

НАУЧНО-ПОПУЛЯРНАЯ ЛИТЕРАТУРА

А. В. СКОРОХОД

Вероятность вокруг нас

КИЕВ «НАУКОВА ДУМКА» 1980

В окружающем нас мире все время происходят явления, которые заранее невозможно предсказать: это и ядерные реакции, и передача наследственных признаков, и солнечные вспышки, и появление новых и сверхновых звезд... Можно ли какими-либо точными методами изучать случайность? Кажется, что одно исключает другое. Однако среди большой семьи математических наук есть одна — теория вероятностей, которая всецело посвящена именно теории случайных явлений. О том, как математика изучает случайные явления, и рассказывается в этой книге.

Адресована широкому кругу читателей — и ученикам старших классов, интересующихся математикой, и нематематикам, желающим получить представление о теории вероятностей.

ОТВЕТСТВЕННЫЙ РЕДАКТОР
ЧЛЕН-КОРРЕСПОНДЕНТ АН УССР *И. П. ГИХМАН*

Редакция научно-популярной литературы

С $\frac{20204-384}{M221(04)-80}$ 574-80. 1702060000

© Издательство «Наукова думка», 1980

ВВЕДЕНИЕ

На первый взгляд математика, известная достоверностью и даже непреложностью своих выводов, никак не может быть применена к случайному. Но именно этой наукой создана числовая мера случайности — вероятность, которая и дает возможность изучать случайно протекающие реальные явления, используя их математические модели. С некоторыми такими моделями читатель и познакомится в этой книге.

И в своей обыденной жизни, и в производственной деятельности (технике), и в процессе познания природы (естествознания) мы непрерывно сталкиваемся со случайностью (недостоверностью, вероятностью). Эта случайность несколько не противоречит закономерности развития мира: случайность и закономерность — диалектические противоположности, закономерность во многих случаях складывается из случайностей (математическим выражением этого факта есть закон больших чисел).

Какова же природа случайности, проявляющейся на различных ступенях реальности? Если исходить из современных физических представлений, то основной источник случайности — квантовый характер явлений микромира. Случайными являются переходы атомов и молекул из одних энергетических состояний в другие, эти переходы сопровождаются поглощением или излучением света. Таким образом, световая стихия, окружающая нас всюду (электромагнитное поле), по своему происхождению случайна. Случайным образом происходят и ядерные реакции. Но именно эти реакции и есть источники энергии звезд и собственного тепла планет; а первая — основа всех жизненных процессов на нашей планете, второе — причина тектонической деятельности Земли.

Второй вид проявления случайности в микромире, влияние которого также очень существенно для большинства процессов, — броуновское движение. Это хаотическое тепловое движение, в котором пребывают ионы плазмы, молекулы газов и жидкостей, электроны в кристаллах. Броуновское движение поддерживается взаимодействиями между отдельными микрочастицами среды, и поскольку такие взаимодействия носят случайный характер, то и само броуновское движение по своей природе случайно.

Роль броуновского движения в окружающих нас явлениях трудно переоценить. Благодаря ему происходят растворение одних веществ в других, химические реакции, все жизненные процессы в живых организмах. Броуновское движение электронов создает электро- и радиопомехи в различных технических устройствах. Ядерные реакции и броуновское движение в солнечной плазме приводят к появлению солнечных вспышек, вызывающих магнитные бури, полярные сияния и оказывающих влияние на погоду.

Случайна и природа более сильных звездных вспышек — появление новых и сверхновых звезд. Такие явления — источник космического излучения, непрерывно приходящего к нам из глубин Вселенной.

Итак, мы окружены явлениями, природа которых случайна. Как же может оказаться, что тем не менее существуют точные законы, которым явления подчиняются? Ведь, несмотря на броуновское движение отдельных молекул газа, мы можем утверждать, что состояние газа подчиняется строгому закону Клапейрона. Чтобы ответить на вопрос, нужно изучить случайные события, отвлекаясь от их конкретных свойств, а лишь имея в виду случайность. Этим занимается теория вероятностей — один из разделов математики. Исследуя общие закономерности случайных явлений, теория вероятностей установила один из фундаментальных законов, которому они подчиняются, — закон больших чисел. Из него можно вывести и закон Клапейрона, и многие другие, относящиеся к большим совокупностям случайных событий.

Современная теория вероятностей — наука, опирающаяся на сложный математический аппарат. Но многие ее идеи и методы можно объяснить, используя лишь школьную математику (с учетом того, что современный школьник имеет понятие о производной и интеграле). Предлагаемая книга посвящена элементарному изложению основ

теории вероятностей. Для лучшего овладения математическими понятиями она содержит большое число примеров из биологии (в основном генетики и экологии), физики (ядерные реакции, энергетические переходы в атомах, броуновское движение), радиотехники (случайные электромагнитные колебания, спектры радиосигналов, фильтрация), теории обслуживания. Некоторые примеры носят чисто иллюстративный характер, другие же являются содержательными и дают понятие о возможных применениях теории вероятностей в различных областях науки.

Книга состоит из трех разделов: случайные события, случайные величины, случайные процессы. При переходе от одного раздела к другому понятие случайности все более усложняется. В первом разделе мы по сути имеем дело с одним или несколькими случайными событиями. Основное здесь — выяснение связей между частотами событий при повторении экспериментов и их вероятностями. При изучении случайных величин во втором разделе мы уже должны одновременно рассматривать бесконечную совокупность случайных событий, заключающихся в том, что наблюдаемая величина приняла заданное значение. Здесь также рассмотрены закон больших чисел и центральная предельная теорема и выяснены их значения для различных приложений. Наконец, в третьем разделе речь идет о случайных величинах, меняющихся во времени, т. е. о бесконечных совокупностях случайных величин. Описаны лишь некоторые классы процессов, служащих вероятностными моделями для реальных явлений в технике, физике, биологии, радиотехнике. Для каждого класса процессов решаются свои задачи, вытекающие из природы описываемых явлений (эргодическая теорема, вырождение и «взрыв», нахождение спектра колебаний и т. п.).

При изложении всех вопросов выводы по возможности подтверждаются математическими доказательствами (если они не выходят за рамки элементарного курса математики); исключение составляют центральная предельная теорема и эргодическая теорема.

В списке литературы, приведенном в конце книги, указаны некоторые руководства, которые можно рекомендовать читателю, желающему более глубоко ознакомиться с различными аспектами рассматриваемого предмета.

1. СЛУЧАЙНЫЕ СОБЫТИЯ

§ 1. Что такое случайное событие?

Понятие случайного события легче уяснить в сопоставлении с противоположным понятием — неслучайного, или детерминированного, события. Последнее — это такое, которое происходит всегда, если выполнены определенные условия. Так, вода, нагретая до 100°C , при нормальном давлении закипает; тяжелое тело, лишенное горизонтальной опоры, будет падать, пока не встретит такую опору; нагретый газ при неизменном давлении увеличит свой объем, причем можно указать на сколько, если известны начальная и конечная температуры газа. Посеяв пшеничное зерно, мы заранее можем предсказать, что из него вырастет пшеница. У кошек рождаются котята, а у свиней — поросята. Число подобных примеров достоверных событий каждый может увеличить неограниченно.

Случайное событие противоположно детерминированному. Это значит, что при одних и тех же условиях оно может как наступать, так и не наступать. Основное его свойство то, что оно непредсказуемо, если известно, что те условия, в которых оно возможно, осуществлены.

Если посеяно зерно пшеницы, то мы не можем сказать заранее, прорастет оно или нет, вырастет ли из него взрослое растение или оно погибнет (будучи затоптанным, побитым градом, вымыто ливнем или от недостатка влаги), сколько зерен принесет взрослое растение, даже о свойствах его семян (как мы увидим далее, рассматривая законы генетики) нельзя заранее сказать определенно. Рассматривая газ, мы не можем предсказать число молекул и энергии этих молекул в выделенном объеме в любой фиксированный момент времени. Случайным образом происходит переход атомов из возбужденного состояния в устойчивое (электромагнитное поле, окружающее нас, обязано своим происхож-

денем именно этому случайному механизму), радиоактивный распад и вообще все ядерные реакции, солнечные вспышки, вызывающие магнитные бури и радиопомехи, вспышки новых звезд, являющиеся источниками космического излучения.

При рассмотрении случайных событий удобно использовать понятие случайного эксперимента, в результате которого и происходят случайные события. Случайный эксперимент характеризуется набором условий, в которых наблюдаются случайные события, и совокупностью событий, фиксируемых в эксперименте. Если эксперимент случаен, то его повторение (рассматриваются лишь эксперименты, которые могут повторяться достаточно много раз) будет сопровождаться, вообще говоря, различными событиями (если бы всегда происходило одно и то же, то событие было бы детерминированным).

Приведем примеры случайных экспериментов.

1. Бросается монета. Результаты: монета легла на пол гербом вверх, монета легла цифрами вверх, монета стала на ребро.

2. Завод электроламп выпускает продукцию партиями по 10 000 штук, изготовленных примерно в одинаковых условиях. Для контроля качества отбирается 100 ламп, которые испытываются при повышенном напряжении. Лампа годна, если она выдерживает 1 час работы в таких условиях. Эксперимент заключается в одновременной проверке 100 ламп. События, которые наблюдаются, — число ламп, сгоревших за 1 час.

3. Запускается космическая ракета в сторону Луны и измеряется наименьшее расстояние, на котором она пройдет от спутника Земли. Недостоверность ожидаемого результата определяется невозможностью: 1) абсолютно точно задать время запуска; 2) абсолютно точно дозировать горючее; 3) полностью учесть влияние атмосферных условий. Поэтому даже событие «ракета достигла Луны» (при отсутствии дополнительной коррекции) является случайным.

4. Происходит самоопыление гибридного растения. Высаживается его семя. Наблюдаются свойства развившегося растения (они могут быть различными комбинациями свойств тех сортов, от скрещивания которых получен гибрид).

5. В урне имеется m черных и n белых шаров. Вынимается наудачу (не глядя) один шар. События, которые в

этом эксперименте могут произойти, — извлечение белого или черного шара.

Если производить многократное повторение случайного эксперимента, то одни события при этом будут происходить чаще других. Ожидать появления первых событий при одном осуществлении эксперимента можно с большим основанием, чем вторых. Тогда мы говорим, что первое событие более вероятно, чем второе, или что его вероятность больше, чем вероятность второго. Например, у кареглазых родителей, если среди их предков были голубоглазые, гораздо чаще рождаются кареглазые дети, чем голубоглазые. Поэтому можно сказать, что кареглазый ребенок у таких родителей вероятнее, чем голубоглазый.

Определим частоту события A в данном эксперименте. Предположим, что эксперимент повторялся n раз и при этом событие A произошло $k_n(A)$ раз. Тогда частота события A в этих n испытаниях есть отношение

$$\frac{k_n(A)}{n}. \quad (1)$$

Пусть эксперимент заключается в проверке качества электроламп (как указано в примере 2). Если из 100 ламп оказалось 6 бракованных, то частота события «лампа из данной партии негодна» будет $\frac{6}{100} = 0,06$.

Частота события, вообще говоря, изменится, если варьировать число наблюдений или взять вторую серию из такого же числа наблюдений. Однако многими экспериментами установлено: при очень большом числе наблюдений n значение частоты устойчиво, т. е. оно мало меняется при увеличении числа наблюдений или при переходе к другой серии наблюдений, если число наблюдений достаточно велико. Так, при бросании правильной монеты (если исключена возможность падения монеты на ребро) частота выпадения герба будет равна примерно $\frac{1}{2}$, и она тем ближе будет приближаться к этому значению, чем больше будет проведено наблюдений.

При математическом изучении случайных событий постулируется существование идеальной частоты события — неизменной частоты при бесконечном числе испытаний, являющейся пределом выражения (1) при неограниченном возрастании числа испытаний. Эта идеальная частота называется *вероятностью* события.

Вероятность события A обычно обозначается $P(A)$. Если для двух событий A и B имеет место $P(A) < P(B)$, то событие B более вероятно, чем A . Так, события «рождение голубоглазого ребенка» и «рождение кареглазого ребенка» в описанных выше условиях имеют соответственно вероятности $\frac{1}{4}$ и $\frac{3}{4}$, так что второе событие более вероятно.

Перечислим некоторые свойства вероятности.

I. Вероятность события A — неотрицательное число, не превосходящее единицы: $0 \leq P(A) \leq 1$. Это свойство очевидным образом выполняется для частоты $\frac{k_n(A)}{n}$ ($k_n(A)$ — число появлений события A в n экспериментах неотрицательно и не превосходит n). Поскольку вероятность рассматривается как предельное значение частоты, то это свойство сохраняется и для вероятности. Заметим, что указанным приемом мы будем пояснять и другие свойства вероятности, т. е. проверять их для частот событий, а затем пользоваться тем, что эти свойства сохраняются при предельном переходе.

Для формулировки других свойств нам потребуются некоторые определения, относящиеся к случайным событиям.

Будем говорить, что некоторое событие U в данном эксперименте *достоверно*, если оно происходит при каждой реализации эксперимента. (Если в примере 4 самоопыляется горох, полученный от скрещивания растений с различной окраской цветов, то достоверно, что из полученного семени также вырастет горох.) Достоверное событие неслучайно в данном эксперименте, однако его рассмотрение удобно в связи с некоторыми операциями со случайными событиями.

II. Вероятность достоверного события равна 1. Это вытекает из того, что его частота равна 1, так как $k_n(U) = n$ для любой серии из n повторений эксперимента.

Событие V называется невозможным в данном эксперименте, если оно не может произойти в нем никогда. (Например, невозможно из урны, содержащей белые и черные шары, извлечь красный шар). Невозможное событие, как и достоверное, неслучайно в данном эксперименте. Оно вводится также ради удобства.

III. Вероятность невозможного события равна 0. (Легко видеть, что $k_n(V) = 0$ для всех n .)

Введем понятие *суммы двух событий*. Если A и B — два события в данном эксперименте, то под $A + B$ понимают следующее событие: произошло или A , или B , либо они произошли одновременно. Другими словами, событие $A + B$ происходит тогда и только тогда, когда происходит по крайней мере одно из событий — или A , или B .

Пусть эксперимент заключается в проверке качества изделий (например, электроламп, как в примере 2). Результат эксперимента — число дефектных изделий в проверяемой партии. Если A — событие «дефектные изделия отсутствуют», B — «обнаружено 1 дефектное изделие», то $A + B$ — событие «число дефектных изделий не превосходит одного».

Обратим внимание на существенное отличие «сложения» событий от сложения чисел. В частности, справедлива формула $A + A = A$ ($A + A$ происходит тогда и только тогда, когда происходит событие A , т. е. оно совпадает с A ; следует сказать, что мы всюду будем отождествлять события, которые обязательно происходят одновременно).

Событие, заключающееся в том, что два события A и B происходят одновременно, называется *произведением событий A и B* и обозначается $A \cdot B$, или AB . Пусть, например, A — рождение мальчика, B — рождение кареглазого ребенка, тогда AB — рождение кареглазого мальчика.

Произведение нескольких событий A_1, \dots, A_k есть событие, заключающееся в том, что все эти события происходят одновременно.

События A и B называются несовместимыми, если они не могут произойти одновременно. Примеры: выпадение герба или выпадение цифр при однократном бросании монеты; рождение девочки или мальчика при рождении одного ребенка; обнаружение 1 или 2 дефектных изделий при контроле партии готовой продукции и т. п.

IV. Если A и B — два несовместимых события, то

$$P(A + B) = P(A) + P(B).$$

Опять проверим это свойство для частот. Если происходит $A + B$ и A и B несовместимы, то происходит или A , или B , а одновременно они произойти не могут. Поэтому, чтобы подсчитать, сколько раз в n экспериментах произойдет $A + B$, нужно знать, сколько раз произошло A , сколько раз — B , а полученные числа сложить. Таким образом,

$$k_n(A + B) = k_n(A) + k_n(B);$$

$$\frac{k_n(A+B)}{n} = \frac{k_n(A)}{n} + \frac{k_n(B)}{n}.$$

Мы получили требуемое равенство для частот.

Можно понятие суммы событий и свойство IV распространить на большее число событий. Пусть в данном эксперименте могут происходить события A_1, A_2, \dots, A_k . Событие $A_1 + \dots + A_k$ (его еще обозначают $\sum_{i=1}^k A_i$) есть событие, заключающееся в том, что происходит по крайней мере одно из событий: или A_1 , или A_2, \dots , или A_k . Сумму событий можно определять последовательно (или, как говорят, индуктивно). Зная определение суммы двух событий, полагаем для трех:

$$A_1 + A_2 + A_3 = (A_1 + A_2) + A_3$$

(событие $A_1 + A_2$ уже известно, к нему прибавляем A_3).

Определив сумму трех событий для четырех, полагаем

$$A_1 + A_2 + A_3 + A_4 = (A_1 + A_2 + A_3) + A_4$$

и т. д.

Назовем события A_1, \dots, A_k *попарно несовместимыми*, если никакие два из этой совокупности событий не могут произойти одновременно. Пусть, например, в урне находятся шары четырех цветов: белые, черные, зеленые, красные. Черные, зеленые, красные шары будем именовать цветными. Событие «извлечен цветной шар» есть сумма трех событий: «извлечен черный шар», «извлечен зеленый шар», «извлечен красный шар».

Из свойства IV вытекает следующее свойство.

V. Если A_1, A_2, \dots, A_k — попарно несовместимые события, то

$$P(A_1 + A_2 + \dots + A_k) = P(A_1) + P(A_2) + \dots + P(A_k).$$

Таким образом, справедливо следующее общее утверждение: вероятность суммы любого числа попарно несовместимых событий равна сумме вероятностей этих событий.

Событие \bar{A} называется противоположным событию A , если оно заключается в том, что A не произошло. Так, если урна содержит белые и черные шары, а эксперимент заключается в извлечении одного шара, то событие «извлечен белый шар» противоположно к событию «извлечен

черный шар». Очевидно, что при каждом повторении эксперимента A или происходит или не происходит, т. е. происходит \bar{A} . Поэтому $A + \bar{A}$ — достоверное событие. Значит, $P(A + \bar{A}) = 1$. С другой стороны, $P(A + \bar{A}) = P(A) + P(\bar{A})$, так как A и \bar{A} несовместимы. Таким образом, выполняется данное свойство.

VI. Если событие \bar{A} противоположно событию A , то

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

Пусть в эксперименте могут происходить два события A и B . Будем говорить, что событие A влечет событие B , если всегда, когда происходит событие A , происходит и событие B . При этом событие B называется следствием события A . Для обозначения отношения « A влечет B » (« B следствие A ») используются обозначения

$$A \subset B, \quad B \supset A.$$

Отметим, что понятие следования, употребляемое по отношению к случайным событиям, отличается от принятого в обыденной жизни, где событие, следующее за A , во-первых, происходит позже по времени, во-вторых, A для него выступает в качестве действительной причины. Здесь же A и B появляются одновременно, и причинная зависимость между ними не обязательна.

В частности, поскольку с каждым событием A наступает и достоверное событие U (оно ведь происходит обязательно, независимо от того, произошло A или нет), то $A \subset U$ для любого события A . Этот пример указывает на необязательность причинной связи для следования событий одного из другого.

Более парадоксальным кажется другое утверждение: всякое событие A следует из невозможного события V . Действительно, поскольку V не происходит никогда, то для любого события A можно утверждать, что оно происходит, если только произошло V (на самом деле мы ничего относительно A не утверждаем, так что наше утверждение не может быть ошибочным). Значит, $V \subset A$.

VII. Если $A \subset B$, то $P(A) \leq P(B)$.

Чтобы проверить это свойство вероятности, рассмотрим событие C , заключающееся в том, что событие B произошло, а A при этом не произошло (C может быть и невозможным событием). Тогда A и C — несовместимые события.

Событие B происходит либо когда происходит A , либо когда A не происходит, т. е. происходит C . Значит, $B = A + C$

$$P(B) = P(A) + P(C) \geq P(A).$$

Заметим, что свойство VII также просто вытекает для частот, поскольку в силу определения при $A \subset B$ событие B происходит не реже, чем A . Однако мы использовали для проверки свойства VII свойства I ($P(C) \geq 0$) и IV.

Вообще мы видели, что из приведенных свойств некоторые являются логическими следствиями других. Это дает возможность выбрать основные свойства вероятности, а остальные получать как их логические следствия. Эти основные свойства называются *аксиомами вероятности*. Только они нуждаются в экспериментальном обосновании, все остальные утверждения выводятся из аксиом.

Сейчас в теории вероятностей принята система аксиом, предложенная известным советским математиком академиком А. Н. Колмогоровым. Мы приведем эти аксиомы в той форме, в которой они будут использоваться в дальнейшем.

Аксиома 1. Вероятность события $P(A) \geq 0$; если U достоверное событие, то

$$P(U) = 1.$$

Аксиома 2. Если A и B несовместимы, то

$$P(A + B) = P(A) + P(B).$$

Как видим, аксиома 1 объединяет свойства I и II, аксиома 2 совпадает со свойством IV. Оказывается, этих двух аксиом достаточно, чтобы чисто логическими методами получить весьма далеко идущие следствия. В частности, как будет показано ниже, из них будет вытекать, что при повторении эксперимента частота появления события A в определенном смысле приближается к его вероятности. Это как раз то фундаментальное свойство, которое первоначально было положено в основу определения вероятности.

§ 2. Эксперименты с конечным числом исходов. Классическое определение вероятности

Наиболее простые случайные эксперименты — это те, в которых число возможных событий конечно. Среди примеров, приведенных в предыдущем параграфе, все, кроме примера 3, обладают этим свойством. Такие случайные эксперименты называются *экспериментами с конечными*

числом исходов. Во множестве событий, которые наблюдаются в них, выделяют «элементарные» события, «кирпичики», из которых могут быть составлены остальные события.

Предположим, бросают монету. В этом эксперименте два элементарных события: выпадение герба, выпадение цифр.

Более сложно обстоит дело с извлечением одного шара из урны, содержащей черные и белые шары (пример 5 в § 1). На первый взгляд кажется, что здесь всего два элементарных события: извлечение белого или извлечение черного шара. Однако на самом деле может быть извлечен каждый из m черных шаров и каждый из n белых. Так что более естественно считать, что имеется $m + n$ элементарных событий: A_1, \dots, A_n — извлечение заданных белых шаров (можем считать, что они занумерованы) и B_1, \dots, B_m — извлечение черных шаров. При этом события A (извлечен белый шар) и события B (извлечен черный шар) оказываются составными:

$$A = A_1 + \dots + A_n, \quad B = B_1 + \dots + B_m.$$

Зачем нам понадобилось разлагать на первый взгляд простые события на еще более простые, мы увидим немного позже. Дадим точное определение элементарного события. Событие $E \neq V$ называется элементарным, если, каково бы ни было событие A — или $E \subset A$, или $E \subset \bar{A}$ (это событие или влечет A , или влечет противоположное к A событие).

Пусть E_1 и E_2 — два элементарных события. Если E_1 влечет E_2 , то E_2 не может вызвать \bar{E}_1 (а в силу своей элементарности E_2 влечет или E_1 , или \bar{E}_1). Значит, тогда и $E_2 \subset E_1$, т. е. события E_1 и E_2 наступают всегда одновременно и $E_1 = E_2$. Поэтому, если $E_1 \neq E_2$, то $E_1 \subset \bar{E}_2$ и $E_1 E_2 = V$.

Различные элементарные события несовместимы. Заметим, что если $A \neq V$, то всегда существует такое элементарное событие E , которое влечет A . Чтобы установить это, отметим следующее определяющее свойство элементарного события: событие E элементарно тогда и только тогда, когда из соотношения $C \subset E$ $C \neq V$ вытекает $E = C$. Действительно, в этом случае не может быть $E \subset \bar{C}$, значит, $E \subset C$ и, следовательно, $E = C$.

Если A само не элементарное событие, то найдется такое A_1 , что $A_1 \neq A$ и $A_1 \subset A$. Если A_1 не элементарно, то найдется $A_2 \neq A_1$ и $A_2 \subset A_1$. Поскольку различных событий конечное число, то, продолжая этот процесс, придем к конечной цепочке $A_k \subset A_{k-1} \subset \dots \subset A_1 \subset A$, и для A_k уже не существует такого $A_{k+1} \neq V$, что $A_{k+1} \subset A_k$ и $A_{k+1} \neq A_k$. Значит, A_k — элементарное событие.

Пусть S — сумма всех элементарных событий. Если $S \neq U$, то $\bar{S} \neq V$ и по доказанному содержало бы элементарное событие. А это противоречило бы определению S как суммы всех элементарных событий. Значит, $S = U$. Таким образом, в эксперименте с конечным числом исходов всегда существует полная группа элементарных событий E_1, \dots, E_n , т. е. такая совокупность попарно непересекающихся событий, сумма которых есть достоверное событие

$$U = E_1 + \dots + E_n. \quad (1)$$

Это означает, что при любом осуществлении эксперимента происходит — и при том лишь одно — из событий E_1, \dots, E_n . Всякое событие, происходящее в эксперименте, есть сумма нескольких элементарных событий, именно событие A есть сумма тех E_k , для которых E_k влечет A : $A = A \cdot U = A \cdot (E_1 + \dots + E_n) = A \cdot E_1 + \dots + A E_n$ — слагаемые в правой части либо невозможные события, если $A \cdot E_k = V$ (это будет, когда $E_k \subset \bar{A}$), или совпадают с самими элементарными событиями: $A \cdot E_k = E_k$ если $E_k \subset A$.

Для того чтобы определить вероятность любого события, достаточно знать вероятности элементарных событий. Пусть $p_k = P(E_k)$. Тогда вероятность события A есть сумма тех p_k , для которых $E_k \subset A$.

Наиболее распространены такие случайные эксперименты с конечным числом исходов, в которых элементарные события равновероятны. Эта равновероятность есть следствие определенной симметрии условий эксперимента по отношению к отдельным элементарным событиям (такая симметрия приводит к тому, что элементарные события выступают в эксперименте как равновозможные). Например, для симметричной монеты равновозможны выпадения обеих сторон, для шаров в урне после тщательного их перемешивания равновозможны извлечения каждого из шаров.

Определение вероятности, основанное на предположении о равновероятности (или равновозможности) элемен-

тарных событий, называется классическим. Чтобы определить в этом случае вероятность одного элементарного события, достаточно знать их число. Если всего n элементарных событий, то исходя из формулы (1), видим

$$1 = P(E_1) + \dots + P(E_n) = nP(E_1).$$

Значит,

$$P(E_1) = \dots = P(E_n) = \frac{1}{n}.$$

Предположим, что событие A есть сумма m различных элементарных событий. Тогда вероятность события A равна сумме вероятностей этих событий, а так как вероятность каждого из элементарных событий есть $\frac{1}{n}$, то

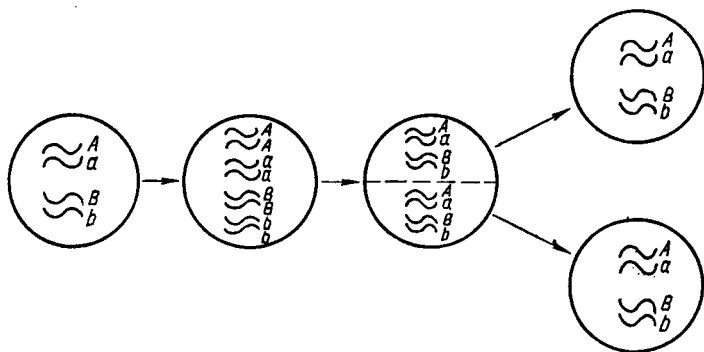
$$P(A) = \frac{m}{n}.$$

Таким образом, при классическом определении вероятности вероятность события A есть отношение числа элементарных событий, при которых наступает A , к числу всех элементарных событий. (Говорят еще так: отношение числа благоприятствующих событию A шансов к числу всех шансов.)

Приведенное определение вероятности показывает, насколько важно в случайном эксперименте правильно определить элементарные события. Если бы в примере с черными и белыми шарами в урне мы бы считали, что есть лишь два различных элементарных события, то вероятности извлечений как белого, так и черного шара были бы равны $\frac{1}{2}$ и не зависели бы от числа тех и других шаров. Если же считать извлечение каждого из шаров равновероятным (так должно быть, если шары тщательно перемешаны), то тогда число элементарных событий равно $n + m$ (общее число шаров), из этих элементарных событий n влекут событие «извлечен белый шар». Значит, вероятность извлечь белый шар равна $\frac{n}{n+m}$. Аналогично находим, что вероятность извлечь черный шар равна $\frac{m}{n+m}$.

Рассмотрим некоторые следствия из приведенного определения для задачи о передаче наследственных признаков от родителей к потомкам. Для этого нам понадобятся некоторые сведения из генетики.

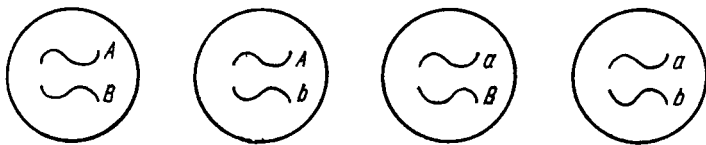
Наследственные признаки передаются с помощью генов, находящихся в хромосомах ядра клетки. Обычные клетки содержат четное число хромосом, которые разбиты на пары аналогичных (исключение представляют хромосомы, отвечающие за пол). При росте организма клетки делятся (это явление называется митозом). Митоз происходит следующим образом: каждая хромосома делится на две хромосомы,



совершенно подобных исходной, которые затем расходятся в две вновь образованные дочерние клетки. Схематически митоз двух пар хромосом представлен на рисунке. Одинаковыми латинскими (большими и малыми) буквами обозначены хромосомы в парах.

Итак, в результате митоза из одной клетки образуются две, у каждой из которых набор хромосом идентичен набору хромосом исходной клетки.

Иначе образуются половые клетки (мейоз). После удвоения числа хромосом клетка делится уже не на две, а на четыре половых клетки с половинным набором хромосом. Это могут быть, например, такие клетки: (они называются га-



метами). При оплодотворении две гаметы (материнская и отцовская) сливаются в одну клетку (зиготу), которая уже содержит полный набор хромосом. Половина хромосом —

отцовская, вторая — материнская. Благодаря этому дочерняя клетка наследует признаки обоих родителей.

В большинстве случаев за некоторый признак отвечают гены, находящиеся во вполне определенном месте хромосомной нити, причем они могут иметь два вида (эти виды в генетике называются аллелями). Например, у определенных видов гороха ген, отвечающий за окраску цветка, имеет два аллеля, соответствующие красному и белому цвету, за цвет горошин — тоже два (зеленый или желтый), за гладкость зерен — столько же (сморщенность или гладкость). У людей существуют два аллеля гена, отвечающего за цвет глаз (голубоглазость или кареглазость), курчавость или гладкость волос; левшой или правой будет человек, также определяют аллели одного гена. Обычно для аллелей одного гена используются те обозначения, которые мы применяли для обозначения хромосом: малые и большие одинаковые латинские буквы.

Весьма существенно, что в аналогичных хромосомах могут находиться различные аллели одного и того же гена.

Набор генов определяет генотип данного индивида. Если ограничиться одним признаком с аллелями гена А и а, то может быть 3 генотипа: АА, Аа, аа. Очень часто для проявления некоторого признака достаточно, чтобы имелся хотя бы один ген с данным аллелем. Тогда внешний вид (фенотип) индивидуумов с генотипами, например, АА и Аа одинаков. В этом случае говорят, что аллель А доминирует над а; соответствующий признак называют доминантным, а второй — рецессивным. Так, у гороха гладкость и желтый цвет зерен — доминантные признаки, а сморщенность и зеленый цвет — рецессивные. У человека карий цвет глаз доминирует над голубым, курчавость над прямизной волос.

Обычно доминантный признак (точнее, соответствующий аллель) обозначают большой буквой, рецессивный — малой. Особи (клетки) с генотипами АА и аа называются гомозиготными, а с генотипом Аа — гетерозиготными.

При мейозе и оплодотворении хромосомы делятся и комбинируются случайным образом. При этом вероятности вычисляются в соответствии с классическим определением.

Рассмотрим сначала наследственную передачу какого-нибудь одного признака. Заметим, что для гомозиготных родителей генотип потомков определяется однозначно.

Используя естественную символику, можем записать

$$AA + AA = AA, aa + aa = aa, AA + aa = Aa.$$

Более сложный результат, если один из родителей гетерозиготен.

Рассмотрим потомство родителей с генотипами AA и Aa. В результате мейоза получатся клетки у первого родителя с генотипами A (четыре), у второго — A (две) и a (две). Возможные результаты их соединений удобно представить в виде таблицы:

	A	A	A	A
A	AA	AA	AA	AA
A	AA	AA	AA	AA
a	Aa	Aa	Aa	Aa
a	Aa	Aa	Aa	Aa

Таким образом, у нас имеется 16 элементарных событий. В половине из них генотип потомка будет AA, в половине — Aa. Следовательно, вероятность генотипа AA есть $\frac{8}{16} = \frac{1}{2}$, точно так и вероятность генотипа Aa равна $\frac{1}{2}$.

Аналогично можно рассмотреть случай потомства гетерозиготных родителей:

	A	A	a	a
A	AA	AA	Aa	Aa
A	AA	AA	Aa	Aa
a	Aa	Aa	aa	aa
a	Aa	Aa	aa	aa

Опять из 16 элементарных событий в 4 случаях будет генотип AA и в 4 — aa, а в 8 случаях — генотип Aa. Таким образом, вероятности гомозиготных потомков — $\frac{1}{4}$, а гетерозиготных — $\frac{1}{2}$.

Эти факты были впервые обнаружены в прошлом столетии экспериментальным путем выдающимся естествоиспытателем Г. Менделем. Он установил, что при скрещивании желтого гороха с зеленым получают желтые горошины,

т. е. желтый цвет доминирует над зеленым. Если полученные в результате скрещивания растения самоопыляются, то среди полученных семян примерно $\frac{3}{4}$ будут желтыми, а $\frac{1}{4}$ — зелеными.

Действительно, первоначально были горошины с генотипами AA и aa (A — желтый, a — зеленый цвет семян). После скрещивания

$$AA + aa = Aa$$

получили семена с генотипом Aa. Они желтого цвета, поскольку он доминирует. При самоопылении растений с генотипом Aa мы будем иметь с вероятностью $\frac{1}{4}$ генотип AA, с вероятностью $\frac{1}{2}$ — генотип Aa, с вероятностью $\frac{1}{4}$ — генотип aa. Только в последнем случае окраска горошин зеленая, при генотипах AA и Aa (т. е. с вероятностью $\frac{3}{4}$) — желтая.

Аналогичные результаты получались при рассмотрении наследования других признаков (например, гладкость или морщинистость гороховых зерен). Так был открыт Менделем один из основных законов генетики: к потомку от родителей переходит с вероятностью $\frac{1}{2}$ каждый из признаков, имеющихся в генотипе родителей.

Этот закон позволяет проверять гомозиготность некоторой популяции по некоторому признаку. Как мы видели, при явлении доминирования по внешнему виду (фенотипу) невозможно отличить особи с генотипом AA от особей с генотипом Aa. Однако при скрещивании последних друг с другом с вероятностью $\frac{1}{4}$ получают потомки с генотипом aa (т. е. примерно в $\frac{1}{4}$ случаев, что можно обнаружить в результате достаточного числа экспериментов), который уже и внешне отличается от родителей. Так, если у кареглазых родителей рождается голубоглазый ребенок, то они оба гетерозиготны по признаку цвета глаз.

§ 3. Комбинаторные методы определения вероятности

При классическом определении вероятностей нужно знать число всех элементарных событий и число элементарных событий, которые влекут данное событие. Только в немногих экспериментах эти числа заданы непосредственно (как, например, при извлечении одного шара из урны, содержащей черные и белые шары). Обычно же приходится подсчитывать и число способов осуществления эксперимента, и число способов осуществления события, вероятность которого мы определяем.

Рассмотрим такой эксперимент. Из урны, содержащей 3 белых и 4 черных шара, извлечено последовательно 2 шара. Определим вероятности событий: A_1 — оба шара белых, A_2 — один черный, один белый, A_3 — оба шара черных. Пронумеруем шары, считая, что номера 1, 2, 3 соответствуют белым, 4, 5, 6, 7 — черным. Будем записывать результаты эксперимента парой чисел: например (3,6) означает, что первый извлеченный шар был с номером 3, второй имел номер 6.

Тогда все возможные результаты эксперимента можно представить в виде такой таблицы:

(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)
(2,1)	(2,3)	(2,4)	(2,5)	(2,6)	(2,7)
(3,1)	(3,2)	(3,4)	(3,5)	(3,6)	(3,7)
(4,1)	(4,2)	(4,3)	(4,5)	(4,6)	(4,7)
(5,1)	(5,2)	(5,3)	(5,4)	(5,6)	(5,7)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,7)
(7,1)	(7,2)	(7,3)	(7,4)	(7,5)	(7,6)

Как видим, всего 42 элементарных события. Событию A_1 благоприятствуют 6 элементарных событий, событию A_2 — 24, событию A_3 — 12. Следовательно,

$$P(A_1) = \frac{6}{42} = \frac{1}{7}, \quad P(A_2) = \frac{24}{42} = \frac{4}{7};$$

$$P(A_3) = \frac{12}{42} = \frac{2}{7}.$$

Если бы мы каждый раз выписывали все элементарные события, а затем из них отбирали те, которые влекут

данное событие, то во многих случаях даже вычислительные машины не помогли бы определить нужные вероятности. Укажем для примера, что при мейозе человеческой клетки, содержащей 23 пары хромосом, она может разделиться на 2 гаметы примерно 10 000 000 способов, значит, будет 100 000 000 000 000 способов образования зиготы.

Комбинаторные методы позволяют непосредственно вычислять число способов, с помощью которых может осуществиться данная комбинация. Обычно комбинаторные задачи формулируются как задачи о выборе из некоторой совокупности предметов.

Перестановки. Имеется n предметов. Сколькими способами их можно расположить в ряд (один за другим) или занумеровать?

Например, 2 предмета a и b можно расположить в ряд двумя способами: a, b или b, a . Число таких расположений n предметов (они называются перестановками) равно $n! = 1 \cdot 2 \cdot \dots \cdot n$ (произведение чисел от 1 до n , оно обозначается $n!$, читается « n -факториал»). Число перестановок из 5 предметов будет $5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$.

Установим общую формулу. Пусть P_n обозначает число перестановок из n предметов a_1, \dots, a_n . Если мы уже знаем, что на первом месте находится a_i , то на остальных местах могут быть остальные $n - 1$ предмета. Число таких перестановок равно числу перестановок из $n - 1$ предмета, т. е. P_{n-1} . Итак, имеется P_{n-1} перестановок, у которых на первом месте находится a_1 , P_{n-1} перестановок, у которых на первом месте a_2 , и т. д., P_{n-1} перестановка, у которых на первом месте a_n . Все перестановки разбиваются на n групп, в i -й группе на первом месте находится a_i , и число таких перестановок P_{n-1} . Значит, $P_n = nP_{n-1}$. Тогда и $P_{n-1} = (n - 1) P_{n-2}, \dots, P_2 = 2P_1$. Поэтому

$$P_n = nP_{n-1} = n(n - 1) P_{n-2} = n(n - 1)(n - 2) P_{n-3} = \\ = \dots = n(n - 1) \dots 2P_1$$

Остается заметить, что $P_1 = 1$.

В качестве примера применения предыдущей формулы рассмотрим такую задачу. В связке имеется n внешне не отличимых ключей, среди них только один, неизвестно какой, подходит к замку. Вы по очереди пробуете ключи. Какова вероятность, что вы добьетесь успеха во время i -го испытания? Будем считать, что ключи пронумерованы, нужный ключ имеет номер 1 (естественно, что номера ключей

чей нам на самом деле неизвестны). Ключи могут проверяться в любом порядке. Тогда общее число способов перебора ключей $n!$ Способы, благоприятствующие нашему событию, — это те перестановки, у которых ключ 1 находится в очереди на проверку i -м. При этом остальные ключи могут быть перепробованы в произвольном порядке, т. е. $(n - 1)!$ Поэтому искомая вероятность

$$\frac{(n - 1)!}{n!} = \frac{1}{n}.$$

Таким образом, с равными вероятностями мы можем отпереть замок на каждом из n испытаний.

Рассмотрим такую же задачу, полагая, что имеется 2 правильных ключа — 1 и 2. Мы добьемся успеха во время i -го испытания, если ключ 1 будет иметь номер i , а ключ 2 — номер $k > i$ (в противном случае мы бы имели успех раньше) или ключ 2 будет иметь номер i , а ключ 1 — номер $k > i$. Пусть A_{ik} событие, заключающееся в том, что ключ 1 имеет номер i , а ключ 2 — номер k . Поскольку остальные ключи могут иметь произвольные $n - 2$ номера, то число перестановок, при которых происходит A_{ik} , равно $(n - 2)!$ Значит,

$$P(A_{ik}) = \frac{(n - 2)!}{n!} = \frac{1}{n(n - 1)}.$$

Но интересующее нас событие есть сумма таких событий:

$$A_{i+1} + A_{i+2} + \dots + A_{in} + A_{i+1i} + A_{i+2i} + \dots + A_{ni}.$$

Всего в этой сумме $2(n - i)$ несовместимых событий, каждое из которых имеет вероятность $\frac{1}{n(n - 1)}$.

Значит, искомая вероятность равна

$$\frac{2(n - i)}{n(n - 1)}.$$

Выборки. Из n предметов выбирается m по очереди. Сколькими способами можно сделать выбор?

Каждый такой выбор будем называть выборкой размера m и n элементов. Из соображений симметрии вытекает, что если считать все перестановки равновероятными, то равновероятны и события, заключающиеся в том, что

заданные m предметов занимают в заданном порядке первые m мест в выборке. Действительно, количество тех перестановок, где зафиксированы предметы на m местах, совпадает с числом перестановок из оставшихся $n - m$ предметов, которые могут произвольным образом располагаться на остальных $n - m$ местах.

Значит, вероятность любой выборки m предметов равна

$$\frac{(n - m)!}{n!} = \frac{1}{n(n - 1) \dots (n - m + 1)}.$$

Число же различных выборок из n элементов размера m есть величина обратная вероятности, так как, если имеется полная группа равновероятных событий, то вероятность одного события есть величина, обратная их числу.

Итак, число выборок из n элементов размера m есть

$$n(n - 1) \dots (n - m + 1) = \frac{n!}{(n - m)!}$$

Выборки с возвращением. Представим себе, что в урне имеются n пронумерованных шаров. Мы извлекаем шар, отмечаем его номер, а затем возвращаем шар в урну, после чего шары опять тщательно перемешиваются. Повторив эту процедуру m раз, мы получим выборку с возвращением (или с повторением): в ней, в отличие от предыдущего случая, один и тот же предмет (точнее, его символ или номер) может встречаться несколько раз. Заметим, что число выборок с возвращением совпадает с числом способов, которыми заданные m предметов могут быть разложены по n ящикам (можно считать, что если при k -м испытании был извлечен шар с номером i , то k -й предмет попал в i -й ящик).

Обозначим через N_m число выборок с возвращением размера m . Пусть $N_m^{(i)}$ — число тех выборок, у которых на первом месте i -й шар. Поскольку на остальных $m - 1$ местах могут быть любые шары, то $N_m^{(i)}$ равно числу выборок без возвращения размера $m - 1$: $N_m^{(i)} = N_{m-1}$.

Значит,

$$N_m = N_m^{(1)} + \dots + N_m^{(n)} = nN_{m-1}.$$

Так как $N_1 = n$, то $N_2 = n \cdot n = n^2$, $N_3 = n \cdot n^2 = n^3$, ..., $N_m = n^m$.

Рассмотрим такую задачу. Какова вероятность, что в группе людей, родившихся в невисокосном году (напри-

мер, в группедетсада или в классе школы), у кого-нибудь совпадают дни рождения? Вероятность этого события зависит от числа людей. Предположим, что их m человек. Число способов, которыми могут расположиться m дней рождения по 365 дням, равно числу выборок с возвращениями из 365 размера m , т. е. 365^m . Число таких расположений, при которых все дни рождения различны, будет равно числу выборок (без повторений) из 365 размера m , т. е.

$$365 \cdot 364 \dots (365 - m + 1).$$

Таким образом, вероятность того, что все дни рождения различны, равна

$$\frac{365 \cdot 364 \dots (365 - m + 1)}{365^m} = \left(1 - \frac{1}{365}\right) \left(1 - \frac{2}{365}\right) \dots \dots \left(1 - \frac{m-1}{365}\right).$$

Вероятность того, что по крайней мере у двух людей из группы дни рождения совпадают, составляет

$$1 - \left(1 - \frac{1}{365}\right) \left(1 - \frac{2}{365}\right) \dots \dots \left(1 - \frac{m-1}{365}\right).$$

Например, при $m = 10$ эта вероятность равна приблизительно 0,12, при $m = 20$ — около 0,5, а при $m = 30$ — уже 0,7. Таким образом, в школьных классах должно довольно часто наблюдаться такое событие, как совпадение дней рождения учеников.

Сочетания. Рассмотрим теперь, сколькими способами из n предметов можно выбрать группу в m предметов (порядок в группе нас не интересует). Такие выборы называются сочетаниями из n элементов по m . Число сочетаний обозначается символом C_n^m (из n элементов по m). Вычислим это число.

Проанализируем вероятностный эксперимент, элементарными событиями у которого являются выборки (без возвращений) из n элементов размера m . Число элементарных событий равно $n(n-1)\dots(n-m+1)$. Пусть A — событие, заключающееся в том, что в результате выборки выбраны заданные m элементов (например, a_1, \dots, a_m) в произвольном порядке. Эти элементы можно выбрать стольким числом способов, сколько будет перестановок

из m элементов, т. е. $m!$ способов. Поэтому вероятность события A — получить заданное сочетание — будет

$$\frac{m!}{n(n-1)\dots(n-m+1)} = \frac{m!(n-m)!}{n!}.$$

Поскольку вероятности всех сочетаний равны, то вероятность некоторого сочетания есть $\frac{1}{C_n^m}$, следовательно,

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

Размещения по ячейкам. В статистической физике состояние систем большого числа частиц характеризуется положением в фазовом пространстве отдельных частиц. Все фазовое пространство разбивается на большое число частей (ячеек), и состояние системы описывается размещением частиц по ячейкам. В зависимости от того, какие размещения считаются равновероятными, получаются различные статистики (другими словами, по-разному вычисляются вероятности различных событий, а следовательно, если эти вероятности соответствуют реальности, то и наблюдаемые частоты событий будут различны). Различные типы частиц подчиняются различным статистикам.

Пусть имеется n частиц и N ячеек.

Статистика Максвелла — Больцмана получается, если считать равновероятными все способы размещения частиц по ячейкам. При этом частицы выступают как индивидуальные (они различимы одна от другой). Число размещений, как указывалось при рассмотрении выборок с возвращением, равно числу выборок с возвращением из N элементов по n , т. е. N^n .

Статистика Максвелла — Больцмана может быть справедлива для макрообъектов. При переходе к микрообъектам предположение о различимости объектов неестественно. Впервые обратил на это внимание Эйнштейн.

Статистика Бозе — Эйнштейна. Будем считать, что равновероятны состояния системы, которые определяются набором чисел n_1, n_2, \dots, n_N , где n_i — число частиц в i -й ячейке (n_i может обращаться в 0, если в i -й ячейке нет частиц, $n_1 + \dots + n_N = n$). Подсчитаем число состояний системы. Удобно каждому состоянию сопоставить последовательность из n нулей и $N - 1$ единицы по следующему правилу: пишем n_1 нулей, затем единицу, n_2 нулей, единицу, ...

и т. д., n_N нулей. Если где-нибудь идут подряд 2 единицы, то соответствующее n_i равно нулю, т. е. i -я ячейка пуста.

Пусть, например, $n = 7$, $N = 4$. Тогда последовательности 0 100 100 001 отвечает такое распределение по ячейкам 1, 2, 4, 0.

Таким образом, число состояний системы равно числу последовательностей из n нулей и $N - 1$ единицы. А их число равно числу способов, которыми можно из $N + n - 1$ элементов (чисел) выбрать $N - 1$ элемент (номера, на которых стоят единицы). Это число равно числу сочетаний из $N + n - 1$ элемента по $N - 1$, т. е.

$$C_{N-n+1}^{N-1} = \frac{(N-n+1)!}{(N-1)!n!}.$$

Приведем для примера сравнение вероятностей различных состояний по статистикам Максвелла — Больцмана и Бозе — Эйнштейна для размещений 5 частиц по 6 ячейкам.

Размещение по ячейкам	Число благоприятных элементарных событий	Вероятность (B_1) по Максвеллу — Больцману	Вероятность (B_2) по Бозе — Эйнштейну	B_1/B_2
111110	$5! = 120$	0,015	0,004	4
211100	60	0,008	0,004	2
221000	30	0,004	0,004	1
311000	20	0,003	0,004	0,7
320000	10	0,001	0,004	0,25
410000	5	0,0006	0,004	0,012
500000	1	0,0001	0,004	0,025

Как видим, вероятности различных состояний по различным статистикам существенно различаются. При статистике Максвелла — Больцмана более вероятны равномерные распределения, а при статистике Бозе — Эйнштейна — неравномерные. Статистике Бозе — Эйнштейна подчиняются фотоны, атомные ядра, содержащие четное число протонов и нейтронов.

Для других типов частиц (электроны, протоны, нейтроны) нужно использовать статистику Ферми — Дирака. В этом случае обязательно $n \leq N$ и возможны лишь такие состояния системы, при которых в каждой ячейке не более одной частицы. Все такие состояния считаются

равновероятными. Число их совпадает с числом возможных выборов из N ячеек n числа ячеек, занятых частицами. Поэтому оно равно числу сочетаний из N по n , т. е. C_N^n ; вероятность одного состояния равна $1/C_N^n$. Например, при $N = 6$, $n = 5$ вероятность одного состояния есть $1/C_6^5 = \frac{1}{6}$. Такую вероятность имеет первое из приведенных в таблице состояний, для остальных вероятности равны нулю. (В таблице указаны далеко не все возможные состояния, а только типичные, характеризующиеся различными распределениями по ячейкам — таких состояний, при которых в одной ячейке 2 частицы, а в трех — по 1, будет при статистике Бозе — Эйнштейна 60.)

Любопытно отметить, что при увеличении N вероятность события «в каждой ячейке не более одной частицы», в случае статистики Бозе — Эйнштейна равная

$$\begin{aligned} \frac{C_N^n}{C_{N+n-1}^{N-1}} &= \frac{N!(N-1)!n!}{n!(N-n)!(N+n-1)!} = \\ &= \frac{(N-1) \dots (N-n+1)}{(N+1) \dots (N+n-2)}, \end{aligned}$$

приближается к единице, и если $N \rightarrow \infty$, то стремится к единице. Поэтому для «разреженных» систем, для которых $\frac{n}{N}$ мало, статистики Бозе — Эйнштейна и Ферми — Дирака практически совпадают.

§ 4. Независимость случайных событий

Понятие независимости — очень важное в теории вероятностей. Оно дает возможность, зная вероятности событий, найти вероятности их совместного осуществления. Это понятие довольно точно отражает понятие независимости некоторых явлений в обычном смысле, т. е. отсутствие влияния одних событий на другие. Хотя, как увидим ниже, обычное понятие независимости и теоретико-вероятностная независимость отличаются друг от друга некоторыми сторонами.

Рассмотрим предварительно следующий простой пример. Пусть имеется урна с белыми и черными шарами. Событие A заключается в том, что в первый раз из урны будет извлечен черный шар, событие B — в том, что во

второй раз будет извлечен черный шар. Если шары извлекаются друг за другом (без возвращения), то событие A существенно влияет на событие B . Действительно, если в первый раз вынули черный шар, то число черных шаров уменьшилось, а значит, соответственно уменьшилась и вероятность события B (например, если число черных шаров равно 1, то после события A событие B не может произойти). Если же первый раз был вынут белый шар, то уменьшилось число белых шаров при том же количестве черных, следовательно, вероятность B увеличилась (так, при одном белом шаре после его извлечения событие B становится достоверным).

Теперь обратимся к другому способу извлечения шаров — извлечению с возвращением: вынутый шар возвращается в урну, после чего шары тщательно перемешивают. Таким образом, перед повторным извлечением шара урна находится точно в таком же состоянии, в каком она была перед первым извлечением, и то, какой шар вынули в первый раз, никак не влияет на возможность появления события B . Поэтому естественно считать, что события A и B независимы.

Рассмотрим два эксперимента, в каждом из которых имеется конечное число равновероятных событий: n_1 — в первом, n_2 — во втором. Будем производить оба эксперимента одновременно (точнее, будем одновременно рассматривать события, которые произошли в каждом эксперименте, они могут быть такими, что один обязательно происходит после другого, как в рассмотренном выше примере повторного извлечения шаров). Пусть $E_1^{(1)}, \dots, E_{n_1}^{(1)}$ — элементарные события в первом эксперименте, $E_1^{(2)}, \dots, E_{n_2}^{(2)}$ — во втором. Тогда в составном эксперименте, состоящем в том, что производятся оба эксперимента, происходят элементарные события $E_k^{(1)} \cdot E_l^{(2)}$ (т. е. $E_k^{(1)}$ произошло в первом, $E_l^{(2)}$ — во втором эксперименте). Если результаты одного эксперимента не влияют на результаты другого, то при любом $E_k^{(1)}$ в первом эксперименте может произойти любое $E_l^{(2)}$ во втором эксперименте. Всего будет $n_1 \cdot n_2$ элементарных событий составного эксперимента, при этом они будут равновероятны. Таким образом, вероятность элементарного события в составном эксперименте будет $\frac{1}{n_1 n_2} = \frac{1}{n_1} \cdot \frac{1}{n_2}$. Первый сомножитель обозначает вероят-

ность элементарного события в первом эксперименте, второй — во втором. Таким образом, имеем

$$P(E_k^{(1)} \cdot E_i^{(2)}) = P(E_k^{(1)}) \cdot P(E_i^{(2)}).$$

Это свойство перемножения вероятностей для произведения событий и кладется в основу определения независимости двух событий.

События A и B называются независимыми, если

$$P(A \cdot B) = P(A)P(B). \quad (1)$$

Покажем, что если событие A наблюдается в первом эксперименте, а событие B — во втором, то они независимы. Пусть A осуществляется, когда происходит одно из m_1 элементарных событий первого эксперимента, а B — одно из m_2 элементарных событий второго эксперимента. Тогда событие $A \cdot B$ в составном эксперименте осуществляется, если происходит одно из элементарных событий $E_k^{(1)} \cdot E_i^{(2)}$, где k принимает m_1 различных значений, а i — m_2 . Поэтому число таких элементарных событий равно $m_1 \cdot m_2$. Учитывая, что общее число элементарных событий составляет $n_1 n_2$, получим в соответствии с классическим определением вероятности

$$P(A \cdot B) = \frac{m_1 m_2}{n_1 n_2} = \frac{m_1}{n_1} \cdot \frac{m_2}{n_2}.$$

Но $\frac{m_1}{n_1}$ и $\frac{m_2}{n_2}$ есть вероятности событий A и B соответственно в первом и втором экспериментах. Значит, действительно выполнено (1), т. е. события A и B независимы. При этом мы также говорим, что A не зависит от B , B не зависит от A .

Покажем, что если A не зависит от B , то A не зависит и от \bar{B} (в обычном понимании независимости так должно быть: A не зависит от того, произошло ли B или не произошло).

Имеем

$$P(A \cdot B) + P(A \cdot \bar{B}) = P(A);$$

$$\begin{aligned} P(A \cdot \bar{B}) &= P(A) - P(A \cdot B) = P(A) - P(A)P(B) = \\ &= P(A)(1 - P(B)) = P(A)P(\bar{B}). \end{aligned}$$

Аналогично устанавливается, что тогда независимы и такие пары событий: \bar{A}, B ; \bar{A}, \bar{B} .

Укажем на кажущееся отличие понятий теоретико-вероятностной независимости от независимости в обычном понимании. Так как при обычном понимании будущее не влияет на прошлое, то, если эксперименты выполняются один за другим во времени, первый кажется независимым от второго, хотя второй и может зависеть от первого. Таким образом, обычное понятие независимости (зависимости) несимметрично, тогда как в теории вероятности оно симметрично. Однако на самом деле здесь нет противоречия, если учесть, что, зная события в будущем, можно получить определенную информацию о том, что происходило в прошлом (на этом основаны некоторые исторические исследования, относящиеся к далекому прошлому).

Проанализируем определение независимости нескольких событий. Здесь мы встретимся с существенным отличием обычного и теоретико-вероятностного понятий. В обычном понимании, если событие A не зависит от B и не зависит от C , то A не зависит и от $B \cdot C$, т. е. от того, что произошли одновременно и B и C .

Приведем пример, показывающий, что для теоретико-вероятностной независимости такое положение вещей не обязательно. Рассмотрим тетраэдр, стороны которого окрашены в разные цвета: первая — в красный, вторая — в зеленый, третья — в синий, а четвертая — во все три. Тетраэдр бросают на плоскость, и он с одинаковыми вероятностями падает на каждую из граней. Обозначим через K, Z, C события: на той грани, на которую упал тетраэдр, имеется красный (зеленый, синий) цвет.

В этом эксперименте 4 элементарных события. Каждое из событий K, Z, C происходит в двух случаях: либо если тетраэдр падает на грань соответствующего цвета, либо когда он падает на трехцветную грань. Поэтому

$$P(K) = P(Z) = P(C) = \frac{2}{4} = \frac{1}{2}.$$

Каждое из событий $K \cdot Z, K \cdot C, Z \cdot C, K \cdot Z \cdot C$ происходит тогда, когда тетраэдр падает на трехцветную грань, т. е. в одном случае из четырех:

$$P(K \cdot Z) = P(K \cdot C) = P(Z \cdot C) = P(K \cdot Z \cdot C) = \frac{1}{4}.$$

Поэтому

$$P(K \cdot Z) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(K) \cdot P(Z),$$

$$P(K \cdot C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(K) P(C).$$

Таким образом, событие K не зависит как от события Z , так и от события C . При этом и события Z и C также независимы:

$$P(Z \cdot C) = \frac{1}{4} = P(Z) \cdot P(C).$$

Но событие K уже зависит от события $Z \cdot C$:

$$P(K \cdot Z \cdot C) = \frac{1}{4} \neq \frac{1}{8} = P(K) \cdot P(Z \cdot C).$$

(Если произошло событие $Z \cdot C$, то тетраэдр упал на трехцветную грань, а значит, событие K также произошло.)

Поэтому независимость даже трех событий нуждается в дополнительном определении. События A , B , C независимы, если независимы каждая пара из этих событий и, кроме того,

$$P(A \cdot B \cdot C) = P(A) \cdot P(B) \cdot P(C). \quad (2)$$

Так что для независимости трех событий, кроме соотношения (2), нужно выполнение еще трех таких соотношений:

$$P(A \cdot B) = P(A) \cdot P(B),$$

$$P(A \cdot C) = P(A) \cdot P(C),$$

$$P(B \cdot C) = P(B) \cdot P(C).$$

Дальнейшее распространение понятия независимости на большее число событий проводится индуктивно: при определении независимости n событий уже используется независимость $n - 1$ события (при определении независимости трех событий использовалась независимость двух).

Так, для четырех событий A_1, A_2, A_3, A_4 : они будут независимы, если независимы каждые из трех событий и

$$P(A_1 \cdot A_2 \cdot A_3 \cdot A_4) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot P(A_4).$$

n событий A_1, \dots, A_n независимы, если независимы каждые из $n - 1$ событий и, кроме того

$$P(A_1 \cdot A_2 \cdot \dots \cdot A_n) = P(A_1) P(A_2) \dots P(A_n).$$

Оказывается, независимость событий A_1, \dots, A_n влечет независимость и такой последовательности событий, которая получается, если некоторые заменены на противоположные. Так, в случае трех событий независимы каждые из трех событий:

$$\begin{aligned} \bar{A}_1, A_2, A_3; \quad A_1, \bar{A}_2, A_3; \quad A_1, A_2, \bar{A}_3; \quad \bar{A}_1, \bar{A}_2, \bar{A}_3; \\ \bar{A}_1, A_2, \bar{A}_3; \quad A_1, \bar{A}_2, \bar{A}_3; \quad \bar{A}_1, \bar{A}_2, \bar{A}_3. \end{aligned}$$

Поскольку любую такую последовательность событий можно получить, заменяя в последовательности событий лишь одно на противоположное (такую операцию можно производить несколько раз), то для доказательства утверждения достаточно показать, что независимы события $\bar{A}_1, A_2, \dots, A_n$. Покажем это для случая $n = 3$, т. е. что из независимости A_1, A_2, A_3 — вытекает независимость \bar{A}_1, A_2, A_3 . События \bar{A}_1 и A_2 , \bar{A}_1 и A_3 независимы, так как независимы A_1 и A_2 , A_1 и A_3 (из независимости трех событий вытекает независимость каждой пары). Далее

$$P(\bar{A}_1 \cdot A_2 \cdot A_3) + P(A_1 \cdot A_2 \cdot A_3) = P(A_2 \cdot A_3),$$

так как события $\bar{A}_1 \cdot A_2 \cdot A_3$ и $A_1 \cdot A_2 \cdot A_3$ несовместимы и

$$\bar{A}_1 \cdot A_2 \cdot A_3 + A_1 \cdot A_2 \cdot A_3 = A_2 \cdot A_3.$$

Значит,

$$\begin{aligned} P(\bar{A}_1 \cdot A_2 \cdot A_3) &= P(A_2 \cdot A_3) - P(A_1 \cdot A_2 \cdot A_3) = \\ &= P(A_2)P(A_3) - P(A_1)P(A_2)P(A_3) = \\ &= (1 - P(A_1))P(A_2)P(A_3) = P(\bar{A}_1)P(A_2)P(A_3). \end{aligned}$$

Рассмотрим наследование двух признаков, определяемых генами, расположенными в различных парах хромосом. Интерес представляет тот случай, когда по обоим признакам родители гетерозиготны. Будем обозначать аллели соответствующих генов A, a (первый признак), B, b (второй признак).

В результате мейоза возможны такие четыре комбинации генов в гаметах:

$$(A; B), \quad (A; b), \quad (a; B), \quad (a; b).$$

Поэтому после слияния гамет получаем следующие 16 возможных видов зигот (на первом месте — ген, переданный отцом, на втором — матерью):

(AA; BB)	(AA; Bb)	(Aa; BB)	(Aa; Bb)
(AA; bB)	(AA; bb)	(Aa; bB)	(Aa; bb)
(aA; BB)	(aA; Bb)	(aa; BB)	(aa; Bb)
(aA; bB)	(aA; bb)	(aa; bB)	(aa; bb)

Таким образом, с вероятностью $\frac{1}{16}$ потомок будет гомозиготным по обоим признакам (генотипы AA; BB, AA; bb, aa; BB, aa; bb), с вероятностью $\frac{1}{4}$ — гетерозиготен по обоим признакам (генотип Aa; Bb), с вероятностями $\frac{1}{8}$ — гомозиготен по одному признаку и гетерозиготен по другому (генотипы AA; Bb, aa; Bb, Aa; BB, Aa; bb). Значения этих вероятностей определяются подсчетом числа представленных в таблице элементарных событий, приводящих к данному генотипу. Легко проверить, что для любой комбинации генов у потомка справедливо равенство: вероятность данного генотипа равна произведению вероятностей комбинаций генов по каждому признаку, например

$$P(AA; BB) = P(AA) \cdot P(BB), P(Aa; Bb) = P(Aa) P(Bb).$$

Здесь $P(AA)$, $P(Bb)$ и т. п. обозначают вероятности того, что у потомка по первому признаку генотип AA или по второму признаку — Bb. (Вероятности таких комбинаций вычислялись в § 2.) В частности, показано, что $P(AA) = P(BB) = \frac{1}{4}$.

Так что

$$P(AA; BB) = P(AA) P(BB) = \frac{1}{16} = \frac{1}{4} \cdot \frac{1}{4}.$$

Там же найдено, что $P(Aa) = \frac{1}{2}$, а значит, и $P(Bb) = \frac{1}{2}$. Отсюда

$$P(Aa; Bb) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(Aa) P(Bb).$$

Итак, мы пришли ко второму закону Менделя: при передаче от родителей к потомкам признаков, определяемых генами в различных парах хромосом, они передаются независимо друг от друга. Это справедливо и для большего числа признаков, чем два.

В качестве примера рассмотрим потомство кареглазых курчавых родителей, если они гетерозиготны по этим признакам (кареглазость доминирует над голубоглазостью, а курчавость — над прямоволосостью). Если А — ген кареглазости, а — голубоглазости, В — курчавости, в — прямоволосости, то потомки по фенотипу разделяются на такие группы: 1) кареглазые курчавые; 2) кареглазые прямоволосые; 3) голубоглазые курчавые; 4) голубоглазые прямоволосые.

Первому фенотипу отвечают генотипы АА; ВВ, Аа; Вв, АА; Вв, Аа; ВВ. Поэтому вероятность появления таких потомков есть

$$P(АА)P(ВВ) + P(Аа)P(Вв) + P(АА)P(Вв) + \\ + P(Аа)P(ВВ) = \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} + \\ + \frac{1}{4} \cdot \frac{1}{2} = \frac{9}{16}.$$

Второму фенотипу — соответствуют генотипы АА; вв, Аа; вв; третьему — аа; ВВ, аа; Вв. Вероятности обоих одинаковы и равны $\frac{1}{16} + \frac{1}{8} = \frac{3}{16}$. Наконец, последний имеет генотип аа; вв, вероятность его $\frac{1}{16}$.

§ 5. Геометрические методы. Задача о встрече

В § 3 были рассмотрены случайные эксперименты, имеющие конечное число исходов. Однако они далеко не охватывают всех возможных случайных экспериментов. Во многих случаях возможные исходы образуют бесконечную непрерывную совокупность.

Самый простой пример — бросаем на пол шарик и смотрим, какой точкой поверхности он прикоснется к поверхности пола. Это элементарное событие в нашем эксперименте. Множество таких элементарных событий совпадает со множеством точек поверхности сферы (шарика).

Можно ли здесь разумно использовать понятие равновероятности для определения вероятности различных событий? Какого рода события нужно рассматривать в этом эксперименте? Можем считать, что шарик окрашен в различные цвета и нас интересует цвет, которым шарик коснется пола впервые.

Давайте изменим эксперимент: шарик будет неподвижен, а на него случайным образом опускается точка. Пусть, например, шарик — наша Земля, на которую падает метеорит. Нас может интересовать, упадет ли он на воду или сушу, если на сушу, то в пустыню или густонаселенную местность. Вместо метеорита это может быть автоматическая космическая станция, случайно вышедшая из строя, или космический корабль инопланетян.

Вернемся к нашему шарiku. Пусть он окрашен в два цвета (черный и белый), причем цветовые пятна произвольной формы. Попробуем определить вероятность того, что шарик коснется пола черной поверхностью. Для этого представим, что шарик равномерно утыкан короткими иголками, окрашенными в цвет точки поверхности, из которой выходит иголка. Пусть общее число иголок $N = n + m$, где n — число черных, m — число белых. При бросании шарик будет касаться пола одной из иголок. Если считать касания любой иголкой равновероятными, то вероятность коснуться пола черной иголкой равна $\frac{n}{N}$. Но при больших N

отношение $\frac{n}{N}$ примерно равно отношению площади поверхности мячика, окрашенной в черный цвет, ко всей площади поверхности шарика. Это равенство будет тем более точно, чем больше N . Но если иголочки будут сплошь заполнять поверхность, то мы получим опять гладкий шарик.

Итак, мы пришли к следующему методу определения вероятности в данном эксперименте: вероятность того, что точка, случайно брошенная на шарик, попадет в заданную часть поверхности шарика, равна отношению площади этой части поверхности ко всей поверхности шарика.

Легко представить и эксперимент, когда элементарные события можно считать точками некоторой линии. Так, при бросании тонкого кольца, которое можно рассматривать как окружность, наблюдается такое событие: кольцо коснулось поверхности данной точкой своей окружности. Здесь элементарные события — точки окружности. Пусть окружность окрашена в черный и белый цвета. С помощью рассуждений, использовавшихся при анализе эксперимента с бросанием шарика, можно установить, что при бросании кольца вероятность того, что оно коснется впервые поверхности точкой, окрашенной в черный цвет, рав-

на отношении суммы длин отрезков дуг окружности, окрашенных в черный цвет, к длине всей окружности.

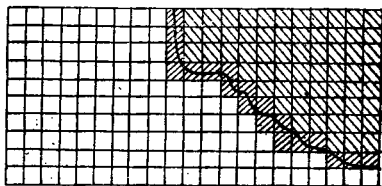
Наконец, обратимся к эксперименту, в котором элементарные события есть точки некоторого тела. Возьмем сосуд, заполненный газом. Пусть среди молекул газа есть одна «меченая». Какова вероятность того, что в определенный момент времени эта молекула попадет в заданную часть объема? Пусть, например, это молекула радиоактивного вещества. Нас интересует событие, заключающееся в том, что радиоактивное превращение произойдет в той части объема, где расположен счетчик таких частиц (только в этом случае данный факт будет зарегистрирован, и, значит, мы узнаем о присутствии в сосуде «меченой» молекулы).

Необходимость рассмотрения таких событий возникает, если для слежения за возможной утечкой некоторого газа в сосуды, содержащие этот газ (газгольдеры), добавляют радиоактивные примеси, а для обнаружения их снаружи сосудов помещают счетчики радиоактивных частиц.

Для подсчета вероятности того, что «меченая» молекула в объеме S окажется в его части S_1 , разобьем весь объем на N ячеек (частей) одинакового объема. Какая бы из статистик (о них шла речь в § 3) не использовалась, вероятность попасть в одну ячейку будет $\frac{1}{N}$. Пусть n_1 — число

ячеек, полностью попавших в S_1 , а \hat{n}_1 — число таких, которые имеют с S_1 общие точки (для пояснения приводится рисунок для плоской ситуации, ячейками служат квадратики, $n_1 = 62$, $\hat{n}_1 = 81$).

Тогда вероятность попасть в S_1 больше, чем вероятность попасть в n_1 квадратиков, и меньше, чем вероятность попасть в \hat{n}_1 квадратиков, т. е.



$$\frac{n_1}{N} \leq P(S_1) \leq \frac{\hat{n}_1}{N}.$$

Если $N \rightarrow \infty$, то оба отношения $\frac{n_1}{N}$ и $\frac{\hat{n}_1}{N}$ стремятся к отношению объема S_1 к объему S . Таким образом, вероят-

ность попадания точки (молекулы) в данную часть S_1 объема S есть отношение этих объемов.

Если эксперименты с конечным числом исходов можно представлять как случайный выбор из конечного числа предметов, то в описанных примерах эксперименты состоят в случайном выборе одной точки из совокупности точек образующих некоторую геометрическую фигуру (линию, плоскую фигуру, пространственную фигуру). Наблюдаемые события состоят в том, что выбранная точка будет принадлежать данной части фигуры. Вероятность такого события определяется как отношение мер части фигуры к всей фигуре (в зависимости от вида фигуры под «мерой» понимается длина, площадь или объем).

Приведем три наиболее характерных эксперимента связанных со случайным выбором точки.

1. Случайный выбор точки из отрезка. Выбирается точка отрезка $[a, b]$. Вероятность того, что выбранная точка попадет в отрезок $[\alpha, \beta]$, лежащий в отрезке $[a, b]$, равна $\frac{\beta - \alpha}{b - a}$.

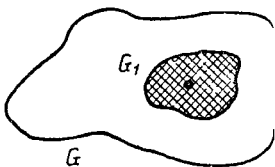
На практике со случайным выбором точки из отрезка мы встречаемся при рассмотрении ошибок измерения.

Каждый измерительный прибор имеет шкалу, (разбитую на отрезки равной длины) и некоторый указатель (стрелку). Этот указатель устанавливается при измерении между двумя делениями шкалы, например между отметками n и $n + 1$. В этом случае положение указателя можно рассматривать как случайный выбор точки из отрезка $[n, n + 1]$.

2. Случайный выбор точки из плоской фигуры. В этом случае задана некоторая ограниченная фигура G , из которой и выбирается точка. Вероятность того, что выбранная точка находится в некоторой части G_1 , образующей фигуру G , определяется так:

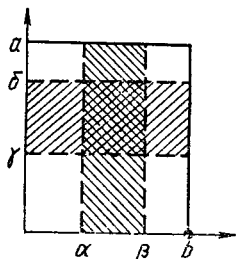
$$\frac{\text{площадь } G_1}{\text{площадь } G}$$

Рассмотрим подробнее случай, когда G — прямоугольник. Будем считать, что длины его сторон a и b . Выберем оси координат так, как указано на рисунке (с. 39). Чтобы выбрать некоторую точку P в прямоугольнике, нужно выбрать ее координат



x и y , при этом координата x есть некоторая точка на отрезке $[0, a]$, координата y — точка на отрезке $[0, b]$.

Покажем, что этот выбор следует производить независимо. Пусть A — событие, заключающееся в том, что точка P выбрана так, что координата x оказалась из отрезка $[\alpha, \beta]$ на оси x , лежащем в отрезке $[0, a]$, событие B заключается в том, что координата y оказалась из отрезка $[\gamma, \delta]$ на оси y , лежащем в отрезке $[0, b]$. Событие A происходит, если точка P попадает в прямоугольник с основанием $[\alpha, \beta]$ высоты b , а событие B , если точка P лежит в прямоугольнике, боковая сторона которого есть отрезок $[\gamma, \delta]$ на оси y , а длина основания a . Наконец, событие $A \cdot B$ происходит, если точка P лежит в пересечении этих двух прямоугольников, т. е. в прямоугольнике с длинами сторон $\beta - \alpha$ и $\delta - \gamma$ (все прямоугольники изображены на рисунке). Имеем

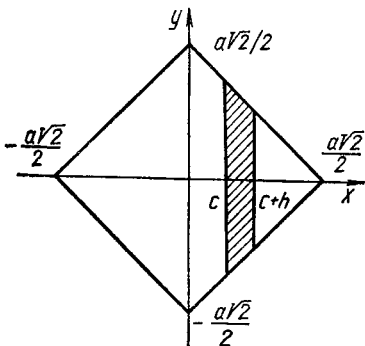


$$P(A) = \frac{(\beta - \alpha)b}{ba} = \frac{\beta - \alpha}{a}, \quad P(B) = \frac{(\delta - \gamma)a}{ab} = \frac{\delta - \gamma}{b}.$$

Эти соотношения показывают, что, выбирая случайно точку P в прямоугольнике, мы случайно выбираем ее координаты на сторонах прямоугольника. Наконец,

$$P(A \cdot B) = \frac{(\beta - \alpha)(\delta - \gamma)}{ab} = \frac{\beta - \alpha}{a} \cdot \frac{\delta - \gamma}{b} = P(A)P(B).$$

Таким образом, события A и B независимы: при случайном выборе точки из прямоугольника мы случайным образом выбираем и ее координаты.



Следует отметить, что этот результат справедлив только для прямоугольника, причем со сторонами, расположенными параллельно осям координат.

Рассмотрим в качестве примера случайный выбор из повернутого на 45° квадрата (см. рисунок). Пусть

сторона квадрата a . Событие, заключающееся в том, что координата x выбранной точки P попадет в отрезок $[c, c + h]$, происходит тогда и только тогда, когда точка P лежит в трапеции, вырезаемой из квадрата прямыми $x = c$ и $x = c + h$. Основания трапеции $2\left(\frac{a\sqrt{2}}{2} - c\right)$ и $2\left(\frac{a\sqrt{2}}{2} - c - h\right)$ (мы считаем $0 \leq c < c + h \leq \frac{a\sqrt{2}}{2}$), высота h , так что площадь равна $(a\sqrt{2} - 2c - h)h$, а вероятность события $\left(\sqrt{2} - \frac{2c}{a} - h\right) \frac{h}{a}$.

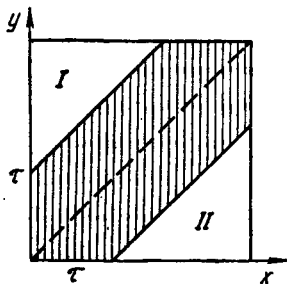
Итак, вероятность того, что координата x попадает в отрезок длины h , не пропорциональна h (если бы она выбиралась случайно из отрезка длины $a\sqrt{2}$, она бы равнялась $\frac{h}{a\sqrt{2}}$), поэтому нельзя считать, что эта координата выбрана случайно из отрезка $\left[-\frac{a\sqrt{2}}{2}, \frac{a\sqrt{2}}{2}\right]$. Очевидно также, что значение координаты x влияет на значение координаты y , так как при заданном x координата y меняется на отрезке $\left[-\left(\frac{a\sqrt{2}}{2} - |x|\right), \frac{a\sqrt{2}}{2} - |x|\right]$.

3. Задача о встрече. Пусть два объекта (субъекта) попадают на место встречи в некоторые случайные моменты времени из отрезка $[0, a]$. Каждый объект проводит там время τ (фиксированное) и затем покидает это место. При этом они прибывают на место встречи независимо один от другого. Нужно определить вероятность того, что объекты встретятся, т. е. в некоторый момент времени окажутся одновременно на месте встречи.

С такой ситуацией мы сталкиваемся, например, при рассмотрении химической реакции двух веществ в растворе в присутствии катализатора: здесь место встречи — активная зона у молекулы катализатора, а должны встретиться молекулы различных веществ. Другой пример — встреча в заданном месте (например, на водопое) двух заданных животных (хищника и жертвы или разнополых особей одного вида).

Будем отмечать моменты попадания объектов на место встречи точками x (первого) и y (второго) на двух взаимно перпендикулярных сторонах квадрата со стороной a (эти стороны можно считать расположенными на осях

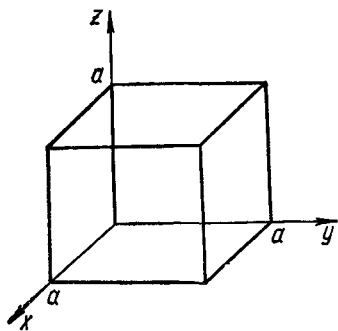
прямоугольных координат, как указано на рисунке). Встреча между объектами состоится, если расстояние между моментами появления объектов не превосходит τ , т. е. $|x - y| \leq \tau$. Множество точек квадрата, имеющих координаты (x, y) , удовлетворяющие этому неравенству, представляет собой шестиугольную фигуру, которая получается в результате пересечения квадрата и полосы, лежащей между прямыми $x - y = \tau$, $x - y = -\tau$. Искомая вероятность равна отношению площади этого шестиугольника к площади квадрата. Заметим, что два треугольника, дополняющих шестиугольник до квадрата (треугольники I и II на рисунке), составляют вместе квадрат со стороной $a - \tau$. Значит, площадь шестиугольника равна $a^2 - (a - \tau)^2$, а вероятность встречи будет



$$\frac{a^2 - (a - \tau)^2}{a^2} = 1 - \left(1 - \frac{\tau}{a}\right)^2.$$

Если $\tau = \frac{a}{3}$, то искомая вероятность равна $1 - \left(1 - \frac{1}{3}\right)^2 = \frac{5}{9}$ (больше половины).

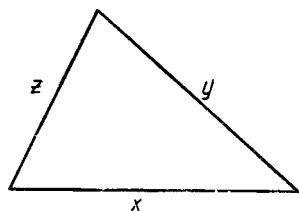
4. Выбор точки из куба. Рассмотрим случайный выбор точки из куба со стороной a . Будем считать, что одна вершина куба лежит в начале координат, а оси координат направлены по ребрам, выходящим из этой вершины. Вероятность того, что точка будет выбрана из некоторой части куба, равна отношению объема этой части к a^3 — объему куба. Для того чтобы выбрать точку внутри куба, нужно указать ее координаты x, y, z , каждая из координат меняется от 0 до a .



Точно так, как при рассмотрении случайного выбора точки из прямоугольника, можем установить, что при случайном выборе точки из куба каждая координата случайно

выбирается из отрезка $[0, a]$, причем выбор разных координат происходит независимым образом. Например, если A есть событие «координата x лежит в отрезке $[\alpha_1, \alpha_2]$ », B — событие «координата y лежит в отрезке $[\beta_1, \beta_2]$ », а

C — «координата z лежит в отрезке $[\gamma_1, \gamma_2]$ », то $A \cdot B \cdot C$ есть событие, при котором одновременно выполняются такие неравенства для координат выбранной точки: $\alpha_1 \leq x \leq \alpha_2$, $\beta_1 \leq y \leq \beta_2$, $\gamma_1 \leq z \leq \gamma_2$. Это означает, что



точка попала в параллелепипед, который вырезают из куба полосы, параллельные координатным граням ширины $\alpha_2 - \alpha_1$, $\beta_2 - \beta_1$, $\gamma_2 - \gamma_1$, т. е. это будут размеры сторон параллелепипеда. Поэтому

$$P(A \cdot B \cdot C) = \frac{(\alpha_2 - \alpha_1)(\beta_2 - \beta_1)(\gamma_2 - \gamma_1)}{a^3} = \\ = \frac{\alpha_2 - \alpha_1}{a} \cdot \frac{\beta_2 - \beta_1}{a} \cdot \frac{\gamma_2 - \gamma_1}{a}.$$

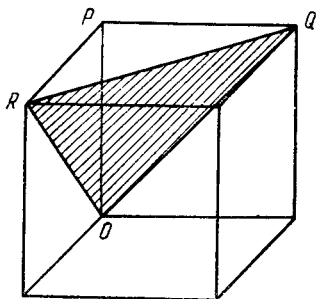
Но $\frac{\alpha_2 - \alpha_1}{a}$ есть вероятность выбрать случайно из отрезка $[0, a]$ точку так, чтобы она попала в $[\alpha_1, \alpha_2]$, аналогично $\frac{\beta_2 - \beta_1}{a}$ и $\frac{\gamma_2 - \gamma_1}{a}$ есть вероятности того, что случайно выбранные точки y и z попадут соответственно в отрезки $[\beta_1, \beta_2]$, $[\gamma_1, \gamma_2]$.

Рассмотрим такую задачу. Из отрезка $[0, a]$ выбираются случайно, независимо одна от другой три точки x, y, z . Какова вероятность, что из отрезков длин x, y, z можно составить треугольник? Из геометрии известно, что такой треугольник можно составить, если выполнены неравенства $x + y > z$, $x + z > y$, $z + y > x$. Обозначим через A событие, которое заключается в том, что выполнены все эти три неравенства. Пусть далее B, C, D — соответственно события: $z \geq x + y$, $y \geq x + z$, $x \geq z + y$. Будем считать, что все числа x, y, z положительны. Тогда события A, B, C, D несовместимы ($\bar{A} = B + C + D$, если происходит B и C , то $x = 0, z = y$). Значит, $P(A) = 1 - P(B) - P(C) - P(D)$.

Из соображений симметрии легко видеть, что $P(B) = P(C) = P(D)$. Вычислим

$P(B)$. $z = x + y$ на плоскости, заштрихованной внутри куба (см. рисунок). Поэтому событие B происходит, если точка попадает или на эту плоскость, или выше нее, т. е. попадает в пирамиду $OPQR$ (она перевернутая) с площадью

основания $\frac{a^2}{2}$ и высотой a . Значит, ее объем $\frac{a^3}{6}$. Поэтому



$$P(A) = 1 - 3P(B) = 1 - 3 \frac{a^3}{6} \cdot \frac{1}{a^3} = \frac{1}{2}.$$

§ 6. Схема независимых испытаний

Рассмотрим некоторый эксперимент и событие A , которое с вероятностью p ($0 < p < 1$) может происходить в этом эксперименте. Будем повторять эксперимент независимо n раз. Сколько раз произойдет событие A при n повторениях этого эксперимента? Обозначим через A_k , $k = 1, 2, \dots, n$ событие, заключающееся в том, что событие A произошло в k -м эксперименте.

Сначала разберем такой важный частный случай общей ситуации: какова вероятность того, что событие A произойдет хотя бы раз в серии из n экспериментов. Это означает, что произойдет хотя бы одно из событий A_1, A_2, \dots, A_n , т. е. произойдет событие $A_1 + \dots + A_n$. Для вычисления $P(A_1 + \dots + A_n)$ (так как указанные события могут происходить одновременно, то нельзя пользоваться формулой сложения вероятностей) заметим, что противоположное событие заключается в том, что A ни разу не произошло, а значит, произошли события $\bar{A}_1, \dots, \bar{A}_n$:

$$\overline{A_1 + \dots + A_n} = \bar{A}_1 \cdot \bar{A}_2 \cdot \dots \cdot \bar{A}_n.$$

Поскольку эксперименты производятся независимо, то события $\bar{A}_1, \dots, \bar{A}_n$ независимы, так что

$$P(\bar{A}_1 \cdot \bar{A}_2 \dots \bar{A}_n) = P(\bar{A}_1)P(\bar{A}_2) \dots P(\bar{A}_n).$$

Но вероятности событий \bar{A}_k одинаковы и равны $1 - p$ ($P(\bar{A}) = 1 - P(A)$). Значит,

$$P(\overline{A_1 + \dots + A_n}) = (1 - p)^n,$$

откуда

$$P(A_1 + \dots + A_n) = 1 - (1 - p)^n. \quad (1)$$

Это и есть формула, дающая вероятность того, что в n независимых испытаниях событие, имеющее вероятность p , произойдет по крайней мере 1 раз.

Из этой формулы вытекает такое важное следствие: как бы мала ни была вероятность события A , при достаточно большом числе испытаний оно произойдет (хотя бы раз) с вероятностью, сколько угодно близкой к единице. Действительно, правая часть равенства (1) при $n \rightarrow \infty$ стремится к 1, так как $1 - p < 1$ и, значит, $(1 - p)^n \rightarrow 0$ при $n \rightarrow \infty$. Пусть, например, $p = 0,0001$. Тогда при $n = 100\,000$ вероятность в правой части (1) равна примерно 0,99995.

Благодаря этому обстоятельству в физике были обнаружены (и продолжают обнаруживать) новые элементарные частицы. Большинство из них возникают в ядерных реакциях с очень малыми вероятностями, кроме того, как правило, они существуют очень короткое время, так что вероятность обнаружить их в единичном эксперименте ничтожна (но положительна!). Повторяя эксперименты достаточно много раз, — а это возможно, так как количество атомов, которые участвуют в экспериментах (при этом то, что происходит с отдельным атомом, можно считать независимым от остальных экспериментов) имеет порядок 10^{20} , — мы тем не менее в состоянии регистрировать такие маловероятные события.

Вычислим теперь вероятность того, что при n повторениях эксперимента событие A произойдет ровно m раз ($m \leq n$). Обозначим это событие через C_m . Если произошло событие C_m , то среди событий A_1, \dots, A_n произошло m событий, а среди событий $\bar{A}_1, \dots, \bar{A}_n$ — $n - m$. Это будут события вида

$$A_1 \cdot A_2 \dots A_m \cdot \bar{A}_{m+1} \dots \bar{A}_n, \quad (2)$$

$$A_1 \cdot A_2 \dots A_{m-1} \cdot \bar{A}_m \cdot A_{m+1} \cdot \bar{A}_{m+2} \dots \bar{A}_n$$

и т. д. Вероятность каждого такого произведения в силу независимости событий, происходящих в различных экспе-

риментах, равна произведению вероятностей отдельных событий, среди которых m имеют вероятность p , а $n - m$ — вероятность $1 - p$. Значит, вероятность отдельного произведения равна

$$p^m (1 - p)^{n-m}.$$

Сколько произведений событий вида (2) входит в C_m ? Каждое такое произведение определяется выбором из n элементов (номеров событий) m (тех, при которых происходит событие A ; в экспериментах с остальными номерами происходит \bar{A}). Но число способов выбрать из n элементов m равно числу сочетаний из n по m , т. е. C_n^m .

Заметим, что различные произведения вида (2) являются несовместимыми событиями. Действительно, для таких двух произведений всегда можно указать такое k , что в одно входит A_k , а во второе \bar{A}_k . Значит, вероятность события C_m равна сумме C_n^m вероятностей событий вида (2), а так как все они равны между собой, то она равна C_n^m , умноженному на вероятность одного события. Таким образом,

$$P(C_m) = C_n^m p^m (1 - p)^{n-m} = \frac{n!}{m!(n-m)!} p^m (1 - p)^{n-m}. \quad (3)$$

Вероятности $P(C_m)$ носят название *биномиальных*. При $m = 0$ и $m = n$ считаем $C_n^m = 1$ (или $0! = 1$). Тогда формула (3) будет справедливой и для $m = 0$ и $m = n$. Например,

$$P(C_0) = P(\bar{A}_1 \dots \bar{A}_n) = (1 - p)^n.$$

Поскольку при проведении n экспериментов произойдет хотя бы одно из событий C_m , $m = 0, 1, \dots, n$, то

$$U = C_0 + C_1 + \dots + C_n.$$

(U — достоверное событие). События в правой части последнего равенства несовместимы, так что

$$1 = P(C_0) + P(C_1) + \dots + P(C_n).$$

Подставляя вместо вероятностей их значения с помощью формулы (3), получим

$$(1 - p)^n + n(1 - p)^{n-1} p + \frac{n(n-1)}{2} (1 - p)^{n-2} p^2 + \dots + \frac{n!}{m!(n-m)!} (1 - p)^{n-m} p^m + \dots + p^n = 1.$$

Это есть формула бинома Ньютона (мы ее получили из вероятностных соображений). Поэтому вероятности (3) и носят название биномиальных.

Подсчитаем вероятности появления события в 5 испытаниях, если вероятность появления его в одном равна $\frac{1}{3}$.

Тогда

$$P(C_0) = \left(\frac{2}{3}\right)^5 = \frac{32}{243} \approx 0,132,$$

$$P(C_1) = 5 \cdot \frac{1}{3} \cdot \left(\frac{2}{3}\right)^4 = \frac{80}{243} \approx 0,329,$$

$$P(C_2) = \frac{5 \cdot 4}{1 \cdot 2} \left(\frac{2}{3}\right)^3 \left(\frac{1}{3}\right)^2 = \frac{80}{243} \approx 0,329,$$

$$P(C_3) = \frac{5 \cdot 4}{1 \cdot 2} \left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^3 = \frac{40}{243} \approx 0,165,$$

$$P(C_4) = 5 \cdot \left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^4 = \frac{10}{243} \approx 0,041,$$

$$P(C_5) = \left(\frac{1}{3}\right)^5 = \frac{1}{243} \approx 0,004.$$

Как видим, событие появляется 1 или 2 раза с вероятностью примерно $\frac{2}{3}$, а более чем 3 раза — с вероятностью меньшей, чем 0,05. Заметим, что если происходит событие C_1 , частота появления события A — $\frac{1}{5}$, а если C_2 , то частота A равна $\frac{2}{5}$. $\frac{1}{5} < \frac{1}{3} < \frac{2}{5}$ — это наиболее близкие возможные частоты к значению вероятности, они оказываются наиболее вероятными.

Рассмотрим пример с гетерозиготными по признакам «кареглазость» и «курчавость» родителями, который уже обсуждался в § 4. Мы видели, что по фенотипу дети таких родителей могут быть четырех типов: 1) кареглазые курчавые с вероятностью $\frac{9}{16}$, 2) кареглазые прямоволосые с вероятностью $\frac{3}{16}$, 3) голубоглазые курчавые с вероятностью $\frac{3}{16}$, 4) голубоглазые прямоволосые с вероятностью $\frac{1}{16}$. Предположим, у них родились пять детей (без двойняшек). В этом случае рождения отдельных детей можно рассматривать как независимые испытания.

Составим таблицу вероятностей для событий, заключающихся в том, что среди детей заданное число будет принадлежать к одному из четырех типов (m — число детей, A_k — тип ребенка, $k = 1, 2, 3, 4$).

$A_k \backslash m$	0	1	2
1	$\left(\frac{7}{16}\right)^5 \approx 0,016$	$5 \cdot \frac{9}{16} \cdot \left(\frac{7}{16}\right)^4 \approx 0,103$	$10 \cdot \left(\frac{9}{16}\right)^2 \left(\frac{7}{16}\right)^3 \approx 0,267$
2,3	$\left(\frac{13}{16}\right)^5 = 0,354$	$5 \cdot \frac{3 \cdot 13^4}{16^5} = 0,409$	$10 \cdot \frac{9 \cdot 13^3}{16^5} = 0,188$
4	$\left(\frac{15}{16}\right)^5 = 0,723$	$5 \cdot \frac{15^4}{16^5} = 0,241$	$10 \cdot \frac{15^3}{16^5} = 0,032$
	3	4	5
	$10 \cdot \left(\frac{9}{16}\right)^3 \left(\frac{7}{16}\right)^2 = 0,34$	$5 \left(\frac{9}{16}\right)^4 \cdot \frac{7}{16} = 0,219$	$\left(\frac{9}{16}\right)^5 = 0,019$
	$10 \frac{27 \cdot 13^2}{16^5} = 0,044$	$5 \cdot \frac{81 \cdot 13}{16^5} = 0,005$	$\left(\frac{3}{16}\right)^5 = 0,0002$
	$10 \cdot \frac{15^2}{16^5} = 0,002$	$5 \cdot \frac{15}{16^5} = 0,0001$	$\left(\frac{1}{16}\right)^5 = 0,000001$

Из этой таблицы, например, вытекает, что вероятность появления среди детей хотя бы одного голубоглазого и прямоволосого равна примерно 0,3. Заметим: вероятность того, что хотя бы один ребенок будет голубоглазым, составит

$$1 - \left(\frac{3}{4}\right)^5 = 1 - \frac{243}{1024} \approx 0,77 -$$

более чем $\frac{3}{4}$.

§ 7. Закон редких событий

Формула биномиальных вероятностей при большом числе испытаний n неудобна для вычисления, поскольку в нее входят трудно считаемые факториалы больших чисел. Сейчас мы получим некоторые приближенные выражения для указанных вероятностей в случае большого n ; эти формулы

будут тем более точными, чем больше будет n . В этом параграфе мы рассмотрим специальный случай, когда вероятность появления события в одном испытании очень мала, а число испытаний так велико, что вероятность появления хотя бы одного события в серии из n испытаний не очень близка к нулю и не очень близка к 1. В этом случае событие A может либо ни одного разу не произойти в серии, либо произойдет конечное число раз (1, 2, 3, ... число раз существенно меньше, чем число испытаний). Такого рода события называются *редкими*. К ним относятся, например, рождение тройни у людей, регистрация космических частиц сверхвысоких энергий, возникновение электрон-позитронной пары, различного рода катастрофы.

Чтобы вывести формулу вероятностей редких событий, нам потребуется один математический факт. В теории пределов устанавливается, что при $x \rightarrow 0$ величина

$$(1 + x)^{1/x}$$

стремится к некоторому пределу, равному приблизительно 2,718281828... и обозначаемому буквой e . Число e играет очень важную роль в математике, оно служит основанием натуральных логарифмов и, кроме того, входит во многие математические и физические формулы, в частности в формулу Пуассона для вероятностей редких событий, которую мы получим ниже.

Если вероятность появления события A в одном испытании равна p , то вероятность того, что оно ни разу не появится в n испытаниях, составляет

$$1 - (1 - p)^n = 1 - (1 - p)^{-\frac{1}{p}(-np)}.$$

Если $p \rightarrow 0$, то и $-p \rightarrow 0$ и, значит,

$$(1 - p)^{-\frac{1}{p}} \rightarrow e.$$

Поэтому при достаточно малых p

$$1 - (1 - p)^n \approx 1 - e^{-np} \approx 1 - e^{-a}, \quad \text{где } a = np.$$

Для того чтобы это выражение было отлично (не приближалось) ни к 0, ни к 1, нужно, чтобы a не стремилось ни к 0, ни к ∞ , т. е. было ограниченным. Это означает, что n должно иметь порядок $\frac{1}{p}$ (или, что то же самое, p имеет порядок $\frac{1}{n}$). Предположим, что $np \sim a$, а число $n \rightarrow \infty$.

Тогда для вероятности того, что произошло m событий, можем записать такое представление:

$$\frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} = \\ = \frac{1}{m!} \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right) n^m p^m (1-p)^{n-m}.$$

Число m у нас фиксировано, $n \rightarrow \infty$, $np \sim a$. Поэтому

$$\left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right) \rightarrow 1,$$

так как каждый сомножитель стремится к 1. Далее $(np)^m$ стремится к a^m , $(1-p)^n \rightarrow e^{-a}$ (это было установлено выше). Наконец, так как $p \rightarrow 0$, $(1-p) \rightarrow 1$, то и $(1-p)^{-m} \rightarrow 1$. Окончательно получаем

$$\frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \rightarrow \frac{a^m}{m!} e^{-a}.$$

Правые части этих соотношений

$$p_m(a) = \frac{a^m}{m!} e^{-a} \quad (1)$$

дают вероятность того, что произойдет m редких событий. Формула (1) носит название формулы Пуассона. Ею пользуются при конечных (но больших) n , подставляя np на место a .

Из формулы (1) можем получить следующую формулу математического анализа:

$$e^a = 1 + a + \frac{a^2}{2} + \cdots + \frac{a^m}{m!} + \cdots \quad (2)$$

И действительно, при неограниченном возрастании n может произойти любое из событий C_m , где C_m — событие, заключающееся в том, что произошло ровно m редких событий. Значит,

$$1 = P(C_0) + P(C_1) + \cdots + P(C_m) + \cdots$$

(справа стоит бесконечная сумма, т. е. сумма ряда). Умножая это соотношение на e^a и заменяя $P(C_m)$ на $p_m(a)$, получим (2).

В истории физики известен случай, когда именно наблюдение редких событий привело к великому открытию. Это было открытие Э. Резерфордом атомного ядра.

Мы несколько схематично опишем опыт Резерфорда. Соответствующие вычисления будут носить поясняющий

характер, поэтому мы будем оперировать весьма приближенными значениями.

В лаборатории Резерфорда наблюдали прохождение потока α -частиц (ядер гелия) через тонкую фольгу. При этом довольно редко отмечалось такое событие, когда α -частицы как бы отскакивали от фольги и летели в обратном направлении. Сотрудники Резерфорда, проводившие опыт, считали, что такого быть не может, ибо противоречит имеющемуся представлению о строении вещества. Поэтому они относили «обратно движущиеся частицы» к каким-то невыясненным помехам. Резерфорд же пришел к выводу, что внутри атомов имеются тяжелые ядра. Вследствие столкновений с ними и получают «обратно движущиеся частицы», а из-за малости размеров ядер (по сравнению с атомами) такие столкновения происходят достаточно редко.

Рассмотрим некоторые числовые характеристики эксперимента. Если в качестве источника α -частиц взять 1 г радия, то каждую секунду он будет излучать около $4 \cdot 10^{10}$ частиц. Из всех этих частиц, равномерно вылетающих по всем направлениям, нужно вырезать достаточно узкий пучок, который затем будет направлен на испытываемую фольгу. Число частиц в таком пучке будет иметь порядок 10^4 в секунду, т. е. 10^8 в сутки. Теперь уже известно, что при диаметре атома порядка 10^{-8} см ядро имеет размеры порядка 10^{-12} см. Поэтому отношение площадей поперечного сечения ядра и атома, равное отношению квадратов их диаметров, имеет порядок 10^{-8} (нужно рассматривать сечение атома, перпендикулярное к направлению движения частиц).

Так как прохождение каждой частицы через фольгу можно рассматривать как независимый эксперимент, вероятность того, что частица при одном эксперименте столкнется с ядром, есть 10^{-8} . В сутки имеем 10^8 таких экспериментов. Вероятность того, что в течение суток будет зарегистрировано m обратных частиц, можно определять с помощью формулы Пуассона, взяв $a = 10^{-8} \cdot 10^8 = 1$. Значит, для вероятностей того, что появится m частиц, можем составить такую таблицу:

m	0	1	2	3	4	5	6
$P_m(1)$	0,37	0,37	0,184	0,061	0,015	0,003	< 0,001

Любопытно сопоставить ее с другой, соответствующей наблюдению в течение двух суток. В этом случае $a = 2$.

m	0	1	2	3	4	5	6	7	8
$p_m(2)$	0,134	0,265	0,265	0,176	0,088	0,035	0,012	0,003	$< 0,001$

За двое суток интересующее нас редкое событие будет наблюдаться с вероятностью, уже довольно близкой к 1 (0,866), а за трое, как показывают вычисления, — с вероятностью, большей 0,95.

§ 8. Закон больших чисел

Пусть производится n экспериментов, в каждом из которых может произойти событие A с вероятностью p . Если событие A наблюдалось в m экспериментах из n , то частота появления события A равна $\frac{m}{n}$. Как связаны частота и вероятность?

Ранее, при введении вероятности, отмечался такой экспериментально наблюдаемый факт, как колебание частоты около некоторого фиксированного значения, которое и принималось за значение вероятности. Имеет ли этот факт объяснение в теории вероятностей? Оказывается, да. В этом параграфе будет разобран закон больших чисел, являющийся точной математической формулировкой упомянутого эмпирического факта.

В описанной выше ситуации, обозначая частоту события A через ν_n , в результате проведения n экспериментов мы будем наблюдать одно из событий B_m : частота $\nu_n = \frac{m}{n}$, $m = 0, 1, \dots, n$. События B_m несовместимы и по формуле биномиальных вероятностей

$$P(B_m) = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}.$$

Пусть $\alpha > 0$. Оценим вероятность события, заключающегося в том, что частота ν_n больше, чем $p + \alpha$. Такое событие происходит, если происходит одно из событий B_m с номерами m , для которых выполнено неравенство

$$\frac{m}{n} > p + \alpha, \quad m > n(p + \alpha).$$

Поэтому искомая вероятность представима в виде

$$P(v_n > p + \alpha) = P(B_{m^*}) + P(B_{m^*+1}) + \dots + P(B_n), \quad (1)$$
 где m^* — наименьшее целое число, для которого выполнено неравенство $m^* > n(p + \alpha)$.

Для оценки суммы вероятностей в правой части (1) рассмотрим отношение

$$\begin{aligned} \frac{P(B_{m+1})}{P(B_m)} &= \frac{n!}{(m+1)!(n-m-1)!} p^{m+1} (1-p)^{n-m-1} \\ \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} &= \frac{m!(n-m)!}{(m+1)!(n-m-1)!} \frac{p}{1-p} = \\ &= \frac{n-m}{m+1} \cdot \frac{p}{1-p}. \end{aligned}$$

Оно имеет смысл при $0 \leq m \leq n-1$ ($m+1 \leq n$) и, как видно, убывает с ростом m (числитель дроби уменьшается, а знаменатель увеличивается).

Найдем то m , начиная с которого отношение меньше 1, т. е.

$$\begin{aligned} \frac{n-m}{m+1} \cdot \frac{p}{1-p} < 1, \quad np - mp < (m+1)(1-p), \\ np < mp + m(1-p) + 1 - p = m + 1 - p, \\ m > np + p - 1. \end{aligned}$$

Если m_1 наименьшее целое, при котором выполнено неравенство

$$m_1 > np + p - 1, \text{ то тогда } m_1 - 1 \leq np + p - 1, \text{ т. е.}$$

$$\begin{aligned} m_1 &\leq np + p \text{ и} \\ m^* - m_1 &\geq n(p + \alpha) - np - p = n\alpha - p. \end{aligned} \quad (2)$$

Так как при $m \geq m^*$

$$\frac{P(B_{m+1})}{P(B_m)} \leq \frac{P(B_{m^*+1})}{P(B_{m^*})} = \frac{n-m^*}{m^*+1} \frac{p}{1-p},$$

то, полагая $\lambda_n = \frac{n-m^*}{m^*+1} \cdot \frac{p}{1-p}$, будем иметь

$$P(B_{m+1}) \leq \lambda_n P(B_m).$$

Значит,

$$\begin{aligned} P(B_{m^*+k}) &\leq \lambda_n P(B_{m^*+k-1}) \leq \lambda_n^2 P(B_{m^*+k-2}) \leq \dots \leq \\ &\leq \lambda_n^k P(B_{m^*}). \end{aligned}$$

Поэтому

$$\begin{aligned} &P(B_{m^*}) + P(B_{m^*+1}) + \dots + P(B_n) \leq \\ &\leq P(B_{m^*}) + \lambda_n P(B_{m^*}) + \lambda_n^2 P(B_{m^*}) + \dots + \lambda_n^{n-m^*} P(B_{m^*}) = \\ &= P(B_{m^*}) [1 + \lambda_n + \lambda_n^2 + \dots + \lambda_n^{n-m^*}]. \end{aligned}$$

Справа в скобках стоит геометрическая прогрессия с первым членом 1 и знаменателем $\lambda_n < 1$ (ибо $m^* > m_1$).
Имеем

$$1 + \lambda_n + \dots + \lambda_n^{n-m^*} = \frac{1 - \lambda_n^{n-m^*+1}}{1 - \lambda_n} \leq \frac{1}{1 - \lambda_n},$$

и, значит, сумма в правой части (1) оценивается выражением

$$\frac{1}{1 - \lambda_n} P(B_{m^*}). \quad (3)$$

Используя значение λ_n , можем записать:

$$\begin{aligned} 1 - \lambda_n &= 1 - \frac{n - m^*}{m^* + 1} \cdot \frac{p}{1 - p} = \\ &= \frac{m^*(1 - p) + (1 - p) - np + m^*p}{(m^* + 1)(1 - p)} = \frac{m^* + (1 - p) - np}{(m^* + 1)(1 - p)} \gg \\ &\gg \frac{n(p + \alpha) + (1 - p) - np}{n(1 - p)} \gg \frac{n\alpha}{n(1 - p)} = \frac{\alpha}{1 - p} \end{aligned}$$

(мы воспользовались неравенствами $m^* > n(p + \alpha)$, $m^* + 1 \leq n$). Значит,

$$\frac{1}{1 - \lambda_n} \leq \frac{1 - p}{\alpha}.$$

Далее $P(B_{m_1}) \geq P(B_{m_1+1}) \geq \dots \geq P(B_{m^*})$. Поэтому

$$\begin{aligned} 1 &\geq P(B_{m_1}) + P(B_{m_1+1}) + \dots + P(B_{m^*}) \geq \\ &\geq P(B_{m^*}) + P(B_{m^*}) + \dots + P(B_{m^*}) = \\ &= [(m^* - m_1) + 1] P(B_{m^*}) \geq \\ &\geq (n\alpha + 1 - p) P(B_{m^*}) \geq n\alpha P(B_{m^*}). \end{aligned}$$

Таким образом, $n\alpha P(B_{m^*}) \leq 1$, $P(B_{m^*}) \leq \frac{1}{n\alpha}$.

Мы получили оценки сверху для обоих сомножителей в (3). Перемножая их, получим число $\frac{1-p}{n\alpha^2}$, которое оценивает сумму вероятностей справа в (1). Итак, имеет место неравенство:

$$P\{v_n > p + \alpha\} \leq \frac{1-p}{n\alpha^2}. \quad (4)$$

Найдем теперь оценку для вероятности события $v_n < p - \alpha$. Оказывается, ее можно получить, используя неравенство (4). Для этого рассмотрим вместо события A

событие \bar{A} . Пусть \bar{p} — вероятность этого события, \bar{v}_n — его частота в n экспериментах.

По доказанному справедливо неравенство

$$P\{\bar{v}_n > \bar{p} + \alpha\} \leq \frac{1 - \bar{p}}{n\alpha^2}.$$

Но $\bar{v}_n + v_n = 1$ и $p + \bar{p} = 1$. Поэтому событие $\{\bar{v}_n > \bar{p} + \alpha\}$ совпадает с событием $\{1 - v_n > 1 - p + \alpha\}$ или $\{v_n < p - \alpha\}$.

Таким образом,

$$P\{v_n < p - \alpha\} \leq \frac{p}{n\alpha^2}.$$

Событие $|v_n - p| > \alpha$ влечет одно из событий $v_n - p > \alpha$ (т. е. $v_n > p + \alpha$) или $v_n - p < -\alpha$ (т. е. $v_n < p - \alpha$).

Эти события несовместимы, так что

$$\begin{aligned} P\{|v_n - p| > \alpha\} &= P\{v_n < p - \alpha\} + P\{v_n > p + \alpha\} \leq \\ &\leq \frac{1 - p}{n\alpha^2} + \frac{p}{n\alpha^2} = \frac{1}{n\alpha^2}. \end{aligned}$$

Или окончательно

$$P\{|v_n - p| > \alpha\} \leq \frac{1}{n\alpha^2}. \quad (5)$$

При $n \rightarrow \infty$ правая часть неравенства (5) стремится к нулю, каково бы ни было $\alpha > 0$. Отсюда и вытекает закон больших чисел, утверждающий, что вероятность того, что частота появления события в n независимых испытаниях будет отличаться от вероятности события (при одном испытании) больше, чем на заданное положительное число α , стремится к нулю, каково бы ни было $\alpha > 0$.

Если подбрасывать симметричную монету (с вероятностью выпадения герба $1/2$) n раз, то, обозначая частоту выпадения герба через v_n , можем, используя неравенство (5), оценить вероятности отклонения частоты от значения $1/2$ при разных n и α таким образом:

$\alpha \backslash n$	1000	5000	10 000	100 000	1 000 000 000
0,1	0,1	0,02	0,01	0,001	0,0000001
0,05	0,4	0,08	0,04	0,004	0,0000004
0,01	—	—	—	0,1	0,0001
0,005	—	—	—	—	0,01

В таблице против соответствующих значений n и α записано выражение $\frac{1}{n\alpha^2}$, оценивающее вероятность события $\left| \nu_n - \frac{1}{2} \right| > \alpha$. Чем точнее мы хотим с помощью частоты оценить вероятность события, тем больше наблюдений нужно сделать. Заметим, что оценка (5) весьма грубая, она нужна лишь для обоснования закона больших чисел.

§ 9. Нормальная аппроксимация

Теперь рассмотрим более точно вероятности для отклонения частоты от вероятности, которые уже рассматривались в предыдущем параграфе. Пусть $\eta_n = \nu_n - p$, где ν_n — частота события A в n экспериментах, p — вероятность события A в одном испытании. Разность η_n будем называть флуктуацией частоты.

Следующий пример поясняет это понятие. Представим себе газ, наполняющий сосуд V (например, комнату). Если выделить в этом сосуде некоторый объем V_1 , то вероятность того, что заданная молекула попадет в V_1 , равна отношению объемов V_1 и сосуда. Каждая молекула попадает в V_1 независимо от других (мы пользуемся статистикой Максвелла — Больцмана).

Пусть в сосуде N молекул и N_1 их в объеме V_1 . Давление в объеме V_1 пропорционально N_1 и обратно пропорционально величине рассматриваемого объема. Если P_1 — давление в V_1 , а P — среднее давление во всем сосуде, то при некотором c

$$P_1 = \frac{cN_1}{\text{объем } V_1}, \quad P = \frac{cN}{\text{объем } V}.$$

Значит,

$$\begin{aligned} P_1 - P &= \frac{cN}{\text{объем } V_1} \left(\frac{N_1}{N} - \frac{\text{объем } V_1}{\text{объем } V} \right) = \\ &= P \frac{\text{объем } V}{\text{объем } V_1} (\nu_N - p), \end{aligned}$$

где ν_N — частота попадания молекул в V_1 в N испытаниях, p — вероятность попадания отдельной молекулы в V_1 . Разность между давлением газа в данном небольшом объеме и средним давлением и есть флуктуация давления. Полученная формула для флуктуации давления показы-

вает, что оно просто выражается через флуктуацию частоты такого события: молекула попала в выделенный объем V_1 .

Уже формула (5) предыдущего параграфа показывает, что при больших n естественно рассматривать лишь величины отклонения (т. е. флуктуации) частоты от вероятности порядка $\frac{1}{\sqrt{n}}$. Действительно, на основании указанной формулы

$$P \left\{ |\eta_n| > \frac{x}{\sqrt{n}} \right\} \leq \frac{1}{\frac{x^2}{n} \cdot n} = \frac{1}{x^2}$$

и при достаточно большом x выражение справа становится сколь угодно малым. Мы найдем точное асимптотическое выражение для вероятности слева для всякого $x > 0$ при $n \rightarrow \infty$. Для этого надо найти предельное значение вероятности

$$P \left\{ \eta_n = \frac{z}{\sqrt{n}} \right\}$$

при $n \rightarrow \infty$. Оказывается, при $n \rightarrow \infty$ для любого z эта вероятность стремится к нулю, но порядок стремления будет $\frac{1}{\sqrt{n}}$, так что после умножения на \sqrt{n} эта вероятность уже будет иметь конечный предел:

$$\lim_{n \rightarrow \infty} \sqrt{np(1-p)} P \left\{ \eta_n = \frac{z \sqrt{p(1-p)}}{\sqrt{n}} \right\} = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}. \quad (1)$$

(Вообще z слева зависит от n , меняясь так, чтобы η_n могло принимать соответствующее значение, т. е. чтобы при некотором целом m $\frac{m}{n} = p + \frac{z \sqrt{p(1-p)}}{\sqrt{n}}$; предполагается, что z стремится к некоторому предельному значению, которое и входит в правую часть (1).)

Формула (1) носит название формулы Муавра — Лапласа. Ее применения весьма распространены. Вывод этой формулы довольно громоздкий. Мы для пояснения рассмотрим простейший случай, когда $p = \frac{1}{2}$ (именно этот случай исследован Муавром, общий случай рассмотрен Лапласом). Нам потребуется одна асимптотическая формула для факториалов больших чисел, так называемая формула Стирлинга:

$$n! \sim \sqrt{2\pi n} n^n e^{-n} \quad (2)$$

(знак \sim означает, что отношение чисел, соединенных этим знаком, стремится к 1, π — отношения длины окружности к диаметру, e — основание натуральных логарифмов, оно уже рассматривалось в § 7). Для $\rho = 1/2$ формула (1) принимает вид

$$\lim_{n \rightarrow \infty} \frac{\sqrt{n}}{2} P \left\{ \eta_n = \frac{z}{2\sqrt{n}} \right\} = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}. \quad (3)$$

Пусть $m = \frac{1}{2}n + \frac{z\sqrt{n}}{2}$ (z таково, что это число целое). Тогда

$$\begin{aligned} P \left\{ \eta_n = \frac{z}{2\sqrt{n}} \right\} &= C_n^m \left(\frac{1}{2} \right)^m \left(1 - \frac{1}{2} \right)^{n-m} = \\ &= C_n^m \frac{1}{2^n} = \frac{n!}{m! (n-m)! 2^n}. \end{aligned}$$

Воспользуемся для факториалов формулой Стирлинга: используя (2), можем записать

$$\begin{aligned} m! &\sim \sqrt{2\pi m} m^m e^{-m}, \quad (n-m)! \sim \\ &\sim \sqrt{2\pi (n-m)} (n-m)^{n-m} e^{-(n-m)}. \end{aligned}$$

Значит,

$$\begin{aligned} &\frac{\sqrt{n}}{2} P \left\{ \eta_n = \frac{z}{2\sqrt{n}} \right\} \sim \\ &\sim \frac{\sqrt{n} \sqrt{2\pi n} n^n e^{-n}}{2 \sqrt{2\pi m} m^m e^{-m} \sqrt{2\pi (n-m)} (n-m)^{n-m} e^{-(n-m)}} = \\ &= \frac{1}{2 \sqrt{2\pi \cdot \frac{m}{n} \cdot \frac{n-m}{n}}} \cdot \frac{n^n}{(2m)^m (2n-2m)^{n-m}}. \end{aligned}$$

Исследуем каждый из двух сомножителей в последнем произведении.

Так как

$$m = \frac{1}{2}n + \frac{z\sqrt{n}}{2}, \quad \frac{m}{n} = \frac{1}{2} + \frac{z}{2\sqrt{n}},$$

то $\frac{m}{n} \rightarrow \frac{1}{2}$ при $n \rightarrow \infty$. Поэтому и $\frac{n-m}{n} = 1 - \frac{m}{n} \rightarrow \frac{1}{2}$ при $n \rightarrow \infty$. Так что

$$2 \sqrt{2\pi \cdot \frac{m}{n} \cdot \frac{n-m}{n}} \rightarrow 2 \sqrt{2\pi \cdot \frac{1}{2} \cdot \frac{1}{2}} = \sqrt{2\pi}. \quad (4)$$

Преобразуем выражение

$$\begin{aligned} \frac{(2m)^m (2n-2m)^{n-m}}{n^n} &= \left(\frac{2m}{n}\right)^m \left(\frac{2n-2m}{n}\right)^{n-m} = \\ &= \left(1 + \frac{z}{\sqrt{n}}\right)^m \left(2 - 1 - \frac{z}{\sqrt{n}}\right)^{n-m} = \\ &= \left(1 + \frac{z}{\sqrt{n}}\right)^m \left(1 - \frac{z}{\sqrt{n}}\right)^{n-m} = \\ &= \left(1 - \frac{z^2}{n}\right)^m \left(1 - \frac{z}{\sqrt{n}}\right)^{n-2m}. \end{aligned}$$

Пределы сомножителей справа можно найти, используя предел для числа e , приведенный в § 7:

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(1 - \frac{z^2}{n}\right)^m &= \lim_{n \rightarrow \infty} \left(1 - \frac{z^2}{n}\right)^{-\frac{n}{z^2} \left(-\frac{mz^2}{n}\right)} = \\ &= e^{-\lim_{n \rightarrow \infty} \frac{mz^2}{n}} = e^{-\frac{1}{2}z^2}, \end{aligned}$$

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(1 - \frac{z}{\sqrt{n}}\right)^{n-2m} &= \lim_{n \rightarrow \infty} \left(1 - \frac{z}{\sqrt{n}}\right)^{-z\sqrt{n}} = \\ &= \left[\lim_{n \rightarrow \infty} \left(1 - \frac{z}{\sqrt{n}}\right)^{-\frac{\sqrt{n}}{z}} \right]^{z^2} = e^{z^2}. \end{aligned}$$

Значит,

$$\lim_{n \rightarrow \infty} \frac{(2m)^m (2n-2m)^{n-m}}{n^n} = e^{-z^2/2} \cdot e^{z^2} = e^{z^2/2}$$

$$\text{и } \frac{n^n}{(2m)^m (2n-2m)^{n-m}} \sim e^{-z^2/2}.$$

Учитывая (4), получаем равенство (3).

Из формулы (1) вытекает следующее важное следствие. Рассмотрим вероятность того, что флуктуация частоты η_n лежит в заданных пределах, например вероятность того, что $a \leq \sqrt{n}\eta_n \leq b$.

Пусть m_1 наименьшее целое число, которое удовлетворяет неравенству:

$$a \leq \frac{m - np}{\sqrt{n}},$$

а m_2 — наибольшее из чисел, для которого $\frac{m - np}{\sqrt{n}} \leq b$.

Если m меняется от m_1 до m_2 , то флуктуация частоты, умноженная на \sqrt{n} , меняется от a до b . Следовательно,

$$P\left\{a \leq \sqrt{n} \eta_n < b\right\} = P\left\{v_n = \frac{m_1}{n}\right\} + P\left\{v_n = \frac{m_1+1}{n}\right\} + \dots + P\left\{v_n = \frac{m_2}{n}\right\}. \quad (5)$$

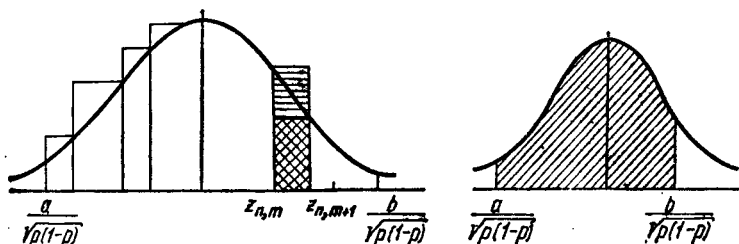
Последнюю сумму можно преобразовать на основании формулы (1), используя то, что

$$P\left\{v_n = \frac{m}{n}\right\} \sim \frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_{n,m}^2},$$

где $z_{n,m} = \frac{m-np}{\sqrt{np(1-p)}}$. Произведение, стоящее в правой части последнего равенства, можно представить так:

$$(z_{n,m+1} - z_{n,m}) \frac{1}{\sqrt{2\pi}} e^{-z_{n,m}^2}.$$

Это площадь прямоугольника, заштрихованного на рисунке.



Сумма справа в (5), таким образом, — сумма прямоугольников, высоты которых есть значение функции $\frac{1}{\sqrt{2\pi}} e^{-z^2/2} = \varphi(z)$ в точках $z_{n,m}$, а основание — отрезок $[z_{n,m}, z_{n,m+1}]$. Если перейти к пределу при $n \rightarrow \infty$, то сумма площадей прямоугольников будет стремиться к площади фигуры, заштрихованной на правом рисунке. Сверху она ограничена графиком функции $\varphi(z)$, снизу — осью z , а с боков — пределами $z_1 = \frac{a}{\sqrt{p(1-p)}}$ и $z_2 = \frac{b}{\sqrt{p(1-p)}}$, к этим значениям стремятся z_{n,m_1} и z_{n,m_2} соответственно.

Эта площадь равна интегралу $\int_{a/\sqrt{\rho(1-\rho)}}^{b/\sqrt{\rho(1-\rho)}} \varphi(z) dz$, что и есть предельное значение вероятности события $a \leq \sqrt{n} \eta_n \leq b$.

Как видно из формулы, более простое выражение получится для вероятности события $a \leq \sqrt{\frac{n}{\rho(1-\rho)}} \eta_n \leq b$. В этом случае пределы интегрирования будут просты a и b .

Таким образом, мы получаем следующее утверждение которое носит название интегральной предельной теоремы Лапласа:

$$\lim_{n \rightarrow \infty} P \left\{ a \leq \sqrt{\frac{n}{\rho(1-\rho)}} \eta_n \leq b \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz. \quad (6)$$

В частности, отсюда вытекает, что

$$\lim_{n \rightarrow \infty} P \left\{ \left| \sqrt{\frac{n}{\rho(1-\rho)}} \eta_n \right| > \alpha \right\} = 1 - \frac{2}{\sqrt{2\pi}} \int_0^\alpha e^{-z^2/2} dz.$$

Для функции, стоящей справа, составлены таблицы, помощью которых можно оценивать точно вероятности отклонения частоты от вероятности. Приведем некоторые характерные значения этой функции:

α	0,5	1,0	1,5	2	2,5	3	3,4	3,9
	0,6171	0,3173	0,1336	0,0455	0,0124	0,0027	0,0007	0,0001

Из этой таблицы видно, что значение флуктуации, превосходящее по абсолютной величине $4 \sqrt{\frac{\rho(1-\rho)}{n}}$, встречается с вероятностью, меньшей 0,0001.

Оценим порядок флуктуации давления в нормальных комнатных условиях. Пусть $\rho = \frac{\text{объем } V_1}{\text{объем } V_2}$ (это малая величина). Тогда, полагая $\nu_N - \rho \sim \sqrt{\frac{\rho(1-\rho)}{N}}$, будем иметь

$$P_1 - P \sim P \cdot \frac{1}{\rho} \sqrt{\frac{\rho(1-\rho)}{N}} = P \frac{1}{\sqrt{N\rho}}.$$

Если $N = 10^{24}$ (такого порядка число молекул в комнате, а $p = 10^{-4}$), флуктуации имеют порядок 10^{-10} . Чем меньше объем, тем значительнее флуктуации: Например, если взять объем порядка кубического миллиметра (при объеме комнаты 150 м^3), то $p = 10^{-11}$, $Np = 10^{13}$, флуктуации имеют порядок 10^{-6} .

§ 10. Оценка неизвестной вероятности события

Во многих случаях нам предварительно может быть и неизвестна вероятность события A , но мы можем, многократно производя эксперимент, в котором событие A происходит, находить его частоту.

Какие достоверные выводы можно сделать о вероятности события A , зная его частоту?

Заметим, что часто нас не интересует вероятность сама по себе, а нужно определить некоторую физическую величину, через которую выражается вероятность. Приведем два простых примера подобных ситуаций.

Пусть имеется большая партия однотипных изделий. N — число изделий, среди них — n негодных. Нас интересует доля брака $\frac{n}{N}$. Предположим, что анализ качества изделия приводит к его разрушению (или просто негодности). Поэтому проверять все изделия не имеет смысла. Производится выборочный контроль случайным образом отобранных M изделий (их число считается существенно меньшим N). Так как M/N — малое число, можно считать, что вероятность извлечь негодное изделие при каждом из M извлечений изделий для выборочного контроля неизменна. Поэтому можно считать, что производится M независимых экспериментов, в каждом из которых с вероятностью $\frac{n}{N}$ появляется событие «выбранное изделие оказалось негодным». Если среди этих M изделий оказалось m негодных, то по частоте $\frac{m}{M}$ нужно оценить вероятность $\frac{n}{N}$, т. е. долю негодных изделий в партии.

Второй пример связан с упоминавшимся в § 7 опытом Резерфорда, в котором атомы бомбардировались α -частицами. Зная вероятность столкновения частицы с ядром атома, можно судить о размерах ядра. Экспериментаторам известны исходные данные, относящиеся к числу α -частиц

в пучке в единицу времени, т. е. известно число экспериментов (здесь в качестве эксперимента выступает пролет через облучаемую фольгу каждой отдельной α -частицы). Наблюдая «отскоки» α -частиц, можно узнать частоту события « α -частица столкнулась с ядром». Для определения размера атомного ядра нужно найти вероятность этого события, так как эта вероятность равна примерно $\frac{r^2}{R^2}$, где R — радиус атома, r — радиус ядра.

Итак, предположим, что событие A имеет неизвестную вероятность p , произведено n экспериментов (n считаем достаточно большим) и частота появления события A в этих экспериментах равна v_n . Тогда из предельной теоремы Лапласа (см. § 9, (6)) вытекает, что вероятность события

$$\sqrt{\frac{n}{p(1-p)}} |v_n - p| \leq \alpha \quad (1)$$

при достаточно больших n не зависит от p . При этом выбором α эту вероятность можно сделать сколько угодно близкой к 1. Например, как вытекает из таблицы, приведенной в § 9, при $\alpha = 2$ вероятность этого события больше 0,95, при $\alpha = 3$ больше 0,997, при $\alpha = 4$ больше 0,9999. Пусть, например, $\alpha = 3$. Тогда, какова бы ни была неизвестная вероятность p , с вероятностью не меньшей, чем 0,997, выполнено неравенство

$$\sqrt{\frac{n}{p(1-p)}} |v_n - p| \leq 3.$$

Из этого неравенства можно получить неравенство для p . Возводя в квадрат и избавляясь от знаменателя, находим

$$\begin{aligned} n(v_n^2 - 2pv_n + p^2) &\leq 3p - 3p^2, \\ (n+3)p^2 - (2nv_n + 3)p + nv_n^2 &\leq 0. \end{aligned}$$

Это квадратическое неравенство относительно p . Для того чтобы квадратный трехчлен был отрицательным (при условии, что коэффициент при p^2 положителен), нужно, чтобы p менялось между корнями этого квадратного трехчлена. Они определяются по формуле

$$\begin{aligned} p_{1,2} &= \frac{(2nv_n + 3) \pm \sqrt{4n^2v_n^2 + 12nv_n + 9 - 4n^2v_n^2 - 12nv_n^2}}{2(n+3)} = \\ &= v_n + \frac{3 - 6v_n}{2n+6} \pm \frac{1}{2(n+3)} \sqrt{12nv_n + 9 - 12nv_n^2}. \end{aligned}$$

Таким образом, с вероятностью, не меньшей 0,997, выполнено неравенство

$$v_n + \frac{3 - 6v_n}{2n + 6} - \frac{1}{2n + 6} \sqrt{12nv_n - 12nv_n^2 + 9} < p < \\ < v_n + \frac{3 - 6v_n}{2n + 6} + \frac{1}{2n + 6} \sqrt{12nv_n - 12nv_n^2 + 9}. \quad (2)$$

Пусть, например, произведено 10^4 испытаний. В них событие A произошло 325 раз. Значит, $v_n = 0,0325$. Вычисляя значения в правой и левой частях неравенства (2), находим для p такие границы:

$$0,0325 + 0,00015 - 0,00307 < p < 0,0325 + \\ + 0,00015 + 0,00307, \\ 0,02958 < p < 0,03572.$$

Указанные интервалы носят название *доверительных*. Вероятность, с которой выполнено неравенство (значения на концах неравенства случайны, они зависят от результатов эксперимента), называется *уровнем доверия*. Формула (2) дает доверительный интервал для p , соответствующий уровню доверия 0,997. Беря в (1) $\alpha = 4$ и производя аналогичные вычисления, можем получить доверительный интервал, соответствующий уровню доверия 0,9999.

Близкой к задаче об оценке неизвестной вероятности есть задача о проверке гипотезы о вероятности события. Иногда можно из косвенных соображений гипотетически определить вероятность некоторого события.

Предположим, что скрещиваются особи, гетерозиготные по двум признакам, т. е. имеющим генотип (Aa, Bb). Если соответствующие гены расположены в различных парах хромосом, то передача признаков происходит независимо и вероятность такого генотипа (aa, bb) в потомстве равна $\frac{1}{16}$. Мы хотим проверить гипотезу, заключающуюся в том, что гены, определяющие указанные признаки, расположены в различных парах хромосом. Для этого нам нужно проверить гипотезу, что вероятность генотипа (aa, bb) в потомстве равна $\frac{1}{16}$.

Пусть проверяется гипотеза, что вероятность события A равна p . Для этого произведено n испытаний и вычислена частота появления A , равная v_n .

Далее рассуждаем так. Если гипотеза верна, то вероятность того, что выполнено неравенство

$$\sqrt{\frac{n}{p(1-p)}} |v_n - p| > \alpha \quad (3)$$

при достаточно больших n определяется с помощью таблицы из § 9.

Если некоторое событие имеет достаточно малую вероятность, то оно практически не происходит при единичном испытании. Например, можем считать, что уже события с вероятностью 0,05 при одном испытании не происходят. Тогда, учитывая, что соответствующее значение $\alpha = 2$, мы вычисляем выражение

$$\sqrt{\frac{n}{p(1-p)}} |v_n - p|.$$

Если оно окажется больше 2, то гипотеза отвергается, если же меньше 2, то принимается.

Выбор α в неравенстве (3) связан с тем, какую вероятность отвергнуть правильную гипотезу мы допускаем. Беря $\alpha = 2$, мы можем, руководствуясь указанным правилом, отвергнуть правильную гипотезу с вероятностью 0,05. Если возьмем $\alpha = 3$, то эта вероятность будет 0,003, если $\alpha = 4$, то 0,0001.

Пусть, например, $p = \frac{1}{16}$, $n = 10\,000$ (большое количество независимых экспериментов получаем у растений). Тогда

$$\sqrt{\frac{n}{p(1-p)}} = \sqrt{\frac{256 \cdot 10\,000}{15}} = \frac{1600}{\sqrt{15}}.$$

При $\alpha = 3$ мы должны отвергнуть гипотезу, если $\left|v_n - \frac{1}{16}\right| > 0,073$, а в противном случае принять.

§ 11. Условные вероятности

В § 4 уже обсуждался вопрос о зависимости событий. С вероятностной точки зрения зависимость событий A и B заключается в том, что осуществление события B влияет на вероятность события A . Так, если эксперимент заключается в последовательном извлечении двух шаров из урны (без возвращения), событие B — первый шар черный, событие

A — второй шар черный, а урна содержит n белых и m черных шаров, то для вычисления вероятности события A , если уже известно, что B произошло, нужно найти вероятность того, что из урны, содержащей n белых и $m - 1$ черных шар, будет извлечен черный шар. Эта вероятность равна $\frac{m-1}{n+m-1}$.

Давайте подсчитаем вероятность того, что при извлечении двух шаров второй шар окажется черным. Выбрать два шара последовательно из $n + m$ шаров можно $(n + m)(n + m - 1)$ способами. Оба черных шара можно выбрать $m(m - 1)$ способами, а белый и черный — nm способами, так что вероятность события A будет

$$\frac{m(n + m - 1)}{(n + m)(n + m - 1)} = \frac{m}{n + m}.$$

Эта вероятность отличается от той, которую мы получили в предположении, что произошло событие B .

В разобранным примере было ясно, как вычислять вероятность события A , если известно, что событие B произошло. Теперь рассмотрим общую ситуацию. Для того чтобы стало понятно, как определить условную вероятность для произвольных событий A и B , вспомним, как понятие вероятности связано с частотой: вероятность — это предельное значение частоты, когда число экспериментов неограниченно возрастает.

Пусть эксперимент был произведен n раз, $k_n(B)$ — число раз, когда наблюдалось событие B . Поскольку нас интересуют только те эксперименты, когда B произошло, то лишь результаты этих $k_n(B)$ экспериментов мы и будем анализировать. Другими словами, вместо исходного эксперимента мы имеем дело с новым (условным) экспериментом: проводим исходный до тех пор, пока не произойдет B , и отмечаем результат этого эксперимента; те эксперименты, где B не происходит, просто игнорируем.

(Так, в приведенном выше примере будем извлекать по два шара, возвращая их после извлечения, до тех пор, пока первый шар не окажется черным; результат этого эксперимента, определяемый цветом второго шара, и будем отмечать).

Таким образом мы имеем $k_n(B)$ повторный условного эксперимента. В условном эксперименте событие A происходит только тогда, когда в исходном эксперименте прои-

зойдут и A и B , т. е. AB . Это значит, что событие A в $k_n(B)$ условных экспериментах произойдет $k_n(A \cdot B)$ раз. Частота появления события в условном эксперименте будет $\frac{k_n(A \cdot B)}{k_n(B)}$. Для того чтобы найти предельное значение этого отношения при неограниченном возрастании числа экспериментов, представим его так:

$$\frac{k_n(A \cdot B)/n}{k_n(B)/n} = \frac{v_n(A \cdot B)}{v_n(B)},$$

где $v_n(A \cdot B)$ и $v_n(B)$ — частоты соответствующих событий.

Поскольку частоты стремятся к вероятностям, то предел этого выражения равен $\frac{P(A \cdot B)}{P(B)}$ (предполагаем $P(B) \neq 0$).

Будем обозначать условную вероятность события A относительно события B через $P(A/B)$. Эта вероятность определяется соотношением:

$$P(A/B) = \frac{P(A \cdot B)}{P(B)}. \quad (1)$$

Из данного выражения вытекает следующая общая формула произведения вероятностей:

$$P(A \cdot B) = P(A/B) \cdot P(B).$$

Найдем связь между условными вероятностями $P(A/B)$ и $P(B/A)$. Имеем

$$P(A \cdot B) = P(B/A) \cdot P(A) = P(A/B) \cdot P(B).$$

Значит,

$$P(B/A) = P(A/B) \frac{P(B)}{P(A)}.$$

Приведем пример «обратного» влияния будущего на прошлое. Опять рассмотрим тот же эксперимент последовательного извлечения двух шаров из урны с n белыми и m черными шарами. События B и A , как и раньше, обозначают извлечение черного шара в первый и второй раз.

$$\text{Мы уже вычислили } P(A/B) = \frac{m-1}{n+m-1}, \quad P(A) = \frac{m}{n+m}.$$

Очевидно, $P(B) = \frac{m}{n+m}$. Поэтому

$$P(B/A) = P(A/B) \cdot \frac{P(B)}{P(A)} = P(A/B) = \frac{m-1}{n+m-1}.$$

Оказывается, событие A , происходящее позже B , влияет на B точно так, как и событие B на A .

Во многих случаях условные вероятности вычисляются достаточно просто. Рассмотрим два наиболее характерных случая.

1. Пусть эксперимент имеет конечное число равновероятных элементарных событий: E_1, \dots, E_n . Если $B = E_{i_1}, \dots, E_{i_l}$, то для вычисления условной вероятности нужно подсчитать, какие из событий E_{i_1}, \dots, E_{i_l} влекут событие A (пусть таких событий k). Тогда

$$P(A/B) = \frac{P(A \cdot B)}{P(B)} = \frac{\frac{k}{n}}{\frac{l}{n}} = \frac{k}{l}.$$

Таким образом, условная вероятность вычисляется точно так, как безусловная, только нужно считать, что в эксперименте имеется l элементарных событий (те, которые составляют B), и подсчитать, сколько из этих событий благоприятствуют A .

Рассмотрим следующий пример. Пусть в семье имеется трое детей. Считаем, что мальчики и девочки равновероятны. Имеем всего 8 равновероятных событий: (М, М, М), (М, М, Д), (М, Д, М) (Д, М, М), (М, Д, Д), (Д, М, Д), (Д, Д, М), (Д, Д, Д) (Д — девочка, М — мальчик). Пусть событие B заключается в том, что в семье есть мальчик; A — число мальчиков больше, чем число девочек. Событию B благоприятствуют 7 элементарных событий, из них при 4 происходит A , значит, $P(A/B) = \frac{4}{7}$. Если спрашивать у мальчиков, кого у них в семье больше — братьев или сестер, они чаще будут отвечать, что братьев. Если тот же самый вопрос задавать девочкам, то они чаще будут говорить, что сестер.

2. Пусть эксперимент заключается в случайном выборе точки из некоторой фигуры G (будем ее для определенности считать плоской), а событие B заключается в том, что точка попала в часть G_1 фигуры G . Пусть, далее, A — событие «точка попала во множество G_2 » (тоже часть G).

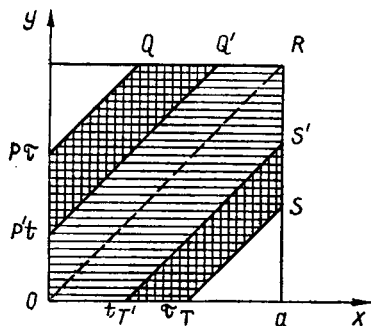
Если известно, что точка лежит в G_1 , то имеет смысл рассматривать лишь те G_2 , которые являются частями G_1 . Тогда

$$P(A/B) = \frac{\text{площадь } G_2}{\text{площадь } G_1} : \frac{\text{площадь } G}{\text{площадь } G_1} = \frac{\text{площадь } G_2}{\text{площадь } G}.$$

Таким образом, условный эксперимент в данном случае есть тоже случайный выбор точки, только уже из области G_1 , состоящей из тех элементарных событий, которые влекут событие B .

Рассмотрим в качестве примера ту же задачу о встрече двух объектов, которая обсуждалась в § 5. Объекты на промежутке $[0, a]$ попадают независимо на место встречи и находятся там время τ . Событие B заключается в том,

что встреча объектов состоялась. Пусть событие A состоит в том, что первому объекту пришлось ожидать второй не меньше, чем t ($t < \tau$). Найдём $P(A/B)$. Воспользуемся той же геометрической интерпретацией, которая была приведена в § 5. На рисунке точки квадрата, соответствующие событиям B и A , отмечены различной штриховкой. Событию B отвечает многоугольник $OP\tau QRSt\tau T$, а событию A — две трапеции $P'tP\tau QQ'$, $tT'S'S\tau T$.



Площадь $OP\tau QRSt\tau T$ вычислена в § 5 и равна $a^2 - (a - \tau)^2$. Чтобы вычислить площадь двух трапеций, заметим, что они вместе с многоугольником $OP'tQ'RS'tT'$ составляют $OP\tau QRSt\tau T$. Площадь $OP'tQ'RS'tT'$ определяется точно так же, как и площадь $OP\tau QRSt\tau T$: она равна $a^2 - (a - t)^2$. Значит, искомая условная вероятность будет

$$\frac{a - (a - \tau)^2 - a^2 - (a - t)^2}{a^2 - (a - \tau)^2} = \frac{(a - t)^2 - (a - \tau)^2}{a^2 - (a - \tau)^2}.$$

Например, при $\tau = \frac{a}{3}$, $t = \frac{a}{6}$ эта вероятность равна

$$\frac{\left(\frac{5}{6}\right)^2 - \left(\frac{2}{3}\right)^2}{1 - \left(\frac{2}{3}\right)^2} = \frac{\frac{25}{36} - \frac{16}{36}}{1 - \frac{4}{9}} = \frac{9}{20}.$$

Формула полной вероятности. Пусть B_1, \dots, B_m — некоторая полная группа событий, т. е. B_k несовместимы, и $B_1 + \dots + B_m$ — достоверное событие. Во многих случаях можно просто вычислить условные вероятности $P(A/B_k)$.

Формула полной вероятности дает выражение вероятности события A через условные вероятности.

Имеем

$$P(A) = P(A \cdot (B_1 + \dots + B_m)) = \\ = P(A \cdot B_1) + P(A \cdot B_2) + \dots + P(A \cdot B_m).$$

Воспользовавшись формулой (1), можем записать

$$P(A \cdot B_k) = P(A/B_k) P(B_k).$$

Поэтому

$$P(A) = P(A/B_1) P(B_1) + P(A/B_2) P(B_2) + \dots \\ \dots + P(A/B_m) P(B_m). \quad (2)$$

Это и есть формула полной вероятности.

Распределение генотипов в популяции. Рассмотрим некоторую совокупность особей какого-нибудь биологического вида. Нас будет интересовать, как часто будут встречаться те или иные генотипы по одному признаку, точнее, как они будут меняться от поколения к поколению.

Возможные генотипы AA , Aa , и aa . Предполагается, что родители комбинируются случайным образом. Пусть далее $p(AA)$, $p(Aa)$, $p(aa)$ — вероятности того, что случайно выбранная особь имеет соответствующий генотип.

Вычислим, пользуясь формулой полной вероятности, значения вероятностей для генотипов потомков этих особей. Будем исходную совокупность особей называть родительской популяцией, их детей — популяцией первого поколения, детей у первого поколения — популяцией второго поколения и т. д. Вероятности для генотипов в k -м поколении обозначим через $p_k(AA)$, $p_k(Aa)$ и $p_k(aa)$.

Для особи первого поколения возможны такие комбинации родителей (первым будем записывать генотип матери): $AA \times AA$, $Aa \times AA$, $aa \times AA$, $AA \times Aa$, $Aa \times Aa$, $aa \times Aa$, $AA \times aa$, $Aa \times aa$, $aa \times aa$. Эти комбинации образуют полную группу событий. Так как родители выбираются независимо, то вероятность каждой комбинации есть произведение соответствующих вероятностей. Скажем,

$$P(AA \times AA) = p(AA) p(AA)$$

($P(AA \times AA)$ есть вероятность того, что мать и отец имеют генотип AA). Поэтому

$$P_1(AA) = P(AA/AA \times AA) p(AA) p(AA) +$$

$$+ P(AA/AA \times Aa) p(AA) p(Aa) + P(AA/Aa \times AA) \times \\ \times p(Aa) p(AA) + P(AA/Aa \times Aa) p(Aa) p(Aa)$$

(при других родителях генотип AA у ребенка появиться не может).

Очевидно,

$$P(AA/AA \times AA) = 1,$$

$$P(AA/AA \times Aa) = P(AA/Aa \times AA) = \frac{1}{2},$$

$$P(AA/Aa \times Aa) = \frac{1}{4}.$$

Следовательно,

$$p_1(AA) = p^2(AA) + p(Aa) p(AA) + \frac{1}{4} p^2(Aa) = \\ = \left(p(AA) + \frac{1}{2} p(Aa) \right)^2.$$

Аналогично

$$p_1(aa) = \left(p(aa) + \frac{1}{2} p(Aa) \right)^2.$$

Поскольку $p_1(AA) + p_1(Aa) + p_1(aa) = 1$, то

$$p_1(Aa) = 1 - \left(p(AA) + \frac{1}{2} p(Aa) \right)^2 - \\ - \left(p(aa) + \frac{1}{2} p(Aa) \right)^2 = (p(AA) + p(Aa) + p(aa))^2 - \\ - \left(p(AA) + \frac{1}{2} p(Aa) \right)^2 - \left(p(aa) + \frac{1}{2} p(Aa) \right)^2 = \\ = 2 \left(p(AA) + \frac{1}{2} p(Aa) \right) \left(p(aa) + \frac{1}{2} p(Aa) \right).$$

При вычислении вероятностей для второго поколения можно использовать полученные выражения, только вместо вероятностей для родительской популяции подставлять вероятности для первого поколения. Поэтому

$$p_2(AA) = \left(p_1(AA) + \frac{1}{2} p_1(Aa) \right)^2 = \\ = \left[\left(p(AA) + \frac{1}{2} p(Aa) \right)^2 + \left(p(AA) + \frac{1}{2} p(Aa) \right) \times \right.$$

$$\begin{aligned} & \times \left(p(aa) + \frac{1}{2} p(Aa) \right) \Big]^2 = \left[\left(p(AA) + \frac{1}{2} p(Aa) \right) \times \right. \\ & \left. \times \left(p(AA) + \frac{1}{2} p(Aa) + p(aa) + \frac{1}{2} p(Aa) \right) \right]^2 = \\ & = \left(p(AA) + \frac{1}{2} p(Aa) \right)^2. \end{aligned}$$

Аналогично получаем

$$\begin{aligned} p_2(aa) &= \left(p(aa) + \frac{1}{2} p(Aa) \right)^2, \\ p_2(Aa) &= 1 - p_2(AA) - p_2(aa) = \\ &= 2 \left(p(AA) + \frac{1}{2} p(Aa) \right) \left(p(aa) + \frac{1}{2} p(Aa) \right). \end{aligned}$$

Следовательно, вероятности различных генотипов уже с первого поколения являются неизменными.

Предположим, что родительская популяция состояла только из гомозиготных особей, причем вероятность доминантного признака AA была p , а вероятность рецессивного — $1 - p$. Тогда для всех последующих поколений вероятность генотипа AA равна p^2 , генотипа Aa — $2p(1 - p)$, генотипа aa — $(1 - p)^2$.

Распределение генотипов в популяции с учетом пола. Будем учитывать теперь и пол входящих в популяцию особей. Передача пола потомкам также обуславливается хромосомным механизмом. В наборе хромосом имеется одна пара, которая и определяет пол. У женщин — это пара двух идентичных хромосом XX; у мужчин хромосомы в паре различаются: XY. Возможные комбинации хромосом у потомков XX (одна X — хромосома от матери, вторая — от отца, ребенок — девочка), XY (X — хромосома от матери, Y — хромосома от отца, ребенок — мальчик). Если интересующий нас признак определяется генами, локализованными на другой паре хромосом, то этот признак передается потомству независимо от пола. Рассмотрим сначала этот случай.

Пусть для особей женского пола вероятности генотипов в родительской популяции будут $p(AA)$, $p(Aa)$, $p(aa)$, а для особей мужского пола — $p'(AA)$, $p'(Aa)$, $p'(aa)$.

Рассуждая точно так, как в предыдущем случае, получаем

$$p_1(AA) = \left(p(AA) + \frac{1}{2} p(Aa) \right) \left(p'(AA) + \frac{1}{2} p'(Aa) \right),$$

$$p_1(aa) = \left(p(aa) + \frac{1}{2} p(Aa) \right) \left(p'(aa) + \frac{1}{2} p'(Aa) \right),$$

$$p_1(Aa) = \left(p(AA) + \frac{1}{2} p(Aa) \right) \left(p'(aa) + \frac{1}{2} p'(Aa) \right) + \\ + \left(p(aa) + \frac{1}{2} p(Aa) \right) \left(p'(AA) + \frac{1}{2} p'(Aa) \right).$$

При этом распределение генотипов в первом поколении не будет зависеть от пола. Со второго поколения вероятности генотипов уже не будут меняться, так как, взяв в качестве родительской популяции популяцию первого поколения, мы приходим к выше рассмотренной задаче. Используя полученные формулы, можем найти вероятности для распределения генотипов, начиная со второй популяции: при $k \geq 2$

$$p_k(AA) = \left(\frac{p(AA) + p'(AA)}{2} + \frac{p(Aa) + p'(Aa)}{4} \right)^2,$$

$$p_k(aa) = \left(\frac{p(aa) + p'(aa)}{2} + \frac{p(Aa) + p'(Aa)}{4} \right)^2,$$

$$p_k(Aa) = 2 \left(\frac{p(AA) + p'(AA)}{2} + \frac{p(Aa) + p'(Aa)}{4} \right) \times \\ \times \left(\frac{p(aa) + p'(aa)}{2} + \frac{p(Aa) + p'(Aa)}{4} \right).$$

Признаки, сцепленные с полом. Это такие признаки, которые определяются генами, расположенными на половых хромосомах. Если это Y-хромосома, то такой признак обязательно передается от отца к сыну, а женщины его не имеют. Так что никакой задачи не возникает.

Рассмотрим признак, определяемый X-хромосомой (примерами таких признаков являются гемофилия, дальтонизм). Они могут наблюдаться как у мужчин, так и у женщин.

Обозначая соответствующие признаку гены через A и a, будем иметь такие генотипы: для женщин — AA, Aa, aa (поскольку у них две X-хромосомы), для мужчин A, a (у них лишь одна X-хромосома). Будем обозначать через $p(A)$ и $p(a)$ условные вероятности генотипов A и a при условии, что индивид имеет мужской пол, а через

$p(AA)$, $p(Aa)$, $p(aa)$ — условные вероятности генотипов при условии, что индивид имеет женский пол.

Существует 6 возможных комбинаций родительских генотипов: $AA \times A$, $AA \times a$, $Aa \times A$, $Aa \times a$, $aa \times A$, $aa \times a$. Имеем $P(A/AA \times A) = \frac{1}{2}$ (это вероятность рождения мальчика); $P(A/AA \times a) = \frac{1}{2}$ (на тех же основаниях); $P(A/Aa \times a) = P(A/Aa \times A) = \frac{1}{4}$ (мальчику с вероятностью $\frac{1}{2}$ попадет от матери ген A). В остальных случаях мальчик с генотипом A появиться не может (X -хромосома к мальчику попадает от матери).

Аналогично $P(a/aa \times a) = \frac{1}{2}$, $P(a/aa \times A) = \frac{1}{2}$, $P(a/Aa \times A) = P(a/Aa \times a) = \frac{1}{4}$. Девочки получают обязательно X -хромосому отца и одну из X -хромосом матери. Поэтому $P(AA/AA \times A) = \frac{1}{2}$, $P(AA/Aa \times A) = \frac{1}{4}$; в других случаях генотип AA у девочки невозможен:

$$P(Aa/AA \times a) = \frac{1}{2}, P(Aa/Aa \times A) = \frac{1}{4},$$

$$P(Aa/Aa \times a) = \frac{1}{4},$$

$$P(aa/aa \times A) = \frac{1}{2}, P(aa/aa \times a) = \frac{1}{2}, P(aa/Aa \times a) = \frac{1}{4}.$$

Будем, как и раньше, условные вероятности генотипов в k -м поколении при условии, что они имеют соответствующий пол, обозначать $p_k(A)$, $p_k(a)$, $p_k(AA)$, $p_k(Aa)$, $p_k(aa)$. Тогда

$$\begin{aligned} p_1(A) &= \\ & \frac{1}{2} p(AA) p(A) + \frac{1}{2} p(AA) p(a) + \frac{1}{4} p(Aa) p(a) + \\ & \quad + \frac{1}{4} p(Aa) p(A) \\ &= \frac{1}{2} \end{aligned}$$

Мы вычисляем условную вероятность генотипа при условии, что родился мальчик, поэтому полученную вероятность генотипа А — а он может быть только у мальчика — делим на $\frac{1}{2}$ — вероятность рождения мальчика. Так как $p(A) + p(a) = 1$, то

$$p_1(A) = p(AA) + \frac{1}{2} p(Aa).$$

Аналогично

$$p_1(a) = p(aa) + \frac{1}{2} p(Aa).$$

Точно так вычисляем вероятности генотипов для девочки в первом поколении:

$$p_1(AA) = \left[p(AA) + \frac{1}{2} p(Aa) \right] p(A),$$

$$p_1(aa) = \left[p(aa) + \frac{1}{2} p(Aa) \right] p(a),$$

$$p_1(Aa) = p(AA) p(a) + p(aa) p(A) + \frac{1}{2} p(Aa).$$

Для того чтобы вычислить вероятности генотипов в k -м поколении, нужно взять за исходные вероятности те, которые будут в $k - 1$ -м поколении, и воспользоваться полученными формулами. Таким образом,

$$p_k(A) = p_{k-1}(AA) + \frac{1}{2} p_{k-1}(Aa),$$

$$p_k(a) = p_{k-1}(aa) + \frac{1}{2} p_{k-1}(Aa),$$

$$p_k(AA) = \left[p_{k-1}(AA) + \frac{1}{2} p_{k-1}(Aa) \right] p_{k-1}(A),$$

$$p_k(aa) = \left[p_{k-1}(aa) + \frac{1}{2} p_{k-1}(Aa) \right] p_{k-1}(a),$$

$$p_k(Aa) = p_{k-1}(AA) p_{k-1}(a) + \\ + p_{k-1}(aa) p_{k-1}(A) + \frac{1}{2} p_{k-1}(Aa).$$

Оказывается, что если вероятности генотипов для особей разного пола связаны определенным образом, то тогда они не меняются в различных поколениях.

Пусть

$$p(A) = p, \quad p(a) = 1 - p, \quad p(AA) = p^2, \quad p(Aa) = 2p(1 - p), \\ p(aa) = (1 - p)^2.$$

Тогда

$$\begin{aligned} p_1(A) &= [p^2 + p(1-p)] = p, & p_1(a) &= (1-p), \\ p_1(AA) &= [p^2 + p(1-p)]p = p^2, & p_1(aa) &= (1-p)^2, \\ p_1(Aa) &= 2p(1-p), \end{aligned}$$

т. е. частота генотипов в первом поколении такая же, как и в родительской популяции. Поэтому она все время остается постоянной. Каковы бы ни были начальные вероятности для генотипов, они через несколько поколений становятся примерно неизменными, так что между вероятностями генотипов у мужчин и женщин имеется указанная выше зависимость:

$$p(AA) = p^2(A), \quad p(Aa) = 2p(A)p(a), \quad p(aa) = p^2(a).$$

Если при передаче признака имеется доминирование (например, дальтонизм рецессивен), то рецессивный признак проявляется гораздо чаще у мужчин, чем у женщин. Так, вероятность дальтонизма для мужчины $\frac{1}{100}$, а для женщины $\left(\frac{1}{100}\right)^2 = 0,0001$.

§ 12. Формула Байесса

Пусть E_1, \dots, E_m — полная группа событий, которые могут происходить в эксперименте, и A — некоторое событие. Если известны вероятности событий E_k и условная вероятность A при условии, что произошло E_k для $k = 1, \dots, m$, то по формуле полной вероятности мы можем определить и вероятность события A , и условные вероятности событий E_k при условии, что произошло A . Имеем

$$\begin{aligned} P(A \cdot E_k) &= P(A/E_k)P(E_k), \\ P(A) &= P(A/E_1)P(E_1) + \dots + P(A/E_m)P(E_m), \\ P(E_k/A) &= \frac{P(A/E_k)P(E_k)}{P(A/E_1)P(E_1) + \dots + P(A/E_m)P(E_m)}. \end{aligned} \quad (1)$$

Формула (1) носит название формулы Байесса.

Эта формула обычно интерпретируется следующим образом. Предположим, что событие A может произойти при одной из m взаимоисключающих гипотез. Событие E_k играет роль k -й гипотезы. Известна вероятность события A при каждой из гипотез. Из априорных соображений гипотезам можно приписать определенные вероятности.

Пусть в результате эксперимента произошло событие A . Условные вероятности гипотез E_k при условии, что наблюдалось A , называются *апостериорными вероятностями*: (они вычисляются с учетом того, что произошло в эксперименте, в отличие от исходных вероятностей, которые называются *априорными*). Формула Байесса дает значения апостериорных вероятностей гипотез.

Рассмотрим сначала схематический пример. Имеются 3 урны, каждая содержит 10 шаров, причем в первой 8 белых и 2 черных, во второй 5 белых и 5 черных, в третьей 2 белых и 8 черных. Эксперимент заключается в том, что выбирается случайно одна урна, а затем из урны извлекается два шара. Наша цель — на основании результатов эксперимента установить, какая из урн была выбрана.

В этом случае имеются 3 гипотезы: E_k — выбрана k -я урна, $k = 1, 2, 3$.

Пусть событие A состоит в том, что извлечены 2 черных шара. Тогда

$$P(A/E_1) = \frac{2}{10 \cdot 9} = \frac{1}{45}, \quad P(A/E_2) = \frac{5 \cdot 4}{10 \cdot 9} = \frac{10}{45},$$

$$P(A/E_3) = \frac{8 \cdot 7}{10 \cdot 9} = \frac{28}{45}.$$

Априорные вероятности гипотез равны $\frac{1}{3}$. Поэтому для апостериорных вероятностей будем иметь следующие значения:

$$P(E_1/A) = \frac{\frac{1}{45} \cdot \frac{1}{3}}{\frac{1}{3} \left(\frac{1}{45} + \frac{10}{45} + \frac{28}{45} \right)} = \frac{1}{39},$$

$$P(E_2/A) = \frac{\frac{10}{45} \cdot \frac{1}{3}}{\frac{1}{3} \left(\frac{1}{45} + \frac{10}{45} + \frac{28}{45} \right)} = \frac{10}{39},$$

$$P(E_3/A) = \frac{28}{39}.$$

Таким образом, с вероятностью примерно 0,72 можно утверждать, что справедлива третья гипотеза.

Пусть B — событие «извлечены черный и белый шары»

(в любом порядке). Тогда

$$P(B/E_1) = \frac{2 \cdot 8}{10 \cdot 9}, \quad P(B/E_2) = \frac{5 \cdot 5}{10 \cdot 9},$$

$$P(B/A_3) = \frac{2 \cdot 8}{10 \cdot 9}$$

и

$$P(E_1/B) = \frac{16}{57}, \quad P(E_2/B) = \frac{25}{57}, \quad P(E_3/B) = \frac{16}{57}.$$

В этом случае апостериори наиболее вероятна вторая гипотеза, хотя ее вероятность и меньше половины.

Если C — событие «извлечены 2 белых шара», то

$$P(E_1/C) = \frac{28}{39}, \quad P(E_2/C) = \frac{10}{39}, \quad P(E_3/C) = \frac{1}{39}.$$

Как видим, если происходит A или C , гипотезы хорошо различаются: с большой вероятностью можно выбрать одну из гипотез. Если же происходит B , то, даже выбирая наиболее вероятную гипотезу, мы с вероятностью больше $\frac{1}{2}$ производим неправильный выбор ($P(E_1 + E_3/B) = \frac{32}{57} \approx \approx 0,56$).

Пусть имеется родительская пара с доминантным фенотипом по некоторому признаку. Нас интересует, являются ли родительские особи гомозиготными или гетерозиготными по этому признаку. Для проверки гипотезы рассматриваются их потомки. Предположим, что их было k . Если среди них есть хотя бы 1 фенотип, который отвечает рецессивному аллелю, то это означает, что оба родителя гетерозиготны (только при генотипе Aa у обоих родителей может появиться потомок с генотипом aa). Поэтому задача имеет нетривиальное решение лишь в том случае, когда у всех k потомков тот же фенотип, что и у родителей.

Есть три гипотезы относительно генотипов родителей: I — $AA \times AA$; II — $AA \times Aa$, III — $Aa \times Aa$. Вероятности этих гипотез определяются вероятностями для различных генотипов в популяции, откуда взяты особи (мы видели в § II, что эти вероятности в популяции не меняются в различных поколениях).

Пусть эти вероятности равны p_I, p_{II}, p_{III} . Если B -событие, что из k потомков ни один не имеет генотип aa , то

$$P(B/I) = 1, \quad P(B/II) = 1, \quad P(B/III) = \left(\frac{3}{4}\right)^k$$

(это условные вероятности при I, II и III гипотезах).

Значит,

$$P(I/B) = \frac{p_I}{p_I + p_{II} + p_{III} \cdot \left(\frac{3}{4}\right)^k},$$

$$P(II/B) = \frac{p_{II}}{p_I + p_{II} + p_{III} \cdot \left(\frac{3}{4}\right)^k},$$

$$P(III/B) = \frac{p_{III} \left(\frac{3}{4}\right)^k}{p_I + p_{II} + p_{III} \cdot \left(\frac{3}{4}\right)^k}.$$

Например, при $p_I = p_{II} = p_{III}$ будем иметь

$$P(I/B) = P(II/B) = \frac{1}{2 + \left(\frac{3}{4}\right)^k}, \quad P(III/B) = \frac{\left(\frac{3}{4}\right)^k}{2 + \left(\frac{3}{4}\right)^k}$$

При больших k первая вероятность равна примерно $\frac{1}{2}$, а $P(III/B) \approx 0$. В общем случае при достаточно больших k

$$P(III/B) = 0, \quad \text{а } P(I/B) = \frac{p_I}{p_I + p_{II}}, \quad P(II/B) = \frac{p_{II}}{p_I + p_{II}}.$$

Т. е. после наблюдения события B мы можем гипотезу III отбросить; при этом вероятности (апостериорные) гипотез I и II одинаково возрастут.

Очевидно, что выбор гипотезы I или II не влияет на фенотип потомства. Поэтому для различения этих гипотез нужно привлекать следующее поколение. Пусть имеются лишь две гипотезы — I и II — и во втором поколении, получаемом случайным скрещиванием из первого, имеется l особей. Если хотя бы у одной был генотип aa , то справедлива гипотеза II. Пусть их фенотип совпадает с фенотипом родительских особей. Тогда, если B — указанное событие, то $P(B/I) = 1$. Для вычисления $P(B/II)$ заметим, что генотип aa во втором поколении возможен с вероятностью

стью $\frac{1}{4}$, если родители имели генотип Aa и Aa, который в первом поколении (у одного родителя) появляется с вероятностью $\frac{1}{2}$. Поэтому вероятность генотипа aa во втором поколении равна $\frac{1}{16}$ и

$$P(B/II) = \left(\frac{15}{16}\right)^l.$$

Отсюда

$$P(I/B) = \frac{p_I}{p_I + \left(\frac{15}{16}\right)^l p_{II}}, \quad P(II/B) = \frac{\left(\frac{15}{16}\right)^l p_{II}}{p_I + \left(\frac{15}{16}\right)^l p_{II}}.$$

II. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

§ 13. Примеры случайных величин

Результаты многих экспериментов выражаются числами, так что сами эксперименты можно рассматривать как изменение некоторых величин. Рассмотрим наиболее характерные примеры.

1. Из партии изделий для контроля выбирается n изделий; в результате их проверки обнаруживается ν дефектных. Это число ν является случайным, оно может иметь значения $0, 1, \dots, m$. В каждом эксперименте ν принимает вполне определенное значение, но при повторении эксперимента число дефектных изделий меняется случайным образом. Случайное число ν полностью характеризуется $m + 1$ событиями: E_0, E_1, \dots, E_m , событие E_k заключается в том, что число ν обнаруженных дефектных изделий равно k . Найдем вероятности событий E_k .

Пусть N — число изделий в партии, среди них n дефектных. Если не обращать внимание на порядок извлечения изделий, то число способов извлечь m изделий из N есть $C_N^m = \frac{N!}{(N-m)!m!}$ (см. § 3). Число способов извлечь k дефектных изделий равно произведению числа способов извлечь k предметов из m на число способов извлечь $m - k$ из $N - n$ (если имеется k дефектных изделий, то будет и $m - k$ годных изделий, которые должны быть выбраны из общего числа $N - n$ годных изделий). Поэтому

$$P(E_k) = \frac{C_n^k C_{N-n}^{m-k}}{C_N^m} = \frac{(N-m)! m! n! (N-n)!}{N! k! (n-k)! (m-k)! (N+k-n-m)!},$$

Или можно так:

$$P\{\nu = k\} = \frac{(N-m)! (N-n)! n! m!}{N! (N+k-n-m)! (n-k)! (m-k)! k!},$$

$$k = 0, \dots, m. \quad (1)$$

Набор этих вероятностей определяет распределение случайной величины v . Зная ее распределение, мы можем судить, как часто случайная величина v принимает те или иные значения.

2. Пусть в эксперименте событие A может происходить с вероятностью p . Эксперимент проводится m раз. Число появлений события A в m экспериментах является случайной величиной μ . Набор вероятностей $P\{\mu = k\}$, $k = 0, \dots, m$ называется распределением этой величины.

Как вытекает из результатов § 6,

$$P\{\mu = k\} = C_m^k p^k (1 - p)^{m-k} = \frac{m!}{k!(m-k)!} p^k (1 - p)^{m-k}. \quad (2)$$

Любопытно отметить, что при $N \rightarrow \infty$ и $\frac{n}{N} \rightarrow p$ правая часть (1) стремится к правой части (2). Это вытекает из того, что

$$\frac{(N-m)!}{N!} = \frac{1}{N(N-1)\dots(N-m+1)} \sim N^{-m},$$

$$\frac{(N-n)!}{(N-n-(m-k))!} = (N-n)(N-n-1)\dots$$

$$\dots (N-n-m+k+1) \sim (N-n)^{m-k},$$

$$\frac{n!}{(n-k)!} = n(n-1)\dots(n-k+1) \sim n^k$$

(здесь знаком \sim соединяются выражения, отношение которых стремится к единице). Поэтому

$$\begin{aligned} & \frac{(N-m)!(N-n)!n!m!}{N!(N+k-n-m)!(n-k)!(m-k)!k!} \sim \\ & \sim C_m^k \frac{n^k (N-n)^{m-k}}{N^m} = C_m^k \left(\frac{n}{N}\right)^k \left(1 - \frac{n}{N}\right)^{m-k}. \end{aligned}$$

Следовательно, при больших N сложное выражение (1) можно заменить более простым (2), полагая $p = \frac{n}{N}$, если только при этом и n достаточно велико.

3. Рассмотрим теперь число появлений некоторого редкого события. Пусть, например, v — число регистраций космических частиц некоторым счетчиком за определенное время. v будет случайной величиной, которая может принимать значения $0, 1, 2, \dots$ Как было показано в § 7, событие

$\{v = m\}$ имеет вероятность вида $\frac{a^m}{m!} e^{-a}$. В отличие от величин, приведенных в примерах 1 и 2, здесь случайная величина v принимает бесконечное число значений. Распределение ее определяется уже бесконечной последовательностью вероятностей

$$P\{v = k\} = \frac{a^k}{k!} e^{-a}, \quad k = 0, 1, 2, \dots \quad (3)$$

($a > 0$ — некоторая величина).

Заметим, что вероятности вида (3) также могут быть получены из (1) с помощью определенного предельного перехода. Пусть $N \rightarrow \infty$, $n \rightarrow \infty$ и $\frac{n}{N} \rightarrow 0$. Если при этом и $m \rightarrow \infty$, так что $mnN^{-1} \rightarrow a$, $m^2N^{-1} \rightarrow 0$, то правая часть (1) стремится к правой части (3). Это устанавливается таким же образом, как в § 7 была получена формула Пуассона.

Величины, рассмотренные в примерах 1—3, принимают целочисленные значения. Однако случайные величины могут принимать и произвольные вещественные значения.

4. Пусть из отрезка $[a, b]$ на вещественной оси выбирается случайно точка ξ . ξ — случайная величина, принимающая произвольное вещественное значение между a и b . Значит, множество возможных значений величины ξ бесконечно. Но, в отличие от случайной величины, рассмотренной в примере 3, каждое свое значение, как мы увидим ниже, она принимает с вероятностью 0. Поэтому нельзя говорить о распределении этой случайной величины в том смысле, как мы делали в примерах 1—3; распределение ее невозможно задать с помощью вероятностей, с которыми она принимает свои возможные значения.

Как же все-таки охарактеризовать с вероятностной точки зрения эту случайную величину? При измерении величины, принимающей произвольные значения, нас в первую очередь интересует, между какими делениями измерительной шкалы прибора, с помощью которого величина измеряется, она окажется. Другими словами, наше внимание привлекают события вида $\alpha < \xi < \beta$ (так, если шкалу прибора можно представить линейкой с отметками через интервалы длины h , то нас интересует, между какими делениями вида kh и $(k+1)h$ величина окажется, т. е.

нас интересует, какое из событий вида $kh < \xi < (k + 1)h$ произошло).

Как мы видели в § 5, если точка выбирается случайно из отрезка $[a, b]$, то вероятность того, что она будет выбрана из отрезка $[\alpha, \beta]$, лежащего в $[a, b]$, равна отношению длины этого отрезка к длине отрезка $[a, b]$. Таким образом,

$$P\{\alpha < \xi < \beta\} = \frac{\beta - \alpha}{b - a}.$$

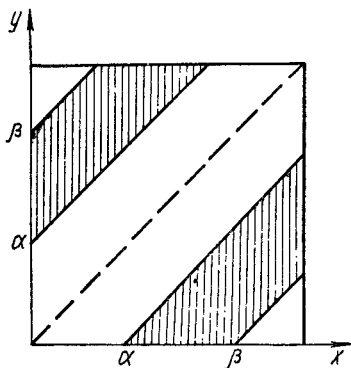
Совокупность этих вероятностей и является аналогом распределения для непрерывно меняющейся величины. Покажем, что при любом x из $[a, b]$ вероятность события $\{\xi = x\}$ равна нулю. Действительно, если это событие произошло, то произошло и событие $\{\alpha < \xi < b\}$, каковы бы ни были α и β , для которых $\alpha < x < \beta$. Значит, событие $\{\xi = x\}$ влечет событие $\{\alpha < \xi < \beta\}$ для α и β указанного вида. Поэтому

$$P\{\xi = x\} \leq P\{\alpha < \xi < \beta\} = \frac{\beta - \alpha}{b - a}.$$

Но можно выбрать $\alpha < x < \beta$ так, чтобы $\beta - \alpha$ было сколь угодно малым. Значит, $P\{\xi = x\}$ меньше произвольного положительного числа, поэтому эта вероятность равна нулю.

5. Рассмотрим задачу о встрече, о которой уже шла речь в § 5. Величина времени, проходящая от момента появления первого объекта на месте встречи до момента появления второго объекта, есть случайная. Обозначим ее через η . Если объекты появляются на месте встречи в течение промежутка времени $[0, a]$, то η может принимать любые вещественные значения из промежутка $[0, a]$. Какова вероятность того, что η примет значение из некоторого промежутка (α, β) , лежащего в $[0, a]$?

В § 5 уже указывалось, что эксперимент в задаче о встрече можно рассматривать как случайный выбор точки из квадрата со стороной a , если выбранная точка будет иметь координаты x и y , то эти числа и будут временами



появления объектов на месте встречи. Так что при выборе точки $(x; y)$ величина η принимает значение $|x - y|$. Для того чтобы значения величины η попали в интервал (α, β) , нужно, чтобы точка $(x; y)$ попала в одну из двух заштрихованных трапеций, симметрично расположенных относительно диагонали квадрата. Основаниями верхней трапеции служат отрезки прямых $y = \alpha + x$, $y = \beta + x$. Вычисляя площадь трапеции как разность площадей треугольников, получаем

$$P\{\alpha < \eta < \beta\} = 2 \left[\frac{(a - \alpha)^2}{2a^2} - \frac{(\alpha - \beta)^2}{2a^2} \right] = \\ = \frac{(\beta - \alpha)[2a - \beta - \alpha]}{a^2}.$$

Это равенство определяет распределение величины η . Если длина $\beta - \alpha \rightarrow 0$, то и вероятность стремится к нулю. Поэтому $P\{\eta = x\} = 0$ для всех x .

Случайные величины, которые каждое свое возможное значение принимают с вероятностью 0, называются *непрерывными*. Примерами непрерывных случайных величин могут служить размеры и веса отдельных представителей данного биологического вида, энергии космических частиц, моменты солнечных вспышек и т. д.

§ 14. Функция распределения

Пусть в некотором эксперименте наблюдается случайная величина ξ . Ее основной вероятностной характеристикой является функция распределения, определяемая равенством

$$F(x) = P\{\xi < x\}$$

для всех вещественных x . Зная функцию распределения, можно определить вероятность того, что величина ξ попала в интервал $[\alpha, \beta)$:

$$P\{\alpha \leq \xi < \beta\} = P\{\xi < \beta\} - P\{\xi < \alpha\} = F(\beta) - F(\alpha).$$

Пусть $\beta_n \rightarrow \alpha$ и $\beta_n > \alpha$. Тогда

$$\lim_{n \rightarrow \infty} [F(\beta_n) - F(\alpha)] = \lim_{n \rightarrow \infty} P\{\alpha \leq \xi < \beta_n\} = P\{\xi = \alpha\}.$$

Последнее свойство вытекает из следующей аксиомы вероятности, которая считается выполненной на вероятностных пространствах с бесконечным числом исходов (до сих

пор эта аксиома нам не была нужна и поэтому не упоминалась).

Аксиома непрерывности. а) Если A_n — последовательность событий, причем A_n для всех n влечет A_{n+1} , а событие A заключается в том, что происходит хотя бы одно из событий A_n , то

$$P(A) = \lim_{n \rightarrow \infty} P(A_n).$$

б) Если A_n такая последовательность событий, что для всех n A_{n+1} влечет A_n , а A — событие, заключающееся в том, что происходят все события A_n , то

$$P(A) = \lim_{n \rightarrow \infty} P(A_n).$$

Пусть β_n , убывая, стремится к α , тогда $\{\alpha \leq \xi < \beta_{n+1}\} \subset \{\alpha \leq \xi < \beta_n\}$ и события $\{\alpha \leq \xi < \beta_n\}$ происходят одновременно, если только $\xi = \alpha$. Итак, зная функцию распределения, можем найти вероятность того, что величина принимает заданное значение. Используя аксиому непрерывности, убеждаемся, что

$$\lim_{n \rightarrow \infty} P\{\alpha_n \leq \xi < \beta\} = \lim_{n \rightarrow \infty} [F(\beta) - F(\alpha_n)] = 0$$

при $\alpha_n \rightarrow \beta$, $\alpha_n < \beta$, так как одновременно события $\{\alpha_n \leq \xi < \beta\}$ для всех n невозможны, а вероятность невозможного события равна нулю. Если величина ξ имеет непрерывное распределение, то функция распределения непрерывна, так как

$$F(x) = \lim_{\substack{h \rightarrow 0 \\ h > 0}} F(x+h),$$

$$P\{\xi = x\} = \lim_{\substack{h \rightarrow 0 \\ h > 0}} [F(x+h) - F(x)] = 0.$$

Если же функция $F(x)$ имеет разрыв (скачок) в точке x_0 , то величина этого скачка равна вероятности того, что величина ξ примет значение x_0 .

Укажем некоторые общие свойства функции распределения.

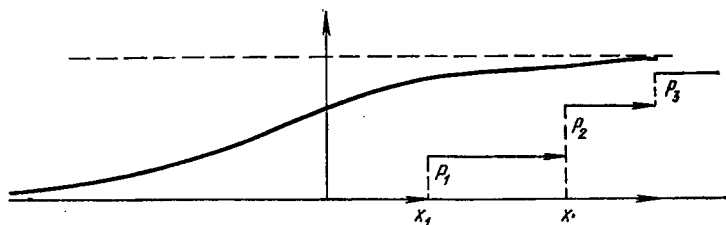
I. Она не убывает: при $x < y$ будет $F(x) \leq F(y)$, так как $F(y) - F(x) = P\{x \leq \xi < y\} \geq 0$.

II. При $x \rightarrow +\infty$ $F(x) \rightarrow 1$. Действительно, при $n \rightarrow \infty$ события $\{\xi < n\}$ влекут события $\{\xi < n+1\}$ и хотя бы одно из этих событий происходит. Поэтому на основании

$$\lim_{n \rightarrow \infty} P \{ \xi < n \} = 1.$$

III. Аналогичные соображения показывают, что $\lim_{x \rightarrow -\infty} F(x) = 0$.

(Примерные графики функций распределения изображены на рисунке.)



Случайная величина называется *дискретной*, если ее возможные значения образуют некоторую последовательность x_1, x_2, \dots (конечную или бесконечную).

Пусть p_k вероятность того, что величина принимает значения x_k . Тогда функция распределения величины так выражается через p_k и x_k : чтобы найти значение $F(x)$, нужно сложить все те значения p_k , для которых $x_k < x$.

Рассмотрим примеры дискретных случайных величин и их функций распределения. Будем пользоваться следующим сокращенным обозначением суммы: если a_1, a_2, \dots — некоторая последовательность чисел, то $\sum_{k < x} a_k$ — сумма $a_1 + \dots + a_m$, где m наибольшее число, для которого $m < x$.

1. *Биномиальное распределение*. Случайная величина принимает значения $0, 1, \dots, n$, при этом

$$P \{ \xi = k \} = C_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n,$$

$0 < p < 1$ — некоторое число. Таким образом, биномиальное распределение зависит от двух параметров: n — целого натурального числа и $p \in (0, 1)$. Функция распределения случайной величины определяется следующим образом: $F(x) = 0$ при $x \leq 0$, $F(x) = 1$ при $x > n$, а при $x \in [0, n]$

$$F(x) = \sum_{k < x} C_n^k p^k (1-p)^{n-k}.$$

Биномиальное распределение имеет число появлений некоторого события в n экспериментах, если вероятность появления его в одном эксперименте равна p .

2. *Геометрическое распределение.* Пусть событие A имеет вероятность p , $0 < p < 1$. Будем повторять эксперимент до тех пор, пока A не произойдет впервые. Число повторений эксперимента v до первого появления события A — случайная величина. Если $v = k$ ($k = 1, 2, \dots$), то в $k-1$ -м эксперименте произошло событие \bar{A} с вероятностью $1 - p$, а в k -м эксперименте — событие A . Значит,

$$P\{v = k\} = p(1 - p)^{k-1}.$$

Случайная величина, принимающая всевозможные натуральные значения с указанными вероятностями, имеет геометрическое распределение (вероятности образуют геометрическую прогрессию). Функция распределения в этом случае имеет вид

$$F(x) = \begin{cases} 0, & x \leq 1; \\ \sum_{k < x} p(1 - p)^k, & x > 1. \end{cases}$$

3. *Равномерное дискретное распределение.* Величина ξ принимает значения $0, 1, \dots, n$ с вероятностями, равными $\frac{1}{n+1}$. Функция распределения величины определяется равенствами

$$F(x) = \begin{cases} 0, & x \leq 0; \\ \frac{k+1}{n+1}, & k < x \leq k+1, \quad k = 0, 1, \dots, n-1; \\ 1, & x > n. \end{cases}$$

Наиболее характерным примером равномерно распределенной дискретной случайной величины служат десятичные знаки для случайно выбранной точки из отрезка $[0, 1]$. При этом $n = 9$. Например, первый десятичный знак числа из отрезка $[0, 1]$ равен k ($k = 0, 1, \dots, 9$), если число лежит в отрезке $\left[\frac{k}{10}, \frac{k+1}{10}\right)$, а вероятность точке попасть в этот отрезок равна его длине, т. е. $\frac{1}{10}$. Второй десятичный знак равен k , если число лежит в одном из

десяти отрезков:

$$\left[\frac{k}{100}, \frac{k+1}{100} \right), \left[\frac{10+k}{100}, \frac{10+k+1}{100} \right), \dots$$

$$\dots, \left[\frac{90+k}{100}, \frac{90+k+1}{100} \right).$$

Длина каждого отрезка равна $\frac{1}{100}$, а поскольку их 10, то вероятность того, что точка попадет в один из них, равна $\frac{1}{10}$.

4. *Распределение Пуассона.* Величина ξ , принимающая значения $0, 1, 2, \dots$, имеет распределение Пуассона, если при некотором $a > 0$ выполняются соотношения

$$P \{ \xi = k \} = \frac{a^k}{k!} e^{-a}, \quad k = 0, 1, \dots$$

Как уже отмечалось, распределение Пуассона будет у числа появлений некоторого редкого события в определенный промежуток времени. Функция распределения имеет вид

$$F(x) = \begin{cases} 0, & x \leq 0, \\ \sum_{k < x} \frac{a^k}{k!} e^{-a}, & x > 0. \end{cases}$$

Эта функция распределения (как и вероятности) зависит от одного параметра a , который называется *параметром распределения Пуассона*.

Среди величин, имеющих непрерывные распределения, практически наиболее важными являются *абсолютно непрерывные распределения*. Это такие распределения, для которых вероятность попадания в малый интервал $[x, x + \Delta x]$ может быть представлена приближенно так:

$$P \{ \xi \in [x, x + \Delta x] \} \sim f(x) \Delta x$$

(как и раньше, знак \sim соединяет величины, отношение которых стремится к 1, в нашем случае при $\Delta x \rightarrow 0$).

Функция $f(x)$ называется *плотностью распределения случайной величины* (или дифференциальной функцией распределения). Зная плотность распределения, можем найти вероятность попадания ξ в некоторый интервал $[a, b]$. Для этого разобьем его на малые интервалы точками $a = x_0 < x_1 < \dots < x_n = b$. Вероятность того, что величина попадет в $[x_k, x_{k+1}]$, равна примерно $f(x_k) \Delta x_k$, $\Delta x_k =$

$= x_{k+1} - x_k$. Складывая эти величины и переходя к пределу при $\Delta x_k \rightarrow 0$, получаем вероятность попадания в $[a, b]$:

$$P\{a < \xi < b\} = \lim \sum f(x_k) \Delta x_k = \int_a^b f(x) \alpha x.$$

Через плотность выражается и функция распределения: нужно в предыдущем равенстве перейти к пределу при $a \rightarrow -\infty$,

$$F(b) = \int_{-\infty}^b f(x) \alpha x.$$

Для непрерывной случайной величины плотность аналогична вероятностям принимать отдельные возможные значения для дискретной величины. Формула, выражающая функцию распределения через интеграл от плотности, является аналогом формулы, выражающей функцию распределения через сумму вероятностей значений для дискретной величины.

Рассмотрим примеры непрерывных случайных величин, их функций распределения и плотностей.

5. Равномерное распределение. Пусть случайная величина ξ равномерно распределена на отрезке $[0, 1]$. Это означает, что точка ξ случайно выбирается из этого отрезка, так что для всякого отрезка $[x, x + \Delta x]$, принадлежащего $[a, b]$, вероятность того, что ξ будет выбрана из этого отрезка, равна отношению длины отрезка к длине отрезка $[a, b]$:

$$P\{x < \xi < x + \Delta x\} = \frac{\Delta x}{b - a}, \text{ если } a \leq x < x + \Delta x < b.$$

Если взять точку x вне отрезка $[a, b]$, то при достаточно малых Δx

$$P\{x < \xi < x + \Delta x\} = 0.$$

Поэтому плотность распределения случайной величины равна

$$f(x) = \begin{cases} 0, & x < a \text{ или } x \geq b, \\ \frac{1}{b - a}, & x \in [a, b]. \end{cases}$$

Функция распределения случайной величины имеет вид

$$F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases}$$

6. *Нормальное распределение.* Случайная величина имеет нормальное (или Гауссово) распределение, если плотность распределения задается формулой

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

или

$$P\{a < \xi < b\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

В § 9 установлено следующее: к интегралу предыдущей формулы справа стремится вероятность того, что величина

$\sqrt{\frac{n}{p(1-p)}} [v_n - p]$, где v_n — частота появления некоторого события, имеющего вероятность p , в n независимых испытаниях, будет принадлежать интервалу (a, b) . Таким образом, нормальное распределение является предельным для флуктуации частоты, умноженной на $\sqrt{\frac{n}{p(1-p)}}$. Функция распределения для нормальной величины имеет вид

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

Как мы увидим, нормальное распределение широко распространено в явлениях природы. Это является следствием центральной предельной теоремы, которая будет рассмотрена в одном из последующих параграфов.

В дальнейших примерах анализируются случайные величины, наблюдающиеся в конкретных экспериментах.

7. Пусть эксперимент заключается в том, что из круга радиуса R случайно выбирается точка. Рассмотрим расстояние данной точки от центра круга. Это случайная величина, принимающая значения от 0 до R . Обозначим через ρ ($\rho > 0$).

Если $0 < x < R$, то

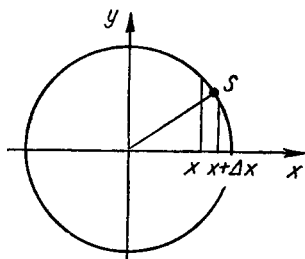
$$P\{\rho < x\} = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}.$$

Вычислим плотность распределения. Имеем

$$\begin{aligned} P\{x < \rho < x + \Delta x\} &= \frac{1}{R^2} [(x + \Delta x)^2 - x^2] = \\ &= \frac{1}{R^2} \Delta x (2x + \Delta x) \sim \frac{2x}{R^2} \Delta x, \end{aligned}$$

если только $0 \leq x < x + \Delta x \leq R$.
Значит, плотность определяется равенством

$$f(x) = \begin{cases} 0, & x \leq 0, \\ \frac{2x}{R^2}, & 0 < x < R, \\ 0, & x \geq R. \end{cases}$$



8. Случайная величина ξ есть координата проекции на ось x точки S , случайно выбранной на окружности радиуса R с центром в начале координат. Найдем ее функцию распределения и плотность распределения. Вероятность того, что ξ лежит в интервале $(x, x + \Delta x)$, равна $\frac{1}{2\pi R}$, умноженной на удвоенную длину дуги верхней полуокружности, проектирующейся на отрезок $[x, x + \Delta x]$. Длина этой дуги равна

$$R \arcsin \frac{x + \Delta x}{R} - R \arcsin \frac{x}{R}.$$

Воспользуемся тем, что для малых дуг $y \sin y \sim y$. Поэтому

$$\begin{aligned} &R \left[\arcsin \frac{x + \Delta x}{R} - \arcsin \frac{x}{R} \right] \sim \\ &\sim R \sin \left[\arcsin \frac{x + \Delta x}{R} - \arcsin \frac{x}{R} \right] = \\ &= R \sin \arcsin \frac{x + \Delta x}{R} \cdot \cos \arcsin \frac{x}{R} - \\ &- R \sin \arcsin \frac{x}{R} \cdot \cos \arcsin \frac{x + \Delta x}{R}. \end{aligned}$$

Поскольку $\sin \arcsin y = y$, $\cos \arcsin y = \sqrt{1 - y^2}$,

$$\begin{aligned} & \arcsin \frac{x + \Delta x}{R} - \arcsin \frac{x}{R} \sim \\ & \sim \frac{x + \Delta x}{R} \sqrt{1 - \frac{x^2}{R^2}} - \frac{x}{R} \sqrt{1 - \frac{(x + \Delta x)^2}{R^2}} = \\ & = \frac{\Delta x}{R} \sqrt{1 - \frac{x^2}{R^2}} + \frac{x}{R} \left(\sqrt{1 - \frac{x^2}{R^2}} - \sqrt{1 - \frac{(x + \Delta x)^2}{R^2}} \right) = \\ & = \frac{\Delta x}{R} \sqrt{1 - \frac{x^2}{R^2}} + \frac{x}{R} \times \\ & \quad \times \frac{\frac{(x + \Delta x)^2}{R^2} - \frac{x^2}{R^2}}{\sqrt{1 - \frac{x^2}{R^2}} + \sqrt{1 - \frac{(x + \Delta x)^2}{R^2}}} \sim \\ & \sim \frac{\Delta x}{R} \left(\sqrt{1 - \frac{x^2}{R^2}} + \frac{\frac{x^2}{R^2}}{\sqrt{1 - \frac{x^2}{R^2}}} \right) = \frac{\Delta x}{\sqrt{R^2 - x^2}}. \end{aligned}$$

Таким образом, плотность распределения имеет вид

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

Функцию распределения определить значительно проще. При $-R < x < R$

$$\begin{aligned} P\{-R < \xi < x\} &= \frac{1}{\pi} \left[\arcsin \frac{x}{R} - \arcsin(-1) \right] = \\ &= \frac{1}{\pi} \arcsin \frac{x}{R} + \frac{1}{2}. \end{aligned}$$

Следовательно,

$$F(x) = \begin{cases} 0, & x \leq -R, \\ \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{R}, & -R \leq x \leq R, \\ 1, & x > R. \end{cases}$$

§ 15. Среднее значение

Как охарактеризовать случайную величину одним числом? Пусть, например, случайная величина есть какой-либо параметр для представителей какого-нибудь биологич

ского вида (вес горохового зерна, рост взрослого человека, длина тела ящерицы, ширина размаха крыльев у вороны и т. п.). Хотя от наблюдения к наблюдению значение случайной величины (параметра) меняется, но существует некоторое «среднее» ее значение: средний рост человека, средний вес зерен гороха и т. д.

В этом параграфе мы займемся определением среднего значения случайной величины.

Предположим сначала, что случайная величина ξ принимает конечное число различных значений x_1, x_2, \dots, x_m с вероятностями p_1, p_2, \dots, p_m соответственно. Повторим независимо n раз эксперимент, в котором наблюдается величина ξ . В каждом эксперименте отметим значение случайной величины. (Например, желая установить длину ящериц определенного вида, будем отлавливать этих ящериц в различных местах расселения вида и измерять их длину.) Обозначим значения величины через y_1, \dots, y_n (среди них могут быть и повторяющиеся). Среднее арифметическое значение из наблюдений $\frac{1}{n} (y_1 + \dots + y_n)$ также величина случайная, она зависит от изменения как серии наблюдений, так и n .

Однако, как и частота появления некоторого события в n экспериментах, среднее значение имеет тенденцию определенной устойчивости: при возрастании n стремится к некоторому неслучайному числу. Это неслучайное предельное значение арифметических средних и называется *средним значением случайной величины*, или ее *математическим ожиданием*.

Свойство устойчивости арифметических средних вытекает из свойства устойчивости частоты события. Обозначим через $A_i, i = 1, \dots, m$ событие, заключающееся в том, что случайная величина приняла значение x_i . Если событие A_i в n экспериментах произошло $k_n(A_i)$ раз, то

$$y_1 + \dots + y_n = k_n(A_1) x_1 + \dots + k_n(A_m) x_m$$

(среди y_1, \dots, y_n значение x_1 встречается столько раз, сколько раз произошло событие A_1 ; тоже можно сказать и относительно других возможных значений случайной величины). Поэтому

$$\begin{aligned} \frac{1}{n} (y_1 + \dots + y_n) &= x_1 \frac{k_n(A_1)}{n} + \dots + x_m \frac{k_n(A_m)}{n} = \\ &= x_1 v_n(A_1) + \dots + x_m v_n(A_m), \end{aligned}$$

здесь $\nu_n(A_i)$ — частота события A_i в n экспериментах. Заменяя частоты предельными значениями — вероятностями событий, получим среднее значение, или математическое ожидание, случайной величины (оно обозначается символом $M\xi$).

Средним значением случайной величины ξ , принимающей значения x_1, \dots, x_m с вероятностями p_1, \dots, p_m , называется число

$$M\xi = x_1 p_1 + \dots + x_m p_m.$$

Рассмотрим примеры на вычисление среднего значения.

1. Пусть ξ принимает значения $0, 1, \dots, m$ с вероятностями $\frac{1}{m+1}$. Тогда

$$\begin{aligned} M\xi &= 0 \cdot \frac{1}{m+1} + 1 \cdot \frac{1}{m+1} + \dots + \frac{m}{m+1} = \\ &= \frac{1}{m+1} \cdot \frac{m(m+1)}{2} = \frac{m}{2}. \end{aligned}$$

2. Случайная величина ξ имеет биномиальное распределение:

$$P\{\xi = k\} = C_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, m.$$

Тогда

$$\begin{aligned} M\xi &= 0 \cdot (1-p)^n + 1 \cdot np(1-p)^{n-1} + \dots + \\ &+ k \frac{n(n-1)\dots(n-k+1)}{1 \dots k} p^k (1-p)^{n-k} + \dots + np^n = \\ &= np \left[(1-p)^{n-1} + \dots + \frac{(n-1)\dots(n-k+1)}{1 \dots (k-1)} p^{k-1} \times \right. \\ &\left. \times (1-p)^{n-k} + \dots + p^{n-1} \right] = np [(1-p) + p]^{n-1} = np. \end{aligned}$$

Для произвольных дискретных случайных величин среднее значение определяется аналогично: это сумма произведений $x_k p_k$, где x_k — возможные значения случайной величины, а p_k — вероятность того, что величина ξ примет данное значение. Символически это записывается так:

$$M\xi = \sum x_k p_k.$$

Если множество возможных значений бесконечно, то под бесконечной суммой понимают предел суммы

$$x_1 p_1 + \dots + x_n p_n,$$

когда $n \rightarrow \infty$.

Рассмотрим примеры на вычисление математического ожидания для величин, принимающих бесконечное множество значений.

3. Пусть величина ξ имеет геометрическое распределение: $P\{\xi = n\} = (1-p)p^{n-1}$, $n = 1, 2, \dots$. Тогда

$$\begin{aligned} M\xi &= 1 \cdot (1-p) + 2(1-p)p + \dots + n(1-p)p^{n-1} + \dots \\ &\dots = (1-p)(1 + 2p + 3p^2 + \dots + np^{n-1} + \dots) = \\ &= (1-p)[1 + p + p^2 + \dots] + (1-p)[p + p^2 + p^3 + \dots] + \\ &\quad + (1-p)[p^2 + p^3 + \dots] + \dots + \\ &\quad + (1-p)[p^n + p^{n+1} + \dots] + \dots = (1-p) \times \\ &\quad \times \left[\frac{1}{1-p} + \frac{p}{1-p} + \frac{p^2}{1-p} + \dots \right] = \frac{1}{1-p}. \end{aligned}$$

4. Величина ξ имеет распределение Пуассона с параметром a , т. е.

$$P\{\xi = k\} = \frac{a^k}{k!} e^{-a}, \quad k = 0, 1, \dots$$

Тогда

$$\begin{aligned} M\xi &= 0 \cdot e^{-a} + 1 \cdot \frac{a}{1!} e^{-a} + \dots + k \frac{a^k}{k!} e^{-a} + \dots = \\ &= a \left[e^{-a} + \dots + \frac{a^{k-1}}{(k-1)!} e^{-a} + \dots \right] = a, \end{aligned}$$

так как выражение в скобках есть сумма вероятностей всех значений для величины, имеющей распределение Пуассона, а значит, равна 1. Мы установили вероятностный смысл параметра a для распределения Пуассона; a есть среднее значение случайной величины с этим распределением. Отметим, что распределение Пуассона возникло как предельное для биномиального при $n \rightarrow \infty$, $np \rightarrow a$, а np есть среднее значение для биномиального распределения, поэтому то, что $M\xi = a$, для распределения Пуассона вполне естественно.

Для величин, принимающих бесконечное множество значений, может оказаться, что среднее значение равно бесконечности. Это кажется парадоксальным: все значения случайной величины конечны, а среднее значение — бесконечно. Однако это не парадокс, а особенность бесконечного множества, одно из отличий его свойств от свойств конечных множеств.

Рассмотрим пример, иллюстрирующий указанную возможность.

5. Случайная величина ξ принимает значения $1, 2, 4, \dots, 2^n, \dots$ с вероятностями $\frac{1}{2}, \frac{1}{4}, \dots, \frac{1}{2^{n+1}}, \dots$ соответственно. Пусть, например, имеется большое число пробирок в которых посеяна некоторая бактериальная культура. Вероятность того, что она окажется жизнеспособной для каждой пробирки, равна $\frac{1}{2}$. За единицу времени, необходимую для исследования содержимого пробирки, количество бактерий в пробирке, где они жизнеспособны, удваивается. Пробирки выбираются по одной и исследуются. Количество бактерий в первой пробирке с жизнеспособными бактериями (за единицу измерения принимается количество бактерий в начальный момент времени) будет случайной величиной. Поскольку вероятность извлечь пробирку с жизнеспособными бактериями равна $\frac{1}{2}$, то вероятность того, что такая пробирка впервые будет исследована при k -м испытании, равняется $\frac{1}{2^k}$. Тогда число бактерий в ней за время $k - 1$, удваиваясь на каждом шаге, будет 2^{k-1} . Для вычисления среднего значения этой величины нужно найти бесконечную сумму

$$\begin{aligned} & \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \dots + \frac{1}{2^k} \cdot 2^{k-1} + \dots = \\ & = \lim_{n \rightarrow \infty} \left(\frac{1}{2} + \frac{2}{4} + \dots + \frac{2^n}{2^{n+1}} \right) = \lim_{n \rightarrow \infty} \frac{n}{2} = \infty. \end{aligned}$$

Рассмотрим теперь определение среднего значения для произвольной случайной величины. Пусть ξ такая величина и $F(x)$ — ее функция распределения. Построим дискретную случайную величину, которая мало отличается от величины ξ . Это можно сделать следующим образом. Возьмем $h > 0$ и будем измерять ξ с точностью до h , полагая $\xi_h = kh$, если $kh \leq \xi < kh + h$. Величина ξ_h отличается от ξ не больше чем на h и $\xi_h \rightarrow \xi$ при $h \rightarrow 0$. Кроме того, величина ξ_h имеет дискретное распределение, она принимает значения лишь вида kh , причем

$$P\{\xi_h = kh\} = F((k+1)h) - F(kh).$$

Следовательно,

$$M\xi_h = \sum kh [F((k+1)h) - F(kh)].$$

Поскольку $\xi_h \rightarrow \xi$ при $h \rightarrow 0$, то естественно предел последнего выражения при $h \rightarrow 0$ назвать *математическим ожиданием* случайной величины:

$$M\xi = \lim_{h \rightarrow 0} h \sum k [F((k+1)h) - F(kh)]$$

(математическое ожидание существует, если существуют $M\xi_h$ и последний предел).

Вычислим математическое ожидание, используя данное выше определение, для непрерывных случайных величин.

6. Пусть величина ξ равномерно распределена на отрезке $[a, b]$. Если $a \leq kh < kh + h < b$, то $F(kh + h) - F(kh) = \frac{h}{b-a}$, при $k_1h \leq a < k_1h + h$, $F(k_1h + h) - F(k_1h) = \frac{k_1h + h - a}{b-a}$, при $k_2h \leq b < k_2h + h$, $F(k_2h + h) - F(k_2h) = \frac{b - k_2h}{b-a}$ (мы использовали выражение для функции распределения равномерно определенной величины, приведенное в § 14). Таким образом,

$$\begin{aligned} M\xi_h &= k_1h \frac{k_1h + h - a}{b-a} + k_2h \frac{b - k_2h}{b-a} + \frac{h^2}{b-a} (k_1 + 1 + \\ &+ k_1 + 2 + \dots + (k_2 - 1)) = k_1h \frac{k_1h + h - a}{b-a} + \\ &+ k_2h \frac{b - k_2h}{b-a} + \frac{h^2}{b-a} \cdot \frac{(k_1 + k_2)(k_2 - k_1 - 1)}{2}. \end{aligned}$$

Из определения k_1 и k_2 вытекает, что $k_1h \rightarrow a$, $k_2h \rightarrow b$ при $h \rightarrow 0$. Поэтому

$$\begin{aligned} h(k_1 + k_2) &\rightarrow a + b, \quad h(k_2 - k_1 - 1) \rightarrow b - a, \\ k_1h + h - a &\rightarrow 0, \quad b - k_2h \rightarrow 0. \end{aligned}$$

Следовательно,

$$M\xi = \lim_{h \rightarrow 0} M\xi_h = \frac{b^2 - a^2}{2(-a)} = \frac{a+b}{2}.$$

Для равномерно распределенной случайной величины среднее значение совпадает с серединой отрезка, в котором величина ξ изменяется.

7. Найдем среднее значение времени ожидания в задаче о встрече. Эта случайная величина η рассматривалась в примере 5 § 13. Для нее была получена такая формула:

$$P\{\alpha \leq \eta < \beta\} = \frac{1}{a^2} (\beta - \alpha) [2a - \beta - \alpha], \quad 0 < \alpha < \beta < a.$$

Поэтому, беря $h = \frac{a}{n}$, будем иметь при $0 \leq k < n$

$$\begin{aligned} F(kh + h) - F(kh) &= P\{kh \leq \eta < kh + h\} = \\ &= \frac{h}{a^2} [2a - 2kh - h]. \end{aligned}$$

Если же k не удовлетворяет приведенному неравенству то $F(kh + h) - F(kh) = 0$. Значит,

$$M\eta_n = \frac{h^2}{a^2} \sum k [2a - 2kh - h],$$

где k меняется от 0 до n . Но

$$\begin{aligned} 1 + 2 + \dots + n &= \frac{n(n+1)}{2}, \quad 1 + 2^2 + \dots + n^2 = \\ &= \frac{n(n+1)(2n+1)}{6}. \end{aligned}$$

Поэтому

$$M\eta_n = \frac{h^2}{a^2} \left[(2a - h) \frac{n(n+1)}{2} - 2h \frac{n(n+1)(2n+1)}{6} \right].$$

Поскольку $nh = a$,

$$M\eta_n = \frac{2a - h}{a^2} \cdot \frac{a^2 + ha}{2} - \frac{2a^3 + 3a^2h + ah^2}{3a^2}$$

и

$$M\eta = \lim_{h \rightarrow 0} M\eta_n = a - \frac{2}{3}a = \frac{a}{3}.$$

8. Пусть ξ имеет нормальное распределение, т. е. плотность распределения равна $\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Тогда

$$F(kh + h) - F(kh) \sim h \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{k^2h^2}{2}}.$$

Поскольку

$$\sum_k k e^{-\frac{k^2h^2}{2}} = \sum_{k=1}^{\infty} k \left(e^{-\frac{k^2h^2}{2}} - e^{-\frac{(k+1)^2h^2}{2}} \right) = 0,$$

то $M_{\xi_h}^2 = 0$ для всех $h > 0$. Значит, $M_{\xi}^2 = 0$.

Приведем два важных свойства математического ожидания, они будут использоваться в дальнейшем:

I. Пусть ξ — некоторая случайная величина, c — постоянная. Тогда $c\xi$ также случайная величина. При этом $M(c\xi) = c M\xi$. Это означает, что при умножении случайной величины на некоторое число на это же число умножается и среднее значение случайной величины.

Проверим это свойство для величин дискретных. Пусть ξ принимает значения x_1, x_2, \dots с вероятностями p_1, p_2, \dots . Величина $c\xi$ принимает значение cx_k , если $\xi = x_k$, поэтому всевозможные значения величины есть последовательность cx_1, cx_2, \dots , при этом величина $c\xi$ принимает значение cx_k с вероятностью p_k . Значит,

$$M(c\xi) = \sum cx_k p_k = c \sum x_k p_k = cM\xi.$$

II. Пусть ξ и η — две случайные величины, которые наблюдаются в одном эксперименте. Тогда $\xi + \eta$ также случайная величина, наблюдаемая в этом эксперименте. Справедливо равенство

$$M(\xi + \eta) = M\xi + M\eta,$$

т. е. среднее значение суммы случайных величин равно сумме средних значений этих величин. (Это справедливо, если средние значения существуют.)

Данное свойство очевидным образом распространяется на сумму нескольких случайных величин.

Остановимся лишь на случае, когда величины принимают конечное число значений. Пусть ξ принимает значения x_1, \dots, x_m , η — значения y_1, \dots, y_l . Если одновременно происходят события $\{\xi = x_k\}$ и $\{\eta = y_j\}$, то $\xi + \eta = x_k + y_j$. Обозначим через $P\{\xi = x_k, \eta = y_j\}$ — вероятность события $\{\xi = x_k\} \cdot \{\eta = y_j\}$. Тогда по определению

$$\begin{aligned} M(\xi + \eta) &= \sum (x_k + y_j) P\{\xi = x_k, \eta = y_j\} = \\ &= \sum x_k P\{\xi = x_k, \eta = y_j\} + \sum y_j P\{\xi = x_k, \eta = y_j\}. \end{aligned} \quad (1)$$

Заметим, что если сложить события $\{\xi = x_k\} \times \{\eta = y_j\}$ по j , то получим событие $\{\xi = x_k\}$, значит,

$$\sum_j P\{\xi = x_k, \eta = y_j\} = P\{\xi = x_k\}$$

(слева отмечено, что вероятность складывается по всем j).

Аналогично

$$\sum_k P \{ \xi = x_k, \eta = y_j \} = P \{ \eta = y_j \}.$$

Поэтому, производя в сумме сначала сложение по j , а затем по k , получаем

$$\sum x_k P \{ \xi = x_k, \eta = y_j \} = \sum x_k P \{ \xi = x_k \} = M\xi.$$

Точно так

$$\sum y_j P \{ \xi = x_k, \eta = y_j \} = \sum y_j P \{ \eta = y_j \} = M\eta.$$

Значит, правая часть формулы (1) есть $M\xi + M\eta$, что и требовалось доказать.

Это свойство иногда облегчает вычисление математического ожидания. Рассмотрим пример.

9. Вычислим математическое ожидание величины ν , рассмотренной в примере 1 § 13. ν — число дефектных изделий в выборке объема m из партии в N изделий, содержащей n дефектных. Вычисления, основанные на вычисленных в § 13 вероятностях, очень сложны. Пусть m изделий выбираются по одному, ν_k — число дефектных изделий, выбранных в k -й раз ($\nu_k = 1$, если выбрано дефектное изделие, $\nu_k = 0$, если выбрано стандартное изделие). Тогда

$$\nu = \nu_1 + \dots + \nu_m, \quad M\nu = M\nu_1 + \dots + M\nu_m.$$

Но $M\nu_k = 1 \cdot P \{ \nu_k = 1 \}$. Из соображений симметрии вытекает, что вероятность вынуть дефектное изделие на каждом шаге одна и та же и равна $\frac{n}{N}$ (вероятность того, что дефектное изделие выбрано на первом шаге). Значит,

$$M\nu_1 = M\nu_2 = \dots = M\nu_m = \frac{n}{N}, \quad M\nu = \frac{mn}{N}.$$

§ 16. Дисперсия случайной величины

Среднее значение случайной величины — это такое значение, возле которого группируются ее наблюдаемые значения. Однако они могут группироваться по-разному: более тесно и, наоборот, более широко.

Проиллюстрируем это на таком примере. Пусть наблюдаемые значения первой величины: 9, 7, 10, 13, 11, 8, 9, 10, 8, 12. Среднее арифметическое наблюдений равно 9,7. Отклонения наблюдаемых значений от среднего арифметического изменяются от $-2,7$ до $3,3$. В такой цепочке на-

блюдений: 11, 9, 8, 15, 9, 6, 11, 8, 10, 10 — среднее арифметическое то же, что и раньше, — 9,7. Но теперь наблюдения более широко расположены около среднего значения, отклонения от среднего меняются от —3,7 до 5,3.

Как охарактеризовать «кучность» наблюдений возле среднего значения? Для этого используется величина, которая называется *средним квадратическим отклонением*.

Если x_1, \dots, x_n — наблюдения, $\bar{x} = \frac{1}{n} (x_1 + \dots + x_n)$ — среднее арифметическое, то квадрат среднего квадратического отклонения s^2 определяется как среднее арифметическое квадратов отклонений наблюдений от среднего арифметического:

$$s^2 = \frac{1}{n} [(x_1 - \bar{x})^2 + \dots + (x_n - \bar{x})^2].$$

Для первой серии наблюдений $s^2 = 3,210$, так что $s \approx 1,8$. Для второй серии наблюдений $s^2 = 5,21$, $s \approx 2,3$. Во втором случае среднее квадратическое отклонение больше.

Пусть x_1, \dots, x_n — наблюдения некоторой величины ξ . Тогда при возрастании n к бесконечности величина \bar{x} приближается к $M\xi$, так что

$$s^2 \sim \frac{1}{n} [(x_1 - M\xi)^2 + \dots + (x_n - M\xi)^2].$$

Очевидно, что $(x_k - M\xi)^2$ будут значениями случайной величины $(\xi - M\xi)^2$, которую можно наблюдать в том же эксперименте, что и величину ξ . Среднее арифметическое из значений этой случайной величины при $n \rightarrow \infty$ стремится к ее математическому ожиданию, т. е. к числу

$$D\xi = M(\xi - M\xi)^2,$$

которое называется *дисперсией случайной величины ξ* . Она является мерой рассеяния значений случайной величины вокруг среднего значения. Поскольку $(\xi - M\xi)^2$ имеет только неотрицательные значения, то и $M(\xi - M\xi)^2 \geq 0$.

При вычислении дисперсии удобно использовать несколько преобразованную формулу. Так как

$$(\xi - M\xi)^2 = \xi^2 - 2\xi M\xi + (M\xi)^2,$$

то на основании свойств математического ожидания

$$M(\xi - M\xi)^2 = M\xi^2 - M[(2M\xi) \cdot \xi] + M(M\xi)^2.$$

Но, очевидно, что математическое ожидание постоянной величины (она с вероятностью 1 принимает свое единст-

венное значение) равно самой этой величине. Кроме того $2M\xi$ — величина постоянная и может быть вынесена за знак математического ожидания. Поэтому

$$M(\xi - M\xi)^2 = M\xi^2 - 2M\xi \cdot M\xi + (M\xi)^2.$$

Окончательно получаем следующее выражение:

$$D\xi = M\xi^2 - (M\xi)^2.$$

Таким образом, дисперсия случайной величины равна среднему значению квадрата величины минус квадрат среднего значения. Величина $\sigma = \sqrt{D\xi}$ называется *стандартным отклонением*, оно имеет ту же размерность, что и сама величина.

Если ξ является дискретной величиной и принимает значения x_1, x_2, \dots с вероятностями p_1, p_2, \dots , то ξ^2 будет принимать значения x_1^2, x_2^2, \dots с теми же вероятностями. Поэтому

$$M\xi^2 = \sum x_k^2 p_k$$

и

$$D\xi = \sum x_k^2 p_k - \left(\sum x_k p_k\right)^2.$$

Для непрерывных распределений можно использовать аппроксимации величин с помощью дискретных, так как это было сделано при вычислении математических ожиданий в предыдущем параграфе. Таким образом, устанавливается формула

$$D\xi = \lim_{h \rightarrow 0} \sum k^2 h^2 (F(kh + h) - F(k)) - (M\xi)^2.$$

Рассмотрим примеры на вычисление дисперсий случайных величин.

1. Величина ξ принимает значение 1 с вероятностью p , и значение 0 с вероятностью $1 - p$. Тогда $\xi^2 = \xi$, $M\xi^2 =$

и

$$D\xi = M\xi^2 - (M\xi)^2 = p - p^2 = p(1 - p).$$

2. Величина ξ имеет биномиальное распределение параметрами n и p , т. е. $P\{\xi = k\} = C_n^k p^k (1 - p)^{n-k}$. Как вычислено в § 15, $M\xi = np$. Имеем

$$M\xi^2 = 1 \cdot np(1 - p)^{n-1} + 4 \frac{n(n-1)}{2} p^2 (1 - p)^{n-2} + \dots + k^2 \frac{n(n-1) \dots (n-k+1)}{1 \cdot 2 \dots k} p^k (1 - p)^{n-k} + \dots + n^2 p^n,$$

k -е слагаемое в этой сумме можно преобразовать так:

$$k(k-1) \frac{n(n-1) \dots (n-k+1)}{1 \cdot 2 \dots k} p^k (1-p)^{n-k} +$$

$$+ k \frac{n(n-1) \dots (n-k+1)}{1 \cdot 2 \dots k} p^k (1-p)^{n-k} =$$

$$= n(n-1) p^2 C_{n-2}^{k-2} p^{k-2} (1-p)^{n-k} + np C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k}.$$

Поэтому

$$M\xi^2 = n(n-1) p^2 [(1-p)^{n-2} + (n-2) p (1-p)^{n-3} + \dots$$

$$\dots + C_{n-2}^{k-2} p^{k-2} (1-p)^{n-k} + \dots + p^{n-2}] +$$

$$+ np [(1-p)^{n-1} + (n-1) p (1-p)^{n-2} + \dots +$$

$$+ C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} + \dots + p^{n-1}] = n(n-1) p^2 + np,$$

так как суммы в скобках являются суммами биномиальных вероятностей и, значит, равны 1. Окончательно получаем

$$D\xi = n(n-1) p^2 + np - (np)^2 = np(1-p).$$

3. Пусть величина ξ имеет равномерное дискретное распределение, т. е. ξ принимает значения $0, 1, \dots, m$ с вероятностями $\frac{1}{m+1}$. Тогда

$$M\xi^2 = \frac{1}{m+1} (1 + \dots + m^2) =$$

$$= \frac{1}{m+1} \cdot \frac{m(m+1)(2m+1)}{6}.$$

Как было вычислено в § 15, $M\xi = \frac{m}{2}$, Значит,

$$D\xi = \frac{m^2}{3} + \frac{m}{6} - \frac{m^2}{4} = \frac{m^2}{12} + \frac{m}{6}.$$

4. Величина ξ имеет распределение Пуассона с параметром a . Пусть

$$M\xi^2 = \sum k^2 \frac{a^k}{k!} e^{-a} = \sum k \frac{a^k}{(k-1)!} e^{-a} =$$

$$= \left[a + 2a^2 + \dots + \frac{ka^k}{(k-1)!} + \dots \right] e^{-a} =$$

$$= a \left[e^{-a} + ae^{-a} + \dots + \frac{a^{k-1}}{(k-1)!} e^{-a} + \dots \right] +$$

$$+ a^2 \left[e^{-a} + ae^{-a} + \dots + \frac{a^{k-2}}{(k-2)!} e^{-a} + \dots \right] = a + a^2$$

(в скобках стоит сумма вероятностей для распределения Пуассона, значит, она равна 1). Таким образом,

$$D\xi = a + a^2 - (M\xi)^2 = a + a^2 - a^2 = a.$$

Поэтому для пуассоновского распределения дисперсия совпадает с математическим ожиданием.

Отметим два простых свойства дисперсии.

I. Если c постоянная, то

$$D(\xi + c) = D\xi.$$

Действительно, $M(\xi + c)^2 = M\xi^2 + 2cM\xi + c^2$, $(M(\xi + c))^2 = (M\xi)^2 + 2cM\xi + c^2$. Вычитая эти два выражения, получаем требуемое.

II. Если c постоянная, то

$$D(c\xi) = c^2 D\xi.$$

Действительно, $M(c\xi)^2 = c^2 M\xi^2$, $(M(c\xi))^2 = (cM\xi)^2 = c^2 (M\xi)^2$. (мы использовали свойство 1 математического ожидания из § 15), так что

$$D(c\xi) = c^2 M\xi^2 - c^2 (M\xi)^2 = c^2 D\xi.$$

Рассмотрим применение этих свойств к вычислению дисперсий непрерывных величин.

5. ξ имеет равномерное распределение на отрезке $[a, b]$. Тогда $\xi - a$ имеет равномерное распределение на $[0, b - a]$. Пусть $h = \frac{b-a}{m+1}$, $\xi_h = kh$ при $[kh, kh + h)$, $k = 0, 1, \dots, m$. Тогда $\xi_h = hv$, где v имеет такое же распределение, как величина в примере 3. Значит,

$$D\xi_h = h^2 \left(\frac{m^2}{12} + \frac{m}{6} \right) = (b-a)^2 \left(\frac{m^2}{(m+1)^2} \cdot \frac{1}{12} + \frac{m}{(m+1)^2} \cdot \frac{1}{6} \right).$$

Переходя к пределу при $h \rightarrow 0$, т. е. $m \rightarrow \infty$, получаем

$$D\xi = \frac{1}{12} (b-a)^2.$$

6. Рассмотрим величину ξ , имеющую нормальное распределение. Для вычисления дисперсии этой величины обратимся к величине η с биномиальным распределением. Как вытекает из результатов § 9 и задачи 6 § 14, величина

$$\frac{\eta - np}{\sqrt{np(1-p)}} = (v_n - p) \sqrt{\frac{n}{p(1-p)}}$$

имеет в пределе при $n \rightarrow \infty$ нормальное распределение. Поэтому, учитывая результаты задачи 2 и свойства I и II, приходим к формуле

$$D\xi = \lim_{n \rightarrow \infty} D\left(\frac{\eta - np}{\sqrt{np(1-p)}}\right) = \lim_{n \rightarrow \infty} \frac{1}{np(1-p)} D\eta = 1.$$

§ 17. Независимые случайные величины

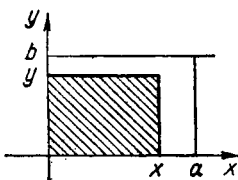
Если две случайные величины ξ и η наблюдаются в разных независимых экспериментах, то события $\{\xi < x\}$ и $\{\eta < y\}$ независимы, каковы бы ни были x и y . Поэтому вероятность одновременного осуществления данных событий равна произведению их вероятностей. Это можно записать следующим образом:

$$P\{\xi < x, \eta < y\} = P\{\xi < x\} P\{\eta < y\}. \quad (1)$$

Величины ξ и η называются независимыми, если для них выполнено равенство (1) для всех x и y . При этом величины ξ и η могут наблюдаться и в одном и том же эксперименте.

Рассмотрим примеры независимых случайных величин.

1. Пусть из прямоугольника, одна из вершин которого расположена в начале координат, а две стороны с размерами a и b на осях координат, выбирается случайно точка. Обозначим координаты этой точки через ξ и η . Если x и y — числа, для которых $0 \leq x \leq a$ и $0 \leq y \leq b$, то событие $\{\xi < x\}$ и $\{\eta < y\}$ происходит, если точка выбрана из прямоугольника с вершиной в начале координат и сторонами x и y . Вероятность этого события равна площади прямоугольника xy , деленной на площадь большого прямоугольника ab (так определяется вероятность при случайном выборе точки из прямоугольника). Значит,



$$P\{\xi < x, \eta < y\} = \frac{xy}{ab}. \quad (2)$$

Так как $\eta < b$, какую бы точку из прямоугольника мы не выбрали, то событие $\{\eta < b\}$ достоверно. Поэтому событие $\{\xi < x\} \cdot \{\eta < b\}$ совпадает с событием $\{\xi < x\}$ и $P\{\xi < x, \eta < b\} = P\{\xi < x\}$. Полагая в (2) $y = b$, будем иметь

$$P\{\xi < x\} = \frac{x}{a}.$$

Аналогично при $x = a$ получим из (2), что $P \{ \eta < y \} = \frac{y}{b}$. Так что для величин ξ и η выполнено равенство (1), если x и y таковы, как было указано выше.

Пусть теперь $x \leq 0$. Тогда событие $\{ \xi < x \}$ невозможно, и в обеих частях равенства (1) получаем 0, оно выполнено и в этом случае. Если $x \geq b$, то событие $\{ \xi < x \}$ достоверно, $P \{ \xi < x \} = 1$ и $\{ \xi < x \} \cdot \{ \eta < y \} = \{ \eta < y \}$. Значит,

$$P \{ \xi < x, \eta < y \} = P \{ \eta < y \} = 1 \cdot P \{ \eta < y \} = \\ = P \{ \xi < x \} \cdot P \{ \eta < y \},$$

т. е. равенство (1) выполнено при всех x .

Аналогично убеждаемся, что y тоже можно выбирать произвольно и при этом равенство (1) сохраняется. Таким образом, величины ξ и η , наблюдаемые в нашем эксперименте, независимы.

2. Пусть эксперимент заключается в случайном выборе точки из круга радиуса R с центром в начале координат. Если ξ и η , как и в первом примере, координаты точки, то (см. рис. на с. 107)

$$P \left\{ \xi < -\frac{R}{\sqrt{2}} \right\} = \frac{1}{R^2} \left(\frac{\pi R^2}{4} - \frac{1}{2} R^2 \right) = \frac{\pi - 2}{2}.$$

Точно так

$$P \left\{ \eta < -\frac{R}{\sqrt{2}} \right\} = \frac{\pi - 2}{2}.$$

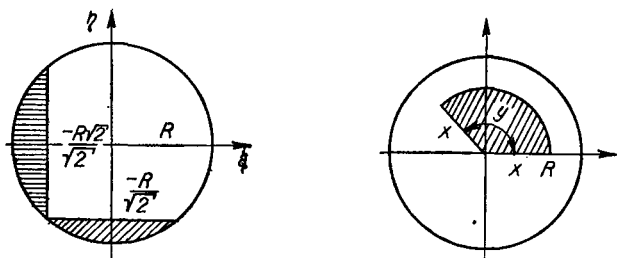
Но одновременно события $\left\{ \xi < -\frac{R}{\sqrt{2}} \right\}$ и $\left\{ \eta < -\frac{R}{\sqrt{2}} \right\}$ произойти не могут, так как те части круга, где для координат точки выполнены указанные неравенства, не пересекаются. Значит,

$$P \left\{ \xi < -\frac{R}{\sqrt{2}}, \eta < -\frac{R}{\sqrt{2}} \right\} = 0 \neq \left(\frac{\pi - 2}{2} \right)^2 = \\ = P \left\{ \xi < -\frac{R}{\sqrt{2}} \right\} P \left\{ \eta < -\frac{R}{\sqrt{2}} \right\}.$$

Величины ξ и η в этом случае не являются независимыми.

Рассмотрим другую пару величин: полярные координаты выбранной точки ρ — расстояние точки от начала координат

нат и φ — угол между положительным направлением оси x и лучом, проведенным из начала координат через точку, угол отсчитывается от оси x против часовой стрелки. ρ и φ — случайные величины, ρ меняется от 0 до R , φ — от 0 до 2π . Пусть $0 \leq x \leq R$, $0 \leq y \leq 2\pi$. Тогда событие $\{\rho < x\} \cdot \{\varphi < y\}$ происходит, если выбранная случайно точка попала в сектор круга радиуса x , вырезаемый углом



величины y , одна сторона которого есть положительная полуось x . Площадь такого сектора равна $\frac{1}{2} x^2 y$. Значит,

$$P \{ \rho < x, \varphi < y \} = \frac{x^2 y}{2\pi R^2} = \frac{x^2}{R^2} \cdot \frac{y}{2\pi}.$$

Полагая поочередно $x = R$ и $\varphi = 2\pi$, находим

$$P \{ \varphi < y \} = \frac{y}{2\pi}, \quad P \{ \rho < x \} = \frac{x^2}{R^2}.$$

Отсюда вытекает независимость величин ρ и φ (при $x < 0$, $y < 0$ или $x > R$, $y > 2\pi$ рассуждаем точно так же, как в предыдущем примере).

Пусть ξ и η — независимые величины, причем каждая из них дискретна. Тогда справедливо равенство

$$P \{ \xi = x, \eta = y \} = P \{ \xi = x \} P \{ \eta = y \}.$$

Для пояснения рассмотрим случай, когда величины ξ и η принимают каждая конечное число значений; значения ξ — x_1, \dots, x_n , значения η — y_1, \dots, y_m . Будем предполагать, что они записаны в порядке возрастания.

Если $x < x'$ и $y < y'$, то

$$\begin{aligned} P \{ x \leq \xi < x', y \leq \eta < y' \} &= P \{ \xi < x', y \leq \eta < y' \} - \\ &\quad - P \{ \xi < x, y \leq \eta < y' \} = P \{ \xi < x', \eta < y' \} - \\ &\quad - P \{ \xi < x', \eta < y \} - P \{ \xi < x, \eta < y' \} + P \{ \xi < x, \eta < y \} = \end{aligned}$$

$$\begin{aligned}
&= P\{\xi < x'\} P\{\eta < y'\} - P\{\xi < x'\} P\{\eta < y\} - \\
&\quad - P\{\xi < x\} P\{\eta < y'\} + P\{\xi < x\} P\{\eta < y\} = \\
&= (P\{\xi < x'\} - P\{\xi < x\})(P\{\eta < y'\} - P\{\eta < y\}) = \\
&\quad = P\{x \leq \xi < x'\} P\{y \leq \eta < y'\}.
\end{aligned}$$

Возьмем x и x' такими, чтобы единственное значение ξ , лежащее между x и x' , было x_k , а y и y' такими, чтобы единственным значением η , лежащим между y и y' , было y_l . Тогда $\{x \leq \xi < x'\} = \{\xi = x_k\}$, $\{y \leq \eta < y'\} = \{\eta = y_l\}$. Поэтому

$$P\{\xi = x_k, \eta = y_l\} = P\{\xi = x_k\} P\{\eta = y_l\}.$$

На данном равенстве основано следующее важное свойство во независимых случайных величин.

I. Если ξ и η независимые величины и существуют $M\xi$ и $M\eta$, то

$$M\xi\eta = M\xi M\eta \quad (3)$$

(математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий).

Докажем это равенство для дискретных величин. Пусть значения ξ будут x_1, \dots, x_n , а значения η — y_1, \dots, y_m . Тогда произведение $\xi\eta$ принимает значение $x_k y_l$, когда происходит событие $\{\xi = x_k\} \cap \{\eta = y_l\}$. Значит,

$$M\xi\eta = \sum x_k y_l P\{\xi = x_k\} P\{\eta = y_l\}.$$

Сумма, стоящая справа, есть произведение двух сумм $(x_1 P\{\xi = x_1\} + \dots + x_n P\{\xi = x_n\})(y_1 P\{\eta = y_1\} + \dots + y_m P\{\eta = y_m\})$.

Выражение в первой скобке есть $M\xi$, а во второй — $M\eta$. Тем самым формула (3) доказана.

Из этого свойства вытекает следующее.

II. Если ξ и η независимы, то справедливо соотношение

$$D(\xi + \eta) = D\xi + D\eta \quad (4)$$

(дисперсия суммы независимых случайных величин равна сумме их дисперсий). Действительно,

$$\begin{aligned}
D(\xi + \eta) &= M(\xi + \eta)^2 - (M\xi + M\eta)^2 = M\xi^2 + 2M\xi\eta + \\
&\quad + M\eta^2 - (M\xi)^2 - 2M\xi M\eta - (M\eta)^2 = D\xi + D\eta + \\
&\quad + 2(M\xi\eta - M\xi M\eta) = D\xi + D\eta,
\end{aligned}$$

так как на основании (3) выражение в скобках в последней строке равно нулю.

Заметим, что если ξ и η не являются независимыми, соотношение (3) может не выполняться. Например, при $\xi = \eta$

$$M\xi\eta = M\xi^2 \neq (M\xi)^2,$$

если только $D\xi \neq 0$. То же касается соотношения (4). Например,

$$D(\xi + \xi) = D(2\xi) = 4 \cdot D\xi,$$

а не $2 \cdot D\xi$, как вытекало бы из формулы (4).

Понятие независимости обобщается на несколько случайных величин.

Случайные величины ξ_1, \dots, ξ_n называются *независимыми*, если, каковы бы ни были x_1, \dots, x_n , события $\{\xi_1 < x_1\}, \dots, \{\xi_n < x_n\}$ независимы. Это эквивалентно соотношению

$$\begin{aligned} P\{\xi_1 < x_1, \dots, \xi_n < x_n\} &= \\ &= P\{\xi_1 < x_1\} \cdot \dots \cdot P\{\xi_n < x_n\}. \end{aligned}$$

Свойство I обобщается так: если существуют $M\xi_k$

$$M\xi_1 \dots \xi_n = M\xi_1 \dots M\xi_n, \quad (5)$$

то соотношение выводится из (3) по индукции.

Величины ξ_1, \dots, ξ_n называются *попарно независимыми*, если ξ_k и ξ_l независимы, каковы бы ни были $k \neq l$.

Свойство II распространяется на попарно независимые величины:

$$D(\xi_1 + \dots + \xi_n) = D\xi_1 + \dots + D\xi_n. \quad (6)$$

Например, для трех величин будем иметь

$$\begin{aligned} D(\xi_1 + \xi_2 + \xi_3) &= M(\xi_1 + \xi_2 + \xi_3)^2 - (M\xi_1 + M\xi_2 + M\xi_3)^2 = \\ &= M\xi_1^2 + M\xi_2^2 + M\xi_3^2 + 2M\xi_1\xi_2 + 2M\xi_1\xi_3 + 2M\xi_2\xi_3 - \\ &- (M\xi_1)^2 - (M\xi_2)^2 - (M\xi_3)^2 - 2M\xi_1M\xi_2 - 2M\xi_1M\xi_3 - \\ &- 2M\xi_2M\xi_3 = D\xi_1 + D\xi_2 + D\xi_3 + 2(M\xi_1\xi_2 - M\xi_1M\xi_2) + \\ &+ 2(M\xi_1\xi_3 - M\xi_1M\xi_3) + 2(M\xi_2\xi_3 - M\xi_2M\xi_3) = \\ &= D\xi_1 + D\xi_2 + D\xi_3, \end{aligned}$$

так как

$$M\xi_1\xi_2 = M\xi_1M\xi_2, \quad M\xi_1\xi_3 = M\xi_1M\xi_3, \quad M\xi_2\xi_3 = M\xi_2M\xi_3.$$

§ 18. Закон больших чисел

При рассмотрении понятия среднего значения в § 15 мы исходили из арифметического среднего независимых наблюдений случайной величины, полученных при повторении эксперимента, в котором эта величина наблюдается. Существование предельного значения у средних арифметических при возрастании числа наблюдений постулировалось на основании эмпирических фактов. Оказывается, сходимость средних можно получить строго математически. Это и составляет содержание закона больших чисел для случайных величин.

Напомним, что аналогичная ситуация была и при определении вероятности события: сначала постулировалось существование предельных значений частот, а затем после введения вероятности доказывался закон больших чисел, утверждающий сходимость в определенном смысле частоты события к его вероятности.

Мы рассмотрим сначала простейший вариант, когда величины ξ_1, ξ_2, \dots являются независимыми наблюдениями некоторой случайной величины, принимающей конечно число значений x_1, \dots, x_m .

Для всякого $\varepsilon > 0$ выполнено соотношение

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_1^n \xi_k - M\xi \right| > \varepsilon \right\} = 0. \quad (1)$$

Это соотношение устанавливается с использованием закона больших чисел для частот событий.

Обозначим через $\nu_n(\xi = x_i)$ частоту события $\{\xi = x_i\}$ в первых n наблюдениях, т. е. число тех величин ξ_k при $k \leq n$, которые принимают значение x_i , деленное на n . Тогда

$$\frac{1}{n} \sum_1^n \xi_k = x_1 \nu_n(\xi = x_1) + \dots + x_m \nu_n(\xi = x_m),$$

$$M\xi = x_1 P\{\xi = x_1\} + \dots + x_m P\{\xi = x_m\},$$

$$\begin{aligned} \left| \frac{1}{n} \sum_1^n \xi_k - M\xi \right| &\leq |x_1| |\nu_n(\xi = x_1) - P\{\xi = x_1\}| + \dots + \\ &+ |x_m| |\nu_n(\xi = x_m) - P\{\xi = x_m\}|. \end{aligned}$$

Пусть $|x_k| \leq c$ для всех k . Тогда

$$\left| \frac{1}{n} \sum_1^n \xi_k - M\xi \right| \leq c (|v_n(\xi = x_1) - P\{\xi = x_2\}| + \dots + |v_n(\xi = x_m) - P\{\xi = x_m\}|).$$

Если $\left| \frac{1}{n} \sum_1^n \xi_k - M\xi \right| > \varepsilon$, то происходит хотя бы одно из событий

$$\left\{ c|v_n(\xi = x_k) - P\{\xi = x_k\}| \geq \frac{\varepsilon}{m} \right\}, \quad k = 1, \dots, m,$$

так как, если для всех k происходят события

$$\left\{ c|v_n(\xi = x_k) - P\{\xi = x_k\}| > \frac{\varepsilon}{m} \right\}, \quad k = 1, \dots, m,$$

то $\left| \frac{1}{n} \sum_1^n \xi_k - M\xi \right| \leq \varepsilon$. Таким образом,

$$P \left\{ \left| \frac{1}{n} \sum_1^n \xi_k - M\xi \right| > \varepsilon \right\} \leq P \left\{ \left\{ c|v_n(\xi = x_1) - P\{\xi = x_1\}| > \frac{\varepsilon}{m} \right\} + \dots + \left\{ c|v_n(\xi = x_m) - P\{\xi = x_m\}| > \frac{\varepsilon}{m} \right\} \right\} \leq P \left\{ |v_n(\xi = x_1) - P\{\xi = x_1\}| > \frac{\varepsilon}{mc} \right\} + \dots + P \left\{ |v_n(\xi = x_1) - P\{\xi = x_1\}| > \frac{\varepsilon}{mc} \right\}.$$

Если $n \rightarrow \infty$, то все слагаемые в правой части на основании закона больших чисел стремятся к нулю. Справедливость (1) доказана.

Прежде чем перейти к общему случаю, разберем важное неравенство, выведенное П. Л. Чебышевым и носящее его имя.

Пусть ξ такая величина, для которой существует дисперсия. Тогда для всякого $\varepsilon > 0$

$$P \{ |\xi - M\xi| \geq \varepsilon \} \leq \frac{D\xi}{\varepsilon^2}. \quad (2)$$

Чтобы получить это неравенство, установим упрощенный вариант неравенства Чебышева. Если η неотрицательная случайная величина, то для всякого $\lambda > 0$

$$P \{ \eta \geq \lambda \} \leq \frac{1}{\lambda} M\eta. \quad (3)$$

Рассмотрим величину η_λ , определенную так: $\eta_\lambda = 0$ при $\eta < \lambda$, $\eta_\lambda = \lambda$ при $\eta \geq \lambda$. Тогда $\eta_\lambda \leq \eta$ и, значит, $M\eta_\lambda \leq M\eta$. Но

$$M\eta_\lambda = 0 \cdot P\{\eta < \lambda\} + \lambda P\{\eta \geq \lambda\} = \lambda P\{\eta \geq \lambda\}.$$

Таким образом,

$$M\eta \geq M\eta_\lambda = \lambda P\{\eta \geq \lambda\}.$$

Отсюда вытекает (3).

Для доказательства (2) рассмотрим неотрицательную величину $\eta = (\xi - M\xi)^2$. Имеем $M\eta = D\xi$. Значит,

$$P\{|\xi - M\xi| \geq \varepsilon\} = P\{(\xi - M\xi)^2 \geq \varepsilon^2\} \leq \frac{1}{\varepsilon^2} M(\xi - M\xi)^2 = \frac{1}{\varepsilon^2} D\xi.$$

Перейдем теперь к формулировке и доказательству обобщенного закона больших чисел, установленного П. Л. Чебышевым.

Пусть $\xi_1, \xi_2, \dots, \xi_n, \dots$ — последовательность попарно независимых случайных величин, для которых существуют математические ожидания и дисперсии, причем для некоторого c $D\xi_k \leq c$ для всех k . Тогда для всякого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_1^n \xi_k - \frac{1}{n} \sum_1^n M\xi_k\right| > \varepsilon\right\} = 0. \quad (4)$$

Это означает, что среднее арифметическое случайных величин мало отличается от среднего арифметического их математических ожиданий.

Для доказательства рассмотрим случайную величину $\eta = \frac{1}{n} \sum_1^n \xi_k$. Для нее справедливы равенства

$$M\eta = \frac{1}{n} \sum_1^n M\xi_k, \quad D\eta = \frac{1}{n^2} D\left(\sum_1^n \xi_k\right) = \frac{1}{n^2} \sum_1^n D\xi_k$$

(мы используем свойство суммы попарно независимых случайных величин). Записывая для величины η неравенство Чебышева по формуле (2), будем иметь

$$P\left\{\left|\frac{1}{n} \sum_1^n \xi_k - \frac{1}{n} \sum_1^n M\xi_k\right| > \varepsilon\right\} \leq \frac{1}{\varepsilon^2} \cdot \frac{1}{n^2} \sum_1^n D\xi_k.$$

Так как $D\xi_k \leq c$, то

$$\sum_1^n D\xi_k \leq nc$$

(сумма n слагаемых, каждое из которых не превосходит c). Следовательно,

$$P \left\{ \left| \frac{1}{n} \sum_1^n \xi_k - \frac{1}{n} \sum_1^n M\xi_k \right| > \varepsilon \right\} \leq \frac{nc}{\varepsilon^2 n^2} = \frac{c}{\varepsilon^2 n}.$$

При $n \rightarrow \infty$ правая часть стремится к нулю, каково бы ни было $\varepsilon > 0$. Соотношение (4) доказано.

Укажем примеры наиболее характерных ситуаций, в которых проявляется или используется закон больших чисел.

1. Давление газа на стенку сосуда определяется импульсами молекул газа, соударяющихся со стенкой. Для простоты возьмем молекулы, летящие перпендикулярно к поверхности стенки (на самом деле нужно рассматривать проекции скоростей на это направление). Если молекула массы m соударяется со стенкой со скоростью v , то в результате упругого удара ее скорость, оставаясь по величине той же, меняет знак. Изменение количества движения этой молекулы равно $2mv$. Если за некоторое время h со стенкой столкнулись n молекул и их скорости были v_1, \dots, v_n , то общее изменение количества движения газа равно $\sum_1^n 2mv_i$. На основании второго закона Ньютона оно равно импульсу силы со стороны стенки на газ: произведению силы на время. Сила воздействия стенки сосуда на газ равна на основании третьего закона Ньютона силе воздействия газа на стенку, т. е. силе давления газа. Если площадь стенки S , давление P , то

$$\sum_1^n 2mv_i = hSP, \quad P = \frac{2m}{hS} \sum_1^n v_i.$$

Скорости отдельных молекул случайны, справа стоит случайная величина. Тем не менее давление газа постоянно. Это результат действия закона больших чисел. Вместо суммы справа можно рассматривать ее математическое ожидание.

2. Ток в проводнике создается движением свободных электронов. Если e — заряд одного электрона, а v — его

скорость, то он создает ток ve . Поэтому суммарный ток, создаваемый n электронами, если их скорости v_1, \dots, v_n , будет

$$I = e(v_1 + \dots + v_n).$$

Здесь справа также стоит случайная величина, так как скорости отдельных электронов случайны. Тем не менее можно говорить о силе тока. И зависит она не от скоростей электронов, а от таких характеристик, как напряжение на концах проводника и его сопротивление. Именно они определяют среднюю скорость электронов, через нее и выражается сила тока. Так что в выражение справа вместо скоростей можно подставлять их среднее значение. Это также проявление закона больших чисел.

3. Исключение случайных ошибок измерений также основано на действии закона больших чисел. Пусть измеряется некоторая величина x . Обычно с помощью измерительных приборов нет возможности определить ее абсолютно точно. Существует некоторая ошибка измерения, скажем, ξ . Если $M\xi \neq 0$, то тогда говорят, что прибор имеет систематическую ошибку. Перестроив шкалу, можно добиться, чтобы систематическая ошибка отсутствовала, т. е. чтобы $M\xi = 0$. Если x_1, \dots, x_n — измерения величины x , а $\xi_1 = x_1 - x, \dots, \xi_n = x_n - x$ — ошибки измерения при различных измерениях, то естественно считать их независимыми случайными величинами. Так как $M\xi_k = 0$, то на основании закона больших чисел

$$P\left\{\left|\frac{1}{n} \sum_1^n \xi_k\right| > \varepsilon\right\} \rightarrow 0$$

при $n \rightarrow \infty$.

Значит, беря среднее значение измерений $\frac{1}{n}(x_1 + \dots + x_n) = x + \frac{1}{n}(\xi_1 + \dots + \xi_n)$, мы можем сколь угодно точно получить значение измеряемой величины, хотя все измерения производились с ошибками.

4. При определении средних характеристик некоторого биологического вида берут некоторое число особей (достаточно большое) и проводят соответствующие измерения. Арифметические средние, в силу закона больших чисел, дают достаточно хорошее приближение для средних по всему виду. Это значит, что для определения, например, сред-

него веса, скажем, особей данного вида не нужно взвешивать всех представителей, достаточно взвесить 100 независимо отобранных особей.

§ 19. Центральная предельная теорема

При конечных значениях n даже в условиях применения закона больших чисел разность между средними арифметическими независимых случайных величин и их математических ожиданий не равна нулю. Эта разность носит название *флуктуации* (точно так мы говорим о флуктуации частоты). Существуют флуктуация давления газа на стенку сосуда, флуктуация силы тока, флуктуация средних от измерения какой-либо характеристики или величины.

Для грубой оценки характера отклонения суммы случайных величин от математического ожидания этой суммы можно использовать неравенство Чебышева. Пусть $\xi_1 + \dots + \xi_n$ — независимые случайные величины, для которых существуют математические ожидания и дисперсии. Тогда на основании неравенства Чебышева

$$P \left\{ \left| \sum_1^n \xi_k - M \sum_1^n \xi_k \right| \geq x \right\} \leq \frac{1}{x^2} \sum_1^n D\xi_k.$$

Подставим вместо x величину $y \sqrt{\sum_1^n D\xi_k}$. Получим

$$P \left\{ \frac{\left| \sum_1^n \xi_k - M \sum_1^n \xi_k \right|}{\sqrt{\sum_1^n D\xi_k}} > y \right\} \leq \frac{1}{y^2}.$$

Это показывает, что отклонения суммы случайных величин от суммы их дисперсий имеют порядок $\sqrt{\sum_1^n D\xi_k}$,

так что величина $\frac{1}{\sqrt{\sum_1^n D\xi_k}} \left(\sum_1^n \xi_k - M \sum_1^n \xi_k \right)$ ограничена

(в вероятностном смысле). Центральная предельная теорема устанавливает, что (при определенных условиях) эта величина в пределе при $n \rightarrow \infty$ имеет нормальное

распределение, т. е. распределение с плотностью

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

Наиболее просто центральная предельная теорема формулируется для того случая, когда ξ_k являются независимыми наблюдениями одной и той же случайной величины. В этом случае у величин ξ_k совпадают функции распределения (такие величины называются *одинаково распределенными*), математические ожидания и дисперсии.

Пусть $M\xi_k = a$, $D\xi_k = b$. Тогда

$$\sum_1^n M\xi_k = na, \quad \sum_1^n D\xi_k = nb.$$

Каковы бы ни были $x_1 < x_2$,

$$\lim_{n \rightarrow \infty} P \left\{ x_1 < \frac{1}{\sqrt{nb}} \left(\sum_1^n \xi_k - na \right) < x_2 \right\} = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx. \quad (1)$$

Доказательство равенства (1) опирается на довольно сложный аналитический аппарат, поэтому мы рассмотрим его в простейшем случае, когда величина ξ принимает всего два значения. Заметим, что если $\tilde{\xi} = \frac{1}{\sqrt{b}} (\xi - a)$, то $M\tilde{\xi} = \frac{1}{\sqrt{b}} (M\xi - a) = 0$, $D\tilde{\xi} = \frac{1}{b} D(\xi - a) = \frac{1}{b} D\xi = 1$.

Далее,

$$\frac{1}{\sqrt{nb}} \left(\sum_1^n \xi_k - na \right) = \frac{1}{\sqrt{n}} \sum_1^n \frac{\xi_k - a}{\sqrt{b}} = \frac{1}{\sqrt{n}} \sum_1^n \tilde{\xi}_k.$$

Поэтому доказательство центральной предельной теоремы достаточно получить для величин с математическим ожиданием 0 и дисперсией 1. Если ξ принимает два значения, то и $\tilde{\xi}$ принимает два значения. Обозначим значения $\tilde{\xi}$ через y_1 и y_2 ,

$P \{ \tilde{\xi} = y_1 \} = p$, $P \{ \tilde{\xi} = y_2 \} = 1 - p$. Тогда

$$M\tilde{\xi} = y_1 p + y_2 (1 - p) = 0.$$

Откуда

$$y_1 = \frac{p-1}{p} y_2,$$

$$1 = D\tilde{\xi} = y_1^2 p + y_2^2 (1-p) = \left[\frac{(1-p)^2}{p} + (1-p) \right] y_2^2 = \\ = \frac{1-p}{p} y_2^2.$$

Значит, $y_1 = \sqrt{\frac{1-p}{p}}$, $y_2 = -\sqrt{\frac{p}{1-p}}$.

Обозначим через A событие $\{\tilde{\xi} = y_1\}$. Если v_n — частота события A в n испытаниях, т. е. среди величин $\tilde{\xi}_1, \dots, \tilde{\xi}_n$ nv_n раз встречается значение y_1 и $n - nv_n$ раз значение y_2 , то

$$\tilde{\xi}_1 + \dots + \tilde{\xi}_n = n \left(v_n \sqrt{\frac{1-p}{p}} + (1-v_n) \sqrt{\frac{p}{1-p}} \right) = \\ = n \left[v_n \left(\sqrt{\frac{1-p}{p}} + \sqrt{\frac{p}{1-p}} \right) - \sqrt{\frac{p}{1-p}} \right] = \\ = \frac{n}{\sqrt{p(1-p)}} (v_n - p).$$

Поэтому

$$\frac{1}{\sqrt{n}} (\tilde{\xi}_1 + \dots + \tilde{\xi}_n) = \sqrt{\frac{n}{p(1-p)}} (v_n - p).$$

Остается воспользоваться результатами § 9, в частности формулой (6), в которую вместо a и b следует подставить x_1 и x_2 .

Центральная предельная теорема дает более точную оценку для флуктуаций, чем та, которая может быть получена на основании неравенства Чебышева. Пусть, например, величины ξ_1, \dots, ξ_n независимы и одинаково распределены, $M\xi_k = 0$, $D\xi_k = 1$. Тогда

$$P \left\{ \left| \frac{1}{n} \sum_1^n \xi_k \right| \geq x \right\} \leq \frac{1}{x^2} D \left(\frac{1}{n} \sum_1^n \xi_k \right) = \frac{1}{nx^2}.$$

На основании центральной предельной теоремы

$$P \left\{ \left| \frac{1}{\sqrt{n}} \sum_1^n \xi_k \right| \geq x \sqrt{n} \right\} = \\ = 1 - P \left\{ -x \sqrt{n} < \frac{1}{\sqrt{n}} \sum_1^n \xi_k < x \sqrt{n} \right\} \sim \\ \sim 1 - \frac{1}{\sqrt{2\pi}} \int_{-x\sqrt{n}}^{x\sqrt{n}} e^{-\frac{u^2}{2}} du.$$

Последнее выражение, как вытекает из таблиц, приведенных в § 9 при $x \sqrt{n} \geq 3$, меньше чем 0,0027, в то время как неравенство Чебышева дает оценку $\frac{1}{9} \approx 0,11$.

Центральная предельная теорема справедлива и для различно распределенных независимых случайных величин при некоторых дополнительных ограничениях. Для этого, например, достаточно, чтобы все величины были ограничены одной постоянной.

Для того чтобы лучше понять возможные следствия из центральной предельной теоремы, нам понадобится понятие об общем Гауссовом распределении. Величина ξ имеет Гауссово распределение, если существуют $M\xi$, $D\xi$ и величина

$$\tilde{\xi} = \frac{1}{\sqrt{D\xi}} (\xi - M\xi)$$

имеет нормальное распределение ($\tilde{\xi}$ имеет математическое ожидание 0 и дисперсию 1).

Найдем плотность распределения величины ξ . Пусть $M\xi = a$, $D\xi = \sigma^2$. Тогда

$$\tilde{\xi} = \frac{1}{\sigma} (\xi - a), \quad \xi = \sigma \tilde{\xi} + a.$$

Вероятность того, что ξ попадет в интервал $[x, x + \Delta x]$, совпадает с вероятностью того, что $\tilde{\xi}$ попадет в интервал $\left[\frac{1}{\sigma} (x - a), \frac{1}{\sigma} (x + \Delta x - a) \right]$, т. е. интервал $[y, y + \Delta y]$,

где $y = \frac{1}{\sigma} (x - a)$, $\Delta y = \frac{1}{\sigma} \Delta x$. По определению плотности нормального распределения эта вероятность эквивалентна выражению $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} y^2} \Delta y = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} \frac{\Delta x}{\sigma}$.

Таким образом,

$$P \{x < \xi < x + \Delta x\} \sim \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} \Delta x.$$

Это означает, что ξ имеет плотность распределения, равную

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

Здесь $a = M\xi$, $\sigma^2 = D\xi$. Гауссово распределение можно непосредственно задавать с помощью плотности вида (2), оно зависит от двух параметров и однозначно определяется заданием среднего значения и дисперсии случайной величины.

Из центральной предельной теоремы вытекает, что сумму большого числа независимых случайных величин имеет приближенно Гауссово распределение. Каждый раз, когда наблюдаемая величина есть сумма большого числа независимых случайных величин, мы можем ожидать, что она имеет Гауссово распределение. Такими будут флуктуации давления газа, силы тока в проводнике (основываясь на соображениях, приведенных в предыдущем параграфе). Гауссово распределение будут иметь ошибки измерения (этот факт был обнаружен К. Гауссом, отсюда и название распределения).

Приведем одно важное следствие центральной предельной теоремы для биологии. Оно относится к размерам различных органов биологических индивидов. Можно исходить из такой схемы роста индивида. Увеличение размера органа в определенный промежуток пропорционально уже достигнутому размеру с коэффициентом, вообще говоря, случайным, зависящим от условий внешней среды. Пусть весь период роста разбит на n промежутков, η_0 — начальное значение размера, ξ_k — коэффициент пропорциональности на k -м промежутке. Естественно предположение, что значения ξ_k на разных промежутках времени независимы. Если η_k — размер органа после истечения k -го промежутка времени, то $\eta_k = \xi_k \eta_{k-1}$. Поэтому $\eta_1 = \xi_1 \eta_0$, $\eta_2 = \xi_2 \eta_1 = \xi_2 \xi_1 \eta_0$, $\eta_n = \eta_0 \xi_1 \dots \xi_n$.

Возьмем логарифм η_n , получим

$$\log \eta_n = \log \eta_0 + \log \xi_1 + \dots + \log \xi_n.$$

Из независимости величин η_0 , ξ_1 , ..., ξ_n вытекает независимость их логарифмов, поэтому $\log \eta_n$ имеет приближенно нормальное распределение.

Таким образом, размеры органов биологических индивидов являются такими случайными величинами, логарифмы которых имеют Гауссово распределение. В этом случае говорят, что сама случайная величина имеет логарифмически Гауссово (логнормальное) распределение. Факт логарифмически Гауссовых распределений для размеров многих органов подтвержден экспериментально.

§ 20. Эмпирическая функция распределения

Случайные величины, наблюдаемые в различных экспериментах, имеют, как правило, неизвестные нам функции распределения. Здесь мы встречаемся с такой же ситуацией, как и при наблюдении случайных событий, вероятности которых неизвестны.

Для решения многих задач важно знать распределение случайных величин, ибо на их основании принимаются те или иные решения. Так, можно пренебрегать маловероятными событиями, однако для оценки степени риска, которой мы подвергаемся при таком пренебрежении, нужно знать вероятности этих событий. В § 10 рассматривалась задача об оценке неизвестной вероятности события по результатам независимых экспериментов, в которых это событие наблюдается. Аналогично, имея независимые наблюдения величины ξ , можно оценить ее функцию распределения. Для этого и используют эмпирическую функцию распределения.

Пусть проведено n независимых экспериментов, в которых наблюдалась одна и та же случайная величина ξ . Обозначим через x_k значение величины в k -м эксперименте. Расположим n значений x_1, x_2, \dots, x_n в порядке возрастания: $x_1^* \leq x_2^* \leq \dots \leq x_n^*$. Эта последовательность называется *вариационным рядом*. Если наблюдения были $-2; 0; 1; 1,5; -1,5; 1; 0; -1; 0; 2$, то соответствующий вариационный ряд будет $-2; -1,5; -1; 0; 0; 0; 1; 1; 1,5; 2$. Как видим, в вариационном ряде значения могут повторяться.

Пусть $k_n(x)$ обозначает число членов вариационного ряда, для которых $x_i^* < x$. Тогда эмпирическая функция распределения определяется равенством

$$F_n^*(x) = \frac{1}{n} k_n(x).$$

Например, для приведенной выше последовательности наблюдений имеем

x	$(-\infty; -2]$	$(-2; -1,5]$	$(-1,5; -1]$	$(-1; 0]$	$(0; 1]$	$(1; 1,5]$	$(1,5; 2]$	$(2; \infty)$
$F_n^*(x)$	0	0,1	0,2	0,3	0,6	0,8	0,9	1

Так как $k_n(x)$ — число появлений событий $\{\xi < x\}$ в n наблюдениях (именно x_1, \dots, x_n), то $F_n^*(x)$ есть частота события $\{\xi < x\}$. Поэтому при $n \rightarrow \infty$ $F_n^*(x)$ на основании закона больших чисел, установленного в § 8, стремится к вероятности этого события, т. е. к $P\{\xi < x\}$. Если обозначить функцию распределения через $F(x)$, то стремление $F_n^*(x)$ к $F(x)$ происходит в следующем смысле: для всякого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\{|F_n^*(x) - F(x)| > \varepsilon\} = 0$$

(заметим, что эмпирическая функция распределения есть случайная величина, поэтому имеет смысл рассматривать указанную вероятность).

Для того чтобы оценить скорость сходимости эмпирической функции распределения к функции распределения, можно использовать те же соображения, которые использовались в § 10 при оценке неизвестной вероятности с помощью частоты. Это можно сделать, так как $F_n^*(x)$ есть частота события $\{\xi < x\}$, а $F(x)$ — вероятность этого события. Поэтому, как вытекает из § 9, для любых $\alpha < \beta$

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left\{\alpha < \sqrt{\frac{n}{F(x)(1-F(x))}} (F_n^*(x) - F(x)) < \beta\right\} = \\ = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{u^2}{2}} du. \end{aligned}$$

Таким образом, отклонение эмпирической функции распределения от функции распределения имеет порядок

$$\sqrt{\frac{1}{n} F(x)(1-F(x))}.$$

Исследование отклонения эмпирической функции распределения позволяет проверять гипотезу о согласии наблюдаемых значений случайной величины с заданной функцией распределения. Эта задача по-разному решается для дискретных и непрерывных распределений, эмпирическую функцию распределения естественно использовать именно для непрерывных величин (в случае дискретных величин задача сводится к проверке гипотезы о вероятности события, которая уже рассматривалась в § 10).

Пусть, исходя из теоретических соображений, можно предположить, что данная величина ξ имеет функцию

распределения $F(x)$. В нашем распоряжении имеются наблюдения x_1, \dots, x_n величины ξ , полученные в независимых экспериментах. Нужно решить, подтверждают или опровергают эти наблюдения гипотезу о том, что ξ имеет распределение $F(x)$.

Для решения обратимся к той же идее, которую использовали в § 10 при проверке гипотезы о том, что вероятность события A равна заданному числу p . Если гипотеза верна, то, скажем, вероятность того, что $\left| \sqrt{\frac{n}{p(1-p)}} |v_n(A) - p| > 3 \right.$ меньше, чем 0,0027, поэтому, отвергая гипотезу, если выполнено это неравенство, и принимая в противоположном случае, мы исключим истинную гипотезу с весьма малой вероятностью. В нашем случае дело усложняется тем, что наблюдается не одно событие, а бесконечное множество событий $\{\xi < x\}$.

Чтобы избежать этой трудности, А. Н. Колмогоров предложил рассматривать максимальное отклонение эмпирической функции распределения от $F(x)$; точнее, он предложил рассматривать величину

$$D_n = \sqrt{n} \max_x |F(x) - F_n^*(x)|$$

(справа стоит умноженное на \sqrt{n} максимальное отклонение по абсолютной величине функции $F_n^*(x)$ от $F(x)$). Оказалось, что величина D_n имеет при $n \rightarrow \infty$ предельное распределение, причем оно не зависит от гипотетической функции распределения $F(x)$.

Н. В. Смирнов обнаружил, что таким же свойством обладает и величина

$$D_n^+ = \sqrt{n} \max_x (F_n^*(x) - F(x))$$

(здесь берется максимальная положительная разность между $F_n^*(x)$ и $F(x)$). Преимущество величины D_n^+ заключается в том, что предельное распределение для нее имеет более простой вид, чем у D_n . Справедливо равенство: при $a > 0$

$$\lim_{n \rightarrow \infty} P \{D_n^+ < a\} = 1 - e^{-2a^2}.$$

Например, при $a = 2$

$$\lim_{n \rightarrow \infty} P \{D_n^+ > a\} < 0,0004.$$

Использовать этот результат можно следующим образом. Строим по наблюдениям $F_n^*(x)$, находим D_n^+ . Если $D_n^+ > 2$, гипотеза отвергается, при $D_n^+ \leq 2$ считаем, что они не противоречат гипотезе, можно продолжить наблюдения для уточнения.

Предположим, проверяется гипотеза о равномерной распределенности величины на $[1, 30]$ (предполагается, что некое событие может произойти с равными вероятностями в любой день месяца, хотя здесь мы имеем дело с дискретным распределением, но в силу малости вероятностей отдельных значений можем с достаточной степенью точности считать его непрерывным). Проведено 20 наблюдений, они записаны в порядке возрастания; далее в таблице приведены значения в этих точках эмпирической и гипотетической функций распределения и их разность, если она положительна.

1	1	3	5	5	5	6	7	10	11	
	0,1	0,15			0,3	0,35	0,4	0,45		
	0,03	0,1			0,17	0,2	0,23	0,33		
	0,07	0,05			0,13	0,15	0,17	0,12		
11	13	13	14	16	18	20	22	24	25	
0,55		0,65	0,7	0,75	0,8	0,85	0,9	0,95	1	
0,36		0,43	0,46	0,53	0,6	0,67	0,73	0,8	0,83	
0,19		0,22	0,24	0,22	0,2	0,18	0,17	0,15	0,17	

Максимальное отклонение будет 0,24. Следовательно,

$$D_{20}^+ = \sqrt{20} \cdot 0,24 = 1,56.$$

Оценим вероятность того, что $D_n^+ > 1,5$ (будем использовать предельное распределение). Имеем

$$P \{D_n^+ > 1,5\} = e^{-2(1,5)^2} = e^{-4,5} < 0,012.$$

Поэтому, отвергая гипотезу о равномерном распределении, мы примем неверную гипотезу с вероятностью меньшей, чем $0,012$.

Кроме эмпирической функции распределения рассматриваются также эмпирическое среднее значение и эмпирическая дисперсия. Они вычисляются по эмпирическому распределению точно так, как математическое ожидание и дисперсия вычисляются по обычному. Эмпирическое распределение — дискретное. Пусть $z_1 < z_2 < \dots < z_m$ — различные члены вариационного ряда. Если среди членов вариационного ряда $k_1 = z_1, k_2 = z_2, \dots, k_m = z_m$ ($k_1 + \dots + k_m = n$), то эмпирическое распределение есть распределение величины, принимающей значения z_1, \dots, z_m с вероятностями $\frac{k_1}{n}, \dots, \frac{k_m}{n}$. Среднее значение для такого распределения будет

$$\bar{x}_n = \frac{k_1}{n} z_1 + \dots + \frac{k_m}{n} z_m = \frac{x_1 + \dots + x_n}{n},$$

т. е. это просто среднее арифметическое из наблюдений.

Из закона больших чисел вытекает, что \bar{x}_n стремится к a , где $a = M\xi$. Величина

$$\sqrt{n}(\bar{x} - a)$$

при $n \rightarrow \infty$ имеет нормальное предельное распределение со средним 0 и дисперсией $b = D\xi$ (это следствие центральной предельной теоремы).

Пусть \bar{s}_n^2 — эмпирическая дисперсия. Тогда она определяется равенством

$$\begin{aligned} \bar{s}_n^2 &= \frac{k_1}{n} (z_1 - \bar{x})^2 + \dots + \frac{k_m}{n} (z_m - \bar{x})^2 = \\ &= \frac{k_1 z_1^2 + \dots + k_m z_m^2}{n} - \bar{x}^2 = \frac{x_1^2 + \dots + x_n^2}{n} - \bar{x}^2 = \\ &= \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x}_n)^2. \end{aligned}$$

Это просто среднее квадратическое отклонение. Так как в силу закона больших чисел для величин x_1^2, \dots, x_n^2 (будем предполагать, что $D\xi^2$ конечна) среднее $\frac{1}{n} (x_1^2 + \dots + x_n^2)$ сходится к $M\xi^2$, а $\bar{x}_n \rightarrow M\xi$, то \bar{s}_n^2 будет сходиться к

$b = D\xi$. Если заранее известно, что распределение величины Гауссово, а неизвестны его параметры a и b , то величины \bar{x}_n и \bar{s}_n^2 используются в качестве их приближенных значений.

§ 21. Корреляция случайных величин

Если случайные величины ξ и η независимы, то $M\xi\eta = M\xi M\eta = 0$. Для произвольных величины ξ и η выражение $M\xi\eta - M\xi M\eta$ может быть отлично от нуля. В этом случае величины ξ и η называются *коррелированными*. Коррелированные величины не являются независимыми.

Рассмотрим две величины ξ и η с конечными дисперсиями. Покажем, что имеет место неравенство

$$|M(\xi - M\xi)(\eta - M\eta)|^2 \leq D\xi \cdot D\eta. \quad (1)$$

Обозначим $\xi - M\xi = \xi'$, $\eta - M\eta = \eta'$. Тогда

$$M(\xi' + t\eta')^2 = M(\xi')^2 + 2tM\xi'\eta' + t^2M(\eta')^2.$$

Справа записан квадратный трехчлен относительно t . Этот квадратный трехчлен неотрицателен для всех t , поэтому его корни либо совпадают, либо мнимы, следовательно,

$$(M\xi'\eta')^2 - M(\xi')^2 M(\eta')^2 \leq 0.$$

Если вспомнить обозначения ξ' и η' , то из последнего неравенства вытекает (1). Заметим, что

$$M(\xi - M\xi)(\eta - M\eta) = M\xi\eta - M(M\xi)\eta - M\xi(M\eta) + M\xi M\eta = M\xi\eta - M\xi M\eta.$$

Из неравенства (1) и последнего равенства вытекает, что величина

$$r_{\xi\eta} = \frac{M\xi\eta - M\xi M\eta}{\sqrt{D\xi D\eta}}$$

удовлетворяет неравенству $|r_{\xi\eta}| \leq 1$. $r_{\xi\eta}$ называется *коэффициентом корреляции* величин ξ и η . Если $r_{\xi\eta} = 0$, то величины называются *некоррелированными* (они будут такими, например, в случае независимых величин).

Рассмотрим некоторые свойства коэффициента корреляции.

I. Пусть величины ξ' и η' выражаются через ξ и η соотношениями $\xi' = a\xi + b$, $\eta' = c\eta + d$, $a > 0$, $c > 0$.

Тогда $r_{\xi\eta} = r_{\xi'\eta'}$. Действительно,

$$M\xi' = aM\xi + b, \quad M\eta' = cM\eta + d,$$

$$M(\xi' - M\xi')(\eta' - M\eta') = M(a\xi - aM\xi)(c\eta - cM\eta) = \\ = acM(\xi - M\xi)(\eta - M\eta),$$

$$D\xi' = a^2D\xi, \quad D\eta' = c^2D\eta.$$

Поэтому

$$r_{\xi'\eta'} = \frac{acM(\xi - M\xi)(\eta - M\eta)}{\sqrt{a^2D\xi \cdot c^2D\eta}} = r_{\xi\eta}.$$

II. Если $|r_{\xi\eta}| = 1$, то ξ и η связаны линейным соотношением

$$\frac{\xi - M\xi}{\sqrt{D\xi}} = r_{\xi\eta} \frac{\eta - M\eta}{\sqrt{D\eta}}. \quad (2)$$

Установим это равенство для случая $r_{\xi\eta} = 1$. Положим

$$\xi' = \frac{\xi - M\xi}{\sqrt{D\xi}}, \quad \eta' = \frac{\eta - M\eta}{\sqrt{D\eta}}. \quad \text{Тогда } M\xi' = M\eta' = 0.$$

Поэтому

$$M(\xi' - \eta')^2 = M(\xi')^2 + M(\eta')^2 - 2M\xi'\eta' = \\ = D\xi' + D\eta' - 2r_{\xi'\eta'} = 0.$$

Поскольку $(\xi' - \eta')^2$ неотрицательная величина и ее математическое ожидание равно 0, то эта величина равна нулю с вероятностью 1. Значит $\xi' = \eta'$, откуда и вытекает (2) для $r_{\xi\eta} = 1$.

III. Предположим, что мы желаем приближенно выразить η через ξ с помощью линейной функции $\eta \approx a\xi + b$. Как выбрать a и b так, чтобы точность этой формулы была наибольшей? Ошибка представления есть $\eta - a\xi - b$. Будем подбирать a и b так, чтобы $M(\eta - a\xi - b)^2$ было наименьшим (чтобы среднее значение квадрата ошибки было минимальным). Имеем

$$M(\eta - a\xi - b)^2 = [M\eta - aM\xi - b]^2 + D(\eta - a\xi - b).$$

Второе слагаемое от b не зависит. А первое имеет вид

$$(M\eta - aM\xi - b)^2.$$

Это неотрицательная величина, она принимает наименьшее значение, если

$$b = M\eta - aM\xi.$$

Второе слагаемое можно записать так:

$$\begin{aligned}
 D(\eta - a\xi) &= D\eta - 2aM(\xi - M\xi)(\eta - M\eta) + a^2D\xi = \\
 &= D\eta - 2ar_{\xi\eta} \sqrt{D\eta} \sqrt{D\xi} + a^2D\xi = D\eta(1 - r_{\xi\eta}^2) + \\
 &\quad + D\eta r_{\xi\eta}^2 - 2ar_{\xi\eta} \sqrt{D\eta} \sqrt{D\xi} + a^2D\xi = \\
 &= D\eta(1 - r_{\xi\eta}^2) + (\sqrt{D\eta} r_{\xi\eta} - a \sqrt{D\xi})^2.
 \end{aligned}$$

Первое слагаемое от a не зависит, второе неотрицательно и обращается в нуль при

$$\sqrt{D\eta} r_{\xi\eta} - a \sqrt{D\xi} = 0, \quad a = r_{\xi\eta} \sqrt{\frac{D\eta}{D\xi}}.$$

Таким образом, наилучшее представление (с наименьшим средним значением квадрата ошибки) величины η с помощью величины ξ имеет вид

$$\eta = r_{\xi\eta} \sqrt{\frac{D\eta}{D\xi}} (\xi - M\xi) + M\eta + \varepsilon, \quad (3)$$

где ε ошибка, при этом

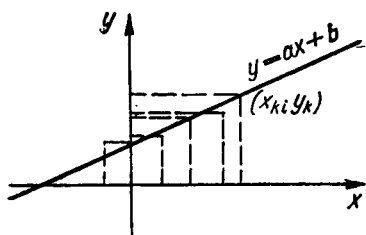
$$M\varepsilon^2 = D\eta(1 - r_{\xi\eta}^2). \quad (4)$$

Чем ближе $|r_{\xi\eta}|$ к 1, тем точнее будет формула после отбрасывания ε . Поэтому $r_{\xi\eta}$ есть мера линейной зависимости между величинами ξ и η .

Свойство III используется при нахождении линейных зависимостей между наблюдаемыми величинами. Например, в биологии можно исследовать зависимость между размерами определенных органов для некоторого вида либо влияния определенных условий на протекание каких-либо физиологических процессов в организме. Можно таким образом изучать зависимость урожайности от проведения тех или других сельскохозяйственных мероприятий, например от количества внесенных удобрений.

Во всех этих случаях имеется такая общая схема. В эксперименте одновременно наблюдаются две величины ξ и η . При повторении эксперимента n раз получаем два ряда наблюдений

$$\begin{array}{c} \xi \mid x_1, \dots, x_n \\ \eta \mid y_1, \dots, y_n \end{array}$$



Будем отмечать наблюдения в системе координат n точками, координаты которых $(x_1, y_1), \dots, (x_n, y_n)$. Тогда прямая $y = ax + b$, где

$$a = r_{\xi\eta} \sqrt{\frac{D\eta}{D\xi}},$$

$$b = M\eta - aM\xi,$$

будет обладать тем свойством, что сумма квадратов расстояний точек, характеризующих наблюдения, от этой прямой при таком выборе a и b будет в среднем наименьшей (т. е. будет обладать наименьшим средним значением).

Если коэффициент корреляции $r_{\xi\eta}$, а также $M\xi$, $M\eta$, $D\xi$, $D\eta$ неизвестны, то вместо них можно использовать их эмпирические аналоги. Для математических ожиданий и дисперсий это средние арифметические и средние квадратические отклонения:

$$\bar{x} = \frac{1}{n} \sum x_i, \quad \bar{y} = \frac{1}{n} \sum y_i, \quad \bar{s}_x^2 = \frac{1}{n} \sum (x_i - \bar{x})^2,$$

$$\bar{s}_y^2 = \frac{1}{n} \sum (y_i - \bar{y})^2.$$

Эмпирический коэффициент корреляции определяется соотношением

$$\hat{r}_{xy} = \frac{1}{\sqrt{\bar{s}_x^2 \bar{s}_y^2}} \left(\frac{1}{n} \sum x_i y_i - \bar{x} \bar{y} \right).$$

На основании закона больших чисел среднее арифметическое из независимых случайных величин $x_1 y_1, \dots, x_n y_n$ (они одинаково распределены так же, как $\xi\eta$) стремится к математическому ожиданию одной величины, т. е. к $M\xi\eta$. Кроме того, \bar{x} стремится к $M\xi$, а \bar{y} к $M\eta$. Точно так же $\bar{s}_x^2 \rightarrow D\xi$, $\bar{s}_y^2 \rightarrow D\eta$. Поэтому \hat{r}_{xy} стремится при $n \rightarrow \infty$ к $r_{\xi\eta}$. Это и оправдывает такой выбор эмпирического коэффициента корреляции.

Наилучшее представление линейной функцией связи между величинами x_1, \dots, x_n и y_1, \dots, y_n будет, если

$$y_i = ax_i + b + \varepsilon_i, \quad \text{где} \quad a = \hat{r}_{xy} \frac{\bar{s}_y}{\bar{s}_x}, \quad b = \bar{y} - a\bar{x}. \quad (5)$$

Здесь ϵ_i — ошибка представления. При этом оказывается, что при подобном выборе a и b $\sum \epsilon_i^2$ является наименьшей, причем

$$\frac{1}{n} \sum \epsilon_i^2 = \bar{s}_y^2 (1 - \hat{r}_{xy}^2). \quad (6)$$

Чтобы убедиться в этом, достаточно рассмотреть величины $\tilde{\xi}$ и $\tilde{\eta}$, для которых $P \{ \tilde{\xi} = x_i, \tilde{\eta} = y_i \} = \frac{1}{n}$ (для простоты считаем, что все пары (x_i, y_i) различны). Тогда эмпирические характеристики совпадают с точными характеристиками для $\tilde{\xi}$ и $\tilde{\eta}$:

$$\bar{x} = M\tilde{\xi}, \quad \bar{y} = M\tilde{\eta}, \quad \bar{s}_x^2 = D\tilde{\xi}, \quad \bar{s}_y^2 = D\tilde{\eta}, \quad \hat{r}_{xy} = r_{\tilde{\xi}\tilde{\eta}}.$$

Поэтому формулы (5) и (6) вытекают из формул (3) и (4), если в последних заменить ξ и η на $\tilde{\xi}$ и $\tilde{\eta}$.

Задачу о приближенном наилучшем линейном представлении одной величины с помощью другой можно распространить на несколько случайных величин. Пусть имеются случайные величины η и ξ_1, \dots, ξ_k . Мы хотим представить η : помощью линейного выражения

$$a_1 \xi_1 + \dots + a_k \xi_k + b,$$

выбирая a_1, \dots, a_k, b так, чтобы среднее квадрата ошибки представления

$$\epsilon = \eta - (a_1 \xi_1 + \dots + a_k \xi_k + b)$$

было минимальным.

Заметим, что здесь величины ξ_1, \dots, ξ_k необязательно независимы.

Так как

$$M\epsilon^2 = D\epsilon + (M\epsilon)^2$$

и $D\epsilon$ от b не зависят, то всегда, не меняя дисперсию, можно выбрать b так, чтобы $M\epsilon = 0$, полагая

$$b = M\eta - a_1 M\xi_1 - \dots - a_k M\xi_k.$$

При подобном выборе b задача сводится к выбору таких a_1, \dots, a_k , чтобы

$$D\eta = M(\tilde{\eta} - (a_1 \tilde{\xi}_1 + \dots + a_k \tilde{\xi}_k))^2 \quad (7)$$

была минимальной, где $\tilde{\eta} = \eta - M\eta$, $\tilde{\xi}_i = \xi_i - M\xi_i$.

Таким образом, задача может быть легко сведена к случаю, когда математические ожидания величин η и ξ_i равны 0. Тогда b нужно также взять равным нулю, а для нахождения чисел a_i нужно минимизировать выражение (7). Это проще всего сделать, когда величины ξ_1, \dots, ξ_k (а значит, и $\tilde{\xi}_1, \dots, \tilde{\xi}_k$) некоррелированы между собой, т. е. $r_{\xi_i \xi_j} = 0$ при $i \neq j$. Такие величины называются *попарно некоррелированными*.

Нам потребуется следующее свойство.

IV. Если величины ξ_1, \dots, ξ_k попарно некоррелированы, то

$$D(\xi_1 + \dots + \xi_k) = D\xi_1 + \dots + D\xi_k.$$

При подсчете дисперсии можно считать, что величины имеют математическое ожидание, равное нулю, так как $D(\xi_1 + \dots + \xi_k) = D(\tilde{\xi}_1 + \dots + \tilde{\xi}_k)$. При этом предположении

$$\begin{aligned} D(\xi_1 + \dots + \xi_k) &= M(\xi_1 + \dots + \xi_k)^2 = \\ &= M\xi_1^2 + \dots + M\xi_k^2 + M\xi_1\xi_2 + \dots + M\xi_{k-1}\xi_k. \end{aligned}$$

Но

$$M\xi_i\xi_j = r_{\xi_i\xi_j} \sqrt{D\xi_i D\xi_j} = 0 \quad \text{при } i \neq j.$$

Поэтому

$$M(\xi_1 + \dots + \xi_k)^2 = M\xi_1^2 + \dots + M\xi_k^2,$$

это и доказывает свойство IV.

Покажем теперь, как выбрать a_i , минимизирующие правую часть (7), в предположении, что величины ξ_i попарно некоррелированы. Возводя в квадрат и беря математическое ожидание почленно, получим

$$\begin{aligned} D\epsilon &= M\tilde{\eta}^2 - 2M\tilde{\eta}(a_1\tilde{\xi}_1 + \dots + a_k\tilde{\xi}_k) + \\ &+ M(a_1\tilde{\xi}_1 + \dots + a_k\tilde{\xi}_k)^2. \end{aligned}$$

Но $M\tilde{\eta}\tilde{\xi}_i = r_{\tilde{\eta}\tilde{\xi}_i} \sqrt{D\tilde{\eta} D\tilde{\xi}_i}$, а так как величины $a_i\tilde{\xi}_i$ и $a_j\tilde{\xi}_j$ некоррелированы при $i \neq j$, то

$$\begin{aligned} M(a_1\tilde{\xi}_1 + \dots + a_k\tilde{\xi}_k) &= D(a_1\tilde{\xi}_1 + \dots + a_k\tilde{\xi}_k) = \\ &= D(a_1\tilde{\xi}_1) + \dots + D(a_k\tilde{\xi}_k) = a_1^2 D\tilde{\xi}_1^2 + \dots + a_k^2 D\tilde{\xi}_k^2. \end{aligned}$$

Таким образом,

$$D\varepsilon = D\eta - 2r_{\eta\xi_1} \sqrt{D\eta} \sqrt{D\xi_1} a_1 - \dots \\ \dots - 2r_{\eta\xi_k} \sqrt{D\eta} \sqrt{D\xi_k} a_k + a_1^2 D\xi_1 + \dots + a_k^2 D\xi_k.$$

Используя преобразование

$$- 2r_{\eta\xi_i} \sqrt{D\eta} \sqrt{D\xi_i} a_i + D\xi_i a_i^2 = \\ = [\sqrt{D\xi_i} a_i - r_{\eta\xi_i} \sqrt{D\eta}]^2 - r_{\eta\xi_i}^2 D\eta,$$

можем записать

$$D\varepsilon = D\eta (1 - r_{\eta\xi_1}^2 - \dots - r_{\eta\xi_k}^2) + (\sqrt{D\xi_1} a_1 - r_{\eta\xi_1} \sqrt{D\eta})^2 + \\ + \dots + (\sqrt{D\xi_k} a_k - r_{\eta\xi_k} \sqrt{D\eta})^2.$$

Очевидно, что это выражение будет наименьшим, если для всех i

$$(\sqrt{D\xi_i} a_i - r_{\eta\xi_i} \sqrt{D\eta})^2 = 0,$$

т. е.

$$a_i = r_{\eta\xi_i} \sqrt{\frac{D\eta}{D\xi_i}}.$$

Среднее значение квадрата ошибки задается в этом случае выражением

$$D\varepsilon = (1 - r_{\eta\xi_1}^2 - \dots - r_{\eta\xi_k}^2) D\eta. \quad (8)$$

Учитывая значение b , имеем такое представление для η :

$$\eta = \sum_{i=1}^k r_{\eta\xi_i} \sqrt{\frac{D\eta}{D\xi_i}} (\xi_i - M\xi_i) + M\eta + \varepsilon. \quad (9)$$

Формулы (8) и (9) аналогичны формулам (4) и (3).

Оказывается, что задачу о наилучшем линейном представлении всегда можно свести к тому случаю, когда величины ξ_1, \dots, ξ_k некоррелированы. Для этого используется следующий прием, называемый *ортогонализацией случайных величин*.

Построим новые величины ζ_1, \dots, ζ_k , которые так связаны с величинами ξ_1, \dots, ξ_k :

$$\zeta_1 = \xi_1, \quad \zeta_2 = \xi_2 + \alpha_1 \xi_1, \quad \zeta_3 = \xi_3 + \beta_1 \xi_1 + \beta_2 \xi_2, \quad \dots, \quad (10) \\ \zeta_k = \xi_k + \theta_1 \xi_1 + \dots + \theta_{k-1} \xi_{k-1}.$$

Задача состоит в таком выборе коэффициентов $\alpha_1, \beta_1, \beta_2, \dots, \theta_1, \dots, \theta_{k-1}$, чтобы величины ζ_1, \dots, ζ_k уже были некоррелированными. Предположим, что эти коэффициенты можно так выбрать.

Тогда на основании формулы (9)

$$\eta = \sum_{i=1}^k r_{\eta\zeta_i} \sqrt{\frac{D\eta}{D\zeta_i}} (\zeta_i - M\zeta_i) + M\eta + \varepsilon \quad (11)$$

является наилучшим линейным представлением η через ζ_1, \dots, ζ_k .

Выражая ζ_i через ξ_1, \dots, ξ_i по формулам (10), получим линейное представление η через ξ_1, \dots, ξ_k . Оно также будет наилучшим.

Действительно, из системы равенств (10) можно последовательно выразить ξ_i через ζ_i :

$$\begin{aligned} \xi_1 &= \zeta_1, & \xi_2 &= \zeta_2 - \alpha\xi_1 \Rightarrow \zeta_2 - \alpha\zeta_1, \\ \xi_3 &= \zeta_3 - \beta_1\zeta_1 - \beta_2(\zeta_2 - \alpha\zeta_1) \end{aligned}$$

и т. д. Поэтому каждое линейное представление величины η через ξ_1, \dots, ξ_k можно переписать как линейное представление через величины ζ_1, \dots, ζ_k , при этом ошибка представления окажется неизменной. Значит, ошибки представления (в том числе и ошибка с минимальным средним значением квадрата) через величины ξ_1, \dots, ξ_k и величины ζ_1, \dots, ζ_k одинаковы, и, взяв представление (11) и заменяя в нем величины ζ_i через ξ_1, \dots, ξ_i , получим наилучшее представление η через ξ_1, \dots, ξ_k .

Будем строить величины ζ_1, \dots, ζ_k последовательно. Как уже отмечалось выше, задачу о линейном представлении нужно уметь решать лишь для величин с математическим ожиданием 0. Будем считать, что это предположение для ξ_i выполнено, тогда и ζ_i , выражаемые через ξ_i по формулам (10), будут иметь математическое ожидание 0.

Величину ζ_2 представим в виде

$$\zeta_2 = \xi_2 + \alpha_1\xi_1.$$

Чтобы величины ζ_1 и ζ_2 были некоррелированными, нужно, чтобы $M\zeta_1\zeta_2 = 0$, т. е.

$$0 = M(\xi_2 + \alpha_1\xi_1)\zeta_1 + r_{\xi_2\zeta_1} \sqrt{D\xi_2 D\zeta_1} + \alpha_1 D\zeta_1,$$

$$\alpha_1 = -r_{\xi_2\zeta_1} \sqrt{\frac{D\xi_2}{D\zeta_1}},$$

$$\zeta_2 = \xi_2 - r_{\xi_2\zeta_1} \sqrt{\frac{D\xi_2}{D\zeta_1}} \zeta_1. \quad (12)$$

Заметим, что поскольку $\zeta_1 = \xi_1$, то

$$r_{\xi_2 \xi_1} = r_{\xi_2 \xi_1}, \quad D\zeta_1 = D\xi_1$$

— известные величины.

Исходя из формулы (12), можем найти $D\zeta_2$ и $r_{\xi_1 \xi_2}$ для $l > 2$:

$$\begin{aligned} D\zeta_2 &= M \left(\xi_2 - r_{\xi_2 \xi_1} \sqrt{\frac{D\xi_2}{D\xi_1}} \xi_1 \right)^2 = D\xi_2 - 2M\xi_2 \xi_1 \times \\ &\times r_{\xi_1 \xi_2} \sqrt{\frac{D\xi_2}{D\xi_1}} + r_{\xi_1 \xi_2}^2 \frac{D\xi_2}{D\xi_1} D\xi_1 = D\xi_2 (1 - r_{\xi_1 \xi_2}^2), \\ r_{\xi_1 \xi_2} \sqrt{D\xi_1 D\zeta_2} &= M\xi_1 \xi_2 - r_{\xi_1 \xi_2} \sqrt{\frac{D\xi_2}{D\xi_1}} r_{\xi_1 \xi_1}. \end{aligned}$$

Пусть уже построены величины ζ_1, \dots, ζ_i . Покажем, как строить величину ζ_{i+1} . Так как ξ_1, \dots, ξ_i выражаются через ζ_1, \dots, ζ_i , то положим

$$\zeta_{i+1} = \xi_{i+1} + \lambda_1 \zeta_1 + \dots + \lambda_i \zeta_i. \quad (13)$$

Должны выполняться соотношения $M\zeta_{i+1}\zeta_j = 0$ для всех $j = 1, \dots, i$. Поэтому

$$\begin{aligned} 0 &= M\xi_{i+1}\zeta_j + M\zeta_j (\lambda_1 \zeta_1 + \dots + \lambda_i \zeta_i) = \\ &= M\xi_{i+1}\zeta_j + \lambda_j D\zeta_j. \end{aligned}$$

Значит, нужно выбрать $\lambda_j = -\frac{M\xi_{i+1}\zeta_j}{D\zeta_j}$.

Так, последовательно определяются ζ_i . Формула (13) позволяет определить $M\xi_j \zeta_{i+1}$ и $D\zeta_{i+1}$. Может оказаться, что некоторое ζ_{i+1} равно 0 ($D\zeta_{i+1} = 0$). Тогда ξ_{i+1} выражается линейно через ξ_1, \dots, ξ_i (через них выражаются ζ_1, \dots, ζ_i), и поэтому в линейном представлении можно исключить ξ_{i+1} .

Рассмотрим простой пример. Пусть величина ξ принимает значения $-2, -1, 0, 1, 2$ с вероятностями $\frac{1}{5}$, $\xi_1 = \xi$, $\xi_2 = \xi^2$, $\xi_3 = \xi^3$, $\eta = 1$, если $\xi > 0$, $\eta = 0$, если $\xi = 0$, $\eta = -1$, если $\xi < 0$. Найдем наилучшее представление величины η через ξ_1, ξ_2, ξ_3 , т. е. представление вида

$$\eta = a\xi^3 + b\xi^2 + c\xi + d.$$

Легко видеть, что $M\eta = M\xi = M\xi^3 = 0$, $M\xi^2 = D\xi = \frac{4+1+1+4}{5} = 2$.

Ортогонализуем величины ξ , $\xi^2 - z$, ξ^3 . Имеем

$$\zeta_1 = \xi, \quad \zeta_2 = \xi^2 - 2 + \alpha\xi, \quad M\zeta_1\zeta_2 = M\xi(\xi^2 - 2 + \alpha\xi) =$$

$$= M\xi^3 - 2M\xi + \alpha M\xi^2 = 2\alpha.$$

Значит, нужно выбрать $\alpha = 0$. Далее,

$$\zeta_3 = \xi^3 + \beta_1\zeta_1 + \beta_2\zeta_2, \quad M\zeta_3\zeta_1 = M\xi(\xi^3 + \beta_1\xi + \beta_2(\xi^2 - 2)) =$$

$$= M\xi^4 + \beta_1 M\xi^2 = M\xi^4 + 2\beta_1.$$

Имеем

$$M\xi^4 = \frac{1}{5}(1 + 16 + 1 + 16) = \frac{34}{5}, \quad \beta_1 = -3,4;$$

$$0 = M(\xi^2 - 2)(\xi^3 + \beta_1\xi + \beta_2(\xi^2 - 2)) =$$

$$= M\xi^5 - 2M\xi^3 + \beta_2 M(\xi^2 - 2)^2.$$

Так как $M\xi^5 = (-1 - 32 + 1 + 32) = 0$, то $\beta_2 = 0$.
Итак, $\zeta_1 = \xi$, $\zeta_2 = \xi^2 - 2$, $\zeta_3 = \xi^3 - \frac{34}{10}\xi$.

На основании формулы (11)

$$\eta = \frac{M\eta\zeta_1}{D\zeta_1} \zeta_1 + \frac{M\eta\zeta_2}{D\zeta_2} \zeta_2 + \frac{M\eta\zeta_3}{D\zeta_3} \zeta_3 + \varepsilon.$$

Далее,

$$M\eta\zeta_1 = M\eta\xi = \frac{1}{5}(1 + 2 + 1 + 2) = \frac{6}{5},$$

$$M\eta\zeta_2 = M\eta(\xi^2 - 2) = 2M\eta = M\eta\xi^2 = \frac{1}{5}(-4 - 1 + 1 + 4) = 0,$$

$$M\eta\zeta_3 = M\eta\xi^3 - \frac{34}{10}M\eta\xi =$$

$$= \frac{1}{5}(8 + 1 + 1 + 8) - \frac{34}{10} \cdot \frac{6}{5} = -0,48,$$

$$D\zeta_1 = 2, \quad D\zeta_2 = \frac{1}{5}(4 + 1 + 4 + 1 + 4) = \frac{14}{5},$$

$$D\zeta_3 = M\xi^6 - \frac{34}{5}M\xi^4 + \left(\frac{34}{10}\right)^2 \cdot M\xi^2 =$$

$$= 26 - \frac{34^2}{50} = \frac{144}{50} = 2,88.$$

Окончательно получаем такое представление:

$$\begin{aligned}\eta &= 0,6\xi - \frac{0,48}{2,88} (\xi^3 - 3,4\xi) + \varepsilon = \\ &= 0,6\xi - \frac{1}{6} \xi^3 + \frac{3,4}{6} \xi + \varepsilon = \frac{7}{6} \xi - \frac{1}{6} \xi^3 + \varepsilon.\end{aligned}$$

Ошибка прогноза равна

$$D\eta \left(1 - \frac{(M_{\xi_1}^r \eta)^2}{D\eta D_{\xi_1}^r} - \frac{(M_{\xi_3}^r \eta)^2}{D\eta D_{\xi_3}^r} \right) = 0.$$

III. СЛУЧАЙНЫЙ ПРОЦЕСС

§ 22. Время ожидания

На практике мы часто имеем дело со случайными событиями, происходящими во времени. К ним относятся, например, регистрация счетчиком космической частицы, солнечные вспышки, переход атома из возбужденного состояния в устойчивое, деление ядра радиоактивного элемента.

Каждое из данных событий можно характеризовать временем от момента начала наблюдения до момента, когда событие произойдет. Это время является случайной величиной. Назовем его *временем ожидания случайного события*.

Изучим при определенных условиях распределение времени ожидания.

Предположим, что условия эксперимента, в котором наблюдается появление некоторого события A , не меняются со временем по крайней мере до тех пор, пока событие A не произойдет. Так, например, пока состояние возбужденного атома не меняется, не изменены и условия эксперимента, в котором наблюдается переход атома из возбужденного состояния, точно так не меняются условия, в которых находится ядро радиоактивного элемента, пока не произойдет ядерное деление. В таком случае вероятность того, что событие A произойдет на отрезке $[t, t + s]$ при условии, что оно не произошло на отрезке $[0, t]$, такова, как и вероятность того, что событие произойдет на отрезке $[0, s]$, поскольку условия эксперимента в момент t и в момент 0 в обоих случаях одинаковы.

Обозначим через $A [t, t + s]$ событие, заключающееся в том, что A не произойдет на $[t, t + s]$.

Тогда

$$\begin{aligned} P \{ \bar{A} [0, t + s] \} &= P \{ \bar{A} [0, t] \} P \{ \bar{A} [t, t + s] / A [0, t] \} = \\ &= P \{ \bar{A} [0, t] \} P \{ \bar{A} [0, s] \}. \end{aligned}$$

выразим время до ожидания события A через τ . Очевидно,

$$P\{\bar{A}[0, t]\} = P\{\tau > t\}.$$

Итак, для вероятности $P\{\tau > t\}$ справедливо соотношение

$$P\{\tau > t + s\} = P\{\tau > t\} P\{\tau > s\}. \quad (1)$$

Будем рассматривать лишь те события, которые не могут произойти мгновенно, т. е. такие, для которых

$$P\{\tau > 0\} = 1.$$

Так как при $t_n \rightarrow 0$

$$\{\tau > 0\} = \cup \{\tau_n > t_n\},$$

то $P\{\tau > t_n\} \rightarrow 1$ при $n \rightarrow \infty$. Значит, для достаточно малых δ $P\{\tau > \delta\} > 0$, поэтому для всех n

$$P\{\tau > n\delta\} = (P\{\tau > \delta\})^n > 0$$

мы использовали равенство (1)). Таким образом, для всех $\delta > 0$ определена функция

$$g(t) = \ln P\{\tau > t\}$$

\ln — натуральный логарифм, т. е. логарифм при основании e).

Логарифмируя равенство (1), получим

$$g(t + s) = g(t) + g(s).$$

Из этого равенства вытекает, что для всех n и t

$$g(nt) = ng(t), \quad g(t) = \frac{1}{n} g(nt), \quad g\left(\frac{t}{n}\right) = \frac{1}{n} g(t).$$

Поэтому

$$g\left(\frac{m}{n} t\right) = mg\left(\frac{t}{n}\right) = \frac{m}{n} g(t).$$

значит, для всякого рационального t

$$g(t) = tg(1). \quad (2)$$

Используя то, что при $\delta \rightarrow 0$ $g(\delta) = \ln P\{\tau > \delta\} \rightarrow +\ln P\{\tau > 0\} = 0$, можно убедиться, что формула (2) справедлива для всех t . Положим $g(1) = -\lambda$, где $\lambda > 0$ так как $g(t)$ есть логарифм числа, меньшего 1, то ($t < 0$). Тогда

$$P\{\tau > t\} = e^{-\lambda t}. \quad (3)$$

Это означает, что время ожидания имеет показательное распределение. λ называется *параметром показательного распределения*.

Рассмотрим пример с радиоактивным распадом. $1 - e^{-\lambda t}$ — есть вероятность того, что ядро за время t распадается. Если N — число ядер, N_t — число ядер, распавшихся за время t , то $\frac{N_t}{N}$ — частота события «за время t произошел радиоактивный распад» в N независимых испытаниях (при естественной радиоактивности превращения ядер происходят независимо). На основании закона больших чисел (N обычно число порядка 10^{20})

$$\frac{N_t}{N} = 1 - e^{-\lambda t}, \quad N_t = N(1 - e^{-\lambda t}).$$

В физике скорость распада характеризуют периодом полураспада. Это такое время T , по истечении которого распадается половина наличного вещества. Значит,

$$e^{-\lambda T} = \frac{1}{2}, \quad \lambda T = \ln 2, \quad T = \frac{\ln 2}{\lambda} = \frac{0,693}{\lambda}.$$

Установим вероятностный смысл параметра λ показательного распределения. Для этого нам для показательной функции потребуется следующее соотношение:

$$\lim_{t \rightarrow 0} \frac{e^t - 1}{t} = 1.$$

Используя его, можем записать

$$P\{\tau < t\} = 1 - e^{-\lambda t} \sim \lambda t \quad (4)$$

при малых t . Таким образом, вероятность того, что событие произойдет на промежутке $[t, t + h]$, при условии, что оно не произошло до момента t при малых h , равна примерно λh , т. е. вероятность появления события A за малый промежуток времени примерно пропорциональна длине промежутка. λ — коэффициент пропорциональности. Это оправдывает название λ как интенсивности появления события A .

Вычислим среднее значение времени ожидания события A . Вероятность того, что τ попадет в интервал $[kh, kh + h]$, есть

$$e^{-\lambda kh} - e^{-\lambda(kh+h)} = e^{-\lambda kh} [1 - e^{-\lambda h}].$$

Значит,

$$M\tau = \lim_{h \rightarrow 0} \sum_{k=1}^{\infty} k h e^{-\lambda k h} [1 - e^{-\lambda h}] = \lim_{h \rightarrow 0} \frac{h}{1 - e^{-\lambda h}} = \frac{1}{\lambda}$$

(мы использовали вычисление среднего для величины, имеющей геометрическое распределение, в § 15, пример 3, а также соотношение (4)). Тем самым установлено, что среднее время ожидания события равно обратной величине интенсивности появления события.

Пусть A_1, A_2, \dots, A_n — независимые события, имеющие интенсивности $\lambda_1, \lambda_2, \dots, \lambda_n$. Рассмотрим событие $A = A_1 + \dots + A_n$. Оно происходит, если осуществляется хотя бы одно из событий A_1, \dots, A_n .

Обозначим через τ_k время ожидания события A_k . Если на промежутке $[0, t]$ не произошло событие A , то это означает, что не произошло ни одно из событий A_1, \dots, A_n , т. е. осуществилось событие

$$\bar{A}_1[0, t] \cdot \bar{A}_2[0, t] \dots \bar{A}_n[0, t].$$

По предположению эти события независимы. Поэтому

$$P(\bar{A}[0, t]) = P(\bar{A}_1[0, t]) \dots P(\bar{A}_n[0, t]) = \\ = e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} \dots e^{-\lambda_n t} = e^{-(\lambda_1 + \dots + \lambda_n)t}.$$

Значит, обозначив через τ время ожидания события A , будем иметь

$$P\{\tau > t\} = e^{-(\lambda_1 + \dots + \lambda_n)t}.$$

Из этой формулы вытекает, что интенсивность суммы независимых событий равняется сумме интенсивностей этих событий.

Рассмотрим вероятность того, что из событий A_1, \dots, A_n раньше всех произойдет A_k . Эту вероятность можно определить как вероятность совместного осуществления событий

$$\{\tau_1 > \tau_k\} \cdot \{\tau_2 > \tau_k\} \dots \{\tau_n > \tau_k\}$$

(среди них, естественно, отсутствует невозможное событие $\{\tau_k > \tau_k\}$).

Пусть $\tau_k^{(h)} = mh$ при $mh \leq \tau_k < (m+1)h$. Тогда $0 \leq \tau_k - \tau_k^{(h)} < h$, значит, при $h \rightarrow 0$ $\tau_k^{(h)} \rightarrow \tau_k$. Поэтому

$$\lim_{h \rightarrow 0} P(\{\tau_1 > \tau_k^{(h)}\} \dots \{\tau_n > \tau_k^{(h)}\}) = P(\{\tau_1 > \tau_k\} \dots \{\tau_n > \tau_k\}).$$

На основании формулы полной вероятности можем записать

$$P \{ \tau_1 > \tau_k^{(h)} \} \dots \{ \tau_n > \tau_k^{(h)} \} = \\ = \sum_m P \{ (\tau_1 > \tau_k^{(h)}) \dots (\tau_n > \tau_k^{(h)}) / \tau_k^{(h)} = mh \} P \{ \tau_k^{(h)} = mh \}.$$

Используя независимость случайных величин $\tau_1, \dots, \dots, \tau_n$, получаем

$$P \{ (\tau_1 > mh) \dots (\tau_n > mh) / (mh \leq \tau_k < mh + h) \} P \{ mh \leq \\ \leq \tau_k < mh + h \} = e^{-\lambda_1 mh} \dots e^{-\lambda_n mh} (e^{-\lambda_k mh} - e^{-\lambda_k (mh+h)})$$

(в произведении перед скобкой отсутствует член $e^{-\lambda_k mh}$).

Таким образом,

$$P \{ (\tau_1 > \tau_k^{(h)}) \dots (\tau_n > \tau_k^{(h)}) \} = \\ = \sum_{m=0}^{\infty} e^{-m(\lambda_1 + \dots + \lambda_n)h} (1 - e^{-\lambda_k h}) = \frac{1 - e^{-\lambda_k h}}{1 - e^{-(\lambda_1 + \dots + \lambda_n)h}}$$

(мы воспользовались тем, что последнее выражение есть сумма членов геометрической прогрессии).

Так как при $h \rightarrow 0$

$$1 - e^{-\lambda_k h} \sim \lambda_k h, \quad 1 - e^{-(\lambda_1 + \dots + \lambda_n)h} \sim (\lambda_1 + \dots + \lambda_n) h,$$

то, переходя к пределу при $h \rightarrow 0$, находим

$$P \{ (\tau_1 > \tau_k) \dots (\tau_n > \tau_k) \} = \frac{\lambda_k}{\lambda_1 + \dots + \lambda_n}.$$

Это можно записать еще так:

$$P(A_k/A) = \frac{\lambda_k}{\lambda_1 + \dots + \lambda_n}.$$

Если A_1, \dots, A_n — независимые события с интенсивностями $\lambda_1, \dots, \lambda_n$, то вероятность того, что произошло событие A_k при условии, что произошло хотя бы одно из этих событий, равна отношению интенсивности события A_k к интенсивности суммы событий.

§ 23. Процесс Пуассона

В § 7 мы уже встречались с распределением Пуассона как законом редких событий. События, наступление которых нужно ожидать некоторое время, рассматривавшиеся в § 22, можно трактовать как редкие, поскольку они происходят с как угодно малыми вероятностями за достаточно

малые промежутки времени. Например, событие с интенсивностью 1000 1/с за промежуток времени 1/1000000 с происходит с вероятностью 0,001. А для многих явлений указанный промежуток времени весьма большой; свет за это время пролетит 30 000 см, это порядка 10^{12} атомных диаметров.

Во многих случаях то, что событие A произошло, не влияет (или влияет незначительно) на условия протекания эксперимента. Так что можно после появления события A снова ожидать наступления этого события, причем интенсивность его не меняется.

Число появлений события A за время $[0, t]$ есть случайная величина. Обозначим ее через $\nu(t)$. Мы имеем в данном случае не одну, а совокупность случайных величин: каждому $t \geq 0$ отвечает случайная величина. Другими словами, у нас определена функция t , значениями которой являются случайные величины. Аргумент t есть время от начала наблюдения. Функции времени, принимающие случайные значения, называются *случайными процессами*. Таким образом, $\nu(t)$ есть случайный процесс.

Приведем схему, в которой событие A может происходить многократно с мало меняющейся интенсивностью. Пусть A_1, \dots, A_n — независимые события, происходящие во времени с равными интенсивностями λ . Каждое из событий может произойти только раз (например, A_k — распад k -го ядра из совокупности n радиоактивных ядер). Обозначим через A сумму событий A_k . Интенсивность события A равна $n\lambda$. После того как наступило A , осталось $n - 1$ событие, которое может произойти после этого (если распалось одно ядро, осталось $n - 1$ ядер, которые способны распасться). Поэтому интенсивность события A (оно происходит, если происходит одно из оставшихся событий) равна $(n - 1)\lambda$. После того как событие A наступило m раз, его интенсивность равна $(n - m)\lambda$. Отношение этой интенсивности к начальной интенсивности равно $1 - \frac{m}{n}$.

Если n достаточно большое, то, пока число появлений события A мало по сравнению с n , можем считать, что интенсивность события A практически неизменна. Так будет, например, при рассмотрении радиоактивных веществ на промежутках времени, малых по сравнению с периодом полураспада. Практически не меняется интенсивность космического излучения (за доступные промежутки наблюдения).

Обратимся к распределению процесса $v(t)$. Обозначим через τ_1 момент наступления события A в первый раз. Поскольку по предположению после первого наступления A условия эксперимента не меняются, то время τ_2 , которое потребуется от момента первого наступления A до того, как оно произойдет во второй раз, имеет такое же распределение, как и τ_1 , и не зависит от τ_1 .

Поясним последнее обстоятельство на дискретной схеме. Пусть наблюдения производятся через отрезки времени h , событие впервые произошло на промежутке $[mh, mh + h]$. Тогда вероятность того, что событие A произошло на отрезке $[mh + h, mh + h + t]$, совпадает с вероятностью того, что A произойдет на отрезке $[0, t]$. Она не зависит от m , которое определяется значением τ_1 .

Точно так же, определяя τ_k как промежуток времени между $k-1$ -м и k -м моментом наступления события A , убеждаемся, что величины $\tau_1, \dots, \tau_k, \dots$ независимы, одинаково распределены. Распределение каждой из величин показательное; если интенсивность события A равна λ , то

$$P\{\tau_k < t\} = \begin{cases} 0, & t \leq 0, \\ 1 - e^{-\lambda t}, & t > 0. \end{cases}$$

Если за время t событие A произошло ровно k раз, то происходит событие

$$\tau_1 + \dots + \tau_k < t < \tau_1 + \dots + \tau_k + \tau_{k+1}$$

($\tau_1 + \dots + \tau_k$ — это момент, когда событие A произошло в k -й раз, что случилось до момента t , в $k+1$ -й раз оно произошло в момент $\tau_1 + \dots + \tau_{k+1}$, он должен быть после момента t). Таким образом,

$$P\{v(t) = k\} = P\{\tau_1 + \dots + \tau_k < t < \tau_1 + \dots + \tau_{k+1}\}, \\ (k \geq 1). \quad (1)$$

Так как при $v(t) = 0$ до момента t не произошло ни одного события, то

$$P\{v(t) = 0\} = P\{\tau_1 > t\} = e^{-\lambda t}.$$

Будем вычислять вероятности (1) последовательно. При $k = 1$ имеем

$$P\{v(t) = 1\} = P\{\tau_1 < t < \tau_1 + \tau_2\} = \lim_{h \rightarrow 0} P\{\tau_1^{(h)} < t < \tau_1^{(h)} + \\ + h + \tau_2\},$$

где $\tau_1^{(m)} = mh$ при $mh \leq \tau_1 < mh + h$. По формуле полной вероятности

$$\begin{aligned}
 P\{\tau_1^{(h)} < t < t_1^{(h)} + h + \tau_2\} &= \sum_{m=0}^{\infty} P\{\tau_1^{(h)} < t < \tau_1^{(h)} + \\
 &+ h + \tau_2 / \tau_1^{(h)} = mh\} P\{\tau_1^{(h)} = mh\} = \\
 &= \sum_{m=0}^{\infty} P\{mh < t < mh + h + \tau_2\} P\{\tau_1^{(h)} = mh\}.
 \end{aligned}$$

Очевидно, суммирование справа происходит лишь по тем m , для которых $mh < t$. При выполнении этого условия отдельное слагаемое преобразуется так:

$$\begin{aligned}
 &P\{\tau_2 > t - mh - h\} P\{mh \leq \tau_1 < mh + h\} = \\
 &= e^{-\lambda(t-mh-h)} (e^{-\lambda(mh+h)} - e^{-\lambda mh}) = e^{-\lambda t} (e^{\lambda h} - 1).
 \end{aligned}$$

Таким образом, все слагаемые одинаковы, а их число — наибольшее целое число, не превосходящее t/h , оно эквивалентно t/h . Поэтому

$$\begin{aligned}
 P\{v(t) = 1\} &= \lim_{h \rightarrow 0} \frac{t}{h} (e^{\lambda h} - 1) e^{-\lambda t} = \\
 &= \lambda t e^{-\lambda t} \lim_{h \rightarrow 0} \frac{e^{\lambda h} - 1}{\lambda h}, \quad (2) \\
 P\{v(t) = 1\} &= \lambda t e^{-\lambda t}.
 \end{aligned}$$

Для вычисления последующих вероятностей установим соотношение, выражающее $P\{v(t) = k + 1\}$ через $P\{v(t) = k\}$. Если $\tau_2 > h$, то при некотором m выполнено событие $\{mh \leq \tau_1 < mh + h, \tau_1 + \tau_2 > mh + h\}$. Поэтому

$$\begin{aligned}
 P(\{v(t) = k + 1\} \cdot \{\tau_2 > h\}) &= \sum_m P(\{v(t) = k + 1\} \times \\
 \times \{\tau_2 > h\} \cdot \{mh \leq \tau_1 < mh + h\}) &\leq \sum_m P(\{v(t) = k + 1\} \times \\
 \times \{mh \leq \tau_1 < mh + h\} \cdot \{\tau_1 + \tau_2 > mh + h\}) &\leq \\
 &\leq P\{v(t) = k + 1\},
 \end{aligned}$$

сумма берется по тем m , для которых $mh + h < t$. Заметим, что

$$\begin{aligned}
 &P(\{v(t) = k + 1\} \cdot \{mh \leq \tau_1 < mh + h\} \times \\
 \times \{\tau_1 + \tau_2 > mh + h\}) &= P(\{v(mh) = 0\} \cdot \{v(mh + h) = \\
 &= v(mh) = 1\} \cdot \{v(t) - v(mh + h) = k\}).
 \end{aligned}$$

Поскольку интенсивность события A не зависит по предположению от того, происходило ли оно раньше и сколько раз, то событие

$$\{v(t) - v(mh + h) = k\},$$

совпадающее с тем, что на $[mh + h, t]$ событие A произошло k раз, не зависит от событий $\{v(mh) = 0\}$ (на $[0, mh]$ событие A не произошло ни разу) и $\{v(mh + h) - v(mh) = 1\}$ (на $[mh, mh + h]$ событие A произошло 1 раз).

Последние два события также независимы. Поэтому

$$P(\{v(t) = k + 1\} \cdot \{mh \leq \tau_1 < mh + h\} \cdot \{\tau_1 + \tau_2 > mh + h\}) = e^{-\lambda mh} \cdot \lambda h e^{-\lambda h} \cdot P\{v(t) - v(mh + h) = k\}.$$

Но вероятность того, что на данном промежутке произойдет k раз событие A , зависит лишь от длины промежутка, так что

$$P\{v(t) - v(mh + h) = k\} = P\{v(t - mh - h) = k\}.$$

Значит,

$$P(\{v(t) = k + 1\} \cdot \{\tau_2 > h\}) \leq \sum_{mh+h < t} \lambda h e^{-\lambda(m+1)h} \times \\ \times P\{v(t - mh - h) = k\} \leq P\{v(t) = k + 1\}.$$

Переходя к пределу при $h \rightarrow 0$, получаем

$$\lim_{h \rightarrow 0} P(\{v(t) = k + 1\} \cdot \{\tau_2 > h\}) \leq \\ \leq \lim_{h \rightarrow 0} \sum_{mh < t} \lambda h e^{-\lambda mh} P\{v(t - mh) = k\} \leq P\{v(t) = k + 1\}.$$

Поскольку левая часть неравенства совпадает с $P\{v(t) = k + 1\}$, то

$$P\{v(t) = k + 1\} = \lim_{h \rightarrow 0} \sum_{mh < t} \lambda h e^{-\lambda mh} P\{v(t - mh) = k\}. \quad (3)$$

Полагая $k = 1$, будем иметь

$$P\{v(t) = 2\} = \lim_{h \rightarrow 0} \sum_{mh < t} \lambda h e^{-\lambda mh} \lambda (t - mh) e^{-\lambda(t-mh)} = \\ = \lim_{h \rightarrow 0} \lambda^2 e^{-\lambda t} h \sum_{mh < t} (t - mh).$$

Пусть $h = \frac{t}{n}$, $t = nh$. Тогда

$$\sum_{mh < t} = h \sum_{m < n} (n - m) = h \frac{n(n-1)}{2}.$$

ледовательно,

$$\lim_{h \rightarrow 0} h \sum_{mh < t} (t - mh) = \lim_{n \rightarrow \infty} \frac{t^2}{n^2} \cdot \frac{n(n-1)}{2} = \frac{t^2}{2},$$

е.

$$P \{v(t) = 2\} = \frac{(\lambda t)^2}{2} e^{-\lambda t}.$$

Установим по индукции, что

$$P \{v(t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \quad (4)$$

Для $k = 0, 1, 2$ эта формула уже доказана. Предположим, что она верна при некотором k , и покажем, что она верна и для $k + 1$. Используя формулу (3) и предположение о виде вероятности $P \{v(t) = k\}$, точно так, как при выводе формулы для $P \{v(t) = 2\}$, можем записать

$$\begin{aligned} P \{v(t) = k + 1\} &= \lim_{n \rightarrow \infty} \sum_{m < n} \frac{\lambda t}{n} e^{-\frac{\lambda m t}{n}} \frac{\left(t - \frac{m}{n} t\right)^k}{k!} e^{-\lambda \left(t - \frac{m}{n} t\right)} = \\ &= \frac{(\lambda t)^{k+1}}{k!} e^{-\lambda t} \lim_{n \rightarrow \infty} \frac{1}{n^{k+1}} \sum_{m < n} (n - m)^k = \\ &= \frac{(\lambda t)^{k+1}}{k!} e^{-\lambda t} \lim_{n \rightarrow \infty} \frac{1}{n^{k+1}} \sum_{m < n} m^k. \end{aligned}$$

Если мы покажем, что

$$\lim_{n \rightarrow \infty} \frac{1}{n^{k+1}} \sum_{m < n} m^k = \frac{1}{k+1}, \quad (5)$$

тем самым установим равенство

$$P \{v(t) = k + 1\} = \frac{(\lambda t)^{k+1}}{k!(k+1)} e^{-\lambda t} = \frac{(\lambda t)^{k+1}}{(k+1)!} e^{-\lambda t},$$

значит, и формулу (4).

Для доказательства (5) используем неравенство

$$m^k < \frac{(m+1)^{k+1} - m^{k+1}}{k+1} < (m+1)^k, \quad (6)$$

получающее из соотношения

$$(m+1)^{k+1} - m^{k+1} = (m+1)^k + (m+1)^{k-1} m + \dots + m^k$$

правой части $k+1$ слагаемых, из них k больше m^k и

k меньше $(m+1)^k$. Из неравенства (6) вытекает, что $1^k + 2^k + \dots + (n-1)^k < \frac{1}{k+1} (2^{k+1} - 1^{k+1} + 3^{k+1} - 2^{k+1} + \dots + n^{k+1} - (n-1)^{k+1}) < 2^k + 3^k + \dots + n^k$,

или

$$0 < \frac{n^{k+1} - 1}{k+1} - (1^k + 2^k + \dots + (n-1)^k) < n^k - 1.$$

Разделив это неравенство на n^{k+1} , получим

$$0 < \frac{1}{k+1} - \frac{1}{(k+1)n^{k+1}} - \frac{1}{n^{k+1}} (1^k + 2^k + \dots + (n-1)^k) < \frac{1}{n} - \frac{1}{n^{k+1}},$$

откуда

$$0 < \frac{1}{k+1} - \frac{1}{n^{k+1}} (1^k + 2^k + \dots + (n-1)^k) < \frac{1}{n}.$$

Переходя к пределу при $n \rightarrow \infty$, убеждаемся, что

$$\lim_{n \rightarrow \infty} \left(\frac{1}{k+1} - \frac{1}{n^{k+1}} (1^k + 2^k + \dots + (n-1)^k) \right) = 0.$$

Отсюда и вытекает соотношение (5). Формула (4) доказана.

Итак, если интенсивность появления события A постоянна и равна λ , то число появлений за время t событий A имеет распределение Пуассона с параметром λt . При этом и $Mv(t) = Dv(t) = \lambda t$. Про события, которые появляются во времени с постоянной интенсивностью, говорят, что они образуют *пуассоновский поток событий*. Телефонные вызовы, поступающие на станцию, регистрация космических частиц, солнечные вспышки, распады радиоактивных ядер — все они образуют пуассоновские потоки событий.

§ 24. Процессы с конечным множеством состояний

Простейшим примером системы с двумя состояниями может служить нуклон. Он имеет два состояния: нейтрон и протон. Состояние нейтрон неустойчиво, через некоторое случайное время он излучает электрон и превращается в протон. При отсутствии внешних воздействий протон стабильно.

лен, но при облучении его электронами он может поглотить один из них и превратиться обратно в нейтрон. Если следить за эволюцией облучаемого таким образом нуклона, мы будем наблюдать превращения нейтрона в протон и обратно.

Другой пример системы с двумя состояниями можно привести из менее абстрактной сферы. Рассмотрим работу автоматического станка. Через определенное время он выходит из строя, после чего его надо наладить. На наладку требуется какое-то, вообще говоря, случайное время, зависящее от характера разладки (поломки). Станок, таким образом, может находиться в двух состояниях: работы и наладки. Переходы из состояния в состояние происходят случайным образом.

Пусть один рабочий-наладчик обслуживает несколько автоматических станков. Если один из станков выходит из строя, он налаживает его, пока не пустит в ход. За это время могут выйти из строя другие станки. Тогда после окончания работы над первым станком он переходит к другим. Если все станки однотипные (а так чаще всего и бывает на производстве), то состояние системы определяется числом работающих станков (важно, не какие станки вышли из строя, а сколько их). Здесь мы уже имеем систему с несколькими состояниями, которая в случайные моменты переходит из одного состояния в другое.

Приведем еще пример из физики системы со многими состояниями. Рассмотрим атом водорода. Он имеет один электрон, который может находиться на различных энергетических уровнях. В устойчивом состоянии атома электрон занимает уровень с наименьшей энергией. Под влиянием внешних воздействий (например, излучения) электрон может перейти на более высокий энергетический уровень, тогда атом становится возбужденным. Из этого состояния он самопроизвольно переходит в состояния с более низкими энергетическими уровнями электрона, излучая при этом квант света. Переходы как на более высокие энергетические уровни, так и на более низкие происходят случайным образом.

Для описания эволюции такого рода систем и служат процессы с конечным множеством состояний. Пусть система может находиться в одном из состояний E_1, \dots, E_n . Характер поведения системы зависит только от того, в каком состоянии она находится. (Если на систему воздей-

ствуют какие-то внешние влияния, то они считаются неизменными.) Покажем, как можно использовать результаты § 22 для излучения данной системы.

Предположим, что система находится в состоянии E_k . Тогда в течение времени возможны такие события: $E_k \rightarrow E_1, \dots, E_k \rightarrow E_n$ ($E_k \rightarrow E_i$ обозначает событие, заключающееся в том, что система перейдет из состояния E_k в E_i). Среди всех перечисленных переходов могут оказаться и невозможные. Но если система когда-либо покинет состояние E_k , то хотя бы один переход осуществим. Такое состояние, из которого система никогда не выйдет, называется *поглощающим*. Для системы, состоящей из нуклона, подобным будет состояние протон, для атома — устойчивое состояние (в обоих случаях при отсутствии внешних воздействий).

Пусть A_k — событие, заключающееся в том, что система вышла из состояния E_k , а A_{ki} ($k \neq i$) — что она перешла в состояние E_i . Тогда $A_k = A_{k1} + \dots + A_{kn}$ (событие A_{kk} не имеет смысла и в сумме отсутствует). Обозначим через a_{ki} интенсивность события A_{ki} (в силу предположения, что пока система находится в состоянии E_k , условия, в которых она находится, остаются неизменными). Из результатов § 22 вытекает, что интенсивность события A_k равна сумме $a_{k1} + \dots + a_{kn}$. Обозначим эту интенсивность λ_k :

$$\lambda_k = \sum_i a_{ki}.$$

В § 22 было установлено, что величина $\frac{a_{ki}}{\lambda_k} = \pi_{ki}$ есть условная вероятность того, что произошло событие A_{ki} при условии, что произошло A_k , т. е. это вероятность перейти системе из состояния E_k в состояние E_i (непосредственно). При $\lambda_k = 0$ состояние E_k поглощающее.

Числа λ_k и π_{ki} дают возможность наглядно описать характер изменения состояния системы во времени (иначе говоря, ее эволюцию). Если система находится в состоянии E_k , то она будет пребывать в этом состоянии случайное время, имеющее показательное распределение с параметром λ_k , по истечении которого перейдет с вероятностью π_{ki} в состояние E_i (i любое число, не равное k). В этом состоянии система пробудет показательное время с параметром λ_i и с вероятностями π_{ij} перейдет в одно из состояний E_j ($j \neq i$) и т. д.

На практике нас интересует вероятность того, что система в момент t будет находиться в некотором состоянии E_i , если известно, что в начальный момент времени она была в состоянии E_k (так как интенсивности переходов не зависят от времени, то указанная вероятность совпадает с вероятностью того, что система, находясь в некоторый момент в состоянии E_k , через время t попадает в E_i).

Заметим, что за время t может осуществиться много переходов из состояния в состояние, поэтому эта вероятность существенно отличается от π_{ki} .

Обозначим через $E_k(t)$ событие «система в момент t находится в состоянии E_k ». Вероятности

$$P\{E_i(t)/E_k(0)\} = p_{ki}(t) \quad (1)$$

называются *вероятностями перехода системы*.

Укажем некоторые их свойства.

I. $0 \leq p_{ki}(t) \leq 1$, поскольку это вероятности.

II. $\sum_i p_{ki}(t) = 1$, так как события $E_i(t)$ несовместимы, $E_1(t) + \dots + E_n(t)$ — достоверное событие (в момент t система находится в одном из состояний E_1, \dots, E_n).

III. При $t > 0$, $s > 0$ выполнено равенство

$$p_{ki}(t+s) = \sum_j p_{kj}(t) p_{ji}(s). \quad (2)$$

Оно называется уравнением Колмогорова — Чепмена и требует более подробного рассмотрения. Имеем равенство

$$\begin{aligned} p_{ki}(t+s) &= \frac{P(E_k(0) \cdot E_i(t+s))}{P(E_k(0))} = \\ &= \frac{\sum_j P(E_k(0) \cdot E_j(t) \cdot E_i(t+s))}{P(E_k(0))}. \end{aligned} \quad (3)$$

Здесь использовано то, что $E_1(t) + \dots + E_n(t)$ — достоверное событие. Но

$$P(E_i(t+s)/E_k(0) \cdot E_j(t)) = p_{ji}(s)$$

(вероятность того, что система, находясь в момент t в состоянии E_j , попадет через время s в состояние E_i ; тот факт, что она в начальный момент была в состоянии E_k , не существен, поскольку эволюция системы на промежутке $[t, t+s]$ определяется ее состоянием в момент t).

Значит,

$$\frac{P(E_k(0) \cdot E_j(t) \cdot E_i(t+s))}{P(E_k(0))} = \frac{P(E_k(0) \cdot E_j(t)) p_{ji}(s)}{P(E_k(0))} = \\ = p_{kj}(t) p_{ji}(s).$$

Подставляя это равенство в правую часть (3), получим (2).

Уравнения (2) используются для того, чтобы связать вероятности перехода процесса с непосредственно заданными интенсивностями λ_k и a_{kj} . Эта связь осуществляется через уравнения Колмогорова (это простейшие дифференциальные уравнения, выражающие производные от вероятностей перехода через сами вероятности перехода). Чтобы их получить, исследуем вероятности перехода за малый промежуток времени.

Мы видели в § 22, что событие с интенсивностью λ происходит на промежутке длины h с вероятностью примерно λh . Два таких независимых события могут произойти одновременно с вероятностью порядка h^2 (произведение вероятностей). Поэтому если считать вероятность перехода с точностью до величин порядка h^2 , то можно предполагать, что произошло не более одного перехода. Значит,

$$p_{kk}(h) = e^{-\lambda_k h} + O(h^2)$$

(то, что система за время h не вышла из состояния E_k имеет вероятность $\lambda_k h e^{-\lambda_k h}$, а вероятность за это время выйти из состояния E_k и вернуться назад имеет порядок h^2 ; такую величину обозначают $O(h^2)$, поскольку при этом совершается более одного перехода).

Точно так

$$p_{kl}(h) = a_{kl}(h) + O(h^2).$$

Значит, выполнены предельные соотношения

$$\lim_{h \rightarrow 0} \frac{p_{kk}(h) - 1}{h} = \lim_{h \rightarrow 0} \frac{e^{-\lambda_k h} - 1}{h} = -\lambda_k, \\ \lim_{h \rightarrow 0} \frac{p_{kl}(h)}{h} = a_{kl}. \quad (4)$$

Используя уравнение (2), можем записать

$$p_{kl}(t+h) = \sum_j p_{kj}(h) p_{jl}(t), \quad (5)$$

$$\frac{p_{ki}(t+h) - p_{ki}(t)}{h} = \frac{1}{h} (p_{kk}(h) - 1) p_{ki}(t) + \sum_{i \neq k} \frac{1}{h} p_{ki}(h) p_{ji}(t).$$

Правая часть этого соотношения имеет в силу (4) предел при $h \rightarrow 0$. Поэтому имеет предел и левая часть, т. е. существует производная у функции $p_{ki}(t)$. Переходя к пределу, получаем равенство

$$p'_{ki}(t) = -\lambda_k p_{ki}(t) + \sum_j a_{kj} p_{ji}(t). \quad (6)$$

Эти равенства выполняются для k и i , меняющихся от 1 до n .

Совокупность равенств (6) называется *первой системой уравнений Колмогорова*. При решении этих уравнений нужно учитывать равенства $p_{ki}(0) = 0$ при $i \neq k$, $p_{kk}(0) = 1$.

Вторая система уравнений Колмогорова получается аналогично, если вместо соотношения (5) использовать равенство

$$p_{ki}(t+h) = \sum_j p_{kj}(t) p_{ji}(h).$$

Она имеет вид

$$p'_{ki}(t) = -\lambda_i p_{ki}(t) + \sum_j a_{ji} p_{kj}(t). \quad (7)$$

Каждая из систем Колмогорова (с учетом значений $p_{ki}(0)$) однозначно определяет вероятности перехода.

В качестве примера вычислим вероятности перехода для системы, состоящей из одного станка, обслуживаемого одним рабочим. Пусть λ — интенсивность события «работающий станок вышел из строя», а μ — интенсивность события «неисправный станок налажен». Если через E_1 обозначить состояние системы, когда станок работает, а E_2 — состояние, когда он налаживается, то $p_{11}(t)$ есть вероятность того, что в момент t станок работал, если он работал в начальный момент времени, $p_{21}(t)$ — вероятность того, что он в момент t налаживается, если он работал в начальное время, $p_{12}(t)$ — вероятность того, что в момент t станок работает, если он в начальный момент налаживается, $p_{22}(t)$ — вероятность того, что он налаживается, если он налаживался и в начальный момент времени.

Для вероятностей перехода выполнены следующие два соотношения:

$$p_{11}(t) + p_{12}(t) = 1, \quad p_{21}(t) + p_{22}(t) = 1.$$

Из сказанного выше вытекает, что λ_1 — интенсивность того, что система покинет состояние E_1 (она совпадает с интенсивностью перехода $E_1 \rightarrow E_2$), равна λ , т. е.

$$\lambda_1 = a_{12} = \lambda.$$

Точно так

$$\lambda_2 = a_{21} = \mu.$$

Запишем вторую систему уравнений Колмогорова

$$\left. \begin{aligned} \dot{p}_{11}(t) &= -\lambda p_{11}(t) + \mu p_{12}(t), \\ \dot{p}_{12}(t) &= -\mu p_{12}(t) + \lambda p_{11}(t), \\ \dot{p}_{21}(t) &= -\lambda p_{21}(t) + \mu p_{22}(t), \\ \dot{p}_{22}(t) &= -\mu p_{22}(t) + \lambda p_{21}(t). \end{aligned} \right\} \quad (8)$$

В первом уравнении заменим $p_{12}(t) = 1 - p_{11}(t)$.
Получим

$$\dot{p}_{11}(t) = -(\lambda + \mu)p_{11}(t) + \mu = -(\lambda + \mu) \left[p_{11}(t) - \frac{\mu}{\lambda + \mu} \right],$$

или

$$\left[p_{11}(t) - \frac{\mu}{\lambda + \mu} \right]' = -(\lambda + \mu) \left[p_{11}(t) - \frac{\mu}{\lambda + \mu} \right].$$

Если функция $v(t)$ удовлетворяет уравнению

$$v'(t) = -kv(t),$$

то $[e^{kt}v(t)]' = e^{kt}v'(t) + ke^{kt}v(t) = 0$.

Значит, $[e^{kt}v(t)] = c$ и $v(t) = ce^{-kt}$, где c — постоянная. Поэтому

$$p_{11}(t) - \frac{\mu}{\lambda + \mu} = ce^{-(\lambda + \mu)t}.$$

Полагая $t = 0$ и учитывая, что $p_{11}(0) = 1$, находим

$$c = 1 - \frac{\mu}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu}. \text{ Значит,}$$

$$p_{11}(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t},$$

$$p_{12}(t) = 1 - p_{11}(t) = \frac{\lambda}{\lambda + \mu} (1 - e^{-(\lambda + \mu)t}).$$

Аналогично из уравнения (8) получаем

$$p_{22}(t) = \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} e^{-(\lambda + \mu)t},$$

$$p_{21}(t) = \frac{\mu}{\lambda + \mu} (1 - e^{-(\lambda + \mu)t}).$$

§ 25. Эргодическая теорема

Рассмотрим опять тот пример из предыдущего параграфа, для которого мы вычислили вероятности перехода. Что происходит с ними по истечении достаточно большого промежутка времени? Так как $(\lambda + \mu) > 0$, то $e^{-(\lambda + \mu)t} \rightarrow 0$ при $t \rightarrow \infty$. Если, например, $\lambda + \mu$ имеет порядок $0,001$ 1/с, то при t порядка $10\,000$ с

$$e^{-(\lambda + \mu)t} \sim e^{-10} < 0,0001.$$

Поэтому при больших t получим

$$p_{11}(t) \sim \frac{\mu}{\lambda + \mu}, \quad p_{21}(t) \sim \frac{\mu}{\lambda + \mu},$$

$$p_{12}(t) \sim \frac{\lambda}{\lambda + \mu}, \quad p_{22}(t) \sim \frac{\lambda}{\lambda + \mu}.$$

Это означает, что вероятность пребывания системы в состояниях E_1 и E_2 для больших t примерно постоянна и не зависит от того, в каком состоянии система находилась в начальный момент времени. Данное свойство называется *эргодичностью*. Из эргодической теоремы вытекает, что таким свойством обладают все физически разумные системы с конечным множеством состояний.

Поясним, что следует понимать под физически разумной системой. Во-первых, система должна в каждом своем состоянии проводить положительное время, так как если она мгновенно его покидает, то такое состояние можно вообще не рассматривать, а считать, что система, минуя его, сразу попала в то состояние, в которое она перешла из мгновенного. Во-вторых, состояния системы невозможно разбить на два класса так, чтобы переходы из класса в класс были невозможны. В последнем случае естественно считать, что у нас имеются по крайней мере две независимые системы.

Для изучения асимптотического поведения вероятностей перехода полезно ввести понятие класса сообщающихся

состояний. Два состояния E_i и E_j называется *сообщающимися*, если для всех $t > 0$ $p_{ij}(t) > 0$ и $p_{ji}(t) > 0$, т. е. при любых t за время t можно с положительной вероятностью перейти из одного состояния в другое. Заметим, что каждое состояние общается само с собой, так как время выхода из состояния E_i имеет показательное распределение с параметром λ_i , значит,

$$p_{ii}(t) \geq e^{-\lambda_i t}.$$

Далее, если состояние E_i общается с E_k и E_j общается с E_k , то E_i и E_j общаются:

$$p_{ij}(t) \geq p_{ik}\left(\frac{t}{2}\right) p_{kj}\left(\frac{t}{2}\right), \quad p_{ji}(t) \geq p_{jk}\left(\frac{t}{2}\right) p_{ki}\left(\frac{t}{2}\right)$$

(мы использовали равенство Колмогорова — Чепмена

$$p_{ij}(t) = \sum_k p_{ik}\left(\frac{t}{2}\right) p_{kj}\left(\frac{t}{2}\right),$$

взяв лишь одно слагаемое справа).

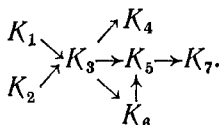
Рассмотрим совокупность состояний, сообщающихся с состоянием E_i . Любые два состояния из этой совокупности общаются. Эта совокупность содержит состояние E_i , и никакое состояние из этой совокупности не общается с состоянием вне этой совокупности. Такую совокупность будем называть *классом состояний*. Все состояния системы разбиваются на классы. Класс называется *существенным*, если система, попав в такой класс, никогда уже из него не выйдет. Остальные классы называются *несущественными*. Из несущественного класса система выходит или в другой несущественный класс, или в существенный класс. Заметим, что система, выйдя из класса, не может больше в него вернуться. Если бы она могла вернуться, то состояние, в которое система перешла после выхода из класса, общалось бы с состояниями класса, а значит, и само должно было бы принадлежать этому классу.

Простейший пример разбиения системы на классы имеет вид

$$K_1 \rightarrow K_2 \rightarrow K_3$$

(возможны переходы из класса K_1 в K_2 , а из K_2 в K_3 , K_1 и K_2 — несущественны, K_3 — существенный класс). Более

сложный случай приведен ниже



В этой схеме классы K_4 и K_7 существенны, остальные несущественны. Через конечное время система попадет в один из существенных классов. Поэтому, если E_i принадлежит несущественному классу, то вероятность пребывания системы в этом состоянии стремится к нулю при $t \rightarrow \infty$, это справедливо, каково бы ни было начальное состояние E_k (может оказаться, что из E_k вообще нельзя попасть в E_i , например в схеме невозможны попадания в классы K_1 и K_2 из других классов). Значит, для всех k

$$\lim_{t \rightarrow \infty} p_{ki}(t) = 0.$$

Поскольку система по истечении достаточно большого промежутка времени попадает в один из существенных классов, а такие классы никак не сообщаются между собой, то они могут рассматриваться как самостоятельные системы, у которых все состояния сообщаются. Для таких систем и справедлива эргодическая теорема.

Теорема. Если все состояния системы сообщаются, то существуют пределы

$$\lim_{t \rightarrow \infty} p_{ik}(t) = p_k, \quad (1)$$

вероятности p_k являются единственным решением системы уравнений

$$\sum_k p_k = 1, \quad \lambda_k p_k = \sum_i p_i a_{ik}, \quad k = 1, \dots, m. \quad (2)$$

Доказательство существования пределов (1) опирается на некоторые специальные математические факты, которые здесь приводить неуместно. Поэтому остановимся на выводе системы уравнений (2). Записывая уравнение Колмогорова — Чепмена

$$p_{ik}(t+s) = \sum_j p_{ij}(t) p_{jk}(s)$$

и переходя в нем к пределу при $t \rightarrow \infty$, получим для всех $s > 0$

$$p_k = \sum_j p_j p_{jk}(s). \quad (3)$$

Это означает, что вероятности p_k обладают следующим свойством. Пусть в начальный момент времени система находится в состоянии E_k с вероятностью p_k (т. е. начальное положение системы также случайно). Тогда, обозначая через $p_k(s)$ вероятность того, что система находится в состоянии E_k в момент s , получим по формуле полной вероятности

$$p_k(s) = \sum_j P\{E_k(s)/E_j(0)\} P\{E_j(0)\} = \sum_j p_j p_{jk}(s) = p_k,$$

т. е. если в начальный момент вероятность находиться в состоянии E_k равнялась p_k , то она останется неизменной в любой последующий момент времени. Поэтому вероятности p_k называются *стационарными для данной системы*.

Рассмотрим теперь вторую систему уравнений Колмогорова (уравнение (7) § 24), умножим k -е уравнение на p_k и сложим все уравнения. Получим

$$\sum_k p_k p'_{ki}(t) = -\lambda_i \sum_k p_k p_{ki}(t) + \sum_j \sum_k p_k p_{kj}(t) a_{ji}. \quad (4)$$

Но

$$\sum_k p_k p_{ki}(t) = p_i, \quad \left(\sum_k p_k p_{ki}(t)\right)' = 0$$

(производная от постоянной). Значит, (4) преобразуется в равенство

$$0 = -\lambda_i p_i + \sum_j p_j a_{ji},$$

из которого и вытекает (3).

Покажем, что при наших предположениях стационарное распределение единственно. Пусть \bar{p}_k — некоторое другое стационарное распределение. Тогда

$$\bar{p}_k = \sum_j \bar{p}_j p_{jk}(s).$$

Перейдем в этом соотношении к пределу при $s \rightarrow \infty$ $p_{jk}(s) \rightarrow p_k$. Значит,

$$\bar{p}_k = p_k \sum_j \bar{p}_j = p_k,$$

так как $\sum_j \bar{p}_j = 1$.

Если вероятности пребывания системы в каждом состоянии неизменны, то говорят, что система находится в стационарном режиме. Из эргодической теоремы вытекает, что каждая система со временем переходит в стационарный режим. Не следует думать, что состояния системы в стационарном режиме остаются неизменными. Система меняет свои состояния, неизменны только вероятности пребывания.

Чтобы пояснить ситуацию, воспользуемся частотной интерпретацией вероятности. Пусть имеется N одинаковых и независимо эволюционирующих систем, причем в начальный момент число систем, находящихся в состоянии E_i , равнялось N_i , $N = N_1 + \dots + N_m$. Если числа N_i большие, то $p_{ik}(t)$ есть доля тех систем, которые перейдут в состояние E_k из E_i , поэтому их число равно примерно $N_i p_{ik}(t)$. Общее число систем, находящихся в состоянии E_k , будет

$$N_1 p_{1k}(t) + \dots + N_m p_{mk}(t).$$

Если $\frac{N_i}{N} \approx p_i$, то последнее выражение равно приближенно

$$N(p_1 p_{1k}(t) + \dots + p_m p_{mk}(t)) = N p_k \approx N_k.$$

Хотя каждая система меняет свое состояние, число систем, находящихся в данном состоянии, остается примерно постоянным.

Рассмотрим, например, систему из станка, обслуживаемого одним рабочим. Если интенсивность выхода из строя станка — λ , а интенсивность наладки станка равна μ , то стационарные вероятности имеют вид

$$p_1 = \frac{\mu}{\lambda + \mu}, \quad p_2 = \frac{\lambda}{\lambda + \mu}$$

(E_1 — станок работает, E_2 — станок палаживается). Если в цехе N станков, то по истечении некоторого времени примерно $N \frac{\mu}{\lambda + \mu}$ из этих станков работает, $N \frac{\lambda}{\lambda + \mu}$ — налаживается.

Заметим, что в последнем случае для стационарных вероятностей можно дать более наглядную интерпретацию. Величина $\frac{1}{\lambda}$ есть среднее значение для показательного распределения с параметром λ , т. е. среднее время работы станка

до разладки, точно так $\frac{1}{\mu}$ — среднее время ремонта станка. Так что

$$p_1 = \frac{1/\lambda}{1/\mu + 1/\lambda}, \quad p_2 = \frac{1/\mu}{1/\mu + 1/\lambda},$$

вероятность пребывания системы в состоянии E_1 пропорциональна среднему времени работы станка, а в E_2 — среднему времени ремонта станка.

Рассмотрим некоторые примеры на вычисление стационарных вероятностей.

1. Обслуживание станков. Имеется m станков, которые обслуживаются одним рабочим. Среднее время работы одного станка до разладки $1/\lambda$ (т. е. интенсивность разладки λ), среднее время ремонта равно $1/\mu$. Система находится в одном из состояний E_0, E_1, \dots, E_m , в состоянии E_k работают k станков, вышедшие из строя станки налаживаются по очереди. Ниже стрелками показаны возможные переходы из состояния в состояние:

$$E_0 \rightleftharpoons E_1 \rightleftharpoons \dots \rightleftharpoons E_m.$$

Интенсивность перехода $E_k \rightarrow E_{k+1}$ всегда равна μ (это интенсивность события «окончен ремонт станка»). Интенсивность перехода $E_k \rightarrow E_{k-1}$ равна $k\lambda$ (это интенсивность суммы событий, каждое из которых имеет интенсивность λ : «хотя бы один из k станков вышел из строя»). Любые другие переходы невозможны: $\lambda_0 = a_{01} = \mu$, $a_{10} = \lambda$, $a_{12} = \mu$, $\lambda_1 = \lambda + \mu$, \dots , $a_{kk-1} = k\lambda$, $a_{kk+1} = \mu$, $\lambda_k = k\lambda + \mu$, $a_{mm-1} = \lambda_m = m\lambda$.

Система уравнений (2) принимает вид

$$\begin{aligned} \mu p_0 &= p_1 \lambda, \\ (\lambda + \mu) p_1 &= p_0 \mu + p_2 (2\lambda), \\ &\dots \dots \dots \\ (k\lambda + \mu) p_k &= p_{k-1} \mu + p_{k+1} (k+1) \lambda, \quad (k < m) \\ m\lambda p_m &= p_{m-1} \mu. \end{aligned}$$

Из нее последовательно получаем

$$\begin{aligned} p_1 &= \frac{\mu}{\lambda} p_0, \quad 2\lambda p_2 = (\lambda + \mu) \frac{\mu}{\lambda} p_0 - \mu p_0 = \frac{\mu^2}{\lambda} p_0, \\ p_2 &= \frac{\mu^2}{2\lambda^2}, \quad 3\lambda p_3 = (2\lambda + \mu) \frac{\mu^2}{2\lambda^2} - \frac{\mu^2}{\lambda} p_0 = \frac{\mu^3}{2\lambda^2} p_0, \end{aligned}$$

$$p_3 = \frac{\mu^3}{\lambda^3 1 \cdot 2 \cdot 3} p_0, \quad p_k = \frac{\mu^k}{\lambda^k k!} p_0, \quad k < m,$$

$$p_m = \frac{\mu}{m\lambda} p_{m-1} = \frac{\mu^m}{m! \lambda^m} p_0.$$

Для определения p_0 воспользуемся тем, что сумма вероятностей равна 1. Значит,

$$1 = p_0 \left(1 + \frac{\mu}{\lambda} + \frac{1}{2!} \left(\frac{\mu}{\lambda} \right)^2 + \dots + \frac{1}{m!} \left(\frac{\mu}{\lambda} \right)^m \right),$$

$$p_k = \frac{\left(\frac{\mu}{\lambda} \right)^k}{k! \left(1 + \frac{\mu}{\lambda} + \frac{1}{2!} \left(\frac{\mu}{\lambda} \right)^2 + \dots + \frac{1}{m!} \left(\frac{\mu}{\lambda} \right)^m \right)}.$$

Из математического анализа известна формула (см. также формулу (2) на стр. 49)

$$e^a = 1 + a + \frac{a^2}{2!} + \dots + \frac{a^m}{m!} + \dots$$

Поэтому для достаточно больших m при ограниченных $\frac{\mu}{\lambda}$ получаем следующую приближенную формулу: $p_k = \frac{a^k}{k!} e^{-a}$, где $a = \frac{\mu}{\lambda}$, т. е. число работающих станков имеет распределение Пуассона.

2. Ионизация газа. Дано N газовых молекул, ионизируемых излучением. Интенсивность ионизации одной молекулы равна λ . Кроме того, возможны рекомбинации положительного и отрицательного ионов в нейтральную молекулу. Интенсивность рекомбинации одного положительного иона равна μ . Состояние системы E_m определяется числом m положительных ионов (число отрицательных ионов равно также m).

Интенсивность перехода

$$E_m \rightarrow E_{m+1}$$

совпадает с интенсивностью суммы $N - m$ событий: одна из $N - m$ имеющихся в наличии нейтральных молекул ионизирована, поэтому она равна $(N - m)\lambda$. Интенсивность перехода $E_m \rightarrow E_{m-1}$ есть интенсивность суммы m событий: «один из m положительных ионов рекомбинирует».

вал», и следовательно, равна $m\mu$. Итак,

$$a_{mm-1} = m\mu \quad (\text{при } m = 0 \quad a_{mm-1} = 0),$$

$$a_{mm+1} = (N - m)\lambda \quad (\text{при } m = N \quad a_{mm+1} = 0),$$

$$\lambda_m = m\mu + (N - m)\lambda.$$

Система уравнений для стационарных вероятностей принимает вид

$$N\lambda p_0 = \mu p_1,$$

.....

$$[(N - m)\lambda + m\mu] p_m = \lambda(N - m + 1) p_{m-1} + (m + 1)\mu p_{m+1}, \quad 0 < m < N,$$

.....

$$N\mu p_N = \lambda p_{N-1}.$$

Из этих уравнений можно последовательно выразить все вероятности через p_0 . Так, из первого уравнения находим

$$p_1 = N \left(\frac{\lambda}{\mu} \right) p_0,$$

из второго —

$$[(N - 1)\lambda + \mu] p_1 = N\lambda p_0 + 2\mu p_2,$$

$$\begin{aligned} p_2 &= \frac{1}{2\mu} \left\{ [(N - 1)\lambda + \mu] N \left(\frac{\lambda}{\mu} \right) p_0 - N\lambda p_0 \right\} = \\ &= \frac{N(N - 1)}{2} \left(\frac{\lambda}{\mu} \right)^2 p_0 = C_N^2 \left(\frac{\lambda}{\mu} \right)^2 p_0. \end{aligned}$$

Покажем, что

$$p_m = C_N^m \left(\frac{\lambda}{\mu} \right)^m p_0. \quad (5)$$

Достаточно проверить выполнение системы уравнений. Имеем

$$\begin{aligned} & [(N - m)\lambda + m\mu] \frac{N!}{m!(N - m)!} \left(\frac{\lambda}{\mu} \right)^m p_0 = \\ &= \frac{N!}{m!(N - m - 1)!} \frac{\lambda^{m+1}}{\mu^m} p_0 + \frac{N!}{(m - 1)!(N - m)!} \frac{\lambda^m}{\mu^{m-1}} p_0 = \end{aligned}$$

$$= \mu (m + 1) C_N^{m+1} \left(\frac{\lambda}{\mu} \right)^{m+1} p_0 + \\ + \lambda (N - m + 1) C_N^{m-1} \left(\frac{\lambda}{\mu} \right)^{m-1} p_0.$$

Формула (5) установлена.

Далее,

$$1 = p_0 + p_1 + \dots + p_N = p_0 \left(1 + C_N^1 \left(\frac{\lambda}{\mu} \right) + \right. \\ \left. + C_N^2 \left(\frac{\lambda}{\mu} \right)^2 + \dots + \left(\frac{\lambda}{\mu} \right)^N \right) = p_0 \left(1 + \frac{\lambda}{\mu} \right)^N.$$

Значит,

$$p_0 = \frac{\mu^N}{(\lambda + \mu)^N}, \\ p_m = C_N^m \left(\frac{\lambda}{\lambda + \mu} \right)^m \left(\frac{\mu}{\lambda + \mu} \right)^{N-m}.$$

Стационарные вероятности таковы, как у биномиального распределения с $p = \frac{\lambda}{\lambda + \mu}$.

§ 26. Ветвящиеся процессы

Рассмотрим колонию бактерий. Пока их количество таково, что на всех хватает питания и они не заражают среду обитания продуктами жизнедеятельности, можно считать, что отдельные особи развиваются независимо. Каждая особь живет определенное время, оно зависит от многих непредсказуемых факторов, поэтому его естественно считать случайным. По истечении этого времени особь либо делением порождает две дочерние особи, либо погибает («исчезает»).

Описывая эволюцию колонии бактерий вероятностными методами, мы можем рассматривать ее как систему, состояние которой определяется числом особей. Изменение состояния системы происходит тогда и только тогда, когда одна из особей или исчезает, или порождает две новые.

Аналогичный характер имеет эволюция системы из совершенно другой области. Известно, что ядра урана-235, поглощая медленные нейтроны, делятся на части; при этом могут выделиться несколько новых медленных нейтронов. Вновь выделенные нейтроны могут опять вступать в реак-

цию. Кроме того, каждый медленный нейтрон может поглощаться примесями или покинуть область, где происходит реакция. Очевидно, что до тех пор, пока число тепловых нейтронов мало по сравнению с числом ядер урана, поведение отдельных нейтронов независимо. Каждый нейтрон имеет определенное «время жизни» в свободном состоянии, после которого он или «исчезает», или порождает несколько новых нейтронов. Состояние системы так же, как в случае с колонией бактерий, характеризуется числом нейтронов, переходы из состояния в состояние происходят случайным образом, и только тогда, когда один из нейтронов «исчезает» или порождает в результате ядерной реакции несколько новых.

Рассмотрим, наконец, распространение эпидемии некоторой инфекционной болезни. Если описывать состояние системы числом больных, то состояние системы меняется тогда, когда один из больных выздоровеет («исчезнет») или заразит несколько здоровых людей. Опять мы имеем дело с системой, состояние которой определяется числом некоторых объектов; эти объекты случайным образом или исчезают, или порождают несколько новых объектов такого же типа.

Для математического описания систем такого вида используются ветвящиеся процессы. Сейчас мы подробно остановимся на ветвящемся процессе с одним типом частиц. *Частицами* называются те объекты, которые определяют состояние процесса (особи, нейтроны, больные в рассмотренных выше примерах). Каждая частица живет определенное (случайное) время, после чего превращается в несколько частиц такого же вида или исчезает. Основное предположение заключается в том, что различные частицы ведут себя независимо. В частности, времена жизни частиц независимы.

Введем основные вероятностные характеристики ветвящегося процесса.

Пусть λ — интенсивность события A «произошло превращение частицы» (т. е. частица или исчезла, или превратилась в несколько). Тогда, как установлено в § 22, время до наступления этого события имеет показательное распределение с параметром λ , т. е.

$$P(\tau > t) = e^{-\lambda t}, \quad t \geq 0.$$

Обозначим через π_k вероятность того, что в момент пре-

зращения частица превратилась в k частиц, $k = 0, 2, 3, \dots$
 π_0 — вероятность того, что частица исчезнет).

Пусть A_k — это событие «данная частица превратилась в k частиц», $k = 0, 2, 3, \dots$. Если λ_k — интенсивность события A , то поскольку

$$A = A_0 + A_2 + \dots + A_k + \dots,$$

то

$$\lambda = \lambda_0 + \lambda_2 + \dots + \lambda_k + \dots$$

и

$$\pi_k = P(A_k/A) = \frac{\lambda_k}{\lambda}$$

(это было установлено в § 22). Таким образом, $\lambda_k = \lambda \pi_k$.

Пусть $p_{kj}(t)$ обозначает вероятность того, что в момент t будет j частиц, если в начальный момент было k частиц (это вероятности перехода для нашей системы). Через $v(t)$ выразим случайную величину — число частиц в момент t . Следовательно,

$$p_{kj}(t) = P\{v(t) = j/v(0) = k\}.$$

Во многих случаях можно ограничиться изучением поведения среднего числа частиц. Рассмотрим эту величину. Будем предполагать, что среднее число частиц, в которое превращается одна частица, конечно, т. е. величина

$$m = 2\pi_2 + 3\pi_3 + \dots + k\pi_k + \dots < \infty.$$

Покажем, что если в начальный момент было k частиц, то

$$Mv(t) = ke^{\lambda(m-1)t}. \quad (1)$$

Обозначим через $v_1(t), \dots, v_k(t)$ число частиц, порожденных соответственно каждой из k начальных частиц. Тогда

$$v(t) = v_1(t) + \dots + v_k(t)$$

и

$$Mv(t) = Mv_1(t) + \dots + Mv_k(t) = kMv(t),$$

так как величины $v_i(t)$ имеют одинаковые распределения. Поэтому формулу (1) достаточно установить для случая $k = 1$.

Положим

$$a(t) = Mv_1(t).$$

Величину $v_1(t+s)$ можно представить как сумму $v_1(s)$ величин

$$v_1(t+s) = \sum_{i=1}^{v_1(s)} \tilde{v}_i(t), \quad (2)$$

которые являются потомствами тех $v_1(s)$ частиц, появившихся к моменту s из начальной частицы: $\tilde{v}_i(t)$ — число частиц, получившихся из одной частицы, находящейся в системе в момент s за время t (т. е. до момента $s+t$).

Величины $\tilde{v}_i(t)$ не зависят от $v_1(s)$. Поэтому

$$M[v_1(t+s)/v_1(s) = j] = jM\tilde{v}_1(t) = ja(t).$$

Если мы умножим это соотношение на $P\{v_1(s) = j\}$ и просуммируем по j , то в левой части получим по формуле полной вероятности $Mv_1(t+s)$. Действительно,

$$M[v_1(t+s)/v_1(s) = j] = \sum_k kp_{jk}(t),$$

$$\sum_j p_{jk}(t) P\{v_1(s) = j\} = P\{v_1(t+s) = k\}.$$

Значит,

$$a(t+s) = a(t) \sum_j jP\{v_1(s) = j\} = a(t) Mv_1(s) = a(t)a(s).$$

Так как за малое время h событие A_k происходит с вероятностью $\lambda\pi_k h$, то

$$P\{v_1(h) = k\} \sim \lambda\pi_k h, \quad k = 0, 2, \dots \quad (3)$$

$$1 - P\{v_1(h) = 1\} \sim \lambda h.$$

Следовательно,

$$a(h) - 1 = \sum_{k=1}^{\infty} kP\{v_1(h) = k\} - 1 \sim$$

$$\sim \lambda h \left[\sum_{k=2}^{\infty} k\pi_k - 1 \right] = \lambda(m-1)h.$$

Имеем

$$a(t+h) - a(t) = a(t)a(h) - a(t) = a(t)[a(h) - 1] \sim \sim \lambda(m-1)a(t),$$

$$\lim_{h \rightarrow 0} \frac{a(t+h) - a(t)}{h} = \lambda(m-1)a(t).$$

Функция $a(t)$, таким образом, удовлетворяет соотношению

$$a'(t) = \lambda(m-1)a(t)$$

и, кроме того, $a(0) = 1$. Поэтому $a(t) = e^{\lambda(m-1)t}$. Тем самым формула (1) установлена.

Из этой формулы вытекает, что поведение среднего значения числа частиц зависит от величины $\lambda(m-1)$. Качественно можно выделить 3 случая.

1. $m > 1$, т. е. среднее значение числа частиц, в которые превращается одна частица в момент превращения, больше 1. Тогда $\lambda(m-1) > 0$ и при $t \rightarrow \infty Mv(t)$ также возрастает до бесконечности. Такой ветвящийся процесс называется *надкритическим*.

2. $m = 1$. Тогда $\lambda(m-1) = 0$ и, следовательно, $Mv(t) = k$, если в начальный момент было k частиц. Другими словами, среднее число частиц остается неизменным. В этом случае ветвящийся процесс называется *критическим*.

3. $m < 1$, $\lambda(m-1) < 0$. Значит, $Mv(t) \rightarrow 0$ при $t \rightarrow \infty$. Так как $v(t)$ — целочисленная случайная величина, то $P\{v(t) \geq 1\} \rightarrow 0$, $P\{v(t) = 0\} \rightarrow 1$. Это означает, что со временем в процессе исчезают все частицы (говорят, что процесс вырождается). Процесс при $m < 1$ называется *докритическим*.

Рассмотрим теперь вопрос о *вырождении процесса*. Поскольку частицы в процессе могут появляться лишь в результате превращения других частиц, то если в некоторый момент все частицы исчезнут, тогда и во все последующие моменты времени частицы в системе будут отсутствовать.

Пусть B — событие, которое заключается в том, что наступит такой момент σ , что $v(\sigma) = 0$. Наступление события B и обозначает вырождение процесса. Найдем вероятность вырождения. Она, очевидно, зависит от числа частиц в начальный момент времени.

Предположим, что

$$P(B/v(0) = k)$$

обозначает вероятность вырождения при условии, что в начальный момент было k частиц. Так как в этом случае $v(t) = v_1(t) + \dots + v_k(t)$, где $v_i(t)$ — величина потомства i -й начальной частицы, $v_i(t)$ — независимые величины,

то

$$P \{v(t) = 0\} = P \{v_1(t) = 0, \dots, v_k(t) = 0\} = \\ = (P \{v_1(t) = 0\})^k. \quad (4)$$

Какова бы ни была последовательность $t_n \rightarrow \infty$, события $\{v(t_n) = 0\}$ монотонно возрастают с t_n ($\{v(t_n) = 0\}$ влечет событие $\{v(t_{n+1}) = 0\}$ при $t_n < t_{n+1}$), а сумма этих событий есть событие B . Поэтому

$$P(B) = \lim_{t_n \rightarrow \infty} P \{v(t_n) = 0\} = \lim_{t \rightarrow \infty} P \{v(t) = 0\}$$

и аналогичное соотношение справедливо для условных вероятностей.

Переходя в (4) к пределу при $t \rightarrow \infty$, получим

$$P(B/v(0) = k) = (P(B/v(0) = 1))^k. \quad (5)$$

Найдем $P(B/v(0) = 1)$. Как было установлено выше,

$$P(B/v(0) = 1) = \lim_{t \rightarrow \infty} p_{10}(t).$$

Воспользуемся равенством Колмогорова — Чепмена:

$$p_{10}(t+h) = \sum_k p_{1k}(h) p_{k0}(t).$$

Переходя к пределу при $t \rightarrow \infty$, находим

$$P(B/v(0) = 1) = \sum_k p_{1k}(h) (P(B/v(0) = 1))^k.$$

Используя равенства (3), может записать

$$P(B/v(0) = 1) (1 - p_{11}(h)) = \sum_{k \neq 1} p_{1k}(h) (P(B/v(0) = 1))^k,$$

$$\lambda P(B/v(0) = 1) = \sum_{k \neq 1} \lambda \pi_k (P(B/v(0) = 1))^k. \quad (6)$$

Введем функцию

$$\varphi(\alpha) = \sum_{k \neq 1} \pi_k \alpha^k - \alpha.$$

Тогда $P(B/v(0) = 1)$ является корнем уравнения $\varphi(\alpha) = 0$. Это уравнение имеет по крайней мере один корень, так как

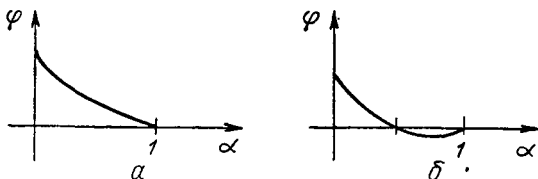
$$\varphi(1) = \sum_{k \neq 1} \pi_k - 1 = 0.$$

Далее, если $\pi_0 = 0$, т. е. вероятность исчезновения частиц равна нулю, то и $P(B) = 0$.

Рассмотрим случай, когда $\pi_0 > 0$. Функция $\varphi(\alpha)$ при $\alpha \in [0, 1]$ выпукла вниз, поскольку

$$\varphi''(\alpha) = \sum_{k=2}^{\infty} \pi_k k(k-1) \alpha^{k-2} \geq 0.$$

Поэтому возможен один из двух видов графика функции $\varphi(\alpha)$:



В случае *a* $\alpha = 1$ является единственным корнем уравнения $\varphi(\alpha) = 0$. Значит, $P(B) = 1$. Так как $\varphi(\alpha)$ при этом убывает, то $\varphi'(\alpha) \leq 0$, т. е.

$$\sum_{k=2}^{\infty} k \pi_k \alpha^{k-1} - 1 \leq 0.$$

Полагая $\alpha = 1$, получаем

$$m - 1 \leq 0.$$

Следовательно, случай *a* имеет место для докритического и критического процессов. При $m > 1$ будет случай *б*. Так как

$$P(B/\nu(0) = 1) = \lim_{t \rightarrow \infty} p_{10}(t),$$

а $P_{10}(0) = 0$, то P_{10} при $t \rightarrow \infty$ непрерывно возрастает от 0 до первого после нуля корня уравнения $\varphi(\alpha) = 0$, поэтому

$$P(B/\nu(0) = 1) = \bar{\alpha},$$

где $\bar{\alpha}$ — меньший из корней уравнения $\varphi(\alpha) = 0$ на отрезке $[0, 1]$. При этом $0 < \bar{\alpha} < 1$. Таким образом, для надкритического процесса (даже при $\pi_0 = 0$)

$$P(B/\nu(0) = k) = \bar{\alpha}^k,$$

где $\bar{\alpha} < 1$ (при $\pi_0 = 0$ будет $\bar{\alpha} = 0$). Если k достаточно большое, то вероятность вырождения сколь угодно мала.

Значит, надкритический процесс, имеющий в начальный момент достаточно большое число частиц, практически не вырождается.

Что же происходит с ветвящимся процессом в надкритическом случае, если он не вырождается? Оказывается, что для любого

$$P \{v(t) > n\} \rightarrow 1 - P(B), \quad (7)$$

т. е. в этом случае $v(t)$ становится неограниченным.

Покажем это. Опять, используя уравнение Колмогорова — Чепмена, запишем для $t > 0$ и $s > 0$

$$p_{k0}(t+s) = p_{k0}(t) + \sum_{i>0} p_{ki}(t) p_{i0}(s) \quad (8)$$

(мы воспользовались тем, что $P_{00}(s) = 1$, так как из состояния 0 система никогда не выйдет). Перейдем к пределу при $s \rightarrow \infty$ и воспользуемся тем, что

$$\lim_{s \rightarrow \infty} p_{i0}(s) = \bar{\alpha}^i.$$

Тогда из (8) получим

$$\bar{\alpha}^k = p_{k0}(t) + \sum_{i>0} p_{ki} \bar{\alpha}^i.$$

При $\bar{\alpha} > 0$ имеем

$$\begin{aligned} p_{k1}(t) + \dots + p_{kn}(t) &\leq \frac{1}{\bar{\alpha}^n} (p_{k1}(t) \bar{\alpha} + \dots + p_{kn} \bar{\alpha}^n) \leq \\ &\leq \frac{1}{\bar{\alpha}^n} \sum_{i>0} p_{ki}(t) \bar{\alpha}^i = \frac{1}{\bar{\alpha}^n} [\bar{\alpha}^k - p_{k0}(t)]. \end{aligned}$$

Поскольку $p_{k0}(t) \rightarrow \bar{\alpha}^k$ при $t \rightarrow \infty$, то

$$\lim_{t \rightarrow \infty} (p_{k1}(t) + \dots + p_{kn}(t)) = 0.$$

Из этого соотношения и вытекает (7). Если же $\bar{\alpha} = 0$, то $p_0 = 0$ и, значит, частицы не исчезают. Тогда $v(t)$ возрастает с t , причем после каждого превращения $v(t)$ увеличится по крайней мере на 1. Поскольку в этом случае обязательно будет бесконечное число превращений (за бесконечное время), то $v(t) \rightarrow +\infty$.

Следовательно, всегда, когда вырождение не происходит, возникает «взрыв» — число частиц неограниченно возрастает.

Что означает вырождение или «взрыв» в тех примерах, которые рассматривались в начале параграфа? Для колонии бактерий вырождение есть гибель всех особей, «взрыв» же — разрастание колонии до размеров, при которых развитие отдельных особей еще возможно (достаточно питательных веществ и среда не сильно отравлена продуктами жизнедеятельности живых и разложения мертвых особей). Такие колонии бактерий получаются при посевах для обнаружения их паличия в некоторой среде. Если колония гибнет, то и при паличии бактерий в среде мы их не обнаружим, если же происходит «взрыв», колония разрастается и присутствие бактерий в среде может быть установлено.

При ядерной реакции с медленными нейтронами «взрыв» приводит к реальному взрыву в результате выделения большого количества энергии, при этом сама система разрушается и процесс прекращается. Вырождение означает, что система находится в стабильном состоянии. Даже притекание небольшого количества нейтронов со стороны (которое всегда имеет место) не приводит к ядерному взрыву.

Если рассматривать как ветвящийся процесс распространение инфекционной болезни, то вырождение означает прекращение эпидемии, а взрыв — почти поголовное заболевание всех предрасположенных к данной болезни.

Проанализируем детальнее процесс, в котором возможны лишь два вида превращения: исчезновение или деление на две частицы. Пусть λ — интенсивность превращений; π_0 и π_2 — вероятности того, что в момент превращения частица исчезнет или превратится в две. В этом случае

$$m = 2\pi_2, \quad m - 1 = 2\pi_2 - 1 = \pi_2 - \pi_0.$$

Значит, при $\pi_2 < \pi_0$ процесс докритический, $\pi_2 = \pi_0$ — критический, $\pi_2 > \pi_0$ — надкритический. Определим вероятность вырождения в надкритическом случае. Функция $\varphi(\alpha)$ имеет вид

$$\begin{aligned} \varphi(\alpha) &= \pi_0 - \alpha + \pi_2 \alpha^2 = \pi_0 - (\pi_0 + \pi_2) \alpha + \pi_2 \alpha^2 = \\ &= \pi_0 (1 - \alpha) - \pi_2 \alpha (1 - \alpha) = (1 - \alpha) (\pi_0 - \pi_2 \alpha). \end{aligned}$$

Так как в надкритическом случае $\pi_2 > \pi_0$, то

$$\bar{\alpha} = \frac{\pi_0}{\pi_2} < 1$$

есть наименьший положительный корень уравнения $\varphi(\alpha) = 0$. Значит, вероятность вырождения равна $\frac{\pi_0}{\pi_2}$.

Среднее значение числа частиц имеет вид

$$Mv(t) = ke^{\mu t},$$

где k — число частиц в начальный момент времени,

$$\mu = \lambda(\pi_2 - \pi_0) = \frac{\pi_2 - \pi_0}{M\tau},$$

где τ — время жизни частицы до превращения; $M\tau$ — среднее значение τ . По величине μ можно определить «время удвоения числа частиц», т. е. такое время T , что $Mv(T) = 2k$. Имеем

$$e^{\mu T} = 2, \quad T = \frac{\ln 2}{\mu} = 0,693 \frac{M\tau}{\pi_2 - \pi_0}.$$

§ 27. Колебания со случайной амплитудой

Колебательные процессы в природе возникают при взаимно обратных переходах двух видов энергии (движения) друг в друга. В механических колебаниях — это потенциальная и кинетическая энергия, в электромагнитных — энергия электрических и магнитных полей. Простейшее колебание — это гармоническое колебание, в котором находится гармонический осциллятор.

Механический гармонический осциллятор получаем при рассмотрении движения материальной точки около положения равновесия под действием сил, которые возвращают ее в положение равновесия. Гармоническое колебание будет получаться, если возвращающая сила пропорциональна отклонению точки от положения равновесия. Такой характер имеют, например, упругие силы в пределах небольших отклонений.

Предположим, что точка движется вдоль оси x , причем $x = 0$ является точкой равновесия. Если $F(x)$ — величина силы, действующей на точку при отклонении x , то $F(x) = -cx$, где c — некоторый коэффициент пропорциональности, при $x = 0$ сила равна нулю, при $x \neq 0$ сила направлена к положению равновесия и пропорциональна отклонению. Пусть $x(t)$ — положение точки в момент t , тогда $x'(t)$ — ее скорость (с учетом знака), $x''(t)$ — ее ускорение.

По второму закону Ньютона

$$mx''(t) = -cx(t)$$

(m — масса точки).

Обозначим $\frac{c}{m} = \omega^2$. Тогда предыдущее уравнение можно переписать в виде

$$x''(t) + \omega^2 x(t) = 0. \quad (1)$$

Это и есть уравнение для гармонических колебаний. Так как для функций $\sin \omega t$ и $\cos \omega t$ справедливы соотношения

$$(\sin \omega t)'' = -\omega^2 \sin \omega t, \quad (\cos \omega t)'' = -\omega^2 \cos \omega t,$$

то функция

$$x(t) = A \cos \omega t + B \sin \omega t \quad (2)$$

удовлетворяет уравнению (1). (Можно показать, что всякая функция $x(t)$, удовлетворяющая соотношению (1), имеет вид (2).)

A и B — постоянные, имеющие простой физический смысл. Полагая $t = 0$, находим $A = x(0)$, т. е. A — начальное положение точки.

Продифференцируем (2) по t :

$$x'(t) = -A\omega \sin \omega t + B\omega \cos \omega t.$$

Полагая здесь $t = 0$, видим

$$B = \frac{1}{\omega} x'(0)$$

($x'(0)$ — начальная скорость точки).

Преобразуем (2) следующим образом. Пусть

$$R = \sqrt{A^2 + B^2},$$

φ такой угол, что

$$B = R \cos \varphi, \quad A = -R \sin \varphi.$$

Тогда

$$\begin{aligned} x(t) &= R [\cos \varphi \cdot \cos \omega t - \sin \varphi \cdot \sin \omega t], \\ x(t) &= R \cos(\omega t + \varphi). \end{aligned} \quad (8)$$

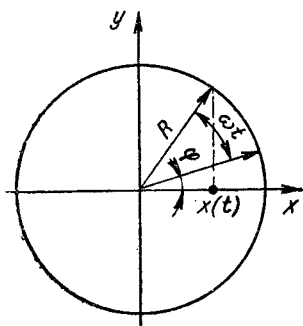
Это и есть общий вид функции, описывающий гармоническое колебание; R — амплитуда колебания; ω — частота (точнее, циклическая частота); φ — начальная фаза.

Рассмотрим энергию колеблющейся точки. Для того чтобы сдвинуть точку из положения 0 в положение x , нужно затратить работу на преодоление возвращающей силы, равную $\frac{c}{2} x^2$. Это потенциальная энергия точки в положении x . Ее кинетическая энергия в момент t равна

$$\frac{m}{2} [x'(t)]^2 = \frac{m}{2} R^2 \omega^2 \sin^2(\omega t + \varphi).$$

Значит, полная энергия будет

$$\frac{c}{2} R^2 \cos^2(\omega t + \varphi) + \frac{m\omega^2}{2} R^2 \sin^2(\omega t + \varphi) = R^2 \frac{c}{2},$$



так как $m\omega^2 = c$. Отсюда видим, что энергия пропорциональна квадрату амплитуды и при заданном c не зависит от m , а значит, и ω .

Заметим, что частота колебаний не меняется при изменении начальных условий, от них зависят лишь амплитуда и начальная фаза.

Гармоническое колебание удобно изображать схематически в виде кругового движения. Пусть точка движется по окружности радиуса R с постоянной угловой скоростью ω против часовой стрелки; φ — начальный угол, отсчитываемый от положительного направления оси x до точки; $\varphi(t)$ — такой угол до положения точки в момент t (при этом $\varphi(t)$ может содержать несколько полных оборотов). Тогда

$$\varphi(t) = \varphi + \omega t$$

и проекция вращающейся точки на ось x будет

$$x(t) = R \cos(\omega t + \varphi),$$

т. е. также имеет вид (3).

Равномерное движение по окружности проще изучать (в силу его равномерности). Заметим, что для важных случаев колебаний имеет физический смысл и проекция кругового движения на ось y

$$y(t) = R \sin(\omega t + \varphi).$$

Так, в электромагнитных колебаниях, если $x(t)$ — напряженность электрического поля, то $y(t)$ будет напряженностью магнитного поля.

Точки плоскости можно изображать с помощью комплексных чисел. Используем представление комплексных чисел в тригонометрической форме в таком виде:

$$z = re^{i\varphi} = r(\cos \varphi + i \sin \varphi), \quad (i = \sqrt{-1})$$

(применяем известную формулу Эйлера $\cos \varphi + i \sin \varphi = e^{i\varphi}$). Здесь r — модуль комплексного числа z ; φ — его аргумент. Тогда, если $z(t)$ — положение вращающейся точки в момент t , то

$$z(t) = Re^{i(\omega t + \varphi)}. \quad (4)$$

Это комплексное представление гармонического колебания. Оно удобно, и мы будем его в дальнейшем использовать.

Предположим, что начальное положение точки, совершающей гармонические колебания, было случайным: случайны положение и скорость в начальный момент времени. Тогда будут случайными амплитуда и начальная фаза, на частоту колебаний это влияния не окажет, она по-прежнему будет равна ω . Положение вращающейся точки на плоскости также будет случайным. Обозначим его через $\zeta(t)$. В соответствии с формулой (4) $\zeta(t)$ можно представить в виде

$$\zeta(t) = Z_0 e^{i\omega t}, \quad (5)$$

где $Z_0 = Re^{i\varphi}$ — случайная величина с комплексными значениями (комплексная амплитуда колебания).

Рассмотрим теперь суперпозицию (наложение) нескольких гармонических колебаний. Наиболее просто такую суперпозицию представить для электромагнитных колебаний. Каждый гармонический осциллятор таких колебаний излучает электромагнитное колебание определенной частоты вида (4), при этом напряженности (и электрическая и магнитная), создаваемые полями отдельных осцилляторов, складываются. Поэтому для нахождения напряженности суммарного поля нужно сложить описывающие отдельные поля комплексные числа вида (4) и найти вещественную часть полученного комплексного числа; мнимая часть этого комплексного числа будет напряженностью магнитного поля.

Следовательно, суммарное поле описывается функцией

$$z(t) = R_1 e^{i(\omega_1 t + \varphi_1)} + \dots + R_n e^{i(\omega_n t + \varphi_n)}$$

(рассматривается суперпозиция n гармонических колебаний вида $R_k e^{i(\omega_k t + \varphi_k)}$). Если каждый осциллятор колеблется случайно, то $Z(t)$ также случайно меняется, является случайным процессом. Представим этот процесс, пользуясь формулой (5), в следующем виде:

$$\zeta(t) = Z_1 e^{i\omega_1 t} + \dots + Z_n e^{i\omega_n t}, \quad (6)$$

где Z_k — комплексные случайные амплитуды отдельных колебаний; ω_k — их частоты.

Будем считать, что процесс (6) определен для всех вещественных t , кроме того, естественно предполагать, что отдельные колебания независимы. Это означает, что независимы случайные величины Z_1, \dots, Z_n .

Величина $Z_k = X_k + iY_k$, где X_k и Y_k — вещественные случайные величины. Будем предполагать, что величины X_k и Y_k таковы, что

$$MX_k = MY_k = 0.$$

Это предположение выполнено, если комплексная амплитуда колебания $R_k e^{i\varphi_k}$ удовлетворяет условию: φ_k и R_k независимы и φ_k равномерно распределено на $[0, 2\pi]$. Действительно, тогда

$$X_k = R_k \cos \varphi_k,$$

$$MX_k = MR_k M \cos \varphi_k = MR_k \int_0^{2\pi} \cos \theta \frac{d\theta}{2\pi} = 0;$$

аналогично устанавливается равенство $MY_k = 0$.

Напомним, что модуль комплексного числа, представимого в виде $Re^{i\varphi}$, равен R , модуль числа Z обозначается $|Z|$, если $Z = X + iY$, то

$$|Z|^2 = X^2 + Y^2 = R^2 \cos^2 \varphi + R^2 \sin^2 \varphi = R^2.$$

Величина

$$M|Z_k|^2 = MR_k^2 = b_k$$

пропорциональна средней энергии колебания k -го осциллятора. Для простоты будем считать, что b_k — средняя энергия колебания.

Рассмотрим некоторые характеристики процесса $\zeta(t)$ при сделанных предположениях. Будем определять математическое ожидание комплексной случайной величины $Z = X + iY$ с помощью формулы

$$MZ = MX + iMY.$$

Найдем $M\zeta(t)$. Поскольку $MZ_k = 0$ для величин Z_k , входящих в формулу (6), то $M\zeta(t) = 0$. Если $\zeta(t)$ — некоторый случайный процесс, то функция $M\zeta(t)$ называется *средним значением процесса*. Для случайного процесса $\zeta(t)$ вида (6) при сделанных предположениях среднее значение равно нулю.

Другой важной характеристикой случайного процесса $\zeta(t)$, принимающего комплексные значения, является *корреляционная функция*. Она определяется равенством

$$B(t, s) = M\zeta(t)\overline{\zeta(s)} - M\zeta(t)M\overline{\zeta(s)}$$

(здесь $\bar{\zeta}$ — комплексное число, сопряженное к ζ , если $\zeta = \xi + i\eta$, то $\bar{\zeta} = \xi - i\eta$).

Вычислим корреляционную функцию для процесса $\zeta(t)$ вида (6) при сделанных выше предположениях. Так как тогда $M\zeta(t) = 0$ и $M\bar{\zeta}(s) = 0$, то

$$B(t, s) = M(Z_1 e^{i\omega_1 t} + \dots + Z_n e^{i\omega_n t}) \times \\ \times \overline{(Z_1 e^{i\omega_1 s} + \dots + Z_n e^{i\omega_n s})}$$

(число во второй скобке берется сопряженным).

Известно, что сопряженное к сумме и произведению комплексных чисел есть сумма и произведение сопряженных комплексных чисел. Поэтому

$$\overline{(Z_1 e^{i\omega_1 s} + \dots + Z_n e^{i\omega_n s})} = \bar{Z}_1 e^{-i\omega_1 s} + \dots + \bar{Z}_n e^{i\omega_n s}$$

(дело в том, что $\cos \varphi + i \sin \varphi = \cos \varphi - i \sin \varphi = \cos(-\varphi) + i \sin(-\varphi) = e^{-i\varphi}$).

Значит,

$$B(t, s) = \sum_{k,j} MZ_k \bar{Z}_j e^{i\omega_k t} \cdot e^{-i\omega_j s}. \quad (7)$$

Из независимости величин Z_k и Z_j при $k \neq j$ вытекает, что $MZ_k \bar{Z}_j = MZ_k \cdot M\bar{Z}_j = 0$. Поэтому в сумме справа

останутся лишь слагаемые, для которых $k = j$. Так что

$$B(t, s) = \sum_k M Z_k \bar{Z}_k e^{i\omega_k(t-s)} = \\ = \sum_k M |Z_k|^2 e^{i\omega_k(t-s)} = \sum_k b_k e^{i\omega_k(t-s)}.$$

Таким образом, корреляционная функция процесса зависит от разности аргументов $t - s$, будем обозначать ее $B(t - s)$.

Случайный процесс $\zeta(t)$, для которого среднее значение постоянно, а корреляционная функция $B(t, s)$ зависит от разности аргументов $t - s$, называется *стационарным*. Это название отражает то свойство случайного процесса, что его характеристики (среднее значение и корреляционная функция) не меняются при сдвиге времени.

Рассмотрим вместо процесса $\zeta(t)$ процесс $\zeta(t + h)$, где h фиксировано. Тогда $M\zeta(t + h) = M\zeta(t)$ и

$$M\zeta(t + h) \overline{\zeta(s + h)} - M\zeta(t + h) \overline{M\zeta(s + h)} = \\ = B(t + h, s + h) = B(t + h - s - h) = B(t - s),$$

т. е. и среднее значение и корреляционная функция процесса $\zeta(t + h)$ такие же, как у $\zeta(t)$. Для стационарного процесса корреляционной функцией называется именно функция одного аргумента

$$B(\tau) = B(t + \tau, t), \quad B(t, s) = B(t - s).$$

Итак, процесс вида (6) при условии, что случайные величины Z_k независимы и $MZ_k = 0$, является стационарным процессом, для которого среднее равно нулю, а для корреляционной функции справедливо представление

$$B(\tau) = \sum_k b_k e^{i\omega_k \tau}, \quad (8)$$

где ω_k — частоты составляющих $\zeta(t)$ гармонических колебаний, а b_k — энергия колебаний с частотой ω_k .

Наборы чисел ω_k и b_k определяют спектр процесса $\zeta(t)$. Зная их, мы можем восстановить отдельные составляющие процесса (точнее — указать их частоты и энергии). Если нам известен спектр процесса, можно определить по формуле (8) корреляционную функцию. Наоборот, по корреляционной функции можно восстановить спектр процесса.

На практике приходится иметь дело с процессами,

представимыми в виде суммы очень большого числа осцилляторов, излучающих на близко расположенных частотах. Например, излучение нагретого тела (газа) есть суммарное излучение отдельных атомов, которые выступают здесь в качестве осцилляторов, их число имеет порядок 10^{20} . Спектр такого случайного процесса уже неудобно описывать с помощью наборов частот и энергий. Для его задания используют спектральную функцию распределения.

Обозначим через $F(\omega)$ сумму энергий тех колебаний, для которых частота ω_k меньше, чем ω :

$$F(\omega) = \sum_{\omega_k < \omega} b_k,$$

ω — любое вещественное число. Тогда энергия колебаний на частотах из интервала $[\omega, \omega + \Delta\omega)$ будет

$$F(\omega + \Delta\omega) - F(\omega).$$

Функция $F(\omega)$ называется *спектральной функцией распределения*.

Корреляционная функция может быть выражена через спектральную функцию распределения следующим образом:

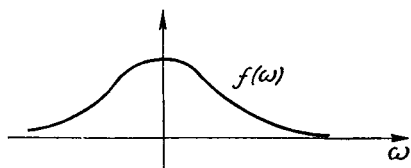
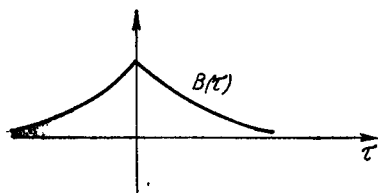
$$B(\tau) = \lim_{h \rightarrow 0} \sum e^{in\hbar\tau} [F((n+1)h) - F(nh)].$$

Предел справа есть по определению интеграл $\int e^{i\omega\tau} dF(\omega)$ (интеграл от $-\infty$ до ∞), так что

$$B(\tau) = \int e^{i\omega\tau} dF(\omega). \quad (9)$$

Для процессов вида (6) формула (9) сводится к формуле (8). Но оказывается, что для всякого стационарного процесса существует спектральная функция распределения (что утверждает теорема Бохнера — Хипчина, доказательство которой выходит за рамки используемых здесь понятий). Это такая неубывающая ограниченная функция (суммарная энергия отдельных колебаний должна быть ограничена) $F(\omega)$, что для корреляционной функции справедливо представление (9). Так как интеграл (9) можно приближенно представить в виде суммы (8), то отсюда вытекает, что каждый стационарный процесс, для которого $M \xi(t) = 0$, можно приближенно представить в виде (6).

Если $F(\omega + \Delta\omega) - F(\omega) \sim f(\omega) \Delta\omega$ для некоторой функции $f(\omega)$, то $f(\omega)$ называется *спектральной*



плотностью процесса. Через спектральную плотность корреляционная функция выражается обычным интегралом

$$B(\tau) = \int e^{i\omega\tau} f(\omega) d\omega. \quad (10)$$

Спектральная плотность является неотрицательной (интегрируемой) функцией. Она может быть вычислена по корреляционной функции с помощью формулы

$$f(\omega) = \frac{1}{2\pi} \int e^{-i\omega\tau} B(\tau) d\tau. \quad (11)$$

Приведем некоторые примеры спектральных плотностей и соответствующих корреляционных функций.

1. Равномерный спектр в некоторой полосе. В этом случае $f(\omega) = c$ при $\omega \in [a, b]$, $f(\omega) = 0$ при $\omega \notin [a, b]$. Корреляционная функция имеет вид

$$\begin{aligned} B(\tau) &= c \int_a^b e^{i\omega\tau} d\omega = \frac{c}{i\tau} [e^{i\tau b} - e^{i\tau a}] = \\ &= \frac{c}{i\tau} e^{i\tau \frac{a+b}{2}} [e^{i\tau \frac{b-a}{2}} - e^{-i\tau \frac{b-a}{2}}] = \\ &= \frac{2c}{\tau} e^{i\tau \frac{b+a}{2}} \sin \tau \frac{b-a}{2}. \end{aligned}$$

Пусть полоса симметрична относительно точки 0 и имеет вид $[-a, a]$. Тогда корреляционная функция будет

$$B(\tau) = \frac{2c}{\tau} \sin a\tau.$$

Заметим, что величина $\int_{-a}^a c d\omega = 2ac$ есть полная энергия всех колебаний. Поэтому, обозначая ее через B , будем иметь

$$B(\tau) = B \frac{\sin a\tau}{a\tau}.$$

2. Пусть спектральная плотность имеет вид

$$f(\omega) = \frac{c}{\pi(a^2 + \omega^2)}.$$

Тогда

$$B(\tau) = \frac{c}{a} e^{-a|\tau|}.$$

Чтобы убедиться в этом, достаточно воспользоваться формулой (11):

$$f(\omega) = \frac{1}{2\pi} \cdot \frac{c}{a} \int e^{-a|\tau| - i\omega\tau} d\tau = \frac{c}{2\pi a} \left(\int_{-\infty}^0 e^{a\tau - i\omega\tau} d\tau + \int_0^{\infty} e^{-a\tau - i\omega\tau} d\tau \right) = \frac{c}{\pi(a^2 + \omega^2)}.$$

На приведенном выше (с. 178) графике изображены: корреляционная функция, вещественная часть наблюдаемой кривой и спектральная плотность для последнего примера.

Можно ли, наблюдая процесс $\zeta(t)$ на достаточно большом интервале, определить его характеристики (среднее значение и корреляционную функцию)? Оказывается, можно, если спектр его непрерывен, т. е. приращения спектральной функции распределения стремятся к нулю: $F(\omega + \Delta\omega) - F(\omega) \rightarrow 0$ при $\Delta\omega \rightarrow 0$.

Пусть

$$\zeta(t) = a + \sum_k Z_k e^{i\omega_k t}, \quad \omega_k \neq 0.$$

Рассмотрим среднее значение по времени от $\zeta(t)$ на отрезке $[0, T]$. Получим

$$\begin{aligned} \frac{1}{T} \int_0^T \zeta(t) dt &= a + \sum_k Z_k \frac{1}{T} \int_0^T e^{i\omega_k t} dt = \\ &= a + \sum_k Z_k \cdot \frac{1}{T} \cdot \frac{e^{i\omega_k T} - 1}{i\omega_k}. \end{aligned}$$

Так как величины $\frac{e^{i\omega_k T} - 1}{i\omega_k}$ ограничены, то при $T \rightarrow \infty$ сумма справа стремится к нулю. Значит,

$$a = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) dt.$$

Для определения корреляционной функции будем считать, что $a = 0$ и рассмотрим среднее значение по времени вида

$$\begin{aligned} \frac{1}{T} \int_0^T \xi(t + \tau) \overline{\xi(t)} dt &= \sum_{k,i} Z_k \overline{Z_i} \frac{1}{T} \int_0^T e^{i\omega_k(t+\tau) - i\omega_i t} dt = \\ &= \sum_k |Z_k|^2 e^{i\omega_k \tau} + \sum_{k \neq i} Z_k \overline{Z_i} e^{i\omega_k \tau} \frac{e^{i(\omega_k - \omega_i)T} - 1}{i(\omega_k - \omega_i) T}. \end{aligned}$$

Переходя к пределу при $T \rightarrow \infty$, найдем

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t + \tau) \overline{\xi(t)} dt = \sum_k |Z_k|^2 e^{i\omega_k \tau}.$$

В том случае, когда спектр непрерывен, $M |Z_k|^2$ будут сколь угодно малыми, и поэтому к сумме независимых величин справа будет применим закон больших чисел, так что эту сумму можно заменить ее средним значением. Но

$$M \sum_k |Z_k|^2 e^{i\omega_k \tau} = \sum_k b_k e^{i\omega_k \tau} = B(\tau).$$

Таким образом, для процессов с непрерывным спектром

$$B(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t + \tau) \overline{\xi(t)} dt.$$

Какие задачи возникают при рассмотрении стационарных процессов и как они решаются, если известен спектр процесса, мы увидим в следующем параграфе.

§ 28. Фильтрация и прогноз

Как правило, при обработке данных о случайных функциях используют лишь значения функции в некоторой дискретной последовательности моментов времени $t = kh$, $k = 0, \pm 1, \pm 2, \dots$, где h — величина промежутка, через который

производятся измерения. Обычно приборы устроены так, что они позволяют замерять лишь такую последовательность, но даже тогда, когда записывающее устройство выдает график функции, для дальнейшей численной обработки с графика снимают лишь значения функции на последовательности моментов времени указанного вида. Не ограничивая общности, можно h выбрать за единицу измерения времени. Тогда вместо непрерывного процесса $\zeta(t)$ мы будем рассматривать последовательность $\zeta(k)$, $k = 0, \pm 1, \pm 2, \dots$. Чтобы не менять терминологию, назовем ее также процессом.

Если $\zeta(t)$ — стационарный процесс, то для представления $B(k)$, корреляционной функции $\zeta(k)$, можно воспользоваться формулой (10) § 27 (в предположении существования спектральной плотности). Однако она допускает некоторое упрощение. Из указанной формулы вытекает

$$B(k) = \int e^{i\omega k} f(\omega) d\omega.$$

Но функция $e^{i\omega k} = \cos k\omega + i \sin k\omega$ является периодической с периодом 2π . Поэтому

$$\int_{2n\pi}^{2(n+1)\pi} e^{i\omega k} f(\omega) d\omega = \int_0^{2\pi} e^{i\omega k} f(\omega + 2n\pi) d\omega.$$

Положим

$$\hat{f}(\omega) = \sum_n f(\omega + 2n\pi).$$

Тогда

$$B(k) = \int_0^{2\pi} e^{i\omega k} \hat{f}(\omega) d\omega. \quad (1)$$

Функция $\hat{f}(\omega)$, входящая в представление (1), называется *спектральной плотностью для процесса* (последовательности) $\zeta(k)$.

Будем считать, что процесс $\zeta(k)$ есть некоторый сигнал (например, радиосигнал). Приемное устройство, принимая сигнал, преобразует его определенным образом. Наиболее широкое применение на практике имеют линейные устройства. Схематически действие устройства таково:

$$\zeta(k) \rightarrow U \rightarrow \eta(k),$$

на устройство поступает сигнал $\zeta(k)$, а $\eta(k)$ — результат преобразования сигнала устройством.

Для линейных устройств справедливо свойство

$$\zeta_1(k) + \zeta_2(k) \rightarrow U \rightarrow \eta_1(k) + \eta_2(k),$$

т. е. сумма двух сигналов преобразуется в сумму преобразований отдельных сигналов. Если значение $\eta(k)$ зависит лишь от значения $\zeta(k)$ (устройство без памяти), то

$$\eta(k) = c\zeta(k),$$

где c — некоторая постоянная (считаем, что устройство не меняется со временем). Общий линейный преобразователь получим, если будем считать, что $\eta(k)$ зависит от всех предыдущих значений $\zeta(k)$ (из физических соображений понятно, что от будущих значений $\zeta(k)$ $\eta(k)$ зависеть не может). Если свойства преобразователя не меняются со временем, то он будет действовать по формуле

$$\eta(k) = \sum_{n \geq 0} c_n \zeta(k - n) \quad (2)$$

(значение процесса $\zeta(t)$ в момент, на n единиц предшествующий текущему моменту k , входит в $\eta(k)$ с коэффициентом c_n). Числа c_n таковы, что сумма в (2) определена.

Оказывается, в результате преобразования (2) мы опять получим стационарную последовательность. Пусть $M\zeta(k) = a$, $B(k)$ — корреляционная функция для $\zeta(k)$. Тогда

$$a_1 = M\eta(k) = a \sum_{n \geq 0} c_n,$$

$$\begin{aligned} M\eta(k+m) \overline{\eta(m)} &= M \sum_{n \geq 0} c_n \zeta(k+m-n) \sum_{l \geq 0} \bar{c}_l \bar{\zeta}(m-l) = \\ &= \sum_{n \geq 0} \sum_{l \geq 0} c_n \bar{c}_l M\zeta(k+m-n) \bar{\zeta}(m-l) = \\ &= \sum_{n \geq 0} \sum_{l \geq 0} c_n \bar{c}_l [B(k-n+l) + |a|^2]; \end{aligned}$$

как видим, выражение справа действительно зависит только от k . Отсюда и вытекает стационарность $\eta(k)$.

Предположим, что $B(k)$ допускает представление (1). Покажем, что тогда и $\eta(k)$ имеет спектральную плотность и найдем ее выражение.

Из приведенных выше вычислений вытекает, что корреляционная функция $B_1(k)$ для процесса $\eta(k)$ имеет вид

$$B_1(k) = \sum_{n \geq 0} \sum_{l \geq 0} c_n \bar{c}_l B(k - n + l).$$

Поэтому

$$B_1(k) = \sum_{n \geq 0} \sum_{l \geq 0} c_n \bar{c}_l \int_{-\infty}^{\infty} e^{tk\omega} e^{-ln\omega} e^{il\omega} \hat{f}(\omega) d\omega = \\ = \int_0^{2\pi} e^{ik\omega} \sum_{n \geq 0} c_n e^{-in\omega} \sum_{l \geq 0} c_l e^{il\omega} \hat{f}(\omega) d\omega.$$

Обозначим $u(\omega) = \sum_{n \geq 0} c_n e^{-in\omega}$,

тогда $\bar{u}(\omega) = \sum_{n \geq 0} \bar{c}_n e^{in\omega}$,

$$B_1(k) = \int_0^{2\pi} e^{tk\omega} |u(\omega)|^2 \hat{f}(\omega) d\omega.$$

Функция $|u(\omega)|^2$ называется *частотной характеристикой преобразователя*. Если $\hat{f}_1(\omega)$ — спектральная плотность $\eta(k)$, то

$$\hat{f}_1(\omega) = |u(\omega)|^2 \hat{f}(\omega), \quad (3)$$

т. е. линейный преобразователь работает очень просто: он умножает спектральную плотность принимаемого сигнала на частотную характеристику.

Преобразователи, которые не меняют одни частоты и гасят другие, называются *фильтрами*. Рассмотрим одну из задач типа фильтрации, т. е. построения фильтра с заданными свойствами. Предположим, что наблюдаемый процесс $\xi(k)$ имеет полезный сигнал, частота которого $\omega_0 \in [0, 2\pi]$ и помехи со спектральной плотностью $\hat{f}(\omega)$ (полезный сигнал будем считать неслучайным).

Идеальный фильтр должен был бы иметь такую частотную характеристику, чтобы $|u(\omega_0)|^2 = 1$ и $u(\omega) = 0$ при $\omega \neq \omega_0$. Однако такой фильтр построить нельзя. Мы сейчас найдем фильтр, у которого $|u(\omega_0)|^2 = 1$ и $u(\omega)$ достаточно мало, если $|\omega - \omega_0| > \Delta$, где Δ — некоторое заданное число.

Пусть

$$|u(\omega)|^2 = \frac{\varepsilon^2}{\varepsilon^2 + 4(1 - \varepsilon) \sin^2\left(\frac{\omega - \omega_0}{2}\right)}. \quad (4)$$

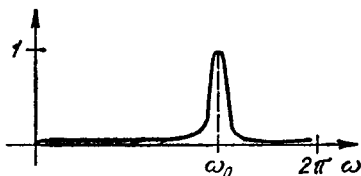


График этой функции имеет примерно такой вид:

(при $|\omega - \omega_0| > \sqrt{\varepsilon}$,

$$|u(\omega)| < \frac{4\varepsilon}{1-\varepsilon}.$$

Найдем вид соответствующего линейного преобразователя. Для этого представим знаменатель в (4) так:

$$\begin{aligned} & \varepsilon^2 + 2(1-\varepsilon)(1 - \cos(\omega - \omega_0)) = \\ & = (1-\varepsilon)^2 + 1 - 2(1-\varepsilon)\cos(\omega - \omega_0) = \\ & = (1-\varepsilon)^2 \cos^2(\omega - \omega_0) + (1-\varepsilon)^2 \sin^2(\omega - \omega_0) + \\ & \quad + 1 - 2(1-\varepsilon)\cos(\omega - \omega_0) = \\ & = [(1-\varepsilon)\cos(\omega - \omega_0) - 1]^2 + (1-\varepsilon)^2 \sin^2(\omega - \omega_0) = \\ & = |(1-\varepsilon)\cos(\omega - \omega_0) - 1 + i(1-\varepsilon)\sin(\omega - \omega_0)|^2 = \\ & = |1 - (1-\varepsilon)e^{i(\omega - \omega_0)}|^2. \end{aligned}$$

Положим
$$u(\omega) = \frac{\varepsilon}{1 - (1-\varepsilon)e^{-i(\omega - \omega_0)}}.$$

Тогда выполнена формула (4). Выражение справа можно представить в виде суммы геометрической прогрессии

$$u(\omega) = \varepsilon \sum_{n \geq 0} (1-\varepsilon)^n e^{in\omega_0} e^{-in\omega}.$$

Значит, искомый фильтр получится, если в преобразовании (2) взять

$$c_n = \varepsilon (1-\varepsilon)^n e^{in\omega_0}.$$

Рассмотрим теперь задачу о прогнозе случайного процесса. Она заключается в наилучшем приближении значений процесса в будущем по его значениям до настоящего момента времени.

Простейшая задача прогноза: как, зная значения процесса ζ_k при $k \leq 0$, определить ζ_1 ? Уточним, во-первых, что нужно понимать под построением некоторой величины по прошлым значениям. Мы уже имели дело с таким построением при знакомстве с линейными преобразователями. Будем и сейчас рассматривать в качестве приближенных значений величины, полученные из ζ_k с помощью линейных преобразователей:

$$\hat{\zeta}_1 = \sum_{n \geq 0} c_n \zeta_{-n}. \quad (5)$$

Здесь $\hat{\zeta}_1$ — искомое приближенное значение ζ_1 . Линейный преобразователь выберем так, чтобы среднее квадратическое отклонение приближения от истинного значения

$$M |\zeta_1 - \hat{\zeta}_1|^2 \quad (6)$$

было минимальным. Если c_n выбраны указанным образом, то величина

$$\hat{\zeta}_{k+1} = \sum_{n \geq 0} c_n \zeta_{k-n}$$

будет наилучшим приближением величины ζ_{k+1} по значениям ζ_l , $l \leq k$.

Установим одно важное свойство наилучшего приближения $\hat{\zeta}_l$: для всех $l \leq 0$

$$M (\zeta_1 - \hat{\zeta}_1) \bar{\zeta}_l = 0. \quad (7)$$

Действительно, при всех комплексных α имеем

$$M |\zeta_1 - \hat{\zeta}_1 - \alpha \zeta_l|^2 \geq M |\zeta_1 - \hat{\zeta}_1|^2$$

(в силу минимальности величины (6)), значит,

$$\begin{aligned} M |\zeta_1 - \hat{\zeta}_1|^2 &\leq M (\zeta_1 - \hat{\zeta}_1 - \alpha \zeta_l) \overline{(\zeta_1 - \hat{\zeta}_1 - \alpha \zeta_l)} = \\ &= M |\zeta_1 - \hat{\zeta}_1|^2 - \bar{\alpha} M (\zeta_1 - \hat{\zeta}_1) \bar{\zeta}_l - \alpha M (\zeta_1 - \hat{\zeta}_1) \zeta_l + \\ &\quad + |\alpha|^2 M |\zeta_l|^2, \end{aligned}$$

$$M |\zeta_l|^2 |\alpha|^2 - \bar{\alpha} M (\zeta_1 - \hat{\zeta}_1) \bar{\zeta}_l - \alpha M (\zeta_1 - \hat{\zeta}_1) \zeta_l \geq 0.$$

Если $M (\zeta_1 - \hat{\zeta}_1) \bar{\zeta}_l = a + ib$, то, взяв вещественное α , найдем

$$\alpha^2 M |\zeta_l|^2 - 2a\alpha \geq 0.$$

Поэтому

$$2a \frac{\alpha}{|\alpha|} \leq |\alpha| M |\zeta_l|^2, \quad |a| \leq \frac{|\alpha| \cdot M |\zeta_l|^2}{2}.$$

В силу произвольности α $|a| = 0$. Аналогично, взяв $\alpha = i\beta$, где β — вещественное, найдем, что $b = 0$. Тем самым (7) установлено. Подставляя в (7) вместо $\hat{\zeta}_1$ выражение (5), будем иметь

$$B(1-l) - \sum_{n \geq 0} c_n B(-n-l) = 0, \quad l \leq 0.$$

Значит,

$$\int_0^{2\pi} \left[e^{l\omega - l\omega l} - \sum_{n \geq 0} c_n e^{-in\omega - l\omega l} \right] \hat{f}(\omega) d\omega = 0.$$

Положим

$$u(\omega) = \sum_{n \geq 0} c_n e^{-in\omega}. \quad (8)$$

Тогда

$$\int_0^{2\pi} [e^{l\omega} - u(\omega)] e^{-l\omega l} \hat{f}(\omega) d\omega = 0, \quad l \leq 0. \quad (9)$$

Нужно найти функцию $u(\omega)$ вида (8) так, чтобы выполнялось соотношение (9). Это будет, если

$$[e^{l\omega} - u(\omega)] \hat{f}(\omega) = \sum_{k=1}^{\infty} d_k e^{ik\omega},$$

поскольку $\int_0^{2\pi} e^{ik\omega - l\omega l} d\omega = 0$ при $k \geq 1, l \leq 0$. Поэтому спектральная плотность будет представима в виде

$$\tilde{f}(\omega) = \frac{\sum_{k=1}^{\infty} d_k e^{ik\omega}}{e^{l\omega} - \sum_{n \geq 0} c_n e^{-ln\omega}}.$$

Пусть, например, $\hat{f}(\omega) = |ae^{l\omega} - 1|^2$ при $0 < |a| < 1$. Тогда

$$\begin{aligned} \hat{f}(\omega) &= (1 - ae^{l\omega})(1 - ae^{-l\omega}) = \frac{1 - ae^{l\omega}}{\sum_{n \geq 0} a^n e^{-ln\omega}} = \\ &= \frac{e^{l\omega} - ae^{l2\omega}}{e^{l\omega} + \sum_{n \geq 0} a^{n+1} e^{-ln\omega}}. \end{aligned}$$

Поэтому

$$u(\omega) = - \sum_{n \geq 0} a^{n+1} e^{-ln\omega}.$$

Наилучший прогноз имеет вид

$$\hat{\xi}_1 = - \sum_{n \geq 0} a^{n+1} \xi_{-n}.$$

§ 29. Броуновское движение

Микроскопические частицы, взвешенные в жидкости или газе, под воздействием молекул среды, находящихся в хаотическом тепловом движении, двигаются случайным, очень нерегулярным образом. Это явление получило название *броуновского движения*. Мы рассмотрим одну вероятностную модель этого явления.

Будем исходить из небольшого числа физически оправданных условий.

1. Поскольку в единицу времени количество соударений частицы с молекулами среды огромно (число молекул имеет порядок 10^{20} , их скорости порядка 10^5 см/с при обычных температурах), молекулы, можно считать, двигаются независимо, то перемещения частицы на двух непересекающихся промежутках времени (и вообще при любой последовательности непересекающихся промежутков времени) независимы, так как они вызваны независимыми воздействиями среды (это либо воздействия различных молекул, либо воздействия одной молекулы, но скорость ее успела измениться в результате столкновений с другими молекулами).

2. Среда, в которой движется частица, однородна в пространстве и времени: ее свойства не меняются со временем и в любой точке среды они одинаковы. Кроме того, предположим, что среда изотропна, т. е. ее свойства одинаковы по любым направлениям. Это будет в жидкости или газе, имеющим всюду одинаковую температуру (она характеризует скорости молекул среды) и давление (от него зависит число молекул в единице объема) при отсутствии течений, которые бы указывали преимущественное направление движения.

3. Двигающаяся частица описывает непрерывную траекторию, т. е. не может мгновенно перескочить из одной точки пространства в другую.

Движение точки в пространстве описывается с помощью координат. Пусть $x(t)$, $y(t)$, $z(t)$ — координаты точки в момент t . Из изотропности среды вытекает, что характер изменения всех координат одинаков. Поэтому рассмотрим только одну координату — $x(t)$. Для каждого t $x(t)$ является случайной величиной, значит, $x(t)$ — случайный процесс. Он обладает следующими свойствами:

1. Если $t_1 < t_2 < \dots < t_k$, то случайные величины $x(t_1)$, $x(t_2) - x(t_1)$, \dots , $x(t_k) - x(t_{k-1})$

независимы, так как в силу условия 1 перемещения частицы на непересекающихся промежутках $[0, t_1]$, $[t_1, t_2]$, ..., $[t_{k-1}, t_k]$ независимы.

II. Распределение величины $x(t+h) - x(t)$ не зависит от t (процесс однороден во времени), это вытекает из условия 2;

III. $x(t)$ — непрерывная функция.

Из свойства II вытекает, что распределение $x(t+h) - x(t)$ совпадает с распределением $x(h) - x(0)$. Рассмотрим математическое ожидание $M(x(t) - x(0))$. Поскольку

$$\begin{aligned} M[x(nh) - x(0)] &= \sum_{k=1}^n M[x(kh) - x((k-1)h)] = \\ &= nM(x(h) - x(0)), \end{aligned}$$

то

$$M(x(t) - x(0)) = ta, \quad a = M(x(h) - x(0)).$$

Если бы $a \neq 0$, то существовало бы течение, уносящее частицы в положительном направлении вдоль оси x при $a > 0$ и в обратном — при $a < 0$. Поэтому $a = 0$.

Из независимости величин

$$x(h) - x(0), \quad x(2h) - x(h), \quad \dots, \quad x(nh) - x((n-1)h)$$

вытекает, что дисперсия их суммы равна сумме дисперсий.

Обозначим

$$b(t) = D[x(t) - x(0)].$$

Тогда

$$D[x(kh) - x((k-1)h)] = D[x(h) - x(0)] = b(h).$$

Значит,

$$b(nh) = \sum_{k=1}^n D[x(kh) - x(kh-h)] = nb(h).$$

Поэтому $b(t)$ — линейная функция: $b(t) = bt$.

Величина $x(t) - x(0)$ представима в виде суммы сколь угодно большого числа малых (в силу условия 3) одинаково распределенных независимых приращений:

$$x(t) - x(0) = \sum_{k=1}^n \left[x\left(\frac{k}{n}t\right) - x\left(\frac{k-1}{n}t\right) \right].$$

Поэтому на основании центральной предельной теоремы можно сделать заключение, что $x(t) - x(0)$ имеет нормаль-

ное распределение. Как вычислено выше, среднее значение этого нормального распределения будет 0, а дисперсия — bt . Значит, плотность распределения величины $x(t) - x(0)$ будет

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

Найдем плотность вероятности перехода для процесса $x(t)$, т. е. плотность условного распределения величины $x(t)$, если задано $x(0) = x_0$. Тогда $x(t) = x_0 + [x(t) - x(0)]$ и из формулы (1) вытекает

$$f_t(x_0, x) = \frac{1}{\sqrt{2\pi bt}} e^{-\frac{(x-x_0)^2}{2bt}} \quad (2)$$

(если $x(t) = x$, то $x(t) - x(0) = x - x_0$, поэтому в формулу (1) нужно подставить вместо x величину $x - x_0$).

Коэффициент b , входящий в формулы (1) и (2), называется *коэффициентом диффузии*, он характеризует средний квадрат смещения частицы (вдоль оси x) за единицу времени и зависит от числа соударений и скорости молекул, т. е. от температуры и плотности среды.

Как указывалось, координаты $y(t)$ и $z(t)$ ведут себя точно так, как и координата $x(t)$. Поэтому для плотности распределения величины $y(t) - y(0)$ (а также $z(t) - z(0)$) справедлива формула (1). Вычислим вероятность того, что частица из точки с координатами x_0, y_0, z_0 за время t попадет в параллелепипед, имеющий проекциями на координатные оси отрезки $[x, x + \Delta x]$, $[y, y + \Delta y]$, $[z, z + \Delta z]$. Будем исходить из того, что величины $x(t) - x(0)$, $y(t) - y(0)$ и $z(t) - z(0)$ независимы. (Независимость вытекает из изотропности, далее будут приведены некоторые соображения, поясняющие этот факт.) Тогда в силу формулы (2)

$$P\{x(t) \in [x, x + \Delta x] | x(0) = x_0\} \sim \frac{1}{\sqrt{2\pi bt}} e^{-\frac{(x-x_0)^2}{2bt}} \Delta x,$$

$$P\{y(t) \in [y, y + \Delta y] | y(0) = y_0\} \sim \frac{1}{\sqrt{2\pi bt}} e^{-\frac{(y-y_0)^2}{2bt}} \Delta y,$$

$$P\{z(t) \in [z, z + \Delta z] | z(0) = z_0\} \sim \frac{1}{\sqrt{2\pi bt}} e^{-\frac{(z-z_0)^2}{2bt}} \Delta z,$$

и значит,

$$P \left\{ \begin{array}{l} x(t) \in [x, x + \Delta x], y(t) \in [y, y + \Delta y], \\ z(t) \in [z, z + \Delta z] \end{array} \right\} \left. \begin{array}{l} x(0) = x_0 \\ y(0) = y_0 \\ z(0) = z_0 \end{array} \right\} = \sim (2\pi bt)^{-\frac{3}{2}} e^{-\frac{r^2}{2bt}} \Delta x \Delta y \Delta z,$$

где $r^2 = x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$ — квадрат расстояния между точками (x_0, y_0, z_0) и (x, y, z) . Функция

$$f_t(x_0, y_0, z_0; x, y, z) = (2\pi bt)^{-\frac{3}{2}} e^{-\frac{r^2}{2bt}} \quad (3)$$

называется *трехмерной плотностью вероятности перехода* для процесса броуновского движения.

Вероятность попадания частицы в малый объем, содержащий точку (x, y, z) , в момент времени t при условии, что в начальный момент времени она находилась в точке (x_0, y_0, z_0) , равна произведению трехмерной плотности вероятности перехода на величину объема. Тот факт, что функция (3) зависит лишь от расстояния между точками (x_0, y_0, z_0) и (x, y, z) , и есть математическое выражение изотропности процесса. Таким образом, именно при предположении о независимости координат мы приходим к изотропному процессу.

Броуновское движение — одно из самых распространенных явлений в природе. В броуновском движении непрерывно находятся молекулы жидкости газа, свободные электроны в проводниках, ионы в плазме, благодаря ему происходит диффузия одного вещества в другом, осуществляются питание клеток в организме и вывод из клеток продуктов жизнедеятельности, происходят химические реакции, вообще возможна жизнь. Именно броуновское движение — источник случайности в явлениях макромира.

При рассмотрении задач, относящихся к броуновскому движению, удобно использовать дискретную аппроксимацию движения. Опишем ее для случая одномерного процесса $x(t)$. Будем считать, что $x(t)$ может принимать лишь значения вида $k\Delta$, где Δ — фиксированное достаточно малое число, а k принимает любые целые значения. Частица меняет свое положение через отрезки времени длины h ; h также малое число. Если частица находится в точке $k\Delta$,

то за время h она переходит либо в точку $(k - 1) \Delta$, либо в точку $(k + 1) \Delta$, оба перехода осуществляются с вероятностями $1/2$. Тогда $x(h) - x(0)$ — величина, принимающая значения $\pm \Delta$ с вероятностями $1/2$. При этом

$$M(x(h) - x(0)) = \frac{1}{2} \Delta - \frac{1}{2} \Delta = 0,$$

$$D(x(h) - x(0)) = \frac{1}{2} \Delta^2 + \frac{1}{2} \Delta^2 = \Delta^2.$$

Поскольку для броуновского движения $D(x(h) - x(0)) = bh$, то между h и Δ должна быть такая связь:

$$\Delta^2 = bh, \quad \Delta = \sqrt{bh}.$$

Найдем вероятность того, что за время t частица сдвинется на x . Пусть $t = nh$, $x = m\Delta$. Тогда из n шагов должно быть $\frac{n+m}{2}$ положительных и $\frac{n-m}{2}$ отрицательных. Поскольку шаги независимы, то вероятность этого события

$$C_n^{\frac{n+m}{2}} \left(\frac{1}{2}\right)^n = \frac{n!}{\left(\frac{n+m}{2}\right)! \left(\frac{n-m}{2}\right)!} \frac{1}{2^n}.$$

Как показано в § 9, последнее выражение эквивалентно

$$\frac{1}{\sqrt{2\pi n}} e^{-\frac{m^2}{2n}}.$$

Подставляя $n = \frac{t}{h} = \frac{bt}{\Delta^2}$, $m = \frac{x}{\Delta}$, получим, что эта вероятность эквивалентна выражению

$$\frac{\Delta}{\sqrt{2\pi bt}} e^{-\frac{x^2}{2bt}},$$

которое при малых Δ есть вероятность того, что $x(t) - x(0)$ попадает в интервал $\left(x - \frac{\Delta}{2}, x + \frac{\Delta}{2}\right)$, а это согласуется с видом плотности распределения (1).

Используем дискретную аппроксимацию для определения среднего времени, которое потребуется частице, чтобы выйти из отрезка (α, β) . Пусть $\alpha = A\Delta$, $\beta = B\Delta$, A, B — целые числа. Обозначим через x точку вида $k\Delta$ из отрезка $[\alpha, \beta]$, а через $m(x)$ — среднее время до попадания на границу отрезка (α, β) при условии, что в начальный

момент времени частица находилась в точке x . Тогда $m(\alpha) = m(\beta) = 0$, так как при условии, что $x(0) = \alpha$ или $x(\beta) = \beta$, частица уже в начальный момент времени находится на границе отрезка (α, β) и время до выхода равно нулю. Если $\alpha < x < \beta$, то для достижения границы частица должна сделать по крайней мере 1 шаг, затратив время h , после этого она с вероятностями $\frac{1}{2}$ попадает в одну из точек $x \pm \Delta$. Из точки $x - \Delta$ она затратит в среднем $m(x - \Delta)$ времени для достижения границы (вероятность этого $\frac{1}{2}$), а из точки $x + \Delta$ — среднее время $m(x + \Delta)$ (тоже с вероятностью $\frac{1}{2}$). Значит, $m(x)$ удовлетворяет соотношению

$$m(x) = h + \frac{1}{2} m(x - \Delta) + \frac{1}{2} m(x + \Delta). \quad (4)$$

Поэтому

$$m(x) - m(x - \Delta) = 2h + [m(x + \Delta) - m(x)],$$

$$\begin{aligned} m(\alpha + k\Delta) - m(\alpha + k\Delta - \Delta) &= \\ &= 2h + [m(\alpha + k\Delta + \Delta) - m(\alpha + k\Delta)]. \end{aligned}$$

Последовательность

$$u_k = m(\alpha + k\Delta) - m(\alpha + (k-1)\Delta), \quad k = 1, 2, \dots$$

образует арифметическую прогрессию, разность которой $-2h$, первый член $m(\alpha + \Delta)$. Поэтому

$$\begin{aligned} m(\alpha + k\Delta) &= u_1 + \dots + u_k = \frac{u_1 + u_k}{2} k = \\ &= \frac{m(\alpha + \Delta) + m(\alpha + \Delta) - 2(k-1)h}{2} k = \\ &= km(\alpha + \Delta) - k(k-1)h. \end{aligned}$$

Если $k = B - A$, то

$$\begin{aligned} 0 &= m(B\Delta) = m(\alpha + (B - A)\Delta) = \\ &= (B - A)[m(\alpha + \Delta) + h] - (B - A)^2 h. \end{aligned}$$

Значит, $m(\alpha + \Delta) = (B - A - 1)h$.

Таким образом,

$$m(x) = m\left(\alpha + \frac{x-\alpha}{\Delta}\Delta\right) = \frac{x-\alpha}{\Delta}(B-A)h - \left(\frac{x-\alpha}{\Delta}\right)^2 h$$

(мы вместо k подставили $\frac{x-\alpha}{\Delta}$). Учитывая равенство $h = \frac{\Delta^2}{b}$, окончательно находим

$$m(x) = \frac{(x-\alpha)(\beta-\alpha) - (x-\alpha)^2}{b} = \frac{(x-\alpha)(\beta-x)}{b}.$$

РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА

- Баручс-Рид А. Т.* Элементы теории марковских процессов и их приложения.— М.: Наука, 1969.— 512 с.
- Гнеденко Б. В., Хинчин А. Я.* Элементарное введение в теорию вероятностей.— 6-е изд.— М.: Наука, 1964.— 144 с.
- Карлин С.* Основы теории случайных процессов.— М.: Мир, 1971.— 536 с.
- Кац М.* Вероятность и смежные вопросы в физике.— М.: Мир, 1965.— 408 с.
- Кац М.* Несколько вероятностных задач физики и математики.— М.: Наука, 1967.— 176 с.
- Нейман Ю.* Вводный курс теории вероятностей и математической статистики.— М.: Наука, 1968.— 448 с.
- Смит Дж.* Математические идеи в биологии.— М.: Мир, 1970.— 180 с.
- Смит Дж.* Модели в экологии.— М.: Мир, 1976.— 180 с.
- Феллер В.* Введение в теорию вероятностей и ее приложения.— М.: Мир, 1964.— 498 с.

ОГЛАВЛЕНИЕ

Введение	3
I. Случайные события	
§ 1. Что такое случайное событие?	6
§ 2. Эксперименты с конечным числом исходов. Классическое определение вероятности	13
§ 3. Комбинаторные методы определения вероятности	21
§ 4. Независимость случайных событий	28
§ 5. Геометрические методы. Задача о встрече	35
§ 6. Схема независимых испытаний	43
§ 7. Закон редких событий	47
§ 8. Закон больших чисел	51
§ 9. Нормальная аппроксимация	55
§ 10. Оценка неизвестной вероятности события	61
§ 11. Условные вероятности	64
§ 12. Формула Байесса	75
II. Случайные величины	
§ 13. Примеры случайных величин	80
§ 14. Функция распределения	84
§ 15. Среднее значение	92
§ 16. Дисперсия случайной величины	100
§ 17. Независимые случайные величины	105
§ 18. Закон больших чисел	110
§ 19. Центральная предельная теорема	115
§ 20. Эмпирическая функция распределения	120
§ 21. Корреляция случайных величин	125
III. Случайный процесс	
§ 22. Время ожидания	136
§ 23. Процесс Пуассона	140
§ 24. Процессы с конечным множеством состояний	146
§ 25. Эргодическая теорема	153
§ 26. Ветвящиеся процессы	161
§ 27. Колебания со случайной амплитудой	170
§ 28. Фильтрация и прогиоз	180
§ 29. Броуновское движение	187
Рекомендованная литература	194

Научно-популярная литература

Анатолий Владимирович Скороход

ВЕРОЯТНОСТЬ ВОКРУГ НАС

Утверждено к печати Редакционной коллегией научно-популярной литературы АН УССР

Редакторы *А. Г. Пеккер, А. М. Азаров*
Оформление художника *Ю. В. Бойченко*
Художественный редактор *Б. И. Прищепя*
Технический редактор *С. Г. Максимова*
Корректоры *Л. Я. Постолова,*
Э. Я. Белокопытова, Л. В. Малюта

Информ. бланк № 3635

Сдано в набор 27.02.80, Подп. в печ. 30. 10. 80.
БФ 00188. Формат: 84×108/32. Бумага типогр. № 3.
Лит. гарн. Выс. печ. Усл. печ. л. 10,29. Уч.-изд. л.
9,34. Тираж 29 000 экз. Заказ 533. Цена 30 коп.

Издательство «Наукова думка». 252601, Киев, ГСП,
Репина, 3.

Отпечатано с матриц республиканского производственного объединения «Полиграфкинг» Госкомиздата УССР, Киев, Довженко, 3 в областной книжной типографии львовского облполиграфиздата, Львов, Стефаника, 11. Зак. 3828.

30 коп.

„ НАУКОВА ДУМКА “

