В. П. ЧИСТЯКОВ

Курс теории вероятностей

Издание третье, исправленное

Допущено Министерством высшего и среднего специального образования СССР в качестве учебника для студентов высших технических учебных заведений

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1987

ББК 22.171
Ч.68
УДК 519.21 (075.8)

Чистяков В. П. **Курс теории вероятностей:** Учеб.— 3-е изд., испр.— М.: Наука. Гл. ред. физ.-мат. лит.— 1987.— 240 с.

В основу положен материал полугодового курса лекций, читавшегося автором в течение ряда лет в МИФИ. Рассматриваемые темы обычны для начального курса теории вероятностей. В конце глав приводятся задачи для практических занятий; имеются задачи, в которых требуется моделировать различные случайные явления.

Расширенные разделы «Математическая статистика» и «Элементы теории случайных процессов» позволяют использовать книгу в вузах, в которых на изучение теории вероятностей отводится более одного семестра. Предполагается знакомство читателей с курсом математического анализа в объеме программ технических вузов.

Ил. 1. Библиогр. 21 назв.

Рецензент
доктор технических наук профессор *Г. П. Башарин*

Владимир Павлович Чистяков

КУРС ТЕОРИИ ВЕРОЯТНОСТЕЙ

Редактор *И. Е. Морозова.*
Художественный редактор *Г. М. Коровина*
Технический редактор *Е. В. Морозова.*
Корректоры *О. М. Березина, М. Л. Медведская*

ИБ № 32402

Сдано в набор 12.12.86. Подписано к печати 27.05.87.
Формат 84×108/32. Бумага тип. № 2.
Гарнитура обыкновенная. Печать высокая. Усл. печ. л. 12,6.
Усл. кр.-отт. 12,81. Уч.-изд. л. 14,01. Тираж 52 500 экз.
Заказ № 115. Цена 60 коп.

Ордена Трудового Красного Знамени издательство «Наука»
Главная редакция физико-математической литературы
117071 Москва В-71, Ленинский проспект, 15

2-я типография издательства «Наука»
121099 Москва Г-99, Шубинский пер., 6

Ч $\dfrac{1702060000-131}{053(02)-87}$ 67-87

Предисловие к третьему изданию

Третье издание отличается от второго лишь исправлением замеченных опечаток и частичным сокращением таблиц случайных чисел. Поэтому остановимся на основных изменениях, внесенных во второе издание, которое по сравнению с первым было существенно переработано и дополнено.

Во многих технических вузах на основе обязательного курса теории вероятностей читаются полугодовые курсы математической статистики. Автор предполагал, что «Курс теории вероятностей», вышедший в 1978 г., обеспечит, с привлечением дополнительной литературы, потребности курса математической статистики. Однако, как выяснилось из отзывов, многим читателям удобнее пользоваться учебником, в котором оба курса излагаются с единой точки зрения и с едиными обозначениями. В связи с этим глава «Элементы математической статистики» для второго издания была написана заново. В ней изложены основные вопросы, включаемые обычно в полугодовой курс математической статистики для технических вузов. Однако не надо считать, что эта глава может заменить специальные учебники математической статистики, в которых изложение ведется более широко и подробно.

В главу «Элементы теории случайных процессов» добавлены сведения о процессах гибели и размножения и их приложения к задачам массового обслуживания, расширено изложение ветвящихся процессов.

Подробнее чем прежде рассматривается многомерное нормальное распределение. В главу «Предельные теоремы» включено обоснование известного в прикладных задачах метода линеаризации.

Во втором издании исключен параграф, относящийся к обоснованию построения вероятности для бесконечных последовательностей независимых испытаний. Понятия σ-алгебры и борелевских множеств в основном тексте в явном виде практически не используются.

Задачи значительно обновлены; при их подборе существенно использовался задачник [16].

В связи с интенсивным развитием вычислительной техники и доступностью больших и малых ЭВМ появилась возможность существенно использовать в процессе обучения задачи с моделированием тех или иных случайных явлений. Решая эти задачи, учащиеся получают возможность сравнивать выводы теории с численными результатами проведенного ими моделирования, что способствует более глубокому усвоению теории.

Во втором издании увеличено число задач, связанных с моделированием. Понятие о случайных числах введено сразу после классического определения вероятности. Далее регулярно даются задачи с моделированием, иллюстрирующие вновь вводимые понятия.

Отметим в заключение, что предлагаемый учебник фактически объединяет в себе учебник по начальному курсу теории вероятностей, учебник по элементарному курсу математической статистики и задачник.

Автор благодарен всем читателям, сообщившим свои замечания,

Из предисловия к первому изданию

В книге дается математическое изложение некоторых разделов теории вероятностей, основанное на обычном курсе математического анализа технических вузов. Теория меры, интегралы Лебега и Стилтьеса не используются. В связи с этим рассматриваются вероятностные пространства, в которых при задании вероятности используются ряды, интегралы Римана и несобственные интегралы. К таким пространствам относятся конечные и дискретные, а также вероятностные пространства, в которых вероятность событий определяется как интеграл Римана от некоторой положительной функции. Последние для удобства ссылок названы абсолютно непрерывными (термин не является общепринятым). Этих пространств достаточно для изучения тем, которые обычно включаются в курс теории вероятностей: «классические» и «геометрические» вероятности; конечные последовательности испытаний (в частности, цепи Маркова, независимые испытания, схема Бернулли); случайные величины и т. д.

Однако в ряде интересных задач нельзя обойтись указанными выше вероятностными пространствами. В задаче о разорении игрока, при изучении времени до первого успеха в схеме Бернулли, в ряде задач, связанных с приложениями формулы полной вероятности, требуется введение более сложных вероятностных пространств. Такие задачи рассматриваются в предположении, что подходящие вероятностные пространства существуют.

Курс предлагаемого типа автор читал много лет. В зависимости от математической подготовки слушателей менялся только объем излагаемого материала и подбор задач; аксиоматическое изложение сохранялось в любом случае.

Другой подход к способу изложения теории вероятностей в технических вузах состоит в использовании интуитивных представлений о вероятности, независимости, случайной величине вместо их точных определений. Это приводит к замене теории вероятностей решением разрозненных задач из анализа. Точную характеристику подхода такого типа дает Феллер в предисловии к своей книге [18]: «При преподавании теории вероятностей существует тенденция возможно быстрее сводить вероятностные задачи к задачам чистого анализа, избегая специфических особенностей самой теории вероятностей. Такое изложение основывается на плохо определяемом понятии случайной величины, вводимом обычно вначале. В противоположность этому настоящая книга строится на понятии пространства элементарных событий, без которого случайные величины остаются искусственной выдумкой».

Следует отметить, что аналогичные вопросы в преподавании математического анализа давно решены. Обычно не возникают сомнения в том, что нужно давать точное определение предела па языке «ε — δ», хотя оно не очень легко согласуется с интуитивным представлением о пределе. Теория вероятностей даже в кратком изложении должна оставаться математической дисциплиной.

По ходу изложения «Курса теории вероятностей» дается решение некоторого количества задач. Это не устраняет необходимости в самостоятельном решении задач, которые приведены в конце глав. Некоторые из задач являются существенным дополнением к изложенному в основном тексте.

При подборе задач были использованы многие задачники, учебники и монографии (в частности, [4], [7], [9], [11], [18], [19]). Приведенный набор задач желательно пополнить задачами, отражающими специфику вуза.

В каждой главе принята своя нумерация формул. При ссылке на формулу из другой главы к номеру добавляется номер главы.

Я пользуюсь случаем, чтобы поблагодарить сотрудников Математического института, с которыми неоднократно обсуждались вопросы преподавания теории вероятностей. Особенно интенсивными были беседы с Л. Н. Большевым, В. К. Захаровым, О. В. Висковым во время работы по составлению новой программы курса теории вероятностей.

Мне были очень полезны критические замечания и советы Б. А. Севастьянова и В. Ф. Колчина, сделанные ими при многочисленных обсуждениях различных разделов курса лекций по теории вероятностей, а также их замечания, связанные непосредственно с рукописью данной книги.

Введение

1. *Случайные явления.* До возникновения теории вероятностей объектом исследования науки были явления или опыты, в которых условия практически однозначно определяют исход. Так, например, если на материальную точку действует сила тяжести и в некоторый момент заданы положение и скорость материальной точки, то ее дальнейшее движение определяется соответствующим дифференциальным уравнением однозначно. Однако эта механическая модель не всегда удовлетворительно описывает реальные физические явления. Если, например, рассматривать движение пули, то ее траектория практически уже не будет определяться однозначно; начальная скорость пули по многим причинам не остается постоянной при различных выстрелах и, следовательно, окажет влияние на неоднозначность исхода всего опыта.

Если колебания значений начальной скорости невелики (например, меньше погрешности при численном интегрировании уравнения), то можно использовать детерминированную механическую модель, в которой движение однозначно определяется начальными условиями.

Неоднозначность исхода при сохранении основных условий опыта наблюдается для широкого круга явлений. При подбрасывании монеты мы не можем предсказать исход: упадет монета гербом вверх или нет.

Результаты нескольких измерений одной и той же величины, полученных одним и тем же прибором в одних и тех же условиях, различны. Влияние очень большого числа разнообразных причин, каждая из которых в отдельности не может повлиять на результат опыта, приводит к тому, что результат опыта не определяется заранее однозначно; говорят, что результат такого опыта случаен.

Примеры случайных явлений можно указать во многих областях науки и техники (например, в физике, биологии, демографии, в массовом производстве, в системах автоматического управления и т. д.).

2. *Вероятность и частота.* В повседневной речи мы часто используем слова «вероятность», «случай», «событие». Интуитивно вероятность некоторого события воспринимается как объективная характеристика возможности его появления.

Экспериментально установлено, что с ростом числа опытов N, проводимых в одинаковых условиях, *частота* появления некоторого случайного события A (отношение числа опытов N_A, в которых событие A появилось, к общему числу опытов N) становится почти постоянной. Таким образом, можно попытаться с каждым случайным событием связать некоторое число P, к которому приближается частота, и считать это число вероятностью события. Очевидно, что приведенные соображения нельзя считать удовлетворительным математическим определением вероятности. Попытки выбирать число P по разным сериям опытов будут приводить обычно к разным, хотя и близким значениям P.

В простейших случаях интуитивные представления о вероятности часто приводят к однозначному ответу. Так, например, никто не сомневается в том, что вероятность выпадения герба при подбрасывании монеты равна $\frac{1}{2}$. Объяснения обычно приводятся достаточно убедительные:

1) возможны два события — выпал герб, выпала решетка; ни одному из них нельзя отдать предпочтения, так как монета симметрична;

2) при многократном повторении опыта частота появления герба близка к $\frac{1}{2}$.

Первая часть этого объяснения является попыткой построить модель случайного явления; вторая часть — экспериментальная проверка соответствия модели реальному явлению.

Однако при незначительном усложнении опыта «повседневная интуиция» или «здравый смысл» могут подводить. Представим себе, что при каждом из 10 подбрасываний монеты выпал герб. Предлагается угадать, какой стороной упадет монета в следующий раз. Довольно распространено мнение, что выпадение решетки более вероятно. Этот «вывод» делается примерно следующим образом: длинные серии подряд идущих гербов встречаются не часто, а если известно, что длинная серия начинает появляться, то она должна быстрее кончиться, и, следовательно, появление решетки более вероятно. Отсутствие длинных серий гербов нетрудно заметить, даже если не записывать результаты опытов. Однако экспериментальная оценка вероятности появления решетки после 10 гербов требует уже фиксирования результатов опыта. Нужно отобрать все цепочки из 10 гербов и посмотреть, сколько раз следующим результатом будет герб и сколько раз — решетка. Ответ $\frac{1}{2}$ в этом более сложном опыте свидетельствует о более развитой интуиции.

В качестве примера совсем неприемлемого рассуждения можно привести следующий: я *не знаю*, пойдет сегодня дождь или нет, следовательно, его вероятность $\frac{1}{2}$. Эта «модель» к дождю вряд ли имеет какое-либо отношение. Отметим для сравнения, что в примере с монетой учитывалась ее объективная характеристика — симметричность.

Многообразие и сложность задач, встречающихся в приложениях, требуют четкого определения понятий, связанных со случайными явлениями.

3. *Математическая модель.* Теория вероятностей, так же как и другие разделы математики, имеет дело не с явлениями окружающего мира непосредственно, а с их математическими моделями.

В математической модели должны быть правильно переданы существенные стороны изучаемого явления, а несущественные — отброшены. Слишком подробное описание изучаемого явления приводит к усложнению математической модели и может значительно затруднить исследование. Излишнее упрощение модели может привести к неверным выводам. Насколько удачно введена модель, можно судить по согласованности теоретических выводов с опытом.

В математических моделях случайных явлений вероятность рассматривается как функция от случайного события. В курсах математического анализа прежде, чем приступить к изучению функции, довольно много внимания уделяется изучению ее аргумента — действительного числа. Аргументом вероятности является случайное событие. Поэтому в гл. 1 мы прежде всего займемся уточнением интуитивного понятия случайного события, а затем введем понятие вероятности.

Аксиоматический подход построения теории вероятностей, предложенный А. Н. Колмогоровым в книге «Основные понятия теории вероятностей», сделал теорию вероятностей математической наукой. Ее аксиомы и теоремы в абстрактной форме отражают закономерности, присущие случайным явлениям. В настоящее время аксиоматический подход является общепринятым.

В других разделах математики аксиоматический подход был принят значительно раньше, чем в теории вероятностей.

Вероятностное пространство

Для избежания неясностей при описании случайных явлений, результатов опыта или наблюдений необходимо формализовать эти описания. С этой целью вводится множество Ω элементарных исходов эксперимента (пространство элементарных событий); выделяется класс событий $\mathfrak{A}$ (алгебра или σ-алгебра событий), рассмотрением которого можно ограничиться в данной задаче; на множестве событий $\mathfrak{A}$ задается функция P (вероятность), удовлетворяющая некоторым условиям. Тройку $(\Omega, \mathfrak{A}, \mathsf{P})$, которая вводится при формализации любой вероятностной задачи, называют *вероятностным пространством*. Дадим теперь формальное определение вероятностного пространства и покажем на примерах, как проводится формализация реальной задачи.

§ 1. Пространство элементарных событий

Произвольное множество Ω назовем *пространством элементарных событий*. Элементы ω этого множества Ω будем называть *элементарными событиями*. Эти понятия являются первоначальными. В реальном опыте элементарным событиям соответствуют взаимоисключающие исходы. Ввиду большого разнообразия случайных явлений нельзя дать более конкретное определение множества элементарных событий. Для описания каждой реальной задачи множество Ω выбирается наиболее подходящим образом. Рассмотрим ряд примеров, поясняющих выбор множества Ω.

1. Подбрасывание монеты один раз. Возможными исходами в этом опыте будут: выпадение монеты гербом вверх (или просто выпадение герба), выпадение решетки. Кроме того, монета, возможно, встанет на ребро, укатится куда-нибудь и т. д. Можно перечислить ряд исключающих друг друга событий, которые могут произойти с реальной монетой. При математиче-

ском описании этого опыта естественно отвлечься от ряда несущественных исходов и ограничиться только двумя: выпадение герба (можно обозначить это событие Г, ω_1 или $\omega_Г$), выпадение решетки (Р, ω_2 или $\omega_Р$). Таким образом, при описании этого опыта мы полагаем

$$\Omega = \{Г, Р\}, \quad \Omega = \{\omega_1, \omega_2\} \text{ или } \Omega = \{\omega_Г, \omega_Р\}.$$

2. Подбрасывание игральной кости один раз. В этом опыте естественно выбрать $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, где ω_k обозначен исход опыта, заключающийся в выпадении k очков. Имеем шесть исключающих друг друга исходов.

3. Подбрасывание монеты n раз. Каждому исходу опыта естественно поставить в соответствие последовательность длины n по следующему правилу: если при k-м подбрасывании монеты выпал герб, то на k-м месте последовательности пишем Г, а при выпадении решетки — Р. Так, последовательность ГГГ...ГГ обозначает исход опыта, заключающийся в том, что каждый раз выпадал герб. При небольших значениях n все элементарные события нетрудно выписать. Например, при $n = 3$

$$\Omega = \{ГГГ, ГГР, ГРГ, РГГ, РРГ, РГР, ГРР, РРР\}.$$

Нетрудно проверить, что число элементарных событий при любом n равно 2^n. Действительно, по первому знаку последовательности множество Ω можно разбить на две группы цепочек вида

$$\{Г\ldots\}, \{Р\ldots\}.$$

Каждую из этих групп, фиксируя второй знак последовательности, можно снова разбить на две группы. Получим $2 \cdot 2$ групп:

$$\{ГГ\ldots\}, \{ГР\ldots\}, \{РГ\ldots\}, \{РР\ldots\}.$$

Фиксируя третий знак, получим $2 \cdot 2 \cdot 2$ групп. Продолжив этот процесс до фиксирования n знаков, мы получим 2^n групп, каждая из которых состоит из одной последовательности.

4. Работа телефонной станции. Предположим, что мы наблюдаем работу телефонной станции в течение четверти часа и нас интересует число поступивших вызовов. Если телефонная станция обслуживает незначительное количество абонентов, то за время наблюдения может не поступить ни одного вызова, может посту-

пить один вызов, два вызова и т. д. Достаточно ясно, что число вызовов будет всегда конечно. Однако разумно установить верхнюю границу числа вызовов довольно затруднительно. Проще не ограничивать возможное число вызовов и считать возможными исходами 0 и все натуральные числа: $\Omega = \{0, 1, 2, \ldots\}$. Это предположение проще, чем довольно искусственный подбор верхней границы числа вызовов. Предположение о возможности любого числа вызовов кажется абсурдным. Однако если окажется, что очень большое число вызовов происходит с очень малой вероятностью, то это будет совместимо с нашим практическим понятием невозможности. Подобные ситуации возникают довольно часто. Например, в страховом деле при расчетах не ограничивается максимальный возраст; при измерении довольно часто предполагается, что величина ошибки может принимать любые значения. Если в рассматриваемой задаче с работой телефонной станции нас интересует не только общее число поступивших вызовов, но и моменты их поступления, то пространство элементарных событий нужно выбрать более детализированным. Можно, например, исход нашего наблюдения описать ступенчатой функцией, постоянной между моментами поступления вызовов и имеющей скачки в моменты поступления вызовов. В этом случае Ω — множество ступенчатых функций.

5. Стрельба по плоской мишени. Введем в плоскости мишени прямоугольную систему координат uOv и каждому исходу опыта (попадание в определенную точку плоскости) поставим в соответствие координаты этой точки. Тогда множеством Ω является вся плоскость или множество всех упорядоченных пар действительных чисел. Этот факт мы запишем следующим образом:

$$\Omega = \{(u, v): \quad -\infty < u < \infty, \ -\infty < v < \infty\},$$

где u, v — действительные числа. В дальнейшем мы будем часто пользоваться такими обозначениями для различных множеств. Например, множество точек замкнутого единичного круга с центром в начале координат будем записывать в виде

$$\{(u, v): \ u^2 + v^2 \leqslant 1\}.$$

6. Броуновское движение. В микроскоп наблюдается движение малой частицы, подвергающейся большому числу ударов со стороны молекул. Наблюде-

ние проводится в промежутке времени $[0, T]$. Исходом этого опыта будет определенная траектория движения частицы. Если интересоваться перемещением частицы вдоль заданного направления, то в каждый момент времени t, $t \in [0, T]$, положение ее проекции на заданное направление будет определяться координатой $x(t)$. В этом случае $\Omega = \{x(t)\}$ — множество непрерывных действительных функций, определенных на отрезке $[0, T]$.

§ 2. Алгебра событий

В реальном опыте кроме взаимоисключающих исходов можно указать много других случайных событий. В примере 2 из § 1 перечислены элементарные исходы для одного бросания игральной кости. В этом же опыте можно говорить, например, о случайном событии, состоящем в том, что выпало четное число очков. Это событие происходит только в том случае, когда происходит одно из трех элементарных событий: ω_2, ω_4 или ω_6. Выпадению нечетного числа очков соответствуют элементарные события ω_1, ω_3, ω_5. Представляется естественным каждое реальное событие A' в математической модели рассматривать как некоторое подмножество A множества Ω, включив в A те и только те элементарные события, при которых происходит A'.

Случайным событием или просто *событием* будем называть любое подмножество множества Ω, если Ω конечно или счетно:

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\} \text{ или } \Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}.$$

В случае произвольного Ω событиями будем называть только подмножества из некоторого класса $\mathfrak{A}$ подмножеств Ω, который будет определен после введения операций над событиями, совпадающих с операциями над множествами.

Суммой двух событий A и B назовем событие $A + B$ (или $A \cup B$), состоящее из всех элементарных событий, принадлежащих по крайней мере одному из событий A или B. Можно сказать, что в реальном опыте событие, соответствующее $A + B$, состоит в том, что произошло по крайней мере одно из событий A или B. Пусть в примере 5 из § 1 событиями A и B являются попадания соответственно в большой и малый круг (рис. 1) . Тогда событием $A + B$ является заштрихованная область на рис. 1, *а*.

Произведением AB (или $A \cap B$) называется событие, состоящее из элементарных событий, принадлежащих и A и B. Событие AB происходит тогда и только тогда, когда происходит и A и B. Для примера 5 событие AB изображено заштрихованной фигурой на рис. 1, *б*.

Разностью $A \setminus B$ называется событие, состоящее из элементов множества A, не принадлежащих B (см. рис. 1, *в*).

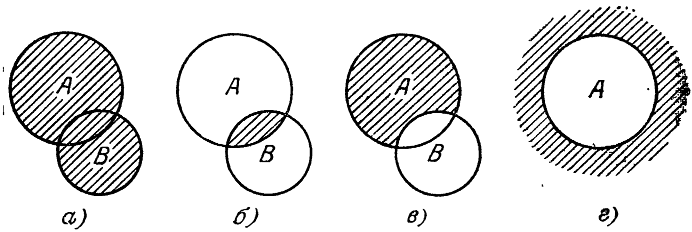

Рис. 1

Событие $A \setminus B$ состоит в том, что A произошло, а B не произошло.

Событие Ω назовем *достоверным*; пустое множество $\varnothing$ назовем *невозможным* событием. Событие $\bar{A} = \Omega \setminus A$ называется *противоположным* событию A (см. рис. 1, *г*). Событие $\bar{A}$ означает, что A не произошло.

События A и B *несовместны*, если $AB = \varnothing$. Тот факт, что A является подмножеством B, будем записывать так: $A \subset B$ (или $B \supset A$). Это значит, что из наступления события A следует наступление B. В примере 5, если $A \subset B$, то попадание в область A, содержащуюся в B, означает также попадание в B. Принадлежность элемента множеству обозначается символом $\in$. Например, $\omega \in \Omega$.

Понятия произведения и суммы событий переносятся на бесконечные последовательности событий. Событие

$$A_1 + A_2 + \ldots + A_n + \ldots = \bigcup_{n=1}^{\infty} A_n$$

состоит из элементарных событий, принадлежащих хотя бы одному из событий A_n, $n = 1, 2, \ldots$ Событие $\bigcap_{n=1}^{\infty} A_n = A_1 A_2 \ldots A_n \ldots$ состоит из элементарных событий, принадлежащих каждому событию A_n, $n = 1, 2, \ldots$

Для произвольных событий непосредственно из определения легко проверить, что

$$AA = A,\ A + A = A,\ A\bar{A} = \varnothing.$$

Часто оказываются полезными следующие равенства:

$$\overline{A + B} = \bar{A}\bar{B},$$

$$(A + B)\,C = AC + BC.$$

Докажем, например, второе. Нужно убедиться, что множества, стоящие в обеих частях равенства, состоят из одних и тех же элементов. Пусть произвольное $\omega \in (A + B)\,C$. Тогда $\omega \in A + B$ и $\omega \in C$. Из $\omega \in A + B$ следует, что ω принадлежит хотя бы одному слагаемому. Пусть, например, $\omega \in A$. Из $\omega \in A$ и $\omega \in C$ следует по определению произведения событий, что $\omega \in AC$, и, следовательно, $\omega \in AC + BC$. Таким образом, любой элемент множества $(A + B)\,C$ является элементом множества $AC + BC$, т. е. $(A + B)\,C \subset AC + BC$. Предположив, что $\omega \in AC + BC$, мы, используя рассуждения, аналогичные приведенным выше, покажем, что любой элемент $AC + BC$ является элементом $(A + B)\,C$. Отсюда следует доказываемое равенство, так как множества его левой и правой частей состоят из одних и тех же элементов. До проведения доказательства равенств полезно, считая A, B, C множествами на плоскости, сделать рисунки множеств, стоящих в левой и правой частях доказываемых равенств.

Перейдем теперь к определению некоторых классов подмножеств множества Ω. В начале главы было отмечено, что вероятность будет рассматриваться как функция от события. Для любых подмножеств множества Ω вероятность не всегда удается определить. В тех случаях, когда класс подмножеств приходится ограничивать, мы будем требовать от него замкнутости относительно введенных выше операций над событиями.

Пусть Ω — произвольное пространство элементарных событий, а $\mathfrak{A}$ — некоторый класс подмножеств множества Ω.

Класс подмножеств $\mathfrak{A}$ называется *алгеброй событий*, если $\Omega \in \mathfrak{A}$ и если $AB \in \mathfrak{A}$, $A + B \in \mathfrak{A}$, $A \setminus B \in \mathfrak{A}$ при любых $A \in \mathfrak{A}$, $B \in \mathfrak{A}$. Отсюда следует, что $\varnothing = \Omega \setminus \Omega \in \mathfrak{A}$. Наименьшей системой подмножеств, являющейся алгеброй, очевидно является система $\mathfrak{A} = \{\varnothing, \Omega\}$. Нетрудно провермть следующее утверждение.

Если $\mathfrak{A}$ — система всех подмножеств множества Ω, то $\mathfrak{A}$ — алгебра. Если Ω — конечное множество, то система всех подмножеств будет также конечным множеством. Для Ω из примера 2 § 1 можно выписать все события алгебры $\mathfrak{A}$, состоящей из всех подмножеств Ω:

$$\varnothing;\ \{\omega_1\},\ \{\omega_2\},\ \ldots,\ \{\omega_6\};$$

$$\{\omega_1,\ \omega_2\},\ \{\omega_1,\ \omega_3\},\ \ldots,\ \{\omega_5,\ \omega_6\};\ \{\omega_1,\ \omega_2,\ \omega_3\},\ \ldots;$$

$$\{\omega_1,\ \omega_2,\ \omega_3,\ \omega_4,\ \omega_5,\ \omega_6\} = \Omega.$$

В этом примере алгебра $\mathfrak{A}$ состоит из $2^6 = 64$ событий. Если множество Ω состоит из N элементов, то число всех подмножеств равно 2^N. Действительно, число последовательностей из 0 и 1 длины N равно 2^N, а между такими последовательностями и подмножествами Ω можно установить взаимно однозначное соответствие по следующему правилу: элемент с номером i из множества Ω включается в подмножество, соответствующее данной последовательности, если на i-м месте последовательности стоит 1.

Приведем еще один пример алгебры событий. Пусть $\Omega = \{(u,\ v)\colon\ 0 < u < 1,\ \ 0 < v < 1\}$ — единичный квадрат в плоскости. В курсе анализа доказано, что объединение, пересечение и разность квадрируемых фигур являются квадрируемыми фигурами. Следовательно, система $\mathfrak{A}$ квадрируемых подмножеств квадрата Ω образует алгебру событий.

Алгебра событий $\mathfrak{A}$ называется *σ-алгеброй* или *борелевской алгеброй*, если из того, что $A_n \in \mathfrak{A}$, $n = 1,\ 2,\ \ldots$, следует

$$\bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}, \qquad \bigcap_{n=1}^{\infty} A_n \in \mathfrak{A}.$$

Рассмотрим следующий пример. Пусть $\Omega = \{u\colon\ -\infty < u < \infty\}$ — числовая прямая. Определим систему множеств $\mathfrak{A}_0$, состоящую из конечных и бесконечных отрезков, интервалов и полуинтервалов:

$$[u_1,\ u_2] = \{u\colon\ u_1 \leqslant u \leqslant u_2\},\quad (u_1,\ u_2] = \{u\colon\ u_1 < u \leqslant u_2\},$$

$$[u_1,\ u_2) = \{u\colon\ u_1 \leqslant u < u_2\},\quad (u_1,\ u_2) = \{u\colon\ u_1 < u < u_2\},$$

где u_1, u_2 — действительные числа ($u_1 \leqslant u_2$) и, кроме того, в строгих неравенствах u_1, u_2 можно заменять на $-\infty$, ∞. Эта система $\mathfrak{A}_0$ не является алгеброй; например, сумма множеств $(-\infty,\ -1) \cup (1,\ \infty)$ не входит в $\mathfrak{A}_0$, хотя каждое слагаемое принадлежит $\mathfrak{A}_0$. Дополним $\mathfrak{A}_0$ всеми конечными суммами множеств из $\mathfrak{A}_0$. Новая, более широкая система множеств $\mathfrak{A}$ является алгеброй.

Назовем σ-алгебру $\mathfrak{A}^*$ *минимальной σ-алгеброй*, порожденной $\mathfrak{A}$, если любая другая σ-алгебра, содержащая множества из $\mathfrak{A}$, содержит также все множества из $\mathfrak{A}^*$. Множества из минимальной σ-алгебры рассматриваемого примера называются *борелевскими*. В частности, борелевским множеством является каждая точка числовой прямой (вырожденный отрезок $\{u_1\} = [u_1, u_1]$). Так как множество рациональных чисел R счетно, то их можно перенумеровать: $R = \{r_1, r_2, \ldots, r_m, \ldots\}$. Тогда $R = \bigcup\limits_{m=1}^{\infty} \{r_m\}$ и, следовательно, R является борелевским множеством. Из $\Omega \in \mathfrak{A}^*$ и $R \in \mathfrak{A}^*$ следует, что множество иррациональных чисел $\Omega \setminus R \in \mathfrak{A}^*$, т. е. тоже является борелевским.

Аналогично определяются борелевская алгебра и борелевские множества на плоскости и в пространстве.

В дальнейшем мы обычно будем рассматривать те задачи, в которых можно ограничиться алгеброй событий.

§ 3. Вероятность

Теперь мы можем ввести понятие вероятности события.

Числовая функция P, *определенная на классе событий* $\mathfrak{A}$, *называется вероятностью, если выполнены следующие условия:*

A1. $\mathfrak{A}$ *является алгеброй событий.*

A2. $\mathsf{P}(A) \geqslant 0$ *для любого* $A \in \mathfrak{A}$.

A3. $\mathsf{P}(\Omega) = 1$.

A4 (*аксиома конечной аддитивности*). *Если* A *и* B *несовместны, то*

$$\mathsf{P}(A + B) = \mathsf{P}(A) + \mathsf{P}(B).$$

Для решения задач, связанных с бесконечными последовательностями событий, требуется дополнить приведенные аксиомы следующей аксиомой непрерывности:

A5. *Для любой убывающей последовательности*

$$A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots \tag{3.1}$$

событий из $\mathfrak{A}$ *такой, что*

$$\bigcap_{n=1}^{\infty} A_n = \varnothing, \tag{3.2}$$

имеет место равенство

$$\lim_{n\to\infty} \mathsf{P}(A_n) = 0. \tag{3.3}$$

Тройку $(\Omega, \mathfrak{A}, \mathsf{P})$, в которой P удовлетворяет A2—A5, а множество $\mathfrak{A}$ не только является алгеброй событий, но еще и замкнуто относительно счетных сумм и произведений (т. е. $\mathfrak{A}$ является σ-алгеброй), называют *вероятностным пространством*.

В дальнейшем мы почти не будем использовать A5 и дополнительное предположение о замкнутости $\mathfrak{A}$ относительно операций в счетном числе.

Если вероятность P удовлетворяет A1—A5 и алгебра $\mathfrak{A}$ не является σ-алгеброй, то функцию P можно доопределить на множествах из минимальной σ-алгебры $\mathfrak{A}^*$, порожденной $\mathfrak{A}$. Приведем без доказательства следующую теорему.

Теорема о продолжении вероятности. *Пусть вероятность* P *удовлетворяет* A1—A5 *и* $\mathfrak{A}^*$ — *минимальная* σ-*алгебра, порожденная* $\mathfrak{A}$. *Тогда существует, и притом единственная, функция* $\mathsf{P}^*(A)$, *которая определена на* $\mathfrak{A}^*$, *удовлетворяет* A2—A5 *и совпадает с* $\mathsf{P}(A)$, *когда* $A \in \mathfrak{A}$.

Таким образом, можно сразу считать, что в A1 система $\mathfrak{A}$ является σ-алгеброй.

Понятие вероятностного пространства содержит лишь самые общие требования, предъявляемые к математической модели случайного явления, и не определяет вероятность однозначно. Дальнейшая конкретизация определения проводится применительно к рассматриваемой реальной задаче.

Рассмотрим два простых примера, которые дают некоторое представление о выборе вероятностной модели на основе понятия вероятностного пространства.

Пусть брошена один раз игральная кость. Для этого опыта в примере 2 § 1 было выбрано $\Omega = \{\omega_1, \omega_2, \ldots, \omega_6\}$. Алгебра событий $\mathfrak{A}$ в данном случае состоит из всех подмножеств $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\} \subset \Omega$ $(k = 1, \ldots, 6;\ i_1 < i_2 < \ldots < i_k,\ i_l = 1, 2, \ldots, 6)$ и пустого множества $\varnothing$. Пусть $p_i \geqslant 0$, $i = 1, \ldots, 6$, — любые числа, для которых $\sum\limits_{i=1}^{6} p_i = 1$. Положим

$$\mathsf{P}(A) = p_{i_1} + p_{i_2} + \ldots + p_{i_k}. \tag{3.4}$$

Отсюда, в частности, получаем

$$\mathsf{P}(\{\omega_i\}) = p_i,\quad i = 1, 2, \ldots, 6. \tag{3.5}$$

Нетрудно проверить, что функция $\mathsf{P}(A)$, определенная формулой (3.4), удовлетворяет аксиомам A2—A4. При различных наборах чисел $(p_1, p_2, \ldots, p_6)$ мы будем получать разные определения вероятности. Очевидно, что посредством математических рассуждений из этого множества определений выбрать нужное мы не сможем.

Если кость симметрична, то представляется естественным дополнительное предположение о равновероятности выпадения различных граней. Тогда из формул (3.5) заключаем, что $p_1 = p_2 = \ldots = p_6 = 1/6$, и, следовательно, вероятность $\mathsf{P}(A)$ определяется однозначно. Окончательное заключение о качестве выбранной модели может быть сделано после экспериментальной проверки.

Если кость несимметрична, то в подходящей модели некоторые или все p_i должны быть отличны от 1/6. Пусть, например, кубик игральной кости склеен из плотной бумаги и к грани, противоположной грани с «6», прикреплен груз. В этом случае будет выпадать всегда «6». В формуле (3.4) нужно положить $p_6 = 1$, $p_1 = p_2 = \ldots$ $\ldots = p_5 = 0$.

Рассмотрим еще две математические модели опыта, заключающегося в подбрасывании двух одинаковых монет.

Модель 1. Положим $\Omega = \{\omega_1, \omega_2, \omega_3\}$, где элементарное событие ω_1 означает, что обе монеты выпали гербами вверх, ω_2 — обе монеты выпали решетками вверх, ω_3 — монеты выпали разными сторонами. Если считать элементарные события равновероятными, то вероятность события $A = \{\omega_1, \omega_2\}$, состоящего в том, что монеты выпали одинаковыми сторонами, равна $\mathsf{P}_1(A) = 2/3$.

Модель 2. Пусть $\Omega = \{ГГ, ГР, РГ, РР\}$, где знак Г на первом или втором месте в обозначении элементарного события означает, что на первой или на второй монете выпал герб; Р аналогично связано с выпадением решетки. Считая элементарные события равновероятными, для события $A = \{ГГ, РР\}$ получим $\mathsf{P}_2(A) = 2/4 = 1/2$.

Получили две разные вероятности одного и того же события. Экспериментальный выбор подходящей модели может быть проведен следующим образом. Подбросим N раз (N — достаточно большое) пару монет; пусть N_A раз произошло событие A. Если частота $N_A/N \approx 2/3$, то считаем подходящей модель 1; если $N_A/N \approx 1/2$, то выбираем модель 2. Эта задача выбора из двух моделей (или двух гипотез) является одной из задач математической статистики. Более подробно она будет рассмотрена в § 6 гл. 9.

В задачах, подобных задаче о подбрасывании двух монет, практически всегда правильный выбор может быть сделан из соображений симметрии. Подбрасываются две физически различные монеты, поэтому естественно считать множество Ω состоящим из четырех элементарных событий, которым естественно приписать одинаковые вероятности. Модель 1 казалась бы более естественной, если бы можно было себе представить монеты физически неразличимыми. Опыт показывает, что в действительности монеты ведут себя как различимые. Однако этот достаточно очевидный факт для монет оказывается неверным для некоторых типов частиц. Бозе и Эйнштейн показали, что некоторые типы частиц ведут себя как неразличимые *).

Рассмотренный пример показывает, что нельзя слишком полагаться на интуицию; подходящей моделью может оказаться совершенно неожиданная.

Укажем несколько простых свойств вероятности, которые непосредственно следуют из аксиом А2 — А4. Из аксиом А3, А4 и равенства $A + \bar{A} = \Omega$ следует, что

$$\mathsf{P}(\bar{A}) = 1 - \mathsf{P}(A). \tag{3.6}$$

Полагая здесь $A = \Omega$, получим $\mathsf{P}(\varnothing) = 0$. Для любых событий A и B имеет место формула

$$\mathsf{P}(A + B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB). \tag{3.7}$$

Действительно, представим события $A + B$ и B в виде $A + B = A + B\bar{A}$ и $B = B\bar{A} + BA$. События в правых частях этих равенств несовместны, и, следовательно,

$$\mathsf{P}(A + B) = \mathsf{P}(A) + \mathsf{P}(B\bar{A}),$$
$$\mathsf{P}(B) = \mathsf{P}(B\bar{A}) + \mathsf{P}(AB).$$

Отсюда легко следует (3.7).

Непосредственно из (3.7) для любых A и B получаем

$$\mathsf{P}(A + B) \leqslant \mathsf{P}(A) + \mathsf{P}(B). \tag{3.8}$$

Нетрудно проверить следующее свойство: если $A \subseteq B$, то

$$\mathsf{P}(A) \leqslant \mathsf{P}(B). \tag{3.9}$$

Можно показать (см. [2], гл. 2, § 1, стр. 30—31), что аксиомы А4 и А5 эквивалентны следующему свойству счетной аддитивности вероятности: *если в последовательности $A_1, A_2, \ldots, A_n, \ldots$ события попарно несовместны (т. е. $A_iA_j = \varnothing$ при $i \neq j$)* и $A = \bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}$, *то*

$$\mathsf{P}(A) = \sum_{n=1}^{\infty} \mathsf{P}(A_n). \tag{3.10}$$

*) Подробнее см. [18], § 5 гл. 2.

Отметим без доказательства еще одно свойство, связанное с аксиомой непрерывности А5.

Если

$$A_1 \subset A_2 \subset \ldots \subset A_n \subset A_{n+1} \subset \ldots \quad и \quad A = \bigcup_{n=1}^{\infty} A_n$$

или

$$A_1 \supset A_2 \supset \ldots \supset A_n \supset A_{n+1} \supset \ldots \quad и \quad A = \bigcap_{n=1}^{\infty} A_n,$$

то

$$\mathsf{P}(A) = \lim_{n \to \infty} \mathsf{P}(A_n). \tag{3.11}$$

Несколько важных частных случаев вероятностных пространств рассматриваются в гл. 2. В дальнейшем вновь вводимые теоретико-вероятностные понятия позволят нам расширить набор частных случаев вероятностных пространств и приемов их построения.

Задачи к главе 1

1. Проверить следующие соотношения между событиями:
1) $A \setminus B = A\bar{B}$;
2) $A \setminus B = A \setminus AB = (A + B) \setminus B$;
3) $\overline{(A + B)} = \bar{A}\bar{B}$, $\overline{AB} = \bar{A} + \bar{B}$;
4) $A\,(B \setminus C) = AB \setminus AC$.

2. Установить, какие из следующих соотношений правильны:
1) $A \setminus (B \setminus C) = (A \setminus B) + C$;
2) $A \setminus (B \setminus C) = (A \setminus B) + AC$;
3) $(A + B) \setminus C = (A \setminus C) + (B \setminus C)$;
4) $(A + B) \setminus C = A + (B \setminus C)$;
5) $A\bar{B}C \subseteq A + B$;
6) $(A \setminus B)\,(C \setminus D) = AC \setminus BD$.

3. Упростить следующие выражения:
1) $A + AB$; 3) $(A \setminus C)\,(B \setminus \bar{C})$;
2) $(A + B)\,(A + \bar{B})$; 4) $(A + B)\,(\bar{A} + B)\,(A + \bar{B})$.

4. Пусть

$$A_n = \left\{x\colon\ a \leqslant x < a + \frac{1}{n}\right\},$$

$$B_n = \left\{x\colon\ a \leqslant x \leqslant b - \frac{1}{n}\right\}.$$

Для событий

$$A = \bigcap_{n=1}^{\infty} A_n, \quad B = \bigcup_{n=1}^{\infty} B_n$$

найти более простые выражения.

5. Какие подмножества множества Ω в примере 3 из § 1 при $n = 3$ соответствуют событиям, состоящим в том, что

1) при первом подбрасывании выпал герб;
2) всего выпало ровно два герба;
3) выпало не более одного герба?

6. Пусть A, B, C — три произвольных события. Найти выражения для событий, состоящих в том, что из A, B, C

1) произошло только A;
2) произошли A и B, C не произошло;
3) все три события произошли;
4) произошло по крайней мере одно из событий;
5) произошло одно и только одно событие;
6) ни одно событие не произошло;
7) произошло не более двух событий.

7. Пусть в примере 3 из § 1 $n = 3$. Является ли алгеброй следующая система подмножеств:

$$\varnothing,\ \Omega\ (ГГГ,\ ГРГ,\ ГГР,\ ГРР),\ (РГГ,\ РРГ,\ РГР,\ РРР)?$$

Простейшие вероятностные схемы и их обобщения

§ 1. Классическое определение вероятности

Пусть $\Omega = \{\omega_1, \ldots, \omega_s\}$ — пространство элементарных событий, а $\mathfrak{A}$ — алгебра событий, содержащая все 2^s подмножеств $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ множества Ω. В *классическом определении вероятности* полагают, что $\mathsf{P}(\omega_i) = 1/s$, $i = 1, 2, \ldots, s$; поэтому вероятность $\mathsf{P}(A)$ события $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ равна отношению числа элементарных событий ω_i, входящих в A, к общему числу элементарных событий в Ω:

$$\mathsf{P}(A) = \frac{|A|}{|\Omega|} = \frac{k}{s}. \tag{1.1}$$

Здесь и в дальнейшем число элементов любого конечного множества M будем обозначать $|M|$.

Определенная равенством (1.1) функция $\mathsf{P}(A)$ удовлетворяет всем аксиомам А1—А5. Действительно, для любого множества Ω система всех его подмножеств является алгеброй. Аксиомы А2—А4 непосредственно следуют из (1.1).

Аксиома А5 для конечных $\mathfrak{A}$ следует из предыдущих аксиом, поскольку в этом случае в (1.3.1) все события, начиная с некоторого n, совпадут с $\varnothing$, а $\mathsf{P}(\varnothing) = 0$.

Таким образом, введенная формулой (1.1) функция $\mathsf{P}(A)$ является вероятностью.

Классическое определение вероятности служит хорошей математической моделью тех случайных явлений, для которых исходы опыта в каком-либо смысле симметричны и поэтому представляется естественным предположение об их равновозможности. Обычно это предположение оправдано в задачах из области азартных игр, лотерей и т. д. Это объясняется тем, что при изготовлении игральных костей, карт и организации лотерей заботятся о соблюдении равновозможности различных исходов. Такие же требования предъявляются к организации выборочного контроля и выборочных статистических исследований.

Дадим описание некоторых часто встречающихся вероятностных схем, в которых детализируется общее классическое определение.

Обозначим $\mathcal{A}$ множество из N чисел: $\mathcal{A} = \{1, 2, \ldots, N\}$; пусть $\omega = (i_1, i_2, \ldots, i_n)$ — упорядоченный набор из n элементов множества $\mathcal{A}$. Вероятностную схему, в которой

$$\Omega = \{\omega = (i_1, \ldots, i_n) \colon i_k \in \mathcal{A},\ k = 1, 2, \ldots, n\} \tag{1.2}$$

и все элементарные события ω равновероятны, называют *схемой случайного выбора с возвращением.*

Схемой случайного выбора без возвращения называют вероятностную схему, в которой

$$\Omega = \{\omega = (i_1, \ldots, i_n) \colon i_k \in \mathcal{A},\ k = 1, \ldots, n,\ i_1, \ldots, i_n \text{ различны}\} \tag{1.3}$$

и элементарные события ω равновероятны.

Эти две схемы, (1.2) и (1.3), можно интерпретировать следующим образом. Пусть в урне имеется N одинаковых шаров, занумерованных числами множества $\mathcal{A}$. Из урны по одному выбирается n шаров, и записываются их номера: $i_1, i_2, \ldots, i_n$.

Если каждый вынутый шар сразу возвращается в урну, то для обозначения множества исходов Ω используется множество (1.2). Если вынимаемые шары в урну не возвращаются, то множеством исходов является (1.3). В случае выбора без возвращения иногда удобнее считать, что шары извлекаются не по одному, а сразу все. В этом случае не нужно учитывать порядок извлекаемых шаров. В (1.3) произойдет «укрупнение» элементарных событий: все события $\omega = (i_1, \ldots, i_n)$, отличающиеся друг от друга только расположением номеров $i_1, \ldots, i_n$, объединяются в одно событие $\widetilde{\omega}$. В результате такого объединения получится более «бедное», чем (1.3), пространство элементарных событий:

$$\widetilde{\Omega} = \{\widetilde{\omega} = (i_1, \ldots, i_n) \colon i_k \in \mathcal{A},\ k = 1, \ldots, n,\ i_1 < \ldots < i_n\}. \tag{1.4}$$

Используя формулу (1.1), нетрудно показать, что любое реальное событие, которое можно описать в обеих схемах (1.3) и (1.4), имеет в этих схемах одинаковые вероятности.

В формулировках задач по теории вероятностей, предназначенных для упражнений, довольно часто приводится только описание опыта или явления, а полная математическая формулировка не дается. Предполагается, что решение должно состоять из двух частей: 1) выбор подходящей модели для описания данного в условии задачи опыта и математическая формулировка задачи; 2) решение математической задачи.

Рассмотрим несколько примеров.

Пример 1. Из урны, содержащей M белых и $N - M$ черных шаров, наудачу извлекается сразу n шаров. Какова вероятность того, что среди выбранных n шаров окажется ровно m белых?

Решение. Слово «наудачу» в описаниях опыта встречается довольно часто. В данной задаче предполагается, что шары были хорошо перемешаны, что все они одного радиуса, одинаково гладкие и отличаются только цветом; выбирающий шаров не видит. В таком случае разумно предположить, что все элементарные исходы опыта равновероятны, и, следовательно, можно воспользоваться классической схемой.

За элементарные события естественно принять любые подмножества по n элементов, выбранные из множества N шаров. Из школьного курса математики известно, что число таких подмножеств равно C_N^n. Таким образом, в формуле (1.1) нужно положить $s = C_N^n$. Каждый набор шаров, входящий в интересующее нас событие (обозначим его A_m), состоит из двух частей: 1) m белых шаров и 2) $n - m$ черных шаров. Все такие наборы можно получить следующим образом. Сначала выберем части наборов из белых шаров; число таких частей C_M^m; затем отдельно составим части наборов из черных шаров; число таких частей C_{N-M}^{n-m}. Объединение любой части набора из белых шаров с любой частью набора из черных шаров дает полный набор шаров, принадлежащий A_m. Следовательно, число элементарных событий в A_m равно $k = C_M^m C_{N-M}^{n-m}$ и по формуле (1.1)

$$\mathsf{P}(A_m) = P_n(m, N, M) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}, \tag{1.5}$$

$$m = m_0, m_0 + 1, \ldots, m_1,$$

где $m_0 = \max(0, M - N + n)$, $m_1 = \min(M, n)$.

Набор чисел (1.5) называют *гипергеометрическим распределением*.

В приложениях к задачам выборочного контроля роль шаров играют N изделий проверяемой партии. Число M бракованных изделий (белых шаров) неизвестно. Может оказаться, что сплошь все изделия проверить нельзя: их слишком много, или проверка приводит к уничтожению изделия (например, потребуется при проверке установить срок службы лампочки). Тогда из всей партии изделий отбирают для проверки небольшую часть из n изделий. Если среди выбранных изделий оказалось m бракованных, то полагают $\frac{M}{N} \approx \frac{m}{n}$. Формула (1.5) используется при оценке отклонения $\frac{m}{n}$ от $\frac{M}{N}$. В рассмотренной ситуации неизвестным параметром является число M.

Рассмотрим пример, когда требуется оценить неизвестный параметр N. Пусть N — неизвестное число рыб в некотором водоеме. Можно провести отлов M рыб, пометить их и пустить обратно. По числу m помеченных рыб в повторном отлове из n рыб можно делать заключания о величине N.

Оценим по формуле (1.5) вероятность получения какого-либо выигрыша в «Спортлото» по одной карточке. Участник лотереи из 49 номеров отмечает 6 (49 шаров, среди которых 6 белых; $N = 49$, $M = 6$). После того как участник сдал карточку, проводится выборка $n = 6$ номеров. Если число m номеров, отмеченных участником и попавших в выборку, оказалось больше 2, то участник получает какой-либо выигрыш. Если событие A — получение выигрыша, то $\bar{A} = A_0 + A_1 + A_2$. По формуле (1.5) находим:

$$\mathsf{P}(A_0) = 0{,}435965, \quad \mathsf{P}(A_1) = 0{,}413019,$$
$$\mathsf{P}(A_2) = 0{,}132378, \quad \mathsf{P}(A_3) = 0{,}017650,$$
$$\mathsf{P}(A_4) = 0{,}000969, \quad \mathsf{P}(A_5) = 0{,}000018,$$
$$\mathsf{P}(A_6) = 0{,}0000007151.$$

Отсюда

$$\mathsf{P}(\bar{A}) = \mathsf{P}(A_0) + \mathsf{P}(A_1) + \mathsf{P}(A_2) = 0{,}981362.$$

Искомую вероятность находим по формуле (1.3.6):

$$\mathsf{P}(A) = 1 - \mathsf{P}(\bar{A}) = 0{,}018638.$$

Пример 2. Ребенок, играя десятью кубиками, на которых написаны буквы М, М, Т, Т, А, А, А, К, И, Е, сложил слово «МАТЕМАТИКА». Можно ли считать, что ребенок грамотный?

Решение. Сначала дадим математическую формулировку задачи. Если допустить, что ребенок неграмотный и родители не научили его складывать единственное слово, то естественно предположить, что расположение кубиков «математика» не более привлекательно по сравнению с остальными. В таком случае можно ожидать, что классическая схема окажется достаточно подходящей. Оценим вероятность события A, состоящего в расположении кубиков «матаматика». Пространством элементарных событий являются всевозможные перестановки 10 кубиков. Таких перестановок 10!. При этом кубики с одинаковыми буквами мы считали различными. Подсчитаем, сколько элементарных событий входит в A. Кубики с буквами, отличными от М, Т, А, должны стоять на определенных местах, 3 кубика с буквой A можно расположить на трех местах $3! = 6$ способами; кубики с буквой М располагаются двумя способами и с буквой Т тоже двумя. Сочетая каждое расположение с каждым, получим, что A состоит из $2 \cdot 2 \cdot 6 = 24$ элементарных событий. По формуле (1.1)

$$\mathsf{P}(A) = \frac{24}{10!} = \frac{1}{15\,120}.$$

Эта вероятность очень мала, и событие A можно считать практически невозможным. Если же оно осуществилось, то следует считать нашу гипотезу о неграмотности ребенка неверной. Однако при большом числе испытаний по данной «программе» даже неграмотный может сложить слово «математика»; в среднем из 15 120 неграмотных 1 будет ошибочно считаться грамотным.

Пример 3. У человека в кармане n ключей, из которых только один подходит к его двери. Ключи последовательно извлекаются (без возвращания) до тех пор, пока не появится нужный ключ. Найти вероятность того, что нужный ключ появится при k-м извлечении.

Решение. Будем продолжать извлекать ключи, после появления нужного ключа, до конца. За множество элементарных событий примем всевозможные последовательности из n ключей. Таких последовательностей $n!$. Последовательностей, у которых нужный ключ находится на определенном месте, очевидно, $(n-1)!$, так как одно место занято нужным ключом, а остальные $n-1$ ключей можно на $n-1$ месте расставить $(n-1)!$ способами. Таким образом, искомая вероятность равна $(n-1)!/n! = 1/n$.

§ 2. Дискретные вероятностные пространства

Пусть $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$ — счетное множество, $\mathfrak{A}$ — набор всех подмножеств Ω. Пусть p_n, $n = 1, 2, \ldots$, — неотрицательные числа, удовлетворяющие условию $\sum_{n=1}^{\infty} p_n = 1$. Для любого события $A \in \mathfrak{A}$ положим

$$\mathsf{P}(A) = \sum_{n \in \{n:\omega_n \in A\}} p_n. \tag{2.1}$$

Нетрудно показать, что функция $\mathsf{P}(A)$, определенная (2.1), удовлетворяет аксиомам А2—А4.

Проверим А5. События A_n в последовательности (1.3.1) запишем в виде

$$A_n = \{\omega_{l_1(n)}, \omega_{l_2(n)}, \ldots\},$$

где $l_1(n) < l_2(n) < \ldots$, причем при любом $n > 1$ последовательность индексов элементарных событий из A_n является подпоследовательностью соответствующей последовательности из A_{n-1}. Из условия (1.3.2) следует, что $l_1(n) \to \infty$ при $n \to \infty$. Если бы это было не так, то произведение в (1.3.2) не было бы пусто. По формуле (2.1) находим

$$\mathsf{P}(A_n) = \sum_{k=1}^{\infty} p_{l_k(n)} \leqslant \sum_{s=l_1(n)}^{\infty} p_s.$$

Отсюда при $n \to \infty$ следует (1.3.3), так как правая часть последнего неравенства является остатком сходящегося ряда. Из свойства счетной аддитивности (1.3.11) следует, что для дискретных вероятностных пространств вероятность $\mathsf{P}(A)$ однозначно определяется ее значениями $p_n = \mathsf{P}(\omega_n)$, $n = 1, 2, \ldots$

Отметим, что при $p_n = 0$ с $n > N$ можно ограничиться конечным пространством элементарных событий $\Omega = \{\omega_1, \ldots, \omega_N\}$. В случае $p_1 = p_2 = \ldots = p_N = 1/N$ получаем классическое определение вероятности.

§ 3. Геометрические вероятности

Классическое определение вероятности нельзя применить к опыту с бесконечным числом «равновероятных» исходов. К описанию такой ситуации приспособлено геометрическое определение вероятности.

Пусть Ω — ограниченное множество n-мерного евклидова пространства. Будем предполагать, что Ω имеет объем *). Рассмотрим систему $\mathfrak{A}$ подмножеств множества Ω, имеющих объем. Для любого $A \in \mathfrak{A}$ положим

$$\mathsf{P}(A) = \frac{\mu(A)}{\mu(\Omega)}, \tag{3.1}$$

где $\mu(C)$ — объем множества $C \in \mathfrak{A}$.

В формулировках задач мы часто будем использовать выражение: «точка равномерно распределена на множестве Ω». Это выражение означает, что вероятность попадания точки в подмножество A нужно вычислять по формуле (3.1).

Отметим, что в двух предыдущих классах вероятностных пространств (§§ 1, 2) в систему $\mathfrak{A}$ входили все подмножества Ω. При геометрическом определении вероятности в качестве $\mathfrak{A}$ уже нельзя взять все подмножества Ω, так как некоторые подмножества не имеют площади или объема.

В схеме геометрических вероятностей выбор модели, подходящей для описания реального явления, менее очевиден, чем в классической схеме. В разных моделях для одного и того же реального события можно получить разные вероятности. На выборе разных математических моделей основан известный парадокс Бертрана (см. задачу 20; более подробно см. [4], гл. I, § 6, стр. 36—37).

Для иллюстрации схемы геометрических вероятностей рассмотрим следующую задачу.

Задача Бюффона. Плоскость расчерчена параллельными прямыми, расстояние между которыми равно $2a$. На плоскость наудачу брошена игла длины $2l$ ($l < a$). Найти вероятность того, что игла пересечет какую-нибудь прямую.

Решение. Дадим сначала определение подходящего пространства элементарных событий. Пусть u — расстояние от центра иглы до ближайшей прямой, а φ — угол, составленный иглой с этой прямой. Пара чисел (φ, u) задает положение иглы с точностью до выбора конкрет-

---

*) Если под объемом множества понимать его меру Лебега, то система $\mathfrak{A}$ окажется σ-алгеброй, и тогда функция $\mathsf{P}(A)$, определенная формулой (3.1), будет удовлетворять и аксиоме А5. Отметим, что система $\mathfrak{A}$ содержит, в частности, все подмножества множества Ω, являющиеся обычными квадрируемыми или кубируемыми фигурами, которые изучаются в любом курсе анализа.

ной прямой. Поскольку нас интересует только взаимное расположение иглы и ближайшей прямой, то можно в качестве Ω выбрать прямоугольник

$$\Omega = \{(\varphi, u)\colon\ 0 \leqslant \varphi \leqslant \pi,\ 0 \leqslant u \leqslant a\}.$$

Пересечение иглы с прямой происходит в том и только в том случае, когда $u \leqslant l \sin \varphi$. Таким образом, интересующее нас событие имеет вид

$$A = \{(\varphi, u)\colon\ 0 \leqslant u \leqslant l \sin \varphi,\ 0 \leqslant \varphi \leqslant \pi\}.$$

Так как

$$\mu(A) = \int\limits_0^\pi l \sin \varphi \, d\varphi, \quad \mu(\Omega) = a\pi,$$

то по формуле (3.1) находим

$$\mathsf{P}(A) = 2l/a\pi. \tag{3.2}$$

О соответствии математической модели опыту можно судить по результатам экспериментов. Пусть игла была брошена n раз и m раз произошло пересечение. При больших n частота $\frac{m}{n} \approx \mathsf{P}(A) = \frac{2l}{a\pi}$. Отсюда можно получить экспериментальную оценку числа $\pi \approx 2\frac{l}{a} \cdot \frac{n}{m}$. Приведем результаты некоторых экспериментов с бросанием иглы, заимствованные из книги [6].

Экспериментатор	l/a	n	m	Оценка π
Вольф, 1850 г.	0,8	5000	2532	3,1596
Рейна, 1925 г.	0,5419	2520	859	3,1795

Схема геометрических вероятностей успешно применяется в астрономии, атомной физике, биологии, кристаллографии.

§ 4. Абсолютно непрерывные вероятностные пространства

Пусть $\Omega = \{(u_1, u_2, \ldots, u_n)\}$ — n-мерное действительное евклидово пространство, $\pi(u_1, u_2, \ldots, u_n)$ — неотрицательная функция, интегрируемая по Риману по любой квадрируемой области из Ω. Будем предполагать, что существует несобственный интеграл по Ω от функции $\pi(u_1, u_2, \ldots, u_n)$ и

$$\int \underset{\Omega}{\cdots} \int \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n = 1.$$

Обозначим $\mathfrak{A}$ алгебру, порожденную квадрируемыми областями в Ω. Для любого $A \in \mathfrak{A}$ положим

$$\mathsf{P}(A) = \int \underset{A}{\cdots} \int \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n. \tag{4.1}$$

Можно рассмотреть более общий случай, когда функция $\pi(u_1, \ldots, u_n)$ не ограничена в конечном числе точек Ω. Тогда интегралы в (4.1) по множествам A, содержащим такие точки, нужно понимать как несобственные. Нетрудно проверить, что функция $\mathsf{P}(A)$, определенная соотношением (4.1), удовлетворяет аксиомам А2—А5 *). Таким образом, мы определили вероятностное пространство. Построенное вероятностное пространство будем называть n-мерным *абсолютно непрерывным вероятностным пространством*.

Отметим, что рассмотренная выше схема геометрических вероятностей является двумерным абсолютно непрерывным вероятностным пространством с $\pi(u_1, u_2) = 1/\mu(\Omega)$, если $(u_1, u_2) \in \Omega$, Ω — квадрируемая фигура, и $\pi(u_1, u_2) = 0$, если $(u_1, u_2) \notin \Omega$.

§ 5. Случайные числа

Рассмотрим один важный частный случай схемы выбора с возвращением (1.2). Пусть $\mathcal{A} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Тогда множеством Ω в (1.2) является множество всех цифровых последовательностей длины n. Так как число элементов множества Ω равно $|\Omega| = 10^n$, то вероятность любой фиксированной последовательности $\omega = (i_1 i_2 \ldots i_n)$ равна 10^{-n}.

Случайными (равномерно распределенными) числами называют цифровую последовательность $(i_1 i_2 \ldots i_n)$, полученную в результате проведения реального опыта, который хорошо описывается рассматриваемой схемой. Такого

*) Из теоремы о продолжении вероятности следует, что формула (4.1) позволяет определить вероятность на минимальной σ-алгебре $\mathfrak{A}^*$, порожденной алгеброй $\mathfrak{A}$.

типа опыты проводят, например, при разыгрывании номеров в «Спортлото». Оценка близости реального опыта и его математической модели является одной из задач математической статистики.

Последовательности случайных чисел можно использовать для получения реализаций различных случайных процессов. Если, например, в последовательности случайных чисел $(i_1 i_2 \ldots i_n)$ четные цифры заменить буквой Г, а нечетные — Р, то полученную последовательность из букв Г и Р можно рассматривать как реализацию опыта, состоящего в n-кратном подбрасывании симметричной монеты. Действительно, каждая последовательность, состоящая из букв Г и Р, может быть получена из одинакового числа равновероятных цифровых последовательностей, и, следовательно, новые последовательности равновероятны.

При расчетах методом Монте-Карло (см. [5]) требуются длинные последовательности случайных чисел, которые в процессе вычислений можно непрерывно вводить в ЭВМ. Методы получения случайных чисел, основанные на реальном извлечении шаров из урны, для этой цели непригодны. Поэтому в практических вычислениях пользуются какими-либо алгоритмами получения последовательности псевдослучайных чисел, близкой по свойствам к последовательности случайных чисел. Если в экспериментах со случайными числами требуется не очень много чисел, то можно воспользоваться таблицами случайных чисел (например, табл. 7). Эксперименты с небольшими массивами случайных чисел часто используются для иллюстрации различных выводов теории.

Задачи к главе 2

1. Брошено две игральных кости. Предполагая, что элементарные события равновероятны, найти вероятности событий:

$$A = \{\text{на 1-й кости выпала «1»}\},\ \bar{A},$$

$$B = \{\text{выпала хотя бы одна «6»}\},\ A\bar{B}.$$

2. На полке в случайном порядке расставлено n книг, среди которых находится двухтомник Д. Лондона. Предполагая, что различные расположения книг равновероятны, найти вероятность того, что оба тома двухтомника расположены рядом.

3. Числа $1, 2, \ldots, n$ расставлены случайным образом. Предполагая, что различные расположения чисел равновероятны, найти вероятность того, что числа 1, 2, 3 расположены в порядке возрастания, но не обязательно рядом.

4. Выписано три случайных числа (см. § 5). Найти вероятности событий:

A = {все выписанные числа одинаковы},

B = {все выписанные числа различны},

C = {среди выписанных чисел ровно два совпадают}.

5. Выписана последовательность из n случайных чисел. Найти вероятности событий:

A = {1-е число — четное},

B = {среди n чисел ровно m делятся на 3},

C = {среди n чисел ровно $m+2$ делятся на 3, и два из них расположены на концах последовательности}.

6. Сравнить вероятности событий:

A = {при одновременном бросании четырех костей выпала хотя бы одна «1»},

B = {при 24 бросаниях двух костей выпали хотя бы один раз две «1»}.

7. В чулане n пар ботинок. Из них случайно выбирается $2r$ ботинок ($2r < n$). Найти вероятность того, что среди выбранных ботинок: а) нет парных; б) имеется ровно одна пара.

8. В партии изделий 90 исправных и 10 бракованных. Найти вероятность того, что среди 10 проданных изделий: а) ровно одно бракованное; б) нет бракованных.

9. По некоторому участку железной дороги за N интервалов времени проходит заданное количество поездов, среди которых M тяжелых. Каждый интервал времени может быть свободен или занят одним поездом. Любое расположение тяжелых поездов по интервалам времени имеет одну и ту же вероятность. Прохождение тяжелых поездов в соседних интервалах времени нежелательно. Оценить сверху вероятность появления хотя бы одной пары соседних интервалов, занятых тяжелыми поездами, если $N = 1000$, $M = 10$.

Указание. В качестве математической модели взять схему случайного выбора без возвращения. Воспользоваться неравенством

$$\mathsf{P}(A_1 + A_2 + \ldots + A_{999}) \leqslant \mathsf{P}(A_1) + \ldots + \mathsf{P}(A_{999}),$$

где A_k = {k-й и $(k+1)$-й интервалы заняты тяжелыми поездами}.

10. В купейный вагон (9 купе по 4 места) семи пассажирам продано семь билетов. Найти вероятности событий:

A = {пассажиры попали в два купе},

B = {пассажиры попали в три купе}.

Рассмотреть два случая:

1) пассажиры покупают билеты в разное время независимо друг от друга (воспользоваться схемой случайного выбора без возвращения);

2) пассажиры едут вместе, и один покупает билеты всей группе (предположить, что номера проданных пассажиру мест идут подряд, а наименьший номер места выбирается случайно из множества номеров $\{1, 2, \ldots, 30\}$).

11. Из множества чисел $\{-n, -n+1, \ldots, -1, 0, 1, \ldots, n\}$ по схеме случайного выбора с возвращением выбирается два числа x и y. Пусть p_n — вероятность того, что $x^2 + y^2 \leqslant n^2$. Найти $\lim\limits_{n\to\infty} p_n$.

12. Найти вероятность того, что дни рождения 12 человек придутся на разные месяцы года.

13. Равновероятной схемой размещения частиц по ячейкам называют схему размещения, в которой номера ячеек, последовательно занимаемых частицами, получают посредством случайного выбора с возвращением.

Обозначим $\mu_r = \mu_r(n, N)$ число ячеек, содержащих ровно по r частиц после размещения n частиц по N ячейкам. Найти вероятности событий:

1) $\mu_0(n, N) > 0$ (при $n = N$);
2) $\mu_0(n, N) = 1$ (при $n = N$).

14. Найти вероятность того, что на две карточки «Спортлото» с отмеченными номерами (4, 12, 38, 20, 41, 46) и (4, 12, 38, 20, 41, 49) будет получено по одному минимальному выигрышу (т. е. угадано ровно по три числа).

15. На отрезок $[0, 1]$ наудачу брошена точка. Предположив, что ее координата ξ равномерно распределена на $[0, 1]$, найти функции $F(x) = \mathsf{P}(\xi < x)$, $F'(x)$ $(-\infty < x < \infty)$.

16. Случайная точка A с координатами (ξ_1, ξ_2) равномерно распределена в квадрате $\Omega = \{(x_1, x_2)\colon 0 \leqslant x_i \leqslant 1,\ i = 1, 2\}$. Найти функции $F(x) = \mathsf{P}(\xi_1 + \xi_2 < x)$, $F'(x)$ $(-\infty < x < \infty)$.

17. На отрезок $[0, 1]$ наудачу брошены 2 точки, разбившие его на 3 отрезка. Какова вероятность того, что из этих отрезков можно построить треугольник? За множество Ω принять значения пары чисел (ξ_1, ξ_2), являющихся координатами точек на отрезке $[0, 1]$; предположить, что точка (ξ_1, ξ_2) равномерно распределена на квадрате Ω.

18. На бесконечную шахматную доску со стороной квадрата a бросается наудачу монета радиуса r, $2r < a$. Найти вероятность того, что: 1) монета целиком попадет внутрь одного квадрата; 2) пересечет не более одной стороны квадрата.

19. Двое договорились встретиться в отрезке времени $[0, T]$. Первый пришедший ждет второго в течение времени l, $l < T$. Найти вероятность того, что встреча произойдет. За множество Ω принять точки квадрата $\{(t_1, t_2)\colon 0 \leqslant t_i \leqslant T,\ i = 1, 2\}$, где t_1 и t_2 — моменты прихода встречающихся.

20. *Парадокс Бертрана*. В круге радиуса R случайно проводится хорда. Обозначим ξ ее длину. Найти вероятность $\mathsf{P}(\xi > R)$ того, что длина хорды больше стороны правильного вписанного шестиугольника, если: а) середина хорды равномерно распределена в круге; б) направление хорды задано, а ее середина равномерно распределена на диаметре, перпендикулярном ее направлению; в) один конец хорды закреплен, а другой равномерно распределен на окружности.

Вероятность $\mathsf{P}(\xi > R)$ зависит от интерпретации слова «случайно», и ее числовые значения различны в случаях а), б), в).

21. В урне M белых и $N-M$ черных шаров. По схеме случайного выбора с возвращением из урны извлекается n шаров. Найти вероятности событий:

A = {при k-м извлечении появился белый шар},

B = {при k-м и l-м извлечениях появились белые шары},

C = {среди n извлеченных шаров ровно m белых}.

22. Решить задачу 21 в случае выбора без возвращения.

23. Пусть ξ_k — k-е ($k = 1, 2, \ldots, n$) число последовательности из n случайных чисел. Найти вероятности событий:

а) $\{\xi_k = i\}$, $k = 1, \ldots, n$; $i = 0, 1, \ldots, 9$;

б) $\{\xi_k = i, \xi_l = j\}$, k, $l = 1, \ldots, n$; $i, j = 0, 1, \ldots, 9$.

24. Используя таблицу случайных чисел, получить реализацию опыта, описанного в задаче 21, если: а) $N = 10$, $M = 3$, $n = 20$; б) $N = 100$, $M = 35$, $n = 20$. В обоих случаях по полученным реализациям найти частоту появления белого шара (отношение числа белых шаров в полученной последовательности к общему числу n извлеченных шаров).

Условные вероятности. Независимость событий

§ 1. Условные вероятности

При изучении реальных случайных явлений иногда возникает или искусственно создается ситуация, когда мы получаем дополнительную информацию о возможных исходах опыта Ω. Например, при исследовании превращений частиц мы можем сначала ограничиться рассмотрением только результатов опыта с превращениями определенного типа.

Остановимся более подробно на следующем примере иллюстративного характера. Допустим, что студент из 30 билетов успел выучить билеты с 1-го по 3-й и с 28-го по 30-й. На экзамен он пришел одиннадцатым, и оказалось, что к его приходу остались только билеты с 1-го по 20-й (событие A). Вероятность события

$$B = \{\text{студент получил выученный билет}\}$$

без дополнительной информации о том, что событие A произошло, может быть вычислена по классическому определению с $\Omega = \{1, 2, \ldots, 30\}$. Согласно формуле (2.1.1) имеем $\mathsf{P}(B) = \frac{|B|}{|\Omega|} = \frac{1}{5}$. При дополнительной информации (событие A произошло) множество возможных исходов A состоит из $|A| = 20$ элементарных исходов, а событие B вместе с A наступает в $|AB| = 3$ случаях. Следовательно, в рассматриваемом примере естественно определить условную вероятность $\mathsf{P}(B \mid A)$ события B при условии, что A произошло, формулой $\mathsf{P}(B \mid A) = |AB|/|A| = 3/20$. Отсюда, поделив числитель и знаменатель на $|\Omega|$, получим

$$\mathsf{P}(B \mid A) = \frac{\mathsf{P}(AB)}{\mathsf{P}(A)}, \tag{1.1}$$

так как согласно классическому определению $\mathsf{P}(AB) = |AB|/|\Omega|$ и $\mathsf{P}(A) = |A|/|\Omega|$. Формула (1.1) принимается за общее определение условной вероятности. *Пусть* $(\Omega, \mathfrak{A}, \mathsf{P})$ — *произвольное вероятностное прост-*

ранство. Если A, $B \in \mathfrak{A}$, $\mathsf{P}(A) > 0$, то условная вероятность события B при условии, что событие A произошло, определяется формулой (1.1), *в правой части которой символ P понимается как вероятность в рассматриваемом вероятностном пространстве.*

Пусть теперь некоторое событие A с $\mathsf{P}(A) > 0$ фиксировано. Нетрудно проверить, что функция $\mathsf{P}_A(B) = \mathsf{P}(AB \mid A) = \mathsf{P}(AB)/\mathsf{P}(A)$, определенная для всех $B \in \mathfrak{A}$, удовлетворяет аксиомам А2—А4. Таким образом, для $\mathsf{P}_A(B)$ справедливы все следствия (1.3.6)—(1.3.11), полученные в § 3 гл. 1 непосредственно из аксиом.

Понятие условной вероятности позволяет естественным образом определить независимость событий. Равенство

$$\mathsf{P}(B \mid A) = \mathsf{P}(B), \quad \mathsf{P}(A) > 0, \tag{1.2}$$

легко согласуется с интуитивным представлением о независимости события B от A. Если $\mathsf{P}(B) > 0$, то из равенства (1.2), используя (1.1), нетрудно получить, что $\mathsf{P}(A \mid B) = \mathsf{P}(A)$. *За определение независимости двух событий A и B принимается более симметричное условие*:

$$\mathsf{P}(AB) = \mathsf{P}(A) \cdot \mathsf{P}(B), \tag{1.3}$$

эквивалентное (1.2), если $\mathsf{P}(A) > 0$.

§ 2. Вероятность произведения событий

Равенство (1.1) можно записать в виде «теоремы умножения»

$$\mathsf{P}(AB) = \mathsf{P}(A)\,\mathsf{P}(B \mid A). \tag{2.1}$$

По индукции из (2.1) легко получить более общую формулу

$$\mathsf{P}(A_1 A_2 \ldots A_n) = \mathsf{P}(A_1)\,\mathsf{P}(A_2 \mid A_1) \ldots \mathsf{P}(A_n \mid A_1 \ldots A_{n-1}). \tag{2.2}$$

При $n = 2$ формула (2.2) совпадает с (2.1). Пусть (2.2) доказано для $n - 1$ сомножителей. Тогда для n сомножителей (2.2) следует из предположения индукции и равенства (2.1), в котором нужно положить $A = A_1 \ldots A_{n-1}$, $B = A_n$.

Если события A и B независимы, то условную вероятность $\mathsf{P}(B \mid A)$ в формуле (2.1) можно заменить на безусловную $\mathsf{P}(B)$. В результате получим (1.3). Возможность замены условных вероятностей безусловными в (2.2) связана с новым свойством совокупности событий $A_1, A_2, \ldots, A_n$. *События* $A_1, \ldots, A_n$ *будем называть взаимно независимыми* (*или независимыми в совокупности, или просто независимыми*), *если для всех комбинаций индексов* $1 \leqslant i_1 < \ldots < i_k \leqslant n$ $(k = 2, \ldots, n)$ *имеем*

$$\mathsf{P}(A_{i_1} A_{i_2} \ldots A_{i_k}) = \mathsf{P}(A_{i_1}) \mathsf{P}(A_{i_2}) \ldots \mathsf{P}(A_{i_k}). \quad (2.3)$$

Если (2.3) выполнено только при $k = 2$, то события называются *попарно независимыми*. Отметим, что из попарной независимости не следует взаимная независимость (см. задачу 7).

Формулы (2.1), (2.2), очевидно, не пригодны для вычисления вероятностей произведений событий, так как правые части этих формул содержат условные вероятности, для вычисления которых нужно знать вероятности произведений. Полезность формул (2.1) и (2.2) обнаруживается при построении математических моделей серий опытов, которые будут рассмотрены в следующей главе. Здесь мы ограничимся разбором отдельных примеров.

В классическом и геометрическом определениях вероятности (см. §§ 1, 3 гл. 2) предполагалась «равновозможность» исходов опыта. Теперь мы обратимся к примерам вероятностных пространств, построение которых основано на знании значений условных вероятностей или знании независимости некоторых событий.

В приложениях часто оказывается, что условные вероятности естественно задаются условиями опыта или определяются приближенно из дополнительных опытов. Например, при превращениях частиц в химических или ядерных реакциях частоты превращений различных типов можно принять за приближенные значения соответствующих условных вероятностей. В сериях опытов с перекладыванием или извлечением шаров из урн можно естественно задать условные вероятности получения выборки определенного типа из данной урны, если состав шаров в ней известен из предыдущих опытов. Отправляясь от заданных значений условных вероятностей, можно вычислить вероятности элементарных событий.

Пусть, например, заданы вероятности

$$\mathsf{P}(A) = a_1, \ \mathsf{P}(B \mid A) = a_{11}, \ \mathsf{P}(B \mid \bar{A}) = a_{21}. \quad (2.4)$$

Требуется построить такое вероятностное пространство $(\Omega, \mathfrak{A}, \mathsf{P})$, что для событий A, B выполнены равенства (2.4.).

Из равенств (2.4) следует, что

$$\mathsf{P}(\bar{A}) = 1 - a_1 = a_2, \quad \mathsf{P}(\bar{B} \mid A) = 1 - a_{11} = a_{12},$$
$$\mathsf{P}(\bar{B} \mid \bar{A}) = 1 - a_{21} = a_{22}.$$

Положим $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Событиями A и B назовем подмножества $A = \{\omega_1, \omega_3\}$, $B = \{\omega_1, \omega_2\}$. Тогда $AB = \{\omega_1\}$. Нетрудно проверить, что

$$\{\omega_2\} = \bar{A}B, \ \{\omega_3\} = A\bar{B}, \ \{\omega_4\} = \bar{A}\bar{B}.$$

Если подходить к определению Ω менее формально, то можно сразу положить

$$\Omega = \{AB, \bar{A}B, A\bar{B}, \bar{A}\bar{B}\}.$$

Распределение вероятностей на всех подмножествах конечного множества однозначно определяется заданием вероятностей всех подмножеств, состоящих из одного элементарного события. Положим

$$\begin{aligned} \mathsf{P}(AB) &= \mathsf{P}(A)\,\mathsf{P}(B \mid A) = a_1 a_{11}, \\ \mathsf{P}(\bar{A}B) &= \mathsf{P}(\bar{A})\,\mathsf{P}(B \mid \bar{A}) = a_2 a_{21}, \\ \mathsf{P}(A\bar{B}) &= \mathsf{P}(A)\,\mathsf{P}(\bar{B} \mid A) = a_1 a_{12}, \\ \mathsf{P}(\bar{A}\bar{B}) &= \mathsf{P}(\bar{A})\,\mathsf{P}(\bar{B} \mid \bar{A}) = a_2 a_{22}. \end{aligned} \tag{2.5}$$

Правые части этих равенств неотрицательны и в сумме дают единицу. Распределение вероятностей задано. Отправляясь от заданного распределения вероятностей, найдем

$$\mathsf{P}(A) = \mathsf{P}(AB) + \mathsf{P}(A\bar{B}) = a_1 a_{11} + a_1 a_{12} = a_1,$$
$$\mathsf{P}(B \mid A) = \frac{\mathsf{P}(AB)}{\mathsf{P}(A)} = \frac{a_1 a_{11}}{a_1} = a_{11},$$
$$\mathsf{P}(B \mid \bar{A}) = \frac{\mathsf{P}(\bar{A}B)}{\mathsf{P}(\bar{A})} = \frac{a_2 a_{21}}{1 - a_1} = a_{21}.$$

Таким образом, мы нашли распределение вероятностей, удовлетворяющее условию (2.3).

Построение вероятностного пространства по условным вероятностям в более общем случае будет рассмотрено в следующей главе.

В примерах, рассматриваемых в этой главе, обычно предполагается, что соответствующие условные вероятности заданы. Построение по ним вероятностного пространства может быть проведено так же, как в рассмотренном выше примере.

Дадим теперь еще одно решение задачи о ключах (см. пример 3 § 1 гл. 2), основанное на использовании понятия условной вероятности. Предлагаемое здесь решение является частичным построением другого вероятностного пространства, описывающего последовательное извлечение ключей. Вероятностные пространства, удобные для описания последовательности испытаний, будут подробно рассмотрены в следующей главе. Событие A_k, состоящее в том, что нужный ключ появится в k-м испытании, можно представить в виде произведения

$$A_k = \bar{A}_1\bar{A}_2 \ldots \bar{A}_{k-1}A_k.$$

Отсюда, используя формулу (2.2), получим

$$\mathsf{P}(A_k) = \mathsf{P}(\bar{A}_1)\,\mathsf{P}(\bar{A}_2 \mid \bar{A}_1) \ldots \mathsf{P}(A_k \mid \bar{A}_1\bar{A}_2 \ldots \bar{A}_{k-1}).$$

Значения сомножителей можно считать заданными в условиях задачи. Действительно, из n ключей только один подходит и $n-1$ не подходят. Следовательно, $\mathsf{P}(\bar{A}_1) = \frac{n-1}{n}$. Если произошло событие $\bar{A}_1$, то осталось $n-1$ ключей, среди которых один подходит и $n-2$ не подходят. Отсюда $\mathsf{P}(\bar{A}_2 \mid \bar{A}_1) = \frac{n-2}{n-1}$ и т. д. Чтобы приписать вероятности $\mathsf{P}(A_k \mid \bar{A}_1 \ldots \bar{A}_{k-1})$ определенное значение, нужно иметь в виду, что к k-му извлечению осталось $n-k+1$ ключей, из которых один подходит к двери. Тогда естественно положить $\mathsf{P}(A_k \mid \bar{A}_1 \ldots \bar{A}_{k-1}) = \frac{1}{n-k+1}$. Окончательно получаем

$$\mathsf{P}(A_k) = \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot \ldots \cdot \frac{n-k+1}{n-k+2} \cdot \frac{1}{n-k+1} = \frac{1}{n}.$$

Результат получился такой же, как в § 1 гл. 2. Если отправляться от распределения вероятностей, введенного для этой задачи в § 1 гл. 2, то можно по формуле (1.1) вычислить условные вероятности

$$\mathsf{P}(\bar{A}_2 \mid \bar{A}_1), \quad \mathsf{P}(\bar{A}_3 \mid \bar{A}_1\bar{A}_2), \ldots$$

Их числовые значения совпадут со значениями, которые были использованы в этом параграфе.

§ 3. Формула полной вероятности

Пусть A — произвольное событие, события $B_1, B_2, \ldots, B_n$ попарно несовместны, $\mathsf{P}(B_k) > 0$, $k = 1, \ldots, n$, и $A \subseteq B_1 + B_2 + \ldots + B_n$. Тогда имеет место следующая формула (*формула полной вероятности*):

$$\mathsf{P}(A) = \sum_{k=1}^{n} \mathsf{P}(B_k)\mathsf{P}(A \mid B_k). \tag{3.1}$$

Для доказательства этой формулы заметим, что A можно представить в виде следующей суммы попарно несовместных событий:

$$A = AB_1 + AB_2 + \ldots + AB_n.$$

Отсюда, воспользовавшись аксиомой А4 и формулой (2.1), получим формулу (3.1):

$$\mathsf{P}(A) = \sum_{k=1}^{n} \mathsf{P}(AB_k) = \sum_{k=1}^{n} \mathsf{P}(B_k)\mathsf{P}(A \mid B_k).$$

Используя (1.3.10), формулу (3.1) можно распространить на случай счетной системы попарно несовместных событий B_k, $k = 1, 2, \ldots, n, \ldots$

Заменив в равенстве

$$\mathsf{P}(B_k \mid A) = \frac{\mathsf{P}(AB_k)}{\mathsf{P}(A)} = \frac{\mathsf{P}(B_k)\mathsf{P}(A \mid B_k)}{\mathsf{P}(A)}$$

вероятность $\mathsf{P}(A)$ по формуле (3.1), получим *формулу Байеса*:

$$\mathsf{P}(B_k \mid A) = \frac{\mathsf{P}(B_k)\mathsf{P}(A \mid B_k)}{\sum_{i=1}^{n} \mathsf{P}(B_i)\mathsf{P}(A \mid B_i)}. \tag{3.2}$$

Формула полной вероятности, так же как формулы (2.1), (2.2), может быть использована для построения подходящего вероятностного пространства по заданным условным вероятностям. Ниже рассматриваются два примера, в которых условные вероятности естественно считать заданными.

Пример 1. На фабрике, изготовляющей болты, первая машина производит 25%, вторая — 35%, третья — 40% всех изделий. Брак в их продукции составляет соответственно 5%, 4%, 2%.

а) Какова вероятность того, что случайно выбранный болт оказался дефектным?

б) Какова вероятность того, что случайно выбранный болт произведен первой, второй и третьей машинами, если он оказался дефектным?

Решение. а) Обозначим через A событие, состоящее в том, что случайно выбранный болт — дефектный, а через B_1, B_2, B_3 — события, состоящие в том, что этот болт произведен соответственно первой, второй и третьей машинами. Очевидно, что формула (3.1) применима. Таким образом, используя условие задачи, получим

$$\mathsf{P}(A) =$$
$$= \mathsf{P}(B_1)\mathsf{P}(A \mid B_1) + \mathsf{P}(B_2)\mathsf{P}(A \mid B_2) + \mathsf{P}(B_3)\mathsf{P}(A \mid B_3) =$$
$$= 0{,}25\cdot 0{,}05 + 0{,}35\cdot 0{,}04 + 0{,}40\cdot 0{,}02 = 0{,}0345.$$

б) К тем же событиям можно применить формулу Байеса (3.2) при $n = 3$ для $k = 1, 2, 3$:

$$\mathsf{P}(B_1 \mid A) = \frac{0{,}25\cdot 0{,}05}{0{,}0345} = \frac{125}{345},$$
$$\mathsf{P}(B_2 \mid A) = \frac{0{,}35\cdot 0{,}04}{0{,}0345} = \frac{140}{345},$$
$$\mathsf{P}(B_3 \mid A) = \frac{0{,}40\cdot 0{,}02}{0{,}0345} = \frac{80}{345}.$$

Пример 2. Из урны, содержавшей M белых и $N - M$ черных шаров, один шар неизвестного цвета утерян. Какова вероятность извлечь наудачу из урны белый шар?

Решение. Пусть B_k — событие, состоящее в том, что утеряно k белых шаров ($k = 0, 1$); A — событие, состоящее в том, что шар, извлеченный из оставшихся шаров, оказался белым. Положим

$$\mathsf{P}(B_0) = \frac{N-M}{N}, \quad \mathsf{P}(B_1) = \frac{M}{N},$$
$$\mathsf{P}(A \mid B_0) = \frac{M}{N-1}, \quad \mathsf{P}(A \mid B_1) = \frac{M-1}{N-1};$$

По формуле полной вероятности

$$\mathsf{P}(A) = \frac{N-M}{N}\cdot\frac{M}{M-1} + \frac{M}{N}\cdot\frac{M-1}{N-1} = \frac{M}{N}.$$

Отметим, что вероятность извлечения белого шара из урны до утери шара тоже равна M/N.

Рассмотрим еще один пример, который дает некоторое представление об использовании формулы полной вероятности при исследовании случайных процессов с непрерывным временем.

Пример 3. Вероятность поступления на телефонную линию одного вызова за время $(t, t+h)$ равна $\alpha h + o(h)$, $h \to 0$ вероятность того, что ни один вызов за время $(t, t+h)$ не поступит, равна $1 - \alpha h + o(h)$. Если линия занята, то вызов теряется. Если в момент t еще продолжается разговор, то за время $(t, t+h)$ он окончится с вероятностью $\beta h + o(h)$, $h \to 0$. Вызовы поступают независимо друг от друга. Найти $P_0(t)$ — вероятность того, что линия в момент t свободна, и вероятность $P_1(t)$ того, что линия занята.

Решение. Предположим сразу, что подходящее вероятностное пространство существует и что для интересующих нас событий

$$B_t^{(0)} = \{\text{линия в момент } t \text{ свободна}\},$$

$$B_t^{(1)} = \{\text{линия в момент } t \text{ занята}\}$$

при каждом t вероятности $P_0(t) = \mathsf{P}(B_t^{(0)})$, $P_1(t) = \mathsf{P}(B_t^{(1)})$ определены и непрерывны по t. Так как

$$B_t^{(0)} + B_t^{(1)} = \Omega, \quad B_t^{(0)} B_t^{(1)} = \varnothing,$$

то по формуле полной вероятности (3.1)

$$\mathsf{P}(B_{t+h}^{(0)}) = \mathsf{P}(B_t^{(0)})\,\mathsf{P}(B_{t+h}^{(0)} \mid B_t^{(0)}) + \mathsf{P}(B_t^{(1)})\,\mathsf{P}(B_{t+h}^{(0)} \mid B_t^{(1)}). \quad (3.3)$$

Свободная в момент времени t линия останется свободной в момент $t+h$, если за время $(t, t+h)$ вызовов не будет. Так как другие события, при которых линия останется свободной в момент $t+h$, имеют вероятность $o(h)$, то

$$\mathsf{P}(B_{t+h}^{(0)} \mid B_t^{(0)}) = 1 - \alpha h + o(h).$$

Занятая в момент t линия будет свободной к моменту $t+h$, если закончится разговор и за время $(t, t+h)$ новых вызовов не будет. Вероятность этого события равна

$$(\beta h + o(h))(1 - \alpha h + o(h)) = \beta h + o(h).$$

Эта вероятность вносит основной вклад в $\mathsf{P}(B_{t+h}^{(0)} \mid B_t^{(1)})$. Сумма остальных слагаемых равна $o(h)$. Таким образом,

$$\mathsf{P}(B_{t+h}^{(0)} \mid B_t^{(1)}) = \beta h + o(h).$$

Подставляя найденные условные вероятности в (3.3) и используя обозначения $P_0(t)$, $P_1(t)$, получим

$$P_0(t+h) = (1-\alpha h)P_0(t) + \beta h P_1(t) + o(h).$$

Отсюда

$$\frac{P_0(t+h)-P_0(t)}{h} = -\alpha P_0(t) + \beta P_1(t) + \frac{o(h)}{h}.$$

Переходя к пределу при $h \to 0$, получим *)

$$\frac{dP_0(t)}{dt} = -\alpha P_0(t) + \beta P_1(t). \tag{3,4}$$

Аналогично найдем уравнение для $P_1(t)$:

$$\frac{dP_1(t)}{dt} = \alpha P_0(t) - \beta P_1(t). \tag{3.5}$$

Полагая $P_0(0) = 1$, $P_1(0) = 0$, найдем решение системы (3.4)—(3.5):

$$\begin{aligned} P_0(t) &= \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}\, e^{-(\alpha+\beta)t}, \\ P_1(t) &= \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}\, e^{-(\alpha+\beta)t}. \end{aligned} \tag{3,6}$$

При $t \to \infty$ вероятность $P_1(t)$ стремится к величине $\alpha/(\alpha+\beta)$. Естественно, что с ростом интенсивности поступления вызовов α вероятность занятости линии увеличивается.

Из формул (3.6) следует, что решение $(P_0(t), P_1(t))$ системы (3.4)—(3.5) удовлетворяет равенству $P_0(t) + P_1(t) = 1$. Это же равенство следует из предположения о существовании вероятностного пространства, в котором определены вероятности $P_0(t)$ и $P_1(t)$ противоположных событий. Если подставить $P_1(t) = 1 - P_0(t)$ в уравнение (3.4), то вместо системы можно решить одно уравнение для $P_0(t)$.

К предположениям о существовании подходящего вероятностного пространства, в котором определены вероятности нужных нам событий, нужно относиться с известной осторожностью. Рассмотрим следующий пример. Пусть состояниями частицы являются числа 0, 1, 2, . . . При $t = 0$ частица находится в состоянии 0. Если частица в момент t находится в состоянии k, то к моменту $t+h$

*) Из существования предела при $h \to 0$ следует существование только левосторонних производных в (3.4) и (3.5). Заменив в предыдущих рассуждениях t и $t+h$ на $t-h$ и t, получим выражения для правосторонних производных, совпадающие с (3.4) и (3.5).

останется в этом состоянии с вероятностью $1-\alpha_k h+o(h)$, $h\to 0$, и перейдет в состояние $k+1$ с вероятностью $\alpha_k h+o(h)$. Пусть $P_k(t)$ — вероятность того, что частица в момент t находится в состоянии k. Рассуждая так же, как в рассмотренном выше примере, для функций $P_k(t)$, $k=0, 1, 2, \ldots$, можно получить бесконечную систему дифференциальных уравнений. Естественно было ожидать, что $\sum\limits_{k=0}^{\infty} P_k(t)=1$. Однако если α_k быстро увеличиваются с ростом k, то $\sum\limits_{k=0}^{\infty} P_k(t)<1$. Более подробно этот пример рассмотрен в книге [18], § 4 гл. 17. Оказывается, что частица за конечное время с положительной вероятностью может пройти по всем состояниям. В этом случае набор чисел $\{\alpha_k\}$ уже не определяет процесс при всех $t>0$.

Задачи к главе 3

1. Брошено две игральных кости. Какова вероятность того, что выпало две «3», если известно, что сумма выпавших очков делится на три?

2. Известно, что при бросании 10 игральных костей появилась по крайней мере одна «1». Какова вероятность того, что появилось две «1» или более?

3. Из множества чисел $\{1, 2, \ldots, N\}$ по схеме случайного выбора без возвращения выбираются три числа. Найти условную вероятность того, что третье число попадет в интервал, образованный первыми двумя, если известно, что первое число меньше второго.

4. Из урны, содержащей 3 белых и 5 черных шаров, последовательно без возвращения извлекают 8 шаров. Пусть $A_0^{(i)}$ $(A_1^{(i)})$ — событие, состоящее в том, что i-й шар был черный (белый). Найти условные вероятности:

1) $\mathsf{P}(A_1^{(5)} \mid A_1^{(1)}A_0^{(2)}A_0^{(3)}A_1^{(4)})$;

2) $\mathsf{P}(A_0^{(4)} \mid A_{\alpha_1}^{(1)}A_{\alpha_2}^{(2)}A_{\alpha_3}^{(3)})$, $(\alpha_1, \alpha_2, \alpha_3)=(0, 0, 1), (0, 1, 0), (1, 0, 0)$.

5. Доказать, что события A и $\bar{B}$ независимы, если независимы события A и B.

6. Случайная точка (ξ_1, ξ_2) имеет равномерное распределение в квадрате $\{(x_1, x_2)\colon 0\leqslant x_1\leqslant 1, 0\leqslant x_2\leqslant 1\}$. При каких значениях r независимы события

$$A_r=\{|\xi_1-\xi_2|\geqslant r\},\quad B_r=\{\xi_1+\xi_2\leqslant 3r\}?$$

7. Пусть

$$C_1=\left\{\xi_1\leqslant\frac{1}{2}\right\},\quad C_2=\left\{\xi_2\leqslant\frac{1}{2}\right\},\quad C_3=\left\{\left(\xi_1-\frac{1}{2}\right)\left(\xi_2-\frac{1}{2}\right)<0\right\},$$

где ξ_1, ξ_2 определены в задаче 6. Показать, что события C_1, C_2, C_3

попарно независимы. Являются ли события C_1, C_2, C_3 взаимно независимыми? Зависимы ли события C_1C_2 и C_3?

8. События A_1, A_2, A_3, A_4 взаимно независимы. Доказать взаимную независимость событий $\bar{A}_1A_2$ и A_3A_4.

9. События A_1, A_2, A_3, A_4 взаимно независимы: $\mathsf{P}(A_k) = p_k$, $k = 1, 2, 3, 4$. Найти вероятности событий: 1) $A_1\bar{A}_3A_4$; 2) $A_1 + A_2$; 3) $(A_1 + A_2)(A_3 + A_4)$.

10. Электрическая цепь составлена из элементов A_k, $k = 1, 2, 3$ (элементы A_1 и A_2 соединены параллельно, а A_3 присоединен к ним последовательно). При выходе из строя любого элемента цепь в месте его включения разрывается. Вероятность выхода из строя за данный период элемента A_k равна p_k, $k = 1, 2, 3$. Предполагается, что элементы выходят или не выходят из строя независимо друг от друга. Найти вероятность того, что за рассматриваемый период по цепи будет проходить ток.

11. Из урны, содержащей 3 белых и 5 черных шаров, два игрока по очереди вытащили по одному шару. Положим $A_k = \{k$-й игрок вытащил белый шар$\}$. Найти вероятности событий: A_1, A_2, A_1A_2.

Указание. Выбрать подходящим образом числовые значения вероятностей $\mathsf{P}(A_1)$, $\mathsf{P}(A_2 \mid A_1)$, $\mathsf{P}(A_2 \mid \bar{A}_1)$. Воспользоваться формулами (2.5).

12. Упрощенная система контроля изделий состоит из двух независимых проверок. В результате k-й проверки ($k = 1, 2$) изделие, удовлетворяющее стандарту, отбраковывается с вероятностью β_k, а бракованное изделие принимается с вероятностью α_k. Изделие принимается, если оно прошло обе проверки. Найти вероятности событий:

1) бракованное изделие будет принято;

2) изделие, удовлетворяющее стандарту, будет отбраковано.

13. В первой урне 2 белых и 4 черных шара, а во второй — 3 белых и 1 черный шар. Из первой урны переложили во вторую два шара. Найти вероятность того, что шар, вынутый из второй урны после перекладывания, окажется белым.

14. Изделия поступают на проверку, описанную в задаче 12. Предполагая, что каждое изделие удовлетворяет стандарту с вероятностью p, найти следующие вероятности:

1) вероятность того, что поступившее на проверку изделие не будет отбраковано;

2) вероятность того, что неотбракованное изделие удовлетворяет стандарту.

15. Предположим, что 5% всех мужчин и 0,25% всех женщин — дальтоники. Наугад выбранное лицо оказалось дальтоником. Какова вероятность, что это мужчина? (Считать, что количество мужчин и женщин одинаково.)

16. По каналу связи может быть передана одна из трех последовательностей букв: АААА, ВВВВ, СССС; известно, что вероятности каждой из последовательностей равны соответственно 0,3; 0,4; 0,3. В результате шумов буква принимается правильно с вероятностью 0,6. Вероятности приема переданной буквы за две другие равны 0,2 и 0,2. Предполагается, что буквы искажаются независимо друг от друга. Найти вероятность того, что передано АААА, если на приемном устройстве получено АВСА.

17. Вероятность того, что молекула, испытавшая в момент $t = 0$ столкновение с другой молекулой и не имевшая других столкновений до момента t, испытает столкновение в промежуток времени $(t, t+h)$, равна $\lambda h + o(h)$, $h \to 0$. Найти вероятность того, что время свободного пробега будет больше t.

18. На одну телефонную линию могут поступать вызовы двух типов: срочные и простые. При поступлении срочного вызова разговор по простому вызову прекращается. Вероятности поступления за время $(t, t+h)$ срочного и обычного вызовов равны соответственно $\alpha_1 h + o(h)$, $\alpha_2 h + o(h)$, $h \to 0$; вероятность прекращения любого разговора за время $(t, t+h)$ равна $\beta h + o(h)$. Пусть $P_0(t)$, $P_1(t)$, $P_2(t)$ — вероятности того, что в момент t линия свободна, занята срочным вызовом, занята простым вызовом. Написать для $P_k(t)$ дифференциальные уравнения и найти $\lim\limits_{t\to\infty} P_k(t) = \pi_k$, $k = 0, 1, 2$.

19. Изменим условия работы телефонной линии, описанной в примере 3 из § 3. Будем считать, что при занятой линии вызовы не теряются, а становятся в очередь. Обозначим $P_k(t)$ вероятность того, что в момент t один вызов обслуживается и $k-1$ образуют очередь ($k \geqslant 1$; $P_0(t)$ — вероятность того, что линия свободна). Составить для $P_k(t)$ дифференциальные уравнения. Найти $\lim\limits_{\to\infty} P_k(t) = \pi_k$, если $\theta = \alpha/\beta < 1$.

Последовательности испытаний

§ 1. Общее определение последовательности испытаний

Дадим сначала описание реальных испытаний в опыте с перекладыванием шаров в урнах, а затем построим математическую модель этих испытаний.

Пусть имеется три урны, содержащих белые и черные шары, одинаковые по форме. Состав шаров в первой урне: 2 белых и 3 черных; во второй — 2 белых и 2 черных; в третьей — 3 белых и 1 черный. Из первой урны случайно выбирается один шар и перекладывается во вторую. После этого из второй урны также случайно один шар перекладывается в третью урну и, наконец, из третьей какой-то из шаров перекладывается в первую. Таким образом, мы имеем три испытания, состоящих в извлечении шаров из трех урн. Предположим, что нас интересует, какой состав шаров в первой урне после перекладываний наиболее вероятен, а также — что вероятнее: изменение состава шаров в первой урне или сохранение.

Построим математическую модель этого опыта. Обозначим A_k реальное событие, состоящее в том, что при k-м перекладывании ($k = 1, 2, 3$) был переложен белый шар. Положим

$$\Omega = \{A_1A_2A_3, \bar{A}_1A_2A_3, A_1\bar{A}_2A_3, A_1A_2\bar{A}_3, \bar{A}_1\bar{A}_2\bar{A}_3, A_1\bar{A}_2\bar{A}_3, \bar{A}_1A_2\bar{A}_3, \bar{A}_1\bar{A}_2A_3\}.$$

Здесь, например, символом $A_1\bar{A}_2\bar{A}_3$ обозначено элементарное событие, соответствующее в реальном опыте перекладыванию из первой урны белого шара, а из двух других — черного. Событию A_1 реального опыта соответствует в математической модели следующее подмножество Ω:

$$A_1 = \{A_1A_2A_3, A_1\bar{A}_2A_3, A_1A_2\bar{A}_3, A_1A_2\bar{A}_3\}.$$

Аналогично определяются события $\tilde{A}_i$, $\bar{\tilde{A}}_i$, соответствующие реальным событиям A_i, $\bar{A}_i$. Ниже мы не будем использовать разные обозначения для реального события и соответствующего ему события модели, т. е. положим $\tilde{A}_i = A_i$. Тогда на введенные формально обозначения для элементарных событий можно смотреть как на произведения соответствующих событий. Если состав шаров в данной урне известен, то мы легко можем задать вероятность события, состоящего в извлечении из этой урны шара определенного цвета. В рассматриваемой задаче состав шаров в урне становится известным, если известно, какой шар в данную урну был переложен. Таким образом, мы можем считать, что заданы вероятности $\mathsf{P}(A_1)$, $\mathsf{P}(\bar{A}_1)$, $\mathsf{P}(A_2 \mid A_1)$, $\mathsf{P}(A_3 \mid A_1\bar{A}_2)$ и т. д. Например,

$$\mathsf{P}(A_1) = \frac{2}{5}, \quad \mathsf{P}(A_2 \mid A_1) = \frac{3}{5}, \quad \mathsf{P}(A_3 \mid A_1A_2) = \frac{4}{5}. \tag{1.1}$$

При определении второй и третьей вероятностей мы использовали то, что во вторую и третью урны был переложен белый шар и их состав стал соответственно: 3 белых, 2 черных; 4 белых, 1 черный. Аналогично можно приписать значения другим условным вероятностям. По заданным условным вероятностям так же, как в § 2 гл. 3 (см. (3.2.5)), мы можем полностью восстановить распределение вероятностей на подмножествах Ω. По формулам (3.2.2) и (1.1) находим

$$\mathsf{P}(A_1A_2A_3) = \mathsf{P}(A_1)\,\mathsf{P}(A_2 \mid A_1)\,\mathsf{P}(A_3 \mid A_1A_2) = \frac{2}{5}\cdot\frac{3}{5}\cdot\frac{4}{5}.$$

Аналогично приписываются вероятности другим элементарным событиям. Окончательно получим

$$\begin{aligned}
\mathsf{P}(A_1A_2A_3) &= \frac{24}{125}, & \mathsf{P}(\bar{A}_1\bar{A}_2\bar{A}_3) &= \frac{18}{125},\\
\mathsf{P}(\bar{A}_1A_2A_3) &= \frac{24}{125}, & \mathsf{P}(A_1\bar{A}_2\bar{A}_3) &= \frac{8}{125},\\
\mathsf{P}(A_1\bar{A}_2A_3) &= \frac{12}{125}, & \mathsf{P}(\bar{A}_1A_2\bar{A}_3) &= \frac{6}{125},\\
\mathsf{P}(A_1A_2\bar{A}_3) &= \frac{6}{125}, & \mathsf{P}(\bar{A}_1\bar{A}_2A_3) &= \frac{27}{125}.
\end{aligned} \tag{1.2}$$

По этим вероятностям однозначно определяется вероятность любого события.

Пусть, например,

$B_l =$ {после перекладываний в первой урне оказалось l белых шаров},

$l = 1, 2, 3$. Тогда

$$B_1 = \{A_1A_2\bar{A}_3,\ A_1\bar{A}_2\bar{A}_3\},$$
$$B_2 = \{A_1A_2A_3,\ \bar{A}_1\bar{A}_2A_3\},$$
$$B_3 = \{A_1A_2A_3,\ A_1\bar{A}_2A_3,\ \bar{A}_1A_2\bar{A}_3,\ \bar{A}_1\bar{A}_2A_3\}$$

и

$$\mathsf{P}(B_1) = \frac{14}{125},\quad \mathsf{P}(B_2) = \frac{60}{125},\quad \mathsf{P}(B_3) = \frac{51}{125}.$$

Дадим определение последовательности испытаний. Пусть множеством элементарных событий является множество

$$\Omega_n = \{(i_1i_2\ldots i_n):\ i_k = 1, 2, \ldots, N;\ k = 1, 2, \ldots, n\}. \tag{1.3}$$

Элементарное событие $\omega = (i_1i_2\ldots i_n)$ интерпретируется как цепочка исходов в n последовательных испытаниях, каждое из которых имеет N несовместных исходов: $1, 2, \ldots, N$. Положим

$$p(\omega) = p(i_1)p(i_2 \mid i_1)\ldots p(i_n \mid i_1\ldots i_{n-1}), \tag{1.4}$$

где $p(i_s \mid i_1\ldots i_{s-1}) \geqslant 0$,

$$\sum_{i_s=1}^{N} p(i_s \mid i_1\ldots i_{s-1}) = 1,\quad s = 1, \ldots, n, \tag{1.5}$$

$$i_k = 1, \ldots, N,\ 1 \leqslant k \leqslant n.$$

Тогда на подмножествах Ω_n однозначно определяется вероятность:

$$\mathsf{P}(A) = \sum_{\omega\in A} p(\omega),\quad A \subset \Omega_n. \tag{1.6}$$

Нужно только проверить (см. § 2 гл. 2), что

$$\sum_{\omega\in\Omega} p(\omega) = 1. \tag{1.7}$$

Для доказательства (1.7) заметим, что

$$\sum_{\omega\in\Omega} p(\omega) = \sum_{i_1\ldots i_n=1}^{N} p(i_1)\,p(i_2\,|\,i_1)\ldots p(i_n\,|\,i_1\ldots i_{n-1}) =$$

$$= \sum_{i_1\ldots i_{n-1}=1}^{N} p(i_1)\ldots p(i_{n-1}\,|\,i_1\ldots i_{n-2}) \sum_{i_n=1}^{N} p(i_n\,|\,i_1\ldots i_{n-1}) =$$

$$= \sum_{i_1,\ldots,i_{n-1}}^{N} p(i_1)\,p(i_2\,|\,i_1)\ldots p(i_n\,|\,i_1\ldots i_{n-1}),$$

так как $\sum_{i_n=1}^{N} p(i_n\,|\,i_1\ldots, i_{n-1}) = 1$ согласно (1.5). Далее мы можем выделить сумму по i_{n-1} и вновь воспользоваться (1.5). Повторив этот процесс n раз, получим (1.7). Таким образом, *мы определили вероятностное пространство, являющееся математической моделью последовательности из n испытаний.* Нетрудно проверить, что число $p(i_k\,|\,i_1\ldots i_{k-1})$ является условной вероятностью появления в k-м испытании исхода i_k при условии, что до этого была получена цепочка исходов $(i_1 i_2 \ldots i_{k-1})$.

Схему случайного выбора без возвращения, введенную в § 1 гл. 2, естественно определять как последовательность испытаний. Для этого условные вероятности $p(i_s\,|\,i_1\ldots i_{s-1})$ нужно задать следующим образом: если известен результат первых $s-1$ испытаний, то при s-м испытании с равными вероятностями может быть получено любое из оставшихся чисел. Модель случайного выбора без возвращения, сформулированная в терминах условных вероятностей, совпадает с классическим определением из § 1 гл. 2.

Решенная в начале параграфа задача является частным случаем общей схемы испытаний с $n = 3$, $N = 2$,

$$p(1) = p(2\,|\,1) = p(1\,|\,2) = p(2\,|\,12) = p(2\,|\,22) = 2/5,$$

$$p(2) = p(1\,|\,1) = p(2\,|\,2) = p(1\,|\,12) = p(1\,|\,22) = 3/5,$$

$$p(1\,|\,11) = p(1\,|\,21) = 4/5, \quad p(2\,|\,11) = p(2\,|\,21) = 1/5,$$

где исходы «1» и «2» соответствуют извлечению белого и черного шара.

В данной задаче оказалось, что вероятности $p(i\,|\,jk)$ не зависят от j. Это естественно, так как на состав последней урны влияет только цвет шара, переложенного из второй урны, и если этот цвет уже известен, то неважно, что было переложено во вторую урну из первой.

Последовательность испытаний, в которой условные вероятности $p(l_t \mid l_1 \ldots l_{t-1})$ не зависят от $l_1, \ldots, l_{t-2}$,

$$p(l_t \mid l_1 \ldots l_{t-1}) = p^{(t)}_{l_{t-1} l_t},$$

называется *цепью Маркова*. Более подробно такие испытания будут рассмотрены в гл. 8. В случае, когда $p(l_t \mid l_1 \ldots l_{t-1})$ не зависят от $l_1, \ldots, l_{t-1}$, последовательность испытаний называется *последовательностью независимых испытаний.*

§ 2. Последовательность независимых испытаний

По определению, данному в конце § 1, в случае независимых испытаний условные вероятности $p(l \mid l_1 \ldots \ldots l_{t-1})$ не зависят от $l_1, \ldots, l_{t-1}$, но могут зависеть от t. Если зависимость от t есть, то независимые испытания называют *неоднородными*, а в противном случае — *однородными*. Таким образом, *последовательность независимых однородных испытаний определяется равенствами* (1.3), (1.6) *и формулой*

$$p(\omega) = p_{i_1} p_{i_2} \ldots p_{i_n}, 1 \leqslant i_k \leqslant N, k = 1, \ldots, n, \qquad (2.1)$$

где $\sum\limits_{l=1}^{N} p_l = 1, p_l \geqslant 0, \; l = 1, 2, \ldots, N.$

Последовательность однородных независимых испытаний является математической моделью серии опытов, повторяющихся в одинаковых условиях. Вероятностную модель, определенную формулами (1.3), (1.6), (2.1), называют также *полиномиальной схемой.* Частный случай полиномиальной схемы с $N = 2$ называют *схемой Бернулли* или *испытаниями Бернулли*. Два исхода каждого испытания схемы Бернулли будем обозначать символами 1 и 0 или называть «успехом» и «неудачей», а соответствующие им вероятности в одном испытании обозначим буквами p и $q = 1 - p$.

Покажем, что любые события, связанные с разными испытаниями полиномиальной схемы, взаимно независимы. Пусть S_t $(t = 1, \ldots, n)$ — подмножества множества исходов отдельного испытания $\{1, 2, \ldots, N\}$. Обозначим $B^{(t)}_{S_t}$ событие, состоящее в том, что исход t-го испытания принадлежит множеству S_t.

Теорема 2.1. *При любых* S_t, $t = 1, \ldots, n$,

$$\mathsf{P}(B^{(1)}_{S_1} B^{(2)}_{S_2} \ldots B^{(n)}_{S_n}) = \mathsf{P}(B^{(1)}_{S_1}) \mathsf{P}(B^{(2)}_{S_2}) \ldots \mathsf{P}(B^{(n)}_{S_n}). \qquad (2.2)$$

Доказательство. Так как $\sum_{i=1}^{N} p_i = 1$,

$$B^{(t)}_{S_{(t)}} = \{(i_1 i_2 \ldots i_n):\ i_t \in S_t\},$$

$$B^{(1)}_{S_1} B^{(2)}_{S_2} \ldots B^{(n)}_{S_n} = \{(i_1 i_2 \ldots i_n):\ i_1 \in S_1, \ldots, i_n \in S_n\},$$

то из формул (1.5) и (2.1) получим

$$\mathsf{P}(B^{(1)}_{S_1} B^{(2)}_{S_2} \ldots B^{(n)}_{S_n}) = \sum_{i_1 \in S_1} \sum_{i_2 \in S_2} \ldots \sum_{i_n \in S_n} p_{i_1} p_{i_2} \ldots p_{i_n} = \\ = \sum_{i_1 \in S_1} p_{i_1} \sum_{i_2 \in S_2} p_{i_2} \ldots \sum_{i_n \in S_n} p_{i_n};$$

$$\mathsf{P}(B^{(1)}_{S_1}) = \sum_{i_1 \in S_1} p_{i_1} \sum_{i_2 = 1}^{N} p_{i_2} \ldots \sum_{i_n = 1}^{N} p_{i_n} = \sum_{i_1 \in S_1} p_{i_1}.$$

Аналогично проверяется, что при любых $B^{(t)}_{S_t}$

$$\mathsf{P}(B^{(t)}_{S_t}) = \sum_{l \in S_t} p_l. \tag{2.3}$$

Из доказанных равенств следует утверждение теоремы.

Положим $S_t = \{l\}$ в формуле (2.3). Тогда в (2.3) сумма сведется к одному слагаемому p_l, которое оказывается вероятностью l-го исхода в t-м испытании. Если в (2.2) положить $S_3 = S_4 = \ldots = S_n = \{1, 2, \ldots, N\}$, то получим равенство $\mathsf{P}(B^{(1)}_{S_1} B^{(2)}_{S_2}) = \mathsf{P}(B^{(1)}_{S_1}) \mathsf{P}(B^{(2)}_{S_2})$, так как $B^{(3)}_{S_3} = \ldots = B^{(n)}_{S_n} = \Omega$ и $B^{(1)}_{S_1} B^{(2)}_{S_2} \ldots B^{(n)}_{S_n} = B^{(1)}_{S_1} B^{(2)}_{S_2}$. Аналогично проверяется, что при любых $1 \leqslant t_1 < t_2 < \ldots < t_k \leqslant n$ имеем

$$\mathsf{P}(B^{(t_1)}_{S_{t_1}} \ldots B^{(t_k)}_{S_{t_k}}) = \mathsf{P}(B^{(t_1)}_{S_{t_1}}) \ldots \mathsf{P}(B^{(t_k)}_{S_{t_k}}). \tag{2.4}$$

Таким образом, мы показали, что при любых $S_1, \ldots, S_n$ события $B^{(1)}_{S_1} \ldots B^{(n)}_{S_n}$ взаимно независимы.

Во многих задачах мы часто каждому элементарному событию ставим в соответствие определенное действительное число, например: величину выигрыша, число исходов данного типа в полиномиальной схеме, число успехов в n испытаниях схемы Бернулли и т. д. Пусть $(\Omega, \mathfrak{A}, \mathsf{P})$ — вероятностное пространство. Назовем случайной величиной действительную функцию от элементарного события: $\xi = \xi(\omega)$, $\omega \in \Omega$. Для дискретных вероятностных пространств случайной величиной мы будем называть произвольную функцию от ω, а в общем случае, рассматриваемом в гл. 5, на функцию $\xi(\omega)$ будут налагаться дополнительные условия.

Для n испытаний схемы Бернулли элементарные события удобно обозначать цепочками длины n, составленными из букв У и Н: $\omega =$ УУН. . .ННУ. Обозначим $\mu_n = \mu_n$ (УУН. . .ННУ) случайную величину, равную числу успехов в первых n испытаниях Бернулли. Так, например, при $n = 4$ имеем

$$\mu_4 = \mu_4\,(\text{УУНУ}) = 3,\ \mu_4 = \mu_4\,(\text{НННН}) = 0$$

и т. д. Найдем вероятность событий

$$\{\mu_n = m\} = \{(\text{УУН}\ldots\text{У}).\ \mu_n\,(\text{УУН}\ldots\text{У}) = m\}. \tag{2.5}$$

Теорема 2.2. *Если μ_n — число успехов в n испытаниях Бернулли, то*

$$\mathsf{P}\,(\mu_n = m) = P_n\,(m) = C_n^m p^m q^{n-m}, \tag{2.6}$$

где $m = 0, 1, 2, \ldots, n$, $q = 1 - p$, p — вероятность успеха в отдельном испытании.

Доказательство. Каждая цепочка исходов, входящая в (2.5), содержит ровно m успехов и $n - m$ неудач. Тогда, воспользовавшись формулой (2.1) с $N = 2$, $p_1 = p$, $p_2 = q$, мы получим, что любая цепочка из множества (2.5) имеет одну и ту же вероятность $p^m q^{n-m}$. Различные цепочки в (2.5) отличаются только расположением успехов и неудач, так как общее число успехов фиксированно. Расположение успехов и неудач однозначно определяется выбором из n мест m мест для успехов. Это можно сделать C_n^m способами. Отсюда и из формулы (2.5) следует утверждение теоремы, так как каждая цепочка в (2.5) имеет вероятность $p^m q^{n-m}$.

Для полиномиальной схемы с произвольным N введем случайные величины ξ_k, равные числам исходов k, $k = 1, 2, \ldots, N$. Покажем, что

$$\mathsf{P}\,(\xi_1 = m_1, \xi_2 = m_2, \ldots, \xi_N = m_N) = \frac{n!}{m_1!\,m_2!\ldots m_N!}\, p_1^{m_1}\ldots p_N^{m_N}, \tag{2.7}$$

где $m_1, m_2, \ldots, m_N$ — неотрицательные целые числа, удовлетворяющие условию $m_1 + m_2 + \ldots + m_N = n$. Для остальных значений m_k левая часть (2.7) равна 0. В каждой цепочке $(l_1 l_2 \ldots l_n)$, удовлетворяющей условию

$(\xi_1 = m_1, \xi_2 = m_2, \ldots, \xi_N = m_N)$, «1» встречается m_1 раз, «2» — m_2 раз, ..., «N» — m_N раз. Следовательно, вероятность каждой такой цепочки равна $p_1^{m_1} p_2^{m_2} \ldots p_N^{m_N}$. Число цепочек заданного вида можно найти следующим образом: $C_n^{m_1}$ способами можно выбрать из n мест m_1 мест для «1»; из оставшихся $n - m_1$ мест можно выбрать $C_{n-m_1}^{m_2}$ способами места для «2» и т. д. Отсюда общее число цепочек нужного нам вида равно

$$C_n^{m_1} C_{n-m_1}^{m_2} \ldots C_{n-m_1-\ldots-m_{N-1}}^{m_N} = \frac{n!}{m_1! m_2! \ldots m_N!}.$$

Из приведенных рассуждений следует формула (2,7). При $N = 2$ формула (2.7) совпадает с (2.6).

Обозначим $A_1^{(t)}$ событие, состоящее в том, что в t-м испытании полиномиальной схемы появился исход N, и $A_0^{(t)} = \bar{A}_1^{(t)}$. Тогда, полагая сначала $S_t = \{N\}$, а затем $S_t = \{1, 2, \ldots, N-1\}$, по формуле (2.3) найдем

$$\mathsf{P}(A_1^{(t)}) = p_N, \quad \mathsf{P}(A_0^{(t)}) = p_1 + \ldots + p_{N-1} = 1 - p_N.$$

Отсюда и из (2.2) получим

$$\mathsf{P}(A_{\varepsilon_1}^{(1)} A_{\varepsilon_2}^{(2)} \ldots A_{\varepsilon_n}^{(n)}) = p_N^m (1 - p_N)^{n-m}, \quad \varepsilon_t = 0, 1,$$

если $\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n = m$. Таким образом, цепочки $A_{\varepsilon_1}^{(1)} A_{\varepsilon_2}^{(2)} \ldots A_{\varepsilon_n}^{(n)}$ можно принять за «элементарные события» в схеме Бернулли с $p_N = p$ и $p_1 + \ldots + p_{N-1} = q$.

При n подбрасываниях игральной кости мы имеем полиномиальную схему с $N = 6$ и $p_1 = p_2 = \ldots = p_6 = 1/6$. Если теперь «укрупнить» элементарные события, например различать в цепочках исходов в каждом испытании только «6» и «не 6», то получим схему Бернулли с $p = 1/6$, $q = 5/6$. Отправляясь от полиномиальной схемы с N исходами, можно при более мелком укрупнении получить новую полиномиальную схему с меньшим числом исходов.

Отметим еще, что полиномиальная схема с $p_1 = p_2 = \ldots = p_N = 1/N$ совпадает со схемой случайного выбора с возвращением, определенной в § 1 гл. 2 на основе классического определения. Таким образом, определенные в § 5 гл. 2 случайные числа можно также рассматривать как реализацию испытаний полиномиальной схемы с равновероятными исходами 0, 1, 2, ..., 9.

§ 3. Предельные теоремы в схеме Бернулли

В приложениях часто приходится вычислять вероятности различных событий, связанных с числом успехов в n испытаниях схемы Бернулли при больших значениях n. В этом случае вычисления по формуле (2.6) становятся затруднительными. Трудности возрастают, когда приходится еще суммировать вероятности (2.6). К суммированию (2.6) сводится вычисление вероятностей событий вида $\{a < \mu_n < b\}$. Действительно,

$$\mathsf{P}(a < \mu_n < b) = \sum_{a<m<b} \mathsf{P}(\mu_n = m) = \sum_{a<m<b} P_n(m). \quad (3.1)$$

Затруднения при вычислениях возникают также при малых значениях p или q.

Иногда при больших n удается заменить формулу (2.6) какой-либо приближенной асимптотической формулой. Приведем три предельные теоремы, содержащие асимптотические формулы для вероятностей (2.6) и (3.1) при $n \to \infty$.

Теорема 3.1 (теорема Пуассона). *Если* $n \to \infty$ *и* $p \to 0$ *так, что* $np \to \lambda$, $0 < \lambda < \infty$, *то*

$$\mathsf{P}(\mu_n = m) = C_n^m p^m q^{n-m} \to p_m(\lambda) = \frac{\lambda^m}{m!} e^{-\lambda}$$

при любом постоянном m, $m = 0, 1, 2, \ldots$

Доказательство. Положив $np = \lambda_n$, представим вероятность $\mathsf{P}(\mu_n = m)$ в виде

$$\mathsf{P}(\mu_n = m) = \frac{n(n-1)\ldots(n-m+1)}{m!}\left(\frac{\lambda_n}{n}\right)^m \left(1 - \frac{\lambda_n}{n}\right)^{n-m} =$$
$$= \frac{\lambda_n^m}{m!}\left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots$$
$$\ldots\left(1 - \frac{m-1}{n}\right)\left(1 - \frac{\lambda_n}{n}\right)^{-m}.$$

Отсюда при $n \to \infty$ получим утверждение теоремы.

Таким образом, при больших n и малых p мы можем воспользоваться приближенной формулой

$$\mathsf{P}(\mu_n = m) \approx \frac{\lambda_n^m}{m!} e^{-\lambda_n}, \ \lambda_n = np. \quad (3.2)$$

Известно (см. [2], гл. 5, § 4, теорема 8, стр. 116), что при любом множестве $B \subset \{0, 1, 2, \ldots\}$

$$\left|\mathsf{P}(\mu_n \in B) - \sum_{m \in B} p_m(np)\right| \leqslant np^2.$$

Приведенное неравенство верно для любых множеств B. При использовании приближенной формулы для индивидуальных вероятностей ошибки могут быть значительно меньше. Для сравнения точных и приближенных значений приведем следующую табличку:

m	0	1	2	3	4	5
$p_m(np)$	0,1353	0,2707	0,2707	0,1805	0,0902	0,0361
$\mathsf{P}\{\mu_{10} = m\}$	0,1074	0,2684	0,3020	0,2013	0,0880	0,0264
$\mathsf{P}\{\mu_{100} = m\}$	0,1326	0,2707	0,2734	0,1823	0,0902	0,0353

При $n = 10$ вероятности $\mathsf{P}\{\mu_{10} = m\}$ вычислены для $p = 1/5$ $(np^2 = 0{,}4)$, а при $n = 100$ — для $p = 1/50$ $(np^2 = 0{,}04)$.

Если мало значение q, то пуассоновским приближением можно воспользоваться для числа неудач. В случае, когда оба параметра p и q заметно отличны от 0, используются теоремы Муавра — Лапласа (локальная и интегральная).

Введем функции $\varphi(x)$ и $\Phi(x)$ следующими равенствами:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \Phi(x) = \int_{-\infty}^{x} \varphi(u)\, du.$$

Теорема 3.2 (локальная теорема Муавра — Лапласа). *Если $n \to \infty$, p $(0 < p < 1)$ постоянно, величина $x_m = (m - np)/\sqrt{npq}$ ограничена равномерно по m и n $(-\infty < a \leqslant x_m \leqslant b < +\infty)$, то*

$$\mathsf{P}(\mu_n = m) = \varphi(x_m)\,(1 + \alpha_n(m)) / \sqrt{npq},$$

где $|\alpha_n| < C/\sqrt{n}$ при $x_m \in [a, b]$, $C > 0$ — постоянная.

Доказательство. Коэффициент C_n^m в (2.6) запишем в виде

$$C_n^m = \frac{n!}{m!\,(n-m)!}. \tag{3.3}$$

Так как

$$m = np + x_m \sqrt{npq}, \quad n - m = nq - x_m \sqrt{npq},$$

то, воспользовавшись формулой Стирлинга

$$\ln n! = \ln \sqrt{2\pi n} + n \ln n - n + O\left(\frac{1}{n}\right), \tag{3.4}$$

получим

$$\begin{aligned} \ln m! &= \ln \sqrt{2\pi m} + (np + x_m \sqrt{npq}) \ln (np + x_m \sqrt{npq}) - \\ &\qquad - np - x_m \sqrt{npq} + O\left(\frac{1}{n}\right), \\ \ln (n-m)! &= \ln \sqrt{2\pi (n-m)} + \\ &\quad + (nq - x_m \sqrt{npq}) \ln (nq - x_m \sqrt{npq}) - \\ &\qquad - nq + x_m \sqrt{npq} + O\left(\frac{1}{n}\right). \end{aligned} \tag{3,5}$$

Логарифмы в (3.5) при помощи формулы $\ln (1 + x) = x - x^2/2 + O(x^3)$ запишем в виде

$$\begin{aligned} \ln (np + x_m \sqrt{npq}) &= \ln np + \ln \left(1 + x_m \sqrt{\frac{q}{np}}\right) = \\ &= \ln np + x_m \sqrt{\frac{q}{np}} - \frac{x_m^2}{2} \frac{q}{np} + O\left(\frac{q^3}{n \sqrt{n}\, p^3}\right), \\ \ln (nq - x_m \sqrt{npq}) &= \\ &= \ln nq - x_m \sqrt{\frac{p}{nq}} - \frac{x_m^2}{2} \frac{p}{nq} + O\left(\frac{p^3}{n \sqrt{n}\, q^3}\right). \end{aligned} \tag{3.6}$$

Утверждение теоремы следует из (2.6), (3.3) — (3.6).

При качественной оценке условий применимости приближенной формулы

$$\mathsf{P}(\mu_n = m) \approx \varphi(x_m)/\sqrt{npq} \tag{3,7}$$

нужно оценить величину остаточных членов в (3.6). При $n \to \infty$ сумма остаточных членов стремится к 0 при любых фиксированных p и q, $0 < p < 1$. Однако при конечных значениях n сумма остаточных членов может быть очень большой, если p или q малы. Хорошие приближения формула (3.7) дает при $p = q = 1/2$. Если в этом случае провести более точную оценку остаточного члена, то в формулировке теоремы можно заменить $|\alpha_n| < C/\sqrt{n}$ на $|\alpha_n| < C/n$. Формулу (3.7) часто используют при $n > 100$ и $npq > 20$. Указания о границах применимости формул

(3.2) и (3.7) являются очень приближенными и носят скорее качественный характер; к ним следует относиться с осторожностью.

Отметим еще, что в условиях теоремы 3.2 из того, что $n \to \infty$, следует стремление к бесконечности m. Значение m не должно отклоняться от np очень сильно; например, для $\mathsf{P}\ (\mu_n = 0)$ локальная теорема дает плохое приближение.

Теорема 3.3 (интегральная теорема Муавра—Лапласа). *Если p, $0 < p < 1$, постоянно, то при $n \to \infty$*

$$\mathsf{P}\left(a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right) - \frac{1}{\sqrt{2\pi}} \int_a^b c^{-\frac{x^2}{2}}\, dx \to 0$$

равномерно по a, b, $-\infty \leqslant a \leqslant b \leqslant +\infty$.

Доказательство. Докажем теорему сначала при фиксированных a, b, $-\infty < a \leqslant b < +\infty$. Очевидно, что вероятность события $\left(a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right)$ можно представить в виде

$$\mathsf{P}\ \left(a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right) = P_n(a, b) = \sum_{a \leqslant x_m \leqslant b} \mathsf{P}(\mu_n = m), \quad (3.8)$$

где $x_m = \frac{m - np}{\sqrt{npq}}$ и суммирование распространяется на значения m, для которых $x_m \in [a, b]$. Применяя к слагаемым (3.8) локальную теорему, получим

$$P_n\ (a,\ b) = S_n + T_n, \qquad (3.9)$$

где

$$S_n = \sum_{x_m \in [a, b]} \varphi(x_m) \frac{1}{\sqrt{npq}},$$

$$T_n = \sum_{x_m \in [a, b]} \alpha_n(m)\, \varphi(x_m) \frac{1}{\sqrt{npq}},$$

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Так как

$$\Delta x_m = x_{m+1} - x_m = \frac{m + 1 - np}{\sqrt{npq}} - \frac{m - np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}},$$

то S_n можно записать в виде суммы

$$S_n = \sum_{x_m \in [a,b]} \varphi(x_m)\,\Delta x_m,$$

которая отличается не более чем двумя слагаемыми от подходящим образом выбранной интегральной суммы, соответствующей интегралу $\int\limits_a^b \varphi(x)dx$. Следовательно,

$$\lim_{n\to\infty} S_n = \int\limits_a^b \varphi(x)\,dx. \tag{3.10}$$

Используя оценку для $\alpha_n(m)$ из локальной теоремы, получим

$$|T_n| \leqslant \sum_{x_m \in [a,b]} \varphi(x_m)\,\Delta x_m\,|\alpha_n(m)| \leqslant \frac{C}{\sqrt{n}} S_n.$$

Таким образом, при $n \to \infty$

$$T_n \to 0. \tag{3.11}$$

Утверждение теоремы для постоянных a, b следует из формул (3.8)—(3.11).

Из доказанного и из теоремы 2.4 гл. 7 следует равномерная сходимость. Приближенная формула

$$\mathsf{P}\left(a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right) \approx \int\limits_a^b \varphi(x)\,dx = \Phi(b) - \Phi(a) \tag{3.12}$$

используется в тех случаях, когда возможно использование (3.7). Численное значение интеграла можно найти, если воспользоваться таблицей 2 для функции

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{u^2}{2}}\,du. \tag{3.13}$$

При небольших значениях npq приближенную формулу (3.12) нужно заменить следующей формулой (см. [18], гл. 7, § 2, стр. 189):

$$\mathsf{P}\left(a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right) \approx \Phi\left(b + \frac{1}{2\sqrt{npq}}\right) - \Phi\left(a - \frac{1}{2\sqrt{npq}}\right). \tag{3.14}$$

Приведем таблички точных и приближенных значений вероятностей событий $\{m_1 \leqslant \mu_n \leqslant m_2\}$. Положим

$$P(m_1, m_2) = P\{m_1 \leqslant \mu_n \leqslant m_2\} = P\left\{x_1 \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant x_2\right\},$$

$$P_0(m_1, m_2) = \Phi(x_2) - \Phi(x_1),$$

$$P_1(m_1, m_2) = \Phi\left(x_2 + \frac{h}{2}\right) - \Phi\left(x_1 - \frac{h}{2}\right),$$

где $h = 1/\sqrt{npq}$, $x_1 = (m_1 - np)h$, $x_2 = (m_2 - np)h$. Используя таблицу 2 для $\Phi_0(x)$ и формулу (2.6), получим:

а) $n = 10$

p	m_1, m_2	0; 2	3; 5	6; 8	9; 10
0,2	$P(m_1, m_2)$ $P_0(m_1, m_2)$ $P_1(m_1, m_2)$	0,6778 0,4429 0,6316	0,3158 0,2059 0,3455	0,0064 0,0007 0,0029	0,0000 0,0000 0,0000
0,5	$P(m_1, m_2)$ $P_0(m_1, m_2)$ $P_1(m_1, m_2)$	0,0547 0,0277 0,0558	0,5684 0,3980 0,5696	0,3662 0,2356 0,3651	0,0107 0,0057 0,0133

б) $n = 100$, $p = 0,2$

m_1, m_2	0; 9	10; 15	16; 20	21; 25	26; 30	31; 40
$P(m_1, m_2)$	0,0023	0,1262	0,4310	0,3531	0,0814	0,0061
$P_0(m_1, m_2)$	0,0030	0,0994	0,3413	0,2957	0,0606	0,0030
$P_1(m_1, m_2)$	0,0020	0,1250	0,4223	0,3480	0,0795	0,0043

в) $n = 100$, $p = 0,5$

m_1, m_2	30; 35	36; 40	41; 45	46; 50
$P(m_1, m_2)$	0,0017	0,0267	0,1557	0,3557
$P_0(m_1, m_2)$	0,0013	0,0202	0,1227	0,2881
$P_1(m_1, m_2)$	0,0018	0,0269	0,1563	0,3558

Формула (3.12) позволяет оценить близость частоты и вероятности. Пусть p — вероятность успеха в схеме Бернулли и μ_n — общее число успехов. Частотой успеха называют отношение μ_n/n. Оценим вероятность события $\left(\left|\frac{\mu_n}{n}-p\right|<\Delta\right)$. Если n достаточно велико, то можно воспользоваться формулой (3.12). Тогда

$$\mathsf{P}\left(\left|\frac{\mu_n}{n}-p\right|<\Delta\right)=\mathsf{P}\left(-\Delta\sqrt{\frac{n}{pq}}<\frac{\mu_n-np}{\sqrt{npq}}<\Delta\sqrt{\frac{n}{pq}}\right)\approx$$

$$\approx\frac{1}{\sqrt{2\pi}}\int\limits_{-\Delta\sqrt{\frac{n}{pq}}}^{\Delta\sqrt{\frac{n}{pq}}} e^{-\frac{x^2}{2}}\,dx=2\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right), \qquad (3.15)$$

так как функция $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ четная. Значение $\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right)$ находится по таблице в конце книги.

Часто возникает обратная задача: сколько нужно провести испытаний, чтобы частота μ_n/n отличалась от вероятности p не больше чем на Δ с вероятностью $1-2\alpha$ (α мало)? Такого типа задачи возникают при использовании метода Монте-Карло (метод статистических испытаний). Идея метода заключается в моделировании случайного процесса или последовательности испытаний, вероятностные характеристики которых просто связаны с подлежащими вычислению величинами. Если много раз бросать иглу, как описано в задаче Бюффона (см. § 3 гл. 2), то частота μ_n/n будет мало отличаться от вероятности p пересечения иглой какой-либо линии. Зная величину отклонения μ_n/n от p, можно оценить ошибку в определении числа π. В таких задачах естественно считать p неизвестным. Тогда, чтобы подобрать наименьшее n, при котором вероятность отклонения будет равна $1-2\alpha$, нужно, согласно (3.14), решить уравнение

$$2\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right)=1-2\alpha.$$

Решение будет зависеть от неизвестного p. От этой зависимости можно избавиться, если потребовать, чтобы

$$\mathsf{P}\left(\left|\frac{\mu_n}{n}-p\right|<\Delta\right)\geqslant 1-2\alpha,$$

Тогда из (3.15), используя неравенство $pq \leqslant 1/4$, получим

$$\mathsf{P}\left(\left|\frac{\mu_n}{n} - p\right| < \Delta\right) \approx 2\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right) \geqslant 2\Phi_0(2\Delta\sqrt{n}) = 1 - 2\alpha$$

и для определения n будем иметь уравнение $\Phi_0(2\Delta\sqrt{n}) = (1-2\alpha)/2$. По таблице 3 можно найти u_α, для которых $\Phi_0(u_\alpha) = (1-2\alpha)/2$. Тогда $2\Delta\sqrt{n} = u_\alpha$ и $n \geqslant u_\alpha^2/4\Delta^2$. Довольно часто используются значения 2α, равные 0,05 и 0,01. Для этих значений соответствующие u_α равны 1,960 и 2,576.

В задаче с иглой вовсе не обязательно проводить реальные подбрасывания иглы. Можно этот процесс моделировать; например, используя случайные или псевдослучайные числа, получить на ЭВМ последовательность величин (φ_k, x_k), $k = 1, 2, \ldots, n$ (см. § 3 гл. 2), задающих положение иглы, и вычислить частоту пересечений.

§ 4. Бесконечные последовательности независимых испытаний

Во многих интересных задачах приходится рассматривать бесконечные последовательности испытаний. Ограничимся рассмотрением последовательности испытаний Бернулли. Успехи и неудачи мы будем обозначать 1 и 0 соответственно. Положим

$$\Omega = \{(i_1 i_2 \ldots i_n \ldots) : i_k \in \{0; 1\},\ k = 1, 2, \ldots\}. \tag{4.1}$$

Обозначим $\mathfrak{A}$ минимальную σ-алгебру событий, порождаемую событиями вида

$$A_{\varepsilon_1 \ldots \varepsilon_n}^{t_1 \ldots t_n} = \{(i_1 i_2 \ldots i_n \ldots) : i_{t_1} = \varepsilon_1, \ldots, i_{t_n} = \varepsilon_n\}. \tag{4.2}$$

В случае независимых испытаний

$$\mathsf{P}\,(A_{\varepsilon_1 \ldots \varepsilon_n}^{t_1 \ldots t_n}) = p_{\varepsilon_1} p_{\varepsilon_2} \ldots p_{\varepsilon_n}, \tag{4.3}$$

где p_0, p_1 — вероятности исходов 0 и 1 в отдельном испытании ($p_0, p_1 \geqslant 0$, $p_0 + p_1 = 1$). Отметим без доказательства *), что равенства (4.3) однозначно определяют вероятность на рассматриваемой σ-алгебре $\mathfrak{A}$, порожденной множествами (4.2). Таким образом, утверждается лишь существование вероятности; явной формулы,

---

*) Построение вероятностных пространств для простейших схем было проведено в гл. 2. Для бесконечных последовательностей соответствующие построения являются более сложными (см. А. Н. Колмогоров [8]).

подобной формулам гл. 2, для вычисления вероятности произвольного события в рассматриваемом случае нет. Вероятности событий A_n, наступление которых определяется не более чем за n испытаний ($n = 1, 2, 3, \ldots$), можно вычислять по формуле (4.3). Вероятности событий, которые можно представить через монотонные последовательности событий A_n, вычисляются по формуле (3.11).

Рассмотрим задачу о случайном блуждании, связанную с бесконечными последовательностями испытаний. Пусть по целым точкам отрезка $[0, n]$ движется частица. Обозначим ξ_t ее координату в момент времени t ($t = 0, 1, 2, \ldots$). Движением частицы управляет бесконечная последовательность независимых испытаний с двумя исходами 1 и -1. Положим

$$\xi_{t+1} = \begin{cases} \xi_t + 1, & \text{если в } t\text{-м испытании появилась } 1 \quad \text{и } \xi_t < n; \\ \xi_t - 1, & \text{если в } t\text{-м испытании появилась } -1 \quad \text{и } \xi_t > 0. \end{cases}$$

Если $\xi_t = 0$ или $\xi_t = n$ при некотором t, то частица навсегда остается в этих точках. Будем предполагать, что $\xi_0 = k$.

Обозначим p и q вероятности исходов 1 и -1 соответственно.

Пусть A — событие, состоящее в том, что частица когда-нибудь попадет в точку n. Нас будет интересовать вероятность $\pi_{kn} = \mathsf{P}(A)$.

Сформулированную задачу можно интерпретировать как задачу о разорении игрока. Пусть в начале игры 1-й игрок имеет k рублей, а второй игрок — $n - k$ рублей. Если при бросании монеты (или кости) выпал герб (или «6»), то первый игрок получает 1 рубль (это соответствует движению частицы вправо), а в противном случае отдает 1 рубль (движение частицы влево). Поглощению частицы на правом конце соответствует выигрыш первого игрока.

Положим $\pi_{kn}(t) = \mathsf{P}(\xi_t = n)$. Для рассматриваемого блуждания довольно очевидно, что условная вероятность события $\{\xi_{t+1} = n\}$ при условии, что первым переходом частицы был переход из k в $k + 1$, равна безусловной вероятности события $\{\xi_t = n\}$ в схеме блуждания, начинающейся из точки $k + 1$, т. е.

$$\mathsf{P}(\xi_{t+1} = n \mid \xi_1 = k + 1) = \pi_{k+1,n}(t), \tag{4.4}$$

и, аналогично,

$$\mathsf{P}(\xi_{t+1} = n \mid \xi_1 = k - 1) = \pi_{k-1,n}(t). \tag{4.5}$$

Докажем, например, равенство (4.4). Обозначим $B(\varepsilon_1\varepsilon_2 \ldots \varepsilon_t)$ событие, состоящее в том, что $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_t$ являются исходами первых t испытаний ($\varepsilon_s = -1; 1$, $s = 1, \ldots, t$). Положим $p(\varepsilon_s) = p$, если $\varepsilon_s = 1$, $p(\varepsilon_s) = q$, если $\varepsilon_s = -1$. Пусть частица начинает блуждание из точки k. Обозначим $C(l, v)$ множество последовательностей $u_1, u_2, \ldots, u_v$, удовлетворяющих условиям: 1) $u_i = -1$ или $u_i = 1$ ($i = 1, \ldots, v$); 2) существует s, $1 \leqslant s \leqslant v$, такое, что $l + u_1 + \ldots + u_s = n$ и $0 < l + u_1 + \ldots + u_{s_1} < n$ ($1 \leqslant s_1 < s$).

В момент времени t частица окажется в точке n, если $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_t) \in C(k, t)$. Отметим еще, что $\mathsf{P}(B(\varepsilon_1 \ldots \varepsilon_t)) = \prod_{s=1}^{t} p(\varepsilon_s)$.

Используя сделанные замечания, для частицы, начинающей блуждания из точки k, получим

$$\mathsf{P}\{(\xi_{t+1}=n)\cap(\xi_1=k+1)\}=\sum\nolimits^{*}\mathsf{P}(B(\varepsilon_2\ldots\varepsilon_{t+1}))=p\sum\nolimits^{*}\prod_{s=2}^{t+1}p(\varepsilon_s),$$

$$\mathsf{P}\{\xi_1=k+1\}=\mathsf{P}\{B(1)\}=p, \tag{4.6}$$

$$\mathsf{P}\{\xi_{t+1}=n\mid\xi_1=k+1\}=\sum\nolimits^{*}\prod_{s=2}^{t+1}p(\varepsilon_s).$$

Здесь символом Σ^* обозначено суммирование по всем цепочкам $(\varepsilon_2,\ldots,\varepsilon_{t+1})\in C(k+1,t)$. Если частица начинает блуждание из $k+1$, то

$$\pi_{k+1,n}(t)=\mathsf{P}\{\xi_t=n\}=\sum\nolimits^{**}\prod_{s=1}^{t}p(\varepsilon_s), \tag{4.7}$$

где символом Σ^{**} обозначено суммирование по всем $(\varepsilon_1,\ldots,\varepsilon_t)\in\in C(k+1,t)$. Очевидно, что правые части (4.6) и (4.7) совпадают. Равенство (4.4) доказано.

По формуле полной вероятности

$$\mathsf{P}(\xi_{t+1}=n)=\mathsf{P}(\xi_1=k+1)\,\mathsf{P}(\xi_{t+1}=n\mid\xi_1=k+1)+\\+\mathsf{P}(\xi_1=k-1)\,\mathsf{P}(\xi_{t+1}=n\mid\xi_1=k-1). \tag{4.8}$$

Так как $\xi_0=k$, то $\mathsf{P}(\xi_1=k+1)=p$, $\mathsf{P}(\xi_1=k-1)=q$. Используя эти равенства и равенства (4.4), (4.5), запишем (4.8) в виде

$$\pi_{kn}(t+1)=p\pi_{k+1,n}(t)+q\pi_{k+1,n}(t),\quad k=1,2,\ldots,n-1. \tag{4.9}$$

Заметим, что

$$(\xi_1=n)\subset(\xi_2=n)\subset\ldots\subset(\xi_t=n)\subset(\xi_{t+1}=n)\subset\ldots$$

и $A=\bigcup_{t=1}^{\infty}(\xi_t=n)$. Таким образом, по формуле (1.3.11)

$$\pi_{kn}=\mathsf{P}(A)=\lim_{t\to\infty}\mathsf{P}(\xi_t=n)=\lim_{t\to\infty}\pi_{kn}(t).$$

Переходя к пределу при $t\to\infty$ в (4.9), получим

$$\pi_{kn}=p\pi_{k+1,n}+q\pi_{k-1,n},\quad k=1,2,\ldots,n-1.$$

Так как π_{kn} — вероятность поглощения на правом конце в процессе, начавшемся из точки k, то $\pi_{0n}=0$, $\pi_{nn}=1$. Таким образом, вероятность π_{kn}, рассматриваемая как функция от k, является решением линейного однородного уравнения в конечных разностях с постоянными коэффициентами

$$pf_{k+1}-f_k+qf_{k-1}=0, \tag{4.10}$$

удовлетворяющим граничным условиям $f_0=0$, $f_1=1$. Теория таких уравнений (см. А. О. Гельфонд [3]) во многом аналогична теории линейных дифференциальных уравнений с постоянными коэффициентами.

Пусть сначала $p\neq q$. Подставив $f_k=\lambda^k$ в уравнение (4.10), получим $p\lambda^{k+1}-\lambda^k+q\lambda^{k-1}=0$ или $p\lambda^2-\lambda+q=0$. Корнями этого уравнения являются $\lambda_1 = 1$, $\lambda_2 = q/p$. Следовательно, функции λ_1^k и λ_2^k удовлетворяют (4.10). Линейная комбинация

$$f_k = C_1\lambda_1^k + C_2\lambda_2^k \tag{4.11}$$

при любых C_1, C_2 тоже является решением. Используя условия $f_0 = 0$ и $f_1 = 1$, из (4.11) при $k = 0$ и $k = n$ получим

$$C_1 + C_2 = 0, \quad C_1 + (q/p)^n C_2 = 1.$$

Отсюда и из (4.11) находим

$$\pi_{kn} = (1 - (q/p)^k)/(1 - (q/p)^n). \tag{4.12}$$

Вероятности поглощения на левом конце π_{k0} тоже удовлетворяют (4.10), но из общего решения (4.11) нужно выделить решение f_k с $f_0 = 1$ и $f_n = 0$. Определяя из этих условий C_1 и C_2, получим

$$\pi_{k0} = ((q/p)^k - (q/p)^n/ (1 - (q/p)^n).$$

Так как $\pi_{kn} + \pi_{k0} = 1$, то блуждание закончится с вероятностью 1.

Пусть теперь $p = q = 1/2$. В этом случае $\lambda_1 = \lambda_2 = 1$ и решение (4.10) нужно искать в виде

$$f_k = C_1 + kC_2.$$

Используя граничные условия, получим

$$\pi_{kn} = \frac{k}{n}, \quad \pi_{k0} = 1 - \frac{k}{n}. \tag{4.13}$$

В схеме блуждания по целым точкам прямой с поглощением только в нуле (в n поглощение не происходит) вероятность попасть когда-нибудь в n равна вероятности π_{kn}, вычисленной для схемы с поглощением в нуле и в n. Вероятность того, что частица побывает во всех точках правее k, равна

$$\pi_k = \lim_{n\to\infty} \pi_{kn} = \begin{cases} 0, & \text{если } q \geqslant p, \\ 1 - (q/p)^k, & \text{если } q < p. \end{cases}$$

Этот результат не противоречит интуитивным представлениям. Если движение вправо более вероятно, чем влево, то с положительной вероятностью частица может уйти вправо; в противном случае с вероятностью 1 колебания ограничены и происходит поглощение в нуле.

Задачи к главе 4

1. Два игрока поочередно извлекают шары (без возвращения) из урны, содержащей 2 белых и 4 черных шара. Выигрывает тот, кто первым вынет белый шар. Найти вероятность выигрыша участника, начавшего игру.

2. Два игрока поочередно извлекают шары (без возвращения) из урны, содержащей 2 белых шара, 4 черных и 1 красный. Выигрывает тот, кто первым вынет белый шар. Если появляется красный шар, то объявляется ничья. Пусть $A_1 = \{$выиграет игрок, начавший игру$\}$, $A_2 = \{$выиграет второй участник$\}$, $B = \{$игра закончится вничью$\}$. Найти $\mathsf{P}(A_1)$, $\mathsf{P}(A_2)$, $\mathsf{P}(B)$.

3. Из урны, содержащей M белых и $N - M$ черных шаров, по одному без возвращения извлекают все шары. Найти вероятности событий:

$$A_k = \{k\text{-й шар белый}\},$$

$$B_{k,l} = \{k\text{-й и } l\text{-й шары белые}\},$$

$$C_{k,l} = \{k\text{-й шар черный, а } l\text{-й — белый}\}.$$

Указание. Пусть $A_0^{(i)}$ $(A_1^{(i)})$ — событие, состоящее в том, что 1-й шар черный (белый). Из равновероятности элементарных событий $A_{\varepsilon_1}^{(1)}A_{\varepsilon_2}^{(2)} \ldots A_{\varepsilon_N}^{(N)}$ $(\varepsilon_1 + \ldots + \varepsilon_N = M,\ \varepsilon_i = 0, 1)$ следует, что $\mathsf{P}(A_k) = \mathsf{P}(A_1)$, $\mathsf{P}(B_{k,l}) = \mathsf{P}(B_{1,2})$, $\mathsf{P}(C_{k,l}) = \mathsf{P}(C_{1,2})$. Воспользоваться равенствами $B_{1,2} = A_1^{(1)}A_1^{(2)}$, $C_{1,2} = A_0^{(1)}A_1^{(2)}$ и определением случайного выбора в терминах условных вероятностей (см. § 1).

4. Проведено 10 независимых испытаний, каждое из которых заключается в одновременном подбрасывании трех игральных костей. Найти вероятность того, что в четырех испытаниях появятся в точности по две «6».

5. При передаче сообщения вероятность искажения одного знака равна 1/100. В предположении независимости искажения знаков найти вероятность того, что сообщение из 5 знаков: а) не будет искажено; б) содержит ровно одно искажение; в) содержит хотя бы два искажения.

6. Сколько нужно взять случайных чисел, чтобы число «6» появилось хотя бы один раз с вероятностью, не меньшей а) 0,7; б) 0,9?

7. В некоторой лотерее вероятность выигрыша на один билет равна 1/5. Предполагая, что выигрыши на различные билеты независимы, определить число билетов n, которое нужно купить, чтобы вероятность $Q(n)$ получения хотя бы одного выигрыша была не меньше а) 0,65; б) 0,9; в) 0,99.

Найти также $Q(n)$ в этих случаях.

8. Среди $5M$ билетов M выигрышных. Найти вероятность $Q(n)$ того, что среди n купленных билетов есть хотя бы один выигрышный. Вычислить $Q(n)$ при 1) $M = 3$; 2) $M = 10$ для n, определенных в задаче 7 в случаях а), б), в).

9. Найти вероятность того, что в n испытаниях схемы Бернулли с вероятностью успеха p появятся $m + l$ успехов, причем l успехов появятся в l последних испытаниях.

10. Двое бросают монету по n раз. Найти вероятность того, что у них выпадет одинаковое число гербов.

11. Из множества $S = \{1, 2, \ldots, n\}$ выбирается подмножество $\mathscr{A}_1$ так, что каждый элемент из S независимо от остальных с вероятностью p включается в множество $\mathscr{A}_1$ и с вероятностью $q = 1 - p$ не включается. Аналогичным образом независимо от $\mathscr{A}_1$ выбирается подмножество $\mathscr{A}_2$. Найти вероятности событий: а) $\mathscr{A}_1 \cap \mathscr{A}_2 = \varnothing$; б) множество $\mathscr{A}_1 \cap \mathscr{A}_2$ состоит из k $(k = 0, 1, 2, \ldots, n)$ элементов (т. е. $|\mathscr{A}_1 \cap \mathscr{A}_2| = k$); в) $|\mathscr{A}_1| > |\mathscr{A}_2|$ при $p = q = 1/2$.

Указание. в) Найти $\mathsf{P}\{|\mathscr{A}_1| = |\mathscr{A}_2|\}$.

12. На отрезок $[0, 10]$ наудачу брошено 5 точек. Найти вероятность того, что две точки попадут в $[0, 2]$, одна — в $[2, 3]$, две — в $[3, 10]$.

13. В круг вписан квадрат. Какова вероятность того, что из 10 точек, брошенных наудачу в круг, четыре попадут в квадрат, три — в один сегмент и по одной — в оставшиеся три сегмента?

14. Пусть ξ_k — число появлений исхода k в n испытаниях полиномиальной схемы. Найти: 1) $\mathsf{P}(\xi_1 = m_1)$; 2) $\mathsf{P}(\xi_2 = m_2, \ldots, \xi_N = m_N \mid \xi_1 = m_1)$, если p_k, $k = 1, \ldots, N$,— вероятность появления исхода k в одном испытании.

Указание. 1) Просуммировать (2.7) или использовать взаимную независимость событий $B_{\varepsilon_1}^{(1)}, B_{\varepsilon_2}^{(2)}, \ldots, B_{\varepsilon_n}^{(n)}$, где $\varepsilon_i = 0, 1$,

$$B_1^{(l)} = \{\text{в } l\text{-м испытании появился исход } 1\},$$

$$B_0^{(l)} = \{\text{в } l\text{-м испытании исход 1 не появился}\},$$

2) Использовать 1) и формулу (2.7).

15. Из множества $S = \{1, 2, \ldots, n\}$ выбираются два подмножества $\mathscr{A}_1$ и $\mathscr{A}_2$ так, что каждый элемент из S независимо от остальных с вероятностью p_1 включается только в $\mathscr{A}_1$, с вероятностью p_2 — в оба подмножества, с вероятностью p_3 — только в $\mathscr{A}_2$ и с вероятностью $p_4 = 1 - p_1 - p_2 - p_3$ не включается ни в одно из подмножеств. Найти вероятности событий: а) $\mathscr{A}_1 \cap \mathscr{A}_2 = \varnothing$; б) $|\mathscr{A}_1 \cap \mathscr{A}_2| = k$; в) $|\mathscr{A}_1| = |\mathscr{A}_2| = 1$.

16. В таблице случайных чисел цифры сгруппировали по две. Найти приближенное значение вероятности того, что среди 200 пар пара 01 встретится не менее трех раз.

17. Брошено 6 правильных игральных костей. Какова вероятность выпадения: 1) хотя бы одной; 2) ровно одной; 3) ровно двух единиц? Найти точные значения и сравнить их со значениями, вычисленными по формуле Пуассона.

18. Сколько изюма должна в среднем содержать булочка, чтобы вероятность иметь хотя бы одну изюмину в булочке была не менее 0,99?

19. Вероятность попадания в цель при каждом выстреле равна 0,001. Найти вероятность попадания в цель двумя и более выстрелами при залпе в 5000 выстрелов.

20. Найти приближенное выражение того, что число выпадений «1» при 12 000 бросаний игральной кости заключено между 1900 и 2150.

21. В поселке A 2500 жителей. Каждый из них примерно 6 раз в месяц ездит на поезде в город B, выбирая дни поездок по случайным мотивам независимо от остальных жителей. Какой наименьшей вместительностью должен обладать поезд, чтобы он переполнялся в среднем не чаще одного раза в 100 дней? (Поезд идет раз в сутки.)

22. Из таблицы случайных чисел отбирают числа, делящиеся на 4, до тех пор, пока не наберется 588. Найти приближенное значение вероятности того, что потребуется таблица, содержащая не меньше 2100 чисел.

Указание. Пусть ν — число чисел в требуемой таблице; μ_n — число чисел, делящихся на 4, среди n чисел таблицы ($n = 2100$). Использовать равенство $\mathsf{P}(\nu \geqslant n) = \mathsf{P}(\mu_n \leqslant 588)$.

23. Рыбак забросил спиннинг 100 раз. Какова вероятность того, что он поймал хотя бы одну рыбу, если одна рыба приходится в среднем на 200 забрасываний?

24. Две монеты бросают до тех пор, пока не выпадет герб хотя бы на одной из них. Найти вероятность того, что будет проведено n бросаний ($n = 1, 2, \ldots$).

25. Монету бросают до тех пор, пока два раза подряд она не выпадет одной и той же стороной. Найти вероятности событий: 1) опыт закончится не более чем за 4 бросания; 2) опыт закончится за четное число бросаний.

26. Две игральные кости бросаются до тех пор, пока впервые на двух костях не выпадет сумма очков, меньшая шести. Какова вероятность того, что при последнем бросании сумма очков не меньше трех?

27. Используя табл. 7, выписать реализацию бросаний симметричной монеты. Длину реализации n подобрать так, чтобы частота выпадений герба отличалась от 1/2 не более чем на $\Delta = 0{,}1$ с вероятностью, примерно равной 0,9. По реализации вычислить частоту выпадений герба.

28. Используя табл. 7, выписать N реализаций блуждания частицы (см. § 4) по целым точкам отрезка $[0, n]$ до ее поглощения в 0 или в n. Величины N, n, k — начальное положение частицы, p — вероятность перехода вправо, $q = 1 - p$ выбрать равными:

$$\text{а) } N = 10,\ n = 6,\ p = 1/6,\ k = 1;$$

$$\text{б) } N = 10,\ n = 6,\ p = 1/2,\ k = 1.$$

Вычислить частоту поглощения частицы в 0 и в n; сравнить с вероятностями соответствующих поглощений (см. (4.12), (4.13)).

Случайные величины

§ 1. Определения и примеры

В § 2 гл. 4 было дано определение случайной величины для дискретных вероятностных пространств. Пусть теперь $(\Omega, \mathfrak{A}, \mathsf{P})$ — произвольное вероятностное пространство. Случайной величиной ξ назовем действительную функцию $\xi = \xi(\omega)$, $\omega \in \Omega$, такую, что при любом действительном x

$$\{\omega\colon\ \xi(\omega) < x\} \in \mathfrak{A}. \tag{1.1}$$

Если в $\mathfrak{A}$ включаются все подмножества Ω, то (1.1) очевидно выполняется. Событие в (1.1) более коротко иногда будем записывать в виде $(\xi < x)$. Так как операции над событиями из $\mathfrak{A}$ не выводят за пределы $\mathfrak{A}$, то из (1.1) следует, что

$$\begin{gathered}(\xi \geqslant x) = \overline{(\xi < x)} \in \mathfrak{A},\\ (x_1 \leqslant \xi < x_2) = (\xi < x_2) \setminus (\xi < x_1) \in \mathfrak{A}, \\ (\xi = x) = \bigcap_{n=1}^{\infty}\left(x \leqslant \xi < x + \frac{1}{n}\right) \in \mathfrak{A}.\end{gathered} \tag{1.2}$$

Таким образом, вероятности этих событий определены. Для вычисления вероятностей событий вида (1.2) достаточно при любом x знать вероятность

$$F_\xi(x) = \mathsf{P}(\xi < x). \tag{1.3}$$

Функция $F_\xi(x)$ действительной переменной x, $-\infty < x < \infty$, называется *функцией распределения* случайной величины ξ. Так как

$$(\xi < x_2) = (x_1 \leqslant \xi < x_2) + (\xi < x_1),$$

то согласно аксиоме А4

$$\mathsf{P}(\xi < x_2) = \mathsf{P}(x_1 \leqslant \xi < x_2) + \mathsf{P}(\xi < x_1).$$

Отсюда и из (1.3) находим

$$\mathsf{P}(x_1 \leqslant \xi < x_2) = F_\xi(x_2) - F_\xi(x_1). \tag{1.4}$$

По формуле (1.3.6) $\mathsf{P}(\xi \geqslant x) = 1 - F_\xi(x)$. Из (1.3.11), (1.4) и (1.2) получим

$$\mathsf{P}(\xi = x) = \lim_{n\to\infty}\left(F_\xi\left(x + \frac{1}{n}\right) - F_\xi(x)\right) = \\ = F_\xi(x+0) - F_\xi(x). \qquad (1.5)$$

По функции распределения можно вычислить вероятности любых событий вида (1.2).

Иногда вместо $F_\xi(x)$ будем писать просто $F(x)$.

Пример 1. Два игрока подбрасывают монету. Если при данном подбрасывании выпал «герб», то первый игрок получает 1 рубль, а если выпала «решетка», то отдает 1 рубль. Для описания данной игры естественно положить $\Omega = \{\Gamma, \mathrm{P}\}$ и $\mathsf{P}(\{\Gamma\}) = \mathsf{P}(\{\mathrm{P}\}) = 1/2$. Случайная величина ξ, равная выигрышу первого игрока, определяется следующим образом: $\xi = \xi(\Gamma) = 1$, $\xi = \xi(\mathrm{P}) = -1$.

Легко вычисляется функция распределения $F_\xi(x)$ величины ξ. Если $x \leqslant -1$, то множество $\{\xi < x\}$ является пустым и его вероятность равна 0. Если $-1 < x \leqslant 1$, то $\{\xi < x\} = \{\mathrm{P}\}$ и, следовательно, $F_\xi(x) = \mathsf{P}(\xi < x) = 1/2$. При $x > 1$ имеем $\{\xi < x\} = \{\Gamma, \mathrm{P}\}$ и $\mathsf{P}(\xi < x) = 1$. Таким образом,

$$F_\xi(x) = \begin{cases} 0, & \text{если} \quad x \leqslant -1, \\ 1/2, & \text{если} \quad -1 < x \leqslant 1, \\ 1, & \text{если} \quad x > 1. \end{cases}$$

Пример 2. Пусть точка A равномерно распределена в квадрате $\Omega = \{(u, v)\colon 0 \leqslant u \leqslant 1, 0 \leqslant v \leqslant 1\}$. Элементарными событиями ω являются точки квадрата Ω; множество событий $\mathfrak{A}$ здесь порождается квадрируемыми подмножествами квадрата. Вероятность события, являющегося квадрируемым подмножеством, равна его площади. Случайными величинами являются, например, функции $\xi = \xi(\omega) = \sqrt{u^2 + v^2}$ — расстояние брошенной точки от начала координат, $\eta = \eta(u, v) = u$ — первая координата брошенной точки и т. д. Найдем функцию распределения величины η. При $0 < x < 1$ имеем

$$\{(u, v)\colon \xi(u, v) < x\} = \{(u, v)\colon u < x\}.$$

Стороны этого прямоугольника равны 1 и x. Таким образом,

$$\mathsf{P}(\{(u, v): u < x\}) = x$$

и

$$F_\eta(x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ x, & \text{если } 0 < x \leqslant 1, \\ 1, & \text{если } x > 1. \end{cases}$$

Пример 3. Рассмотрим следующий опыт. Пусть один раз подбрасывается монета. Если выпал «герб», то на этом опыт заканчивается. Если выпала «решетка», то на отрезок [0, 1] наудачу бросается точка. Для описания этого опыта введем следующее пространство элементарных событий:

$$\Omega = \{\text{Г}; (\text{Р}, u)\},$$

где $0 \leqslant u \leqslant 1$; σ-алгебра $\mathfrak{A}$ пусть порождается событиями

$$\{\text{Г}\}; \{(\text{Р}, u): a \leqslant u < b\},$$

где $0 \leqslant a \leqslant b \leqslant 1$. Положим

$$\mathsf{P}(\{\text{Г}\}) = \frac{1}{2}, \ \mathsf{P}(\{(\text{Р}, u): a \leqslant u < b\}) = \frac{b-a}{2}.$$

По вероятностям этих событий однозначно определяется вероятность на $\mathfrak{A}$. Рассмотрим случайную величину ζ, определяемую равенствами

$$\zeta = \zeta(\text{Г}) = -1, \ \zeta = \zeta(\text{Р}, u) = u.$$

Так как

$$(\zeta < x) = \begin{cases} \varnothing, & \text{если } x \leqslant -1, \\ \{\text{Г}\}, & \text{если } -1 < x \leqslant 0, \\ \{\text{Г}\} + \{(\text{Р}, u): 0 \leqslant u < x\}, & \text{если } 0 < x < 1, \end{cases}$$

$$F_\zeta(x) = \begin{cases} 0, & \text{если } x \leqslant -1, \\ \dfrac{1}{2}, & \text{если } -1 < x \leqslant 0, \\ \dfrac{x+1}{2}, & \text{если } 0 < x \leqslant 1, \\ 1, & \text{если } x > 1. \end{cases}$$

§ 2. Свойства функции распределения

Пусть $F(x)$ — функция распределения некоторой случайной величины ξ. Формула (1.3) определяет $F(x)$ при любом действительном x. При $x = +\infty$ и при $x = -\infty$ положим $F(-\infty) = 0$, $F(+\infty) = 1$.

Теорема 2.1. *Функция распределения $F(x)$ обладает следующими свойствами:*

1. *Если $x_1 \leqslant x_2$, то $F(x_1) \leqslant F(x_2)$.*
2. $\lim\limits_{x\to-\infty} F(x) = F(-\infty) = 0$; $\lim\limits_{x\to+\infty} F(x) = F(+\infty) = 1$.
3. $\lim\limits_{x\to x_0-0} F(x) = F(x_0)$ (*непрерывность слева*).

Доказательство.

1. Так как $(\xi < x_1) \subset (\xi < x_2)$ при $x_1 \leqslant x_2$, то неравенство $F(x_1) \leqslant F(x_2)$ следует из (1.3.9).

2. Так как последовательность событий $\{\xi < -n\}$, $n = 1, 2, \ldots$, монотонно убывает (т. е. $\{\xi < -n\} \supset \supset \{\xi < -n-1\}$, $n = 1, 2, \ldots$) и $\bigcap\limits_{n=1}^{\infty} \{\xi < -n\} = \varnothing$, то по формуле (1.3.3) получим

$$\lim_{n\to+\infty} F(-n) = \lim_{n\to+\infty} \mathsf{P}(\xi < -n) = \mathsf{P}(\varnothing) = 0.$$

Отсюда с учетом доказанной монотонности функции $F(x)$ получаем, что $\lim\limits_{x\to-\infty} F(x) = 0$. Второе соотношение во втором свойстве доказывается аналогично с использованием монотонной последовательности $(\xi < n)$, $n = 1, 2, \ldots$

3. Пусть числовая последовательность y_n возрастает и $\lim\limits_{n\to\infty} y_n = x_0$. Тогда

$$(\xi < y_n) \subset (\xi < y_{n+1}), \quad \bigcup_{n=1}^{\infty} (\xi < y_n) = (\xi < x_0)$$

и по формуле (1.3.11)

$$\lim_{n\to\infty} \mathsf{P}(\xi < y_n) = \mathsf{P}(\xi < x_0).$$

Отсюда с учетом монотонности $F(x)$ получим третье свойство. Теорема доказана.

Иногда, вместо (1.3), функцию $F(x)$ определяют по формуле $F(x) = \mathsf{P}(\xi \leqslant x)$. При таком определении функция $F(x)$ непрерывна справа: $\lim\limits_{x\to x_0+0} F(x) = F(x_0)$.

Любая функция $G(x)$, обладающая тремя свойствами, указанными в теореме 2.1, является функцией распределения некоторой случайной величины, т. е. можно построить вероятностное пространство $(\Omega, \mathfrak{A}, \mathsf{P})$ и определить на нем случайную величину ξ такую, что $F_\xi(x) = G(x)$. Положим $\Omega = \{u\colon -\infty < u < \infty\}$. Обозначим $\mathfrak{A}_0$ алгебру, порожденную полуинтервалами $[u_1, u_2)$. На этих полуинтервалах зададим вероятность с помощью равенства

$$\mathsf{P}([u_1, u_2)) = G(u_2) - G(u_1). \tag{2.1}$$

Эта формула однозначно определяет вероятность на алгебре $\mathfrak{A}_0$*). Если мы теперь положим $\xi = \xi(u) = u$, то $F_\xi(x) = \mathsf{P}(\xi < x) = \mathsf{P}\{(-\infty, x)\} = \lim\limits_{n\to\infty}[G(x) - G(-n)] = G(x)$. При изучении отдельной случайной величины можно заменить исходное вероятностное пространство $(\Omega, \mathfrak{A}, \mathsf{P})$ на вероятностное пространство, в котором множеством элементарных событий является числовая прямая $\widetilde{\Omega}$ (т. е. множество значений случайной величины), а вероятность $\widetilde{\mathsf{P}}$ на подмножествах $\widetilde{\Omega}$ определяется формулой (2.1) с $G(x) = F_\xi(x)$. Вероятность $\widetilde{\mathsf{P}}$ называется *законом распределения случайной величины* или просто *распределением* ξ.

В приложениях иногда оказывается удобным описывать результаты опытов или измерений в терминах случайных величин. В этом случае при построении математической модели в качестве пространства элементарных событий можно выбрать множество значений случайной величины, и если случайная величина одна, то, как было отмечено выше, достаточно задать ее функцию распределения.

§ 3. Дискретные и абсолютно непрерывные распределения

Отметим два важных частных случая законов распределений. Закон распределения случайной величины ξ называется *дискретным*, если существует конечное или счетное множество чисел $x_1, x_2, \ldots, x_n \ldots$ (без точек накопления) таких, что

$$\mathsf{P}(\xi = x_n) = p_n > 0, \quad n = 1, 2, \ldots; \ \sum_{n=1}^{\infty} p_n = 1.$$

---

*) По теореме о продолжении вероятности мы можем единственным образом продолжить заданную на $\mathfrak{A}_0$ вероятность на минимальную σ-алгебру $\mathfrak{A}$, порожденную $\mathfrak{A}_0$.

Дискретный закон распределения полностью определяется указанием значений x_n, $n = 1, 2, \ldots$, и вероятностей p_n, с которыми эти значения принимает случайная величина. Случайная величина, имеющая дискретный закон распределения, тоже называется *дискретной*. Отметим также, что для любого конечного или счетного набора пар (x_n, π_n), $n = 1, 2, \ldots$, где x_n — попарно различные действительные числа и

$$\pi_n > 0, \quad n = 1, 2, \ldots; \ \sum_{n=1}^{\infty} \pi_n = 1,$$

можно указать дискретную случайную величину ξ такую, что $\mathsf{P}(\xi = x_n) = \pi_n$.

Действительно, пусть $\Omega = \{x_1, x_2, \ldots, x_n, \ldots\}$. Для любого подмножества $A \subset \Omega$ положим $\mathsf{P}(A) = \sum\limits_{n:\, x_n \in A} \pi_n$. Тогда для $\xi = \xi(x_n) = x_n$ имеем $\mathsf{P}(\xi = x_n) = \pi_n$.

Закон распределения случайной величины ξ называется *абсолютно непрерывным*, если существует неотрицательная функция $p_\xi(x)$ такая, что при любом x

$$F_\xi(x) = \mathsf{P}(\xi < x) = \int_{-\infty}^{x} p_\xi(u)\, du.$$

Случайная величина, имеющая абсолютно непрерывное распределение, называется *абсолютно непрерывной*.

Ниже мы будем всегда предполагать, что $p_\xi(x)$ непрерывна всюду, за исключением конечного числа точек. Функция $p_\xi(x)$ называется *плотностью распределения вероятностей*. Очевидно,

$$\mathsf{P}(a \leqslant \xi < b) = \int_a^b p_\xi(x)\, dx,$$

$$\mathsf{P}(\xi = a) = \lim_{n\to\infty} \mathsf{P}\left(a \leqslant \xi < a + \frac{1}{n}\right) = \lim_{n\to\infty} \int_a^{a+1/n} p_\xi(x)\, dx = 0. \tag{3.1}$$

Отсюда следует, что для абсолютно непрерывных величин

$$\mathsf{P}(a \leqslant \xi \leqslant b) = \mathsf{P}(a < \xi \leqslant b) = \mathsf{P}(a \leqslant \xi < b) = \mathsf{P}(a < \xi < b).$$

Если x является точкой непрерывности $p_\xi(x)$, то при $\Delta x \to 0$

$$\mathsf{P}(x < \xi < x + \Delta x) = p_\xi(x)\, \Delta x + o(\Delta x).$$

Это равенство следует из (3.1). Плотность распределения, обладает следующими свойствами:

1) $p_\xi(x) \geqslant 0, \; -\infty < x < \infty$;

2) $\int\limits_{-\infty}^{\infty} p_\xi(x)\,dx = 1$;

3) $F'_\xi(x) = p_\xi(x)$ в точках непрерывности $p_\xi(x)$.

Плотность распределения полностью определяет распределение случайной величины. Функция распределения абсолютно непрерывной величины, очевидно, непрерывна.

Пример 3 из § 1 показывает, что существуют распределения, не принадлежащие ни одному из указанных типов.

Нетрудно проверить, что любая неотрицательная функция $p(x)$, интеграл от которой по всей числовой прямой равен 1, является плотностью распределения некоторой случайной величины. Действительно, если в определении абсолютно непрерывного вероятностного пространства (см. § 4 гл. 2) положить $n = 1$, $\pi(u_1) = p(u_1)$, то плотность распределения случайной величины $\xi = \xi(u_1) = u_1$ совпадает с $p(x)$.

Приведем несколько часто встречающихся законов распределения. Сначала перечислим некоторые дискретные распределения.

1. *Вырожденное распределение.*

$$\mathsf{P}(\xi = a) = 1, \; a \text{ — постоянная.}$$

2. *Гипергеометрическое распределение* (N, M, n — натуральные числа, $M \leqslant N$, $n \leqslant N$):

$$\mathsf{P}(\xi = m) = C_M^m C_{N-M}^{n-m} / C_N^n, \; m = 0, 1, \ldots, \min(M, n).$$

3. *Биномиальное распределение* (n — натуральное число, $0 < p < 1$):

$$\mathsf{P}(\xi = k) = C_n^k p^k (1-p)^{n-k}, \; k = 0, 1, \ldots, n.$$

4. *Распределение Пуассона с параметром* $\lambda > 0$:

$$\mathsf{P}(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots$$

5. *Геометрическое распределение* ($0 < p < 1$):

$$\mathsf{P}(\xi = k) = (1-p)^{k-1} p, \; k = 1, 2, \ldots$$

Перечислим теперь некоторые абсолютно непрерывные распределения, указав их плотность $p(x)$.

1. *Равномерное распределение на отрезке* $[a, b]$, $a < b$:

$$p(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

2. *Нормальное* (или *гауссовское*) *распределение с параметрами* (a, σ^2) $(\sigma > 0, \ -\infty < a < \infty)$:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Нормальное распределение с параметрами $(0, 1)$ называют *стандартным нормальным распределением*.

3. *Показательное распределение с параметром* $\lambda > 0$:

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

Случайные величины с указанными выше распределениями часто появляются естественным образом в разных задачах. Так, в примере 1 (см. § 1 гл. 2) случайная величина, равная числу шаров в выборке, имеет гипергеометрическое распределение (см. (2.1.5)). В § 2 гл. 4 была введена случайная величина μ_n, равная числу успехов в n первых испытаниях схемы Бернулли. Эта величина имеет биномиальное распределение (см. (4.2.6)).

Для бесконечной последовательности испытаний Бернулли (см. § 4 гл. 4, (4.4.1)—(4.4.3)) определим случайную величину τ, равную числу испытаний до первого успеха (или до первого исхода 1) включительно. Тогда $(\tau = k) = A_0^{(1)} A_0^{(2)} \ldots A_0^{(k-1)} A_1^{(k)}$, где события $A_{\varepsilon_t}^{(t)}$ определены формулой (4.4.2). Отсюда и из формулы (4.4.3) получим $\mathsf{P}(\tau = k) = (1-p)^{k-1} p$. Таким образом, величина τ имеет геометрическое распределение.

Нормальное распределение и распределение Пуассона в § 3 гл. 4 были использованы в качестве распределений, приближающих биномиальное распределение.

§ 4. Совместные распределения нескольких случайных величин

Пусть на вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$ заданы случайные величины

$$\xi_1 = \xi_1(\omega), \ \xi_2 = \xi_2(\omega), \ \ldots, \ \xi_r = \xi_r(\omega), \ \omega \in \Omega.$$

Каждому ω эти величины ставят в соответствие r-мерный вектор. *Совместной функцией распределения* (или *многомерной функцией распределения*) величин $\xi_1, \ldots, \xi_r$ (или случайного вектора $\xi = (\xi_1, \ldots, \xi_r)$) называется вероятность

$$F_\xi(x) = F_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r) = \mathsf{P}(\xi_1 < x_1, \ldots, \xi_r < x_r),$$

рассматриваемая как функция точки $x = (x_1, \ldots, x_r)$ r-мерного евклидова пространства R^r. Функция распределения $F_\xi(x)$ однозначно определяет вероятности $\mathsf{P}(\xi \in B)$ для любых параллелепипедов $B \subset R^r$, а следовательно, и для достаточно широкого класса подмножеств R^r. Так, при $r = 2$ и $B = \{(x_1, x_2)\colon a_1 \leqslant x_1 < a_2, b_1 \leqslant x_2 < b_2\}$

$$\mathsf{P}\{(\xi_1, \xi_2) \in B\} = \\ = F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1), \quad (4.1)$$

где $F(x_1, x_2) = F_{\xi_1\xi_2}(x_1, x_2)$. Действительно, по формуле (1.3.7)

$$\mathsf{P}\{(\xi_1 < a_1, \xi_2 < b_2) \cup (\xi_1 < a_2, \xi_2 < b_1)\} = \\ = \mathsf{P}\{\xi_1 < a_1, \xi_2 < b_2\} + \mathsf{P}\{\xi_1 < a_2, \xi_2 < b_1\} - \\ - \mathsf{P}\{\xi_1 < a_1, \xi_2 < b_1\} = F(a_1, b_2) + F(a_2, b_1) - F(a_1, b_1).$$

Отсюда и из равенства

$$\{\xi_1 < a_2, \xi_2 < b_2\} = \{(\xi_1, \xi_2) \in B\} \cup \\ \cup (\xi_1 < a_1, \xi_2 < b_2\} \cup \{\xi_1 < a_2, \xi_2 < b_1\}$$

следует (4.1).

По многомерным функциям распределения можно найти одномерные распределения отдельных координат. Например, так же, как в доказательстве теоремы 2.1, проверяется, что

$$\lim_{y \to +\infty} F_{\xi_1\xi_2}(x, y) = F_{\xi_1\xi_2}(x, +\infty) = F_{\xi_1}(x),$$
$$\lim_{y \to +\infty} F_{\xi_1\xi_2}(y, x) = F_{\xi_1\xi_2}(+\infty, x) = F_{\xi_2}(x).$$

Аналогично одномерному случаю определяется r-мерное дискретное распределение. Случайный вектор $\xi = (\xi_1, \ldots, \xi_r)$ называется *дискретным*, если существует множество $\{x(1), x(2), \ldots, x(k), \ldots\}$, $x(k) \in R^r$, не имеющее точек накопления, такое, что

$$\mathsf{P}(\xi = x(k)\} = p_k > 0, \quad \sum_{k=1}^{\infty} p_k = 1.$$

Тогда для любого множества $B \subset R^r$

$$\mathsf{P}\{\xi \in B\} = \sum_{x(k) \in B} \mathsf{P}\{\xi = x(k)\}. \qquad (4.2)$$

Случайный вектор называется *абсолютно непрерывным*, если существует такая неотрицательная функция $p_\xi(x) = p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r)$, называемая *совместной* (или *r-мерной*) *плотностью распределения*, что для любого параллелепипеда B

$$\mathsf{P}\{\xi \in B\} = \int \underset{B}{\ldots} \int p_\xi(x_1, \ldots, x_r)\, dx_1 \ldots dx_r. \qquad (4.3)$$

Формула (4.3) сохраняется для любых квадрируемых или кубируемых множеств $B \subset R^r$. Из формулы (4.3) следует, что интеграл от плотности $p_\xi(x)$ по всему пространству R^r равен 1.

Найдем одномерные распределения для двумерных абсолютно непрерывных и дискретных векторов. Аналогичные формулы могут быть получены для векторов произвольной размерности.

Пусть (ξ_1, ξ_2) — абсолютно непрерывный вектор. Тогда, полагая в (4.3) $r = 2$ и

$$B = \{(u, v)\colon -\infty < u < x, \ -\infty < v < \infty\},$$

получим

$$F_{\xi_1}(x) = \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} p_{\xi_1\xi_2}(u, v)\, dv \right) du = \int_{-\infty}^{x} p_{\xi_1}(u)\, du,$$

где

$$p_{\xi_1}(x) = \int_{-\infty}^{\infty} p_{\xi_1\xi_2}(x, v)\, dv \qquad (4.4)$$

— плотность распределения величины ξ_1. Аналогично находим, что

$$p_{\xi_2}(x) = \int_{-\infty}^{\infty} p_{\xi_1\xi_2}(u, x)\, du. \qquad (4.5)$$

Пусть (ξ_1, ξ_2) — дискретный вектор и

$$\mathsf{P}\,(\xi_1 = x_{1i}, \xi_2 = x_{2j}) = p_{ij} \geqslant 0;\ i, j = 1, 2, \ldots;$$

$$\sum_{i,\,j=1}^{\infty} p_{ij} = 1. \tag{4.6}$$

Тогда

$$\mathsf{P}\,\{\xi_1 = x_{1i}\} = \sum_{j=1}^{\infty} \mathsf{P}\,\{\xi_1 = x_{1i}, \xi_2 = x_{2j}\} = \sum_{j=1}^{\infty} p_{ij}.$$

Пусть

$$p_{i\cdot} = \sum_{j=1}^{\infty} p_{ij}, \quad p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij}. \tag{4.7}$$

В этих обозначениях

$$\mathsf{P}\,\{\xi_1 = x_{1i}\} = p_{i\cdot}, \quad \mathsf{P}\,\{\xi_2 = x_{2j}\} = p_{\cdot j}. \tag{4.8}$$

Рассмотрим два примера, показывающих, что по одномерным распределениям (4.7) без дополнительной информации нельзя восстановить совместное распределение случайных величин (4.6).

Пример 1. Пусть $\Omega = \{(1, 1), (-1, 1), (1, -1), (-1, -1)\}$ и все элементарные события равновероятны. Положим

$$\xi_1 = \xi_1\,(i, j) = i, \quad \xi_2 = \xi_2\,(i, j) = j, \quad i, j = -1, 1.$$

Очевидно, при любой паре (i, j)

$$\mathsf{P}\,\{\xi_1 = i,\ \xi_2 = j\} = 1/4 \tag{4.9}$$

и

$$\mathsf{P}\,\{\xi_1 = i\} = \mathsf{P}\,\{\xi_1 = i,\ \xi_2 = -1\} + \\ + \mathsf{P}\,\{\xi_1 = i,\ \xi_2 = 1\} = 1/4 + 1/4 = 1/2,\ i = -1, 1.$$

Аналогично получим $\mathsf{P}\,\{\xi_2 = j\} = 1/2$, $j = -1, 1$.

Пример 2. Пусть Ω, ξ_1 и ξ_2 такие же, как в примере 1, а

$$\mathsf{P}\,\{(1, 1)\} = \mathsf{P}\,\{(-1, -1)\} = 1/2,$$
$$\mathsf{P}\,\{(-1, 1)\} = \mathsf{P}\,\{(1, -1)\} = 0.$$

Тогда, например, $\mathsf{P}\,\{\xi_1 = -1,\ \xi_2 = 1\} = 0$, и, следовательно, совместное распределение ξ_1, ξ_2 не совпадает с (4.9). Одномерные распределения

$$\mathsf{P}\,\{\xi_1 = -1\} = \mathsf{P}\,\{\xi_1 = -1,\ \xi_2 = -1\} = 1/2,$$
$$\mathsf{P}\,\{\xi_1 = 1\} = \mathsf{P}\,\{\xi_1 = 1,\ \xi_2 = 1\} = 1/2$$

и

$$\mathsf{P}\{\xi_2 = -1\} = \mathsf{P}\{\xi_2 = 1\} = 1/2$$

совпадают с соответствующими распределениями предыдущего примера.

§ 5. Независимость случайных величин

Случайные величины $\xi_1, \xi_2, \ldots, \xi_r$ называются *независимыми*, если для любых действительных $x_1, x_2, \ldots, x_r$

$$F_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r) = F_{\xi_1}(x_1) \ldots F_{\xi_r}(x_r). \tag{5.1}$$

Часто удобнее использовать следующее эквивалентное определение независимости: для любых событий $\{\xi_k \in B_k\}$, $k = 1, \ldots, r$, где $B_1, \ldots, B_r$ — подмножества числовой прямой, имеет место равенство

$$\mathsf{P}\{\xi_1 \in B_1, \ldots, \xi_r \in B_r\} = \mathsf{P}\{\xi_1 \in B_1\} \ldots \mathsf{P}\{\xi_r \in B_r\}. \tag{5.2}$$

Если положить $B_k = (-\infty, x_k)$, то (5.1) следует из (5.2). Покажем, что при $r = 2$ из (5.1) следует (5.2) для любых полуинтервалов $B_1 = [a_1, a_2)$, $B_2 = [b_1, b_2)$.

Теорема 5.1. *Если при любых* x_1, x_2

$$F_{\xi_1\xi_2}(x_1, x_2) = F_{\xi_1}(x_1)\, F_{\xi_2}(x_2), \tag{5.3}$$

то при любых $a_1 < a_2$, $b_1 < b_2$

$$\mathsf{P}\{\xi_1 \in [a_1, a_2),\ \xi_2 \in [b_1, b_2)\} = \\ = \mathsf{P}\{\xi_1 \in [a_1, a_2)\}\, \mathsf{P}\{\xi_2 \in [b_1, b_2)\}. \tag{5.4}$$

Доказательство. Заменим в правой части формулы (4.1) совместные функции распределения по формуле (5.3): $F_{\xi_1\xi_2}(a_i, b_j) = F_{\xi_1}(a_i)\, F_{\xi_2}(b_j)$, $i, j = 1, 2$. Разложив полученное выражение на множители, найдем

$$\mathsf{P}\{\xi_1 \in [a_1, a_2), \xi_2 \in [b_1, b_2)\} = \\ = (F_{\xi_1}(a_2) - F_{\xi_1}(a_1))\,(F_{\xi_2}(b_2) - F_{\xi_2}(b_1)).$$

Отсюда и из формулы (4.1) следует (5.1). Теорема доказана.

Можно показать *), что из равенства (5.4), доказанного для полуинтервалов, следует равенство (5.2) при $r = 2$ для любых событий $\{\xi_k \in B_k\}$, $k = 1, 2$.

---

*) Для этого нужно воспользоваться теоремой о продолжении вероятности (см., § 3 гл. 1).

Условие независимости (5.4) удобно использовать для установления условий независимости дискретных и абсолютно непрерывных распределений. Ограничимся установлением условий независимости в двух частных случаях.

Теорема 5.2. *Пусть распределение величин* ξ_1, ξ_2 *дискретно*, $\sum_{i,\,j=1}^{\infty} \mathsf{P}\{\xi_1 = x_{1i},\ \xi_2 = x_{2j}\} = 1$ *и множества* $(x_{11}, x_{12}, \ldots)$ *и* $(x_{21}, x_{22}, \ldots)$ *не имеют предельных точек на числовой прямой. Случайные величины* ξ_1, ξ_2 *независимы тогда и только тогда, когда при любых* i, j

$$\mathsf{P}\{\xi_1 = x_{1i},\ \xi_2 = x_{2j}\} = \mathsf{P}\{\xi_1 = x_{1i}\}\,\mathsf{P}\{\xi_2 = x_{2j}\}. \quad (5.5)$$

Доказательство. *Необходимость.* Пусть для величин ξ_1, ξ_2 выполняется условие (5.4). Для каждой точки $(x_{1i},\ x_{2j})$ выберем прямоугольник $\{(x_1,\ x_2)\colon\ a_1 < x_1 < a_i',\ b_j < x_2 < b_j'\}$ такой, что $x_{1i} \in (a_i,\ a_i')$, $x_{2j} \in (b_j,\ b_j')$, а другие значения величин ξ_1, ξ_2 не входят в эти интервалы. Тогда из условия (5.4) следует (5.5), так как $\mathsf{P}\{a_i < \xi_1 < a_i',\ \ b_j < \xi_2 < b_j'\} = \mathsf{P}\{\xi_1 = x_{1i},\ \ \xi_2 = x_{2j}\}$, $\mathsf{P}\{a_i < \xi_1 < a_i'\} = \mathsf{P}\{\xi_1 = x_{1i}\}$, $\mathsf{P}\{b_j < \xi_2 < b_j'\} = \mathsf{P}\{\xi_2 = x_{2j}\}$.

Достаточность. Пусть выполнено условие (5.5). В полуинтервалы $[a_1,\ a_2)$ и $[b_1,\ b_2)$ попадут некоторые из чисел $\{x_{11},\ x_{12}, \ldots\}$, $\{x_{21},\ x_{22}, \ldots\}$. Положим

$$X_1 = \{i\colon\ x_{1i} \in [a_1,\ a_2)\},\quad X_2 = \{i\colon\ x_{2i} \in [b_1,\ b_2)\}.$$

Из формулы (4.2), используя условие (5.5), получим

$$\begin{aligned}
\mathsf{P}\{\xi_1 \in [a_1, a_2), \xi_2 \in [b_1, b_2)\} &= \\
&= \sum_{i\in X_1,\ j\in X_2} \mathsf{P}\{\xi_1 = x_{1i}, \xi_2 = x_{2j}\} = \\
&= \sum_{i\in X_1,\ j\in X_2} \mathsf{P}\{\xi_1 = x_{1i}\}\,\mathsf{P}\{\xi_2 = x_{2j}\} = \\
&= \sum_{i\in X_1} \mathsf{P}\{\xi_1 = x_{1i}\} \sum_{j\in X_2} \mathsf{P}\{\xi_2 = x_{2j}\} = \\
&= \mathsf{P}\{\xi_1 \in [a_1, a_2)\}\,\mathsf{P}\{\xi_2 \in [b_1, b_2)\},
\end{aligned}$$

т. е. выполнено общее достаточное условие независимости (5.4). Теорема доказана.

В примере 1 из § 4 случайные величины ξ_1, ξ_2 независимы, так как равенства $\mathsf{P}\{\xi_1 = i,\ \xi_2 = j\} = \mathsf{P}\{\xi_1 = i\}\,\mathsf{P}\{\xi_2 = j\}$ выполняются при любых i, j. В примере 2 того же параграфа величины ξ_1, ξ_2 зависимы, так как, например,

$$0 = \mathsf{P}\{\xi_1 = 1, \xi_2 = -1\} \neq \mathsf{P}\{\xi_1 = 1\}\mathsf{P}\{\xi_2 = -1\} = \frac{1}{4}.$$

Теорема 5.3. *Пусть $p_{\xi_1\xi_2}(x_1, x_2)$ — плотность распределения случайных величин ξ_1, ξ_2. Случайные величины ξ_1, ξ_2 независимы тогда и только тогда, когда во всех точках непрерывности функций $p_{\xi_1\xi_2}(x_1, x_2)$, $p_{\xi_1}(x_1)$, $p_{\xi_2}(x_2)$ имеем*

$$p_{\xi_1\xi_2}(x_1, x_2) = p_{\xi_1}(x_1)\, p_{\xi_2}(x_2). \tag{5.6}$$

Доказательство. *Необходимость.* Пусть выполнено условие (5.4). Выражая левую и правую части (5.4) через плотности распределения вероятностей, при любом $B = \{(x_1, x_2)\colon a_1 \leqslant x_1 < a_2,\ b_1 \leqslant x_2 < b_2\}$ получим

$$\iint\limits_B p_{\xi_1\xi_2}(x_1, x_2)\, dx_1\, dx_2 = \int\limits^{a_2} p_{\xi_1}(x_1)\, dx_1 \int\limits^{b_2} p_{\xi_2}(x_2)\, dx_2. \tag{5.7}$$

Так как

$$\int\limits_{a_1}^{a_2} p_{\xi_1}(x_1)\, dx_1 \int\limits_{b_1}^{b_2} p_{\xi_2}(x_2)\, dx_2 = \iint\limits_B p_{\xi_1}(x_1)\, p_{\xi_2}(x_2)\, dx_1\, dx_2,$$

то из (5.7) найдем

$$\iint\limits_B [p_{\xi_1\xi_2}(x_1, x_2) - p_{\xi_1}(x_1)\, p_{\xi_2}(x_2)]\, dx_1\, dx_2 = 0. \tag{5.8}$$

Пусть (x_1, x_2), x_1, x_2 — точки непрерывности функций $p_{\xi_1\xi_2}(x_1, x_2)$, $p_{\xi_1}(x_1)$, $p_{\xi_2}(x_2)$. Нужно показать, что в точке (x_1, x_2) имеет место (5.6). Пусть это не так. Тогда в силу непрерывности подынтегральной функции в (5.8) найдется окрестность точки (x_1, x_2), в которой функция отлична от нуля и сохраняет знак. Выбрав прямоугольник B, целиком лежащий внутри этой окрестности, получим противоречие с (5.8).

Достаточность. Пусть выполнено (5.6). Тогда при любом B верно (5.8) и, следовательно, имеют место (5.7) и (5.4). Теорема доказана.

В случае нескольких величин можно проверить, что определению независимости (5.1) или (5.2) равносильны следующие условия:

1) *Абсолютно непрерывные распределения.* Для любых $x = (x_1, \ldots, x_r) \in R^r$

$$p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r) = \prod_{k=1}^{r} p_{\xi_k}(x_k). \tag{5.9}$$

2) *Дискретные распределения.* Для любых $x = (x_1, \ldots, x_r) \in R^r$

$$\mathsf{P}\{\xi_1 = x_1, \ldots, \xi_r = x_r\} = \prod_{k=1}^{r} \mathsf{P}\{\xi_k = x_k\}. \tag{5.10}$$

Таким образом, если известны одномерные распределения величин $\xi_1, \ldots, \xi_r$ и известно, что эти величины независимы, то по каким-либо из формул (5.1), (5.2), (5.9), (5.10) можно найти совместное распределение этих величин.

Будем говорить, что независимые случайные величины $\xi_1, \ldots, \xi_r$ имеют *многомерное нормальное распределение*, если каждая из этих величин имеет одномерное нормальное распределение. Пусть, например, ξ_k, $k = 1, \ldots, r$, имеют нормальное распределение с параметрами (a_k, σ_k^2). Тогда по формуле (5.10), используя формулу плотности одномерного распределения, получим

$$p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r) = \frac{1}{(2\pi)^{r/2}\sigma_1 \ldots \sigma_r} \exp\left\{-\frac{1}{2}\sum_{k=1}^{r}\frac{(x_k - a_k)^2}{\sigma_k^2}\right\}. \tag{5.11}$$

Общее определение нормального распределения будет дано в § 7 гл. 6.

Рассмотрим еще один пример независимых величин. *Индикатором произвольного события* A называется случайная величина

$$I_A = I_A(\omega) = \begin{cases} 1, & \text{если } \omega \in A, \\ 0, & \text{если } \omega \in \bar{A}. \end{cases} \tag{5.12}$$

Число успехов μ_n в n первых испытаниях схемы Бернулли можно представить в виде

$$\mu_n = \xi_1 + \xi_2 + \ldots + \xi_n, \tag{5.13}$$

где $\xi_t = I_{A_1^{(t)}}$, $t = 1, 2, \ldots, n$, $A_1^{(t)} = \{$испытание с номером t закончилось успехом$\}$. Покажем, что индикаторы

ξ_t независимы. Из определения индикатора (5.12) следует, что $\{\xi_t = 1\} = A_1^{(t)}$, $t = 1, \ldots, n$. Положим $\bar{A}_1^{(t)} = A_0^{(t)}$. Выбирая в теореме 2.1 § 2 гл. 4 в качестве событий $B_{s_1}^{(1)}, \ldots, B_{s_n}^{(n)}$ события

$$A_{\varepsilon_1}^{(1)} = \{\xi_1 = \varepsilon_1\}, \ldots, A_{\varepsilon_n}^{(n)} = \{\xi_n = \varepsilon_n\}$$

при любых $\varepsilon_k = 0, 1$, $k = 1, \ldots, n$, получим равенство

$$\mathsf{P}\{\xi_1 = \varepsilon_1, \ldots, \xi_n = \varepsilon_n\} = \mathsf{P}\{\xi_1 = \varepsilon_1\} \ldots \mathsf{P}\{\xi_n = \varepsilon_n\},$$

равносильное определению независимости дискретных величин $\xi_1, \ldots, \xi_n$ (см. (5.10)). Таким образом, в равенстве (5.13) слагаемые независимы. Представление (5.13) будет в дальнейшем часто использоваться.

Рассмотрим теперь последовательность индикаторов, связанную с выбором n шаров по схеме случайного выбора без возвращения из урны, содержащей M белых и $N - M$ черных шаров (см. § 2 гл. 2). Положим $\xi_k = I_{A_k}$, где A_k — событие, состоящее в том, что k-й шар выборки оказался белым. Случайные величины $\xi_1, \ldots, \xi_n$ не являются независимыми, так как, например,

$$\mathsf{P}\{\xi_1 = \xi_2 = \ldots = \xi_n = 1\} = \frac{M^{[n]}}{N^{[n]}} \neq \frac{M^n}{N^n} = \prod_{k=1}^{n} \mathsf{P}\{\xi_k = 1\}.$$

§ 6. Функции от случайных величин

Пусть $(\Omega, \mathfrak{A}, \mathsf{P})$ — произвольное вероятностное пространство и $\xi = \xi(\omega)$, $\omega \in \Omega$, — некоторая случайная величина. Суперпозицией действительной функции ξ, заданной на Ω, и функции $\varphi(x)$, заданной на действительной прямой, является функция $\eta = \varphi[\xi(\omega)] = \eta(\omega)$, заданная на Ω. Для дискретных вероятностных пространств функция η является случайной величиной, так как никаких ограничений на функцию $\eta(\omega)$ не налагается. Для произвольных вероятностных пространств требуется, чтобы

$$(\eta < x) \in \mathfrak{A} \tag{6.1}$$

при любом x *).

*) При произвольной функции φ условие (6.1) может оказаться невыполненным. Однако можно показать, что (6.1) имеет место, если при любом борелевском множестве B множество $\{u\colon \varphi(u) \in B\}$ является борелевским множеством на числовой прямой (т. е. полный прообраз $\varphi^{-1}(B)$ множества B при отображении, осуществляемом функцией φ, является борелевским множеством). Этим свой-

Справедлива следующая теорема.

Теорема 6.1. *Если случайные величины* ξ_1 *и* ξ_2 *независимы, то независимы и случайные величины* $\eta_1 = \varphi_1(\xi_1)$, $\eta_2 = \varphi_2(\xi_2)$.

Доказательство. Пусть $\eta_1 \in B_1$, $\eta_2 \in B_2$ (B_1, B_2 — подмножества числовой прямой) — произвольные события. Тогда

$$\mathsf{P}\{\eta_1 \in B_1, \eta_2 \in B_2\} = \mathsf{P}\{\varphi_1(\xi_1) \in B_1, \varphi_2(\xi_2) \in B_2\} = \\ = \mathsf{P}\{\xi_1 \in \overset{-1}{\varphi_1}(B_1), \xi_2 \in \overset{-1}{\varphi_2}(B_2)\}. \quad (6.2)$$

Так как случайные величины ξ_1 и ξ_2 независимы, то согласно определению независимости в форме (5.2) имеем

$$\mathsf{P}(\xi_1 \in \overset{-1}{\varphi_1}(B_1), \xi_2 \in \overset{-1}{\varphi_2}(B_2)) = \\ = \mathsf{P}(\xi_1 \in \overset{-1}{\varphi_1}(B_1))\,\mathsf{P}(\xi_2 \in \overset{-1}{\varphi_2}(B_2)). \quad (6.3)$$

Очевидно, что

$$(\xi_i \in \overset{-1}{\varphi_i}(B_i)) = (\varphi_i(\xi_i) \in B_i) = (\eta_i \in B_i).$$

Отсюда и из равенств (5.2) и (5.3) получаем

$$\mathsf{P}(\eta_1 \in B_1, \eta_2 \in B_2) = \mathsf{P}(\eta_1 \in B_1)\,\mathsf{P}(\eta_2 \in B_2)$$

для любых событий $\eta_1 \in B_1$, $\eta_2 \in B_2$. Теорема доказана.

Если на исходном вероятностном пространстве задано несколько случайных величин $(\xi_1, \xi_2, \ldots, \xi_n)$, то сложная функция $\eta = \varphi(\xi_1, \ldots, \xi_n)$ также является случайной величиной для достаточно «хороших» функций $\varphi(u_1, u_2, \ldots, u_n)$; например, достаточно потребовать, чтобы $\varphi(u_1, u_2, \ldots, u_n)$ была непрерывной. Можно доказать следующее утверждение, обобщающее теорему 6.1: если $\xi_1, \xi_2, \ldots, \xi_n, \eta_1, \eta_2, \ldots, \eta_m$ — независимые случайные величины, то случайные величины

$$\zeta_1 = \varphi_1(\xi_1, \ldots, \xi_n), \quad \zeta_2 = \varphi_2(\eta_1, \ldots, \eta_m)$$

независимы.

Закон распределения новых случайных величин, являющихся функциями от старых, очевидно, определен, так как они заданы на том же вероятностном пространстве. Можно найти функцию распределения новой случайной

ством обладают, например, функции с конечным числом точек разрыва. Таким образом, в практически важных случаях условие (6.1) оказывается выполненным. В дальнейшем мы будем использовать выражения типа «случайная величина $\eta = \varphi(\xi)$», не оговаривая каждый раз, что $\varphi(x)$ удовлетворяет условию (6.1).

величины $\eta = \varphi(\xi)$. Для этого достаточно знать только функцию распределения ξ. Действительно,

$$F_\eta(x) = \mathsf{P}\{\varphi(\xi) < x\} = \mathsf{P}\{\xi \in \varphi^{-1}(-\infty, x)\}. \quad (6.4)$$

Вероятность в правой части (6.4) может быть вычислена по $F_\xi(x)$. Аналогично находится функция распределения или плотность распределения (если она существует) величины $\eta = g(\xi_1, \ldots, \xi_r)$. Рассмотрим один важный частный случай.

Теорема 6.2. *Пусть* $\eta = (\eta_1, \ldots, \eta_r) = g(\xi)$, *где* $\xi = (\xi_1, \ldots, \xi_r)$ — *случайный абсолютно непрерывный вектор с непрерывной плотностью распределения* $p_\xi(x)$, $x = (x_1, \ldots, x_r)$, $g(x) = (g_1(x), \ldots, g_r(x))$. *Если отображение* $y = g(x)$ *взаимно однозначно, непрерывно дифференцируемо и якобиан*

$$J(x) = \frac{D(g_1, \ldots, g_r)}{D(x_1, \ldots, x_r)} \neq 0,$$

то распределение вектора η *абсолютно непрерывно и*

$$p_\eta(x) = p_\xi(g^{-1}(x))\,|J(g^{-1}(x))|^{-1}, \quad (6.5)$$

где $g^{-1}(x)$ — *прообраз* x.

Доказательство. Пусть B — произвольная область пространства R^r с кусочно-гладкой границей. Тогда

$$\mathsf{P}\{\eta \in B\} = \mathsf{P}\{\xi \in g^{-1}(B)\} = \int\limits_{g^{-1}B}\!\!\ldots\!\int p_\xi(x)\,dx_1 \ldots dx_r.$$

Переходя в последнем интеграле к новым переменным $y = g(x)$, получим

$$\mathsf{P}\{\eta \in B\} = \int\limits_B\!\ldots\!\int p_\xi(g^{-1}(y))\,|J(g^{-1}(y))|^{-1}\,dy.$$

Отсюда согласно определению (4.3) следует, что подынтегральная функция является плотностью распределения вектора η, так как область B может быть, в частности, любым параллелепипедом. Теорема доказана.

Отметим, что в доказанной теореме требования непрерывности $p_\xi(x)$ и дифференцируемости $g(x)$ являются лишними. Они были необходимы только для того, чтобы воспользоваться теоремой о замене переменных, известной из курса математического анализа.

Приведем две теоремы, в которых устанавливается закон распределения функции от случайных величин в двух часто встречающихся случаях.

Теорема 6.3. *Если случайная величина ξ имеет нормальное распределение с параметрами (a, σ^2), то случайная величина $\eta = A\xi + B$, $A \neq 0$, имеет нормальное распределение с параметрами $(Aa + B, A^2\sigma^2)$.*

Доказательство. Пусть сначала $A > 0$. Тогда

$$F_\eta(x) = \mathsf{P}(A\xi + B < x) = \mathsf{P}\left(\xi < \frac{x-B}{A}\right) = \int_{-\infty}^{(x-B)/A} p_\xi(u)\,du.$$

Отсюда, так как $p_\xi(x)$ непрерывна при любом x, следует, что существует

$$p_\eta(x) = F'_\eta(x) = p_\xi\left(\frac{x-B}{A}\right)\cdot\left(\frac{x-B}{A}\right)'_x = \frac{1}{A}\,p_\xi\left(\frac{x-B}{A}\right). \qquad (6.6)$$

Если $A < 0$, то

$$F_\eta(x) = \mathsf{P}(A\xi + B < x) = \mathsf{P}\left(\xi > \frac{x-B}{A}\right) = \int_{(x-B)/A}^{+\infty} p_\xi(u)\,du$$

и

$$p_\eta(x) = -\frac{1}{A}\,p_\xi\left(\frac{x-B}{A}\right).$$

Объединяя последнюю формулу с (6.6) и используя формулу плотности нормального распределения, получим

$$p_\eta(x) = \frac{1}{|A|}\,p_\xi\left(\frac{x-B}{B}\right) =$$

$$= \frac{1}{|A|}\,\frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{\left(\frac{x-B}{A}-a\right)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\,\sigma_1}\,e^{-\frac{(x-a_1)^2}{2\sigma_1^2}},$$

где $a_1 = Aa + B$, $\sigma_1 = |A|\,\sigma$.

Теорема 6.4. *Если случайные величины ξ_1 и ξ_2 абсолютно непрерывны и независимы, то случайная величина $\xi_1 + \xi_2$ тоже абсолютно непрерывна и ее плотность распределения определяется по формуле*

$$p_{\xi_1+\xi_2}(x) = \int_{-\infty}^{\infty} p_{\xi_1}(u)\,p_{\xi_2}(x-u)\,du. \qquad (6,7)$$

Доказательство. Найдем сначала

$$F_{\xi_1+\xi_2}(x) = \mathsf{P}(\xi_1 + \xi_2 < x) = \mathsf{P}((\xi_1, \xi_2) \in D_x);$$

где $D_x = \{(u, v)\colon u + v < x\}$. Так как по условию теоремы величины ξ_1 и ξ_2 независимы, то их плотность совместного распределения согласно (5.6) равна

$$p_{\xi_1\xi_2}(u, v) = p_{\xi_1}(u)\, p_{\xi_2}(v),$$

и, следовательно, по формуле (4.3) найдем

$$F_{\xi_1+\xi_2}(x) = \iint_{D_x} p_{\xi_1}(u)\, p_{\xi_2}(v)\, du\, dv = \int_{-\infty}^{\infty} p_{\xi_1}(u) \left(\int_{-\infty}^{-u+x} p_{\xi_2}(v)\, dv \right) du.$$

Отсюда, полагая $v = z - u$, получим

$$F_{\xi_1+\xi_2}(x) = \int_{-\infty}^{\infty} p_{\xi_1}(u)\, du \int_{-\infty}^{x} p_{\xi_2}(z - u)\, dz = \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} p_{\xi_1}(u)\, p_{\xi_2}(z - u)\, du \right) dz = \int_{-\infty}^{x} p_{\xi_1+\xi_2}(z)\, dz,$$

где

$$p_{\xi_1+\xi_2}(x) = \int_{-\infty}^{\infty} p_{\xi_1}(u)\, p_{\xi_2}(x - u)\, du.$$

Теорема доказана.

Отметим, что вместе с (6.7) верна формула

$$p_{\xi_1+\xi_2}(x) = \int_{-\infty}^{\infty} p_{\xi_1}(x - u)\, p_{\xi_2}(u)\, du.$$

Используя формулу (6.7), можно показать, что сумма независимых нормально распределенных случайных величин имеет нормальное распределение. В гл. 7 этот факт будет установлен при помощи характеристических функций.

Рассмотрим следующий пример.

Пример 1. Пусть случайные величины ξ_1 и ξ_2 независимы и имеют распределения соответственно равномерное и показательное:

$$p_{\xi_1}(x) = \begin{cases} 1, & \text{если } x \in [0, 1], \\ 0, & \text{если } x \notin [0, 1], \end{cases}$$

$$p_{\xi_2}(x) = \begin{cases} e^{-x}, & \text{если } x \geqslant 0, \\ 0, & \text{если } x < 0. \end{cases}$$

Требуется найти $p_{\xi_1+\xi_2}(x)$.

По формуле (6.7) получим

$$p_{\xi_1+\xi_2}(x) = \int_{-\infty}^{\infty} p_{\xi_1}(u)\, p_{\xi_2}(x-u)\, du = \int_0^1 p_{\xi_2}(x-u)\, du.$$

Подынтегральная функция $p_{\xi_2}(x-u) > 0$, если $0 < u < x$, и $p_{\xi_2}(x-u) = 0$ в остальных случаях. Таким образом, при $0 < x \leqslant 1$

$$p_{\xi_1+\xi_2}(x) = \int_0^x e^{-(x-u)}\, du = 1 - e^{-x}$$

и при $x > 1$

$$p_{\xi_1+\xi_2}(x) = \int_0^1 e^{-(x-u)}\, du = (e-1)\, e^{-x}.$$

Окончательно получим

$$p_{\xi_1+\xi_2}(x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ 1 - e^{-x}, & \text{если } 0 < x \leqslant 1, \\ (e-1)\, e^{-x}, & \text{если } x > 1. \end{cases}$$

Задачи к главе 5

1. Плотность распределения ξ задана формулами

$$p_\xi(x) = C/x^4 \ (x \geqslant 1), \quad p_\xi(x) = 0 \ (x < 1).$$

Найти постоянную C, плотность распределения величины $\eta = \ln \xi$, $\mathsf{P}(0{,}5 < \eta < 0{,}75)$.

2. Случайная величина ξ равномерно распределена на отрезке $[0, 1]$. Найти плотности распределения величин: а) $\eta_1 = 2\xi + 1$; б) $\eta_2 = -\ln(1-\xi)$.

3. Случайная величина ξ имеет показательное распределение с плотностью распределения $p_\xi(x) = \alpha e^{-\alpha x} (x > 0)$. Найти плотности распределения случайных величин: а) $\eta_1 = \sqrt{\xi}$; б) $\eta_2 = \xi^2$; в) $\eta_3 = \frac{1}{\alpha} \ln \xi$; г) $\eta_4 = 1 - e^{-\alpha\xi}$.

4. Случайная величина ξ распределена нормально с параметрами $a = 0$, $\sigma^2 = 1$. Найти плотности распределения величин: а) $\eta_1 = \xi^2$; б) $\eta_2 = e^\xi$ (логарифмически нормальное распределение).

5. Точка P равномерно распределена на единичном квадрате $ABCD$. Найти плотность распределения площади ξ прямоугольника $AB'PD'$, где B' и D' — основания перпендикуляров, опущенных из точки P на стороны AB и AD соответственно.

6. Функция распределения $F(x)$ величины ξ строго монотонна и непрерывна. Найти закон распределения величины $\eta = F(\xi)$.

7. Случайные величины ξ_1 и ξ_2 независимы и имеют равномерное распределение на отрезке $[0, 1]$. Найти плотности распределения величин: а) $\xi_1 + \xi_2$; б) $\xi_1 - \xi_2$; в) ξ_1/ξ_2.

8. Случайные величины ξ_1 и ξ_2 независимы, одинаково распределены и имеют показательное распределение: $p_{\xi_1}(x) = p_{\xi_2}(x) = \alpha e^{-\alpha x}$, $x > 0$. Найти плотность распределения их суммы.

9. Случайные величины ξ_1 и ξ_2 независимы и нормально распределены с параметрами $(0, 1)$. Найти плотности распределения величин: а) $\eta_1 = \xi_1^2 + \xi_2^2$; б) $\eta_2 = \operatorname{arctg} \xi_1/\xi_2$; в) совместную плотность распределения (η_1, η_2).

10. Найти совместное распределение величин $\eta_1 = \sqrt{-2\ln \xi_1} \times \cos(2\pi\xi_2)$, $\eta_2 = \sqrt{-2\ln \xi_1} \sin(2\pi\xi_2)$, где ξ_1 и ξ_2 распределены так же, как в задаче 7.

11. Совместное распределение случайных величин ξ_1, ξ_2 задано таблицей

ξ_1 \ ξ_2	-1	0	1
-1	$1/8$	$1/12$	$/24$
1	$5/24$	$1/6$	$1/8$

в которой на пересечении i-й строки и j-го столбца ($i = -1, 1$, $j = -1, 0, 1$) приведена вероятность $p_{ij} = \mathsf{P}\{\xi_1 = i, \xi_2 = j\}$. Найти: а) одномерные законы распределения ξ_1 и ξ_2; б) закон распределения $\eta_1 = \xi_1 + \xi_2$; в) закон распределения $\eta_2 = \xi_2^2$; г) $\mathsf{P}(\eta_1 = 0, \eta_2 = 1)$.

12. Величины ξ_1 и ξ_2 независимы; $\mathsf{P}\{\xi_1 = 0\} = \mathsf{P}\{\xi_2 = 1\} = 1/2$; ξ_2 равномерно распределена на отрезке $[0, 1]$. Найти закон распределения $\xi_1 + \xi_2$.

13. Найти закон распределения суммы двух независимых случайных величин ξ_1 и ξ_2, каждая из которых имеет распределение Пуассона с параметром λ_1 и λ_2 соответственно.

14. Обозначим τ число испытаний в схеме Бернулли до появления первого успеха включительно. Найти закон распределения τ.

15. Величина $\tau^{(1)}$ равна числу испытаний в схеме Бернулли до первого успеха включительно, $\tau^{(2)}$ — число испытаний, прошедших после первого успеха до второго успеха. Найти совместное распределение $\tau^{(1)}$, $\tau^{(2)}$. Являются ли $\tau^{(1)}$ и $\tau^{(2)}$ независимыми?

16. Найти закон распределения величины $\tau^{(m)}$, равной числу испытаний в схеме Бернулли до появления m-го успеха.

17. Игральная кость бросается до тех пор, пока впервые не выпадет меньше пяти очков. Обозначим θ число очков, выпавших при последнем бросании игральной кости, и через ν — число бросаний кости. Найти совместное распределение θ и ν. Являются ли θ и ν независимыми?

18. На n станках одновременно началась обработка n деталей. Предполагая, что времена обработки деталей независимы и имеют показательные распределения с параметром α, найти распределение времени: а) до получения первой обработанной детали; б) до окончания обработки всех деталей.

19. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — независимые, одинаково распределенные, абсолютно непрерывные случайные величины; $\mathsf{P}\{\xi_i < x\} = F(x)$, $F'(x) = p(x)$. При каждом ω расположим числа $\xi_k(\omega)$, $k = 1, 2, \ldots, n$, в порядке возрастания и перенумеруем: $\eta_1 \leqslant \eta_2 \leqslant \ldots \leqslant \eta_n$. Таким образом, $\eta_1(\omega)$ является наименьшим из чисел $(\xi_1(\omega), \ldots, \xi_n(\omega))$, $\eta_n(\omega)$ — наибольшим из тех же чисел и т. д. Найти плотности распределений: 1) η_1; 2) η_n; 3) η_m, $1 < m < n$; 4) (η_m, η_k).

20. Машина состоит из 10 000 деталей. Каждая деталь независимо от других оказывается неисправной с вероятностью p_i, причем для $n_1 = 1000$ деталей $p_1 = 0{,}0003$; для $n_2 = 2000$ деталей $p_2 = 0{,}0005$ и для $n_3 = 7000$ деталей $p_3 = 0{,}0001$. Машина не работает, если в ней неисправны хотя бы две детали. Найти приближенное значение вероятности того, что машина не будет работать.

Указание. Воспользоваться теоремой Пуассона.

21. Совместное распределение величин ξ_1, ξ_2 является равномерным в круге $\{(x_1, x_2)\colon x_1^2 + x_2^2 < 1\}$. Найти вероятность

$$\mathsf{P}\left\{|\xi_1| < \frac{1}{\sqrt{2}},\ |\xi_2| < \frac{1}{\sqrt{2}}\right\}.$$

Являются ли величины ξ_1, ξ_2 независимыми?

22. Случайная величина ξ с равномерным распределением на $[0, 1]$ записывается в виде бесконечной десятичной дроби: $\xi = \sum_{n=1}^{\infty} \xi_n \cdot 10^{-n}$, $0 \leqslant \xi_n \leqslant 9$. Найти совместные и одномерные распределения величин ξ_1, ξ_2. Являются ли ξ_1, ξ_2 независимыми?

23. Используя табл. 7 выписать 10 реализаций испытаний, описанных в задаче 17. Найти частоты событий: $\{\theta = k\}$, $k = 1, 2, 3, 4$; $\{\nu = l\}$, $l = 1, 2, 3, \ldots$

24 Используя табл. 7, выписать реализацию 50 независимых случайных величин, равномерно распределенных на отрезке $[0, 1]$. Числовые значения взять с точностью до 10^{-3}.

25. Пусть $F(x)$ — непрерывная строго возрастающая функция распределения и $F^{-1}(x)$ — обратная к ней функция. Показать, что случайная величина $\eta = F^{-1}(\xi)$, где ξ равномерно распределена на отрезке $[0, 1]$, имеет своей функцией распределения $F(x)$. Указанное свойство позволяет из реализаций равномерно распределенных величин получать реализации величин с функцией распределения $F(x)$. (См. также задачу 10, в которой содержится метод получения нормально распределенных величин.)

Математическое ожидание

§ 1. Определения

Функция распределения полностью определяет закон распределения случайной величины. Однако часто можно ограничиться более простой характеристикой. Одной из важнейших характеристик случайной величины является ее среднее значение, или математическое ожидание. Математическое ожидание легко вычисляется и обладает полезным свойством: арифметическое среднее одинаково распределенных случайных величин близко к математическому ожиданию. Определение математического ожидания связано с обычным понятием о среднем значении. Пусть, например, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$; $\mathsf{P}(\omega_k) = 1/N$, $k = 1, \ldots, N$; $\xi = \xi(\omega_k) = x_k$. Тогда математическое ожидание $\mathsf{M}\xi$ случайной величины ξ определяется формулой $\mathsf{M}\xi = (x_1 + \ldots + x_N)/N = \frac{1}{N} x_1 + \ldots + \frac{1}{N} x_N$. Если вероятности $\mathsf{P}(\omega_k) = p_k$ отличны от $1/N$, то среднее значение ξ естественно вычислять по формуле $\mathsf{M}\xi = p_1x_1 + p_2x_2 + \ldots + p_Nx_N$.

Наиболее удобное и простое определение математического ожидания произвольной случайной величины вводится при помощи интеграла Лебега.

Математическим ожиданием случайной величины ξ, заданной на вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$, называется число

$$\mathsf{M}\xi = \int_{\Omega} \xi(\omega)\, \mathsf{P}(d\omega), \tag{1.1}$$

если интеграл Лебега, стоящий в правой части равенства, существует (см. [2], стр. 74, 319—324).

Мы не будем пользоваться теорией интеграла Лебега и ограничимся определением математического ожидания для случайных величин, заданных на дискретном и абсолютно непрерывном вероятностных пространствах (см. §§ 2, 4 гл. 2). Чтобы получить интересующие нас определения для указанных пространств, нужно в дискретном случае заменить в (1.1) символ $\int$ на $\sum$, а ω и $\mathsf{P}(d\omega)$ на ω_k и p_k; в абсолютно непрерывном случае нужно в (1.1) положить $\mathsf{P}(d\omega) = \pi(u_1, \ldots, u_n)\,du_1, \ldots, du_n$, $\omega = (u_1, \ldots, u_n)$.

Приведем полные формулировки определений.

Математическим ожиданием $\mathsf{M}\xi$ *случайной величины* $\xi = \xi(\omega_k)$, *заданной на дискретном вероятностном пространстве с* $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$, $\mathsf{P}(\omega_k) = p_k$, $k = 1, \ldots, n, \ldots$ (см. § 2 гл. 2), *называется число*

$$\mathsf{M}\xi = \sum_{k=1}^{\infty} \xi(\omega_k)\, p_k, \tag{1.2}$$

если ряд абсолютно сходится. Если же ряд (1.2) не сходится абсолютно, то говорят, что математическое ожидание случайной величины ξ не существует.

Математическим ожиданием $\mathsf{M}\xi$ *случайной величины* $\xi = \xi(u_1, u_2, \ldots, u_n)$, *заданной на абсолютно непрерывном вероятностном пространстве* $(\Omega, \mathfrak{A}, \mathsf{P})$ (см. § 2 гл. 2), *называется число*

$$\mathsf{M}\xi = \int \underset{\Omega}{\ldots} \int \xi(u_1, \ldots, u_n)\, \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n, \tag{1.3}$$

если интеграл абсолютно сходится.

Если интеграл (1.3) не сходится абсолютно, то говорят, что математическое ожидание случайной величины ξ не существует.

П р и м е р 1. Найдем математическое ожидание выигрыша первого игрока в примере 1 § 1 гл. 5. В этом примере

$$\Omega = \{\Gamma, \text{Р}\},\ \mathsf{P}(\{\Gamma\}) = \mathsf{P}(\{\text{Р}\}) = 1/2,$$
$$\xi(\Gamma) = 1,\quad \xi(\text{Р}) = -1.$$

По формуле (1.2)

$$\mathsf{M}\xi = \xi(\Gamma)\,\mathsf{P}(\Gamma) + \xi(P)\,\mathsf{P}(\text{Р}) = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot(-1) = 0.$$

П р и м е р 2. Пусть в дискретном вероятностном пространстве $\Omega = \{1, 2, \ldots, k, \ldots\}$ и $\mathsf{P}(k) = p_k = 1/k\,(k+1)$. Непосредственно проверяется, что

$$\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty}\left(\frac{1}{k} - \frac{1}{k+1}\right) = 1.$$

Положим $\xi_1 = \xi_1(k) = (-1)^k$, $\xi_2 = \xi_2(k) = (-1)^k k$. Так как ряд $\sum\limits_{k=1}^{\infty}(-1)^k/k(k+1)$ сходится абсолютно, то $\mathsf{M}\xi_1$ существует. По формуле (1.2)

$$\mathsf{M}\xi_1 = \sum_{k=1}^{\infty} \xi_1(k)\,p_k = \sum_{k=1}^{\infty}(-1)^k\left(\frac{1}{k} - \frac{1}{k+1}\right) =$$
$$= \sum_{k=1}^{\infty}\frac{(-1)^k}{k} + \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k+1} = 1 + 2\sum_{k=1}^{\infty}\frac{(-1)^k}{k}.$$

Отсюда, используя разложение $-\ln(1+x) = = \sum\limits_{k=1}^{\infty}(-1)^k x^k/k$ при $x = 1$, получим $\mathsf{M}\xi_1 = 1 - 2\ln 2$. Однако $\mathsf{M}\xi_2$ по определению не существует, так как ряд $\sum\limits_{k=1}^{\infty}|(-1)^k k|\dfrac{1}{k(k+1)}$ расходится.

Пример 3. Пусть в абсолютно непрерывном вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$ (см. § 4 гл. 2) $n = 1$,

$$\Omega = \{u\colon 0 \leqslant u \leqslant 1\},\quad \pi(u) = 1\ (0 \leqslant u \leqslant 1).$$

Рассмотрим случайные величины: $\xi_1(u) = u$ $(0 \leqslant u \leqslant \leqslant 1)$; $\xi_2(u) = k\left(\frac{k-1}{3} \leqslant u < \frac{k}{3}\right)$, $k = 1, 2, 3$; $\xi_3(u) = u^2$ $(0 \leqslant u < 1/2)$, $\xi_3(u) = 1$ $(1/2 \leqslant u \leqslant 1)$. По формуле (1.3) $\mathsf{M}\xi_l = \int\limits_{\Omega}\xi_l(u)\,\pi(u)\,du$. Отсюда

$$\mathsf{M}\xi_1 = \int_0^1 u\,du = \frac{1}{2};\quad \mathsf{M}\xi_2 = \int_0^{1/3} du + \int_{1/3}^{2/3} 2\,du + \int_{2/3}^{1} 3\,du = 2;$$
$$\mathsf{M}\xi_3 = \int_0^{1/2} u^2\,du + \int_{1/2}^{1} du = 13/24.$$

Отметим, что в данном примере величина ξ_1 абсолютно непрерывна, ξ_2 — дискретна, а ξ_3 — смешанного типа.

Если известна плотность распределения случайной величины или заданы вероятности значений дискретной величины, то может оказаться, что математическое ожидание удобнее вычислять не по сформулированным выше определениям, а по формулам, которые мы получим в качестве частных случаев следующих двух теорем.

Теорема 1.1. *Пусть* $\xi = (\xi_1, \xi_2, \ldots, \xi_r)$ — *дискретный случайный вектор, для которого*

$$\mathsf{P}\{\xi = x(k)\} = p_k > 0, \quad x(k) = (x_{k1}, \ldots, x_{kr}) \in R_r,$$

$$\sum_{k=1}^{\infty} p_k = 1.$$

Если ряд $\sum_{k=1}^{\infty} |g(x(k)| p_k$ *сходится, то случайная величина* $\eta = g(\xi) = g(\xi_1, \ldots, \xi_r)$ *имеет математическое ожидание*

$$\mathsf{M}\eta = \sum_{k=1}^{\infty} g(x_{k1}, \ldots, x_{kr})\, p_k. \tag{1.4}$$

Теорема 1.2. *Пусть* $\xi = (\xi_1, \ldots, \xi_r)$ — *абсолютно непрерывный случайный вектор с плотностью распределения* $p_\xi(x) = p_\xi(x_1, \ldots, x_r)$. *Если интеграл*

$$\int_{R^r} \cdots \int |g(x_1, \ldots, x_r)|\, p_\xi(x_1, \ldots, x_r)\, dx_1 \ldots dx_r$$

сходится, то математическое ожидание случайной величины $\eta = g(\xi_1, \ldots, \xi_r)$ *существует и*

$$\mathsf{M}\eta = \int_{R^r} \cdots \int g(x_1, \ldots, x_r)\, p_\xi(x_1, \ldots, x_r)\, dx_1 \ldots dx_r. \tag{1.5}$$

Отсюда, полагая $n = 1$ и $g(x) = x$, получим формулы для вычисления математического ожидания случайной величины по плотности распределения или по вероятностям значений дискретной случайной величины.

Если $\sum_{k=1}^{\infty} \mathsf{P}(\xi = x_k) = 1$, то из формулы (1.4) получим

$$\mathsf{M}\xi = \sum_{k=1}^{\infty} x_k \mathsf{P}(\xi = x_k). \tag{1.6}$$

Если $p_\xi(x)$ — плотность распределения ξ, то из формулы (1.5) следует, что *)

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x p_\xi(x)\, dx. \tag{1.7}$$

---

*) Используя интеграл Стилтьеса, можно записать общую формулу для вычисления $\mathsf{M}\xi$, аналогичную (1.6) и (1.7), в виде

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x\, dF_\xi(x).$$

Доказательство теоремы 1.1. Проведем доказательство для случая, когда случайные величины $\xi_1, \xi_2, \ldots, \xi_r$ заданы на дискретном вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$ (см. § 2 гл. 2). По определению случайный вектор называется дискретным, если существует такое множество точек $x(k) = (x_{k1}, x_{k2}, \ldots, x_{kr})$, $k = 1, 2, \ldots$, что

$$\sum_{k=1}^{\infty} \mathsf{P}(\xi_1 = x_{k1}, \ldots, \xi_r = x_{kr}) = 1. \tag{1.8}$$

Записывая математическое ожидание случайной величины $\eta = \eta(\omega_m) = g(\xi_1(\omega_m), \ldots, \xi_r(\omega_m))$ по формуле (1.2), получим

$$\mathsf{M}\eta = \sum_{m=1}^{\infty} g(\xi_1(\omega_m), \ldots, \xi_r(\omega_m))\, \mathsf{P}(\omega_m). \tag{1.9}$$

Преобразование (1.9) к виду (1.4) основано на «приведении подобных членов». Введем множества

$$A_k = \{m\colon \xi(\omega_m) = x(k)\}. \tag{1.10}$$

Очевидно, что любая пара A_k и A_l $(l \neq k)$ не имеет общих элементов и что $\bigcup\limits_{k=1}^{\infty} A_k = \{1, 2, 3, \ldots, m, \ldots\}$. Если $\mathsf{M}\eta$ существует, то ряд (1.9) сходится абсолютно и его члены можно произвольно переставлять. Тогда (1.9) можно записать в виде

$$\mathsf{M}\eta = \sum_{k=1}^{\infty} \sum_{m \in A_k} g(\xi(\omega_m))\, \mathsf{P}(\omega_m).$$

Отсюда, учитывая (1.10), находим

$$\mathsf{M}\eta = \sum_{k=1}^{\infty} (g(x(k)) \sum_{m \in A_k} \mathsf{P}(\omega_m)).$$

В предположении существования $\mathsf{M}\eta$ формула (1.4) доказана, так как $\sum\limits_{m \in A_k} \mathsf{P}(\omega_m) = \mathsf{P}(\xi = x(k)) = p_k$. Если абсолютно сходится ряд в (1.4), то, начав с (1.4), можно аналогичными рассуждениями получить формулу для $\mathsf{M}\eta$ в виде (1.9). Теорема 1.1 доказана в случае, когда вектор ξ задан на дискретном вероятностном пространстве.

Доказательство (1.4) для случая, когда дискретный случайный вектор задан на абсолютно непрерывном вероятностном пространстве, а также доказательство теоремы 1.2 приведены не будут. Эти доказательства при условии использования только фактов из курса обычного анализа потребовали бы введения дополнительных лишних ограничений на рассматриваемые случайные величины. Приведенное для дискретных пространств доказательство дает достаточно хорошее представление о связи формул (1.2), (1.3) с (1.4), (1.5).

Воспользуемся формулами (1.6) и (1.7) для вычисления математических ожиданий случайных величин, имеющих распределения, приведенные в § 3 гл. 5.

1. *Нормальное распределение.* Подставляя в формулу (1.7) плотность нормального распределения, получим

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}} dy + \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy.$$

Так как функция $ye^{-\frac{y^2}{2}}$ нечетная, а $\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ является плотностью нормального распределения с параметрами (0, 1), то первое слагаемое в правой части последнего равенства равно 0, а второе a. Таким образом, если случайная величина ξ распределена нормально с параметрами (a, σ^2), то $\mathsf{M}\xi = a$.

2. *Показательное распределение*:

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x p_\xi(x)\, dx = \int_0^{\infty} \alpha x e^{-\alpha x}\, dx = \frac{1}{\alpha}.$$

3. *Равномерное распределение*:

$$\mathsf{M}\xi = \frac{1}{b-a} \int_a^b x\, dx = \frac{a+b}{2}.$$

Найдем теперь по формуле (1.6) математические ожидания некоторых дискретных случайных величин.

4. *Биномиальное распределение.* Из формулы

$$\mathsf{M}\xi = \sum_{m=0}^{n} m C_n^m p^m (1-p)^{n-m},$$

воспользовавшись равенством $mC_n^m = nC_{n-1}^{m-1}$, получим

$$\mathsf{M}\xi = np \sum_{m=1}^{n} C_{n-1}^{m-1} p^{m-1} (1-p)^{n-m} = np\,(p + (1-p))^{n-1}.$$

Таким образом, $\mathsf{M}\xi = np$.

5. *Пуассоновское распределение.* Так как

$$\mathsf{M}\xi = \sum_{m=0}^{n} m \frac{\lambda^m}{m!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = \lambda e^{-\lambda} \cdot e^{\lambda},$$

то $\mathsf{M}\xi = \lambda$.

§ 2. Свойства математического ожидания

Приведем основные свойства математического ожидания.

Теорема 2.1.

1°. *Если C — постоянная, то $\mathsf{M}C = C$.*

2°. *Если C — постоянная, то $\mathsf{M}(C\xi) = C\mathsf{M}\xi$.*

3°. *Для любых величин ξ*

$$|\mathsf{M}\xi| \leqslant \mathsf{M}|\xi|.$$

4°. *Для любых случайных величин ξ_1 и ξ_2*

$$\mathsf{M}(\xi_1 + \xi_2) = \mathsf{M}\xi_1 + \mathsf{M}\xi_2.$$

Если из участвующих в равенстве математических ожиданий какие-нибудь два существуют, то существует третье.

5°. *Если случайные величины ξ_1 и ξ_2 независимы, то $\mathsf{M}\xi_1\xi_2 = \mathsf{M}\xi_1 \cdot \mathsf{M}\xi_2$. Из существования любых двух математических ожиданий следует существование третьего.*

Доказательство. Свойства 1°—4° следуют из соответствующих свойств интеграла или ряда. Докажем, например, свойство 4° для величин, определенных в абсолютном непрерывном вероятностном пространстве (см. § 4 гл. 2). Пусть ξ_1, ξ_2 заданы на вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$. Тогда сумма $\xi_1 + \xi_2$ также является случайной величиной, определенной на том же вероятностном пространстве, и по формуле (1.3)

$$\mathsf{M}(\xi_1 + \xi_2) = \int \underset{\Omega}{\ldots} \int (\xi_1(u) + \xi_2(u))\, \pi(u)\, du_1 \ldots du_r =$$

$$= \int \underset{\Omega}{\ldots} \int \xi_1(u)\, \pi(u)\, du_1 \ldots du_r + \int \underset{\Omega}{\ldots} \int \xi_2(u) \pi(u)\, du_1 \ldots du_r =$$

$$= \mathsf{M}\xi_1 + \mathsf{M}\xi_2 \quad (u = (u_1, \ldots, u_r)).$$

Здесь мы воспользовались тем, что интеграл от суммы двух функций равен сумме интегралов.

Докажем свойство 5°. Пусть, например, ξ_1 и ξ_2 — абсолютно непрерывные величины и $p_{\xi_1\xi_2}(u, v)$ — их плотность распределения. Так как ξ_1 и ξ_2 независимы, то $p_{\xi_1\xi_2}(u, v) = p_{\xi_1}(u)\, p_{\xi_2}(v)$. По формуле (1.7) с $n = 2$ и $g(x_1, x_2) = x_1 x_2$ получим

$$\mathsf{M}\xi_1\xi_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv p_{\xi_1}(u)\, p_{\xi_2}(v)\, du\, dv =$$
$$= \int_{-\infty}^{\infty} u p_{\xi_1}(u)\, du \cdot \int_{-\infty}^{\infty} v p_{\xi_2}(v)\, dv = \mathsf{M}\xi_1 \cdot \mathsf{M}\xi_2.$$

Из свойств 2° и 4° по индукции следует, что

$$\mathsf{M}(C_1\xi_1 + \ldots + C_n\xi_n) = C_1\mathsf{M}\xi_1 + \ldots + C_n\mathsf{M}\xi_n. \quad (2.1)$$

Приведенные выше свойства помогают при вычислении математических ожиданий. Рассмотрим два примера.

Пример 1. Пусть ξ имеет нормальное распределение с параметрами (a, σ^2). В § 1 было показано, что $\mathsf{M}\xi = a$. Найдем математическое ожидание величины $\eta = A\xi + B$:

$$\mathsf{M}\eta = A\mathsf{M}\xi + B = Aa + B.$$

Таким образом, формула для $\mathsf{M}\eta$ дает простое объяснение формулы для параметра a_1.

Пример 2. Пусть μ_n — число успехов в n испытаниях Бернулли с вероятностью p успеха в отдельном испытании. Эта величина имеет биномиальное распределение. В § 1 было показано, что случайная величина с биномиальным распределением имеет математическое ожидание, равное np. Покажем, как этот результат можно получить при помощи представления μ_n в виде суммы индикаторов (5.5.13):

$$\mu_n = \xi_1 + \xi_2 + \ldots + \xi_n,$$

где ξ_k — независимые одинаково распределенные случайные величины с $\mathsf{P}(\xi_k = 1) = p$, $\mathsf{P}(\xi_k = 0) = 1 - p$. По формуле (2.1)

$$\mathsf{M}\mu_n = \mathsf{M}\xi_1 + \mathsf{M}\xi_2 + \ldots + \mathsf{M}\xi_n.$$

Слагаемые легко вычисляются по формуле (1.6):

$$\mathsf{M}\xi_k = 1 \cdot p + 0 \cdot (1 - p) = p, \quad k = 1, 2, \ldots, n.$$

Таким образом, $\mathsf{M}\mu_n = np$.

Пример 3. Из урны, содержащей M белых и $N-M$ черных шаров, по схеме случайного выбора без возвращения извлекается n шаров (см. § 2 гл. 2). Пусть ξ — число белых шаров в выборке. Найдем $\mathsf{M}\xi$. Обозначим A_k событие, состоящее в том, что k-й шар в выборке оказался белым. Представим ξ в виде суммы индикаторов $\xi_k = I_{A_k}$, $k = 1, \ldots, n$:

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_n. \tag{2.2}$$

Так как $\mathsf{P}(\xi_k = 1) = M/N$ (см. задачу 22 гл. 2), то $\mathsf{M}\xi_k = M/N$ и $\mathsf{M}\xi = n \cdot M/N$.

В качестве еще одного примера использования свойств математического ожидания докажем следующую теорему.

Теорема 2.2. *Если $A_1, A_2, \ldots, A_n$ — произвольные события, то*

$$\mathsf{P}\left(\bigcup_{m=1}^{n} A_m\right) = \sum_{m=1}^{n} (-1)^{m+1} \sum_{1 \leqslant k_1 < k_2 < \ldots < k_m \leqslant n} \mathsf{P}(A_{k_1} A_{k_2} \ldots A_{k_m}).$$

Доказательство. Из равенства $\overline{\bigcup_{m=1}^{n} A_m} = \bigcap_{m=1}^{n} \bar{A}_m$ следует, что

$$\mathsf{P}\left(\bigcup_{m=1}^{n} A_m\right) = 1 - \mathsf{P}\left(\bigcap_{m=1}^{n} \bar{A}_m\right). \tag{2.3}$$

Пусть $\eta_k = I_{A_k}$ $(k = 1, \ldots, n)$ — индикатор события A_k. Тогда случайная величина $\eta = \prod_{k=1}^{n} (1 - \eta_k)$ принимает два значения: значение 1, если все $\eta_k = 0$, и значение 0 в остальных случаях. Событие $\{\eta_1 = \eta_2 = \ldots = \eta_n = 0\} = \bar{A}_1 \bar{A}_2 \ldots \bar{A}_n$. Следовательно,

$$\mathsf{M}\eta = \mathsf{P}(\bar{A}_1 \bar{A}_2 \ldots \bar{A}_n). \tag{2.4}$$

Так как

$$\eta = 1 - \sum_{m=1}^{n} (-1)^{m+1} \sum_{1 \leqslant k_1 < \ldots < k_m \leqslant n} \eta_{k_1} \ldots \eta_{k_m},$$

то

$$\mathsf{M}\eta = 1 - \sum_{m=1}^{n} (-1)^{m+1} \sum_{1 \leqslant k_1 < \ldots < k_m \leqslant n} \mathsf{M}\eta_{k_1} \ldots \eta_{k_m}. \tag{2.5}$$

Очевидно, что

$$\eta_{k_1}\eta_{k_2}\ldots\eta_{k_m}=I_{A_{k_1}\ldots A_{k_m}}=\begin{cases}1, & \text{если } \omega\in A_{k_1}\ldots A_{k_m},\\ 0, & \text{если } \omega\in\overline{A_{k_1}\ldots A_{k_m}}.\end{cases}$$

Поэтому

$$\mathsf{M}\eta_{k_1}\eta_{k_2}\ldots\eta_{k_m}=\mathsf{P}(A_{k_1}A_{k_2}\ldots A_{k_m}). \tag{2.6}$$

Утверждение теоремы следует теперь из формул (2.3)—(2.6).

Приведенная в теореме формула обобщает формулу (1.3.7).

§ 3. Дисперсия

Дисперсией $\mathsf{D}\xi$ случайной величины ξ называется число

$$\mathsf{D}\xi=\mathsf{M}(\xi-\mathsf{M}\xi)^2, \tag{3.1}$$

если математическое ожидание в правой части (3.1) существует. Дисперсия является мерой рассеяния значений случайной величины около ее математического ожидания. Величину $\sqrt{\mathsf{D}\xi}$ называют средним квадратическим отклонением.

Если воспользоваться свойствами математического ожидания, то правую часть (3.1) можно привести к следующему виду:

$$\mathsf{M}(\xi-\mathsf{M}\xi)^2=\mathsf{M}(\xi^2-2\xi\mathsf{M}\xi+(\mathsf{M}\xi)^2)=$$
$$=\mathsf{M}\xi^2-2\mathsf{M}\xi\cdot\mathsf{M}\xi+(\mathsf{M}\xi)^2.$$

Отсюда и из (3.1) следует, что

$$\mathsf{D}\xi=\mathsf{M}\xi^2-(\mathsf{M}\xi)^2. \tag{3.2}$$

Если случайная величина ξ абсолютно непрерывна, то, полагая в формуле (1.5) $n=1$ и $g(x_1)=(x_1-\mathsf{M}\xi)^2$, получим

$$\mathsf{D}\xi=\int_{-\infty}^{\infty}(x-\mathsf{M}\xi)^2p_\xi(x)\,dx. \tag{3.3}$$

Для дискретной величины ξ из (1.6) найдем

$$\mathsf{D}\xi=\sum_{k=1}^{\infty}(x_k-\mathsf{M}\xi)^2\mathsf{P}(\xi=x_k), \tag{3.4}$$

где $\sum_{k=1}^{\infty}\mathsf{P}(\xi=x_k)=1$.

Приведем свойства дисперсии.

Теорема 3.1.

1°. *Для любой случайной величины* ξ *имеем* $\mathsf{D}\xi \geqslant 0$.

2°. *Если* c — *постоянная, то* $\mathsf{D}c = 0$.

3°. *Если* c — *постоянная, то* $\mathsf{D}(c\xi) = c^2\mathsf{D}\xi$.

4°. *Если случайные величины* ξ_1 *и* ξ_2 *независимы, то*

$$\mathsf{D}(\xi_1 + \xi_2) = \mathsf{D}\xi_1 + \mathsf{D}\xi_2. \tag{3.5}$$

Доказательство. Свойства 1°—3° следуют непосредственно из определения и свойств математического ожидания. Докажем свойство 4°. По определению (3.1)

$$\begin{aligned}\mathsf{D}(\xi_1 + \xi_2) &= \mathsf{M}[(\xi_1 + \xi_2) - \mathsf{M}(\xi_1 + \xi_2)]^2 = \\ &= \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1) + (\xi_2 - \mathsf{M}\xi_2)]^2 = \\ &= \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)^2 + (\xi_2 - \mathsf{M}\xi_2)^2 + \\ &\qquad + 2(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)]. \end{aligned} \tag{3.6}$$

Отсюда, так как случайные величины $\xi_1 - \mathsf{M}\xi_1$, $\xi_2 - \mathsf{M}\xi_2$ независимы и

$$\begin{aligned}\mathsf{M}(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2) &= \mathsf{M}(\xi_1 - \mathsf{M}\xi_1)\cdot\mathsf{M}(\xi_2 - \mathsf{M}\xi_2) = \\ &= (\mathsf{M}\xi_1 - \mathsf{M}\xi_1)(\mathsf{M}\xi_2 - \mathsf{M}\xi_2) = 0,\end{aligned}$$

следует, что

$$\mathsf{D}(\xi_1 + \xi_2) = \mathsf{M}(\xi_1 - \mathsf{M}\xi_1)^2 + \mathsf{M}(\xi_2 - \mathsf{M}\xi_2)^2 = \mathsf{D}\xi_1 + \mathsf{D}\xi_2.$$

Вычислим дисперсии некоторых случайных величин.

1. *Нормальное распределение.* По формуле (3.3) находим

$$\mathsf{D}\xi = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}(x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx.$$

Отсюда, полагая $x = \sigma y + a$, получим

$$\mathsf{D}\xi = \sigma^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 e^{-y^2/2}\,dy = -\sigma^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y\,d\left(e^{-y^2/2}\right).$$

Применив формулу интегрирования по частям, находим $\mathsf{D}\xi = \sigma^2$.

В § 6 гл. 5 было показано, что линейная функция $\eta = A\xi + B$ от нормально распределенной случайной величины ξ имеет нормальное распределение. В § 2 был указан способ вычисления $\mathsf{M}\eta$, не связанный с доказательством нормальности η (см. формулу (2.2)). Найдем теперь $\mathsf{D}\eta$. Нетрудно показать, что случайная величина, являющаяся постоянной, взаимно независима с любой случайной величиной. Поэтому при вычислении $\mathsf{D}\eta = \mathsf{D}(A\xi + B)$ можно воспользоваться формулой (3.5). Тогда

$$\mathsf{D}\eta = \mathsf{D}(A\xi + B) = \mathsf{D}(A\xi) + \mathsf{D}B = A^2\mathsf{D}\xi = A^2\sigma^2.$$

2. *Равномерное распределение.* По формуле (3.3)

$$\mathsf{D}\xi = \int_{-\infty}^{\infty}\left(x - \frac{a+b}{2}\right)^2 p_\xi(x)\,dx = \frac{1}{b-a}\int_a^b \left(x - \frac{a+b}{2}\right)^2 dx.$$

Отсюда $\mathsf{D}\xi = \dfrac{(b-a)^2}{12}$.

3. *Пуассоновское распределение.* Найдем сначала $\mathsf{M}[\xi(\xi-1)]$. По формуле (1.4) с $r = 1$ и $g(x) = x(x-1)$ получим

$$\mathsf{M}\xi(\xi-1) = \sum_{m=0}^{\infty} m(m-1)\frac{\lambda^m}{m!}e^{-\lambda} = \lambda^2\sum_{m=2}^{\infty}\frac{\lambda^{m-2}}{(m-2)!}e^{-\lambda} = \lambda^2.$$

Так как $\mathsf{M}\xi(\xi-1) = \mathsf{M}\xi^2 - \mathsf{M}\xi$ и $\mathsf{M}\xi = \lambda$, то $\mathsf{M}\xi^2 = \lambda^2 + \lambda$. Подставляя это выражение в (3.2), получим $\mathsf{D}\xi = \lambda$.

4. *Биномиальное распределение.* Так же как в примере 2 из § 2, воспользуемся тем, что число успехов μ_n представимо в виде суммы независимых индикаторов (5.5.13). Тогда $\mathsf{D}\mu_n = \mathsf{D}\xi_1 + \ldots + \mathsf{D}\xi_n$ и $\mathsf{D}\xi_k = \mathsf{M}\xi_k^2 - (\mathsf{M}\xi_k)^2 = p - p^2 = p(1-p)$. Таким образом, $\mathsf{D}\mu_n = np(1-p)$. Отметим, что в формулировку теоремы Муавра — Лапласа (теорема 3.3 гл. 4) входила случайная величина $(\mu_n - np)/\sqrt{npq}$. Теперь мы можем эту же величину записать в виде $(\mu_n - \mathsf{M}\mu_n)/\sqrt{\mathsf{D}\mu_n}$. Такое линейное преобразование случайной величины используется довольно часто. Если ξ — произвольная случайная величина, то для случайной величины $\eta = (\xi - \mathsf{M}\xi)/\sqrt{\mathsf{D}\xi}$ имеем

$$\mathsf{M}\eta = 0,\ \mathsf{D}\eta = 1.$$

В приложениях для оценки вероятности отклонения случайной величины от своего математического ожидания часто используют «правило трех сигм», согласно которому событие $|\xi - \mathsf{M}\xi| > 3\sqrt{\mathsf{D}\xi}$ практически невозможно, т. е. его вероятность очень мала. Это действительно так, если ξ распределена нормально. В этом случае

$$\mathsf{P}\left\{|\xi-\mathsf{M}\xi|>3\sqrt{\mathsf{D}\xi}\right\}=$$

$$=\mathsf{P}\left\{\left|\frac{\xi-\mathsf{M}\xi}{\sqrt{\mathsf{D}\xi}}\right|>3\right\}=\frac{2}{\sqrt{2\pi}}\int_3^\infty e^{-\frac{x^2}{2}}\,dx<0{,}003.$$

Однако если распределение ξ отлично от нормального, то возможны большие значения вероятности такого отклонения (см. задачу 19).

Часто изучаемую случайную величину удается представить в виде суммы более простых, возможно зависимых, случайных величин. Пусть, например,

$$\eta_n=\xi_1+\xi_2+\ldots+\xi_n,\tag{3.7}$$

где случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ зависимы и каждая из них принимает значения 0 и 1. В этом случае можно воспользоваться формулой

$$\mathsf{M}\eta_n=\mathsf{M}\xi_1+\mathsf{M}\xi_2+\ldots+\mathsf{M}\xi_n.$$

Однако $\mathsf{D}\eta_n$ уже не равна сумме дисперсий $\mathsf{D}\xi_k$. Для вычисления $\mathsf{D}\eta_n$ можно использовать формулу (3.2). Так как $\xi_k^2(\omega)=\xi_k(\omega)$, $\omega\in\Omega$ (действительно, $\xi_k(\omega)=1$ или $\xi_k(\omega)=0$), то

$$\eta_n^2=\sum_{k=1}^n\xi_k^2+\sum_{k\neq l}\xi_k\xi_l=\sum_{k=1}^n\xi_k+\sum_{k\neq l}\xi_k\xi_l=\eta_n+\sum_{k\neq l}\xi_k\xi_l.\tag{3.8}$$

Таким образом,

$$\mathsf{D}\eta_n=\mathsf{M}\eta_n^2-(\mathsf{M}\eta_n)^2=\sum_{k\neq l}\mathsf{M}\xi_k\xi_l+\mathsf{M}\eta_n-(\mathsf{M}\eta_n)^2.$$

5. *Гипергеометрическое распределение.* Воспользуемся формулой (3.8) для вычисления дисперсии величины ξ, имеющей гипергеометрическое распределение. Для индикаторов ξ_k в сумме (2.2) имеем (см. задачу 22 гл. 2)

$$\mathsf{M}\xi_k\xi_l=\mathsf{P}\{\xi_k=\xi_l=1\}=\frac{M(M-1)}{N(N-1)}.$$

Отсюда и из (3.8) находим

$$\mathsf{M}\xi^2=\mathsf{M}\xi+\sum_{k\neq l}\mathsf{M}\xi_k\xi_l=n\frac{M}{N}+n(n-1)\frac{M(M-1)}{N(N-1)}.$$

Следовательно,

$$\mathsf{D}\xi = n(n-1)\frac{M(M-1)}{N(N-1)} + n\frac{M}{N} - n^2\frac{M^2}{N^2}.$$

После несложных преобразований получим

$$\mathsf{D}\xi = n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}. \tag{3.9}$$

§ 4. Ковариация. Коэффициент корреляции

При доказательстве формулы (3.5) нам потребовалось вычислить $\mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)]$. Это число называется *ковариацией* случайных величин ξ_1, ξ_2 и обозначается $\operatorname{cov}(\xi_1, \xi_2)$. Таким образом,

$$\operatorname{cov}(\xi_1, \xi_2) = \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)]. \tag{4.1}$$

Отсюда, используя свойства математического ожидания, легко получить следующую формулу:

$$\operatorname{cov}(\xi_1, \xi_2) = \mathsf{M}\xi_1\xi_2 - \mathsf{M}\xi_1 \cdot \mathsf{M}\xi_2. \tag{4.2}$$

Очевидно, что

$$\operatorname{cov}(\xi, \xi) = \mathsf{D}\xi.$$

Нетрудно проверить также следующие равенства:

$$\operatorname{cov}(\xi_1, \xi_2) = \operatorname{cov}(\xi_2, \xi_1), \quad \operatorname{cov}(c\xi_1, \xi_2) = c\operatorname{cov}(\xi_1, \xi_2),$$

где c — любая постоянная.

Из равенств (3.6) и определения (4.1) вытекает формула для дисперсии суммы двух произвольных случайных величин:

$$\mathsf{D}(\xi_1 + \xi_2) = \mathsf{D}\xi_1 + \mathsf{D}\xi_2 + 2\operatorname{cov}(\xi_1, \xi_2). \tag{4.3}$$

Теорема 4.1. *Если для случайных величин* $\xi_1, \xi_2, \ldots, \xi_n$ *существуют* $\operatorname{cov}(\xi_i, \xi_j) = \sigma_{ij}$, $i, j = 1, \ldots, n$, *то при любых постоянных* $c_1, c_2, \ldots, c_n$ *имеем*

$$\mathsf{D}(c_1\xi_1 + c_2\xi_2 + \ldots + c_n\xi_n) = \sum_{i,\,j=1}^{n} \sigma_{ij}c_ic_j. \tag{4.4}$$

Доказательство. Положим

$$\eta_n = c_1\xi_1 + c_2\xi_2 + \ldots + c_n\xi_n.$$

Нетрудно проверить, что

$$\eta_n - \mathsf{M}\eta_n = \sum_{i=1}^{n} c_i(\xi_i - \mathsf{M}\xi_i)$$

и

$$(\eta_n - M\eta_n)^2 = \sum_{i,\, j=1}^{n} c_i c_j (\xi_i - \mathsf{M}\xi_i)(\xi_j - \mathsf{M}\xi_j).$$

Вычисляя математическое ожидание от обеих частей последнего равенства, получим утверждение теоремы.

Правую часть (4.4) можно рассматривать как квадратичную форму от переменных $c_1, c_2, \ldots, c_n$. Так как при любых $c_1, c_2, \ldots, c_n$ дисперсия в левой части (4.4) неотрицательна, то квадратичная форма в правой части (4.4) неотрицательно определена. Квадратичная форма неотрицательно определена тогда и только тогда, когда неотрицательны все главные миноры матрицы, составленной из ее коэффициентов. Таким образом, из теоремы 4.1 получили следующее утверждение: определитель $|\mathsf{D}[\xi]|$ матрицы $\mathsf{D}[\xi] = \mathsf{D}[(\xi_1, \ldots, \xi_m)] = \| \operatorname{cov}(\xi_i, \xi_j) \|$ при любом $m = 1, 2, \ldots$ для любых случайных величин $\xi_1, \xi_2, \ldots$ $\ldots, \xi_m$ удовлетворяет неравенству

$$|\mathsf{D}[(\xi_1, \ldots, \xi_m)]| \geqslant 0. \tag{4.5}$$

При $m = 2$ неравенство (4.5) имеет вид

$$\begin{vmatrix} \mathsf{D}\xi_1 & \operatorname{cov}(\xi_1, \xi_2) \\ \operatorname{cov}(\xi_1, \xi_2) & \mathsf{D}\xi_2 \end{vmatrix} = \mathsf{D}\xi_1 \mathsf{D}\xi_2 - \operatorname{cov}(\xi_1, \xi_2) \geqslant 0.$$

Отсюда

$$|\operatorname{cov}(\xi_1, \xi_2)| \leqslant \sqrt{\mathsf{D}\xi_1 \mathsf{D}\xi_2}. \tag{4.6}$$

В доказательстве формулы (3.5) было попутно получено, что для независимых случайных величин ξ_1, ξ_2 имеет место равенство

$$\operatorname{cov}(\xi_1, \xi_2) = 0. \tag{4.7}$$

Таким образом, если $\operatorname{cov}(\xi_1, \xi_2) \neq 0$, то величины ξ_1 и ξ_2 зависимы. В качестве количественной характеристики степени зависимости случайных величин ξ_1 и ξ_2 используется *коэффициент корреляции* $\rho(\xi_1, \xi_2)$, определяемый следующим равенством:

$$\rho(\xi_1, \xi_2) = \frac{\operatorname{cov}(\xi_1, \xi_2)}{\sqrt{\mathsf{D}\xi_1 \mathsf{D}\xi_2}}. \tag{4.8}$$

Свойства коэффициента корреляции:

1°. $|\rho(\xi_1, \xi_2)| \leqslant 1$.

2°. *Если* ξ_1 *и* ξ_2 *независимы, то* $\rho(\xi_1, \xi_2) = 0$.

3°. *Если* $\xi_2 = A\xi_1 + B$, *где* A *и* B — *постоянные, то*

$$|\rho(\xi_1, \xi_2)| = 1.$$

Доказательство. Свойство 1° следует из (4.8) и (4.6); свойство 2° следует из (4.8) и (4.7). Докажем свойство 3°. Положим $\mathsf{M}\xi_1 = a$, $\mathsf{D}\xi_1 = \sigma^2$. Тогда

$$\mathsf{M}\xi_2 = Aa + B, \ \mathsf{D}\xi_2 = A^2\sigma^2,$$

$$\operatorname{cov}(\xi_1, \xi_2) = \mathsf{M}[(\xi_1 - a)(\xi_2 - \mathsf{M}\xi_2)] =$$
$$= \mathsf{M}[(\xi_1 - a)(A(\xi_1 - a))] = A\mathsf{D}\xi_1 = A\sigma^2$$

и, следовательно,

$$\rho(\xi_1, \xi_2) = \frac{A\sigma^2}{\sqrt{A^2\sigma^2\cdot\sigma^2}} = \frac{A}{|A|}.$$

Таким образом, $|\rho(\xi_1, \xi_2)| = 1$.

Отметим, что равенство нулю коэффициента корреляции не является достаточным условием независимости случайных величин. Из равенства $\rho(\xi_1, \xi_2) = 0$ не следует независимость случайных величин (см. задачу 10 этой главы).

Наряду с рассмотренными выше числовыми характеристиками случайных величин часто используются моменты более высоких порядков. *Моментом порядка k* случайной величины ξ называется число $\mathsf{M}\xi^k$. Число $\mathsf{M}(\xi - \mathsf{M}\xi)^k$ называется *центральным моментом порядка k*.

Пусть задан случайный вектор $(\xi_1, \xi_2, \ldots, \xi_n)$. Величины

$$\mathsf{M}\xi_1^{k_1}\xi_2^{k_2}\ldots\xi_n^{k_n}, \quad \mathsf{M}(\xi_1 - \mathsf{M}\xi_1)^{k_1}\ldots(\xi_n - \mathsf{M}\xi_n)^{k_n}$$

называются соответственно *смешанным моментом порядка $k = k_1 + \ldots + k_n$*, и *смешанным центральным моментом порядка k*. Вычислять моменты более высоких порядков можно по формулам (1.4), (1.5). Например,

$$\mathsf{M}\xi_1^{k_1}\xi_2^{k_2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1^{k_1}x_2^{k_2}p_{\xi_1\xi_2}(x_1, x_2)\,dx_1\,dx_2.$$

Отметим также, что из существования момента $\mathsf{M}\xi^m$ вытекает существование моментов $\mathsf{M}\xi^k$, $k = 1, \ldots, m - 1$. Это утверждение следует из неравенств

$$|\xi(\omega)|^k \leqslant |\xi(\omega)|^m + 1, \ \omega \in \Omega, \ k = 1, 2, \ldots, m - 1.$$

§ 5. Закон больших чисел

Во введении был отмечен экспериментальный факт, состоящий в том, что в длинной серии опытов частота появления события A сближается с определенным числом, которое можно рассматривать как вероятность события A. В математической модели серии опытов этот факт доказан. Сначала получим некоторые оценки распределений случайных величин.

Теорема 5.1. *Пусть* $\xi = \xi(\omega) \geqslant 0$ *при любом* $\omega \in \Omega$. *Если* $\mathsf{M}\xi$ *существует, то при любом* $\varepsilon > 0$

$$\mathsf{P}(\xi \geqslant \varepsilon) \leqslant \frac{\mathsf{M}\xi}{\varepsilon}. \tag{5.1}$$

Доказательство. Проведем доказательство в случае, когда ξ задана в абсолютно непрерывном вероятностном пространстве (см. § 4 гл. 2). По определению математического ожидания имеем

$$\mathsf{M}\xi = \int \underset{\Omega}{\cdots} \int \xi(u_1, \ldots, u_n)\, \pi(u_1, \ldots, u_n)\, du_1\, du_2 \ldots du_n.$$

Пусть

$$\Omega_\varepsilon = \{(u_1, \ldots, u_n):\ \xi(u_1, \ldots, u_n) \geqslant \varepsilon\} \subseteq \Omega.$$

Введем случайную величину

$$\eta = \begin{cases} 0, & \text{если } (u_1, \ldots, u_n) \in \Omega \setminus \Omega_\varepsilon, \\ \varepsilon, & \text{если } (u_1, \ldots, u_n) \in \Omega_\varepsilon. \end{cases}$$

При любом $(u_1, \ldots, u_n) \in \Omega$ имеем $\xi \geqslant \eta$. Умножим обе части этого неравенства на $\pi(u_1, \ldots, u_n)$ и проинтегрируем по Ω. Получим, что $\mathsf{M}\xi \geqslant \mathsf{M}\eta$. Отсюда следует утверждение теоремы, так как

$$\mathsf{M}\eta = \varepsilon \mathsf{P}(\Omega_\varepsilon) = \varepsilon \mathsf{P}(\xi \geqslant \varepsilon).$$

Доказанная теорема позволяет легко получить неравенство Чебышева.

Теорема 5.2 (неравенство Чебышева). *Если случайная величина* ξ *имеет дисперсию, то при любом* $\varepsilon > 0$

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geqslant \varepsilon) \leqslant \frac{\mathsf{D}\xi}{\varepsilon^2}.$$

Доказательство. Случайная величина $\eta = (\xi - \mathsf{M}\xi)^2 \geqslant 0$ при всех $\omega \in \Omega$, и $\mathsf{M}\eta = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \mathsf{D}\xi$ конечно. Следовательно, можно воспользоваться неравенством (5.1). Таким образом,

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geqslant \varepsilon) = \mathsf{P}(\eta \geqslant \varepsilon^2) \leqslant \frac{\mathsf{M}\eta}{\varepsilon^2} = \frac{\mathsf{D}\xi}{\varepsilon^2},$$

Неравенство Чебышева позволяет оценивать вероятности отклонений значений случайной величины от своего математического ожидания.

Пусть, например, нужно оценить долю бракованных изделий в партии, содержащей N изделий. Обозначим число бракованных изделий M. По схеме случайного выбора без возвращения отберем n изделий. Пусть среди отобранных изделий ξ бракованных. Случайная величина ξ имеет гипергеометрическое распределение (см. § 1 гл. 2). В §§ 2 и 3 было показано, что

$$\mathsf{M}\xi = n\frac{M}{N}, \quad \mathsf{D}\xi = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}.$$

Отсюда

$$\mathsf{M}\left(\frac{\xi}{n}\right) = \frac{M}{N}, \quad \mathsf{D}\left(\frac{\xi}{n}\right) = \frac{1}{n}\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}.$$

Воспользовавшись неравенством Чебышева, получим

$$\mathsf{P}\left\{\left|\frac{\xi}{n} - \frac{M}{N}\right| \geqslant \Delta\right\} \leqslant \frac{1}{n\Delta^2}\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}.$$

Другим примером полезного применения неравенства Чебышева является оценка ошибки приближенного значения измеряемой величины. Пусть проводится n независимых измерений некоторой неизвестной величины a. Ошибки измерения $\delta_1, \delta_2, \ldots, \delta_n$ будем считать случайными величинами. Предположим, что $\mathsf{M}\delta_k = 0$, $k = 1, 2, \ldots, n$. Это условие можно рассматривать как отсутствие систематической ошибки. Пусть еще $\mathsf{D}\delta_k = b^2$. За значение неизвестной величины a принимают обычно среднее арифметическое результатов измерений. Тогда ошибка в определении числа a будет равна

$$\eta_n = \frac{\delta_1 + \delta_2 + \ldots + \delta_n}{n}$$

и

$$\mathsf{D}\eta_n = \frac{1}{n^2}(\mathsf{D}\delta_1 + \ldots + \mathsf{D}\delta_n) = \frac{b^2}{n}, \quad \mathsf{M}\eta_n = 0.$$

Предположим, что нам нужно, чтобы ошибка η_n не превосходила Δ с достаточно большой вероятностью. Например,

$$\mathsf{P}(|\eta_n| < \Delta) > 0{,}99.$$

Это неравенство можно записать в эквивалентном виде

$$\mathsf{P}(|\eta_n| \geqslant \Delta) \leqslant 0{,}01. \qquad (5.2)$$

По неравенству Чебышева имеем

$$\mathsf{P}(|\eta_n| \geqslant \Delta) \leqslant \frac{\mathsf{D}\eta_n}{\Delta^2} = \frac{b^2}{n\Delta^2}.$$

Следовательно, (5.2) будет выполнено, если

$$\frac{b^2}{n\Delta^2} \leqslant 0{,}01, \text{ или } n \geqslant 100\frac{b^2}{\Delta^2}.$$

Таким образом, мы получили оценку числа измерений, необходимого для получения заданной точности. В рассматриваемой задаче оценка для n является завышенной. Ее можно улучшить, если воспользоваться тем, что η_n является суммой независимых случайных величин. Будет показано (см. теорему 4.1 гл. 7), что при больших n величина η_n имеет распределение, близкое к нормальному. Однако если о случайной величине ничего не известно, кроме математического ожидания и дисперсии, то оценку, которую дает неравенство Чебышева, улучшить нельзя. Укажем распределение случайной величины, для которой неравенство Чебышева при данных $b^2 = \mathsf{D}\xi$ и $\varepsilon > b$ обращается в равенство. Пусть

$$\mathsf{P}(\xi = 0) = 1 - \frac{b^2}{\varepsilon^2}, \quad \mathsf{P}(\xi = -\varepsilon) = \mathsf{P}(\xi = \varepsilon) = \frac{b^2}{2\varepsilon^2}.$$

Тогда

$$\mathsf{M}\xi = 0, \quad \mathsf{D}\xi = \mathsf{M}\xi^2 = (-\varepsilon)^2\frac{b^2}{2\varepsilon^2} + \varepsilon^2\frac{b^2}{2\varepsilon^2} = b^2$$

и

$$\mathsf{P}(|\xi| \geqslant \varepsilon) = \frac{b^2}{\varepsilon^2}.$$

Нетрудно получить еще ряд полезных неравенств типа неравенства Чебышева. Пусть $f(x)$ — неубывающая неотрицательная функция. Если существует $\mathsf{M}f(\xi)$, то при любом $\varepsilon > 0$

$$\mathsf{P}(\xi \geqslant \varepsilon) \leqslant \frac{\mathsf{M}f(\xi)}{f(\varepsilon)}. \qquad (5.3)$$

Доказательство проводится так же, как доказательство теоремы 5.1.

Теорема 5.3. *Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ попарно независимы и*

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} \mathsf{D}\xi_k = 0, \tag{5.4}$$

то для любого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{\xi_1+\xi_2+\ldots+\xi_n}{n} - \frac{\mathsf{M}\xi_1+\mathsf{M}\xi_2+\ldots+\mathsf{M}\xi_n}{n}\right| < \varepsilon\right) = 1.$$

Доказательство. Положим

$$\eta_n = \frac{\xi_1+\xi_2+\ldots+\xi_n}{n}.$$

Утверждение теоремы равносильно тому, что при любом $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}(|\eta_n - \mathsf{M}\eta_n| \geqslant \varepsilon) = 0. \tag{5.5}$$

Так как случайные величины $\xi_1, \ldots, \xi_n$ попарно независимы, то

$$\mathsf{D}\eta_n = \frac{1}{n^2} \sum_{k=1}^{n} \mathsf{D}\xi_k. \tag{5.6}$$

По неравенству Чебышева

$$\mathsf{P}(|\eta_n - \mathsf{M}\eta_n| \geqslant \varepsilon) \leqslant \frac{\mathsf{D}\eta_n}{\varepsilon^2}.$$

Отсюда, воспользовавшись (5.6) и (5.4), получим (5.5). Теорема доказана.

Случайные величины $\xi_1, \ldots, \xi_n, \ldots$ называют *некоррелированными*, если $\operatorname{cov}(\xi_i, \xi_j) = 0$ при любых $i, j, i \neq j$.

В условии теоремы 5.3 можно вместо попарной независимости величин $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ потребовать, чтобы они были некоррелированными, так как для некоррелированных величин сохранится формула (5.6). Если для величин $\xi_1, \ldots, \xi_n, \ldots$ выполнено утверждение теоремы 5.3, то говорят, что к ним применим закон больших чисел.

Отметим некоторые частные случаи этой теоремы.

Теорема 5.4 (теорема Чебышева). *Если случайные величины* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *попарно независимы и*

$$\mathsf{D}\xi_k \leqslant C, \quad k = 1, 2, \ldots, \tag{5.7}$$

где C *— некоторая постоянная, то при любом* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{\left|\frac{\xi_1 + \ldots + \xi_n}{n} - \frac{\mathsf{M}\xi_1 + \ldots + \mathsf{M}\xi_n}{n}\right| < \varepsilon\right\} = 1.$$

Теорема 5.4 следует из теоремы 5.3, так как из условия (5.7) следует (5.4).

Теорема 5.5. *Если случайные величины* $\xi_1, \ldots, \xi_n, \ldots$ *одинаково распределены, попарно независимы и имеют конечные дисперсии, то при любом* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{\left|\frac{\xi_1 + \ldots + \xi_n}{n} - a\right| < \varepsilon\right\} = 1,$$

где $a = \mathsf{M}\xi_k$.

Теорема 5.5 следует из теоремы 5.4. Действительно, $\mathsf{D}\xi_k$, $k = 1, 2, \ldots$, существуют и равны между собой. Следовательно, выполнено условие (5.7). В гл. 7 теорема 5.5 при помощи характеристических функций будет доказана без предположения о существовании дисперсии.

Теорема 5.6 (теорема Бернулли). *Пусть* μ_n *— число успехов в* n *испытаниях Бернулли и* p *— вероятность успеха в отдельном испытании. Тогда при любом* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right\} = 1.$$

Доказательство. Воспользуемся для μ_n представлением

$$\mu_n = \xi_1 + \xi_2 + \ldots + \xi_n,$$

где ξ_k, $k = 1, 2, \ldots, n$, независимы и $\mathsf{P}(\xi_k = 1) = p$, $\mathsf{P}(\xi_k = 0) = 1 - p = q$. Очевидно, что $\mathsf{M}\xi_k = p$ и дисперсии величин ξ_k конечны.

Таким образом, доказываемая теорема сразу следует из теоремы 5.5.

Схема Бернулли является математической моделью серии опытов, повторяющихся в неизменных условиях. В каждом опыте может произойти событие A, которое мы назвали успехом. Согласно теореме Бернулли частота μ_n/n наступления события A сближается с вероятностью p. Этот же факт установлен экспериментально, Математической моделью последовательности из n измерений неизвестной величины a является случайный вектор $(\xi_1, \xi_2, \ldots, \xi_n)$, где ξ_k, $k = 1, 2, \ldots, n$, независимы, одинаково распределены, $\mathsf{M}\xi_k = a$, $\mathsf{D}\xi_k = b^2$. По теореме 5.5 среднее арифметическое $(\xi_1 + \ldots + \xi_n)/n$ при больших n мало отличается от измеряемой величины с вероятностью, близкой к 1.

§ 6. Условные распределения и условные математические ожидания

Пусть в пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$ определен случайный вектор (ξ, η). Введем понятие условного распределения величины ξ при условии, что задано значение η. Рассмотрим сначала дискретный случай. Пусть

$$\mathsf{P}(\xi = x_i, \eta = y_j) = p_{ij} \geqslant 0, \quad \sum_{i,\, j=1}^{\infty} p_{ij} = 1,$$

$$\mathsf{P}(\xi = x_i) = \sum_{j=1}^{\infty} p_{ij} = p_{i\cdot} \geqslant 0, \quad \mathsf{P}(\eta = y_j) = \sum_{i=1}^{\infty} p_{ij} = p_{\cdot j} > 0.$$

В § 1 гл. 3 было дано определение условной вероятности. По этому определению

$$\mathsf{P}(\xi = x_i \mid \eta = y_j) = \frac{\mathsf{P}(\xi = x_i, \eta = y_j)}{\mathsf{P}(\eta = y_j)} = \frac{p_{ij}}{p_{\cdot j}}. \tag{6.1}$$

Если фиксировать y_j (или j), то вероятности (6.1) можно рассматривать как *условное распределение величины* ξ *при условии, что* $\eta = y_j$. *Условным математическим ожиданием случайной величины* ξ *при условии, что* $\eta = y_j$, называется число

$$\mathsf{M}(\xi \mid \eta = y_j) = \sum_{i=1}^{\infty} x_i \frac{p_{ij}}{p_{\cdot j}}. \tag{6.2}$$

Если левые части (6.1) и (6.2) рассматривать как функции от y_j, то можно считать условное распределение и условное математическое ожидание случайными величинами, определенными в исходном вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$. Тогда условное распределение и условное математическое ожидание ξ при условии η определяются соответственно формулами

$$\mathsf{P}(\xi = x_i \mid \eta) = \mathsf{P}(\xi = x_i \mid \eta = y_j), \quad \text{если} \quad \omega \in \{\eta = y_j\},\ j = 1, 2, \ldots;$$

$$\mathsf{M}(\xi \mid \eta) = \mathsf{M}(\xi \mid \eta = y_j), \quad \text{если} \quad \omega \in \{\eta = y_j\},\ j = 1, 2, \ldots,$$

где $\mathsf{P}(\xi = x_i \mid \eta = y_j)$, $\mathsf{M}(\xi \mid \eta = y_j)$ определены формулами (6.1) и (6.2). Справедлива следующая формула (*формула полного математического ожидания*):

$$\mathsf{M}\xi = \mathsf{M}[\mathsf{M}(\xi \mid \eta)]. \tag{6.3}$$

Здесь условное математическое ожидание в квадратных скобках рассматривается как случайная величина. Применяя формулу (1.6) к случайной величине $\mathsf{M}(\xi \mid \eta)$, можно (6.3) записать в следующей эквивалентной форме:

$$\mathsf{M}\xi = \sum_{j=1}^{\infty} \mathsf{P}(\eta = y_j)\, \mathsf{M}(\xi \mid \eta = y_j). \tag{6.4}$$

Докажем формулу (6.4). Воспользовавшись определением (6.2), получим

$$\sum_{j=1}^{\infty} \mathsf{P}(\eta = y_j)\, \mathsf{M}(\xi \mid \eta = y_j) = \sum_{j=1}^{\infty} \mathsf{P}(\eta = y_j) \sum_{i=1}^{\infty} x_i \frac{\mathsf{P}(\xi = x_i, \eta = y_j)}{\mathsf{P}(\eta = y_j)} =$$

$$= \sum_{i,\, j=1}^{\infty} x_i \mathsf{P}(\xi = x_i, \eta = y_j) = \sum_{i=1}^{\infty} x_i \left(\sum_{j=1}^{\infty} \mathsf{P}(\xi = x_i, \quad \eta = y_j) \right).$$

Так как $\sum\limits_{j=1}^{\infty} \mathsf{P}(\xi = x_i, \eta = y_j) = \mathsf{P}(\xi = x_i)$, то правая часть последнего равенства равна $\mathsf{M}\xi$. Покажем, что из формул (6.3) и (6.4) следует формула полной вероятности. Пусть

$$\xi = \begin{cases} 1, & \text{если } \omega \in A, \\ 0, & \text{если } \omega \in \bar{A}, \end{cases}$$

$$\eta = y_j, \text{ если } \omega \in B_j,\ j = 1, 2, \ldots,$$

где $B_i B_j = \varnothing$, $i \neq j$, $\bigcup\limits_{j=1}^{\infty} B_j = \Omega$. Тогда

$$\mathsf{M}\xi = \mathsf{P}(A),\ \ \mathsf{P}(\eta = y_j) = \mathsf{P}(B_j),\ \ \mathsf{M}(\xi \mid \eta = y_j) = \mathsf{P}(A \mid B_j).$$

Подставляя эти выражения в (6.4), получим формулу полной вероятности со счетной системой событий $B_1, B_2, \ldots, B_n, \ldots$:

$$\mathsf{P}(A) = \sum_{n=1}^{\infty} \mathsf{P}(B_k)\, \mathsf{P}(A \mid B_k). \tag{6.5}$$

Положим в (6.5) $A = \{\xi \in C\}$, $B_k = \{\eta = y_k\}$, где $C \subset \{x_1, x_2, \ldots$ $\ldots, x_n, \ldots\}$, (x_i, y_j) — значения (ξ, η). Тогда при любом C

$$\mathsf{P}(\xi \in C) = \sum_{k=1}^{\infty} \sum_{x_i \in C} \mathsf{P}(\eta = y_k)\, \mathsf{P}(\xi = x_i \mid \eta = y_k). \tag{6.6}$$

Свойства условного математического ожидания:

1°. $\mathsf{M}[\varphi(\eta) \mid \eta] = \varphi(\eta)$. (6.7)

2°. $\mathsf{M}[\varphi(\eta) \cdot \xi \mid \eta] = \varphi(\eta)\, \mathsf{M}(\xi \mid \eta)$. (6.8)

3°. $\mathsf{M}(\xi_1 + \xi_2 \mid \eta) = \mathsf{M}(\xi_1 \mid \eta) + \mathsf{M}(\xi_2 \mid \eta)$. (6.9)

4°. *Если* ξ *и* η *независимы, то*

$$\mathsf{M}(\xi \mid \eta) = \mathsf{M}\xi. \tag{6.10}$$

Равенства (6.7) — (6.10) верны при любом $\omega \in \Omega$. Равенства (6.7), (6.8), (6.10) следуют непосредственно из определения, аналогичного (1.2), (1.3).

Перейдем к рассмотрению абсолютно непрерывного вектора (ξ, η). Так как в этом случае $\mathsf{P}(\eta = y) = 0$ при любом y, то мы не сможем воспользоваться определением условной вероятности (2.1.1), как это мы сделали в дискретном случае. Назовем *условной плотностью распределения вероятностей величины* ξ *при условии, что* $\eta = y$, следующую функцию:

$$p_\xi(x \mid \eta = y) = \frac{p_{\xi,\eta}(x, y)}{\int\limits_{-\infty}^{\infty} p_{\xi,\eta}(u, y)\, du}. \tag{6.11}$$

Условное математическое ожидание ξ *при условии, что* $\eta = y$, *определяется формулой*

$$\mathsf{M}(\xi \mid \eta = y) = \int\limits_{-\infty}^{\infty} x p_\xi(x \mid \eta = y)\, dx. \tag{6.12}$$

Из (6.11) нетрудно получить, что при любых a и b

$$\mathsf{P}(a \leqslant \xi \leqslant b) = \int\limits_{-\infty}^{\infty} p_\eta(y) \left(\int\limits_a^b p_\xi(x \mid \eta = y)\, dx\right) dy. \tag{6.13}$$

Легко также проверяется формула, аналогичная (6.4):

$$\mathsf{M}\xi = \int\limits_{-\infty}^{\infty} p_\eta(y)\, \mathsf{M}(\xi \mid \eta = y)\, dy. \tag{6.14}$$

Если рассматривать (6.11) и (6.12) как случайные величины

$$p_\xi(x \mid \eta) = p_\xi(x \mid \eta = y), \text{ если } \omega \in \{\eta = y\}, \quad y \in (-\infty, \infty);$$

$$\mathsf{M}(\xi \mid \eta = y), \text{ если } \omega \in \{\eta = y\}, \ \mathsf{M}(\xi \mid \eta) = y \in (-\infty, \infty),$$

то формулы (6.13) и (6.14) можно записать в виде

$$\mathsf{P}(a \leqslant \xi \leqslant b) = \mathsf{M}\left(\int\limits_a^b p_\xi(x \mid \eta)\, dx\right)$$

и в виде (6.3) соответственно. Для условного математического ожидания абсолютно непрерывных величин сохраняются свойства (6.7) — (6.10).

§ 7. Многомерное нормальное распределение

Пусть $\xi = (\xi_1, \ldots, \xi_r)'$ — r-мерный случайный вектор-столбец *). Положим

$$\eta_i = \sum_{l=1}^{r} c_{il}\xi_l + a_i, \quad i = 1, \ldots, m, \tag{7.1}$$

или в векторной форме:

$$\eta = C\xi + a, \tag{7.2}$$

где $\eta = (\eta_1, \ldots, \eta_m)'$, $a = (a_1, \ldots, a_r)'$ и $C = \| c_{ij} \|$ —$(m \times r)$-матрица. Здесь вектор a и матрица C постоянны. Как и в § 4, будем обозначать $D[\xi] = D[\xi_1, \ldots \xi_r]$ ковариационную матрицу $\| \operatorname{cov}(\xi_i, \xi_j) \|$. Следующая теорема является обобщением теоремы 4.1.

Теорема 7.1. *При любой постоянной матрице C в* (7.2) *имеем*

$$\mathsf{D}[\eta] = C\mathsf{D}[\xi]\, C'. \tag{7.3}$$

Доказательство. Из формулы (7.1) следует, что

$$\eta_i - \mathsf{M}\eta_i = \sum_{l=1}^{r} c_{il}(\xi_l - \mathsf{M}\xi_l)$$

и

$$(\eta_i - \mathsf{M}\eta_i)(\eta_j - \mathsf{M}\eta_j) = \sum_{k,\, l=1}^{r} c_{il}c_{jk}(\xi_l - \mathsf{M}\xi_l)(\xi_k - \mathsf{M}\xi_k).$$

Вычисляя от обеих частей последнего равенства математическое ожидание, получим

$$\operatorname{cov}(\eta_i, \eta_j) = \sum_{k,\, l=1}^{r} c_{il} \operatorname{cov}(\xi_l, \xi_k)\, c'_{kj},$$

где $c'_{kj} = c_{jk}$. Теорема доказана.

Многомерным нормальным распределением m случайных величин назовем распределение вектора $\eta = (\eta_1, \ldots, \eta_m)'$, где η_i определены формулой (7.1), *а случайные величины $\xi_1, \xi_2, \ldots, \xi_r$ независимы и каждая из них распределена нормально с параметрами* (0, 1).

Если $m = r$ и определитель $|C| \neq 0$, то нормальное распределение назовем *невырожденным*. В других случаях

---

*) Мы будем r-мерные векторы рассматривать как $(r \times 1)$-матрицы. Знаком $'$ будем обозначать транспонирование.

можно показать, что распределение η сосредоточено на подпространстве размерности, меньшей m. Отметим некоторые общие свойства многомерного распределения.

Пусть вектор $\eta = (\eta_1, \ldots, \eta_m)'$ *распределен нормально. Тогда:*

1°. *Одномерные распределения координат* η_i *являются нормальными, если* $\mathsf{D}\eta_i > 0$.

2°. *Любая линейная функция* $\eta = A_1\eta_1 + \ldots + A_m\eta_m$ *имеет нормальное распределение, если* $\mathsf{D}\eta > 0$.

3°. *Любое линейное преобразование* $\zeta = A\eta$, *где* A — *постоянная матрица, имеет многомерное нормальное распределение.*

Все эти свойства непосредственно следуют из определения и того факта, что сумма независимых нормально распределенных величин распределена нормально. Нормальность распределения суммы может быть доказана при помощи формулы (5.6.7) или при помощи характеристических функций (см. § 2 гл. 7).

Теорема 7.2. *Невырожденное r-мерное нормальное распределение является абсолютно непрерывным, и его плотность распределения имеет вид*

$$p_\eta(x_1, \ldots, x_r) = \frac{1}{(2\pi)^{r/2}\sqrt{|B|}}\, e^{-\frac{1}{2}Q(x-a)}, \tag{7.4}$$

где $B = \| b_{kl} \|$ — *невырожденная* $(r \times r)$-*матрица*, $b_{kl} = \operatorname{cov}(\eta_k, \eta_l)$, $|B|$ — *определитель матрицы* B, $x = (x_1, \ldots, x_r)$, $a = (a_1, \ldots, a_r)$, $a_k = \mathsf{M}\eta_k$, $Q(x) = \sum b^*_{kl}x_kx_l$ — *квадратичная форма, коэффициенты которой образуют матрицу* $\| b^*_{kl} \|$, *обратную к матрице* B.

Доказательство. Плотность распределения величины $\xi = (\xi_1, \ldots, \xi_r)$ имеет вид

$$p_\xi(x_1, \ldots, x_r) = \frac{1}{(\sqrt{2\pi})^r}\, e^{-\frac{1}{2}\sum\limits_{k=r}^{r} x_k^2}.$$

Так как $|C| \neq 0$, то преобразование $y = g(x)$, где $y = (y_1, \ldots, y_r)$, $x = (x_1, \ldots, x_r)$,

$$y_i = g_i(x_1, \ldots, x_r) = \sum_{l=1}^{r} c_{il}x_l + a_i,$$

является взаимно однозначным и для вычисления плотности распределения η можно воспользоваться формулой

(5.6.5). Поскольку $\frac{\partial g_i}{\partial x_l} = c_{il}$, то якобиан преобразования J имеет вид $J = |C|$ и $x_i = \sum_{l=1}^{r} c_{il}^{-1}(y_l - a_l)$, где c_{il}^{-1} — элементы матрицы, обратной к C. Из формулы (5.6.5) получим

$$p_\eta(y_1, \ldots, y_r) = |J|^{-1} p_\xi(g^{-1}(y)) = \\ = (\sqrt{2\pi})^{-r} |J|^{-1} e^{-\frac{1}{2}Q(y-a)}, \qquad (7,5)$$

где

$$Q(y-a) = \sum_{k=1}^{r} x_k^2 = \sum_{k=1}^{r} \Big(\sum_{l=1}^{r} c_{kl}^{-1}(y_l - a_l) \Big)^2 = \\ = \sum_{l_1, l_2=1}^{r} (y_{l_1} - a_{l_1})(y_{l_2} - a_{l_2}) \Big(\sum_{k=1}^{r} c_{kl_1}^{-1} c_{kl_2}^{-1} \Big). \qquad (7.6)$$

Отсюда следует утверждение теоремы.

Действительно, $\mathsf{D}[\xi] = E$ (E — единичная матрица), и согласно теореме 7.1 имеем $B = C\mathsf{D}[\xi]C' = CC'$. Следовательно, $\|b_{l_1 l_2}^*\| = B^{-1} = (C')^{-1}(C)^{-1} = \Big\| \sum_{k=1}^{r} c_{kl_1}^{-1} c_{kl_2}^{-1} \Big\|$ и $|B| = |C|^2$. Заменяя в (7.5) и (7.6) $|J| = \sqrt{|B|}$ и $\sum_{k=1}^{r} c_{kl_1}^{-1} c_{kl_2}^{-1}$ найденными выражениями, получим (7.4).

Таким образом, параметры невырожденного многомерного нормального распределения определяются первыми и вторыми моментами. В § 5 было показано, что равенство нулю ковариации является необходимым условием независимости случайных величин. Если $b_{kl} = \operatorname{cov}(\eta_k, \eta_l) = 0$, $k \neq l$, то (7.4) распадается на произведение одномерных плотностей распределения. Таким образом, равенства $b_{kl} = 0$, $k \neq l$, необходимы и достаточны для независимости координат случайных векторов, имеющих плотности распределения (7.4).

При небольших значениях r можно явно выразить величины b_{kl}^* через b_{kl}. Пусть, например, $r = 2$. Тогда

$$B = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad B^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{pmatrix},$$

$|B| = (1-\rho^2)\sigma_1^2\sigma_2^2$ и, следовательно,

$$p_{\eta_1\eta_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(x_1-a_1,\, x_2-a_2)}, \qquad (7,7)$$

где

$$Q(y_1, y_2) = \frac{1}{1-\rho^2}\left(\frac{y_1^2}{\sigma_1^2} + \frac{y_2^2}{\sigma_2^2} - \frac{2\rho}{\sigma_1\sigma_2} y_1 y_2\right).$$

Уже отмечалось (свойства 1°, 3°), что как одномерные распределения величин η_1 и η_2, так и совместное распределение величин

$$\zeta_1 = c_{11}\eta_1 + c_{12}\eta_2, \quad \zeta_2 = c_{21}\eta_1 + c_{22}\eta_2$$

нормальны. По теореме 7.2 имеем $\mathsf{M}\eta_1 = a_1$, $\mathsf{D}\eta_1 = \sigma_1^2$, $\mathsf{M}\eta_2 = a_2$, $\mathsf{D}\eta_2 = \sigma_2^2$, $\operatorname{cov}(\eta_1, \eta_2) = \rho\sigma_1\sigma_2$, и, следовательно, можно выписать формулы для одномерных плотностей ζ_1 и ζ_2. Используя свойства математических ожиданий, нетрудно найти параметры распределения вектора (ζ_1, ζ_2).

Действительно,

$$\mathsf{M}\zeta_1 = c_{11}a_1 + c_{12}a_2,$$

$$\mathsf{M}\zeta_2 = c_{21}a_1 + c_{22}a_2,$$

$$\mathsf{D}\zeta_1 = c_{11}^2\sigma_1^2 + c_{12}^2\sigma_2^2 + 2c_{11}c_{12}\rho\sigma_1\sigma_2,$$

$$\mathsf{D}\zeta_2 = c_{21}^2\sigma_1^2 + c_{22}^2\sigma_2^2 + 2c_{21}c_{22}\rho\sigma_1\sigma_2,$$

$$\operatorname{cov}(\zeta_1, \zeta_2) = c_{11}c_{21}\sigma_1^2 + c_{22}c_{12}\sigma_2^2 + (c_{11}c_{22} + c_{21}c_{12})\rho\sigma_1\sigma_2.$$

Если случайные величины η_1, η_2 независимы, $\sigma_1 = \sigma_2$ и матрица C ортогональна, то величины ζ_1, ζ_2 независимы, так как $\rho = 0$,

$$c_{11}^2 + c_{12}^2 = c_{21}^2 + c_{22}^2 = 1, \quad c_{11}c_{21} + c_{22}c_{12} = 0$$

и, следовательно, $\operatorname{cov}(\zeta_1, \zeta_2) = 0$.

Можно найти и условную плотность распределения η_2 при условии, что фиксировано η_1. Из формулы

$$p_{\eta_1}(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x_1-a_1)^2}{2\sigma_1^2}}$$

и формул (7.7), (6.11) получим

$$p_{\eta_2}(x \mid \eta_1) = p_{\eta_1\eta_2}(\eta_1, x)/p_{\eta_1}(\eta_1) =$$
$$= \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma^2}} \exp\left\{-\frac{1}{2(1-\rho^2)\sigma_2^2}\left(x - a_2 - \rho\frac{\sigma_2}{\sigma_1}(\eta_1 - a_1)\right)^2\right\}.$$

Таким образом, условное распределение оказалось нормальным и

$$\mathsf{M}(\eta_2 \mid \eta_1) = a_2 + \rho \frac{\sigma_2}{\sigma_1}(\eta_1 - a_1),$$

$$\mathsf{D}(\eta_2 \mid \eta_1) = (1 - \rho^2)\,\sigma_2^2.$$

Задачи к главе 6

1. Найти математическое ожидание величины τ, определенной в задаче 14 гл. 5.

2. Обозначим ξ номер испытания, в котором появился нужный ключ (см. пример 3 из § 1 гл. 2). Найти $\mathsf{M}\xi$.

3. Решить задачу 2 в случае с возвращением ключей.

4. Найти $\mathsf{M}\tau^{(m)}$ величины $\tau^{(m)}$, определенной в задаче 16 гл. 5.

5. В задаче 17 гл. 3 обозначим τ время свободного пробега молекулы. Найти $\mathsf{M}\tau$, $\mathsf{D}\tau$.

6. Найти $\mathsf{M}(\xi_1 + \xi_2)$ и $\mathsf{D}(\xi_1 + \xi_2)$, где ξ_1, ξ_2 определены в задаче 22 гл. 5.

7. Пусть ξ — число комбинаций НУ в $n + 1$ испытаниях схемы Бернулли. Найти $\mathsf{M}\xi$, $\mathsf{D}\xi$.

8. Из 100 карточек с числами 00, 01, 02, . . ., 98, 99 наудачу вынимается одна. Пусть η_1, η_2 — соответственно сумма и произведение цифр на вынутой карточке. Найти $\mathsf{M}\eta_1$, $\mathsf{M}\eta_2$, $\mathsf{D}\eta_1$, $\mathsf{D}\eta_2$.

9. Для величин ξ_1, ξ_2 определенных в задаче 11 гл. 5, найти $\mathsf{M}\xi_1$, $\mathsf{M}\xi_2$, $\mathsf{D}\xi_1$, $\mathsf{D}\xi_2$, $\operatorname{cov}(\xi_1, \xi_2)$.

10. Совместное распределение величин ξ_1, ξ_2 определяется формулами $\mathsf{P}(\xi_1 = 0,\ \xi_2 = 1) = \mathsf{P}(\xi_1 = 0,\ \xi_2 = -1) = \mathsf{P}(\xi_1 = 1,\ \xi_2 = 0) = \mathsf{P}(\xi_1 = -1,\ \xi_2 = 0) = 1/4$. Найти $\mathsf{M}\xi_1$, $\mathsf{M}\xi_2$, $\mathsf{D}\xi_1$, $\mathsf{D}\xi_2$, $\operatorname{cov}(\xi_1, \xi_2)$. Являются ли ξ_1, ξ_2 независимыми величинами?

11. Случайные величины ξ_1, ξ_2, ξ_3, ξ_4, ξ_5 независимы; $\mathsf{D}\xi_i = \sigma^2$. Найти коэффициент корреляции величин а) $\xi_1 + \xi_2$, $\xi_3 + \xi_4 + \xi_5$; б) $\xi_1 + \xi_2 + \xi_3$, $\xi_3 + \xi_4 + \xi_5$.

12. Пусть $(\xi_1, \xi_2, \ldots, \xi_N)$ — дискретный случайный вектор с полиномиальным распределением (4.2.7). Найти $\mathsf{M}\xi_k$, $\operatorname{cov}(\xi_k, \xi_l)$, $k, l = 1, \ldots, N$.

13. Для величины μ_0, определенной в задаче 13 гл. 2, найти $\mathsf{P}(\mu_0 = 0)$.

Указание. Пусть $A_k = \{k\text{-я ячейка осталась пустой}\}$, $k = 1, \ldots, N$. Использовать равенство $\{\mu_0 > 0\} = A_1 + \ldots + A_N$.

14. Для величины μ_0, определенной в задаче 13 гл. 2, найти $\mathsf{M}\mu_0$, $\mathsf{D}\mu_0$. Найти асимптотические формулы при $N \to \infty$, $\dfrac{n}{N} = \alpha = \text{const}$.

15. В задаче 9 гл. 2 оценить с точностью до 0,004 вероятность p появления хотя бы один раз тяжелых поездов в двух соседних интервалах.

Указание. Доказать, что

$$\sum_{k=1}^{r} \mathsf{P}(A_k) - \sum_{1 \leqslant k < l \leqslant r} \mathsf{P}(A_k A_l) \leqslant \mathsf{P}(A_1 + \ldots + A_r) \leqslant \sum_{k=1}^{r} \mathsf{P}(A_k).$$

16. По n конвертам случайно разложили n писем различным адресатам. Найти вероятность того, что хотя бы одно письмо попадет своему адресату. Найти предел этой вероятности при $n \to \infty$.

17. В задаче 16 найти математическое ожидание и дисперсию числа ξ писем, попавших своему адресату.

Указание. Представить ξ в виде суммы индикаторов.

18. В N ящиков случайно и независимо друг от друга бросают шары, пока не останется пустых ящиков. Обозначим ν число брошенных шаров. Найти $\mathsf{M}\nu$.

Указание. Представить ν в виде суммы времен между заполнениями новых ящиков.

19. Найти $\mathsf{P}\{|\xi - \mathsf{M}\xi| > 3\sqrt{\mathsf{D}\xi}\}$, если ξ имеет: а) нормальное распределение; б) показательное распределение; в) равномерное распределение на отрезке $[-1, 1]$; г) $\mathsf{P}(\xi = -1) = \mathsf{P}(\xi = 1) = \frac{1}{18}$, $\mathsf{P}(\xi = 0) = \frac{8}{9}$; д) распределение Пуассона с $\mathsf{M}\xi = 0{,}09$.

20. Иногда в приложениях сложную функцию от случайных аргументов $(\xi_1, \xi_2, \ldots)$ заменяют линейной частью ее разложения по формуле Тейлора в окрестности точки $(\mathsf{M}\xi_1, \mathsf{M}\xi_2, \ldots)$. Пусть случайная величина ξ распределена нормально с параметрами $\mathsf{M}\xi = 0$, $\mathsf{D}\xi = \sigma^2$ и $f(\xi) = \xi + k\xi^2$, $k > 0$. Сравнить $\mathsf{P}(|\xi| < u_\alpha\sigma) = 1 - 2\alpha$ и $\mathsf{P}(|f(\xi)| < u_\alpha\sigma)$:

1) Вычислить вторую из этих вероятностей, если а) $\alpha = 0{,}05$; $\sigma = 2$, $k = 0{,}5$; б) $\alpha = 0{,}05$, $\sigma = 0{,}1$, $k = 0{,}25$.

2) Предположив, что $k\sigma$ мало, найти приближенное решение неравенства $|\xi + k\xi^2| < u_\alpha\sigma$ и вероятность $\mathsf{P}(|f(\xi)| < u_\alpha\sigma)$.

21. Точки $c_1, c_2, \ldots, c_n$ независимы и имеют равномерное распределение в круге K единичного радиуса. Пусть случайное множество A состоит из тех и только тех точек круга, которые находятся ближе к центру круга, чем к границе и к любой из точек $c_1, \ldots, c_n$. Найти математическое ожидание площади ξ множества A.

Указание. Пусть $K = \{(x, y): x^2 + y^2 \leqslant 1\}$. Положим $\chi(x, y) = 1$, если $(x, y) \in A$; и $\chi(x, y) = 0$, если $(x, y) \in K \setminus A$. Использовать равенство

$$\xi = \iint\limits_K \chi(x, y)\, dx\, dy.$$

22. Пусть $\mathsf{M}\xi_1 = a_1$, $\mathsf{M}\xi_2 = a_2$, $\mathsf{D}\xi_1 = \sigma_{11} > 0$, $\mathsf{D}\xi_2 = \sigma_{22} > 0$, $\operatorname{cov}(\xi_1, \xi_2) = \sigma_{12}$. Найти α и β, при которых $\mathsf{M}(\xi_2 - \alpha\xi_1 - \beta)^2$ минимально.

23. Пусть случайная величина ξ имеет нормальное распределение с параметрами (a, σ^2). Оценить по неравенству Чебышева $\mathsf{P}(|\xi - a| > 2\sigma)$. Сравнить с точным значением этой вероятности.

24. Применим ли закон больших чисел к последовательности независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_k, \ldots$, если

$$\mathsf{P}(\xi_k = \sqrt{k}) = \mathsf{P}(\xi_k = -\sqrt{k}) = \frac{1}{2\sqrt{k}},$$

$$\mathsf{P}(\xi_k = 0) = 1 - \frac{1}{\sqrt{k}}?$$

25. Применим ли закон больших чисел к последовательности независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_k, \ldots$, если они нормально распределены с $\mathsf{M}\xi_k = 0$, $\mathsf{D}\xi_k = ck^\alpha$, $c > 0$, $\alpha > 0$ — некоторые постоянные?

26. Доказать, что к последовательности $\xi_1, \xi_2, \ldots, \xi_k, \ldots$ применим закон больших чисел, если $|\operatorname{cov}(\xi_k, \xi_l)| \leqslant C$ для всех $k, l = 1, 2, \ldots$ и $\operatorname{cov}(\xi_k, \xi_l) \to 0$ при $|k - l| \to \infty$.

27. Случайные величины $\xi_1, \xi_2, \ldots, \xi_N$ независимы и распределены по закону Пуассона с параметрами $\lambda_1, \lambda_2, \ldots, \lambda_N$. Найти:

а) $\mathsf{P}(\xi_1 + \ldots + \xi_k = m \mid \xi_1 + \ldots + \xi_N = n)$;

б) $\mathsf{M}(\xi_1 + \ldots + \xi_k \mid \xi_1 + \xi_2 + \ldots + \xi_N = n)$.

28. Случайные величины ξ_1, ξ_2, ξ_3 независимы и распределены нормально с параметрами (0, 1). Найти плотность распределения величины $\eta = (\xi_1 + \xi_2\xi_3)/\sqrt{1 + \xi_3^2}$.

Указание. Найти условную плотность распределения η при условии, что фиксировано ξ_3.

29. Доказать, что $\mathsf{D}\xi = \mathsf{M}(\mathsf{D}(\xi \mid \eta)) + \mathsf{D}(\mathsf{M}(\xi \mid \eta))$.

30. В задаче о случайном блуждании, рассмотренной в § 4 гл. 4, обозначим τ время до поглощения в точках 0 или n. Найти $m_k = \mathsf{M}(\tau \mid \xi_0 = k)$, где ξ_t — координата частицы в момент времени t.

Указание. Составить уравнение в конечных разностях для m_k

31. Предполагается провести 10 измерений $x_1, x_2, \ldots, x_{10}$ неизвестной величины a. Считая $x_1, \ldots, x_{10}$ независимыми нормально распределенными случайными величинами с $\mathsf{M}x_k = a$, $\mathsf{D}x_k = 0{,}01$, подобрать Δ так, чтобы

$$\mathsf{P}\left(\left|\frac{x_1 + x_2 + \ldots + x_{10}}{10} - a\right| < \Delta\right) = 0{,}99.$$

32. Случайные величины ξ_1 и ξ_2 независимы и распределены нормально с параметрами (0, 1). Являются ли независимыми величины $\eta_1 = \xi_1 + \xi_2$, $\eta_2 = \xi_1 - \xi_2$?

33. Случайная величина ξ распределена нормально с параметрами (0, 1). Положим $\eta = \xi$, если $|\xi| \leqslant 1$, и $\eta = -\xi$, если $|\xi| > 1$. а) Найти закон распределения η. б) Является ли величина $\xi + \eta$ нормально распределенной?

34. Найти среднее арифметическое реализации 50 равномерно распределенных случайных величин, полученной в задаче 24 гл. 5. Сравнить с математическим ожиданием случайной величины, равномерно распределенной на отрезке (0, 1).

35. Получить 10 реализаций случайного блуждания частицы по целым точкам отрезка $[0, n]$ (см. § 4 гл. 4). Найти частоты поглощения частицы в точках 0 и n. Сравнить с вероятностями, найденными в § 4 гл. 4. Найти среднее арифметическое времен блуждания до поглощения. Сравнить с $\mathsf{M}(\tau \mid \xi_0 = k)$, найденным в задаче 30. Положить $k = 1$, $n = 10$. Рассмотреть случаи: а) $a = 1/6$, $q = 5/6$; б) $p = q = 1/2$.

36. Для студентов группы найти число месяцев, на которые не приходится ни одного дня рождения. Найти математическое ожидание и дисперсию числа таких месяцев. (Воспользоваться результатом задачи 14.)

Предельные теоремы

Практическое значение предельных теорем состоит в том, что они позволяют аппроксимировать распределения допредельных величин предельными распределениями, аналитическая запись которых часто оказывается проще выражений для допредельных функций распределения. В главе 4 мы прямым вычислением получили предельную теорему Пуассона и теоремы Муавра—Лапласа, которые использовались для аппроксимации распределения числа успехов в схеме Бернулли. В данной главе мы воспользуемся для получения предельных теорем аналитическими методами, основанными на использовании свойств производящих или характеристических функций.

§ 1. Производящие функции

Если случайная величина ξ принимает только целые неотрицательные значения *), т. е.

$$\sum_{k=0}^{\infty} \mathsf{P}(\xi = k) = 1, \tag{1.1}$$

то *производящей функцией распределения* ξ называется функция

$$\varphi_\xi(x) = \mathsf{M}x^\xi = \sum_{k=0}^{\infty} x^k \mathsf{P}(\xi = k), \tag{1.2}$$

где x ($|x| \leqslant 1$) — действительная или комплексная переменная.

Теорема 1.1. *Пусть* $\varphi_\xi(x)$ — *производящая функция, определенная формулой* (1.2). *Тогда*:

1°. $\varphi_\xi(x)$ *определена в каждой точке отрезка* $[-1, 1]$.

2°. $\varphi_\xi(1) = 1$.

3°. *Соответствие, устанавливаемое формулой* (1.2), *между множеством производящих функций* $\varphi_\xi(x)$ *и мно-*

*) Такие случайные величины называют обычно целочисленными.

жеством распределений $\{p_k\}$ является взаимно однозначным.

Доказательство. Первое утверждение теоремы следует из того, что степенной ряд в (1.2) мажорируется при $|x| \leqslant 1$ сходящимся рядом в левой части (1.1). Равенство $\varphi_\xi(1) = 1$ совпадает с (1.1). Третье утверждение теоремы является следствием единственности разложения функции в ряд Тейлора. Теорема доказана.

Используя формулы для коэффициентов ряда Тейлора, можно явно указать распределение $\{p_k\}$, соответствующее вероятностной производящей функции $\varphi_\xi(x)$. Имеем

$$p_k = \frac{1}{k!}\varphi_\xi^{(k)}(0), \quad k = 0, 1, 2, \ldots$$

Пример 1. Пусть случайная величина ξ имеет биномиальное распределение

$$\mathsf{P}(\xi = m) = C_n^m p^m q^{n-m}, \; m = 0, 1, \ldots, n.$$

По формуле (1.2)

$$\varphi_\xi(x) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} x^k = (px + q)^n. \tag{1.3}$$

По производящей функции легко определить моменты случайной величины. Особенно просто находится математическое ожидание

$$\mathsf{M}\xi^{[k]}, \; k = 0, 1, 2, \ldots, \tag{1.4}$$

где $\xi^{[k]} = \xi(\xi - 1)\ldots(\xi - k + 1)$. Математическое ожидание (1.4) будем называть k-м *факториальным моментом*. Зная факториальные моменты первого и второго порядков, можно найти дисперсию по следующей формуле:

$$\mathsf{D}\xi = \mathsf{M}\xi(\xi - 1) + \mathsf{M}\xi - (\mathsf{M}\xi)^2. \tag{1.5}$$

Теорема 1.2. *Если конечен k-й факториальный момент, то существует левосторонняя производная $\varphi_\xi^{(k)}(1)$ и*

$$\mathsf{M}\xi^{[k]} = \varphi_\xi^{(k)}(1), \tag{1.6}$$

в частности,

$$\mathsf{M}\xi = \varphi_\xi'(1). \tag{1.7}$$

Доказательство. При любом $|x| < 1$ функцию $\varphi_\xi(x)$ можно дифференцировать сколько угодно раз.

Таким образом, при любом k определена k-я производная

$$\varphi_\xi^{(k)}(x) = \sum_{m=0}^{\infty} m^{[k]} x^{m-k} \mathsf{P}(\xi = m). \tag{1.8}$$

По условию теоремы конечен k-й факториальный момент

$$\mathsf{M}\xi^{[k]} = \sum_{m=0}^{\infty} m^{[k]} \mathsf{P}(\xi = m),$$

являющийся суммой ряда (1.8) в точке $x = 1$. Следовательно, по теореме Абеля $\varphi_\xi^{(k)}(x)$ непрерывна в точке $x = 1$ и

$$\lim_{x \to 1-0} \varphi_\xi^{(k)}(x) = \sum_{m=0}^{\infty} m^{[k]} \mathsf{P}(\xi = m) = \varphi_\xi^{[k]}(1).$$

П р и м е р 2. Найдем $\mathsf{M}\xi$ и $\mathsf{D}\xi$ случайной величины ξ, распределенной по биномиальному закону. Дифференцируя (1.3) два раза по x, получим

$$\varphi_\xi'(x) = np\,(px+q)^{n-1},$$

$$\varphi_\xi''(x) = n\,(n-1)\,p^2\,(px+q)^{n-2}.$$

Отсюда, воспользовавшись (1.6), при $x = 1$ получим

$$\mathsf{M}\xi = \varphi_\xi'(1) = np\,(p+q)^{n-1} = np,$$

$$\mathsf{M}\xi\,(\xi - 1) = \varphi_\xi''(1) = n\,(n-1)\,p^2\,(p+q)^{n-2} = n\,(n-1)\,p^2.$$

По формуле (1.7)

$$\mathsf{D}\xi = n\,(n-1)\,p^2 + np - n^2p^2 = np\,(1-p).$$

Легко вычисляется k-й факториальный момент биномиального распределения. Так как $\varphi_\xi^{(k)}(x) = n^{[k]}(px+q)^{n-k}$, то

$$\mathsf{M}\xi^{[k]} = n^{[k]} p^k. \tag{1.9}$$

Применение производящих функций к изучению сумм независимых целочисленных случайных величин основано на следующей теореме.

Т е о р е м а 1.3. *Если случайные величины* $\xi_1, \ldots, \xi_n$ *независимы, то*

$$\varphi_{\xi_1+\ldots+\xi_n}(x) = \varphi_{\xi_1}(x) \ldots \varphi_{\xi_n}(x).$$

Доказательство. Воспользуемся представлением производящей функции в виде математического ожидания (1.2). Тогда

$$\varphi_{\xi_1+\ldots+\xi_n} = \mathsf{M}x^{\xi_1+\ldots+\xi_n} = \mathsf{M}(x^{\xi_1}\ldots x^{\xi_n}). \qquad (1.10)$$

Случайные величины $x^{\xi_1}, x^{\xi_2}, \ldots, x^{\xi_n}$ независимы, как функции от независимых случайных величин. Следовательно,

$$\mathsf{M}(x^{\xi_1}\ldots x^{\xi_n}) = \mathsf{M}x^{\xi_1}\ldots \mathsf{M}x^{\xi_n}.$$

Отсюда и из (1.10) следует утверждение теоремы, так как $\mathsf{M}x^{\xi_k} = \varphi_{\xi_k}(x)$.

Пример 3. Найдем производящую функцию биномиального распределения при помощи теоремы 1.3. Число успехов μ_n в схеме Бернулли имеет биномиальное распределение. Представим μ_n в виде суммы независимых слагаемых:

$$\mu_n = \xi_1 + \xi_2 + \ldots + \xi_n,$$

где

$$\mathsf{P}(\xi_k = 1) = 1 - \mathsf{P}(\xi_k = 0) = p, \quad k = 1, 2, \ldots, n.$$

По теореме 1.3

$$\varphi_{\mu_n}(x) = \varphi_{\xi_1}(x)\,\varphi_{\xi_2}(x)\ldots\varphi_{\xi_n}(x).$$

Производящие функции слагаемых получим по формуле (1.2):

$$\varphi_{\xi_k}(x) = xp + x^0q = px + q, \quad k = 1, \ldots, n.$$

Таким образом,

$$\varphi_{\mu_n}(x) = (px+q)^n.$$

Найдем производящую функцию случайного числа случайных слагаемых.

Теорема 1.4. *Пусть целочисленные величины* $\nu, \xi_1, \xi_2, \ldots, \xi_n$ *независимы при любом* $n = 1, 2, 3, \ldots$; ξ_k, $k = 1, 2, \ldots$, *одинаково распределены. Положим*

$$\zeta_\nu = \xi_1 + \xi_2 + \ldots + \xi_\nu, \quad \zeta_0 = 0.$$

Тогда

$$\varphi_{\zeta_\nu}(x) = \varphi_\nu[\varphi_{\xi_1}(x)]. \qquad (1.11)$$

Доказательство*). Вероятность события $(\zeta_\nu = m)$ представим в виде

$$\mathsf{P}(\zeta_\nu = m) = \sum_{k=0}^{\infty} \mathsf{P}(\nu = k, \zeta_\nu = m).$$

Так как $(\nu = k, \zeta_\nu = m) = (\nu = k, \zeta_k = m)$ и события $(\nu = k)$, $(\zeta_k = m)$ независимы, то

$$\mathsf{P}(\nu = k, \zeta_\nu = m) = \mathsf{P}(\nu = k, \zeta_k = m) = \\ = \mathsf{P}(\nu = k, \xi_1 + \ldots + \xi_k = m) = \\ = \mathsf{P}(\nu = k)\,\mathsf{P}(\xi_1 + \ldots + \xi_k = m).$$

Тогда

$$\mathsf{P}(\zeta_\nu = m) = \sum_{k=0}^{\infty} \mathsf{P}(\nu = k)\,\mathsf{P}(\xi_1 + \xi_2 + \ldots + \xi_k = m).$$

Умножив обе части этого равенства на x^m и просуммировав по $m = 0, 1, 2, \ldots$, получим

$$\varphi_{\zeta_\nu}(x) = \sum_{k=0}^{\infty} \mathsf{P}(\nu = k) \Big(\sum_{m=0}^{\infty} x^m \mathsf{P}(\xi_1 + \ldots + \xi_k = m) \Big).$$

Ряд в круглых скобках является производящей функцией распределения суммы $\xi_1 + \ldots + \xi_k$. Так как слагаемые ξ_l, $l = 1, \ldots, k$, независимы и одинаково распределены, то

$$\sum_{m=0}^{\infty} x^m \mathsf{P}(\xi_1 + \ldots + \xi_k = m) = \varphi_{\xi_1 + \ldots + \xi_k}(x) = [\varphi_{\xi_1}(x)]^k.$$

*) Если воспользоваться условными математическими ожиданиями и формулой полного математического ожидания (6.6.3), то можно дать более короткое доказательство теоремы 1.4. Действительно, по формуле (6.6.3)

$$\varphi_{\zeta_\nu}(x) = \mathsf{M} x^{\zeta_\nu} = \mathsf{M}(\mathsf{M}(x^{\zeta_\nu} \mid \nu)).$$

Так как $\mathsf{M}(x^{\zeta_\nu} \mid \nu) = \mathsf{M}(x^{\xi_1} \ldots x^{\xi_\nu} \mid \nu) = [\varphi_{\xi_1}(x)]^\nu$, то

$$\varphi_{\zeta_\nu}(x) = \mathsf{M}[(\varphi_{\xi_1}(x))^\nu] = \varphi_\nu[\varphi_{\xi_1}(x)].$$

В намеченном доказательстве основная трудность заключается в проверке равенства $\mathsf{M}(x^{\xi_1 + \ldots + \xi_\nu} \mid \nu) = (\mathsf{M} x^{\xi_1})^\nu$ или эквивалентного ему равенства $\mathsf{M}(x^{\xi_1 + \ldots + \xi_\nu} \mid \nu = k) = \mathsf{M} x^{\xi_1 + \ldots + \xi_k}$, которое интуитивно не вызывает сомнений.

Следовательно,

$$\varphi_{\zeta_\nu}(x)=\sum_{k=0}^{\infty}\mathsf{P}(\nu=k)\,[\varphi_{\xi_1}(x)]^k=\varphi_\nu[\varphi_{\xi_1}(x)].$$

Теорема доказана.

При доказательстве предельных теорем часто используется свойство непрерывности соответствия множества производящих функций множеству распределений.

Теорема 1.5. *Пусть при любом фиксированном n, $n=1, 2, \ldots$, последовательность $\{p_k(n)\}$ является распределением вероятностей, т. е.*

$$p_k(n)\geqslant 0,\quad \sum_{k=0}^{\infty}p_k(n)=1. \tag{1.12}$$

Для того чтобы при любом фиксированном k

$$\lim_{n\to\infty}p_k(n)=p_k \tag{1.13}$$

и

$$\sum_{k=0}^{\infty}p_k=1, \tag{1.14}$$

необходимо и достаточно, чтобы при любом $x\in[0, 1)$

$$\lim_{n\to\infty}\varphi_n(x)=\varphi(x), \tag{1.15}$$

где

$$\varphi_n(x)=\sum_{k=0}^{\infty}p_k(n)\,x^k,\quad \varphi(x)=\sum_{k=0}^{\infty}p_kx^k,\quad \varphi(1)=1.$$

Доказательство. Пусть выполнены (1.13) и (1.14). Представим разность $\varphi_n(x)-\varphi(x)$ в виде

$$\varphi_n(x)-\varphi(x)=\\ =\sum_{k=0}^{N}(p_k(n)-p_k)\,x^k+\sum_{k=N+1}^{\infty}p_k(n)\,x^k-\sum_{k=N+1}^{\infty}p_kx^k,$$

где N — некоторое целое число. Пусть $x\in[0, 1)$ фиксирован. Докажем (1.15). Так как $0\leqslant p_k(n)\leqslant 1$, $0\leqslant p_k\leqslant\leqslant 1$, то для любого фиксированного $\varepsilon>0$ можно подобрать N так, чтобы при любом n

$$\sum_{k=N+1}^{\infty}p_k(n)\,x^k\leqslant\sum_{k=N+1}^{\infty}x^k<\frac{\varepsilon}{3},\quad \sum_{k=N+1}^{\infty}p_kx^k<\frac{\varepsilon}{3}.$$

Тогда при данном N

$$|\varphi_n(x)-\varphi(x)|\leqslant\sum_{k=0}^{N}|p_k(n)-p_k|\,x^k+\frac{2}{3}\varepsilon.$$

Сумма в правой части этого неравенства может быть сделана меньше $\varepsilon/3$ при достаточно больших n, так как содержит конечное число стремящихся к нулю слагаемых. Равенство $\varphi(1)=1$ следует из (1.14).

Пусть теперь выполнено (1.15). Докажем (1.13) от противного. Предположим, что (1.13) не выполнено. Тогда можно показать, что найдутся две последовательности n'_m и n''_m для которых

$$\lim_{n'_m\to\infty} p_k(n'_m)=\tilde{p}_k,\quad \lim_{n''_m\to\infty} p_k(n''_m)=\tilde{\tilde{p}}_k, \tag{1.16}$$

причем $\{\tilde{p}_k\}$ и $\{\tilde{\tilde{p}}_k\}$ не совпадают. Тогда по доказанной части теоремы из (1.16) следует, что

$$\lim_{n'_m\to\infty}\varphi_{n'_m}(x)=\tilde{\varphi}(x)=\sum_{k=0}^{\infty}\tilde{p}_k x^k,$$

$$\lim_{n''_m\to\infty}\varphi_{n''_m}(x)=\tilde{\tilde{\varphi}}(x)=\sum_{k=0}^{\infty}\tilde{\tilde{p}}_k x^k$$

и $\tilde{\varphi}(x)\not\equiv\tilde{\tilde{\varphi}}(x)$. Это невозможно, так как предел (1.15) существует. Теорема доказана.

П р и м е р 4. Пусть μ_n — число успехов в n испытаниях Бернулли и p_n — вероятность успеха в одном испытании. Будем предполагать, что существует $\lim\limits_{n\to\infty} np_n=\lambda$, $0<\lambda<\infty$. Воспользуемся теоремой 1.5 для вычисления $\lim\limits_{n\to\infty}\mathsf{P}(\mu_n=m)$. Положим $\lambda_n=np_n$. По формуле (1.3)

$$\varphi_{\mu_n}(x)=\left(\frac{n\lambda}{n}x+1-\frac{\lambda_n}{n}\right)^n=\left[1+\frac{\lambda_n}{n}(x-1)\right]^n.$$

Следовательно,

$$\lim_{n\to\infty}\varphi_{\mu_n}(x)=e^{\lambda(x-1)}=\sum_{m=0}^{\infty}\frac{\lambda^m}{m!}e^{-\lambda}x^m,$$

Отсюда по теореме 1.5

$$\lim_{n\to\infty} \mathsf{P}(\mu_n = m) = \frac{\lambda^m}{m!} e^{-\lambda}.$$

Таким образом, получили новое доказательство теоремы Пуассона.

§ 2. Характеристические функции

Производящие функции определены для целочисленных случайных величин. Для исследования распределений произвольных случайных величин вводятся характеристические функции. Комплекснозначной случайной величиной будем называть функцию $\xi_1(\omega) + i\xi_2(\omega)$, где $\omega \in \Omega$, (ξ_1, ξ_2) — случайный вектор. По определению положим

$$\mathsf{M}(\xi_1 + i\xi_2) = \mathsf{M}\xi_1 + i\mathsf{M}\xi_2. \tag{2.1}$$

Для математического ожидания от комплекснозначной случайной величины легко проверяются свойства 1°—4°, приведенные в § 2 гл. 6. В дальнейшем мы ими будем пользоваться без дополнительных оговорок. Понятие независимости для комплекснозначных случайных величин не будет введено, и свойство 5° использоваться не будет.

Характеристической функцией действительной случайной величины ξ называется

$$f_\xi(t) = \mathsf{M}e^{it\xi}, \tag{2.2}$$

где t — действительное число, $-\infty < t < \infty$.

Если случайная величина ξ дискретна, то по теореме 1.1 гл. 6

$$\mathsf{M}(\cos t\xi) = \sum_{k=1}^{\infty} \cos(tx_k)\,\mathsf{P}(\xi = x_k),$$

$$\mathsf{M}(\sin t\xi) = \sum_{k=1}^{\infty} \sin(tx_k)\,\mathsf{P}(\xi = x_k).$$

Отсюда и из (2.1) получим

$$f_\xi(t) = \sum_{k=1}^{\infty} e^{itx_k}\mathsf{P}(\xi = x_k). \tag{2.3}$$

Используя теорему 1.2 гл. 6, для характеристической функции абсолютно непрерывной величины ξ будем иметь

$$f_\xi(t)=\int_{-\infty}^{\infty} e^{itx} p_\xi(x)\,dx. \tag{2.4}$$

Если случайная величина ξ определена на дискретном или абсолютно непрерывном вероятностном пространстве, то

$$f_\xi(t)=\sum_{k=1}^{\infty} e^{it\xi(\omega_k)} p_k, \quad \mathsf{P}(\{\omega_k\})=p_k \tag{2.5}$$

и

$$f_\xi(t)=\int\underset{\Omega}{\cdots}\int e^{it\xi(u_1,\ldots,u_n)}\pi(u_1,\ldots,u_n)\,du_1\ldots du_n. \tag{2.6}$$

Теорема 2.1. *Пусть $f_\xi(t)$ — характеристическая функция случайной величины ξ. Тогда:*

1°. *$f_\xi(t)$ определена при любом $t\in(-\infty,\infty)$.*

2°. *$f_\xi(0)=1$, $|f_\xi(t)|\leqslant 1$.*

3°. *Если $\eta=a\xi+b$, где a и b — постоянные, то $f_\eta(t)=e^{itb}f_\xi(at)$.*

4°. *Соответствие, устанавливаемое формулой (2.2) между множеством характеристических функций $f_\xi(t)$ и множеством функций распределения, является взаимно однозначным.*

Докажем 1°—3°. Так как при любом действительном t

$$|e^{itx}|\leqslant 1, \quad -\infty<x<\infty,$$

то доказательство первых двух утверждений легко следует из формул (2.3)—(2.6). Третье утверждение получим из следующих равенств:

$$f_\eta(t)=\mathsf{M}e^{it(a\xi+b)}=e^{itb}\mathsf{M}e^{it\xi a}=e^{itb}f_\xi(at).$$

Для дискретных и абсолютно непрерывных величин по функции распределения определяются соответственно вероятности значений и плотность распределения. Тогда по формулам (2.3) и (2.4) однозначно определяется $f_\xi(t)$. Обратное утверждение следует из формулы обращения

$$F_\xi(x_2)-F_\xi(x_1)=\frac{1}{2\pi}\lim_{N\to\infty}\int_{-N}^{N}\frac{e^{-itx_1}-e^{-itx_2}}{it}f_\xi(t)\,dt. \tag{2.7}$$

Доказательство этой формулы приводится в более полных курсах теории вероятностей (см., например, [2]).

Найдем характеристические функции некоторых распределений.

Пример 1. Пусть ξ — целочисленная случайная величина с производящей функцией $\varphi_\xi(x)$. Очевидно (см. (1.2) и (2.3)), что $f_\xi(t) = \varphi_\xi(e^{it})$.

Пример 2. Если $\mathsf{P}(\xi = a) = 1$, то по формуле (2.3) находим: $f_\xi(t) = e^{ita}$.

Пример 3. Пусть ξ нормально распределена с параметрами (0, 1). Тогда

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad f_\xi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx - \frac{x^2}{2}}\, dx.$$

Так как при любом t в силу нечетности подынтегральной функции

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \sin tx\, dx = 0,$$

то $f_\xi(t)$ имеет действительные значения и

$$f_\xi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \cos tx\, dx.$$

При формальном дифференцировании этого равенства по t справа получается интеграл, сходящийся равномерно по $t \in (-\infty, \infty)$. Следовательно,

$$f_\xi'(t) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} \sin tx\, dx.$$

Отсюда интегрированием по частям получим

$$f_\xi'(t) = \frac{1}{\sqrt{2\pi}} \int \sin tx\, de^{-\frac{x^2}{2}} = -t f_\xi(t).$$

Таким образом, функция $f_\xi(t)$ удовлетворяет уравнению

$$\frac{df_\xi(t)}{dt} = -t f_\xi(t)$$

и начальному условию $f_\xi(0) = 1$. Отсюда

$$f_\xi(t) = e^{-\frac{t^2}{2}}.$$

Пример 4. Пусть случайная величина ξ распределена нормально с параметрами (a, σ^2). Найдем $f_\xi(t)$. Положим $\eta = \frac{\xi - a}{\sigma}$. Случайная величина η имеет нормальное распределение с параметрами $(0, 1)$, и, следовательно, $f_\eta(t) = e^{-\frac{t^2}{2}}$. Тогда по утверждению 3° теоремы 2.1 получим

$$f_\xi(t) = f_{\sigma\eta+a}(t) = e^{ita} f_\eta(\sigma t) = e^{ita - \frac{\sigma^2 t^2}{2}}.$$

Таким образом,

$$f_\xi(t) = e^{ita - \frac{\sigma^2 t^2}{2}}. \tag{2.8}$$

Зная характеристическую функцию, можно легко найти моменты случайной величины.

Теорема 2.2. *Если существует k-й момент $\mathsf{M}|\xi|^k < \infty$, $k \geqslant 1$, то существует непрерывная k-я производная $f_\xi(t)$ и*

$$f_\xi^{(k)}(0) = i^k \mathsf{M}\xi^k.$$

Доказательство. Докажем теорему, например, для абсолютно непрерывных величин. Если существует k-й момент, то существуют все моменты меньшего порядка. Так как

$$\left| \int_{-\infty}^{\infty} ixe^{itx} p_\xi(x)\,dx \right| \leqslant \int_{-\infty}^{\infty} |x|\, p_\xi(x)\,dx = \mathsf{M}|\xi| < \infty,$$

то интеграл в левой части неравенства сходится равномерно по t. Следовательно, можно дифференцировать под знаком интеграла:

$$f'_\xi(t) = i \int_{-\infty}^{\infty} xe^{itx} p_\xi(x)\,dx, \quad f'_\xi(0) = i\mathsf{M}\xi.$$

Пусть теперь существует производная порядка l, $l < k$, и

$$f_\xi^{(l)}(t) = i^l \int_{-\infty}^{\infty} x^l e^{itx} p_\xi(x)\,dx.$$

Отсюда

$$f_\xi^{(l+1)}(t) = i^{l+1} \int_{-\infty}^{\infty} x^{l+1} e^{itx} p_\xi(x)\,dx,$$

так как интеграл в правой части последнего равенства сходится равномерно по t. Таким образом,

$$f_\xi^{(l+1)}(0) = i^{l+1}\mathsf{M}\xi^{l+1}.$$

Теорема доказана.

Пример 5. Пусть ξ имеет нормальное распределение с параметрами (a, σ^2). Тогда характеристическая функция величины $\eta = \xi - a$ равна

$$f_\eta(t) = e^{-\frac{t^2}{2}\sigma^2} = \sum_{k=0}^{\infty}(-1)^k\left(\frac{\sigma^{2k}(2k)!}{2^k k!}\right)\frac{t^{2k}}{(2k)!}.$$

Так как $f_\eta^{(2k+1)}(0) = 0$ и

$$f_\eta^{(2k)}(0) = \frac{(2k)!}{2^k k!}\sigma^{2k}(-1)^k = (-1)^k(2k-1)(2k-3)\ldots 3\cdot 1\sigma^{2k},$$

то по теореме 2.2 получаем центральные моменты ξ:

$$\mathsf{M}(\xi - a)^{2k+1} = 0,$$

$$\mathsf{M}(\xi - a)^{2k} = (2k-1)(2k-3)\ldots 3\cdot 1\cdot\sigma^{2k}.$$

В следующей главе нам потребуются разложения характеристической функции по степеням t в окрестности точки $t = 0$. Если существует $\mathsf{M}\xi$, то

$$f_\xi(t) = 1 + it\mathsf{M}\xi + t\varepsilon(t), \qquad (2.9)$$

где $\varepsilon(t) \to 0$ при $t \to 0$. Для доказательства (2.9) представим $f_\xi(t)$ в виде $f'_\xi(t) = u(t) + iv(t)$. Из существования $\mathsf{M}\xi$ следует существование $f'_\xi(0)$, а также $u'(0)$ и $v'(0)$. По формуле Тейлора с остаточным членом в форме Пеано

$$u(t) = u(0) + tu'(0) + t\varepsilon_1(t),$$

$$v(t) = v(0) + tv'(0) + t\varepsilon_2(t),$$

где $\varepsilon_1(t)$, $\varepsilon_2(t) \to 0$ при $t \to 0$. Складывая два последних равенства, получим (2.9), так как $f'_\xi(0) = i\mathsf{M}\xi$. Аналогично проверяется следующее утверждение. Если $\mathsf{M}\xi^2$ существует, то

$$f_\xi(t) = 1 + it\mathsf{M}\xi - \frac{t^2}{2}\mathsf{M}\xi^2 + t^2\varepsilon(t), \qquad (2.10)$$

где $\varepsilon(t) \to 0$ при $t \to 0$.

В предположении существования $\mathsf{M}\xi^3$ дадим другую оценку остаточного члена в (2.10).

Из равенства

$$e^{ix}-1=\int_0^x ie^{iu}\,du$$

следует, что $|e^{ix}-1|\leqslant\int_0^{|x|}du=|x|$, так как $|ie^{iu}|\leqslant 1$. Используя найденную оценку и равенства

$$e^{ix}-1-ix=\int_0^x (e^{iu}-1)\,du,$$

$$e^{ix}-1-ix+\frac{x^2}{2}=i\int_0^x (e^{iy}-1-iy)\,dy, \ldots,$$

получим

$$|e^{ix}-1-ix|\leqslant\frac{|x|^2}{2}, \qquad \left|e^{ix}-1-ix+\frac{x^2}{2}\right|\leqslant\frac{|x|^3}{6}, \ldots$$

Отсюда, в частности, следует, что

$$e^{itx}=1+itx-\frac{t^2x^2}{2}+R_1(t, x), \qquad |R_1(t, x)|\leqslant\frac{1}{6}|tx|^3.$$

Заменим здесь x на ξ и вычислим математическое ожидание. Получим следующее разложение:

$$f_\xi(t)=1+it\mathsf{M}\xi-\frac{t^2}{2}\mathsf{M}\xi^2+R(t), \qquad |R(t)|\leqslant\frac{|t|^3}{6}\mathsf{M}|\xi|^3. \tag{2.11}$$

Для характеристических функций имеет место теорема, аналогичная теореме 1.3.

Теорема 2.3. *Если случайные величины* $\xi_1, \xi_2, \ldots, \xi_n$ *независимы, то*

$$f_{\xi_1+\xi_2+\ldots+\xi_n}(t)=f_{\xi_1}(t)\,f_{\xi_2}(t)\ldots f_{\xi_n}(t).$$

Доказательство. Достаточно доказать теорему для $n=2$, так как общее утверждение можно будет получить по индукции. По определению характеристической функции

$$f_{\xi_1+\xi_2}(t)=\mathsf{M}e^{it(\xi_1+\xi_2)}=\mathsf{M}e^{it\xi_1}\cdot e^{it\xi_2}=$$
$$=\mathsf{M}(\cos t\xi_1+i\sin t\xi_1)(\cos t\xi_2+i\sin t\xi_2). \tag{2.12}$$

Дальше нужно сделать следующее: 1) перемножить выражения, стоящие в скобках, и перейти к сумме математических ожиданий; 2) математические ожидания от произведений заменить на произведения математических ожиданий (это возможно, так как функции от независимых случайных величин являются независимыми случайными величинами); 3) полученное выражение вновь разложить на множители. В результате этих преобразований знак математического ожидания в правой части (2.12) появится перед каждым слагаемым в скобках. Таким образом,

$$f_{\xi_1+\xi_2}(t) = (\mathsf{M}\cos t\xi_1 + i\mathsf{M}\sin t\xi_1)(\mathsf{M}\cos t\xi_2 + i\mathsf{M}\sin t\xi_2) = \mathsf{M}e^{it\xi_1}\cdot\mathsf{M}e^{it\xi_2} = f_{\xi_1}(t)\cdot f_{\xi_2}(t).$$

Теорема доказана.

Пример 6. Пусть ξ_1 и ξ_2 независимы и нормально распределены с параметрами (a_1, σ_1^2), (a_2, σ_2^2). Найдем распределение суммы $\xi_1 + \xi_2$. По формуле (2.8)

$$f_{\xi_1}(t) = e^{ia_1t - \frac{\sigma_1^2t^2}{2}}, \quad f_{\xi_2}(t) = e^{ia_2t - \frac{\sigma_2^2t^2}{2}}.$$

Отсюда по теореме 2.3 получим

$$f_{\xi_1+\xi_2}(t) = f_{\xi_1}(t)f_{\xi_2}(t) = e^{iat - \frac{\sigma^2t^2}{2}},$$

где $a = a_1 + a_2$, $\sigma^2 = \sigma_1^2 + \sigma_2^2$. Так как полученная характеристическая функция является характеристической функцией нормального распределения, то по теореме 2.1 (4°) сумма $\xi_1 + \xi_2$ распределена нормально.

Пусть задана последовательность функций распределения $F_n(x)$, $n = 1, 2, \ldots$ Последовательность $\{F_n(x)\}$ *слабо сходится* к функции распределения $F(x)$, если

$$\lim_{n\to\infty} F_n(x) = F(x)$$

при любом x, являющемся точкой непрерывности $F(x)$.

Теорема 2.4. *Если последовательность функций распределения $F_n(x)$, $n = 1, 2, \ldots$, слабо сходится к непрерывной функции распределения $F(x)$, то $F_n(x) \to F(x)$ при $n \to \infty$ равномерно по $x \in (-\infty, +\infty)$.*

Доказательство. Из монотонности и ограниченности функции $F(x)$ следует, что для любого $\varepsilon > 0$ среди точек непрерывности $F(x)$ можно выбрать конечное число точек $x_1 < x_2 < \ldots < x_N$ так, чтобы на каждом из следующих множеств

$$(-\infty = x_0, x_1),\ [x_2, x_3), \ldots, [x_N, x_{N+1} = +\infty)$$

приращение функции $F(x)$ не превосходило ε. Пусть $x \in [x_k, x_{k+1})$. Так как

$$F_n(x) - F(x) = \\ = (F_n(x) - F_n(x_k)) + (F_n(x_k) - F(x_k)) + (F(x_k) - F(x))$$

и

$$|F_n(x) - F_n(x_k)| \leqslant F_n(x_{k+1}) - F_n(x_k) \leqslant \\ \leqslant |F_n(x_{k+1}) - F(x_{k+1})| + |F(x_k) - F_n(x_k)| + F(x_{k+1}) - F(x_k),$$

$$|F(x_k) - F(x)| \leqslant F(x_{k+1}) - F(x_k)$$

в силу монотонности $F_n(x)$ и $F(x)$, то

$$|F_n(x) - F(x)| \leqslant |F_n(x_{k+1}) - F(x_{k+1})| + \\ + 2|F_n(x_k) - F(x_k)| + 2(F(x_{k+1}) - F(x_k)). \qquad (2.13)$$

По условию теоремы $F_n(x_k) \to F(x_k)$ при $n \to \infty$. Следовательно, из (2.13) с учетом выбора $[x_k, x_{k+1})$

$$|F_n(x) - F(x)| < 5\varepsilon$$

при $n > n_0(k)$. Полагая $n_0 = \max\limits_{0 \leqslant k \leqslant N} n_0(k)$, получим $|F_n(x) - F(x)| < 5\varepsilon$ при $n > n_0$ и любых x. Теорема доказана.

Приведем формулировку теоремы о непрерывности соответствия множества характеристических функций множеству функций распределения. Пусть $f_n(t)$, $n = 1, 2, \ldots$, — последовательность характеристических функций и $F_n(x)$, $n = 1, 2, \ldots$,— последовательность соответствующих функций распределения.

Теорема 2.5. *Если $f_n(t) \to f(t)$ при $n \to \infty$ для любого t и $f(t)$ непрерывна при $t = 0$, то*

1) $f(t)$ — характеристическая функция, соответствующая некоторой функции распределения $F(x)$;

2) $F_n(x)$ слабо сходится к $F(x)$ при $n \to \infty$.

Если $F_n(x)$ слабо сходится к функции распределения $F(x)$, то $f_n(x) \to f(t)$, где $f(t)$ — характеристическая функция, соответствующая функции распределения $F(x)$.

Эта теорема будет использована для доказательства центральной предельной теоремы.

Приведем еще две теоремы, которые часто используются при доказательстве предельных теорем. Пусть

$$F_n(x) = \mathsf{P}(\xi_n < x),\ n = 1, 2, \ldots,$$

— последовательность функций распределения. Обозначим $m_n^{(k)} = \mathsf{M}\xi_n^k$.

Теорема 2.6. *Если при всех k, $k = 0, 1, 2, \ldots$, $\lim\limits_{n\to\infty} m_n^{(k)} = m^{(k)} < \infty$, то существует функция распределения $F(x) = \mathsf{P}(\xi < x)$ такая, что $m^{(k)} = \mathsf{M}\xi^k$. Если этому условию удовлетворяет единственная функция $F(x)$, то $F_n(x)$ при $n \to \infty$ слабо сходится к $F(x)$.*

Иногда последовательность случайных величин ζ_n ($n = 1, 2, \ldots$), для которой исследуется сходимость функций распределения, удается представить в виде $\zeta_n = \xi_n + \eta_n$; при этом сходимость функций распределения величин ξ_n уже известна или ее можно легко установить, а величинами η_n можно пренебречь. Следующая теорема дает условия, при которых можно воспользоваться указанным представлением.

Теорема 2.7. *Пусть $\zeta_n = \xi_n + \eta_n$, $n = 1, 2, \ldots$ Если $\lim\limits_{n\to\infty} \mathsf{P}\{|\eta_n| > \varepsilon\} = 0$ для любого $\varepsilon > 0$ и последовательность функций распределения $F_n(x) = \mathsf{P}(\xi_n < x)$ слабо сходится к функции распределения $F(x)$, то последовательность функций распределения $G_n(x) = \mathsf{P}(\zeta_n < x)$ тоже слабо сходится к $F(x)$.*

Доказательство. Пусть x — точка непрерывности функции $F(x)$. Положим

$$A_n = \{\xi_n + \eta_n < x\},\ B_n = \{|\eta_n| < \varepsilon\}.$$

Так как $A_n = A_nB_n + A_n\bar{B}_n$ и $A_n\bar{B}_n \subset \bar{B}_n$, то

$$\mathsf{P}(A_nB_n) \leqslant \mathsf{P}(A_n) \leqslant \mathsf{P}(A_nB_n) + \mathsf{P}(\bar{B}_n).$$

Отсюда и из соотношений

$$\{(\xi_n < x - \varepsilon) \cap B_n\} \subset A_nB_n \subset \{\xi_n < x + \varepsilon\}$$

получим

$$\mathsf{P}\{(\xi_n < x - \varepsilon) \cap B_n\} \leqslant \mathsf{P}(A_n) \leqslant \\ \leqslant \mathsf{P}\{\xi_n < x + \varepsilon\} + \mathsf{P}(\bar{B}_n). \qquad (2.14)$$

Из неравенств (2.14) и неравенства

$$\mathsf{P}\{(\xi_n < x - \varepsilon) \cap B_n\} \geqslant \mathsf{P}\{\xi_n < x - \varepsilon\} - \mathsf{P}(\bar{B}_n)$$

следует, что

$$F_n(x-\varepsilon) - \mathsf{P}(\bar{B}_n) \leqslant \mathsf{P}(A_n) \leqslant F_n(x+\varepsilon) + \mathsf{P}(\bar{B}_n).$$

Отсюда, так как по условию теоремы $\mathsf{P}(\bar{B}_n) \to 0$ при $n \to \infty$, находим

$$F(x-\varepsilon) \leqslant \varliminf_{n\to\infty} \mathsf{P}(\zeta_n < x) \leqslant \varlimsup_{n\to\infty} \mathsf{P}(\zeta_n < x) \leqslant F(x+\varepsilon). \quad (2.15)$$

Из (2.15) при $\varepsilon \to 0$ получим утверждение теоремы, так как x является точкой непрерывности $F(x)$.

Для исследования распределений векторных случайных величин, так же как и в одномерном случае, полезно использовать производящие и характеристические функции. *Производящие и характеристические функции векторных случайных величин* $\xi = (\xi_1, \ldots, \xi_k)' \in R^k$ определяются соответственно формулами

$$\varphi_\xi(z_1, \ldots, z_k) = \mathsf{M} z_1^{\xi_1} \ldots z_k^{\xi_k} \quad (|z_i| \leqslant 1, \quad i = 1, \ldots, k),$$
$$f_\xi(t) = f_\xi(t_1, \ldots, t_k) = \mathsf{M} \exp\{it'\xi\}, \quad (2.16)$$

где $t = (t_1, \ldots, t_k)'$, $t'\xi = t_1\xi_1 + \ldots + t_k\xi_k$, $-\infty < t_s < \infty$, $s = 1, \ldots, k$. Их свойства аналогичны свойствам производящих и характеристических функций одномерных случайных величин. Например, если $\mathsf{M}|\xi_1|^{r_1} \ldots |\xi_k|^{r_k} < \infty$ для целых $r_1, \ldots, r_k \geqslant 0$ *), то

$$\left.\frac{\partial^{r_1+\ldots+r_k}}{\partial z_1^{r_1} \ldots \partial z_k^{r_k}} \varphi_\xi(z_1, \ldots, z_k)\right|_{z_1=\ldots=z_k=1} = \mathsf{M}\xi_1^{[r_1]} \ldots \xi_k^{[r_k]},$$

$$\left.\frac{\partial^{r_1+\ldots+r_k}}{\partial t_1^{r_1} \ldots \partial t_k^{r_k}} f_\xi(t_1, \ldots, t_k)\right|_{t_1=\ldots=t_k=0} = i^{r_1+\ldots+r_k}\mathsf{M}\xi_1^{r_1} \ldots \xi_k^{r_k}.$$

Если $\eta = C\xi + a$, где $\xi = (\xi_1, \ldots, \xi_r)'$, $\eta = (\eta_1, \ldots, \eta_m)'$, $a = (a_1, \ldots, a_r)'$, $C = \| c_{ij} \|$ — $(m \times r)$-матрица (a и C постоянны), то

$$f_\eta(t) = e^{it'a} f_\xi(t'C). \quad (2,17)$$

Найдем характеристическую функцию нормально распределенного вектора $\eta = (\eta_1, \ldots, \eta_m)'$. По определению

*) Отметим, что для смешанных моментов из существования момента порядка $r_1 + \ldots + r_k$ не следует существование момента меньшего порядка.

$\eta_k = c_{k1}\xi_1 + c_{k2}\xi_2 + \ldots + c_{kr}\xi_r + a_k, \quad k = 1, \ldots, m,$ где $\xi_1, \ldots, \xi_r$ — независимые нормально распределенные величины с параметрами (0, 1). По определению (2.16)

$$f_\eta(t) = \mathsf{M}\exp(it'\eta) = \mathsf{M}\exp\left\{i\sum_{l=1}^{r}\left(\sum_{k=1}^{m} t_k c_{kl}\right)\xi_l + i\sum_{k=1}^{m} t_k a_k\right\}.$$

Отсюда, используя независимость $\xi_1, \ldots, \xi_r$, теоремы 2.1, 2.3 и формулу (2.8), получим

$$f_\eta(t) = \exp\left\{i\sum_{k=1}^{m} t_k a_k\right\}\cdot\prod_{l=1}^{r}\mathsf{M}\exp\left\{i\left(\sum_{k=1}^{m} t_k c_{kl}\right)\xi_l\right\} =$$

$$= \exp(it'a)\prod_{l=1}^{r}\exp\left\{-\frac{1}{2}\left(\sum_{k=1}^{m} t_k c_{kl}\right)^2\right\} =$$

$$= \exp(it'a)\exp\left\{-\frac{1}{2}\sum_{k_1, k_2=1}^{m} t_{k_1}t_{k_2}\left(\sum_{l=1}^{r} c_{k_1 l}c_{k_2 l}\right)\right\}.$$

Так как

$$b_{k_1 k_2} = \mathrm{cov}(\eta_{k_1}, \eta_{k_2}) = \sum_{l=1}^{r} c_{k_1 l}c_{k_2 l},$$

то окончательно получаем

$$f_\eta(t) = \exp\left\{it'a - \frac{1}{2}\sum_{k_1, k_2=1}^{m} b_{k_1 k_2}t_{k_1}t_{k_2}\right\}. \tag{2.18}$$

§ 3. Закон больших чисел

В этом параграфе будет доказана теорема 5.5 гл. 6 без предположения о конечности дисперсии. Случайные величины бесконечной последовательности $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ называются независимыми, если при любом n независимы величины $\xi_1, \xi_2, \ldots, \xi_n$.

Теорема 3.1. *Пусть случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ независимы, одинаково распределены и имеют математическое ожидание $\mathsf{M}\xi_k = a$, $k = 1, 2, \ldots$ Тогда при любом $\varepsilon > 0$*

$$\lim_{n\to\infty}\mathsf{P}\left(\left|\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} - a\right| < \varepsilon\right) = 1.$$

Доказательство. Характеристические функции $f_{\xi_k}(t)$, $k = 1, 2, \ldots$, одинаковы. Поэтому можно положить $f_{\xi_k}(t) = f(t)$. Из существования $\mathsf{M}\xi_k$ следует, что верно разложение (2.9):

$$f(t) = 1 + ita + t\varepsilon(t), \tag{3.1}$$

где $\varepsilon(t) \to 0$ при $t \to 0$. Положим $\eta_n = (\xi_1 + \ldots + \xi_n)/n$. Так как случайные величины $\xi_1, \ldots, \xi_n$ независимы, то по теореме 2.3

$$f_{\xi_1+\ldots+\xi_n}(t) = [f(t)]^n = (1 + ita + t\varepsilon(t))^n.$$

Отсюда, используя теорему 2.1 (3°), при $n \to \infty$ получим

$$f_{\eta_n}(t) = \left(1 + i\frac{t}{n}a + \frac{t}{n}\varepsilon\left(\frac{t}{n}\right)\right)^n \to e^{ita}, \tag{3.2}$$

так как $\varepsilon(t/n) \to 0$ при $n \to \infty$. Таким образом, при любом t последовательность $f_{\eta_n}(t)$ сходится к функции e^{ita}, являющейся характеристической функцией постоянной величины a. Функция распределения $F_a(x)$ постоянной a равна 1 при $x > a$ и 0 при $x \leqslant a$. По теореме 2.5 $\lim\limits_{n\to\infty} F_{\eta_n}(x) = F_a(x)$ при любом $x \neq a$, так как $x = a$ — единственная точка разрыва функции $F_a(x)$. Пусть задано $\varepsilon > 0$. Тогда

$$\mathsf{P}(|\eta_n - a| < \varepsilon) = \mathsf{P}(a - \varepsilon < \eta_n < a + \varepsilon) \geqslant$$
$$\geqslant \mathsf{P}\left(a - \frac{\varepsilon}{2} \leqslant \eta_n < a + \varepsilon\right) = F_{\eta_n}(a + \varepsilon) - F_{\eta_n}\left(a - \frac{\varepsilon}{2}\right).$$

Точки $x = a + \varepsilon$ и $x = a - \varepsilon/2$ являются точками непрерывности функции $F_a(x)$. Следовательно, при $n \to \infty$

$$F_{\eta_n}(a + \varepsilon) - F_{\eta_n}\left(a - \frac{\varepsilon}{2}\right) \to F_a(a + \varepsilon) - F_a\left(a - \frac{\varepsilon}{2}\right) = 1.$$

Отсюда и из неравенств

$$1 \geqslant \mathsf{P}(|\eta_n - a| < \varepsilon) \geqslant F_{\eta_n}(a + \varepsilon) - F_{\eta_n}\left(a - \frac{\varepsilon}{2}\right)$$

следует утверждение теоремы.

§ 4. Центральная предельная теорема

В § 3 гл. 4 была доказана теорема Муавра — Лапласа, согласно которой число успехов μ_n в n испытаниях схемы Бернулли при больших n имеет распределение, близкое к нормальному. Если воспользоваться тем, что μ_n представляется в виде суммы независимых слагаемых (5.5.13), то теорему Муавра — Лапласа можно сформулировать в следующем виде.

Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *независимы*, $\mathsf{P}(\xi_n = 1) = 1 - \mathsf{P}(\xi_n = 0) = p$, *то при* $n \to \infty$, $0 < p < 1$

$$\mathsf{P}\left(\frac{\xi_1 + \xi_2 + \ldots + \xi_n - n\mathsf{M}\xi_1}{\sqrt{n\mathsf{D}\xi_1}} < x\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du \quad (4.1)$$

равномерно по $x \in (-\infty, \infty)$.

Утверждение (4.1) сохраняется при достаточно общих предположениях о законе распределения слагаемых ξ_k. Докажем следующую теорему.

Теорема 4.1. *Если случайные величины* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *независимы, одинаково распределены и имеют конечную дисперсию, то при* $n \to \infty$ *равномерно по* $x \in (-\infty, \infty)$

$$\mathsf{P}\left(\frac{\xi_1 + \xi_2 + \ldots + \xi_n - na}{\sigma\sqrt{n}} < x\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du,$$

где $a = \mathsf{M}\xi_n$, $\sigma^2 = \mathsf{D}\xi_n > 0$.

Доказательство. Положим

$$\eta_n = \frac{\xi_1 + \xi_2 + \ldots + \xi_n - na}{\sigma\sqrt{n}} = \sum_{k=1}^{n} \frac{\xi_k - a}{\sigma\sqrt{n}}.$$

Разложим характеристическую функцию величин $\xi_k - a$ по формуле (2.10):

$$f_{\xi_k - a}(t) = 1 + it\mathsf{M}(\xi_k - a) - \frac{t^2}{2}\mathsf{M}(\xi_k - a)^2 + t^2\varepsilon(t),$$

где $\varepsilon(t) \to 0$ при $t \to 0$. Так как $\mathsf{M}(\xi_k - a) = 0$ и $\mathsf{M}(\xi_k - a)^2 = \sigma^2$, то

$$f_{\xi_k - a}(t) = 1 - \frac{t^2\sigma^2}{2} + t^2\varepsilon(t). \quad (4.2)$$

Отметим, что в этом равенстве функция $\varepsilon(t)$ не зависит от k. По теореме 2.3

$$f_{\sum\limits_{k=1}^{n}(\xi_k - a)}(t) = \left(1 - \frac{t^2\sigma^2}{2} + t^2\varepsilon(t)\right)^n.$$

Отсюда

$$f_{\eta_n}(t) = f_{\sum\limits_{k=1}^{n}(\xi_k - a)}\left(\frac{t}{\sigma\sqrt{n}}\right) = \left(1 - \frac{t^2}{2}\cdot\frac{1}{n} + \frac{t^2}{n^2\sigma^2}\varepsilon\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n,$$

и, следовательно, $f_{\eta_n}(t) \to e^{-t^2/2}$ при $n \to \infty$ для любого фиксированного t. Предельная характеристическая функция является характеристической функцией нормального распределения с параметрами (0, 1). Отсюда по теореме 2.5 следует слабая сходимость $F_{\eta_n}(x)$. Из слабой сходимости по теореме 2.4 следует равномерная сходимость, так как предельная функция распределения непрерывна. Теорема 4.1 доказана.

Условия сходимости функций распределения разнораспределенных слагаемых к нормальному закону содержатся в теореме Ляпунова. Приведем без доказательства ее упрощенную формулировку.

Теорема 4.2. *Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — независимые случайные величины, имеющие конечный третий абсолютный момент. Положим*

$$a_n = \mathsf{M}\xi_n, \qquad b_n^2 = \mathsf{D}\xi_n, \qquad c_n^3 = \mathsf{M}\,|\,\xi_n - a_n\,|^3,$$

$$A_n = \sum_{k=1}^{n} a_k, \qquad B_n^2 = \sum_{k=1}^{n} b_k^2, \qquad C_n^3 = \sum_{k=1}^{n} c_k^3.$$

Если при $n \to \infty$

$$\frac{C_n}{B_n} \to 0, \tag{4.3}$$

то при $n \to \infty$

$$\mathsf{P}\left(\frac{\xi_1 + \ldots + \xi_n - A_n}{B_n} < x\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du.$$

При ссылках на теоремы о сходимости сумм распределений к нормальному закону удобно использовать понятие асимптотической нормальности. Если функции распределения последовательности случайных величин $(\eta_n - A_n)/B_n$ сходятся при $n \to \infty$ к функции $\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}\exp\left(-\frac{u^2}{2}\right)du$, то говорят, что *случайная величина η_n при $n \to \infty$ асимптотически нормальна с параметрами (A_n, B_n^2).*

Приведем без доказательства обобщение теоремы 4.1 на случай r-мерных величин. Случайные r-мерные векторы последовательности $\xi_n = (\xi_{n1}, \xi_{n2}, \ldots, \xi_{nr})$, $n = 1, 2, \ldots$, будем называть независимыми, если при любом n и для любых r-мерных прямоугольников B_k, $k = 1, \ldots, n$,

$$\mathsf{P}(\xi_1 \in B_1, \ldots, \xi_n \in B_n) = \mathsf{P}(\xi_1 \in B_1) \ldots \mathsf{P}(\xi_n \in B_n). \quad (4.4)$$

Теорема 4.3. *Если случайные векторы* $\xi_n = (\xi_{n1}, \ldots, \xi_{nr})$, $k = 1, 2, \ldots$, *независимы, одинаково распределены и имеют конечные*

$$a_l = M\xi_{1l}, \quad b_{l,k} = \operatorname{cov}(\xi_{1,l}, \xi_{1,k}),$$

то r-мерная функция распределения

$$\mathsf{P}(\eta_{n1} < x_1, \ldots, \eta_{nr} < x_r), \quad \textit{где} \quad \eta_{nl} = \frac{\xi_{1l} + \ldots + \xi_{nl} - na_l}{\sqrt{n}}$$

($l = 1, \ldots, r$), *при* $n \to \infty$ *и при любом* $x = (x_1, \ldots, x_r)$ *сходится к* $\mathsf{P}(\zeta_1 < x_1, \ldots, \zeta_r < x_r)$, *где* $\zeta = (\zeta_1, \ldots, \zeta_r)$ — *r-мерный нормальный вектор с* $\mathsf{M}\zeta_l = 0$, $\operatorname{cov}(\zeta_l, \zeta_k) = b_{k,l}$.

§ 5. Вычисление интегралов методом Монте-Карло

В качестве примера применения теоремы 4.1 оценим число испытаний в методе Монте-Карло, необходимое для вычисления кратного интеграла с заданной точностью. Пусть функция $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_s)$ определена на s-мерном единичном кубе V. Требуется вычислить интеграл

$$a = \int\limits_V \ldots \int f(x)\, dx.$$

Пусть известна постоянная C такая, что $|f(\mathbf{x})| \leqslant C$, $\mathbf{x} \in V$. Обозначим $\xi = (\xi_1, \ldots, \xi_n)$ случайный вектор, равномерно распределенный на V. Тогда $p_\xi(x_1, \ldots, x_s) = 1$, если $x \in V$, и $p_\xi(x_1, \ldots, x_s) = 0$ в противном случае. Математическое ожидание случайной величины $\eta = f(\xi)$ найдем по формуле (5.1.5):

$$\mathsf{M}\eta = \int\limits_\Omega \ldots \int f(x_1, \ldots, x_s)\, p_\xi(x_1, \ldots, x_s)\, dx_1 \ldots dx_s = \int\limits_V \ldots \int f(\mathbf{x})\, d\mathbf{x} = a.$$

Таким образом, $\mathsf{M}\eta$ совпадает со значением вычисляемого интеграла. Так как $|f(\mathbf{x})| \leqslant C$, то

$$\sigma^2 = \mathsf{D}\eta = \int_V \dots \int (f(\mathbf{x}) - a)^2 \, d\mathbf{x} \leqslant 4C^2. \tag{5.1}$$

Пусть теперь случайные векторы $\xi_k = (\xi_{k1}, \dots, \xi_{ks})$, $k = 1, \dots, n$, независимы и распределены равномерно на единичном кубе V. Тогда случайные величины $\eta_k = f(\xi_k)$, $k = 1, \dots, n$, независимы и одинаково распределены. По закону больших чисел случайная величина

$$\zeta_n = \frac{\eta_1 + \dots + \eta_n}{n}$$

при больших n близка к постоянной $a = \mathsf{M}\eta_k$. Предположим, что нужно вычислить a с точностью Δ. Оценим вероятность

$$\mathsf{P}(|\zeta_n - a| < \Delta) = \mathsf{P}\left(\left|\frac{\zeta_n - a}{\sigma/\sqrt{n}}\right| < \frac{\Delta\sqrt{n}}{\sigma}\right).$$

Так как $\sigma < 2C$, то

$$\mathsf{P}(|\zeta_n - a| < \Delta) \geqslant \mathsf{P}\left(\left|\frac{\zeta_n - a}{\sigma/\sqrt{n}}\right| < \frac{\Delta\sqrt{n}}{2C}\right)$$

и при больших n

$$\mathsf{P}\left(\left|\frac{\zeta_n - a}{\sigma/\sqrt{n}}\right| < \frac{\Delta\sqrt{n}}{2C}\right) \approx 2\Phi_0\left(\frac{\Delta\sqrt{n}}{2C}\right).$$

По заданной малой вероятности нежелательного события $|\zeta_n - a| \geqslant \Delta$ можно, так же как в § 3 гл. 4, найти n.

Оценка дисперсии (5.1) является обычно очень завышенной. В качестве приближенного значения σ можно использовать выражение

$$\sigma^* = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(\eta_k - \zeta_n)^2}.$$

Для оценки вероятности можно воспользоваться тем, что величина $(\zeta_n - na)\sqrt{n}/\sigma^*$ при $n \to \infty$ является асимптотически нормальной с параметрами (0, 1) *).

---

*) Для доказательства этого утверждения нужно показать, что $\lim\limits_{n\to\infty} \mathsf{P}\left\{\left|\frac{\sigma^*}{\sigma} - 1\right| > \varepsilon\right\} = 0$ для любого $\varepsilon > 0$, и воспользоваться результатом задачи 17.

Применение метода Монте-Карло к вычислению интегралов, а также сравнение его с другими методами даются в книге [5], гл. 4.

§ 6. Прием линеаризации

В прикладных задачах часто требуется знать распределение функции от случайных величин. Обычно в таких случаях в качестве приближенного распределения принимается распределение линейной части разложения функции по формуле Тейлора в окрестности математического ожидания аргументов, а распределение линейной части (поскольку линейная часть является суммой) считается нормальным. По-видимому, можно надеяться на близость рассматриваемого распределения к нормальному в случае, когда аргументы распределены нормально, а их дисперсии малы. Более сомнительно использование нормальности, которое основывается на центральной предельной теореме, примененной к линейной части функции. Приведем теорему, которая дает некоторое представление о возможности использования нормального приближения в случае малых дисперсий нормально распределенных аргументов.

Т е о р е м а 6.1. *Пусть* $\xi_n = (\xi_{n1}, \ldots, \xi_{nr})$, $n = 1, 2, \ldots$, — *последовательность асимптотически нормальных векторов с конечными моментами*

$$\mathsf{M}\xi_{nl} = a_{nl} \quad (l = 1, \ldots, r),$$

$$\operatorname{cov}(\xi_{nl}, \xi_{nk}) = \frac{b_{kl}(n)}{n} \quad (k, l = 1, \ldots, r),$$

где $(a_{nl} - a_l)\sqrt{n} \to 0$ *и* $b_{kl}(n) \to b_{kl}$ *при* $n \to \infty$; *функция* $f(x) = f(x_1, \ldots, x_r)$ *определена в окрестности точки* $a = (a_1, \ldots, a_r)$ *и имеет в этой окрестности непрерывные частные производные второго порядка. Если* $\eta_n = f(\xi_{n1}, \ldots, \xi_{nr})$ *и величина* $\sigma^2 = \sum_{k,l=1}^{r} f_k f_l b_{kl}$, *где* $f_i = \left.\frac{\partial f}{\partial x_i}\right|_{x=a}$, *положительна, то при* $n \to \infty$

$$\mathsf{P}\left\{\frac{\sqrt{n}\,(\eta_n - f(a))}{\sigma} < x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du.$$

Доказательство. Пусть

$$R_n = \eta_n - f_0^* - \sum_{l=1}^{r} f_l^* (\xi_{nl} - a_{nl}),$$

где $f_i^* = \left.\frac{\partial f}{\partial x_i}\right|_{x=a_n}$, $f_0^* = f(a_n)$, $a_n = (a_{n1}, \ldots, a_{nr})$. Воспользовавшись теоремой 2.7, покажем, что предельное распределение $\sqrt{n}\,(\eta_n - f(a_n))$ совпадает с предельным распределением $\zeta_n^* = \sqrt{n} \sum_{l=1}^{r} f_l^* \cdot (\xi_{nl} - a_{nl})$. Положим

$A_n = \bigcap_{l=1}^{r} A_{nl}$, где $A_{nl} = \left\{\omega\colon |\xi_{nl} - a_{nl}| < \frac{1}{n^{3/8}}\right\}$. Используя неравенство Чебышева, получим

$$\mathsf{P}(\bar{A}_n) = \mathsf{P}\left(\bigcup_{l=1}^{r} \bar{A}_{nl}\right) \leqslant \sum_{l=1}^{r} \mathsf{P}(\bar{A}_{nl}) \leqslant n^{3/4} \sum_{l=1}^{r} \frac{b_{ll}(n)}{n} \to 0 \tag{6.1}$$

при $n \to \infty$. Отметим, что на множестве A_n величину R_n, используя формулу Тейлора с остаточным членом в форме Лагранжа, можно представить в виде

$$R_n = \frac{1}{2} \sum_{k,l=1}^{r} \tilde{f}_{kl} (\xi_{nk} - a_{nk})(\xi_{nl} - a_{nl}),$$

где $\tilde{f}_{kl}$ — значение производной $\partial^2 f/\partial x_k \partial x_l$ в точке $(a_{n1} + \theta(\xi_{n1} - a_{n1}), \ldots, a_{nr} + \theta(\xi_{nr} - a_{nr}))$, $0 < \theta < 1$. Так как $\tilde{f}_{kl}$ ограничены и на множестве A_n выполнены неравенства $|\xi_{nk} - a_{nk}| < 1/n^{3/8}$, $k = 1, \ldots, r$, то

$$|\sqrt{n}\, R_n| \leqslant \sqrt{n} K \frac{1}{n^{3/4}} = \frac{K}{\sqrt[4]{n}}, \quad \omega \in A_n. \tag{6.2}$$

Для произвольного $\varepsilon > 0$ имеем

$$\mathsf{P}(|\sqrt{n}\, R_n| > \varepsilon) =$$
$$= \mathsf{P}\{(|\sqrt{n}\, R_n| > \varepsilon) \cap A_n\} + \mathsf{P}\{(|\sqrt{n} R_n| > \varepsilon) \cap \bar{A}_n\}. \tag{6.3}$$

Так как $\varepsilon > 0$ фиксировано и $|\sqrt{n} R_n| \to 0$ в силу (6.2) при $n \to \infty$ равномерно по $\omega \in A_n$, то первое слагаемое в правой части (6.3) стремится к 0. Второе слагаемое в

(6.3) оценивается сверху вероятностью (6.1). Таким образом, $\mathsf{P}(|\sqrt{n}R_n| > \varepsilon) \to 0$ при $n \to \infty$. Следовательно, предельное распределение $(\eta_n - f(a_n))\sqrt{n}$ согласно теореме 2.7 совпадает с предельным распределением ζ_n^*. Так как величина ζ_n^* распределена асимптотически нормально и $(\sigma^*)^2 = \sum_{k,l=1}^{r} f_k^* f_l^* b_{kl}(n) \to \sigma^2$ при $n \to \infty$, то предельным распределением ζ_n^* является нормальное с параметрами $(0, \sigma^2)$. Нетрудно проверить, что предельное распределение $\sqrt{n}(\eta_n - f(a))$ совпадает с предельным распределением $\sqrt{n}(\eta_n - f(a_n))$. Теорема доказана.

Можно сформулировать теорему об асимптотической нормальности $\eta_n = f_n(\xi_{n1}, \ldots, \xi_{nr})$, где $r = r(n) \to \infty$ при $n \to \infty$, без предположения об асимптотической нормальности аргументов. Однако потребуются не очень естественные ограничения на последовательность функций f_n.

Отметим, что условия теоремы 6.1 не позволяют сделать вывод о том, что нормирующие и центрирующие постоянные в формулировке этой теоремы являются главной частью в асимптотических формулах для $\mathsf{M}\eta_n$, $\mathsf{D}\eta_n$. Асимптотические формулы для $\mathsf{M}\eta_n$, $\mathsf{D}\eta_n$, совпадающие с соответствующими формулами для моментов линейного приближения, получаются при дополнительных предположениях (см. [10], § 27.7, стр. 388).

Задачи к главе 7

1. Пусть ξ — неотрицательная целочисленная величина с производящей функцией $\varphi(x)$. Найти: а) $\mathsf{M}x^{2\xi+1}$; б) $\sum_{n=0}^{\infty} x^n \mathsf{P}(\xi \geqslant n)$; в) $\sum_{n=0}^{\infty} x^n \mathsf{P}(\xi \leqslant n)$; г) $\sum_{n=0}^{\infty} x^n \mathsf{P}(\xi = 2n)$.

2. Найти производящие функции величин τ, τ_m, определенных в задачах 14, 16 гл. 5. Найти $\mathsf{M}\tau$, $\mathsf{D}\tau$, $\mathsf{M}\tau_m$, $\mathsf{D}\tau^{(m)}$.

3. Найти производящую функцию величины ν, определенной в задаче 18 гл. 6.

4. Найти производящую функцию случайной величины, распределенной по закону Пуассона с параметром λ. Доказать, что сумма независимых пуассоновских величин имеет пуассоновское распределение.

5. Случайные величины ν, ξ_1, ξ_2, ..., ξ_n независимы при любом $n = 1, 2, \ldots$; $\mathsf{P}(\xi_k = 1) = 1 - \mathsf{P}(\xi_k = 0) = p$; ν имеет распределение Пуассона с параметром λ. Найти производящую функцию величины $\eta = \xi_1 + \ldots + \xi_\nu$ $(\nu > 0)$, $\eta = 0$ $(\nu = 0)$.

6. Вычислить характеристические функции следующих законов распределения: а) биномиального; б) Пуассона; в) показательного; г) равномерного на отрезке $[-1, 1]$ (см. § 3 гл. 5).

7. Найти законы распределения, соответствующие характеристическим функциям: а) $\cos t$; б) $\cos^2 t$; в) $\sum\limits_{k=1}^{\infty}\left(\frac{1}{2}\right)^k \cos kt$.

8. Величины ξ_1, ξ_2 независимы и одинаково распределены, их характеристическая функция равна $f(t)$. Найти характеристическую функцию величины $\xi_1 - \xi_2$.

9. Величины ξ_1, ξ_2, ξ_3 независимы и нормально распределены с параметрами $(1, 1)$, $(0, 4)$, $(-1, 1)$. Найти: а) $\mathsf{P}(\xi_1 + \xi_2 + \xi_3 < 0)$; б) $\mathsf{P}(|2\xi_1 - \xi_2 + \xi_3| < 3)$.

10. Доказать теорему Пуассона при помощи теоремы 2.6.

11. Складывается 10^4 чисел, каждое из которых округлено с точностью до 10^{-m}. Предполагая, что ошибки от округления независимы и в интервале $(-0{,}5\cdot 10^{-m},\ 0{,}5\cdot 10^{-m})$ распределены равномерно, найти пределы, в которых с вероятностью, не меньшей 0,99, будет лежать суммарная ошибка.

12. Получить теорему Муавра — Лапласа в качестве следствия теорем 4.1 и 4.2.

13. Пусть случайная величина ξ_λ распределена по закону Пуассона с параметром λ. Найти

$$\lim_{\lambda\to\infty} \mathsf{P}\left(\frac{\xi_\lambda - \lambda}{\sqrt{\lambda}} < x\right).$$

14. Пусть случайные величины $\xi_1, \ldots, \xi_n$ независимы; $\mathsf{P}(\xi_k = 1) = 1 - \mathsf{P}(\xi_k = 0) = p_k(n)$, $k = 1, 2, \ldots$ Найти $\lim\limits_{n\to\infty} \mathsf{P}(\xi_1 + \ldots + \xi_n = m)$, если $p_1(n) + \ldots + p_n(n) \to \lambda$ и $\max\limits_{1\leqslant k\leqslant n} p_k(n) \to 0$ при $n \to \infty$.

15. Плотность распределения $p_\xi(x)$ случайной величины ξ непрерывна и ограничена на отрезке $[a, b]$ и равна 0 вне $[a, b]$. Положим $\eta_n = \{n\xi\}$, где $\{x\}$ — дробная часть числа x. Найти $\lim\limits_{n\to\infty} \mathsf{P}(\eta_n < x)$, $0 < x < 1$.

16. Из урны, содержащей M белых шаров и $N - M$ черных, по схеме случайного выбора без возвращения вынимается n шаров. Обозначим ξ_n число белых шаров в выборке из n шаров. Найти:

а) $\lim\limits_{N\to\infty} \mathsf{P}(\xi_n = m)$, если $\dfrac{M}{N} \to p$, $n = \mathrm{const}$;

б) $\lim\limits_{\substack{n\to\infty\\ M\to\infty}} \mathsf{P}(\xi_n = m)$, если $\dfrac{Mn}{N} \to \lambda$.

17. Пусть последовательность случайных величин ξ_n удовлетворяет условию: $\mathsf{P}\{|\xi_n - 1| < \varepsilon\} \to 1$ при $n \to \infty$ и любом $\varepsilon > 0$; функции распределения $\mathsf{P}(\eta_n < x)$ слабо сходятся к $F(x)$ при $n \to \infty$. Доказать, что $\mathsf{P}(\eta_n \xi_n < x)$ слабо сходятся к $F(x)$.

Указание. Доказательство проводится аналогично доказательству теоремы 2.7.

18. Используя решение задачи 25 гл. 5 и таблицы случайных чисел, получить 10 независимых реализаций случайной величины ξ с функцией распределения $F_\xi(x) = 1 - x^{-n}$ $(x \geqslant 1)$, если: а) $n = 2$; б) $n = 1$. В обоих случаях найти среднее арифметическое полученных реализаций и в случае а) сравнить среднее арифметическое с $M\xi$.

19. Пусть τ_2 — число испытаний в схеме Бернулли, прошедшее от начала испытаний до появления второго успеха. Получить десять независимых реализаций величины τ_2, если вероятность успеха p в отдельном испытании: а) $p = 0{,}5$; б) $p = 0{,}75$. Найти средние арифметические и сравнить с $\mathsf{M}\tau_2$.

20. По 100 реализациям равномерно распределенной случайной величины (см. § 5) вычислить методом Монте-Карло интеграл $a = \int_0^1 e^x dx$. Полученное значение a^* сравнить с точным значением a. Теоретически найти Δ такое, чтобы $\mathsf{P}(|a - a^*| < \Delta) \approx 0{,}99$.

Указание. Использовать точное значение $\mathsf{D}a^*$ и центральную предельную теорему.

Цепи Маркова

§ 1. Определение

В главе 4 было дано определение цепи Маркова как частного случая общей схемы испытаний. Дадим теперь определение цепи Маркова в терминах случайных величин.

Последовательность случайных величин ξ_t, $t = 0, 1, 2, \ldots$, называется *цепью Маркова* с состояниями $\mathcal{N} = \{1, 2, \ldots, N\}$, если

$$\sum_{k=1}^{N} \mathsf{P}(\xi_t = k) = 1, \quad t = 0, 1, 2, \ldots,$$

и при любых $0 \leqslant t_1 < t_2 < \ldots < t_n < s < t$ ($n = 1, 2, \ldots$), любых $i, j \in \mathcal{N}$ и любых подмножествах $B_1, \ldots, B_n$ множества $\mathcal{N}$ выполняется равенство

$$\mathsf{P}(\xi_t = j \mid \xi_{t_1} \in B_1, \ldots, \xi_{t_n} \in B_n, \xi_s = i) = \\ = \mathsf{P}(\xi_t = j \mid \xi_s = i). \tag{1.1}$$

В приложениях значения случайных величин ξ_t интерпретируются как номера состояний изучаемой системы, которая в дискретные моменты времени t ($t = 0, 1, 2, \ldots$) меняет свое состояние. Свойство (1.1) означает, что при фиксированном положении системы в данный момент времени s будущее поведение системы ($t > s$) не зависит от поведения системы в прошлом ($\xi_{t_1} \in B_1, \ldots, \xi_{t_n} \in B_n$), или более кратко: при фиксированном настоящем будущее не зависит от прошлого. Свойство (1.1) называют *свойством марковости.*

Будем называть цепь Маркова ξ_t *однородной*, если при любых $i, j \in \mathcal{N}$ вероятности

$$\mathsf{P}(\xi_{t+1} = j \mid \xi_t = i) = p_{ij}, \quad t = 0, 1, 2, \ldots, \tag{1.2}$$

не зависят от t. Матрицу $P = \| p_{ij} \|$, элементами которой являются вероятности (1.2), называют *матрицей вероятностей перехода*, а вектор

$$q = (q_1, q_2, \ldots, q_N), \tag{1.3}$$

где $q_i = \mathsf{P}(\xi_0 = i)$, $i = 1, 2, \ldots, N$, — *вектором начальных вероятностей*. Очевидно, что числа p_{ij} и q_i удовлетворяют условиям

$$p_{ij} \geqslant 0, q_i \geqslant 0, \sum_{j=1}^{N} q_j = \sum_{j=1}^{N} p_{ij} = 1. \tag{1.4}$$

Любые матрицы $P = \| p_{ij} \|$, элементы которых удовлетворяют (1.4), называют *стохастическими*. Покажем, что матрица вероятностей перехода и вектор начальных вероятностей однозначно определяют совместные распределения величин $\xi_0, \xi_1, \ldots, \xi_t$ при любом t. По формуле (3.2.2) получаем

$$\mathsf{P}(\xi_0 = i_0, \ \xi_1 = i_1, \ \xi_2 = i_2, \ \ldots, \ \xi_{t-1} = i_{t-1}, \ \xi_t = i_t) =$$
$$= \mathsf{P}(\xi_0 = i_0)\, \mathsf{P}(\xi_1 = i_1 \mid \xi_0 = i_0) \times$$
$$\times \mathsf{P}(\xi_2 = i_2 \mid \xi_0 = i_0, \ \xi_1 = i_1) \ldots$$
$$\ldots \mathsf{P}(\xi_t = i_t \mid \xi_0 = i_0, \ldots, \xi_1 = i_1, \ldots, \xi_{t-1} = i_{t-1}). \tag{1.5}$$

Воспользуемся теперь следующим частным случаем равенства (1.1):

$$\mathsf{P}(\xi_s = i_s \mid \xi_0 = i_0, \ \xi_1 = i_1, \ \ldots, \ \xi_{s-1} = i_{s-1}) =$$
$$= \mathsf{P}(\xi_s = i_s \mid \xi_{s-1} = i_{s-1}), \ s = 1, 2, \ldots; \tag{1.6}$$
$$i_k \in \mathcal{N} \ (k = 0, 1, 2, \ldots).$$

Согласно (1.2) для однородных цепей Маркова правая часть (1.6) равна p_{i_{s-1}, i_s}. Заменив этими величинами и величинами (1.3) соответствующие сомножители в правой части (1.5), получим совместное распределение величин $\xi_0, \xi_1, \ldots, \xi_t$:

$$\mathsf{P}(\xi_0 = i_0, \ \xi_1 = i_1, \ \xi_2 = i_2, \ \ldots, \ \xi_t = i_t) =$$
$$= q_{i_0} p_{i_0, i_1} p_{i_1, i_2} \ldots p_{i_{t-1}, i_t}. \tag{1.7}$$

Мы определили цепь Маркова, сформулировав свойства, которыми должна обладать последовательность случайных величин ξ_t. При таком подходе остается открытым вопрос о существовании такой последовательности случайных величин. Покажем теперь, что при любом T, любом векторе $q = (q_1, \ldots, q_N)$ $\left(q_i \geqslant 0, \sum_{i=1}^{N} q_i = 1\right)$ и любой стохастической матрице $P = \| p_{ij} \|$ можно определить вероятностное пространство и случайные величины ξ_t, $t = 0, 1, 2, \ldots, T$, являющиеся цепью Маркова с вектором начальных вероятностей q и матрицей вероятностей перехода P. Пусть $\Omega = \{\omega\}$, где $\omega = (i_0, i_1, \ldots, i_T)$, $i_0, i_1, \ldots, i_T \in \mathcal{N}$. Положим

$$p(\omega) = q_{i_0} p_{i_0, i_1} p_{i_1, i_2} \ldots p_{i_{T-1}, i_T}. \tag{1.8}$$

Равенство (1.8) однозначно определяет вероятность на всех подмножествах $\mathfrak{A}$ множества Ω. Построенное вероятностное пространство в гл. 4 мы назвали цепью Маркова. Осталось только связать это определение с определением, данным в этой главе в терминах случайных величин. Положим

$$\xi_t(\omega) = \xi_t(i_0 i_1 \ldots i_T) = i_t, \quad 0 \leqslant t \leqslant T. \tag{1.9}$$

Для определенных таким образом случайных величин легко проверяется равенство (1.7). Используя (1.7), нетрудно вычислить условные вероятности в обеих частях равенств (1.7) или (1.1) и убедиться в том, что эти равенства верны. Указанные вычисления приводят к доказательству того, что случайные величины (1.9) являются цепью Маркова, у которой q — вектор начальных вероятностей и P — матрица вероятностей перехода.

Очевидно, что равенство (1.6) — это частный случай соотношения (1.1). С другой стороны, как отмечалось ранее, из (1.6) следует (1.7), а из (1.7) вытекает условие марковости (1.1). Таким образом, верно следующее утверждение: равенство (1.6) равносильно условию марковости (1.1).

Заметим также, что для однородной цепи Маркова ξ_t при любом s выполняется равенство

$$\mathsf{P}(\xi_{t+s} = j \mid \xi_s = i) = \mathsf{P}(\xi_t = j \mid \xi_0 = i), \quad i, j \in \mathcal{N}. \tag{1.10}$$

Это соотношение доказывается прямым вычислением условных вероятностей при помощи формулы (1.8). (См. задачи 2 и 3 гл. 8.) Поскольку вероятность (1.10) не зависит от s, то можно положить

$$\mathsf{P}(\xi_{t+s} = j \mid \xi_s = i) = P_{ij}(t). \tag{1.11}$$

Функции $P_{ij}(t)$, $i, j \in \mathcal{N}$, называют вероятностями перехода из состояния i в состояние j за время t.

Рассмотрим несколько примеров.

Пример 1. Блуждание частицы по целым точкам отрезка $[0, n]$, описанное в § 4 гл. 4, является цепью Маркова, в которой $q_l = 0$ $(l \neq k)$, $q_k = 1$; $p_{n,l} = 0$ $(l = 0, 1, \ldots, n-1)$; $p_{0,l} = 0$ $(l = 1, \ldots, n)$; $p_{0,0} = p_{n,n} = 1$; $p_{l,l+1} = p$ и $p_{l,l-1} = 1 - p$ $(l = 1, \ldots, n-1)$. Равенства (4.4.4) и (4.4.5) являются следствием однородности цепи Маркова (см. (1.10)).

Пример 2. Пусть

$$P=\begin{pmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ 0 & \frac{1}{2} & \frac{1}{2}\\ 0 & \frac{1}{2} & \frac{1}{2}\end{pmatrix},\ q_1=1,\quad q_2=q_3=0.$$

Так как переходы из 2-го и 3-го состояний в 1-е не происходят, то

$$\mathsf{P}(\xi_t=1)=\mathsf{P}(\xi_0=1,\ \xi_1=1,\ \ldots,\ \xi_t=1)=$$
$$=\mathsf{P}(\xi_0=1)\,\mathsf{P}(\xi_1=1\mid\xi_0=1)\ldots\mathsf{P}(\xi_t=1\mid\xi_{t-1}=1)=3^{-t}.$$

Событие A, состоящее в том, что цепь всегда будет находиться в 1-м состоянии, можно представить в виде $A=\bigcap_{t=1}^{\infty}\{\xi_t=1\}$, где $\{\xi_t=1\}$ — монотонно убывающая последовательность событий. Следовательно, по формуле (1.3.11) находим

$$\mathsf{P}(A)=\lim_{t\to\infty}\mathsf{P}(\xi_t=1)=\lim_{t\to\infty}3^{-t}=0.$$

Состояние i цепи Маркова называется *несущественным*, если существуют состояние j и число t_0 такие, что $P_{ij}(t_0)>0$, и $P_{ji}(t)=0$ при любом t. В противном случае состояние называется *существенным*.

В примере 1 существенными состояниями являются 0 и n, а остальные — несущественными. В примере 2 состояние 1 — несущественное, а 2 и 3 — существенные. Можно доказать, что система, описываемая цепью Маркова, уходит из несущественных состояний с вероятностью 1.

Пример 3. Положим

$$P=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad q_1=1,\quad q_2=0.$$

Если в момент $t=0$ система находилась в состоянии 1, то в нечетные моменты времени система будет находиться в состоянии 2, а в четные — в состоянии 1. Таким образом,

$$P_{12}(t)=\frac{1-(-1)^t}{2},\quad P_{11}(t)=\frac{1+(-1)^t}{2}.$$

Пример 4. Рассмотрим $T+1$ испытаний Бернулли; вероятность успеха в испытании с номером t, $t=0,1,$

2, ..., T, равна p $(0 < p < 1)$. Пусть случайная величина $\xi_t = 1$, если общее число успехов в испытаниях с номерами $s = 0, 1, 2, \ldots, t$ было нечетным, и $\xi_t = 2$ в противном случае.

Равенство

$$\mathsf{P}(\xi_{t+1} = j \mid \xi_0 = i_0, \ \xi_1 = i_1, \ \ldots, \ \xi_{t-1} = i_{t-1}, \ \xi_t = i) = \\ = \mathsf{P}(\xi_{t+1} = j \mid \xi_t = i), \ i, j = 1, 2, \quad (1.12)$$

достаточно очевидно: если в момент t четность числа успехов известна, то четность в следующий момент определяется только результатом очередного испытания, а испытания по условию независимы. Равенство (1.12) легко доказать, если воспользоваться тем, что событие

$$(\xi_0 = i_0, \ \xi_1 = i_1, \ \ldots, \ \xi_t = i, \ \xi_{t+1} = j) \qquad (1.13)$$

однозначно определяет исходы в испытаниях Бернулли с номерами 0, 1, ..., $t+1$, а событие $(\xi_t = i, \ \xi_{t+1} = j)$ является конечной суммой событий вида (1.13). Из (1.12) следует, что ξ_t является цепью Маркова. Поскольку состояние цепи ξ_t меняется лишь в случае успеха в очередном испытании, то непосредственно из определения ξ_t находим

$$p_{11} = \mathsf{P}(\xi_1 = 1 \mid \xi_0 = 1) = 1 - p = q,$$

$$p_{22} = \mathsf{P}(\xi_1 = 2 \mid \xi_0 = 2) = q.$$

Аналогично получим $p_{12} = p_{21} = p$, $q_1 = p$, $q_2 = q$. Нетрудно вычислить вероятности перехода $P_{ij}(t)$. Например,

$$P_{21}(t) = P(\xi_t = 1 \mid \xi_0 = 2) = \sum_{k=0}^{[(t-1)/2]} C_t^{2k+1} p^{2k+1} q^{t-2k-1}.$$

§ 2. Уравнения для вероятностей перехода

Кроме вероятностей «сплошных» цепочек исходов (1.7), нередко приходится вычислять вероятности цепочек вида

$$(\xi_{t_1} = l_1, \ \xi_{t_2} = l_2, \ \ldots, \ \xi_{t_s} = l_s), \qquad (2.1)$$

где моменты времени $t_1, \ldots, t_s$ уже не обязательно являются соседними. Вероятность события (2.1) можно выразить через вероятности перехода $P_{ij}(t)$. По формуле полной вероятности

$$\mathsf{P}(\xi_{t_1}=l_1, \ldots, \xi_{t_s}=l_s) = \\ = \sum_{k=1}^{N} P(\xi_0=k)\, P(\xi_{t_1}=l_1, \ldots, \xi_{t_s}=l_s \mid \xi_0=k). \tag{2.2}$$

Преобразуя второй сомножитель под знаком суммы по формуле (3.2.2) и используя условия марковости (1.1) и условия однородности (1.10) (1.11), получим

$$\mathsf{P}(\xi_{t_1}=l_1, \ldots, \xi_{t_s}=l_s) = \\ = \sum_{k=1}^{N} q_k P_{k, l_1}(t_1)\, P_{l_1, l_2}(t_2-t_1) \ldots P_{l_{s-1}, l_s}(t_s-t_{s-1}). \tag{2.3}$$

Для вычислений по формуле (2.3) нужно уметь находить $P_{ij}(t)$.

Теорема 2.1. *При любых s, t*

$$P_{ij}(t+s) = \sum_{k=1}^{N} P_{ik}(s)\, P_{kj}(t), \quad i, j = 1, 2, \ldots, N. \tag{2.4}$$

Доказательство. Вычислим вероятность $\mathsf{P}(\xi_{t+s}=j \mid \xi_0=i)$ по формуле полной вероятности (3.3.1), положив $B_k = (\xi_s = k)$:

$$\mathsf{P}(\xi_{t+s}=j \mid \xi_0=i) = \\ = \sum_{k=1}^{N} \mathsf{P}(\xi_s=k \mid \xi_0=i)\, \mathsf{P}(\xi_{t+s}=j \mid \xi_0=i, \xi_s=k). \tag{2.5}$$

Из равенств (1.1) и (1.10) следует

$$\mathsf{P}(\xi_{t+s}=j \mid \xi_0=i, \ \xi_s=k) = \\ = \mathsf{P}(\xi_{t+s}=j \mid \xi_s=k) = \mathsf{P}(\xi_t=j \mid \xi_0=k).$$

Отсюда и из равенств (2.5) и (1.12) получаем утверждение теоремы.

Определим матрицу $P(t) = \| P_{ij}(t) \|$. В матричной записи (2.4) имеет вид

$$P(t+s) = P(s)\, P(t). \tag{2.6}$$

Так как $P_{ij}(1) = p_{ij}$, то $P(1) = P$, где P — матрица вероятностей перехода. Из (2.6) следует

$$P(t) = (P(1))^t = P^t. \tag{2.7}$$

Результаты, полученные в теории матриц, позволяют по формуле (2.7) вычислять $P_{ij}(t)$ и исследовать их поведение при $t \to \infty$. Некоторые теоремы о неотрицательных матрицах приведены в монографии [15]; подробное исследование стохастических матриц и цепей Маркова содержится в книге [14].

§ 3. Стационарное распределение. Теорема о предельных вероятностях

Распределение вероятностей ξ_t в произвольный момент времени t можно найти, воспользовавшись формулой полной вероятности

$$\mathsf{P}(\xi_t = j) = \sum_{k=1}^{N} q_k P_{kj}(t). \tag{3.1}$$

Может оказаться, что $\mathsf{P}(\xi_t = j)$ не зависит от времени. Назовем *стационарным распределением* вектор

$q^* = (q_1^*, \ldots, q_N^*)$, удовлетворяющий условиям

$$q_k^* \geqslant 0, \quad k = 1, \ldots, N;$$

$$\sum_{k=1}^{N} q_k^* = 1, \quad \sum_{k=1}^{N} q_k^* p_{kj} = q_j^*, \quad j = 1, 2, \ldots, N, \tag{3.2}$$

где p_{ij} — вероятности перехода (1.2).

Если в цепи Маркова $\mathsf{P}(\xi_0 = k) = q_k = q_k^*$, то при любом t

$$\mathsf{P}(\xi_t = k) = q_k^*, \; k = 1, 2, \ldots, N.$$

Это утверждение следует по индукции из (3.1) и (3.2).

Приведем формулировку теоремы о предельных вероятностях для одного важного класса цепей Маркова.

Теорема 3.1. *Если при некотором $t_0 > 0$ все элементы матрицы P^{t_0} положительны, то для любых $i, j = 1, 2, \ldots, N$ при $t \to \infty$*

$$P_{ij}(t) = q_j^* + \varepsilon_{ij}(t), \tag{3.3}$$

где $q^* = (q_1^*, \ldots, q_N^*)$ — стационарное распределение с $q_k^* > 0$, $k = 1, \ldots, N$, $\varepsilon_{ij}(t) = O(h^t)$, а h — некоторая постоянная, удовлетворяющая неравенствам $0 < h < 1$.

Доказательство этой теоремы, опирающееся на факты из теории матриц, содержится в книге [14]; прямое доказательство имеется в книгах [4] и [13]. В § 4 приведено прямое доказательство.

Так как $P(t_0) = P^{t_0}$, то по условию теоремы из любого состояния можно попасть в любое за время t_0 с положительной вероятностью. Условия теоремы исключают цепи, являющиеся в некотором смысле периодическими (см. пример 3 из § 1).

Если выполнены условия теоремы 3.1, то вероятность того, что система находится в некотором состоянии j, в пределе не зависит от начального распределения. Действительно, из (3.3) и (3.1) следует, что при любом начальном распределении $q_i = \mathsf{P}(\xi_0 = i)$, $i = 1, \ldots, N$,

$$\lim_{t\to\infty} \mathsf{P}(\xi_t = j) = q_j^*, \quad j = 1, 2, \ldots, N.$$

Рассмотрим несколько примеров цепей Маркова, для которых условия теоремы 3.1 не выполнены. Нетрудно проверить, что такими примерами являются примеры 1, 2, 3 из § 1. В примере 1 при $t \to \infty$ вероятности перехода имеют пределы, но эти пределы зависят от начального состояния. В частности, при $p = q = 1/2$ (см. § 4. гл. 4)

$$\lim_{t\to\infty} P_{kl}(t) = 0, \quad 0 < l < n, \quad k = 0, 1, \ldots, n,$$

$$\lim_{t\to\infty} P_{k0}(t) = 1 - \frac{k}{n}, \quad \lim_{t\to\infty} P_{kn}(t) = \frac{k}{n},$$

$$k = 0, 1, \ldots, n.$$

В примере 3 пределы вероятностей $P_{12}(t)$, $P_{11}(t)$ при $t \to \infty$, очевидно, не существуют.

Найдем стационарное распределение в примере 2. Нужно найти вектор $q^* = (q_1^*, q_2^*, q_3^*)$, удовлетворяющий условиям (3.2):

$$\begin{cases} q_1^* \dfrac{1}{3} & = q_1^*, \\ q_1^* \dfrac{1}{3} + q_2^* \dfrac{1}{2} + q_3^* \dfrac{1}{2} & = q_2^*, \\ q_1^* \dfrac{1}{3} + q_2^* \dfrac{1}{2} + q_3^* \dfrac{1}{2} & = q_3^*; \end{cases}$$

$$q_k^* \geqslant 0, \quad k = 1, 2, 3, \quad q_1^* + q_2^* + q_3^* = 1.$$

Отсюда $q_1^* = 0$, $q_2^* = q_3^* = 1/2$. Таким образом, стационарное распределение существует, но не все координаты вектора q^* положительны.

В примере 4 из § 1 условия теоремы 3.1 выполнены уже при $t_0 = 1$. Стационарное распределение (q_1^*, q_2^*) удовлетворяет уравнениям

$$q_1^* + q_2^* = 1, \quad q_1^* p_{11} + q_2^* p_{21} = q_1^*, \quad q_1^* p_{12} + q_2^* p_{22} = q_2^*,$$

где $p_{11} = p_{22} = q$, $p_{12} = p_{21} = p$. Отсюда $q_1^* = q_2^* = 1/2$ и по теореме 3.1 $\lim\limits_{t\to\infty} P_{ij}(t) = 1/2$, $i, j = 1, 2$.

Во всех разобранных примерах оказалось, что система однородных уравнений

$$(p_{jj} - 1) q_j^* + \sum_{k \neq j} q_k^* p_{kj} = 0, \quad j = 1, 2, \ldots, N, \tag{3.4}$$

относительно неизвестных $q_1^*, \ldots, q_N^*$ имеет ненулевое решение. Нетрудно проверить, что определитель матрицы коэффициентов системы (3.4) равен нулю (если все столбцы определителя сложить с первым, то в силу свойства (1.4) получим столбец из нулей).

В § 2 гл. 4 для полиномиальной схемы были введены случайные величины, равные числу исходов данного типа. Введем аналогичные величины для цепей Маркова. Пусть $\tau_k(t)$, $k = 1, \ldots, N$ — число попаданий системы в состояние k за время t. Тогда $\tau_k(t)/t$ — частота попаданий системы в состояние k. Используя формулы (3.3), можно доказать, что частота $\tau_k(t)/t$ при $t \to \infty$ сближается с q_k^*. Для этого нужно получить асимптотические формулы для $\mathsf{M}\tau_k(t)$ и $\mathsf{D}\tau_k(t)$ и воспользоваться неравенством Чебышева. Приведем вывод формулы для $\mathsf{M}(\tau_j(t) \mid \xi_0 = i) = m_{ij}(t)$. Представим $\tau_j(t)$ в виде

$$\tau_j(t) = \eta_0 + \eta_1 + \ldots + \eta_t, \tag{3.5}$$

где $\eta_s = 1$, если $\xi_s = j$, и $\eta_s = 0$ в противном случае. Так как

$$\mathsf{M}(\eta_s \mid \xi_0 = i) = P_{ij}(s),$$

то, воспользовавшись свойствами математического ожидания и формулой (3.3), получим

$$m_{ij}(t) = \sum_{s=0}^{t} P_{ij}(s) = q_j^* \cdot t + \sum_{s=0}^{t} \varepsilon_{ij}(s).$$

Второе слагаемое в правой части этого равенства в силу теоремы 3.1 является частной суммой сходящегося ряда. Положив $\alpha_{ij} = \sum\limits_{s=0}^{\infty} \varepsilon_{ij}(s)$, получим

$$m_{ij}(t) = q_j^* \cdot t + \alpha_{ij} + O(h^t), \quad 0 < h < 1, \tag{3.6}$$

поскольку

$$\sum_{s=0}^{t} \varepsilon_{ij}(s) = \alpha_{ij} - \sum_{s=t+1}^{\infty} \varepsilon_{ij}(s), \quad |\varepsilon_{ij}(s)| < Kh^s.$$

Из формулы (3.6), в частности, следует, что

$$\mathsf{M}\left(\frac{\tau_j(t)}{t}\,\middle|\,\xi_0 = i\right) \to q_j^*, \quad j = 1, 2, \ldots, N,$$

при $t \to \infty$. Аналогично можно получить формулу для $\mathsf{M}\,[\tau_j(t)(\tau_j(t) - 1) \mid \xi_0 = i]$, которая используется для вычисления дисперсии.

§ 4. Доказательство теоремы о предельных вероятностях в цепи Маркова

Докажем сначала две леммы. Положим

$$v_l(t) = \min_i P_{il}(t), \quad V_l(t) = \max_j P_{jl}(t).$$

Лемма 4.1. *При любых l существуют пределы*

$$\lim_{t\to\infty} v_l(t) = q_l, \quad \lim_{t\to\infty} V_l(t) = Q_l$$

и

$$q_l \leqslant Q_l.$$

Доказательство. Используя уравнение (2.4) с $s = 1$, получим

$$v_j(t+1) = \min_i \sum_{k=1}^{N} p_{ik} P_{kj}(t) \geqslant \min_i \sum_{k=1}^{N} p_{ik} v_j(t) = v_j(t),$$

$$V_j(t+1) = \max_i \sum_{k=1}^{N} p_{ik} P_{kj}(t) \leqslant \max_i \sum_{k=1}^{N} p_{ik} V_j(t) = V_j(t).$$

Таким образом, последовательности $v_j(t)$ и $V_j(t)$ монотонны и ограничены. Отсюда следует утверждение леммы 4.1.

Лемма 4.2. *Если выполнены условия теоремы* 3.1, *то существуют постоянные C, δ $(C > 0,\ 0 < \delta < 1)$ такие, что*

$$V_j(nt_0) - v_j(nt_0) \leqslant C\delta^n, \quad n = 1, 2, \ldots$$

Доказательство. Для любых i, j

$$0 = \sum_{k=1}^{N} P_{ik}(t_0) - \sum_{k=1}^{N} P_{jk}(t_0) = \sum_k{}^{+} \Delta_k(i, j) + \sum_k{}^{-} \Delta_k(i, j),$$

где $\Delta_k(i, j) = P_{ik}(t_0) - P_{jk}(t_0)$, $\sum^{+}$ означает суммирование по всем k, при которых $\Delta_k(i, j)$ положительны, а $\sum^-$ — суммирование по остальным k. Отсюда

$$\sum_k^+ \Delta_k(i, j) = -\sum_k^- \Delta_k(i, j). \tag{4.1}$$

Так как в условиях теоремы 3.1 вероятности перехода $P_{ij}(t_0) > 0$ при всех i, j, то при любых i, j

$$\sum_k^+ \Delta_k(i, j) = \sum_k^+ (P_{ik}(t_0) - P_{jk}(t_0)) < \sum_{k=1}^N P_{ik}(t_0) = 1 \tag{4.2}$$

и в силу конечности числа состояний

$$\delta = \max_{i,j} \sum_k^+ (P_{ik}(t_0) - P_{jk}(t_0)) < 1. \tag{4.3}$$

Оценим теперь разность $V_j(nt_0) - v_j(nt_0)$. Используя уравнение (2.4) с $s = t_0$, $t = nt_0$, получим

$$V_l((n+1)t_0) - v_l((n+1)t_0) = \max_{i,j} [P_{il}((n+1)t_0) - P_{jl}((n+1)t_0)] =$$

$$= \max_{i,j} \sum_{k=1}^N [P_{ik}(t_0) - P_{jk}(t_0)] P_{kl}(nt_0) = \max_{i,j} \sum_{k=1}^N \Delta_k(i, j) P_{kl}(nt_0) \leqslant$$

$$\leqslant \max_{i,j} \left\{ \sum_k^+ \Delta_k(i, j) V_l(nt_0) + \sum_k^- \Delta_k(i, j) v_l(nt_0) \right\}.$$

Отсюда, используя (4.1), (4.2) и (4.3), найдем

$$V_l((n+1)t_0) - v_l((n+1)t_0) \leqslant$$

$$\leqslant \max_{i,j} \left\{ (V_l(nt_0) - v_l(nt_0)) \sum_k^+ \Delta_k(i, j) \right\} = \delta (V_l(nt_0) - v_l(nt_0)).$$

Объединяя это соотношение с неравенством (4.3), получаем утверждение леммы.

Перейдем к доказательству теоремы. Так как последовательности $v_l(t)$, $V_l(t)$ монотонны, то

$$0 < v_l(t) \leqslant q_l \leqslant Q_l \leqslant V_l(t). \tag{4.4}$$

Отсюда и из леммы 4.2 находим

$$0 \leqslant Q_l - q_l \leqslant V_l(nt_0) - v_l(nt_0) \leqslant C\delta^n.$$

Следовательно, при $n \to \infty$ получим $Q_l = q_l = q_l^*$ и

$$|P_{il}(t) - q_l^*| \leqslant V_l(t) - v_l(t) \leqslant V_l\left(\left[\frac{t}{t_0}\right] t_0\right) - v_l\left(\left[\frac{t}{t_0}\right] t_0\right) \leqslant C\delta^{\frac{t}{t_0} - 1}. \tag{4.5}$$

Положительность q_l^* следует из неравенств (4.4). Переходя к пределу при $t \to \infty$ и $s = 1$ в уравнении (2.4), получим, что q_l^* удовлетворяют уравнению (3.2). Таким образом, теорема доказана.

Задачи к главе 8

1. Матрица вероятностей перехода цепи Маркова имеет вид

$$P = \begin{pmatrix} 0,1 & 0,5 & 0,4 \\ 0,6 & 0,2 & 0,2 \\ 0,3 & 0,4 & 0,3 \end{pmatrix}.$$

Распределение по состояниям в момент $t=0$ определяется вектором $q = (0,7;\ 0,2;\ 0,1)$. Найти: 1) распределение по состояниям в момент $t = 2$; 2) вероятность того, что в моменты $t = 0, 1, 2, 3$ состояниями цепи будут соответственно 1, 3, 3, 2; 3) стационарное распределение.

2. Пусть ξ_t, $t = 0, 1, 2, \ldots$,— цепь Маркова с матрицей переходных вероятностей $P = \| p_{ij} \|$ и начальным распределением по состояниям $q = (q_1, \ldots, q_N)$. Выразить в виде сумм вероятностей вида (1.7) следующие вероятности: а) $\mathsf{P}(\xi_s = i)$; б) $\mathsf{P}(\xi_s = i, \xi_{t+s} = j)$.

3 (продолжение задачи 2). Доказать, что

$$\mathsf{P}(\xi_{t+s} = j \mid \xi_s = i) = \mathsf{P}(\xi_t = j \mid \xi_0 = i).$$

Указание. Вычислить обе условные вероятности, воспользовавшись результатом задачи 2.

4. Пусть ξ_t, $t = 1, 2, \ldots, T$,— независимые одинаково распределенные случайные величины, $\mathsf{P}(\xi_t = 1) = p$, $\mathsf{P}(\xi_t = -1) = 1 - p$. Являются ли цепью Маркова величины: а) $\eta_t = \xi_t \xi_{t+1}$; б) $\eta_t = \xi_1 \xi_2 \ldots \xi_t$; в) $\eta_t = \varphi(\xi_t, \xi_{t+1})$, где $\varphi(-1, -1) = 1$, $\varphi(-1, 1) = 2$, $\varphi(1, -1) = 3$, $\varphi(1, 1) = 4$? Для цепей Маркова найти матрицы вероятностей перехода за один шаг.

5. Пусть ξ_t — номер состояния в цепи Маркова в момент времени t, $\mathsf{P}(\xi_0 = 1) = 1$ и матрица вероятностей перехода за единицу времени равна

$$\begin{pmatrix} 3/7 & 3/7 & 1/7 \\ 1/11 & 2/11 & 8/11 \\ 1/11 & 4/11 & 8/11 \end{pmatrix}; \qquad \eta_t = \begin{cases} 1, & \text{если } \xi_t = 1, \\ 2, & \text{если } \xi_t \neq 1. \end{cases}$$

Показать, что последовательность η_t является цепью Маркова. Найти соответствующую ей матрицу вероятностей перехода.

6. В N ячейках последовательно независимо друг от друга равновероятно размещают частицы. Пусть $\mu_0(n)$ — число ячеек, оставшихся пустыми после размещения n частиц. Показать, что последовательность $\mu_0(n)$, $n = 0, 1, 2, \ldots$ является цепью Маркова. Найти вероятности перехода.

7. Пусть $\eta_1, \eta_2, \ldots$— последовательность независимых одинаково распределенных случайных величин и $f(x, y)$ — функция, принимающая значения в множестве $\{1, \ldots, N\}$; $x = 1, 2, \ldots, N$, множество значений y совпадает с множеством значений η_1. Является ли последовательность случайных величин

$$\xi_0 \, (\mathsf{P}(\xi_0 = k) = p_k^{(0)}, \qquad k = 0, 1, 2, \ldots, N),$$

$$\xi_{t+1} = f(\xi_t, \eta_{t+1}), \quad t = 0, 1, 2, \ldots,$$

цепью Маркова?

8. Игральная кость все время перекладывается случайным образом с одной грани равновероятно на любую из четырех соседних граней независимо от предыдущего. К какому пределу стремит-

ся при $t \to \infty$ вероятность того, что в момент времени t кость лежит на грани «6», если в момент $t = 0$ она находилась в этом же положении ($t = 0, 1, 2, \ldots$)?

9. По двум урнам разложено N черных и N белых шаров так, что каждая урна содержит N шаров. Число черных шаров в первой урне определяет состояние системы. В каждый момент времени выбирают случайно по одному шару из урны и выбранные шары меняют местами. Найти вероятности перехода и показать, что стационарные вероятности $q_k^* = C_N^k C_N^{N-k}/C_{2N}^N$, $k = 1, \ldots, N$.

10. Матрица вероятностей перехода P и вектор q начального распределения по состояниям имеют вид: $q = (1/2,\ 1/2,\ 0, 0, 0, 0)$,

$$P = \begin{pmatrix} 3/12 & 2/12 & 1/12 & 3/12 & 1/12 & 2/12 \\ 1/12 & 1/12 & 3/12 & 1/12 & 4/12 & 2/12 \\ 0 & 0 & 3/4 & 1/4 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 2/3 \\ 0 & 0 & 0 & 0 & 2/3 & 1/3 \end{pmatrix}.$$

Найти: а) несущественные состояния; б) математическое ожидание τ — времени до выхода из несущественных состояний; в) вероятности $p_i^{(\alpha)}$, $p_i^{(\beta)}$ попадания в множества состояний $\alpha = \{3, 4\}$, $\beta = \{5, 6\}$, если начальным состоянием является $i \in \{1, 2\}$; г) предельное при $t \to \infty$ распределение по состояниям, т. е. величины $\pi_k = \lim\limits_{t\to\infty} \mathsf{P}\,(\xi_t = k)$.

Указание. б) Используя формулу полного математического ожидания, составить систему линейных уравнений для $m_i = \mathsf{M}\,(\tau \mid \xi_0 = i)$, $i = 1, 2$. Воспользоваться формулой $\mathsf{M}\tau = \mathsf{P}\,(\xi_0 = 1)\, m_1 + \mathsf{P}\,(\xi_0 = 2)\, m_2$. в) Используя формулу полной вероятности, составить системы линейных уравнений для $p_1^{(\alpha)}$, $p_2^{(\alpha)}$ и для $p_1^{(\beta)}$, $p_2^{(\beta)}$. г) Найти предельное распределение по состояниям, когда начальным состоянием является состояние из множества α и из множества β. Использовать формулу полной вероятности и решение в).

11. Матрица вероятностей перехода $P = \| p_{ij} \|$ цепи Маркова ξ_t определяется формулами $p_{11} = 1 - \alpha$, $p_{12} = \alpha$, $p_{21} = \beta$, $p_{22} = 1 - \beta$. Доказать, что

$$\begin{pmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{pmatrix} = \frac{1}{\alpha+\beta}\begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + \frac{(1-\alpha-\beta)^t}{\alpha+\beta}\begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix}.$$

Найти стационарные вероятности.

12. Для цепи Маркова, определенной в задаче 11, обозначим τ_t число попаданий в состояние 1 за время t. Доказать, что

$$\lim_{t\to\infty} \mathsf{P}\left(\left|\frac{\tau_t}{t} - \frac{\beta}{\alpha+\beta}\right| < \varepsilon\right) = 1 \text{ при любом } \varepsilon > 0.$$

Указание. Показать, что $\mathsf{M}\tau_t = \dfrac{\beta}{\alpha+\beta}\, t + O(1)$ и $\mathsf{D}\tau_t = o(t^2)$ при $t \to \infty$. Использовать представление τ_t в виде суммы (3.5) и решение задачи 11.

Элементы математической статистики

§ 1. Задачи математической статистики

Пусть требуется измерить некоторую величину a. Результаты измерений $x_1(\omega), x_2(\omega), \ldots, x_n(\omega)$ естественно рассматривать как значения случайных величин $x_1, x_2, \ldots$ $\ldots, x_n$, полученные в данном опыте с исходом ω. Если измерительный прибор не дает систематической ошибки, то можно предположить, что $\mathsf{M}x_k = a$. Таким образом, по результатам наблюдений $x_1, x_2, \ldots, x_n$ нужно определить неизвестный параметр a. Это типичная *задача оценки неизвестных параметров*. Общая ошибка измерения часто складывается из большого числа ошибок, каждая из которых невелика. В такой ситуации на основании центральной предельной теоремы становится правдоподобным следующее предположение (гипотеза): случайные величины x_k имеют нормальное распределение. Таким образом, мы пришли к *задаче статистической проверки гипотезы о законе распределения*.

В § 3 гл. 1 рассматривались две различные модели (модель 1 и модель 2) подбрасывания двух монет. Вероятность выпадения монет одинаковыми сторонами в модели 1 оказались равной 2/3, а в модели 2 — равной 1/2. По результатам подбрасываний двух монет требуется выбрать одну из двух моделей (или гипотез).

К задачам оценки параметров часто сводятся задачи, в которых нужно установить зависимость между переменными. Пусть, например, из некоторых соображений известно, что переменная y линейно зависит от переменных $x_1, x_2, \ldots, x_k$: $y = A_0 + A_1x_1 + \ldots + A_kx_k$. Коэффициенты $A_0, A_1, \ldots, A_k$ неизвестны. При различных наборах $(x_{i1}, x_{i2}, \ldots, x_{ik})$, $i = 1, \ldots, n$, измерены значения $y_i = A_0 + A_1x_{i1} + \ldots + A_kx_{ik} + \delta_i$, где δ_i — ошибки измерения y при наборе $(x_{i1}, x_{i2}, \ldots, x_{ik})$. По значениям $(y_i, x_{i1}, \ldots, x_{ik})$ требуется оценить коэффициенты $A_0, A_1, \ldots, A_k$. Задачи такого типа называют *регрессионными*.

В перечисленных выше задачах оценка неизвестных параметров или выбор гипотезы проводится по результатам наблюдений или измерений, т. е. на основе эмпирических данных. В следующем параграфе приводится описание математической модели независимых измерений.

§ 2. Понятие выборки. Выборочные распределения

Чтобы сделать формальное определение выборки более понятным, рассмотрим несколько примеров. Пусть нам требуется оценить неизвестную долю белых шаров в урне, содержащей белые и черные шары. По схеме случайного выбора с возвращением отберем n шаров. При больших n доля белых шаров в полученной выборке близка к доле белых шаров в урне. Указанная процедура может быть описана в терминах случайных величин. Пусть $x_k = 1$, если k-й шар выборки оказался белым, и $x_k = 0$ в противном случае ($k = 1, 2, \ldots, n$). Тогда случайные величины $x_1, x_2, \ldots, x_n$ независимы и одинаково распределены. Доля белых шаров $\bar{x}$ в выборке может быть вычислена по формуле $\bar{x} = (x_1 + x_2 + \ldots + x_n)/n$.

При измерении числовых характеристик часто оказывается, что результат каждого измерения можно рассматривать как значение $\xi(\omega)$ некоторой случайной величины ξ, которое эта величина принимает в данном исходе опыта ω. Функция распределения $F_\xi(x)$ величины ξ характеризует используемый для измерений прибор. Несколько независимых измерений, проводимых в одинаковых условиях, можно описать соответствующим количеством случайных величин, каждая из которых имеет функцию распределения $F_\xi(x)$.

Назовем *случайной выборкой объема n* (или просто *выборкой*) случайный вектор $(x_1, x_2, \ldots, x_n)$, где x_k, $k = 1, \ldots, n$, независимы и одинаково распределены с $\mathsf{P}(x_k < x) = F_\xi(x)$. Случайная выборка является математической моделью независимых измерений, проводимых в одинаковых условиях.

Иногда в задачах, связанных с измерениями, оказывается, что величины $x_1, \ldots, x_n$ можно считать нормально распределенными. В задаче с подбрасыванием двух монет (см. § 3 гл. 1) имеем дискретные величины $\mathsf{P}(x_k = 1) = p_l$, $\mathsf{P}(x_k = 0) = 1 - p_l$, $l = 1, 2$, где $x_k = 1$, если в k-м испытании монеты выпали одинаковыми сторонами; $p_1 = 2/3$, $p_2 = 1/2$. При оценке доли брака в партии изделий по схеме случайного выбора без возвращения величины, входящие в выборку, оказываются зависимыми. Обычно мы будем рассматривать случай независимых измерений.

Одной из основных характеристик выборки является *эмпирическая функция распределения* $F_n^*(x)$, определяемая формулой

$$F_n^*(x) = \frac{\mu_n(x)}{n}, \tag{2.1}$$

где $\mu_n(x)$ — число значений среди $x_1, \ldots, x_n$, меньших x. Таким образом, при любом x величины $\mu_n(x)$, а следовательно, и $F_n^*(x)$ являются случайными.

Случайные величины выборки $x_1, x_2, \ldots, x_n$, расположенные при каждом ω в порядке возрастания их значений:

$$x_{(1)} \leqslant x_{(2)} \leqslant \ldots \leqslant x_{(n)},$$

определяют новые случайные величины, называемые *вариационным рядом*. В частности, $x_{(1)} = \min\{x_1, \ldots, x_n\}$, $x_{(n)} = \max\{x_1, x_2, \ldots, x_n\}$.

Покажем, что при больших n эмпирическая функция распределения $F_n^*(x)$ близка к функции распределения $F(x)$.

Теорема 2.1 *Для любого x $(-\infty < x < \infty)$ и любого $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathsf{P}\{|F_n^*(x) - F(x)| < \varepsilon\} = 1.$$

Доказательство. С каждым x_k $(k = 1, \ldots, n)$ можно связать два события: $\{x_k < x\}$ и $\{x_k \geqslant x\}$. Вероятности этих событий, очевидно, равны

$$p = \mathsf{P}\{x_k < x\} = F(x), \quad q = \mathsf{P}\{x_k \geqslant x\} = 1 - F(x).$$

Если событие $\{x_k < x\}$ назвать успехом, то $\mu_n(x)$ является числом успехов в серии из n независимых испытаний. Так как в рассматриваемой последовательности испытаний $F_n^*(x) = \mu_n(x)/n$ является частотой успеха, то утверждение доказываемой теоремы сразу следует из теоремы 5.6 гл. 6.

Иногда для наглядного представления выборки используется *гистограмма*, получаемая следующим образом. Числовая ось разбивается на несколько непересекающихся полуинтервалов: $(-\infty, \infty) = \bigcup\limits_{k=0}^{r} [z_k, z_{k+1})$, где $-\infty = z_0 < z_1 < z_2 < \ldots < z_r < z_{r+1} = \infty$. Далее вычисляются частоты $p_k^* = F_n^*(z_{k+1}) - F_n^*(z_k)$ попадания элементов выборки в эти полуинтервалы. Затем над каждым отрезком $[z_k, z_{k+1}]$ строится прямоугольник, площадь которого пропорциональна частоте p_k^*. Так же как при доказательстве теоремы 2.1, можно показать, что p_k^* при больших n близки к $p_k = F(z_{k+1}) - F(z_k)$ или (если $F(x)$ имеет плотность распределения $p(x)$) к $p_k = \int\limits_{z_k}^{z_{k+1}} p(x)\,dx$. Таким образом, при удачном выборе ширины интервалов гистограмма может напоминать график плотности распределения $p(x)$.

Рассмотрение графиков эмпирической функции распределения и гистограммы может дать некоторое предварительное представление о неизвестной функции распределения $F(x)$.

Более точные выводы о неизвестной функции распределения $F(x)$ могут быть сделаны на основе следующей теоремы А. Н. Колмогорова.

Т е о р е м а 2.2. *Если функция $F(x)$ непрерывна, то при $n \to \infty$*

$$\mathsf{P}\{\sqrt{n}\,(\sup_{-\infty<x<\infty} |F_n^*(x) - F(x)|) < z\} \to K(z) = \begin{cases} 0 & \text{при } z \leqslant 0, \\ \sum\limits_{k=-\infty}^{\infty} (-1)^k e^{-2k^2z^2} & \text{при } z > 0. \end{cases}$$

Теорема А. Н. Колмогорова позволяет указать границы, в которых с большой вероятностью будет заключена неизвестная функция распределения. Если z_α подобрано так, что $1 - K(z_\alpha) = \alpha$, то неравенства

$$F_n^*(x) - \frac{z_\alpha}{\sqrt{n}} < F(x) < F_n^*(x) + \frac{z_\alpha}{\sqrt{n}}$$

выполняются сразу для всех x с вероятностью, близкой к $1 - \alpha$.

§ 3. Выборочные моменты

3.1. Математическое ожидание и дисперсия выборочных моментов. Для каждой реализации измерений $x_1(\omega)$, $x_2(\omega), \ldots, x_n(\omega)$ эмпирическая функция распределения

$F_n^*(x)$ является функцией распределения некоторой дискретной случайной величины, принимающей n значений: $x_1(\omega)$, $x_2(\omega)$, ..., $x_n(\omega)$ — с вероятностями, равными $1/n$*). При различных ω соответствующие функции $F_n^*(x)$ различны. При каждом ω можно ввести различные числовые характеристики соответствующего данному ω закона распределения, определяемого $F_n^*(x)$. Эти характеристики будем называть *выборочными*. *Выборочные моменты* и *центральные выборочные моменты* порядка ν вычисляются по формулам

$$a_\nu = \frac{1}{n}\sum_{k=1}^{n} x_k^\nu, \quad m_\nu = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^\nu, \tag{3.1}$$

где

$$\bar{x} = a_1 = \frac{1}{n}\sum_{k=1}^{n} x_k \tag{3.2}$$

— *выборочное математическое ожидание* (или *выборочное среднее*).

Можно показать, что при достаточно общих ограничениях на неизвестную функцию распределения $F_\xi(x)$ выборочные характеристики (3.1) близки к соответствующим теоретическим характеристикам $F_\xi(x)$:

$$\alpha_\nu = \mathsf{M}\xi^\nu, \; \mu_\nu = \mathsf{M}(\xi - \mathsf{M}\xi)^\nu. \tag{3.3}$$

Найдем сначала математическое ожидание и дисперсию выборочных моментов. Так как x_k независимы и распределены так же, как ξ, то

$$\mathsf{M}\bar{x} = \frac{1}{n}\sum_{k=1}^{n} \mathsf{M}x_k = \mathsf{M}\xi = \alpha_1,$$

$$\mathsf{D}\bar{x} = \frac{1}{n^2}\sum_{k=1}^{n} \mathsf{D}x_k = \frac{1}{n}\mathsf{D}\xi = \mu_2/n.$$

Таким образом,

$$\mathsf{M}\bar{x} = \alpha_1, \quad \mathsf{D}\bar{x} = \frac{\mu_2}{n}. \tag{3.4}$$

Аналогично находим

$$\mathsf{M}a_\nu = \alpha_\nu, \quad \mathsf{D}a_\nu = \frac{\alpha_{2\nu} - \alpha_\nu^2}{n}. \tag{3.5}$$

---

*) Если какое-либо значение встретится в выборке k раз, то этому значению соответствует вероятность k/n.

Несколько сложнее вычисляются математическое ожидание и дисперсия центральных выборочных моментов. Ограничимся вычислением $\mathsf{M}m_2$ и $\mathsf{D}m_2$. Положив $y_k = x_k - \mathsf{M}x_k$, $k = 1, \ldots, n$, запишем вторую формулу (3.1) при $\nu = 2$ в виде

$$m_2 = \frac{1}{n}\sum_{k=1}^{n}(y_k - \bar{y})^2 = \frac{1}{n}\sum_{k=1}^{n} y_k^2 - \bar{y}^2, \tag{3.6}$$

где $\bar{y} = (y_1 + \ldots + y_n)/n$. Поскольку $\mathsf{M}y_k = 0$, $\mathsf{M}y_k^2 = \mu_2$ и $\mathsf{M}y_k y_l = \mathsf{M}y_k \cdot \mathsf{M}y_l = 0$ $(k \neq l)$, то

$$\mathsf{M}\bar{y}^2 = \frac{1}{n^2}\sum_{k,\,l=1}^{n} \mathsf{M}y_k y_l = \frac{\mu_2}{n}.$$

Отсюда и из (3.6) находим

$$\mathsf{M}m_2 = \frac{n-1}{n}\mu_2. \tag{3.7}$$

В формуле (3.5) математическое ожидание a_ν совпало с соответствующим теоретическим моментом. В (3.7) правая часть смещена относительно μ_2. Если вместо m_2 ввести величину

$$s^2 = \frac{n}{n-1}\, m_2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2, \tag{3.8}$$

то для нее получим

$$\mathsf{M}s^2 = \mu_2. \tag{3.9}$$

Перейдем к вычислению $\mathsf{D}m_2$. Из (3.6) находим

$$m_2^2 = \frac{1}{n^2}\left(\sum_{k=1}^{n} y_k^2\right)^2 - \frac{2}{n}\bar{y}^2 \sum_{k=1}^{n} y_k^2 + \bar{y}^4. \tag{3.10}$$

Так как случайные величины y_k, $k = 1, \ldots, n$, независимы и $\mathsf{M}y_k = 0$, то в правой части равенства

$$\mathsf{M}\bar{y}^4 = \frac{1}{n^4}\sum_{l_1,l_2,l_3,l_4=1}^{n} \mathsf{M}y_{l_1}y_{l_2}y_{l_3}y_{l_4}$$

отличны от нуля только n слагаемых $\mathsf{M}y_l^4 = \mu_4$ и $C_n^2 C_4^2 = 3n(n-1)$ слагаемых $\mathsf{M}y_l^2 y_k^2 = \mu_2^2$, $l \neq k$. Поэтому

$\mathsf{M}y^4 = (\mu_4 + 3(n-1)\mu_2^2)/n^3$. Аналогично находим

$$\mathsf{M}\left[\left(\sum_{k=1}^{n} y_k^2\right)^2 \Big/ n\right] = n^2 \mathsf{M}\left[\frac{1}{n}\bar{y}^2\left(\sum_{k=1}^{n} y_k^2\right)\right] = \mu_4 + (n-1)\mu_2^2.$$

Отсюда и из (3.10) по формуле $\mathsf{D}m_2 = \mathsf{M}m_2^2 - (\mathsf{M}m_2)^2$ получим

$$\mathsf{D}m_2 = \frac{\mu_4 - \mu_2^2}{n} - \frac{2(\mu_4 - 2\mu_2^2)}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}. \tag{3.11}$$

Можно показать (соответствующие вычисления см. в книге [10], § 27.5), что для любого ν при $n \to \infty$

$$\mathsf{M}m_\nu = \mu_\nu + O\left(\frac{1}{n}\right), \ \mathsf{M}(m_\nu - \mu_\nu)^{2k} = O\left(\frac{1}{n^k}\right). \tag{3.12}$$

Вторая из этих формул верна, если конечен момент $\mu_{2k\nu}$. Из неравенства Чебышева и формул (3.4), (3.9) и (3.11) следует, что для любого $\varepsilon > 0$ при $n \to \infty$

$$\mathsf{P}\{|\bar{x} - \alpha_1| < \varepsilon\} \to 1, \ \mathsf{P}\{|s^2 - \mu_2| < \varepsilon\} \to 1, \tag{3.13}$$

где величина s^2 определена формулой (3.8). Аналогичные утверждения верны и для a_ν, m_ν при любых $\nu \geqslant 2$.

3.2. Асимптотические распределения. Найдем сначала асимптотические при $n \to \infty$ распределения выборочных моментов a_ν. Величина $na_\nu = \sum_{k=1}^{n} x_k^\nu$ является суммой независимых одинаково распределенных случайных величин. Если конечен момент $\alpha_{2\nu} = \mathsf{M}\xi^{2\nu}$, то к этой сумме можно применить теорему 4.1 гл. 7. Так как $\mathsf{M}x_k^\nu = \alpha_\nu$, $\mathsf{D}x_k^\nu = \alpha_{2\nu} - \alpha_\nu^2$, то величина

$$(na_\nu - n\alpha_\nu)/\sqrt{(\alpha_{2\nu} - \alpha_\nu^2)\,n} = (a_\nu - \alpha_\nu) \Big/ \sqrt{\frac{\alpha_{2\nu} - \alpha_\nu^2}{n}}$$

асимптотически нормальна с параметрами (0, 1). Таким образом, справедлива следующая теорема.

Теорема 3.1. *Если конечен теоретический момент $\alpha_{2\nu}$, то при $n \to \infty$ выборочный момент a_ν асимптотически нормален с параметрами* $\left(\alpha_\nu, \dfrac{\alpha_{2\nu} - \alpha_\nu^2}{n}\right)$.

Несколько сложнее доказывается асимптотическая нормальность центральных выборочных моментов m_ν. Ограничимся лишь случаем $\nu = 2$.

Теорема 3.2. *Если конечен теоретический момент* μ_4, *то при* $n \to \infty$ *выборочная дисперсия* m_2 *асимптотически нормальна с параметрами* $\left(\mu_2, \frac{\mu_4 - \mu_2^2}{n}\right)$.

Доказательство. Используя формулу (3.6), получим

$$\zeta_n = \sqrt{n}\,(m_2 - \mu_2) = \xi_n + \eta_n,$$

где

$$\xi_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (y_k^2 - \mathsf{M} y_k^2), \quad \eta_n = -\bar{y}^2 \sqrt{n}, \qquad (3.14)$$

y_k^2, $k = 1, \ldots, n$,— независимые одинаково распределенные величины. Так как $\mathsf{M}\bar{y}^2 = \mu_2/n$ и, следовательно, $\mathsf{M}\,|\,\eta_n\,| = \mu_2/\sqrt{n}$, то, согласно лемме 5.1 гл. 6, для любого $\varepsilon > 0$

$$\mathsf{P}\{|\,\eta_n\,| > \varepsilon\} \leqslant \frac{\mathsf{M}\,|\,\eta_n\,|}{\varepsilon} = \frac{\mu_2}{\varepsilon\sqrt{n}} \to 0$$

при $n \to \infty$. Отсюда и из теоремы 2.7 гл. 7 вытекает, что предельное распределение ζ_n совпадает с предельным распределением ξ_n. В соответствии с теоремой 4.1 гл. 7 случайная величина ξ_n асимптотически нормальна с параметрами $(0, \mu_4 - \mu_2^2)$. Теорема доказана.

Иногда требуется определить точное или хотя бы приближенное распределение некоторой функции от выборочных моментов. В качестве примера, иллюстрирующего исследование предельного распределения в таких задачах, найдем предельное распределение функции от выборочного среднего $\bar{x}$ и выборочной дисперсии m_2.

Теорема 3.3. *Пусть* $\eta_n = g(\bar{x}, m_2)$, *функция* $g(z_1, z_2)$ *определена в окрестности точки* (α_1, μ_2) *и имеет в этой окрестности непрерывные частные производные второго порядка. Если конечен момент* μ_4 *и* $\sigma^2 > 0$, *то величина* η_n *при* $n \to \infty$ *асимптотически нормальна с параметрами* $(g_0, \sigma^2/n)$, *где* $g_0 = g(\alpha_1, \mu_2)$, $\sigma^2 = g_{11}\mu_2 + 2g_{12}\mu_3 + g_{22}(\mu_4 - \mu_2^2)$, *а* $g_{ij} = \frac{\partial g}{\partial z_i} \frac{\partial g}{\partial z_j}$ $(i, j = 1, 2)$ *в точке* (α_1, μ_2).

Доказательство. Покажем, что из условий теоремы 3.3 следуют условия теоремы 6.1 гл. 7 при $r = 2$, $\xi_{n1} = \bar{x}$, $\xi_{n2} = m_2$. Действительно, из формул (3.4), (3.7), (3.11) при $n \to \infty$ получаем

$$a_{n1} = \mathsf{M}\xi_{n1} = \alpha_1, \quad a_{n2} = \mathsf{M}\xi_{n2} = \mu_2 + O\left(\frac{1}{n}\right),$$
$$\mathsf{D}\xi_{n1} = \frac{\mu_2}{n}, \quad \mathsf{D}\xi_{n2} = \frac{\mu_4 - \mu_2^2}{n} + O\left(\frac{1}{n^2}\right). \tag{3.15}$$

Кроме того, воспользовавшись формулой (3.6), найдем

$$\operatorname{cov}(\xi_{n1}, \xi_{n2}) = \operatorname{cov}(\bar{x}, m_2) = \mathsf{M}\left[\bar{y}\left(\frac{1}{n}\sum_{k=1}^{n} y_k^2 - \bar{y}^2 - \mu_2\right)\right] =$$
$$= \frac{1}{n}\sum_{k=1}^{n} \mathsf{M}\bar{y}y_k^2 - \mathsf{M}\bar{y}^3 = \frac{\mu_3}{n} + O\left(\frac{1}{n^2}\right).$$

Отсюда и из (3.15) следует выполнение условий теоремы 6.1 гл. 7. Теорема доказана.

Отметим, что из условий доказанной теоремы еще не следует, что параметры предельного нормального распределения являются главными членами асимптотических формул для $\mathsf{M}\eta_n$, $\mathsf{D}\eta_n$ (см. также [10], § 27.7, стр. 388).

3.3. Точные выборочные распределения. Допредельные распределения различных выборочных характеристик могут заметно отличаться от предельных. Отыскание точных распределений является обычной задачей нахождения распределения функции от случайных величин (см. § 6 гл. 5). Однако найти точные выборочные распределения в удобной для приложений форме в общем случае не удается.

Как уже отмечалось в § 2, величины x_k, $k = 1, \ldots, n$, часто оказываются нормально распределенными. Здесь мы ограничимся рассмотрением выборок $x_1, x_2, \ldots, x_n$ с нормально распределенными x_k ($k = 1, \ldots, n$). Для нормально распределенных x_k распределения многих выборочных характеристик выражаются через небольшое число распределений, для которых имеются таблицы. Дадим еще несколько определений часто встречающихся в статистических исследованиях распределений.

Пусть $\xi_0, \xi_1, \ldots, \xi_n$ — независимые нормально распределенные случайные величины с параметрами (0, 1). Распределение величины

$$\chi_n^2 = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2 \tag{3.16}$$

называется *распределением* χ^2 (*хи-квадрат*) *с* n *степеня-*

ми свободы. Распределение величины

$$\tau_n = \xi_0 \Big/ \sqrt{\frac{1}{n}\chi_n^2} \tag{3.17}$$

называется *распределением Стьюдента с n степенями свободы*. Обозначим $\chi_{n_1}^2$ и $\chi_{n_2}^2$, независимые случайные величины, имеющие распределение χ^2 с n_1 и n_2 степенями свободы соответственно. Распределение величины

$$F_{n_1, n_2} = \frac{\chi_{n_1}^2/n_1}{\chi_{n_2}^2/n_2} \tag{3.18}$$

называется *F-распределением* (или *распределением Фишера*) *с n_1 и n_2 степенями свободы*.

Плотности распределения этих величин могут быть найдены в явном виде (см. [10] и [16]); формулы плотностей распределений приведены перед таблицами 4, 5, 6.

Найдем теперь совместное распределение выборочного среднего $\bar{x}$ и выборочной дисперсии m_2. Следующая теорема о совместном распределении $\bar{x}$ и m_2 была доказана Р. Фишером.

Теорема 3.4. *Если элементы выборки x_k, $k = 1, \ldots, n$, независимы и распределены нормально с параметрами (a, σ^2), то $\bar{x}$ и m_2 независимы, причем $\bar{x}$ распределено нормально с параметрами $(a, \sigma^2/n)$, а nm_2/σ^2 имеет распределение χ^2 с $n-1$ степенями свободы.*

Доказательство. Обратимся к соотношению (3.6). В предположениях доказываемой теоремы $y_1, y_2, \ldots, y_n$ независимы и каждая величина y_k распределена нормально с параметрами $(0, 1)$. Введем новые случайные величины $\xi = (\xi_1, \ldots, \xi_n)' = \frac{1}{\sigma} Cy$, где $y = (y_1, \ldots y_n)'$, $c_{1k} = \frac{1}{\sqrt{n}}$, $k = 1, 2, \ldots, n$, а остальные строки c_{lk}, $k = 1, \ldots, n$, $l \geqslant 2$, выбраны так, что матрица C ортогональна. При таком выборе C

$$\xi_1 = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} y_k = \frac{\sqrt{n}}{\sigma}\bar{y}, \quad \mathsf{M}\xi_k = 0, \quad \mathsf{D}\xi_k = 1$$

$$(k = 1, \ldots, n),$$

$$\operatorname{cov}(\xi_k, \xi_l) = \\ = \frac{1}{\sigma^2} \sum_{i,j=1}^{n} c_{ki} c_{lj} \mathsf{M} y_i y_j = \frac{1}{\sigma^2} \sum_{i,j=1}^{n} c_{ki} c_{lj} = 0 \quad (k \neq l)$$

Отсюда следует, что $\xi_1, \ldots, \xi_n$ независимы и каждая величина распределена нормально с параметрами (0, 1). Выразим теперь m_2 через новые случайные величины. Поскольку сумма квадратов инвариантна относительно ортогональных преобразований, то

$$\frac{nm_2}{\sigma^2}=\frac{1}{\sigma^2}\sum_{k=1}^{n}y_k^2-\left(\frac{\bar{y}}{\sigma}\right)^2=\sum_{k=1}^{n}\xi_k^2-\xi_1^2=\xi_2^2+\ldots+\xi_n^2. \quad (3.19)$$

Отсюда и из независимости $\xi_1, \ldots, \xi_n$ следует независимость $\frac{nm_2}{\sigma^2}$ и $\bar{x}=\frac{1}{\sqrt{n}}\sigma\xi_1+a$. Нормальная распределенность $\bar{x}$ следует из нормальной распределенности ξ_1. Из (3.16) и (3.19) вытекает, что nm_2/σ^2 имеет распределение χ^2 с $n-1$ степенями свободы. Теорема доказана.

При помощи теоремы 3.4 можно найти распределение еще одной часто используемой величины. Величина $\xi_1 = (\bar{x}-a)/(\sigma/\sqrt{n})$ имеет нормальное распределение. Кроме того, ξ_1 и nm_2/σ^2 независимы. Следовательно, по определению (3.17) величина

$$\tau_{n-1}=\frac{\xi_1}{\sqrt{\dfrac{\xi_2^2+\ldots+\xi_n^2}{n-1}}}=\frac{(\bar{x}-a)/(\sigma/\sqrt{n}\,)}{\sqrt{nm_2/\sigma^2\,(n-1)}}=\frac{\bar{x}-a}{\sqrt{m^2}}\sqrt{n-1}$$

имеет распределение Стьюдента с $n-1$ степенями свободы. Таким образом, доказано следующее утверждение.

Теорема 3.5. *Если элементы выборки $x_1, \ldots, x_n$ независимы и каждая из этих величин распределена нормально с параметрами (a, σ^2), то величина $\frac{\bar{x}-a}{\sqrt{m_2}}\sqrt{n-1}$ имеет распределение Стьюдента с $n-1$ степенями свободы.*

Распределение Фишера будет использоваться в § 8. Там же нам понадобится следующая теорема, касающаяся распределения величины, пропорциональной отношению двух выборочных дисперсий, вычисленных по двум независимым выборкам.

Теорема 3.6. *Пусть $x_1, \ldots, x_{n_1}$ и $x'_1, \ldots, x'_{n_2}$ — две независимые выборки и любая из этих n_1+n_2 независимых величин имеет нормальное распределение с параметрами (a, σ^2). Тогда величина $\frac{m_2}{m'_2}\cdot\frac{n_1}{n_2}\cdot\frac{n_2-1}{n_1-1}$, где m_2 — выборочная дисперсия выборки $x_1, \ldots, x_{n_1}$, а m'_2 — выбо-*

речная дисперсия выборки $x'_1, \ldots, x'_{n_2}$, имеет F-распределение с $n_1 - 1$ и $n_2 - 1$ степенями свободы.

Утверждение этой теоремы непосредственно следует из теоремы 3.4 и определения величины (3.18).

§ 4. Точечные оценки

4.1. Определения и примеры. Предположим, что функция распределения, соответствующая выборке $x_1, x_2, \ldots, x_n$, зависит от неизвестного параметра θ:

$$\mathsf{P}\,(x_k < x) = F_\xi\,(x;\,\theta).$$

Оценкой θ_n^* параметра θ называется произвольная функция $\theta_n^* = \theta_n^*\,(x_1, x_2, \ldots, x_n)$. Таким образом, θ_n^* является случайной величиной. Естественно потребовать, чтобы значения оценки в большинстве опытов были близки к значению оцениваемого параметра. Будем называть оценку θ_n^* *несмещенной* оценкой параметра θ, если при любом n

$$\mathsf{M}\theta_n^* = \theta. \tag{4.1}$$

Оценка θ_n^* называется *состоятельной*, если для любого $\varepsilon > 0$ при $n \to \infty$

$$\mathsf{P}\,\{|\,\theta_n^* - \theta\,| < \varepsilon\} \to 1. \tag{4.2}$$

Условиям (4.1) и (4.2) может удовлетворять несколько разных оценок одного и того же параметра. Тогда из них естественно выбрать ту, у которой дисперсия наименьшая. Для широкого класса распределений можно указать точную нижнюю грань дисперсий оценок одного и того же параметра (см. п. 4.3).

Рассмотрим два примера.

Пример 1. Пусть $\theta = \alpha_1 = \mathsf{M}x_k$. Положим $\theta_n^* = \theta_n^*\,(x_1, \ldots, x_n) = \bar{x}$, $\tilde{\theta}_n^* = x_1$, где $\bar{x}$ определено формулой (3.2). Из определения несмещенности и из формулы (3.4) следует, что обе оценки θ_n^* и $\tilde{\theta}_n^*$ являются несмещенными. Оценка $\tilde{\theta}_n^*$, очевидно, не является состоятельной. Состоятельность оценки $\bar{x}$ была доказана в § 3 (см. (3.13)).

Пример 2. Пусть $\theta = \mu_2 = \mathsf{M}\,(x_k - Mx_k)^2$. Положим $\theta_n^* = s^2$, где s^2 определено формулой (3.8). Из соотношений (3.9) и (3.13) следует несмещенность и состоятельность оценки θ_n^*. Выборочная дисперсия m_2 не обладает свойством несмещенности (см. (3.7)).

4.2. Достаточные статистики. Может оказаться, что вся полезная информация о неизвестных параметрах содержится в небольшом числе функций, зависящих от выборки $(x_1, \ldots, x_n)$. В этом случае, особенно при больших объемах выборки, задача построения оценок может значительно упроститься. Различные функции от выборки обычно называют *статистиками*. Обозначим $\mathsf{P}_\theta(A)$ закон распределения вероятностей случайного вектора $(x_1, x_2, \ldots, x_n)$ с $\mathsf{P}(x_k < x) = F(x, \theta)$. Статистика $X = X(x_1, \ldots$ $\ldots, x_n)$, скалярная или векторная, называется *достаточной* для параметра θ, если для любого события A условная вероятность $\mathsf{P}_\theta(A \mid X)$ не зависит от θ. Определенная таким образом достаточная статистика X подытоживает все существенные сведения о параметре θ. Действительно, вероятность любого события, которое может произойти при фиксированном X, не зависит от θ, и, следовательно, оно не может содержать дополнительной информации о неизвестном параметре θ.

Очевидно, что $X = (x_1, \ldots, x_n)$, т. е. сама выборка является достаточной статистикой. (Событие A является подмножеством множества значений $(x_1, \ldots, x_n)$, и, следовательно, $\mathsf{P}_\theta(A \mid X)$ при любом θ есть либо 0, либо 1.) Больший интерес представляют случаи, когда размерность X меньше n. Приведем теорему, позволяющую практически находить достаточные статистики.

Теорема 4.1. *Пусть $p_\theta(u)$, $u = (u_1, \ldots, u_n)$,— совместная плотность распределения случайной выборки $x_1, x_2, \ldots, x_n$, где $x_k = x_k(u) = u_k$, $k = 1, \ldots, n$. Статистика $X = X(u)$ достаточна для параметра θ тогда и только тогда, когда*

$$p_\theta(u) = g_\theta(X(u)) \cdot h(u), \tag{4.3}$$

где g_θ, h — неотрицательные функции и h не зависит от θ. Для дискретного случайного вектора $x = (x_1, \ldots, x_n)$ утверждение теоремы сохранится, если $p_\theta(u)$ понимать как вероятность события $\{x = u\}$.

Доказательство приведем только для дискретного вектора x. Пусть $\mathscr{U} = \{u^{(1)}, u^{(2)}, \ldots, u^{(k)}, \ldots\}$ — множество значений x; $X = X(u)$ $(u \in \mathscr{U})$ — некоторая достаточная статистика. Выберем какое-нибудь $u^{(k)} \in \mathscr{U}$ и положим $v = X(u^{(k)})$. Тогда

$$\mathsf{P}_\theta(x = u^{(k)}) = \mathsf{P}\{x = u^{(k)}, X(u) = v\} = \\ = \mathsf{P}_\theta(X = v) \cdot \mathsf{P}(x = u^{(k)} \mid X = v).$$

Первый множитель в правой части зависит только от v и неизвестного параметра θ, а второй множитель по определению достаточной статистики (в качестве A нужно взять событие $\{x = u^{(k)}\}$) не зависит от θ. Таким образом, получили (4.3). Пусть теперь выполнено (4.3) и X — некоторая статистика. Покажем, что X — достаточная статистика. Найдем условную вероятность $\mathsf{P}_\theta(A \mid X = v)$. Используя (4.3), получим

$$\mathsf{P}(X = v) = \sum_{u^{(k)} \in \{X = v\}} p_\theta(u^{(k)}) = \sum_{u^{(k)} \in \{X = v\}} g_\theta(X(u^{(k)}))\, h(u^{(k)}) = \\ = g_\theta(v) \sum_{u^{(k)} \in \{X = v\}} h(u_k),$$

$$\mathsf{P}(A \cap \{X = v\}) = \sum_{u^{(k)} \in B} g_\theta(X(u^{(k)}))\, h(u^{(k)}) = g_\theta(v) \sum_{u^{(k)} \in B} h(u^{(k)}),$$

где $B = A \cap \{X = v\}$. Отсюда находим, что условная вероятность

$$\mathsf{P}(A \mid X = v) = \left(\sum_{u^{(k)} \in \{X = v\}} h(u^{(k)}) \right) \Big/ \left(\sum_{u^{(k)} \in B} h(u^{(k)}) \right)$$

и, следовательно, не зависит от θ. Теорема для дискретного случая доказана.

Рассмотрим два примера.

Пример 3. Пусть $x_1, x_2, \ldots, x_n$ — независимая выборка с $\mathsf{P}(x_k = v) = \theta^v (1-\theta)^{1-v}$ $(v = 0;\ 1)$, $k = 1, \ldots, n$. Тогда

$$p_\theta(u) = \mathsf{P}\{x_1 = u_1, \ldots, x_n = u_n\} =$$

$$= \prod_{k=1}^{n} \theta^{u_k} (1-\theta)^{1-u_k} = \theta^{\sum\limits_{k=1}^{n} u_k} (1-\theta)^{n - \sum\limits_{k=1}^{n} u_k}.$$

Это равенство для статистики $\bar{x} = \frac{1}{n} \sum\limits_{k=1}^{n} x_k$ имеет вид (4.3) с $h(u) \equiv 1$, $g_\theta(v) = \theta^{nv} (1-\theta)^{n-nv}$. Рассмотренная в данном примере выборка определяет n испытаний схемы Бернулли с вероятностью появления исхода 1 в каждом испытании, равной θ. Таким образом, показано, что достаточной статистикой для вероятности успеха в схеме Бернулли является частота $\bar{x}$.

Пример 4. Пусть $x_1, x_2, \ldots, x_n$ — независимая выборка, в которой x_k $(k = 1, \ldots, n)$ имеет нормальное распределение с параметрами (a, σ^2). Будем искать достаточную статистику для двумерного параметра $\theta = (a, \sigma^2)$. Плотность совместного распределения выборки равна

$$p_\theta(u) = \frac{1}{(\sqrt{2\pi}\,\sigma)^n} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{k=1}^{n} (u_k - a)^2 \right\} =$$

$$= (\sqrt{2\pi}\,\sigma)^{-n} \exp\left\{ -\frac{na^2}{2\sigma^2} - \frac{1}{2\sigma^2} \sum_{k=1}^{n} u_k^2 + \frac{a}{\sigma^2} \sum_{k=1}^{n} u_k \right\}.$$

Отсюда, так же как в предыдущем примере, нетрудно получить, что достаточной статистикой является двумерная статистика $\left\{ \sum\limits_{k=1}^{n} x_k,\ \sum\limits_{k=1}^{n} x_k^2 \right\}$. Так как обе функции этой статистики можно выразить через $\bar{x}$ и s^2 без использования a и σ^2, то $(\bar{x}, s^2)$ тоже достаточная статистика.

Приведем теперь частный случай теоремы Рао — Блэкуэлла — Колмогорова. Эта теорема позволяет по данной оценке найти оценку с меньшей дисперсией, если известна достаточная статистика.

Теорема 4.2. *Пусть* $x_1, \ldots, x_n$ — *независимая выборка с* $\mathsf{P}\{x_k < x\} = F_\xi(x, \theta)$, θ_1^* — *несмещенная оценка параметра*

θ и X — достаточная статистика. Тогда дисперсии оценок θ_1^* и $\theta^* = \mathsf{M}(\theta_1^* \mid X)$ удовлетворяют неравенству $\mathsf{D}\theta_1^* \geqslant \mathsf{D}\theta^*$.

Доказательство проведем для случая, когда θ_1^* и X имеют дискретное распределение: $\mathsf{P}\{\theta_1^* = u_k,\ X = v_l\} = p_{k,l}$, $k, l = 1, 2, \ldots$ Из формулы полного математического ожидания (6.6.3) следует, что $\mathsf{M}\theta^* = \mathsf{M}[\mathsf{M}(\theta_1^* \mid X)] = \mathsf{M}\theta_1^* = \theta$. Поэтому для доказательства нужного неравенства достаточно показать, что

$$\mathsf{M}(\theta^*)^2 = \mathsf{M}\{[\mathsf{M}(\theta_1^* \mid X)]^2\} \leqslant \mathsf{M}(\theta_1^*)^2. \tag{4.4}$$

Пусть $p_{\cdot l} = \sum\limits_k p_{k,l}$. Тогда $\mathsf{M}(\theta_1^*)^2 = \sum\limits_{k,l} u_k^2 p_{k,l}$ и $\mathsf{M}\{[\mathsf{M}(\theta_1^* | X)]^2\} = \sum\limits_l \left(\sum\limits_k u_k \frac{p_{k,l}}{p_{\cdot l}}\right)^2 p_{\cdot l}$. По неравенству Коши — Буняковского имеем

$$\left(\sum_k u_k \frac{p_{k,l}}{p_{\cdot l}}\right)^2 \leqslant \left(\sum_k u_k^2 \frac{p_{k,l}}{p_{\cdot l}}\right)\left(\sum_k \frac{p_{k,l}}{p_{\cdot l}}\right) = \sum_k u_k^2 \frac{p_{k,l}}{p_{\cdot l}}.$$

Умножив обе части этого неравенства на $p_{\cdot l}$ и просуммировав по всем l, получим (4.4). Теорема доказана.

4.3. Неравенство Рао—Крамера. Пусть $x = (x_1, \ldots, x_n)$ — независимая выборка, в которой x_k ($k = 1, \ldots, n$) имеет плотность распределения $p(v, \theta)$, где θ — неизвестный параметр, а v — переменная плотности распределения. Положим $G = \{v\colon p(v, \theta) > 0\}$ и

$$G_n = \{u = (u_1, \ldots, u_n)\colon u_i \in G,\ i = 1, \ldots, n\}.$$

Совместная плотность распределения $p_\theta(u) = p(u_1, \theta) \ldots p(u_n, \theta)$ выборки $(x_1, \ldots, x_n)$ отлична от 0 на множестве G_n. Пусть $\theta^* = \theta^*(x_1, x_2, \ldots, x_n)$ — несмещенная оценка параметра θ. Тогда $\mathsf{M}\theta^* = \theta$. Это равенство можно записать в виде

$$\theta = \int\limits_{G_n} \theta^*(u)\, p_\theta(u)\, du, \quad u = (u_1, \ldots, u_n). \tag{4.5}$$

Оценим $\mathsf{D}\theta^*$ в следующих *условиях регулярности*:

1°. Множество $G = \{v\colon p(v, \theta) > \theta\}$ не зависит от θ.

2°. Выражение в правой части (4.5) и выражение $\int\limits_{-\infty}^{\infty} p(u, \theta)\, du$ можно дифференцировать по θ под знаком интеграла.

3°. Выражение

$$I(\theta)=\int\limits_G\left(\frac{\partial \ln p(v,\theta)}{\partial\theta}\right)^2 p(v,\theta)\,dv=\mathsf{M}\left[\frac{\partial \ln p(x_k,\theta)}{\partial\theta}\right]^2 \tag{4.6}$$

положительно.

Так как $p_\theta(u)$ — плотность распределения, то

$$1=\int\limits_{G_n} p_\theta(u)\,du,\quad u=(u_1,\ldots,u_n). \tag{4.7}$$

Продифференцировав (4.5) и (4.7) по θ, найдем

$$1=\int\limits_{G_n}\theta^*(u)\frac{\partial p_\theta(u)}{\partial\theta}\,du,\quad 0=\int\limits_{G_n}\frac{\partial p_\theta(u)}{\partial\theta}\,du. \tag{4.8}$$

Умножим теперь второе равенство в (4.8) на θ и вычтем из первого. В результате выполнения указанных преобразований получим

$$1=\int\limits_{G_n}(\theta^*(u)-\theta)\frac{\partial p_\theta(u)}{\partial\theta}\,du. \tag{4.9}$$

Так как на множестве G плотность $p_\theta(u)>0$, то

$$\frac{\partial p_\theta(u)}{\partial\theta}=\frac{\partial \ln p_\theta(u)}{\partial\theta}\,p_\theta(u).$$

Подставляя это выражение в (4.9) и используя неравенство Коши — Буняковского, находим

$$1=\left(\int\limits_{G_n}(\theta_n^*(u)-\theta)\frac{\partial \ln p_\theta(u)}{\partial\theta}\,p_\theta(u)\,du\right)^2\leqslant$$

$$\leqslant\int\limits_{G_n}(\theta^*-\theta)^2 p_\theta(u)\,du\cdot\int\limits_{G_n}\left(\frac{\partial \ln p_\theta(u)}{\partial\theta}\right)^2 p_\theta(u)\,du,$$

или

$$1\leqslant \mathsf{D}\theta^*\cdot\mathsf{M}\left[\left(\frac{\partial \ln p_\theta(x)}{\partial\theta}\right)^2\right]. \tag{4.10}$$

Выразим второй сомножитель в (4.10) через величину $I(\theta)$, определенную (4.6). Из равенств (4.8) при $n=1$ следует, что

$$\mathsf{M}\left[\frac{\partial \ln p(x_k,\theta)}{\partial\theta}\right]=0. \tag{4.11}$$

Так как $\ln p_\theta(x) = \sum_{k=1}^{n} \ln p(x_k, \theta)$, то

$$\left(\frac{\partial \ln p_\theta(x)}{\partial\theta}\right)^2 = \sum_{k,l=1}^{n} \frac{\partial \ln p(x_k\theta)}{\partial\theta} \frac{\partial \ln\cdot p(x_l, \theta)}{\partial\theta}.$$

Отсюда, учитывая независимость сомножителей при $k \neq l$ и равенство (4.11), находим

$$\mathsf{M}\left(\frac{\partial \ln p_\theta(x)}{\partial\theta}\right)^2 = n\mathsf{M}\left(\frac{\partial \ln p(x_1, \theta)}{\partial\theta}\right)^2 = nI(\theta).$$

Подставляя это выражение в (4.10), получим окончательный результат:

$$\mathsf{D}\theta_n^* \geqslant \frac{1}{nI(\theta)}, \tag{4.12}$$

где $I(\theta)$ определено формулой (4.6). Это неравенство называют *неравенством Рао — Крамера*. Аналогичное доказательство может быть проведено и для дискретных выборок. В этом случае нужно использовать представление $I(\theta)$ в форме математического ожидания *).

Назовем *эффективностью* несмещенной оценки θ_n^*, удовлетворяющей условиям регулярности 1°—3°, величину

$$e(\theta_n^*) = (nI(\theta)\,\mathsf{D}\theta_n^*)^{-1}. \tag{4.13}$$

Из неравенства (4.12) следует, что $0 \leqslant e(\theta_n^*) \leqslant 1$. Оценку θ_n^* называют *эффективной*, если $e(\theta_n^*) = 1$. Оценка θ_n^* параметра θ может оказаться асимптотически нормальной с параметрами $(\theta, \sigma^2/n)$ при $n \to \infty$ даже в тех случаях, когда $\mathsf{D}\theta_n^*$ не существует. *Асимптотической эффективностью* назовем величину $e_0(\theta_n^*) = (\sigma^2 I(\theta))^{-1}$. Оценку θ_n^* называют *асимптотически эффективной*, если $e_0(\theta_n^*) = 1$.

Вычислим эффективность нескольких оценок.

П р и м е р 5. Пусть $x_1, x_2, \ldots, x_n$ — независимая выборка, соответствующая нормальному распределению с параметрами (a, σ^2). В качестве оценок a и σ^2 возьмем выборочное среднее $\bar{x}$ и величину s^2, определенную формулой (3.8). Обе оценки являются несмещенными и состоятель-

*) Можно показать, что неравенство (4.12) обращается в равенство, если оценка θ_n^* является достаточной статистикой.

ными (см. примеры 1 и 2):

$$\mathsf{M}\bar{x} = a, \quad \mathsf{M}s^2 = \sigma^2,$$

$$\mathsf{D}\bar{x} = \frac{\sigma^2}{n}, \quad \mathsf{D}s^2 = \frac{\mu_4 - \mu_2^2}{n} + O\left(\frac{1}{n^2}\right).$$

В примере 5 § 2 гл. 7 были найдены центральные моменты нормально распределенной случайной величины; в частности, $\mathsf{M}(x_k - a)^4 = 3\sigma^4$. Подставляя это значение в формулу (3.11), получим точную формулу для $\mathsf{D}m_2 = 2(n-1)\sigma^4/n^2$. Отсюда

$$\mathsf{D}s^2 = \mathsf{D}\left[\frac{n}{n-1} m_2\right] = 2\sigma^4/(n-1).$$

Для вычисления эффективности нужно еще найти величину $I(\theta)$ для $\theta = a$ и $\theta = \sigma^2$. По формуле (4.6), полагая

$$p(u, \theta) = p(u, a, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\},$$

находим

$$I(a) = \int_{-\infty}^{\infty} \left(\frac{x-a}{\sigma^2}\right)^2 p(u, a, \sigma^2)\, du = \frac{1}{\sigma^2},$$

$$I(\sigma^2) = \int_{-\infty}^{\infty} \left(\frac{(x-a)^2}{2\sigma^4} - \frac{1}{2\sigma^2}\right)^2 p(u, a, \sigma^2)\, du = \frac{1}{2\sigma^4}.$$

По формуле (4.13) получаем

$$e(\bar{x}) = 1, \quad e(s^2) = \frac{n-1}{n} < 1.$$

Итак, оценка $\bar{x}$ является эффективной, а оценка s^2 асимптотически эффективна.

Интересно отметить, что при нарушении условий регулярности (см. задачу 12; в условиях этой задачи нарушено условие 1°) дисперсия оценки при $n \to \infty$ может быть меньше $1/nI(\theta)$ даже по порядку.

4.4. Метод наибольшего правдоподобия для нахождения оценок параметров. Метод моментов. Если функция распределения $F(x, \theta) = \mathsf{P}(x_k < x)$, $k = 1, \ldots, n$, имеет плотность $p(x, \theta)$, то функция

$$L = L(x_1, \ldots, x_n, \theta) = p(x_1, \theta)\, p(x_2, \theta) \ldots p(x_n, \theta) \tag{4.14}$$

называется *функцией правдоподобия*. Для выборки, состоящей из дискретных величин x_k, $k = 1, \ldots, n$, с распределением $\mathsf{P}(x_k = x) = p_x(\theta)$ функция правдоподобия определяется равенством

$$L = L(x_1, \ldots, x_n, \theta) = p_{x_1}(\theta)\, p_{x_2}(\theta) \ldots p_{x_n}(\theta). \quad (4.15)$$

При фиксированных $x_1, \ldots, x_n$ функцию L будем рассматривать как функцию параметра θ. По методу наибольшего правдоподобия за оценку параметра θ принимается значение аргумента θ, при котором L имеет максимальное значение. При фиксированном θ функция правдоподобия L ((4.14) или (4.15)) задает совместное распределение выборки, если рассматривать $x_1, \ldots, x_n$ в (4.14) и (4.15) как независимые переменные. Таким образом, по методу наибольшего правдоподобия выбирается значение θ, при котором вероятность получения данных значений выборки максимальна. Поскольку $\ln L$ при фиксированных $x_1, \ldots, x_n$ достигает максимума при том же значении θ, что и L, то для нахождения оценки можно решать уравнение правдоподобия

$$\frac{\partial \ln L}{\partial \theta} = 0. \quad (4.16)$$

Решением (4.16) будем называть только корни, зависящие от $x_1, x_2, \ldots, x_n$. Каждое решение уравнения правдоподобия (4.16) будем называть *оценкой наибольшего правдоподобия для* θ.

При некоторых достаточно общих условиях уравнение (4.16) *имеет решение* θ_n^*, *являющееся состоятельной оценкой параметра* θ; θ_n^* *при* $n \to \infty$ *является асимптотически нормальной и асимптотически эффективной оценкой* θ. Условия, при которых доказывается это утверждение (см. [10], гл. 33, стр. 544), заключаются в существовании нескольких производных по θ от $\ln L$ и в существовании моментов этих производных.

Если $\theta = (\theta_1, \ldots, \theta_s)$, то оценками наибольшего правдоподобия параметров $\theta_1, \ldots, \theta_s$ являются зависящие от $(x_1, \ldots, x_n)$ решения системы уравнений

$$\frac{\partial \ln L}{\partial \theta_k} = 0, \quad k = 1, 2, \ldots, s.$$

Пример 6. Пусть величины x_k, $k = 1, 2, \ldots, n$, имеют нормальное распределение. Неизвестными параметрами являются $a = \mathsf{M}x_k$, $b = \sigma^2 = \mathsf{D}x_k$. Найдем их оценки наибольшего правдоподобия. По формуле (4.14)

$$L = L(x_1, \ldots, x_n, a, b) = \left(\frac{1}{\sqrt{2\pi b}}\right)^n \exp\left(-\sum_{k=1}^{n} \frac{(x_k - a)^2}{2b}\right)$$

и

$$\ln L = -\frac{n}{2}(\ln 2\pi + \ln b) - \frac{1}{2b}\sum_{k=1}^{n}(x_k - a)^2.$$

Отсюда для оценок a^* и b^* получим систему

$$\begin{cases} \dfrac{\partial \ln L}{\partial a} = \dfrac{1}{b^*}\displaystyle\sum_{k=1}^{n}(x_k - a^*)^2 = 0, \\ \dfrac{\partial \ln L}{\partial b} = -\dfrac{n}{2b^*} + \dfrac{1}{2(b^*)^2}\displaystyle\sum_{k=1}^{n}(x_k - a^*)^2 = 0. \end{cases}$$

Из первого уравнения $a^* = \frac{1}{n}\sum_{k=1}^{n} x_k = \bar{x}$. Подставляя это значение во второе уравнение, найдем

$$b^* = m_2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2.$$

Асимптотическая нормальность и асимптотическая эффективность оценок $\bar{x}$, m_2 параметров a, σ^2 следуют из результатов п. 4.3 и теоремы 3.2.

Пример 7. Найдем оценку наибольшего правдоподобия для вероятности успеха p в схеме Бернулли. По формуле (4.15)

$$L = L(x_1, \ldots, x_n; p) = \prod_{k=1}^{n} p^{x_k}(1-p)^{1-x_k},$$

$x_k = 0, 1$ ($x_k = 1$, если в испытании k был успех, и $x_k = 0$ в противном случае), и

$$\ln L = \sum_{k=1}^{n}(x_k \ln p + (1 - x_k)\ln(1-p)).$$

Отсюда

$$\frac{\partial \ln L}{\partial p} = \sum_{k=1}^{n}\left(\frac{x_k}{p^*} - \frac{1-x_k}{1-p^*}\right) = \frac{\bar{x}n}{p^*} - \frac{n}{1-p^*} - \frac{n\bar{x}}{1-p^*} = 0$$

и $p^* = \bar{x}$. Так как для числа успехов μ_n имеем равенство $\mu_n = x_1 + x_2 + \ldots + x_n$, то $p^* = \mu_n/n$. Оценка p^* является асимптотически нормальной (следствие теоремы Муавра — Лапласа) и эффективной.

Другим методом, который иногда используется для оценки параметров, является метод моментов. Если функция распределения $F_\xi(x, \theta) = \mathsf{P}\{x_k < x\}$ зависит от s параметров $\theta = (\theta_1, \ldots, \theta_s)$, то теоретические моменты, например $\alpha_k = \mathsf{M}\xi^k$, являются функциями от параметров $(\theta_1, \ldots, \theta_s)$. Предположим, что систему уравнений $\alpha_k = \alpha_k(\theta_1, \ldots, \theta_s)$, $k = 1, \ldots, s$, можно разрешить относительно θ_i: $\theta_i = \theta_i(\alpha_1, \ldots, \alpha_s)$, $i = 1, \ldots, s$. Тогда в качестве оценки параметра θ_i можно взять величину $\theta_i^* = \theta_i(a_1, \ldots, a_s)$, где a_i — выборочные моменты, соответствующие α_i. При достаточно общих условиях выборочные моменты близки к теоретическим, функция $\theta_i(a_1, \ldots, a_s)$ от этих моментов асимптотически нормальна, а математическое ожидание предельного распределения равно $\theta_i = \theta_i(\alpha_1, \ldots, \alpha_s)$. В таком случае оценки θ_i^* оказываются состоятельными.

§ 5. Интервальные оценки

Точечные оценки дают приближенное значение неизвестного параметра. Если известен закон распределения оценки или ее дисперсия, то можно указать пределы, в которых с большой вероятностью находится неизвестное значение параметра.

Эти пределы легко выражаются через дисперсию оценки. Однако иногда дисперсия может зависеть от неизвестного параметра. В этом случае границы, в которых лежит неизвестный параметр, тоже зависят от значений этого параметра и, следовательно, пользоваться такими границами нельзя.

Рассмотрим выборку $x_1, x_2, \ldots, x_n$ с $\mathsf{P}(x_i < x) = F_\xi(x, \theta)$, где θ — неизвестный параметр. Предположим, что нам удалось найти функции $\underline{\theta} = \underline{\theta}(x_1, \ldots, x_n)$ и $\theta = \bar{\theta}(x_1, \ldots, x_n)$ такие, что $\underline{\theta} < \bar{\theta}$ при всех $x_1, \ldots, x_n$ и при любых значениях θ:

$$\mathsf{P}(\underline{\theta}(x_1, \ldots, x_n) < \theta < \bar{\theta}(x_1, \ldots, x_n)) = 1 - 2\alpha,$$

т. е. вероятность того, что случайный интервал $(\underline{\theta}, \bar{\theta})$ накроет неизвестный параметр θ, не зависит от параметра θ, В этом случае интервал $(\underline{\theta}, \overline{\theta})$ называют *доверительным интервалом* для неизвестного параметра θ, соответствующим *доверительной вероятности* $1 - 2\alpha$. В ряде случаев функции $\underline{\theta}$ и $\overline{\theta}$, обладающие указанными свойствами, можно найти.

Пусть имеется выборка $x_1, \ldots, x_n$, причем величины x_k распределены нормально с параметрами (a, σ^2), а параметр σ известен. Найдем доверительный интервал для a. Случайная величина $\bar{x} = (x_1 + \ldots + x_n)/n$ имеет нормальное распределение и $\mathsf{M}\bar{x} = a$, $\mathsf{D}\bar{x} = \sigma^2/n$. Следовательно, величина $\dfrac{\bar{x} - a}{\sigma/\sqrt{n}}$ распределена нормально с параметрами $(0, 1)$ и ее распределение не зависит от a. Отсюда

$$\mathsf{P}\left\{\left|\frac{\bar{x} - a}{\sigma/\sqrt{n}}\right| < u_\alpha\right\} = 1 - 2\alpha,$$

или

$$\mathsf{P}\left\{\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}} < a < \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}}\right\} = 1 - 2\alpha, \tag{5.1}$$

где u_α определяется как решение уравнения

$$\frac{1}{\sqrt{2\pi}} \int_{u_\alpha}^{\infty} e^{-\frac{x^2}{2}}\, dx = \alpha. \tag{5.2}$$

Таким образом, мы нашли доверительный интервал $\left(\bar{x} - u_\alpha \dfrac{\sigma}{\sqrt{n}},\ \bar{x} + u_\alpha \dfrac{\sigma}{\sqrt{n}}\right)$ для параметра a.

Более естественной является ситуация, когда оба параметра, a и σ, неизвестны. В этом случае воспользуемся теоремой 3.5, согласно которой величина $\tau_{n-1} = \dfrac{\bar{x} - a}{\sqrt{m_2}}\sqrt{n-1}$ имеет распределение Стьюдента с $n - 1$ степенями свободы. Отсюда, определяя $t_{\alpha, n-1}$ как решение уравнения

$$\mathsf{P}\{|\tau_{n-1}| < t_{\alpha, n-1}\} = 1 - 2\alpha, \tag{5.3}$$

получим, что

$$\mathsf{P}\left\{\bar{x} - t_{\alpha, n-1}\sqrt{\frac{m_2}{n-1}} < a < \bar{x} + t_{\alpha, n-1}\sqrt{\frac{m_2}{n-1}}\right\} = 1 - 2\alpha. \tag{5.4}$$

Найдем теперь доверительные интервалы для параметра σ^2. При известном a воспользуемся тем, что величина $\chi_n^2 =$

$= S^2/\sigma^2$, где $S^2 = \sum_{k=1}^{n} (x_k - a)^2$, имеет распределение χ^2 с n степенями свободы. Пусть $\chi^2_{\alpha, n}$ является решением уравнения

$$\mathsf{P}\,\{\chi^2_n > \chi^2_{\alpha, n}\} = \alpha. \tag{5.5}$$

Тогда $\mathsf{P}\left\{\chi^2_{1-\alpha, n} < \frac{S^2}{\sigma^2} < \chi^2_{\alpha, n}\right\} = 1 - 2\alpha$ и, следовательно,

$$\mathsf{P}\,\{S/\chi_{\alpha, n} < \sigma < S/\chi_{1-\alpha, n}\} = 1 - 2\alpha. \tag{5.6}$$

При неизвестном a воспользуемся теоремой 3.4. Поскольку величина nm_2/σ^2 имеет распределение χ^2 с $n - 1$ степенями свободы, то

$$\mathsf{P}\,\{\sqrt{nm_2}/\chi_{\alpha, n-1} < \sigma < \sqrt{nm_2}/\chi_{1-\alpha, n-1}\} = 1 - 2\alpha. \tag{5.7}$$

Среди прикладных задач встречаются такие, при решении которых приходится сравнивать средние двух независимых выборок. Для такого сравнения полезно иметь доверительный интервал для разности соответствующих выборкам математических ожиданий. Пусть $x_{i1}, x_{i2}, \ldots, x_{in_i}$, $i = 1, 2$, независимы, нормально распределены и

$$\mathsf{M}x_{ik} = a_i,\ \mathsf{D}x_{ik} = \sigma^2,\ i = 1, 2,\ k = 1, \ldots, n_i.$$

Найдем доверительный интервал для $a_1 - a_2$. Положим

$$\bar{x}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} x_{ik}, \quad m_{2i} = \frac{1}{n_i} \sum_{k=1}^{n_i} (x_{ik} - \bar{x}_i)^2, \quad i = 1, 2.$$

Из теоремы 3.4 и из независимости выборок следует, что $\bar{x}_1, \bar{x}_2, m_{21}, m_{22}$ независимы. Кроме того, величина $n_i m_{2i}/\sigma^2$ имеет распределение χ^2 с $n_i - 1$ степенями свободы. Следовательно, величина $(n_1 m_{21} + n_2 m_{22})/\sigma^2$ имеет распределение χ^2 с $n_1 + n_2 - 2$ степенями свободы. Таким образом, величина

$$\tau_{n_1+n_2-2} = \frac{(\bar{x}_1 - \bar{x}_2 - (a_1 - a_2))/\sigma}{\sqrt{\dfrac{n_1 m_{21} + n_2 m_{22}}{(n_1 + n_2 - 2)\,\sigma^2}}} = \frac{\bar{x}_1 - \bar{x}_2 - a_1 + a_2}{\sqrt{n_1 m_{21} + n_2 m_{22}}} \sqrt{n_1 + n_2 - 2} \tag{5.8}$$

имеет распределение Стьюдента с $n_1 + n_2 - 2$ степенями свободы. Используя величину (5.8), легко выписать доверительный интервал для $a_1 - a_2$.

При построении доверительных интервалов (5.1), (5.4), (5.6), (5.7) и величины (5.8) предполагалось, что элементы выборок имеют нормальное распределение. Пусть теперь $x_1, x_2, \ldots, x_n$ — произвольная выборка. При больших n можно построить простые приближенные доверительные интервалы без предположения о нормальности x_k. Найдем, например, доверительный интервал для параметра $a_1 = \mathsf{M}x_k$. Пусть $\sigma^2 = \mathsf{D}x_k$. Тогда по центральной предельной теореме распределение величины

$$\frac{x_1 + x_2 + \ldots + x_n - na}{\sqrt{\sigma^2/n}} = \frac{\bar{x} - a}{\sigma/\sqrt{n}}$$

близко к нормальному распределению с параметрами (0, 1). Если σ_n^2 — состоятельная оценка параметра σ^2, то распределение величины $(\bar{x} - a)\sqrt{n}/\sigma_n$ тоже близко к нормальному с параметрами (0, 1) и, следовательно,

$$\lim_{n\to\infty} \mathsf{P}\left\{\left|\frac{\bar{x} - a}{\sigma_n/\sqrt{n}}\right| < u_\alpha\right\} = 1 - 2\alpha,$$

или

$$\mathsf{P}\left\{\bar{x} - u_\alpha \frac{\sigma_n}{\sqrt{n}} < a < \bar{x} + u_\alpha \frac{\sigma_n}{\sqrt{n}}\right\} \approx 1 - 2\alpha \qquad (5.9)$$

при $n \to \infty$.

Рассмотрим асимптотическое равенство (5.9) в двух частных случаях. Пусть $x_1, \ldots, x_n$ независимы и имеют распределение Пуассона с параметром $\lambda = \mathsf{M}x_k = \mathsf{D}x_k$. Величина $\bar{x}$ является состоятельной оценкой λ. Следовательно, при $n \to \infty$

$$\mathsf{P}\left\{\bar{x} - u_\alpha \sqrt{\frac{\bar{x}}{n}} < \lambda < \bar{x} + u_\alpha \sqrt{\frac{\bar{x}}{n}}\right\} \to 1 - 2\alpha. \qquad (5.10)$$

Пусть μ_n — число успехов в n испытаниях Бернулли с вероятностью успеха в каждом испытании, равной p. Величина μ_n/n является состоятельной оценкой p. Следовательно, при $n \to \infty$

$$\mathsf{P}\left\{\frac{\mu_n}{n} - \frac{u_\alpha}{\sqrt{n}}\sigma_n < p < \frac{\mu_n}{n} + \frac{u_\alpha}{\sqrt{n}}\sigma_n\right\} \to 1 - 2\alpha, \quad (5.11)$$

где $\sigma_n = \sqrt{\frac{\mu_n}{n}\left(1 - \frac{\mu_n}{n}\right)}$.

Таким же способом можно найти асимптотические доверительные интервалы и для других моментов. При помощи теоремы 3.3 и аналогичных ей можно также получать асимптотические доверительные интервалы для функций от неизвестных параметров.

§ 6. Статистическая проверка гипотез

В прикладных задачах часто требуется по эмпирическим данным проверить то или иное предположение. Например: 1) используя полученную выборку, нужно установить, имеет ли некоторый параметр определенное значение; 2) для различающихся выборочных характеристик, соответствующих одинаковым теоретическим, нужно решить, следует ли их различие считать случайным или его следует признать значимым, и т. д.

Правила, согласно которым проверяемые предположения (гипотезы) принимаются или отвергаются, называют *критериями значимости*. Рассмотрим некоторые из них.

6.1. Критерии значимости, основанные на интервальных оценках. Пусть $x_1, x_2, \ldots, x_n$ — независимые нормально распределенные случайные величины с параметрами (a, σ^2). Используя эту выборку, необходимо принять или отвергнуть гипотезу, состоящую в том, что $a=a_0$, где a_0— некоторое заданное число. Если эта гипотеза является верной, то из равенства (5.4) следует, что $\mathsf{P}\left\{|\bar{x}-a_0|> t_{\alpha, n-1}\sqrt{\frac{m_2}{n-1}}\right\}=2\alpha$. При достаточно малом 2α событие $|\bar{x}-a_0|>t_{\alpha, n-1}\sqrt{\frac{m_2}{n-1}}$ практически невозможно. Если для данной выборки обнаруживается, что $|\bar{x}-a_0|>t_{\alpha, n-1}\sqrt{\frac{m_2}{n-1}}$, то это отклонение считается значимым для выбранного уровня значимости 2α и гипотеза о равенстве $a = a_0$ отвергается. Вероятность ошибочного решения (отвергнуть гипотезу $a = a_0$, если она верна) равна 2α.

Полагая в (5.8) $a_1 = a_2$, можно по аналогии с критерием для проверки равенства $a = a_0$ построить критерий для проверки предположения о равенстве математических ожиданий, соответствующих двум независимым выборкам.

6.2. Критерий χ^2. Рассмотрим сначала следующую задачу. По выборке $x_1, x_2, \ldots, x_n$ нужно решить, является ли заданная функция $F(x)$ функцией распределения случайной величины x_k, $k = 1, \ldots, n$. Разобьем числовую ось точками $z_1 < z_2 < \ldots < z_r$ на $r+1$ непересекающихся интервалов и полуинтервалов:

$$(-\infty, z_1),\ [z_1, z_2), \ldots, [z_r, +\infty). \tag{6.1}$$

Если величины x_k имеют своей функцией распределения $F(x)$, то можно найти вероятности:

$$p_1 = \mathsf{P}\{x_k \in (-\infty, z_1)\} = F(z_1),$$

$$p_l = \mathsf{P}\{x_k \in [z_l, z_{l+1})\} = F(z_{l+1}) - F(z_l), \quad l = 1, 2, \ldots, r,$$

$$p_{r+1} = \mathsf{P}\{x_k \in [z_r, +\infty)\} = 1 - F(z_r).$$

Со случайными величинами $x_1, x_2, \ldots, x_n$ естественно связана полиномиальная схема с n испытаниями, в которой результатом k-го испытания является попадание значения x_k в какой-либо интервал (6.1). Обозначим $m_l = m_l(x_1, \ldots, x_n)$ число значений среди $x_1(\omega), x_2(\omega), \ldots, x_n(\omega)$, попавших в $[z_l, z_{l+1})$. Если наше предположение о законе распределения x_k верно, то m_l должны быть близки к $\mathsf{M}m_l = np_l$, $l = 1, 2, \ldots, r+1$. Общее отклонение всех m_l можно измерять различными способами. Чаще всего в качестве меры отклонения используют величину

$$\eta_{n,r} = \sum_{l=1}^{r+1} \frac{(m_l - np_l)^2}{np_l}. \tag{6.2}$$

Оказывается (см. [10]), что для любого x при $n \to \infty$ и постоянных $p_i > 0$ $(i = 1, 2, \ldots, r+1)$

$$\mathsf{P}\{\eta_{n,r} < x\} \to \mathsf{P}\{\chi_r^2 < x\}, \tag{6.3}$$

где случайная величина χ_r^2 имеет распределение χ^2 с r степенями свободы. Используя (6.3), можно выбрать «критическое» значение C так, чтобы при нашей гипотезе вероятность события

$$\eta_{n,r} > C \tag{6.4}$$

была мала и, таким образом, это событие можно было бы считать практически невозможным. Если же событие (6.4) фактически наблюдается, то говорят, что выборка обнаруживает значимое отклонение от проверяемой гипотезы. Это указывает на несовместимость нашей гипотезы с наблюденными значениями $x_1(\omega), x_2(\omega), \ldots, x_n(\omega)$. Таким образом, правило проверки, или *статистический критерий, состоит в том, что гипотеза отвергается, ес-*

ли произошло событие (6.4), *и гипотеза не противоречит наблюдениям, если произошло противоположное событие.*

Величина $\alpha = \mathsf{P}\{\eta_{n,r} > C\}$ в случае, когда гипотеза верна, является вероятностью отвергнуть правильную гипотезу. Если значение α задано, то ввиду (6.3) число C можно вычислить приближенно, используя уравнение $\alpha = \mathsf{P}\{\chi_r^2 > C\}$. В конце книги приведены таблицы чисел $\chi^2_{\alpha, m}$, для которых $\mathsf{P}\{\chi_n^2 > \chi^2_{\alpha, m}\} = \alpha$. Выбор α зависит от рассматриваемой практической задачи. Чаще всего полагают $\alpha = 0{,}01$ или $\alpha = 0{,}05$. При $\alpha = 0{,}01$ примерно в 1% случаев будет отвергаться правильная гипотеза.

Если функция распределения $F(x)$ непрерывна, то проверку гипотезы о соответствии $F(x)$ полученной выборке можно проводить, используя величину $\zeta_n = \sqrt{n} \sup\limits_{-\infty<x<\infty} |F_n^*(x) - F(x)|$, где $F_n^*(x)$ — эмпирическая функция распределения. Зададим некоторый уровень значимости α. Тогда для достаточно больших n можно подобрать, воспользовавшись теоремой А. Н. Колмогорова, число C так, чтобы $\mathsf{P}\{\zeta_n > C\} = \alpha$.

Указанные критерии применимы лишь тогда, когда закон распределения элементов выборки точно определен. Однако во многих случаях бывает известен только тип распределения (например, нормальное), но неизвестны некоторые параметры распределения $\theta_1, \ldots, \theta_s$. В этих условиях вероятности $p_l = p_l(\theta_1, \ldots, \theta_s)$ в (6.2) будут содержать неизвестные параметры и, следовательно, получить числовые значения p_l нельзя. Обозначим $\eta^*_{n,r}$ величину, вычисленную по формуле (6.2), в которой при нахождении p_l в качестве числовых значений неизвестных параметров $\theta_1, \ldots, \theta_s$ были использованы их оценки $\theta_1^*, \ldots, \theta_s^*$. Будем отвергать проверяемую гипотезу, если произойдет событие $\eta^*_{n,r} > C$. Для вычисления уровня значимости можно воспользоваться предельным при $n \to \infty$ распределением $\eta^*_{n,r}$. Оказывается (см. [10], гл. 30, § 30.3), что если оценки $\theta_1^*, \ldots, \theta_s^*$ получены из уравнения правдоподобия или минимизируют величину $\eta_{n,r} = \eta_{n,r}(\theta_1, \ldots, \theta_s)$, то предельным распределением $\eta^*_{n,r}$ является распределение χ^2. В этом случае число степеней свободы предельного распределения χ^2 равно $r - s$, т. е. меньше числа степеней свободы предельного распределения $\eta_{n,r}$ на число параметров, замененных оценками, Рассмотрим еще два применения критерия χ^2. Пусть по элементам выборки образована $(r \times s)$-таблица с двумя входами по двум переменным признакам A и B, принимающим значения $A_1, \ldots, A_r$, $B_1, \ldots, B_s$; (i, j)-й элемент таблицы ν_{ij} является числом элементов выборки, обладающих значениями признаков A_i и B_j. Пусть p_{ij} — вероятность того, что элемент выборки обладает значениями A_i и B_j. Тогда гипотеза о независимости признаков A и B равносильна тому, что существует $r+s$ постоянных $p_{i\cdot}$ $(i = 1, \ldots, r)$, $p_{\cdot j}$ $(j = 1, \ldots, s)$ таких, что $p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$, $\sum\limits_i p_{i\cdot} = \sum\limits_j p_{\cdot j} = 1$. Для применения критерия χ^2 нужно вычислить величину

$$\sum \frac{(\nu_{ij} - np_{i\cdot}p_{\cdot j})^2}{np_{i\cdot}p_{\cdot j}} \qquad \left(n = \sum_{i,j} \nu_{ij}\right),$$

заменив в ней вероятности $p_{i\cdot}$ и $p_{\cdot j}$ их оценками $p^*_{i\cdot} = \frac{\nu_{i\cdot}}{n}$, $p^*_{\cdot j} = \frac{\nu_{\cdot j}}{n}$,

где $\nu_{i\cdot} = \sum\limits_j \nu_{ij}$, $\nu_{\cdot j} = \sum\limits_i \nu_{ij}$. Полученная величина имеет вид

$$\eta^*_{n,\,rs} = n \sum_{i,j} \frac{\left(\nu_{ij} - \frac{\nu_{i\cdot}\nu_{\cdot j}}{n}\right)^2}{\nu_{i\cdot}\nu_{\cdot j}} = n\left(\sum_{i,j} \frac{\nu_{ij}^2}{\nu_{i\cdot}\nu_{\cdot j}} - 1\right).$$

Число независимых параметров, замененных оценками, равно $r + s - 2$, и, следовательно, распределение величины $\eta^*_{n,\,rs}$ сходится при $n \to \infty$ к распределению χ^2 с $rs - (r + s - 2) - 1 = (r-1)(s-1)$ степенями свободы. Это предельное распределение используется для приближенного вычисления уровня значимости. Большое значение величины $\eta^*_{n,\,rs}$ указывает на то, что отклонение от гипотезы о независимости значимо.

Перейдем к рассмотрению другого примера. Пусть имеется s независимых выборок. Требуется проверить гипотезу о том, что эти выборки однородны, т. е. что их элементы имеют одинаковое распределение. Разобьем каждую выборку по значениям какого-либо признака (одного для всех выборок) на r групп. Обозначим ν_{ij} число элементов j-й выборки, обладающих i-м значением признака; p_{ij} — вероятность того, что элемент j-й выборки обладает i-м значением признака. При использовании величин ν_{ij} гипотеза об однородности означает, что вероятности $p_{ij} = p_i$, т. е. одинаковы для всех выборок. Для применения критерия χ^2 нужно вычислить величину $\sum\limits_{i,j} (\nu_{ij} - n_j p_i)^2 / n_j p_i$, где n_j — объем j-й выборки. Заменяя неизвестные параметры p_i их оценками $p_i^* = \left(\sum\limits_j \nu_{ij}\right) \Big/ \left(\sum\limits_j n_j\right)$, получим величину, предельным распределением которой является распределение χ^2 с $(r-1)(s-1)$ степенями свободы (см. [10], § 30.6).

6.3. Общие понятия о статистической проверке гипотез. Критерии, рассмотренные в пп. 6.1, 6.2, формулируются в виде неравенств для некоторой функции от выборки. Эту функцию обычно называют статистикой критерия. Неравенства для статистик критерия в пространстве значений выборки $x = (x_1, \ldots, x_n)$ определяют некоторую *критическую область* S, зависящую от вида статистик и уровня значимости. Если реализация выборки попадает в критическую область, то проверяемая гипотеза отвергается. Очевидно, что даже при фиксированной вероятности $\alpha = \mathsf{P}_H\{x \in S\}$*), где P_H — распределение вероятностей, соответствующее гипотезе H, выбор критической области S не всегда однозначен. Таким образом, среди многих критериев нужно отобрать один каким-либо разумным способом.

Будем называть гипотезу H *простой*, если она однозначно определяет распределение выборки; в других случаях гипотеза H называется *сложной*. Обычно сложной гипотезой является предположение о распределении выборки, зависящем от некоторых параметров, значения которых неизвестны. Предположим, что требуется по выборке проверить некоторую простую гипотезу H_0. Пусть S — критическая область рассматриваемого критерия: если $x \in S$, то гипотеза H_0 отвергается. Одной из характеристик такого критерия является *ошибка 1-го рода* $\alpha = \mathsf{P}_{H_0}(x \in S)$. Таким образом, ошибка 1-го рода — это вероятность отвергнуть гипотезу H_0, когда она верна. Однако одной этой характеристики еще недостаточно для выбора критерия. Может возникнуть ошибка другого типа: H_0 не верна, но $x \in \bar{S}$, и принимается H_0, т. е. не отвергается ложная гипотеза. Вероятность этого события можно вычислить, если известно, какое распределение выборки в этом случае истинно. Предположим, что альтернативная (или конкурирующая) гипотеза H_1 простая. Тогда она однозначно определяет распределение выборки. *Ошибкой 2-го рода* называют вероятность $\beta = \mathsf{P}_{H_1}(\bar{S})$, т. е. вероятность принять основную гипотезу H_0, когда верна конкурирующая. Вероятность $1 - \beta$ называют *мощностью критерия*. При заданной ошибке 1-го рода α нужно среди всех критериев выбирать *наиболее мощный критерий*, т. е. критерий, имеющий наибольшую мощность $1 - \beta$ или наименьшую ошибку 2-го рода β.

*) Для дискретных распределений не всегда можно подобрать S так, чтобы $\mathsf{P}_H\{x \in S\} = \alpha$. В этом случае требуют, чтобы выполнялось неравенство $\mathsf{P}_H\{x \in S\} \leqslant \alpha$.

Для иллюстрации введенных характеристик критерия рассмотрим два примера.

Пример 1. Пусть в случайной выборке $x_1, x_2, \ldots, x_n$ величины x_k распределены нормально с параметрами (a, σ^2). Параметр σ известен, а относительно a имеется два предположения, или две гипотезы H_0 и H_1, согласно которым $a = a_0$ и $a = a_1$ $(a_0 < a_1)$ соответственно. Несмещенной и состоятельной оценкой математического ожидания a является $\bar{x} = (x_1 + \ldots + x_n)/n$. Таким образом, в рассматриваемой задаче естественно выбрать ту гипотезу, к параметру a которой ближе $\bar{x}$. При любой из гипотез (H_0 или H_1) $\bar{x}$ имеет нормальное распределение с $\mathsf{D}\bar{x} = \sigma^2/n$; $\mathsf{M}\bar{x} = a$ зависит от гипотезы. Следовательно, распределения $\bar{x}$ при H_0 и H_1 различны. Вероятности событий, вычисляемые при гипотезах H_0 и H_1, будем обозначать соответственно P_0 и P_1.

Выберем некоторое число C, $a_0 < C < a_1$. Критерий можно сформулировать следующим образом: если $\bar{x} > C$, то принимается H_1, а если $\bar{x} < C$, то принимается H_0.

Если верна гипотеза H_0 и произошло событие $\bar{x} > C$, то принимается H_1. Вероятность отвергнуть H_0, когда она верна (ошибка 1-го рода), равна $\alpha = \mathsf{P}_0(\bar{x} > C)$.

Если верна гипотеза H_1 и произошло событие $\bar{x} < C$, то принимается H_0. Вероятность принять H_0, когда верна H_1 (ошибка 2-го рода), равна $\beta = \mathsf{P}_1(\bar{x} < C)$.

Пусть α задано. Тогда

$$\alpha = \mathsf{P}_0(\bar{x} > C) = \mathsf{P}_0\left\{\frac{\bar{x} - a_0}{\sigma/\sqrt{n}} > (C - a_0)\frac{\sqrt{n}}{\sigma}\right\}.$$

Так как $\sqrt{n}(\bar{x} - a_0)/\sigma$ имеет нормальное распределение с параметрами (0, 1), то $(C - a_0)\frac{\sqrt{n}}{\sigma} = u_\alpha$, где u_α определено уравнением (5.2). Отсюда

$$C = a_0 + u_\alpha \frac{\sigma}{\sqrt{n}}. \tag{6.5}$$

Следовательно, ошибка 2-го рода β-удовлетворяет соотношению

$$\beta = \mathsf{P}_1(\bar{x} < C) = \mathsf{P}_1\left\{\frac{\bar{x} - a_1}{\sigma/\sqrt{n}} < (C - a_1)\frac{\sqrt{n}}{\sigma}\right\}.$$

Поэтому

$$(C - a_1)\frac{\sqrt{n}}{\sigma} = u_{1-\beta} = -u_\beta,$$

Подставляя в это равенство значение C из (6.5), получим

$$u_\alpha + u_\beta = \frac{a_1 - a_0}{\sigma} \sqrt{n}. \tag{6.6}$$

Из (6.6) при заданных α и β нетрудно найти n. Если n менять нельзя, то (6.6) используется для выбора α и β. Если же α и n фиксированы, то величина β определяется однозначно. Величина α и β выбираются в зависимости от степени нежелательности связанных с ними ошибочных решений. Иногда «степень нежелательности» можно выразить довольно точно. Пусть, например, при некоторой упрощенной проверке бракованное изделие может быть пропущено с вероятностью α и хорошее изделие принято за бракованное с вероятностью β. Если бракованное изделие продано, то за его гарантийный ремонт нужно платить P рублей. Если хорошее изделие забраковано, то теряется его стоимость Q рублей. Пусть в проверяемой партии из N изделий примерно M бракованных. Тогда средние потери при контроле этой партии равны

$$(M\alpha)\, P + (N - M)\, \beta Q. \tag{6.7}$$

Для выбора α и β нужно найти минимум (6.7) при наличии связи (6.6).

Пример 2 (см. § 3 гл. 1). Перейдем теперь к задаче выбора модели опыта, состоящего в подбрасывании двух монет. Обозначим μ_n число выпадений двух монет одинаковой стороной в n опытах. Вместо моделей 1 и 2 мы будем теперь говорить о гипотезах H_1 и H_0. Рассмотрим величину μ_n/n. Так как μ_n/n имеет при гипотезе H_l, $l = 0, 1$, биномиальное распределение, то

$$\mathsf{M}\frac{\mu_n}{n} = a_l, \quad \mathsf{D}\frac{\mu_n}{n} = \frac{\sigma_l^2}{n},$$

где

$$a_1 = \frac{2}{3}, \quad \sigma_1^2 = \frac{2}{3}\cdot\frac{1}{3}; \quad a_0 = \frac{1}{2}, \quad \sigma_0^2 = \frac{1}{4}. \tag{6.8}$$

Если $\mu_n/n > C$, то будем принимать H_1, а в противном случае H_0. Если биномиальное распределение заменить аппроксимирующим его нормальным, то для вычисления C сохранится формула (6.5) с $\sigma = \sigma_0$, а вместо (6.6) получим

$$\sigma_0 u_\alpha + \sigma_1 u_\beta = (a_1 - a_0)\sqrt{n}, \tag{6.9}$$

При $\alpha = \beta = 0{,}05$ имеем $u_\alpha = u_\beta = 1{,}65$. Подставляя эти значения и величины (6.8) в (6.9), получим

$$n \geqslant \frac{(\sigma_0 u_\alpha + \sigma_1 u_\beta)^2}{(a_1 - a_0)^2} \approx \left(\frac{1{,}65 \left(\frac{1}{2} + \frac{\sqrt{2}}{3} \right)}{\frac{2}{3} - \frac{1}{3}} \right)^2 \approx 79{,}2,$$

и, следовательно, можно положить $n = 80$. По формуле (6.5)

$$C = 0{,}5 + 1{,}65 \frac{1}{2\sqrt{80}} \approx 0{,}61.$$

Таким образом, гипотеза H_1 принимается, если $\mu_n/n > 0{,}61$, в противном случае принимается H_0.

Покажем теперь, как найти наиболее мощный критерий. Пусть по выборке $x = (x_1, \ldots, x_n)$ нужно различить две простые гипотезы H_0 и H_1, согласно которым выборка имеет соответственно плотности распределения $p_0(u)$, $p_1(u)$, $u = (u_1, \ldots, u_n)$. Рассмотрим критические множества вида

$$S_c = \{u = (u_1, \ldots, u_n)\colon p_1(u) \geqslant c p_0(u)\}. \tag{6.10}$$

Ограничимся рассмотрением случаев, когда для любого α существует постоянная c такая, что $\mathsf{P}_{H_0}(S_c) = \alpha$. Следующий результат был получен Нейманом и Пирсоном.

Теорема 6.1. *Среди всех критериев, различающих гипотезы H_0 и H_1 с заданной ошибкой 1-го рода α, наиболее мощным является критерий, определяемый критическим множеством* (6.10).

Доказательство. Пусть S — критическое множество произвольного критерия с $\mathsf{P}_{H_0}(S) = \alpha$. Тогда

$$\mathsf{P}_{H_0}(S_c \setminus S_c S) = \alpha - \mathsf{P}_{H_0}(S_c S) = \mathsf{P}_{H_0}(S \setminus S S_c).$$

Используя это равенство и (6.10), находим

$$\mathsf{P}_{H_1}(S_c \setminus S_c S) = \int\limits_{S_c \setminus S S_c} p_1(u)\,du \geqslant$$

$$\geqslant c \int\limits_{S_c \setminus S S_c} p_0(u)\,du = c\mathsf{P}_{H_0}(S_c \setminus S S_c) = c\mathsf{P}_{H_0}(S \setminus S S_c) =$$

$$= c \int\limits_{S \setminus S S_c} p_0(u)\,du \geqslant \int\limits_{S \setminus S S_c} p_1(u)\,du = \mathsf{P}_{H_1}(S \setminus S S_c),$$

так как $p_1(u) < cp_0(u)$ на множестве $S \setminus SS_c$. Прибавляя к обеим частям последнего неравенства $\mathsf{P}_{H_1}(SS_c)$, получим $\mathsf{P}_{H_1}(S) \geqslant \mathsf{P}_{H_1}(S_c)$. Теорема доказана.

Используя доказанную теорему, нетрудно проверить, что критерий в примере 1 является наиболее мощным. Так как совместная плотность распределения выборки согласно гипотезе H_i $(i = 0, 1)$ равна

$$p_i(u_1, \ldots, u_n) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{k=1}^{n} (u_k - a_i)^2\right\},$$

то неравенство в (6.10) равносильно неравенству

$$\sum_{k=1}^{n} (u_k - a_0)^2 \leqslant \sum_{k=1}^{n} (u_k - a_1)^2 + c_1$$

или неравенству

$$\bar{u} = \frac{1}{n} \sum_{k=1}^{n} u_k > C, \tag{6.11}$$

где c_1 и C — некоторые постоянные. Попадание выборки $x = (x_1, \ldots, x_n)$ в критическое множество, определяемое (6.11), очевидно равносильно наступлению события $\{\bar{x} > C\}$. Таким образом, доказано, что критерий примера 1 является наиболее мощным.

Пусть теперь конкурирующая гипотеза H_1 является сложной и определяемое ею распределение выборки P_θ зависит от параметра $\theta \in \Theta$; при основной гипотезе H_0 выборка имеет распределение P_{θ_0}. В такой ситуации естественно выбрать критическое множество S таким, чтобы построенный на его основе критерий был *несмещенным*, т. е. чтобы при заданной ошибке 1-го рода α неравенства $\mathsf{P}_\theta(S) \geqslant \alpha$ выполнялись при всех $\theta \in \Theta$.

§ 7. Регрессионный анализ

Задачи регрессионного типа были указаны в § 1. Рассмотрим сначала простейший случай. Пусть требуется найти коэффициенты a и b линейной функции $y = ax + b$. Допустим, что при любом значении x мы можем измерить с некоторой ошибкой значение y.

Будем предполагать, что в выборке $(y_1, \ldots, y_n)$ величины y_i представимы в виде

$$y_i = ax_i + b + \delta_i, \tag{7.1}$$

где δ_i независимы и нормально распределены, $\mathsf{M}\delta_i = 0$, $\mathsf{D}\delta_i = \sigma^2$, x_i не случайны и их значения известны. Для оценки неизвестных параметров a, b, σ^2 можно воспользоваться методом наибольшего правдоподобия. Функция правдоподобия выборки $y = (y_1, \ldots, y_n)$ имеет следующий вид:

$$L(y, a, b, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - ax_i - b)^2\right\}.$$

Отсюда получим систему уравнений

$$\begin{gathered}
\frac{\partial \ln L}{\partial a} = \frac{1}{\sigma^2} \sum_{i=1}^{n} x_i (y_i - ax_i - b) = 0, \\
\frac{\partial \ln L}{\partial b} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (y_i - ax_i - b) = 0, \\
\frac{\partial \ln L}{\partial (\sigma^2)} = -\frac{n}{2} \cdot \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (y_i - ax_i - b)^2 = 0.
\end{gathered} \tag{7.2}$$

Из (7,2) в предположении, что $\sum_{i=1}^{n} x_i = 0$, находим оценки

$$\begin{gathered}
a^* = \sum_{i=1}^{n} x_i y_i \Big/ \left(\sum_{i=1}^{n} x_i^2\right), \quad b^* = \frac{1}{n} \sum_{i=1}^{n} y_i, \\
\sigma^{*2} = \frac{1}{n} \sum_{i=1}^{n} (y_i - a^* x_i - b^*)^2.
\end{gathered} \tag{7.3}$$

Заменив в первых двух формулах (7,3) y_i по формуле (7,1), получим

$$a^* = a + \frac{\sum_{i=1}^{n} x_i \delta_i}{\sum_{i=1}^{n} x_i^2}, \quad b^* = b + \frac{1}{n} \sum_{i=1}^{n} \delta_i.$$

Из этих равенств следует, что

$$\mathsf{M}a^* = a, \quad \mathsf{M}b^* = b, \quad \mathsf{D}a^* = \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}, \quad \mathsf{D}b^* = \frac{\sigma^2}{n}.$$

Таким образом, непосредственно проверяется, что оценки a^*, b^* являются несмещенными и состоятельными.

Отметим, что в случае нормального распределения y_i оценки параметров a и b, полученные методом наибольшего правдоподобия, совпадают с оценками этих параметров, полученными минимизированием выражения

$$Q(a, b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2.$$

Метод получения оценок, основанный на минимизации суммы квадратов отклонений измерений от теоретических значений, называют *методом наименьших квадратов*.

Для иллюстрации метода наименьших квадратов рассмотрим следующую задачу. Пусть функция $y = \theta_1 x_1 + \theta_2 x_2 + \ldots + \theta_p x_p$ измерена в точках $(x_{1i}, x_{2i}, \ldots, x_{pi})$, $i = 1, 2, \ldots, n$, и

$$y_i = \theta_1 x_{1i} + \ldots + \theta_p x_{pi} + \delta_i, \tag{7.4}$$

где δ_i независимы и $\mathsf{D}\delta_i = \sigma^2$. Будем предполагать, что ранг матрицы $X = \| x_{ij} \|$ равен p. Используя метод наименьших квадратов, найдем по выборке $(y_1, \ldots, y_n)$ оценки параметров $\theta_1, \ldots, \theta_p$. Положим

$$Q(\theta_1, \ldots, \theta_p) = \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} \theta_k x_{ki} \right)^2.$$

Приравнивая к 0 производные Q по θ_l, $l = 1, \ldots, p$, получим систему уравнений

$$\frac{\partial Q}{\partial \theta_l} = -2 \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} \theta_k x_{ki} \right)^2 x_{li} = 0, \quad l = 1, \ldots, p;$$

отсюда

$$\sum_{i=1}^{n} \sum_{k=1}^{p} x_{li} x_{ki} \theta_k = \sum_{i=1}^{n} x_{li} y_i, \quad l = 1, \ldots, p, \tag{7.5}$$

или в матричной форме:

$$S\theta = Xy, \tag{7.6}$$

где $\theta = (\theta_1, \ldots, \theta_p)'$, $y = (y_1, \ldots, y_n')$, $X = \| x_{ij} \|$, $S = XX'$. Если определитель $| S | \neq 0$ *), то (7.6) имеет един-

*) Можно показать, что $| S | \neq 0$, если ранг X равен p.

ственное решение $\theta^* = (\theta_1^*, \ldots, \theta_p^*)'$, причем

$$\theta^* = (\theta_1^*, \ldots, \theta_p^*)' = S^{-1}Xy. \tag{7.7}$$

Отсюда, используя теорему 7.1 гл. 6, получаем ковариационную матрицу вектора θ^*:

$$\mathsf{D}\,[\theta^*] = (S^{-1}X)\,\mathsf{D}\,[y]\,(S^{-1}X)'.$$

Так как S симметрична, то симметрична и S^{-1}, поэтому $\mathsf{D}\,[\theta^*] = \sigma^2 S^{-1}XX'S^{-1}$ и, следовательно,

$$\mathsf{D}\,[\theta^*] = \sigma^2 S^{-1}. \tag{7.8}$$

Из формул (7.4) и (7.7), опираясь на свойства математического ожидания, получаем, что $\mathsf{M}\theta^* = S^{-1}X\mathsf{M}y = S^{-1}XX'\theta = \theta$. Эти соотношения доказывают несмещенность линейной оценки (7.7).

Теорема 7.1. *Если* $\tilde{\theta}^* = (\tilde{\theta}_1^*, \ldots, \tilde{\theta}_p^*)'$ — *любая линейная* (*относительно* $(y_1, \ldots, y_n)$ *в* (7.4)) *несмещенная оценка параметров* $\theta = (\theta_1, \ldots, \theta_p)$, *то* $\mathsf{D}\theta_k^* \leqslant \mathsf{D}\tilde{\theta}_k^*$, $k = 1, \ldots, p$, *где* $(\theta_1^*, \ldots, \theta_p^*)$ *определены формулой* (7.7).

Доказательство. Пусь $\tilde{\theta}^* = Cy$ — произвольная несмещенная оценка θ. Так как $\theta = \mathsf{M}\tilde{\theta}^* = CX'\theta$, то $CX' = E$, где E — единичная матрица. По теореме 7.1 гл. 6 находим $\mathsf{D}\,[\tilde{\theta}^*] = C\mathsf{D}\,[y]\,C' = \sigma^2 CC'$. Сравним матрицу CC' с матрицей S^{-1} в (7.8). Используя равенства $CX' = XC' = E$, нетрудно проверить, что

$$CC' = S^{-1} + (C - S^{-1}X)\,(C - S^{-1}X)'.$$

Отсюда и следует утверждение теоремы, поскольку диагональные элементы матрицы $(C - S^{-1}X)(C - S^{-1}X)'$ неотрицательны.

§ 8. Дисперсионный анализ

Дисперсионным анализом называют статистический метод анализа результатов измерений, зависящих от различных одновременно действующих факторов. Ограничимся рассмотрением простейшего случая, когда действует один фактор. Пусть, например, выборка разбита на r групп, причем i-я группа содержит n_i величин $x_{i1}, x_{i2}, \ldots, x_{in_i}$. Предполагается, что все указанные величины распределены нормально и $\mathsf{M}x_{ij} = a_i$, $\mathsf{D}x_{ij} = \sigma^2$, $j = 1, \ldots, n_i$, $i = 1, \ldots, r$. Нам нужно проверить гипотезу, согласно которой $a_1 = \ldots = a_r = a$. В физической постановке эта задача выглядит так. Одна и та же величина a измеряется r различными приборами, имеющими одинаковую точность. Нас интересует, имеют ли приборы различные систематические ошибки. В рассматриваемом примере исследуется влияние одного фактора (прибора) на погрешность измерения. Введем следующие обозначения:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{r} \sum_{j=1}^{n_i} x_{ij}, \quad \bar{x}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}, \quad n = n_1 + \ldots + n_r.$$

Групповые средние $\bar{x}_{i\cdot}$ являются, очевидно, несмещенными и состоятельными оценками величин a_i. Если все a_i одинаковы, то общее среднее не должно сильно отличаться от групповых. В противном случае разброс $\bar{x}_{i\cdot}$ относительно $\bar{x}$ должен быть более значительным. Представим общую, или полную, сумму квадратов отклонений

$$Q = \sum_{i=1}^{r} \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2 \qquad (8.1)$$

в следующем виде:

$$Q = Q_1 + Q_2, \qquad (8.2)$$

где

$$Q_1 = \sum_{i=1}^{r} n_i (\bar{x}_{i\cdot} - \bar{x})^2, \quad Q_2 = \sum_{i=1}^{r} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i\cdot})^2. \qquad (8.3)$$

Равенство (8.2) легко следует из (8.1), если воспользоваться формулами

$$(x_{ij} - \bar{x})^2 = [(x_{ij} - \bar{x}_{i\cdot}) + (\bar{x}_{i\cdot} - \bar{x})]^2 =$$
$$= (x_{ij} - \bar{x}_{i\cdot})^2 + (\bar{x}_{i\cdot} - \bar{x})^2 + 2(x_{ij} - \bar{x}_i)(\bar{x}_{i\cdot} - \bar{x}),$$

$$\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i\cdot}) = 0.$$

Сумму Q_1 в (8.2) называют суммой квадратов отклонений «между группами», а Q_2 — суммой квадратов отклонений «внутри групп».

Из теоремы 3.4 следует, что величина $\frac{1}{\sigma^2} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i\cdot})^2$ имеет распределение χ^2 с $n_i - 1$ степенями свободы и, следовательно, Q_2 имеет распределение χ^2 с $\sum\limits_{i=1}^{r}(n_i-1)=$ $=n-r$ степенями свободы. Можно показать (см. [10], гл. 36, § 36.2), что если $a_1=\ldots=a_r=a$, то Q_1 и Q_2 независимы и Q_1 имеет распределение χ^2 с $r-1$ степенями свободы. Следовательно, при $a_1=\ldots=a_r$ величина

$$F_{r-1,\,n-r}=\frac{Q_1/(r-1)}{Q_2/(n-r)} \tag{8.4}$$

имеет распределение Фишера с $r-1$, $n-r$ степенями свободы. Величина (8.4) может быть использована для проверки гипотезы о совпадении математических ожиданий: $a_1=\ldots=a_r=a$. Если эта гипотеза верна, то $\bar{x}_{i\cdot}$ и $\bar{x}$ являются состоятельными оценками одной и той же величины a и, следовательно, близки между собой, а величина Q_1 мала. Если a_i различны, то $\bar{x}_{i\cdot}$ и $\bar{x}$ сближаются с разными математическими ожиданиями:

$$\mathsf{M}\bar{x}_{i\cdot}=a_i,\qquad \mathsf{M}\bar{x}=\sum_{i=1}^{r}\frac{n_i}{n}a_i,$$

и, следовательно, сумма Q_1 должна принимать большие значения. Независимо от предположения о равенстве a_i знаменатель в (8.4) остается оценкой σ^2. Это, говоря нестрого, означает, что при увеличении расхождения между a_i величина (8.4) в среднем должна принимать большие значения. Статистический критерий формулируется следующим образом: *если $F_{r-1,\,n-r}>C$, то гипотеза, состоящая в том, что $a_1=\ldots=a_r$, отвергается.* При заданном уровне значимости $\alpha=0{,}05$ постоянную C нетрудно определить по таблице 6. Отметим, что при $r=2$ статистический критерий для проверки рассматриваемой гипотезы может быть получен на основе величины (5.8).

Задачи к главе 9

1. По выборке $x_1,\ldots,x_n$, полученной в задаче 24 гл. 5, найти: а) вариационный ряд $x_{(1)}\leqslant x_{(2)}\leqslant\ldots\leqslant x_{(n)}$; б) эмпирическую функцию распределения (построить ее график и график теоретической функции распределения); в) $\bar{x}=(x_1+\ldots+x_n)/n$ (сравнить с $\mathsf{M}x_x$);

г) $s^2=\dfrac{1}{n-1}\sum\limits_{k=1}^{n}(x_k-\bar{x})^2$ (сравнить с $\mathsf{D}x_k$).

2. Используя таблицу нормальных случайных чисел, получить реализацию выборки $x_1, \ldots, x_n$, где x_k нормально распределены с $\mathsf{M}x_k = 0{,}5$, $\mathsf{D}x_k = 1$; $n = 22$. Для полученной выборки выполнить задание, указанное в задаче 1.

3. Пусть $(x_1, y_1), \ldots, (x_n, y_n)$ — независимые одинаково распределенные двумерные величины. Положим

$$m_{11} = \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})(y_k - \bar{y}),$$

где $\bar{x} = (x_1 + \ldots + x_n)/n$, $\bar{y} = (y_1 + \ldots + y_n)/n$. Найти $\mathsf{M}m_{11}$ и показать, что $\mathsf{D}m_{11} = O\left(\frac{1}{n}\right)$ при $n \to \infty$.

4. Показать, что величина m_{11}, определенная в задаче 3, при $n \to \infty$ асимптотически нормальна.

5. Выборочным коэффициентом корреляции называется величина

$$r = m_{11}/\sqrt{m_{20}m_{02}},$$

где $m_{20} = \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^2$, $m_{02} = \frac{1}{n} \sum_{k=1}^{n} (y_k - \bar{y})^2$, а m_{11}, $\bar{x}$, $\bar{y}$ определены в задаче 3. Величина r используется в качестве критерия зависимости координат наблюдаемых двумерных величин. Рассуждая так же, как при доказательстве теоремы 3.3, показать, что при $n \to \infty$ величина r асимптотически нормальна.

6. Рассмотрим последовательность испытаний Бернулли с вероятностью успеха p в отдельном испытании. Пусть μ_n — число успехов в n первых испытаниях. Воспользовавшись схемой доказательства теоремы 3.3, показать, что при $n \to \infty$ величина $\eta_n = 2 \arcsin \sqrt{\frac{\mu_n}{n}}$ асимптотически нормальна с параметрами $(A, B/n)$. Найти A и B.

7. По выборке $x_1, \ldots, x_n$, где x_k, $k = 1, \ldots, n$, независимы, $\mathsf{M}x_k = a$, $\mathsf{D}x_k = \sigma_k^2$ (σ_k известны), найти несмещенную, линейную относительно $x_1, \ldots, x_n$ оценку a^* параметра a с наименьшей дисперсией.

8. Пусть $x_{i1}, \ldots, x_{in_i}$ $(i = 1, \ldots, I)$ — независимые нормально распределенные величины с параметрами (a, σ^2);

$$\bar{x}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} x_{ik}, \quad s_i^2 = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (x_{ik} - \bar{x}_i)^2.$$

Показать, что величина $\bar{x} = \sum_{i=1}^{I} c_i \bar{x}_i$, где $c_i = (1/s_i^2)/(1/s_1^2 + \ldots + 1/s_I^2)$, является несмещенной оценкой a.

9. Пусть $x_1, \ldots, x_n$ — независимые нормально распределенные случайные величины с параметрами (0, 1). Положим

$$\bar{x}_k = \frac{1}{k}\sum_{l=1}^{k} x_l, \quad S_k = \sum_{l=1}^{k}(x_l - \bar{x}_k)^2, \quad k = 1, 2, \ldots, n.$$

Доказать, что а) величины $x_{k+1} - \bar{x}$ и $\bar{x}_k$ независимы; б) величина S_n имеет распределение χ^2 с $n-1$ степенями свободы.

Указание. а) Найти $\operatorname{cov}(x_{k+1} - \bar{x}, \bar{x}_k)$; б) провести доказательство индукцией по k, воспользовавшись равенством

$$S_{k+1} = S_k + \frac{k}{k+1}(x_{k+1} - \bar{x}_k)^2.$$

10. Используя метод наибольшего правдоподобия, найти по выборке $x_1, \ldots, x_n$, где $\mathsf{P}\{x_k = m\} = \dfrac{\lambda^m}{m!}e^{-\lambda}$, $m = 0, 1, \ldots$, оценку λ^* параметра λ. Будет ли эта оценка несмещенной и состоятельной? Найти $\mathsf{M}\lambda^*$, $\mathsf{D}\lambda^*$.

11. Пусть $x_{(1)} \leqslant \ldots \leqslant x_{(n)}$ — вариационный ряд, построенный по выборке $x_1, x_2, \ldots, x_n$. Положим

$$\theta_1^* = \bar{x} = \frac{x_1 + \ldots + x_n}{n}, \quad \theta_2^* = \frac{x_{(1)} + x_{(n)}}{2}.$$

Найти $\mathsf{M}\theta_k^*$, $\mathsf{D}\theta_k^*$ $(k = 1, 2)$, если а) величины x_k равномерно распределены на отрезке $[a, b]$; б) $p_{x_k}(x) = \alpha e^{-\alpha x}$, $x > 0$.

12. Пусть $x_{(1)} \leqslant \ldots \leqslant x_{(n)}$ — вариационный ряд, построенный по выборке $x_1, \ldots, x_n$, где x_k имеет плотность распределения, равную e^{c-x} при $x \geqslant c > 0$. Положим $c^* = x_{(1)} - \dfrac{1}{n}$. Найти $\mathsf{M}c^*$, $\mathsf{D}c^*$.

13. По выборке, полученной в задаче 2, построить доверительные интервалы для a с доверительной вероятностью 0,95 и для σ^2 с доверительной вероятностью 0,94.

14. Используя критерий χ^2, проверить гипотезу о том, что выборка, полученная в задаче 24 гл. 5, соответствует равномерному распределению на отрезке [0, 1]. Уровень значимости $\alpha = 0{,}05$.

15. Найти статистику наиболее мощного критерия, различающего по выборке $x_1, \ldots, x_n$ гипотезы

$$H_0\colon \mathsf{P}\{x_k = i\} = p_i^{(0)} > 0, \ i = 1, 2, \ldots, N;$$

$$H_1\colon \mathsf{P}\{x_k = i\} = p_i^{(1)} > 0, \ i = 1, 2, \ldots, N;$$

$$\sum_{i=1}^{N} p_i^{(0)} = \sum_{i=1}^{N} p_i^{(1)} = 1.$$

16. Для величины σ^{*2}, определенной формулой (7.3), найти $\mathsf{M}\sigma^{*2}$.

17. В случае произвольных $a_1, a_2, \ldots, a_r$ найти $\mathsf{M}Q_1$ и $\mathsf{M}Q_2$, где Q_1 и Q_2 определены формулами (8.3).

18. Пусть y_i — значения функции $f(x) = ax^2 + bx + c$, измеренные в точках $x_i = 0{,}2 + 0{,}5\,(i-1)$, $i = 1, 2, \ldots, 10$. При $a = 1$, $b = 2$, $c = -1$ найти реализацию выборки $y_i = x_i^2 + 2x_i - 1 + \delta_i$, $i = 1, 2, \ldots, 10$, где δ_i независимы и нормально распределены с $\mathsf{M}\delta_i = 0$, $\mathsf{D}\delta_i = 0{,}06$. Методом наименьших квадратов получить оценки a^*, b^*, c^* параметров a, b, c. Сравнить эти оценки с известными истинными значениями. Построить графики функций $y = x^2 + 2x - 1$, $y = a^*x^2 + b^*x + c^*$; отметить точки (x_i, y_i).

19. Заменить в задаче 17 реализацию $\delta_1, \ldots, \delta_{10}$ на реализацию этих величин, соответствующую равномерному распределению с $\mathsf{M}\delta_i = 0$, $\mathsf{D}\delta_i = 0{,}06$. Выполнить для новой реализации задание, указанное в задаче 17.

20. Пусть $x_1, x_2, \ldots, x_{15}$ — выборка, для которой $\mathsf{P}\{x_k = l\} = p_l$, $l = 1, 2, 3$. Получить 10 реализаций этой выборки с $p_l = p_l^{(0)} = 1/3$, $l = 1, 2, 3$, и 10 реализаций с $p_1 = p_1^{(1)} = 0{,}40$, $p_2 = p_2^{(1)} = 0{,}42$, $p_3 = p_3^{(1)} = 0{,}18$. Для каждой выборки, используя наиболее мощный критерий, найденный в задаче 15, выбрать одну из двух гипотез. Для вычисления ошибок 1-го и 2-го рода α и β использовать нормальное приближение. Выбрать $\alpha = \beta$. Найти частоты ошибок 1-го и 2-го рода.

Элементы теории случайных процессов

§ 1. Понятие о случайных процессах

Случайные процессы являются моделями многих реальных процессов. Два примера таких реальных процессов (работа телефонной станции и броуновское движение) обсуждались в § 1 гл. 1.

Случайным процессом называется семейство случайных величин $\xi_t = \xi_t(\omega)$, заданных на вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$. Параметр t обычно интерпретируется как время. Если $t \in (0, 1, 2, \ldots)$, то говорят, что ξ_t — процесс с дискретным временем; если же $t \in [0, T]$, то ξ_t — процесс с непрерывным временем.

Действительную функцию $\xi_t(\omega_0)$ при фиксированном ω_0 называют *реализацией* или *траекторией случайного процесса*. Если фиксировать $t = t_0$, то $\xi_{t_0}(\omega)$ является обычной случайной величиной. Функция распределения

$$F_t(x) = \mathsf{P}(\xi_t < x)$$

задает распределение значений процесса в момент времени t. Зная только $F_t(x)$ при всех t, мы еще ничего не можем сказать о зависимости значений процесса в какие-либо фиксированные моменты времени. Ввиду этого необходимо указать всевозможные совместные распределения

$$F_{t_1 t_2 \ldots t_n}(x_1, x_2, \ldots, x_n) = \mathsf{P}(\xi_{t_1} < x_1, \ldots, \xi_{t_n} < x_n), \quad (1.1)$$

где $0 \leqslant t_1 < t_2 < \ldots < t_n$, $n = 1, 2, \ldots$ В гл. 5 мы показали, что можно построить вероятностное пространство и определить на нем случайную величину с заданной функцией распределения. Аналогичная задача построения подходящего вероятностного пространства для семейства случайных величин ξ_t с заданными распределениями (1.1) является значительно более сложной задачей и выходит за рамки настоящей книги (см. А. Н. Колмогоров [8]).

В следующих параграфах мы опишем несколько типов случайных процессов и вычислим для них вероятности отдельных событий. Тип процесса будет определяться свойствами, которыми должны обладать величины ξ_t. Построения подходящих вероятностных пространств, на которых можно определить ξ_t с заданными свойствами, приведены не будут.

§ 2. Пуассоновский процесс

Случайный процесс ξ_t с непрерывным временем называется *процессом с независимыми приращениями*, если при любых $0 \leqslant t_1 < t_2 < \ldots < t_n$, $t_1, \ldots, t_n \in [0, \infty)$, $n = 1, 2, \ldots$, случайные величины

$$\xi_{t_1}, \xi_{t_2} - \xi_{t_1}, \ldots, \xi_{t_n} - \xi_{t_{n-1}}$$

независимы.

Пуассоновским процессом называется процесс ξ_t, удовлетворяющий следующим условиям:

1° ξ_t — *процесс с независимыми приращениями.*

2°. *При любых $t_1 < t_2$, s приращения $\xi_{t_2} - \xi_{t_1}$, $\xi_{t_2+s} - \xi_{t_1+s}$ одинаково распределены (однородность по времени).*

3°. $\xi_0(\omega) = 0$, $\omega \in \Omega$.

4°. *При* $h \to 0$

$$\mathsf{P}(\xi_h = 0) = 1 - \lambda h + o(h), \quad \mathsf{P}(\xi_h \geqslant 2) = o(h),$$

$$\mathsf{P}(\xi_h = 1) = \lambda h + o(h), \quad 0 < \lambda < \infty.$$

Число λ будем называть параметром пуассоновского процесса.

Пуассоновский процесс ξ_t можно задать на вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$, где множество Ω совпадает с множеством ступенчатых функций, у которых имеются только единичные скачки, а моменты времени, соответствующие скачкам, не имеют точек накопления.

Теорема 2.1. *Если ξ_t — пуассоновский процесс, то*

$$\mathsf{P}(\xi_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \ldots$$

Доказательство. Пусть $\varphi_t(x) = \mathsf{M}x^{\xi_t}$ — производящая функция ξ_t. По условию 1° величины $\xi_{t+h} - \xi_t$, ξ_t независимы. Следовательно,

$$\varphi_{t+h}(x) = \mathsf{M}x^{\xi_{t+h}} = \mathsf{M}x^{(\xi_{t+h} - \xi_t) + \xi_t} = \varphi_t(x)\, \mathsf{M}x^{\xi_{t+h} - \xi_t}.$$

Так как по свойству 2° величины $\xi_{t+h} - \xi_t$ и $\xi_h - \xi_0 = \xi_h$ одинаково распределены, то

$$\mathsf{M}x^{\xi_{t+h}-\xi_t} = \mathsf{M}x^{\xi_h}$$

и при $h \to 0$ согласно 4°

$$\mathsf{M}x^{\xi_h} = 1 - \lambda h + x\lambda h + o(h), \quad |x| \leqslant 1.$$

Таким образом,

$$\varphi_{t+h}(x) = \varphi_t(x)\,(1 - \lambda h + x\lambda h + o(h)).$$

Отсюда

$$\frac{\varphi_{t+h}(x) - \varphi_t(x)}{h} = -\lambda(x-1)\,\varphi_t(x) + o(1).$$

Переходя к пределу при фиксированном x и $h \to 0$, получим

$$\frac{d\varphi_t(x)}{dt} = -\lambda(x-1)\,\varphi_t(x).$$

Решение этого уравнения с начальным условием $\varphi_0(x) = 1$ определяется формулой

$$\varphi_t(x) = e^{\lambda t(x-1)} = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} x^k.$$

Полученная производящая функция является производящей функцией распределения Пуассона. Теорема доказана.

Используя доказанную теорему, можно найти совместные распределения $\xi_{t_1}, \ldots, \xi_{t_n}$ при любых $t_1 < \ldots < t_n$. Очевидно, что

$$\{\xi_{t_1} = k_1, \xi_{t_2} = k_2, \ldots, \xi_{t_n} = k_n\} =$$
$$= \{\xi_{t_1} = k_1, \xi_{t_2} - \xi_{t_1} = k_2 - k_1, \ldots, \xi_{t_n} - \xi_{t_{n-1}} = k_n - k_{n-1}\}, \tag{2.1}$$

где $k_1 \leqslant k_2 \leqslant \ldots \leqslant k_n$. Отсюда, учитывая независимость и однородность приращений, получим

$$\mathsf{P}(\xi_{t_1} = k_1, \ldots, \xi_{t_n} = k_n) =$$
$$= \frac{(\lambda t_1)^{k_1}}{k_1!} e^{-\lambda t_1} \cdot \frac{(\lambda(t_2 - t_1))^{k_2-k_1}}{(k_2-k_1)!} e^{-\lambda(t_2-t_1)} \ldots$$
$$\ldots \frac{(\lambda(t_n - t_{n-1}))^{k_n-k_{n-1}}}{(k_n - k_{n-1})} e^{-\lambda(t_n-t_{n-1})}, \tag{2.2}$$

Ввиду условия 4° случайная величина ξ_t при любом фиксированном t равна числу скачков траектории процесса ξ_s на отрезке $[0, t]$. Рассмотрим теперь величины τ_k ($k = 0, 1, 2, \ldots$), равные промежуткам времени между скачками, т. е. положим $\tau_k = \theta_k - \theta_{k-1}$, где θ_k — момент k-го скачка траектории процесса ($k = 1, 2, \ldots$), $\theta_0 = 0$.

Теорема 2.2. *При любом n случайные величины τ_1, $\tau_2, \ldots, \tau_n$ независимы и каждая из них имеет показательное распределение с параметром λ.*

Доказательство. Пусть $0 = t_0 < t_1 < \ldots < t_n$ и отрезки $\Delta_k = [t_k, t_k + h_k]$, $h_k > 0$ ($k = 1, \ldots, n$) не пересекаются. Так как:

$$\{\theta_1 \in \Delta_1, \theta_2 \in \Delta_2, \ldots, \theta_n \in \Delta_n\} = \\ = \bigcap_{k=1}^{n} \{(\xi_{t_k} - \xi_{t_{k-1}} = 0) \cap (\xi_{t_k+h_k} - \xi_{t_k} = 1)\},$$

то

$$\mathsf{P}\{\bigcap_{k=1}^{n} \theta_k \in \Delta_k\} = \prod_{k=1}^{n} (e^{-\lambda(t_k - t_{k-1})} \lambda h_k e^{-\lambda h_k}) = \\ = e^{-\lambda t_n} \lambda^n e^{-\lambda(h_1 + \ldots + h_n)} h_1 \ldots h_n = \\ = \int_{\Delta} \ldots \int p_{\theta_1 \ldots \theta_n}(t_1, \ldots, t_n)\, dt_1, \ldots, dt_n,$$

где $\Delta = \Delta_1 \times \ldots \times \Delta_n$ — n-мерный параллелепипед. Отсюда при $h_1, \ldots, h_n \to 0$ получим

$$\int_{\Delta} \ldots \int p_{\theta_1 \ldots \theta_n}(t_1, \ldots, t_n)\, dt_1 \ldots dt_n = \\ = \lambda^n e^{-\lambda t_n} h_1 \ldots h_n (1 + o(1)),$$

и, следовательно,

$$p_{\theta_1 \ldots \theta_n}(t_1, \ldots, t_n) = e^{-\lambda t_n} \lambda^n, \; 0 \leqslant t_1 < t_2 < \ldots < t_n. \tag{2.3}$$

Для вычисления совместной плотности распределения величин $(\tau_1, \ldots, \tau_n) = (g_1(\theta_1, \ldots, \theta_n), \ldots, g_n(\theta_1, \ldots, \theta_n))$, где $g_k(\theta_1, \ldots, \theta_n) = \theta_k - \theta_{k-1}$, воспользуемся теоремой 6.2 гл. 5. Отображение $s_k = g_k(t_1, \ldots, t_k) = t_k - t_{k-1}$, $k = 1, \ldots, n$, взаимно однозначно, и якобиан этого отображения равен 1. Кроме того, $t_n = s_1 + s_2 + \ldots + s_n$. Отсюда и из формул (2.2) и (5.6.5) получим

$$p_{\tau_1 \ldots \tau_n}(s_1, \ldots, s_n) = \lambda^n e^{-\lambda(s_1 + \ldots + s_n)} = \prod_{k=1}^{n} p_{\tau_k}(s_k),$$

где $p_{\tau_k}(s_k) = \lambda e^{-\lambda s_k}$, $s_k > 0$ $(k = 1, \ldots, n)$. Теорема доказана.

Поскольку траектория пуассоновского процесса определяется моментами скачков, то можно получить эквивалентное определение этого процесса, приняв независимые показательно распределенные величины $\tau_1, \tau_2, \ldots, \tau_n, \ldots$ за промежутки времен между скачками.

§ 3. Винеровский процесс

В § 2 был рассмотрен процесс, в котором изменения происходят скачками. Здесь мы определим процесс, имеющий непрерывные траектории.

Винеровским процессом называется случайный процесс ξ_t, *удовлетворяющий условиям*:

1°. ξ_t — *процесс с независимыми приращениями.*

2°. *При любых* $t_1 < t_2$, s *приращения* $\xi_{t_2} - \xi_{t_1}$, $\xi_{t_2+s} - \xi_{t_1+s}$ *одинаково распределены.*

3°. $\xi_0(\omega) = 0$, $\omega \in \Omega$.

4°. *При* $h \to 0$.

$$\mathsf{M}\xi_h = ah + o(h), \quad \mathsf{M}\,|\,\xi_n\,|^3 = o(h),$$

$$\mathsf{M}\xi_n^2 = bh + o(h), \quad -\infty < a < \infty, \ 0 < b < \infty.$$

Теорема 3.1. *Если* ξ_t — *винеровский процесс, то*

$$\mathsf{P}(\xi_t < x) = \frac{1}{\sqrt{2\pi bt}} \int\limits_{-\infty}^{x} e^{-\frac{(u-at)^2}{2bt}}\,du.$$

Доказательство. Положим $f_t(s) = \mathsf{M}e^{is\xi_t}$. Здесь аргумент характеристической функции обозначен буквой s. Пусть s фиксировано. Так же, как при доказательстве теоремы 2.1, найдем

$$f_{t+h}(s) = f_t(s)\,\mathsf{M}e^{is(\xi_{t+h}-\xi_t)} = f_t(s)\,\mathsf{M}e^{is\xi_h}.$$

По формуле (7.2.11)

$$\mathsf{M}e^{is\xi_h} = 1 + is\mathsf{M}\xi_h - \frac{s^2}{2}\mathsf{M}\xi_h^2 + O(|\,s\,|^3\,\mathsf{M}\,|\,\xi_h\,|^3).$$

Отсюда и из свойства 4° следует

$$\mathsf{M}e^{is\xi_h} = 1 + is(ah) - \frac{s^2bh}{2} + o(h).$$

Таким образом,

$$f_{t+h}(s)=\left(1+is(ah)-\frac{s^2bh}{2}+o(h)\right)f_t(s)$$

и

$$\frac{f_{t+h}(s)-f_t(s)}{h}=\left(isa-\frac{s^2b}{2}\right)f_t(s)+o(1).$$

Переходя в этом равенстве к пределу при $h\to 0$, получим

$$\frac{df_t(s)}{dt}=\left(isa-\frac{s^2b}{2}\right)f_t(s).$$

Отсюда, учитывая начальное условие $f_0(s)=1$, найдем

$$f_t(s)=e^{is(at)-\frac{(bt)s^2}{2}}.$$

Полученная характеристическая функция величины ξ_t соответствует нормальному распределению с параметрами $(at,\ bt)$. Теорема доказана.

Найдем совместное распределение величин $\xi_{t_1},\ldots,\xi_{t_n}$. В силу теоремы 3.1 и условия 1° совместное распределение приращений $\eta_k=\xi_{t_{k+1}}-\xi_{t_k}$, $k=1,2,\ldots,n$ $(\xi_{t_0}=0)$, легко выписывается. Тогда, используя равенства

$$\xi_{t_1}=\eta_1,\ \xi_{t_2}=\eta_1+\eta_2,\ldots,\ \xi_{t_n}=\eta_1+\eta_2+\ldots+\eta_n,$$

получим

$$\mathsf{P}(\xi_{t_1}<x_1,\xi_{t_2}<x_2,\ldots,\xi_{t_n}<x_n)=$$
$$=\mathsf{P}(\eta_1<x_1,\eta_1+\eta_2<x_2,\ldots,\eta_1+\ldots+\eta_n<x_n)=$$
$$=\int\limits_G\!\cdots\!\int\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi b\,(t_k-t_{k-1})}}\,e^{-\frac{[u_k-a(t_k-t_{k-1})]^2}{2b(t_k-t_{k-1})}}\,du_1\ldots du_n,$$

где

$$G=\{(u_1,\ldots,u_n):\ u_1<x_1,\ u_1+u_2<x_2,\ldots,\ u_1+u_2+\ldots+u_n<x_n\}.$$

§ 4. Ветвящийся процесс

Ветвящийся процесс — это процесс размножения и превращения частиц, в котором частицы развиваются независимо друг от друга. Дадим определение ветвящегося процесса с дискретным временем. Пусть

$$\xi_k(t),\ k = 1, 2, \ldots, n, \ldots;\ t = 0, 1, 2, \ldots,$$

— независимые одинаково распределенные целочисленные случайные величины с общей производящей функцией

$$\varphi(x) = \sum_{m=0}^{\infty} x^m p_m.$$

Величина $\xi_k(t)$ будет интерпретироваться как число непосредственных потомков частицы с номером k, существовавшей в момент t. Случайный процесс μ_t, определяемый равенствами

$$\mu_0 = 1,\ \mu_{t+1} = \begin{cases} \xi_1(t) + \xi_2(t) + \ldots + \xi_{\mu_t}(t), & \text{если } \mu_t > 0, \\ 0, & \text{если } \mu_t = 0, \end{cases}$$

называется *ветвящимся*.

При исследовании ветвящихся процессов часто используются производящие функции. Положим

$$\Phi_t(x) = \mathsf{M} x^{\mu_t}.$$

Отметим, что

$$\Phi_0(x) = x,\ \Phi_1(x) = \varphi(x).$$

Так как μ_{t+1} является суммой случайного числа независимых слагаемых, то, положив в теореме 1.4 гл. 7

$$\varphi_{\zeta_\nu}(x) = \Phi_{t+1}(x),\ \varphi_\nu(x) = \Phi_t(x),\ \varphi_{\xi_1}(x) = \varphi(x),$$

получим

$$\Phi_{t+1}(x) = \Phi_t(\varphi(x)). \tag{4.1}$$

Это функциональное уравнение позволяет найти производящую функцию $\Phi_t(x)$ при любом t. Нетрудно проверить, полагая $t = 1, 2, \ldots$ в (4.1), что

$$\Phi_2(x) = \varphi(\varphi(x)),\ \Phi_3(x) = \varphi(\varphi(\varphi(x))), \ldots$$

Таким образом, $\Phi_t(x)$ является t-й итерацией функции $\varphi(x)$: $\Phi_t(x) = \varphi(\varphi(\ldots\varphi(x))\ldots)$.

Найдем $A_t = \mathsf{M}\mu_t$. Продифференцировав (4.1) по x, придем к соотношению

$$\frac{d\Phi_{t+1}(x)}{dx} = \frac{d\Phi_t}{d\varphi} \cdot \frac{d\varphi}{dx}.$$

Последнее равенство, полагая $x = 1$ и $a = A_1 = \varphi'(1)$, можно записать в виде $A_{t+1} = A_t a$. Следовательно, $A_t = a^t$.

Поведение A_t при $t \to \infty$ качественно различается в следующих трех случаях: $a < 1$, $a = 1$, $a > 1$. Если $t \to \infty$, то при $a < 1$ среднее число частиц стремится к 0, $A_t = 1$ при $a = 1$, и $A_t \to \infty$ в случае $a > 1$. Ветвящиеся процессы с $a < 1$, $a = 1$ и $a > 1$ называют соответственно *докритическими*, *критическими* (если $\varphi(x) \neq x$) и *надкритическими*. Асимптотические свойства ветвящихся процессов в этих трех случаях существенно различны.

Исследуем условия вырождения процесса. Обозначим C событие, состоящее в том, что $\mu_t = 0$ при некотором t. Очевидно, что

$$(\mu_t = 0) \subset (\mu_{t+1} = 0),\ t = 1, 2, \ldots \tag{4.2}$$

Событие C можно представить в виде $C = \bigcup_{t=1}^{\infty} (\mu_t = 0)$. Согласно (1.3.11)

$$\lambda = P(C) = \lim_{t \to \infty} \mathsf{P}(\mu_t = 0).$$

Если $\lambda = 1$, то процесс называется *вырождающимся*. Докритические процессы ($a < 1$) являются вырождающимися, так как при $t \to \infty$

$$1 - \mathsf{P}(\mu_t = 0) = \mathsf{P}(\mu_t > 0) = \mathsf{P}\left(\mu_t > \frac{1}{2}\right) \leqslant 2\mathsf{M}\mu_t = 2a^t \to 0.$$

Покажем, что вероятность λ удовлетворяет уравнению

$$\varphi(x) = x. \tag{4.3}$$

Так как $\Phi_t(x)$ является t-й итерацией $\varphi(x)$, то наряду с (4.1) имеет место равенство

$$\Phi_{t+1}(x) = \varphi(\Phi_t(x)). \tag{4.4}$$

Очевидно, что $\Phi_t(0) = \mathsf{P}(\mu_t = 0)$. Полагая $x = 0$ в (4.4), получим

$$\mathsf{P}(\mu_{t+1} = 0) = \varphi(\mathsf{P}(\mu_t = 0)). \tag{4.5}$$

Отсюда при $t \to \infty$ найдем

$$\lambda = \varphi(\lambda). \tag{4.6}$$

Таким образом, λ является корнем уравнения (4.3). В тех случаях, когда (4.3) имеет единственный корень на отрезке $[0, 1]$, этот корень совпадает с λ. Пусть $\varphi(x) \neq x$. При $x \in [0, 1]$ имеем $\varphi'(x) > 0$, $\varphi''(x) > 0$. Следовательно, $\varphi(x)$ на отрезке $[0, 1]$ возрастает и обращена выпуклостью книзу. Параметр $a = \varphi'(1)$ определяет угловой коэффициент касательной к графику $y = \varphi(x)$ в точке $x = 1$. Следовательно, при $a = 1$ график $y = \varphi(x)$ касается $y = x$ в точке $x = 1$; при $a < 1$ касательная к $y = \varphi(x)$ проходит выше $y = x$. Таким образом, при $a \leqslant 1$ уравнение (4.3) имеет на отрезке $[0, 1]$ единственный корень $x = 1$, и, следовательно, ветвящийся процесс вырождается, если $a \leqslant 1$.

При $a > 1$ касательная к графику $y = \varphi(x)$ проходит ниже $y = x$ и уравнение (4.3) имеет на отрезке $[0, 1]$ два корня x_1, x_2 $(0 \leqslant x_1 < x_2 = 1)$. Покажем, что $\lambda = x_1$. Для этого достаточно показать, что $\mathsf{P}(\mu_t = 0) \leqslant x_1$ при любом t. Проведем доказательство по индукции. При $t = 1$

$$\mathsf{P}(\mu_1 = 0) = p_0 \leqslant p_0 + p_1 x_1 + p_2 x_1^2 + \ldots = \varphi(x_1) = x_1,$$

так как x_1 — корень уравнения (4.3). Предположим, что $\mathsf{P}(\mu_t = 0) \leqslant x_1$. Тогда, воспользовавшись (4.5), получим

$$\mathsf{P}(\mu_{t+1} = 0) = \varphi(\mathsf{P}(\mu_t = 0)) \leqslant \varphi(x_1) = x_1.$$

Таким образом, при любом t

$$\mathsf{P}(\mu_t = 0) \leqslant x_1.$$

Переходя к пределу в этом неравенстве, получим $\lambda \leqslant x_1$. Отсюда следует, что $\lambda = x_1$, так как λ и x_1 — корни уравнения (4.3) и x_1 — наименьший корень.

Таким образом, в случае надкритических процессов $\lambda = P(C) < 1$, и, следовательно, надкритические процессы являются вырождающимися.

Для вырождающихся процессов можно ввести случайную величину τ — время до вырождения процесса. Функция распределения τ есть $\mathsf{P}(\tau < t) = \mathsf{P}(\mu_t = 0)$. Найдем асимптотическую формулу при $t \to \infty$ для

$$Q(t) = \mathsf{P}(\tau \geqslant t) = 1 - \mathsf{P}(\mu_t = 0).$$

Уравнение (4.5) для $Q(t)$ запишется в виде

$$Q(t + 1) = 1 - \varphi(1 - Q(t)). \tag{4.7}$$

Рассмотрим сначала докритический случай. Для докритических процессов $a = \varphi'(1) < 1$. Предположим еще, что конечен второй факториальный момент

$$b = \mathsf{M}\mu_1(\mu_1 - 1) = \varphi''(1). \tag{4.8}$$

Используя формулу Тейлора, запишем (4.7) в виде

$$Q(t+1) = 1 - \varphi(1) + \varphi'(1)Q(t) - \\ - \frac{Q^2(t)}{2}\varphi''(1-\theta Q(t)), \quad 0<\theta<1.$$

Отсюда

$$Q(t+1) = aQ(t) - \frac{b_t}{2}Q^2(t), \tag{4.9}$$

где

$$b_t = \varphi''(1-\theta Q(t)), \ 0<\theta<1, \tag{4.10}$$

и

$$\frac{Q(s+1)}{Q(s)} = a\left(1 - \frac{b_s}{2a}Q(s)\right), \quad s = 1, 2, \ldots$$

Перемножив эти равенства от $s = t_0$ до $s = t-1$ $(0 < t_0 < t-1)$, получим

$$Q(t) = K_t a^t,$$

где

$$K_t = Q(t_0)\prod_{s=t_0}^{t-1}\left(1 - \frac{b_s}{2a}Q(s)\right). \tag{4.11}$$

Так как

$$Q(s) = \mathsf{P}\left(\mu_s \geqslant \frac{1}{2}\right) \leqslant 2\mathsf{M}\mu_s = 2a^s, \quad 0<a<1, \tag{4.12}$$

и $b_s \to b$ при $s \to \infty$, то при достаточно большом t_0 сомножители в (4.11) положительны. Из неравенства (4.12) следует сходимость ряда

$$\sum_{s=t_0}^{\infty}\ln\left(1 - \frac{b_s}{2a}Q(s)\right) = \ln K, \quad 0<K<\infty,$$

и оценка

$$\sum_{s=t}^{\infty}\ln\left(1 - \frac{b_s}{2a}Q(s)\right) = O(a^t), \quad t\to\infty.$$

Отсюда при $t \to \infty$

$$K_t = K(1 + O(a^t)).$$

Таким образом, *если* $0<a<1$ *и конечен момент* (4.8), *то при* $t\to\infty$

$$Q(t) = Ka^t(1 + O(a^t)).$$

Рассмотрим критический процесс. В этом случае $a = 1$, $b > 0$ и уравнение (4.9) запишется в виде

$$Q(t+1) = Q(t) - \frac{b_t}{2} Q^2(t), \tag{4.13}$$

где $b_t \to b$ при $t \to \infty$. Отсюда

$$\frac{Q(t+1)}{Q(t)} = 1 - \frac{b_t}{2} Q(t) \to 1, \qquad t \to \infty. \tag{4.14}$$

Равенство (4.13) запишем в виде

$$Q(t+1) = Q(t) - \frac{b}{2} Q(t) Q(t+1) + \varepsilon(t), \tag{4.15}$$

где

$$\varepsilon(t) = \frac{b}{2} Q(t) Q(t+1) - \frac{b_t}{2} Q^2(t).$$

Используя (4.14), при $t \to \infty$ получим

$$\frac{\varepsilon(t)}{Q(t) Q(t+1)} = \frac{b}{2} - \frac{b_t}{2} \frac{Q(t)}{Q(t+1)} \to 0. \tag{4.16}$$

Отсюда и из (4.15) вытекает, что

$$\frac{1}{Q(s+1)} = \frac{1}{Q(s)} + \frac{b}{2} + \delta(s),$$

где $\delta(s) \to 0$ при $s \to \infty$.

Суммируя последнее равенство от $s = 1$ до $s = t - 1$, при $t \to \infty$ получим

$$\frac{1}{Q(t)} = 1 + \frac{bt}{2} + \sum_{s=1}^{t-1} \delta(s) = \frac{bt}{2}\left(1 + \frac{2}{bt} + \frac{2}{bt} \sum_{s=1}^{t-1} \delta(s)\right) = \frac{bt}{2}(1 + o(1)).$$

Таким образом, *для критических ветвящихся процессов с конечным моментом* (4.8) *при* $t \to \infty$ *имеем*

$$Q(t) = \frac{2}{bt}(1 + o(1)).$$

Мы рассматривали ветвящийся процесс, начинающийся с одной частицы. Общее число частиц μ_t в процессе, начавшемся с n частиц, можно представить в виде

$$\mu_t = \mu_t^{(1)} + \mu_t^{(2)} + \ldots + \mu_t^{(n)}, \quad \mu_0 = n,$$

где случайные величины $\mu_t^{(k)}$, $k = 1, \ldots, n$ независимы и имеют такое же распределение, как рассмотренная выше величина μ_t с $\mu_0 = 1$. Величину $\mu_t^{(k)}$ можно интерпретировать как число потомков k-й начальной частицы в момент t. Нетрудно найти $\mathsf{M}\mu_t$:

$$\mathsf{M}\mu_t = \mathsf{M}\mu_t^{(1)} + \ldots + \mathsf{M}\mu_t^{(n)} = n\mathsf{M}\mu_t^{(1)} = na^t,$$

где

$$a = \mathsf{M}\mu_1^{(1)} = \varphi'(1).$$

Событие $(\mu_t > 0)$ можно представить в виде

$$(\mu_t > 0) = \bigcup_{k=1}^{n} (\mu_t^{(k)} > 0) = \overline{\bigcap_{k=1}^{n} (\mu_t^{(k)} = 0)}.$$

Отсюда, учитывая независимость величин $\mu_t^{(k)}$, $k = 1, \ldots, n$, находим

$$\mathsf{P}(\mu_t > 0) = 1 - \prod_{k=1}^{n} \mathsf{P}(\mu_t^{(k)} = 0) = 1 - (1 - Q(t))^n,$$

где

$$Q(t) = \mathsf{P}(\mu_t^{(1)} > 0).$$

Таким образом, зная поведение μ_t с $\mu_0 = 1$, нетрудно исследовать μ_t с $\mu_0 = n$.

Мы рассмотрели некоторые свойства наиболее простого ветвящегося процесса. Систематическое изложение теории ветвящихся процессов дается в книге [15].

§ 5. Процессы гибели и размножения

Случайные процессы ξ_t, рассмотренные в §§ 2—4, обладают одним общим свойством: при любом n и любых $t_1 < t_2 < \ldots < t_n < t_{n+1}$ равенство

$$\mathsf{P}(\xi_{t_{n+1}} \in B_{n+1} \mid \xi_{t_1} \in B_1, \ldots, \xi_{t_{n-1}} \in B_{n-1}, \xi_{t_n} = x) = \\ = \mathsf{P}(\xi_{t_{n+1}} \in B_{n+1} \mid \xi_{t_n} = x) \qquad (5.1)$$

выполняется для любых событий $\xi_{t_1} \in B_1, \ldots, \xi_{t_{n-1}} \in B_{n-1}$, $\xi_{t_n} = x$. Это свойство называют *марковским*; (5.1) является обобщением свойства (8.1.1), сформулированного для целочисленных моментов времени и подмножеств B_i множества натуральных чисел. Процессы, удовлетворяющие (5.1), называют *марковскими*. Марковский процесс ξ_t называется *процессом гибели и размножения*, если при $h \to 0$ выполняются условия

$$
\begin{aligned}
&\mathsf{P}\,(\xi_{t+h} = k+1 \mid \xi_t = k) = \lambda_k h + o\,(h),\\
&\mathsf{P}\,(\xi_{t+h} = k-1 \mid \xi_t = k) = \mu_k h + o\,(h),\\
&\mathsf{P}\,(\xi_{t+h} = k \mid \xi_t = k) = 1 - (\lambda_k + \mu_k)\,h + o\,(h),
\end{aligned}
\tag{5.2}
$$

где

$$\lambda_k,\ \mu_k \geqslant 0\ (k > 0),\quad \mu_k = 0\ (k \leqslant 0).$$

Получим теперь, аналогично тому как это делалось в § 3 гл. 3, систему дифференциальных уравнений для вероятностей $P_k\,(t)$ того, что в момент t состоянием процесса является k. Для этого запишем вероятность события $\{\xi_{t+h} = k\}$ по формуле полной вероятности, выбирая в качестве системы несовместных событий состояния процесса в момент t. В результате при $k \geqslant 1$ получим

$$
\begin{aligned}
P_k\,(t+h) = P_k\,(t)\,(1 - (\lambda_k + \mu_k)\,h) +{}&\\
&+ \lambda_{k-1} h P_{k-1}\,(t) + \mu_{k+1} h P_{k+1}\,(t) + o\,(h).
\end{aligned}
$$

Перенеся $P_k\,(t)$ в левую часть этого равенства и поделив обе части полученного соотношения на h, при $h \to 0$ находим

$$P_k'\,(t) = -(\lambda_k + \mu_k)\,P_k\,(t) + \lambda_{k-1} P_{k-1}\,(t) + \mu_{k+1} P_{k+1}\,(t). \tag{5.3}$$

Аналогичные выкладки показывают, что

$$P_0'\,(t) = -\lambda_0 P_0\,(t) + \mu_1 P_1\,(t). \tag{5.4}$$

Если задать начальное распределение $\{P_k\,(0)\}$, то бесконечная система дифференциальных уравнений (5.3)—(5.4) имеет единственное решение лишь в случае, когда λ_k, μ_k ограничены или возрастают достаточно медленно (см. [18], гл. XVII, §§ 4, 5). В этом случае $\sum\limits_{k=0}^{\infty} P_k\,(t) = 1$, а марковский процесс ξ_t, удовлетворяющий условиям (5.2) и имеющий заданное начальное распределение вероятностей $\mathsf{P}\,\{\xi_0 = k\} = P_k\,(0)$, $k = 0,\ 1,\ 3,\ \ldots$, является единственным.

Рассмотрим теперь несколько частных видов общего процесса гибели и размножения.

5.1. Процесс чистого размножения. Пусть $\mu_k = 0$, $k = 0, 1, 2, \ldots$ При этом условии переходы из состояния k

возможны только в состояние $k+1$. Вероятности выхода $\lambda_k h + o(h)$ $(h \to 0)$ из данного состояния k в процессе чистого размножения зависят в общем случае от номера состояния. Система уравнений (5.3)—(5.4) для процесса чистого размножения имеет вид

$$P_0'(t) = -\lambda_0 P_0(t),$$

$$P_k'(t) = -\lambda_k P_k(t) + \lambda_{k-1} P_{k-1}(t), \quad k \geqslant 1.$$

Процесс чистого размножения с одинаковыми λ_k является процессом Пуассона.

5.2. Система массового обслуживания с потерями. Пусть требования, поступающие на n обслуживающих устройств, образуют пуассоновский процесс с параметром λ. Время обслуживания любого требования любым устройством имеет показательное распределение с параметром μ и не зависит от работы других обслуживающих устройств и от поступающих новых требований. Если все устройства заняты, то вновь поступающее требование теряется. Число требований в системе ξ_t в момент t может принимать только значения $0, 1, \ldots, n$. Вероятность перехода системы из состояния k в состояние k за время $h \to 0$ отличается на $o(h)$ от вероятности произведения независимых событий: {за время h не закончится обслуживание ни одного из k требований} $\cap$ {за время h не поступит новых требований}. Следовательно, $\mathsf{P}\{\xi_{t+h} = k \mid \xi_t = k\} = (1 - \mu h + o(h))^k \ (1 - \lambda h + o(h)) + o(h)$, или $\mathsf{P}\{\xi_{t+h} = k \mid \xi_t = k\} = 1 - (k\mu + \lambda)h + o(h)$, $0 < k < n$. Аналогично находим

$$\mathsf{P}\{\xi_{t+h} = k+1 \mid \xi_t = k\} = \lambda h + o(h),$$

$$\mathsf{P}\{\xi_{t+h} = k-1 \mid \xi_t = k\} = k\mu h + o(h), \quad 0 < k < n.$$

Так же рассматриваются переходы из состояний $k = 0$ и $k = n$. Сравнив вычисленные вероятности перехода с (5.2), получим, что рассматриваемый процесс является процессом гибели и размножения с $\lambda_k = \lambda$ $(k = 0, 1, \ldots, n-1)$, $\lambda_k = 0$ $(k \geqslant n)$, $\mu_k = k\mu$ $(1 \leqslant k \leqslant n)$, $\mu_k = 0$ $(k > n$, $k = 0)$. Для этого процесса система (5.3)—(5.4) имеет вид

$$P_0'(t) = -\lambda P_0(t) + \mu P_1(t),$$

$$P_k'(t) = -(\lambda + k\mu) P_k(t) + \lambda P_{k-1}(t) + (k+1)\mu P_{k+1}(t), \quad 0 < k < n,$$

$$P_n'(t) = -(\lambda + n\mu) P_n(t) + \lambda P_{n-1}(t).$$

Непосредственно проверяется, что эта система имеет следующее стационарное решение:

$$P_0=\left(\sum_{l=0}^{n}\frac{1}{l!}\,\theta^l\right)^{-1}, \qquad P_k=\frac{1}{k!}\,\theta^k P_0, \qquad 0\leqslant k\leqslant n, \tag{5.5}$$

где $\theta = \lambda/\mu$. Формулы (5.5) называют *формулами Эрланга*. Полагая в (5.5) $k = n$, получим формулу для вероятности потери требования. Эта вероятность является важной характеристикой системы обслуживания.

5.3. Ветвящийся процесс. Один частный случай ветвящегося процесса с непрерывным временем является процессом гибели и размножения. Предположим, что каждая частица независимо от своего прошлого и от поведения других частиц за время h с вероятностью $\lambda h + o(h)$ делится на две, с вероятностью $\mu h + o(h)$ погибает и с вероятностью $1 - (\lambda + \mu)\, h + o(h)$ не изменяется. Пусть ξ_t — число частиц в момент времени t. Непосредственно проверяется, что при $h \to 0$

$$\mathsf{P}\,(\xi_{t+h} = k+1 \mid \xi_t = k) = \lambda kh + o\,(h),$$

$$\mathsf{P}\,(\xi_{t+h} = k-1 \mid \xi_t = k) = \mu kh + o\,(h),$$

$$\mathsf{P}\,(\xi_{t+h} = k \mid \xi_t = k) = 1 - (\lambda+\mu)\,kh + o\,(h),\ k>0.$$

Рассматриваемый ветвящийся процесс является процессом гибели и размножения с $\lambda_k = k\lambda\ (k\geqslant 0)$, $\mu_k = k\mu\ (k\geqslant 0)$. Решение $P_k(t) = \mathsf{P}\{\xi_t = k \mid \xi_0 = 1\}$ системы уравнений (5.3)—(5.4) с начальными условиями $\mathsf{P}\{\xi_0 = 1\} = 1$, $\mathsf{P}\{\xi_0 = k\} = 0\ (k \geqslant 1)$ определяется формулами

$$P_0(t) = 1 - \frac{e^{at}}{\lambda a^{-1}(e^{at}-1)+1} \qquad (a\neq 0),$$

$$P_0(t) = 1 - (1+\lambda t)^{-1} \qquad (a=0),$$

$$P_k(t) = \frac{1}{\frac{\lambda}{a}\left[(e^{at}-1)+1\right]^2}\left[\frac{\frac{\lambda}{a}(e^{at}-1)}{1-\frac{\lambda}{a}(e^{at}-1)}\right]^{k-1} \qquad (a\neq 0,\ k\geqslant 1),$$

$$P_k(t) = \left(\frac{1}{1+\lambda t}\right)^2\left(\frac{\lambda t}{1+\lambda t}\right)^{k-1} \qquad (a=0,\ k\geqslant 1),$$

где $a = \lambda - \mu$.

Задачи к главе 10

1. Найти $\operatorname{cov}(\xi_t, \xi_{t+s})$, если: а) ξ_t — пуассоновский процесс с параметром λ; б) ξ_t — винеровский процесс с $a = 0$.

Указание. Воспользоваться независимостью и однородностью приращений.

2. Доказать, что для любого промежутка времени $\tau_k = \theta_k - \theta_{k-1}$, где θ_k — момент k-го скачка процесса Пуассона, выполняется равенство $\mathsf{P}\{\tau_k > t + s \mid \tau_k > s\} = \mathsf{P}\{\tau_k > t\}$.

3. Пусть ν, $\eta_1, \eta_2, \ldots$ — независимые случайные величины; $\mathsf{P}(\nu = k) = (\lambda T)^k e^{-\lambda T}/k!$; величины η_l $(l = 1, 2, \ldots)$ равномерно распределены на отрезке $[0, T]$. Обозначим ξ_t число величин η_l, удовлетворяющих неравенству $\eta_l < t$, $l = 1, 2, \ldots, \nu$, если $\nu > 0$, и $\xi_t = 0$ при $\nu = 0$. Найти вероятность

$$\mathsf{P}\{\xi_{t_1} = k_1,\ \xi_{t_2} - \xi_{t_1} = k_2 - k_1,\ \xi_T - \xi_{t_2} = k_3 - k_2\}.$$

4. Пусть μ_t — ветвящийся процесс с дискретным временем; $\varphi(x) = \mathsf{M}x^{\mu_1} = px^2 + 1 - p$. Найти: а) $\lambda = \lim\limits_{t\to\infty} \mathsf{P}\{\mu_t = 0\}$, $p > 1/2$; б) $B_t = \mathsf{M}\mu_t(\mu_t - 1)$ и $\mathsf{D}\mu_t$ при $p = 1/2$.

5. Пусть μ_t — ветвящийся процесс с непрерывным временем, определенный в § 5. Найти: а) дифференциальное уравнение для $F(t, x) = \mathsf{M}\{x^{\mu_t} \mid \mu_0 = 1\}$; б) явное выражение для $F(t, x)$; в) $A(t) = \mathsf{M}\{\mu_t \mid \mu_0 = 1\}$.

Указание. Использовать формулу полного математического ожидания, рассмотрев возможные превращения начальной частицы за промежуток времени (o, h).

6. Случайные величины ξ_t $(t = 1, 2, \ldots)$ независимы и $\mathsf{P}(\xi_t = 1) = p$, $\mathsf{P}(\xi_t = -1) = 1 - p = q$. Положим

$$\eta_{t+1} = \begin{cases} \eta_t + \xi_{t+1}, & \text{если } \eta_t \neq 0, \\ a, & \text{если } \eta_t = 0,\ t = 0, 1, 2, \ldots; \end{cases}$$

$\eta_0 = k$. Найти: 1) вероятность того, что процесс η_t когда-либо попадет в состояние 0, если $a = 0$; 2) стационарное распределение вероятностей $\pi_l = \lim\limits_{t\to\infty} \mathsf{P}(\eta_t = l)$, если $a = 1$, $q > p$.

7. Пусть ξ_t — процесс гибели и размножения. Положим $m_k = \mathsf{M}(\tau \mid \xi_0 = k)$, $0 \leqslant k \leqslant n$, где τ — время до первого достижения состояния n. Найти: 1) математическое ожидание времени пребывания в состоянии k до первого выхода из него; 2) вероятность того, что первым переходом из k будет переход в $k + 1$ $(k - 1)$. Составить систему уравнений для m_k.

8. Марковский случайный процесс ξ_t с непрерывным временем называют цепью Маркова, если множество его состояний конечно или счетно. Пусть ξ_t — цепь Маркова с двумя состояниями (1 и 2) и при $h \to 0$

$$\mathsf{P}(\xi_h = 2 \mid \xi_0 = 1) = \alpha h + o(h),\quad \mathsf{P}(\xi_h = 1 \mid \xi_0 = 2) = \beta h + o(h).$$

Найти $P_{ij}(t) = \mathsf{P}(\xi_t = j \mid \xi_0 = i)$, $i, j = 1, 2$.

9. Пусть $\tau_k(t)$, $k = 1, 2,$ — суммарная длительность пребывания в состоянии k за время t цепи Маркова, определенной в за-

даче 8. Составить дифференциальное уравнение для

$$f_k(t, s) = \mathsf{M}\left(e^{is(\tau_1(t)-\tau_2(t))} \mid \xi_0 = k\right).$$

10. Найти $\lim\limits_{t\to\infty} \frac{1}{t} \mathsf{M}\{\tau_1(t) \mid \xi_0 = k\}$, где $\tau_1(t)$ определено в задаче 9.

Указание. Использовать формулу $\tau_1(t) = \int\limits_0^t \delta_{1,\,\xi_u}\,du$, где $\delta_{1,1} = 1$, $\delta_{1,2} = 0$.

11. Движение точки по прямой управляется цепью Маркова ξ_t, определенной в задаче 8. Если $\xi_t = 1$, то точка движется вправо со скоростью v_1, а если $\xi_t = 2$, то влево со скоростью v_2. Пусть η_t — координата точки в момент t. Найти $\lim\limits_{t\to\infty} \frac{1}{t} M_k(t, x)$, где

$$M_k(t, x) = \mathsf{M}\{\eta_t \mid \eta_0 = x, \xi_0 = k\}, \quad k = 1, 2.$$

Указание. Составить дифференциальное уравнение для $M_k(t, x)$; использовать равенство $M_k(t, x) = x + M_k(t, 0)$.

12. Решить задачу 11, используя равенство $\eta_t = v_1\tau_1(t) - v_2\tau_2(t) + x$, где $\tau_k(t)$ определены в задаче 9, и решение задачи 10.

13. Движение точки по прямой управляется цепью Маркова ξ_t, определенной в задаче 8. Если $\xi_t = 1$, то точка движется вправо со скоростью v, а при $\xi_t = 2$ точка движется влево с постоянным ускорением a (при начале движения с ускорениям начальная скорость считается равной 0). Пусть η_t — координата точки в момент t. Является ли процесс η_t марковским? Составить интегральное уравнение дл $M_k(t, x) = \mathsf{M}\{\eta_t \mid \eta_0 = x, \xi_0 = k\}$, $k = 1, 2$. Найти $\lim\limits_{t\to\infty} \frac{1}{t} M_k(t, x)$.

Указание. Использовать равенство $M_k(t, x) = x + M_k(t, 0)$.

14. Пусть ξ_t $(t = \ldots -2, -1, 0, 1, 2, \ldots)$ — последовательность независимых случайных величин с $\mathsf{M}\xi_t = a$, $\mathsf{D}\xi_t = \sigma^2$. Положим $\eta_t = c_0\xi_t + c_1\xi_{t-1} + c_2\xi_{t-2}$, $c_0 + c_1 + c_2 = 1$. Найти $\mathsf{M}\eta_t$, $\mathsf{D}\eta_t$, $\operatorname{cov}(\eta_{t_1}, \eta_{t_2})$. Проверить, что $\operatorname{cov}(\eta_{t_1}, \eta_{t_2}) = R(|t_1 - t_2|)$.

15. Пусть η_t — процесс, определенный в задаче 14. Положим $a_t^* = (\eta_1 + \eta_2 + \ldots + \eta_t)/t$. Найти $\mathsf{M}a_t^*$ и $\mathsf{D}a_t^*$ Является ли a_t^* несмещенной и состоятельной оценкой a?

16. Пусть η_t — процесс, определенный в задаче 14. Показать, что $R_t^*(k) = \frac{1}{t}\sum\limits_{s=1}^{t}(\eta_s - a_t^*)(\eta_{s+k} - a_t^*)$, $k = 0, 1, 2$, является асимптотически несмещенной оценкой $R(k)$.

17. При $t = 1, 2, \ldots, 100$ получить реализацию процесса η_t, определенного в задаче 14, если ξ_t распределены равномерно на отрезке $[-1, 1]$. Значения получить с точностью до 0,01. По реализации вычислить оценки a_t^* и $R_t^*(k)$, определенные в задачах 15, 16. Сравнить с теоретическими.

Таблицы

Распределение Пуассона

Значения функции $p_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$

Таблица 1

k \ λ	0,1	0,2	0,3	0,4	0,5
0	0,90484	0,81873	0,74082	0,67032	0,60653
1	0,09048	0,16375	0,22225	0,26813	0,30327
2	0,00452	0,01638	0,03334	0,05363	0,07582
3	0,00015	0,00109	0,00333	0,00715	0,01264
4		0,00006	0,00025	0,00072	0,00158
5			0,00002	0,00006	0,00016
6					0,00001

k \ λ	0,6	0,7	0,8	0,9
0	0,54881	0,49659	0,44933	0,40657
1	0,32929	0,34761	0,35946	0,36591
2	0,09879	0,12166	0,14379	0,16466
3	0,01976	0,02839	0,03834	0,04940
4	0,00296	0,00497	0,00767	0,01112
5	0,00036	0,00070	0,00123	0,00200
6	0,00004	0,00008	0,00016	0,00030
7		0,00001	0,00002	0,00004

k \ λ	1,0	2,0	3,0	4,0	5,0
0	0,36788	0,13534	0,04979	0,01832	0,00674
1	0,36788	0,27067	0,14936	0,07326	0,03369
2	0,18394	0,27067	0,22404	0,14653	0,08422
3	0,06131	0,18045	0,22404	0,19537	0,14037
4	0,01533	0,09022	0,16803	0,19537	0,17547
5	0,00307	0,03609	0,10082	0,15629	0,17547
6	0,00051	0,01203	0,05041	0,10419	0,14622
7	0,00007	0,00344	0,02160	0,05954	0,10445
8	0,00001	0,00086	0,00810	0,02977	0,06528
9		0,00019	0,00270	0,01323	0,03627
10		0,00004	0,00081	0,00529	0,01813
11		0,00001	0,00022	0,00193	0,00824
12			0,00006	0,00064	0,00343
13			0,00001	0,00020	0,00132
14				0,00006	0,00047
15				0,00002	0,00016
16					0,00005
17					0,00001

Нормальное распределение

Значения функции $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}}\,du$

Таблица 2

∞	сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0200	0,0239	0,0279	0,0319	0,0359
0,1	398	438	478	517	557	596	636	675	714	753
0,2	793	832	871	910	948	987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	406	443	480	517
0,4	554	591	628	664	700	736	772	808	844	879
0,5	915	950	985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	357	389	422	454	486	517	549
0,7	580	611	642	673	703	734	764	794	823	852
0,8	881	910	939	967	995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	289	315	340	365	389
1,0	413	437	461	485	508	531	554	577	599	621
1,1	643	665	686	708	729	749	770	790	810	830
1,2	849	869	888	907	925	944	962	980	997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	192	207	222	236	251	265	279	292	306	319
1,5	332	345	357	370	382	394	406	418	429	441
1,6	452	463	474	484	495	505	515	525	535	545
1,7	554	564	573	582	591	599	608	616	625	633
1,8	641	649	656	664	671	678	686	693	699	706
1,9	713	719	726	732	738	744	750	756	761	767

Продолжение табл. 2

x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	821	826	830	834	838	842	846	850	854	857
2,2	861	864	868	871	875	878	881	884	887	890
2,3	893	896	898	901	904	906	909	911	913	916
2,4	918	920	922	925	927	929	931	932	934	936
2,5	938	940	941	943	945	946	948	949	951	952
2,6	953	955	956	957	959	960	961	962	963	964
2,7	965	966	967	968	969	970	971	972	973	974
2,8	974	975	976	977	977	978	979	979	980	981
2,9	981	982	982	983	984	984	985	985	985	986
3,0	987	987	987	988	988	989	989	989	990	990

Значения функции u_α

Функция u_α определяется равенством $\alpha = \frac{1}{\sqrt{2\pi}} \int_{u_\alpha}^{\infty} e^{-\frac{x^2}{2}} dx$

Таблица 3

α	0,001	0,005	0,010	0,015	0,020	0,025	0,030	0,035	0,040	0,045	0,050
u_α	3,0902	2,5758	2,3263	2,1701	2,0537	1,9600	1,8808	1,8119	1,7507	1,6954	1,6449

Распределение Стьюдента

Значения функции $t_{\alpha, n}$

Функция $t_{\alpha, n}$ определяется равенством $\mathsf{P}\,(|\tau_n| < t_{\alpha, n}) = 1 - 2\alpha$, где случайная величина τ_n имеет распределение Стьюдента с n степенями свободы. Плотность распределения τ_n равна

$$p_{\tau_n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \cdot \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

Таблица 4

n \ 2α	0,10	0,05	0,02	0,01
5	2,015	2,571	3,365	4,032
6	1,943	2,447	3,143	3,707
7	1,895	2,365	2,998	3,499
8	1,860	2,306	2,896	3,355
9	1,833	2,262	2,821	3,250
10	1,812	2,228	2,764	3,169
12	1,782	2,179	2,681	3,055
14	1,761	2,145	2,624	2,977
16	1,746	2,120	2,583	2,921
18	1,734	2,101	2,552	2,878
20	1,725	2,086	2,528	2,845
22	1,717	2,074	2,508	2,819
30	1,697	2,042	2,457	2,750
∞	1,645	1,960	2,326	2,576

χ^2-распределение

Значения функции $\chi^2_{\alpha, m}$

Функция $\chi^2_{\alpha, m}$ определяется равенством $\mathsf{P}\left(\chi^2_m > \chi^2_{\alpha, m}\right) = \alpha$, где случайная величина χ^2_m имеет χ^2-распределение с m степенями свободы. Плотность распределения χ^2_m равна

$$p_{\chi^2_m}(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{n/2}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \quad x > 0.$$

Таблица 5

m \ α	0,99	0,10	0,05	0,02	0,01	0,005
1	0,00016	2,7	3,8	5,4	6,6	7,9
2	0,020	4,6	6,0	7,8	9,2	11,6
3	0,115	6,3	7,8	9,8	11,3	12,8
4	0,30	7,8	9,5	11,7	13,3	14,9
5	0,55	9,2	11,1	13,4	15,1	16,3
6	0,87	10,6	12,6	15,0	16,8	18,6
7	1,24	12,0	14,1	16,6	18,5	20,3
8	1,65	13,4	15,5	18,2	20,1	21,9
9	2,09	14,7	16,9	19,7	21,7	23,6
10	2,56	16,0	18,3	21,2	23,2	25,2
11	3,1	17,3	19,7	22,6	24,7	26,8
12	3,6	18,5	21,0	24,1	26,2	28,3
13	4,1	19,8	22,4	25,5	27,7	29,8
14	4,7	21,1	23,7	26,9	29,1	31
15	5,2	22,3	25,0	28,3	30,6	32,5
16	5,8	23,5	26,3	29,6	32,0	34
17	6,4	24,8	27,6	31,0	33,4	35,5
18	7,0	26,0	28,9	32,3	34,8	37
19	7,6	27,2	30,1	33,7	36,2	38,5
20	8,3	28,4	31,4	35,0	37,6	40
21	8,9	29,6	32,7	36,3	38,9	41,5
22	9,5	30,8	33,9	37,7	40,3	42,5
23	10,2	32,0	35,2	39,0	41,6	44,0
24	10,9	33,2	36,4	40,3	43,0	45,5
25	11,5	34,4	37,7	41,6	44,3	47

F-распределение

Значения функции $F_{\alpha;\, n_1,\, n_2}$

Функция $F_{\alpha;\, n_1,\, n_2}$ определяется равенством $\mathsf{P}\{F_{n_1,\, n_2} > F_{\alpha;\, n_1,\, n_2}\} = \alpha$, где случайная величина $F_{n_1,\, n_2}$ имеет F-распределение с n_1 и n_2 степенями свободы. Плотность распределения $F_{n_1,\, n_2}$ равна

$$p_{F_{n_1,\, n_2}}(x) = \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \frac{x^{\frac{n_1}{2}-1}}{\left(\frac{n_1}{n_2}x+1\right)^{\frac{n_1+n_2}{2}}}, \quad x > 0.$$

$$\alpha = 0{,}05$$

Таблица 6

$n_2 \backslash n_1$	10	20	30	40	50	100	∞
10	2,97	2,77	2,70	2,67	2,64	2,59	2,54
15	2,55	2,33	2,25	2,21	2,18	2,12	2,07
20	2,35	2,12	2,04	1,99	1,96	1,90	1,84
30	2,16	1,93	1,84	1,79	1,76	1,69	1,62
40	2,07	1,84	1,74	1,69	1,66	1,59	1,51
50	2,02	1,78	1,69	1,63	1,60	1,52	1,44
100	1,92	1,68	1,57	1,51	1,48	1,39	1,28
∞	1,83	1,57	1,46	1,40	1,35	1,24	1,00

Случайные числа

Равномерно распределенные случайные числа

Приведенные в таблице цифры можно рассматривать как реализации независимых и одинаково распределенных случайных величин, принимающих значения 0, 1, ..., 9 с одной и той же вероятностью, равной 0,1.

Таблица 7

10	09	73	25	33	76	52	01	35	86	34	67	35	48	76
37	54	20	48	05	64	89	47	42	96	24	80	52	40	37
08	42	26	89	53	19	64	50	93	03	23	20	90	25	60
99	01	90	25	29	09	37	67	07	15	38	31	13	11	65
12	80	79	99	70	80	15	73	61	47	64	03	23	66	53

Продолжение табл. 7

80	95	90	91	17	39	29	27	49	45	66	06	57	47	17
20	63	61	04	02	00	82	29	16	65	31	06	01	08	05
15	95	33	47	64	35	08	03	36	06	85	26	97	76	02
88	67	67	43	97	04	43	62	76	59	63	57	33	21	35
98	95	11	68	77	12	17	17	68	33	73	79	64	57	53
34	07	27	68	50	36	69	73	61	70	65	81	33	98	85
45	57	18	24	06	35	30	34	26	14	86	79	90	74	39
02	05	16	56	92	68	66	57	48	18	73	05	38	52	47
05	32	54	70	48	90	55	35	75	48	28	46	82	87	09
03	52	96	47	78	35	80	83	42	82	60	93	52	03	44
11	19	92	91	70	98	52	01	77	67	14	90	56	86	07
23	40	30	97	32	11	80	50	54	31	39	80	82	77	32
18	62	38	85	79	83	45	29	96	34	06	28	89	80	83
83	49	12	56	24	88	68	54	02	00	86	50	75	84	01
35	27	38	84	35	99	59	46	73	48	87	51	76	49	69

Нормально распределенные случайные числа

Приведенные в таблице числа можно рассматривать как реализации независимых нормально распределенных величин с параметрами $a = 0$, $\sigma^2 = 1$.

Таблица 8

0,464	0,137	2,455	−0,323	−0,068	0,296	−0,288	1,298
0,060	−2,256	−0,531	−0,194	0,543	−1,558	0,187	−1,190
1,486	−0,354	−0,634	0,697	0,926	1,375	0,785	−0,963
1,022	−0,472	1,279	3,521	0,571	−1,851	0,194	1,192
1,394	−0,555	0,046	0,321	2,945	1,974	−0,258	0,412
0,906	−0,513	−0,525	0,595	0,881	−0,934	1,579	0,161
1,179	−1,055	0,007	0,769	0,971	0,712	1,090	−0,631
−1,501	−0,488	−0,162	−0,136	1,033	0,303	0,448	0,748
−0,690	0,756	−1,618	−0,345	−0,511	−2,051	−0,457	−0,218
1,372	0,225	0,378	0,761	0,181	−0,736	0,960	−1,530
−0,482	1,678	−0,057	−1,229	−0,486	0,856	−0,491	−1,983
−1,376	−0,150	1,356	−0,561	−0,256	−0,212	0,219	0,779
−1,010	0,598	−0,918	1,598	1,065	0,415	−0,169	0,313
−0,005	−0,899	0,012	−0,725	0,147	−0,121	1,096	0,181
1,393	−1,163	−0,911	1,231	−0,199	−0,246	1,239	−2,574

Ответы к задачам

ГЛАВА 1

2. 1), 4) — 6) неверны; 2), 3), 5) верны.
3. 1) A; 2) A; 3) $\varnothing$; 4) AB.
4. $A = \{a\}$, $B = \{x\colon a \leqslant x < b\}$.
5. 1) {ГГГ, ГРГ, ГГР, ГРР}; 2) {ГГР, ГРГ, РГГ}; 3) {PPP, ГРР, РГР, РРГ}.
6. 1) $A\bar{B}\bar{C}$; 2) $AB\bar{C}$; 3) ABC; 4) $A + B + C$; 5) $A\bar{B}\bar{C} + \bar{A}B\bar{C} + \bar{A}\bar{B}C$; 6) $\bar{A}\bar{B}\bar{C}$; 7) $\bar{A} + \bar{B} + \bar{C}$.
7. Да.

ГЛАВА 2

1. 1/6, 5/6, 11/36, 5/36. 2. $2/n$. 3. 1/6.
4. 0,01; 0,72; 0,27.
5. 1/2; $C_n^m (2/5)^m (3/5)^{n-m}$; $C_{n-2}^m (2/5)^{m+2} (3/5)^{n-m-2}$.
6. $\mathsf{P}(A) = 1 - (5/6)^4 = 0{,}5177\ldots > \mathsf{P}(B) = 1 - (35/36)^{24} = 0{,}4914\ldots$
7. а) $2^{2r} C_n^{2r}/C_{2n}^{2r}$; б) $n 2^{2r-2} C_{n-1}^{2r-2}/C_{2n}^{2r}$.
8. а) 0,4080; б) 0,3305. 9. 0,09.
10. 1) $\mathsf{P}(A) = 1/28985 = 0{,}0000345\ldots$, $\mathsf{P}(B) = 224/28985 = 0{,}0077281\ldots$; 2) $\mathsf{P}(A) = 8/15 = 0{,}5333\ldots$, $\mathsf{P}(B) = 7/15 = 0{,}4666\ldots$
11. $\pi/4 = 0{,}7853981\ldots$ 12. $12!/12^{12} = 0{,}0000537$.
13. 1) $1 - (n!/n^n)$; 2) $n!\, C_n^2/n^n$.
14. $61705/6991908 = 0{,}0088252\ldots$
15. $F'(x) = 1$ $(0 \leqslant x \leqslant 1)$.
16. $F'(x) = x$ $(0 \leqslant x \leqslant 1)$, $F'(x) = 2 - x$ $(1 \leqslant x \leqslant 2)$.
17. 1/4. 18. 1) $(1 - (2r/a))^2$; 2) $1 - (4r^2/a^2)$.
19. $1 - (1 - l/T)^2$.
20. а) 3/4; б) $\sqrt{3}/2 = 0{,}8660254\ldots$; в) 2/3.
21. $\mathsf{P}(A) = M/N$, $\mathsf{P}(B) = (M/N)^2$, $\mathsf{P}(C) = C_n^m (M/N)^m (1 - M/N)^{n-m}$.
22. $\mathsf{P}(A) = M/N$, $\mathsf{P}(B) = M(M-1)/N(N-1)$, $\mathsf{P}(C) = C_M^m C_{N-M}^{n-m}/C_N^n = C_n^m \dfrac{M^{[m]}(N-M)^{[n-m]}}{N^{[n]}}$.
23. а) 1/10; б) 1/100.

ГЛАВА 3

1. 1/12. 2. 0,61477...
3. 1/3. 4. 1) 1/4; 2) 3/5.

6. $\mathsf{P}(A_rB_r) = \mathsf{P}(A_r)\,\mathsf{P}(B_r)$ только при $r \leqslant 0$, $r \geqslant 2/3$ и $r = 1/3$.

7. C_1, C_2, C_3 не являются взаимно независимыми; C_1C_2 и C_3 зависимы.

9. 1) $p_1(1-p_3)p_4$; 2) $p_1+p_2-p_1p_2$; 3) $(p_1+p_2-p_1p_2)(p_3+p_4-p_3p_4)$.

10. $(1-p_3)(1-p_1p_2)$.

11. $\mathsf{P}(A_1) = \mathsf{P}(A_2) = 3/8$, $\mathsf{P}(A_1A_2) = 3/28$.

12. 1) $\alpha_1\alpha_2$; 2) $\beta_1+\beta_2-\beta_1\beta_2$. 13. 11/18.

14. 1) $p(1-\beta_1)(1-\beta_2)+(1-p)\alpha_1\alpha_2$;

$$2)\quad \frac{p(1-\beta_1)(1-\beta_2)}{p(1-\beta_1)(1-\beta_2)+(1-p)\alpha_1\alpha_2}.$$

15. 20/21. 16. 9/16. 17. $e^{-\lambda t}$.

18. $\pi_0 = \dfrac{\beta}{\alpha_1+\alpha_2+\beta}$, $\pi_1 = \dfrac{\alpha_1}{\alpha_1+\beta}$, $\pi_2 = \dfrac{\alpha_2}{\alpha_1+\beta}\pi_0$.

19. $\pi_0 = 1-\theta$, $\pi_k = (1-\theta)\,\theta^k$, $k \geqslant 1$.

ГЛАВА 4

1. 3/5.

2. $\mathsf{P}(A_1) = 44/105 = 0{,}4190476\ldots$, $\mathsf{P}(A_2) = 26/105 = 0{,}247619\ldots$, $\mathsf{P}(B) = 1/3$.

3. $\mathsf{P}(A_k) = M/N$, $\mathsf{P}(B_{k,l}) = M(M-1)/N(N-1)$, $\mathsf{P}(C_{k,l}) = M(N-M)/N(N-1)$.

4. $C_{10}^4(5/72)^4(1-5/72)^6 = 0{,}00317\ldots$

5. а) $0{,}95099\ldots$; б) $0{,}0480298\ldots$; в) $0{,}0009802\ldots$

6. а) 12; б) 22.

7. а) 5, $Q(5) = 0{,}67232\ldots$; б) 11, $Q(11) = 0{,}9141008\ldots$; в) 21, $Q(21) = 0{,}9907768\ldots$

8. 1) $Q(5) = 0{,}73626\ldots$, $Q(11) = 0{,}9912088\ldots$, $Q(21) = 1$; 2) $Q(5) = 0{,}6894374\ldots$, $Q(11) = 0{,}9381106\ldots$, $Q(21) = 0{,}9980502\ldots$, $Q(n) = 1-((4M)^{[n]}/(5M)^{[n]})$.

9. $C_{n-l}^m p^{m+l}(1-p)^{n-m-l}$. 10. $C_{2n}^n 2^{-2n}$.

11. а) $(1-p^2)^N$; б) $C_n^k p^{2k}(1-p^2)^{n-k}$; в) $(1-2^{-2n}C_{2n}^n)/2$.

12. 0,0588. 13. $393{,}75\,(1/\pi)^4(1-2/\pi)^6 = 0{,}0093063\ldots$

14. 1) $C_n^{m_1}p_1^{m_1}(1-p_1)^{n-m_1}$;

$$2)\quad \frac{(n-m_1)!}{m_2!\ldots m_N!}\left(\frac{p_2}{1-p_1}\right)^{m_2}\cdots\left(\frac{p_N}{1-p_N}\right)^{m_N}.$$

15. в) $np_4^{n-2}(n-1)\,p_1p_3+p_2p_4)$. 16. 0,32332.

17. 1) 0,6651, ..., 0,6321; 2) 0,40187..., 0,3679...; 3) 0,2009..., 0,1839...

18. 5. 19. 0,95957. 20. 0,99. 21. 547. 22. 0,0228.

23. 0,39347. 24. $(1/4)^{n-1}3/4$. 25. 1) 7/8; 2) 2/3. 26. 9/10.

ГЛАВА 5

1. $C = 3$; $p_\eta(x) = 3e^{-3x}$ $(x \geqslant 0)$; $\mathsf{P}\{0{,}5 < \eta < 0{,}75\} = e^{-1,5} - e^{-2,25} = 0{,}117731\ldots$

2. а) $1/2$ $(x \in [1, 3])$; б) e^{-x} $(x > 0)$.

3. а) $2\alpha x e^{-\alpha x^2}$ $(x > 0)$; б) $\alpha e^{-\alpha\sqrt{x}}/2\sqrt{x}$ $(x > 0)$; в) $\alpha^2 e^{-\alpha(e^{\alpha x} - x)}$ $(-\infty < x < \infty)$; г) 1 $(x \in [0, 1])$.

4. а) $\dfrac{1}{\sqrt{2\pi x}}\, e^{-x/2}$ $(x > 0)$; б) $\dfrac{1}{x\sqrt{2\pi}}\, e^{-\ln^2 x/2}$.

5. $p_\xi(x) = -\ln x$ $(0 < x < 1)$.

6. $p_\eta(x) = 1$ $(0 \leqslant x \leqslant 1)$.

7. а) $1 - |x - 1|$ $(0 \leqslant x \leqslant 2)$; б) $1 - |x|$ $(|x| \leqslant 1)$; в) $1/2$ $(0 \leqslant x \leqslant 1)$; $1/2x^2$ $(x \geqslant 1)$; 0 $(x < 0)$.

8. $\alpha^2 x e^{-\alpha x}$ $(x > 0)$.

9. а) $p_{\eta_1}(x) = \dfrac{1}{2} e^{-\frac{x}{2}}$ $(x > 0)$; б) $p_{\eta_2}(x) = \dfrac{1}{\pi}\left(-\dfrac{\pi}{2} \leqslant x \leqslant \dfrac{\pi}{2}\right)$; в) $p_{\eta_1, \eta_2}(x_1, x_2) = p_{\eta_1}(x_1)\, p_{\eta_2}(x_2)$.

10. $p_{\eta_1, \eta_2}(x_1, x_2) = \dfrac{1}{2\pi} e^{-(x_1^2 + x_2^2)/2}$.

11. а) $\mathsf{P}(\xi_1 = -1) = \mathsf{P}(\xi_1 = 1) = 1/2$; $\mathsf{P}(\xi_2 = -1) = 1/3$, $\mathsf{P}(\xi_2 = 0) = 1/4$; $\mathsf{P}(\xi_2 = 1) = 5/12$; б) $\mathsf{P}(\eta_1 = 2) = \mathsf{P}(\eta_1 = -2) = 1/8$; $\mathsf{P}(\eta_1 = -1) = 1/12$, $\mathsf{P}(\eta_1 = 0) = 1/2$; $\mathsf{P}(\eta_1 = 1) = 1/6$; в) $\mathsf{P}(\eta_2 = 0) = 1/4$, $\mathsf{P}(\eta_2 = 1) = 3/4$; г) $1/2$.

12. $p_{\xi_1 + \xi_2}(x) = 1/2$ $(0 \leqslant x \leqslant 2)$.

13. Распределение Пуассона с параметром $\lambda = \lambda_1 + \lambda_2$.

14. $\mathsf{P}(\tau = k) = (1 - p)^{k-1} p$ $(k \geqslant 1)$.

15. $\mathsf{P}(\tau^{(1)} = k, \tau^{(2)} = l) = (1 - p)^{k-1}(1 - p)^{l-1} p^2$ $(k \geqslant 1,\ l \geqslant 1)$.

16. $\mathsf{P}(\tau_m = k) = C_{k-1}^{m-1} p^m q^{k-m}$ $(k = m, m + 1, \ldots)$.

17. $\mathsf{P}(\theta = l, \ \nu = n) = (1/3)^{n-1}\, 1/6$ $(l = 1, 2, 3, 4;\ \ n = 1, 2, \ldots)$; величины ν и θ независимы.

18. а) $\alpha n e^{-\alpha n x}$ $(x > 0)$; б) $\alpha n (1 - e^{-\alpha x})^{n-1} e^{-\alpha x}$ $(x > 0)$.

19. 1) $n(1 - F(x))^{n-1} p(x)$; 2) $n(F(x))^{n-1} p(x)$; 3) $\dfrac{n!}{(m-1)!\,(n-m)!} \times (F(x))^{m-1}(1 - F(x))^{n-m} p(x)$; 4) $\dfrac{n!}{(m-1)!\,(k-m-1)!\,(n-k)!} \times (F(x_1))^{m-1}(F(x_2) - F(x_1))^{k-m-1}(1 - F(x_2))^{n-k} p(x_1)\, p(x_2)$ $(x_1 < x_2;$ $m < k)$.

20. $\lambda = n_1 p_1 + n_2 p_2 + n_3 p_3 = 2$; $1 - e^{-\lambda}(1 + \lambda) = 0{,}59399\ldots$

21. $2/\pi = 0{,}6366197\ldots$; величины ξ_1, ξ_2 зависимы.

22. $\mathsf{P}(\xi_1 = i, \ \ \xi_2 = j) = 0{,}01$; $i, j = 0, 1, \ldots, 9$. Величины ξ_1, ξ_2 независимы.

ГЛАВА 6

1. $1/p$. **2.** $(n+1)/2$.

3. $\mathsf{P}(\xi=k)=\left(1-\frac{1}{n}\right)^{k-1}\frac{1}{n}$, $\mathsf{M}\xi=n$.

4. m/p. **5.** $\mathsf{M}\tau=1/\lambda$, $\mathsf{D}\tau=1/\lambda^2$. **6.** 9; 16,5.

7. $\mathsf{M}\xi=npq$, $\mathsf{D}\xi=npq\,(p^3+q^3)+2p^2q^2$.

8. $\mathsf{M}\eta_1=9$, $\mathsf{M}\eta_2=20{,}25$, $\mathsf{D}\eta_1=16{,}5$, $\mathsf{D}\eta_2=402{,}1875$.

9. $\mathsf{M}\xi_1=0$, $\mathsf{M}\xi_2=1/12$, $\mathsf{D}\xi_1=1$, $\mathsf{D}\xi_2=107/144$, $\operatorname{cov}(\xi_1,\xi_2)=-1/4$.

10. $\mathsf{M}\xi_1=\mathsf{M}\xi_2=\operatorname{cov}(\xi_1,\xi_2)=0$, $\mathsf{D}\xi_1=\mathsf{D}\xi_2=0{,}5$. Случайные величины зависимы.

11. а) 0; б) 1/3.

12. $\mathsf{M}\xi_k=np_k$, $\mathsf{D}\xi_k=np_k(1-p_k)$, $\operatorname{cov}(\xi_k,\xi_l)=-np_kp_l\ (k\neq l)$.

13. $\mathsf{P}(\mu_0=0)=\sum\limits_{l=0}^{N}(-1)^l C_N^l\left(1-\frac{l}{N}\right)^n$.

14. $\mathsf{M}\mu_0=N\left(1-\frac{1}{N}\right)^n\sim Ne^{-\alpha}$; $\mathsf{D}\mu_0=\mathsf{M}\mu_0+N(N-1)\left(1-\frac{2}{N}\right)^n-$ $-(\mathsf{M}\mu_0)^2\sim Ne^{-\alpha}(1-(1+\alpha)e^{-\alpha})$, $N\to\infty$, $\frac{n}{N}\to\alpha$.

15. $0{,}08675<p<0{,}09$.

16. $\sum\limits_{k=1}^{n}(-1)^{k+1}\frac{1}{k!}\to 1-e^{-1}=0{,}63212\ldots$

17. $\mathsf{M}\xi=\mathsf{D}\xi=1$. **18.** $\mathsf{M}\nu=N\sum\limits_{k=1}^{N}\frac{1}{k}$.

19. а) 0,00269...; б) 0,01831...; в) 0; г) $1/9=0{,}111\ldots$; д) 0,08606...

20. 1) $\mathsf{P}\{|\xi|<u_\alpha\sigma\}=0{,}90$; а) $\mathsf{P}\{|f(\xi)|<u_\alpha\sigma\}=0{,}7805\ldots$; б) $\mathsf{P}\{|f(\xi)|<u_\alpha\sigma\}=0{,}9014\ldots$; 2) $\mathsf{P}\{|f(\xi)|<u_\alpha\sigma\}\approx\mathsf{P}\{|\xi|<u_\alpha\sigma\}+$ $+\mathsf{P}\left\{\left|\xi+\frac{1}{k}\right|<u_\alpha\sigma\right\}$.

21. $\frac{\pi}{n+1}\left(1-\left(\frac{3}{4}\right)^{n+1}\right)$.

22. $\alpha=\sigma_{12}/\sigma_{11}$, $\beta=a_2-(a_1\sigma_{12}/\sigma_{11})$.

23. 0,25; 0,0455... **24.** Да.

25. При $\alpha<1$ применим; при $\alpha\geqslant 1$ не применим.

27. а) $C_n^m(\Lambda_k/\Lambda_N)^m(1-(\Lambda_k/\Lambda_N))^{n-m}$, $\Lambda_k=\lambda_1+\ldots+\lambda_k$; б) $n\Lambda_k/\Lambda_N$.

28. η распределено нормально с параметрами (0, 1).

30. $M_k=\frac{k}{q-p}-\frac{n}{q-p}\cdot\frac{1-(q/p)^k}{1-(q/p)^n}$, $p\neq q$; $M_k=k(n-k)$, $p=$ $=q=1/2$.

31. 0,08145... **32.** Да.
33. а) η распределено нормально; б) нет.

ГЛАВА 7

1. а) $x\varphi(x^2)$; б) $(1 - x\varphi(x))/(1-x)$; в) $\varphi(x)/(1-x)$; г) $(\varphi(\sqrt{x}) + \varphi(-\sqrt{x}))/2$.

2. $px/(1-qx)$; $\mathsf{M}\tau = 1/p$, $\mathsf{D}\tau = q/p^2$, $q = 1-p$; $\left(\frac{px}{1-qx}\right)^m$; $\mathsf{M}\tau_m = m/p$, $\mathsf{D}\tau_m = mq/p^2$.

3. $\prod_{k=1}^{N}\left(\left(1 - \frac{k-1}{N}\right)\cdot\frac{x}{1 - \frac{k-1}{N}x}\right)$. **4.** $e^{\lambda(x-1)}$. **5.** $e^{\lambda p(x-1)}$.

6. а) $(pe^{it} + q)^n$; б) $e^{\lambda(e^{it}-1)}$; в) $1/(1 - (it/\alpha))$; г) $\sin t/t$.
7. а) $\mathsf{P}(\xi = 1) = \mathsf{P}(\xi = -1) = 1/2$; б) $\mathsf{P}(\xi = 2) = \mathsf{P}(\xi = -2) = 1/4$, $\mathsf{P}(\xi = 0) = 1/2$; в) $\mathsf{P}(\xi = k) = \mathsf{P}(\xi = -k) = 2^{-k-1}$, $k = 1, 2, \ldots$
8. $|f(t)|^2$. **9.** а) 0,5; б) 0,6826.
11. $0{,}75\cdot 10^{-m+2}$.
13. Нормальное распределение с параметрами (0, 1).
14. $\frac{\lambda^m}{m!}e^{-\lambda}$. **15.** $x\ (0 < x < 1)$.
16. а) $C_n^m p^m (1-p)^{n-m}$; б) $\frac{\lambda^m}{m!}e^{-\lambda}$.
20. $a = 1{,}71828\ldots$, $\mathsf{D}a^* = 0{,}00242036\ldots$, $\Delta = 0{,}1267203\ldots$

ГЛАВА 8

1. 1) (0,385; 0,336; 0,279); 2) 0,0336; 3) (16/47, 17/47, 14/47).

2. $\mathsf{P}(\xi_s = i) = \sum_{i_0,\ldots,i_{s-1}} q_{i_0}p_{i_0i_1}\cdots p_{i_{s-1}i}$, $\mathsf{P}(\xi_s = i,\quad \xi_{t+s} = j) = \mathsf{P}(\xi_s = i)\sum_{l_1,\ldots,l_{t-1}} p_{il_1}\cdots p_{l_{t-1}j}$.

4. а) Нет, если $p \neq q$; да, если $p = q = 1/2$. б) Да, $\mathsf{P}(\eta_{t+1} = 1 \mid \eta_t = 1) = p$, $\mathsf{P}(\eta_{t+1} = 1 \mid \eta_t = -1) = 1 - p$. в) Да, $p_{11} = p_{23} = p_{31} = p_{43} = 1 - p$, $p_{12} = p_{24} = p_{32} = p_{44} = p$.

5. $\begin{pmatrix} 3/7 & 4/7 \\ 1/11 & 10/11 \end{pmatrix}$.

6. $\mathsf{P}(\mu_0(n+1) = k \mid \mu_0(n) = k) = \frac{k}{N}$; $\mathsf{P}(\mu_0(n+1) = k-1 \mid \mu_0(n) = k) = \frac{N-k}{N}$.

7. Является. **8.** 1/6.

10. а) $\{1, 2\}$; б) $\frac{138}{97} = 1{,}42268\ldots$; в) $p_1^{(\alpha)} = \frac{52}{97} = 0{,}5360\ldots$,

$p_2^{(\alpha)} = \frac{40}{97} = 0{,}4123\ldots,\ p_j^{(\beta)} = 1 - p_j^{(\alpha)},\ j = 1, 2;$ г) $\pi_1 = \pi_2 = 0,\ \pi_3 = \frac{92}{291} = 0{,}3161\ldots,\ \pi_4 = \frac{46}{291} = 0{,}1580\ldots,\ \pi_5 = \pi_6 = \frac{51}{194} = 0{,}2628\ldots$

11. $(\beta/(\alpha+\beta),\ \alpha/(\alpha+\beta))$.

ГЛАВА 9

3. $\mathsf{M}m_{11} = \frac{n}{n-1}\operatorname{cov}(x_1, y_1)$. 6. $A = 2\arcsin\sqrt{p},\ B = 1$.

7. $a^* = \sum_{k=1}^{n} c_k x_k$, где $c_k = (1/\sigma_k^2)\Big/\left(\sum_{l=1}^{n} 1/\sigma_l^2\right)$.

11. а) $\mathsf{M}\theta_1^* = \mathsf{M}\theta_2^* = \frac{a+b}{2},\ \mathsf{D}\theta_1^* = \frac{(b-a)^2}{12n},\ \mathsf{D}\theta_2^* = \frac{(b-a)^2}{2(n+1)(n+2)}$;
б) $\mathsf{M}\theta_1^* = 1/\alpha,\ \mathsf{D}\theta_1^* = 1/\alpha^2 n,\ \mathsf{M}\theta_2^* = \frac{1}{2\alpha}\left(1 + \frac{1}{2} + \ldots + \frac{1}{n-1} + \frac{2}{n}\right)$,
$\mathsf{D}\theta_2^* = \frac{1}{4\alpha^2}\left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{(n-1)^2} + \frac{4}{n^2}\right)$.

12. $\mathsf{M}c^* = c,\ \mathsf{D}c^* = 1/n^2$.

15. $\eta_n = \sum_{k=1}^{N} \xi_k \ln \frac{p_k^{(0)}}{p_k^{(1)}}$, где ξ_k — число исходов полиномиальной схемы с номером k.

16. $\frac{n-2}{n}\sigma^2$.

17. $\mathsf{M}Q_1 = (r-1)\sigma^2 + n\left[\sum_{i=1}^{r} a_i^2 \frac{n_i}{n} - \left(\sum_{i=1}^{r} a_i \frac{n_i}{n}\right)^2\right]$, $\mathsf{M}Q_2 = (n-r)\sigma^2$.

ГЛАВА 10

1. а) λt; б) bt.
3. См. (10.2.2) с $n = 3,\ t_3 = T$.
4. а) $\lambda = (1-p)/p$; б) $\mathsf{D}\mu_t = B_t = t,\ \mathsf{M}\mu_t = 1$.
5. а) $\frac{\partial F}{\partial t} = \mu - (\lambda + \mu)F + \lambda F^2$;

б) $F(t, x) = 1 - \frac{e^{(\lambda-\mu)t}(1-x)}{\frac{\lambda}{\lambda-\mu}\left(e^{(\lambda-\mu)t} - 1\right)(1-x) + 1}\ (\lambda \ne \mu),\ F(t, x) = 1 - \frac{1-x}{\lambda t(1-x) + 1}\ (\lambda = \mu)$; в) $A(t) = e^{(\lambda-\mu)t}$.

6. 1) 1, если $q \geqslant p$; $(q/p)^k$, если $q < p$; 2) $\pi_0 = \frac{q-p}{1+q-p}$; $\pi_k = \frac{\pi_0}{p}(p/q)^k,\ k \geqslant 1$.

7. 1) $1/(\lambda_k+\mu_k)$; 2) $\lambda_k/(\lambda_k+\mu_k)$, $(\mu_k/(\lambda_k+\mu_k))$, $(\lambda_k+\mu_k)\,m_k=$ $=1+\mu_k m_{k-1}+\lambda_k m_{k+1}$, $0<k<n$; $m_0=m_1+\lambda_0^{-1}$, $m_n=0$.

8. $\|P_{ij}(t)\|=\dfrac{1}{\alpha+\beta}\begin{pmatrix}\beta & \alpha\\ \beta & \alpha\end{pmatrix}+\dfrac{e^{-(\alpha+\beta)t}}{\alpha+\beta}\begin{pmatrix}\alpha & -\alpha\\ -\beta & \beta\end{pmatrix}$.

9. $\dfrac{\partial f_1}{\partial t}=-(\alpha-is)\,f_1+\alpha f_2$, $\dfrac{\partial f_2}{\partial t}=\beta f_1-(\beta+is)\,f_2$, $f_1(0,\ s)=$ $=f_2(0,s)=1$.

10. $\beta/(\alpha+\beta)$.

11. $\dfrac{\partial M_1}{\partial t}-v_1\dfrac{\partial M_1}{\partial x}=-\alpha M_1+\alpha M_2$, $\dfrac{\partial M_2}{\partial t}+v_2\dfrac{\partial M_2}{\partial x}=\beta M_1-\beta M_2$,

$$\begin{cases}\dfrac{dm_1}{dt}=-\alpha m_1+\alpha m_2+v_1, & m_k(t)=M_k(t,0),\\[2mm] \dfrac{dm_2}{dt}=\beta m_1-\beta m_2-v_2, & m_1(0)=m_2(0)=0;\end{cases}$$

$$\lim_{t\to\infty}\frac{1}{t}M_k(t,\ x)=(v_1\beta-v_2\alpha)/(\alpha+\beta).$$

13. $$\begin{cases}M_1(t,\ x)=e^{-\alpha t}(x+vt)+\displaystyle\int_0^t \alpha e^{-\alpha u}M_2(t-u,\ x+vu)\,du,\\[2mm] M_2(t,\ x)=e^{-\beta t}\left(x-\dfrac{at^2}{2}\right)+\displaystyle\int_0^t \beta e^{-\beta u}M_1\left(t-u,\ x-\dfrac{au^2}{2}\right)du;\end{cases}$$

$$\lim_{t\to\infty}\frac{1}{t}M_k(t,\ x)=v\,\frac{\beta}{\alpha+\beta}-\frac{a}{\beta}\,\frac{\alpha}{\alpha+\beta}.$$

14. $\mathsf{M}\eta_t=a$, $\mathsf{D}\eta_t=(c_0^2+c_1^2+c_2^2)\,\sigma^2=R(0)$, $R(1)=c_1(c_0+$ $+c_2)\,\sigma^2$, $R(2)=c_0c_2\sigma^2$, $R(k)=0$, $k\geqslant 3$.

15. $\mathsf{M}a_t^*=a$, $\mathsf{D}a_t^*=\dfrac{\sigma^2}{t^2}\,[t\,(c_0^2+c_1^2+c_2^2)+2\,(t-1)\,c_1(c_0+$ $+c_2)+2\,(t-2)\,c_0c_2]$; оценка a_t^* является несмещенной и состоятельной оценкой a.

Список литературы

[1] Большев Л. Н., Смирнов Н. В. Таблицы математической статистики.— 3-е изд.— М.: Наука, 1983.

[2] Боровков А. А. Теория вероятностей.— М.: Наука, 1976.

[3] Гельфонд А. О. Исчисление конечных разностей.— 3-е изд.— М.: Наука, 1967.

[4] Гнеденко Б. В. Курс теории вероятностей.— 5-е изд.— М.: Наука, 1969.

[5] Ермаков С. М. Метод Монте-Карло и смежные вопросы.— М.: Наука, 1971.

[6] Кендалл М., Моран П. Геометрические вероятности/Пер. с англ.— М.: Наука, 1972.

[7] Коваленко И. Н., Филиппова А. А. Теория вероятностей и математическая статистика.— М.: Высшая школа, 1973.

[8] Колмогоров А. Н. Основные понятия теории вероятностей.— М.: Наука, 1974.

[9] Колчин В. Ф., Севастьянов Б. А., Чистяков В. П. Случайные размещения.— М.: Наука, 1976.

[10] Крамер Г. Математические методы статистики.— 2-е изд./Пер. с англ.— М.: Мир, 1975.

[11] Мешалкин Л. Д. Сборник задач по теории вероятностей.— М.: Изд-во МГУ, 1963.

[12] Прохоров Ю. В., Розанов Ю. А. Теория вероятностей (Основные понятия, предельные теоремы, случайные процессы). — 2-е изд. — М.: Наука, 1973.

[13] Розанов Ю. А. Случайные процессы.— 2-е изд.— М.: Наука, 1979.

[14] Романовский В. И. Дискретные цепи Маркова.— М.: Гостехиздат, 1949.

[15] Севастьянов Б. А. Ветвящиеся процессы.— М.: Наука, 1971.

[16] Севастьянов Б. А., Чистяков В. П., Зубков А. М. Сборник задач по теории вероятностей.— М.: Наука, 1980.

[17] Смирнов Н. В., Дунин-Барковский И. В. Курс теории вероятностей и математической статистики для технических приложений.— 3-е изд.— М.: Наука, 1969.

[18] Феллер В. Введение в теорию вероятностей и ее приложения. Т. 1./Пер. с англ.— М.: Мир, 1967.

[19] Ченцов Н. Н., Чистяков В. П. Теория вероятностей (методические указания).— М.: Изд-во МГУ, 1961.

[20] Чистяков В. П. Задачник по теории вероятностей (Задачи и эксперименты со случайными числами).— М.: Изд-во МИФИ, 1980.

[21] Шеффе Г. Дисперсионный анализ.— 2-е изд./Пер. с англ.— М.: Наука, 1980.

Оглавление

В.П.ЧИСТЯКОВ

Курс теории вероятностей

Курс теории вероятностей

В.П.ЧИСТЯКОВ