат $0v_2$ и имеет длину $2^{1/2}$. Если $\alpha = \pi/4$, то приведенная на рис. 9.6, *в* область $D^2 = \{(v_1, v_2)\colon -2^{-1} \leqslant v_1, v_2 \leqslant 2^{-1}\}$ является квадратом с центром в начале координат, стороны которого параллельны осям $0v_1$ и $0v_2$ и имеют единичную длину.

Непосредственно из (9.13) находим, что при $0 < \alpha \leqslant \pi/4$ формулы

$$\begin{aligned} \xi_1 &= t_1(\eta_1, \eta_2) = \eta_1 \cos\alpha - \eta_2 \sin\alpha, \\ \xi_2 &= t_2(\eta_1, \eta_2) = \eta_1 \sin\alpha + \eta_2 \cos\alpha \end{aligned} \tag{9.14}$$

дают взаимно однозначное обратное преобразование двумерной случайной величины $\boldsymbol{\eta} = (\eta_1, \eta_2)$ в случайный вектор $\boldsymbol{\xi} = (\xi_1, \xi_2)$. Функциональная зависимость (9.14) позволяет вычислить якобиан

$$J(v_1, v_2) = \begin{vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{vmatrix} = 1 \neq 0.$$

Поэтому случайный вектор $\boldsymbol{\eta} = (\eta_1, \eta_2)$ будет непрерывным, а его плотность распределения $f_\eta(v_1, v_2)$ равна $f_\xi(t_1(v_1, v_2), t_2(v_1, v_2))\,|J(v_1, v_2)| = 1$ в области D^2 и равна нулю при $(v_1, v_2) \notin D^2$. Следовательно, двумерная случайная величина (η_1, η_2) равномерно распределена в области D^2 на плоскости $v_1 0 v_2$. Используя формулы (7.15) и уравнения

$$\begin{aligned} (\sin\alpha + \cos\alpha)\, v_1 + (\cos\alpha - \sin\alpha)\, v_2 &= 2^{-1/2}, \\ (\sin\alpha - \cos\alpha)\, v_1 + (\sin\alpha + \cos\alpha)\, v_2 &= 2^{-1/2}, \\ -v_1 \sin\alpha - v_1 \cos\alpha + (\sin\alpha - \cos\alpha)\, v_2 &= 2^{-1/2}, \\ (\cos\alpha - \sin\alpha)\, v_1 - (\sin\alpha + \cos\alpha)\, v_2 &= 2^{-1/2} \end{aligned}$$

для границы области D^2, отображенной на рис. 9.6, *б*, найдем плотности $f_{\eta_1}(v_1)$ и $f_{\eta_2}(v_2)$ отдельных случайных величин η_1 и η_2:

$$f_{\eta_1}(v_1) = \int\limits_{-\infty}^{\infty} f_{\boldsymbol{\eta}}(v_1, v_2)\, dv_2 =$$

$$= \begin{cases} (\cos^2\alpha - \sin^2\alpha)^{-1}(2v_1 + \sqrt{2}\,\cos\alpha) \\ \qquad\qquad \text{при} \ \ -(\sqrt{2}\,)^{-1}\cos\alpha \leqslant v_1 \leqslant -(\sqrt{2}\,)^{-1}\sin\alpha; \\ \sqrt{2}\,(\sin\alpha + \cos\alpha)^{-1} \\ \qquad\qquad \text{при} \ \ -(\sqrt{2}\,)^{-1}\sin\alpha \leqslant v_1 \leqslant (\sqrt{2}\,)^{-1}\sin\alpha; \\ (\cos^2\alpha - \sin^2\alpha)^{-1}(-2v_1 + \sqrt{2}\,\cos\alpha) \\ \qquad\qquad \text{при} \ \ (\sqrt{2}\,)^{-1}\sin\alpha \leqslant v_1 \leqslant (\sqrt{2}\,)^{-1}\cos\alpha, \end{cases}$$

$$f_{\eta_2}(v_2) = \int\limits_{-\infty}^{\infty} f_{\boldsymbol{\eta}}(v_1, v_2)\, dv_1 =$$

$$= \begin{cases} (\cos^2\alpha - \sin^2\alpha)^{-1}(2v_2 + \sqrt{2}\cos\alpha) \\ \qquad\qquad \text{при} \ -(\sqrt{2}\,)^{-1}\cos\alpha \leqslant v_2 \leqslant -(\sqrt{2}\,)^{-1}\sin\alpha; \\ \sqrt{2}\,(\sin\alpha + \cos\alpha)^{-1} \\ \qquad\qquad \text{при} \ -(\sqrt{2}\,)^{-1}\sin\alpha \leqslant v_2 \leqslant (\sqrt{2}\,)^{-1}\sin\alpha; \\ (\cos^2\alpha - \sin^2\alpha)^{-1}(-2v_2 + \sqrt{2}\cos\alpha) \\ \qquad\qquad \text{при} \ (\sqrt{2}\,)^{-1}\sin\alpha \leqslant v_2 \leqslant (\sqrt{2}\,)^{-1}\cos\alpha, \end{cases}$$

если $0 < \alpha < \pi/4$. Ясно, что $f_{\eta_1}(v_1) = 0$ и $f_{\eta_2}(v_2) = 0$ при $|v_1|, |v_2| \geqslant 2^{-1/2}\cos\alpha$. Из полученных формул для $f_\eta(v_1, v_2)$, $f_{\eta_1}(v_1)$ и $f_{\eta_2}(v_2)$ мы заключаем, что преобразование (9.13) при $0 < \alpha < \pi/4$ переводит равномерный случайный вектор $\boldsymbol{\xi} = (\xi_1, \xi_2)$ в равномерно распределенную двумерную случайную величину $\boldsymbol{\eta} = (\eta_1, \eta_2)$. На рис. 9.8 приведены графики функций $f_{\eta_1}(v_1)$ и $f_{\eta_2}(v_2)$.

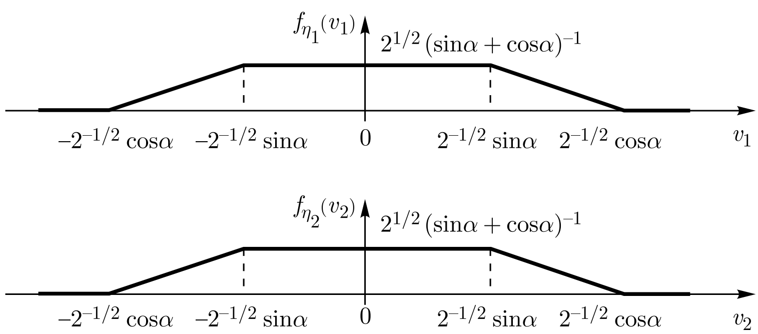


Рис. 9.8

Далее, каждая из одномерных случайных величин η_1 и η_2 не распределена равномерно. Наконец, случайные величины η_1 и η_2 являются статистически зависимыми, так как $f_\eta(v_1, v_2) \neq f_{\eta_1}(v_1) \times f_{\eta_2}(v_2)$ в области D^2, представленной на рис. 9.6, *б*).

Если $\alpha = \pi/4$, то, учитывая уравнения $v_1 = 1/2$, $v_2 = 1/2$, $v_1 = -1/2$, $v_2 = -1/2$ границы области D^2, приведенной на рис. 9.6, *в*), получим, что $f_{\eta_1}(v_1) = \int_{-\infty}^{+\infty} f_\eta(v_1, v_2)\, dv_2 = 1$ при $-0{,}5 \leqslant v_1 \leqslant 0{,}5$ и $f_{\eta_2}(v_2) = \int_{-\infty}^{+\infty} f_\eta(v_1, v_2)\, dv_1 = 1$ при $-0{,}5 \leqslant v_2 \leqslant 0{,}5$. Очевидно, что $f_{\eta_1}(v_1) = 0$ и $f_{\eta_2}(v_2) = 0$ при $|v_1|, |v_2| > 0{,}5$. Согласно этим соотношениям проверим, что $f_\eta(v_1, v_2) = f_{\eta_1}(v_1) \times f_{\eta_2}(v_2)$ при всех $-\infty < v_1, v_2 < +\infty$. Итак, с помощью простого отображения (9.13) при $\alpha = \pi/4$ равномерно распределенная двумерная случайная величина

$\boldsymbol{\xi} = (\xi_1, \xi_2)$ с зависимыми компонентами преобразована в систему равномерно распределенных независимых случайных величин η_1 и η_2.

Во многих практических задачах теории вероятностей необходимо получать значения n-мерной случайной величины с заданной интегральной функцией распределения, если предварительно имеется возможность вычислять значения вектора с n независимыми и равномерно распределенными на промежутке $[0,1]$ компонентами. Ради простоты пусть $n = 2$, а случайный вектор $\boldsymbol{\xi} = (\xi_1, \xi_2)$ имеет интегральную функцию распределения $F_{\xi}(x_1, x_2)$, для которой одномерная интегральная функция распределения $F_{\xi_1}(x_1)$ случайной величины ξ_1 и условная интегральная функция распределения $F_{\xi_2}(x_2 \mid x_1)$ случайной величины ξ_2 являются непрерывными и строго возрастающими. Введем теперь одномерные случайные величины η_1 и η_2 с помощью следующих функциональных зависимостей:

$$\eta_1 = g_1(\xi_1) = F_{\xi_1}(\xi_1), \quad \eta_2 = g_2(\xi_1, \xi_2) = F_{\xi_2}(\xi_2 \mid \xi_1). \tag{9.15}$$

Принимая во внимание решение задачи имитации и свойства функции распределения $F_{\xi_2}(x_2 \mid x_1)$, выводим, что независимо от значения x_1 случайной величины ξ_1 условное распределение $F_{\eta_2}(y_2 \mid x_1)$ случайной величины η_2 совпадает с равномерным на отрезке $[0,1]$. Отсюда получаем, что случайные величины ξ_1 и η_2 статистически независимы и случайная величина η_2 распределена равномерно на промежутке $[0,1]$. Очевидно, что случайные величины $\eta_1 = g_1(\xi_1)$, $\eta_2 = \eta_2$ независимы как функции независимых случайных величин ξ_1, η_2. Более того, случайная величина $\eta_1 = F_{\xi_1}(\xi_1)$ распределена равномерно на промежутке $[0,1]$, так как $F_{\xi_1}(x_1)$ является непрерывной и строго возрастающей интегральной функцией распределения (см. решение задачи имитации — пример 9.2). Разрешая формулы (9.15) относительно ξ_1 и ξ_2, найдем взаимно однозначное обратное преобразование: $\xi_1 = t_1(\eta_1, \eta_2) = F_{\xi_1}^{-1}(\eta_1)$, $\xi_2 = t_2(\eta_1, \eta_2) = F_{\xi_2}^{-1}(\eta_2 \mid F_{\xi_1}^{-1}(\eta_1))$, которое вычисляет компоненты двумерной случайной величины $\boldsymbol{\xi}$ через независимые равномерно распределенные на отрезке $[0,1]$ случайные величины η_1 и η_2. Следовательно, здесь решена задача имитации случайного вектора $\boldsymbol{\xi} = (\xi_1, \xi_2)$.

Аналогичным способом можно решить задачу имитации случайного вектора $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$, если заданы интегральная функция распределения $F_{\xi_1}(x_1)$ случайной величины ξ_1 и условные функции распределения $F_{\xi_2}(x_2 \mid x_1)$, $F_{\xi_3}(x_3 \mid x_1, x_2), \ldots$, $F_{\xi_n}(x_n \mid x_1, x_2, \ldots, x_{n-1})$ случайных величин ξ_2, $\xi_3, \ldots$, ξ_n. Предполагается, что указанные функции распределения являются непрерывными и строго возрастающими. Рассмотрим теперь пример решения задачи имитации для $n = 2$.

Пример 9.7. Двумерная случайная величина $\boldsymbol{\xi} = (\xi_1, \xi_2)$ равномерно распределена в квадрате C^2 (см. рис. 9.9, *а*).

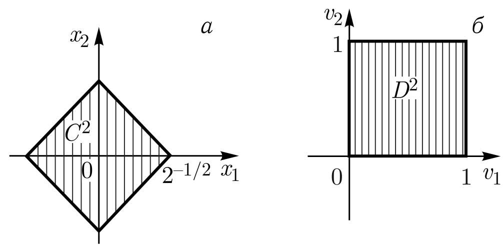


Рис. 9.9

Найти такое преобразование случайного вектора $\boldsymbol{\eta} = (\eta_1, \eta_2)$ с независимыми и равномерно распределенными на отрезке $[0, 1]$ компонентами (см. рис. 9.9, *б*), в результате которого получается вектор $\boldsymbol{\xi} = (\xi_1, \xi_2)$.

Решение. Используя формулу для плотности $f_{\xi_1}(x_1)$ и ее график на рис. 9.7, которые были получены при решении примера 9.6, равенство для функции $F_{\xi_2}(x_2 \mid x_1) = \left(\int_{-\infty}^{x_2} f(x_1, v)\, dv\right) / f_{\xi_1}(x_1)$ при $|x_1| < \sqrt{0{,}5}$, легко найдем

$$v_1 = g_1(x_1) = F_{\xi_1}(x_1) = \int\limits_{-\infty}^{x_1} f_{\xi_1}(u)\, du =$$

$$= \begin{cases} 0 & \text{при } x_1 < -\sqrt{0{,}5}\,; \\ (x_1 + \sqrt{0{,}5}\,)^2 & \text{при } -\sqrt{0{,}5} \leqslant x_1 < 0; \\ 1 - (x_1 - \sqrt{0{,}5}\,)^2 & \text{при } 0 \leqslant x_1 < \sqrt{0{,}5}\,; \\ 1 & \text{при } x_1 \geqslant \sqrt{0{,}5}\,; \end{cases}$$

$$v_2 = g_2(x_1, x_2) = F_{\xi_2}(x_2 \mid x_1) =$$

$$= \begin{cases} 0 \text{ при } -\sqrt{0{,}5} < x_1 < 0,\ x_2 \leqslant -x_1 - \sqrt{0{,}5}\,, \\ \qquad \text{или } 0 \leqslant x_1 < \sqrt{0{,}5}\,,\ x_2 \leqslant x_1 - \sqrt{0{,}5}\,; \\ (x_2 + x_1 + \sqrt{0{,}5}\,)(\sqrt{2} + 2x_1)^{-1} \text{ при } -\sqrt{0{,}5} < x_1 < 0, \\ \qquad -x_1 - \sqrt{0{,}5} \leqslant x_2 \leqslant x_1 + \sqrt{0{,}5}\,; \\ (x_2 - x_1 + \sqrt{0{,}5}\,)(\sqrt{2} - 2x_1)^{-1} \text{ при } 0 \leqslant x_1 < \sqrt{0{,}5}\,, \\ \qquad x_1 - \sqrt{0{,}5} \leqslant x_2 \leqslant -x_1 + \sqrt{0{,}5}\,; \\ 1 \text{ при } -\sqrt{0{,}5} < x_1 < 0,\ x_2 \geqslant x_1 + \sqrt{0{,}5}\,, \\ \qquad \text{или } 0 \leqslant x_1 < \sqrt{0{,}5}\,,\ x_2 \geqslant -x_1 + \sqrt{0{,}5}\,. \end{cases} \tag{9.16}$$

К формулам (9.16) следует добавить, что при $|x_1| \geqslant \sqrt{0{,}5}$ функция $f_{\xi_1}(x_1) = 0$, а функция $g_2(x_1, x_2) = F_{\xi_2}(x_2 \mid x_1)$ определяется совершенно произвольным образом. Например, можно положить: $0 \leqslant v_2 = g_2(x_1, x_2) = F_{\xi_2}(x_2) \leqslant 1$ для всех $|x_1| \geqslant \sqrt{0{,}5}$ и $-\infty < x_2 < +\infty$.

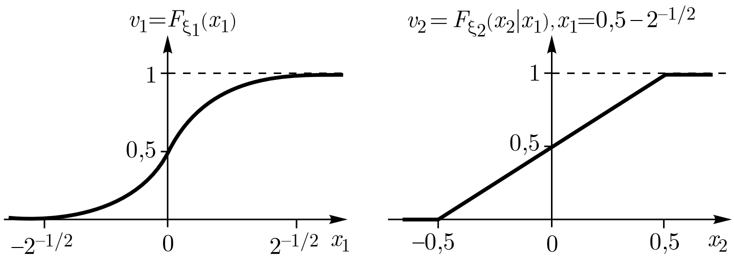


Рис. 9.10

На рис. 9.10 показаны графики интегральной функции распределения $F_{\xi_1}(x_1)$ и условной функции распределения $F_{\xi_2}(x_2 \mid x_1)$ при заданном значении величины $x_1 = 0{,}5 - \sqrt{0{,}5}$.

Отметим, что условная интегральная функция распределения $F_{\xi_2}(x_2 \mid x_1)$ при заданном значении x_1 является функцией равномерного распределения на промежутке $[-x_1 - \sqrt{0{,}5}\,, x_1 + \sqrt{0{,}5}\,]$ для $-\sqrt{0{,}5} \leqslant x_1 < 0$ и, естественно, на промежутке $[x_1 - \sqrt{0{,}5}\,, -x_1 + \sqrt{0{,}5}\,]$ для $0 \leqslant x_1 < \sqrt{0{,}5}$.

Формулы (9.16) осуществляют преобразование всей плоскости $x_1 0 x_2$ в заштрихованный квадрат D^2 на плоскости $v_1 0 v_2$ (см. рис. 9.9). Однако функции $v_1 = F_{\xi_1}(x_1)$ и $v_2 = F_{\xi_2}(x_2 \mid x_1)$ сохраняют постоянные значения на некоторых промежутках. Поэтому нарушается взаимно однозначное соответствие между переменными x_1, x_2 и v_1, v_2. Так, например, при $v_1 = 1/2$, $v_2 = 0$ из (9.16) находим, что $x_1 = 0$, $x_2 \leqslant -\sqrt{0{,}5}$. Далее можно условиться, что на рис. 9.9, а) точка $(x_1, x_2) \in C^2$. Такое ограничение не является существенным, так как из условия задачи событие $\{\omega\colon (\xi_1, \xi_2) \notin C^2\}$ является невозможным. Более того, чтобы получить все точки области D^2 плоскости $v_1 0 v_2$, достаточно ограничиться значениями $(x_1, x_2) \in C^2$. Теперь каждой точке (v_1, v_2) из области D^2 плоскости $v_1 0 v_2$ отвечает только одна точка (x_1, x_2) из множества C^2 плоскости $x_1 0 x_2$. При этом величины x_1 и x_2 непосредственно определяются из (9.16) с помощью формул:

$$x_1 = t_1(v_1) = \begin{cases} \sqrt{v_1} - \sqrt{0{,}5} & \text{при } 0 \leqslant v_1 \leqslant 0{,}5; \\ -\sqrt{1 - v_1} + \sqrt{0{,}5} & \text{при } 0{,}5 \leqslant v_1 \leqslant 1, \end{cases}$$

$$x_2 = t_2(v_1, v_2) = \begin{cases} \sqrt{v_1}\,(2v_2 - 1) & \text{при } 0 \leqslant v_1 \leqslant 0{,}5,\ 0 \leqslant v_2 \leqslant 1; \\ \sqrt{1 - v_1}\,(2v_2 - 1) & \text{при } 0{,}5 \leqslant v_1 \leqslant 1,\ 0 \leqslant v_2 \leqslant 1. \end{cases} \tag{9.17}$$

Ясно, что в данном примере для преобразования (9.16) области C^2 и для его обратного преобразования (9.17) области D^2 нарушаются условия непрерывности и взаимно однозначного соответствия в точках $(-\sqrt{0{,}5}\,, 0)$, $(\sqrt{0{,}5}\,, 0)$ границы области C^2 или в точках границы вида

$\{(v_1, v_2)\colon v_1 = 0,\ 0 \leqslant v_2 \leqslant 1\} \cup \{(v_1, v_2)\colon v_1 = 1,\ 0 \leqslant v_2 \leqslant 1\}$ области D^2. Например, неустранимое нарушение однозначности соответствия связано с граничной точкой $(-\sqrt{0{,}5}\,, 0)$ области C^2. В самом деле, для точки $(-\sqrt{0{,}5}\,, 0)$ области C^2 из соотношения (9.16) получаем точки множества $\{(v_1, v_2)\colon v_1 = 0,\ 0 \leqslant v_2 \leqslant 1\}$ на плоскости $v_1 0 v_2$. Якобиан $J(v_1, v_2)$ преобразования (9.17), если исключить $v_1 = 0$ и $v_1 = 1$, есть

$$\begin{vmatrix} \left(2\sqrt{1-v_1}\,\right)^{-1} & 0 \\ \left(-2\sqrt{1-v_1}\,\right)^{-1}(2v_2 - 1) & 2\sqrt{1-v_1} \end{vmatrix} = 1 \quad \text{при } 0 < v_1 \leqslant 0{,}5,\ 0 \leqslant v_2 \leqslant 1$$

и

$$\begin{vmatrix} \left(2\sqrt{v_1}\,\right)^{-1} & 0 \\ -\left(2\sqrt{1-v_1}\,\right)^{-1}(2v_2 - 1) & 2\sqrt{1-v_1} \end{vmatrix} = 1 \text{ при } 0{,}5 \leqslant v_1 < 1,\ 0 \leqslant v_2 \leqslant 1.$$

Используя метод, который был обоснован и апробирован при решении примера 9.6, в конечном счете, найдем простое искомое преобразование

$$\xi_1 = t_1(\eta_1) = \begin{cases} \sqrt{\eta_1} - \sqrt{0{,}5} & \text{при } \ 0 \leqslant \eta_1 \leqslant 0{,}5; \\ -\sqrt{1-\eta_1} + \sqrt{0{,}5} & \text{при } \ 0{,}5 \leqslant \eta_1 \leqslant 1, \end{cases}$$

$$\xi_2 = t_2(\eta_1, \eta_2) = \begin{cases} \sqrt{\eta_1}\,(2\eta_2 - 1) & \text{при } \ 0 \leqslant \eta_1 \leqslant 0{,}5,\ 0 \leqslant \eta_2 \leqslant 1; \\ \sqrt{1-\eta_1}\,(2\eta_2 - 1) & \text{при } \ 0{,}5 \leqslant \eta_1 \leqslant 1,\ 0 \leqslant \eta_2 \leqslant 1 \end{cases}$$

случайного вектора $\boldsymbol{\eta} = (\eta_1, \eta_2)$ с независимыми и равномерно распределенными на отрезке $[0, 1]$ компонентами в двумерную случайную величину $\boldsymbol{\xi} = (\xi_1, \xi_2)$. При этом плотность $f_\xi(x_1, x_2)$ вектора $\boldsymbol{\xi} = (\xi_1, \xi_2)$, очевидно, равна $f_\eta(g_1(x_1), g_2(x_1, x_2))\,|J(x_1, x_2)| = f_\eta(g_1(x_1), g_2(x_1, x_2))(J(v_1, v_2))^{-1} = 1$ в области $C^2 \backslash \{(x_1, 0)\colon x_1 = \pm\sqrt{0{,}5}\,\}$ и равна нулю при $(x_1, x_2) \notin C^2$. Далее для определенности полагаем, что $f_\xi(-\sqrt{0{,}5}\,, 0) = f_\xi(\sqrt{0{,}5}\,, 0) = 1$, т. е. двумерная случайная величина $\boldsymbol{\xi} = (\xi_1, \xi_2)$ равномерно распределена в области C^2 на плоскости $x_1 0 x_2$.

§ 2. Элементы теории корреляций

2.1. Линейная статистическая зависимость двух случайных величин. Числовая характеристика $\mathsf{cov}(\xi, \eta)$ отвечает не только за смешанный разброс одномерных случайных величин ξ и η, но и в значительной степени решает вопрос об их статистической зависимости или независимости. Для подтверждения этого докажем следующее простое утверждение.

Лемма 9.1. *Если ξ и η являются независимыми случайными величинами, для которых существуют $\mathsf{M}\xi$ и $\mathsf{M}\eta$, то числовые характеристики $\mathsf{cov}(\xi,\eta)$ и $\mathsf{corr}(\xi,\eta)$ равны нулю.*

Доказательство. Из определения корреляционного момента и свойства математического ожидания для независимых случайных величин следует, что $\mathsf{cov}(\xi,\eta) = \mathsf{M}[(\xi - \mathsf{M}\xi) \times (\eta - \mathsf{M}\eta)] = \mathsf{M}(\xi - \mathsf{M}\xi) \times \times \mathsf{M}(\eta - \mathsf{M}\eta) = 0$. Поэтому $\mathsf{corr}(\xi,\eta) = \mathsf{cov}(\xi,\eta)/(\boldsymbol{\sigma}\xi \times \boldsymbol{\sigma}\eta) = 0$.

Итак, если случайные величины ξ и η независимы, то их ковариация или корреляционный момент равна нулю. Отсюда вытекает, что эти величины обязательно статистически зависимы при $\mathsf{cov}(\xi,\eta) \neq 0$.

Замечание 9.1. Можно сказать, что неравенство $\mathsf{cov}(\xi,\eta) \neq 0$ является достаточным условием зависимости случайных величин ξ и η. Значит, ненулевые значения числовых характеристик $\mathsf{cov}(\xi,\eta)$ и $\mathsf{corr}(\xi,\eta)$ характеризуют некоторую степень вероятностной зависимости между случайными величинами. К сожалению, некоррелированность (в смысле рассматриваемого здесь пирсоновского линейного коэффициента корреляции $\mathsf{corr}(\xi,\eta) = 0$) в общем случае не влечет независимость случайных величин ξ и η. Другими словами, для зависимых случайных величин в некоторых случаях может оказаться $\mathsf{cov}(\xi,\eta) = 0$, так как это обстоятельство зависит от соотношений (8.21)–(8.23), например, если величины ξ и η связаны нелинейной зависимостью. Для подтверждения этого факта достаточно рассмотреть следующий пример.

Пример 9.8. Закон распределения случайного вектора (ξ_1, ξ_2) непрерывного типа определяется двумерной плотностью вероятности $f(x,y)$. При этом функция $f(x,y) = (6\pi)^{-1}$, если $(x,y) \in D^2 = \{(x,y)\colon 9^{-1}x^2 + 4^{-1}y^2 \leqslant 1\}$, и $f(x,y) = 0$ в остальной части плоскости. Доказать, что ξ_1 и ξ_2 — некоррелированные и зависимые случайные величины.

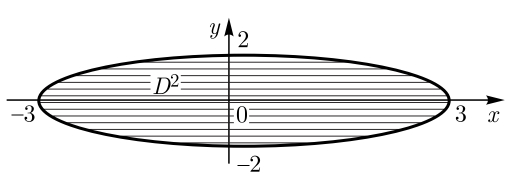


Рис. 9.11

Решение. На рис. 9.11 представлена область D^2. Проверим, что $f(x,y)$ действительно является плотностью некоторой двумерной случайной величины (ξ_1, ξ_2):

$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)\,dx\,dy = \iint\limits_{D^2} f(x,y)\,dx\,dy = \int\limits_{-3}^{+3}\left(\int\limits_{-2\sqrt{1-9^{-1}x^2}}^{2\sqrt{1-9^{-1}x^2}} (6\pi)^{-1}dy\right)dx =$$

$$= 2(9\pi)^{-1} \int\limits_{-3}^{+3} \sqrt{9 - x^2}\, dx =$$

$$= 2(9\pi)^{-1} \left[2^{-1} x\sqrt{9 - x^2} + 9\,(2^{-1}) \arcsin(x\, 3^{-1})\right]\Big|_{-3}^{+3} = 1.$$

Этот же результат можно получить другим способом. Интеграл вида $\iint_{D^2} dx\, dy$ равен площади области D^2. Так как граница области D^2 является эллипсом с полуосями, равными числу три и числу два, то площадь этой области равна $\pi \times 3 \times 2 = 6\,\pi$. Поэтому $\iint_{D^2} f(x, y)\, dx\, dy = 1$. Найдем плотности $f_{\xi_1}(x)$ и $f_{\xi_2}(y)$ для каждой из случайных величин ξ_1 и ξ_2:

$$f_{\xi_1}(x) = \int\limits_{-\infty}^{+\infty} f(x, y)\, dy = (6\pi)^{-1} \int\limits_{-2\sqrt{1-9^{-1}x^2}}^{2\sqrt{1-9^{-1}x^2}} dy = (9\pi)^{-1} 2\sqrt{9 - x^2}$$

при $|x| \leqslant 3$,

$$f_{\xi_2}(y) = \int\limits_{-\infty}^{+\infty} f(x, y)\, dx = (6\pi)^{-1} \int\limits_{-3\sqrt{1-4^{-1}y^2}}^{3\sqrt{1-4^{-1}y^2}} dx = (2\pi)^{-1} \sqrt{4 - y^2}$$

при $|y| \leqslant 2$.

Легко видеть, что $f_{\xi_1}(x) = 0$ при $|x| > 3$ и $f_{\xi_2}(y) = 0$ при $|y| > 2$.

Отсюда непосредственной проверкой убеждаемся, что $f_{\xi_1}(x)\, f_{\xi_2}(y) \neq f(x, y)$ при $(x, y) \in D^2$, т. е. случайные величины зависимы. Графики функций $f_{\xi_1}(x) = f_1(x)$ и $f_{\xi_2}(y) = f_2(y)$ изображены на рис. 9.12, a) и 9.12, b).

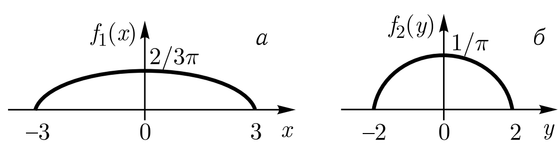


Рис. 9.12

Для вычисления корреляционного момента $\mathsf{cov}(\xi, \eta)$, прежде всего, найдем $\mathsf{M}\xi_1$ и $\mathsf{M}\xi_2$:

$$\mathsf{M}\xi_1 = \int\limits_{-\infty}^{+\infty} x f_{\xi_1}(x)\, dx = 2(9\pi)^{-1} \int\limits_{-3}^{+3} x\sqrt{9 - x^2}\, dx =$$

$$= -2(27\pi)^{-1} \sqrt{(9 - x^2)^3}\,\Big|_{-3}^{+3} = 0,$$

$$
\mathsf{M}\xi_2 = \int\limits_{-\infty}^{+\infty} y f_{\xi_2}(y)\,dy = (2\pi)^{-1}\int\limits_{-2}^{+2} y\sqrt{4-y^2}\,dy =
$$
$$
= -2(12\pi)^{-1}\sqrt{(4-y^2)^3}\,\Big|_{-2}^{+2} = 0.
$$

Затем, используя формулу (8.22), получим, что

$$
\mathsf{cov}(\xi_1,\xi_2) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x-\mathsf{M}\xi_1)\,(y-\mathsf{M}\xi_2)\,f(x,y)\,dx\,dy =
$$
$$
= (6\pi)^{-1}\iint\limits_{D^2} xy\,dx\,dy = (6\pi)^{-1}\int\limits_{-3}^{+3}\!\!\int x\left(\int\limits_{-2\sqrt{1-9^{-1}x^2}}^{2\sqrt{1-9^{-1}x^2}} y\,dy\right)dx = 0.
$$

Поэтому случайные величины ξ_1 и ξ_2 некоррелированы.

Заметим, что для некоррелированных случайных величин ξ_1 и ξ_2 из утверждений лемм 8.14 и 8.15 следуют очевидные равенства

$$
\mathsf{M}(\xi_1\times\xi_2) = \mathsf{M}\xi_1\times\mathsf{M}\xi_2 \quad \text{и} \quad \mathsf{D}(\xi_1+\xi_2) = \mathsf{D}\xi_1 + \mathsf{D}\xi_2,
$$

которые также справедливы при независимых случайных величинах. Так как из некоррелированности случайных величин не следует их независимость, то эти равенства справедливы при более слабых условиях.

Две случайные величины могут быть связаны либо функциональной зависимостью, либо зависимостью другого рода, называемой вероятностной (статистической), либо, наконец, случайные величины могут быть статистически независимыми. Строгая функциональная зависимость реализуется редко. В этом случае имеет место самая жесткая связь между ξ_1 и ξ_2, например, $\xi_2 = g(\xi_1)$, в частности, $\xi_2 = (\xi_1)^2$, $\xi_2 = a\xi_1 + b$ и т. д. *Вероятностной зависимостью* называют такой факт, когда изменение одной из случайных величин влечет за собой изменение законов распределения других случайных величин. В этом случае, например, для непрерывных двумерных случайных величин $\boldsymbol{\xi} = (\xi_1,\xi_2)$ выполняется неравенство вида $f_{\xi_1}(x)\,f_{\xi_2}(y) \neq f_{\boldsymbol{\xi}}(x,y)$. В результате решения примера 9.8 было установлено, что некоррелированность не влечет независимость случайных величин ξ_1 и ξ_2. Более того, можно привести примеры, для которых случайные величины связаны друг с другом функциональной, т. е. самой жесткой зависимостью, и, тем не менее, некоррелированы (см. также рисунки 9.19, *д* и 9.19, *е*).

Пример 9.9. Пусть плотность $f_{\xi_1}(x)$ случайной величины ξ_1 является симметричной функцией относительно оси ординат и моменты $\beta_2\xi_1$, $\beta_3\xi_1$ конечны. Определить корреляционный момент $\mathsf{cov}(\xi_1,\xi_2)$, если $\xi_2 = (\xi_1)^2$.

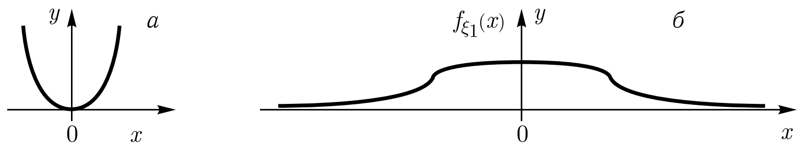


Рис. 9.13

Решение. На рис. 9.13 приведены сами наблюдения и качественный график плотности распределения вероятностей или дифференциальной функции распределения $f_{\xi_1}(x)$ для случайной величины ξ_1.

Тогда $\mathsf{M}\xi_1 = 0$ и $\xi_1 = \overset{\circ}{\xi}_1$, где $\overset{\circ}{\xi}_1$ — центрированная случайная величина (см. п. 2.1 гл. 8). Для реализации условий задачи пусть случайные величины ξ_1 и ξ_2 будут связаны функциональной зависимостью вида $\xi_2 = \xi_1^2$. Поэтому все возможные значения точки двумерной случайной величины (ξ_1, ξ_2) лежат на параболе $y = x^2$ (рис. 9.13, a). Здесь имеет место самая жесткая зависимость между величинами ξ_1 и ξ_2, так как значение величины ξ_1 полностью определяет значение $\xi_2 = \xi_1^2$. Найдем теперь корреляционный момент: $\mathsf{cov}(\xi_1, \xi_2) = \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)] =$ $= \mathsf{M}[\overset{\circ}{\xi}_1(\xi_2 - \mathsf{M}\xi_2)] = \mathsf{M}[\overset{\circ}{\xi}_1(\xi_1^2 - \mathsf{M}\xi_1^2)] = \mathsf{M}[\overset{\circ}{\xi}_1(\overset{\circ}{\xi}{}_1^2 - \mathsf{M}\overset{\circ}{\xi}{}_1^2)] = \mathsf{M}[\overset{\circ}{\xi}{}_1^3] - \mathsf{D}\xi_1\mathsf{M}\overset{\circ}{\xi}_1$. Отсюда $\mathsf{cov}(\xi_1, \xi_2) = 0$, так как в силу симметрии распределения случайной величины ξ_1 все ее центральные моменты нечетного порядка будут равны нулю. Итак, случайные величины ξ_1 и ξ_2 некоррелированы, хотя между ними имеется функциональная зависимость. Решения простых примеров 9.8 и 9.9 показывают, что условие $\mathsf{cov}(\xi_1, \xi_2) = 0$ не является достаточным критерием независимости случайных величин.

Выводы из решений примеров 9.8 и 9.9 на содержательном уровне можно пояснить следующими двумя обстоятельствами. Во-первых, введенный корреляционный момент характеризует степень взаимного рассеивания случайных величин: $\mathsf{cov}(\xi_1, \xi_2) = \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)]$. Поэтому рассеивание одной случайной величины может компенсировать разброс другой случайной величины так, что значение корреляционного момента будет равно нулю. Во-вторых, ковариация характеризует лишь один частный случай вероятностной зависимости между случайными величинами— так называемую линейную статистическую (корреляционную) зависимость. В частности, в рассмотренном выше примере 9.9 отсутствует линейная зависимость между случайными величинами ξ_1 и ξ_2. Более того, рассмотрим теперь линейную функциональную зависимость между случайными величинами, при которой корреляционный момент всегда будет отличен от нуля. В связи с этим покажем следующее утверждение.

Лемма 9.2. *Пусть случайные величины ξ_1 и ξ_2 связаны функциональной линейной зависимостью вида $\xi_2 = c\xi_1 + b$, причем $c \neq 0$. Тогда корреляционный момент $\mathsf{cov}(\xi_1, \xi_2) = c\,(\boldsymbol{\sigma}\xi_1)^2 \neq 0$.*

Доказательство. Из линейной зависимости случайных величин ξ_1 и ξ_2 легко получить, что $\mathsf{M}\xi_2 = c\,\mathsf{M}\xi_1 + b$, $\mathsf{D}\xi_2 = c^2\mathsf{D}\xi_1$, $\sigma\xi_2 = |c|\,\sigma\xi_1$. Вычислим теперь корреляционный момент: $\mathsf{cov}(\xi_1, \xi_2) = \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)] = \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(c\,\xi_1 + b - c\,\mathsf{M}\xi_1 - b)] = c\,\mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)^2] = c\,\mathsf{D}\xi_1 = c\,(\boldsymbol{\sigma}\xi_1)^2 \neq 0$. Отсюда выводим, что коэффициент корреляции $\mathsf{corr}(\xi_1, \xi_2) = c\,(\boldsymbol{\sigma}\xi_1)^2 \times (|c|\,\boldsymbol{\sigma}\xi_1\boldsymbol{\sigma}\xi_1)^{-1} = |c|\,c^{-1} = \pm 1$. Имеет место и обратное утверждение.

Замечение 9.2. Независимость случайных величин ξ и η из равенства нулю пирсоновского коэффициента линейной корреляции следует тогда и только тогда, когда случайные величины ξ и η распределены по нормальному закону.

Лемма 9.3. *Если коэффициент корреляции случайных величин* ξ_1 *и* ξ_2 *по абсолютной величине равен единице, т. е.* $|\mathsf{corr}(\xi_1, \xi_2)| = 1$, *то между случайными величинами имеется функциональная линейная зависимость.*

Доказательство. Для случайных величин ξ_1 и ξ_2 проведем вычисления:

$$\begin{aligned}
&\mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\boldsymbol{\sigma}\xi_1)^{-1} + (\xi_2 - \mathsf{M}\xi_2)(\boldsymbol{\sigma}(\xi_2))^{-1}]^2 = \\
&\qquad = \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)/\boldsymbol{\sigma}\xi_1]^2 + \mathsf{M}[(\xi_2 - \mathsf{M}\xi_2)/\boldsymbol{\sigma}\xi_2]^2 + \\
&\qquad\qquad + 2\,(\boldsymbol{\sigma}\xi_1\boldsymbol{\sigma}\xi_2)^{-1}\mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)] = \\
&= \mathsf{D}\xi_1(\boldsymbol{\sigma}\xi_1)^{-2} + \mathsf{D}\xi_2(\boldsymbol{\sigma}\xi_2)^{-2} + 2\,\mathsf{cov}(\xi_1, \xi_2)/(\boldsymbol{\sigma}\xi_1\boldsymbol{\sigma}\xi_2) = 2 + 2\mathsf{corr}(\xi_1, \xi_2).
\end{aligned}$$

Аналогичным способом получим, что

$$\mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\boldsymbol{\sigma}\xi_1)^{-1} - (\xi_2 - \mathsf{M}\xi_2)(\boldsymbol{\sigma}\xi_2)^{-1}]^2 = 2 - 2\,\mathsf{corr}(\xi_1, \xi_2).$$

Если предположить, что $\mathsf{corr}(\xi_1, \xi_2) = 1$, тогда находим

$$\mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\boldsymbol{\sigma}\xi_1)^{-1} - (\xi_2 - \mathsf{M}\xi_2)(\boldsymbol{\sigma}\xi_2)^{-1}]^2 = 2 - 2\,\mathsf{corr}(\xi_1, \xi_2) = 0.$$

Отсюда

$$(\xi_1 - \mathsf{M}\xi_1)(\boldsymbol{\sigma}\xi_1)^{-1} - (\xi_2 - \mathsf{M}\xi_2)(\boldsymbol{\sigma}\xi_1)^{-1} = 0,$$

$$\xi_2 = \mathsf{M}\xi_2 + (\xi_1 - \mathsf{M}\xi_1)\,(\boldsymbol{\sigma}\xi_2)(\boldsymbol{\sigma}\xi_1)^{-1}.$$

Если теперь $\mathsf{corr}(\xi_1, \xi_2) = -1$, то

$$\mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\boldsymbol{\sigma}\xi_1)^{-1} + (\xi_2 - \mathsf{M}\xi_2)(\boldsymbol{\sigma}\xi_2)^{-1}]^2 = 2 + 2\,\mathsf{corr}(\xi_1, \xi_2) = 0$$

и

$$(\xi_1 - \mathsf{M}\xi_1)(\boldsymbol{\sigma}\xi_1)^{-1} + (\xi_2 - \mathsf{M}\xi_2)(\boldsymbol{\sigma}\xi_1)^{-1} = 0.$$

Отсюда имеем линейную зависимость:

$$\xi_2 = \mathsf{M}\xi_2 - (\xi_1 - \mathsf{M}\xi_1)\,(\boldsymbol{\sigma}\xi_2)(\boldsymbol{\sigma}\xi_1)^{-1}.$$

Лемма 9.3 доказана.

Приведем теперь поучительный пример.

Пример 9.10. Наудачу подбрасываются один раз две монеты разной симметрии. Герб появляется с вероятностью p_1 на первой монете и с вероятностью p_2 на второй. Пусть ξ_1 определяет число выпавших гербов на первой монете и ξ_2 — на второй. Суммарное число выпавших на обеих монетах гербов обозначим через η_1, а решеток — через η_2. Найти коэффициент корреляции случайных величин η_1 и η_2.

Решение. Из определения случайных величин η_1 и η_2 последовательно имеем: $\eta_1 = \xi_1 + \xi_2$, $\eta_2 = (1 - \xi_1) + (1 - \xi_2)$, $\eta_1 + \eta_2 = 2$. Следовательно, между случайными величинами η_1 и η_2 существует линейная зависимость вида $\eta_2 = 2 - \eta_1$. Используя результаты доказательства леммы 9.2, получим, что $\mathsf{corr}(\eta_1, \eta_2) = -1$. Интересно на этом простом примере еще раз проследить факт существования случайных величин, которые одновременно являются некоррелированными и статистически зависимыми. Действительно, пусть случайная величина $\eta_3 = \xi_2 - \xi_1$ определяет разность чисел выпавших гербов на второй и первой монетах. Тогда ковариация для случайных величин η_3 и η_1 вычисляется следующим образом:

$$\begin{aligned}\mathsf{cov}(\eta_3, \eta_1) &= \mathsf{M}[(\eta_3 - \mathsf{M}\eta_3)\,(\eta_1 - \mathsf{M}\eta_1)] = \\ &= \mathsf{M}[((\xi_2 - \mathsf{M}\xi_2) - (\xi_1 - \mathsf{M}\xi_1)) \times ((\xi_2 - \mathsf{M}\xi_2) + (\xi_1 - \mathsf{M}\xi_1))] = \\ &= \mathsf{M}[(\xi_2 - \mathsf{M}\xi_2)^2 - (\xi_1 - \mathsf{M}\xi_1)^2] = \mathsf{D}\xi_2 - \mathsf{D}\xi_1 = p_2(1 - p_2) - p_1(1 - p_1).\end{aligned}$$

Если дополнительно $p_2 = 1 - p_1$ или $p_2 = p_1$, то корреляционный момент $\mathsf{cov}(\eta_3, \eta_1) = 0$. Легко вычислим вероятность $\mathsf{P}(\{\omega\colon \eta_1(\omega) = 2\}) = \mathsf{P}(\{\omega\colon \eta_1(\omega) = 2,\ \eta_3(\omega) = 0\}) = p_1p_2$ и $\mathsf{P}(\{\omega\colon \eta_3(\omega) = 0\}) = p_1p_2 + (1 - p_1) \times (1 - p_2)$. Отсюда находим, что при $0 < p_1, p_2 < 1$ произведение вида $\mathsf{P}(\{\omega\colon \eta_1(\omega) = 2\}) \times \mathsf{P}(\{\omega\colon \eta_3(\omega) = 0\})$ не будет равно $\mathsf{P}(\{\omega\colon \eta_1(\omega) = 2,\ \eta_3(\omega) = 0\})$. Значит, при любых $0 < p_1, p_2 < 1$ случайные величины η_3, η_1 будут зависимыми. В то же время при $p_2 = 1 - p_1$ или при $p_2 = p_1$ ковариация вида $\mathsf{cov}(\eta_3, \eta_1) = 0$, например, для симметричных монет, и, следовательно, случайные величины η_3, η_1 будут некоррелированными.

Лемма 9.4. *Пусть ξ_1 и ξ_2 — случайные величины и* $\mathsf{corr}(\xi_1, \xi_2)$ *есть коэффициент корреляции. Тогда имеет место неравенство* $|\mathsf{corr}(\xi_1, \xi_2)| \leqslant 1$.

Доказательство. Рассмотрим новые случайные величины вида: $\xi_3 = (\boldsymbol{\sigma}\xi_2)\,\xi_1 + (\boldsymbol{\sigma}\xi_1)\,\xi_2$, $\xi_4 = (\boldsymbol{\sigma}\xi_2)\,\xi_1 - (\boldsymbol{\sigma}\xi_1)\,\xi_2$. Сосчитаем дисперсии $\mathsf{D}\xi_3$ и $\mathsf{D}\xi_3$: $\mathsf{D}\xi_3 = (\boldsymbol{\sigma}\xi_2)^2\mathsf{D}\xi_1 + (\boldsymbol{\sigma}\xi_1)^2\mathsf{D}\xi_2 + 2\,(\boldsymbol{\sigma}\xi_2)\,(\boldsymbol{\sigma}\xi_1)\,\mathsf{cov}(\xi_1,\xi_2)$, $\mathsf{D}\xi_4 = (\boldsymbol{\sigma}\xi_2)^2\mathsf{D}\xi_1 + (\boldsymbol{\sigma}\xi_1)^2\mathsf{D}\xi_2 - 2\,(\boldsymbol{\sigma}\xi_2)\,(\boldsymbol{\sigma}\xi_1)\,\mathsf{cov}(\xi_1,\xi_2)$. Учитывая неравенства $\mathsf{D}\xi_3 \geqslant 0$ и $\mathsf{D}\xi_4 \geqslant 0$, преобразуем эти выражения: $(\boldsymbol{\sigma}\xi_2)^2\,(\boldsymbol{\sigma}\xi_1)^2 + (\boldsymbol{\sigma}\xi_1)^2\,(\boldsymbol{\sigma}\xi_2)^2 + 2\,(\boldsymbol{\sigma}\xi_2)\,(\boldsymbol{\sigma}\xi_1)\,\mathsf{cov}(\xi_1,\xi_2) \geqslant 0$, $(\boldsymbol{\sigma}\xi_2)^2\,(\boldsymbol{\sigma}\xi_1)^2 + (\boldsymbol{\sigma}\xi_1)^2\,(\boldsymbol{\sigma}\xi_2)^2 - 2\,(\boldsymbol{\sigma}\xi_2)\,(\boldsymbol{\sigma}\xi_1)\,\mathsf{cov}(\xi_1,\xi_2) \geqslant 0$, или $(\boldsymbol{\sigma}\xi_2)\,(\boldsymbol{\sigma}\xi_1) + \mathsf{cov}(\xi_1,\xi_2) \geqslant 0$, $(\boldsymbol{\sigma}\xi_2)\,(\boldsymbol{\sigma}\xi_1) - \mathsf{cov}(\xi_1,\xi_2) \geqslant 0$. Отсюда следуют неравенства $1 + \mathsf{corr}(\xi_1,\xi_2) \geqslant 0$, $1 - \mathsf{corr}(\xi_1,\xi_2) \geqslant 0$. Если $\mathsf{corr}(\xi_1,\xi_2) \geqslant 0$, то из $1 - \mathsf{corr}(\xi_1,\xi_2) \geqslant 0$ получаем, что $\mathsf{corr}(\xi_1,\xi_2) \leqslant 1$. Если $\mathsf{corr}(\xi_1,\xi_2) \leqslant 0$, то из соотношения $1 + \mathsf{corr}(\xi_1,\xi_2) \geqslant 0$ получаем, что $\mathsf{corr}(\xi_1,\xi_2) \geqslant -1$. Следовательно, $|\mathsf{corr}(\xi_1,\xi_2)| \leqslant 1$. Итак, лемма 9.4 доказана.

Поясним еще раз тезис о том, что введенный корреляционный момент и, значит, коэффициент корреляции характеризуют особый случай вероятностной зависимости между случайными величинами, а именно, случай *линейной статистической зависимости*. Для этого рассмотрим двумерную случайную величину (ξ_1,ξ_2), для которой $0 < \mathsf{D}\xi_1 < \infty$, $0 < \mathsf{D}\xi_2 < \infty$. Решим теперь задачу наилучшего линейного приближения компоненты ξ_2 с помощью случайной величины ξ_1. Эта задача возникает, когда при проведении эксперимента E на основании лишь измерения значения случайной величины ξ_1 с использованием простейших вычислений требуется наилучшим образом предсказать значение величины ξ_2. В этом случае значение случайной величины ξ_2 заменяется на значение случайной величины $b\,\xi_1 + c$. При этом предполагается, что точность такого приближения измеряется величиной $\mathsf{M}(\xi_2 - (b\,\xi_1 + c))^2$, т. е. чем меньше число $\mathsf{M}(\xi_2 - (b\,\xi_1 + c))^2$, тем точнее выполняется приближенное соотношение $\xi_2 \approx b\,\xi_1 + c$. Для точности приближения $\mathsf{M}(\xi_2 - (b\,\xi_1 + c))^2$ получим следующие равенства:

$$\begin{aligned}
&\mathsf{M}(\xi_2 - (b\,\xi_1 + c))^2 = \mathsf{M}(\xi_2 - b\,\xi_1 - c + b\,\mathsf{M}\xi_1 - b\,\mathsf{M}\xi_1 + \mathsf{M}\xi_2 - \mathsf{M}\xi_2)^2 = \\
&\quad = \mathsf{M}((\xi_2 - \mathsf{M}\xi_2) - b\,(\xi_1 - \mathsf{M}\xi_1) + (\mathsf{M}\xi_2 - b\,\mathsf{M}\xi_1 - c))^2 = \\
&\quad = \mathsf{M}(\xi_2 - \mathsf{M}\xi_2)^2 + b^2\mathsf{M}(\xi_1 - \mathsf{M}\xi_1)^2 + (\mathsf{M}\xi_2 - b\,\mathsf{M}\xi_1 - c)^2 - \\
&\qquad - 2b\,\mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)\,(\xi_2 - \mathsf{M}\xi_2)] + 2(\mathsf{M}\xi_2 - b\,\mathsf{M}\xi_1 - c)\mathsf{M}(\xi_2 - \mathsf{M}\xi_2) - \\
&\qquad - 2b\,(\mathsf{M}\xi_2 - b\,\mathsf{M}\xi_1 - c)\mathsf{M}(\xi_1 - \mathsf{M}\xi_1) = \\
&\quad = (\boldsymbol{\sigma}\xi_2)^2 + b^2(\boldsymbol{\sigma}\xi_1)^2 + (\mathsf{M}\xi_2 - b\,\mathsf{M}\xi_1 - c)^2 - 2b\,\boldsymbol{\sigma}(\xi_1)\boldsymbol{\sigma}(\xi_2)\,\mathsf{corr}(\xi_1,\xi_2) + \\
&\qquad + (\boldsymbol{\sigma}\xi_2)^2\mathsf{corr}^2(\xi_1,\xi_2) - (\boldsymbol{\sigma}\xi_2)^2\mathsf{corr}^2(\xi_1,\xi_2) = \\
&\quad = (\boldsymbol{\sigma}\xi_2)^2(1 - \mathsf{corr}^2(\xi_1,\xi_2)) + (\boldsymbol{\sigma}(\xi_2)\,\mathsf{corr}(\xi_1,\xi_2) - b\,\boldsymbol{\sigma}\xi_1)^2 + \\
&\qquad\qquad + (\mathsf{M}\xi_2 - b\,\mathsf{M}\xi_1 - c)^2.
\end{aligned}$$

Поэтому точность приближения $\mathsf{M}(\xi_2 - (b\,\xi_1 + c))^2$ принимает наименьшее значение при $\boldsymbol{\sigma}\xi_2\,\mathsf{corr}(\xi_1,\xi_2) - b\,\boldsymbol{\sigma}\xi_1 = 0$ и $\mathsf{M}\xi_2 - b\,\mathsf{M}\xi_1 - c = 0$.

Отсюда получим, что наименьшее значение для точности приближения, равное $(1 - \mathsf{corr}^2(\xi_1, \xi_2))\,\mathsf{D}\xi_2$, достигается при $b = (\boldsymbol{\sigma}\xi_1)^{-1}\boldsymbol{\sigma}\xi_2\,\mathsf{corr}(\xi_1, \xi_2)$ и $c = \mathsf{M}\xi_2 - (\boldsymbol{\sigma}\xi_1)^{-1}(\boldsymbol{\sigma}\xi_2)\mathsf{corr}(\xi_1, \xi_2)\,\mathsf{M}\xi_1$. Итак, в классе линейных функций от ξ_1 наилучшее приближение случайной величины ξ_2 дается величиной $\xi_2' = (\mathsf{D}\xi_1)^{-1}\mathsf{cov}(\xi_1, \xi_2)(\xi_1 - \mathsf{M}\xi_1) + \mathsf{M}\xi_2$.

Будем теперь наряду с двумерной случайной величиной (ξ_1, ξ_2) рассматривать ее приближение (ξ_1, ξ_2'). Так как имеем равенство $\mathsf{M}\xi_2 = \mathsf{M}\xi_2'$, то центр рассеивания двумерной случайной величины (ξ_1, ξ_2) и центр рассеивания двумерного случайного вектора (ξ_1, ξ_2') удачно совпали. За точность приближения двумерной случайной величины (ξ_1, ξ_2) случайным вектором (ξ_1, ξ_2') естественно принять число $(1 - \mathsf{corr}^2(\xi_1, \xi_2))\,\mathsf{D}\xi_2$. Отсюда, чем больше модуль коэффициента корреляции между случайными величинами ξ_1 и ξ_2, тем ближе случайные точки (ξ_1, ξ_2) располагаются к точкам (ξ_1, ξ_2') или к прямой $y = (\mathsf{D}\xi_1)^{-1}\mathsf{cov}(\xi_1, \xi_2)(x - \mathsf{M}\xi_1) + \mathsf{M}\xi_2$. В частном случае, если $\mathsf{corr}(\xi_1, \xi_2) = 1$, то $\xi_2' = \mathsf{M}\xi_2 + (\xi_1 - \mathsf{M}\xi_1)\,\boldsymbol{\sigma}\xi_2/\boldsymbol{\sigma}\xi_1$. С другой стороны, при доказательстве леммы 9.3 для $\mathsf{corr}(\xi_1, \xi_2) = 1$ было получено, что $\xi_2 = \mathsf{M}\xi_2 + (\xi_1 - \mathsf{M}\xi_1)\,\boldsymbol{\sigma}\xi_2/\boldsymbol{\sigma}\xi_1$, т. е. $(\xi_1, \xi_2) = (\xi_1, \xi_2')$. Аналогичная ситуация имеет место при $\mathsf{corr}(\xi_1, \xi_2) = -1$. Наконец, если коэффициента корреляции $\mathsf{corr}(\xi_1, \xi_2) = \mathsf{cov}(\xi_1, \xi_2) = 0$, т. е. случайные величины не являются коррелированными, то независимо от значений величины ξ_1 наилучшим линейным приближением для величины ξ_2 будет постоянная $\xi_2' = (\mathsf{D}\xi_1)^{-1}\mathsf{cov}(\xi_1, \xi_2)(\xi_1 - \mathsf{M}\xi_1) + \mathsf{M}\xi_2 = \mathsf{M}\xi_2$. На содержательном уровне это означает, что случайные величины ξ_1 и ξ_2 в линейном приближении являются статистически несвязанными. Подводя итоги этого абзаца, можем сказать, что ковариация является простой и удобной мерой линейной статистической (корреляционной) зависимости. Этим можно объяснить частое применение корреляционного момента в инженерной практике.

2.2. Условное математическое ожидание. Если числовая характеристика $|\mathsf{corr}(\xi_1, \xi_2)|$ равна единице, т. е. принимает наибольшее значение, то между случайными величинами ξ_1 и ξ_2 существует жесткая линейная зависимость. Пусть теперь $|\mathsf{corr}(\xi_1, \xi_2)| \neq 1$. Тогда из вида формул (8.21)–(8.23) для ковариации случайных величин дискретного, непрерывного и соответственно произвольного типа на содержательном уровне можно сделать такой вывод. Если корреляционный момент $\mathsf{cov}(\xi_1, \xi_2) > 0$, то с ростом одной случайной величины другая случайная величина или точно возрастает, или проявляет тенденцию к возрастанию. Если корреляционный момент $\mathsf{cov}(\xi_1, \xi_2) < 0$, то это означает, что с ростом одной случайной величины другая случайная величина или точно убывает, или проявляет тенденцию к убыванию. Поэтому для коррелированных случайных величин математическое ожидание одной случайной величины меняется с изменением значения другой случайной величины. Для пояснения этого факта рассмотрим иллюстративный пример.

Пример 9.11. Пусть ξ_2 — полученный урожай зерна в некоторых единицах с каждого из трех одинаковых по площади участков земли, ξ_1 — количество внесенных удобрений в некоторых единицах на каждый из участков. Пусть при $\xi_1 = 1$ случайная величина ξ_2 приняла следующие значения: 5 на первом участке, 6 на втором участке и 10 на третьем участке. При $\xi_1 = 2$ случайная величина ξ_2 приняла значения 4, 8 и 11 соответственно на первом, втором и третьем участках. Таким образом, с одинаковых по площади участков земли при равных количествах внесенных удобрений снимают различные урожаи, т. е. ξ_2 не является жестко связанной с ξ_1 или, другими словами, ξ_2 не является функцией от ξ_1. Это объясняется влиянием таких случайных факторов, как осадки, температура воздуха и др. Вместе с тем, как показывает опыт, средний урожай является функцией от количества внесенных удобрений, т. е. случайные величины ξ_1 и ξ_2 скорее всего корреляционно связаны.

Итак, целесообразно изучить характер возможных зависимостей математического ожидания одной случайной величины от значения другой, с ней коррелированной. С другой стороны, случайные величины могут быть некоррелированными и одновременно статистически зависимыми. В таких случаях линейная ковариация теряет смысл количественной меры статистической зависимости. Поэтому возникает задача о другом характере измерения степени статистической зависимости между произвольными случайными величинами. Такой мерой статистической зависимости между случайными величинами являются их условные математические ожидания. Переходим теперь к определению условного математического ожидания случайной величины на основе хорошо известного понятия условного эксперимента и условного вероятностного пространства.

Пусть $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ есть априорное вероятностное пространство для эксперимента E, на котором определена случайная величина $\xi(\omega)$. Выберем случайное событие $B \in \mathcal{F}$, для которого вероятность $\mathsf{P}(B) > 0$. Рассмотрим теперь условное вероятностное пространство $(\Omega, \mathcal{F}, \mathsf{P}(\cdot \mid B))$. Здесь $\mathsf{P}(A \mid B) = \mathsf{P}(AB)/\mathsf{P}(B)$ есть условная вероятность события $A \in \mathcal{F}$. Условное вероятностное пространство обладает всеми свойствами априорного вероятностного пространства. Значит, функция $\xi(\omega)$ будет случайной величиной на $(\Omega, \mathcal{F}, \mathsf{P}(\cdot \mid B))$ с условной функцией распределения $F(x|B) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\} \mid B) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap B)/\mathsf{P}(B)$. Поэтому в пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot \mid B))$ можно применить все те определения и доказать свойства различных числовых характеристик случайной величины $\xi(\omega)$, которые были рассмотрены относительно пространства $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. Например, мы можем ввести понятие условного математического ожидания и обозначить его через символ $\mathsf{M}(\xi \mid B)$. Имеет место утверждение.

Лемма 9.5. *Если* $\mathsf{M}|\xi| = \int_{-\infty}^{+\infty} |x|\,dF(x) < +\infty$ и $\mathsf{P}(B) > 0$, *то условное математическое ожидание случайной величины* $\xi(\omega)$ *существует и* $\mathsf{M}(\xi \mid B) = \int_{-\infty}^{+\infty} x\,dF(x \mid B)$.

Доказательство. Так как функции $F(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap B) + \mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap \overline{B})$, $\mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap B)$, $\mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap \overline{B})$ являются неубывающими по переменной x, то, используя свойства интеграла Римана–Стилтьеса, получаем, что

$$\int_{-\infty}^{+\infty} |x|\,dF(x) = \int_{-\infty}^{+\infty} |x|\,d\big[\mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap B) + \mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap \overline{B})\big] =$$

$$= \int_{-\infty}^{+\infty} |x|\,d\mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap B) + \int_{-\infty}^{+\infty} |x|\,d\mathsf{P}(\{\omega\colon \xi(\omega) < x\} \cap \overline{B}) < +\infty.$$

Отсюда $\int_{-\infty}^{+\infty} |x|\,dF(x \mid B) = (\mathsf{P}(B))^{-1} \int_{-\infty}^{+\infty} |x|\,d\mathsf{P}(\{\omega\colon \xi_1(\omega) < x\} \cap B) < +\infty$. Значит, мы можем найти условное математическое ожидание $\mathsf{M}(\xi \mid B)$ величины ξ относительно события B по формуле $\mathsf{M}(\xi \mid B) = \int_{-\infty}^{+\infty} x\,dF(x \mid B)$, если $\mathsf{P}(B)$ — произвольная случайная величина. При этом свойства интеграла Римана–Стилтьеса позволяют вычислить условное математическое ожидание $\mathsf{M}(\xi \mid B)$ по более простым формулам. А именно, по формуле $\mathsf{M}(\xi \mid B) = \int_{-\infty}^{+\infty} x\,f(x \mid B)\,dx$, если ξ есть непрерывная случайная величина, и по формуле $\mathsf{M}(\xi \mid B) = \sum_{i=1}^{\infty} x_i'\,\mathsf{P}(\xi_1 = x_i' \mid B)$, если ξ есть дискретная случайная величина. Лемма 9.5 доказана.

Итак, при $\mathsf{M}|\xi| < +\infty$ и $\mathsf{P}(B) > 0$ условные законы распределения случайной величины ξ позволяют вычислить условное математическое ожидание $\mathsf{M}(\xi \mid B)$ относительно пространства $(\Omega, \mathcal{F}, \mathsf{P}(\cdot \mid B))$ и без всяких существенных изменений решить проблему свойств условных математических ожиданий. При этом свойства условных математических ожиданий случайных величин полностью повторяют свойства безусловных (априорных) математических ожиданий.

Вычислим связь между безусловным математическим ожиданием и условными математическими ожиданиями случайной величины ξ в простейшем случае. Пусть $\mathsf{M}|\xi| < +\infty$ и B_1, $B_2, \ldots$ есть последовательность, быть может, конечная, попарно несовместных событий, такая, что $\cup_k B_k = \Omega$ и $\mathsf{P}(B_k) > 0$ для всех $k = 1, 2, \ldots$. Так как события $\mathsf{P}(B_k)$, $k = 1, 2, \ldots$, образуют полную группу попарно несовместимых событий, то по формуле полной вероятности имеем: $F(x) = \sum_k \mathsf{P}(B_k)\,F(x \mid B_k)$. Вероятности $\mathsf{P}(B_k)$, $k = 1, 2, \ldots$, положительны, а каждая из функций $F(x \mid B_k)$ является неубывающей по переменной x. Отсюда, используя свойства интеграла Римана–Стилтьеса, найдем безусловное математическое ожидание:

$$\text{а) } \mathsf{M}\xi = \int\limits_{-\infty}^{+\infty} x\,dF(x) = \int\limits_{-\infty}^{+\infty} x\,d\Big(\sum_k \mathsf{P}(B_k)\,F(x \mid B_k)\Big) = \\ = \sum_k \mathsf{P}(B_k) \int\limits_{-\infty}^{+\infty} x\,dF(x \mid B_k) = \sum_k \mathsf{P}(B_k)\,\mathsf{M}(\xi \mid B)$$

для произвольной случайной величины;

$$\text{б) } \mathsf{M}\xi = \int\limits_{-\infty}^{+\infty} x\,f(x)\,dx = \int\limits_{-\infty}^{+\infty} x\,dF(x) = \int\limits_{-\infty}^{+\infty} x\,d\Big(\sum_k \mathsf{P}(B_k)\,F(x \mid B_k)\Big) = \\ = \sum_k \mathsf{P}(B_k) \int\limits_{-\infty}^{+\infty} x\,dF(x \mid B_k) = \sum_k \mathsf{P}(B_k) \int\limits_{-\infty}^{+\infty} x\,f(x \mid B_k)\,dx = \\ = \sum_k \mathsf{P}(B_k)\mathsf{M}(\xi \mid B)$$

для непрерывной случайной величины;

$$\text{в) наконец, } \mathsf{M}\xi = \sum_{i=1}^{\infty} x_i'\mathsf{P}(\xi = x_i') = \\ = \sum_{i=1}^{\infty} x_i' \sum_k \mathsf{P}(B_k)\,\mathsf{P}(\{\omega\colon \xi(\omega) = x_i'\} \mid B_k) = \\ = \sum_k \mathsf{P}(B_k) \sum_{i=1}^{\infty} x_i'\,\mathsf{P}(\{\xi = x_i'\} \mid B_k) = \sum_k \mathsf{P}(B_k)\,\mathsf{M}(\xi \mid B)$$

для дискретной случайной величины.

Итак, в каждом из этих трех случаев имеет место формула $\mathsf{M}\xi = \sum_k \mathsf{P}(B_k)\,\mathsf{M}(\xi \mid B)$. Рассмотрим пример на использование понятия условного математического ожидания при вычислении априорного математического ожидания.

Пример 9.12. Пусть n населенных пунктов расположены вдоль прямолинейной дороги. Расстояние в километрах между любыми соседними пунктами равно b. Эти пункты обслуживает автомашина медицинской помощи. Каждый очередной вызов больного с равной вероятностью возникает в любом из n пунктов. Этот вызов сообщается непосредственно на машину, которая ждет его в том пункте, где находился предыдущий больной. Чему равна средняя длина пробега машины за один рейс?

Первый способ решения. Используя числа от единицы до n, перенумеруем все населенные пункты так, что номера соседних пунктов отличаются на единицу. Для данного эксперимента обозначим через $A_{i,j}$ произвольный элементарный исход, который заключается в том, что предыдущий вызов был в пункте с номером i, а последующий будет в пункте с номером j. В качестве описания такого элементарного исхода выберем символ $\omega_{i,j}$. Значит, для этого опыта достоверный исход $\Omega = \{\omega_{i,j}\colon i, j = 1, 2, \ldots, n\}$, $\mathcal{F} = \{A\colon A \subset \Omega\}$ и $\mathsf{P}(A_{i,j}) = n^{-2}$. Пусть теперь случайная величина $\xi(\omega_{i,j}) = b\,|i-j|$ равна длине пробега машины за один рейс. Отсюда, используя известную формулу $\sum_{i=1}^{n} i^2 = 6^{-1}n(n+1)(2n+1)$, легко найти среднюю длину пробега машины за рейс:

$$\begin{aligned}
\mathsf{M}\xi &= b\sum_{i=1}^{n}\sum_{j=1}^{n}|i-j|\,n^{-2} = \\
&= b\,n^{-2}[\,0\cdot n + 1\cdot 2\cdot(n-1) + 2\cdot 2\cdot(n-2) + \ldots + (n-1)\cdot 2\cdot 1] = \\
&= 2b\,n^{-2}[1\cdot(n-1) + 2\cdot(n-2) + \ldots + (n-1)\cdot 1] = \\
&= 2\,b\,n^{-2}\sum_{i=1}^{n-1} i(n-i) = 2b\,n^{-2}\Big[n\sum_{i=1}^{n-1} i - \sum_{i=1}^{n-1} i^2\Big] = \\
&= 2b\,n^{-2}[n^2(n-1)\,2^{-1} - (n-1)\,n\,(2n-1)\,6^{-1}] = \\
&= 2\,b\,n^{-2}[(n^3 2^{-1} - n^2 2^{-1} - (n^2-n)(2n-1)6^{-1}] = \\
&= 3^{-1}b\,n^{-2}(n^3 - 3n^2 + n^2 + 2n^2 - n) = b\,(n^3-n)/3n^2 = b\,(n^2-1)/3n.
\end{aligned}$$

Второй способ решения. Рассмотрим теперь способ решения, который использует понятие условного математического ожидания. Пусть событие $B_i = \{\omega_{i,j}\colon i, j = 1, 2, \ldots, n\}$ заключается в том, что больного обслужили в пункте с номером i. Очевидно, что последовательность B_1, $B_2, \ldots$, B_n образует полную группу попарно несовместимых событий и вероятность $\mathsf{P}(B_i) = 1/n$ при каждом $i = 1, 2, \ldots, n$. Предположим, что при поступлении очередного вызова автомашина находится в пункте с номером i, т.е. произошло событие B_i. Вычислим при $i = 1, 2, \ldots, n$ условное математическое ожидание:

$$\begin{aligned}
\mathsf{M}(\xi \mid B) = b\sum_{i=1}^{n}|i-j|/n = b\,n^{-1}\Big[n\sum_{i=1}^{n-1}(i-j) - \sum_{j=i+1}^{n}(j-i)\Big] = \\
= 2^{-1}b\,n^{-1}[i(i-1) + (n-i)(n-i+1)].
\end{aligned}$$

Найдем теперь значение априорного математического ожидания длины пробега машины за один рейс:

$$\begin{aligned}
\mathsf{M}\xi &= \sum_{i=1}^{n}\mathsf{P}(B_i)\,\mathsf{M}(\xi \mid B_i) = \\
&= 2^{-1}b\,n^{-1}\sum_{i=1}^{n} n^{-1}[i(i-1) + (n-i)(n-i+1)] = b\,n^{-2}\sum_{i=1}^{n} i(i-1) =
\end{aligned}$$

$$= b\,n^{-2}\Big(\sum_{i=1}^{n} i^2 - \sum_{i=1}^{n} i\Big) = b\,n^{-2}6^{-1}n(n+1)(2n+1) - b\,n^{-2}2^{-1}n(n+1) =$$

$$= b(n^2-1)/3n = 3^{-1}b(n-1)(1+n^{-1}).$$

Если число n населенных пунктов достаточно большое, то средняя длина пробега машины скорой помощи за один рейс приблизительно в три раза меньше расстояния между первым и последним населенным пунктами. Отметим, что в некоторых случаях второй способ решения аналогичной задачи является существенно более простым по сравнению с первым.

Если $\mathsf{P}(B) = 0$, то понятие условного распределения и определение условного математического ожидания рассматривают относительно семейства событий и используют для этого сложный аппарат теории меры. Однако на данный момент перед нами стоит более простая задача определения условного математического ожидания случайной величины ξ_1 относительно семейства событий в случае, когда эти события порождаются другой случайной величиной ξ_2. Как правило, вероятности некоторых событий из указанного семейства оказываются равными нулю. Более того, если случайная величина ξ_2 является непрерывной с плотностью $f_{\xi_2}(y)$, то для каждого заданного $y \in R$ вероятность случайного события вида $C_y = \{\omega\colon \xi_2(\omega) = y\}$ равна нулю. Другими словами, имеем равенство $\mathsf{P}(C_y) = 0$.

Напомним, что в разделе об условных законах распределения из гл. 7 получены формулы для $f_{\xi_1}(x \mid C_y) = f_{\xi_1}(x \mid y)$, $F_{\xi_1}(x \mid C_y) = F_{\xi_1}(x \mid y)$, когда (ξ_1, ξ_2) является двумерной непрерывной случайной величиной, и формулу для условной интегральной функции $F_{\xi_1}(x \mid C_y) = F_{\xi_1}(x \mid y)$, если ξ_2 является дискретной случайной величиной. К сожалению, получить аналогичные формулы в общем случае не представляется возможным. Однако важно, что при фиксированном $y \in R$ функция $F_{\xi_1}(x \mid y)$ от переменной x является интегральной функцией распределения для случайной величины ξ_1. Поэтому по отношению к функции $F_{\xi_1}(x \mid y)$ при некоторых дополнительных условиях, которые в реальных задачах выполняются и являются естественными, можно определить условное математическое ожидание, условную дисперсию и другие условные числовые характеристики случайной величины ξ_1 относительно значений любой другой случайной величины ξ_2. Например, при дополнительном ограничении вида $\mathsf{M}|\xi_1| < +\infty$ условным математическим ожиданием случайной величины ξ_1 относительно значений случайной величины ξ_2 называют функцию $\int_{-\infty}^{+\infty} x\,dF_{\xi_1}(x \mid y)$ от переменной $y \in R$ и обозначают ее через $\mathsf{M}(\xi_1 \mid \xi_2 = y)$. Так как условное распределение $F_{\xi_1}(x \mid y)$ вычисляется не однозначно, то условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ определяется однозначно лишь с точностью до некоторого множества $B_0 \subset R$, для которого вероятность $\mathsf{P}(\{\omega\colon \xi_2(\omega) \in B_0\}) = \mathsf{P}_{\xi_2}(B_0) = 0$. Доказательство существования условного математического ожидания $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ при $\mathsf{M}|\xi_1| < +\infty$ для произвольных двумерных случайных величин (ξ_1, ξ_2) является довольно громоздким. Поэтому мы установим этот факт для частных случаев, которые важны для прикладной теории вероятностей.

Покажем сначала существование условного математического ожидания $\mathsf{M}(\xi_1 \mid \xi_2 = y) = \int_{-\infty}^{+\infty} x\,dF_{\xi_1}(x \mid y)$ для всех $y \in R$ в случае, когда ξ_2 является дискретной случайной величиной и $\mathsf{M}|\xi_1| < +\infty$. Если $y \in R$ и вероятность $\mathsf{P}(\{\omega\colon \xi_2(\omega) = y\}) > 0$, то существование условного математического ожидания $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ следует из леммы 9.5. Используя формулу для условного закона распределения $F_{\xi_1}(x \mid y)$, получим следующее равенство:

$$\mathsf{M}(\xi_1 \mid \xi_2 = y) = (\mathsf{P}(\{\omega\colon \xi_2(\omega){=}y\}))^{-1} \int_{-\infty}^{+\infty} x\,d(\mathsf{P}(\{\omega\colon \xi_1(\omega){<}x,\ \xi_2(\omega){=}\,y\})).$$

Если $y \in R$ и $\mathsf{P}(\{\omega\colon \xi_2(\omega) = y\}) = 0$, то $F_{\xi_1}(x \mid y) = F_{\xi_1}(x)$ и, значит, условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2 = y) = \mathsf{M}\xi_1 = \int_{-\infty}^{+\infty} x\,dF_{\xi_1}(x)$.

Рассмотрим теперь вопрос о существовании условного математического ожидания $\mathsf{M}(\xi_1 \mid \xi_2 = y)$, когда случайный вектор (ξ_1, ξ_2) обладает совместной плотностью вероятности $f(x, y)$ и $\mathsf{M}|\xi_1| < +\infty$. В этом случае

$$\mathsf{M}|\xi_1| = \int_{-\infty}^{+\infty} |x|\,f_{\xi_1}(x)\,dx = \int_{-\infty}^{+\infty} |x| \Big(\int_{-\infty}^{+\infty} f(x,y)\,dy\Big)\,dx = \int_{-\infty}^{+\infty} \Big(\int_{-\infty}^{+\infty} |x|\,f(x,y)\,dx\Big)\,dy < +\infty.$$

Отсюда получаем, что

$$\int_{-\infty}^{+\infty} |x|\,f(x,y)\,dx < +\infty \quad \text{и} \quad \Big|\int_{-\infty}^{+\infty} x\,f(x,y)\,dx\Big| < +\infty.$$

Используя свойства интеграла Римана–Стилтьеса, формулы (7.20), (7.15) и обозначение $B_0 = \{y\colon f_{\xi_2}(y) = 0\}$, находим, что

$$\mathsf{M}(\xi_1 \mid \xi_2 = y) \int_{-\infty}^{+\infty} x\,dF_{\xi_1}(x \mid y) = \int_{-\infty}^{+\infty} x\,f_{\xi_1}(x \mid y)\,dx = \Big(\int_{-\infty}^{+\infty} f(u,y)\,du\Big)^{-1} \Big(\int_{-\infty}^{+\infty} x\,f(x,y)\,dx\Big)$$

при $y \in R\ B_0$ и величина $\mathsf{M}(\xi_1 \mid \xi_2 = y) = \mathsf{M}(\xi_1)$ при $y \in B_0$.

Итак, в дальнейшем при рассмотрении условных математических ожиданий будем предполагать, что соответствующие случайные величины имеют конечные безусловные математические ожидания. Если при некотором фиксированном $y \in R$ случайное событие $\{\omega\colon \xi_2(\omega) = y\} \neq \varnothing$, то число $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ называется условным математическим ожиданием случайной величины ξ_1 при условии, что случайная величина $\xi_2(\omega)$ приняла значение $y \in R$. Используя понятие условного математического ожидания случайных величин, можно естественным образом определить условную дисперсию и другие условные числовые характеристики. Например, условная дисперсия величины ξ_1 относительно значений другой случайной величины ξ_2 равна $\mathsf{D}(\xi_1 \mid y) = \mathsf{M}[(\xi_1 - \mathsf{M}(\xi_1 \mid \xi_2 = y))^2 \mid y]$ и является функцией от переменной $y \in R$. Наряду с условными характеристиками $\mathsf{M}(\xi_1 \mid \xi_2 = y)$, $\mathsf{D}(\xi_1 \mid \xi_2 = y)$ мы будем рассматривать, и вычислять условное математическое ожидание $\mathsf{M}(\xi_2 \mid \xi_1 = y)$ и условную дисперсию $\mathsf{D}(\xi_2 \mid \xi_1 = y)$ случайной величины ξ_2 относительно значений величины ξ_1.

2.3. Общие свойства условного математического ожидания. Пусть число $a > 0$ и $\Pi_i = \{\Pi_{i,1}, \Pi_{i,2}, \ldots, \Pi_{i,r(i)}\}$ является i-м элементом st-допустимого класса разбиений промежутка $[-a, +a) \subset R$, когда используется функция распределения $F_{\xi_1}(x \mid y)$. Используя определение интеграла Стилтьеса, найдем условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ с помощью следующих равенств:

$$\int\limits_{-\infty}^{+\infty} x\, dF_{\xi_1}(x \mid y) = \lim_{a\to\infty} \int\limits_{-a}^{+a} x\, dF_{\xi_1}(x \mid y) = \lim_{a\to+\infty} \lim_{i\to+\infty} \sum_{k=1}^{r(i)} v'_{i,k} \Delta F_{\xi_1}(\Pi_{i,k} \mid y),$$

где число $v'_{i,k} \in \Pi_{i,k}$ и $\Delta F_{\xi_1}(\Pi_{i,k} \mid y)$ есть собственное приращение по аргументу x функции $F_{\xi_1}(x \mid y)$ на промежутке $\Pi_{i,k} \subset [-a, +a)$. Значит, $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ вычисляется с помощью арифметических операций и предельных переходов. Напомним, что для каждого фиксированного значения $x \in R$ функция $F_{\xi_1}(x \mid y)$ от переменной y является борелевской. Далее, арифметические операции и предельный переход, будучи применены к борелевским функциям, не выводят нас за пределы этого класса. Отсюда следует, что условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ является борелевской функцией $g(y)$ от переменной y. Это обстоятельство позволяет определить случайную величину $\mathsf{M}(\xi_1 \mid \xi_2 = y) = g(\xi_2(\omega))$ как неслучайную функцию $g(\cdot)$ от случайного аргумента ξ_2, которую называют *условным математическим ожиданием случайной величины ξ_1 относительно случайной величины ξ_2*.

Рассмотрение более сложного математического объекта $\mathsf{M}(\xi_1 \mid \xi_2)$ по сравнению с $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ позволяет в некоторых случаях доказать важные свойства условного математического ожидания. При этом свойства условных математических ожиданий и условных дисперсий при каждом фиксированном $y \in R$ аналогичны свойствам безусловных математических ожиданий и безусловных дисперсий. Однако при формулировке и доказательстве этих свойств необходимо учитывать следующие два важных факта. Во-первых, большинство свойств для $\mathsf{M}(\xi_1 \mid \xi_2 = y) = g(\xi_2(\omega))$ выполняются при всех $\omega \in \Omega \backslash \{\omega\colon \xi_2(\omega) \in B_0\}$, а свойства для $\mathsf{M}(\xi_1 \mid \xi_2 = y) = g(y)$ справедливы при всех $y \in R \backslash B_0$, где B_0 — борелевское множество и имеет место $\mathsf{P}(\Omega \backslash \{\omega\colon \xi_2(\omega) \in B_0\}) = \mathsf{P}_{\xi_2}(R \backslash B_0) = 1$. Иначе говоря, свойства для условных математических ожиданий верны лишь почти всюду или с вероятностью единица. Во-вторых, арифметические операции выполняются теперь над функциями, а не над числами. Поэтому свойства условных математических ожиданий носят функциональный характер. Перечислим некоторые наиболее важные для практических целей свойства условного математического ожидания.

УМ.1. *Если c — постоянная, $\xi_1(\omega) \equiv c$ для всех $\omega \in \Omega$ и ξ_2 — произвольная случайная величина, то с вероятностью единица* $\mathsf{M}(\xi_1 \mid \xi_2) = c$.

УМ.2. *Если $\mathsf{M}|\xi_1| < +\infty$ и случайные величины ξ_1, ξ_2 независимы, то в качестве одного из возможных вариантов условного математического ожидания $\mathsf{M}(\xi_1 \mid \xi_2)$ можно выбрать величину* $\mathsf{M}\xi_1$.

УМ.3. *При $\mathsf{M}|\xi_1| < +\infty$ имеет место равенство* $\mathsf{M}(\mathsf{M}(\xi_1 \mid \xi_2)) = \mathsf{M}\xi_1$.

УМ.4. *Пусть $\mathsf{M}|\xi_1| < +\infty$, $\mathsf{M}|\xi_2| < +\infty$. Если b и c являются постоянными, а ξ_1, ξ_2 случайными величинами, то с вероятностью единица имеет место равенство* $\mathsf{M}[b\,\xi_1 + c\,\xi_2 \mid \eta] = b\,\mathsf{M}(\xi_1 \mid \eta) + c\,\mathsf{M}(\xi_2 \mid \eta)$.

УМ.5. *Если $g(x)\colon R \to R$ и $h(y)\colon R \to R$ являются борелевскими функциями и $\mathsf{M}|g(\xi_1)| < +\infty$, то $\mathsf{M}[g(\xi_1) \cdot h(\xi_2) \mid \xi_2] = h(\xi_2) \cdot \mathsf{M}[g(\xi_1) \mid \xi_2]$ с вероятностью единица.*

Эти свойства непосредственно следуют из определений условных законов распределения, условного математического ожидания и свойств интеграла Римана–Стилтьеса. В качестве иллюстрации докажем первые четыре из указанных свойств.

Рассмотрим первое свойство. Если $\xi_1(\omega) \equiv c$ для всех $\omega \in \Omega$, то функция $F(x, y) = \mathsf{P}(\{\omega\colon \xi_1(\omega) < x,\, \xi_2(\omega) < y\}) = 0$ при $x \leqslant c$, $y \in R$ и равна $\mathsf{P}(\{\omega\colon \xi_2(\omega) < y\})$ при $x > c$, $y \in R$. Подставляя функцию $F(x, y)$ в левую часть равенства (7.19), получим, что интеграл $\int_{-\infty}^{+\infty} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v)$ равен нулю при $x \leqslant c$, $y \in R$ и равен $\mathsf{P}(\{\omega\colon \xi_2(\omega) < y\})$ при $x > c$, $y \in R$. Отсюда $\int_{-\infty}^{+\infty} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v) = 0$ при $x \leqslant c$ и $\int_{-\infty}^{+\infty} (1 - F_{\xi_1}(x \mid v))\, dF_{\xi_2}(v) = 0$ при $x > c$. При любом фиксированном $x \in R$ функции вида $F_{\xi_1}(x \mid y)$ и $1 - F_{\xi_1}(x \mid y)$ от переменной y являются борелевскими. Поэтому на выборочном вероятностном пространстве $(R, \mathcal{B}, \mathsf{P}_{\xi_2}(\cdot))$, которое соответствует случайной величине ξ_2, математические объекты $F_{\xi_1}(x \mid y)$ и $1 - F_{\xi_1}(x \mid y)$ при любом фиксированном $x \in R$ будут случайными величинами, где $y \in R$. По отношению к вероятностному пространству $(R, \mathcal{B}, \mathsf{P}_{\xi_2}(\cdot))$ математические ожидания $\mathsf{M}(F_{\xi_1}(x \mid y))$ и $\mathsf{M}(1 - F_{\xi_1}(x \mid y))$ соответственно равны величинам $\int_{-\infty}^{+\infty} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v)$ и $\int_{-\infty}^{+\infty} (1 - F_{\xi_1}(x \mid v))\, dF_{\xi_2}(v)$. Учитывая теперь соотношение $0 \leqslant F_{\xi_1}(x \mid y) \leqslant 1$, равенство $\mathsf{M}(F_{\xi_1}(x \mid y)) = 0$ при $x \leqslant c$ и равенство $\mathsf{M}(1 - F_{\xi_1}(x \mid y)) = 0$ при $x > c$, на основании леммы 8.7 найдем, что значение функции $F_{\xi_1}(x \mid y)$ равно нулю при $x \leqslant c$, $y \in R \backslash B$ и равно единице при $x > c$, $y \in R \backslash B$. Здесь B — некоторое борелевское множество и $\mathsf{P}_{\xi_2}(R \backslash B) = 1$. Положим теперь $F_{\xi_1}(x \mid y) = F(x)$ при всех $y \in B$, где $F(x)$ — некоторая функция распределения и $\int_{-\infty}^{+\infty} |x|\, dF(x) < \infty$. Тогда $\mathsf{M}(\xi_1 \mid \xi_2 = y) = \int_{-\infty}^{+\infty} x\, dF_{\xi_1}(x|y) = c$ при всех $y \in R \backslash B$ и $\mathsf{M}(\xi_1 \mid \xi_2 = y) = \int_{-\infty}^{+\infty} x\, dF(x)$ при всех $y \in B$, или условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2(\omega)) = c$ при $\omega \in \Omega \backslash \{\omega \colon \xi_2(\omega) \in B\}$ и, естественно, $\mathsf{M}(\xi_1 \mid \xi_2(\omega)) = \int_{-\infty}^{+\infty} x\, dF(x)$ при всех $\omega \in \{\omega \colon \xi_2(\omega) \in B\}$, где $\mathsf{P}_{\xi_2}(B) = \mathsf{P}(\{\omega \colon \xi_2(\omega) \in B\}) = 0$. Итак, $\mathsf{M}(\xi_1 \mid \xi_2 = y) = c$ почти всюду на R и условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2(\omega)) = c$ почти для всех точек Ω. В качестве одного из вариантов условного интегрального закона распределения $F_{\xi_1}(x \mid y)$ случайной величины $\xi_1(\omega) \equiv c$ можно взять, например, функцию вида:

$$F_{\xi_1}(x \mid y) = \begin{cases} 0 & \text{при } \ x \leqslant c, \ y \in R, \\ 1 & \text{при } \ x > c, \ y \in R. \end{cases} \tag{9.18}$$

В самом деле, при каждом фиксированном $y \in R$ функция $F_{\xi_1}(x \mid y)$ является интегральной функцией распределения. Далее, при любом фиксированном $x \in R$ функция $F_{\xi_1}(x \mid y)$ от переменной y равна постоянной, и, значит, является борелевской. Непосредственные вычисления правой части равенства (7.19) с учетом формулы (9.18) для $F_{\xi_1}(x \mid y)$ и свойств интеграла Римана–Стилтьеса приводят к следующему соотношению:

$$\int_{-\infty}^{y} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v) =$$

$$= \begin{cases} 0 & \text{при } x \leqslant c, \ y \in R, \\ \displaystyle\int_{-\infty}^{y} dF_{\xi_2}(\nu) = \mathsf{P}(\{\omega \colon \xi_2(\omega) < y\}) & \text{при } x > c, \ y \in R. \end{cases}$$

Следовательно, функция $F_{\xi_1}(x \mid y)$ от переменных x и y удовлетворяет соотношению (7.19). Теперь из определения условного математического ожидания $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ и свойств интеграла Римана–Стилтьеса получим, что $\mathsf{M}(\xi_1 \mid \xi_2 = y) = \int_{-\infty}^{+\infty} x\, dF_{\xi_1}(x \mid y) = c$ для $y \in R$. Поэтому условное математическое ожидание вида $\mathsf{M}(\xi_1 \mid \xi_2 = y) = \mathsf{M}(c \mid \xi_2) = c$ для всех $\omega \in \Omega$.

Аналогичным способом доказывается второе свойство. Действительно, если случайные величины ξ_1 и ξ_2 независимы, то для $x \in R$ и $y \in R$ в качестве одного из возможных вариантов условного интегрального закона распределения $F_{\xi_1}(x \mid y)$ случайной величины ξ_1 можно взять функцию $F_{\xi_1}(x)$. Очевидно, что при любом фиксированном $x \in R$ функция $F_{\xi_1}(x \mid y)$ от переменной y равна постоянной $F_{\xi_1}(x)$, и, значит, является борелевской. Более того, при каждом фиксированном $y \in R$ функция $F_{\xi_1}(x \mid y) = F_{\xi_1}(x)$ является интегральной функцией распределения случайной величины ξ_1. Наконец, учитывая свойство независимости $F(x, y) = F_{\xi_1}(x) \cdot F_{\xi_2}(y)$, $x \in R$, $y \in R$, случайных величин ξ_1 и ξ_2, легко проверяется соотношение (7.19):

$$F_{\xi_1}(x) \cdot F_{\xi_2}(y) = \int_{-\infty}^{y} F_{\xi_1}(x \mid v)\, dF_{\xi_2}(v),$$

$$F_{\xi_1}(x) \cdot F_{\xi_2}(y) = \int_{-\infty}^{y} F_{\xi_1}(x)\, dF_{\xi_2}(v),$$

$$F_{\xi_1}(x) \cdot F_{\xi_2}(y) = F_{\xi_1}(x) \int_{-\infty}^{y} dF_{\xi_2}(v),$$

$$F_{\xi_1}(x) \cdot F_{\xi_2}(y) = F_{\xi_1}(x) \cdot F_{\xi_2}(v).$$

Откуда, согласно определению условного математического ожидания, имеем: $\mathsf{M}(\xi_1 \mid \xi_2 = y) = \int_{-\infty}^{y} x\, dF_{\xi_1}(x \mid y) = \int_{-\infty}^{y} x\, dF_{\xi_1}(x) = \mathsf{M}\xi_1$ для всех $y \in R$. Значит, в качестве $\mathsf{M}(\xi_1 \mid \xi_2)$ можно взять постоянную величину $\mathsf{M}\xi_1$.

Докажем третье свойство. При каждом $y \in R$ функция $F_{\xi_1}(x \mid y)$ от x удовлетворяет условиям теоремы о выборочном вероятностном пространстве из гл. 6, и, значит, является интегральной функцией для некоторой случайной величины. Функция $F_{\xi_1}(x \mid y)$ удовлетворяет ограничению 2) в определении 7.6. Поэтому функция $F_{\xi_1}(x \mid y)$ от переменной x является равномерно ограниченной и неубывающей при каждом фиксированном $y \in R$. Применяя теперь при $\mathsf{M}|\xi_1| < +\infty$ формулу вычисления безусловного математического ожидания, связь между одномерной и двумерной интегральными функциями распределения, соотношение (7.19) при $y = +\infty$, формулу интегрирования по параметру для интеграла Стилтьеса и, наконец, определение условного математического ожидания, получим

$$\begin{aligned}
\mathsf{M}\xi_1 &= \int\limits_{-\infty}^{+\infty} x\,dF_{\xi_1}(x) = \int\limits_{-\infty}^{+\infty} x\,dF(x,+\infty) = \\
&= \int\limits_{-\infty}^{+\infty} x\,d\Big(\int\limits_{-\infty}^{+\infty} F_{\xi_1}(x\mid v)\,dF_{\xi_2}(v)\Big) = \int\limits_{-\infty}^{+\infty}\Big(\int\limits_{-\infty}^{+\infty} x\,dF_{\xi_1}(x\mid v)\Big)\,dF_{\xi_2}(v) = \\
&= \int\limits_{-\infty}^{+\infty} \mathsf{M}(\xi_1\mid \xi_2 = v)\,dF_{\xi_2}(v) = \int\limits_{-\infty}^{+\infty} g(v)\,dF_{\xi_2}(v) = \\
&= \mathsf{M}(g(\xi_2)) = \mathsf{M}(\mathsf{M}(\xi_1\mid\xi_2)).
\end{aligned}$$

Из этого свойства вытекает следующий дополнительный вывод. Так как математическое значение случайной величины ξ_1 принимает конечное значение и $\mathsf{M}\xi_1 = \int_{-\infty}^{+\infty} g(v)\,dF_{\xi_2}(v)$, то условное математическое ожидание $\mathsf{M}(\xi_1\mid\xi_2 = y) = g(y)$ является ограниченной функцией.

Ради простоты рассуждений докажем четвертое свойство при условии, что все рассматриваемые в этом абзаце случайные величины являются непрерывными. Предварительно установим некоторое свойство для двумерной непрерывной случайной величины (ξ,η) с плотностью $f(x,y)$, компонента ξ которой удовлетворяет неравенству $\mathsf{M}|\xi_1| < +\infty$. Пусть теперь $U(y)\colon R\to R$ есть любая ограниченная борелевская функция. Тогда для математических объектов (ξ,η), $U(\eta)$, $\mathsf{M}(\xi\mid\eta = y) = g_0(y)$ имеем

$$\begin{aligned}
\mathsf{M}(\xi\,U(\eta)) &= \int\limits_{-\infty}^{+\infty}\Big(\int\limits_{-\infty}^{+\infty} x\,U(y)\,f(x,y)\,dx\Big)dy = \\
&= \int\limits_{-\infty}^{+\infty}\Big(\int\limits_{-\infty}^{+\infty} x\,U(y)\,f_\eta(y)\,f_\xi(x\mid y)\,dx\Big)\,dy = \\
&= \int\limits_{-\infty}^{+\infty} U(y)\,f_\eta(y)\Big(\int\limits_{-\infty}^{+\infty} xf_\xi(x\mid y)\,dx\Big)dy = \int\limits_{-\infty}^{+\infty} U(y)\,f_\eta(y)\,\mathsf{M}(\xi\mid\eta = y)\,dy = \\
&= \int\limits_{-\infty}^{\infty} U(y)\,g_0(y)\,f_\eta(y)\,d) = \mathsf{M}(g_0(\eta)\,U(\eta)).
\end{aligned}$$

Следовательно, условное математическое ожидание $\mathsf{M}(\xi\mid\eta) = g_0(\eta(\omega))$ удовлетворяет соотношению

$$\mathsf{M}(\xi\,U(\eta)) = \mathsf{M}(g_0(\eta)\,U(\eta)) \tag{9.19}$$

для произвольной ограниченной борелевской функции $U(y)\colon R \to R$. Если теперь существует другая борелевская функция $g^*(y)\colon R \to R$, которая удовлетворяет соотношению (9.19) при всех ограниченных борелевских функций $U(y)\colon R \to R$, то

$$\begin{gathered} \mathsf{M}(\xi\, U(\eta)) = \mathsf{M}(g_0(\eta)\, U(\eta)) = \mathsf{M}(g^*(\eta)\, U(\eta)), \\ \mathsf{M}(g_0(\eta)\, U(\eta)) - \mathsf{M}(g^*(\eta)\, U(\eta)) = \mathsf{M}((g_0(\eta) - g^*(\eta))\, U(\eta)) = 0. \end{gathered} \tag{9.20}$$

Пусть при $n = 1, 2, \ldots$ функция $I_{Y_n}(y)\colon R \to \{0, 1\}$ является индикатором множества $Y_n = \{y\colon |g_0(y)| \leqslant n,\ |g^*(y)| \leqslant n\}$, иначе: функция $I_{Y_n}(y)$ равна единице при $y \in Y_n$ и равна нулю при $y \in R \backslash Y_n$. Тогда специально подобранная функция $U(y) = I_{Y_n}(y)(g_0(y) - g^*(y))$ является борелевской и ограниченной. Подставляя $U(\eta) = I_{Y_n}(\eta)(g_0(\eta) - g^*(\eta))$ в последнее равенство соотношения (9.20), мы легко получим, что $\mathsf{M}[I_{Y_n}(\eta)(g_0(\eta) - g^*(\eta))^2] = 0$. Отсюда, используя лемму 8.7, имеем при каждом $n = 1, 2, \ldots$ равенство $I_{Y_n}(\eta)|g_0(\eta) - g^*(\eta)| = 0$ с вероятностью единица. Воспользуемся этим равенством и устремим переменную величину n к $+\infty$, тогда получим $g_0(\eta) = g^*(\eta)$ с вероятностью единица. В силу такого рода однозначности условное математическое ожидание $\mathsf{M}(\xi \mid \eta = y) = g_0(y)$ можно определить из функционального уравнения (9.19). Для завершения доказательства четвертого свойства введем следующие обозначения: $\mathsf{M}(\xi_1 \mid \eta = y) = g_1(y)$, $\mathsf{M}(\xi_2 \mid \eta = y) = g_2(y)$, $\mathsf{M}(b\,\xi_1 + c\,\xi_2 \mid \eta) = g_+(y)$. Тогда из (9.19) легко найдем:

$$\begin{aligned} &\mathsf{M}((b\,\xi_1 + c\,\xi_2)\, U(\eta)) = \mathsf{M}(g_+(\eta)\, U(\eta)), \\ &\mathsf{M}((b\,\xi_1 + c\,\xi_2)\, U(\eta)) = \mathsf{M}(b\,\xi_1\, U(\eta)) + \mathsf{M}(c\,\xi_2\, U(\eta)) = \\ &\quad = b\mathsf{M}(\xi_1\, U(\eta)) + c\,\mathsf{M}(\xi_2\, U(\eta)) = b\,\mathsf{M}(g_1(\eta)\, U(\eta)) + c\,\mathsf{M}(g_2(\eta)\, U(\eta)) = \\ &\quad = \mathsf{M}(b\, g_1(\eta)\, U(\eta)) + \mathsf{M}(c\, g_2(\eta)\, U(\eta)) = \mathsf{M}((b\, g_1(\eta) + g_2(\eta))\, U(\eta)). \end{aligned}$$

Из этих равенств находим, что $g_+(\eta) = bg_1(\eta) + cg_2(\eta)$ с вероятностью единица. Отсюда, используя введенные выше обозначения, имеем равенство вида $\mathsf{M}(b\,\xi_1 + c\,\xi_2 \mid \eta) = b\,\mathsf{M}(\xi_1 \mid \eta) + c\,\mathsf{M}(\xi_1 \mid \eta)$ с вероятностью единица.

Условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2) = g(\xi_2)$ играет фундаментальную роль в проблеме предсказания (прогноза, оценивания) случайной величины ξ_1 по наблюдениям за случайной величиной ξ_2. Пусть H_1 множество всех борелевских функций $h(x)\colon R \to R$ и $\mathsf{M}(h(\xi_2))^2 < \infty$. Тогда величину $h(\xi_2)$ называют *оценкой* или *прогнозом случайной величины* ξ_1 *по случайной величине* ξ_2 *в классе* H_1. Чтобы определить точность оценки, предполагаем выполнение неравенства $\mathsf{M}(\xi_1)^2 < \infty$. Тогда при $\mathsf{M}(\xi_1)^2 < \infty$ и $\mathsf{M}(h(\xi_2))^2 < \infty$ имеем $|\xi_1 h(\xi_2)| \leqslant$ $\leqslant 2^{-1}[(\xi_1)^2 + (h(\xi_2))^2]$. Учитывая это обстоятельство и свойства безусловного математического ожидания, получим следующие соотношения: $\mathsf{M}|\xi_1 h(\xi_2)| < \infty$, $\mathsf{M}[\xi_1 - h(\xi_2)]^2 = \mathsf{M}[(\xi_1)^2 + (h(\xi_2))^2 - 2\xi_1 h(\xi_2)] = \mathsf{M}(\xi_1)^2 + \mathsf{M}(h(\xi_2))^2 - 2\,\mathsf{M}(\xi_1 h(\xi_2)) < \infty$. Поэтому неотрицательное число $\mathsf{M}[\xi_1 - h(\xi_2)]^2$ определяет среднеквадратическую ошибку предложенной оценки или точность приближенного равенства $\xi_1 \approx h(\xi_2)$. Условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2) = g(\xi_2)$ для $h(x) \in H_1$ удовлетворяет условию $\mathsf{M}(\xi_1 - h(\xi_2))^2 \geqslant \mathsf{M}(\xi_1 - g(\xi_2))^2$. Иначе говоря, $g(\xi_2)$ является наилучшей (оптимальной) оценкой случайной величины ξ_1 по случайной величине ξ_2 в классе H_1. Действительно, используя свойства УМ.5 и УМ.4 условного математического ожидания, найдем: $\mathsf{M}((\xi_1 - g(\xi_2))(g(\xi_2) - h(\xi_2) \mid \xi_2) = (g(\xi_2) - h(\xi_2)) \cdot \mathsf{M}((\xi_1 - g(\xi_2) \mid \xi_2) = (g(\xi_2) - h(\xi_2))(\mathsf{M}(\xi_1 \mid \xi_2) - g(\xi_2)) = 0$. По свойству УМ.3 условного математического ожидания легко получим:

$$\mathsf{M}\{\mathsf{M}((\xi_1 - g(\xi_2))(g(\xi_2) - h(\xi_2) \mid \xi_2)\} = \mathsf{M}((\xi_1 - g(\xi_2))(g(\xi_2) - h(\xi_2)) = 0.$$

Так как $(\xi_1 - g(\xi_2))^2 \geqslant 0$, $(g(\xi_2) - h(\xi_2))^2 \geqslant 0$, то математические ожидания вида $\mathsf{M}(\xi_1 - g(\xi_2))^2$, $\mathsf{M}(g(\xi_2) - h(\xi_2))^2$ существуют и будут неотрицательными. Используя это и равенство $\mathsf{M}((\xi_1 - g(\xi_2))(g(\xi_2) - h(\xi_2))) = 0$, легко получим, что

$$\begin{aligned}\mathsf{M}(\xi_1 - h(\xi_2))^2 &= \mathsf{M}((\xi_1 - g(\xi_2) + (g(\xi_2) - h(\xi_2)))^2 = \\ &= \mathsf{M}(\xi_1 - g(\xi_2))^2 + \mathsf{M}(g(\xi_2) - h(\xi_2))^2 + 2\,\mathsf{M}((\xi_1 - g(\xi_2)(g(\xi_2) - h(\xi_2))) = \\ &= \mathsf{M}(\xi_1 - g(\xi_2))^2 + \mathsf{M}(g(\xi_2) - h(\xi_2))^2 \geqslant \mathsf{M}(\xi_1 - g(\xi_2))^2\end{aligned}$$

для всех функций $h(x) \in H_1$.

Более того, покажем, что $\mathsf{M}(\xi_1 \mid \xi_2) \in H_1$. В самом деле, так как $\mathsf{M}(\xi_1 - h(\xi_2))^2 < \infty$, то $\mathsf{M}(\xi_1 - g(\xi_2))^2 < \infty$. Используя это, неравенство $\mathsf{M}(\xi_1)^2 < \infty$ и обозначение $\xi_1(\xi_1 - g(\xi_2)) = 0$, получим, что $\mathsf{M}|\xi_1(\xi_1 - g(\xi_2))| = \mathsf{M}|\theta| < \infty$. Так как $\xi_1 g(\xi_2) = (\xi_1)^2 - \theta$, то математическое ожидание $\mathsf{M}(\xi_1 g(\xi_2))$ равно $\mathsf{M}(\xi_1)^2 - \mathsf{M}\theta$ и является конечным числом. С другой стороны, применяя свойство 5° для условных математических ожиданий, найдем, что $\mathsf{M}(\xi_1 g(\xi_2)) = \mathsf{M}\{\mathsf{M}(\xi_1 g(\xi_2)) \mid \xi_2)\} = \mathsf{M}\{g(\xi_2)\,\mathsf{M}(\xi_1 \mid \xi_2)\} = \mathsf{M}(\mathsf{M}(\xi_1 \mid \xi_2))^2$. Итак, условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2) \in H_1$. Рассмотрим несколько удивительный пример, на котором продемонстрируем неожиданный результат оценивания случайной величины ξ_1 по наблюдениям за величиной ξ_2.

Пример 9.13. Пусть $\boldsymbol{\xi} = (\xi_1, \xi_2)$ равномерно распределена в представленном на рис. 9.6, *а* квадрате C^2. Стороны квадрата составляют углы $\pi/4$ с осями координат и равны по единице. Определить в классе H_1 оптимальную оценку $g(\xi_2)$ для случайной величины ξ_1 по случайной величине ξ_2.

Решение. При рассмотрении примера 9.6 для статистически зависимых случайных величин ξ_1 и ξ_2 были получены формулы для плотностей $f_{\xi_1}(x)$ и $f_{\xi_2}(y)$. Из условия задачи получаем, что $f_\xi(x,y)=1$ при $(x,y)\in C^2$ и $f_\xi(x,y)=0$ при $(x,y)\notin C^2$. Отсюда, используя соотношение (7.20), легко получим, что условная плотность распределения $f_{\xi_1}(x\mid y)$ случайной величины ξ_1 при $-2^{-1/2}<y\leqslant 0$ равна нулю, если $|x|\geqslant y+2^{-1/2}$, и равна функции $(2^{1/2}-2\,y)^{-1}$, если $|x|<y+2^{-1/2}$. Далее, условная плотность распределения $f_{\xi_1}(x\mid y)$ случайной величины ξ_1 при $0\leqslant y<2^{-1/2}$ равна нулю, если $|x|\geqslant -y+2^{-1/2}$, и равна функции $(2^{1/2}-2\,y)^{-1}$, если $|x|<-y+2^{-1/2}$. Так как функция $f_{\xi_2}(y)=0$ при $|y|>2^{-1/2}$, то для условной плотности $f_{\xi_1}(x\mid y)$ случайной величины ξ_1 можно принять равенство $f_{\xi_1}(x\mid y)=f_{\xi_1}(x)$ при $|y|>2^{-1/2}$ и $-\infty<x<+\infty$.

Рассмотрим случаи, когда мы наблюдали событие $\{\omega\colon \xi_2(\omega)=$ $=2^{-1/2}\}$ или событие $\{\omega\colon \xi_2(\omega)=-2^{-1/2}\}$. Для каждого из этих условий величина ξ_1 будет вырожденной со значением, равным нулю.

Из рассуждений предыдущего абзаца легко видеть, что при каждом фиксированном значении $y\neq 2^{-1/2}$, $-2^{-1/2}$ условная плотность $f_{\xi_1}(x\mid y)$ является четной функцией по x. Более того, условное распределение случайной величины ξ_1 при $y=2^{-1/2}$, $-2^{-1/2}$ является вырожденным, сосредоточенным в нулевой точке. Поэтому выводим, что условное математическое ожидание $\mathsf{M}(\xi_1\mid\xi_2=y)$ случайной величины ξ_1 равно нулю для всех $y\in R$. Значит, оптимальная оценка $\mathsf{M}(\xi_1\mid\xi_2)=g(\xi_2(\omega))$ случайной величины ξ_1 по ξ_2 в классе H_1 равна нулю для всех $\omega\in\Omega$. Еще раз отметим, что безусловная плотность $f_{\xi_1}(x)$ случайной величины ξ_1 является четной функцией по x. Тогда $\mathsf{M}\xi_1=0$ и, следовательно, оптимальная оценка $\mathsf{M}\xi_1$ случайной величины ξ_1 в классе $H_0\subset H_1$ также равна нулю. Итак, расширение класса оценок для такой случайной величины не приводит к улучшению оптимальной оценки. Точность каждой из таких оценок в классе H_0 или в классе H_1 равна $\mathsf{M}(\xi_1-\mathsf{M}\xi_1)^2=\mathsf{M}(\xi_1-\mathsf{M}g(\xi_2))^2=\mathsf{M}(\xi_1)^2=\mathsf{D}\xi_1=\int_{-2^{-1/2}}^{2^{-1/2}}x^2 f_{\xi_1}(x)\,dx=$ $=\int_{-2^{-1/2}}^{0}x^2(2x+2^{1/2})\,dx+\int_0^{2^{-1/2}}x^2(2^{1/2}-2x)\,dx=1/12$.

2.4. Регрессия случайных величин. Пусть ξ_1 и ξ_2 — две зависимые случайные величины. Из определения условного математического ожидания случайной величины ξ_2 относительно значений другой случайной величины ξ_1 получаем, что величина $\mathsf{M}(\xi_2\mid\xi_1=x)=$ $=\int_{-\infty}^{+\infty}y\,dF_{\xi_2}(y\mid x)$ зависит только от значения переменной $x\in R$. Если $X=\{\xi_1(\omega)\colon \omega\in\Omega\}$ есть множество возможных значений случайной величины ξ_1 и $x\in X\subset R$, то будем обозначать $\mathsf{M}(\xi_2\mid\xi_1=x)$ через $R_{\xi_2|\xi_1}(x)$. Область определения функции $R_{\xi_2|\xi_1}(x)$ совпадает с множеством X возможных значений случайной величины ξ_1. Другими словами, функция $R_{\xi_2|\xi_1}(x)$ есть сужение функции $\mathsf{M}(\xi_2\mid\xi_1=x)$ на X. На рис. 9.14 представлен типичный пример графика функции $R_{\xi_2|\xi_1}(x)$

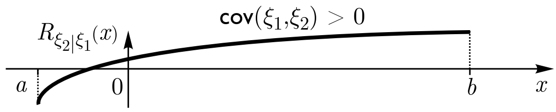


Рис. 9.14

при $X = [a, b]$. График функции $R_{\xi_2|\xi_1}(x)$, который показывает изменение условного математического ожидания случайной величины ξ_2 в зависимости от значения, принятого другой случайной величиной ξ_1, называется регрессией ξ_2 по ξ_1.

В частном случае, когда множество X всех значений случайной величины ξ_1 является счетным, график функции $R_{\xi_2|\xi_1}(x)$ состоит из отдельных точек на координатной плоскости. Например, если двумерная случайная величина (ξ_1, ξ_2) является дискретной, т.е. $X = X_1 = \{x'_1, x'_2, \ldots, x'_m\}$, $Y = Y_1 = \{y_1, y_2, \ldots, y_s\}$ и введено $r_{i,j} = \mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i,\, \xi_2(\omega) = y'_j\})$ при $i \in \{1, 2, \ldots, m\}$, $j \in \{1, 2, \ldots, s\}$, то с учетом неравенства $\mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i\}) > 0$, результатов леммы 9.5 и формул (7.18), (7.10) последовательно имеем для каждого $x = x'_i \in X$ следующее соотношение:

$$R_{\xi_2|\xi_1}(x'_i) = \mathsf{M}(\xi_2 \mid \xi_1 = x'_i) = \sum_{j=1}^{s} y'_j \mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j\} \mid \{\omega\colon \xi_1(\omega) = x'_i\}) =$$
$$= \sum_{j=1}^{s} y'_j \mathsf{P}\big(\{\omega\colon \xi_2(\omega) = y'_j,\, \xi_1(\omega) = x'_i\}\big) \times \big(\mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i\})\big)^{-1} =$$
$$= \Big(\sum_{j=1}^{s} r_{i,j}\Big)^{-1} \sum_{j=1}^{s} y'_i r_{i,j}.$$

С другой стороны, используя вычисление условного математического ожидания $\mathsf{M}(\xi_2 \mid \xi_1 = x)$ с помощью интеграла Стилтьеса $\int_{-\infty}^{+\infty} y\, dF_{\xi_2}(y \mid x)$, формулу для условной интегральной функции распределения, которая была обоснована в гл. 7, и свойства интеграла Стилтьеса найдем

$$\mathsf{M}(\xi_2 \mid \xi_1 = x) = \int\limits_{-\infty}^{+\infty} y\, dF_{\xi_2}(y \mid x) =$$
$$= \int\limits_{-\infty}^{+\infty} y\, d\Big(\sum_{y'_j < y} (\mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i\}))^{-1} \mathsf{P}(\{\omega\colon \xi_2(\omega) = y'_j, \xi_1(\omega) = x'_i\})\Big) =$$
$$= (\mathsf{P}(\{\omega\colon \xi_1(\omega) = x'_i\}))^{-1} \int\limits_{-\infty}^{+\infty} y\, d\Big(\sum_{y'_j < y} r_{i,j}\Big) = \Big(\sum_{j=1}^{s} r_{i,j}\Big)^{-1} \sum_{j=1}^{s} y'_j r_{i,j}$$

при $x = x'_i \in X$ и

$$\mathsf{M}(\xi_2 \mid \xi_1 = x) = \int\limits_{-\infty}^{+\infty} y\,dF_{\xi_2}(y \mid x) = \int\limits_{-\infty}^{+\infty} y\,dF_{\xi_2}(y) = \mathsf{M}\xi_2 = \sum_{j=1}^{s} y_j \sum_{i=1}^{m} r_{i,j}$$

при $x \in R \backslash X$.

Для регрессии случайной величины ξ_2 по ξ_1 получаем тот же результат: $R_{\xi_2|\xi_1}(x'_i) = \left(\sum_{j=1}^{s} r_{i,j}\right)^{-1} \sum_{j=1}^{s} y'_i r_{i,j}$. На рис. 9.15 отображен график функции $R_{\xi_2|\xi_1}(x)$ в виде отдельных черных точек.

Рис. 9.15

Пусть теперь двумерная случайная величина (ξ_1, ξ_2) является непрерывной с плотностью распределения $f(x, y)$. Непосредственно из (7.21) находим условную плотность распределения случайной величины ξ_2 при условии, что случайная величина ξ_1 приняла значение $x \in X$, а именно:

$$f_{\xi_2}(y \mid x) = \begin{cases} f(x,y)/f_{\xi_1}(x) & \text{при } x \in X \cap \{x\colon f_{\xi_1}(x) > 0\}, \\ f_{\xi_2}(y) & \text{при } x \in X \cap \{x\colon f_{\xi_1}(x) = 0\}. \end{cases}$$

Используя равенство $\mathsf{M}(\xi_2 \mid \xi_1 = x) = \int_{-\infty}^{+\infty} y\, f_{\xi_2}(y \mid x)\, dy$, получаем, что

$$R_{\xi_2|\xi_1}(x) = \begin{cases} \int\limits_{-\infty}^{+\infty} y\,(f_{\xi_1}(x))^{-1} f(x,y)\,dy & \text{при } x \in X \cap \{x\colon f_{\xi_1}(x) > 0\}, \\ \int\limits_{-\infty}^{+\infty} y\, f_{\xi_2}(y)\,dy & \text{при } x \in X \cap \{x\colon f_{\xi_1}(x) = 0\}. \end{cases}$$

Эти формулы дают возможность вычисления кривой регрессии случайной величины ξ_2 по ξ_1.

Аналогично можно построить регрессию случайной величины ξ_1 по ξ_2, которая для всех $y \in Y = \{\xi_2(\omega)\colon \omega \in \Omega\}$ определяется соотношением $R_{\xi_1|\xi_2}(y) = \mathsf{M}(\xi_1 \mid \xi_2 = y) = \int_{-\infty}^{+\infty} x\,dF_{\xi_1}(x \mid y)$. Иллюстративный пример кривой регрессии величины ξ_1 по ξ_2 или графика функции $R_{\xi_1|\xi_2}(y)$ изображен на рис. 9.16 при $Y = [\,c, d]$.

Из рисунков 9.14–9.16 следует, что с ростом значений одной случайной величины растет значение условного математического ожида-

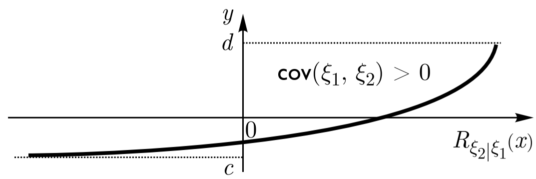


Рис. 9.16

ния другой случайной величины. Как правило, корреляционный момент для этого случая принимает строго положительное значение. Поэтому на этих рисунках отмечено значение $\mathsf{cov}(\xi_1, \xi_2) > 0$. Если случайные величины независимые, то кривые регрессии имеют вид прямых, каждая из которых параллельна соответствующей оси координат (см. рис. 9.17).

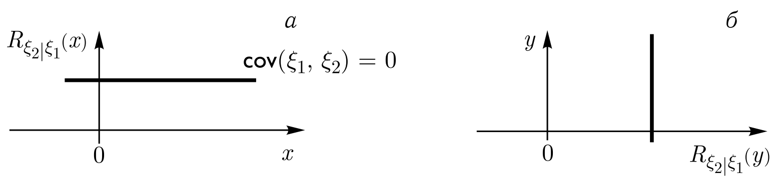


Рис. 9.17

В этом случае условное математическое ожидание одной случайной величины остается неизменным при любых значениях другой случайной величины и всегда выполняется равенство $\mathsf{cov}(\xi_1, \xi_2) = 0$. Заметим, что в некоторых случаях даже для зависимых случайных величин значения условного математического ожидания одной случайной величины остаются постоянными при изменении значений другой случайной величины. Для такого типа статистически зависимых случайных величин каждая кривая или линия регрессии, естественно, параллельна соответствующей оси координат. Если для зависимых случайных величин значения условного математического ожидания одной случайной величины будут равны нулю при изменении значений другой случайной величины, то каждая линия регрессии совпадает с соответствующей осью координат. Изучим пример такой ситуации.

Пример 9.14. Рассмотрим задачу о гончарном круге или задачу Бертрана (см. В.6 и пример 2.37) в случае, когда непреднамеренное появление трещины на гончарном круге не связано ни с закреплением одного из ее концов, ни с фиксированием ее направления. Пусть ξ_1 и ξ_2 суть расстояния от середины хорды до оси абсцисс и до оси ординат соответственно, если центр круга радиуса r и центр прямоугольной системы координат совпадают. Значит, двумерная случайная величина

(ξ_1, ξ_2) определяет среднюю точку трещины в задаче о гончарном круге или центр хорды в задаче Бертрана. В результате решения примеров 7.9 и 7.11 для вектора (ξ_1, ξ_2), компоненты которого являются статистически зависимыми случайными величинами, были получены следующие законы распределения:

1) $f_{\xi_1}(x) = 2\sqrt{r^2 - x^2}/\pi r^2$, $f_{\xi_2}(y) = 2\sqrt{r^2 - y^2}/\pi r^2$ при $|x| < r$, $|y| < r$ и $f_{\xi_1}(x) = 0$, $f_{\xi_2}(y) = 0$ при $|x| \geqslant r$, $|y| \geqslant r$;

2) $f_{\xi_1}(x \mid y) = 2^{-1}(r^2 - y^2)^{-1/2}$ при $|y| < r$, $|x| < (r^2 - y^2)^{1/2}$ и $f_{\xi_1}(x \mid y) = 0$ при $|y| < r$, $|x| \geqslant (r^2 - y^2)^{1/2}$, наконец, $f_{\xi_1}(x \mid y) = f_{\xi_1}(x)$ при $|y| > r$;

3) $f_{\xi_2}(y) \mid x) = 2^{-1}(r^2 - x^2)^{-1/2}$ при $|x| < r$, $|y| < (r^2 - x^2)^{1/2}$ и $f_{\xi_2}(y \mid x) = 0$ при $|x| < r$, $|y| \geqslant (r^2 - x^2)^{1/2}$, наконец, $f_{\xi_2}(y \mid x) = f_{\xi_2}(y)$ при $|x| > r$.

Отсюда видно, что при каждом фиксированном значении $y \neq \neq r, -r$ условная плотность $f_{\xi_1}(x \mid y)$ является четной функцией по x, а при каждом фиксированном значении $x \neq r, -r$ условная плотность $f_{\xi_2}(y \mid x)$ будет четной функцией по y. Поэтому выводим, что условные математические ожидания $\mathsf{M}(\xi_1 \mid \xi_2 = y)$ и $\mathsf{M}(\xi_2 \mid \xi_1 = x)$ тождественно равны нулю. Следовательно, линия регрессии $R_{\xi_1|\xi_2}(y)$ совпадает с осью $0y$, а линия регрессии $R_{\xi_2|\xi_1}(x)$ — с осью $0x$. Для такого рода примеров целесообразно изучить свойства условных дисперсий $\mathsf{D}(\xi_1 \mid \xi_2 = y)$ и $\mathsf{D}(\xi_2 \mid \xi_1 = x)$. Например, если наблюдаемое значение случайной величины ξ_2 равно y, где $|y| < r$, то с учетом равенства $\mathsf{M}(\xi_1 \mid \xi_2 = y) = 0$ и вида функции $f_{\xi_1}(x \mid y)$ легко находим

$$\mathsf{D}(\xi_1 \mid \xi_2 = y) = \int\limits_{-\infty}^{+\infty} x^2 f_{\xi_1}(x \mid y)\,dx =$$

$$= \int\limits_{-(r^2-y^2)^{1/2}}^{(r^2-y^2)^{1/2}} x^2(2^{-1}(r^{\ 2} - y^2)^{-1/2} dx = (r^2 - y^2)/3.$$

Аналогичным способом вычислим условную дисперсию $\mathsf{D}(\xi_2 \mid \xi_1 = x) = = (r^2 - x^2)/3$ при $|x| < r$. При решении примера 7.11 было установлено, что случайные величины ξ_1 и ξ_2 становятся вырожденными при $|y| = r$ и соответственно при $|x| = r$. Поэтому $\mathsf{D}(\xi_1 \mid \xi_2 = r) = \mathsf{D}(\xi_1 \mid \xi_2 = -r) = = \mathsf{D}(\xi_2 \mid \xi_1 = r) = \mathsf{D}(\xi_2 \mid \xi_1 = -r) = 0$. Итак, для этой задачи значение условной дисперсии одной случайной величины существенно зависит от того, какое значение приняла другая случайная величина. При этом условная дисперсия случайной величины ξ_1 или случайной величины ξ_2 принимает значение в отрезке $[0, r^2/3]$.

Если имеет место строго линейная зависимость, то линии регрессии представляют собой прямые линии с углом наклона, тангенс которого

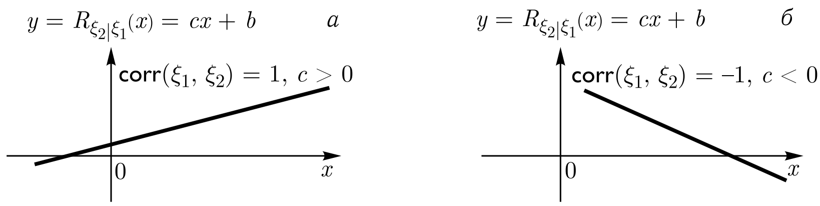


Рис. 9.18

больше нуля, если величина $\mathsf{corr}(\xi_1, \xi_2) = 1$, и тангенс наклона меньше нуля при $\mathsf{corr}(\xi_1, \xi_2) = -1$. На рис. 9.18 отображена такая ситуация.

О характере корреляционной зависимости между двумя случайными величинами ξ_1 и ξ_2 можно судить из наблюдений картины рассеивания при многократном проведении эксперимента. Для этого над системой случайных величин проделывают большое число измерений, а результаты измерений x и y, которые отмечены символом «*» на рисунках 9.19, *а*, 9.19, *б*, 9.19, *в* и 9.19, *г*, фиксируют на плоскости. На рисунках 9.19, *д* и *е* представлены ситуации, в которых нет линейной зависимости, но очевидна нелинейная. В этом случае работают другие

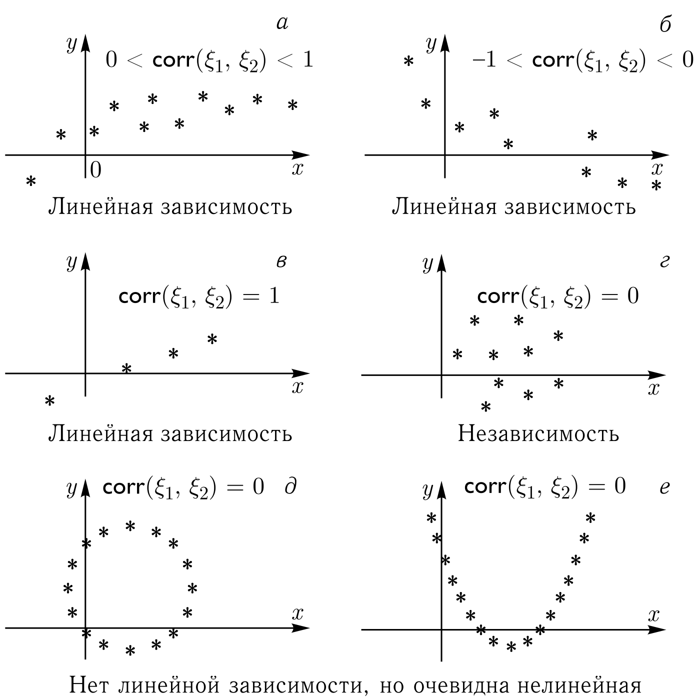


Рис. 9.19

коэффициенты корреляции и соответствующие методы, составляющие корреляционный анализ (см., например, *Кобзарь А. И.* Прикладная математическая статистика. Для инженеров и научных работников. — М.: ФИЗМАТЛИТ, 2006, 2012).

Приведем два простых примера коррелированных случайных величин. Пусть ξ_1 — ошибка прицельного устройства и ξ_2 — отклонение точки попадания снаряда от точки прицеливания. Результаты многочисленных опытов показывают, что с увеличением ошибки прицельного устройства увеличивается условное среднее отклонение точки попадания снаряда от точки прицеливания. Эту ситуацию можно проиллюстрировать рис. 9.19, *а*, для которого имеем неравенство $0 < \mathsf{corr}(\xi_1, \xi_2) < 1$.

Рассмотрим теперь другой пример. Пусть ξ_2 есть наибольшее ускорение, сообщаемое ракете двигателем, и ξ_1 есть время работы двигателя до полного сгорания топлива. В этом эксперименте, наоборот, с уменьшением времени работы двигателя до полного сгорания топлива увеличивается условное среднее ускорение ракеты. Этот эксперимент соответствует рис. 9.19, *б*, когда коэффициент корреляции удовлетворяет соотношению $-1 < \mathsf{corr}(\xi_1, \xi_2) < 0$.

Краткий обзор

В главе 9 рассматривается вопрос о функциональной зависимости между семействами случайных величин. Решается общая задача вычисления неизвестных законов распределения одного случайного вектора через известные законы распределения другого случайного вектора, если задана между ними функциональная зависимость. Для большого числа частных и важных для практики случаев (дискретные одномерные случайные величины, непрерывные одномерные случайные величины с монотонной и дифференцируемой функциональной зависимостью, непрерывные одномерные случайные величины с кусочно-монотонной и дифференцируемой функциональной зависимостью, дискретные многомерные случайные величины, непрерывные многомерные случайные величины с взаимно однозначной функциональной зависимостью и с заданной непрерывной плотностью распределения) получены формулы для вычисления неизвестных законов распределения случайных величин. Подробно излагается задача композиции законов распределения и задача имитационного моделирования одномерных и многомерных случайных величин на компьютере. Рассмотрены элементы теории корреляции и линейной статистической зависимости случайных величин. Решена простейшая задача линейного прогноза или наилучшего линейного приближения одной случайной величины с помощью другой случайной величины. Дано общее определение условного математического ожидания случайной величины на основе понятия условного эксперимента и условного вероятностного пространства. Доказаны свойства условного математического ожидания.

Найдена связь между безусловным математическим ожиданием и условными математическими ожиданиями. Установлено, что условное математическое ожидание играет фундаментальную роль в проблеме нелинейного оценивания случайной величины ξ_1 по наблюдениям за случайной величиной ξ_2. Обсуждаются вопросы использования условных математических ожиданий для измерения степени статистической зависимости между произвольными случайными величинами. Рассмотрены простейшие способы построения линий регрессий случайных величин, которые являются функциональной мерой статистической зависимости между случайными величинами.

Контрольные вопросы и упражнения

1. В чем заключается основное ограничение на функциональные зависимости между семействами случайных величин?
2. Доказать, что в случае дискретной случайной величины ξ при любой однозначной функции $g(x)\colon R \to R$ математический обект $g(\xi)$ является случайной величиной.
3. Как можно вычислить значение случайной величины η, для которой интегральная функция распределения $F_\eta(x)$ является непрерывной и строго возрастающей?
4. Площадь квадрата имеет равномерное распределение на интервале $(1, 4)$. Вычислить законы распределения длины стороны квадрата.
5. Случайная величина ξ распределена равномерно в сегменте $(-\pi/2, \pi/2)$. Определить законы распределения случайной величины $\eta = \sin\xi$.
6. Найти закон распределения случайной величины $\eta(\omega) = (\xi(\omega))^2 + 2$, если случайная величина ξ имеет плотность $f_\xi(x) = \sigma^{-1}(2\pi)^{-1/2}\exp\{-x^2/(2\sigma^2)\}$, $\sigma > 0$, $-\infty < x < +\infty$.
7 Пусть ξ и η независимые и одинаково распределенные положительные случайные величины с непрерывной плотностью рапсределения $f(x)$. Доказать, что $f(x) = \lambda\exp\{-\lambda x\}$ при $x \geqslant 0$ и $f(x) = 0$ при $x \leqslant 0$, если $\lambda > 0$ и случайные величины $(\xi - \eta)$ и $\min\{\xi, \eta\}$ независимы.
8. Определить закон распределения суммы двух независимых и равномерно распределенных на отрезке $[0, 1]$ случайных величин.
9. Приведите пример функционально зависимых и вместе с тем некоррелированных случайных величин.

10. Пусть c — постоянная, $\xi_1(\omega) \equiv c$ для всех $\omega \in \Omega$ и ξ_2 — произвольная дискретная случайная величина, которая принимает значения из множества $\{y'_1, y'_2, \ldots\}$. Показать, что условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2 = y'_j) = c$ для всех $y'_j \in \{y'_1, y'_2, \ldots\}$.
11. Пусть $\mathsf{M}|\xi_1| < +\infty$ и дискретные случайные величины ξ_1, ξ_2 независимы. Случайная величина ξ_1 принимает значения из множества $\{x'_1, x'_2, \ldots\}$, а случайная величина ξ_2 — из множества $\{y'_1, y'_2, \ldots\}$. Доказать, что условное математическое ожидание $\mathsf{M}(\xi_1 \mid \xi_2 = y'_j) = \mathsf{M}(\xi_1)$ для всех $y'_j \in \{y'_1, y'_2, \ldots\}$.
12. В чем заключается отличие условного математического ожидания $\mathsf{M}(\xi_2 \mid \xi_1 = x)$ от функции регрессии $R_{\xi_2|\xi_1}(x)$?
13. Пусть двумерная случайная величина $\boldsymbol{\xi} = (\xi_1, \xi_2)$ равномерно распределена в квадрате C^2, стороны которого составляют углы $\pi/4$ с осями координат и равны по длине единице. Найти функцию регрессии $R_{\xi_2|\xi_1}(x)$ случайной величины ξ_2 по ξ_1.

Глава 10

НЕКОТОРЫЕ НАИБОЛЕЕ РАСПРОСТРАНЕННЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ

§ 1. Тестовые дискретные случайные величины

Большое число различных классов случайных величин широко используется для построения вероятностных моделей измерителей элементарных исходов статистически устойчивых экспериментов. При этом очень часто случайные величины, которые имеют определенный закон распределения, связаны с различными статистически устойчивыми экспериментами. Поэтому целесообразно определять такого рода случайные величины и изучать их вероятностные свойства, не указывая конкретно на их происхождение от реального эксперимента. Ссылка на статистически устойчивый эксперимент будет исключительно вызвана применением и важностью того или иного класса случайных величин. Прежде всего, перейдем к изучению различных классов дискретных случайных величин.

1.1. Биномиальная случайная величина. Биномиальная случайная величина является наиболее распространенной случайной величиной дискретного типа. Она возникает, например, при реализации схемы независимых испытаний Бернулли, когда один и тот же статистически устойчивый опыт E_0 повторяется неоднократно, скажем, n раз. Схема Бернулли была подробно изучена в гл. 4. Напомним, что в результате каждого такого опыта E_0 может появиться некоторое заранее фиксированное событие A с вероятностью p или не появиться с вероятностью $q = 1 - p$. Причем эти вероятности не зависят от того, как наблюдаемые исходы эксперимента E_0 появлялись в остальных опытах. В схеме Бернулли нас интересует не исход отдельного опыта, а результат всей серии независимых испытаний из n опытов. В схеме Бернулли основным количественным признаком является случайная величина ξ, которая подсчитывает число появлений события A в серии из n независимых испытаний. Возможными значениями этой величины будут числа $0, 1, \ldots, n$. При этом вероятность вида

$\mathsf{P}(\{\omega\colon \xi(\omega) = m\}) = C_n^m p^m q^{n-m} = P_n(m)$, где $m = \overline{0, n}$, называется *биномиальной*.

Определение 10.1. Случайная величина называется *биномиальной*, если ее ряд распределения, обозначаемый через $b(n, p)$, имеет вид, представленный в табл. 10.1.

Таблица 10.1

ξ	0	1	...	m	...	n
$\mathsf{P}(\cdot)$	$P_n(0)$	$P_n(1)$	...	$P_n(m)$	...	$P_n(n)$

Случайная величина с таким рядом распределения называется биномиальной потому, что $P_n(m) = C_n^m p^m q^{n-m}$ и разложение бинома $(p+q)^n$ по формуле Ньютона имеет вид: $(p+q)^n = \sum_{m=0}^{n} C_n^m p^m q^{n-m} = 1$. Интегральная функция распределения для биномиальной случайной величины равна $F(x) = \sum_{m<x} C_n^m p^m q^{n-m}$.

Перейдем теперь к вычислению основных числовых характеристик биномиальной случайной величины ξ, непосредственно используя для этого вид ее распределения.

Математическое ожидание биномиальной случайной величины равно

$$\begin{aligned}\mathsf{M}\xi &= \sum_{m=0}^{n} m C_n^m p^m q^{n-m} = \sum_{m=1}^{n} n!\, m\,[m!\,(n-m)!]^{-1} p^m q^{n-m} = \\ &= np \sum_{m=1}^{n} (n-1)!\,[(m-1)!\,(n-m)!]^{-1} p^{m-1} q^{n-m} = \\ &= np \sum_{m=0}^{n-1} (n-1)!\,[(n-1-m)!]^{-1} p^m q^{n-1-m} = np\,(p+q)^{n-1} = np.\end{aligned}$$

Заметим, что при $n = 1$ биномиальная случайная величина ξ является индикатором события A или случайной величиной с распределением Бернулли. Дисперсия $\mathsf{D}\xi$ индикатора события A равна pq. Если $n \geqslant 2$, то дисперсию $\mathsf{D}\xi$ можно вычислить следующим простым способом:

$$\begin{aligned}\mathsf{D}\xi &= \mathsf{M}(\xi^2) - (\mathsf{M}\xi)^2 = \sum_{m=0}^{n} m^2 C_n^m p^m q^{n-m} - (np)^2 = \\ &= \sum_{m=1}^{n} n!\,(m-1+1)[(m-1)!\,(n-m)!]^{-1} p^m q^{n-m} - (np)^2 = \\ &= \sum_{m=2}^{n} n!\,[(m-2)!\,(n-m)!]^{-1} p^m q^{n-m} + \\ &\qquad + np \sum_{m=1}^{n} (n-1)!\,[(m-1)!\,(n-m)!]^{-1} p^{m-1} q^{n-m} - (np)^2 =\end{aligned}$$

$$= n(n-1)\,p^2 \sum_{m=0}^{n-2} (n-2)!\,[m!(n-2-m)!]^{-1} p^m q^{n-2-m} + np - (np)^2 =$$
$$= n(n-1)\,p^2 + np - (np)^2 = np(1-p) = npq.$$

Поэтому среднее квадратическое отклонение $\sigma(\xi) = (\mathsf{D}\xi)^{1/2}$ для биномиальной величины ξ, распределение $b(n,p)$ которой зависит от двух параметров: $n \geqslant 1$ и $0 < p < 1$, равно $(npq)^{1/2}$.

Моду $\mathsf{Mo}\,\xi$ и наивероятнейшее значение $\mathsf{Mo}^*\xi$ биномиальной случайной величины ξ можно найти, используя непосредственно экстремальное свойство вероятностей в схеме Бернулли и определение этих числовых характеристик. Действительно, распределение для биномиальной случайной величины ξ совпадает с множеством $\{C_n^m p^m q^{n-m}\colon m = 0, 1, \ldots, n\}$. Тогда на основании результатов четвертой главы получаем, что наивероятнейшее значение $\mathsf{Mo}^*\xi$ равно $[(n+1)\,p]$, если число $(n+1)\,p$ не является целым, и $\mathsf{Mo}^*\xi \in \{np - q, (n+1)\,p\}$ в противном случае. Здесь $[(n+1)\,p]$ означает целую часть числа $(n+1)\,p$. В зависимости от значений параметров n и p можно выделить семь различных вариантов для моды $\mathsf{Mo}\,\xi$ и наивероятнейшего значения $\mathsf{Mo}^*\xi$ биномиальной случайной величины ξ.

1. Пусть $n = 1$ и $(n+1)\,p = 1$. В этом простейшем случае наивероятнейшее значение $\mathsf{Mo}^*\xi \in \{0, 1\}$ и мода $\mathsf{Mo}\,\xi$ не существует.

2. Если $(n+1)\,p < 1$ и $n \geqslant 1$, то $\mathsf{Mo}^*\xi = 0$, а $\mathsf{Mo}\,\xi$ не существует.

3. При $(n+1)\,p = 1$, $n > 1$ мода $\mathsf{Mo}\,\xi = 1$ и $\mathsf{Mo}^*\xi \in \{0, 1\}$, а распределение биномиальной случайной величины ξ является унимодальным.

4. Предположим, что $1 < (n+1)\,p < n$ и число $(n+1)\,p$ не является целым. Тогда мода и наивероятнейшее значение биномиальной случайной величины совпадают, т. е. $\mathsf{Mo}^*\xi = \mathsf{Mo}\,\xi = [(n+1)\,p]$. Значит, распределение биномиальной случайной величины ξ является унимодальным.

5. Если выполняется неравенство $1 < (n+1)\,p < n$ и число $(n+1)\,p$ оказывается целым, то величины $\mathsf{Mo}^*\xi$ и $\mathsf{Mo}\,\xi$ принимают значения из множества $\{(n+1)\,p - 1, (n+1)\,p\}$. Поэтому распределение биномиальной случайной величины ξ будет мультимодальным.

6. При $(n+1)\,p = n$ и $n > 1$ мода $\mathsf{Mo}\,\xi = n - 1$, наивероятнейшее значение $\mathsf{Mo}^*\xi \in \{n-1, n\}$. Отсюда следует, что распределение биномиальной случайной величины ξ является унимодальным.

7. Наконец, если $(n+1)\,p > n \geqslant 1$, то $\mathsf{Mo}^*\xi = n$, и $\mathsf{Mo}\,\xi$ не существует.

Еще более значительные трудности вызывают вычисления и изучение свойств остальных числовых характеристик биномиальной случайной величины ξ, если использовать для этого непосредственно вид ее распределения. В этом случае для исследования вероятностных свойств целочисленных неотрицательных случайных величин и вычисления их числовых характеристик удобно применять так называемый аппарат производящих функций.

Введем теперь понятие производящей функции для произвольной неотрицательной целочисленной случайной величины η и ее дискретного распределения. Итак, пусть дана некоторая дискретная случайная величина η с своим рядом распределения (табл. 10.2)

Таблица 10.2

η	0	1	...	m	...	n	...
$\mathsf{P}(\cdot)$	p_0	p_1	...	p_m	...	p_n	...

Определение 10.2. *Производящей функцией* для неотрицательной целочисленной случайной величины η, имеющей распределение $\{\mathsf{P}(\{\omega\colon \eta(\omega) = m\}) = p_m;\ m = 0, 1, \ldots\}$, называется ряд $\sum_{m=0}^{\infty} z^m p_m$, обозначаемый через $\Pi_\eta(z)$.

Так как для распределения $\{p_m\colon m = 0, 1, \ldots\}$ всегда выполняется условие нормировки $\sum_{m=0}^{\infty} p_m = 1$, то $\Pi_\eta(z)$, очевидно, абсолютно сходится в области $|z| \leqslant 1$. Если z — действительное число, то производящая функция $\Pi_\eta(z) = \mathsf{M}(z^\eta)$. Если мы знаем производящую функцию $\Pi_\eta(z)$ случайной величины η, то коэффициенты разложения этой функции в ряд Тейлора в окрестности нулевой точки задают распределение случайной величины η. Другими словами, имеет место соотношение $p_m = (m!)^{-1}(d^m\Pi_\eta(z)/dz^m)\big|_z = 0,\ m = 0, 1, \ldots$. Можно сказать, что между производящей функцией $\Pi_\eta(z)$ и законом распределения $\{p_m\colon m = 0, 1, \ldots\}$ случайной величины η установлено взаимно однозначное соответствие. Значит, вероятностные свойства величины η задаются либо распределением, либо производящей функцией. Более того, в некоторых случаях производящая функция случайной величины вычисляется проще, чем ее распределение. В качестве подтверждения этих слов приведем доказательство следующей леммы.

Лемма 10.1. *Если случайные величины η_1, $\eta_2, \ldots$, η_n являются независимыми и принимают целочисленные неотрицательные значения, то производящая функция $\Pi_\eta(z)$ суммы $\eta = \eta_1 + \eta_2 + \ldots + \eta_n$ равна произведению $\Pi_{\eta_1}(z) \times \Pi_{\eta_2}(z) \times \ldots \times \Pi_{\eta_n}(z)$ производящих функций слагаемых.*

Доказательство. Из определения η следует, что она принимает целочисленные неотрицательные значения. Следовательно, производящая функция случайной величины η равна $\Pi_\eta(z) = \mathsf{M}(z^\eta) = \mathsf{M}(z^{\eta_1+\eta_2+\ldots+\eta_n}) = \mathsf{M}(z^{\eta_1} \times z^{\eta_2} \times \ldots \times z^{\eta_n})$. Так как η_1, $\eta_2, \ldots$, η_n — независимые величины, то по теореме 7.2 независимыми будут и величины z^{η_1}, $z^{\eta_2}, \ldots$, z^{η_n}. В силу свойств математического ожидания от произведения независимых случайных величин вытекает, что

$\Pi_\eta(z) = \mathsf{M}(z^{\eta_1}) \times \mathsf{M}(z^{\eta_2}) \times \ldots \times \mathsf{M}(z^{\eta_n}) = \Pi_{\eta_1}(z) \times \Pi_{\eta_2}(z) \times \ldots \times \Pi_{\eta_n}(z)$. Лемма 10.1 доказана.

Далее покажем, что числовые характеристики случайных величин можно вычислять с помощью метода производящих функций. Продемонстрируем этот метод на примере вычисления начальных моментов $\alpha_1\eta$, $\alpha_2\eta$, $\alpha_3\eta$, $\alpha_4\eta$ не выше четвертого порядка, математического ожидания $\mathsf{M}\eta$ и дисперсии $\mathsf{D}\eta$ случайной величины η. Для этого проделаем с производящей функцией следующие очевидные цепочки преобразований:

$$\begin{aligned}
(d\Pi_\eta(z)/dz)\big|_{z=1} &= \Big(d\Big(\sum_{m=0}^{\infty} z^m p_m\Big)/dz\Big)\Big|_{z=1} = \sum_{m=0}^{\infty} m p_m = \alpha_1\eta,\\
(d^2\Pi_\eta(z)/dz^2)\big|_{z=1} &= \Big(d^2\Big(\sum_{m=0}^{\infty} z^m p_m\Big)/dz^2\Big)\Big|_{z=1} =\\
&= \sum_{m=0}^{\infty} m(m-1)\,p_m = \alpha_2\eta - \alpha_1\eta,\\
(d^3\Pi_\eta(z)/dz^3)\big|_{z=1} &= \Big(d^3\Big(\sum_{m=0}^{\infty} z^m p_m\Big)/dz^3\Big)\Big|_{z=1} =\\
&= \sum_{m=0}^{\infty} m(m-1)(m-2)\,p_m = \alpha_3\eta - 3\alpha_2\eta - 2\alpha_1\eta,\\
(d^4\Pi_\eta(z)/dz^4)\big|_{z=1} &= \Big(d^4\Big(\sum_{m=0}^{\infty} z^m p_m\Big)/dz^4\Big)\Big|_{z=1} =\\
&= \sum_{m=0}^{\infty} m(m-1)(m-2)(m-3)\,p_m =\\
&= \alpha_4\eta - 6\alpha_3\eta + 11\alpha_2\eta - 6\alpha_1\eta.
\end{aligned}$$

Отсюда легко получаем формулы для вычисления начальных моментов $\alpha_1\eta$, $\alpha_2\eta$, $\alpha_3\eta$, $\alpha_4\eta$, математического ожидания $\mathsf{M}\eta$ и дисперсии $\mathsf{D}\eta$ в терминах производящей функции $\Pi_\eta(z)$:

$$\begin{aligned}
\alpha_1\eta &= \mathsf{M}\eta = (d\Pi_\eta(z)/dz)\big|_{z=1},\\
\alpha_2\eta &= (d^2\Pi_\eta(z)/dz^2)\big|_{z=1} + (d\Pi_\eta(z)/dz)\big|_{z=1},\\
\mathsf{D}\eta &= \alpha_2\eta - (\alpha_1\eta)^2 = (d^2\Pi_\eta(z)/dz^2)\big|_{z=1} + (d\Pi_\eta(z)/dz)\big|_{z=1} -\\
&\qquad\qquad - ((d\Pi_\eta(z)/dz)\big|_{z=1})^2, \qquad (10.1)\\
\alpha_3\eta &= (d^3\Pi_\eta(z)/dz^3)\big|_{z=1} + 3(d^2\Pi_\eta(z)/dz^2)\big|_{z=1} + (d\Pi_\eta(z)/dz)\big|_{z=1},\\
\alpha_4\eta &= (d^4\Pi_\eta(z)/dz^4)\big|_{z=1} + 6(d^3\Pi_\eta(z)/dz^3)\big|_{z=1} +\\
&\qquad + 7(d^2\Pi_\eta(z)/dz^2)\big|_{z=1} + (d\Pi_\eta(z)/dz)\big|_{z=1}.
\end{aligned}$$

Продемонстрируем эффективность способа производящих функций на выводе формул для числовых характеристик биномиальной случайной величины ξ. С этой целью получим производящую функцию для биномиальной случайной величины ξ. Так как при фиксированном $m = 0, 1, \ldots, n$ вероятность $p_m = C_n^m p^m q^{n-m}$, то производящая функция для величины ξ с биномиальным законом распределения равна $\Pi_\xi(z) = \sum_{m=0}^n z^m C_n^m p^m q^{n-m} = \sum_{m=0}^n C_n^m (pz)^m q^{n-m} = (pz+q)^n$. Дифференцируя по z производящую функцию $\Pi_\xi(z) = (pz+q)^n$, полагая затем $z = 1$ и принимая во внимание равенство $p + q = 1$, получим

$$(d^k \Pi_\xi(z)/dz^k)\big|_{z=1} = n(n-1) \times \ldots \times (n-k+1)\, p^k \tag{10.2}$$

для любого фиксированного $k = 0, 1, \ldots, n$. Используя соотношение (10.1) и равенство (10.2) для производных, легко найдем формулы для начальных моментов $\alpha_1\xi$, $\alpha_2\xi$, $\alpha_3\xi$, $\alpha_4\xi$ биномиального распределения и еще раз повторим результат для математического ожидания $\mathsf{M}\xi$ и дисперсии $\mathsf{D}\xi$, т. е.

$$\alpha_1\xi = \mathsf{M}\xi = np, \qquad \alpha_2\xi = n(n-1)\,p^2 + np,$$
$$\mathsf{D}\xi = n(n-1)\,p^2 + np - (np)^2 = npq,$$
$$\alpha_3\xi = n(n-1)(n-2)\,p^3 + 3n(n-1)\,p^2 + np,$$
$$\alpha_4\xi = n(n-1)(n-2)(n-3)\,p^4 + 6n(n-1)(n-2)\,p^3 + 7n(n-1)\,p^2 + np.$$

Эти равенства позволяют вычислить третий центральный момент $\beta_3\xi$, коэффициент асимметрии $\mathsf{Ka}\,\xi$, четвертый центральный момент $\beta_4\xi$ и, наконец, эксцесс $\mathsf{Э}\xi$ для случайной величины ξ с биномиальным законом распределения:

$$\beta_3\xi = \sum_{k=0}^{3} C_3^k (-1)^{3-k} (\alpha_1\xi)^{3-k} \alpha_k\xi = npq(q-p),$$
$$\mathsf{Ka}\,\xi = \beta_3\xi/(\boldsymbol{\sigma}\xi)^3 = (q-p)(npq)^{-1/2},$$
$$\beta_4\xi = \sum_{k=0}^{4} C_4^k (-1)^{4-k} (\alpha_1\xi)^{4-k} \alpha_k\xi = 3(npq)^2 - 6n(pq)^2 + npq,$$
$$\mathsf{Э}\xi = \beta_4\xi \times (\boldsymbol{\sigma}\xi)^{-4} - 3 = (npq)^{-1} - 6n^{-1}.$$

Рассмотрим еще один и гораздо более простой способ нахождения характеристик $\mathsf{M}\xi$ и $\mathsf{D}\xi$ для биномиальной случайной величины. Для этого введем n одинаково распределенных независимых случайных величин ξ_1, $\xi_2, \ldots,$ ξ_n. Каждая случайная величина ξ_i считает число наступления события A в опыте с номером $i = 0, 1, \ldots, n$ для схемы независимых испытаний Бернулли. Так как вероятности $\mathsf{P}(\{\omega\colon \xi_i(\omega) = 0\}) = q$, $\mathsf{P}(\{\omega\colon \xi_i(\omega) = 1\}) = p$, то $\mathsf{M}\xi_i = p$ и $\mathsf{D}\xi_i = pq$. Из определения случайных величин ξ_1, $\xi_2, \ldots,$ ξ_n следует, что биномиальная случайная величина $\xi = \xi_1 + \xi_2 + \ldots + \xi_n$. Учитывая теперь свойства математического ожидания и дисперсии для независимых бернуллиевских случайных величин ξ_1, $\xi_2, \ldots, \xi_n$, сосчитаем, что $\mathsf{M}\xi = \mathsf{M}\Big(\sum_{m=1}^n \xi_i\Big) = \sum_{m=1}^n \mathsf{M}\xi_i = np$, $\mathsf{D}\xi = \mathsf{D}\Big(\sum_{m=1}^n \xi_i\Big) = \sum_{m=1}^n \mathsf{D}\xi_i = npq$.

Используя свойства биномиальной случайной величины ξ в схеме Бернулли, можно подтвердить на содержательном уровне основной постулат о статистической вероятности $\mathsf{P}^*(A, n)$ наблюдаемого исхода A статистически устойчивого эксперимента E_0. Этот постулат заключается в том, что относительная частота $n^{-1}\mu(A, n) = \mathsf{P}^*(A, n) = n^{-1}\xi$ появления события A в серии из n независимых испытаний опыта E_0 приближенно совпадает с теоретической вероятностью $\mathsf{P}(A) = p$ по мере увеличения числа n. Так как вероятность $\mathsf{P}(\{\omega\colon \xi(\omega)\, n^{-1} = mn^{-1}\}) = \mathsf{P}(\{\omega\colon \xi(\omega) = m\}) = C_n^m p^m q^{n-m}$ для фиксированного значения $m = 0, 1, \ldots, n$, то математическое ожидание, наивероятнейшее значение и дисперсия относительной частоты $n^{-1}\mu(A, n)$ равны $\mathsf{M}(n^{-1}\xi) = p$, $\mathsf{Mo}^*(n^{-1}\xi) = n^{-1}[(n+1)\,p]$ и $\mathsf{D}(n^{-1}\xi) = n^{-1}pq$. Отсюда и из очевидных неравенств

$$0 \leqslant (n+1)\,p - [(n+1)\,p] < 1,$$
$$p + n^{-1}p - n^{-1} < n^{-1}[(n+1)\,p] \leqslant p + n^{-1}p,$$
$$p + n^{-1}p - n^{-1} < p < p + n^{-1}p$$

непосредственно вытекает, что при достаточно большом n относительная частота события A колеблется около вероятности $\mathsf{P}(A)$ с разбросом $n^{-1}pq$, принимает значение равное $\mathsf{Mo}^*(n^{-1}\xi) = n^{-1}[(n+1)\,p] \approx p$ с максимальной вероятностью по сравнению с другими ее значениями и $|\mathsf{Mo}^*(n^{-1}\xi) - p| < n^{-1}$.

Следовательно, для любого сколь угодно малого наперед заданного числа $\varepsilon > 0$ при $n > \varepsilon^{-1}$ разброс колебания относительной частоты события A около вероятности $\mathsf{P}(A)$ будет меньше ε и наивероятнейшее значение относительной частоты $n^{-1}\mu(A, n)$ отличается от вероятности $\mathsf{P}(A)$ сколь угодно мало. Теперь становится более понятным смысл приближенного совпадения относительной частоты и теоретической вероятности случайного события A статистически устойчивого эксперимента E_0.

Биномиальная случайная величина играет значительную роль при решении различных проблем в теории управляемых систем массового обслуживания, в теории надежности, в практике статистического контроля качества выпускаемой продукции, в теории стрельбы и т. д. В заключение приведем решение типичных примеров, в которых применяются свойства случайной величины с биномиальным законом распределения.

Пример 10.1. Мастер спорта по шахматам проводит сеанс одновременной игры со школьниками. Вероятность выигрыша у мастера любого из n школьников равна $p = 0,01$. Мастер выиграет сеанс, если он не проиграет ни одной партии. Сколько потребуется школьников, чтобы вероятность поражения мастера была больше b?

Решение. Пусть в сеансе участвует n школьников, и случайная величина ξ подсчитывает число всех школьников, которые выиграют у мастера. Из условий данного эксперимента следует, что случайная величина ξ является биномиальной. Вероятность того, что выиграет хотя бы один из школьников, равна $\mathsf{P}(\{\omega\colon \xi(\omega) > 0\}) = 1 - \mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) = 1 - q^n = 1 - (0{,}99)^n$. Затем из требования $\mathsf{P}(\{\omega\colon \xi(\omega) > 0\}) > b$ данной задачи последовательно шаг за шагом находим, что $1 - (0,99)^n > b$, $(0,99)^n < 1 - b$, $n\lg(99/100) < \lg(1-b)$, $n(\lg 99 - \lg 100) < \lg(1-b)$, $\lg(1-b) < n(2 - \lg 99)$, $n > -(2 - \lg 99)^{-1}\lg(1-b)$. Пусть, например, вероятность поражения мастера в таком сеансе должна быть больше 0,26. Тогда из соотношения $n > -(2 - \lg 99)^{-1}\lg(1-b)$ при $b = 0,26$ находим, что число школьников должно быть более 30.

1.2. Пуассоновская случайная величина. Рассмотрим дискретную случайную величину ξ, которая может принимать только целые неотрицательные значения 0, 1,…, m,….

Определение 10.3. Будем говорить, что дискретная случайная величина ξ *распределена по закону Пуассона*, если вероятность того, что ξ примет определенное значение m, равна $\mathsf{P}(\{\omega\colon \xi(\omega) = m\}) = (\lambda^m \exp\{-\lambda\})/m!$, где $m = 0, 1, \ldots$ и λ — положительный параметр этого закона распределения.

Вероятность $(\lambda^m \exp\{-\lambda\})/m!$, как и в гл. 4, будем обозначать через $\pi(m;\lambda)$, а распределение Пуассона через $\pi(\lambda)$. Приведем примеры встречающихся на практике такого рода случайных величин:

1) число ξ частиц, испускаемых радиоактивным изотопом за единицу времени;

2) число ξ вызовов, поступающих на телефонную станцию за единицу времени;

3) число ξ элементов сложного устройства (атомная электростанция, космический корабль и т. д.), отказавших в течение заданного отрезка времени;

4) число ξ электронов, вылетевших с катода электронной лампы за определенный промежуток времени;

5) число ξ автомашин, прибывающих к изолированному перекрестку за единицу времени.

Ряд распределения пуассоновской случайной величины ξ, который определяет все ее вероятностные свойства, имеет вид, представленный в табл. 10.3.

Таблица 10.3

ξ	0	1	2	...	m	...
$\mathsf{P}(\cdot)$	p_0	p_1	p_2	...	p_m	...

Здесь вероятность $p_m = \pi(m;\lambda) = (\lambda^m \exp\{-\lambda\})/m!$, $m = 0, 1, \ldots$. Нетрудно показать, что имеет место условие нормировки:

$$\sum_{m=0}^{\infty} (\lambda^m \exp\{-\lambda\})/m! = \exp\{-\lambda\} \sum_{m=0}^{\infty} \lambda^m/m! = \exp\{-\lambda\} \exp\{\lambda\} = 1.$$

Так как вероятности $p_m = (\lambda^m \exp\{-\lambda\})/m! > 0$ при всех $m = 0, 1, \ldots$ и имеет место условие нормировки, то приведенная табл. 10.3, действительно, может служить рядом распределения некоторой дискретной случайной величины. Функция распределения для закона Пуассона $F(x) = \sum_{m<x}(\lambda^m \exp\{-\lambda\})/m!$ Далее, $F(1) = \mathsf{P}(\{\omega\colon \xi(\omega) < 1\}) = \mathsf{P}(\{\omega\colon \xi(\omega) = 0\}) \exp\{-\lambda\}$ и вероятность $\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant 1\}) = 1 - F(1) = 1 - \exp\{-\lambda\}$. Зная ряд распределения ξ, построим многоугольник распределения. На рис. 10.1 представлены многоугольники для случайных величин с параметрами $\lambda = 0{,}5$, $\lambda = 1$ и $\lambda = 3{,}5$.

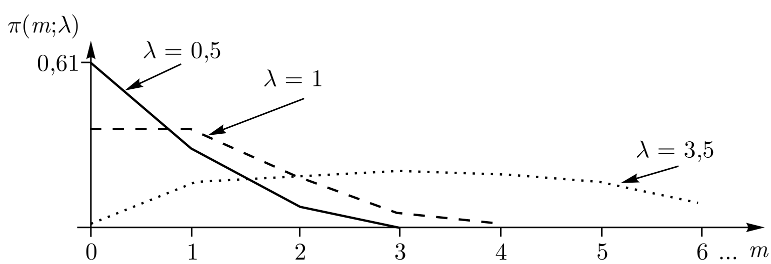


Рис. 10.1

Определим теперь основные числовые характеристики случайной величины ξ, распределенной по закону Пуассона. Математическое ожидание

$$\begin{aligned} \mathsf{M}\xi = \sum_{m=0}^{\infty} m\pi(m;\lambda) = \sum_{m=0}^{\infty} m(\lambda^m \exp\{-\lambda\})/m! = \\ = \sum_{m=1}^{\infty} m(\lambda^m \exp\{-\lambda\})/m! = \lambda \exp\{-\lambda\} \sum_{m=1}^{\infty} m\lambda^{m-1}/m! = \\ = \lambda \exp\{-\lambda\} \sum_{m=1}^{\infty} \lambda^{m-1}/(m-1)! = \lambda \exp\{-\lambda\} \exp\{\lambda\} = \lambda. \end{aligned}$$

Итак, $\mathsf{M}\xi = \lambda$. Сосчитаем теперь дисперсию $\mathsf{D}\xi$:

$$\begin{aligned}
\mathsf{D}\xi &= \alpha_2\xi - (\mathsf{M}\xi)^2,\\
\alpha_2\xi &= \sum_{m=0}^{\infty} m^2(\lambda^m \exp\{-\lambda\})/m! =\\
&= \lambda \exp\{-\lambda\} \sum_{m=1}^{\infty} m\lambda^{m-1}/(m-1)! =\\
&= \lambda \exp\{-\lambda\} \sum_{m=1}^{\infty} (m-1+1)\lambda^{m-1}/(m-1)! =\\
&= \lambda \exp\{-\lambda\} \sum_{m=1}^{\infty} (m-1)\lambda^{m-1}/(m-1)! +\\
&\qquad + \lambda \exp\{-\lambda\} \sum_{m=1}^{\infty} \lambda^{m-1}/(m-1)! =\\
&= \lambda^2 \exp\{-\lambda\} \sum_{m=2}^{\infty} \lambda^{m-2}/(m-2)! + \lambda \exp\{-\lambda\} \exp\{\lambda\} =\\
&= \lambda^2 + \lambda = \lambda(\lambda+1), \quad \mathsf{D}\xi = \lambda^2 + \lambda - \lambda^2 = \lambda.
\end{aligned}$$

Эти вычисления показывают особое свойство закона Пуассона $\pi(\lambda)$, состоящее в том, что математическое ожидание $\mathsf{M}\xi$ и дисперсия $\mathsf{D}\xi$ случайной величины, распределенной по этому закону, численно совпадают. Далее, имеем $\boldsymbol{\sigma}(\xi) = \lambda^{1/2}$.

Найдем теперь $\mathsf{Mo}^*\xi$ или наивероятнейшее значение m_0 пуассоновской случайной величины ξ, для которого имеет место равенство $\pi(m_0;\lambda) = \max\{\pi(m;\lambda)\colon m = 0,1,\ldots\}$. С этой целью рассмотрим отношение последующего члена к предыдущему члену последовательности вида $\pi(0;\lambda)$, $\pi(1;\lambda),\ldots$, $\pi(m-1;\lambda)$, $\pi(m;\lambda),\ldots$. Нетрудно установить равенства вида: $\pi(m;\lambda)/\pi(m-1;\lambda) = [\lambda^m \exp\{-\lambda\}/m!]/[\lambda^{m-1}\exp\{-\lambda\}/(m-1)!] = \lambda/m$. Отсюда, при $m < \lambda$ имеем $\pi(m;\lambda)/\pi(m-1;\lambda) > 1$ или $\pi(m;\lambda) > \pi(m-1;\lambda)$, а при $m > \lambda$ будет $\pi(m;\lambda)/\pi(m-1;\lambda) < 1$ или $\pi(m;\lambda) < \pi(m-1;\lambda)$.

Итак, для бесконечной последовательности $\pi(0;\lambda)$, $\pi(1;\lambda),\ldots$ выполняется соотношение $\pi(0;\lambda) < \pi(1;\lambda) < \ldots < \pi(m-1;\lambda) \leqslant \pi(m_0;\lambda) > \pi(m_0+1;\lambda) > \ldots$. Поэтому при $m = [\lambda]$ имеем максимальное значение $\pi(m;\lambda)$. Здесь символ $[\lambda]$ означает целую часть числа λ. Более того, получаем единственное значение $[\lambda]$ для $\mathsf{Mo}^*\xi = m_0$ пуассоновской случайной величины, если λ есть дробное число, и два значения: λ и $\lambda-1$, если λ целое. При $\lambda < 1$ имеем $\mathsf{Mo}^*\xi = m_0 = [\lambda] = 0$, $\pi(0;\lambda) > \pi(1;\lambda) > \pi(2;\lambda) > \ldots$ и, значит, мода $\mathsf{Mo}\,\xi$ пуассоновской случайной величины не существует. Если параметр $\lambda = 1$, то $\mathsf{Mo}^*\xi = m_0 \in \{0,1\}$, $\pi(0;\lambda) = \pi(1;\lambda) > \pi(2;\lambda) > \ldots$ и, следовательно, мода $\mathsf{Mo}\,\xi$ пуассоновской случайной величины равна единице. Итак, при $\lambda = 1$ распределение пуассоновской случайной величины является унимодальным. Наконец, при $\lambda > 1$ получаем, что $\mathsf{Mo}^*\xi = \mathsf{Mo}\,\xi$. Например, если параметр λ равен трем, то $\pi(0;3) < \pi(1;3) < \pi(2;3) = \pi(3;3) > \pi(4;3) > \ldots$. Отсюда находим, что $\mathsf{Mo}\,\xi$ равна двум или трем и совпадает с $\mathsf{Mo}^*\xi$. Поэтому при $\lambda = 3$ распределение пуассоновской случайной величины является мультимодальным.

Заметим, что пуассоновская случайная величина принимает только целочисленные неотрицательные значения с положительными вероятностями. Поэтому для исследования ее вероятностных свойств и вычисления числовых характеристик естественно применять аппарат производящих функций. Получим теперь производящую функцию для случайной величины ξ с распределением Пуассона. Так как при фиксированном $m = 0, 1, \ldots$ вероятность $p_m = (\lambda^m \exp\{-\lambda\})/m!$, то производящая функция для закона Пуассона равна

$$\Pi_\xi(z) = \exp\{-\lambda\} \sum_{m=0}^{\infty} z^m \lambda^m / m! = \exp\{-\lambda(1-z)\}.$$

Отсюда для любого фиксированного $k = 1, 2, \ldots$ имеем:

$$(d^k \Pi_\xi(z)/dz^k)|_{z=1} = (d^k \exp\{-\lambda(1-z)\}/dz^k)|_{z=1} = \lambda^k.$$

Используя это равенство для производных и соотношение (10.1), получим начальные моменты $\alpha_1\xi$, $\alpha_2\xi$, $\alpha_3\xi$, $\alpha_4\xi$ пуассоновской случайной величины и, естественно, тот же результат для математического ожидания $\mathsf{M}\xi$ и дисперсии $\mathsf{D}\xi$, а именно: $\alpha_1\xi = \mathsf{M}\xi = \lambda$, $\alpha_2\xi = \lambda^2 + \lambda$, $\mathsf{D}\xi = \lambda^2 + \lambda - \lambda^2 = \lambda$, $\alpha_3\xi = \lambda^3 + 3\lambda^2 + \lambda$, $\alpha_4\xi = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$. Найдем третий центральный момент $\beta_3\xi$, коэффициент асимметрии $\mathsf{Ka}\,\xi$, четвертый центральный момент $\beta_4\xi$ и эксцесс $\mathsf{Э}\xi$ для закона Пуассона:

$$\beta_3\xi = \sum_{k=0}^{3} C_3^k (-1)^{3-k} (\alpha_1\xi)^{3-k} \alpha_k\xi = \lambda,$$

$$\mathsf{Ka}\,\xi = \beta_3\xi / (\boldsymbol{\sigma}\xi)^3 = \lambda^{-1/2},$$

$$\beta_4\xi = \sum_{k=0}^{4} C_4^k (-1)^{4-k} (\alpha_1\xi)^{4-k} \alpha_k\xi = 3\lambda^2 + \lambda,$$

$$\mathsf{Э}\xi = (\boldsymbol{\sigma}\xi)^{-4} \beta_4\xi - 3 = \lambda^{-1}.$$

Применяя инструмент производящих функций, покажем, что сумма $\xi_1 + \xi_2 = \xi$ двух независимых пуассоновских случайных величин ξ_1 и ξ_2 с параметрами λ_1 и λ_2 соответственно распределена по закону Пуассона с параметром $\lambda = \lambda_1 + \lambda_2$. Так как случайные величины ξ_1 и ξ_2 являются независимыми, то, используя утверждение леммы 10.1 и вид производящих функций $\Pi_{\xi_1}(z) = \exp\{-\lambda_1(1-z)\}$, $\Pi_{\xi_2}(z) = \exp\{-\lambda_2(1-z)\}$ для величин ξ_1 и ξ_2, найдем:

$$\Pi_\xi(z) = \Pi_{\xi_1}(z) \times \Pi_{\xi_2}(z) = \exp\{-\lambda_1(1-z)\} \times \exp\{-\lambda_2(1-z)\} =$$
$$= \exp\{-(\lambda_1+\lambda_2)(1-z)\} = \exp\{-(\lambda_1+\lambda_2)\} \sum_{m=0}^{\infty} (m!)^{-1} z^m (\lambda_1+\lambda_2)^m.$$

Следовательно, $\mathsf{P}(\{\omega\colon \xi(\omega) = m\}) = (m!)^{-1}(\lambda_1+\lambda_2)^m \exp\{-(\lambda_1+\lambda_2)\}$ для всех $m = 0, 1, \ldots, n$. Итак, случайная величина $\xi = \xi_1 + \xi_2$ распределена по закону Пуассона с параметром $\lambda = \lambda_1 + \lambda_2$. Это свойство легко обобщается на случай конечного числа независимых пуассоновских случайных величин.

Математические модели большого числа физических, экономических, технических и социальных явлений приводят к рассмотрению пуассоновских случайных величин. Приведем пример такой ситуации.

Пример 10.2 (*задача о свободном движении машин по магистрали*). Поучительной иллюстрацией удачного применения закона Пуассона может служить задача об оптимальном управлении транспортом на перекрестке, который расположен при въезде машин в большой город. Условием оптимального управления, как правило, является достижение минимума средних задержек автомобилей перед стоп-линией данного перекрестка за некоторый промежуток времени. Для решения этой трудной задачи, прежде всего, необходимо изучить вероятностные свойства движения потока машин по магистрали. Будем предполагать, что на некотором промежутке времени погодные и дорожные условия обеспечивают:

— во-первых, свободное движение каждой машины по магистрали без препятствий со стороны других автомобилей;

— во-вторых, одиночное движение машин по магистрали, когда исключается групповое следование автомобилей за лидером;

— в-третьих, стационарное движение машин, при котором статистические и, естественно, средние характеристики транспортного потока (например, интенсивность движения) не изменяются во времени.

Как правило, первые два ограничения хорошо соответствуют эмпирическим наблюдениям при относительно слабой интенсивности движения машин по магистрали. Третье ограничение наиболее обременительно, так, например, интенсивность движения машин по магистрали меняется по часам суток, дням недели и месяцам года. Однако с большой степенью достоверности последнее ограничение выполняется в течение некоторого относительно короткого промежутка времени суток. В частности, среднее число поступлений машин в единицу времени (например, в секунду) на перекресток будет постоянной величиной. Ради простоты обозначим такой промежуток через $[0,T]$, где величина T для реальных задач равна около трех часов. Пусть теперь $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ является вероятностной моделью эволюционного эксперимента $E = \{E_t\colon 0 \leqslant t \leqslant T\}$, связанного с наблюдением во времени t процесса движения машин по магистрали и процесса пересечения ими перекрестка. Отметим, что в этой задаче мы не уточняем очень сложную структуру вероятностной модели $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, так как будем изучать только вероятностные свойства количества пересечений машинами стоп-линии перекрестка за некоторые промежутки времени. С этой целью обозначим через $\xi(t)$ случайное число поступлений машин на перекресток за промежуток времени $[0,t)$, где $0 < t \leqslant T$. При этом машина, которая возможно поступит на перекресток в момент t, случайной величиной $\xi(t)$ не учитывается и регистрируется только в следующем промежутке. Значит, при любых $0 \leqslant t_1 < t_2 \leqslant T$ случайная величина $\xi(t_2) - \xi(t_1)$ определяет количество поступивших машин за $[t_1, t_2)$. Для определенности можно положить, что $\xi(0) \equiv 0$. Тогда с помощью семейства $\{\xi(t)\colon 0 \leqslant t \leqslant T\}$ случайных величин, которые заданы на основном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$, указанное движение машин по магистрали математически можно формализовать следующим образом:

1) для любых конечных целых $k > 1$ и $0 < t_1 < t_2 < \ldots < t_k \leqslant T$ величины $\xi(t_1)$, $\xi(t_2) - \xi(t_1), \ldots$, $\xi(t_k) - \xi(t_{k-1})$ независимы в совокупности;

2) при малом $\Delta t > 0$ вероятности $\mathsf{P}(\{\omega\colon \xi(t+\Delta t) - \xi(t) = 0\}) = 1 - \lambda\Delta t + o(\Delta t)$ и $\mathsf{P}(\{\omega\colon \xi(t+\Delta t) - \xi(t) = 1\}) = \lambda\Delta t + o(\Delta t)$, где $0 \leqslant t \leqslant T - \Delta t$, λ — положительная константа и $o(\Delta t)$ есть бесконечно малая величина по сравнению с Δt при $\Delta t \to 0$;

3) при любых $T \geqslant t_2 > t_1 \geqslant 0$ вероятность $\mathsf{P}(\{\omega\colon \xi(t_2) - \xi(t_1) = m\})$ не зависит от t_1, а зависит только от разности $t_2 - t_1$ и m, т. е. имеем равенство $\mathsf{P}(\{\omega\colon \xi(t_2) - \xi(t_1) = m\}) = \mathsf{P}(\{\omega\colon \xi(t_2 - t_1) = m\})$ для $m \geqslant 0$.

Так как $\sum_{m=0}^{\infty} \mathsf{P}(\{\omega\colon \xi(t+\Delta t) - \xi(t) = m\}) = 1$ при $t \geqslant 0$ и $\Delta t > 0$, то из условий $o(\Delta t) \geqslant 0$, 2) и 3) вытекает, что

$$\mathsf{P}(\{\omega\colon \xi(t+\Delta t) - \xi(t) \geqslant 2\}) = \mathsf{P}(\{\omega\colon \xi(\Delta t) \geqslant 2\}) = o(\Delta t). \tag{10.3}$$

Покажем теперь, что при $t > 0$ и $m = 0, 1, \ldots$ вероятность $\mathsf{P}(\{\omega\colon \xi(t) = m\})$, которую будем обозначать через $p(t,m)$, равна $(\lambda t)^m (m!)^{-1} \exp\{-\lambda t\}$. Действительно, имеем:

$$\{\omega\colon \xi(t+\Delta t) = 0\} = \{\omega\colon \xi(t) = 0\} \cap \{\omega\colon \xi(t+\Delta t) - \xi(t) = 0\},$$

$$\{\omega\colon \xi(t+\Delta t) = m\} = \\ = \bigcup_{k=0}^{m} (\{\omega\colon \xi(t) = k\} \cap \{\omega\colon \xi(t+\Delta t) - \xi(t) = m-k\}), \quad m \geqslant 1.$$

Из этих равенств в событиях непосредственно получаем равенства в вероятностях:

$$\mathsf{P}(\{\omega\colon \xi(t+\Delta t)=0\})=\mathsf{P}(\{\omega\colon \xi(t)=0, \xi(t+\Delta t)-\xi(t)=0\}),$$
$$\mathsf{P}(\{\omega\colon \xi(t+\Delta t)=m\})=$$
$$=\sum_{k=0}^{m}\mathsf{P}(\{\omega\colon \xi(t)=k, \xi(t+\Delta t)-\xi(t)=m-k\}),\quad m\geqslant 1.$$

Используя условия 1)–3), соотношение (10.3) и обозначение $p(t,m)$, последовательно получим

$$p(t+\Delta t,0)=p(t,0)\,\mathsf{P}(\{\omega\colon \xi(t+\Delta t)-\xi(t)=0\}),$$
$$p(t+\Delta t,m)=\sum_{k=0}^{m}p(t,k)\,\mathsf{P}(\{\omega\colon \xi(t+\Delta t)-\xi(t)=m-k\}),\quad m\geqslant 1;$$
$$p(t+\Delta t,0)=p(t,0)\,\mathsf{P}(\{\omega\colon \xi(\Delta t)=0\}),$$
$$p(t+\Delta t,m)=\sum_{k=0}^{m}p(t,k)\,\mathsf{P}(\{\omega\colon \xi(\Delta t)=m-k\}),\quad m\geqslant 1;$$
$$p(t+\Delta t,0)=p(t,0)\,p(\Delta t,0),$$
$$p(t+\Delta t,m)=\sum_{k=0}^{m}p(t,k)\,p(\Delta t,m-k),\quad m\geqslant 1;$$
$$p(t+\Delta t,0)=p(t,0)\,(1-\lambda\Delta t+o(\Delta t)),$$
$$p(t+\Delta t,m)=p(t,m)\,(1-\lambda\Delta t+o(\Delta t))+$$
$$+\,p(t,m-1)(\lambda\Delta t+o(\Delta t))+o(\Delta t),\quad m\geqslant 1.$$

Из последней системы равенств в пределе при $\Delta t\to 0$ и $t>0$ получаем бесконечную систему линейных дифференциальных уравнений $dp(t,0)/dt=-\lambda p(t,0)$, $dp(t,m)/dt=-\lambda p(t,m)+\lambda p(t,m-1)$, $m\geqslant 1$, с начальными условиями $p(0,0)=1$, $p(0,m)=0$, $m\geqslant 1$. Умножая m-е уравнение этой системы на z^m и суммируя по $m=0,1,\ldots$, легко найдем дифференциальное уравнение $\partial\Pi_{\xi(t)}(t,z)/\partial t=\lambda(z-1)\Pi_{\xi(t)}(t,z)$ для производящей функции $\Pi_{\xi(t)}(t,z)=\sum_{m=0}^{\infty}z^m p(t,m)$ случайной величины $\xi(t)$. Если записать это дифференциальное уравнение в виде $\partial\ln\Pi_{\xi(t)}(t,z)/\partial t=\lambda(z-1)$ и учесть начальное условие $\Pi_{\xi(t)}(0,z)=p(0,0)=1$, то получим: $\Pi_{\xi(t)}(t,z)=\exp\{\lambda t(z-1)\}=\exp\{-\lambda t\}\exp\{\lambda tz\}=$ $=\exp\{-\lambda t\}\sum_{m=0}^{\infty}(\lambda tz)^m(m!)^{-1}$. Вспоминая определение производящей функции $\Pi_{\xi(t)}(t,z)$ для случайной величины $\xi(t)$, мы без труда видим, что $p(t,m)=(\lambda t)^m(m!)^{-1}\exp\{-\lambda t\}=\pi(m;\lambda t)$. Итак, число поступлений машин на перекресток за промежуток времени $[0,t)$ распределено по закону Пуассона с параметром λt. Поэтому $\mathsf{M}\xi(t)=\lambda t$

и, значит, среднее число поступивших машин на перекресток за единицу времени или интенсивность потока движения машин по магистрали равна λ. В силу ограничения 3) найдем распределение величины $\xi(t_2) - \xi(t_1)$ числа поступлений за время $[t_1, t_2)$ по формуле вида $\mathsf{P}(\{\omega\colon \xi(t_2) - \xi(t_1) = m\}) = \mathsf{P}(\{\omega\colon \xi(t_2 - t_1) = m\}) = (\lambda(t_2 - t_1))^m (m!)^{-1} \exp\{-\lambda(t_2 - t_1)\}$, $m = 0, 1, \ldots$.

Эта формула была получена при выполнении ограничений 1), 2) и 3). Покажем, что, в некотором смысле, имеет место обратное утверждение. Пусть теперь имеет место ограничение 1), а вместо ограничений 2) и 3) потребуем выполнение равенства $\mathsf{P}(\{\omega\colon \xi(t_2 - t_1) = m\}) = (\lambda t)^m (m!)^{-1} \exp\{-\lambda t\}$ при любом $t > 0$. Если теперь при любых $0 \leqslant t_1 < t_2 \leqslant T$ случайная величина $\xi(t_1, t_2) = \xi(t_2) - \xi(t_1)$ определяет количество поступивших машин за время $[t_1, t_2)$, то справедливо очевидное равенство $\xi(t_2) = \xi(t_1) + \xi(t_1, t_2)$. В силу ограничения 1) для движения машин по магистрали случайные величины $\xi(t_1)$ и $\xi(t_1, t_2)$ будут независимыми. Тогда на основании леммы 10.1 для производящих функций $\Pi_{\xi(t_2)}(z)$, $\Pi_{\xi(t_1)}(z)$ и $\Pi_{\xi(t_1,t_2)}(z)$ случайных величин $\xi(t_2)$, $\xi(t_1)$ и $\xi(t_1, t_2)$ получим, что $\Pi_{\xi(t_2)}(z) = \Pi_{\xi(t_1)}(z)\,\Pi_{\xi(t_1,t_2)}(z)$. Отсюда, используя полученные равенства $\Pi_{\xi(t_2)}(z) = \exp\{\lambda t_2(z-1)\}$ и $\Pi_{\xi(t_1)}(z) = \exp\{\lambda t_1(z-1)\}$, найдем $\Pi_{\xi(t_1,t_2)}(z) = \exp\{\lambda(t_2 - t_1)(z-1)\}$. Следовательно, количество поступивших машин на перекресток за промежуток времени $[t_1, t_2)$ является пуассоновской случайной величиной с параметром $\lambda(t_2 - t_1)$. В этом случае будем говорить, что транспортный поток машин на магистрали является пуассоновским процессом. Теперь с полным основанием можно считать ограничения 1), 2) и 3) ограничениями пуассоновского процесса $\{\xi(t)\colon 0 \leqslant t \leqslant T\}$, который может иметь самый разнообразный физический смысл. Приведем пример такого факта.

Пример 10.3. Среднее число вызовов, поступающих на АТС в одну минуту, равно двум. Поток вызовов является пуассоновским. Используя приложение 1 из [5], найти вероятность того, что за 5 минут поступят: а) ровно два вызова; б) менее двух вызовов; в) не менее двух вызовов.

Решение. По условию задачи среднее число вызовов, поступающих за одну минуту, есть $\lambda = 2$. Следовательно, для этого опыта при любом $t_1 \geqslant 0$ и $t_2 = t_1 + 5$ случайная величина $\xi(t_1, t_1 + 5)$, которая считает случайное число вызовов за 5 минут, распределена по закону Пуассона с параметром $\lambda(t_2 - t_1) = 10$. Отсюда получаем, что:

а) $\mathsf{P}(\xi(t_1, t_1 + 5) = 2) = \pi(2; 10) = (10^2 \exp\{-10\})/2! = 0{,}002270$;

б) $\mathsf{P}(\xi(t_1, t_1 + 5) < 2) = \pi(0; 10) + \pi(1; 10) = \exp\{-10\} + (10^1 \exp\{-10\})/1! = 0{,}000495$;

в) $\mathsf{P}(\xi(t_1, t_1 + 5) \geqslant 2) = 1 - \mathsf{P}(\xi(t_1, t_1 + 5) < 2) = 1 - 0{,}000495 = 0{,}999505$.

1.3. Дискретная случайная величина с распределением по закону Бартлетта. Рассмотрим дискретную случайную величину ξ, множество значений которой совпадает с множеством $\{1, 2, \ldots\}$ натуральных чисел.

Определение 10.4. Будем говорить, что *дискретная случайная величина распределена по закону Бартлетта*, если ее распределение равно

$$\mathsf{P}(\{\omega\colon \xi(\omega) = 1\}) = 1 - p, \quad \mathsf{P}(\{\omega\colon \xi(\omega) = m\}) = p\,(1-q)\,q^{m-2}, \quad m \geqslant 2, \tag{10.4}$$

где параметры p и q удовлетворяют условию $0 < p, q < 1$.

Сумма вероятностей возможных значений этой случайной величины равна $1 - p + \sum_{m=2}^{\infty} p\,(1-q)\,q^{m-2} = 1 - p + p\,(1-q)\,(1-q)^{-1} = 1$. Приведем сначала несколько примеров, в которых для изучения вероятностных свойств некоторых количественных признаков реальных экспериментов используются случайные величины с распределением Бартлетта (10.4).

Пример 10.4. Самой характерной особенностью современной управляющей системы является ее сложность. Как правило, сложность означает, что управляющая система состоит из большого числа составляющих ее элементов. При функционировании управляющей системы возможны сбои (отказы) отдельных ее элементов, которые имеют случайный характер. Любой вышедший из строя элемент заменяется новым из некоторого резерва. По признаку влияния на надежность все элементы системы разделяются на главные и второстепенные. При неблагоприятных и экстремальных условиях внешней среды выход из строя одного из главных элементов существенно уменьшает надежность системы и порождает эпидемию отказов второстепенных элементов. Изучение вероятностных свойств случайного числа таких отказов позволяет решить проблему профилактики системы с целью повышения ее надежности. Для многих управляющих систем случайное число отказов, которое порождается одним из главных элементов, распределено по закону Бартлетта. При этом, естественно, учитывается отказ такого главного элемента.

Пример 10.5. Начиная с середины XX в., мир шагнул в сферу электронных финансовых расчетов с помощью сети банкоматов, системы авторизации и индивидуальных пластиковых карт. Используя технологию электронных надписей на счетах клиентов различных банков и инструменты удаленного компьютерного доступа к этим счетам, общество получает существенную экономию времени. В связи с этим возникает трудная задача определения распределения случайного числа запросов (транзакций) на обслуживание клиентов за некоторые случайные промежутки времени. Эти промежутки времени задаются последовательными моментами поступления в систему авторизации запросов специального вида. Такие запросы называются стробирующими. Случайное число всех типов транзакций, поступивших в систему авторизации за любой промежуток времени между двумя последовательными стробирующими запросами, распределено по закону Бартлетта. Решение этой задачи позволяет организовать оптимальную работу центра авторизации по критерию экономии времени доступа к личным счетам.

Пример 10.6. Распределение Бартлетта успешно используется [20] для определения вероятностных свойств при нелокальном описании:

а) потока ошибок при передаче сообщений по телефону в течение достаточно большого промежутка времени;

б) потока запросов пациентов на пункт скорой помощи во время эпидемии некоторого заболевания в большом городе;

в) потока лыжников во время гонки на сверхдлинную дистанцию по неоднородной трассе, и т. д.

Вернемся теперь к изучению свойств случайной величины Бартлетта. Из определения случайной величины Бартлетта следует, что ее интегральная функция $F(x)$ равна нулю при $x \leqslant 1$, равна $1-p$ при $1 < x \leqslant 2$ и, наконец, равна $1 - p + \sum_{2 \leqslant m < x} p\,(1-q)\,q^{m-2}$ при $x > 2$. Вычислим теперь значение интегральной функции $F(x)$ при $x = 2, 3, \ldots$. Из общей формулы для интегральной функции закона Бартлетта имеем: $F(k+1) = 1 - p + \sum_{2 \leqslant m < k+1} p\,(1-q)\,q^{m-2} = 1 - {}$ $- p + \sum_{m=2}^{k} p\,(1-q)\,q^{m-2} = 1 - p + p\,(1-q)\,(1 + q + q^2 + \ldots + q^{k-2}) = {}$ $= 1 - p + p\,(1 - q^{k-1}) = 1 - pq^{k-1}$, $k = 1, 2, \ldots$. Ряд распределения случайной величины Бартлетта имеет вид, представленный в табл. 10.4.

Таблица 10.4

ξ	1	2	3	…	m	…
$\mathsf{P}(\cdot)$	$1-p$	$p\,(1-q)$	$p\,(1-q)q$	…	$p\,(1-q)q^{m-2}$	…

Производящую функцию для случайной величины Бартлетта можно вычислить следующим образом:

$$\Pi_\xi(z) = z(1-p) + \sum_{m=2}^{\infty} z^m p\,(1-q)\,q^{m-2} = p\,(1-q)(1-qz)^{-1} z^2 + (1-p)\,z. \tag{10.5}$$

Дифференцируя четыре раза по z эту функцию и полагая затем $z = 1$ в полученных равенствах для производных, найдем $(d\Pi_\xi(z)/dz)\big|_{z=1} = 1 + p\,(1-q)^{-1}$, $(d^2\Pi_\xi(z)/dz^2)\big|_{z=1} = 2\,p\,(1-q)^{-2}$, $(d^3\Pi_\xi(z)/dz^3)\big|_{z=1} = 6\,pq\,(1-q)^{-3}$, $(d^4\Pi_\xi(z)/dz^4)\big|_{z=1} = 24\,pq^2(1-q)^{-4}$. Отсюда, учитывая соотношение (10.1), получим простые формулы для вычисления начальных моментов $\alpha_1\xi = \mathsf{M}\xi$, $\alpha_2\xi$, $\alpha_3\xi$, $\alpha_4\xi$, дисперсии $\mathsf{D}\xi = \beta_2\xi$, среднего квадратического отклонения $\boldsymbol{\sigma}(\xi)$, центральных моментов $\beta_3\xi$, $\beta_4\xi$, коэффициента асимметрии $\mathsf{Ka}\,\xi$ и, наконец, эксцесса $\mathsf{Э}\xi$ случайной величины Бартлетта:

$$\begin{aligned}
\alpha_1\xi &= \mathsf{M}\xi = p\,(1-q)^{-1} + 1,\\
\alpha_2\xi &= 2p\,(1-q)^{-2} + p\,(1-q)^{-1} + 1,\\
\alpha_3\xi &= 6pq(1-q)^{-3} + 6p\,(1-q)^{-2} + p\,(1-q)^{-1} + 1,\\
\alpha_4\xi &= 24pq^2(1-q)^{-4} + 36pq(1-q)^{-3} + 14p\,(1-q)^{-2} + p\,(1-q)^{-1} + 1,\\
\mathsf{D}\xi &= \beta_2\xi = p\,(1+q-p)(1-q)^{-2},\\
\boldsymbol{\sigma}\xi &= p^{1/2}(1+q-p)^{1/2}(1-q)^{-1},\\
\beta_3\xi &= p\,(2p^2 - 3(1+q)\,p + q^2 + 4q + 1)(1-q)^{-3},\\
\beta_4\xi &= p\,(1-q)^{-4}\times\\
&\quad\times(-3p^3 + 6p^2(q+1) - 4p\,(q^2+4q+1) + q^3 + 11q^2 + 11q + 1),\\
\mathsf{Э}\xi &= p^{-1}(1+q-p)^{-2}\times\\
&\quad\times(-6p^3 + 12p^2(q+1) - p\,(7q^2+22q+7) + q^3 + 11q^2 + 11q + 1),\\
\mathsf{Ka}\,\xi &= (2p^2 - 3(1+q)\,p + q^2 + 4q + 1)(1+q-p)^{-3/2}p^{-1/2}.
\end{aligned} \tag{10.6}$$

Используя непосредственно ряд распределения случайной величины Бартлетта, найдем моду $\mathsf{Mo}\,\xi$ и наивероятнейшее значение $\mathsf{Mo}^*\xi$. Легко проверить, что при $p \leqslant 2^{-1}$ или при $p > 2^{-1}$ и $q > 2 - p^{-1}$ имеет место соотношение $\mathsf{P}(\{\omega\colon \xi(\omega) = m-1\}) > \mathsf{P}(\{\omega\colon \xi(\omega) = m\})$ для всех $m = 2, 3, \ldots$. Значит, в этих двух случаях наивероятнейшее значение $\mathsf{Mo}^*\xi$ случайной величины Бартлетта равно единице, а мода $\mathsf{Mo}\,\xi$ не существует. Далее, при $p > 2^{-1}$ и $q < 2 - p^{-1}$ выполняются следующие неравенства: $\mathsf{P}(\{\omega\colon \xi(\omega) = 1\}) < \mathsf{P}(\{\omega\colon \xi(\omega) = 2\})$, $\mathsf{P}(\{\omega\colon \xi(\omega) = m\}) > \mathsf{P}(\{\omega\colon \xi(\omega) = m+1\})$, $m = 2, 3, \ldots$. В этом случае наивероятнейшее значение и мода случайной величины Бартлетта совпадают и равны двум. Наконец, если $p > 2^{-1}$ и $q = 2 - p^{-1}$, то $\mathsf{P}(\{\omega\colon \xi(\omega) = 1\}) = \mathsf{P}(\{\omega\colon \xi(\omega) = 2\}) > \mathsf{P}(\{\omega\colon \xi(\omega) = 3\}) > \ldots$. Поэтому $\mathsf{Mo}\,\xi = 2$ и $\mathsf{Mo}^*\xi \in \{1, 2\}$. В последних двух случаях распределение случайной величины Бартлетта является унимодальным.

Если параметры p и q совпадают, то ряд распределения случайной величины Бартлетта примет следующий вид (табл. 10.5).

Таблица 10.5

ξ	1	2	3	…	m	…
$\mathsf{P}(\cdot)$	$1-p$	$(1-p)\,p$	$(1-p)\,p^2$	…	$(1-p)\,p^{m-1}$	…

Случайная величина с таким рядом распределения в теории вероятностей называется *геометрической*. Случайная величина имеет такое название потому, что элементы нижней строки табл. 10.5 образуют убывающую геометрическую прогрессию со знаменателем $0 < p < 1$. Сумма такой прогрессии, очевидно, равна единице. Итак, при $p = q$ распределение Бартлетта совпадает с геометрическим. Геометрическое распределение появляется, например, в следующей хорошо известной задаче. Рассмотрим произвольный статистически устойчивый опыт E_0, в котором интересуемся только исходом A (успехом) и противоположным исходом $\overline{A}$ (неудачей). При этом успех достигается с вероятностью $\mathsf{P}(A) = p$, а неудача появляется с вероятностью $\mathsf{P}(\overline{A}) = 1 - p$. Пусть опыт E_0 независимым образом повторяется до появления первой неудачи. Ясно, что опыт E_0 будет проведен всего один раз с вероятностью $1 - p$, или два раза с вероятностью $p\,(1 - p)$, или три раза с вероятностью $p^2(1 - p)$ и т. д. Тогда случайное число ξ всех испытаний до появления первой неудачи, включая испытание с первой неудачей, имеет геометрическое распределение. Так как опыт E_0 является произвольным, то случайная величина с геометрическим распределением встречается довольно часто на практике, и поэтому является важным частным вариантом случайной величины Бартлетта.

Перейдем к изучению свойств геометрической случайной величины.

Среди всех дискретных случайных величин только величина ξ с геометрическим распределением обладает уникальным свойством:

$$
\begin{aligned}
&\mathsf{P}(\{\omega\colon \xi(\omega) = m + k\} \mid \{\omega\colon \xi(\omega) > k\}) = \\
&\qquad = \mathsf{P}(\{\omega\colon \xi(\omega) = m + k, \xi(\omega) > k\})(\mathsf{P}(\{\omega\colon \xi(\omega) > k\}))^{-1} = \\
&\qquad = \mathsf{P}(\{\omega\colon \xi(\omega) = m + k\})(\mathsf{P}(\{\omega\colon \xi(\omega) > k\}))^{-1} = \\
&\qquad = p^{m+k-1}(1 - p)\Big(\sum_{n=k+1}^{\infty}(1 - p)\,p^{n-1}\Big)^{-1} = \\
&\qquad = p^{m+k-1}(1 - p)(p^k)^{-1} = p^{m-1}(1 - p) = \mathsf{P}(\{\omega\colon \xi(\omega) = m\}).
\end{aligned}
$$

Равенство $\mathsf{P}(\{\omega\colon \xi(\omega) = m + k\} \mid \{\omega\colon \xi(\omega) > k\}) = \mathsf{P}(\{\omega\colon \xi(\omega) = m\})$ можно проинтерпретировать следующим образом. Условная вероятность того, что неудача впервые наступит при испытании с номером $m + k$, если все первые k испытаний были успешными, совпадает с безусловной вероятностью наблюдения первой неудачи при испытании с номером m. На содержательном уровне можно сказать, что закон распределения оставшейся части числа испытаний не зависит от числа уже проведенных опытов. В теории вероятностей это свойство называется отсутствием последействия или памяти.

Интегральная функция распределения геометрической случайной величины равна $F(x) = \sum_{m<x} p^{m-1}(1 - p)$. Если же $x = k + 1$, то $F(k + 1) = \sum_{m=1}^{k} p^{m-1}(1 - p) = (1 - p)(1 + p + p^2 + \ldots + p^{k-1}) = 1 - p^k$ для всех $k = 0, 1, \ldots$. Непосредственно из равенств (10.5) и (10.6)

при $p = q$ получим для случайной величины ξ с геометрическим распределением производящую функцию $\Pi_\xi(z) = p\,(1-q)(1-qz)^{-1}z^2 + (1-p)z = (1-p)z(1-pz)^{-1}$ и формулы для вычисления начальных моментов $\alpha_1\xi = \mathsf{M}\xi$, $\alpha_2\xi$, $\alpha_3\xi$, $\alpha_4\xi$, дисперсии $\mathsf{D}\xi = \beta_2\xi$, среднего квадратического отклонения $\boldsymbol{\sigma}\xi$, центральных моментов $\beta_3\xi$, $\beta_4\xi$, коэффициента асимметрии $\mathsf{Ka}\,\xi$ и, наконец, эксцесса $\mathsf{Э}\xi$:

$$
\begin{aligned}
&\alpha_1\xi = \mathsf{M}\xi = (1-p)^{-1}, && \alpha_2\xi = (1+p)(1-p)^{-2},\\
&\alpha_3\xi = (p^2+4p+1)(1-p)^{-3}, && \alpha_4\xi = (p^3+11p^2+11p+1)(1-p)^{-4},\\
&\mathsf{D}\xi = \beta_2\xi = p\,(1-p)^{-2}, && \boldsymbol{\sigma}\xi = p^{1/2}(1-p)^{-1},\\
&\beta_3\xi = p\,(1+p)(1-p)^{-3}, && \beta_4\xi = p\,(p^2+7p+1)(1-p)^{-4},\\
&\mathsf{Ka}\,\xi = (p+1)\,p^{-1/2}, && \mathsf{Э}\xi = (p^2+4p+1)\,p^{-1}.
\end{aligned}
\tag{10.7}
$$

Так как геометрическое распределение получается из распределения Бартлетта при $p = q$ и выполняется неравенство $p > 2 - p^{-1}$, то для геометрической случайной величины мода $\mathsf{Mo}\,\xi$ не существует, а наивероятнейшее значение $\mathsf{Mo}^*\xi = 1$. Рассмотрим типичный пример вопросов, которые возникают при изучении величины с геометрическим распределением.

Пример 10.7. Бригада рабочих ежедневно и только один раз на вертолете облетает трассу нефтепровода с целью обнаружения возможного его повреждения и последующего оперативного и быстрого ремонта. За сутки на нефтепроводе может произойти не более одного повреждения. Вероятность возникновения повреждения нефтепровода в течение каждых суток равна 0,01 и не зависит от последовательности предыдущих аварий. Найти вероятностные законы случайного числа ξ всех вылетов бригады до первого обнаружения повреждения. При этом ξ учитывает вылет, при котором было обнаружено повреждение. Вычислить вероятность обнаружения первой аварии на нефтепроводе в течение года. Определить основные числовые характеристики случайной величины ξ.

Решение. Событие $\{\omega\colon \xi(\omega) = 1\}$ означает, что авария будет обнаружена сразу при первом вылете. По условию задачи $\mathsf{P}(\{\omega\colon \xi(\omega) = 1\}) = 0{,}01$. При $m = 2, 3, \ldots$ событие $\{\omega\colon \xi(\omega) = m\}$ произойдет, если в течение первых $m-1$ дней зафиксировано рабочее состояние нефтепровода, а при вылете в следующий день был обнаружен дефект. В силу условий независимости возникновения повреждений на данном нефтепроводе вероятность $\mathsf{P}(\{\omega\colon \xi(\omega) = m\})$ равна $(0{,}99)^{m-1} \times 0{,}01$ для всех $m \geqslant 2$. Значит, для $\xi(\omega)$ получили геометрическое распределение с параметром $p = 0{,}99$. Воспользовавшись этим, представим интегральную функцию распределения в виде $F(x) = 1 - (0{,}99)^k$ при $k < x \leqslant k+1$ и $k = 0, 1, \ldots$. Отсюда, если календарный год содержит 365 дней, то вероятность обнаружения первой аварии в течение года, очевидно, равна $\mathsf{P}(\{\omega\colon 1 \leqslant \xi(\omega) < 366\}) = F(366) = 1 - (0{,}99)^{365} \approx 1 - 0{,}026 = 0{,}974$. Интегральная функция $F(x)$, ряд распределения и многоугольник распределения для этой случайной величины изображены на рисунках 10.2 и 10.3 и в виде табл. 10.6.

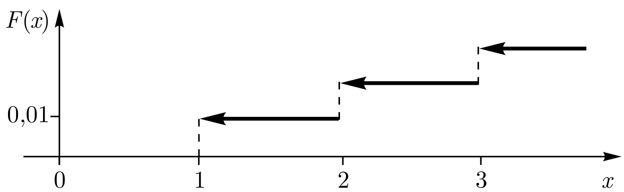


Рис. 10.2

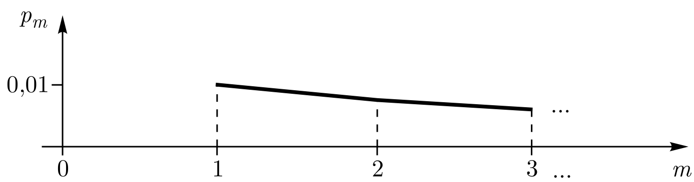


Рис. 10.3

Таблица 10.6

ξ	1	2	3	$\ldots$	m	$\ldots$
$\mathsf{P}(\cdot)$	0,01	0,0099	0,0098	$\ldots$	$(0{,}99)^{m-1} \times 0{,}01$	$\ldots$

Используя формулы (10.7), вычислим основные числовые характеристики числа всех вылетов бригады до первого обнаружения повреждения на трассе: $\mathsf{M}\xi = (1-p)^{-1} = 100$, $\mathsf{D}\xi = p\,(1-p)^{-2} = 9900$, $\boldsymbol{\sigma}\xi \approx 99{,}4987$, $\mathsf{Ka}\,\xi = (p+1)\,p^{-1/2} \approx 2$, $\mathsf{Э}\xi = (p^2 + 4\,p + 1)\,p^{-1} \approx 6$.

На примере этой модельной задачи можно пояснить, когда те или иные события реального эксперимента объявляются практически достоверными или практически невозможными. Вероятность 0,01 возникновения повреждения нефтепровода в течение каждых суток, на первый взгляд, нам кажется весьма малой. Однако обнаружение первой аварии на нефтепроводе в течение года является важным событием с точки зрения возможных последствий для экологии. В силу этого вероятность обнаружения первой аварии на нефтепроводе в течение года, которая равна $1 - (0{,}99)^{365}$, целесообразно считать близкой к единице. Итак, на содержательном уровне можно сказать, что событие обнаружения первой аварии на нефтепроводе в течение года почти обязательно происходит или является практически достоверным, а событие возникновения повреждения нефтепровода в течение каждых суток нельзя считать практически невозможным. Далее, рассмотрим другую задачу с аналогичными численными данными, а именно, задачу о маршрутном такси. В Нижнем Новгороде на остановках указывается интервал движения маршрутного такси. Предположим, что маршрутное такси нарушает расписание в течение каждых суток с вероятностью 0,01. Тогда вероятность появления первого сбоя в расписании маршрутного такси в течение года, приблизительно равная 0,974, с точки зрения практической важности не является катастрофически большой. Значит, вероятность 0,01 нарушения в расписании маршрутным такси в течение каждых суток можно считать малой или, другими словами, событие ежесуточного возникновения сбоя в расписании является для пассажиров практически невозможным. Вернемся снова к задаче о повреждении на нефтепроводе. Пусть теперь за счет повышения качества изготовления труб удалось значительно уменьшить вероятность возникновения повреждения нефтепровода в течение каждых суток до величины 0,001. Тогда вероятность обнаружения первой аварии на таком нефтепроводе в течение года равна $1 - (0{,}999)^{365} \approx 0{,}329$. В этом случае функционирование нефтепровода на практике считается вполне надежной. Поэтому вероятность 0,001 возникновения повреждения нефтепровода в течение каждых суток следует признать малой или, иначе, событие возникновения повреждения нефтепровода в течение каждых суток можно считать практически невозможным.

Для иллюстрации более тесной связи между законом распределения Бартлетта и геометрическим законом распределения решим следующую задачу, которая часто встречается на практике.

Пример 10.8 (*задача о групповом движении машин по магистрали*). При удовлетворительном состоянии дорожного полотна и хороших метеорологических условиях движение автомобилей по магистрали может оказаться беспрепятственным и пуассоновским. При плохих погодных условиях (туман, снег, гололед и т. д.) обгон быстрыми машинами медленных является уже рискованным, зависимым и занимает значительное время. В этом случае на интенсивных магистралях будут возникать автоколонны (группы) машин или транспортные пачки, т. е. транспортные потоки уже не будут пуассоновскими. С такой ситуацией впервые столкнулся в 1963 г. Бартлетт при наблюдении за движением машин вблизи Лондона. Для такого типа транспортного потока Бартлетту и другим исследователям не удалось найти подходящего закона распределения для зависимых интервалов между двумя последовательными автомобилями. Автор настоящего учебника приблизительно в это же время решал задачу об оптимальном управлении транспортными потоками в городе Горьком (Нижний Новгород). По наблюдениям за движением машин на магистралях вблизи Горького и других крупных городов было подмечено, что транспортная пачка состоит из головной машины с медленным движением и очереди из быстрых машин, которые догнали медленную и ожидают возможности обгона. В силу этого был предложен следующий механизм образования транспортных пачек при движении машин в плохих погодных условиях.

Для большого числа магистралей оказалось, что быстрые машины поступают в транспортную пачку или, другими словами, догоняют медленную машину по закону Пуассона с интенсивностью $\lambda > 0$. Это означает, что быстрые машины осуществляют относительно свободное движение на протяженных участках дороги, где нет медленных машин. Подробно такой транспортный поток был изучен в задаче о свободном движении автомобилей по магистрали. В соответствии с этой задачей обозначим через $\xi(\omega;\, t,\, \Delta t)$ случайное число быстрых машин, которые поступают по закону Пуассона в транспортную пачку за промежуток времени $[t,\, t + \Delta t)$. Каждую машину с медленным движением можно интерпретировать как обслуживающий прибор для машин с быстрым движением. При этом под временем обслуживания, естественно, понимается случайное время обгона; среднее время обслуживания (обгона) существенно зависит от числа машин в транспортной пачке. Пусть случайная величина $\chi(\omega; t)$ измеряет число всех типов машин в транспортной пачке в момент времени $t \geqslant 0$. Обозначим через $\eta(\omega;\, t,\, \Delta t)$ случайное число быстрых машин, которые совершают обгон медленной машины за промежуток времени $[t,\, t + \Delta t)$. Очевидно, что $0 \leqslant \eta(\omega;\, t,\, \Delta t) \leqslant \chi(\omega; t) + \xi(\omega;\, t,\, \Delta t) - 1$. Вполне естественно предположить, что при малых значениях $\Delta t > 0$ и $k \geqslant 3$ условные вероятности событий, которые порождаются дискретной случайной величиной $\eta(\omega;\, t,\, \Delta t)$, определяются следующими соотношениями:

$$\begin{aligned}
&\mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) = 0\} \mid \{\omega\colon \chi(\omega; t) = 2,\, \xi(\omega;\, t,\, \Delta t) = 0\}) = \\
&\qquad\qquad\qquad\qquad = 1 - \mu_1 \Delta t + o(\Delta t),\\
&\mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) = 0\} \mid \{\omega\colon \chi(\omega; t) = k,\, \xi(\omega;\, t,\, \Delta t) = 0\}) = \\
&\qquad\qquad\qquad\qquad = 1 - \mu_2 \Delta t + o(\Delta t),\\
&\mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) = 1\} \mid \{\omega\colon \chi(\omega; t) = k,\, \xi(\omega;\, t,\, \Delta t) = 0\}) = \\
&\qquad\qquad\qquad\qquad = \mu_2 \Delta t + o(\Delta t).
\end{aligned} \tag{10.8}$$

В равенствах (10.8) каждый из параметров μ_1^{-1} и μ_2^{-1} задает среднее время обгона быстрой машиной медленную в случае, когда транспортная пачка состоит из двух машин и соответственно из трех или большего числа машин. Эти параметры будем называть интенсивностями обгона. Таким способом моделируется зависимость среднего времени обгона от числа машин в транспортной пачке. Из соотношений (10.8), используя при $k \geqslant 1$, $n = 0,\, 1$ условие нормировки для $\eta(\omega;\, t,\, \Delta t)$ вида $\sum_{s=0}^{k+n-1} \mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) = s\} \mid \{\omega\colon \chi(\omega; t) = k,\, \xi(\omega;\, t,\, \Delta t) = n\}) = 1$

и пуассоновское поступление потока быстрых машин в транспортную пачку, получаем при малых значениях $\Delta t > 0$ равенства

$$\begin{aligned}
&\mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) = 0\} \mid \{\omega\colon \chi(\omega; t) = 1,\, \xi(\omega;\, t,\, \Delta t) = 0\}) = 1,\\
&\mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) = 1\} \mid \{\omega\colon \chi(\omega; t) = 2,\, \xi(\omega;\, t,\, \Delta t) = 0\}) =\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = \mu_1 \Delta t + o(\Delta t),\\
&\mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) \geqslant 2\} \mid \{\omega\colon \chi(\omega; t) = k+2,\, \xi(\omega;\, t,\, \Delta t) = 0\}) = o(\Delta t)
\end{aligned} \tag{10.9}$$

и

$$\begin{aligned}
&\mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) = 0\} \mid \{\omega\colon \chi(\omega; t) = k,\, \xi(\omega;\, t,\, \Delta t) = 1\}) = 1 - O(\Delta t),\\
&\mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) = 1\} \mid \{\omega\colon \chi(\omega; t) = k,\, \xi(\omega;\, t,\, \Delta t) = 1\}) = O(\Delta t),\\
&\mathsf{P}(\{\omega\colon \eta(\omega;\, t,\, \Delta t) \geqslant 2\} \mid \{\omega\colon \chi(\omega; t) = k+1,\, \xi(\omega;\, t,\, \Delta t) = 1\}) = o(\Delta t).
\end{aligned} \tag{10.10}$$

Из равенств (10.8)–(10.10) следует, что при заданном размере транспортной пачки условная вероятность того, что за (где угодно расположенный) промежуток Δt по меньшей мере две машины обгонят медленную, есть величина бесконечно малая по сравнению с Δt. Обозначим теперь через $Q(t,k)$ вероятность $\mathsf{P}(\{\omega\colon \chi(\omega;t) = k\})$ при $t \geqslant 0$ и $k = 1, 2, \ldots$.

Рассуждениями, аналогичными тем, которые были приведены в задаче о свободном движении машин по магистрали, мы можем получить бесконечную систему дифференциальных уравнений для вероятностей вида $Q(t,k)$, $t \geqslant 0$, $k = 1, 2, \ldots$. В самом деле, равенства $\{\omega\colon \chi(\omega; t+\Delta t) = m\} = \cup_{k=1}^{\infty} \cup_{n=0}^{\infty} \{\omega\colon \chi(\omega;t) = k,\, \xi(\omega;\, t,\, \Delta t) = n,\, \eta(\omega;\, t,\, \Delta t) = k + n - m\}$, $m \geqslant 1$, и соотношения (10.8)–(10.10) позволяют написать равенства:

$$\begin{aligned}
Q(t+\Delta t, 1) &= Q(t,1)(1-\lambda\Delta t) + Q(t,2)\,\mu_1\Delta t + o(\Delta t),\\
Q(t+\Delta t, 2) &= Q(t,1)\lambda\Delta t + Q(t,2)(1-(\lambda+\mu_1)\Delta t) +\\
&\qquad\qquad + Q(t,3)\,\mu_2\Delta t + o(\Delta t),\\
Q(t+\Delta t, m) &= Q(t,m-1)\lambda\Delta t + Q(t,\, m)(1-(\lambda+\mu_2)\Delta t) +\\
&\qquad\qquad + Q(t,m+1)\,\mu_2\Delta t + o(\Delta t), \quad m = 3, 4, \ldots
\end{aligned}$$

Отсюда при $t > 0$ имеем

$$\begin{aligned}
(Q(t+\Delta t, 1) - Q(t,1))/\Delta t &= -\lambda Q(t,1) + Q(t,2)\,\mu_1 + o(\Delta t)/\Delta t,\\
(Q(t+\Delta t, 2) - Q(t,2))/\Delta t &= \lambda Q(t,1) - (\lambda+\mu_1)\,Q(t,2) +\\
&\qquad\qquad + Q(t,3)\,\mu_2 + o(\Delta t)/\Delta t,\\
(Q(t+\Delta t, m) - Q(t,m))/\Delta t &= \lambda Q(t,m-1) - (\lambda+\mu_2)\,Q(t,m) +\\
&\qquad\qquad + Q(t,m+1)\,\mu_2 + o(\Delta t)/\Delta t, \quad m = 3, 4, \ldots
\end{aligned}$$

Так как при $\Delta t \to 0$ предел правых частей этих равенств существует, то в пределе для $t > 0$ получаем бесконечную систему линейных дифференциальных уравнений

$$\begin{gathered} dQ(t,1)/dt = -\lambda Q(t,1) + \mu_1 Q(t,2), \\ dQ(t,2)/dt = \lambda Q(t,1) - (\lambda + \mu_1)\, Q(t,2) + \mu_2 Q(t,3), \\ dQ(t,m)/dt = \lambda Q(t,m-1) - (\lambda + \mu_2)\, Q(t,m) + \mu_2 Q(t,m+1), \quad m \geqslant 3. \end{gathered} \tag{10.11}$$

Будем дополнительно предполагать, что в момент $t = 0$ число машин в транспортной пачке равно i. Тогда динамика распределения числа машин в транспортной пачке определяется решением системы дифференциальных уравнений (10.11) с начальными условиями $Q(0,i) = 1$, $Q(0,m) = 0$ при $m \neq i$. Явное решение системы дифференциальных уравнений (10.11) может быть найдено с помощью вывода и последующего решения дифференциального уравнения в частных производных для производящей функции $\Pi_\chi(t)(t,z) = \sum_{m=1}^{\infty} z^m Q(t,m)$, как это было сделано в задаче о свободном движении машин по магистрали.

Однако в этом случае процесс решения является очень громоздким и далеко нетривиальным. К счастью, нам потребуется здесь только свойства распределения числа машин в транспортной пачке при $t \to \infty$, т. е. некоторые свойства решений системы (10.11) при $t \to \infty$. Общие свойства решений такого рода систем дифференциальных уравнений детально изучены в работах Колмогорова, Феллера, Ледермана, Карлина, Кларка, Мак-Гре́гора, Рейтера, Федоткина. В частности, если $\lambda < \mu_2$, то существует единственное решение системы уравнений (10.11) и оно удовлетворяет при $m = 1, 2, \ldots$ следующим условиям: $\sum_{m=1}^{\infty} Q(t,m) = 1$, $\lim\limits_{t\to\infty} Q(t,m) = Q(m) > 0$, $\sum_{m=1}^{\infty} Q(m) = 1$. При этом указанные пределы не зависят от начальных условий, могут быть получены путем решения системы линейных алгебраических уравнений

$$\begin{gathered} 0 = -\lambda Q(1) + \mu_1 Q(2), \quad 0 = \lambda Q(1) - (\lambda + \mu_1)\, Q(2) + \mu_2 Q(3), \\ 0 = \lambda Q(m-1) - (\lambda + \mu_2)\, Q(m) + \mu_2 Q(m+1), \quad m = 3, 4, \ldots . \end{gathered} \tag{10.12}$$

Распределение $\{Q(m);\ m = 1, 2, \ldots\}$ называется *предельным* или *эргодическим* для числа всех типов машин в транспортной пачке. Это распределение характеризует так называемый установившийся или стационарный режим движения машин по магистрали. На содержательном уровне условие $\lambda < \mu_2$ означает, что интенсивность, с которой быстрые машины догоняют медленную, должна быть меньше интенсивности обгона. При $\lambda < \mu_2$ система (10.12) получается с помощью предельного перехода при $t \to \infty$ одновременно во всех уравнениях системы (10.11) с учетом равенств $\lim\limits_{t\to\infty} (dQ(t,m)/dt) = 0$, $\lim\limits_{t\to\infty} Q(t,m) = Q(m)$, $m \geqslant 1$. Отметим, что при каждом фиксированном

$m \geqslant 1$ предел $\lim\limits_{t\to\infty}(dQ(t,m)/dt)$ существует. Более того, каждый такой предел равен нулю. В противном случае модуль величины $Q(t,m)$ при $t \to \infty$ возрастал бы неограниченно, что невозможно в силу смысла величины $Q(t,m)$ как вероятности. Перейдем теперь к решению системы (10.12).

Пусть при $m = 2, 3, \ldots$ величина $u_m = -\lambda Q(m) + \mu_2 Q(m+1)$. Из первых двух уравнений системы (10.12) получим, что $Q(2) = \lambda\mu_1^{-1}Q(1)$ и $Q(3) = \lambda^2\mu_1^{-1}\mu_2^{-1}Q(1)$. Поэтому $u_2 = -\lambda Q(2) + \mu_2 Q(3) = 0$. Используя все равенства системы (10.12) за исключением первых двух, найдем $u_m - u_{m-1} = 0$, $m \geqslant 3$. Отсюда вытекает, что $u_m = 0$ для всех $m \geqslant 2$. Значит, для каждого фиксированного $m = 2, 3, \ldots$ вероятность $Q(m) = \lambda\mu_1^{-1}(\lambda\mu_2^{-1})^{m-2}Q(1)$. Учитывая условие нормировки $\sum_{m=1}^{\infty} Q(m) = 1$ и неравенство $\lambda < \mu_2$, окончательно получим следующие формулы:

$$\begin{gathered} Q(1) = \mu_1(\mu_2 - \lambda)(\mu_1\mu_2 + \lambda\mu_2 - \lambda\mu_1)^{-1}, \\ Q(m) = \lambda(\mu_2 - \lambda)(\mu_1\mu_2 + \lambda\mu_2 - \lambda\mu_1)^{-1}(\lambda\mu_2^{-1})^{m-2}, \quad m \geqslant 2. \end{gathered} \tag{10.13}$$

Сравнивая формулы (10.4) для распределения Бартлетта и формулы (10.13), придем к соотношениям вида

$$p = \lambda\mu_2(\mu_1\mu_2 + \lambda\mu_2 - \lambda\mu_1)^{-1}, \quad q = \lambda\mu_2^{-1}. \tag{10.14}$$

Итак, доказано, что предельное распределение числа всех типов машин в транспортной пачке является распределением Бартлетта. Вид формул (10.4), (10.13) и (10.14) позволяет дать следующий простой смысл параметров p и q для задачи о групповом движении машин по магистрали. Параметр p определяет вероятность появления транспортной пачки из быстрых машин для стационарного режима. Параметр q задает степень насыщения транспортной магистрали быстрыми машинами и его можно назвать коэффициентом заполнения или загрузкой транспортной магистрали. В самом деле, если $\mu_2 > \lambda$ и $\mu_2 \to \lambda$, то из (10.14) легко выводим, что параметры $p \to 1$, $q \to 1$. Поэтому при $\mu_2 > \lambda$ и $\mu_2 \to \lambda$ математическое ожидание $p\,(1-q)^{-1} + 1$ числа всех типов машин в транспортной пачке для стационарного режима растет неограниченно. Это обстоятельство постоянно наблюдают автомобилисты при существенном ухудшении погодных условий, когда на дорогах образуются транспортные заторы значительной протяженности.

В заключение рассмотрения этой задачи необходимо сделать одно важное замечание. Выше было показано, что при $p = q$ распределение Бартлетта совпадает с геометрическим распределением. Из (10.14) находим, что это возможно только при $\mu_2(\mu_1\mu_2 + \lambda\mu_2 - \lambda\mu_1)^{-1} = \mu_2^{-1}$. Непосредственно отсюда получаем условие $\mu_2^2 - (\mu_1 + \lambda)\,\mu_2 + \lambda\mu_1 = 0$

или $(\mu_2 - \lambda)(\mu_2 - \mu_1) = 0$. Так как $\mu_2 > \lambda$, то только при $\mu_2 = \mu_1$ предельное распределение числа всех типов машин в транспортной пачке является геометрическим. Другими словами, если среднее время обгона каждой быстрой машиной медленную не зависит от числа машин в транспортной пачке, то мы снова имеем дело с геометрическим законом распределения.

1.4. Случайная величина с гипергеометрическим распределением. Пусть параметры n, r, s являются натуральными числами, для которых имеет место соотношение $n \geqslant r$, $n \geqslant s$. Случайная величина ξ, которая с вероятностью $\mathsf{P}(\{\omega\colon \xi(\omega) = m\}) = C_r^m C_{n-r}^{s-m}/C_n^s > 0$ принимает значение m, где $\max\{0, s + r - n\} \leqslant m \leqslant \min\{r, s\}$, называется *гипергеометрической*. Легко показать, что выполняется условие нормировки $\sum_m C_r^m C_{n-r}^{s-m}/C_n^s = 1$. Здесь суммирование ведется по всем возможным значениям параметра m. Для доказательства условия нормировки запишем следующее очевидное тождество по переменной величине z: $(1+z)^n = (1+z)^r(1+z)^{n-r}$. Применяя сначала разложение бинома Ньютона, а затем, приравнивая коэффициенты многочленов в левой и правой части указанного тождества при одинаковых степенях переменной z, например, при z^s, последовательно будем иметь:

$$\sum_{j=0}^{n} C_n^j z^j = \Big(\sum_{i=0}^{r} C_r^i z^i\Big) \times \Big(\sum_{k=0}^{n-r} C_{n-r}^k z^k\Big), \qquad C_n^s = \sum_m C_r^m C_{n-r}^{s-m}.$$

Отсюда вытекает условие нормировки.

Гипергеометрическая случайная величина является измерителем элементарных исходов для широкого круга реальных экспериментов. Необходимым условием каждого такого эксперимента является наличие некоторой совокупности из n элементов, среди которых имеется r элементов первого типа и $n-r$ другого типа. Далее, случайным образом отбирают последовательно для обследования s элементов один за другим без возвращения. Будем говорить, что в этом эксперименте имеется s случайных попыток для получения выборки. Основной интерес представляют свойства случайной величины ξ, которая подсчитывает число элементов первого типа в отобранной совокупности. Примерами применения такой схемы можно найти в задачах:

1) оценки качества выпускаемой промышленной продукции, когда проверка отобранной продукции связана с ее полным или частичным разрушением;
2) обработки экономических и социальных данных;
3) изучения популяций в биологии;
4) анализа литературных текстов в математической лингвистике и т. д.

Покажем, что случайная величина ξ распределена по гипергеометрическому закону.

Для определенности пометим элементы первого типа номерами 1, 2,..., r и элементы второго типа номерами $r+1$, $r+2,\ldots$, n. При $i = 1, 2, \ldots, s$ обозначим через e_i номер элемента, который был отобран в i-й попытке. Произвольный элементарный исход этого эксперимента будем описывать вектором $\omega = (e_1, e_2, \ldots, e_s)$, для которого в силу отбора элементов без возвращения имеем $e_i \neq e_j$ при $i \neq j$. Тогда для этого эксперимента достоверное событие $\Omega = \{\omega\colon \omega = (e_1, e_2, \ldots, e_s),\ e_i \neq e_j,\ i \neq j;\ e_i = 1, 2, \ldots, n\}$, а множество всех наблюдаемых исходов имеет вид $\mathcal{F} = \{A\colon A \subset \Omega\}$. Легко подсчитать, что число всех элементарных исходов равно числу размещений из n элементов по s, т. е. $A_n^s = n \times (n-1) \times \ldots \times (n-s+1)$. В силу случайного механизма отбора элементов можно считать элементарные исходы равновозможными. Поэтому вероятность $\mathsf{P}(\{\omega\}) = (A_n^s)^{-1}$. Итак, имеем вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ для этого опыта и дискретную случайную величину $\xi(\omega)\colon \Omega \to \{m\colon \max\{0, s+r-n\} \leqslant m \leqslant \min\{r, s\}\}$.

Событие вида $\{\omega\colon \xi(\omega) = m\}$ заключается в том, что в отобранной выборке объема s имеется m элементов первого типа и $s{-}m$ элементов второго типа. Пусть описание $\omega_0 \in \{\omega\colon \xi(\omega) = m\}$. Тогда элементарный исход $\{\omega_0\}$ произойдет, если сначала отбираем упорядоченную совокупность в количестве m элементов первого типа (первое действие), затем отбираем упорядоченную совокупность в количестве $s{-}m$ элементов второго типа (второе действие) и, наконец, выбираем m мест из s для первой отобранной совокупности (третье действие). Легко видеть, что первое действие можно выполнить $A_r^m = r(r-1) \times \ldots \times (r-m+1)$ различными способами, второе $A_{n-r}^{s-m} = (n-r)(n-r-1) \times \ldots \times (n+m+1-r-s)$ различными способами и, наконец, третье действие можно выполнить $C_s^m = s(s-1) \times \ldots \times (s-m+1)(m!)^{-1}$ различными способами. При этом изменение результата любого из указанных действий неизбежно ведет к наблюдению другого элементарного исхода, при котором случайная величина ξ принимает значение m. Тогда по правилу умножения комбинаторики получаем, что событию $\{\omega\colon \xi(\omega) = m\}$ благоприятствуют $A_r^m\, A_{n-r}^{s-m}\, C_s^m$ элементарных исходов. Отсюда вычисляем вероятность

$$p_m = \mathsf{P}(\{\omega\colon \xi(\omega) = m\}) = A_r^m\, A_{n-r}^{s-m}\, C_s^m / A_n^s = C_r^m\, C_{n-r}^{s-m} / C_n^s, \quad (10.15)$$

где $\max\{0, s+r-n\} \leqslant m \leqslant \min\{r, s\}$ и натуральные числа n, $r{\leqslant}n$, $s{\leqslant}n$ являются параметрами распределения. Следовательно, случайная величина ξ в таких задачах распределена по гипергеометрическому закону.

Вычисления математического ожидания и дисперсии такой случайной величины, используя для этого вид (10.15) гипергеометрического распределения или его производящую функцию, являются слишком громоздкими. Выбор вероятностной модели $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ позволяет решить эту задачу достаточно простыми средствами. С этой целью введем s случайных величин ξ_1, $\xi_2, \ldots$, ξ_s. Случайная величина ξ_v подсчитывает число отобранных элементов первого типа в попытке с номером $v = 1, 2, \ldots, s$ и $\xi_v(\omega) \in \{0, 1\}$. Значит, $\xi = \xi_1 + \xi_2 + \ldots + \xi_s$. Покажем, что величины ξ_1, $\xi_2, \ldots$, ξ_s имеют одинаковое распределение и являются зависимыми. Случайное событие $\{\omega\colon \xi_v(\omega) = 1\} = \{\omega = (e_1, e_2, \ldots, e_{v-1}, e_v, e_{v+1}, \ldots, e_s)\colon e_i \neq e_j,\ i \neq j;\ e_i = 1, 2, \ldots, n;$ $e_v = 1, 2, \ldots, r\}$ содержит $r \times (n-1) \times (n-2) \times \ldots \times (n-s+1)$ описаний всех тех элементарных исходов, которые происходят с этим событием. Откуда следует, что $\mathsf{P}(\{\omega\colon \xi_v(\omega) = 1\}) = rn^{-1}$.

Пусть $v \neq u$, $1 \leqslant v, u \leqslant s$. Событие $\{\omega\colon \xi_v(\omega) = 1,\, \xi_u(\omega) = 1\}$ означает, что в попытках с номерами v и u были отобраны элементы первого типа. Тогда $\{\omega\colon \xi_v(\omega) = 1,\, \xi_u(\omega) = 1\} = \{\omega = (e_1, e_2, \ldots, e_{v-1}, e_v, e_{v+1}, \ldots, e_{u-1}, e_u, e_{u+1}, \ldots, e_s)\colon e_i \neq e_j,\ \ i \neq j;$ $e_i = 1, 2, \ldots, n;\ e_v, e_u = 1, 2, \ldots, r\}$ и, значит, этому событию благоприятствует $r \times (r-1) \times (n-2)(n-3) \times \ldots \times (n-s+1)$ элементарных исходов. Отсюда имеем $\mathsf{P}(\{\omega\colon \xi_v(\omega) = 1\, \xi_u(\omega) = 1\}) = r(r-1)/n(n-1) \neq \neq \mathsf{P}(\{\omega\colon \xi_v(\omega) = 1\}) \times \mathsf{P}(\{\omega\colon \xi_u(\omega) = 1\})$. Очевидно, что случайная величина $\xi_v \xi_u$ принимает значение ноль или единица и $\mathsf{P}(\{\omega\colon \xi_v(\omega)\xi_u(\omega) = 1\}) = \mathsf{P}(\{\omega\colon \xi_v(\omega) = 1,\, \xi_u(\omega) = 1\}) = r(r-1)/n(n-1)$. Отсюда имеем $\mathsf{M}(\xi_v \xi_u) = r(r-1)/n(n-1)$. Далее, используя равенства

$$\xi = \sum_{v=1}^{s} \xi_v, \quad \mathsf{M}(\xi_v) = rn^{-1}, \quad \mathsf{D}(\xi_v) = r(n-r)\, n^{-2},$$
$$\mathsf{M}(\xi_v \xi_u) = r(r-1)/n(n-1)$$

и свойства

$$\mathsf{M}\Big(\sum_{v=1}^{s} \xi_v\Big) = \sum_{v=1}^{s} \mathsf{M}(\xi_v),$$

$$\mathsf{D}\Big(\sum_{v=1}^{s} \xi_v\Big) = \sum_{v=1}^{s} \mathsf{D}(\xi_v) + 2 \sum_{v<u} \mathsf{cov}(\xi_v, \xi_u),$$

$\mathsf{cov}(\xi_v, \xi_u) = \mathsf{M}(\xi_v \xi_u) - \mathsf{M}(\xi_v)\, \mathsf{M}(\xi_u)$, для математического ожидания, ковариации, дисперсии и среднего квадратического отклонения получим

$$\mathsf{M}\xi = sr/n, \quad \mathsf{cov}(\xi_v, \xi_u) = r(r-n)/n^2(n-1),$$
$$\mathsf{D}\xi = sr(n-r)(n-s)\, n^{-2}(n-1)^{-1},$$
$$\sigma(\xi) = (sr(n-r)(n-s))^{1/2} n^{-1} (n-1)^{-1/2}.$$

Найдем теперь $\mathsf{Mo}^*\xi$ или наивероятнейшее значение m_0 гипергеометрически распределеной случайной величины ξ из условия вида:

$$\max\{C_r^m\, C_{n-r}^{s-m}/C_n^s\colon \max\{0, s+r-n\} \leqslant m \leqslant \min\{r,s\}\} = \\ = C_r^{m_0}\, C_{n-r}^{s-m_0}/C_n^s.$$

Для этого при каждом $m \geqslant \max\{0, s+r-n\}+1$ и $m \leqslant \min\{r,s\}$ вычислим отношение

$$\begin{aligned} p_m/p_{m-1} &= [C_r^m C_{n-r}^{s-m}/C_n^s]/[C_r^{m-1}C_{n-r}^{s-m+1}/C_n^s] = \\ &= (s-m+1)(r-m+1)[m(n-r-s+m)]^{-1} = \\ &= 1+[(s+1)(r+1)-(2+n)m]m^{-1}(n-r-s+m)^{-1} = \\ &= 1+[(s+1)(r+1)(2+n)^{-1}-m](2+n)m^{-1}(n-r-s+m)^{-1}. \end{aligned}$$

Так как при $m \geqslant \max\{0, s+r-n\}+1$ имеет место $n-r-s+m>0$, то при любом $m<(s+1)(r+1)(n+2)^{-1}$ выполняется неравенство $p_m>p_{m-1}$, а при $m>(s+1)(r+1)(n+2)^{-1}$ будет $p_m<p_{m-1}$. Методом математической индукции по n легко доказать, что $\max\{0, s+r-n\}+1<(s+1)(r+1)(n+2)^{-1}<\min\{r+1, s+1\}$. Следовательно, максимальное значение для вероятностей p_m, $\max\{0, s+r-n\}+1<(s+1)(r+1)(n+2)^{-1}<\min\{r+1, s+1\}$, достигается при одном значении $m=m_0=[(s+1)(r+1)(n+2)^{-1}]$, если число $(s+1)(r+1)(n+2)^{-1}$ является дробным, или при двух значениях $m=m_0=(s+1)(r+1)(n+2)^{-1}$ и $m=m_0-1$, если $(s+1)(r+1)(n+2)^{-1}$ будет целым.

Имеем $(s+1)(r+1)(n+2)^{-1}<1$ при r, $s<(n+2)^{1/2}-1$ и $n \geqslant 2$. Более того, для фиксированного $t=2,3,\ldots$ существуют такие значения параметров n, r и s, например, $s=tr-1$ и $n \leqslant t(r+1)-2$, что имеет место неравенство $(s+1)(r+1)(n+2)^{-1} \geqslant r$. Отсюда следует, что для наивероятнейшего значения $\mathsf{Mo}^*\xi$ и для моды $\mathsf{Mo}\,\xi$ случайной величины ξ с гипергеометрическим распределением в зависимости от значений параметров n, r и s получаем те же семь различных вариантов, которые были приведены для биномиальной случайной величины. Перечислим некоторые из этих вариантов. Например, при $(s+1)(r+1)(n+2)^{-1}<1$ имеем $\mathsf{Mo}^*\xi=m_0=0$, $p_m>p_{m+1}$ для всех $0 \leqslant m<\min\{r,s\}$ и, значит, мода такой случайной величины ξ не существует. Если $(s+1)(r+1)(n+2)^{-1}=1$, то $\mathsf{Mo}^*\xi=m_0 \in \{0,1\}$, $p_0=p_1$, $p_m>p_{m+1}$ для всех $1 \leqslant m<\min\{r,s\}$ и, следовательно, мода $\mathsf{Mo}\,\xi$ случайной величины ξ с гипергеометрическим распределением равна единице. В этом случае распределение гипергеометрической случайной величины является унимодальным.

Пусть верны неравенства $\min\{r,s\}>(s+1)(r+1)(n+2)^{-1}<1$. Тогда получаем, что $\mathsf{Mo}^*\xi=\mathsf{Mo}\,\xi$. Более того, если при этом произведение $(s+1)(r+1)(n+2)^{-1}$ является целым числом, то гипергеометрическое распределение будет мультимодальным. Рассмотрим применение гипергеометрического распределения.

Пример 10.9 (*задача о выборочном контроле продукции.*) Пусть автомобильный завод каждый раз по заказу спортивных клубов изготовляет партию гоночных автомобилей в количестве 49 штук. За счет технологии и ручной сборки в каждой такой партии имеется 6 дефектных гоночных автомобилей. Спортивные клубы для контроля этой продукции случайным образом отбирают шесть автомобилей и подвергают их экстремальным испытаниям. Согласно условиям контракта, спортивные клубы отказываются от покупки заказанной партии гоночных автомобилей, считая ее бракованной, если среди отобранных шести автомобилей испытания пройдут не более двух автомобилей. Определить вероятность покупки партии гоночных автомобилей, математическое ожидание, дисперсию и наивероятнейшее значение числа бракованных среди отобранных автомобилей.

Решение. Если случайная величина ξ определяет число дефектных автомобилей из случайно отобранных шести, то искомая вероятность равна

$$\sum_{m=0}^{3} \mathsf{P}(\{\omega\colon \xi(\omega)=m\}) = \sum_{m=0}^{3} C_6^m C_{43}^{6-m}/C_{49}^6 \approx 0{,}99901.$$

Заметим, что партия гоночных автомобилей будет забракована спортивными клубами с вероятностью, приблизительно равной 0,00099. Наконец, при $n=49$, $r=6$, $s=6$ математическое ожидание $\mathsf{M}\xi = sr/n \approx \approx 0{,}7347$, дисперсия $\mathsf{D}\xi = sr(n-r)(n-s)/n^2(n-1) \approx 0{,}5776$ и наивероятнейшее значение $\mathsf{Mo}^*\xi = [(s+1)(r+1)(n+2)^{-1}] = [49/51] = 0$.

В пунктах а), б), в) и г) подробно изложена методика, по которой можно изучать свойства дискретных случайных величин. В качестве примера применения этой методики можно предложить изучение свойств так называемой случайной величины ξ с распределением Паскаля. Это распределение, зависящее от двух числовых параметров: $r \in \{1,2,\ldots\}$ и $0<p<1$, имеет следующий вид: $p_m = \mathsf{P}(\{\omega\colon \xi(\omega)=m\}) = C_{r+m-1}^m p^r q^m$, $m=0,1,\ldots$. Здесь $q=1-p$. Случайная величина ξ с распределением Паскаля подсчитывает число неудач до появления r-го успеха в схеме независимых испытаний Бернулли с вероятностью успеха в одном испытании, равной p. Приведем окончательные формулы для вычисления основных числовых характеристик такой случайной величины: $\alpha_1\xi = \mathsf{M}\xi = rq/p$, $\mathsf{D}\xi = \beta_2\xi = rq/p^2$, $\sigma(\xi) = p^{-1}(rq)^{1/2}$, $\beta_3\xi = rq(1+q)/p^3$, $\beta_4\xi = rqp^{-4}(3rq+6q+p^2)$, $\mathsf{Ka}\,\xi = (1+q)(rq)^{-1/2}$, $\text{Э}\xi = 6r^{-1} + p^2(rq)^{-1}$.

§ 2. Тестовые непрерывные случайные величины

2.1. Нормальный закон распределения (закон Гаусса). Этот закон играет исключительно важную роль в теории вероятностей и занимает среди других законов распределения особое положение (*Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 1, 2.— М.: Мир, 1984). Такое обстоятельство вызвано тем, что известно очень много часто встречающихся на практике экспериментов, в которых этот закон позволяет адекватно описывать вероятностные свойства их количественных характеристик. Приведем следующие простые примеры:

1) ошибки измерений, ошибки стрельбы по мишени — это случайные величины, которые подчинены нормальному или близкому к нему закону;
2) координаты точек попадания при обстреле цели в одинаковых условиях из одного и того же орудия, если прицел постоянен, подчиняются нормальному закону;
3) проекции на оси координат скорости, с которой движется молекула свободного газа, также приближенно подчиняются нормальному закону;
4) сумма достаточно большого числа независимых слагаемых случайных величин, подчиненных каким угодно законам распределения (при соблюдении некоторых весьма нежестких ограничений), приближенно подчиняется нормальному закону;
5) наиболее существенные количественные характеристики однотипных изделий при их массовом производстве являются случайными величинами, вероятностные свойства которых определяются законом Гаусса.

Главная особенность, выделяющая нормальный закон среди других законов непрерывного типа, состоит в том, что он является предельным, к которому приближаются другие законы распределений при весьма часто встречающихся типичных условиях. Перейдем непосредственно к определению нормального закона распределения.

Определение 10.5. Случайная величина ξ *подчинена нормальному закону распределения* или *закону Гаусса*, если ее плотность $f_\xi(x) = f_\xi(x; a, \sigma) = (2\pi)^{-1/2}\sigma^{-1}\exp\{-(x-a)^2/(2\sigma^2)\}$, где $-\infty < a < +\infty$ и $0 < \sigma < +\infty$ суть параметры распределения. Будем также говорить, что непрерывная величина ξ является (a, σ)-нормальной или имеет $N(a, \sigma)$-распределение.

Следовательно, кривая распределения для закона Гаусса полностью определяется параметрами $-\infty < a < +\infty$ и $0 < \sigma < +\infty$ и имеет симметричный относительно прямой $x = a$ холмообразный вид или вид сечения церковного колокола. Максимальная ордината кривой распре-

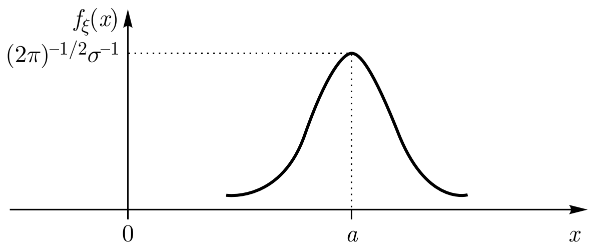


Рис. 10.4

деления равна $(2\pi)^{-1/2}\sigma^{-1}$, и она соответствует такой точке на оси абсцисс, которая удалена на расстоянии $x = a$ от начала координат. Ось абсцисс является асимптотой для обеих ветвей кривой $f_\xi(x)$, расположенных справа и слева относительно прямой $x = a$, так как имеем равенство $\lim\limits_{|x|\to\infty} f_\xi(x) = 0$. График этой кривой представлен на рис. 10.4.

Проверим, что $f_\xi(x)$ обладает свойствами плотности распределения вероятностей. Во-первых, функция $f_\xi(x)$ строго положительна, так как параметр $\sigma > 0$ и функция $\exp\{-(x-a)^2/(2\sigma^2)\}$ также всегда строго положительна. Во-вторых, выполняется условие нормировки $\int_{-\infty}^{+\infty} f_\xi(x)\,dx = 1$. Действительно, интеграл в левой части равенства с помощью замены

$$x = 2^{1/2}\sigma t + a, \quad dx = 2^{1/2}\sigma dt, \quad t = 2^{-1/2}\sigma^{-1}(x-a), \quad -\infty < t < +\infty, \tag{10.16}$$

приводится к виду

$$\int\limits_{-\infty}^{+\infty} f_\xi(x)\,dx = \int\limits_{-\infty}^{+\infty} (2\pi)^{-1/2}\sigma^{-1}\exp\{-(x-a)^2/(2\sigma^2)\}dx =$$
$$= (2\pi)^{-1/2}\sigma^{-1}\int\limits_{-\infty}^{+\infty} 2^{1/2}\sigma\exp\{-t^2\}dt =$$
$$= \pi^{-1/2}\int\limits_{-\infty}^{+\infty}\exp\{-t^2\}dt = \pi^{-1/2}\pi^{1/2} = 1.$$

так как интеграл Эйлера–Пуассона $\int_{-\infty}^{+\infty}\exp\{-t^2\}\,dt$ равен $\pi^{1/2}$.

Выясним теперь смысл параметров a и σ. Для этого найдем последовательно $\mathsf{M}\xi$ и $\mathsf{D}\xi$ случайной величины ξ, распределенной по закону Гаусса с параметрами a и σ. Из определений математического ожидания и дисперсии непрерывной случайной величины, используя замену (10.16), получим, что

$$\mathsf{M}\xi = \int\limits_{-\infty}^{+\infty} x f_\xi(x)\,dx = (2\pi)^{-1/2}\sigma^{-1}\int\limits_{-\infty}^{+\infty} x\exp\{-(x-a)^2/(2\sigma^2)\}dx =$$

$$
\begin{aligned}
&= (2\pi)^{-1/2}\sigma^{-1}\int\limits_{-\infty}^{+\infty}(2^{1/2}\sigma t + a)\exp\{-t^2\}2^{1/2}\sigma\,dt = \\
&= 2^{1/2}\sigma\pi^{-1/2}\int\limits_{-\infty}^{+\infty} t\exp\{-t^2\}\,dt + a\pi^{-1/2}\int\limits_{-\infty}^{+\infty}\exp\{-t^2\}\,dt = \\
&= -\sigma(2\pi)^{-1/2}\exp\{-t^2\}\big|_{-\infty}^{+\infty} + a\pi^{-1/2}\int\limits_{-\infty}^{+\infty}\exp\{-t^2\}\,dt = \\
&= a\pi^{-1/2}\pi^{1/2} = a.
\end{aligned}
$$

Итак, $\mathsf{M}\xi = a$. Аналогичным способом подсчитаем дисперсию:

$$
\begin{aligned}
\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \mathsf{M}(\xi - a)^2 &= \int\limits_{-\infty}^{+\infty}(x-a)^2 f_\xi(x)\,dx = \\
&= (2\pi)^{-1/2}\sigma^{-1}\int\limits_{-\infty}^{+\infty}(x-a)^2\exp\{-(x-a)^2/(2\sigma^2)\}\,dx = \\
&= (2\pi)^{-1/2}\sigma^{-1}\int\limits_{-\infty}^{+\infty}2t^2\sigma^2 2^{1/2}\sigma\exp\{-t^2\}\,dt = \\
&= 2\sigma^2\pi^{-1/2}\int\limits_{-\infty}^{+\infty}t^2\exp\{-t^2\}\,dt.
\end{aligned}
$$

Далее интегрируем по частям и находим:

$$
\begin{aligned}
&\mathsf{D}\xi = -2\sigma^2\pi^{-1/2}\int\limits_{-\infty}^{+\infty}(t/2)\,d(\exp\{-t^2\}) = \\
&= \sigma^2\pi^{-1/2}\Big(-t\exp\{-t^2\}\big|_{-\infty}^{+\infty} + \int\limits_{-\infty}^{+\infty}\exp\{-t^2\}\,dt\Big) = \sigma^2\pi^{-1/2}\pi^{1/2} = \sigma^2,
\end{aligned}
$$

$$\boldsymbol{\sigma}(\xi) = \sigma.$$

Значит, параметры a и σ представляют собой соответственно математическое ожидание и среднее квадратическое отклонение случайной величины ξ, распределенной по нормальному закону. Легко видеть, что наивероятнейшее значение $\mathsf{Mo}^*\xi$ случайной величины с распределением Гаусса равно параметру a. При этом наивероятнейшее значение совпадает с модой, медианой и математическим ожиданием. В самом деле, так как $\max\limits_x f_\xi(x) = f_\xi(a) = (2\pi)^{-1/2}\sigma^{-1}$ и $F(a) = \int_{-\infty}^a f_\xi(x)\,dx = 1/2$, то $\mathsf{Mo}^*\xi = \mathsf{Mo}\,(\xi) = \mathsf{Me}(\xi) = \mathsf{M}\xi = a$.

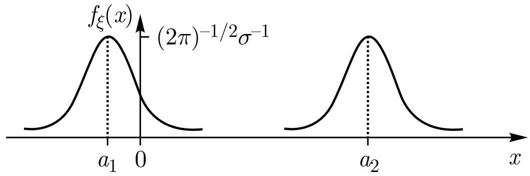


Рис. 10.5

Выясним теперь поведение кривой распределения вероятностей для случайной величины ξ, распределенной по закону Гаусса, в связи с изменением ее параметров a и σ (рис. 10.5). При изменении математического ожидания от значения a_1 до значения a_2 кривая распределения сдвигается вдоль оси абсцисс, однако, вид ее сохраняется. Это объясняется тем, что математическое ожидание характеризует центр рассеивания. Поэтому величина a называется параметром «сдвига» распределения.

При изменении второго параметра, а именно среднего квадратического отклонения σ, сдвига кривой не происходит. Однако кривая становится либо более пологой при увеличении разброса σ, либо более крутой при уменьшении разброса σ $(\sigma_1 > \sigma_2)$. Качественный вид этих графиков можно объяснить тем, что максимальная ордината $f_\xi(x)$ равна $(2\pi)^{-1/2}\sigma^{-1}$ и зависит обратно пропорционально от параметра σ, а площадь под кривой распределения всегда равна единице. Величина σ называется параметром «масштаба» распределения. На рис. 10.6 приведен случай для разброса σ_1 и на рис. 10.7 рассмотрен случай, когда второй параметр равен σ_2.

Вычислим вспомогательные числовые характеристики. Прежде всего, найдем среднее отклонение и центральные моменты для величи-

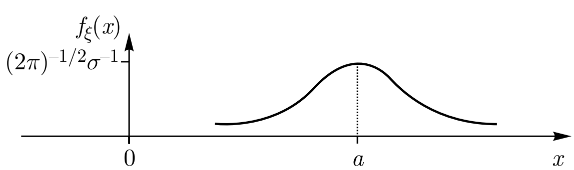


Рис. 10.6

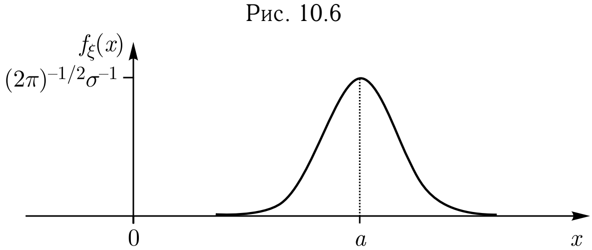


Рис. 10.7

ны ξ, распределенной по закону Гаусса. Используя определение среднего отклонения, формулу для математического ожидания непрерывной случайной величины, свойства интеграла Римана и замену (10.16), получим, что

$$\mathsf{E}\xi = \mathsf{M}(|\xi - \mathsf{Me}\,\xi|) = (2\pi)^{-1/2}\sigma^{-1}\int\limits_{-\infty}^{+\infty}|x-a|\exp\{-(x-a)^2(2\sigma^2)^{-1}\}\,dx =$$

$$= (2\pi)^{-1/2}\sigma^{-1}\Big[\int\limits_{-\infty}^{a}(a-x)\exp\{-(x-a)^2/2\sigma^2\}\,dx +$$

$$+ \int\limits_{a}^{+\infty}(x-a)\exp\{-(x-a)^2/2\sigma^2\}\,dx\Big] =$$

$$= 2^{1/2}\pi^{-1/2}\sigma\Big[-\int\limits_{-\infty}^{0}t\exp\{-t^2\}dt + \int\limits_{0}^{+\infty}t\exp\{-t^2\}dt\Big] =$$

$$= 2^{3/2}\pi^{-1/2}\sigma\int\limits_{0}^{+\infty}t\exp\{-t^2\}\,dt =$$

$$= -2^{1/2}\sigma\pi^{-1/2}\exp\{-t^2\}\Big|_0^{+\infty} = 2^{1/2}\pi^{-1/2}\sigma.$$

Используя определение центрального момента k-го порядка и замену (10.16), получим, что

$$\beta_k(\xi) = (2\pi)^{-1/2}\sigma^{-1}\int\limits_{-\infty}^{+\infty}(x-a)^k\exp\{-(x-a)^2/(2\sigma^2)\}\,dx =$$

$$= (2\pi)^{-1/2}\sigma^{-1}\int\limits_{-\infty}^{+\infty}(2^{1/2}\sigma t)^k\exp\{-t^2\}2^{1/2}\sigma\,dt =$$

$$= (2^{1/2}\sigma)^k\pi^{-1/2}\int\limits_{-\infty}^{+\infty}t^k\exp\{-t^2\}\,dt =$$

$$= -(2^{1/2}\sigma)^k\pi^{-1/2}2^{-1}\int\limits_{-\infty}^{+\infty}t^{k-1}d(\exp\{-t^2\}) =$$

$$= (2^{1/2}\sigma)^k\pi^{-1/2}2^{-1}\Big[-t^{k-1}\exp\{-t^2\}\Big|_{-\infty}^{+\infty} + (k-1)\int\limits_{-\infty}^{+\infty}t^{k-2}\exp\{-t^2\}\,dt\Big] =$$

$$= (k-1)(2^{1/2}\sigma)^k\pi^{-1/2}2^{-1}\int\limits_{-\infty}^{+\infty}t^{k-2}\exp\{-t^2\}\,dt,$$

где $k \geqslant 2$.

Так как $\beta_{k-2}(\xi) = (2^{1/2}\sigma)^{k-2}\pi^{-1/2}\int_{-\infty}^{+\infty} t^{k-2}\exp\{-t^2\}\,dt$, то $\beta_k(\xi) = \sigma^2(k-1)\,\beta_{k-2}(\xi)$. Применяя равенства $\beta_0(\xi) = 1$, $\beta_1(\xi) = 0$ и рекуррентное по $k \geqslant 2$ выражение $\beta_k(\xi) = \sigma^2(k-1)\,\beta_{k-2}(\xi)$, найдем формулы для центральных моментов $\beta_{2k+1}(\xi) = 0$, $\beta_{2k}(\xi) = (2k-1)!!\,\sigma^{2k}$, $k \geqslant 1$. Здесь символ $(2k-1)!! = 1 \times 3 \times \ldots \times (2k-1)$ есть нечетный факториал. Поэтому, коэффициент асимметрии и эксцесс нормальной случайной величины равны $\mathsf{Ka}\,\xi = \beta_3\xi(\boldsymbol{\sigma}\xi)^{-3} = 0$, $\mathsf{Э}\xi = \beta_4\xi(\boldsymbol{\sigma}\xi)^{-4} - 3 = 0$. Последние равенства вполне соответствуют физическому смыслу каждой из этих числовых характеристик. Например, коэффициент асимметрии равен нулю в силу симметрии кривой распределения нормальной случайной величины относительно ее математического ожидания. Далее, в определении эксцесса используется постоянная величина, равная трем. В силу этого эксцесс нормального распределения оказался равным нулю. Теперь с полным основанием можно утверждать, что эксцесс случайной величины характеризует крутизну наклона графика ее плотности вероятностей относительно кривой гауссовского распределения.

В качестве другой вспомогательной характеристики рассеивания нормального распределения специально введена такая числовая характеристика, как срединное (вероятное) отклонение, обозначаемое через $C(\xi)$.

Определение 10.6. *Срединным отклонением* $C(\xi)$ называется половина длины интервала с центром в точке с абсциссой $\mathsf{M}\xi$, вероятность попадания в который для случайной величины ξ равна 1/2 (см. [4]).

Для вычисления значения $C(\xi)$ необходимо определить вероятность $\mathsf{P}(\{\omega\colon |\xi(\omega) - a| < C(\xi)\}) = \mathsf{P}(\{\omega\colon a - C(\xi) \leqslant \xi(\omega) < a + C(\xi)\})$. Вначале рассмотрим самый общий случай, а именно, для $-\infty < c < d < +\infty$ найдем вероятность вида

$$\begin{aligned}\mathsf{P}(\{\omega\colon c \leqslant \xi(\omega) < d\}) &= \mathsf{P}(\{\omega\colon c - a \leqslant \xi(\omega) - a < d - a\}) = \\ &= \mathsf{P}(\{\omega\colon (c-a)\,\sigma^{-1} \leqslant (\xi(\omega) - a)\,\sigma^{-1} < (d-a)\,\sigma^{-1}\}).\end{aligned}$$

Так как $(\xi(\omega) - a)\,\sigma^{-1}$ является центрированной и нормированной случайной величиной, то ее математическое ожидание равно нулю, а дисперсия равна единице. Случайная величина $\eta(\omega) = (\xi(\omega) - a)\,\sigma^{-1}$ выражается через величину $\xi(\omega)$ линейной функциональной зависимостью $y = g(x) = (x-a)\,\sigma^{-1}$, для которой обратная функция есть $x = w(y) = \sigma y + a$ и $w'(y) = dw/dy = \sigma$. Поэтому плотность $\eta(\omega)$ равна $f_\eta(y) = f_\xi(w(y))\,w'(y) = \sigma(2\pi)^{-1/2}\sigma^{-1}\exp\{-(\sigma y + a - a)^2/(2\sigma^2)\} = (2\pi)^{-1/2}\exp\{-y^2/2\}$ и случайная величина $\eta(\omega) = (\xi(\omega) - a)\,\sigma^{-1}$ распределена по закону Гаусса и является $(0,1)$-нормальной. Очень часто $(0,1)$-нормальную случайную величину называют *стандартной*.

Дифференцируя плотность распределения
$$f(x) = (2\pi)^{-1/2}\exp\{-x^2/2\}$$
дважды, получим
$$d^2f(x)/dx^2 = (2\pi)^{-1/2}\exp\{-x^2/2\} + \\ + x^2(2\pi)^{-1/2}\exp\{-x^2/2\} = f(x)(x^2-1).$$

Так как плотность распределения $f(x) \neq 0$, то $d^2f(x)/dx^2 = 0$ при $x_1 = 1$ или при $x_2 = -1$. Поэтому точки вида $(1, (2\pi)^{-1/2}\exp\{-1/2\})$, $(-1, (2\pi)^{-1/2}\exp\{-1/2\})$ на плоскости являются точками перегиба кривой плотности распределения стандартной нормальной случайной величины. Заметим, что $f(0) = (2\pi)^{-1/2} \approx 0{,}3989$, $f(1) = (2\pi)^{-1/2}\exp\{-1/2\} \approx 0{,}2419$, $f(2) \approx 0{,}0539$, $f(3) \approx 0{,}0044$ и $f(4) \approx 0{,}0001$. Отсюда видно, что кривая $f(x)$ с ростом x по абсолютной величине весьма быстро приближается к оси абсцисс. На рис. 10.8 приведены графики плотности $f(x) = \varphi(x)$ и интегральной функции $F(x) = \Phi_0(x)$ стандартной нормальной величины. Таковы простейшие свойства стандартной нормальной случайной величины.

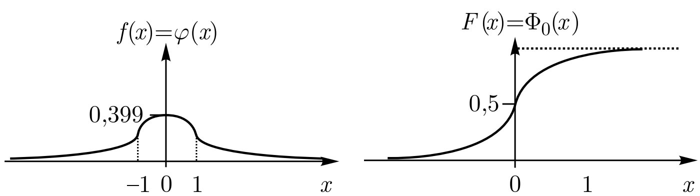


Рис. 10.8

Продолжая преобразования для вероятности
$$\mathsf{P}(\{\omega\colon c \leqslant \xi(\omega) < d\}) = \mathsf{P}(\{\omega\colon (c-a)\sigma^{-1} \leqslant \eta(\omega) < (d-a)\sigma^{-1}\}) = \\ = (2\pi)^{-1/2}\int\limits_{(c-a)/\sigma}^{(d-a)/\sigma} \exp\{-t^2/2\}\,dt = \\ = (2\pi)^{-1/2}\int\limits_{0}^{(d-a)/\sigma} \exp\{-t^2/2\}\,dt - (2\pi)^{-1/2}\int\limits_{0}^{(c-a)/\sigma} \exp\{-t^2/2\}\,dt.$$

Отсюда ясно, что для вычислений вероятностей событий, которые порождаются нормальными случайными величинами, часто приходится использовать определенные интегралы с различными пределами интегрирования от функции $\exp\{-t^2/2\}$. Как правило, такого рода интегралы не вычисляются в элементарных функциях. Поэтому удобно использовать в дальнейшем функцию Лапласа $\Phi(x) = \dfrac{\sqrt{2}}{\sqrt{\pi}}\int_0^x \exp\{-t^2/2\}\,dt$, свойства которой были изучены при выводе приближенных формул для биномиальных вероятностей в схеме независимых испытаний. После этих замечаний вероятность $\mathsf{P}(\{\omega\colon c \leqslant \xi(\omega) < d\}) = 2^{-1}[\Phi((d-a)\,\sigma^{-1}) - \Phi((c-a)\,\sigma^{-1})]$. Если $b > 0$ и $c = a - b$, $d = a + b$, то легко найдем $\mathsf{P}(\{\omega\colon a - b \leqslant \xi(\omega) < a + b\}) = \mathsf{P}(\{\omega\colon |\xi(\omega) - a| < b\}) = 2^{-1}[\Phi(b/\sigma) - \Phi(-b/\sigma)] = 2^{-1}[\Phi(b/\sigma) + \Phi(b/\sigma)] = \Phi(b/\sigma)$. Отсюда непосредственно вычислим значение срединного отклонения $C(\xi)$: $\mathsf{P}(\{\omega\colon |\xi(\omega) - a| < C(\xi)\}) = \Phi(C(\xi)/\sigma) = 1/2$. Используя таблицы функции Лапласа, находим связь между срединным отклонением $C(\xi)$ и средним квадратическим отклонением σ: $C(\xi) \approx 0{,}675\,\sigma$.

Для нормальной случайной величины ξ справедливо так называемое «правило трех сигм». Если случайная величина распределена по нормальному закону, то с вероятностью, близкой к единице, она принимает значения в интервале $(a - 3\sigma, a + 3\sigma)$, т. е.

$$\mathsf{P}(\{\omega\colon |\xi(\omega) - a| < 3\sigma\}) \approx 1.$$

Имеем

$$\mathsf{P}(\{\omega\colon |\xi(\omega) - a| < 3\sigma\}) = \Phi(3\sigma/\sigma) = \Phi(3)$$

и по таблицам для функции Лапласа находим, что $\Phi(3) \approx 0{,}9973 \approx 1$. Отсюда

$$\mathsf{P}(\{\omega\colon |\xi(\omega) - a| \geqslant 3\sigma\}) = 1 - \mathsf{P}(\{\omega\colon |\xi(\omega) - a| < 3\sigma\}) \approx 0{,}0027.$$

Если использовать неравенство Чебышева, то эта вероятность оценивается таким образом:

$$\mathsf{P}(\{\omega\colon |\xi(\omega) - a| \geqslant 3\sigma\}) \leqslant \mathsf{D}\xi/(3\sigma)^2 = \sigma^2/(9\sigma^2) = 1/9.$$

Итак, неравенство Чебышева позволяет установить следующий факт. Нормальная случайная величина отклоняется по модулю от своего математического ожидания на величину, которая не меньше 3σ, с вероятностью, не большей $1/9$. Однако, используя индивидуальные свойства случайной величины, распределенной по закону Гаусса, мы существенно уточнили значение оценки для этой вероятности. Пусть теперь в условиях реального эксперимента мы можем пренебречь возможностью появления некоторого события, вероятность которого меньше величины $0{,}003$. Тогда можно сказать, что в интервале вида $(a - 3\sigma, a + 3\sigma)$ практически достоверно сосредоточены все значения случайной величины, распределенной по нормальному закону. Другими словами, среднее квадратическое отклонение и математическое ожидание нормальной случайной величины ξ, ориентировочно, но все же очень точно, позволяет указать интервал ее практически достоверных значений. И наоборот, если мы знаем из опытов величину практически достоверно возможного отклонения по модулю нормальной случайной величины от ее среднего, то эта величина дает возможность приближенно вычислить значение параметра σ. Таково практическое значение принципа «правило трех сигм».

В заключение свойств нормального закона, используя равенство вида $\mathsf{P}(\{\omega\colon c \leqslant \xi(\omega) < d\}) = 2^{-1}[\Phi((d-a)\,\sigma^{-1}) - \Phi((c-a)\,\sigma^{-1})]$ и $\Phi(-\infty) = -1$, получим, что

$$\begin{aligned} F_\xi(x) = \mathsf{P}(\{\omega\colon -\infty < \xi < x\}) = \\ &= 2^{-1}[\Phi((x-a)/\sigma) - \Phi((-\infty - a)/\sigma)] = \\ &= 2^{-1}[\Phi((x-a)/\sigma) - \Phi(-\infty)] = 2^{-1}[\Phi((x-a)/\sigma) + 1]. \end{aligned}$$

Значит, с помощью таблицы функции Лапласа можно вычислять значения интегральной функции $F_\xi(x)$ для (a,σ)-нормальной случайной величины. В частном случае, если нормальная случайная величина является стандартной и ее интегральная функция обозначена через $\Phi_0(x)$, то $F_\xi(x) = 2^{-1}(1+\Phi(x)) = \Phi_0(x)$ и $\Phi(x) = 2\,\Phi_0(x) - 1$.

Рассмотрим ряд тестовых примеров, иллюстрирующих свойства нормального закона распределения.

Пример 10.10. Случайная величина ξ представляет собой ошибку измерения диаметра некоторой детали. Предполагается, что ξ распределена по нормальному закону с параметрами $\mathsf{M}\xi = 1{,}2$ см (происходит систематическое завышение в измерении). Стандарт ошибки равен 0,8 см. Найти вероятность того, что $-1{,}6$ см $\leqslant \xi < 1,6$ см, т. е. отклонение измеренного значения от истинного не превзойдет по абсолютной величине 1,6 см.

Решение. Ради простоты опускаем размерность заданных в данной задаче величин. Используя полученные формулы для распределения Гаусса и таблицы функции Лапласа, найдем

$$\begin{aligned} \mathsf{P}(\{\omega\colon -1{,}6 \leqslant \xi(\omega) < 1{,}6\}) = \\ &= 2^{-1}[\Phi((1{,}6 - 1{,}2)/0{,}8) - \Phi((-1{,}6 - 1,2)/0{,}8)] = \\ &= 2^{-1}[\Phi(0{,}5) + \Phi(3{,}5)] \approx 2^{-1}[\,0{,}3829 + 0{,}9995] = \\ &= 2^{-1} \times 1{,}3824 = 0{,}6912 \approx 0{,}7. \end{aligned}$$

Решим несколько иную задачу, когда систематической ошибки нет, т. е. $\mathsf{M}\xi = 0$. Тогда $\mathsf{P}(\{\omega\colon -1{,}6 \leqslant \xi(\omega) < 1{,}6\}) = \mathsf{P}(\{\omega\colon |\xi(\omega)| < 1{,}6\}) = \Phi(1{,}6/0{,}8) = \Phi(2) = 0{,}9545 \approx 0{,}95$.

Результаты этих решений можно прокомментировать следующим образом. С этой целью обозначим через b стандартную величину измеряемого параметра выпускаемой продукции, например, диаметра детали. Тогда результаты многократного измерения продукции можно считать значениями некоторой случайной величиной η. Известно, что ошибка $\xi = (\eta - b)$ любого измерения равна сумме так называемой случайной ошибки $(\eta - \mathsf{M}\eta)$ и систематической ошибки $(\mathsf{M}\eta - b)$. Пусть теперь контроль выпускаемой промышленной продукции осуществляется путем современной технологии и новейших измерительных приборов. В этом случае, как правило, систематическая ошибка $(\mathsf{M}\eta - b)$ измерения равна или близка к нулю. Поэтому $\mathsf{M}\xi = 0$ и $\mathsf{M}\eta = b$, а вероятность того, что ошибка измерения будет находиться в заданных допустимых пределах, практически равна единице. В приведенном примере эта вероятность приблизительно равна 0,95 при допустимом промежутке $[-1{,}6; 1{,}6)$ для ошибки измерения. Если отклонение измеренного значения некоторого параметра продукции от стандартного превзойдет допустимые пределы, то этот факт, скорее всего, можно объяснить браком выпускаемой продукции. Наоборот, при контроле выпускаемой продукции с помощью старой технологии и изношенных измерительных приборов вероятность того, что случайная ошибка измерения будет находиться вне допустимых пределов, существенно отлична от нуля. В нашем иллюстративном примере эта вероятность равна около 0,3. Тогда у руководителей такого предприятия появляется возможность объяснения указанных больших отклонений некоторой характеристики от стандарта не за счет брака выпускаемой продукции, а по причине случайных ошибок в измерениях. Этим обстоятельством частично можно объяснить наличие на государственных предприятиях в отсутствии конкуренции устаревшей технологии, изношенных орудий производства и измерительных приборов.

Пример 10.11. По цели, имеющей вид полосы (автострада), ширина которой равна 20 м, ведется стрельба в направлении, перпендикулярном автостраде. Прицеливание ведется по средней линии автострады. Среднее квадратическое отклонение в направлении стрельбы $\sigma = 8$ м. Имеется систематическая ошибка в направлении стрельбы, т. е. недолет 3 м. Найти вероятность попадания в автостраду при одном выстреле.

Решение. Имеются следующие исходные данные задачи: $a = -3$ м, $\sigma = 8$ м; $c = -10$ м, $d = 10$ м. Ради простоты будем опускать размерность величин при решении этого примера. Найдем

$$\mathsf{P}(\{\omega\colon -10 \leqslant \xi(\omega) < 10\}) = 2^{-1}[\Phi((10+3)/0{,}8) - \Phi((-10+3)/0{,}8)] = \\ = 2^{-1}[\Phi(1{,}625) + \Phi(0{,}875)] \approx 2^{-1}[0{,}8904 + 0{,}6319] \approx 0{,}76.$$

Пример 10.12. Известно, что случайная величина ξ распределена по нормальному закону с заданной плотностью распределения $f(x) = (2\pi)^{-1/2}\exp\{-(x-2)^2/2\}$. Найти интегральную функцию $F(x)$ и вероятность $\mathsf{P}(\{\omega\colon 0 \leqslant \xi(\omega) < 3\})$.

Решение. Параметры распределения Гаусса в этой задаче можно определить из вида функции $f(x)$, т. е. $a = 2$, $\sigma = 1$. Теперь можем определить $\mathsf{P}(\{\omega\colon 0 \leqslant \xi(\omega) < 3\})$ и $F(x)$ для $(2,1)$-нормальной случайной величины:

$$\mathsf{P}(\{\omega\colon 0 \leqslant \xi(\omega) < 3\}) = 2^{-1}[\Phi(3-2) - \Phi(0-2)] = \\ = 2^{-1}[\Phi(1) + \Phi(2)] \approx 2^{-1}[\,0{,}6807 + 0{,}9545] = 0{,}8186,$$

$$F(x) = 2^{-1}[\Phi(x-2) + 1].$$

Пример 10.13. При массовом производстве заводом шариков для подшипников контролировался их диаметр. Диаметр каждого шарика должен соответствовать стандарту, который равен 12 мм. При высокоточном измерении диаметра каждого шарика было установлено, что в 997 измерениях из 1000 значений диаметра были расположены в интервале от 11,4 мм до 12,6 мм. Найти приближенное значение среднего квадратического отклонения σ диаметра шариков.

Решение. Будем опускать размерность всех величин, которые рассматриваются в данной задаче. Изготовление шариков связано с большим числом операций, каждая из которых выполняется с незначительной ошибкой. Поэтому считаем, что диаметр ξ шариков распределен по нормальному закону с параметром $a = 12$. В условиях массового производства, когда имеется возможность проведения достаточно большого числа измерений, частоту 99,7 % события $\{\omega\colon 11{,}4 < \xi(\omega) < 12{,}6\}$ можно заменить вероятностью $\mathsf{P}(\{\omega\colon 11{,}4 < \xi(\omega) < 12{,}6\}) = 0{,}997$. Более того, в этих условиях можно утверждать, что в интервале вида $(12 - 0{,}6,\ 12 + 0{,}6) \approx (12 - 3\,\sigma,\ 12 + 3\,\sigma)$ практически достоверно сосредоточены все значения диаметра шариков. Отсюда получаем, что $3\,\sigma \approx 0,6$ или $\sigma \approx 0{,}2$.

Результаты этого раздела позволяют рассмотреть несколько удивительные и важные свойства так называемой двумерной нормальной или гауссовской случайной величины $\boldsymbol{\xi} = (\xi_1, \xi_2)$, для которой плотность распределения

$$f(x,y) = f_{\xi}(x,y;a_1,a_2,\sigma_1,\sigma_2,b) = \\ = (2\pi\sigma_1\sigma_2)^{-1}(1-b)^{-1/2}\exp\{-(2-2b)^{-1}[\sigma_1^{-2}(x-a_1)^2 + \\ + \sigma_2^{-2}(y-a_2)^2 - 2b^{1/2}(\sigma_1\sigma_2)^{-1}(x-a_1)(y-a_2)]\}$$

и зависит от пяти параметров: $-\infty < a_1 < +\infty$, $-\infty < a_1 < +\infty$, $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 < b < 1$. Используя формулы (7.15), замену типа (10.16) для переменных x и y, вычислим одномерные плотности $f_{\xi_1}(x)$ и $f_{\xi_2}(x)$ для случайных величин ξ_1 и ξ_2: $f_{\xi_1}(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = (2\pi)^{-1/2}\sigma_1^{-1}\exp\{-(x-a_1)^2/(2\sigma_1^2)\}$, $f_{\xi_2}(x) = \int_{-\infty}^{+\infty} f(x,y)\,dx = (2\pi)^{-1/2}\sigma_2^{-1}\exp\{-(x-a_2)^2/(2\sigma_2^2)\}$. Отсюда сразу следует, что ξ_1 является нормальной с математическим ожиданием $\mathsf{M}\xi_1 = a_1$ и дисперсией $\mathsf{D}\xi_1 = \sigma_1^2$. Аналогично утверждаем, что ξ_2 является гауссовской с математическим ожиданием $\mathsf{M}\xi_2 = a_2$ и дисперсией $\mathsf{D}\xi_2 = \sigma_2^2$. Применяя формулу (8.22) и методику из [4], найдем ковариацию величин ξ_1 и ξ_2: $\mathsf{cov}(\xi_1, \xi_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x - \mathsf{M}\xi_1)(y - \mathsf{M}\xi_2)f(x,y)\,dx\,dy = b^{1/2}\sigma_1\sigma_2$. Значит, коэффициентом корреляции $\mathsf{corr}(\xi_1, \xi_2)$ случайных величин ξ_1 и ξ_2 равен $b^{1/2}$. Эти подсчеты раскрывают смысл всех параметров a_1, a_2, σ_1, σ_2, b двумерной гауссовской плотности. При $b \neq 0$ плотность $f(x,y) \neq f_{\xi_1}(x)\,f_{\xi_2}(y)$. Тогда случайные величины являются зависимыми и коррелированными. Пусть теперь $b = 0$, т. е величины ξ_1 и ξ_2 являются некоррелированными. Легко проверить, что в этом случае $f(x,y) = f_{\xi_1}(x)\,f_{\xi_2}(y)$, и, следовательно, случайные величины ξ_1 и ξ_2 являются независимыми. Поэтому некоррелированность и независимость двух нормальных случайных величин ξ_1, ξ_2 являются эквивалентными понятиями.

2.2. Равномерный закон распределения. Рассмотрим теперь наиболее простой класс непрерывных случайных величин, который часто используется в практических задачах.

Определение 10.7. Случайная величина ξ называется *равномерной*, а соответствующее распределение — *равномерным на конечном промежутке* $(a, b]$, если $f_\xi(x) = 0$ при $x \notin (a, b]$ и $f_\xi(x) = (b - a)^{-1}$ при $x \in (a, b]$.

Приведем примеры равномерно распределенных случайных величин:

1) ошибка, вызванная разными способами округления показаний измерительного прибора до целых делений шкалы, например, за приближенное значение измеряемой характеристики выбирается ближайшее большее целое, ближайшее меньшее целое и, наконец, ближайшее целое;
2) расстояние между электростанцией и местом обрыва кабеля, по которому передается электрическая энергия потребителю;
3) время ожидания пассажиром прибытия автобуса при регулярном расписании на маршруте и случайном прибытии пассажира на остановку;
4) ошибка при компьютерных расчетах, когда вычисления проводятся с точностью, например, до второго знака после запятой.

Простой аналитический вид $f_\xi(x)$ позволяет определить интегральную функцию равномерной случайной величины: $F_\xi(x) = 0$ при $x \leqslant a$ и $F_\xi(x) = \int_{-\infty}^{x} f_\xi(t)\,dt = (b - a)^{-1}\int_a^x dt = (x - a)(b - a)^{-1}$ при $a < x \leqslant b$, наконец, $F_\xi(x) = 1$ при $x > b$. Определим основные числовые характеристики равномерной случайной величины ξ. Легко получим, что $\mathsf{M}\xi = \int_{-\infty}^{x} x\,f_\xi(x)\,dx = (b - a)^{-1}\int_a^b x\,dx = (b^2 - a^2)\,2^{-1}(b - a)^{-1} = 2^{-1}(b + a)$. Нетрудно видеть, что $\mathsf{Me}\,\xi = (a + b)2^{-1} = \mathsf{M}\xi$. Отметим также, что любое число из промежутка $(a,b]$ является модой $\mathsf{Mo}\,\xi$ и наивероятнейшим значением $\mathsf{Mo}^*\xi$. Поэтому использование таких числовых характеристик, как моды и наивероятнейшего значения для равномерной случайной величины, теряет всякий смысл. Вычислим дисперсию и среднее квадратическое отклонение для величины ξ:

$$\mathsf{D}\xi = \mathsf{M}(\xi^2) - (\mathsf{M}\xi)^2,$$

$$\mathsf{M}(\xi^2) = (b-a)^{-1}\int\limits_a^b x^2dx = (b^3-a^3)3^{-1}(b-a)^{-1} = (a^2+ab+b^2)\,3^{-1},$$

$$\mathsf{D}\xi = (a^2+ab+b^2)3^{-1} - (a+b)^2\,4^{-1} = (b-a)^2/12, \sigma\xi = (b-a)/2\sqrt{3}\,.$$

График плотности распределения $f_\xi(x)$ и график интегральной функции $F_\xi(x)$ равномерной случайной величины представлены соответственно на рис. 10.9 и на рис. 10.10.

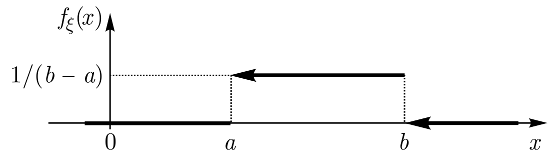


Рис. 10.9

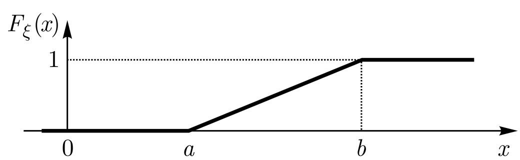


Рис. 10.10

С помощью формул для среднего отклонения, центральных моментов k-го порядка, коэффициента асимметрии и эксцесса найдем теперь вспомогательные числовые характеристики равномерной величины:

$$\mathsf{E}\xi = \mathsf{M}(|\xi - \mathsf{Me}\,\xi|) = \int\limits_{-\infty}^{+\infty} |x-(a+b)2^{-1}|f_\xi(x)\,dx =$$

$$= (b-a)^{-1}\int\limits_a^{(a+b)/2}(a+b-2x)2^{-1}\,dx + \\ + (b-a)^{-1}\int\limits_{(a+b)/2}^{b}(2x-a-b)2^{-1}\,dx =$$

$$= 2^{-1}(b-a)^{-1}\int\limits_0^{b-a} t\,dt = (b-a)/4,$$

$$\beta_k(\xi) = \int\limits_{-\infty}^{+\infty} (x - \mathsf{M}\xi)^k f_\xi(x)\,dx = (b-a)^{-1}\int\limits_a^b (x - (a+b)\,2^{-1})^k\,dx =$$
$$= 2^{-1-k}(b-a)^{-1}\int\limits_a^b (2x - a - b)^k\,d(2x-a-b) =$$
$$= 2^{-1-k}(b-a)^{-1}\int\limits_{a-b}^{b-a} t^k\,dt = 2^{-1\,-\,k}(b-a)^{-1}(k+1)^{-1}t^{k+1}\Big|_{a-b}^{b-a}.$$

Отсюда получаем, что центральный момент k-го порядка $\beta_k(\xi) = 0$ при k нечетном и $2^{-1}(k+1)^{-1}(b-a)^k$ при k четном. Поэтому $\text{Э}(\xi) = (\sigma\xi)^{-4}\beta_4 - 3 = 2^{-4}5^{-1}(b-a)^4((b-a)/2\sqrt{3}\,)^{-4} = 5^{-1}9 - 3 = -6/5$ и $\mathsf{Ka}(\xi) = \beta_3\xi/(\sigma\xi)^3 = 0$. На содержательном уровне равенства $\mathsf{Ka}(\xi) = 0$ и $\text{Э}(\xi) = -6/5$ имеют следующий смысл. График равномерного распределения симметричен относительно центра рассеивания и, значит, коэффициент асимметрии должен быть равен нулю. Далее, так как эксцесс равномерной случайной величины является отрицательным числом, то кривая нормального распределения имеет большую крутизну наклона по сравнению с кривой равномерного распределения. Рассмотрим решение некоторых вопросов с применением равномерного распределения.

Пример 10.14. Маршрутные такси прибывают регулярно на остановку с интервалом 3 минуты. Некто приходит на остановку в случайный и не связанный с расписанием момент времени. Найти среднее квадратическое отклонение $\boldsymbol{\sigma}\xi$ времени ξ ожидания пассажиром такси. Вычислить вероятность того, что пассажир будет ожидать такси не менее $\mathsf{M}\xi + 1{,}6\,\boldsymbol{\sigma}\xi$.

Решение. В дальнейшем для упрощения мы опускаем в записи размерность времени. Элементарное случайное событие, которое заключается в том, что пассажир приходит на остановку в момент времени t, будем обозначать через $\{t\}$. Так как маршрутные такси прибывают на остановку по жесткому расписанию, то в качестве пространства описаний всех элементарных исходов можно считать множество $\Omega = \{\omega = t\colon 0 \leqslant t \leqslant 3\}$. Случайным событием в этом эксперименте является борелевское множество на отрезке $[\,0,3\,]$. Тогда множество $\mathcal{F}$ всех наблюдаемых исходов совпадает с борелевской σ-алгеброй на $[\,0,3\,]$. Из условия задачи следует, что каждый момент прихода пассажира на остановку одинаково возможен. Поэтому вероятности случайных событий для этого опыта определяются на основании геометрического подхода. Следовательно, мы построили вероятностную модель $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ случайного прибытия пассажира на остановку. Если теперь $\eta(\omega) = \omega = t$ есть время прихода пассажира на остановку, то множество

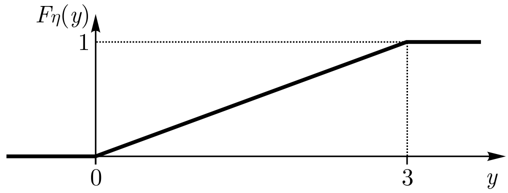


Рис. 10.11

вида $\{\omega\colon \eta(\omega) < y\}$ равно $\varnothing \in \mathcal{F}$ при $y \leqslant 0$, равно $[0, y) \in \mathcal{F}$ при $0 < y \leqslant 3$ и, наконец, равно $\Omega \in \mathcal{F}$ при $y > 3$. Отсюда вытекает, что $\eta(\omega)$ является равномерной случайной величиной, для которой интегральная функция $F_\eta(y)$ представлена на рис. 10.11.

Если η — время прихода пассажира на остановку, то его время ожидания такси равно $\xi = 3 - \eta(\omega)$. Поэтому интегральная функция

$$F_\xi(x) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\}) = \mathsf{P}(\{\omega\colon 3 - \eta(\omega) < x\}) = \\ = 1 - \mathsf{P}(\{\omega\colon \eta(\omega) < 3 - x\}) = 1 - F_\eta(3 - x).$$

Отсюда $F_\xi(x) = 0$ при $x \leqslant 0$, $F_\xi(x) = x/3$ при $0 < x \leqslant 3$ и, наконец, $F_\xi(x) = 1$ при $x > 3$. Итак, время прихода пассажира на остановку и время его ожидания такси являются равномерными случайными величинами на отрезке $[0, 3]$. Отсюда $\mathsf{M}\xi = 1{,}5$ и $\boldsymbol{\sigma}\xi = 2^{-1} \times 3^{1/2}$. Поэтому вероятность того, что пассажир будет ожидать такси не менее $\mathsf{M}\xi + 1{,}6\,\boldsymbol{\sigma}\xi = 1{,}5 + 0{,}8\sqrt{3}$, равна

$$1 - F_\xi(1{,}5 + 0{,}8\sqrt{3}\,) = 1 - 3^{-1}(1{,}5 + 0{,}8\sqrt{3}\,) \approx 0{,}0381.$$

В заключение рассмотрения свойств равномерной случайной величины укажем на ее исключительное место, которое она занимает среди всех других случайных величин. Это связано с имитацией сложных реальных экспериментов на современных компьютерах. В этом случае из равномерного распределения на отрезке $[0, 1]$ можно получить любые другие распределения, которые определяют вероятностные свойства измерителей случайных экспериментов. Решение этого вопроса было дано при рассмотрении задачи имитации в гл. 9.

2.3. Распределение хи-квадрат с r степенями свободы. Сначала рассмотрим задачу, при решении которой возникает непрерывная случайная величина с таким типом распределения.

Пример 10.15. Пусть $\xi_1, \xi_2, \ldots, \xi_r$ являются независимыми случайными величинами, каждая из которых распределена по стандартному нормальному закону. Найти плотность распределения суммы квадратов этих случайных величин или случайной величины $\chi_r^2 = \xi_1^2 + \xi_2^2 + \ldots + \xi_r^2$.

Решение. Пусть $\Gamma(c) = \int_0^\infty x^{c-1} \exp\{-x\}\, dx$ есть гамма-функция Эйлера с параметром $c > 0$. Из курса математического анализа известно, что $\Gamma(1) = 1$, $\Gamma(1/2) = \pi^{1/2}$, $\Gamma(c+1) = c\,\Gamma(c)$. Методом математической индукции по r докажем, что плотность $f_{\chi_r^2}(y)$ случайной величины χ_r^2 равна нулю при $y \leqslant 0$ и равна $(2^{r/2}\Gamma(r/2))^{-1} y^{r/2-1} \exp\{-y/2\}$ при $y > 0$. Рассматривая решение примера 9.2 при $\sigma = 1$ и учитывая равенство $\Gamma(1/2) = \pi^{1/2}$, получаем, что формула для плотности $f_{\chi_r^2}(y)$ справедлива при $r = 1$. Если формула для плотности $f_{\chi_r^2}(y)$ имеет место для некоторого натурального числа $r > 1$, то, используя равенство $\chi_{r+1}^2 = \chi_r^2 + \xi_{r+1}^2$, независимость случайных величин χ_r^2 и ξ_{r+1}^2, вид плотности распределения для каждой из случайных величин χ_r^2, ξ_{r+1}^2 и, наконец, замену $xy^{-1} = t$, последовательно при $y > 0$ находим:

$$
\begin{aligned}
f_{\chi_{r+1}^2}(y) &= \int\limits_{-\infty}^{+\infty} f_{(\xi_{r+1})^2}(y-x) f_{\chi_r^2}(x)\, dx = \\
&= \int\limits_0^y (2\pi(y-x))^{-1/2} (2^{r/2}\Gamma(r/2))^{-1} x^{(r/2)-1} \times \\
&\qquad\qquad \times \exp\{-(y-x)/2\} \exp\{-x/2\}\, dx = \\
&= (2\pi)^{-1/2} (2^{r/2}\Gamma(r/2))^{-1} \exp\{-y/2\} \int\limits_0^y (y-x)^{-1/2} x^{(r/2)-1}\, dx = \\
&= (2\pi)^{-1/2} (2^{r/2}\Gamma(r/2))^{-1} y^{[(r+1)/2]-1} \times \\
&\qquad\qquad \times \exp\{-y/2\} \int\limits_0^1 (1-t)^{-1/2} t^{(r/2)-1}\, dt.
\end{aligned}
$$

Напомним, что интеграл Эйлера первого рода $\int_0^1 t^{u-1}(1-t)^{v-1}dt$ или бэта-функция с параметрами u, v удовлетворяет равенству $\int_0^1 t^{u-1}(1-t)^{v-1}dt = \Gamma(u)\,\Gamma(v)/\Gamma(u+v)$. Поэтому при $u = r/2$ и $v = 1/2$ интеграл $\int_0^1 t^{(r/2)-1}(1-t)^{-1/2}dt = \Gamma(r/2)\,\Gamma(1/2)/\Gamma((r+1)\,2^{-1})$. Отсюда получаем

$$
f_{\chi_{r+1}^2}(y) = (2^{(r+1)/2}\Gamma(2^{-1}(r+1)))^{-1} y^{[(r+1)/2]-1} \exp\{-y/2\} \quad \text{при} \quad y > 0.
$$

Так как $\chi_{r+1}^2 = \chi_r^2 + \xi_{r+1}^2 \geqslant 0$, то $f_{\chi_{r+1}^2}(y) = 0$ при $y \leqslant 0$. Следовательно, формула для плотности $f_{\chi_r^2}(y)$ случайной величины χ_r^2 установлена.

Случайная величина χ_r^2 играет большую роль в теории вероятностей и математической статистике. Заметим, что χ_r^2 часто называется случайной величиной типа хи-квадрат с r степенями свободы. Так как $\mathsf{M}\xi_1^2 = \mathsf{M}\xi_2^2 = \ldots = \mathsf{M}\xi_r^2 = \mathsf{D}\xi_1 = 1$ и $\mathsf{M}(\xi_1^2 - 1)^2 = \mathsf{D}\xi_1^2 = \mathsf{D}\xi_2^2 =$

$= \ldots = \mathsf{D}\xi_r^2 = \mathsf{M}(\xi_1^4 - 2\,(\xi_1)^2 + 1) = \beta_4\xi_1 - 2\,\beta_2\xi_1 + 1 = 3 - 2 + 1 = 2$, то математическое ожидание $\mathsf{M}(\chi_r^2) = r$ и дисперсия $\mathsf{D}(\chi_r^2) = 2r$.

Легко видеть, что мода и наивероятнейшее значение для случайной величины типа хи-квадрат с $r = 1$, 2 степенями свободы не существует.

Пусть $r > 2$ и $x > 0$. В этом случае значение производной $df_{\chi_r^2}(x)/dx = x^{r/2-2}(2^{r/2}\Gamma(r/2))^{-1}(2^{-1}r - 1 - 2^{-1}x)\exp\{-x/2\}$ равно нулю только в точке с абсциссой $x = r - 2$. Значит, для случайной величины χ_r^2 при $r > 2$ наивероятнейшее значение $\mathsf{Mo}^*(\chi_r^2) = \mathsf{Mo}\,(\chi_r^2) = r - 2$. На рис. 10.12 приведены графики плотности $f_{\chi_r^2}(y)$ для $r = 1$, 2 и 6.

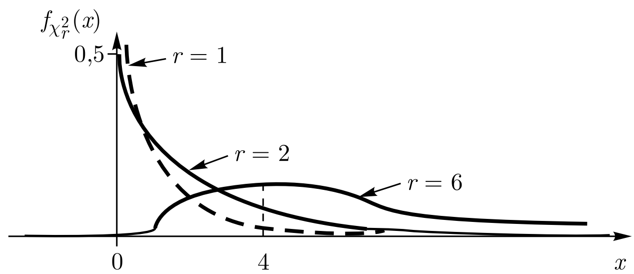


Рис. 10.12

Из определения начального момента k-го порядка и равенства вида $\alpha(\chi_r^2) = \mathsf{M}(\chi_r^2) = r$ для всех $k = 1, 2, \ldots$ непосредственно получаем

$$\begin{aligned}
\alpha_{k+1}(\chi_r^2) &= \int\limits_{-\infty}^{+\infty} x^{k+1} f_{\chi_r^2}(x)\,dx = \\
&= \int\limits_{0}^{+\infty} x^{k+1}(2^{r/2}\Gamma(r/2))^{-1}x^{r/2-1}\exp\{-x/2\}\,dx = \\
&= -2(2^{r/2}\Gamma(r/2))^{-1}\int\limits_{0}^{+\infty} x^{k+1}x^{r/2-1}d(\exp\{-x/2\}) = \\
&= -2x^{k+1}(2^{r/2}\Gamma(r/2))^{-1}x^{r/2-1}\exp\{-x/2\}\big|_0^{\infty} + \\
&+\ 2\,(k + 2^{-1}r)\int\limits_{0}^{+\infty} x^k x^{r/2-1}(2^{r/2}\Gamma(r/2))^{-1}\exp\{-x/2\}\,dx = \\
&= \alpha_k(\chi_r^2)\,2\,(k + 2^{-1}r) = 2^{k+1}\prod_{i=0}^{k}(i + 2^{-1}r).
\end{aligned}$$

Отсюда для случайной величины χ_r^2 легко найдем коэффициент асимметрии $\mathsf{Ka}(\chi_r^2)$ и эксцесс $\mathsf{Э}(\chi_r^2)$. Итак, для распределения хи-квадрат с r степенями свободы имеем формулы:

$$\mathsf{Ka}\,\chi_r^2 = (\sigma\chi_r^2)^{-3}\sum_{k=0}^{3} C_3^k(-1)^{3-k}(\alpha_1(\chi_r^2))^{3-k}\alpha_k(\chi_r^2) = 2^{3/2}r^{-1/2},$$

$$\mathsf{Э}\chi_r^2 = (\boldsymbol{\sigma}\chi_r^2)^{-4}\sum_{k=0}^{4} C_4^k(-1)^{4-k}(\alpha_1(\chi_r^2))^{4-k}\alpha_k(\chi_r^2) - 3 = 12/r.$$

2.4. Показательный, или экспоненциальный, закон распределения. Среди непрерывных случайных величин важное место занимают так называемые экспоненциальные случайные величины.

Определение 10.8. Будем называть случайную величину ξ *распределенной по показательному или экспоненциальному закону*, если ее плотность $f(x)$ равна нулю при $x < 0$ и равна $\lambda\exp\{-\lambda x\}$ при $x \geqslant 0$, где $\lambda = \mathrm{const} > 0$.

Найдем интегральную функцию для показательной случайной величины: $F(x) = 0$ при $x \leqslant 0$ и $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_0^x \lambda\exp\{-\lambda t\}\,dt = 1 - \exp\{-\lambda x\}$ при $x > 0$. Отсюда условие нормировки можно проверить следующим образом: $F(+\infty) = 1 = \int_0^{+\infty}\lambda\exp\{-\lambda t\}\,dt = \int_{-\infty}^{+\infty} f(x)\,dx$. На рис. 10.13 приведены графики плотности $f(x)$ и интегральной функции $F(x)$.

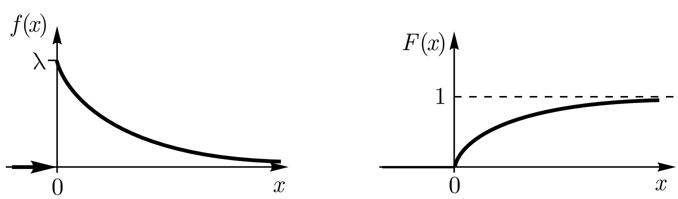


Рис. 10.13

Показательное распределение определяется одним параметром λ. Это обстоятельство очень важно с вычислительной точки зрения, когда требуется оценить параметр закона распределения по наблюдениям за экспоненциальной случайной величиной. Приведем примеры реальных экспериментов, в которых встречается экспоненциальная случайная величина как результат показаний измерителей:

1) время между двумя соседними отказами в процессе эксплуатации некоторого агрегата, который состоит из большого числа элементов;
2) промежуток времени между двумя последовательными поступлениями вызовов от абонентов на телефонную станцию;
3) продолжительность безотказной работы (надежность) многих радиоэлементов;

4) промежуток времени между двумя последовательными распадами атомов некоторого количества радиоактивного вещества;
5) время между двумя последовательными прибытиями автомобилей к стоп-линии перекрестка, который расположен на окраине города;
6) время между двумя последовательными падениями метеоритов на территории Нижегородской области;
7) промежуток времени между двумя последовательными катастрофами на угольных шахтах Азии;
8) промежуток времени между двумя последовательными поступлениями вызовов больных на пункт скорой медицинской помощи;
9) величина времени свободного пробега молекул газа при их движении с заданной скоростью.

Широкое применение показательного закона распределения на практике в значительной мере вызвано его удивительным свойством отсутствия последействия. Отсутствие последействия или памяти на содержательном уровне означает, что распределение оставшейся части показательной случайной величины подчиняется тому же закону и, более того, с тем же параметром. Это свойство, как правило, заметно облегчает теоретические выкладки и вычислительные расчеты при решении важных прикладных задач, в которых используется показательный закон распределения, например, при исследовании систем массового обслуживания, информационных и вычислительных сетей. Перейдем к строгой математической формулировке этого свойства. Если ξ — показательная случайная величина, то $\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u+t\} \mid \{\omega\colon \xi(\omega) \geqslant u\}) = \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant t\}$ при любых значениях $u \geqslant 0$, $t \geqslant 0$. Действительно, условная вероятность вида

$$\begin{aligned}
&\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u+t\}|\{\omega\colon \xi(\omega) \geqslant u\}) = \\
&\qquad = \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u+t\} \cap \{\omega\colon \xi(\omega) \geqslant u\})(\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u\})^{-1} = \\
&\qquad = (\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u\})^{-1}\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u+t\}) = \\
&\qquad = \exp\{-\lambda u\}^{-1}\exp\{-\lambda(u+t)\} = \exp\{-\lambda t\} = \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant t\}).
\end{aligned}$$

Полученный результат можно интерпретировать с помощью большого числа экспериментов. Например, пусть ξ определяет время безотказной работы некоторого узла космической станции, и узел проработал в течение некоторого промежутка времени u. При этих условиях вероятность безотказной работы в течение следующего промежутка времени, длительность которого равна t, та же самая, как если бы узел начал только что работать. Другими словами, оставшееся время безотказной работы узла имеет показательное распределение с тем же параметром λ. Рассмотрим обратное утверждение.

Лемма 10.2. *Пусть ξ — случайная величина с непрерывной плотностью $f(x)$ при $x > 0$ и при любых значениях $u \geqslant 0$, $t \geqslant 0$ имеет место равенство*

$$\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u+t\} \mid \{\omega\colon \xi(\omega) \geqslant u\}) = \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant t\}. \tag{10.17}$$

Тогда случайная величина ξ распределена по показательному закону.

Доказательство. Из (10.17), используя равенство

$$\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u+t\}) \times (\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u\}))^{-1} = \\ = \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u+t\} \mid \{\omega\colon \xi(\omega) \geqslant u\}),$$

получим, что $\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u+t\}) = \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant u\})\,\mathsf{P}(\{\omega\colon \xi(\omega) \geqslant t\})$. Если функция $Q(t) = 1 - F(t) = \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant t\})$, то последнее равенство при любых $u \geqslant 0$, $t \geqslant 0$ можно записать в следующем виде: $Q(u+t) = Q(u)\,Q(t)$. Так как $Q(0) \neq 0$ и $Q(0+0) = Q(0)\,Q(0)$, то $Q(0) = 1$. Дифференцируя равенство $Q(u+t) = Q(u)\,Q(t)$ по u и полагая затем $u = 0$, $(dQ(u)/du)\big|_{u=0} = -\lambda < 0$, найдем $dQ(t)/dt = -\lambda Q(t)$. Отсюда с учетом $Q(0) = 1$ выводим, что $Q(t) = \exp\{-\lambda t\}$ при любом $t \geqslant 0$. Значит, интегральная функция $F(t) = 1 - \exp\{-\lambda t\}$ при $t \geqslant 0$ и $F(t) = 0$ при $t < 0$. Лемма 10.2 доказана.

Теперь можно утверждать, что равенство (10.17) или свойство отсутствия последействия является характеристическим для показательного распределения.

Напомним, что среди всех дискретных случайных величин свойством отсутствия последействия обладает только случайная величина с геометрическим распределением. Более того, геометрическое распределение можно считать дискретным аналогом показательного распределения. Действительно, пусть каждое испытание некоторого опыта E_0 занимает h единиц времени и его удачный исход A наступает с вероятностью p. Опыт E_0 независимым образом повторяется до появления первого неудачного исхода $\overline{A}$. Тогда случайное число η всех испытаний имеет геометрическое распределение: $\mathsf{P}(\{\omega\colon \eta(\omega) = m\}) = p^{m-1}(1-p)$, $m = 1, 2, \ldots$. Отсюда легко вычислить, что $\mathsf{P}(\{\omega\colon \eta(\omega) > k\}) = \sum_{m=k+1}^{\infty} p^{m-1}(1-p) = p^k$, где $k = 1, 2, \ldots$. Очевидно, что случайная величина $\xi_h = \eta\, h$ определяет время всех испытаний над опытом E_0 и принимает значение h, $2h, \ldots$. Имеем: $\mathsf{P}(\{\omega\colon \eta(\omega) > k\}) = \mathsf{P}(\{\omega\colon \eta(\omega)\, h > kh\}) = \mathsf{P}(\{\omega\colon \xi_h > kh\}) = \mathsf{P}(\{\omega\colon \xi_h > t\}) = p^k = p^{t/h} = \exp\{-\lambda t\}$, где $t = kh$ является дискретным временем, $p = \exp\{-\lambda h\}$ и постоянная величина $\lambda = -h^{-1} \ln p > 0$.

Рассмотрим теперь произвольное неотрицательное число t. Тогда существует такое натуральное число n, что при заданном значении h будут выполняться неравенства $(n-1)\,h \leqslant t < nh$. Отсюда выводим соотношение $\{\omega\colon \xi_h(\omega) > (n-1)\,h\} \supset \{\omega\colon \xi_h(\omega) > t\} \supset \{\omega\colon \xi_h(\omega) > nh\}$. Это соотношение позволяет написать следующие неравенства:

$$\mathsf{P}(\{\omega\colon \xi_h(\omega) > (n-1)\,h\}) \geqslant \mathsf{P}(\{\omega\colon \xi_h(\omega) > t\}) \geqslant \mathsf{P}(\{\omega\colon \xi_h(\omega) > nh\}).$$

Следовательно, вероятность $\mathsf{P}(\{\omega\colon \xi_h(\omega) > t\})$ удовлетворяет неравенствам

$$\exp\{-\lambda(n-1)\,h\} \geqslant \mathsf{P}(\{\omega\colon \xi_h(\omega) > t\}) \geqslant \exp\{-\lambda nh\}. \tag{10.18}$$

Пусть теперь h стремится к нулю и, значит, вероятность $\mathsf{P}(\overline{A}) = 1 - p$ наступления неудачного исхода $\overline{A}$ в каждом испытании убывает по закону $1 - \exp\{-\lambda h\} = \lambda h + o(h)$. Иначе, при бесконечно малой величине h вероятность наступления неудачного исхода в каждом испытании асимптотически пропорциональна h с коэффициентом пропорциональности λ. Постоянная λ может быть также интерпретирована как среднее число наступления неудачных исходов в каждом испытании за единицу времени при бесконечно малой величине h. Так как $\lim\limits_{h\to 0}(n-1)\,h = \lim\limits_{h\to 0} nh = t$, то из (10.18) при $h \to 0$ сразу находим $\lim\limits_{h\to 0} \mathsf{P}(\{\omega\colon \xi_h(\omega) > t\}) = \exp\{-\lambda t\}$ для всех чисел $t \geqslant 0$. На содержательном уровне можно сказать, что показательное распределение является предельным для геометрического.

Для определения числовых характеристик $\mathsf{M}\xi$, $\mathsf{D}\xi$, $\boldsymbol{\sigma}\xi$, $\mathsf{Ka}\,\xi$ и $\mathsf{Э}\xi$ случайной величины ξ, распределенной по показательному закону, сначала вычислим начальные моменты k-го порядка. Так как

$$\begin{aligned} \alpha_k\xi = \mathsf{M}\xi^k &= \int\limits_{-\infty}^{+\infty} x^k f(x)\,dx = \int\limits_{0}^{+\infty} \lambda x^k \exp\{-\lambda x\}\,dx = \\ &= -\int\limits_{0}^{+\infty} x^k d(\exp\{-\lambda x\}) = -x^k \exp\{-\lambda x\}\Big|_0^{\infty} + \\ &\quad + k\lambda^{-1} \int\limits_{0}^{+\infty} x^{k-1}\lambda \exp\{-\lambda x\}\,dx = k\lambda^{-1}\alpha_{k-1}\xi \end{aligned}$$

при $k \geqslant 1$ и $\alpha_0\xi = 1$, то имеет место равенство $\alpha_k\xi = k!\,\lambda^{-k}$. Теперь последовательно выводим:

$$\mathsf{M}\xi = \alpha_1\xi = 1/\lambda, \quad \mathsf{D}\xi = \alpha_2\xi - (\mathsf{M}\xi)^2 = 2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2,$$

$$\boldsymbol{\sigma}\xi = 1/\lambda, \quad \beta_3\xi = \sum_{k=0}^{3} C_3^k(-1)^{3-k}(\alpha_1\xi)^{3-k}\alpha_k\xi = 2/\lambda^3,$$

$$\mathsf{Ka}\,\xi = (\boldsymbol{\sigma}\xi)^{-3}\beta_3\xi = 2, \quad \beta_4\xi = \sum_{k=0}^{4} C_4^k(-1)^{4-k}(\alpha_1\xi)^{4-k}\alpha_k\xi = 9/\lambda^4,$$

$$\mathsf{Э}\xi = (\boldsymbol{\sigma}\xi)^{-4}\beta_4\xi - 3 = 6.$$

Дифференцируя плотность $f(x)$, получим, что значение производной $df(x)/dx$ равно нулю при $x < 0$ и равно $(-\lambda^2)\exp\{-\lambda x\} < 0$ при $x > 0$. Производная в точке $x = 0$ не существует. Поэтому для показательной случайной величины ξ наивероятнейшее значение $\mathsf{Mo}^*\xi = \mathsf{Mo}\,\xi = 0$. Медиана показательной случайной величины вычисляется из условия $F(\mathsf{Me}\,\xi) = 1 - \exp\{-\lambda\,\mathsf{Me}\,\xi\} = 2^{-1}$. Решая это простое уравнение, получим, что $\mathsf{Me}\,\xi = -\lambda^{-1}\ln 2$.

Для некоторых реальных экспериментов медиана $\mathsf{Me}\,\xi$ показательной случайной величины имеет вполне содержательный физический смысл. Например, если случайную величину ξ трактовать как время распада радиоактивного атома, то величина $\mathsf{Me}\,\xi = \lambda^{-1}\ln 2$ в физике называется периодом полураспада. Такое название величины $\mathsf{Me}\,\xi = \lambda^{-1}\ln 2$ можно пояснить следующими обстоятельствами. Для непрерывной случайной величины вероятность $\mathsf{P}(\{\omega\colon \xi(\omega) < \mathsf{Me}\,\xi\}) = \mathsf{P}(\{\omega\colon \xi(\omega) > \mathsf{Me}\,\xi\}) = 2^{-1}$. Поэтому вероятность того, что отдельный атом не распадется за время $\lambda^{-1}\ln 2$, равна 2^{-1}. Естественно считать, что в начальный момент $t = t_0$ в некотором количестве радиоактивного вещества содержится большое число n атомов. За время $\lambda^{-1}\ln 2$ каждый из n атомов независимым образом не распадается с вероятностью $p = 2^{-1}$ и распадается с вероятностью $q = 2^{-1}$. Отсюда следует, что процесс распада n атомов радиоактивного вещества за время $\lambda^{-1}\ln 2$ является схемой независимых испытаний с параметрами $p = 2^{-1}$ и n. Значит, число η атомов, которые не распадутся, является биномиальной случайной величиной. Эта случайная величина при достаточно больших n с максимальной вероятностью принимает значение $\mathsf{Mo}^*\eta = [(n+1)\,p]$, которое равно $2^{-1}n = \mathsf{M}\eta$ при четном n и равно величине $2^{-1}(n+1) = \mathsf{M}\eta + 2^{-1}$ при нечетном n. Таким образом, при достаточно больших n за время $\lambda^{-1}\ln 2$ полураспада каждого атома с максимальной вероятностью не распадается приблизительно половина того количества радиоактивного вещества, которое было в начальный момент. Очевидно, что число $(n - \eta)$ атомов, которые распадутся, является также биномиальной случайной величиной. Так как в этой задаче $p = q = 2^{-1}$, то, другими словами, можно сделать и такой вывод: вероятнее всего при достаточно больших значениях n за время $\lambda^{-1}\ln 2$ распадается половина первоначального количества радиоактивного вещества.

Рассмотрим простейшие примеры применения некоторых свойств показательного закона.

Пример 10.16. Изотоп радия ^{223}Ra с массовым числом 223 испытывает самопроизвольный распад с образованием изотопа радона ^{219}Rn с массовым числом 219 и испусканием альфа-частицы ^{4}He. Среднее время жизни ядра ^{223}Ra до его распада равно 16,883 суток. Период полураспада ядра ^{223}Ra равен 11,7 суток. Наблюдается процесс распада образца радия ^{223}Ra, масса которого в начальный момент равна 373,079 г. Определить вероятность того, что через 11,7 суток масса этого образца примет значение от $(186{,}5395 - 746{,}158 \times 10^{-12})$ г до $(186{,}5395 + 746{,}158 \times 10^{-12})$ г.

Решение. Пусть случайная величина ξ определяет время распада радиоактивного ядра атома ^{223}Ra. Современная физика предполагает, что время ξ распада радиоактивного ядра атома ^{223}Ra является экспоненциальной случайной величиной. Отсюда, используя условие задачи, получаем, что $\mathsf{M}\xi = \lambda^{-1} = 16{,}883$ суток. Поэтому период полураспада атома ^{223}Ra равен $\mathsf{M}\xi = \lambda^{-1}\ln 2 = 11{,}7$ суток. Так как масса протона равна $1{,}673 \times 10^{-24}$ г, то масса атома ^{223}Ra равна $223 \times 1{,}673 \times 10^{-24} = 373{,}079 \times 10^{-24}$ г. Рассматриваемый образец радия ^{223}Ra, масса которого в начальный момент равна 373,079 г, содержит $n = 373{,}079 \times (373{,}079)^{-1} \times 10^{24} = 10^{24}$ атомов. Значит, число η атомов, которые не распадутся за период полураспада, является биномиальной случайной величиной с параметрами $p = 2^{-1}$ и $n = 10^{24}$. Применяя приближенную формулу (4.14) для биномиальных вероятностей $\mathsf{P}\Big(\cup_{m=k}^{s} B_m\Big)$ при $k = np - 4(npq)^{1/2}$ и $s = np + 4(npq)^{1/2}$, легко найдем

$$\begin{aligned}\mathsf{P}(\{\omega\colon np - 4(npq)^{1/2} \leqslant \eta(\omega) \leqslant np + 4(npq)^{1/2} \approx \\ \approx 2^{-1}\Phi((np + 4(npq)^{1/2} - np)(npq)^{-1/2}) - \\ - 2^{-1}\Phi((np - 4(npq)^{1/2} - np)/(npq)^{-1/2}) = \Phi(4) \approx 0{,}999936.\end{aligned}$$

При $\varepsilon = 4(npq)^{1/2} \times 373{,}079 \times 10^{-24} = 746{,}158 \times 10^{-12}$ г, $p = 2^{-1}$ и $n = 10^{24}$ легко вычислить, что $373{,}079 \times 10^{-24}(np - 4(npq)^{1/2}) = 186{,}5395 \times 746{,}158 \times 10^{-12}$, $373{,}079 \times 10^{-24}(np + 4(npq)^{1/2}) = 186{,}5395 + 746{,}158 \times 10^{-12}$. Случайная величина $\eta_1 = \eta(\omega) \times 373{,}079 \times 10^{-24}$ г определяет массу образца через 11,7 суток. Далее, ради простоты, опускаем размерность физических величин. Учитывая все это, можем написать следующее соотношение:

$$\begin{aligned}\mathsf{P}(\{\omega\colon np - 4(npq)^{1/2} \leqslant \eta(\omega) \leqslant np + 4(npq)^{1/2}\}) = \\ = \mathsf{P}(\{\omega\colon 186{,}5395 - \varepsilon \leqslant \eta(\omega) \times 373{,}079 \times 10^{-24} \leqslant 186{,}5395 - \varepsilon\}) = \\ = \mathsf{P}(\{\omega\colon 186{,}5395 - \varepsilon \leqslant \eta_1(\omega) \leqslant 186{,}5395 - \varepsilon\}) \approx 0{,}999936.\end{aligned}$$

Из соотношения $186{,}5395 > \varepsilon = 746{,}158 \times 10^{-12}$ и приведенных ранее численных результатов можно сделать следующий вывод. Практически является достоверным, что образец ^{223}Ra массой 373,079 г будет содержать 186,5395 г изотопа радия ^{223}Ra через время полураспада или, другими словами, за это время распадается половина первоначальной массы образца.

Пример 10.17. Время ξ безотказной работы электрической лампочки имеет показательное распределение с параметром λ. Определить вероятность того, что в результате эксперимента время безотказной работы электронной лампы отклонится по абсолютной величине от своего среднего значения $\mathsf{M}\xi$ не меньше, чем на $3\boldsymbol{\sigma}\xi$.

Решение. Из свойств показательного распределения имеем равенство $\mathsf{M}\xi = \boldsymbol{\sigma}\xi = \lambda^{-1}$. Отсюда, легко получаем, что искомая вероятность вида $\mathsf{P}(\{\omega\colon |\xi(\omega) - \mathsf{M}\xi| \geqslant 3\boldsymbol{\sigma}\xi\}) = \mathsf{P}(\{\omega\colon \xi(\omega) \leqslant -2\boldsymbol{\sigma}\xi\}) + \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant 4\boldsymbol{\sigma}\xi\}) = \mathsf{P}(\{\omega\colon \xi(\omega) \geqslant 4\boldsymbol{\sigma}\xi\}) = \exp\{-4\} \approx 0{,}0183$. Как правило, в условиях работы электрической лампочки мы можем пренебречь возможностью появления события, вероятность которого меньше величины 0,019. Поэтому на содержательном уровне результат решения этой задачи означает, что практически достоверно лампа выйдет из строя в течение времени $4\,\mathsf{M}\xi$. Это дает возможность в целях бесперебойного освещения грубо рассчитать количество запасных электрических лампочек на некоторый промежуток времени, например, на год.

2.5. Связь между пуассоновской случайной величиной дискретного типа и показательной случайной величиной непрерывного типа. Пусть имеется эволюционный эксперимент $E = \{E_t\colon 0 \leqslant t < \infty\}$, в котором некоторая его количественная характеристика является пуассоновской случайной величиной. Покажем, что для этого эксперимента можно рассмотреть другую случайную величину (количественную характеристику) с показательным законом распределения, и наоборот. С этой целью предположим, что при проведении эксперимента $E = \{E_t\colon 0 \leqslant t < \infty\}$ мы интересуемся наступлениями некоторого его исхода или события A. Ради определенности интерпретации исходов эволюционного эксперимента E в качестве примера такого эксперимента можно взять процесс свободного движения машин по магистрали или процесс распада некоторого количества радиоактивного вещества с течением времени. В последнем случае событие A означает распад атома. Пусть теперь случайная величина $\xi(t)$ подсчитывает число наблюдений события A за промежуток времени $[0, t)$ и, значит, является количественной характеристикой эволюционного эксперимента $E = \{E_t\colon 0 \leqslant t < \infty\}$. Другими словами, $\xi(t)$ означает число наблюдений события A, предшествующих моменту времени t. Тогда случайная величина $\xi(t_1, t_2) = \xi(t_2) - \xi(t_1)$, которая зависит от двух параметров $0 \leqslant t_1 < t_2$, показывает сколько раз произошло событие A за любой фиксированный промежуток времени $[t_1, t_2)$. Очевидно, что

$\xi(0, t_2) = \xi(t_2)$ при $\xi(0) \equiv 0$. Для большого числа реальных эволюционных экспериментов, в частности для процесса распада некоторого количества радиоактивного вещества, оказывается, что при любых значениях $t_2 > t_1 \geqslant 0$ величина $\xi(t_1, t_2)$ распределена по закону Пуассона:

$$\mathsf{P}(\{\omega\colon \xi(t_1, t_2) = m)\} = \lambda^m (t_2 - t_1)^m (m!)^{-1} \exp\{-\lambda(t_2 - t_1)\}, \quad m = 0, 1, \ldots. \tag{10.19}$$

Здесь λ означает среднее число наблюдений событий A за единицу времени и величина $\lambda(t_2 - t_1)$ определяет среднее число наблюдений событий A за время $(t_2 - t_1)$. Более того, для таких экспериментов часто выполняется условие независимости величин $\xi(t_1)$, $\xi(t_2) - \xi(t_1), \ldots,$ $\xi(t_k) - \xi(t_{k-1})$ при любых конечных целых $k > 1$ и $0 < t_1 < t_2 < \ldots < t_k \leqslant T$, которое называется свойством отсутствия последействия для семейства $\{\xi(t)\colon t \geqslant 0\}$ случайных величин. Иначе: в теории вероятностей математический объект $\{\xi(t)\colon t \geqslant 0\}$ называют *семейством с независимыми приращениями*. В этих предположениях семейство $\{\xi(t)\colon t \geqslant 0\}$ случайных величин является процессом Пуассона с параметром λ (см. задачу о свободном движении машин по магистрали).

Для процесса Пуассона $\{\xi(t)\colon t \geqslant 0\}$ непосредственно из равенств (10.19) при $t_2 = t_1 + \Delta t$ и $\Delta t > 0$ получим так называемое свойство ординарности следующего вида: $\mathsf{P}(\{\omega\colon \xi(t_1, t_1 + \Delta t) > 1)\} = 1 - \mathsf{P}(\{\omega\colon \xi(t_1, t_1 + \Delta t) = 0)\} - \mathsf{P}(\{\omega\colon \xi(t_1, t_1 + \Delta t) = 1)\} = 1 - \exp\{-\lambda\Delta t\} - (\lambda\Delta t)\exp\{-\lambda\Delta t\} = o(\Delta t)$. Заметим, что свойство ординарности в виде равенства (10.3) уже встречалось при рассмотрении задачи о свободном движении машин по магистрали. На содержательном уровне оно означает, что практически невозможно наблюдение не менее двух событий A в один и тот же момент времени. Для подтверждения этого факта зафиксируем любой промежуток времени $[0, t)$ и разобьем его на n частей, длина каждого из которых равна $\Delta t = t/n$. Пусть событие $B_t = \cup_{k=1}^n \{\omega\colon \xi(t(k-1)n^{-1}, tkn^{-1}) > 1\}$ означает наступление двух или более событий A, по крайней мере, в одном из промежутков времени $[t(k-1)n^{-1}, tkn^{-1})$, $k = 1, 2, \ldots, n$. Отсюда, используя свойство ординарности процесса Пуассона, получим, что вероятность

$$\mathsf{P}\Big(\bigcup_{k=1}^{n} \{\omega\colon \xi(t(k-1)n^{-1}, tkn^{-1}) > 1\}\Big) \leqslant \\ \leqslant \sum_{k=1}^{n} \mathsf{P}(\{\omega\colon \xi(t(k-1)n^{-1}, tkn^{-1}) > 1\}) = \\ = n\, o(tn^{-1}) = t(\Delta t)^{-1} o(\Delta t) \to 0$$

для фиксированного $t > 0$, если $n \to \infty$. Ясно, что при $n \to \infty$ длина Δt каждого из промежутков $[t(k-1)n^{-1}, tkn^{-1})$ стремится к нулю.

Обозначим через τ_1, $\tau_2, \ldots$ абсциссы последовательных моментов, в которые можем наблюдать событие A после выбранного начального момента с абсциссой $t = 0$. Здесь τ_1 есть момент времени первого наблюдения события A, τ_2 есть момент времени второго наблюдения события A и т. д. В силу приведенного в предыдущем абзаце свойства будем отмечать на оси времени $0t$ эти моменты разными точками (рис. 10.14).

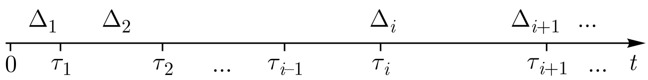


Рис. 10.14

Случайные моменты τ_1, $\tau_2, \ldots$ задают последовательность случайных величин $\Delta_1 = \tau_1$, $\Delta_2 = \tau_2 - \tau_1$, $\Delta_3 = \tau_3 - \tau_2, \ldots$. Случайная величина вида $\Delta_i = \tau_i - \tau_{i-1}$ при каждом $i > 1$ определяет промежуток времени между двумя последовательными и ближайшими наблюдениями события A. При этом интервал Δ_1 несколько отличается от остальных, так как в начальный момент могло и не быть наблюдения события A. Следующие два утверждения непосредственно устанавливают связь между пуассоновской случайной величиной и показательной случайной величиной.

Лемма 10.3. *Если семейство $\{\xi(t)\colon t \geqslant 0\}$ случайных величин является процессом Пуассона с параметром λ, то при любом целом числе $k > 1$ случайные величины Δ_1, $\Delta_2, \ldots, \Delta_k$ независимы в совокупности и распределены по показательному закону с параметром λ.*

Доказательство. Для каждого $i = 1, 2, \ldots$ вероятностные свойства случайной величины Δ_i полностью определяется ее интегральной функцией распределения $F_i(t) = \mathsf{P}(\{\omega\colon \Delta_i(\omega) < t\})$. Из физических соображений ясно, что $\Delta_i \geqslant 0$ при всех $i \geqslant 1$. Поэтому $F_i(t) = \mathsf{P}(\{\omega\colon \Delta_i(\omega) < t\}) = \mathsf{P}(\varnothing) = 0$ при любом $t \leqslant 0$. Рассмотрим сначала величину $\Delta_1 = \tau_1$ и определим ее закон распределения. Так как событие $\{\omega\colon \Delta_i(\omega) \geqslant t\} = \{\omega\colon \xi(t) = 0\}$ при любом $t > 0$, то функция $F_1(t) = \mathsf{P}(\{\omega\colon \Delta_i(\omega) < t\}) = 1 - \mathsf{P}(\{\omega\colon \xi(t) = 0\}) = 1 - \exp\{-\lambda t\}$. Следовательно, случайная величина $\Delta_1 = \tau_1$ распределена по показательному закону с параметром λ.

Перейдем теперь к рассмотрению вероятностных свойств случайной величины Δ_i при каждом $i > 1$. При любом $t_{i-1} > 0$ условная вероятность $\mathsf{P}(\{\omega\colon \Delta_i(\omega) < t\} \mid \{\omega\colon \tau_{i-1} - t_{i-1}\})$ равна вероятности того, что за промежуток $[t_{i-1}, t_{i-1} + t)$ событие A не наступало, или равна вероятности события $\{\omega\colon \eta(t_{i-1}, t_{i-1} + t) = 0\}$. Значит, $\mathsf{P}(\{\omega\colon \Delta_i(\omega) < t\} \mid \{\omega\colon \tau_{i-1} - t_{i-1}\}) = \exp\{-\lambda t\}$ при любом $t_{i-1} > 0$.

Поэтому априорная вероятность $\mathsf{P}(\{\omega\colon \Delta_i(\omega) < t\}) = \exp\{-\lambda t\}$. Отсюда при любом $t > 0$ вычисляем $F_i(t) = \mathsf{P}(\{\omega\colon \Delta_i(\omega) < t\}) = 1 - \mathsf{P}(\{\omega\colon \Delta_i(\omega) \geqslant t\}) = 1 - \exp\{-\lambda t\}$. Итак, случайная величина Δ_i является показательной с параметром λ при $i \geqslant 1$, значит, плотность $f_i(t)$ случайной величины Δ_i равна $\lambda\exp\{-\lambda t\}$ при $t > 0$ и равна нулю при $t \leqslant 0$.

Для доказательства независимости величин Δ_1, $\Delta_2, \ldots, \Delta_k$ рассмотрим последовательность чисел $0 < t_1 - \delta_1 < t_1 < t_1 + \theta_1 < t_2 - \delta_2 < t_2 < t_2 + \theta_2 < \ldots < t_k - \delta_k < t_k < t_k + \theta_k$, где величины δ_i, θ_i для всех $i = 1, 2, \ldots, k$ являются бесконечно малыми. На рис. 10.15 отмечены моменты, абсциссы которых равны элементам этой последовательности.

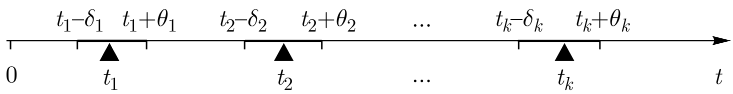


Рис. 10.15

При $t_0 = \theta_0 = 0$ имеем: $\cap_{i=1}^k\{\omega\colon \xi(t_{i-1} + \theta_{i-1}, t_i - \delta_i) = 0, \xi(t_i - \delta_i, t_i + \theta_i) = 1\} = \cap_{i=1}^k\{\omega\colon t_i - \delta_i \leqslant \tau < t_i + \theta_i\}$. Отсюда

$$\mathsf{P}\Big(\bigcap_{i=1}^k\{\omega\colon t_i - \delta_i \leqslant \tau < t_i + \theta_i\}\Big) = \\ = \mathsf{P}\Big(\bigcap_{i=1}^k\{\omega\colon \xi(t_{i-1} + \theta_{i-1}, t_i - \delta_i) = 0,\ \xi(t_i - \delta_i, t_i + \theta_i) = 1\}\Big).$$

Так как при $i = 1, 2, \ldots, k$ величины δ_i, θ_i являются бесконечно малыми и все промежутки $[t_{i-1} + \theta_{i-1}, t_i - \delta_i)$, $[t_i - \delta_i, t_i + \theta_i)$ не пересекаются, то при $m = 0$, $m = 1$ получим

$$\mathsf{P}\Big(\bigcap_{i=1}^k\{\omega\colon \xi(t_{i-1} + \theta_{i-1}, t_i - \delta_i) = 0, \xi(t_i - \delta_i, t_i + \theta_i) = 1\}\Big) = \\ = \prod_{i=1}^k(\exp\{-\lambda(t_i - t_{i-1} - \delta_i - \theta_{i-1})\} \times \lambda(\delta_i + \theta_i)\exp\{-\lambda(\delta_i + \theta_i)\}) = \\ = \exp\{-\lambda\theta_k\}\prod_{i=1}^k(\lambda(\delta_i + \theta_i)\exp\{-\lambda(t_i - t_{i-1})\}) = \\ = \prod_{i=1}^k(\lambda(\delta_i + \theta_i)\exp\{-\lambda(t_i - t_{i-1})\}) + o\Big(\prod_{i=1}^k(\delta_i + \theta_i)\Big).$$

Следовательно, получим, что: $\mathsf{P}\big(\cap_{i=1}^k\{\omega\colon t_i - \delta_i \leqslant \tau_i < t_i + \theta_i\}\big) = \prod_{i=1}^k(\lambda(\delta_i + \theta_i)\exp\{-\lambda(t_i - t_{i-1})\}) + o(\prod_{i=1}^k(\delta_i + \theta_i))$. В силу этого равенства многомерная случайная величина $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_k)$ имеет плотность $f_\tau(t_1, t_2, \ldots, t_k) = \prod_{i=1}^k(\lambda\exp\{-\lambda(t_i - t_{i-1})\})$ при выполнении условий $0 < t_1 < t_2 < \ldots < t_k < \infty$ и $f_\tau(t_1, t_2, \ldots, t_k) \equiv 0$

в противном случае. Здесь произвольное значение случайной величины τ_i обозначено через t_i для всех $i = 1, 2, \ldots, k$. Так как случайные величины $\Delta_1 = \tau_1$, $\Delta_i = \tau_i - \tau_{i-1}$ при каждом $i > 1$, то $\tau_i = \sum_{s=1}^{i} \Delta_s$ для всех $i = 1, 2, \ldots, k$. Произвольное значение случайной величины Δ_i обозначим через z_i для всех $i = 1, 2, \ldots, k$. Теперь, используя функциональные зависимости $t_i = t_i(z_1, z_2, \ldots, z_k) = t_i(\cdot) = \sum_{s=1}^{i} x_s$, $i = 1, 2, \ldots, k$, которые определяют преобразование случайных величин Δ_1, $\Delta_2, \ldots$, Δ_k в случайные величины τ_1, $\tau_2, \ldots$, τ_k, вычислим его якобиан $J(z_1, z_2, \ldots, z_k) = \dfrac{D(t_1(\cdot), t_2(\cdot), \ldots, t_k(\cdot))}{D(z_1, z_2, \ldots, z_k)} = 1 \neq 0$. Поэтому плотность $f_\Delta(z_1, z_2, \ldots, z_k)$ многомерной случайной величины $\Delta = (\Delta_1, \Delta_2, \ldots, \Delta_k)$ равна $f_\tau(t_1(\cdot), t_2(\cdot), \ldots, t_k(\cdot))\,|J| = \prod_{i=1}^{k} (\lambda \exp\{-\lambda(t_i(\cdot) - t_{i-1}(\cdot))\}) = \prod_{i=1}^{k} \lambda \exp\{-\lambda z_i\}$ при $z_1 > 0$, $z_2 > 0, \ldots$, $z_k > 0$ и равна нулю в противном случае. Используя это и показательное распределение каждой случайной величины Δ_i с параметром λ, легко получаем равенство $f_\Delta(z_1, z_2, \ldots, z_k) = \prod_{i=1}^{k} f_i(z_i)$, где $(z_1, z_2, \ldots, z_k) \in R^k$. Итак, доказана статистическая независимость величин Δ_1, $\Delta_2, \ldots$, Δ_k и, тем самым, установлено утверждение леммы 10.3.

Рассмотрим теперь обратное утверждение.

С этой целью на оси времени $0t$ произвольные точки пометим заштрихованными треугольниками (рис. 10.16). Абсциссы этих точек удовлетворяют условию $0 \leqslant t_0 < t_1 < t_1 + \delta_1 < t_2 < t_2 + \delta_2 < \ldots < t_k < t_k + \delta_k < t_0 + t$. При этом величины $\delta_1 < \delta_2 < \ldots < \delta_k$ являются бесконечно малыми и случайные моменты времени $\tau_1 = \Delta_1$, $\tau_2 = \Delta_1 + \Delta_2, \ldots$ появления события A определяются последовательностью случайных величин Δ_1, $\Delta_2, \ldots$.

Пусть величина $\xi(t_0, t_0 + t)$ подсчитывает случайное число наступлений событие A за промежуток времени $[t_0, t_0 + t)$. Имеет место утверждение.

Лемма 10.4. *Пусть при целом $k > 1$ случайные величины Δ_1, $\Delta_2, \ldots$, Δ_k независимы в совокупности и распределены по показательному закону с параметром λ. Тогда для любых фиксированных чисел $t_0 \geqslant 0$, $t > 0$ величина $\xi(t_0, t_0 + t)$ распределена по закону Пуассона с параметром λt.*

Доказательство. Из рис. 10.16 легко получить, что при $\tau_0 \equiv t_0$ событие $B = \left(\bigcap_{i=1}^{k} \{\omega\colon t_i \leqslant \tau_i < t_i + \delta_i\}\right) \cap \{\omega\colon \tau_{k+1} \geqslant t_0 + t\}$, которое

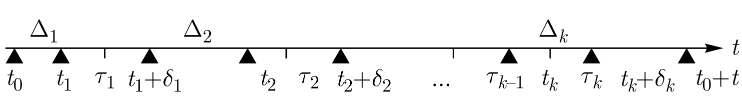


Рис. 10.16

означает наблюдение случайного события A в каждом из промежутков $[t_1, t_1+\delta_1)$, $[t_2, t_2+\delta_2), \ldots$, $[t_k, t_k+\delta_k)$ и отсутствие событий A в остальных точках заданного промежутка $[t_0, t_0+t)$, можно представить в таком виде: $B = \left(\cap_{i=1}^{k}\{\omega\colon t_i - \tau_{i-1} \leqslant \Delta_i < t_i - \tau_{i-1} + \delta_i\}\right) \cap \cap \{\omega\colon \Delta_{k+1} \geqslant t_0 + t - \tau_k\}$. Введем в рассмотрение события

$$B^- = \left(\cap_{i=1}^{k}\{\omega\colon t_i - t_{i-1} \leqslant \Delta_i < t_i - t_{i-1} + \delta_i - \delta_{i-1}\}\right) \cap \\ \cap \{\omega\colon \Delta_{k+1} \geqslant t_0 + t - \tau_k\},$$
$$B^+ = \left(\cap_{i=1}^{k}\{\omega\colon t_i - t_{i-1} - \delta_{i-1} \leqslant \Delta_i < t_i - t_{i-1} + \delta_i\}\right) \cap \\ \cap \{\omega\colon \Delta_{k+1} \geqslant t_0 + t - \tau_k - \delta_k\},$$

где $\delta_0 = 0$. Так как для B^+, B и B^- имеет место соотношение $B^+ \supset B \supset B^-$, то $\mathsf{P}(B^+) \geqslant \mathsf{P}(B) \geqslant \mathsf{P}(B^-)$. Случайные величины Δ_1, $\Delta_2, \ldots$, Δ_{k+1} независимы в совокупности, распределены по показательному закону с параметром λ и величины δ_1, $\delta_2, \ldots$, δ_k являются бесконечно малыми. Поэтому вероятность $\mathsf{P}(B^+)$ равна

$$\exp\{-\lambda(t_0 + t - t_k - \delta_k)\} \times \\ \times \prod_{i=1}^{k}[\lambda(\delta_i + \delta_{i-1})\exp\{-\lambda(t_i - t_{i-1} - \delta_{i-1})\} + o((\delta_i + \delta_{i-1}))] = \\ = \lambda^k \exp\{-\lambda t\}\prod_{i=1}^{k}(\lambda\, dt_i) + o\Big(\prod_{i=1}^{k} dt_i\Big),$$

а вероятность $\mathsf{P}(B^-)$ равна

$$\exp\{-\lambda(t_0 + t - t_k)\} \times \\ \times \prod_{i=1}^{k}[\lambda(\delta_i - \delta_{i-1})\exp\{-\lambda(t_i - t_{i-1})\} + o((\delta_i - \delta_{i-1}))] = \\ = \lambda^k \exp\{-\lambda t\}\prod_{i=1}^{k}(\lambda\, dt_i) + o\Big(\prod_{i=1}^{k} dt_i\Big),$$

где через символы dt_1, $dt_2, \ldots$, dt_k обозначены новые бесконечно малые величины. Отсюда, используя неравенство $\mathsf{P}(B^+) \geqslant \mathsf{P}(B) \geqslant \mathsf{P}(B^-)$, непосредственно получаем, что вероятность

$$\mathsf{P}(B) = \mathsf{P}\Big(\Big(\bigcap_{i=1}^{k}\{\omega\colon t_i \leqslant \tau_i < t_i + \delta_i\}\Big) \cap \{\omega\colon \tau_{k+1} \geqslant t_0 + t\}\Big) = \\ = \lambda^k \exp\{-\lambda t\} \times \prod_{i=1}^{k}(\lambda\, dt_i) + o\Big(\prod_{i=1}^{k} dt_i\Big).$$

Так как dt_1, $dt_2, \ldots, dt_k$ являются бесконечно малыми величинами и имеет место следующее соотношение $0 \leqslant t_0 < t_1 < t_2 < t_2 < \ldots < t_k < t_0 + t$, то вероятность вида

$$\mathsf{P}\Big(\Big(\bigcap_{i=1}^{k}\{\omega\colon t_{i-1} \leqslant \tau_i < t_0 + t\}\Big) \cap \{\omega\colon \tau_{k+1} \geqslant t_0 + t\}\Big) =$$
$$= \lambda^k \exp\{\lambda t\} \int\limits_{t_0}^{t_0+t} \int\limits_{t_1}^{t_0+t} \ldots \int\limits_{t_{k-1}}^{t_0+t} dt_k\, dt_{k-1} \ldots dt_1 = (\lambda t)^k (k!)^{-1} \exp\{-\lambda t\}.$$

Вычисление приведенного выше многократного интеграла можно выполнить, например, с помощью метода математической индукции. Случайное событие $\big(\cap_{i=1}^{k}\{\omega\colon t_{i-1} \leqslant \tau_i < t_0 + t\}\big) \cap \{\omega\colon \tau_{k+1} \geqslant t_0 + t\}$ означает наблюдение k раз события A в какие-то моменты из промежутка $[t_0, t_0 + t)$. Поэтому $\big(\cap_{i=1}^{k}\{\omega\colon t_{i-1} \leqslant \tau_i < t_0 + t\}\big) \cap \cap \{\omega\colon \tau_{k+1} \geqslant t_0 + t\} = \{\omega\colon \xi(t_0, t_0 + t) = k\}$ и, следовательно, имеем $\mathsf{P}(\{\omega\colon \xi(t_0, t_0 + t) = k\}) = (\lambda t)^k (k!)^{-1} \exp\{-\lambda t\}$. Лемма 10.4 доказана.

2.6. Распределение Эрланга. Показательная случайная величина входит в так называемое семейство эрланговских случайных величин. Это семейство может быть определено в результате решения простого примера.

Пример 10.18. Найти законы распределения суммы $\varkappa_r = \xi_1 + \xi_2 + \ldots + \xi_r$ независимых показательных случайных величин ξ_1, $\xi_2, \ldots, \xi_r$, каждая из которых имеет математическое ожидание, равное λ^{-1}.

Решение. Если ξ_1 и ξ_2 суть независимые одинаково распределенные случайные величины, имеющие экспоненциальное распределение с параметром λ, то плотность распределения $\varkappa_2 = \xi_1 + \xi_2$ равна свертке функций $f_{\xi_1}(x)$ и $f_{\xi_2}(x)$, т. е.

$$f_{\varkappa_2}(y) = \int\limits_{-\infty}^{+\infty} f_{\xi_1}(x)\, f_{\xi_2}(y - x)\, dx = \int\limits_{0}^{y} f_{\xi_1}(x)\, f_{\xi_2}(y - x)\, dx =$$
$$= \int\limits_{0}^{y} \lambda \exp\{-\lambda x\}\, \lambda \exp\{-\lambda(y - x)\}\, dx = \lambda^2 y \exp\{-\lambda y\}$$

при $y \geqslant 0$ и $f_{\varkappa_2}(y) = 0$ при $y < 0$. Методом математической индукции по r установим, что для $\varkappa_r$ плотность $f_{\varkappa_r}(y) = 0$ при $y < 0$ и $f_{\varkappa_r}(y) = \lambda^r y^{r-1}((r-1)!)^{-1} \exp\{-\lambda y\}$ при $y \geqslant 0$. Эта формула верна для $r = 2$. Пусть теперь равенство для плотности вероятностей $f_{\varkappa_r}(y)$ выполняется для некоторого натурального $r > 2$. Так как случайная величина $\varkappa_{r+1} = \xi_1 + \xi_2 + \ldots + \xi_{r+1} = \varkappa_r + \xi_{r+1}$ и случайные величины $\varkappa_r$ и ξ_{r+1} являются независимыми, то плотность распределения $\varkappa_{r+1}$

равна свертке функций $f_{\varkappa_r}(x)$ и $f_{\xi_r}(x)$. Отсюда легко получаем соотношение

$$f_{\varkappa_{r+1}}(y) = \int\limits_{-\infty}^{+\infty} f_{\varkappa_r}(x)\, f_{\xi_{r+1}}(y-x)\, dx = \int\limits_0^y f_{\varkappa_r}(x)\, f_{\xi_{r+1}}(y-x)\, dx =$$
$$= \lambda^{r+1}((r-1)!)^{-1} \int\limits_0^y x^{r-1} \exp\{-\lambda x\} \exp\{-\lambda(y-x)\}\, dx =$$
$$= \lambda^{r+1} y^r (r!)^{-1} \exp\{-\lambda y\}$$

для $y \geqslant 0$ и $f_{\varkappa_{r+1}}(y) = 0$ при $y < 0$. Формула для $f_{\varkappa_r}(y)$ доказана.

Для определения интегральной функции распределения $F_{\varkappa_r}(y)$ случайной величины $\varkappa_r$ вычислим сначала интеграл $I_r(y) = \lambda^r((r-1)!)^{-1} \int_0^y x^{r-1} \exp\{-\lambda x\}\, dx$. Применяя правило интегрирования по частям, легко получим рекуррентное по r соотношение $I_r(y) = -\lambda^{r-1} y^{r-1}((r-1)!)^{-1} \exp\{-\lambda y\} + I_{r-1}(y)$. Так как интеграл $I_r(y) = \int_0^y \lambda \exp\{-\lambda x\}\, dx = 1 - \exp\{-\lambda y\}$, то при любом $r = 1, 2, \ldots$ интеграл $I_r(y) = 1 - \exp\{-\lambda y\} \sum_{i=1}^r \lambda^{i-1} y^{i-1}((i-1)!)^{-1}$. Отсюда получим формулу для интегральной функции: $F_{\varkappa_r}(y) = 0$ при $y < 0$ и $F_{\varkappa_r}(y) = I_r(y)$ при $y \geqslant 0$. В качестве простой проверки выполненных вычислений можно продифференцировать равенство для $F_{\varkappa_r}(y)$, и в результате сразу получим формулу для $f_{\varkappa_r}(y)$. На рис. 10.17 приведены графики функций $f_{\varkappa_r}(y)$ и $F_{\varkappa_r}(y)$ при $\lambda = 1$, $r = 2$.

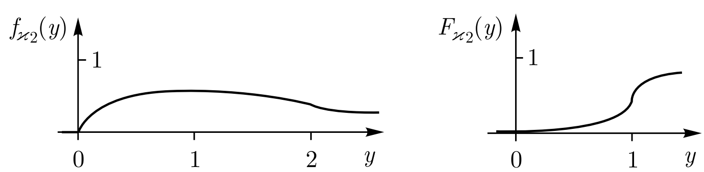


Рис. 10.17

Случайная величина $\varkappa_r$, плотность распределения которой определяется формулой для $f_{\varkappa_r}(y)$, называется *эрланговской*. Распределение Эрланга зависит от двух параметров: λ и r. При $r = 1$ распределение Эрланга совпадает с показательным. Поэтому показательное распределение включается в семейство распределений Эрланга. Распределение Эрланга играет важную роль в теории массового обслуживания и теории надежности.

Заметим, что лемму 10.4 можно доказать более простым способом.

Действительно, сумма $\Delta_1 + \Delta_2 + \ldots + \Delta_r = \varkappa_r$ распределена по закону Эрланга с параметрами λ и r, и событие $\{\omega\colon \xi(t_0, t_0 + t) < r\} = \{\omega\colon \varkappa_r \geqslant t\}$. Так как величина $\xi(t_0, t_0 + t)$ является дискретной, то вероятность $\mathsf{P}(\{\omega\colon \xi(t_0, t_0+t) = r\}) = \mathsf{P}(\{\omega\colon \xi(t_0, t_0+t) < r+1\}) - \mathsf{P}(\{\omega\colon \xi(t_0, t_0+t) < r\})$. Отсюда, используя формулу для $f_{\varkappa_r}(y)$ при $y = t \geqslant 0$, легко получаем, что

$$\begin{aligned}
&\mathsf{P}(\{\omega\colon \xi(t_0, t_0+t) = r\}) = \\
&\quad = \mathsf{P}(\{\omega\colon \xi(t_0, t_0+t) < r+1\}) - \mathsf{P}(\{\omega\colon \xi(t_0, t_0+t) < r\}) = \\
&\quad = \mathsf{P}(\{\omega\colon \varkappa_{r+1} \geqslant t\}) - \mathsf{P}(\{\omega\colon \varkappa_r \geqslant t\}) = \\
&\quad = 1 - \sum_{i=1}^{r+1} \lambda^{i-1} t^{i-1} ((i-1)!)^{-1} \exp\{-\lambda t\} - \\
&\qquad\qquad - \Bigl(1 - \sum_{i=1}^{r} \lambda^{i-1} t^{i-1} ((i-1)!)^{-1} \exp\{-\lambda t\}\Bigr) = \\
&\quad = (\lambda t)^r (r!)^{-1} \exp\{-\lambda t\}.
\end{aligned}$$

Применяя теперь с небольшими изменениями методы вычисления числовых характеристик для показательной случайной величины, найдем аналогичные числовые характеристики для любой эрланговской случайной величины $\varkappa$ с параметрами λ и r. В самом деле, так как $\alpha_0 \varkappa_r = 1$ и

$$\begin{aligned}
\alpha_k \varkappa &= \mathsf{M}\varkappa^k = \int_{-\infty}^{+\infty} x^k f_{\varkappa}(x)\,dx = \int_0^{+\infty} x^k \lambda^r x^{r-1} ((r-1)!)^{-1} \exp\{-\lambda x\}\,dx = \\
&= -\lambda^{r-1} ((r-1)!)^{-1} \int_0^{+\infty} x^{k+r-1}\, d(\exp\{-\lambda x\}) = \\
&= (-\lambda^{r-1} ((r-1)!)^{-1} x^{k+r-1} \exp\{-\lambda x\})\Big|_0^{\infty} + (k+r-1)\lambda^{-1} \times \\
&\times \int_0^{+\infty} x^{k-1} \lambda^{r-1} ((r-1)!)^{-1} x^{r-1} \exp\{-\lambda x\}\,dx = (k+r-1)\lambda^{-1} \alpha_{k-1}\varkappa,
\end{aligned}$$

то начальный момент k-го порядка равен $\alpha_k \varkappa_r = \lambda^{-k} \prod_{i=0}^{k-1} (r+i)$. Отсюда получаем, что

$$\mathsf{M}\varkappa = \alpha_1 \varkappa = \lambda^{-1} r, \quad \mathsf{D}\varkappa = \alpha_2 \varkappa - (\mathsf{M}\varkappa)^2 = \lambda^{-2} r(r+1) - \lambda^{-2} r^2 = \lambda^{-2} r,$$

$$\boldsymbol{\sigma}\varkappa = \lambda^{-1} r^{1/2}, \qquad \beta_3 \varkappa = \sum_{k=0}^{3} C_3^k (-1)^{3-k} (\alpha_1 \varkappa)^{3-k} \alpha_k \varkappa = 2\lambda^{-3} r,$$

$$\mathsf{Ka}\varkappa = (\boldsymbol{\sigma}\varkappa)^{-3} \beta_3 \varkappa = 2 r^{-1/2},$$

$$\beta_4 \varkappa = \sum_{k=0}^{4} C_4^k (-1)^{4-k} (\alpha_1 \varkappa)^{4-k} \alpha_k \varkappa = \lambda^{-4} 3r(r+2),$$

$$\mathsf{Э}\varkappa = (\boldsymbol{\sigma}\varkappa)^{-4} \beta_4 \varkappa - 3 = 6/r.$$

Производная от плотности $f_\varkappa(x)$, равная нулю при всех $x < 0$ и равная $(r-1-\lambda x)\,\lambda^r x^{r-2}((r-1)!)^{-1}\exp\{-\lambda x\}$ при $x \geqslant 0$, меняет знак плюс на минус при переходе значения аргумента через $x = (r-1)\,\lambda^{-1}$. Следовательно, для случайной величины $\varkappa$ с распределением Эрланга наивероятнейшее значение $\mathsf{Mo}^*\varkappa = \mathsf{Mo}\,\varkappa = (r-1)\,\lambda^{-1}$.

2.7. Специальные законы распределения. В этом пункте в основном конспективно рассмотрим свойства еще нескольких весьма распространенных распределений непрерывных случайных величин.

Распределение Вейбулла–Гнеденко. Будем говорить, что случайная величина ξ *распределена по закону Вейбулла–Гнеденко*, если ее интегральная функция $F(x)$ равна нулю при всех $x \leqslant 0$ и равна $1 - \exp\{-(\lambda x)^r\}$ при $x > 0$. Это распределение зависит от параметра масштаба $\lambda > 0$ и параметра формы $r > 0$. Случайная величина ξ с распределением такого вида определяет время безотказной работы большого числа технических элементов и систем или время жизни разнообразных биологических индивидуумов. Поэтому это распределение широко применяется в теории надежности и в демографии. Если параметр формы $r = 1$, то распределение Вейбулла–Гнеденко совпадает с показательным распределением. При $r = 2$ распределение Вейбулла–Гнеденко называется *распределением Рэлея*. Приведем формулы для основных числовых характеристик распределения Вейбулла–Гнеденко: $\alpha_k\xi = \mathsf{M}(\xi^k) = \lambda^{-k}\Gamma(1 + kr^{-1})$, $\mathsf{M}\xi = \alpha_1\xi = \lambda^{-1}\Gamma(1 + r^{-1})$, $\mathsf{D}\xi = \lambda^{-2}[\Gamma(1 + 2r^{-1}) - \Gamma(1 + r^{-1})^2]$, где гамма-функция Эйлера $\Gamma(c) = \int_0^{+\infty} x^{c-1}\exp\{-x\}\,dx$ и параметр $c > 0$. Наконец, наивероятнейшее значение $\mathsf{Mo}^*\xi = \mathsf{Mo}\,\xi = \lambda^{-1}(1 - r^{-1})^{1/r}$ при $r \geqslant 1$ и $\mathsf{Mo}^*\xi = \mathsf{Mo}\,\xi = 0$ при $r < 1$.

Распределение Парето. Случайная величина ξ имеет распределение Парето с параметрами $b > 0$ и $r > 0$, если плотность распределения ξ равна $f(x) = 0$ при $x < b$ и $f(x) = rb^{-1}(b/x)^{r+1}$ при $x \geqslant b$. Случайная величина с таким распределением принимает значения из промежутка $[b, +\infty)$. Распределение Парето часто используется для адекватного описания вероятностных свойств так называемых усеченных количественных характеристик в экономике. Например, налоговая служба обычно интересуется теми гражданами страны, годовой доход ξ каждого из которых не меньше некоторого заданного уровня b. При этом уровень b определяется законом о налогообложении. Закон Парето широко применяется при статистическом анализе распределения заработной платы работников в определенной отрасли народного хозяйства, например, в промышленности, в сельском хозяйстве и т. д. В этом случае, как правило, не учитывается зарплата такого работника, для которого она меньше некоторого заданного порога. При определенных ограничениях на параметр формы r формулы для основных числовых характеристик распределения Парето имеют следущий вид: $\alpha_k\xi = \mathsf{M}(\xi^k) = r(r-k)^{-1}b^k$ при $r > k$, $\mathsf{M}\xi = \alpha_1\xi = r(r-1)^{-1}b$ при $r > 1$, $\mathsf{D}\xi = r(r-1)^{-2}(r-2)^{-1}b^2$ при $r > 2$, $\mathsf{Mo}^*\xi = \mathsf{Mo}\,\xi = b$, $\mathsf{Me}\,\xi = 2^{1/r}b$.

Гамма-распределение. Величина ξ имеет гамма-распределение $\Gamma(\lambda, c)$ с параметром масштаба $\lambda > 0$ и параметром формы $c > 0$, если плотность $f(x) = 0$ при $x < 0$ и $f(x) = \lambda^c x^{c-1}(\Gamma(c))^{-1}\exp\{-\lambda x\}$ при $x \geqslant 0$. Иногда параметр λ заменяют параметром $b = \lambda^{-1}$. Здесь гамма-функция Эйлера $\Gamma(c) = \int_0^{+\infty} x^{c-1}\exp\{-x\}\,dx$. Так как гамма-распределение является двухпараметрическим законом, то с помощью подбора параметров λ и c удается достаточно хорошо описать вероятностные свойства количственных характеристик большого числа реальных экспериментов. Например, с помощью этого закона удается описать распределение времен безотказной работы технических систем при так называемой схеме накопления импульсных воздействий, распределение доходов и сбережений некоторых слоев населения страны в нестабильных экономических ситуациях и т. д. Более того, при $c = 1$ гамма-распределение превращается в показательное с параметром λ и при натуральном значении $c > 1$ оно является распределением Эрланга. Наконец, если $\lambda = 2^{-1}$, $s = 1, 2, \ldots$ и $c = s/2$, то гамма-распределение будет распределением хи-квадрат с s степенями свободы. Следовательно, гамма-распределение естественным образом связано не только с экспоненциальным, нормальным и распределением хи-квадрат, но и с пуассоновским через показательное распределение. Эта удивительная особенность гамма-распределения позволяет его использовать для вычисления различных законов распределения. Для определения основных числовых характеристик гамма-распределения можно использовать формулы $\alpha_k\xi = \mathsf{M}(\xi^k) = \lambda^{-k}\prod_{i=0}^{k-1}(c+i)$, $\mathsf{M}\xi = \alpha_1\xi = \lambda^{-1}c$, $\mathsf{D}\xi = \lambda^{-2}c$, $\mathsf{Mo}^*\xi = \mathsf{Mo}\,\xi = \lambda^{-1}(c-1)$ при $c \geqslant 1$, $\mathsf{Ka}\,\xi = 2\,c^{-1/2}$, $\mathsf{Э}\xi = 6/c$.

Краткий обзор

В гл. 10 детально изучаются различные классы тестовых дискретных и непрерывных случайных величин, которые широко используются для построения вероятностных моделей измерителей элементарных исходов реальных статистически устойчивых экспериментов. Для изучения специальных свойств дискретных случайных величин используется аппарат производящих функций. Из наиболее распространенных дискретных случайных величин подробно рассмотрены: биномиальная случайная величина, пуассоновская случайная величина, распределение Бартлетта, геометрическое распределение, гипергеометрическое распределение, распределение Паскаля. Приводятся эксперименты, в которых встречаются такого типа случайные величины. В частности, подробно исследованы:

1) задача о свободном движении машин по магистрали;
2) задача о групповом движении машин по магистрали;
3) задача о выборочном контроле выпускаемой продукции.

Среди класса непрерывных случайных величин изучены нормальные, равномерные и показательные случайные величины, распределения хи-квадрат, Эрланга, Вейбулла–Гнеденко, Рэлея, Парето и гамма-распределение. В главе рассмотрен широкий круг вопросов, связанных с распределением Гаусса. Устанавливается связь пуассоновской случайной величины дискретного типа и экспоненциальной случайной величины непрерывного типа.

Контрольные вопросы и упражнения

1. Перечислите все различные варианты для моды $\mathsf{Mo}\,\xi$ и наивероятнейшего значения $\mathsf{Mo}^*\xi$ биномиальной случайной величины ξ.
2. На пути движения автомашины 5 светофоров. Каждый из них независимым образом с вероятностью 0,9 либо разрешает, либо с вероятностью 0,1 запрещает на некоторое время автомашине дальнейшее движение. Найти среднее количество остановок автомобиля.
3. Вероятность поражения цели при одиночном выстреле любого из n одинаковых орудий равна $p = 1 - 2^{-a}$, где $a = \mathrm{const} > 0$. Сколько потребуется орудий, чтобы вероятность поражения цели при одном залпе была больше $1 - 2^{-b}$, где $b = \mathrm{const} > 0$?
4. Пусть случайные величины ξ_1, $\xi_2, \ldots$, ξ_n являются независимыми и принимают целочисленные неотрицательные значения. Производящая функция случайной величины ξ_i равна $\Pi_{\xi_i}(z)$, где $i = 1, 2, \ldots, n$. Найти производящую функцию $\Pi_\xi(z)$ суммы $\xi = \xi_1 + 2\,\xi_2 + \ldots + n\,\xi_n$.
5. Используя табл. П. 3, построить графики интегральных функций распределения для трех пуассоновских случайных величин с параметрами $\lambda = 2$, $\lambda = 0,1$ и $\lambda = 2,5$.
6. Количество сбоев за сутки при работе компьютерной сети подчиняется распределению Пуассона с некоторым параметром λ, называемым интенсивностью потока сбоев. Среднее число сбоев за сутки равно 0,1. Используя табл. П.4, найти вероятности того, что за неделю работы сети произойдет не менее трех сбоев.
7. Не используя метод производящих функций, показать, что сумма $\xi_1 + \xi_2 = \xi$ двух независимых пуассоновских случайных величин ξ_1 и ξ_2 с параметрами λ_1 и λ_2 соответственно распределена по закону Пуассона с параметром $\lambda = \lambda_1 + \lambda_2$.

8. Объясните главную причину, которая выделяет нормальный закон среди других законов непрерывного типа. Приведите формулировку правила «трех сигм» и поясните его практическое значение.
9. Какая связь существует между математическим ожиданием и средним квадратическим отклонением показательной случайной величины?
10. Предположим, что рост мужчины определенной возрастной группы есть нормально распределенная случайная величина ξ с заданными математическим ожиданием $a = 173$ и дисперсией $\sigma^2 = 36$. Определить следующие характеристики: а) доли костюмов четвертого роста (176–182 см) и третьего роста (170–176 см), которые нужно предусмотреть в общем объеме производства для данной возрастной группы; б) квантиль уровня 0,7 для случайной величины ξ; в) сформулировать правило «трех сигм» для случайной величины ξ.
11. Назовите основное практическое значение равномерного закона распределения.
12. Случайные величины ξ и η являются независимыми и каждая из них имеет равномерное распределение на отрезке $[0, 1]$. Пусть случайные величины $\zeta_1 = \xi$ и $\zeta_2 = \xi + \eta$. Определить условную плотность вероятностей случайной величины ζ_1 при условии, что случайная величина ζ_2 приняла значение 1/2.
13. Случайная величина ξ распределена по экспоненциальному (показательному) закону с параметром $\lambda > 0$. Найти вероятность $\mathsf{P}(\{\omega\colon |\xi(\omega) - \mathsf{M}\xi| < 3(\mathsf{D}\xi)^{1/2}\})$.
14. Маршрутные такси прибывают регулярно на остановку с интервалом пять минут. Некто приходит на остановку в случайный и не связанный с расписанием момент времени. Найти среднее квадратическое отклонение $\sigma\xi$ времени ξ опоздания пассажиром такси. Вычислить вероятность того, что пассажир будет опережать такси на величину не менее $\mathsf{M}\xi$.

Глава 11

АППРОКСИМАЦИЯ СЛУЧАЙНЫХ ВЕЛИЧИН И ИХ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ

§ 1. Различные типы сходимости последовательностей случайных величин

1.1. Массовые случайные явления и их предсказание. Напомним, что одной из главных проблем теории вероятностей является построение математических моделей массовых явлений, которые порождаются статистически устойчивыми экспериментами при их многократном проведении или наблюдении. При этом математические законы теории вероятностей получаются на основании рационального абстрагирования и адекватного описания реальных статистических закономерностей, свойственных массовым случайным явлениям. Одну из таких групп математических законов или предельных теорем в форме законов больших чисел на содержательном уровне можно описать следующим образом. Пусть имеется однопараметрическое семейство $\{\xi_n; n \geqslant 1\}$ случайных величин с неизвестными законами распределения, и это семейство зависит от натурального параметра n. Каждая случайная величина из этого семейства является количественной характеристикой некоторого реального эксперимента. Оказывается, что при некоторых общих условиях и при достаточно большом значении параметра n случайные величины могут вести себя как некоторая хорошо известная случайная величина или, более того, как постоянная величина. Это замечательное свойство позволяет почти с полной достоверностью предсказывать так называемые предельные значения случайных величин из указанного семейства. Это, в свою очередь, дает возможность не только осуществлять научные прогнозы в области массовых случайных явлений, но и оценивать точность этих прогнозов. Поясним эту ситуацию на следующих двух простых примерах.

При каждом единичном проведении или наблюдении статистически устойчивого эксперимента E в общем случае мы не можем утверждать, что произойдет или не произойдет некоторый его конкретный результат. Заметим, что в частном случае этот результат может быть порожден некоторым количественным признаком эксперимента E или случайной величиной. Однако при большом числе опытов над экспериментом E мы имеем в наличии совокупность его исходов или массовых случайных явлений, относительно которых можно высказать вполне определенные суждения. Так, например, при однократном подбрасывании симметричной игральной кости мы не можем утверждать, что выпадет четное число очков. Если у нас имеется возможность повторить этот опыт с игральной костью 1000 раз, то практически можно быть уверенным в том, что четное число очков появится приблизительно в половине случаев. Этот факт давно известен из практики. Рассмотрим несколько другой пример. Пусть ξ_1, $\xi_2, \ldots$, ξ_n являются независимыми измерениями неизвестного расстояния b между двумя населенными пунктами. В силу разного рода ошибок, которые мы неизбежно допускаем при измерении, величины ξ_1, $\xi_2, \ldots$, ξ_n будут случайными. На практике за приближенное значение расстояния между этими населенными пунктами принимается среднее арифметическое $n^{-1}(\xi_1 + \xi_2 + \ldots + \xi_n)$ результатов n измерений. При этом из большого числа опытов подмечено, что случайная величина $n^{-1}(\xi_1 + \xi_2 + \ldots + \xi_n)$ приблизительно равна постоянному числу b.

Другая группа математических законов в виде так называемых центральных предельных теорем теории вероятностей уже позволяет с достаточной степенью точности аппроксимировать как вероятности некоторых случайных событий, так и законы распределения случайных величин. С некоторыми из такого рода предельных теорем мы встречались в четвертой главе. В этой главе подробно была рассмотрена схема независимых испытаний Бернулли, для которой при больших значениях n, m, k и s вычисление вероятностей $\mathsf{P}(\boldsymbol{B}_m) = C_n^m p^m q^{n-m}$ и $\mathsf{P}\big(\cup_{m=k}^{s} \boldsymbol{B}_m\big) = \sum_{m=k}^{s} C_n^m p^m q^{n-m}$ представляет непростую задачу. Эта задача впервые была решена Муавром, Лапласом и Пуассоном в результате доказательства центральных предельных теорем в схеме независимых испытаний Бернулли. Формулировки, доказательства и применение этих предельных теорем приводятся в четвертой главе. Далее, существует большое число центральных предельных теорем, которые позволяют в некотором смысле аппроксимировать законы распределения случайных величин. При этом точное определение этих законов практически не представляется возможным. Прежде всего, эти теоремы являются основанием методов суммирования случайных величин и, в частности, решают вопрос о приближенном вычислении закона распределения суммы независимых случайных величин при неограниченном увеличении числа слагаемых. Известно, что распределение суммы $\xi_1 + \xi_2$ двух независимых и непрерывных случайных величин определяется композицией или сверткой плотностей распределения случайных величин ξ_1 и ξ_2. Последовательно применяя это правило, можно найти распределение суммы $\xi_1 + \xi_2 + \ldots + \xi_n$ из n независимых случайных величин. Однако этот относительно простой и в то же время очень громоздкий прием приводит к непреодолимым вычислительным трудностям. Наиболее эффективный прием определения приближенного закона распределения суммы $\xi_1 + \xi_2 + \ldots + \xi_n$ случайных величин при достаточно большом значении n или, иначе, при $n \to \infty$ основан на использовании центральных предельных теорем.

Итак, разного рода предельные теоремы позволяют приближенно вычислять значения случайных величин и аппроксимировать их законы распределения. Так как мы имеем дело со случайными объектами, то возникает проблема о математическом смысле приближенного вычисления значений случайных величин и приемлемой аппроксимации их законов распределения. Точный смысл того или иного приближения в теории вероятностей основан на понятии и определении различных видов сходимости последовательностей случайных величин.

1.2. Сходимость по вероятности и сходимость с вероятностью единица. Сначала перейдем к рассмотрению двух типов сходимости последовательностей случайных величин. В этом случае будем предполагать, что случайная величина ξ и последовательность случайных величин вида ξ_1, $\xi_2, \ldots$ заданы на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. Напомним теперь определение обычного предела числовой последовательности из курса математического анализа. Постоянное число c называется пределом числовой последовательности u_1, $u_2, \ldots$, если для любого сколь угодно малого положительного числа ε найдется такое натуральное число $N(\varepsilon)$, зависящее от ε, что при всех значениях $n \geqslant N(\varepsilon)$ будет выполняться неравенство $|u_n - c| < \varepsilon$, или $c - \varepsilon < u_n < c + \varepsilon$. Этот факт записывается в виде предельного равенства $\lim\limits_{n\to\infty} u_n = c$.

На это определение можно посмотреть несколько иначе. Заданное число u_n можно определить как вырожденную случайную величину ξ_n, принимающую одно и то же постоянное значение u_n при любых элементарных исходах эксперимента E, т. е. $\xi_n(\omega) = u_n$ для всех $\omega \in \Omega$. Итак, формально мы имеем последовательность случайных величин ξ_1, $\xi_2, \ldots$ специального вида, определенных на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. Распределение такой случайной величины $\xi_n(\omega)$ является вырожденным и имеет следующий вид: $\mathsf{P}(\{\omega\colon \xi_n(\omega) = u_n\}) = 1$. Аналогичным образом введем в рассмотрение случайную величину $\xi(\omega)$, которая принимает значение c для всех $\omega \in \Omega$. Тогда $\xi(\omega)$ имеет вырожденное распределение $\mathsf{P}(\{\omega\colon \xi(\omega) = c\}) = 1$. При каждом $n = 1, 2, \ldots$ обозначим через A_n событие вида $\{\omega\colon |\xi_n - \xi| < \varepsilon\}$. Если предел последовательности такого рода вырожденных случайных величин ξ_1, $\xi_2, \ldots$ существует, то событие A_n, начиная с номера $n \geqslant N(\varepsilon)$, равно достоверному. Следовательно, $\mathsf{P}(A_n) = 1$ при всех $n \geqslant N(\varepsilon)$. Итак, если существует предел $\lim\limits_{n\to\infty} \xi_n(\omega) = \lim\limits_{n\to\infty} u_n = \xi(\omega) \equiv c$, то для любого $\varepsilon > 0$ найдется такое число $N(\varepsilon)$, что для всех $n \geqslant N(\varepsilon)$ справедливо равенство $\mathsf{P}(\{\omega\colon |\xi_n - \xi| < \varepsilon\}) = 1$. Тогда, очевидно, $\lim\limits_{n\to\infty} \mathsf{P}(A_n) = 1$. Этот замечательный факт положим в основу определения первого типа сходимости последовательности случайных величин, иначе: сходимости по вероятности.

Усложним несколько задачу и рассмотрим последовательность произвольных случайных величин ξ_1, $\xi_2, \ldots$ и постоянную величину $\xi(\omega) \equiv c$, заданные на основном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. Тогда под сходимостью по вероятности (по мере) случайных величин ξ_1, $\xi_2, \ldots$ к величине $\xi(\omega) \equiv c$ понимается тот факт, что при любом $\varepsilon > 0$ имеет место предельное соотношение $\lim\limits_{n\to\infty} \mathsf{P}(\{\omega\colon |\xi_n - c| < \varepsilon\}) = \lim\limits_{n\to\infty} \mathsf{P}(A_n) = 1$. Смысл этого предельного соотношения состоит в том, что для любых сколь угодно малых положительных ε и δ можно найти такое натуральное число $N(\varepsilon, \delta)$, что при всех $n \geqslant N(\varepsilon, \delta)$ имеем $|\mathsf{P}(\{\omega\colon |\xi_n - c| < \varepsilon\}) - 1| < \delta$, или $1 - \delta < \mathsf{P}(\{\omega\colon |\xi_n - c| < \varepsilon\}) < 1 + \delta$.

Итак, если для детерминированных величин неравенство $|\xi_n - c| = |u_n - c| < \varepsilon$ при $n \geqslant N(\varepsilon)$ удовлетворяется с вероятностью, равной единице, то для случайных величин неравенство $|\xi_n - c| < \varepsilon$ при $n \geqslant N(\varepsilon, \delta)$ удовлетворяется с вероятностью, сколь угодно близкой к единице, т. е. практически достоверно. Поэтому понятие сходимости по вероятности последовательности случайных величин ξ_1, $\xi_2, \ldots$ можно считать прямым обобщением сходимости обычного предела числовой последовательности u_1, $u_2, \ldots$. Наконец, можем дать определение сходимости по вероятности.

Определение 11.1. Последовательность случайных величин ξ_1, $\xi_2, \ldots$ называется *сходящейся по вероятности* к случайной величине ξ, если для любого $\varepsilon > 0$ имеет место равенство

$$\lim_{n\to\infty} \mathsf{P}(\{\omega\colon |\xi_n - \xi| < \varepsilon\}) = 1.$$

Это означает, что для любых заданных чисел $\varepsilon, \delta > 0$ существует такое $N(\varepsilon, \delta)$, что $1 - \delta < \mathsf{P}(\{\omega\colon |\xi_n - \xi| < \varepsilon\}) \leqslant 1 + \delta$ при всех $n \geqslant N(\varepsilon, \delta)$, или, что то же самое, $1 - \delta < \mathsf{P}(\{\omega\colon |\xi_n - \xi| < \varepsilon\}) \leqslant 1$. Иначе можно сказать, что последовательность случайных величин ξ_1, $\xi_2, \ldots$ называется *сходящейся по вероятности* к величине ξ, если для любого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}(\{\omega\colon |\xi_n - \xi| \geqslant \varepsilon\}) = 0. \tag{11.1}$$

На основе этого типа сходимости формулируются так называемые законы больших чисел. В ряде вопросов полезно знать необходимые условия сходимости по вероятности. Одно из простейших таких условий можно сформулировать в виде следующего утверждения.

Лемма 11.1. *Если последовательность случайных величин ξ_1, $\xi_2, \ldots$ сходится по вероятности к величине ξ, то при любом $\varepsilon > 0$ и $n \to \infty$, $m \to \infty$ имеет место предельное соотношение* $\mathsf{P}(\{\omega\colon |\xi_n - \xi| \geqslant \varepsilon\}) \to 0$.

Доказательство. Из очевидных соотношений $\{\omega\colon |\xi_n - \xi_m| \geqslant \varepsilon\} = \{\omega\colon |(\xi_n - \xi) + (\xi - \xi_m)| \geqslant \varepsilon\} \subset \{\omega\colon |\xi_n - \xi| + |\xi - \xi_m| \geqslant \varepsilon\} \subset \{\omega\colon |\xi_n - \xi_m| \geqslant \varepsilon/2\} \cup \{\omega\colon |\xi_m - \xi| \geqslant \varepsilon/2\}$ получаем неравенство $\mathsf{P}(\{\omega\colon |\xi_n - \xi_m| \geqslant \varepsilon\}) \leqslant \mathsf{P}(\{\omega\colon |\xi_n - \xi| \geqslant \varepsilon/2\}) + \mathsf{P}(\{\omega\colon |\xi_m - \xi| \geqslant \varepsilon/2\})$. Применяя предельное равенство (11.1), получаем доказательство леммы 11.1.

Сходимость почти наверное (почти всюду, или с вероятностью единица) определяется следующим образом.

Определение 11.2. Последовательность случайных величин ξ_1, $\xi_2, \ldots$ называется *сходящейся с вероятностью единица* к случайной величине ξ, если имеет место равенство $\mathsf{P}(\{\omega\colon \lim\limits_{n\to\infty} \xi_n(\omega) = \xi(\omega)\}) = 1$.

Другими словами, вероятность появления последовательностей $\xi_1(\omega)$, $\xi_2(\omega), \ldots$ возможных значений этих случайных величин, не сходящихся к $\xi(\omega)$, равна нулю. Следует отметить, что множество

$$\{\omega\colon \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\} = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geqslant N} \{\omega\colon |\xi_n(\omega) - \xi(\omega)| < k^{-1}\}. \quad (11.2)$$

Действительно, если $\omega_0 \in \{\omega\colon \lim\limits_{n\to\infty} \xi_n(\omega) = \xi(\omega)\}$, то для любого $k = 1, 2, \ldots$ существует такое натуральное число $N(\omega_0, k)$, зависящее от ω_0 и k, что для всех $n \geqslant N(\omega_0, k)$ будет выполняться неравенство $|\xi_n(\omega_0) - \xi(\omega_0)| < k^{-1}$. Итак, отсюда следует, что $\omega_0 \in \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geqslant N} \{\omega\colon |\xi_n(\omega) - \xi(\omega)| < k^{-1}\}$. Обратно, пусть теперь $\omega_1 \in \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geqslant N} \{\omega\colon |\xi_n(\omega) - \xi(\omega)| < k^{-1}\}$. Тогда для любого $k = 1, 2, \ldots$ существует такое натуральное число $N(\omega_1, k)$, зависящее от ω_1, k, что для всех значений $n \geqslant N(\omega_1, k)$ будет выполняться неравенство $|\xi_n(\omega_1) - \xi(\omega_1)| < k^{-1}$. Отсюда получаем, что предел $\lim\limits_{n\to\infty} \xi_n(\omega_1)$ существует и равен $\xi(\omega_1)$. Итак, доказано равенство (11.2), и поэтому множество $\{\omega\colon \lim\limits_{n\to\infty} \xi_n(\omega) = \xi(\omega)\}$ является случайным событием. На основе этого типа сходимости формулируются так называемые усиленные законы больших чисел. Рассмотрим достаточное условие для сходимости с вероятностью единица последовательности случайных величин.

Лемма 11.2. *Если ряд $\sum_{n=1}^{\infty} \mathsf{P}(\{\omega\colon |\xi_n - \xi| \geqslant \varepsilon\}) < \infty$ для любого $\varepsilon > 0$, то при $n \to \infty$ последовательность случайных величин ξ_1, $\xi_2, \ldots$ сходится с вероятностью единица к случайной величине ξ.*

Доказательство. Ради простоты выберем в качестве $\varepsilon = k^{-1}$. При фиксированном значении $k = 1, 2, \ldots$ введем в рассмотрение последовательность событий $A_{k,n} = \{\omega\colon |\xi_n - \xi| \geqslant k^{-1}\}$, $n = 1, 2, \ldots$, и событие вида $A_k = \cap_{N=1}^{\infty} \cap_{n \geqslant N} A_{k,n}$. Напомним, что A_k означает наступление бесконечного числа событий из последовательности $\{A_{k,n};\ n = 1, 2, \ldots\}$. Из первой леммы Бореля–Кантелли следует, что $\mathsf{P}(A_k) = 0$, $k \geqslant 1$. Эти равенства эквивалентны соотношению $\mathsf{P}(\overline{A}_k) = 1$ для всех $k = 1, 2, \ldots$. В силу закона де Моргана событие $\overline{A}_k = \cup_{N=1}^{\infty} \cap_{n \geqslant N} \{\omega\colon |\xi_n - \xi| < k^{-1}\}$. При этом событие $\overline{A}_k$ состоит в том, что произойдет только конечное число событий из последовательности $\{A_{k,n};\ n = 1, 2, \ldots\}$. Легко видеть, что $\overline{A}_1 \supset \overline{A}_2 \supset \ldots$, откуда $\cap_{k=1}^{\infty} \overline{A}_k = \lim\limits_{k\to\infty} \overline{A}_k$ и вероятность $\mathsf{P}\big(\cap_{k=1}^{\infty} \overline{A}_k\big) = \mathsf{P}\big(\lim\limits_{k\to\infty} \overline{A}_k\big) = \lim\limits_{k\to\infty} \mathsf{P}(\overline{A}_k) = 1$. Используя это и определение $\overline{A}_k$, найдем

$$\mathsf{P}\Big(\bigcap_{k=1}^{\infty} \overline{A}_k\Big) = \mathsf{P}\Big(\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n \geqslant N} \{\omega\colon |\xi_n(\omega) - \xi(\omega)| < k^{-1}\}\Big) = 1.$$

Результат леммы теперь вытекает из соотношения (11.2).

1.3. Сходимость в среднеквадратическом и сходимость по распределению. Отношения между различными видами сходимости. Третий тип сходимости последовательностей случайных величин к случайной величине называется сходимостью в среднеквадратическом. При этом предполагается, что величины $\xi_1, \xi_2, \ldots$ и величина ξ заданы на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ и имеют конечные моменты второго порядка, т. е. $\mathsf{M}(\xi^2) < +\infty$, $\mathsf{M}(\xi_n^2) < +\infty$ для всех $n \geqslant 1$. Сходимость в среднеквадратическом определяется следующим образом.

Определение 11.3. Последовательность случайных величин ξ_1, $\xi_2, \ldots$ называется *сходящейся в среднеквадратическом* к случайной величине ξ, если при $n \to \infty$ предел $\lim\limits_{n\to\infty} \mathsf{M}(\xi_n - \xi)^2 = 0$.

Наконец, рассмотрим четвертый тип сходимости. Пусть случайные величины ξ, ξ_1, $\xi_2, \ldots$ заданы не обязательно на одном вероятностном пространстве. В этом непростом случае всегда можем определить интегральную функцию распределения $F_{\xi}(x)$ случайной величины ξ и интегральную функцию распределения $F_{\xi_n}(x)$ случайной величины ξ_n, где $n \geqslant 1$.

Определение 11.4. Последовательность ξ_1, $\xi_2, \ldots$ случайных величин называется *сходящейся по распределению* или *сходящейся слабо* к случайной величине ξ, если $\lim\limits_{n\to\infty} F_{\xi_n}(x) = F_{\xi}(x)$ в каждой точке x, являющейся точкой непрерывности функции $F_{\xi}(x)$.

Если последовательность $(\xi_1 - a_1)/\sigma_1$, $(\xi_2 - a_2)/\sigma_2, \ldots$ сходится слабо к $(0,1)$-нормальной случайной величине, то последовательность ξ_1, $\xi_2, \ldots$ называется *асимптотически нормальной с параметрами* (a_n, σ_n).

Отметим, что каждая из четырех типов сходимости определяет ту или иную характерную близость последовательности случайных величин к некоторой случайной величине. Существуют последовательности случайных величин, которые на содержательном уровне и с очевидностью близки во всех четырех смыслах к некоторой случайной величине. Например, последовательность $\{\xi_n = \xi - 2^{-n};\ n = 1, 2, \ldots\}$ должна естественно сходится к случайной величине ξ при $n \to \infty$ во всех четырех случаях. Из равенства $\xi_n = \xi - 2^{-n}$ видно, что для случайных величин ξ_1, $\xi_2, \ldots$ предел $\lim\limits_{n\to\infty} F_{\xi_n}(x) = \lim\limits_{n\to\infty} \mathsf{P}(\{\omega\colon \xi_n < x_0\}) = \lim\limits_{n\to\infty} \mathsf{P}(\{\omega\colon \xi < x_0 + 2^{-n}\}) = \lim\limits_{n\to\infty} F_\xi(x_0 + 2^{-n}) = F_\xi(x_0 + 0)$ не совпадает с $F_\xi(x_0)$, если только функция $F_\xi(x)$ не является непрерывной при $x = x_0$. Этот пример показывает, почему в определении сходимости по распределению вполне разумно требовать выполнение предельного равенства $\lim\limits_{n\to\infty} F_{\xi_n}(x) = F_\xi(x)$ только в точках непрерывности $F_\xi(x)$.

Приведем еще один простой пример, когда последовательность $F_{\xi_1}(x)$, $F_{\xi_2}(x), \ldots$ интегральных функций распределения случайных величин ξ_1, $\xi_2, \ldots$ сходится в каждой точке x к некоторой функции $F(x)$. Однако последовательность случайных величин ξ_1, $\xi_2, \ldots$ не сходится по распределению, поскольку предельная функция $F(x)$ не является интегральной.

Пусть $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ — вероятностное пространство и случайная величина $\xi_n(\omega) \equiv 2^n$ при $\omega \in \Omega$, где $n = 1, 2, \ldots$. В этом тривиальном случае имеем предельное равенство $\lim\limits_{n\to\infty} F_{\xi_n}(x) = F_\xi(x) \equiv 0$ в каждой точке x. Поскольку $\lim\limits_{x\to+\infty} F(x) \neq 1$, то предельная функция $F(x)$ для этого примера не будет интегральной. Поэтому в определении слабой сходимости необходимо требовать, чтобы предельная функция $F_\xi(x)$ являлась интегральной.

Между различными видами сходимости последовательностей случайных величин верны следующие высказывания:

— из сходимости с вероятностью единица следует сходимость по вероятности; обратное утверждение, вообще говоря, неверно;

— из сходимости по вероятности следует сходимость по распределению; обратное утверждение, вообще говоря, неверно;

— из сходимости в среднеквадратическом следует сходимость по вероятности; обратное утверждение, вообще говоря, неверно.

Теорема 11.1. *Если последовательность случайных величин* ξ_1, $\xi_2, \ldots$ *сходится с вероятностью единица к случайной величине* ξ, *то она сходится по вероятности к случайной величине* ξ.

Доказательство. Так как выполняется равенство (11.2) и вероятность

$$\mathsf{P}\Big(\bigcap_{k=1}^{\infty}\bigcup_{N=1}^{\infty}\bigcap_{n\geqslant N}\{\omega\colon |\xi_n(\omega)-\xi(\omega)|<k^{-1}\}\Big)=1,$$

то для $k=1,2,\ldots$ имеем

$$\mathsf{P}\Big(\bigcup_{N=1}^{\infty}\bigcap_{n\geqslant N}\{\omega\colon |\xi_n(\omega)-\xi(\omega)|<k^{-1}\}\Big)=1.$$

При $k\geqslant 1$ верно соотношение

$$\bigcap_{n\geqslant 1}\{\omega\colon |\xi_n(\omega)-\xi(\omega)|<k^{-1}\}\subset\bigcap_{n\geqslant 2}\{\omega\colon |\xi_n(\omega)-\xi(\omega)|<k^{-1}\}\subset\ldots\,.$$

Тогда

$$\begin{aligned}\mathsf{P}\Big(\bigcup_{N=1}^{\infty}\bigcap_{n\geqslant N}\{\omega\colon |\xi_n(\omega)-\xi(\omega)|<k^{-1}\}\Big)&=\\ =\mathsf{P}\Big(\lim_{N\to\infty}\bigcap_{n\geqslant N}\{\omega\colon |\xi_n(\omega)-\xi(\omega)|<k^{-1}\}\Big)&=\\ =\lim_{N\to\infty}\mathsf{P}\Big(\bigcap_{n\geqslant N}\{\omega\colon |\xi_n(\omega)-\xi(\omega)|<k^{-1}\}\Big)&=1\end{aligned}$$

при любом фиксированном $k=1,2,\ldots$. Из

$$\{\omega\colon |\xi_N(\omega)-\xi(\omega)|<k^{-1}\}\supset\bigcap_{n\geqslant N}\{\omega\colon |\xi_n(\omega)-\xi(\omega)|<k^{-1}\}$$

получаем, что

$$\lim_{N\to\infty}\mathsf{P}(\{\omega\colon |\xi_N(\omega)-\xi(\omega)|<k^{-1}\})=1,$$

т. е. последовательность $\xi_1,\ \xi_2,\ldots$ сходится по вероятности к величине ξ. Теорема 11.1 доказана.

Теорема 11.2. *Пусть последовательность случайных величин ξ_1, $\xi_2,\ldots$ сходится по вероятности к случайной величине ξ. Тогда она сходится по распределению к случайной величине ξ.*

Доказательство. Используя формулу полной вероятности, получим, что для любого $x\in R$ и любого $\varepsilon>0$ справедливы равенства

$$\begin{aligned}F_\xi(x-\varepsilon)&=\mathsf{P}(\{\omega\colon \xi<x-\varepsilon\})=\\ &=\mathsf{P}(\{\omega\colon \xi<x-\varepsilon,\ |\xi_n-\xi|<\varepsilon\})+\mathsf{P}(\{\omega\colon \xi<x-\varepsilon,\ |\xi_n-\xi|\geqslant\varepsilon\}),\end{aligned}$$

равенства

$$F_\xi(x+\varepsilon) = \mathsf{P}(\{\omega\colon \xi < x+\varepsilon\}) = \\ = \mathsf{P}(\{\omega\colon \xi < x+\varepsilon,\ |\xi_n - \xi| < \varepsilon\}) + \mathsf{P}(\{\omega\colon \xi < x+\varepsilon,\ |\xi_n - \xi| \geqslant \varepsilon\}),$$

наконец, равенства

$$F_{\xi_n}(x) = \mathsf{P}(\{\omega\colon \xi_n < x\}) = \\ = \mathsf{P}(\{\omega\colon \xi_n < x,\ |\xi_n - \xi| < \varepsilon\}) + \mathsf{P}(\{\omega\colon \xi_n < x,\ |\xi_n - \xi| \geqslant \varepsilon\}).$$

Из соотношений $\{\omega\colon \xi < x-\varepsilon,\ |\xi_n - \xi| \geqslant \varepsilon\} \subset \{\omega\colon |\xi_n - \xi| \geqslant \varepsilon\}$, $\{\omega\colon \xi < x+\varepsilon,\ |\xi_n - y\xi| \geqslant \varepsilon\} \subset \{\omega\colon |\xi_n - \xi| \geqslant \varepsilon\}$, $\{\omega\colon \xi_n < x,\ |\xi_n - \xi| \geqslant \varepsilon\} \subset \subset \{\omega\colon |\xi_n - \xi| \geqslant \varepsilon\}$ и сходимости по вероятности последовательности случайных величин $\xi_1, \xi_2, \ldots$ к случайной величине ξ непосредственно следуют предельные равенства:

$$\lim_{n\to\infty} \mathsf{P}(\{\omega\colon \xi < x-\varepsilon,\ |\xi_n - \xi| \geqslant \varepsilon\}) = 0,$$
$$\lim_{n\to\infty} \mathsf{P}(\{\omega\colon \xi < x+\varepsilon,\ |\xi_n - \xi| \geqslant \varepsilon\}) = 0,$$
$$\lim_{n\to\infty} \mathsf{P}(\{\omega\colon \xi_n < x,\ |\xi_n - \xi| \geqslant \varepsilon\}) = 0.$$

Принимая все это во внимание, получаем справедливые неравенства

$$|F_\xi(x-\varepsilon) - \mathsf{P}(\{\omega\colon \xi < x-\varepsilon,\ |\xi_n - \xi| < \varepsilon\})| < \delta,$$
$$|F_\xi(x+\varepsilon) - \mathsf{P}(\{\omega\colon \xi < x+\varepsilon,\ |\xi_n - \xi| < \varepsilon\})| < \delta$$

и следующее соотношение: $|F_{\xi_n}(x) - \mathsf{P}(\{\omega\colon \xi_n < x,\ |\xi_n - \xi| < \varepsilon\})| < \delta$ для сколь угодно малого значения $\delta > 0$ и при всех достаточно больших n. Так как для случайных событий $\{\omega\colon \xi < x-\varepsilon,\ |\xi_n - \xi| < \varepsilon\}$, $\{\omega\colon \xi_n < x,\ |\xi_n - \xi| < \varepsilon\}$, $\{\omega\colon \xi < x+\varepsilon,\ |\xi_n - \xi| < \varepsilon\}$ верно соотношение

$$\{\omega\colon \xi < x-\varepsilon,\ |\xi_n - \xi| < \varepsilon\} \subset \\ \subset \{\omega\colon \xi_n < x,\ |\xi_n - \xi| < \varepsilon\} \subset \{\omega\colon \xi < x+\varepsilon,\ |\xi_n - \xi| < \varepsilon\},$$

то для вероятностей этих событий верны неравенства:

$$\mathsf{P}(\{\omega\colon \xi < x-\varepsilon,\ |\xi_n - \xi| < \varepsilon\}) \leqslant \\ \leqslant \mathsf{P}(\{\omega\colon \xi_n < x,\ |\xi_n - \xi| < \varepsilon\}) \leqslant \mathsf{P}(\{\omega\colon \xi < x+\varepsilon,\ |\xi_n - \xi| < \varepsilon\}).$$

Учитывая полученные оценки для функций $F_\xi(x-\varepsilon)$, $F_\xi(x+\varepsilon)$ и $F_{\xi_n}(x)$, последние неравенства можно переписать в таком виде: $F_\xi(x-\varepsilon) - 2\,\delta < F_{\xi_n}(x) < F_\xi(x+\varepsilon) + 2\,\delta$ для любого $\varepsilon > 0$. Если теперь значение x является абсциссой точки непрерывности функции $F_\xi(x)$, то окончательно находим: $\lim\limits_{n\to\infty} F_{\xi_n}(x) = F_\xi(x)$. Теорема 11.2 доказана.

Далее покажем, что из сходимости в среднеквадратическом следует сходимость по вероятности. С этой целью запишем первое неравенство Чебышева для величины $\xi_n - \xi$: $\mathsf{P}(\{\omega\colon |\xi_n - \xi| \geqslant \varepsilon\}) \leqslant \mathsf{M}(\xi_n - \xi)^2/\varepsilon^2$, $\varepsilon > 0$. Так как $\lim\limits_{n\to\infty} \mathsf{M}(\xi_n - \xi)^2 = 0$, то для заданных чисел $\varepsilon > 0$ и $\delta > 0$ можно найти такое $N(\delta, \varepsilon)$, что при $n \geqslant N(\delta, \varepsilon)$ будет выполняться неравенство $\mathsf{P}(\{\omega\colon |\xi_n - \xi| \geqslant \varepsilon\}) < \delta$. Значит, $1 - \delta < \mathsf{P}(\{\omega\colon |\xi_n - \xi| < \varepsilon\}) \leqslant 1$ при $n \geqslant N(\delta, \varepsilon)$ и последовательность случайных величин ξ_1, $\xi_2, \ldots$ сходится по вероятности к случайной величине ξ.

Однако из сходимости по вероятности в общем случае не следует сходимость в среднеквадратическом.

В самом деле, рассмотрим эксперимент, который описан в примере 1.11. Для этого эксперимента описание ω — абсцисса произвольно поставленной точки на отрезке $[0, 1]$ оси $0x$, $\Omega = \{\omega\colon 0 \leqslant \omega \leqslant 1\}$, $\mathcal{F}$ есть σ-алгебра борелевских подмножеств на $[0, 1]$ и вероятность $\mathsf{P}(A)$ любого борелевского подмножества (случайного события) $A \subset \Omega$ равна его длине. Определим поточечно на Ω для каждого натурального значения n случайную величину $\xi_n(\omega)$, которая равна 2^n при $0 < \omega < n^{-1}$ и равна нулю в противном случае. Дискретная случайная величина ξ_n имеет распределение: $\mathsf{P}(\{\omega\colon \xi_n = 0\}) = n^{-1}(n-1)$, $\mathsf{P}(\{\omega\colon \xi_n = 2^n\}) = n^{-1}$. Тогда для случайной величины ξ_n получаем, что при любом $\varepsilon > 0$ вероятность вида

$$\mathsf{P}(\{\omega\colon |\xi_n| < \varepsilon\}) \geqslant \mathsf{P}(\{\omega\colon |\xi_n| = 0\}) = n^{-1}(n-1).$$

Значит, последовательность $\{\xi_n;\, n = 1, 2, \ldots\}$ сходится по вероятности к нулю при $n \to \infty$. В то же время математическое ожидание $\mathsf{M}\xi_n^2$, которое равно $n^{-1}2^{2n}$, не стремится к нулю при неограниченном увеличении значения n. Заметим, что для этой последовательности $\{\omega\colon \lim\limits_{n\to\infty} \xi_n(\omega) = 0\} = \Omega$, т. е. последовательность ξ_1, $\xi_2, \ldots$ сходится с вероятностью единица к нулю. Этот пример также показывает, что из сходимости с вероятностью единица в общем случае не следует сходимость в среднеквадратическом.

Приведем простой пример, который иллюстрирует, что из сходимости в среднеквадратическом не следует сходимость с вероятностью единица. Пусть дана последовательность $\{\xi_n;\, n = 1, 2, \ldots\}$ из независимых случайных величин, для которых $\mathsf{P}(\{\omega\colon \xi_n = 0\}) = n^{-1}(n-1)$, $\mathsf{P}(\{\omega\colon \xi_n = c\}) = n^{-1}$, где некоторая постоянная $c > 1$. Легко установить, что $\lim\limits_{n\to\infty} \mathsf{M}(\xi_n - 0)^2 = \lim\limits_{n\to\infty} (c^2/n) = 0$, т. е. эта последовательность сходится к нулю в среднеквадратическом. Отсюда следует сходимость этой последовательности к нулю по вероятности. При фиксированных k, $n = 1, 2, \ldots$ обозначим через $A_{k,n}$ событие $A_{k,n} = \{\omega\colon \xi_n(\omega) \geqslant 1/k\}$. Так как вероятность $\mathsf{P}(A_{k,n}) = 1/n$ и числовой ряд $\sum_{n=1}^{\infty} \mathsf{P}(A_{k,n}) = \infty$, то по лемме Бореля–Кантелли вероятность события $A_k = \cap_{i=1}^{\infty} \cup_{n \geqslant i} A_{k,n}$ равна единице для всех $k = 1, 2, \ldots$. Очевидно, что $A_1 \subset A_2 \subset \ldots$ и в силу равенств

$$\lim_{n\to\infty} A_k = \bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} \bigcap_{i=1}^{\infty} \bigcup_{n\geqslant i} A_{k,n}$$

и непрерывности вероятностной функции, находим, что

$$\mathsf{P}\Big(\bigcup_{k=1}^{\infty} \bigcap_{i=1}^{\infty} \bigcup_{n\geqslant i} A_{k,n}\Big) = \mathsf{P}\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \mathsf{P}\big(\lim_{k\to\infty} A_k\big) = \lim_{k\to\infty} \mathsf{P}(A_k) = 1.$$

Если событие $B = \cup_{k=1}^{\infty} \cap_{i=1}^{\infty} \cup_{n\geqslant i} A_{k,n}$, то

$$\overline{B} = \bigcap_{k=1}^{\infty} \bigcup_{i=1}^{\infty} \bigcap_{n\geqslant i} \{\omega\colon \xi_n(\omega) < 1/k\} = \{\omega\colon \lim_{n\to\infty} \xi_n = 0\}.$$

Следовательно, вероятность $\mathsf{P}(\{\omega\colon \lim\limits_{n\to\infty} \xi_n = 0\}) = 0$, т. е. последовательность $\{\xi_n;\, n = 1, 2, \ldots\}$ не сходится с вероятностью единица к нулю. Ясно, что этот пример так же показывает, что из сходимости по вероятности не следует сходимость с вероятностью единица.

Наконец, покажем на примере, что некоторая последовательность сходится по распределению, но не сходится по вероятности. Пусть эволюционный эксперимент E заключается в последовательном и независимом подбрасывании симметричной игральной кости. При броске с нечетным номером $2n-1$ случайная величина $\xi_{2n-1} = 1$, если выпадает нечетное число очков, и $\xi_{2n-1} = 0$ в противном случае. Далее, при броске с четным номером $2n$ случайная величина $\xi_{2n} = 1$, если выпадает четное число очков, и $\xi_{2n-1} = 0$ в противном случае. Легко видеть, что случайные величины ξ_1, $\xi_2, \ldots$ независимы в совокупности и имеют одно и то же распределение вида: $\mathsf{P}(\{\omega\colon \xi_n = 0\}) = 2^{-1}$, $\mathsf{P}(\{\omega\colon \xi_n = 1\}) = 2^{-1}$. Если обозначить интегральную функцию такого распределения через $F(x)$, то $\lim\limits_{n\to\infty} F_{\xi_n}(x) = F(x)$, и, значит, последовательность ξ_1, $\xi_2, \ldots$ сходится по распределению. Однако при любом $0 < \varepsilon \leqslant 1$, принимая во внимание независимость в совокупности случайных величин ξ_1, $\xi_2, \ldots$ и их распределение, получим

$$\begin{aligned}
&\mathsf{P}(\{\omega\colon |\xi_n - \xi_m| \geqslant \varepsilon\}) = \\
&\qquad = \mathsf{P}(\{\omega\colon \xi_n = 0,\, \xi_m = 1)\} + \mathsf{P}(\{\omega\colon \xi_n = 1,\, \xi_m = 0)\} = \\
&= \mathsf{P}(\{\omega\colon \xi_n = 0\})\, \mathsf{P}(\{\omega\colon \xi_m = 1\}) + \mathsf{P}(\{\omega\colon \xi_n = 1\})\, \mathsf{P}(\{\omega\colon \xi_m = 0\}) = 2^{-1}.
\end{aligned}$$

Теперь на основании леммы 11.1 утверждаем, что построенная таким способом последовательность ξ_1, $\xi_2, \ldots$ не сходится по вероятности.

§ 2. Предельные теоремы для последовательностей случайных величин

2.1. Классификация предельных теорем. Теоремы, устанавливающие сходимость по вероятности некоторых последовательностей случайных величин к случайным величинам, получили название законов больших чисел. Утверждения, устанавливающие сходимость с вероятностью единица (почти наверное) некоторых последовательностей случайных величин к случайным величинам, получили название усиленных законов больших чисел. Теоремы, которые устанавливают условия, при которых закон распределения суммы большого числа случайных величин неограниченно приближается к нормальному, называются центральными предельными теоремами. Различные формы законов больших чисел вместе с различными формами центральных предельных теорем образуют совокупность так называемых предельных теорем теории вероятностей. Прежде чем перейти к рассмотрению простейших предельных теорем, докажем еще раз другим способом известное неравенство Чебышева для дискретных и непрерывных случайных величин. Это неравенство и предложенный способ его доказательства с использованием метода различных оценок лежат в основе доказательства многих предельных теорем.

Теорема 11.3 (неравенство Чебышева). *Пусть имеется случайная величина* ξ *на вероятностном пространстве* $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ *с конечными* $\mathsf{M}\xi$ *и* $\mathsf{D}\xi$, *тогда для любого* $\varepsilon > 0$ *имеем* $\mathsf{P}(\{\omega\colon |\xi - \mathsf{M}\xi| \geqslant \varepsilon\}) \leqslant \leqslant \mathsf{D}\xi/\varepsilon^2$.

Доказательство. Рассмотрим сначала случай дискретной случайной величины ξ с рядом распределения (табл. 11.1)

Таблица 11.1

ξ	x'_1	x'_2	$\ldots$	x'_m
$\mathsf{P}(\cdot)$	p_1	p_2	$\ldots$	p_m

Вычислим вероятность $\mathsf{P}(\{\omega\colon |\xi - \mathsf{M}\xi| \geqslant \varepsilon\}) = \sum_{i\colon |x'_i - \mathsf{M}\xi| \geqslant \varepsilon} p_i = p_1 + p_2 + \ldots + p_{k-1} + p_{s+1} + p_{s+2} + \ldots + p_m$, иллюстрируя это равенство на рис. 11.1.

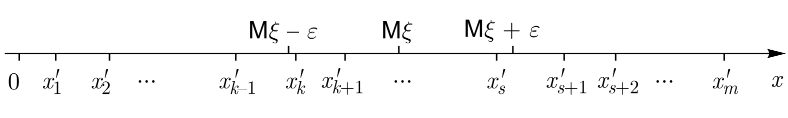


Рис. 11.1

Оценим дисперсию

$$\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \sum_{i=1}^{m}(x_i - \mathsf{M}\xi)^2 p_i = \sum_{i=1}^{m}|x_i - \mathsf{M}\xi|^2 p_i \geqslant$$
$$\geqslant \sum_{i:\,|x_i'-\mathsf{M}\xi|\geqslant\varepsilon}(x_i - \mathsf{M}\xi)^2 p_i \geqslant \varepsilon^2 \sum_{i:\,|x_i'-\mathsf{M}\xi|\geqslant\varepsilon} p_i = \varepsilon^2 \mathsf{P}(\{\omega\colon |\xi - \mathsf{M}\xi| \geqslant \varepsilon\}),$$

откуда и следует утверждение $\mathsf{P}(\{\omega\colon |\xi - \mathsf{M}\xi| \geqslant \varepsilon\}) \leqslant \mathsf{D}\xi/\varepsilon^2$.

Рассмотрим теперь это утверждение для непрерывной случайной величины ξ с плотностью $f(x)$. При этом имеем:

$$\mathsf{P}(\{\omega\colon |\xi - \mathsf{M}\xi| \geqslant \varepsilon\}) = \int\limits_{x:\,|x-\mathsf{M}\xi|\geqslant\varepsilon} f(x)\,dx,$$

$$\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \int\limits_{-\infty}^{+\infty}(x - \mathsf{M}\xi)^2 f(x)\,dx = \int\limits_{-\infty}^{+\infty}|x - \mathsf{M}\xi|^2 f(x)\,dx \geqslant$$
$$\geqslant \int\limits_{\{x:\,|x-\mathsf{M}\xi|\geqslant\varepsilon\}}|x - \mathsf{M}\xi|^2 f(x)\,dx \geqslant \varepsilon^2 \int\limits_{\{x:\,|x-\mathsf{M}\xi|\geqslant\varepsilon\}} f(x)\,dx.$$

Из этих соотношений получаем, что $\mathsf{D}\xi \geqslant \varepsilon^2\mathsf{P}(\{\omega\colon |\xi - \mathsf{M}\xi| \geqslant \varepsilon\})$. Отсюда и выводим искомое утверждение $\mathsf{P}(\{\omega\colon |\xi - \mathsf{M}\xi| \geqslant \varepsilon\}) \leqslant \mathsf{D}\xi/\varepsilon^2$.

Перейдем теперь к иллюстрации важной роли неравенства Чебышева при доказательстве простейших предельных теорем.

Рассмотрим, например, последовательность ξ_1, $\xi_2, \ldots$ попарно некоррелированных случайных величин с конечными математическими ожиданиями $\mathsf{M}\xi_1$, $\mathsf{M}\xi_2, \ldots$ и равномерно ограниченными дисперсиями $\mathsf{D}\xi_1 \leqslant h$, $\mathsf{D}\xi_2 \leqslant h, \ldots$ некоторой постоянной h. Построим новую последовательность $\{\eta_n;\ n = 1, 2, \ldots\}$ случайных величин, где $\eta_n = n^{-1}\sum_{i=1}^{n}(\xi_i - \mathsf{M}\xi_i)$. Вычислим математическое ожидание $\mathsf{M}\eta_n$ и дисперсию $\mathsf{D}\eta_n$ величины η_n:

$$\mathsf{M}\eta_n = \mathsf{M}\Big[n^{-1}\sum_{i=1}^{n}(\xi_i - \mathsf{M}\xi_i)\Big] = n^{-1}\sum_{i=1}^{n}\mathsf{M}(\xi_i - \mathsf{M}\xi_i) = 0,$$
$$\mathsf{D}\eta_n = \mathsf{D}\Big[n^{-1}\sum_{i=1}^{n}(\xi_i - \mathsf{M}\xi_i)\Big] = n^{-2}\mathsf{D}\Big(\sum_{i=1}^{n}\xi_i\Big) = n^{-2}\sum_{i=1}^{n}\mathsf{D}\xi_i \leqslant n^{-1}h.$$

Заметим, что при вычислении дисперсии центрированной случайной величины η_n мы использовали известное равенство $\mathsf{D}\big(\sum_{i=1}^{n}\xi_i\big) = \sum_{i=1}^{n}\mathsf{D}\xi_i$ для попарно некоррелированных случайных величин ξ_1, $\xi_2, \ldots$. Следовательно, дисперсия случайной величины η_n равна $n^{-2}\sum_{i=1}^{n}\mathsf{D}\xi_i$ и убывает с ростом n. На содержательном уровне это означает, что случайная величина η_n с ростом n ведет себя все более и более как неслучайная величина, т. е. приближается к некоторой константе. Теперь легко доказывается следующая теорема.

Теорема 11.4. *Пусть ξ_1, $\xi_2, \ldots$ являются попарно некоррелированными величинами с конечными математическими ожиданиями $\mathsf{M}\xi_1$, $\mathsf{M}\xi_2, \ldots$ и равномерно ограниченными дисперсиями $\mathsf{D}\xi_1 \leqslant h$, $\mathsf{D}\xi_2 \leqslant h, \ldots$ некоторой постоянной h. Тогда последовательность $\{\eta_n = n^{-1}\sum_{i=1}^{n}(\xi_i - \mathsf{M}\xi_i);\ n \geqslant 1\}$ случайных величин $\eta_1, \eta_2, \ldots$ сходится по вероятности к нулю при $n \to \infty$.*

Доказательство. Для доказательства теоремы требуется для любого $\varepsilon > 0$ установить предельное равенство $\lim\limits_{n\to\infty} \mathsf{P}(\{\omega\colon |\eta_n| < \varepsilon\}) = 1$. Иначе: для любых $\varepsilon, \delta > 0$ достаточно найти натуральное число $N(\varepsilon, \delta)$, для которого выполняется неравенство $1 - \delta < \mathsf{P}(\{\omega\colon |\eta_n| < \varepsilon\}) < 1 + \delta$ при всех значениях $n \geqslant N(\varepsilon, \delta)$. Второе неравенство Чебышева существенно применяется для доказательства заключительной части теоремы. Используя второе неравенство Чебышева и полученную ранее оценку $\mathsf{D}\eta_n \leqslant n^{-1}h$ для дисперсии величины η_n, находим $\mathsf{P}(\{\omega\colon |\eta_n| \geqslant \varepsilon\}) \leqslant \varepsilon^{-2}\mathsf{D}\eta_n \leqslant \varepsilon^{-2}n^{-1}h < \delta$, если взять n достаточно большое, например, $n \geqslant N(\varepsilon, \delta) \geqslant h(\delta\varepsilon^2)^{-1} + 1$. Отсюда, переходя к противоположному событию, получаем для $\varepsilon, \delta > 0$ и $n \geqslant N(\varepsilon, \delta) \geqslant h(\delta\varepsilon^2)^{-1} + 1$ соотношение $1 - \delta < \mathsf{P}(\{\omega\colon |\eta_n| < \varepsilon\}) < 1 + \delta$. Теорема 11.4 доказана.

Так как из попарной независимости случайных величин следует их попарная некоррелированность, то имеет место следствие.

Следствие 11.1. *Предположим, что случайные величины ξ_1, $\xi_2, \ldots$ являются попарно независимыми, имеют конечные математические ожидания и равномерно ограниченные дисперсии. Тогда последовательность $\{\eta_n = n^{-1}\sum_{i=1}^{n}(\xi_i - \mathsf{M}\xi_i);\ n \geqslant 1\}$ сходится по вероятности к нулю при $n \to \infty$.*

2.2. Законы больших чисел в форме Чебышева. В теореме 11.4 показывается сходимость по вероятности специально построенной последовательности случайных величин к нулю. Поэтому эту теорему можно считать простейшим вариантом закона больших чисел. На содержательном уровне законы больших чисел устанавливают важную связь между средним арифметическим $n^{-1}\sum_{i=1}^{n}\xi_i$ однотипных в некотором смысле случайных величин ξ_1, $\xi_2, \ldots$ и их математическими ожиданиями.

Теорема 11.5. *Если ξ_1, $\xi_2, \ldots$ являются попарно независимыми случайными величинами, для которых существуют конечные $\mathsf{M}\xi_i = a$ и $\mathsf{M}\xi_i = \sigma^2$ при всех $i = 1, 2, \ldots$. Тогда последовательность $\{\eta_n = n^{-1}\sum_{i=1}^{n}\xi_i;\ n \geqslant 1\}$ из средних арифметических случайных величин $\xi_1, \xi_2, \ldots$ сходится по вероятности к $\mathsf{M}\xi_i = a$ при $n \to \infty$.*

Доказательство. Так как математическое ожидание и дисперсия случайной величины η_n соответственно равны

$\mathsf{M}\eta_n = \mathsf{M}\big(n^{-1}\sum_{i=1}^n \xi_i\big) = \mathsf{M}\xi_i = a$ и $\mathsf{D}\eta_n = \mathsf{D}\big(n^{-1}\sum_{i=1}^n \xi_i\big) = n^{-2}\mathsf{D}\big(\sum_{i=1}^n \xi_i\big) = n^{-2}\sum_{i=1}^n \mathsf{D}\xi_i = n\sigma^2/n^2 = \sigma^2/n$. Применяя второе неравенство Чебышева для случайной величины η_n, находим $\mathsf{P}(\{\omega\colon |\eta_n - a| \geqslant \varepsilon\}) \leqslant \mathsf{D}\eta_n/\varepsilon^2 = \sigma^2/(n\varepsilon^2) < \delta$, если $n \geqslant N(\varepsilon,\delta) \geqslant \sigma^2(\delta\varepsilon^2)^{-1} + 1$. Это означает, что последовательность $\{\eta_n;\, n = 1,2,\ldots\}$ из средних арифметических случайных величин ξ_1, $\xi_2,\ldots$ сходится по вероятности к постоянной величине $\mathsf{M}\xi_i = a$ при $n \to \infty$. Отметим, что условия теоремы 11.5 сильнее ограничений следствия 11.1. Более того, при $\mathsf{M}\xi_i = a$ последовательность $\{n^{-1}\sum_{i=1}^n(\xi_i - \mathsf{M}\xi_i);\, n = 1,2,\ldots\}$ совпадает с последовательностью $\{n^{-1}\sum_{i=1}^n \xi_i - a;\, n \geqslant 1\}$. Поэтому последовательность $\{n^{-1}\sum_{i=1}^n \xi_i - a;\, n \geqslant 1\}$ сходится по вероятности к нулю при $n \to \infty$. Другими словами, последовательностью $\{n^{-1}\sum_{i=1}^n \xi_i;\, n = 1,2,\ldots\}$ сходится по вероятности к постоянной $\mathsf{M}\xi_i = a$ при $n \to \infty$, т. е. теорема 11.5 вытекает из следствия 11.1.

Приведем теперь простейший вариант закона больших чисел в форме Чебышева, который очень часто применяется на практике.

Следствие 11.2. *Пусть ξ_1, $\xi_2,\ldots$ являются независимыми и одинаково распределенными случайными величинами, каждая из которых имеет конечное математическое ожидание $\mathsf{M}\xi_i = a$ и конечную дисперсию σ^2. Тогда последовательность $\{n^{-1}\sum_{i=1}^n \xi_i;\, n \geqslant 1\}$ из средних арифметических величин ξ_1, $\xi_2,\ldots$ сходится по вероятности к $\mathsf{M}\xi_i = a$ при $n \to \infty$.*

В условиях теоремы 11.5 имеет место более сильное и важное для практики утверждение, а именно, усиленный закон больших чисел.

Теорема 11.6. *Пусть ξ_1, $\xi_2,\ldots$ являются попарно независимыми случайными величинами, для которых существуют конечные $\mathsf{M}\xi_i = a$ и $\mathsf{D}\xi_i = \sigma^2$ при всех $i = 1,2,\ldots$. Тогда последовательность $\{\eta_n = n^{-1}\sum_{i=1}^n \xi_i;\, n \geqslant 1\}$ из средних арифметических случайных величин ξ_1, $\xi_2,\ldots$ сходится с вероятностью единица к $\mathsf{M}\xi_i = a$ при $n \to \infty$.*

Доказательство. Справедливы следующие очевидные соотношения: $\{\omega\colon \lim\limits_{n\to\infty} \eta_n(\omega) = a\} = \{\omega\colon \lim\limits_{n\to\infty} (\eta_n(\omega) - a) = 0\}$, $\mathsf{M}(\xi_i - a) = 0$, $\mathsf{D}(\xi_i - a) = \sigma^2$, $\eta_n(\omega) - a = n^{-1}\sum_{i=1}^n \xi_i - a = n^{-1}\sum_{i=1}^n(\xi_i - a)$. Поэтому при $a \neq 0$ и $i, n = 1,2,\ldots$ вместо случайных величин ξ_i, η_n можно рассматривать соответственно случайные величины $\xi_i - a$, $\eta_n - a$. Итак, не ограничивая общности, в дальнейшем полагаем, что $\mathsf{M}\xi_i = 0$ для всех $i = 1,2,\ldots$. Следовательно, $\mathsf{M}\eta_n = 0$ и $\mathsf{D}\eta_n = \sigma^2/n$. Используя теперь второе неравенство Чебышева для случайной величины η_n, получим при $n = k^2$ и любом $\varepsilon > 0$ следующее неравенство: $\mathsf{P}(\{\omega\colon |\eta_{k^2}| \geqslant \varepsilon\}) \leqslant \mathsf{D}(\eta_{k^2})/\varepsilon^2 = \sigma^2/(k^2\varepsilon^2)$.

При $k = 1, 2, \ldots$ введем случайные величины

$$\theta_{k^2} = \max\{\xi_{k^2+1}, \xi_{k^2+1} + \xi_{k^2+2}, \ldots, \xi_{k^2+1} + \xi_{k^2+2}, \ldots, \xi_{(k+1)^2-1}\}.$$

Для каждого натурального числа n определим натуральное число $k = k(n) = [n^{1/2}]$, равное целой части действительного числа $n^{1/2}$. Тогда для всех n из промежутка $k^2 \leqslant n < (k+1)^2$ выполняются следующие соотношения:

$$\sum_{i=1}^{n} \xi_i = \sum_{i=1}^{k^2} \xi_i + \sum_{i=k^2+1}^{n} \xi_i,$$

$$\Big|\sum_{i=1}^{n} \xi_i\Big| \leqslant \Big|\sum_{i=1}^{k^2} \xi_i\Big| + \Big|\sum_{i=k^2+1}^{n} \xi_i\Big| \leqslant \Big|\sum_{i=1}^{k^2} \xi_i\Big| + \Big|\theta_{k^2}\Big|,$$

$$\Big|n^{-1}\sum_{i=1}^{n} \xi_i\Big| \leqslant \Big|k^{-2}\sum_{i=1}^{k^2} \xi_i\Big| + \Big|k^{-2}\theta_{k^2}\Big|, \quad \Big|\eta_n\Big| \leqslant \Big|\eta_{k^2}\Big| + \Big|k^{-2}\theta_{k^2}\Big|.$$

Заметим, что в этих соотношениях при $n = k^2$ полагаем $\sum_{i=k^2+1}^{n} \xi_i = 0$. Следовательно, для каждого n из промежутка $k^2 \leqslant n < (k+1)^2$ для случайных величин η_n, η_{k^2}, θ_{k^2} имеет место неравенство $|\eta_n| \leqslant |\eta_{k^2}| + |k^{-2}\theta_{k^2}|$. Далее, из определения случайной величины θ_{k^2} следует, что при любом $\varepsilon > 0$ событие $\{\omega\colon |\theta_{k^2}| \geqslant k^2\varepsilon\}$ можно представить в виде $\cup_{m=1}^{2k}\{\omega\colon \big|\sum_{r=1}^{m} \xi_{k^2+r}\big| \geqslant k^2\varepsilon\}$. Отсюда, используя сначала свойство вероятностной функции, затем попарную независимость случайных величин ξ_{k^2+1}, $\xi_{k^2+2}, \ldots$, ξ_{k^2+m} для вычисления дисперсии $\mathsf{D}\big(\sum_{r=1}^{m} \xi_{k^2+r}\big) = m\sigma^2$ и, наконец, второе неравенство Чебышева, последовательно найдем:

$$\begin{aligned} \mathsf{P}(\{\omega\colon |\theta_{k^2}| \geqslant k^2\varepsilon\}) &= \mathsf{P}\Big(\bigcup_{m=1}^{2k} \{\omega\colon \Big|\sum_{r=1}^{m} \xi_{k^2+r}\Big| \geqslant k^2\varepsilon\}\Big) \leqslant \\ &\leqslant \sum_{m=1}^{2k} \mathsf{P}\Big(\{\omega\colon \Big|\sum_{r=1}^{m} \xi_{k^2+r}\Big| \geqslant k^2\varepsilon\}\Big) \leqslant \sum_{m=1}^{2k} m\sigma^2 k^{-4}\varepsilon^{-2} = \\ &= k(2k+1)\sigma^2 k^{-4}\varepsilon^{-2} \leqslant 3\sigma^2(k\varepsilon)^{-2}. \end{aligned}$$

Значит, при $\varepsilon > 0$ имеем $\mathsf{P}(\{\omega\colon |k^{-2}\theta_{k^2}| \geqslant \varepsilon\}) \leqslant 3\sigma^2(k\varepsilon)^{-2}$. Отсюда и из неравенства $\mathsf{P}(\{\omega\colon |\eta_{k^2}| \geqslant \varepsilon\}) \leqslant \sigma^2(k\varepsilon)^{-2}$ получим, что при $\varepsilon > 0$ ряды $\sum_{k=1}^{\infty} \mathsf{P}(\{\omega\colon |k^{-2}\theta_{k^2}| \geqslant \varepsilon\})$, $\sum_{k=1}^{\infty} \mathsf{P}(\{\omega\colon |\eta_{k^2}| \geqslant \varepsilon\})$, члены которых неотрицательны, сходятся. Поэтому по лемме 11.2 каждая из последовательностей $\{\eta_{k^2};\ k \geqslant 1\}$, $\{k^{-2}\theta_{k^2};\ k \geqslant 1\}$ сходится с вероятностью единица к нулю. Отсюда, принимая во внимание доказанное ранее неравенство вида $|\eta_n| \leqslant |\eta_{k^2}| + |k^{-2}\theta_{k^2}|$, получаем, что последовательность $\{\eta_n;\ n \geqslant 1\}$ сходится при $n \to \infty$ к нулю с вероятностью единица. Итак, доказан усиленный закон больших чисел в простейшем варианте.

Доказательство законов больших чисел на основе второго неравенства Чебышева по существу является элементарным, так как используется жесткое требование равномерной ограниченности дисперсий. Колмогоров доказал усиленный закон больших чисел, не требуя этих ограничений.

Теорема 11.7 (теорема Колмогорова). *Если величины* $\xi_1, \xi_2, \ldots$ *независимы, одинаково распределены и существует конечное* $\mathsf{M}\xi_i = a$, *то последовательность* $\{n^{-1}\sum_{i=1}^{n}\xi_i;\ n = 1, 2, \ldots\}$ *из средних арифметических случайных величин* ξ_1, $\xi_2, \ldots$ *при* $n \to \infty$ *сходится с вероятностью единица к* $\mathsf{M}\xi_i = a$.

Вместо доказательства этой теоремы, которое приводится в учебниках по теории вероятностей [1, 6, 23], рассмотрим подробно важный и поучительный пример применения простейшего варианта закона больших чисел в форме Чебышева (следствие 11.2) и теоремы Колмогорова.

Пример 11.1 (*задача о приближенных измерениях*). Пусть случайная величина ξ с интегральной функцией распределения $F(x)$, для которой математическое ожидание $\mathsf{M}\xi$ и дисперсия $\mathsf{D}\xi$ конечны, определяет некоторый количественный признак эксперимента E_0. Проведем последовательность независимых испытаний над экспериментом E_0, в результате которых измеряется случайная величина ξ. Обозначим через ξ_1 случайную величину, которая соответствует величине ξ в первом опыте $E_1 = E_0$, через ξ_2 — случайную величину, соответствующую величине ξ во втором опыте $E_2 = E_0$, и т. д. Другими словами, случайная величина ξ_i характеризует количественный признак ξ на i-том испытании эксперимента $E_0 = E_i$, где $i \geqslant 1$. При любом фиксированном $i = 1, 2, \ldots$ обозначим через $(\Omega_i, \mathcal{F}_i, \mathsf{P}_i(\cdot))$ вероятностную модель эксперимента E_i, каждая из которых естественно совпадает с вероятностной моделью $(\Omega_0, \mathcal{F}_0, \mathsf{P}_0(\cdot))$ эксперимента E_0. Здесь следует заметить, что каждая случайная величина $\xi_i(\omega_i)\colon \Omega_i \to R$ определена на своем вероятностном пространстве $(\Omega_i, \mathcal{F}_i, \mathsf{P}_i(\cdot))$. Поэтому для применения следствия 11.2 вместо последовательности ξ_1, $\xi_2, \ldots$ требуется построить новую последовательность независимых и одинаково распределенных случайных величин ς_1, $\varsigma_2, \ldots$ на некотором вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. Причем интегральная функция распределения каждой случайной величины ς_i должна совпадать с функцией $F(x)$.

Для решения этой непростой задачи рассмотрим эволюционный эксперимент $E = \{E_1, E_2, \ldots\}$, для которого в первой главе была построена теоретико-множественная модель $(\Omega_0^\infty, \mathcal{F}_0^\infty)$. Напомним, что достоверный исход Ω_0^∞ эволюционного эксперимента E совпадает с пространством упорядоченных последовательностей вида $\omega = (\omega_1, \omega_2, \ldots)$, где элемент $\omega_i \in \Omega_0 = \Omega_i$ для всех $i = 1, 2, \ldots$. Множество $\mathcal{F}_0^\infty$ наблюдаемых исходов эволюционного эксперимента E

равно наименьшей σ-алгебре, которая содержит все цилиндрические множества

$$\mathbf{Z}(A_1 \times A_2 \times \ldots \times A_n) = \{\omega = (\omega_1, \omega_2, \ldots)\colon \omega_1 \in A_1, \omega_2 \in A_2, \ldots, \omega_n \in A_n\},$$

где $A_1 \in \mathcal{F}_1$, $A_2 \in \mathcal{F}_2, \ldots$, $A_n \in \mathcal{F}_n$ и $n = 1, 2, \ldots$. Определим вероятность $\mathsf{P}(\mathbf{Z}(A_1 \times A_2 \times \ldots \times A_n))$ на произвольном цилиндрическом множестве $\mathbf{Z}(A_1 \times A_2 \times \ldots \times A_n)$ по формуле $\mathsf{P}(\mathbf{Z}(A_1 \times A_2 \times \ldots \times A_n)) = \mathsf{P}_1(A_1) \times \mathsf{P}_2(A_2) \times \ldots \times \mathsf{P}_n(A_n)$. Известно, что вероятность на цилиндрических множествах может быть однозначно продолжена на σ-алгебру $\mathcal{F}_0^\infty$. Доказательство этого факта можно найти в различных учебниках и монографиях по теории вероятностей (*Лоэв М.* Теория вероятностей. — М.: ИЛ, 1962). Построенное таким способом вероятностное пространство $(\Omega_0^\infty, \mathcal{F}_0^\infty, \mathsf{P}_0(\cdot))$ называется прямым произведением вероятностных пространств $(\Omega_i, \mathcal{F}_i, \mathsf{P}_i(\cdot))$, $i = 1, 2, \ldots$. В качестве искомого вероятностного пространства $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ теперь можно взять вероятностное пространство $(\Omega_0^\infty, \mathcal{F}_0^\infty, \mathsf{P}_0(\cdot))$. При этом каждая случайная величина $\varsigma_i(\omega)\colon \Omega \to R$ определяется на новом вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ равенством $\varsigma_i(\omega) = \xi_i(\omega_i)$, где $\omega = (\omega_1, \omega_2, \ldots)$.

Покажем, что случайные величины ς_1, $\varsigma_2, \ldots$ являются независимыми в совокупности и имеют одинаковую функцию распределения $F(x)$. В самом деле, пусть B_1, $B_2, \ldots$ являются произвольными борелевскими множествами на действительной прямой R. Тогда $A_1 = \{\omega_1\colon \xi_1(\omega_1) \in B_1\} \in \mathcal{F}_1$, $A_2 = \{\omega_2\colon \xi_2(\omega_2) \in B_2\} \in \mathcal{F}_2, \ldots$. Отсюда при любом $n \geqslant 1$ получаем

$$\begin{aligned}\mathsf{P}(\{\omega\colon \varsigma_n(\omega) \in B_n\}) &= \mathsf{P}(\{\omega = (\omega_1, \omega_2, \ldots)\colon \xi_n(\omega_n) \in B_n\}) = \\ &= \mathsf{P}(\{\omega = (\omega_1, \omega_2, \ldots)\colon \omega_n \in A_n\}) = \mathsf{P}_n(A_n),\end{aligned}$$

$$\begin{aligned}\mathsf{P}(\{\omega\colon \varsigma_1(\omega) \in B_1, \varsigma_2(\omega) \in B_2, \ldots, \varsigma_n(\omega) \in B_n\}) &= \mathsf{P}\Big(\bigcap_{i=1}^n \{\omega\colon \varsigma_i(\omega) \in B_i\}\Big) = \\ &= \mathsf{P}(\bigcap_{i=1}^n \{\omega = (\omega_1, \omega_2, \ldots)\colon \xi_i(\omega_i) \in B_i\}) = \\ &= \mathsf{P}\Big(\bigcap_{i=1}^n \{\omega = (\omega_1, \omega_2, \ldots)\colon \omega_i \in A_i\}\Big) = \\ &= \mathsf{P}(\{\omega = (\omega_1, \omega_2, \ldots) : \omega_1 \in A_1, \omega_2 \in A_2, \ldots, \omega_n \in A_n\}) = \\ &= \prod_{i=1}^n \mathsf{P}_n(A_n) = \prod_{i=1}^n \mathsf{P}(\{\omega : \varsigma_n(\omega) \in B_n\}).\end{aligned}$$

Наконец, если $B_n = (-\infty, x)$, то $\mathsf{P}(\{\omega\colon \varsigma_n(\omega) < x\}) = \mathsf{P}_n(\{\omega_n\colon \xi_n(\omega_n) < x\}) = F(x)$, $\mathsf{M}\varsigma_n = \mathsf{M}\xi$ и $\mathsf{D}\varsigma_n = \mathsf{D}\xi$ при всех $n = 1, 2, \ldots$ Итак, построенная последовательность ς_1, $\varsigma_2, \ldots$ на $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$ удовлетворяет условиям следствия 11.2. Поэтому для последовательности $\{n^{-1}\sum_{i=1}^{n}\varsigma_i;\ n=1,2,\ldots\}$ при любых $\varepsilon,\delta>0$ верно неравенство $\mathsf{P}(\{\omega\colon \left|n^{-1}\sum_{i=1}^{n}\varsigma_i - \mathsf{M}\xi\right| < \varepsilon\}) > 1-\delta$, если $n \geqslant N(\varepsilon,\delta) = (\delta\varepsilon^2)^{-1}\mathsf{D}\xi + 1$. Более того, по теореме Колмогорова имеет место более сильное утверждение, которое заключается в том, что почти для всех $\omega\in\Omega$ справедливо соотношение $\lim\limits_{n\to\infty} n^{-1}\sum_{i=1}^{n}\varsigma_i(\omega) = \mathsf{M}\xi$. Таков точный смысл приближенного равенства $n^{-1}\sum_{i=1}^{n}\varsigma_i(\omega)\approx\mathsf{M}\xi$. Так как при любом $\omega=(\omega_1,\omega_2,\ldots)\in\Omega$ имеет место соотношение $n^{-1}\sum_{i=1}^{n}\varsigma_i(\omega)\equiv n^{-1}\sum_{i=1}^{n}\xi_i(\omega_i)$, то можно сказать, что при достаточно большом числе n проведения независимых опытов среднее арифметическое наблюдаемых значений случайной величины приблизительно равно $\mathsf{M}\xi$. Этим фактом пользуются на практике при приближенном вычислении некоторой неизвестной величины, например, расстояния между населенными пунктами.

Теорема 11.8 (теорема Бернулли). *Если* $(\Omega_0,\mathcal{F}_0,\mathsf{P}_0(\cdot))$ *есть вероятностная модель эксперимента* E_0, *то при неограниченном увеличении числа* n *независимых опытов над этим экспериментом относительная частота* $n^{-1}\mu(A,n)$ *случайного события* $A\in\mathcal{F}_0$ *сходится по вероятности к вероятности* $\mathsf{P}_0(A)$.

Доказательство. Обозначим вероятность $\mathsf{P}_0(A)$ через p. Пусть ξ_1 — число появлений события A в первом опыте $E_1=E_0$, ξ_2 — число появлений события A во втором опыте $E_2=E_0$ и т. д. Будем теперь использовать для этого частного случая все обозначения и методику решения задачи о приближенных измерениях. Ряд распределения каждой случайной величины $\xi_i(\omega_i)\colon\Omega_i\to\{0,1\}$, которая определяется на $(\Omega_i,\mathcal{F}_i,\mathsf{P}_i(\cdot))$ и называется *бернуллиевской*, имеет следующий вид (табл. 11.2):

Таблица 11.2

ξ_i	0	1
$\mathsf{P}_i(\cdot)$	$1-p$	p

Математическое ожидание и дисперсия для случайной величины ξ_i соответственно равны $\mathsf{M}\xi_i=p<1$ и $\mathsf{D}\xi_i=p\,(1-p)<1$. Относительная частота $n^{-1}\mu(A,n)$ или статистическая вероятность события A за n опытов может быть вычислена следующим образом: $n^{-1}\mu(A,n)=n^{-1}\sum_{i=1}^{n}\xi_i(\omega_i)=n^{-1}\sum_{i=1}^{n}\varsigma_i(\omega)$, где $\omega=(\omega_1,\omega_2,\ldots)\in\Omega$, $\varsigma_i(\omega)=\xi_i(\omega_i)$ и, значит, случайная величина $\varsigma_i(\omega)\colon\Omega\to\{0,1\}$ определена на вероятностном пространстве $(\Omega,\mathcal{F},\mathsf{P}(\cdot))$. На основании решения задачи о приближенных измерениях (см. пример 11.1) получаем, что $\mathsf{P}(\{\omega\colon |n^{-1}\mu(A,n)-p|<\varepsilon\})>1-\delta$ при $\varepsilon,\delta>0$ и $n\geqslant(\delta\varepsilon^2)^{-1}p\,(1-p)+1$, или $\lim\limits_{n\to\infty}\mathsf{P}(\{\omega\colon |n^{-1}\mu(A,n)-p|<\varepsilon\})=1$.

Итак, теорема Бернулли доказана и здесь имеет место закон больших чисел.

Используя рассуждения из доказательства теоремы Бернулли и утверждение теоремы 11.6, выводим, что относительная частота $n^{-1}\mu(A,n)$ случайного события $A \in \mathcal{F}_0$ при $n \to \infty$ сходится с вероятностью единица к вероятности $\mathsf{P}_0(A) = p$. Другими словами, в схеме Бернулли для относительного числа успехов $n^{-1}\mu(A,n)$ имеет место усиленный закон больших чисел. При этом надо иметь в виду следующее обстоятельство.

Если A не является достоверным исходом экспериментом E_0, то существуют такие описания $\omega_0 = (\omega_{10}, \omega_{20}, \ldots) \in \Omega$, например, $\omega_{10} \in A$, $\omega_{20} \in A, \ldots$, что $n^{-1}\mu(A,n) = n^{-1}\sum_{n=1}^{\infty} \varsigma_i(\omega_0) = 1$ для всех $n = 1, 2, \ldots$. Ясно, что это соответствует такой логической возможности, когда в схеме Бернулли постоянно наблюдается событие A. Поэтому для такого случая $\lim\limits_{n\to\infty} n^{-1}\mu(A,n) = 1 \neq \mathsf{P}_0(A) = p$. Ясно, что существуют и другие описания $\omega = (\omega_1, \omega_2, \ldots) \in \Omega$, для которых указанный предел либо не существует, либо существует, но не совпадает с $\mathsf{P}_0(A) = p$. Поэтому предельное поведение относительной частоты $n^{-1}\mu(A,n)$ случайного события A нельзя использовать для строгого и логического определения его вероятности. С другой стороны, усиленный закон больших чисел для относительных частот $n^{-1}\mu(A,n)$, $n \geqslant 1$, гарантирует выполнение соотношения $\mathsf{P}(\{\omega\colon \lim\limits_{n\to\infty} n^{-1}\mu(A,n) \nrightarrow p\}) = 0$. Это равенство вполне соответствует интуитивному представлению о вероятности события A и оправдывает статистический подход для приближенного вычисления $\mathsf{P}_0(A)$. На самом деле, мы пренебрегаем случаями, когда $\lim\limits_{n\to\infty} n^{-1}\mu(A,n) \nrightarrow p$ и адекватно говорим об устойчивости или о надежности относительной частоты $n^{-1}\mu(A,n)$ случайного события A.

Используя теорему Бернулли и интегральную предельную теорему Муавра–Лапласа в схеме независимых испытаний, еще раз поясним целесообразность требования равенства $\lim\limits_{n\to\infty} F_{\xi_n}(x) = F_\xi(x)$ лишь в точках непрерывности $F_\xi(x)$ при определении сходимости по распределению последовательности ξ_1, $\xi_2, \ldots$ к величине ξ. По теореме Бернулли последовательность $\{\eta_n = n^{-1}\sum_{i=1}^{n} \varsigma_i;\ n \geqslant 1\}$ сходится по вероятности к постоянной величине $\mathsf{P}_0(A) = p$. Поэтому эта последовательность случайных величин сходится по распределению к вырожденной случайной величине η, для которой интегральная функция $F_\eta(x) = 0$ при $x \leqslant p$ и $F_\eta(x) = 1$ при $x > p$.

Рассмотрим теперь свойства интегральных функций $F_{\eta_n}(x)$, $n \geqslant 1$. Так как случайная величина η_n принимает неотрицательные значения, то функция $F_{\eta_n}(x) = 0$ при $x \leqslant 0$. Пусть теперь $x > 0$, а величина $\langle nx \rangle$ равна такому наибольшему целому числу, что $\langle nx \rangle < nx$. Из определения величин η_n, ς_1, $\varsigma_2, \ldots$, ς_n, $\mu(A,n)$ при $q = 1 - p$ вытекают соотношения:

$$\begin{aligned} F_{\eta_n}(x) &= \mathsf{P}(\{\omega\colon \eta_n < x\}) = \mathsf{P}\Big(\Big\{\omega\colon -\infty < n^{-1}\sum_{i=1}^{n}\varsigma_i < x\Big\}\Big) = \\ &= \mathsf{P}(\{\omega\colon -\infty < n^{-1}\mu(A,n) < x\}) = \\ &= \mathsf{P}(\{\omega\colon -\infty < \mu(A,n) < nx\}) = \sum_{m=0}^{\langle nx\rangle} C_n^m p^m q^{n-m}. \end{aligned}$$

Применяя интегральную предельную теорему Муавра–Лапласа, получим, что $F_{\eta_n}(x) = 2^{-1}\{\Phi((\langle nx\rangle - np)(npq)^{-1/2}) + \Phi(np\,(npq)^{-1/2})\} + \alpha_n$, где $\lim\limits_{n\to\infty} \alpha_n = 0$. Непосредственно из такого представления функции $F_{\eta_n}(x)$ находим, что предел $\lim\limits_{n\to\infty} F_{\eta_n}(x)$ равен нулю при $x < p$, равен единице при $x > p$ и, наконец, равен 0,5 при $x = p$. Значит, $\lim\limits_{n\to\infty} F_{\eta_n}(x) = F_\eta(x)$ для всех действительных чисел $x \neq p$. Заметим, что только при $x = p$ функция $F_\eta(x)$ терпит разрыв и имеет место $\lim\limits_{n\to\infty} F_{\eta_n}(p) \neq F_\eta(p)$. Даже этот простой пример, который связан со схемой испытаний Бернулли, показывает, что при определении слабой сходимости не следует требовать поточечную сходимость последовательности функций распределения.

2.3. Характеристические функции. В десятой главе была показана большая роль производящих функций для изучения распределений целочисленных неотрицательных случайных величин и вычисления их числовых характеристик. Аналогичную роль для исследования вероятностных свойств произвольных случайных величин выполняют так называемые характеристические функции. Особенно удобным оказался аппарат характеристических функций для доказательства центральных предельных теорем, где необходимо изучить свойства распределений сумм независимых случайных величин. Для определения характеристической функции потребуется рассмотрение комплексных случайных величин.

Пусть (ξ, η) — двумерная случайная величина. Каждому описанию $\omega \in \Omega$ элементарного исхода поставим в соответствие комплексное число вида $\xi(\omega) + i\eta(\omega)$, где $i = (-1)^{1/2}$ является мнимой единицей. Если математические ожидания $\mathsf{M}\xi$ и $\mathsf{M}\eta$ существуют, то математическое ожидание $\mathsf{M}(\xi + i\,\eta)$ комплекснозначной случайной величины $\xi + i\,\eta$ по определению равно комплексному числу $\mathsf{M}\xi + i\,\mathsf{M}\eta$. Будем полагать, что комплексные случайные величины $\xi_1 + i\,\xi_2$ и $\eta_1 + i\,\eta_2$ независимы, если независимы двумерные случайные величины (ξ_1, ξ_2) и (η_1, η_2). Теперь легко проверяются основные свойства математического ожидания (например, свойство аддитивности) для комплекснозначных случайных величин. В качестве примера приведем доказательство для независимых комплексных случайных величин $\xi_1 + i\,\xi_2$ и $\eta_1 + i\,\eta_2$ свойства мультипликативности вида $\mathsf{M}((\xi_1 + i\,\xi_2)(\eta_1 + i\,\eta_2) = \mathsf{M}(\xi_1 + i\,\xi_2) \times \mathsf{M}(\eta_1 + i\,\eta_2)$. Преобразуя левую и правую часть этого равенства, найдем

$$\begin{aligned}
\mathsf{M}((\xi_1 + i\xi_2)(\eta_1 + i\eta_2)) &= \mathsf{M}(\xi_1\eta_1 + i\xi_1\eta_2 + i\xi_2\eta_1 - \xi_2\eta_2) = \\
&= \mathsf{M}((\xi_1\eta_1 - \xi_2\eta_2) + i(\xi_1\eta_2 + \xi_2\eta_1)) = \\
&= \mathsf{M}(\xi_1\eta_1 - \xi_2\eta_2) + i\mathsf{M}(\xi_1\eta_2 + \xi_2\eta_1)) = \\
&= \mathsf{M}\xi_1\mathsf{M}\eta_1 - \mathsf{M}\xi_2\mathsf{M}\eta_2 + i\mathsf{M}\xi_1\mathsf{M}\eta_2 + i\mathsf{M}\xi_2\mathsf{M}\eta_1, \\
\mathsf{M}(\xi_1 + i\xi_2) \times \mathsf{M}(\eta_1 + i\eta_2) &= (\mathsf{M}\xi_1 + i\mathsf{M}\xi_2) \times (\mathsf{M}\eta_1 + i\mathsf{M}\eta_2) = \\
&= \mathsf{M}\xi_1\mathsf{M}\eta_1 - \mathsf{M}\xi_2\mathsf{M}\eta_2 + i\mathsf{M}\xi_1\mathsf{M}\eta_2 + i\mathsf{M}\xi_2\mathsf{M}\eta_1.
\end{aligned}$$

Эти преобразования справедливы при $\mathsf{M}|\xi_1| < \infty$, $\mathsf{M}|\xi_2| < \infty$, $\mathsf{M}|\eta_1| < \infty$, $\mathsf{M}|\eta_2| < \infty$ и доказывают мультипликативное свойство.

Определение 11.5. *Характеристической функцией* действительной случайной величины ξ с интегральной функцией распределения $F_\xi(x)$ называется комплекснозначная функция $\psi_\xi(t)$ переменной $t \in R$, которая равна математическому ожиданию $\mathsf{M}(\exp\{it\,\xi\})$.

Используя формулу Эйлера, определение математического ожидания от комплексной случайной величины и определение интеграла Стилтьеса от функции с комплексными значениями, получим:

$$\begin{aligned}
\psi_\xi(t) = \mathsf{M}(\exp\{it\xi\}) = \mathsf{M}(\cos(t\xi) + i\sin(t\xi)) = \mathsf{M}\cos(t\xi) + i\mathsf{M}\sin(t\xi)) = \\
= \int\limits_{-\infty}^{+\infty} \cos(tx)\,dF_\xi(x) + +i\int\limits_{-\infty}^{+\infty} \sin(tx)\,dF_\xi(x) = \\
= \int\limits_{-\infty}^{+\infty} (\cos(tx) + i\sin(tx))\,dF_\xi(x) = \int\limits_{-\infty}^{+\infty} \exp\{itx\}\,dF_\xi(x).
\end{aligned}$$

Применяя свойства интеграла Стилтьеса от функции с комплексными значениями и учитывая известное равенство $|\exp\{itx\}| = 1$, получаем, что $\psi_\xi(t)| = \left|\int_{-\infty}^{+\infty} \exp\{itx\}\,dF_\xi(x)\right| \leqslant \int_{-\infty}^{+\infty} |\exp\{itx\}|\,dF_\xi(x) = = \int_{-\infty}^{+\infty} dF_\xi(x) = 1$. Следовательно, характеристическую функцию можно определить для любой действительной случайной величины ξ. В частности, если случайная величина ξ является дискретной с распределением $p_k = \mathsf{P}(\{\omega\colon \xi(\omega) = x'_k\}) > 0$, $k \geqslant 1$, то

$$\psi_\xi(t) = \int\limits_{-\infty}^{+\infty} \exp\{itx\}\,dF_\xi(x) = \sum_{k=1}^{\infty} \exp\{itx'_k\}\,p_k.$$

Для непрерывной случайной величины ξ с плотностью $f_\xi(x)$ ее характеристическая функция $\psi_\xi(t)$ равна

$$\int\limits_{-\infty}^{+\infty} \exp\{itx\}\, dF_\xi(x) = \int\limits_{-\infty}^{+\infty} \exp\{itx\}\, f_\xi(x)\, dx,$$

т. е. функция $(2\pi)^{-1/2}\psi_\xi(t)$ совпадает с преобразованием Фурье от плотности $f_\xi(x)$. Отметим основные свойства характеристических функций, некоторые из них (свойства X.4, X.6) приведем без доказательства. Простые доказательства свойств X.4 и X.6 можно найти, например, в [1] и [23].

X.1. Так как $\int_{-\infty}^{+\infty} dF_\xi(x) = 1$ и

$$\mathsf{M}(\exp\{it(b\xi + c)\}) = \exp\{itc\}\, \mathsf{M}(\exp\{itb\,\xi\}) = \exp\{itc\}\, \psi_\xi(bt),$$

то для любой случайной величины ξ выполняются следующие равенства: $\psi_\xi(0) = 1$, $\psi_{b\,\xi+c}(t) = \exp\{itc\}\, \psi_\xi(bt)$.

X.2. Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ являются независимыми и $\eta_n = \xi_1 + \xi_2 + \ldots + \xi_n$, то $\psi_{\eta_n}(t) = \psi_{\xi_1}(t) \times \psi_{\xi_2}(t) \times \ldots \times \psi_{\xi_n}(t)$. В самом деле, из независимости случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ следует независимость случайных величин $\exp\{it\,\xi_1\}, \exp\{it\,\xi_2\}, \ldots, \exp\{it\,\xi_n\}$. Используя свойство математического ожидания от произведения независимых комплексных случайных величин, получим

$$\begin{aligned}
\psi_{\eta_n}(t) &= \mathsf{M}(\exp\{it(\xi_1 + \xi_2 + \ldots + \xi_n)\}) = \\
&= \mathsf{M}(\exp\{it\,\xi_1\} \times \exp\{it\,\xi_2\} \times \ldots \times \exp\{it\,\xi_n\}) = \\
&= \mathsf{M}(\exp\{it\,\xi_1\}) \times \mathsf{M}(\exp\{it\,\xi_2\}) \times \ldots \times \mathsf{M}(\exp\{it\,\xi_n\}) = \\
&= \psi_{\xi_1}(t) \times \psi_{\xi_2}(t) \times \ldots \times \psi_{\xi_n}(t).
\end{aligned}$$

X.3. Пусть абсолютный момент n-го порядка случайной величины ξ удовлетворяет условию $\mathsf{M}|\xi|^n < +\infty$. Тогда характеристическая функция $\psi_\xi(t)$ дифференцируема n раз и ее производная порядка k, обозначаемая через $\psi_\xi^{(k)}(t)$, в точке $t = 0$ удовлетворяет равенству $\psi_\xi^{(k)}(0) = i^k \mathsf{M}\xi^k$ для всех $k = 1, 2, \ldots, n$. Более того, для характеристической функции $\psi_\xi(t)$ имеет место соотношение вида

$$\psi_\xi(t) = 1 + it\mathsf{M}\xi + (it)^2\mathsf{M}\xi^2(2!)^{-1} + \ldots + (it)^n\mathsf{M}\xi^n(n!)^{-1} + o(t^n). \quad (11.3)$$

Действительно, так как $\mathsf{M}|\xi|^n = \int_{-\infty}^{+\infty} |x|^n dF_\xi(x) < +\infty$, то, используя свойства интеграла Стилтьеса, легко выводим

$$
\begin{aligned}
\mathsf{M}|\xi|^k &= \int\limits_{-\infty}^{+\infty} |x|^k dF_\xi(x) = \\
&= \int\limits_{-\infty}^{-1-0} |x|^k dF_\xi(x) + \int\limits_{-1-0}^{1+0} |x|^k dF_\xi(x) + \int\limits_{1+0}^{+\infty} |x|^k dF_\xi(x) \leqslant \\
&\leqslant \int\limits_{-1-0}^{1+0} dF_\xi(x) + \int\limits_{-\infty}^{-1-0} |x|^n dF_\xi(x) + \int\limits_{1+0}^{+\infty} |x|^n dF_\xi(x) < +\infty
\end{aligned}
$$

для всех $k = 1, 2, \ldots, n$. Значит, если абсолютный момент n-го порядка случайной величины ξ является действительным числом, то для ξ существует конечный абсолютный момент любого порядка $k = 1, 2, \ldots, n$.

Отсюда, учитывая равенство вида $|x^k \exp\{itx\}| = |x^k|$, получим $\int_{-\infty}^{+\infty} |x^k \exp\{itx\}|\, dF_\xi(x) = \int_{-\infty}^{+\infty} |x^k|\, dF_\xi(x) < +\infty$, $k \leqslant n$. При этом интеграл $\int_{-\infty}^{+\infty} |x^k|\, dF_\xi(x)$ не зависит от переменной $t \in R$. Тогда интеграл $\int_{-\infty}^{+\infty} x^k \exp\{itx\}\, dF_\xi(x)$ равномерно сходится относительно параметра $t \in R$. Это обстоятельство гарантирует дифференцирование n раз функции $\int_{-\infty}^{+\infty} \exp\{itx\}\, dF_\xi(x)$ по параметру t под знаком интеграла. Другими словами, характеристическая функция $\psi_\xi(t)$ допускает k-кратное дифференцирование по параметру $t \in R$, и, следовательно, $\psi_\xi^{(k)}(t) = i^k \int_{-\infty}^{+\infty} x^k \exp\{itx\}\, dF_\xi(x)$, $k = 1, 2, \ldots, n$. Полагая значение аргумента $t = 0$, вычислим $\psi_\xi^{(k)}(0) = i^k \int_{-\infty}^{+\infty} x^k\, dF_\xi(x) = i^k \mathsf{M}\xi^k$ для всех $k = 1, 2, \ldots, n$. Теперь с учетом равенства $\psi_\xi(0) = 1$ можно утверждать, что для функции $\psi_\xi(t)$ в окрестности точки $t = 0$ справедливо разложение (11.3) по формуле Тейлора с дополнительным членом в форме Пеано.

X.4. Интегральная функция $F_\xi(x)$ величины ξ однозначно определяется своей характеристической функцией $\psi_\xi(t)$ с помощью формулы обращения:

$$
\begin{aligned}
&F_\xi(x_2) - F_\xi(x_1) = \\
&\qquad = (2\pi)^{-1} \lim_{y \to +\infty} \int\limits_{-y}^{y} (it)^{-1} (\exp\{-itx_1\} - \exp\{-itx_2\})\, \psi_\xi(t)\, dt,
\end{aligned}
$$

где $x_2 > x_1$ любые значения аргумента функции $F_\xi(x)$, в которых она непрерывна. Пусть теперь $x_2 = v$ и $x_1 = u$ являются такими значениями аргумента функции $F_\xi(\cdot)$, в которых она непрерывна, и $u \to -\infty$. Тогда формула обращения позволяет найти явное выражение для функции вида

$$
F_\xi(v) = (2\pi)^{-1} \lim_{u \to -\infty} \lim_{y \to +\infty} \int\limits_{-y}^{y} (it)^{-1} (\exp\{-itu\} - \exp\{-itv\})\, \psi_\xi(t)\, dt,
$$

где внешний предел берется по таким значениям u, в которых интегральная функция распределения непрерывна. Отсюда и из свойства непрерывности слева интегральной функции распределения $F_\xi(x)$ непосредственно получаем равенство $F_\xi(x) = \lim\limits_{v \to x-0} F_\xi(v)$ для любого числа x, где предел снова берется по таким значениям v, в которых интегральная функция распределения непрерывна.

X.5. Функция распределения $F_\xi(x)$ случайной величины ξ называется *симметричной*, если для любых $x \in R$ имеет место равенство $\mathsf{P}(\{\omega\colon \xi(\omega) < x\}) = \mathsf{P}(\{\omega\colon \xi(\omega) > -x\})$. Для абсолютно непрерывной и симметричной функции $F_\xi(x)$ последовательно имеем

$$\begin{aligned}
&F_\xi(x) = 1 - \mathsf{P}(\{\omega\colon \xi(\omega) \leqslant -x\}),\\
&F_\xi(x) = 1 - \mathsf{P}(\{\omega\colon \xi(\omega) < -x\}),\\
&F_\xi(x) = 1 - F_\xi(-x), \quad f_\xi(x) = f_\xi(-x).
\end{aligned}$$

Другими словами, плотность распределения $f_\xi(x)$ является четной функцией и, значит, симметричной. Покажем, что характеристическая функция $\psi_\xi(t)$ является действительной тогда и только тогда, когда функция распределения $F_\xi(x)$ симметрична. Действительно, если распределение $F_\xi(x)$ симметрично, то $F_\eta(x) = \mathsf{P}(\{\omega\colon \eta(\omega) < x\}) = \mathsf{P}(\{\omega\colon -\xi(\omega) < x\}) = \mathsf{P}(\{\omega\colon \xi(\omega) > -x\}) = \mathsf{P}(\{\omega\colon \xi(\omega) < x\}) = F_\xi(x)$ для случайной величины $\eta = -\xi$. Поэтому $\psi_\xi(t) = \psi_\eta(t) = \mathsf{M}\cos(t\,\xi)$, так как $\psi_\xi(t) = \psi_\eta(t)$ и, значит,

$$\begin{aligned}
&\mathsf{M}(\exp\{it\,\xi\}) = \mathsf{M}(\exp\{it\,\eta\}),\\
&\mathsf{M}\cos(t\xi) + i\,\mathsf{M}\sin(t\xi) = \mathsf{M}\cos(t\eta) + i\,\mathsf{M}\sin(t\eta),\\
&\mathsf{M}\cos(t\xi) + i\,\mathsf{M}\sin(t\xi) = \mathsf{M}\cos(-t\xi) + i\,\mathsf{M}\sin(-t\xi),\\
&\mathsf{M}\cos(t\xi) + i\,\mathsf{M}\sin(t\xi) = \mathsf{M}\cos(t\xi) - i\,\mathsf{M}\sin(t\xi),\\
&\mathsf{M}\sin(t\xi) = 0.
\end{aligned}$$

Обратно, если функция $\psi_\xi(t)$ является вещественной, то $\mathsf{M}\sin(t\,\xi) = 0$ и, значит, $\psi_\xi(t) = \mathsf{M}\cos(t\,\xi)$, а $\psi_\eta(t) = \mathsf{M}\cos(-t\,\xi) + i\,\mathsf{M}\sin(-t\,\xi) = \mathsf{M}\cos(t\,\xi) = \psi_\xi(t)$. Из четвертого свойства характеристических функций непосредственно следует, что $F_\xi(x) = F_\eta(x)$. Отсюда вероятность вида $\mathsf{P}(\{\omega\colon \xi(\omega) < x\}) = \mathsf{P}(\{\omega\colon \eta(\omega) < x\}) = \mathsf{P}(\{\omega\colon -\xi(\omega) < x\}) = \mathsf{P}(\{\omega\colon \xi(\omega) > -x\})$. Это означает, что $F_\xi(x)$ является симметричной.

X.6. Пусть $\psi_{\xi_n}(t)$ и $F_{\xi_n}(t)$ являются характеристической функцией и интегральной функцией распределения случайной величины ξ_n для каждого натурального значения n. Далее, обозначим через $\psi_\xi(t)$ и $F_\xi(x)$ характеристическую функцию и интегральную функцию распределения некоторой случайной величины ξ. Тогда для слабой сходимости последовательности случайных величин $\xi_1,\ \xi_2, \ldots$ к случайной величине ξ необходимо и достаточно, чтобы $\lim\limits_{n\to\infty}\psi_{\xi_n}(t)=\psi_\xi(t)$ при каждом действительном значении аргумента t. Заметим, что доказать предельное равенство $\lim\limits_{n\to\infty}\psi_{\xi_n}(t)=\psi_\xi(t)$, как правило, гораздо проще, чем установить слабую сходимость последовательности случайных величин $\xi_1,\ \xi_2,\ldots$.

Приведем примеры на вычисление характеристических функций.

Пример 11.2. Найти характеристическую функцию биномиальной случайной величины ξ с параметрами $n\geqslant 1,\ 0<p<1$, распределение которой при $q=1-p$ имеет следующий вид: $\mathsf{P}(\{\omega\colon \xi(\omega)=m\})=$ $=C_n^m p^m q^{n-m},\ m=0,1,\ldots,n$.

Решение. Для этой дискретной случайной величины ξ последовательно имеем

$$\psi_\xi(t)=\sum_{m=0}^{n} C_n^m p^m q^{n-m}\exp\{itm\}=$$

$$=\sum_{m=0}^{n} C_n^m (p\,\exp\{it\})^m q^{n-m}=(p\,\exp\{it\}+q)^n.$$

Этот же результат можно получить, если биномиальную случайную величину ξ представить в виде суммы $(\xi_1+\xi_2+\ldots+\xi_n)$ независимых бернуллиевских случайных величин $\xi_k,\ k=1,2,\ldots,n$, каждая из которых имеет одно и то же распределение следующего вида: $\mathsf{P}(\{\omega\colon \xi_k(\omega)=1\})=p,\ \mathsf{P}(\{\omega\colon \xi_k(\omega)=0\})=q$. Для бернуллиевской случайной величины ξ_k легко получим, что $\psi_{\xi_k}(t)=p\exp\{it\}+q$. Используя теперь второе свойство характеристических функций находим, что $\psi_{\xi_k}(t)=(p\exp\{it\}+q)^n$.

Пример 11.3. Пусть ξ — пуассоновская случайная величина с параметром $\lambda>0$ и распределением $\mathsf{P}(\{\omega\colon \xi(\omega)=m\})=$ $=(m!)^{-1}\lambda^m\exp\{-\lambda\}),\ m=0,1,\ldots$. Тогда для случайной величины ξ характеристическая функция

$$\psi_\xi(t)=\sum_{m=0}^{\infty}\exp\{itm\}(m!)^{-1}\lambda^m\exp\{-\lambda\}=$$

$$=\exp\{-\lambda\}\sum_{m=0}^{\infty}(m!)^{-1}(\lambda\exp\{it\})^m=$$

$$=\exp\{-\lambda\}\exp\{\lambda\exp\{it\}\}=\exp\{\lambda(e^{it}-1)\}.$$

Пример 11.4. Если величина ξ распределена по закону Бартлетта, т.е. $\mathsf{P}(\{\omega\colon \xi(\omega)=1\})=1-p,\ \mathsf{P}(\{\omega\colon \xi(\omega)=m\})=p\,(1-q)\,q^{m-2}$, где $m=2,3,\ldots$ и параметры p и q удовлетворяют условию $0<p,q<1$, то характеристическая функция

$$\psi_\xi(t)=(1-p)\,e^{it}+\sum_{m=2}^{\infty}e^{itm}p\,(1-q)\,q^{m-2}=$$

$$= (1-p)\,e^{it} + p\,(1-q)\,e^{2it} \sum_{m=2}^{\infty} (qe^{it})^{m-2} =$$

$$= (1-p)\,e^{it} + p\,(1-q)\,e^{2it}(1-qe^{it})^{-1}.$$

В частном случае, при $p = q$ получаем характеристическую функцию $\psi_\xi(t) = (1-p)\,e^{it} + p\,(1-p)\,e^{2it}(1-pe^{it})^{-1} = (1-p)\,e^{it}(1-pe^{it})^{-1}$ геометрического распределения.

Пример 11.5. Пусть ξ — равномерная величина, для которой плотность $f_\xi(x) = (b-a)^{-1}$ при $a < x \leqslant b$ и $f_\xi(x) = 0$ при $x \notin (a, b]$. Тогда характеристическая функция $\psi_\xi(t) = (b-a)^{-1}\int_a^b \exp\{itx\}\,dx = (e^{itb} - e^{ita})(it(b-a))^{-1}$. Если $a = -c < 0$ и $b = c$, то плотность распределения равномерной случайной величины будет симметричной относительно оси ординат, а характеристическая функция $\psi_\xi(t) = (e^{itc} - e^{-itc})(2itc)^{-1} = (tc)^{-1}\sin(tc)$. В этом важном для практики случае $\psi_\xi(t)$ является действительной функцией.

Пример 11.6. Рассмотрим $(0,1)$-нормальную случайную величину, или случайную величину ξ, распределенную по стандартному нормальному закону. Для непрерывной случайной величины, плотность распределения которой $f_\xi(x) = (2\pi)^{-1/2}\exp\{-x^2/2\}$, $-\infty < x < +\infty$, характеристическая функция $\psi_\xi(t) = (2\pi)^{-1/2}\int_{-\infty}^{+\infty}\exp\{itx\}\,\exp\{-x^2/2\}\,dx$. Дифференцируя это равенство по t и, затем, интегрируя по частям, легко найдем, что

$$d\psi_\xi(t)/dt = (2\pi)^{-1/2}\int_{-\infty}^{+\infty} ix\exp\{itx\}\exp\{-x^2/2\}\,dx =$$

$$= -(2\pi)^{-1/2}\int_{-\infty}^{+\infty} i\exp\{itx\}\,d\exp\{-x^2/2\} =$$

$$= -(2\pi)^{-1/2} i\exp\{itx\}\exp\{-x^2/2\}\Big|_{-\infty}^{+\infty} -$$

$$-\,(2\pi)^{-1/2}\int_{-\infty}^{+\infty} t\exp\{itx\}\exp\{-x^2/2\}\,dx = -t\psi_\xi(t).$$

Отсюда получим дифференциальное уравнение $d\ln\psi_\xi(t)/dt = -t$ с начальным условием $\psi_\xi(0) = 1$. Решая это дифференциальное уравнение, легко выводим, что $\psi_\xi(t) = \exp\{-t^2/2\}$. Так как распределение $(0,1)$-нормальной случайной величины симметрично относительно оси ординат, то, естественно, $\psi_\xi(t)$ является действительной функцией.

Если теперь η является (a,σ)-нормальной случайной величиной, то случайная величина $\xi = (\eta - a)\,\sigma^{-1}$ является $(0,1)$-нормальной. Следовательно, $\eta = \sigma\xi + a$, и согласно первому свойству характеристической функции имеем $\psi_\xi(t) = \exp\{ita\}\exp\{-(\sigma t)^2/2\} = \exp\{iat - 2^{-1}\sigma^2 t^2\}$. Эта формула и свойства характеристических функций позволяют легко установить, что сумма $\xi = \xi_1 + \xi_2$ двух независимых нормально распределенных случайных величин ξ_1 и ξ_2 также распределена по нормальному закону. В самом деле, пусть $\mathsf{M}\xi_1 = a_1$, $\mathsf{D}\xi_1 = \sigma_1^2$ и $\mathsf{M}\xi_2 = a_2$, $\mathsf{D}\xi_2 = \sigma_2^2$. Тогда, применяя второе свойство характеристических функций, легко получим, что

$$\begin{aligned}
\psi_{\xi_1}(t) &= \exp\{ia_1 t - 2^{-1}\sigma_1^2 t^2\},\\
\psi_{\xi_2}(t) &= \exp\{ia_2 t - 2^{-1}\sigma_2^2 t^2\},\\
\psi_{\xi}(t) &= \exp\{ia_1 t - 2^{-1}\sigma_1^2 t^2\}\exp\{ia_2 t - 2^{-1}\sigma_2^2 t^2\} =\\
&= \exp\{i(a_1 + a_2)\,t - 2^{-1}(\sigma_1^2 + \sigma_2^2)\,t^2\}.
\end{aligned}$$

Отсюда, в силу четвертого свойства характеристических функций, выводим, что ξ будет гауссовской случайной величиной с математическим ожиданием $\mathsf{M}\xi_1 = a_1 + a_2$ и дисперсией $\mathsf{D}\xi = \sigma_1^2 + \sigma_2^2$.

Пример 11.7. Пусть ξ распределена по экспоненциальному закону с параметром $\lambda > 0$, т. е. $f_\xi(x) = \lambda\exp\{-\lambda x\}$ при $x \geqslant 0$ и $f_\xi(x) = 0$ при $x < 0$. Тогда для ξ характеристическая функция $\psi_\xi(t) = \int_0^{+\infty}\lambda\exp\{itx\}\exp\{-\lambda x\}\,dx = \lambda\int_0^{+\infty}\exp\{(it - \lambda)\,x\}\,dx = \lambda(\lambda - it)^{-1}$. Отсюда, используя свойство $\psi_\xi^{(k)}(0) = i^k\mathsf{M}\xi^k$, другим приемом находим формулу для начального момента $\alpha_k\xi = \mathsf{M}\xi^k = i^{-k}(d^k(\lambda(\lambda - it)^{-1})/dt^k)\big|_{t=0} = k!\,\lambda^{-k}$.

Пример 11.8. Найти характеристическую функцию случайной величины χ_r^2, которая распределена по закону хи-квадрат с r степенями свободы.

Решение. Напомним, что случайная величина $\chi_r^2 = \xi_1^2 + \xi_2^2 + \ldots + \xi_r^2$, где $\xi_1^2, \xi_2^2, \ldots, \xi_r^2$ являются независимыми случайными величинами с одной и той же плотностью распределения $f(x) = (2\pi)^{-1/2}x^{-1/2}e^{-x/2}$ при $x > 0$ и $f(x) = 0$ при $x \leqslant 0$. Характеристическая функция $\psi_\xi(t)$, которая соответствует плотности $f(x)$, вычисляется следующим образом:

$$\begin{aligned}
\psi(t) = (2\pi)^{-1/2}\int_0^{+\infty}\exp\{itx\}x^{-1/2}\exp\{-x/2\}\,dx &=\\
= (2\pi)^{-1/2}\int_0^{+\infty}x^{-1/2}\exp\{-x(2^{-1} - it)\}\,dx &=\\
= (\pi(1 - 2it))^{-1/2}\int_0^{+\infty}y^{-1/2}\exp\{-y\}\,dy &= (1 - 2it)^{-1/2},
\end{aligned}$$

где $x = y(2^{-1} - it)^{-1}$, $dx = (2^{-1} - it)^{-1}dy$, $x^{-1/2} = y^{-1/2}(2^{-1} - it)^{1/2}$, $\int_0^{+\infty}y^{-1}\exp\{-y\}\,dy = \Gamma(1/2) = \pi^{1/2}$. В силу второго свойства характеристических функций случайная величина χ_r^2 будет иметь характеристическую функцию вида $\psi_{\chi_r^2}(t) = (1 - 2it)^{-r/2}$.

Рассмотрим теперь применение характеристических функций при доказательстве центральных предельных теорем. Существует большое количество вариантов такого рода теорем. Здесь приведем две теоремы, которые наиболее часто применяются в теории вероятностей, математической статистике и в различных приложениях.

2.4. Центральные предельные теоремы. Нередко приходится иметь дело со случайной величиной, которая получается как сумма весьма большого числа малых по абсолютной величине и не зависимых между собой случайных величин, законы распределения которых неизвестны. Такой случайной величиной будет, например, ошибка при измерениях, являющаяся суммой большого числа малых по величине и не зависимых между собой ошибок, порождаемых различными факторами:

1) ошибки, вызванные состоянием измерительного прибора, которое может изменяться под влиянием механических, температурных или атмосферных причин;

2) ошибки наблюдателя, вносимые особенностями его зрения, слуха и т. д. Наблюдатель имеет дело лишь с суммарным действием отдельных и сравнительно малых факторов, и его интересует закон распределения «суммарной ошибки».

Приведем еще такой пример. Пусть производится стрельба по некоторой плоской мишени. Отклонение точки попадания от центра мишени можно объяснить

1) ошибкой наводки;
2) ошибкой в определении дальности до цели;
3) непреднамеренной вибрацией орудия и всевозможных установок при стрельбе;
4) ошибкой при изготовлении снаряда;
5) атмосферными условиями и т. д.

В приведенных примерах суммарная ошибка ξ равна простой сумме элементарных ошибок ξ_1, $\xi_2, \ldots$, ξ_n, т. е. $\xi = \xi_1 + \xi_2 + \ldots + \xi_n$. Ответ на важный вопрос о поведении интегральной функции распределения суммы случайных величин дают различные центральные предельные теоремы.

Теорема 11.9 (центральная предельная теорема Линдеберга–Леви). *Если* $\xi_1, \xi_2, \ldots$ — *независимые случайные величины, имеющие один и тот же закон распределения с конечными математическим ожиданием* a *и дисперсией* $\sigma^2 > 0$, *то при неограниченном увеличении* n *закон распределения центрированной и нормированной суммы* $\gamma_n = \left(\sum_{k=1}^n \xi_k - an\right)/\sigma n^{1/2}$ *неограниченно приближается* *к стандартному нормальному закону, т.е.*

$$\lim_{n\to\infty} \mathsf{P}(\{\omega\colon \Big(\sum_{k=1}^{n} \xi_k - an\Big)/\sigma n^{1/2} < x\}) = (2\pi)^{-1/2}\int\limits_{-\infty}^{x} \exp\{-t^2/2\}\,dt, \quad (11.4)$$

при этом сходимость равномерна по x, $-\infty < x < +\infty$.

Доказательство. Обозначим через $\psi_{\theta_k}(t)$ характеристическую функцию случайной величины $\theta_k = (\xi_k - a)/\sigma$. Легко вычислить $\mathsf{M}\theta_k = 0$ и $\mathsf{D}\theta_k = 1$. Отсюда, используя равенство (11.3), легко найдем в окрестности $t = 0$ разложение для характеристической функции $\psi_{\theta_k}(t) = 1 + it\,\mathsf{M}\theta_k + (it)^2\mathsf{M}\theta_k^2(2!)^{-1} + o(t^2) = 1 - 2^{-1}t^2 + r(t)\,t^2$, где величина $r(t)$ в общем случае является комплексной и $\lim\limits_{t\to 0} r(t) = 0$. Так как $\gamma_n = n^{-1/2}\sum_{k=1}^{n}\theta_k$, то, в силу первых двух свойств характеристических функций, при любом фиксированном t имеет место равенство $\psi_{\gamma_n}(t) = [1 - t^2 2^{-1}n^{-1} + t^2 n^{-1} r(tn^{-1/2})]^n = (1 + z_n n^{-1})^n$, где $z_n = -t^2 2^{-1} + t^2 r(tn^{-1/2})$. Непосредственно из равенства $\lim\limits_{n\to\infty} z_n = -t^2/2$ и из сходимости последовательности $\{(1 + z_n/n)^n;\, n \geqslant 1\}$, вообще говоря, комплексных чисел следует, что предел $\lim\limits_{n\to\infty}\psi_{\gamma_n}(t) = \lim\limits_{n\to\infty}(1 + z_n/n)^n = \exp\{-t^2/2\}$ при любом фиксированном t. Значит, последовательность $\{\psi_{\gamma_n}(t);\, n \geqslant 1\}$ из характеристических функций случайных величин γ_n, $n \geqslant 1$, при $n \to \infty$ сходится к характеристической функции $\psi_{\gamma_n}(t) = \exp\{-t^2/2\}$ случайной величины γ, распределенной по стандартному гауссовскому закону. Отсюда, используя свойство X.6 характеристических функций, следует предельное равенство (11.4). Более того, интегральная функция распределения $F_\gamma(x)$ случайной величины γ является непрерывной и строго возрастающей функцией.

Наконец, для завершения доказательства теоремы Линдеберга–Леви установим, что последовательность интегральных функций $F_{\gamma_n}(x)$, $n \geqslant 1$, при $n \to \infty$ сходится равномерно к интегральной функции $F_\gamma(x)$ по переменной $x \in (-\infty, +\infty)$. Очевидно, что при $\varepsilon > 5/2$, $n \geqslant 1$ и любых x выполняются неравенства $|F_{\gamma_n}(x) - F_\gamma(x)| \leqslant |F_{\gamma_n}(x)| + |F_\gamma(x)| \leqslant 2 < \varepsilon$. Поэтому в дальнейшем указанную равномерную сходимость будем доказывать для любого фиксированного ε из промежутка $0 < \varepsilon \leqslant 5/2$. С этой целью разобьем всю действительную прямую на непересекающиеся промежутки $\Delta_0 = (x_0, x_1)$, $\Delta_1 = [x_1, x_2), \ldots, \Delta_k = [x_k, x_{k+1})$. Здесь $x_0 = -\infty$, $x_{k+1} < +\infty$ и величина k равна такому наибольшему целому числу $\langle 5/\varepsilon\rangle$, что $\langle 5/\varepsilon\rangle < 5/\varepsilon$. При этом числа x_1, $x_2, \ldots, x_k$ в силу непрерывности, строгой монотонности и предельных свойств $\lim\limits_{x\to-\infty} F_\gamma(x) = 0$, $\lim\limits_{x\to+\infty} F_\gamma(x) = 1$ функции $F_\gamma(x)$ однозначно определяются из соотношений $F_\gamma(x_m) = F_\gamma(x_{m-1}) + \varepsilon/5$, $m = \overline{1, k}$.

Отсюда, с учетом $k = \langle 5/\varepsilon \rangle$, легко проверить, что $F_\gamma(x_1) = \varepsilon/5$, $1 > F_\gamma(x_k) = k\varepsilon/5 = \langle 5/\varepsilon \rangle \varepsilon/5 \geqslant 1 - \varepsilon/5$.

Следовательно, на каждом построенном промежутке Δ_0, $\Delta_1, \ldots$, Δ_k приращение функции $F_\gamma(x)$ не превосходит $\varepsilon/5$. Если теперь x — произвольное действительное число, то $x \in \Delta_j$ при некотором $j \in \{0, 1, \ldots, k\}$. Далее, имеем легко проверяемое соотношение

$$F_{\gamma_n}(x) - F_\gamma(x) = (F_{\gamma_n}(x) - F_{\gamma_n}(x_j)) + \\ + (F_{\gamma_n}(x_j) - F_\gamma(x_j)) + (F_\gamma(x_j) - F_\gamma(x)). \quad (11.5)$$

В силу монотонности $F_{\gamma_n}(x)$ и $F_\gamma(x)$ выводим оценки

$$|F_{\gamma_n}(x) - F_{\gamma_n}(x_j)| \leqslant |F_{\gamma_n}(x_{j+1}) - F_{\gamma_n}(x_j)| \leqslant \\ \leqslant |F_{\gamma_n}(x_{j+1}) - F_\gamma(x_{j+1})| + |F_\gamma(x_j) - F_{\gamma_n}(x_j)| + |F_\gamma(x_{j+1}) - F_\gamma(x_j)|, \\ |F_\gamma(x_j) - F_\gamma(x)| = |F_\gamma(x) - F_\gamma(x_j)| \leqslant |F_\gamma(x_{j+1}) - F_\gamma(x_j)|.$$

Используя соотношение (11.5) и полученные оценки, найдем, что

$$|F_{\gamma_n}(x) - F_\gamma(x)| + |F_{\gamma_n}(x) - F_{\gamma_n}(x_j)| + |F_{\gamma_n}(x_j) - F_\gamma(x_j)| + \\ + |F_\gamma(x_j) - F_\gamma(x)| \leqslant |F_{\gamma_n}(x_{j+1}) - F_\gamma(x_{j+1})| + \\ + 2|F_{\gamma_n}(x_j) - F_\gamma(x_j)| + 2|F_\gamma(x_{j+1}) - F_{.\gamma}(x_j)|. \quad (11.6)$$

Так как множество $\{x_0, x_1, \ldots, x_{k+1}\}$ содержит конечное число элементов и $\lim\limits_{n\to\infty} F_{\gamma_n}(x_j) = F_\gamma(x_j)$, то найдется такое $N(\varepsilon/5)$, что при $n \geqslant$ $\geqslant N(\varepsilon/5)$ будет выполняться неравенство $|F_{\gamma_n}(x_j) - F_\gamma(x_j)| \leqslant \varepsilon/5$ для всех $j = 0, 1, \ldots, k+1$. Принимая во внимание это и выбор промежутков Δ_0, $\Delta_1, \ldots$, Δ_k, из (11.6) получим неравенство $|F_{\gamma_n}(x) - F_\gamma(x)| \leqslant \varepsilon$ для всех $n \geqslant N(\varepsilon/5)$ и любых x. Итак, центральная предельная теорема Линдеберга–Леви доказана.

Частным случаем этой теоремы является предельная интегральная теорема Муавра–Лапласа в схеме Бернулли с n независимыми испытаниями эксперимента E_0 и с вероятностью $0 < p < 1$ появления некоторого события A в каждом испытании. Пусть случайная величина ξ_k определяет число появлений события A в k-м испытании. На основании решения задачи о приближенных измерениях вместо случайных величин ξ_1, $\xi_2, \ldots$ можно рассматривать бесконечную последовательность независимых и одинаково распределенных бернуллиевских случайных величин ζ_1, $\zeta_2, \ldots$ на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P}(\cdot))$. При этом абсолютная частота наступления события A за n опытов $\mu(A, n) = \sum_{k=1}^n \zeta_k$, а величины ζ_1, $\zeta_2, \ldots$ имеют одно и то же распределение $\mathsf{P}(\{\omega\colon \zeta_k(\omega) = 0\}) = 1 - p = q$, $\mathsf{P}(\{\omega\colon \zeta_k(\omega) = 1\}) = p$, для которого $\mathsf{M}\zeta_k = p$, $\mathsf{D}\zeta_k = pq$. Используя теперь центральную предельную теорему Линдеберга–Леви для случайных величин $\zeta_1, \zeta_2, \ldots$ и учитывая равенство $\mu(A, n) = \sum_{k=1}^{n} \zeta_k$, найдем, что

$$\lim_{n\to\infty} \mathsf{P}(\{\omega\colon (\mu(A,n) - np)(npq)^{-1/2} < x\}) = (2\pi)^{-1/2} \int\limits_{-\infty}^{x} \exp\{-t^2/2\}\,dt,$$

причем стремление к пределу равномерно по $x \in (-\infty, +\infty)$. Если принять во внимание справедливые для схемы n независимых испытаний равенства

$$x_m = (m - np)(npq)^{-1/2},$$

$$\begin{aligned} \mathsf{P}\Big(\Big\{\omega\colon a \leqslant \Big(\sum_{k=1}^{n} \zeta_k - np\Big)(npq)^{-1/2} < b\Big\}\Big) &= \\ = \mathsf{P}(\{\omega\colon a \leqslant (\mu(A,n) - np)(npq)^{-1/2} < b\}) &= \\ = \mathsf{P}(\{\omega\colon np + a(npq)^{1/2} \leqslant \mu(A,n) < np + b(npq)^{1/2}\}) &= \\ = \mathsf{P}\Big(\bigcup_{\{m\colon a\leqslant x_m<b\}} \boldsymbol{B}_m\Big) = \sum_{\{m\colon\ a\leqslant x_m< b\}} C_n^m p^m q^{n-m}&, \end{aligned}$$

то предельная интегральная теорема Муавра–Лапласа может быть сформулирована в следующей эквивалентной форме: *равномерно относительно a и b $(-\infty \leqslant a < b \leqslant +\infty)$ имеет место предельное соотношение вида*

$$\lim_{n\to\infty} \sum_{\{m\colon a\leqslant x_m<b\}} C_n^m p^m q^{n-m} = (2\pi)^{-1/2} \int\limits_{a}^{b} \exp\{-t^2/2\}\,dt.$$

Напомним, что интегральная теорема Муавра–Лапласа была доказана в гл. 4 путем преобразований вероятности $\mathsf{P}(\sum_{\{m\colon a\leqslant x_m<b\}} \boldsymbol{B}_m) = \sum_{\{m\colon a\leqslant x_m<b\}} C_n^m p^m q^{n-m}$ с использованием формулы Стирлинга.

Одно из первых утверждений типа центральных предельных теорем, в которых не предполагается обременительное ограничение об одинаковом распределении случайных величин, является теорема Ляпунова. Приведем без доказательства эту теорему, так как метод ее доказательства [6] такой же, как и теоремы Линдеберга–Леви.

Теорема 11.10 (центральная предельная теорема Ляпунова). *Пусть* $\xi_1, \xi_2, \ldots$ — *независимые случайные величины, для которых* $\mathsf{M}\xi_k$, $\mathsf{D}\xi_k$, $\mathsf{M}|\xi_k - \mathsf{M}\xi_k|^3$ *конечны при* $k \geqslant 1$ *и* $a_n = \sum_{k=1}^{n} \mathsf{M}\xi_k$, $b_n = \big(\sum_{k=1}^{n} \mathsf{D}\xi_k\big)^{1/2}$, $c_n = \big(\sum_{k=1}^{n} \mathsf{M}|\xi_k - \mathsf{M}\xi_k|^3\big)^{1/3}$, $\lim\limits_{n\to\infty} c_n/b_n = 0$.*Тогда последовательность* $\big\{\gamma_n = \big(\sum_{k=1}^{n} \xi_k - a_n\big)/b_n;\ n \geqslant 1\big\}$ *сходится по распределению к стандартному нормальному закону равномерно на действительной прямой.*

На практике, когда n невелико, центральную предельную теорему Линдеберга–Леви применять нецелесообразно.

Если же n достаточно большое, обычно $n \geqslant 10$, то случайная величина $\gamma_n = \left(\sum_{k=1}^n \xi_k - an\right)/\sigma n^{1/2}$ имеет распределение, близкое к нормальному. Этот замечательный факт можно использовать для приближенного вычисления $\mathsf{P}\left(c \leqslant \sum_{k=1}^n \xi_k < d\right)$, где $-\infty < c < d < +\infty$. Действительно, используя функцию Лапласа $\Phi(x)$, легко получим, что вероятность

$$\mathsf{P}\Big(c \leqslant \sum_{k=1}^{n} \xi_k < d\Big) = \mathsf{P}\Big(c - na \leqslant \sum_{k=1}^{n} \xi_k - na < d - na\Big) =$$
$$= \mathsf{P}\Big((c-na)/\sigma n^{1/2} \leqslant \Big(\sum_{k=1}^{n} \xi_k - na\Big)/\sigma n^{1/2} < (d-na)/\sigma n^{1/2}\Big) \approx$$
$$\approx 2^{-1}\big[\Phi((d-na)/\sigma n^{1/2}) - \Phi((c-na)/\sigma n^{1/2})\big].$$

Пример 11.9. Найти приближенное аналитическое выражение для плотности вероятности $f_\xi(x)$ случайной величины $\xi = \sum_{k=1}^{24} \xi_k$ и вычислить вероятность попадания этой случайной величины на участок $[6,\ 14)$, если $\xi_1,\ \xi_2, \ldots,\ \xi_{24}$ — независимые случайные величины, имеющие один и тот же равномерный закон распределения на отрезке $[0, 1]$, и $n = 24$.

Решение. Подсчитаем математическое ожидание и дисперсию каждой из случайных величин ξ_k. Имеем $\mathsf{M}\xi_k = (0+1)/2 = 1/2$; $\mathsf{D}\xi_k = (1-0)^2/12 = 1/12$. Математическое ожидание и среднее квадратическое отклонение для случайной величины ξ соответственно равны $\mathsf{M}\xi = n\,\mathsf{M}\xi_k = 12$ и $\sigma(\xi) = \sqrt{2}$. Используя центральную предельную теорему Линдеберга–Леви, можно сказать, что распределение случайной величины ξ близко к нормальному с параметрами $a = 12$ и $\sigma = \sqrt{2}$. Поэтому плотность распределения $f_\xi(x) \approx 2^{-1}\pi^{-1/2}\exp\{(x-12)^2/4\}$, а искомую вероятность можно определить приближенно: $\mathsf{P}(6 \leqslant \xi < 14) \approx 2^{-1}[\Phi((14-12)\,2^{-1/2}) - \Phi((6-12)\,2^{-1/2})] = 2^{-1}[\Phi(\sqrt{2}\,) + \Phi(3\sqrt{2}\,)] \approx 0,92$.

Пример 11.10. Используя центральную предельную теорему Линдеберга–Леви, найти предел вида $\lim\limits_{n\to\infty} \exp\{-n\}\sum_{k=0}^{n} n^k/k!$.

Решение. Пусть при любом $k = 1, 2, \ldots$ случайная величина ξ_k является пуассоновской с параметром $\lambda = 1$, а последовательность $\{\xi_k;\ k \geqslant 1\}$ составлена из независимых случайных величин. Отсюда $\mathsf{M}\xi_k = 1$, $\mathsf{D}\xi_k = 1$ и по центральной предельной теореме Линдеберга–Леви, при $n \to \infty$

$$\lim_{n\to\infty} \mathsf{P}\Big(\Big\{\omega\colon \Big(\sum_{k=1}^{n} \xi_k - n\Big)\, n^{-1/2} < x\Big\}\Big) = (2\pi)^{-1/2} \int\limits_{-\infty}^{x} \exp\{-t^2/2\}\, dt.$$

Если $x = 0$, то

$$\lim_{n\to\infty} \mathsf{P}\Big(\Big\{\omega\colon \Big(\sum_{k=1}^{n} \xi_k - n\Big)\, n^{-1/2} < 0\Big\}\Big) = (2\pi)^{-1/2} \int\limits_{-\infty}^{x} \exp\{-t^2/2\}\, dt = 0{,}5.$$

Так как величина $\xi_1 + \xi_2 + \ldots + \xi_n$ распределена по закону Пуассона с параметром $\lambda = n$, то легко выводим:

$$\exp\{-n\} \sum_{k=0}^{n} n^k/k! - \exp\{-n\}\, n^n/n! = \mathsf{P}\Big(\Big\{\omega\colon \sum_{k=1}^{n} \xi_k < n\Big\}\Big) = \\ = \mathsf{P}\Big(\Big\{\omega\colon \sum_{k=1}^{n} \xi_k - n < 0\Big\}\Big) = \mathsf{P}\Big(\Big\{\omega\colon \Big(\sum_{k=1}^{n} \xi_k - n\Big)\, n^{-1/2} < 0\Big\}\Big).$$

В силу равенств $\lim\limits_{n\to\infty} \mathsf{P}\big(\{\omega\colon \big(\sum_{k=1}^{n} \xi_k - n\big)\, n^{-1/2} < 0\}\big) = 0{,}5$, $\lim\limits_{n\to\infty} n^n (n!)^{-1} \exp\{-n\} = 0$ получаем предельное соотношение: $\lim\limits_{n\to\infty} \exp\{-n\} \sum_{k=0}^{n} n^k/k! = 0{,}5$.

Краткий обзор

Вычисление биномиальной вероятности $P_n(m)$ по формуле $P_n(m) = C_n^m p^m q^{n-m}$ при больших значениях n и m представляет значительные трудности. Поэтому в гл. 4 приводятся локальные и интегральные предельные теоремы Муавра–Лапласа и Пуассона, которые позволяют приближенно и с достаточной степенью точности определять биномиальные вероятности при больших значениях n и m. Вероятности некоторых событий в схеме независимых испытаний Бернулли, вычисленные по соответствующим приближенным формулам, почти не отличаются от единицы. В силу этого можно предсказывать такого рода события в схеме независимых испытаний Бернулли почти с полной определенностью. В общем случае теория вероятностей дает возможность не только осуществлять научные прогнозы в области случайных явлений, но и оценивать точность этих прогнозов. Математическое обоснование научных прогнозов случайных явлений основано на понятии различных видов сходимостей в теории вероятностей и на утверждениях различных предельных теорем. Здесь уделяется большое внимание определению следующих видов сходимости последовательности случайных величин к некоторой случайной величине: 1) по вероятности; 2) с вероятностью единица; 3) в среднеквадратическом; 4) по распределению. Проведена классификация предельных теорем. Рассмотрены простейшие формы предельных теорем, доказательство которых основано на известном неравенстве Чебышева и с использованием аппарата характеристических функций. При этом определение характеристических функций и изучение их свойств требует рассмотрения комплекснозначной случайной величины.

Контрольные вопросы и упражнения

1. Показать, что понятия сходимости по вероятности и с вероятностью единица последовательности случайных величин является прямым обобщением понятия сходимости обычного предела числовой последовательности из курса математического анализа.
2. Дать геометрическую интерпретацию понятия сходимости по вероятности последовательности случайных величин.
3. Дать геометрическую интерпретацию понятия сходимости с вероятностью единица последовательности случайных величин.
4. Доказать, что неравенство Чебышева дает неулучшаемую оценку для вероятности события $\{\omega\colon |\xi(\omega) - \mathsf{M}\xi| \geqslant \varepsilon\}$.
5. Доказать, что среднее арифметическое достаточно большого числа наблюдаемых значений случайной величины сходится по вероятности к ее математическому ожиданию.
6. Дать интерпретацию предельной теоремы Бернулли в связи с понятием статистической устойчивости случайного эксперимента.
7. В чем заключается основное значение центральных предельных теорем Линдеберга–Леви и А. М. Ляпунова?
8. В чем заключается отличие центральных предельных теорем Линдеберга–Леви и А.М. Ляпунова?
9. Используя аппарат характеристических функций, доказать центральную предельную теорему А. М. Ляпунова.
10. Вероятность рождения мальчика равна 0,5. Применяя центральную предельную теорему Линдеберга–Леви, найти вероятность того, что среди 900 новорожденных детей будет от 435 до 465 мальчиков.
11. Кинотеатр вмещает 768 зрителей. Найти вероятность того, что среди такого количества зрителей летом родились от 180 до 216 детей.
12. Предположим, что случайная величина имеет плотность распределения вероятности $f(x) = 2^{-1}\exp\{-|x|\}$. Используя неравенство Чебышева, оценить вероятность того, что эта случайная величина примет значения по модулю не меньше трех. Найти ошибку этой оценки.
13. Построить последовательность $F_{\xi_1}(x)$, $F_{\xi_2}(x), \ldots$ интегральных функций распределения случайных величин ξ_1, $\xi_2, \ldots$, которая сходится в каждой точке x к некоторой функции $G(x)$ и не сходится по распределению.
14. Предположим, что случайные величины ξ_1, $\xi_2, \ldots$ являются попарно независимыми случайными величинами, имеют конечные математические ожидания и равномерно ограниченные дисперсии. Используя второе неравенство Чебышева, показать, что по-

следовательность $\left\{n^{-1}\sum_{i=1}^{n}(\xi_i - \mathsf{M}\xi_i);\ n = 1, 2, \ldots\right\}$ сходится по вероятности к нулю при $n \to \infty$.

15. Пусть ξ_1 и ξ_2 являются независимыми случайными величинами, которые распределены по закону Пуассона с параметрами λ_1 и λ_2. Используя свойства характеристических функций, доказать, что случайная величина вида $\xi_1 + \xi_2$ распределена по закону Пуассона с параметром $\lambda_1 + \lambda_2$.
16. Для определения расстояния s между городами было проделано 100 случайных измерений ξ_1, $\xi_2, \ldots$, ξ_{100}. Пусть случайные величины ξ_1, $\xi_2, \ldots$, ξ_{100} независимы и одинаково распределены с математическим ожиданием s и дисперсией, равной единице. Вычислить вероятность того, что среднее арифметическое измеренное расстояние $(\xi_1 + \xi_2 + \ldots + \xi_{100})/100$ будет отличаться от истинного расстояния s не более, чем на величину 0,5.

Заключение

Хотя традиционный детерминированный взгляд на эксперимент часто оказывается полезным, материал этого учебника убедительно показывает, что природа реального мира в большей степени является стохастической. Поэтому вероятностный подход к построению математических моделей процессов, как правило, более реалистичен по сравнению с детерминированным подходом. За последнее время использование вероятностных моделей в самых разнообразных областях науки, техники, производства, экономики, медицины и социологии можно обосновать следующими обстоятельствами.

Во-первых, современные измерительные устройства позволяют с достаточной степенью точности регистрировать реализации случайных отклонений различных количественных характеристик конкретного эксперимента от их средних значений. Отказ от изучения роли такого рода отклонений приводит к авариям атомных реакторов, разрушению мостов, плотин, промышленных сооружений и гражданских зданий, поломкам самолетов и кораблей, транспортным катастрофам в больших городах, выпуску некачественной и ненадежной продукции, экономическим и природным катаклизмам и т. п.

Во-вторых, современная микропроцессорная техника и новые информационные технологии позволяют находить, хранить и обрабатывать большие массивы экспериментальной информации о реальных объектах.

Используя все это, вероятностные модели дают более ясную картину причинных связей между реальными процессами, позволяют найти закономерности природы там, где детерминированный подход оказывается бессильным. Такую ситуацию можно наблюдать в теории ошибок разного рода измерений, в молекулярной и статистической физике, в биологии, в медицине, в рыночной экономике, в телефонии и процессах обслуживания, в процессах адаптивного управления и принятия решений, при управлении конфликтными транспортными потоками на магистралях крупных городов и т. д.

С целью подтверждения такого рода высказываний рассмотрим пример построения и изучения математической модели движения транспорта в плохих погодных условиях на магистрали с относительно небольшой плотностью быстрых машин и малым средним временем обгона. Обратимся снова к постановке и решению задачи о групповом движении машин по магистрали из гл. 10. Следует заметить, что для транспортного потока важно изучить не только вероятностные свойства последовательности из моментов пересечения машинами так называемой виртуальной стоп-линии, но также необходимо определить свойства случайного расположения автомобилей на автомагистрали. Другими словами, математическая модель движения транспортных потоков должна учитывать как пространственный, так и временной процесс. Если интенсивность λ поступления по закону Пуассона быстрых машин в транспортную пачку существенно меньше интенсивностей μ_1 и μ_2 обгона быстрыми машинами медленной машины, то образуются транспортные пачки относительно небольшого размера. В силу этого будем предполагать, что случайное число $\chi(\omega;t)$ всех типов машин в транспортной пачке в момент времени $t \geqslant 0$ принимает значения из конечного множества $\{1, 2, \ldots, N\}$. Только этим предположением данная задача отличается от задачи о групповом движении машин по магистрали из гл. 10, где случайная величина $\chi(\omega;t) \in \{1, 2, \ldots\}$. Поэтому необходимо дополнительно определить случайный механизм обгона быстрыми машинами медленную машину в случае, когда транспортная пачка достигла размера в N машин.

Итак, будем считать, что каждая быстрая машина, заставшая в автоколонне максимально возможное число N машин, присоединится к автоколонне. В этом случае машина, движущаяся непосредственно за медленной машиной, мгновенно совершает обгон. Это позволяет не увеличивать максимальный размер пачки до величины $N+1$. При этом не происходит потеря машин в транспортном потоке. Напомним (см. гл. 10), что случайные величины $\xi(\omega;\, t,\, \Delta t)$ и $\eta(\omega;\, t,\, \Delta t)$ определяют число быстрых машин, которые на промежутке времени $[t,\, t+\Delta t)$ поступают в транспортную пачку и соответственно ее покидают. Здесь $0 \leqslant \eta(\omega;\, t,\, \Delta t) \leqslant \chi(\omega;t) + \xi(\omega;\, t,\, \Delta t) - 1$. Из предположения механизма обгона в условиях максимального размера транспортной пачки для условных вероятностей событий, которые порождаются величиной $\eta(\omega;\, t,\, \Delta t)$, получаем равенства:

$$\begin{aligned}
&\mathsf{P}\big(\{\omega\colon \eta(\omega;t,\Delta t) = 0\} \mid \{\omega\colon \chi(\omega;t)\big) = \\
&\qquad\qquad\qquad = N, \xi(\omega;t,\Delta t) = 0\}) = 1 - \mu_2\Delta t + (\Delta t), \\
&\mathsf{P}\big(\{\omega\colon \eta(\omega;t,\Delta t) = 1\} \mid \{\omega\colon \chi(\omega;t) = \\
&\qquad\qquad\qquad = N, \xi(\omega;t,\Delta t) = 0\}\big) = \mu_2\Delta t - o(\Delta t), \\
&\mathsf{P}\big(\{\omega\colon \eta(\omega;t,\Delta t) = 1\} \mid \{\omega\colon \chi(\omega;t) = N, \xi(\omega;t,\Delta t) = 1\}\big) = 1,
\end{aligned} \tag{3.1}$$

Используя при $k = \overline{1, N}$ равенство в событиях $\{\omega\colon \chi(\omega;t+\Delta t) = = k\} = \cup_{s=1}^{N} \cup_{n=0}^{\infty} \{\omega\colon \chi(\omega;t) = s,\, \xi(\omega;\,\Delta t) = n,\, \eta(\omega;\,\Delta t) = s+n-k\}$, соотношения (10.8)–(10.10) при $\chi(\omega;t) \in \{1, 2, \ldots, N-1\}$, дополнительное равенство (3.1) и, наконец, методику из гл. 10, легко получим конечную систему дифференциальных уравнений для вероятностей вида

$$\mathsf{P}\big(\{\omega\colon \chi(\omega;t) = k\} = Q(t,k), \quad t \geqslant 0, \quad k = 1, 2, \ldots, N,$$

а именно:

$$
\begin{gathered}
dQ(t,1)/dt = -\lambda Q(t,1) + \mu_1 Q(t,2),\\
dQ(t,2)/dt = \lambda Q(t,1) - (\lambda+\mu_1)\,Q(t,2) + \mu_2 Q(t,3),\\
dQ(t,k)/\Delta t = \lambda Q(t,k-1) - (\lambda+\mu_2)\,Q(t,\ k) + \mu_2 Q(t,k+1),\\
3 \leqslant k \leqslant N-1,\\
dQ(t,N)/dt = \lambda Q(t,N-1) - \mu_2 Q(t,N).
\end{gathered} \tag{З.2}
$$

Тогда динамика распределения числа машин в транспортной пачке определяется решением системы линейных однородных дифференциальных уравнений (З.2), например, с начальными условиями $Q(0,i)=1$ при некотором фиксированном значении $i=1,2,\ldots,N$ и $Q(0,k)=0$ для всех $k\neq i$. Явное решение системы дифференциальных уравнений (З.2) может быть найдено с помощью различных методов, которые применяются в теории линейных однородных дифференциальных уравнений с постоянными коэффициентами. Известно, что в этом случае процесс решения системы (З.2) является очень громоздким. К счастью, нам потребуется здесь только свойства распределения числа машин в транспортной пачке при $t\to\infty$, т.е. некоторые свойства решений системы (З.2) при $t\to\infty$.

Рассмотрим, ради простоты, очень важный для практики частный случай, когда в транспортной пачке число машин не может быть больше двух, т.е. $N=2$. В этом случае $\mu_1=\mu_2$ и система линейных однородных дифференциальных уравнений (З.2) примет вид

$$
\begin{aligned}
dQ(t,1)/dt &= -\lambda Q(t,1) + \mu_1 Q(t,2),\\
dQ(t,2)/dt &= \lambda Q(t,1) - \mu_1 Q(t,2).
\end{aligned} \tag{З.3}
$$

Неизвестные функции $Q(t,1)$ и $Q(t,2)$, которые удовлетворяют системе дифференциальных уравнений (З.3), могут быть получены с помощью следующего простого способа. В самом деле, так как при любом значении $t\geqslant 0$ выполняется условие нормировки $Q(t,1)+Q(t,2)=1$, то вместо системы (З.3) можно рассмотреть одно неоднородное линейное дифференциальное уравнение первого порядка с постоянными коэффициентами:

$$
dQ(t,1)/dt = -(\lambda+\mu_1)\,Q(t,1) + \mu_1 \tag{З.4}
$$

с произвольным начальным условием $Q(0,1)=p$, где $0\leqslant p\leqslant 1$. Решая это простое уравнение, получим, что

$$
Q(t,1) = \mu_1(\lambda+\mu_1)^{-1} + (p - \mu_1(\lambda+\mu_1)^{-1})\exp\{-(\lambda+\mu_1)\,t\}.
$$

Так как $Q(t,2)=1-Q(t,1)$, то для функции $Q(t,2)$ имеем

$$
Q(t,2) = 1 - Q(t,1) = \lambda(\lambda+\mu_1)^{-1} + (\mu_1(\lambda+\mu_1)^{-1} - p)\exp\{-(\lambda+\mu_1)\,t\}.
$$

Отсюда легко вычислить пределы

$$\lim_{t\to\infty} Q(t,1) = Q(1) = \mu_1(\lambda+\mu_1)^{-1},$$
$$\lim_{t\to\infty} Q(t,2) = Q(2) = \lambda(\lambda+\mu_1)^{-1}.$$

Предельное распределение $\{Q(k);\ k=1,2\}$ числа $\chi(\omega)$ всех типов машин в транспортной пачке характеризует так называемый установившийся или стационарный режим движения минимального размера автоколонн по магистрали. Эргодическое распределение $\{Q(k);\ k=1,2\}$ не зависит от начальных условий $Q(0,1)$, $Q(0,2)$ и может быть также получено путем решения простого уравнения:

$$-(\lambda+\mu_1)\,Q(1)+\mu_1=0.$$

Это уравнение получается с помощью предельного перехода при $t\to\infty$ в правой и левой части уравнения (3.4) с учетом очевидного равенства $\lim\limits_{t\to\infty}(dQ(t,1)/dt)=0$.

В реальном транспортном потоке плотность медленных машин (среднее число медленных машин на единичном по длине участке автомагистрали) значительно меньше плотности быстрых машин. В то же время можно допустить, что движение медленных машин в стационарном режиме происходит независимым образом друг от друга и практически с постоянной скоростью. В связи с этим было принято, что транспортный поток из медленных машин будет пуассоновским с интенсивностью λ_0. Пусть максимальная длина участка дороги, на котором располагается среднее число автомобилей в транспортной пачке, много меньше среднего расстояния между соседними автомобилями с медленным движением. Среднее число автомобилей в транспортной пачке для стационарного режима равно

$$\mathsf{M}\chi(\omega) = \mu_1(\lambda+\mu_1)^{-1} + 2\lambda(\lambda+\mu_1)^{-1} = (2\lambda+\mu_1)(\lambda+\mu_1)^{-1}.$$

Тогда можно считать, что все машины каждой движущейся автоколонны в стационарном режиме пересекают некоторую поперечную линию автомагистрали практически одновременно. Поэтому транспортный поток движущихся в стационарном режиме всех типов автомобилей на магистрали с достаточно хорошим приближением будет потоком Гнеденко–Коваленко [7], в котором каждый вызывающий момент совпадает с моментом появления медленной машины. Сравнивая эти два потока, получаем, что в каждый вызывающий момент поступает одна заявка (машина) с вероятностью $p=\mu_1(\lambda+\mu_1)^{-1}$ или поступают две заявки (машины) с вероятностью $q=\lambda(\lambda+\mu_1)^{-1}$, $p+q=1$. Теперь можно дать простой смысл параметра $\lambda(\lambda+\mu_1)^{-1}$ для задачи о групповом движении машин по магистрали при $N = 2$. Этот параметр определяет для стационарного режима вероятность того, что в транспортной пачке будет быстрая машина. Очевидно, что эта вероятность увеличивается с увеличением интенсивности λ поступления быстрых машин в транспортную пачку и уменьшается с уменьшением среднего времени μ_1^{-1} обгона каждой быстрой машиной медленную машину.

Для потока Гнеденко–Коваленко обозначим через $\varkappa(\omega;\, t)$ случайное число всех типов машин, которые пересекают некоторую поперечную линию автомагистрали за промежуток $[0,\, t)$.

При любом $k = 0, 1, \ldots$ вероятность $\mathsf{P}(\{\omega\colon \varkappa(\omega;\, t) = k\})$ будем обозначать через функцию $P_k(t)$. Рассмотрим задачу отыскания одномерных распределений $P_k(t)$, $k = 0, 1, \ldots$, стационарного потока без последействия или неординарного пуассоновского потока $\{\varkappa(\omega;\, t)\colon t \geqslant 0\}$ с параметрами λ_0 и $p = \mu_1(\lambda + \mu_1)^{-1}$. Применяя методику, которая была подробно рассмотрена в гл. 10 при решении задачи о свободном движении машин по магистрали, в работе [7] для производящей функции $\Pi_{\varkappa(t)}(t, z) = \sum_{k=0}^{\infty} P_k(t)\, z^k$ случайной величины $\varkappa(t)$ была получена следующая формула:

$$\Pi_{\varkappa(t)}(t, z) = \exp\{\lambda t(pz + qz^2 - 1)\}.$$

Тогда при $|z| \leqslant 1$ найдем

$$\begin{aligned}\Pi_{\varkappa(t)}(t, z) &= \exp\{\lambda t(pz + qz^2 - 1)\} = \exp\{-\lambda t\} \exp\{\lambda t p z\} \exp\{\lambda t q z^2\} = \\ &= \exp\{-\lambda t\} \sum_{m=0}^{\infty} (\lambda t p z)^m (m!)^{-1} \sum_{n=0}^{\infty} (\lambda t q z^2)^n (n!)^{-1} = \\ &= \exp\{-\lambda t\} \sum_{k=0}^{\infty} z^k \sum_{i=0}^{[k/2]} C_{k-i}^{i} p^{k-2i} q^i (\lambda t)^{k-i} ((k-i)!)^{-1}.\end{aligned}$$

Отсюда имеем

$$P_k(t) = \exp\{-\lambda t\} \sum_{i=0}^{[k/2]} C_{k-i}^{i} p^{k-2i} q^i (\lambda t)^{k-i} ((k-i)!)^{-1}, \qquad k \geqslant 0.$$

Используя равенство $\Pi_{\varkappa(t)}(t, z) = \exp\{\lambda t(pz + qz^2 - 1)\}$, соотношение (10.1) и связь между начальными и центральными моментами вычислим математическое ожидание $\mathsf{M}\varkappa(t)$ и дисперсию $\mathsf{D}\varkappa(t)$:

$$\mathsf{M}\varkappa(t) = \big(d(\exp\{\lambda t(pz + qz^2 - 1)\})/dz\big)\big|_{z=1} = \lambda t\,(1 + q),$$

$$\mathsf{D}\varkappa(t) = \big(d^2(\exp\{\lambda t(pz + qz^2 - 1)\})/dz^2\big)\big|_{z=1} - (\lambda t\,(1 + q))^2 = \lambda t(3q + 1).$$

Значит, при увеличении параметров λ, t и q увеличивается как среднее значение, так и дисперсия случайной величины $\varkappa(t)$.

Аналогичным способом найдем начальные и центральные моменты $\alpha_3(\varkappa(t))$, $\alpha_4(\varkappa(t))$, $\beta_3(\varkappa(t))$, $\beta_4(\varkappa(t))$, коэффициент асимметрии $\mathsf{Ka}(\varkappa(t))$ и эксцесс $\mathsf{Э}(\varkappa(t))$:

$$\begin{aligned}
\alpha_3(\varkappa(t)) &= (\lambda t)^3(1+q)^3 + (\lambda t)^2(9q^2+12q+3) + \lambda t(1+7q),\\
\alpha_4(\varkappa(t)) &= (\lambda t)^4(1+q)^4 + (\lambda t)^3(18q^3+42q^2+30q+6) +\\
&\qquad\qquad + (\lambda t)^2(55q^2+50q+7) + (\lambda t)(15q+1),\\
\beta_3(\varkappa(t)) &= \lambda t(1+7q),\\
\beta_4(\varkappa(t)) &= (\lambda t)^2(27q^2+18q+3) + \lambda t(15q+1),\\
\mathsf{Ka}(\varkappa(t)) &= (1+7q)(\lambda t)^{-1/2}(1+3q)^{-3/2},\\
\mathsf{Э}(\varkappa(t)) &= (1+15q)(\lambda t)^{-1}(1+3q)^{-2}.
\end{aligned}$$

Предложенная здесь вероятностная модель движения машин по магистрали позволяет решать задачу [20] оптимального управления конфликтными транспортными потоками на пересечении магистралей. Этот и многие другие примеры наглядно показывают, что методы построения и изучения вероятностных моделей реальных процессов и явлений должны быть отнесены к числу основных общеобразовательных дисциплин, которые определяют современный профессиональный уровень выпускников вузов по самым различным специальностям. Так, на различных факультетах Нижегородского госуниверситета в рамках общего курса по теории вероятностей и математической статистики читается в течение одного семестра курс по вероятностному моделированию.

В настоящее время при подготовке общих курсов, учебников и задачников по теории вероятностного моделирования различными коллективами в учебных заведениях нашей страны и за рубежом существуют два противоположных подхода. Первый и наиболее распространенный подход [1, 9, 10, 15, 18, 23, 24], который называется теоретико-множественным, заключается в широком использовании теории вероятностей, теории меры и функционального анализа. При таком подходе, как правило, формулируются и решаются только такие проблемы и задачи, в которых имеется неоправданная математическая формализация, и почти отсутствуют способы построения вероятностных моделей. Студенты, хорошо усвоившие такой курс, как правило, совершенно беспомощны в формулировании и решении конкретных практических задач. Второй подход [8, 12] основан на интерпретации основных положений теории вероятностей с использованием большого числа, как правило, простых модельных экспериментов (бросание монет и игральных костей, игры в карты, выборки шаров из урн, схема независимых испытаний, бросание точек на отрезок, задачи лотерей и рулеток, задачи стрельб и т. д.). При втором подходе часто полагаются на интуицию, которая сильно зависит от исследователя и может быть неодинаковой. На этом строятся различные парадоксы в теории вероятностей, которые порождают сомнения в объективности результатов этой теории.

На кафедре прикладной теории вероятностей Нижегородского госуниверситета накоплен богатый опыт апробирования другого подхода в преподавании трехсеместрового общего курса по основам вероятностного моделирования, теории вероятностей, математической статистике и теории случайных процессов. Этот подход минимально и по необходимости использует результаты и приемы абстрактной теории вероятностей, теории меры и функционального анализа для построения и изучения свойств вероятностных моделей реальных процессов и явлений стохастического характера. Наряду с вынужденной математической строгостью в программе первой части трехсеместрового курса предусмотрено решение значительного числа прикладных задач на непосредственное построение и изучение вероятностных моделей, которые пробуждают и развивают интуицию вероятностно-статистического мировоззрения на мир. Такой подход помогает понять теоретические основы этого курса и дать целостную картину изучения реальных случайных экспериментов от постановки задач на содержательном уровне до получения вероятностных моделей и окончательного ответа. Особенно этот подход становится актуальным в наше время, когда эффективное планирование деятельности государственных предприятий, прогнозирование ситуаций частных компаний на финансовых и товарных рынках, проведение избирательных компаний в условиях жесткой конкуренции требует вероятностно-статистического анализа данных, надежных и обоснованных выводов и прогнозов. Все это предъявляет новые требования к методам и уровню подготовки специалистов в области вероятностного моделирования, прикладной теории вероятностей и математической статистики. Эти требования диктуют вполне определенный выбор следующих основных задач и целей предлагаемого учебника:

— знакомство с методами построения и анализа адекватных стохастических моделей реальных процессов и явлений простейшего типа;
— изложение основ современной прикладной теории вероятностей;
— развитие интуиции вероятностного мировоззрения на мир;
— критическое знакомство с решениями конкретных задач с целью усвоения основных понятий, положений и идей вероятностного моделирования и прикладной теории вероятностей;
— приобретение навыков и умения имитационного моделирования простейших ситуаций стохастического характера с использованием современных информационных технологий.

Для освоения материала курса необходимы знания математики в объеме университетской или вузовской программы. Кроме того, следует особо выделить следующие разделы знаний математики:

— элементы функционального анализа и теории меры;
— начальные сведения по комбинаторному анализу, дискретной математике и теории множеств;
— навыки математического моделирования и программирования;
— основы численных методов;
— элементы компьютерных технологий для использования статистических пакетов.

Суммируя все вышесказанное, отметим, что учебник посвящен изучению фундаментальных и прикладных основ современной теории вероятностного моделирования. При этом основное внимание уделяется: проблеме задания и классификации реальных экспериментов, интуитивным понятиям допустимых, элементарных и наблюдаемых исходов, качественным и количественным признакам статистически устойчивых экспериментов, вопросу статистической зависимости и ее математической формализации, методам построения и анализа адекватных стохастических моделей реальных процессов и явлений. Это определило его структуру и изложение на основе теоретико-множественного подхода с минимально возможным применением теории меры и функционального анализа. Любая самая простая и сложная математическая конструкция в учебнике имеет реальную интерпретацию и обоснование. В силу этого настоящий учебник по основам теории вероятностного моделирования значительно отличается от известных учебников по теории вероятностей, так как содержит большое число конкретных задач с подробными решениями, замечаниями и указаниями на типичные ошибки.

Приложение

Таблица П.1. Значения функции $\varphi(x) = (2\pi)^{-1/2} \exp\{-x^2/2\}$

а)

x	0	1	2	3	4	5	6	7	8	9
0,0	0,39894	39892	39886	39876	39862	39844	39822	39797	39767	39733
0,1	39695	39654	39608	39559	39505	39448	39387	39322	39253	39181
0,2	39104	39024	38940	38853	38762	38667	38568	38466	38361	38251
0,3	38139	38023	37903	37780	37654	37524	37391	37255	37115	36973
0,4	36827	36678	36526	36371	36213	36053	35889	35723	35553	35381
0,5	35207	35029	34849	34667	34482	34294	34105	33912	33718	33521
0,6	33322	33121	32918	32713	32506	32297	32086	31874	31659	31443
0,7	31225	31006	30785	30563	30339	30114	29887	29659	29431	29202
0,8	28969	28737	28504	28269	28034	27798	27562	27324	27086	26848
0.9	26609	26369	26129	25888	25647	25406	25164	24923	24681	24439
1,0	24197	23955	23713	23471	23230	22988	22747	22506	22265	22025
1,1	21785	21546	21307	21069	20831	20594	20357	20121	19886	19652
1,2	19419	19186	18954	18724	18494	18265	18037	17810	17585	17360
1,3	17137	16915	16694	16474	16256	16038	15822	15608	15395	15183
1,4	14973	14764	14556	14350	14146	13943	13742	13542	13344	13147
1,5	12952	12758	12566	12376	12188	12001	11816	11632	11450	11270
1,6	11092	10915	10741	10567	10396	10226	10059	09893	09728	09566
1,7	09405	09246	09089	08933	08780	08628	08478	08329	08183	08038
1,8	07895	07754	07614	07477	07341	07206	07074	06943	06814	06687
1,9	06562	06438	06316	06195	06077	05959	05844	05730	05618	05508
2,0	05399	05292	05186	05082	04980	04879	04780	04682	04586	04491
2,1	04398	04307	04217	04128	04041	03955	03871	03788	03706	03626
2,2	03547	03470	03394	03319	03246	03174	03103	03034	02965	02898
2,3	02833	02768	02705	02643	02582	02522	02463	02406	02349	02294
2,4	02239	02186	02134	02083	02033	01984	01936	01888	01842	01797
2,5	01753	01709	01667	01625	01585	01545	01506	01468	01431	01394
2,6	01358	01323	01289	01256	01223	01191	01160	01130	01100	01071
2,7	01042	01014	00987	00961	00935	00909	00885	00861	00837	00814
2,8	00792	00770	00748	00727	00707	00687	00668	00649	00631	00613

Продолжение табл. П.1

x	0	1	2	3	4	5	6	7	8	9
2,9	00595	00578	00562	00545	00530	00514	00499	00485	00470	00457
3,0	00443	00430	00417	00405	00393	00381	00370	00358	00348	00337
3,1	00327	00317	00307	00298	00288	00279	00271	00262	00254	00246
3,2	00238	00231	00224	00216	00210	00203	00196	00190	00184	00178
3,3	00172	00167	00161	00156	00151	00146	00141	00136	00132	00127
3,4	00123	00119	00115	00111	00107	00104	00100	00097	00094	00090
3,5	00087	00084	00081	00079	00076	00073	00071	00068	00066	00063
3,6	00061	00059	00057	00055	00053	00051	00049	00047	00046	00044
3,7	00042	00041	00039	00038	00037	00035	00034	00033	00031	00030
3,8	00029	00028	00027	00026	00025	00024	00023	00022	00021	00021
3,9	00020	00019	00018	00018	00017	00016	00016	00015	00014	00014

б)

x	0	2	4	6	8
4,	0,0001338	0,0000589	0,0000249	0,0000101	0,0000040
5,	0,0000015	—	—	—	—

Таблица П.2. Значения функции $2^{-1}\Phi(x) = (2\pi)^{-1/2}\int_0^x \exp\{-x^2/2\}dx$

а)

x	0	1	2	3	4	5	6	7	8	9
0,0	0,00000	00399	00798	01197	01595	01994	02392	02790	03188	03586
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214
1,1	36433	36550	36864	37076	37286	37493	37698	37900	38100	38298

Продолжение табл. П.2

x	0	1	2	3	4	5	6	7	8	9
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147
1,3	40320	40490	40658	40824	40988	41149	41308	41466	41621	41774
1,4	41924	42073	42220	42364	42507	42647	42786	42992	43056	43189
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861
3.0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997

б)

x	0	2	4	6	8
4,	0,4999683	0,4999867	0,4999946	0,4999979	0,4999992
5,	0,4999997	—	—	—	—

Таблица П.3. Значения функции $\pi(m;\,\lambda) = \lambda^m (m!)^{-1} \exp\{-\lambda\}$

а)

m	λ									
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
0	0,9048	8187	7408	6703	6065	5488	4965	4493	4065	3678
1	0904	1637	2222	2681	3032	3292	3476	3594	3659	3678
2	0045	0163	0333	0536	0758	0987	1216	1437	1646	1839
3	0001	0010	0033	0071	0126	0197	0283	0383	0493	0613
4	—	0000	0002	0007	0015	0029	0049	0076	0111	0153

б)

m	λ									
	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5	6,0
0	0,2231	1353	0820	0497	0301	0183	0111	0067	0040	0024
1	3346	2706	2052	1493	1056	0732	0499	0336	0224	0148
2	2510	2706	2565	2240	1849	1465	1124	0842	0618	0446
3	1255	1804	2137	2240	2157	1953	1687	1403	1133	0892
4	0470	0902	1336	1680	1888	1953	1898	1754	1558	1338
5	0141	0360	0668	1008	1321	1562	1708	1754	1714	1606
6	0035	0120	0278	0504	0770	1041	1281	1462	1571	1606
7	0007	0034	0099	0216	0385	0595	0823	1044	1234	1376
8	0001	0008	0031	0081	0168	0297	0463	0652	0848	1032
9	0000	0001	0008	0027	0065	0132	0231	0362	0518	0688

Таблица П.4. Значения функции $\sum_{m=s}^{\infty} \pi(m;\,\lambda)$

а)

s	λ									
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
0	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
1	0951	1812	2591	3296	3934	4511	5034	5506	5934	6321
2	0046	0175	0369	0615	0902	1219	1558	1912	2275	2642
3	0001	0011	0036	0079	0143	0231	0341	0474	0628	0803
4	—	0000	0002	0007	0017	0033	0057	0090	0134	0189

б)

s	λ				s	λ			
	2	4	6	8		2	4	6	8
0	1,0000	1,0000	1,0000	1,0000	5	0,0526	0,3711	0,7149	0,9003
1	8646	9816	9975	9996	6	0165	2148	5543	8087
2	5939	9084	9826	9969	7	0045	1106	3937	6866
3	3233	7619	9380	9862	8	0011	0511	2560	5470
4	1428	5665	8488	9576	9	0002	0213	1527	4074

Таблица П.5. Значения $100\,\alpha$-процентных точек $c_{1-\alpha;\,r}$ для хи-квадрат распределения с r степенями свободы $(\mathsf{P}(\chi_r^2 \geqslant c_{1-\alpha;\,r}) = \alpha)$

r	α									
	0,995	0,990	0,975	0,950	0,900	0,1	0,050	0,025	0,010	0,005
1	0,0000	0,0001	0,0009	0,0039	0,0157	2,7055	3,8414	5,0238	6,6349	7,8794
2	0,0100	0,0201	0,0506	0,1025	0,2107	4,6051	5,9914	7,3777	9,2103	10,5966
3	0,0717	0,1148	0,2157	0,3518	0,5843	6,2513	7,8147	9,3484	11,3449	12,8381
4	0,2069	0,2971	0,4844	0,7107	1,0636	7,7794	9,4877	11,1433	13,2767	14,8602
5	0,4117	0,5543	0,8312	1,1454	1,6103	9,2363	11,0705	12,8325	15,0863	16,7496
6	0,6757	0,8720	1,2373	1,6353	2,2041	10,6446	12,5916	14,4494	16,8119	18,5476
7	0,9892	1,2390	1,6898	2,1673	2,8331	12,0170	14,0671	16,0128	18,4753	20,2777
8	1,3444	1,6464	2,1797	2,7326	3,4895	13,3616	15,5073	17,5346	20,0902	21,9550
9	1,7349	2,0879	2,7003	3,3251	4,1681	14,6837	16,9190	19,0228	21,6660	23,5893
10	2,1558	2,5582	3,2469	3,9403	4,8651	15,9871	18,3070	20,4831	23,2093	25,1882
11	2,6032	3,0534	3,8157	4,5748	5,5777	17,2750	19,6751	21,9200	24,7250	26,7569
12	3,0738	3,5705	4,4037	5,2260	6,3038	18,5494	21,0261	23,3367	26,2170	28,2995
13	3,5650	4,1069	5,0087	5,8918	7,0415	19,8119	22,3621	24,7356	27,6883	29,8194
14	4,0746	4,6604	5,6287	6,5706	7,7895	21,0642	23,6848	26,1190	29,1413	31,3193
15	4,6009	5,2293	6,2621	7,2609	8,5467	22,3072	24,9958	27,4884	30,5779	32,8013
16	5,1422	5,8122	6,9076	7,9616	9,3122	23,5418	26,2962	28,8454	31,9999	34,2672
17	5,6972	6,4077	7,5641	8,6717	10,0852	24,7690	27,5871	30,1910	33,4087	35,7185
18	6,2648	7,0149	8,2307	9,3904	10,8649	25,9894	28,8693	31,5264	34,8053	37,1564
19	6,8439	7,6327	8,9065	10,1170	11,6509	27,2036	30,1435	32,8523	36,1908	38,5822
20	7,4338	8,2604	9,5908	10,8508	12,4426	28,4120	31,4104	34,1696	37,5662	39,9968
21	8,0336	8,8972	10,2829	11,5913	13,2396	29,6151	32,6705	35,4789	38,9321	41,4010
22	8,6427	9,5424	10,9823	12,3380	14,0415	30,8133	33,9244	36,7807	40,2894	42,7956
23	9,2604	10,1956	11,6885	13,0905	14,8479	32,0069	35,1725	38,0757	41,6384	44,1813
24	9,8862	10,8564	12,4011	13,8484	15,6587	33,1963	36,4151	39,3641	42,9798	45,5585

Продолжение табл. П.5

r	α									
	0,995	0,990	0,975	0,950	0,900	0,1	0,050	0,025	0,010	0,005
25	10,5197	11,5240	13,1197	14,6114	16,4734	34,3816	37,6525	40,6465	44,3141	46,9278
26	11,1603	12,1981	13,8439	15,3791	17,2919	35,5631	38,8852	41,9232	45,6417	48,2899
27	11,8076	12,8786	14,5733	16,1513	18,1138	36,7412	40,1133	43,1944	46,9630	49,6449
28	12,4613	13,5648	15,3079	16,9279	18,9392	37,9159	41,3372	44,4607	48,2782	50,9933
29	13,1211	14,2565	16,0471	17,7083	19,7677	39,0875	42,5569	45,7222	49,5879	52,3356
30	13,7867	14,9535	16,7908	18,4926	20,5992	40,2560	43,7729	46,9792	50,8922	53,6720
31	14,458	15,655	17,539	19,281	21,434	41,422	44,985	48,232	52,191	55,003
32	15,134	16,362	18,291	20,072	22,271	42,585	46,194	49,480	53,486	56,328
33	15,815	17,073	19,047	20,867	23,110	43,745	47,400	50,725	54,776	57,648
34	16,501	17,789	19,086	21,664	23,952	44,903	48,602	51,966	56,061	58,964
35	17,192	18,509	20,569	22,465	24,797	46,059	49,802	53,203	57,342	60,275
36	17,887	19,233	21,336	23,269	25,643	47,212	50,998	54,437	58,619	61,581
37	18,586	19,960	22,106	24,075	26,492	48,363	52,192	55,668	59,892	62,882
38	19,289	20,691	22,878	24,884	27,343	49,513	53,384	56,895	61,162	64,181
39	19,996	21,426	23,654	25,695	28,196	50,660	54,572	58,120	62,428	65,476
40	20,707	22,164	24,433	26,509	29,051	51,805	55,758	59,342	63,691	66,766

Таблица П.6. Значения $100\,\alpha$-процентных точек $z_{1-\alpha}(m,n)$ для случайной величины Фишера–Снедекора $\chi_{m,n}$ с m, n степенями свободы и плотностью вероятностей $f_{\chi_{u,v}}(x) =$
$= (u/v)^{u/2}(\Gamma(u/2)\,\Gamma(v/2))^{-1}\Gamma((u+v)/2)\,x^{(u-2)/2}(1+v^{-1}ux)^{-(u+v)/2}$,
$x > 0$, где $m,n = 1,2,\ldots$ и $\mathsf{P}(\chi_{m,n} \geqslant z_{1-\alpha}(m,n)) = \alpha$

$\alpha = 0{,}1$

n	m														
	1	2	3	4	5	6	7	8	9	10	20	30	40	60	120
1	39,86	49,50	53,59	55,83	57,24	58,20	58,21	59,44	59,86	60,19	61,74	62,26	62,53	62,79	63,06
2	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39	9,44	9,46	9,47	9,47	9,48
3	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23	5,18	5,17	5,16	5,15	5,14
4	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92	3,84	3,82	3,80	3,79	3,78
5	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,21	3,17	3,16	3,14	3,12
6	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,84	2,80	2,78	2,76	2,74
7	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	3,72	2,70	2,59	2,56	2,54	2,51	2,49
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,42	2,38	2,36	2,34	2,32
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,30	2,25	2,23	2,21	2,18
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,20	2,16	2,13	2,11	2,08

$\alpha = 0{,}05$

n	m														
	1	2	3	4	5	6	7	8	9	10	20	30	40	60	120
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	248,0	250,1	251,1	252,2	253,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,45	19,46	19,47	19,48	19,49
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,66	8,62	8,59	8,57	8,55
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,80	5,75	5,72	5,69	5,66
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,56	4,50	4,46	4,43	4,40
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,87	3,81	3,77	3,74	3,70
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,44	3,38	3,34	3,30	3,27
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,15	3,08	3,04	3,01	2,97
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	2,94	2,86	2,83	2,79	2,75
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,77	2,70	2,66	2,62	2,58

$\alpha = 0{,}01$

n	m														
	1	2	3	4	5	6	7	8	9	10	20	30	40	60	120
1	4052	4999,5	5403	5625	5764	5859	5928	5981	6022	6056	6209	6261	6287	6313	6339
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,45	99,47	99,47	99,48	99,49
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	26,69	26,50	26,41	26,32	26,22
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,02	13,84	13,75	13,65	13,56
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,55	9,38	9,29	9,20	9,11
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,40	7,23	7,14	7,06	6,97
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,16	5,99	5,91	5,82	5,74
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,36	5,20	5,12	5,03	4,95
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	4,81	4,65	4,57	4,48	4,40
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,41	4,25	4,17	4,08	4,00

Таблица П.7. Значения $100\,\alpha$-процентных точек $t_{1-\alpha;r}$ для случайной величины Стьюдента St_r с $r = 1, 2, \ldots$ степенями свободы и плотностью вероятностей

$$f_{\mathrm{St}_r}(x) = (\pi r)^{-1/2}\Gamma((r+1)/2)(\Gamma(r/2))^{-1}(1 + x^2 r^{-1})^{-(r+1)/2},$$

где $\mathsf{P}(\mathrm{St}_r \geqslant t_{1-\alpha;\,r}) = \alpha$

r	α							
	0,4	0,25	0,1	0,05	0,025	0,01	0,005	0,0025
1	0,325	1,000	3,078	6,314	12,706	31,821	63,657	127,321
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	14,089

Продолжение табл. П.7

r	α							
	0,4	0,25	0,1	0,05	0,025	0,01	0,005	0,0025
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	7,453
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	5,598
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	4,773
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	4,317
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	4,029
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	3,833
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	3,690
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	3,581
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106	3,497
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	3,428
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012	3,372
14	0,258	0,692	1,345	1,761	2,145	2,694	2,977	3,326
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	3,286
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	3,252
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,222
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,197
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,174
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,153
21	0,267	0,686	1,323	1,721	2,080	2,518	2,831	3,135
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819	3,119
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807	3,104
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797	3,091
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787	3,078
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779	3,067
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771	3,057
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763	3,047
29	0,256	0,683	1,311	1,699	2,045	2,462	2,756	3,038
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750	3,030
32	0,255	0,682	1,310	1,694	2,037	2,450	2,740	3,015
34	0,255	0,682	1,310	1,691	2,032	2,441	2,728	3,002
36	0,255	0,681	1,310	1,688	2,028	2,434	2,720	2,990
38	0,255	0,681	1,304	1,686	2,024	2,429	2,712	2,980

Продолжение табл. П.7

r	α							
	0,4	0,25	0,1	0,05	0,025	0,01	0,005	0,0025
40	0,255	0,681	1,303	1,684	2,021	2,423	2,704	2,971
42	0,255	0,680	1,302	1,682	2,020	2,418	2,698	2,963
44	0,255	0,680	1,301	1,680	2,020	2,414	2,692	2,956
46	0,255	0,680	1,300	1,680	2,013	2,410	2,690	2,949
48	0,255	0,680	1,300	1,680	2,011	2,407	2,682	2,943
50	0,255	0,679	1,299	1,680	2,009	2,403	2,678	2,937

Список литературы

1. *Боровков А. А.* Теория вероятностей. — М.: Эдиториал УРСС, 1999.
2. *Бочаров П. П., Печинкин А. В.* Теория вероятностей. Математическая статистика. — М.: ФИЗМАТЛИТ, 2005.
3. *Ватутин В. А. и др.* Теория вероятностей и математическая статистика в задачах. — М.: Дрофа, 2003.
4. *Вентцель Е. С.* Теория вероятностей. — М.: Академия, 2003.
5. *Вентцель Е. С., Овчаров Л. А.* Задачи и упражнения по теории вероятностей. — М.: Академия, 2003.
6. *Гнеденко Б. В.* Курс теории вероятностей. — М.: Эдиториал УРСС, 2005.
7. *Гнеденко Б. В., Коваленко И. Н.* Введение в теорию массового обслуживания. — М.: ЛКИ, 2007.
8. *Гнеденко Б. В., Хинчин А. Я.* Элементарное введение в теорию вероятностей. — М.: Эдиториал УРСС, 2003.
9. *Козлов М. В.* Элементы теории вероятностей в примерах и задачах. — М.: МГУ, 1990.
10. *Колмогоров А. Н.* Основные понятия теории вероятностей. — М.: ФАЗИС, 1998.
11. *Колмогоров А. Н., Фомин С. В.* Элементы теории функций и функционального анализа. — М.: ФИЗМАТЛИТ, 2003.
12. *Коршунов Д. А., Фосс С. Г., Эйсымонт И. М.* Сборник задач и упражнений по теории вероятностей. — Санкт-Петербург – Москва – Краснодар: Лань, 2004.
13. *Кочетков Е. С., Смерчинская С. О.* Теория вероятностей в задачах и упражнениях. — М.: Форум – Инфра-М, 2005.
14. *Кремер Н. Ш.* Теория вероятностей и математическая статистика. — М.: ЮНИТИ, 2000.
15. *Прохоров А. В., Ушаков В. Г., Ушаков Н. Г.* Задачи по теории вероятностей. — М.: Наука, 1986.
16. *Пугачев В. С.* Теория вероятностей и математическая статистика. — М.: ФИЗМАТЛИТ, 2002.
17. *Севастьянов Б. А.* Курс теории вероятностей и математической статистики. — М. – Ижевск: Ин-т компьютерных исслед., 2004.
18. *Тутубалин В. Н.* Теория вероятностей и случайных процессов. — М.: МГУ, 1992.
19. *Федоткин М. А.* Основы прикладной теории вероятностей и статистики. — М.: Высшая школа, 2006.
20. *Федоткин М. А.* Процессы обслуживания и управляющие системы // Матем. вопр. киберн. Вып. 6. — М.: Наука, 1996. — С. 51–70.
21. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 1, 2. — М.: Мир, 1984.
22. *Чистяков В. П.* Курс теории вероятностей. — СПб.: Лань, 2003.
23. *Ширяев А. Н.* Вероятность — 1, 2. — М.: МЦНМО, 2004.
24. *Ширяев А. Н.* Задачи по теории вероятностей. — М.: МЦНМО, 2006.

Предметный указатель