Министерство Российской Федерации
по связи и информатизации

Сибирский государственный университет
телекоммуникаций и информатики

Н. И. Чернова

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Учебное пособие

Новосибирск
2009

УДК 519.2

Доцент, канд. физ.-мат. наук Н. И. Чернова. Теория вероятностей: Учебное пособие / СибГУТИ.— Новосибирск, 2009.— 128 с.

Учебное пособие содержит семестровый курс лекций по теории вероятностей для студентов экономических специальностей. Учебное пособие соответствует требованиям Государственного образовательного стандарта к профессиональным образовательным программам по специальности 080116 — «Математические методы в экономике».

Кафедра ММБП

Табл. 3, рисунков — 21, список лит. — 9 наим.

Рецензенты: А. П. Ковалевский, канд. физ.-мат. наук, доцент кафедры высшей математики НГТУ
В. И. Лотов, д-р физ.-мат. наук, профессор кафедры теории вероятностей и математической статистики НГУ

Для специальности 080116 — «Математические методы в экономике»

Утверждено редакционно-издательским советом СибГУТИ в качестве учебного пособия

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Учебное пособие содержит полный курс лекций по теории вероятностей для студентов Сибирского государственного университета телекоммуникаций и информатики, обучающихся по специальности «Математические методы в экономике». Содержание курса полностью соответствует образовательным стандартам подготовки бакалавров по указанной специальности.

Курс теории вероятностей опирается на стандартные курсы математического анализа, алгебры и дискретной математики, и служит основой для следующего за ним курса математической статистики. Оба они составляют главную часть необходимой базы для основных дисциплин специализации экономистов-математиков — эконометрики, многомерных статистических методов и теории риска.

Задача курса — обучить студентов математическим методам исследования вероятностных экономико-математических моделей, предоставить им необходимый математический аппарат, привить вероятностный взгляд на мир.

Структура курса является общепринятой, содержание более всего соответствует учебному пособию П. П. Бочарова и А. В. Печинкина [1], которую читатель может использовать для более полного знакомства с предметом.

Автор старался не уделять излишнего внимания комбинаторным задачам и аксиоматике. Последняя дана в виде перечня базовых требований к конструкции вероятностных моделей. В дальнейшем измеримость функций и множеств предполагается по умолчанию и нигде отдельно не обсуждается.

Задание вероятностей на дискретном и на «непрерывном» вероятностных пространствах в курсе изучается дважды: при описании простейших вероятностных моделей и при описании распределений случайных величин. Автор надеется, что это облегчит понимание понятий «случайная величина» и «распределение случайной величины».

Многие формулировки теорем и доказательства в курсе изложены недостаточно строго, чтобы не усложнять изложения. Так, свойства, выполняющиеся «почти наверное», сформулированы как выполненные всегда: можно считать, что любые равенства случайных величин в данном курсе понимаются в смысле совпадения на множестве единичной вероятности. Доказательство предельных свойств функций распределения неявно использует свойство непрерывности меры, отсутствующее в курсе. Математическое ожидание вводится не через интеграл Стилтьеса, а отдельно в дискретном и абсолютно непрерывном случаях. Соответственно, корректные доказательства большинства свойств математического ожидания становятся невозможными и опущены. Центральная предельная теорема и теоремы Муавра — Лапласа оставлены без доказательства, поскольку в рамках этого курса любые их доказательства (будь то доказательство методом композиций, методом характеристических функций или производящих функций моментов) сместят акценты в изучении предмета в сторону излишних технических подробностей.

В курс не вошли такие полезные разделы, как условные распределения, линейные преобразования многомерных нормальных распределений и случайные процессы с непрерывным временем. Эти разделы выходят за рамки образовательного стандарта и изучаются в последующих курсах случайных процессов, эконометрики и многомерного статистического анализа.

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ

Учебное пособие предназначено для самостоятельной работы студентов и является дополнением к курсу лекций и практических занятий. Оно будет полезно также любому читателю, желающему освоить предмет самостоятельно. Предполагается, что читатель владеет основами математического анализа, алгебры, исчисления высказываний и дискретной математики.

Курс разбит на 9 глав. Первая глава знакомит читателя с основными понятиями и аксиомами теории вероятностей. Во второй главе рассматриваются примеры дискретных и «непрерывных» вероятностных пространств. Третья глава посвящена условной вероятности и её следствиям. Здесь же обсуждается понятие независимости событий. Формулы для расчёта вероятностей в последовательностях независимых событий предложены в четвёртой главе. Здесь же встречаются первые предельные теоремы.

Пятая глава знакомит читателя с понятием случайной величины и её распределения. Рассматриваются только дискретные и абсолютно непрерывные распределения случайных величин. Здесь же дан перечень стандартных семейств распределений и изучаются их свойства. Глава завершается изучением линейных преобразований случайных величин.

В шестой главе рассмотрены совместные распределения случайных величин. Особый акцент сделан на влиянии совместного распределения на распределения функций от случайных векторов. Независимость случайных величин вводится как частный случай совместного распределения, позволяющий обойтись знанием маргинальных распределений. Здесь же подробно рассмотрено суммирование независимых случайных величин, формула свёртки и следствия из неё.

Глава седьмая посвящена математическим ожиданиям, дисперсиям и другим важным числовым характеристикам распределений. Математические ожидания вводятся отдельно для дискретного и для абсолютно непрерывного распределения, свойства математических ожиданий изложены без строгих доказательств. Здесь же определяются моменты и дисперсия, рассматриваются их свойства, вычисляются моменты основных распределений, в том числе с помощью производящей функции моментов. Далее в этой главе рассмотрен ряд полезных в статистике числовых характеристик распределений и совместных распределений: коэффициенты эксцесса и асимметрии, квантили, ковариация и коэффициент корреляции.

Глава восьмая вновь возвращает читателя к предельным теоремам теории вероятностей, с которыми читатель впервые встретился в четвёртой главе на примере схемы Бернулли. Изложены слабые законы больших чисел в формах П. Л. Чебышёва и Я. Бернулли, центральная предельная теорема для независимых и одинаково распределённых слагаемых, ЦПТ для разнораспределённых слагаемых в форме А. М. Ляпунова.

В последней главе рассмотрены цепи Маркова с дискретным временем.

Приложение содержит таблицы с перечнем основных характеристик дискретных и абсолютно непрерывных распределений, таблицу значений функции распределения стандартного нормального закона.

В конце книги приведен подробный предметный указатель. В списке литературы перечислены учебники, которые можно использовать в дополнение к курсу, и сборники задач для практических занятий.

Изложение материала сопровождается большим числом примеров. Не рекомендуется пропускать эти примеры при чтении, поскольку каждый такой пример помогает понять теоретический материал. Читателю, желающему освоить курс, стоит выполнять все содержащиеся в тексте упражнения и отвечать на заданные вопросы. В конце каждой главы приводится список контрольных вопросов. Вопросы подобраны так, чтобы акцентировать внимание читателя на важных моментах, которые могут остаться незамеченными при первом чтении. Если читателю не удаётся дать ответ на какой-либо вопрос, ему стоит повторно разобрать теоретический материал. Читатель должен также помнить, что освоение предмета невозможно без опыта решения задач. Поэтому параллельно с изучением теоретического материала следует решать задачи из любого сборника задач в списке литературы.

Нумерация параграфов в каждой главе отдельная. Формулы, примеры, утверждения и т. п. имеют сквозную нумерацию. При ссылке на объект из другой главы для удобства читателя указан номер страницы, на которой содержится объект. При ссылке на объект из той же главы приводится только номер формулы, примера, утверждения. Окончание доказательств отмечено значком $\square$.

ВВЕДЕНИЕ

Человек повсюду встречается с процессами или явлениями, результаты которых не могут быть предсказаны заранее. Мы не можем предугадать точно оценку контрольной работы или урожай картофеля, курс доллара через неделю или объём продаж магазина, погоду в течение дня или результат спортивного состязания, выигрыш в лотерею или срок службы электрической лампочки. Основываясь на своих представлениях или на предыдущем опыте, мы пытаемся всякий раз прогнозировать результат таких «опытов». Чаще всего наш прогноз будущего носит вероятностный характер: «завтра, скорее всего, будет солнечно», «рост курса доллара на этой неделе маловероятен», «шансы на выигрыш в лотерею малы».

Теория вероятностей изучает закономерности, возникающие в случайных экспериментах (точнее, в их математических моделях). Но результат случайного эксперимента нельзя предсказать заранее. О каких же закономерностях может идти речь в условиях непредсказуемости результата?

Давайте подбросим монету. Мы не сможем ответить на вопрос, выпадет ли монета вверх гербом или решкой. Результат индивидуального опыта непредсказуем. Но подбросим монету снова и снова. Конечно, мы не сможем сказать заранее, сколько гербов выпадет, если мы будем бросать монету 1 000 раз. Но количество выпавших гербов не должно очень отличаться от 500 в силу полной симметрии монеты: герб имеет столько же шансов выпасть, сколько и решка. Более того, частота выпадения герба (отношение количества выпавших гербов к числу всех испытаний) при продолжении испытаний должна всё меньше отличаться от $0{,}5$, т. е. от вероятности выпадения герба.

Так, Г. Бюффон в XVIII веке провёл 4040 подбрасываний монеты. Герб выпал 2048 раз, и частота выпадения герба составила 0,508. К. Пирсон в XIX веке бросал монету 24000 раз, при этом герб выпал в 12012 случаях. Частота выпадения герба равна 0,5005[1].

Такая закономерность наблюдается не только при бросании монеты: вероятность любого события можно трактовать как долю случаев, когда это событие происходит при многократном повторении опыта.

[1] Данные заимствованы из [3].

Подведём итог. Не все случайные явления (эксперименты) можно изучать методами теории вероятностей, а лишь те, которые могут быть воспроизведены в одних и тех же условиях, хотя бы мысленно, какое угодно число раз. Случайность и хаос — не одно и то же. В случайных экспериментах можно обнаружить закономерности, главная из которых — свойство «статистической устойчивости»: если некоторое событие A может произойти в результате эксперимента, то доля $n(A)/n$ экспериментов, в которых данное событие произошло, с ростом общего числа экспериментов приближается к некоторому числу $\mathsf{P}(A)$. Это число служит объективной характеристикой «степени возможности» событию A произойти.

Возникает соблазн принять за вероятность события предел частоты этого события при повторении опытов. Но частотное определение (его связывают с именем Р. Мизеса) неудобно по многим причинам. Прежде всего потому, что частота появления события в большой, но конечной, серии меняется от серии к серии испытаний. А бесконечную последовательность испытаний, в которой частота наконец перешла бы в точную вероятность, реализовать невозможно.

В 1900 г. на II Международном Математическом Конгрессе в Париже Д. Гильберт сформулировал проблемы, стоящие перед математикой XX века. Шестая из двадцати трёх знаменитых ныне проблем Гильберта касалась теории вероятностей: «аксиоматизировать те физические науки, в которых математика играет важную роль, в первую очередь теорию вероятностей и механику». Эта проблема (в части, касающейся теории вероятностей) была решена А. Н. Колмогоровым в 1923 г. Предложенная им аксиоматика трактует вероятность как функцию, заданную на подмножествах некоторого множества и обладающую рядом свойств. С аксиоматикой теории вероятностей мы познакомимся в первой главе.

Сформулируем вывод: теория вероятностей изучает не реальные эксперименты, а лишь их математические модели. Вопросы определения вероятностей событий в реальном мире, т. е. выбора математической модели данного случайного явления, лежат за рамками теории вероятностей как математической науки. Такие вопросы решает, в частности, математическая статистика, к изучению которой мы приступим, как только освоим её инструментарий — теорию вероятностей.

ГЛАВА I

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Изучение закономерностей в случайных экспериментах должно начинаться с построения математической модели этих экспериментов. В теории вероятностей любая модель начинается с задания вероятностного пространства. Сначала следует выяснить, каково множество всех возможных результатов данного эксперимента. Затем решить, какие подмножества этого множества интересуют нас в качестве событий, и как мы будем определять степень возможности этих событий. Поговорим об всех этих этапах подробнее.

§ 1. Пространство элементарных исходов

Определение 1. *Пространством элементарных исходов* называется множество Ω, содержащее все возможные взимоисключающие результаты данного случайного эксперимента. Элементы множества Ω называются *элементарными исходами* и обозначаются буквой ω.

Отметим сразу, что *любое* множество Ω можно считать пространством элементарных исходов какого-то случайного эксперимента.

Определение 2. *Событиями* называются подмножества множества Ω. Говорят, что *произошло событие* A, если эксперимент завершился одним из элементарных исходов, входящих в множество A.

Итак, элементарный исход — это мельчайший неделимый результат эксперимента, а событие может состоять из одного или сразу нескольких исходов.

Напомним, что конечные и счётные множества удобно задавать перечислением их элементов. Например, $\Omega = \{1, 2, \ldots, 100\}$ — множество, состоящее из первых ста натуральных чисел. Несчётные множества обычно задают указанием свойства, которым обладают все элементы множества. Так, $\Omega = \{\omega \in \mathbb{R} \mid \omega^2 < 9\}$ — множество действительных чисел из интервала $(-3, 3)$.

Пример 1. Один раз подбрасывают игральную кость. Рассмотрим пространство элементарных исходов $\Omega = \{1, 2, 3, 4, 5, 6\} =$ {⚀, ⚁, ⚂, ⚃, ⚄, ⚅}. Элементарные исходы здесь соответствуют числу выпавших очков.

Событие $A = \{1, 2\} =$ {⚀, ⚁} произойдёт, если выпадет одно или два очка; событие $B = \{1, 3, 5\} =$ {⚀, ⚂, ⚄} означает, что выпадет нечётное число очков. Событие $C = \{6\} = \{⚅\}$ состоит из одного элементарного исхода и означает появление шести очков.

Пример 2. Подбрасываются две игральные кости. Будем считать их различимыми и назовём одну из них первой, другую — второй. Пространством элементарных исходов является множество пар чисел $(i,\, j)$, где i — число очков, выпавших на первой кости, j — на второй. В этом множестве $6 \times 6 = 36$ элементарных исходов:

$$\begin{array}{cccc} (1,1) & (1,2) & \cdots & (1,6) \\ (2,1) & (2,2) & \cdots & (2,6) \\ \multicolumn{4}{c}{\cdots\cdots\cdots} \\ (6,1) & (6,2) & \cdots & (6,6) \end{array} \tag{1}$$

Заметим, что для симметричных костей все эти 36 исходов равновозможны: ни одна из этих комбинаций не имеет больше шансов выпасть, чем другая. Действительно, на первой кости с равными шансами выпадает любая грань. Это означает, что результат бросания двух костей имеет столько же шансов оказаться в первой строке матрицы (1), что и во второй, в третьей и т. д. Но на второй кости снова с одинаковыми шансами выпадает любая грань, поэтому и каждое место в строке равновозможно.

Событие «на первой кости выпадет одно очко» можно записать так: $A = \{(1,1),\, (1,2),\, (1,3),\, (1,4),\, (1,5),\, (1,6)\}$; событие «на второй кости выпадет одно очко» запишется так: $B = \{(1,1),\, (2,1),\, (3,1),\, (4,1),\, (5,1),\, (6,1)\}$; событие $C = \{(2,2),\, (3,1),\, (1,3)\}$ означает, что сумма выпавших очков равна четырём; событие $D = \{(1,1),\, (2,2),\, (3,3),\, (4,4),\, (5,5),\, (6,6)\}$ произойдёт, если на костях выпадет одинаковое число очков.

Пример 3. Подбрасываются две *неразличимые* игральные кости. Элементарными исходами будем считать пары чисел $(i,\, j)$, где $i \leqslant j$. Например, элементарный исход $(1,2)$ случается, если на одной из костей выпадает одно очко, на другой — два очка. В множестве Ω двадцать один исход:

$$\begin{array}{cccc} (1,1) & (1,2) & \cdots & (1,6) \\ & (2,2) & \cdots & (2,6) \\ & & \cdots\cdots & \\ & & & (6,6) \end{array}$$

Для симметричных костей эти исходы равновозможными уже не будут: например, исход $(1,2)$ имеет вдвое больше шансов появиться, чем исход $(1,1)$. Мы просто перестали различать исходы из примера 2, симметричные друг другу относительно главной диагонали матрицы (1).

Теперь событие «сумма выпавших очков равна четырём» состоит из двух элементарных исходов $(2,2)$ и $(1,3)$. Событие «на костях выпадет одинаковое число очков» по-прежнему включает шесть исходов. Слова «на первой кости выпадет одно очко» никакого события уже не описывают, а событие $A = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)\}$ означает, что хотя бы на одной из костей выпало одно очко (ср. с примером 2).

Пример 4. На поверхность стола бросается монета. Результатом эксперимента можно считать положение центра монеты. Пространство элементарных исходов такого эксперимента есть множество всех точек стола. Оно бесконечно и несчётно. Событием можно назвать, например, попадание центра монеты на лист бумаги, лежащий на столе.

Пример 5. Монета подбрасывается до тех пор, пока не выпадет вверх гербом. Пространство элементарных исходов является бесконечным, но счётным множеством: $\Omega = \{\ \textit{г, рг, ррг, рррг, ррррг, рррррг,}\ \ldots\}$, где $\textit{р}$ означает выпадение решки, а $\textit{г}$ — выпадение герба при одном подбрасывании. Событие «герб выпал при броске с чётным номером» выглядит так:

$$A = \{\ \textit{рг, рррг, рррррг,}\ \ldots\ \}.$$

Пример 6. В коробке лежат один чёрный и два белых шара. Из коробки достают наугад один шар.

Можно определить два разных пространства элементарных исходов. Первое из них состоит из двух исходов $\Omega_1 = \{\textit{б},\ \textit{ч}\}$ — мог появиться белый шар или чёрный. Эти исходы, очевидно, не будут равновозможными: появление белого шара вдвое вероятнее, чем появление чёрного.

Если мы хотим иметь дело с равновозможными элементарными исходами, шары следует занумеровать (или различать как-то иначе). Тогда множество $\Omega_2 = \{\textit{б}_1,\ \textit{б}_2,\ \textit{ч}\}$ будет состоять из трёх элементарных исходов.

Пример 7. В коробке лежат один чёрный и два белых шара. Из коробки достают наугад два шара. Пусть порядок следования шаров нам безразличен.

Занумеруем шары, чтобы элементарные исходы были равновозможными. Пространство элементарных исходов состоит из 3 элементов:

$$\Omega = \{(\textit{б}_1, \textit{б}_2),\ (\textit{б}_1, \textit{ч}),\ (\textit{б}_2, \textit{ч})\}.$$

Событие «вынуты два белых шара» включает один исход $\omega_1 = (\textit{б}_1, \textit{б}_2)$, а событие «вынуты разноцветные шары» — два исхода: $\omega_2 = (\textit{б}_1, \textit{ч})$, $\omega_3 = (\textit{б}_2, \textit{ч})$.

Можно, как в примере 6, рассмотреть пространство элементарных исходов, состоящее из двух элементов: $\Omega_1 = \{(\textit{б}, \textit{б}),\ (\textit{б}, \textit{ч})\}$ — вынуты два белых шара или шары разных цветов. Но в таком пространстве второй исход имеет вдвое больше шансов случиться, чем первый.

Операции над событиями. В теории вероятностей рассматривают те же операции над событиями (множествами), что и в теории множеств. Дадим определения новым событиям — результатам этих операций.

Объединением $A \cup B$ событий A и B называется событие, состоящее в том, что из двух событий A и B случилось хотя бы одно. Это событие включает как элементарные исходы из множества A, так и элементарные исходы из множества B.

Пересечением $A \cap B$ событий A и B называется событие, состоящее в том, что произошли сразу оба события A и B. Это событие содержит элементарные исходы, каждый из которых принадлежит и множеству A, и множеству B. Вместо $A \cap B$ часто пишут просто AB.

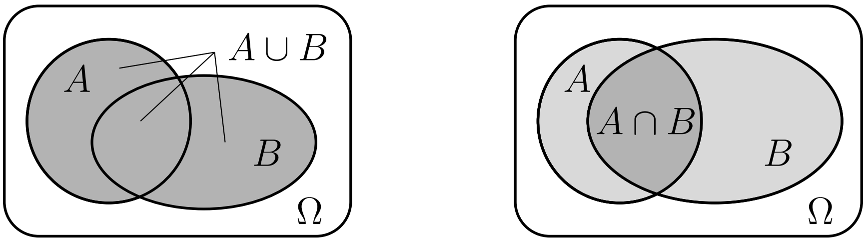

Рис. 1. Объединение и пересечение событий

Дополнением $A \setminus B$ события B до A называется событие, состоящее в том, что произошло событие A, но не произошло B. Событие $A \setminus B$ содержит элементарные исходы, входящие в множество A, но не входящие в B.

Противоположным (или дополнительным) к событию A называется событие $\overline{A} = \Omega \setminus A$, состоящее в том, что событие A не произошло. Событие $\overline{A}$ состоит из элементарных исходов, не входящих в множество A.

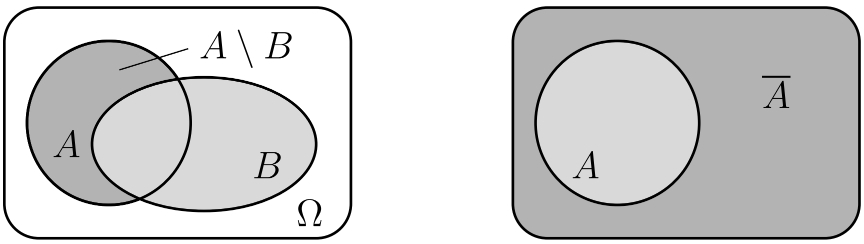

Рис. 2. Дополнение и противоположное событие

Выделим среди подмножеств Ω два особых события.

Достоверным называется событие, которое обязательно происходит в результате эксперимента, т. е. единственное событие, включающее все элементарные исходы — событие Ω.

Невозможным называется событие, которое не может произойти в результате эксперимента, т. е. событие, не содержащее ни одного элементарного исхода — «пустое множество» $\varnothing$.

Очевидно, что $\overline{\Omega} = \varnothing$, $\overline{\varnothing} = \Omega$, $A \cap \overline{A} = \varnothing$, $A \cup \overline{A} = \Omega$. Объединение множеств аналогично сложению чисел, пересечение — умножению. Свойства операций над событиями похожи на свойства умножения и сложения чисел:

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

Первое из этих равенств очевидно, а второе, хоть и противоречит интуиции, всё же верно. Кроме того, объединение и пересечение событий связаны очень важными соотношениями двойственности:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}.$$

П р и м е р 8. Пусть событие A_i означает, что i-я деталь бракованная, где $1 \leqslant i \leqslant 3$ — номер детали. Запишем с помощью операций над событиями событие A — «ровно две из трёх деталей бракованные»:

$$A = \left(A_1 \cap A_2 \cap \overline{A_3}\right) \cup \left(A_1 \cap \overline{A_2} \cap A_3\right) \cup \left(\overline{A_1} \cap A_2 \cap A_3\right).$$

Выше записано буквально следующее: *либо* первая и вторая детали бракованные, а третья годная, *либо* первая и третья детали бракованные, а вторая годная, *либо* вторая и третья детали бракованные, а первая годная.

Событие $B = \overline{A_1} \cap \overline{A_2} \cap \overline{A_3}$ означает «все три детали годные».

Событие «хотя бы одна деталь из трёх бракованная» можно записать двумя способами: $C = A_1 \cup A_2 \cup A_3 = \overline{B}$.

Отношения между событиями. Множества могут пересекаться или не пересекаться, быть включены одно в другое или не быть. В теории вероятностей эти отношения событий носят особые названия.

События A и B называются *несовместными*, если они не могут произойти одновременно: $A \cap B = \varnothing$.

События $A_1, \ldots, A_n$ называются *попарно несовместными*, если несовместны любые два из них, т. е. $A_i \cap A_j = \varnothing$ для любых $i \neq j$.

Говорят, что событие A *влечёт* событие B, и пишут $A \subseteq B$, если всегда, как только происходит событие A, происходит и событие B. Это означает, что любой элементарный исход, входящий в множество A, одновременно входит и в множество B, т. е. A содержится в B.

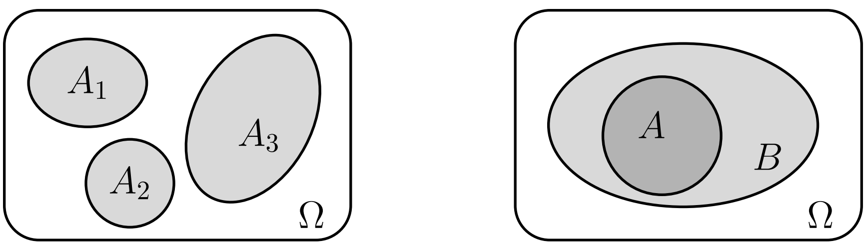

Рис. 3. Попарно несовместные и вложенные события

Два события *совпадают*, если они состоят из одних и тех же элементарных исходов. Можно выразить это иначе: $A = B$, если $A \subseteq B$ и $B \subseteq A$.

Пример 9. При бросании двух игральных костей события «сумма очков равна четырём» и «на первой кости выпало шесть очков» несовместны.

Событие «сумма очков равна двум» влечёт за собой событие «на костях выпало одинаковое число очков». Действительно, сумма очков равна двум лишь при выпадении двух единиц. Но тогда на костях выпадет одинаковое число очков. Обратное включение неверно: не всегда, когда на костях выпадает одинаковое число очков, сумма этих очков равна двум.

Событие «сумма очков меньше пяти» влечёт за собой событие «сумма очков меньше семи».

§ 2. Поле событий

Множество событий может состоять из всех мыслимых подмножеств Ω, но может быть и более узким набором подмножеств. Сужение набора событий бывает необходимо, если Ω является несчётным множеством, и число всех его подмножеств слишком велико: тогда возможны проблемы с корректным определением их вероятностей. Мы не будем подробно касаться этих проблем, но всё же выделим специальное множество событий, т. е. набор таких подмножеств Ω, вероятности которых мы зададим впоследствии.

Итак, *событиями* мы будем называть не любые подмножества Ω, а лишь элементы некоторого множества $\mathcal{F}$. Необходимо, чтобы это множество $\mathcal{F}$ было *замкнуто* относительно всех операций над событиями, т. е. чтобы объединение, пересечение, дополнение событий снова давало событие.

Определение 3. Множество $\mathcal{F}$, состоящее из подмножеств множества Ω, называется *полем событий*, если выполнены следующие условия:

(A1) $\mathcal{F}$ содержит достоверное и невозможное событие: $\Omega \in \mathcal{F}$, $\varnothing \in \mathcal{F}$;

(A2) вместе с любым событием $\mathcal{F}$ содержит и противоположное к нему: если $A \in \mathcal{F}$, то $\overline{A} \in \mathcal{F}$;

(A3) вместе с любым конечным или счётным набором событий $\mathcal{F}$ содержит их объединение и пересечение: если $A_1, A_2, \ldots \in \mathcal{F}$, то

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F} \quad \text{и} \quad \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}.$$

Пример 10. Пусть $\Omega = \{1, 2, 3, 4, 5, 6\}$ — пространство элементарных исходов (например, при бросании игрального кубика). Следующие наборы подмножеств Ω являются полями событий (*доказать!*):

1. $\mathcal{F} = \big\{\Omega, \varnothing\big\} = \big\{\{1, 2, 3, 4, 5, 6\}, \varnothing\big\}$ — *тривиальное* поле событий.

2. $\mathcal{F} = \bigl\{\Omega,\ \varnothing,\ \{1\},\ \overline{\{1\}}\bigr\} = \bigl\{\{1,2,3,4,5,6\},\ \varnothing,\ \{1\},\ \{2,3,4,5,6\}\bigr\}$.
3. $\mathcal{F} = \bigl\{\Omega,\ \varnothing,\ \{1,2,3\},\ \overline{\{1,2,3\}}\bigr\} = \bigl\{\Omega,\ \varnothing,\ \{1,2,3\},\ \{4,5,6\}\bigr\}$.
4. $\mathcal{F}$ — множество всех подмножеств Ω.

Упражнение. Доказать, что если Ω состоит из n элементов, то в множестве *всех его подмножеств* ровно 2^n элементов.

Итак, мы определили специальный класс $\mathcal{F}$ подмножеств пространства элементарных исходов Ω, названный полем событий и обладающий важным свойством: применение счётного числа любых операций к множествам из $\mathcal{F}$ снова дает множество из $\mathcal{F}$ (не выводит за рамки этого класса).

Определим теперь *вероятность* как числовую функцию, определённую на множестве событий, т. е. функцию, которая каждому событию ставит в соответствие число — его вероятность.

§ 3. Вероятность и её свойства

Пусть даны пространство элементарных исходов Ω и поле событий $\mathcal{F}$.

Определение 4. *Вероятностью* или *вероятностной мерой* называется функция P, действующая из $\mathcal{F}$ в множество действительных чисел $\mathbb{R}$ и обладающая свойствами (их называют «аксиомами вероятности»):

(P1) $0 \leqslant \mathsf{P}(A) \leqslant 1$ для любого события A;

(P2) Для любого конечного или счётного набора *попарно несовместных* событий $A_1,\ A_2,\ A_3,\ \ldots$ имеет место равенство
$$\mathsf{P}\left(\bigcup A_i\right) = \sum \mathsf{P}(A_i)\,;$$

(P3) Вероятность достоверного события равна единице: $\mathsf{P}(\Omega) = 1$.

Напомним, что события попарно несовместны, если никакие два из них не пересекаются. Свойство (P2) для совместных событий может быть не выполнено: например, для событий $A_1 = A_2 = \Omega$
$$\mathsf{P}(A_1 \cup A_2) = \mathsf{P}(\Omega) = 1 \neq 2 = \mathsf{P}(\Omega) + \mathsf{P}(\Omega) = \mathsf{P}(A_1) + \mathsf{P}(A_2).$$

Свойство (P1) означает неотрицательность вероятности, свойство (P3) — нормированность. Свойство (P2) называют свойством *счётной аддитивности* вероятности или *расширенной аксиомой сложения.* Можно переформулировать определение так: *вероятностью называется неотрицательная, нормированная, счётно-аддитивная функция событий.*

Определение 5. Тройка объектов $\langle\Omega, \mathcal{F}, \mathsf{P}\rangle$, в которой Ω — пространство элементарных исходов, $\mathcal{F}$ — поле событий и P — вероятностная мера, называется *вероятностным пространством*.

Задание вероятностного пространства — первый необходимый шаг при построении математической модели любого эксперимента.

Теорема 1. *Вероятность обладает следующими свойствами.*

1. $\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A)$, $\mathsf{P}(\varnothing) = 0$.
2. *Если* $A \subseteq B$, *то* $\mathsf{P}(A) \leqslant \mathsf{P}(B)$ (*монотонность вероятности*).
3. $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$.
4. $\mathsf{P}(A \cup B) \leqslant \mathsf{P}(A) + \mathsf{P}(B)$.
5. *Для любых* $A_1, \ldots, A_n$ *имеет место формула включения-исключения*:

$$\mathsf{P}(A_1 \cup \ldots \cup A_n) = \sum_{i=1}^{n} \mathsf{P}(A_i) - \sum_{i<j} \mathsf{P}(A_i A_j) + \\ + \sum_{i<j<m} \mathsf{P}(A_i A_j A_m) - \ldots + (-1)^{n-1} \mathsf{P}(A_1 A_2 \ldots A_n). \quad (2)$$

Доказательство. 1. Достоверное событие $\Omega = A \cup \overline{A}$ есть объединение двух несовместных событий A и $\overline{A}$. Из аксиом (P2) и (P3) получим

$$1 = \mathsf{P}(\Omega) = \mathsf{P}(A) + \mathsf{P}(\overline{A}), \quad \mathsf{P}(\varnothing) = 1 - \mathsf{P}(\Omega) = 0.$$

2. Представим событие B в виде объединения двух несовместных событий: $B = A \cup (B \setminus A)$. По аксиомам (P2) и (P1),

$$\mathsf{P}(B) = \mathsf{P}(A) + \mathsf{P}(B \setminus A) \geqslant \mathsf{P}(A),$$

поскольку $\mathsf{P}(B \setminus A) \geqslant 0$.

3. Событие $A \cup B$ можно разложить в объединение двух несовместных событий $A \cup B = A \cup (B \setminus AB)$. Событие $B = (B \setminus AB) \cup AB$ тоже складывается из двух несовместных событий. По аксиоме (P2) получаем, что $\mathsf{P}(B) = \mathsf{P}(B \setminus AB) + \mathsf{P}(AB)$ и

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B \setminus AB) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB).$$

4. Доказательство очевидно:

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB) \leqslant \mathsf{P}(A) + \mathsf{P}(B).$$

5. Докажем свойство 5 для $n = 3$. Трижды воспользуемся свойством 3:

$$\mathsf{P}(A \cup B \cup C) = \mathsf{P}\big((A \cup B) \cup C)\big) = \mathsf{P}(A \cup B) + \mathsf{P}(C) - \mathsf{P}\big((A \cup B) \cap C\big) = \\ = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB) + \mathsf{P}(C) - \mathsf{P}\big(AC \cup BC\big) = \\ = \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(AB) - \mathsf{P}(AC) - \mathsf{P}(BC) + \mathsf{P}(ABC).$$

Читатель докажет свойство 5 для произвольного n методом математической индукции. □

§ 4. Контрольные вопросы

1. Чем случайный эксперимент отличается от детерминированного?

2. Что такое пространство элементарных исходов?

3. Могут ли два элементарных исхода одновременно случиться в опыте?

4. Игральную кость подбрасывают один раз. Перечислить все элементарные исходы эксперимента.

5. Игральную кость подбрасывают дважды. Перечислить все элементарные исходы эксперимента.

6. Монету подбрасывают трижды. Перечислить все элементарные исходы эксперимента.

7. Из четырёх разных книг на полке берут две. Перечислить все элементарные исходы эксперимента.

8. Что такое событие? Достоверное событие? Невозможное событие?

9. Что такое объединение двух событий? Пересечение?

10. Записать событие, состоящее в том, что из событий A, B, C произошло хотя бы одно.

11. Записать событие, состоящее в том, что из событий A, B, C не произошло хотя бы одно.

12. Записать событие, состоящее в том, что случились все три события A, B, C одновременно.

13. Записать событие, состоящее в том, что событие A произошло, а события B и C не произошли.

14. Когда дополнение события B до события A является невозможным событием? Совпадает с A?

15. Что дают в объединении событие и противоположное к нему? В пересечении?

16. Пусть $\overline{A} = \Omega$. Назовите A.

17. Пусть $\overline{A} = \varnothing$. Назовите A.

18. Что такое формулы двойственности?

19. Чему равно дополнение к объединению событий?

20. В каком случае два события несовместны?

21. Когда говорят про попарную несовместность? Что это за свойство?

22. Чему равно пересечение трёх попарно несовместных событий?

23. Что больше: объединение или пересечение событий?

24. Объединение двух событий влечёт их пересечение или наоборот?

25. $A = \{1, 2\}, B = \{1\}$. Какое из отношений верно: $A \subset B$ или $B \subset A$?

26. Дайте определение поля событий.

27. Является ли множество всех подмножеств полем событий?

28. Может ли поле событий состоять из одного события? Из двух? Из трёх? Из четырёх? Если может, привести примеры.

29. Докажите, что $A \setminus B \in \mathfrak{F}$, если A, $B \in \mathfrak{F}$.

30. Укажите область определения и область значений вероятностной меры.

31. Каких значений не может принимать вероятность?

32. Чему равна вероятность достоверного события? Невозможного?

33. Что такое нормированность вероятности?

34. Что такое неотрицательность вероятности?

35. Что такое счётная аддитивность вероятности?

36. Зачем в свойстве счётной аддитивности требуется попарная несовместность событий?

37. Чему равна вероятность объединения двух произвольных событий?

38. Что такое монотонность вероятности?

39. Как связаны вероятности прямого и противоположного событий?

40. Что больше: вероятность объединения или сумма вероятностей? Когда эти числа равны?

ГЛАВА II

ПРИМЕРЫ ВЕРОЯТНОСТНЫХ ПРОСТРАНСТВ

В этой главе мы опишем все практически значимые виды вероятностных пространств и способы задания вероятности на них. Вероятность по своим свойствам подобна массе. Килограмм массы можно распределить, либо поместив положительную массу в каждую точку некоторого дискретного множества точек, либо «размазав» весь килограмм по некоторому «непрерывному» множеству точек.

§ 1. Дискретное пространство элементарных исходов

Пространство элементарных исходов называется *дискретным*, если множество Ω конечно или счётно: $\Omega = \{\omega_1,\, \omega_2,\, \ldots,\, \omega_n,\, \ldots\}$.

Так, все эксперименты из примеров § 1 главы I, кроме примера 4, приводят к дискретным пространствам элементарных исходов.

Событием на таком пространстве будем считать *любое* подмножество Ω. Чтобы определить вероятность события, присвоим вероятность каждому элементарному исходу в отдельности. Иначе говоря, снабдим вероятностями мельчайшие «кирпичики» — элементарные исходы, из которых составляется любое событие. Тогда вероятность любого события определяется как сумма вероятностей входящих в него элементарных исходов.

Определение 6. Сопоставим каждому исходу $\omega_i \in \Omega$ число $p_i \in [0, 1]$ так, чтобы $p_1 + p_2 + \ldots = 1$. Вероятностью события A называется число

$$\mathsf{P}(A) = \sum_{\omega_i \in A} p_i,$$

равное сумме вероятностей элементарных исходов, входящих в множество A. В случае $A = \varnothing$ положим $\mathsf{P}(A) = 0$.

Приведём пример задания вероятностей на дискретном пространстве.

Пример 11. В эксперименте из примера 5 монета подбрасывается до первого выпадения герба. Присвоим элементарным исходам следующие вероятности:

$$\begin{array}{llllll} \omega_i: & \textit{г}, & \textit{рг}, & \textit{ррг}, & \textit{рррг}, & \ldots \\ p_i: & \dfrac{1}{2}, & \dfrac{1}{4}, & \dfrac{1}{8}, & \dfrac{1}{16}, & \ldots \end{array}$$

Легко проверить, что сумма вероятностей элементарных исходов равна 1: по формуле суммы бесконечной убывающей геометрической прогрессии,

$$p_1 + p_2 + \ldots = \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1/2}{1-1/2} = 1.$$

Вероятность события $A = \{\omega_2, \omega_4, \ldots\}$ (герб выпал при броске с чётным номером) равна:

$$\mathsf{P}(A) = p_2 + p_4 + \ldots = \sum_{i=1}^{\infty} \frac{1}{2^{2i}} = \frac{1/4}{1-1/4} = \frac{1}{3}.$$

Заданные выше вероятности соответствуют, как мы увидим в дальнейшем, подбрасыванию правильной монеты. Можно было задать вероятности как-нибудь иначе: например, $p_i = 2^{i-1}/3^i$. Такие вероятности отвечали бы бросанию утяжелённой монеты, герб на которой выпадает в среднем в одном случае из трёх.

Внимательный читатель уже заметил, что если множество Ω счётно, но не конечно, присвоить всем элементарным исходам одну и ту же вероятность нельзя (*почему*?). Для конечного же множества Ω всегда возможно задать одинаковые вероятности исходов, что мы сейчас и сделаем.

Классическая вероятность. Частным, но часто встречающимся в жизни случаем дискретного вероятностного пространства является классическая вероятностная схема.

Предположим, что мы имеем дело с пространством элементарных исходов, состоящим из конечного числа элементов: $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$, и из каких-то соображений можем считать элементарные исходы *равновозможными*. Равновозможность возникает обычно в силу симметрии в эксперименте (симметричная монета, хорошо перемешанная колода карт, правильная игральная кость, отсутствие оснований предпочесть один результат другому).

Говорят, что эксперимент описывается *классической вероятностной моделью*, если пространство его элементарных исходов состоит из конечного числа равновозможных исходов.

Тогда вероятность любого элементарного исхода равна $\frac{1}{N}$. Если событие $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ состоит из k элементарных исходов, то вероятность этого события равна отношению $\frac{k}{N}$:

$$\mathsf{P}(A) = p_{i_1} + \ldots + p_{i_k} = k \cdot \frac{1}{N} = \frac{|A|}{|\Omega|}. \tag{3}$$

Здесь символом $|A|$ обозначено число элементов конечного множества A.

Формулу $\mathsf{P}(A) = \dfrac{|A|}{|\Omega|}$ называют *классическим определением вероятности* и читают так: «вероятность события A равна отношению числа исходов, *благоприятствующих* событию A, к общему числу равновозможных исходов». Полезно сравнить это определение с формулировкой автора определения, Я. Бернулли: «Вероятность есть степень достоверности и отличается от неё как часть от целого» («Ars Conjectandi», 1713 г.)

Мы видим, что вычисление вероятности в классической схеме сводится к подсчёту общего числа «шансов» и числа шансов, благоприятствующих событию. Число шансов считают с помощью формул комбинаторики. Читатель, желающий освежить в памяти эти формулы, может сейчас же обратиться к следующему параграфу, где объяснены основные принципы и формулы комбинаторики, после чего вернуться к нам.

А мы пока рассмотрим стандартные урновые схемы: из n шаров выбирают k шаров. Будем исходить из предположения о том, что появление любого шара равновозможно. Тогда три схемы: схема выбора с возвращением и с учётом порядка, выбора без возвращения и с учётом порядка, а также выбора без возвращения и без учёта порядка, описываются классической вероятностной моделью. Общее число равновозможных элементарных исходов в этих схемах равно соответственно n^k, A_n^k и C_n^k.

Как показывает следующий пример, последняя схема — схема выбора с возвращением и без учёта порядка — имеет неравновозможные исходы. Поэтому классическое определение вероятности для неё не применимо.

Пример 12. Рассмотрим выбор двух шариков из двух или, что то же самое, дважды подбросим монету. Если учитывать порядок, то исходов получится четыре, и все они равновозможны, т.е. имеют вероятность по $\dfrac{1}{4}$:

(герб, герб), (решка, решка), (решка, герб), (герб, решка).

Если порядок не учитывать, то следует объявить два последних исхода одним и тем же результатом эксперимента и получить не четыре, а три исхода:

(два герба), (две решки), (один герб и одна решка).

Первые два исхода имеют вероятности по $\dfrac{1}{4}$, а вероятность последнего равна $\dfrac{1}{4} + \dfrac{1}{4} = \dfrac{1}{2}$. Видим, что при выборе с возвращением и без учёта порядка элементарные исходы оказываются неравновозможными.

Упражнение. Сравнить примеры 2 и 3. В каком из них перечислены равновозможные элементарные исходы? Найти вероятности всех элементарных исходов в примере 3. Равны ли они $\dfrac{1}{21}$?

В следующем примере разобрана классическая задача, приводящая к так называемому *гипергеометрическому распределению*.

П р и м е р 13. Из урны, в которой K белых и $N-K$ чёрных шаров, наудачу и без возвращения вынимают n шаров, где $n \leqslant N$. Термин «наудачу» означает, что появление любого набора из n шаров равновозможно. Найти вероятность того, что будет выбрано k белых и $n-k$ чёрных шаров.

Р е ш е н и е. Результат эксперимента — какой-то набор из n шаров. Можно не учитывать порядок следования шаров в наборе. Общее число элементарных исходов есть число n-элементных подмножеств множества, состоящего из N элементов: $|\Omega| = C_N^n$. Обозначим через A_k событие, состоящее в том, что в наборе окажется k белых шаров и $n-k$ чёрных. Пусть $k \leqslant K$ и $n-k \leqslant N-K$, иначе $\mathsf{P}(A_k) = 0$.

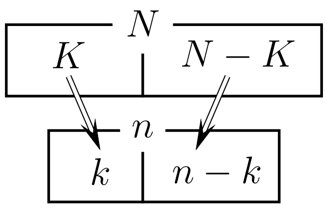

Рис. 4. Выбор n шаров из N

Есть ровно C_K^k способов выбрать k белых шаров из K и C_{N-K}^{n-k} способов выбрать $n-k$ чёрных шаров из $N-K$. Каждый возможный набор выбранных белых шаров можно комбинировать с каждым возможным набором чёрных. Поэтому число благоприятных исходов равно $|A_k| = C_K^k\, C_{N-K}^{n-k}$,

$$\mathsf{P}(A_k) = \frac{|A_k|}{|\,\Omega\,|} = \frac{C_K^k\, C_{N-K}^{n-k}}{C_N^n}\,. \tag{4}$$

Вычисляя вероятность событий A_k, мы сопоставили каждому набору из k белых и $n-k$ чёрных шаров вероятность получить этот набор при выборе шаров из урны. Набор вероятностей (4) называется *гипергеометрическим распределением* вероятностей.

Здесь мы в первый, но далеко не в последний раз встретились с термином «распределение» вероятностей. Это слово всегда обозначает некий способ разделить (распределить) общую единичную вероятность между какими-то точками или множествами на числовой прямой.

П р и м е р 14. Из полной колоды в 52 карты наудачу выбирают 6 карт. Найти вероятность того, что будут выбраны хотя бы две бубновых карты.

Р е ш е н и е. Элементарными исходами будут всевозможные наборы по 6 карт. Их количество равно $|\Omega| = C_{52}^6$.

Обозначим через A событие «среди выбранных карт окажутся хотя бы две бубновых карты». Это событие можно представить в виде объединения попарно несовместных событий $A_2, \ldots, A_6$, где событие A_i означает, что среди выбранных карт окажется ровно i карт бубновой масти. Вероятность каждого из событий A_i ищется по формуле (4): $\mathsf{P}(A_i) = C_{13}^{i}\, C_{52-13}^{6-i} \,/\, C_{52}^{6}$. По аксиоме (P2), вероятность события A равна сумме вероятностей событий A_i:

$$\mathsf{P}(A) = \mathsf{P}(A_2 \cup \ldots \cup A_6) = \frac{C_{13}^{2}\, C_{39}^{4}}{C_{52}^{6}} + \frac{C_{13}^{3}\, C_{39}^{3}}{C_{52}^{6}} + \frac{C_{13}^{4}\, C_{39}^{2}}{C_{52}^{6}} + \frac{C_{13}^{5}\, C_{39}^{1}}{C_{52}^{6}} + \frac{C_{13}^{6}}{C_{52}^{6}}\,.$$

Есть более простое решение: $\mathsf{P}(A) = 1 - \mathsf{P}(\overline{A}\,)$, где событие $\overline{A}$ означает, что выбрано не более одной карты масти бубей. Его вероятность равна

$$\mathsf{P}(\overline{A}\,) \;=\; \mathsf{P}(A_0 \cup A_1) \;=\; \mathsf{P}(A_0) + \mathsf{P}(A_1) \;=\; \frac{C_{39}^{6}}{C_{52}^{6}} + \frac{C_{13}^{1}\, C_{39}^{5}}{C_{52}^{6}}\,.$$

П р и м е р 15. На пяти карточках написаны буквы А, А, Л, М, П. Найти вероятность того, что при случайной расстановке этих карточек в ряд получится слово ЛАМПА.

Р е ш е н и е. Всего возможно $|\Omega| = 5!$ перестановок карточек. Заметим, что перестановка двух карточек с буквой А не меняет слова. Поэтому есть два благоприятных исхода: ЛА$_1$МПА$_2$ и ЛА$_2$МПА$_1$. Вероятность получить нужное слово равна $\dfrac{2}{5!} = \dfrac{1}{60}$.

П р и м е р 16. Игральная кость подбрасывается трижды. Найти вероятность получить в сумме 5 очков.

Р е ш е н и е. Общее число равновозможных элементарных исходов есть $|\Omega| = 6^3$. Сумма очков равна 5, если на двух костях выпали двойки, и на одной — единица, либо на двух костях выпали единицы и на одной — тройка. Каждому из этих событий благоприятствуют 3 исхода. Например, первое событие состоит из исходов $(2, 2, 1)$, $(2, 1, 2)$, $(1, 2, 2)$. Поэтому есть всего 6 благоприятных исходов, и искомая вероятность равна $\dfrac{6}{6^3} = \dfrac{1}{36}$.

Приведём пример задачи, для решения которой проще всего воспользоваться формулой включения-исключения (2).

П р и м е р 17 (задача о рассеянной секретарше). Есть n писем и n подписанных конвертов. Письма раскладываются в конверты наудачу по одному. Найти вероятность того, что хотя бы одно письмо попадёт в нужный конверт.

Р е ш е н и е. Пусть событие A_i, $i = 1, \ldots, n$, означает, что i-е письмо попало в свой конверт. Тогда

$$A = \{\textit{хотя бы одно письмо попало в свой конверт}\} = A_1 \cup \ldots \cup A_n.$$

События $A_1, \ldots, A_n$ совместны, поэтому вычислим вероятность их объединения по формуле (2). Сначала найдём вероятности всех событий A_i и их пересечений. Элементарными исходами будут всевозможные $n!$ размещений n писем по n конвертам. Событию A_i благоприятны $(n-1)!$ из них — всевозможные размещения всех писем, кроме i-го, уже лежащего в своём конверте. Поэтому $\mathsf{P}(A_i) = \dfrac{(n-1)!}{n!} = \dfrac{1}{n}$ — одна и та же для всех i. Аналогично получим, что для любых пар и троек конвертов

$$\mathsf{P}(A_i A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}, \qquad \mathsf{P}(A_i A_j A_m) = \frac{1}{n(n-1)(n-2)}.$$

Точно так же посчитаем вероятности пересечений любого другого числа событий, в том числе $\mathsf{P}(A_1 \ldots A_n) = \dfrac{1}{n!}$.

Вычислим количество слагаемых в каждой сумме в формуле (2). Например, сумма по $1 \leqslant i < j < m \leqslant n$ состоит из C_n^3 слагаемых — ровно столько троек индексов можно образовать из n номеров событий. Подставляя все вероятности в формулу, получим:

$$\begin{aligned} \mathsf{P}(A) &= n \cdot \frac{1}{n} - C_n^2 \cdot \frac{1}{n(n-1)} + C_n^3 \cdot \frac{1}{n(n-1)(n-2)} - \ldots + (-1)^{n-1}\frac{1}{n!} = \\ &= 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{n-1}\frac{1}{n!}. \end{aligned}$$

Упражнение. Выписать разложение e^{-1} в ряд Тейлора и убедиться в том, что $\mathsf{P}(A) \longrightarrow 1 - e^{-1}$ при $n \to \infty$.

Пример 18 (задача про дни рождения). Найти вероятность того, что в группе из 23 человек хотя бы у двоих совпадают дни рождения. Предполагается, что день рождения человека приходится на любой из 365 дней года с равной вероятностью.

Решение. День рождения каждого из данных 23 человек — любой из 365 возможных. Всего элементарных исходов $|\Omega| = 365^{23}$. Действительно, для первого дня рождения есть 365 вариантов, при любом их них для второго дня рождения снова 365 вариантов, и так 23 раза.

Противоположное событие $\overline{A}$ означает, что все 23 дня рождения пришлись на разные дни года. Исходов, благоприятных этому событию, имеется ровно $|\overline{A}| = 365 \cdot 364 \cdot \ldots \cdot 343 = A_{365}^{23}$. Искомая вероятность равна

$$\mathsf{P}(A) = 1 - \mathsf{P}(\overline{A}) = 1 - \frac{365 \cdot \ldots \cdot 343}{365^{23}} = 1 - \frac{364}{365} \cdot \frac{363}{365} \cdot \ldots \cdot \frac{343}{365} \approx 0,5073.$$

Итак, с вероятностью не менее 0,5 в группе из 23 человек найдутся совпадающие дни рождения. Читатель вычислит ту же вероятность для группы из 22 человек и убедится, что она ещё не превышает половины.

§ 2. Основные формулы комбинаторики

В этом параграфе изложены основные принципы и формулы комбинаторики. Именно комбинаторика занимается подсчётом числа способов проделать некоторое действие с несколькими возможными результатами: выбрать что-либо, разбить множество на части, составить набор объектов и т. п.

Принцип перемножения. Основной принцип комбинаторики заключается в следующем: если первый элемент можно выбрать k способами, а второй элемент — m способами, то упорядоченную пару элементов можно составить km способами.

Теорема 2. *Пусть множество $A = \{a_1, \ldots, a_k\}$ состоит из k элементов, а множество $B = \{b_1, \ldots, b_m\}$ — из m элементов. Тогда можно образовать ровно km пар (a_i, b_j), взяв первый элемент из множества A, а второй — из множества B.*

Доказательство. С элементом a_1 мы можем образовать m пар: (a_1, b_1), $(a_1, b_2), \ldots, (a_1, b_m)$. Столько же пар можно составить с элементом a_2, столько же — с элементом a_3 и с любым другим из k элементов множества A. Т. е. всего возможно km пар, в которых первый элемент выбран из множества A, а второй — из множества B. □

Упражнение. С помощью теоремы 2 доказать, что:

а) при подбрасывании трёх монет возможно $2 \cdot 2 \cdot 2 = 8$ результатов;

б) бросая дважды игральную кость, получим $6 \cdot 6 = 36$ результатов;

в) трёхзначных чисел бывает $9 \cdot 10 \cdot 10 = 900$;

г) трёхзначных чисел, все цифры которых различны, существует $9 \cdot 9 \cdot 8$;

д) чётных трёхзначных чисел возможно $9 \cdot 10 \cdot 5$;

Урновые схемы. Есть урна (ящик), содержащая n пронумерованных объектов (шаров). Мы выбираем из урны k шаров; результат этого выбора — набор из k шаров. Нас интересует, сколькими способами можно выбрать k шаров из n, или сколько *различных результатов* возможно.

На этот вопрос нельзя дать однозначный ответ, пока мы не определимся: а) с тем, как организован выбор (можно ли шары возвращать в урну); б) с тем, что понимать под *различными* результатами выбора.

Рассмотрим следующие возможные способы выбора.

1. Выбор *с возвращением*: каждый вынутый шар возвращается в урну, каждый следующий шар выбирается из полной урны. В полученном наборе из k номеров шаров могут встречаться одни и те же номера.

2. Выбор *без возвращения*: вынутые шары в урну *не* возвращаются, и в полученном наборе *не могут* встречаться одни и те же номера.

Условимся, какие результаты выбора (наборы из k номеров шаров) мы будем считать различными. Есть ровно две возможности.

1. Выбор *с учётом порядка*: два набора номеров шаров считаются различными, если они отличаются составом или порядком номеров. Так, наборы $(1, 5, 2)$, $(2, 5, 1)$ и $(4, 4, 5)$ считаются разными, если порядок учитывается.

2. Выбор *без учёта порядка*: два набора номеров шаров считаются различными, если они отличаются составом. Наборы, отличающиеся лишь порядком следования номеров, считаются одинаковыми. Так, наборы $(1, 5, 2)$ и $(2, 5, 1)$ не различаются, если порядок не учитывается.

Подсчитаем, сколько возможно различных результатов для каждой из четырёх схем выбора.

Выбор без возвращения и с учётом порядка. Справдливо утверждение.

Теорема 3. *Общее количество различных наборов при выборе k элементов из n без возвращения и с учётом порядка равняется*

$$A_n^k = n(n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}.$$

Число A_n^k называется *числом размещений* из n элементов по k элементов, а сами результаты выбора — размещениями (без повторений).

Доказательство. Первый шар можно выбрать n способами. При любом выборе первого шара есть $n-1$ способ выбрать второй шар. При любом выборе первых двух шаров есть $n-2$ способа выбрать третий шар и т. д. Применяя последовательно теорему 2 (принцип перемножения), получим, что общее число возможных наборов из k шаров равно произведению k сомножителей $n(n-1)\ldots(n-k+1)$. В этом произведении последний сомножитель $n-k+1$ есть число способов выбора k-го шара, когда уже выбраны предыдущие. □

Следствие 1. *В множестве из n элементов возможно ровно $n!$ перестановок этих элементов.*

Доказательство. Перестановка — результат выбора без возвращения и с учётом порядка n элементов из n. Их число равно $A_n^n = n!$ □

Упражнение. Найти, сколько всего возможно различных результатов в следующих экспериментах:

а) из колоды в 36 карт выдают по карте троим игрокам;

б) Вася, Петя, Оля и Лена выбирают четыре из десяти разных учебников;

в) из различных цифр, не равных нулю, составляется трёхзначное число;

г) 36 карт в колоде перемешиваются и выкладываются на стол в ряд.

Выбор без возвращения и без учёта порядка. Следующее утверждение даёт число результатов этой схемы выбора.

Теорема 4. *Общее количество различных наборов при выборе k элементов из n без возвращения и без учёта порядка равняется*

$$C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$$

Число C_n^k называется *числом сочетаний* из n элементов по k элементов, а сами результаты выбора — сочетаниями.

Доказательство. По следствию 1, k различных номеров шаров можно упорядочить $k!$ способами. Поэтому из каждого *сочетания* можно перестановками образовать $k!$ *размещений*. Следовательно, число наборов, порядок в которых не учитывается (сочетаний), в $k!$ раз меньше числа наборов, отличающихся ещё и порядком (размещений). □

Упражнение. Найти, сколько всего возможно различных результатов в следующих экспериментах:

а) из колоды в 36 карт выдают три карты одному игроку;

б) из двадцати учеников класса выбирают троих дежурных.

Выбор с возвращением и с учётом порядка. Результаты такой схемы выбора называют размещениями с повторением.

Теорема 5. *Общее количество различных наборов при выборе k элементов из n с возвращением и с учётом порядка равняется n^k.*

Доказательство. Первый шар можно выбрать n способами. При каждом из этих способов второй шар можно выбрать также n способами, и так k раз. Общее число наборов равно $n \cdot n \cdot \ldots \cdot n = n^k$. □

Упражнение. Найти, сколько всего возможно различных результатов в следующих экспериментах:

а) монета подбрасывается пять раз;

б) пятизначное число составляется из одних нечётных цифр.

Выбор с возвращением и без учёта порядка. Сформулируем без доказательства утверждение.

Теорема 6. *Общее количество различных наборов при выборе k элементов из n с возвращением и без учёта порядка равняется*

$$C_{n+k-1}^k = C_{n+k-1}^{n-1}.$$

Читателю полезно вернуться к примеру 11, где показана неравновозможность элементарных исходов при выборе с возвращением и без учёта порядка.

Упражнение. Найти число возможных результатов подбрасывания двух игральных костей, если кости считаются неразличимыми. То же самое для трёх игральных костей.

§ 3. Вероятность на числовой прямой и плоскости

Результаты многих экспериментов нельзя описать дискретным множеством точек. Например, бросание монеты на стол в примере 4 приводит к пространству элементарных исходов, совпадающему с множеством точек стола. Дальность броска копья спортсменом — величина с положительными значениями на числовой прямой, и т. д.

Рассмотрим пространство элементарных исходов $\Omega = \mathbb{R}^k$. Нас будут особо интересовать случаи, когда Ω есть множество действительных чисел $\mathbb{R}$ или множество точек плоскости $\mathbb{R}^2$. Эти пространства не являются дискретными — они состоят из несчётного множества точек, и присвоить положительную вероятность каждой точке нельзя. Как же задавать вероятности событий на таком пространстве элементарных исходов?

Например, так. В случае $\Omega = \mathbb{R}$ рассмотрим произвольную неотрицательную интегрируемую функцию f, обладающую тем свойством, что

$$\int\limits_{-\infty}^{\infty} f(x)\,dx = 1. \tag{5}$$

Интеграл в формуле (5) равен площади области под графиком функции f выше оси абсцисс. Определим вероятность любого интервала или отрезка $A = [a,\, b] \subseteq \Omega$ как площадь криволинейной трапеции под графиком функции f над этим интервалом или отрезком:

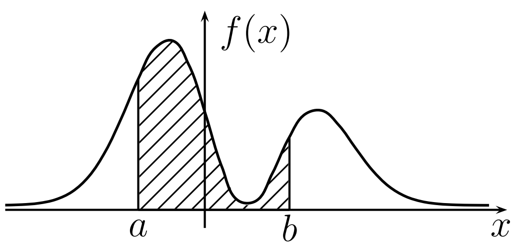

Рис. 5. Плотность в $\mathbb{R}$

$$A = [a,\, b], \qquad \mathsf{P}(A) = \int\limits_a^b f(x)\,dx. \tag{6}$$

Условие (5) означает просто нормированность вероятности $\mathsf{P}(\Omega) = \mathsf{P}(\mathbb{R}) = 1$.

Функцию f называют *плотностью* распределения вероятностей. Можно представлять себе, что с такой *плотностью* общая единичная масса (вероятность) «размазана» по прямой, подобно маслу на бутерброде. Бóльшие значения плотности означают бóльшую массу соответствующего участка прямой, нулевые значения плотности на отрезке отвечают отсутствию массы.

Аддитивность интеграла позволяет нам вычислять вероятности любых конечных и счётных объединений отрезков. Интересно, что вероятность любой точки на прямой — одного элементарного исхода — оказывается *нулевой*, т. к. площадь трапеции «шириной в точку» под графиком f равна нулю.

П р и м е р 19. Пусть $\Omega = \mathbb{R}$, а плотность распределения имеет вид

$$f(x) = \begin{cases} 5e^{-5x}, & \text{если } x \geqslant 0, \\ 0, & \text{если } x < 0. \end{cases}$$

Вычислим по данной плотности f (читатель проверит выполнение условия (5)) вероятности отрезков $[0,\, 1]$ и $[1,\, 2]$:

$$\mathsf{P}\big([0,\, 1]\big) = \int\limits_0^1 5e^{-5x}dx = 1 - e^{-5}, \quad \mathsf{P}\big([1,\, 2]\big) = \int\limits_1^2 5e^{-5x}dx = e^{-5} - e^{-10}.$$

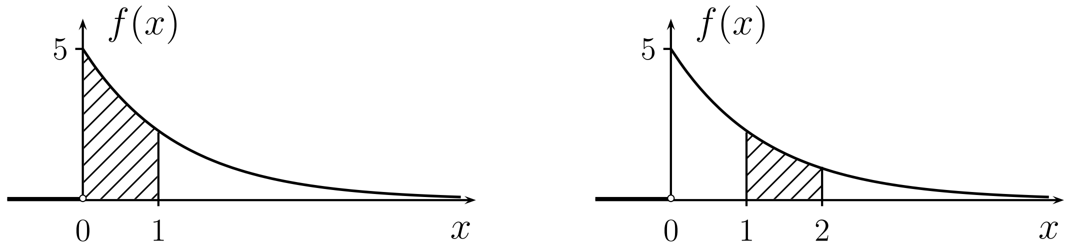

Рис. 6. Вероятности отрезков в примере 19

Вероятности двух отрезков, имеющих одну и ту же длину, оказываются различными. А возможно ли задать плотность распределения так, чтобы вероятность зависела лишь от длины отрезка?

П р и м е р 20. Рассмотрим плотность распределения

$$f(x) = \begin{cases} 0{,}2, & \text{если } 0 \leqslant x \leqslant 5, \\ 0, & \text{если } x \notin [0,\, 5]. \end{cases}$$

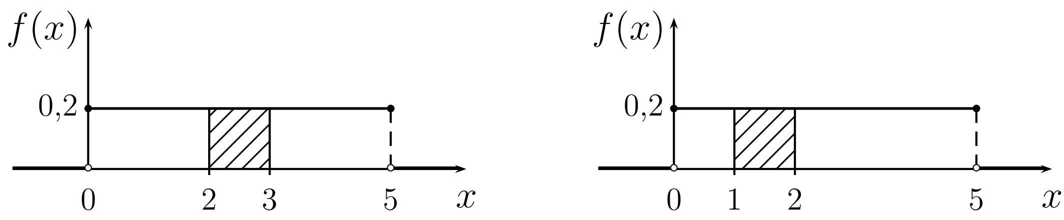

Рис. 7. Вероятности отрезков в примере 20

Выполнение условия (5) здесь очевидно: площадь прямоугольника со сторонами 5 и 0,2 равна 1. При такой плотности вероятность любого отрезка зависит не от положения этого отрезка внутри $[0,\, 5]$, а лишь от его длины и пропорциональна ей. Заметим также, что вероятность любого множества, не пересекающегося с отрезком $[0,\, 5]$, равна нулю. Поэтому множество Ω можно сузить до $[0,\, 5]$.

В случае $\Omega = \mathbb{R}^2$ задать вероятности подмножеств плоскости можно с помощью неотрицательной интегрируемой функции двух переменных $f(x, y)$, нормированной так, что

$$\int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{\infty} f(x, y)\, dy = 1.$$

Подобно одномерному случаю, вероятности прямоугольников и многих более сложных областей на плоскости определяются как объём криволинейного цилиндра над соответствующей областью под графиком функции:

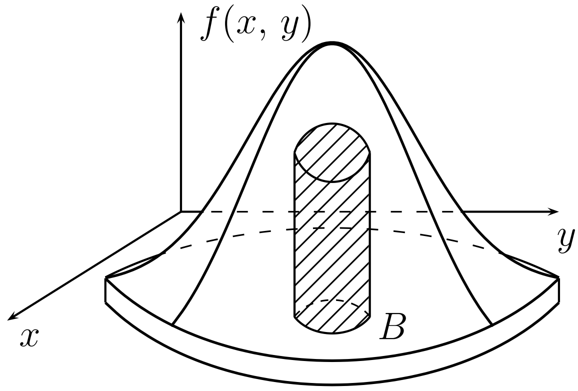

Рис. 8. Плотность в $\mathbb{R}^2$

$$\mathsf{P}(B) = \iint\limits_{B} f(x, y)\, dx\, dy.$$

Если множество Ω совпадает с пространством $\mathbb{R}^3$, вероятность можно задать с помощью неотрицательной интегрируемой функции трёх переменных. Наконец, для $\Omega = \mathbb{R}^k$ нам потребуется функция k переменных.

Геометрическое определение вероятности. Рассмотрим отдельно задание вероятности на ограниченном подмножестве прямой, плоскости или пространства, когда вероятностная масса распределена по области Ω *равномерно,* как в примере 20.

Предположим, что мера области Ω (длина на прямой, площадь на плоскости, объём в пространстве) конечна, а эксперимент состоит в бросании точки в эту область *наудачу*. Термин «наудачу» означает, что для любого множества $A \subseteq \Omega$ с конечной мерой $\mu(A)$ вероятность точке попасть в это множество не зависит от его формы или его расположения внутри Ω, а зависит лишь от его меры и пропорциональна ей.

Для такого эксперимента вероятности определяются согласно *геометрическому определению вероятности*:

$$\mathsf{P}(A) = \frac{\mu(A)}{\mu(\Omega)}\,. \tag{7}$$

Если точка в области Ω выбирается наудачу, то говорят, что она *равномерно распределена* в области Ω. Так, в примере 20 точка равномерно распределена на отрезке $[0, 5]$. А чтобы, скажем, получить равномерное распределение в круге, следует рассмотреть плотность распределения $f(x, y)$, равную нулю вне круга и постоянную внутри него. Значение этой постоянной зависит от площади круга и обратно пропорционально ей.

П р и м е р 21 (задача о встрече). Два человека X и Y условились встретиться в определённом месте между двумя и тремя часами дня. Пришедший первым ждет другого в течение 10 минут, после чего уходит. Чему равна вероятность встречи этих лиц, если каждый из них может прийти в любое время в течение указанного часа независимо от другого?

Р е ш е н и е. Будем считать интервал от двух до трёх часов дня отрезком $[0,\ 1]$ длиной в один час. Обозначим через $\xi \in [0,\ 1]$ и $\eta \in [0,\ 1]$ моменты прихода X и Y в течение этого часа. Результатами эксперимента являются пары точек $(\xi,\ \eta)$ из единичного квадрата:

$$\Omega = \{(\xi, \eta) \mid\ 0 \leqslant \xi \leqslant 1,\ 0 \leqslant \eta \leqslant 1\}.$$

Благоприятными исходами будут точки заштрихованного на рисунке множества A:

$$A = \{(\xi, \eta) \mid\ |\xi - \eta| \leqslant 1/6\}.$$

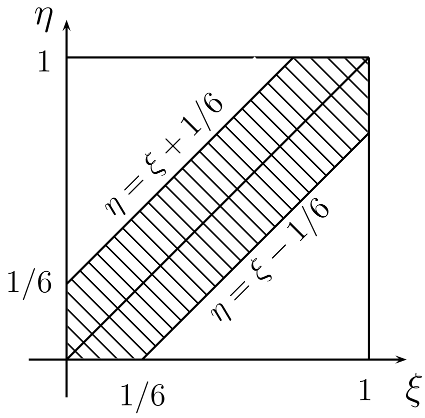

Рис. 9. Задача о встрече

Попадание в множество A наудачу брошенной в квадрат точки означает, что X и Y встретятся. Тогда вероятность встречи равна

$$\mathsf{P}(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{1 - \left(\frac{5}{6}\right)^2}{1} = \frac{11}{36}.$$

П р и м е р 22. Точка наудачу бросается в круг с единичным радиусом. Найти вероятность того, что расстояние ρ до точки от центра круга будет меньше заданного числа $r \in (0,\ 1)$.

Р е ш е н и е. Интересующее нас событие $\{\rho < r\}$ происходит, когда точка попадает во внутренний круг с радиусом r и тем же центром. По формуле (7), вероятность этого события равна отношению площадей кругов:

$$\mathsf{P}(\rho < r) = \frac{\pi r^2}{\pi} = r^2.$$

Заметим, что расстояние ρ до брошенной в круг точки распределено не равномерно на отрезке $[0,\ 1]$. Для равномерного распределения мы получили бы вероятность $\mathsf{P}(\rho < r) = r$, а не r^2.

§ 4. Контрольные вопросы

1. Что такое дискретное пространство элементарных исходов?

2. Как задаётся вероятность на дискретном пространстве?

3. Можно ли на дискретном пространстве взять все $p_i = 0$?

4. Задать вероятность на $\Omega = \{a, b, c\}$ как на дискретном пространстве.

5. Задать вероятность на множестве натуральных чисел $\Omega = \mathbb{N}$ как на дискретном пространстве.

6. Пусть $\Omega = \mathbb{N}$. Можно ли задать вероятность так, чтобы все p_i были одинаковы?

7. Как вычисляется вероятность события в классической схеме?

8. Что больше: общее число исходов или число исходов, благоприятных некоторому событию?

9. Какие вероятности вычисляют по формуле гипергеометрического распределения?

10. Сколько равновозможных элементарных исходов возникает при бросании трёх правильных игральных костей?

11. Если пуговиц 3, а бусинок 5, сколько различных пар «пуговица, бусинка» можно образовать, выбирая одну пуговицу и одну бусинку?

12. Что такое выбор с возвращением? Без возвращения? С учётом порядка? Без учёта порядка?

13. В какой схеме больше исходов: с учётом или без учёта порядка?

14. Сколько исходов возможно при выборе без возвращения и с учётом порядка? Что такое число размещений?

15. Сколько перестановок в множестве из n элементов возможно?

16. Сколько исходов возможно при выборе без возвращения и без учёта порядка? При выборе с возвращением и с учётом порядка?

17. Пусть $\Omega = \mathbb{R}$ — числовая прямая. Как можно задать вероятность на таком пространстве?

18. Что такое плотность распределения вероятностей?

19. Как по плотности распределения вероятностей вычислять вероятности отрезков? Других событий?

20. Может ли плотность быть всюду постоянной, если $\Omega = \mathbb{R}$?

21. Что такое геометрическое определение вероятности?

22. Точка наудачу брошена на отрезок. Какова вероятность ей оказаться на левой половине отрезка?

Г Л А В А III

УСЛОВНАЯ ВЕРОЯТНОСТЬ

Об условной вероятности говорят, когда имеется какое-то знание о результате эксперимента, позволяющее переоценить шансы каждого события с учётом этого знания. С этим понятием тесно связано понятие независимости двух событий, когда появление в опыте одного события не влияет на шансы второго.

§ 1. Понятие условной вероятности

П р и м е р 23. Игральная кость подбрасывается один раз. *Известно*, что выпало более трёх очков. Какова *при этом* вероятность того, что выпало нечётное число очков?

Пусть событие $B = \{4,\ 5,\ 6\}$ означает, что выпало более трёх очков, событие $A = \{1,\ 3,\ 5\}$ — выпало нечётное число очков. Как понимать вероятность события A, если известно, что B случилось? Знаем, что произошло событие B, но всё равно не знаем, что именно выпало на кости. Однако теперь возможностей осталось только *три*: могло выпасть 4, 5 или 6 очков. Событию A из этих равновозможных исходов благоприятен *единственный* исход: выпадение пяти очков. Поэтому искомая вероятность равна $1/3$.

Итак, при вычислении условной вероятности события A при случившемся событии B мы ищем *долю исходов, благоприятствующих A, среди всех исходов события B*. Эту условную вероятность будем обозначать $\mathsf{P}(A \mid B)$.

О п р е д е л е н и е 7. Пусть $\mathsf{P}(B) \neq 0$. *Условной вероятностью* события A при условии, что произошло событие B, называется число

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}\ .$$

В случае $\mathsf{P}(B) = 0$ условная вероятность $\mathsf{P}(A \mid B)$ не определена.

Важно отличать условную вероятность, т. е. вероятность одного события при осуществлении другого, от вероятности им одновременно произойти.

П р и м е р 24. Имеется 1 000 урн, среди которых 999 содержат только чёрные шары, а одна — только белые. Сначала выберем наугад урну, а затем из неё достанем один за другим два шара.

Пусть событие B означает, что первый вынутый шар оказался белым, событие A — второй шар белый. Вероятность этим событиям случиться одновременно мала: $\mathsf{P}(A \cap B) = \mathsf{P}(A) = \mathsf{P}(B) = 0{,}001$, так как только в одном случае из $1\,000$ мы выберем урну с белыми шарами.

Однако условная вероятность $\mathsf{P}(A \mid B)$ равна единице: если первый шар белый, то мы *уже* выбрали урну с одними белыми шарами, и второй обязательно будет белым. При вычислении условной вероятности $\mathsf{P}(A \mid B)$ не играет никакой роли то, насколько мала или велика вероятность события B.

Определение 7 бывает полезно использовать для последовательного вычисления вероятности нескольким событиям случиться одновременно, если известны соответствующие условные вероятности. А именно, справедливы следующие «теоремы умножения вероятностей».

Т е о р е м а 7. *Если* $\mathsf{P}(B) \neq 0$ *и* $\mathsf{P}(A) \neq 0$, *то*

$$\mathsf{P}(A \cap B) \;=\; \mathsf{P}(B)\,\mathsf{P}(A \mid B) \;=\; \mathsf{P}(A)\,\mathsf{P}(B \mid A).$$

Т е о р е м а 8. *Для любых событий* $A_1, \ldots, A_n$ *верно равенство:*

$$\mathsf{P}(A_1 \ldots A_n) = \mathsf{P}(A_1)\,\mathsf{P}(A_2 \mid A_1)\,\mathsf{P}(A_3 \mid A_1A_2) \cdot \ldots \cdot \mathsf{P}(A_n \mid A_1 \ldots A_{n-1}),$$

если все участвующие в нём условные вероятности определены.

У п р а ж н е н и е. Доказать теорему 8 методом математической индукции.

П р и м е р 25. Студент приходит на экзамен, изучив 20 из 25 вопросов программы. На экзамене он наугад выбирает три вопроса. Найти вероятность того, что попадутся только заранее изученные вопросы.

Р е ш е н и е. Эту вероятность легко найти с помощью гипергеометрического распределения из примера 13. С другой стороны, очевидно равенство

$$\mathsf{P}(\text{ответы на 1-й, 2-й, 3-й вопросы студент знает}) = \frac{20}{25} \cdot \frac{19}{24} \cdot \frac{18}{23}.$$

Обратимся к теореме 8. Пусть событие A_i означает, что ответ на i-й вопрос студент знает. Тогда $\mathsf{P}(A_1) = \mathsf{P}(A_2) = \mathsf{P}(A_3) = \dfrac{20}{25}$. Например, третьим вопросом билета может оказаться любой из 25 вопросов программы, и 20 из них благоприятны событию A_3 («хорошие» вопросы).

А что за вероятность $\dfrac{19}{24}$? Это *условная* вероятность второму вопросу оказаться «хорошим», если первый был «хорошим», и «хороших» вопросов осталось 19 из 24 возможных. Далее, $\dfrac{18}{23}$ — условная вероятность третьему вопросу оказаться «хорошим», если *и первый, и второй* были «хорошими», и «хороших» вопросов осталось 18 из 23 возможных. Таким образом,

$$\mathsf{P}(A_1A_2A_3) = \mathsf{P}(A_1)\,\mathsf{P}(A_2 \mid A_1)\,\mathsf{P}(A_3 \mid A_1A_2) = \frac{20}{25} \cdot \frac{19}{24} \cdot \frac{18}{23} \approx 0{,}496.$$

§ 2. Формула полной вероятности

Начнём с простого примера «на проценты».

П р и м е р 26. Есть три завода, производящих одну и ту же продукцию. Первый завод производит 25%, второй завод — 35% и третий — 40% всей продукции. Брак составляет 5% от продукции первого завода, 3% от продукции второго и 4% от продукции третьего завода. Вся продукция смешивается и поступает в продажу. Найти: а) вероятность купить бракованное изделие; б) условную вероятность того, что купленное изделие изготовлено первым заводом, если это изделие оказалось бракованным.

Р е ш е н и е. Первая вероятность равна доле брака в объёме всей продукции: $0{,}05 \cdot 0{,}25 + 0{,}03 \cdot 0{,}35 + 0{,}04 \cdot 0{,}4$. Вторая вероятность равна доле брака первого завода среди всего брака:

$$\frac{0{,}05 \cdot 0{,}25}{0{,}05 \cdot 0{,}25 + 0{,}03 \cdot 0{,}35 + 0{,}04 \cdot 0{,}4}.$$

В этом примере мы вычисляли проценты и не использовали ни одной формулы теории вероятностей, кроме классического определения вероятности как доли нужных исходов среди всех. Попробуем обобщить полученные ответы. Ответ на первый вопрос в общем случае даёт формула полной вероятности, ответ на второй — формула Байеса.

О п р е д е л е н и е 8. Конечный или счётный набор попарно несовместных (взаимоисключающих) событий H_1, H_2, $\ldots$, объединение которых есть всё Ω, называется *полной группой событий* или *разбиением* пространства Ω.

События H_1, H_2, $\ldots$, образующие полную группу событий, называют *гипотезами*. Обычно для некоторого события A можно вычислить $\mathsf{P}(A \mid H_i)$ (вероятность событию A произойти при выполнении «гипотезы» H_i) и $\mathsf{P}(H_i)$ (вероятность выполнения «гипотезы» H_i). Как, используя эти данные, посчитать вероятность события A?

Т е о р е м а 9. *Пусть $H_1, H_2, \ldots$ — полная группа событий. Тогда вероятность любого события A может быть вычислена по формуле:*

$$\mathsf{P}(A) = \sum_i \mathsf{P}(H_i) \cdot \mathsf{P}(A \mid H_i).$$

Д о к а з а т е л ь с т в о. Разложим событие A в сумму попарно несовместных событий: $A = A \cap \Omega = A \cap (H_1 \cup H_2 \cup \ldots) = AH_1 \cup AH_2 \cup \ldots$ Далее применим свойство вероятности (P2) и теорему 7:

$$\mathsf{P}(A) = \sum_i \mathsf{P}(AH_i) = \sum_i \mathsf{P}(H_i) \cdot \mathsf{P}(A \mid H_i). \qquad \square$$

Пример 27. Два игрока по очереди подбрасывают правильную игральную кость. Выигрывает тот, кто первым выкинет шесть очков. Найти вероятность победы игрока, начинающего игру.

Решение. События $H_1 = \{$при первом броске выпало 6 очков$\}$ и противоположное к нему H_2 образуют полную группу событий. Пусть событие A означает победу игрока, начинающего игру. Обозначим через x вероятность события A. Для игрока, бросающего кость вторым, вероятность одержать победу во всей игре равна $1 - x$.

Тогда $\mathsf{P}(A \mid H_1) = 1$, $\mathsf{P}(A \mid H_2) = 1 - x$. Действительно, если при первом броске не выпало 6 очков, право броска переходит ко второму игроку. Игрок, начинавший игру, оказывается теперь в положении второго, и его вероятность выиграть становится равной $1 - x$. По формуле полной вероятности

$$x = \mathsf{P}(A) = \mathsf{P}(H_1)\mathsf{P}(A \mid H_1) + \mathsf{P}(H_2)\mathsf{P}(A \mid H_2) = \frac{1}{6} + \frac{5}{6}\,(1 - x) = 1 - \frac{5}{6}\,x.$$

Отсюда $x = \dfrac{6}{11}$ — вероятность победы первого игрока.

Мы воспользовались выше тем очевидным фактом, что вероятности выигрыша первого и второго игроков дают в сумме единицу. Игра не приведёт к победе ни одного из игроков, если шесть очков так и не выпадет, а это событие имеет нулевую вероятность. Однако доказать этот факт мы сможем лишь в § 2 следующей главы.

§ 3. Формула Байеса

Теорема 10. *Пусть $H_1, H_2, \ldots$ — полная группа событий. Тогда условная вероятность события H_k, если в результате эксперимента наблюдалось событие A, может быть вычислена по формуле:*

$$\mathsf{P}(H_k \mid A) = \frac{\mathsf{P}(H_k)\mathsf{P}(A \mid H_k)}{\sum \mathsf{P}(H_i)\mathsf{P}(A \mid H_i)}.$$

Доказательство. По определению условной вероятности,

$$\mathsf{P}(H_k \mid A) = \frac{\mathsf{P}(H_k \cap A)}{\mathsf{P}(A)} = \frac{\mathsf{P}(H_k)\mathsf{P}(A \mid H_k)}{\sum \mathsf{P}(H_i)\mathsf{P}(A \mid H_i)}. \qquad \square$$

Вероятности $\mathsf{P}(H_i)$, вычисленные заранее, до проведения эксперимента, называют *априорными* вероятностями (a'priori — «до опыта»). Условные вероятности $\mathsf{P}(H_i \mid A)$ называют *апостериорными* вероятностями (a'posteriori — «после опыта»). Формула Байеса позволяет переоценить заранее известные вероятности после того, как получено знание о результате эксперимента. Эта формула находит многочисленные применения в экономике, статистике, социологии и т. п.

Пример 28. Два стрелка подбрасывают монетку и выбирают, кто из них будет стрелять по мишени (одной пулей). Первый стрелок попадает по мишени с вероятностью 1, второй стрелок — с вероятностью 10^{-5}.

Можно сделать два предположения об эксперименте: H_1 — стреляет 1-й стрелок (выпал герб) и H_2 — стреляет 2-й стрелок (выпала решка). Априорные вероятности этих гипотез одинаковы: $\mathsf{P}(H_1) = \mathsf{P}(H_2) = \dfrac{1}{2}$.

Как изменятся вероятности гипотез после проведения опыта? Рассмотрим событие A — пуля попала в мишень. Известно, что

$$\mathsf{P}(A \mid H_1) = 1, \quad \mathsf{P}(A \mid H_2) = 10^{-5}.$$

По формуле полной вероятности, вероятность пуле попасть в мишень равна

$$\mathsf{P}(A) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 10^{-5}.$$

Предположим, что событие A произошло. Тогда по формуле Байеса

$$\mathsf{P}(H_1 \mid A) = \frac{\frac{1}{2} \cdot 1}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 10^{-5}} = \frac{1}{1 + 10^{-5}} \approx 0{,}999\,99,$$

$$\mathsf{P}(H_2 \mid A) = \frac{\frac{1}{2} \cdot 10^{-5}}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 10^{-5}} = \frac{10^{-5}}{1 + 10^{-5}} \approx 0{,}000\,01.$$

Попадание пули в мишень сделало выпадение герба в 10^5 раз более вероятным, чем выпадение решки.

§ 4. Независимость событий

Определение 9. События A и B называются *независимыми*, если

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

Естественно считать события A и B независимыми, когда условная вероятность A, когда B произошло, остаётся такой же, как и безусловная. Убедимся, что этим свойством обладают события, независимые согласно определению 9. Пусть $\mathsf{P}(B) \neq 0$ и события A и B независимы. Тогда

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)} = \frac{\mathsf{P}(A) \cdot \mathsf{P}(B)}{\mathsf{P}(B)} = \mathsf{P}(A).$$

Независимые события возникают, например, при повторении испытаний. Выпадение герба и выпадение решки при двух разных бросках монеты независимы. Любые события, относящиеся к двум разным подбрасываниям игральной кости, независимы.

А вот несовместные (взаимоисключающие) события являются, вообще говоря, зависимыми: если одно из двух несовместных событий случилось, то второе случиться не может. Но в общем случае зависимость двух событий вовсе не обязана быть причинно-следственной. Вероятностная зависимость (отсутствие независимости) событий A и B означает лишь, что вероятность события A изменится, если произойдёт событие B.

Пример 29. Из колоды в 36 карт наугад берут одну. Независимы ли события «вынут туз» и «вынута пиковая карта»?

Решение. Вероятность вытянуть туза равна $\mathsf{P}(A) = \frac{4}{36} = \frac{1}{9}$. Вероятность вытянуть пиковую карту равна $\mathsf{P}(B) = \frac{1}{4}$. Пересечение этих событий означает появление туза пик и имеет вероятность $\mathsf{P}(AB) = \frac{1}{36}$. События A и B независимы, т.к. $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$. Это может показаться странным, ведь эти события относятся к одной и той же карте.

Пример 30. В примере 25 события A_1 и A_2 (первые два вопроса «хорошие») зависимы: $\mathsf{P}(A_2 \,|\, A_1) = \frac{19}{24}$ отличается от $\mathsf{P}(A_2) = \frac{20}{25}$.

Свойство 1. *Если события A и B независимы, то независимы и события A и $\overline{B}$, $\overline{A}$ и B, $\overline{A}$ и $\overline{B}$.*

Доказательство. Так как $A = AB \cup A\overline{B}$, и события AB и $A\overline{B}$ несовместны, то $\mathsf{P}(A) = \mathsf{P}(AB) + \mathsf{P}(A\overline{B})$. Поэтому

$$\mathsf{P}(A\overline{B}) = \mathsf{P}(A) - \mathsf{P}(AB) = \mathsf{P}(A) - \mathsf{P}(A)\mathsf{P}(B) = \mathsf{P}(A)\mathsf{P}(\overline{B}).$$

Независимость A и $\overline{B}$ доказана, откуда сразу же следует независимость для всех остальных пар. □

Пусть есть не два, а n событий, $n > 2$. Тогда из равенства $\mathsf{P}(A_1 \ldots A_n) = \mathsf{P}(A_1) \cdot \ldots \cdot \mathsf{P}(A_n)$ не следует даже попарная независимость этих событий. Например, при таком равенстве события A_1 и A_2 вполне могут оказаться зависимыми, как в наборе событий A, A, …, A, $\varnothing$, где $\mathsf{P}(A)$ не равно 0 и 1. Пересечение всех этих событий пусто, его вероятность равна нулю и равна произведению вероятностей, а событие A от себя зависит. Действительно,

$$\mathsf{P}(A \cap A) = \mathsf{P}(A) \neq \mathsf{P}(A) \cdot \mathsf{P}(A).$$

Хотелось бы независимостью набора событий считать такое свойство, когда любые комбинации событий оказываются независимыми между собой. Просто потребуем этого в определении.

Определение 10. События $A_1, \ldots, A_n$ называются *независимыми в совокупности*, если вероятность пересечения любых событий из этих n равна произведению их вероятностей.

Если n событий независимы в совокупности, то любые два из них независимы. Обратное, как показывает следующий пример, неверно.

П р и м е р 31 (пример Бернштейна). Рассмотрим правильный тетраэдр, три грани которого окрашены соответственно в красный, синий, зелёный цвета, а четвёртая грань содержит все три цвета. Событие A (B, C) означает, что выпала грань, содержащая красный (синий, зелёный) цвета.

Вероятность каждого из этих событий равна $\frac{1}{2}$, так как каждый цвет есть на двух гранях из четырёх. Вероятность пересечения любых двух событий равна $\frac{1}{4}$, так как только одна грань из четырёх содержит два цвета. Поэтому любые два события из трёх независимы, так как $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$.

Но вероятность события ABC (на грани есть все три цвета) тоже равна $\frac{1}{4}$, а не $\frac{1}{8}$, т. е. события *не* являются независимыми в совокупности. Вероятность пересечения оказалась равна произведению вероятностей для пар событий A и B, A и C, B и C, но не для тройки событий A, B, C.

П р и м е р 32. Пусть события $A_1, \ldots, A_n$ независимы в совокупности. Найти вероятность того, что из этих событий произойдёт хотя бы одно.

Р е ш е н и е. Нас интересует вероятность события $A = A_1 \cup \ldots \cup A_n$. Воспользуемся независимостью событий $A_1, \ldots, A_n$. Противоположные события также независимы, поэтому дополнение $\overline{A} = \overline{A}_1 \cap \ldots \cap \overline{A}_n$ имеет вероятность

$$\mathsf{P}(\overline{A}) = \mathsf{P}(\overline{A}_1) \cdot \ldots \cdot \mathsf{P}(\overline{A}_n) = (1 - \mathsf{P}(A_1)) \cdot \ldots \cdot (1 - \mathsf{P}(A_n)).$$

Тогда $\mathsf{P}(A) = 1 - (1 - \mathsf{P}(A_1)) \cdot \ldots \cdot (1 - \mathsf{P}(A_n))$.

П р и м е р 33. Три стрелка, попадающих в мишень с вероятностями 0,3, 0,8 и 0,9 независимо друг от друга, делают по одному выстрелу по мишени. Найти вероятность того, что в мишень попадут ровно две пули.

Р е ш е н и е. Пусть события A, B и C означают попадание первого, второго и третьего стрелков соответственно. Событие D «попали ровно две пули» запишем как объединение попарно несовместных событий:

$$D = AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC.$$

По аксиоме (P2), вероятность D равна сумме их вероятностей:

$$\mathsf{P}(D) = \mathsf{P}\big(AB\overline{C}\big) + \mathsf{P}\big(A\overline{B}C\big) + \mathsf{P}\big(\overline{A}BC\big).$$

Но в силу независимости событий A, B и C каждое слагаемое равно произведению вероятностей:

$$\mathsf{P}(D) = 0{,}3 \cdot 0{,}8 \cdot (1-0{,}9) + 0{,}3 \cdot (1-0{,}8) \cdot 0{,}9 + (1-0{,}3) \cdot 0{,}8 \cdot 0{,}9.$$

§ 5. Контрольные вопросы

1. Что такое условная вероятность?
2. Может ли условная вероятность равняться безусловной?
3. Может ли условная вероятность равняться единице? Нулю?
4. Чему равна вероятность пересечения двух произвольных событий? Нескольких событий?
5. Сформулировать теорему умножения вероятностей.
6. Что такое полная группа событий?
7. Чему равна сумма вероятностей событий из полной группы?
8. Сформулировать формулу полной вероятности.
9. Привести пример задачи на формулу полной вероятности.
10. Для чего нужна формула Байеса?
11. Что такое априорная вероятность? Апостериорная?
12. Какие события называют независимыми?
13. Могут ли несовместные события быть независимы?
14. Привести пример двух независимых событий.
15. Зависит ли событие от самого себя?
16. Зависит ли невозможное событие от самого себя?
17. Зависит ли достоверное событие от самого себя?
18. Независимы ли события, противоположные к независимым?
19. Что такое независимость нескольких событий в совокупности?
20. Почему для независимости в совокупности недостаточно, чтобы вероятность пересечения всех событий равнялась произведению их вероятностей?
21. Достаточно ли попарной независимости (любых пар) событий для независимости в совокупности?
22. Из колоды в 36 карт выбирают наугад одну. Независимы ли события «выбрана пика» и «выбран туз»?
23. Из колоды в 36 карт выбирают наугад одну. Независимы ли события «выбрана пика» и «выбрана бубна»?
24. Дважды бросают правильную монету. Независимы ли события «при первом броске выпал герб» и «при втором броске выпала решка»?
25. Дважды бросают правильную монету. Независимы ли события «при первом броске выпал герб» и «при первом броске выпала решка»?

ГЛАВА IV
СХЕМА БЕРНУЛЛИ

Предположим, что в некотором эксперименте возможно событие A. Проведя такой эксперимент один раз, мы увидим только, произошло A или нет. Чтобы судить о вероятности события A, потребуется повторять этот эксперимент снова и снова, но так, чтобы результаты уже проведённых испытаний не влияли на последующие. Пусть мы провели его пять раз, и всё время случалось событие A. Насколько это возможно, если вероятность события A равна 0,1? Пусть после тысячи экспериментов событие A случилось 430 раз. Можно ли считать вероятность события A примерно равной 0,43, и насколько приблизительно это «примерно»? Чтобы научиться отвечать на эти и другие вопросы, рассмотрим подробно схему независимых испытаний с двумя возможными результатами.

§1. Распределение числа успехов в n испытаниях

Определение 11. *Схемой Бернулли* называется последовательность независимых испытаний, в каждом из которых возможны лишь два исхода — «успех» и «неудача», при этом успех в каждом испытании происходит с одной и той же вероятностью $p \in (0,1)$, а неудача — с вероятностью $q = 1 - p$.

Обозначим через ν_n число успехов, случившихся в n испытаниях схемы Бернулли. Эта (случайная) величина может принимать целые значения от нуля до n в зависимости от результатов испытаний. Например, если все n испытаний завершились неудачей, то величина ν_n равна нулю.

Теорема 11 (формула Бернулли). *Для любого $k = 0, 1, \ldots, n$ вероятность получить в n испытаниях k успехов равна*

$$\mathsf{P}(\nu_n = k) = C_n^k \, p^k \, q^{n-k}. \tag{8}$$

Набор вероятностей (8) называется *биномиальным распределением* вероятностей.

Доказательство. Событие $A = \{\nu_n = k\}$ означает, что в n испытаниях схемы Бернулли произошло ровно k успехов. Рассмотрим один элементарный исход из события A:

$$(\underbrace{у,\ у,\ \ldots,\ у}_{k},\ \underbrace{н, н,\ \ldots,\ н}_{n-k}),$$

когда первые k испытаний завершились успехом, остальные неудачей. Поскольку испытания независимы, вероятность такого элементарного исхода равна $p^k(1-p)^{n-k}$. Другие элементарные исходы из события A отличаются лишь расположением k успехов на n местах. Есть ровно C_n^k способов расположить k успехов на n местах. Поэтому событие A состоит из C_n^k элементарных исходов, вероятность каждого из которых равна $p^k q^{n-k}$. $\square$

П р и м е р 34. Правильная монета подбрасывается 10 раз. Найти вероятность того, что герб выпадет от 4 до 6 раз.

Р е ш е н и е. Вычислим отдельно вероятности получить 4, 5 и 6 гербов после десяти подбрасываний монеты.

$$\mathsf{P}(\nu_{10}=4)=C_{10}^4\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^{10-4}=C_{10}^4\cdot\frac{1}{2^{10}}=\frac{10!}{4!\,6!}\cdot\frac{1}{1024}\approx 0{,}205;$$

$$\mathsf{P}(\nu_{10}=5)=C_{10}^4\cdot\frac{1}{2^{10}}\approx 0{,}246;\quad \mathsf{P}(\nu_{10}=6)=C_{10}^6\cdot\frac{1}{2^{10}}\approx 0{,}205;$$

Сложим вероятности несовместных событий:

$$\mathsf{P}(4\leqslant\nu_{10}\leqslant 6)=\mathsf{P}(\nu_{10}=4)+\mathsf{P}(\nu_{10}=5)+\mathsf{P}(\nu_{10}=6)\approx 0{,}656.$$

§ 2. Номер первого успешного испытания

Рассмотрим схему Бернулли с вероятностью успеха p в одном испытании и обозначим через τ номер первого успешного испытания. Величина τ может принимать любые значения из множества натуральных чисел. Событие $\{\tau=5\}$ означает, что первые четыре испытания закончились неудачей, а в пятом испытании произошёл успех.

Т е о р е м а 12. *Вероятность того, что первый успех произойдёт в испытании с номером* $k\in\mathbb{N}=\{1, 2, 3, \ldots\}$, *равна* $\mathsf{P}(\tau=k)=pq^{k-1}$.

Д о к а з а т е л ь с т в о. Вероятность первым $k-1$ испытаниям завершиться неудачей, а последнему — успехом, равна

$$\mathsf{P}(\tau=k)=\mathsf{P}(н,\ \ldots,\ н,\ у)=pq^{k-1}. \qquad \square$$

Набор вероятностей $\{pq^{k-1}\}$, где k принимает любые значения из множества натуральных чисел, называется *геометрическим распределением* вероятностей. Геометрическое распределение вероятностей обладает интересным свойством *отсутствия последействия*, означающим «нестарение» устройства, время жизни которого подчинено геометрическому распределению.

Т е о р е м а 13. *Пусть* $\mathsf{P}(\tau=k)=pq^{k-1}$ *для любого* $k\in\mathbb{N}$. *Тогда для любых неотрицательных целых* n *и* k *имеет место равенство:*

$$\mathsf{P}(\tau>n+k\mid\tau>n)=\mathsf{P}(\tau>k).$$

Если, например, считать величину τ временем безотказной работы некоторого устройства (измеряемым целым числом часов), то данному равенству можно придать следующее звучание: вероятность работающему в данный момент устройству проработать ещё k часов не зависит от того, сколько уже работает устройство. Эта вероятность такая же, как для нового устройства.

Доказательство. По определению условной вероятности,

$$\mathsf{P}(\tau > n+k \mid \tau > n) = \frac{\mathsf{P}(\tau > n+k, \tau > n)}{\mathsf{P}(\tau > n)} = \frac{\mathsf{P}(\tau > n+k)}{\mathsf{P}(\tau > n)}. \tag{9}$$

Последнее равенство верно в силу того, что событие $\{\tau > n+k\}$ влечёт событие $\{\tau > n\}$, поэтому их пересечением будет событие $\{\tau > n+k\}$. Найдём для целого $m \geqslant 0$ вероятность $\mathsf{P}(\tau > m)$: событие $\{\tau > m\}$ означает, что в схеме Бернулли первые m испытаний завершились «неудачами», т. е. его вероятность равна q^m. Возвращаясь к (9), получим

$$\mathsf{P}(\tau > n+k \mid \tau > n) = \frac{\mathsf{P}(\tau > n+k)}{\mathsf{P}(\tau > n)} = \frac{q^{n+k}}{q^n} = q^k = \mathsf{P}(\tau > k). \qquad \square$$

Пример 35. Два игрока по очереди подбрасывают правильную игральную кость. Выигрывает тот, кто первым выкинет шесть очков. Найти вероятность победы игрока, начинающего игру.

Решение. Мы уже решили эту задачу в примере 27. Решим её иначе. Шесть очков может впервые выпасть при первом, втором, и т. д. бросках кости. Первый игрок побеждает, если это случится при броске с нечётным номером, второй — с чётным.

Пусть событие A_k состоит в том, что что шесть очков впервые выпадет в испытании с номером k. По теореме 12, $\mathsf{P}(A_k) = (1/6) \cdot (5/6)^{k-1}$.

События A и B, означающие победу первого и второго игроков соответственно, представимы в виде объединения взимоисключающих событий:

$$A = A_1 \cup A_3 \cup A_5 \cup \ldots, \qquad B = B_2 \cup B_4 \cup B_6 \cup \ldots$$

Вероятности этих объединений равны суммам вероятностей слагаемых:

$$\mathsf{P}(A) = \frac{1}{6} + \frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 + \frac{1}{6} \cdot \left(\frac{5}{6}\right)^4 + \ldots = \frac{1/6}{1-(5/6)^2} = \frac{6}{11},$$

$$\mathsf{P}(B) = \frac{1}{6} \cdot \frac{5}{6} + \frac{1}{6} \cdot \left(\frac{5}{6}\right)^3 + \frac{1}{6} \cdot \left(\frac{5}{6}\right)^5 + \ldots = \frac{5/36}{1-(5/6)^2} = \frac{5}{11}.$$

Поскольку $\mathsf{P}(A) + \mathsf{P}(B) = 1$, вероятность ничьей оказалсь равна нулю, как мы и ожидали в примере 27.

§ 3. Теорема Пуассона для схемы Бернулли

П р и м е р 36. Запишем вероятность получить не менее пяти успехов в тысяче испытаний схемы Бернулли с вероятностью успеха $0{,}003$:

$$\sum_{k=5}^{1\,000} C_{1\,000}^{k} \cdot 0{,}003^{k} \cdot 0{,}997^{1\,000-k} = 1 - \sum_{k=0}^{4} C_{1\,000}^{k} \cdot 0{,}003^{k} \cdot 0{,}997^{1\,000-k}.$$

Сформулируем теорему, позволяющую приближённо вычислять подобные вероятности.

Т е о р е м а 14 (теорема Пуассона). *Пусть в схеме Бернулли число испытаний n велико, а вероятность успеха p мала. Тогда вероятность получить k успехов в n испытаниях можно вычислить по приближённой формуле Пуассона*

$$\mathsf{P}(\nu_n = k) \approx \frac{\lambda^k}{k!}\, e^{-\lambda}, \quad \textit{где} \quad \lambda = np.$$

Набор вероятностей $\{\lambda^k e^{-\lambda} / k!\}$, где k принимает значения $0, 1, 2, \ldots$, называется *распределением Пуассона* с параметром $\lambda > 0$.

Обычно рекомендуют использовать приближение Пуассона при $npq \leqslant 10$. В каждом конкретном случае полезно оценить погрешность приближения, т. е. разницу между точным и приближённым значением вероятности. В зависимости от полученного значения погрешности можно выбирать, использовать ли формулу Пуассона или вычислять вероятности как-нибудь иначе. Так, например, *при вычислении любых вероятностей по теореме Пуассона погрешность не превысит наименьшей из двух величин: p и np^2.*

Вернёмся к примеру 36 и вычислим искомую вероятность по теореме 14. Число испытаний $n = 1\,000$ велико, а вероятность успеха $p = 0{,}003$ мала. Взяв $\lambda = np = 3$, можно записать приближённое равенство:

$$\begin{aligned}\mathsf{P}(\nu_n \geqslant 5) \;&=\; 1 - \mathsf{P}(\nu_n \leqslant 4) \;=\; 1 - \sum_{k=0}^{4} \mathsf{P}(\nu_n = k) \;\approx \\ &\approx\; 1 - \sum_{k=0}^{4} \frac{3^k}{k!}\, e^{-3} \;=\; 1 - e^{-3} \cdot \left(1 + 3 + \frac{9}{2} + \frac{27}{6} + \frac{81}{24}\right) \approx 0{,}1847.\end{aligned}$$

Насколько «приблизителен» ответ $0{,}1847$? Погрешность приближения не превышает $\min\{p,\, np^2\} = \min\{0{,}003;\, 0{,}009\} = 0{,}003$.

На самом деле точность гораздо выше. Правильное значение искомой вероятности равно $0{,}18448$ и лишь в четвёртом знаке отличается от ответа $0{,}1847$, полученного по формуле Пуассона.

§ 4. Предельная теорема Муавра — Лапласа

Мы только что научились вычислять вероятности в схеме Бернулли при очень маленькой вероятности успеха и большом числе испытаний. А что делать в следующей задаче?

Пример 37. Правильную монету подбрасывают $10\,000$ раз. Найти вероятность того, что герб выпадет от $4\,900$ до $5\,100$ раз.

Запишем эту вероятность:

$$\mathsf{P}(4\,900 \leqslant \nu_n \leqslant 5\,100) = \sum_{k=4\,900}^{5\,100} C_{10\,000}^{k} \cdot 0{,}5^{10\,000}.$$

Здесь вероятность успеха не мала, но число испытаний настолько велико, что вычислить напрямую эту сумму крайне сложно. Следующая теорема позволяет приближённо вычислять подобные вероятности.

Пусть число испытаний n велико, а вероятность успеха p не слишком мала (достаточно, чтобы $npq \geqslant 10$).

Теорема 15 (интегральная теорема Муавра — Лапласа). *Вероятность получить от k_1 до k_2 успехов в n испытаниях схемы Бернулли можно вычислить по следующей приближённой формуле:*

$$\mathsf{P}(k_1 \leqslant \nu_n \leqslant k_2) \;\approx\; \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right),$$

где $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2}\,dt.$

Значения функции $\Phi(x)$ находятся по таблицам (стр. 121). В дальнейшем мы подробно изучим свойства этой функции, а также познакомимся с центральной предельной теоремой, следствием которой является теорема 15.

Вернёмся к примеру 37 и вычислим искомую вероятность по теореме Муавра — Лапласа:

$$\mathsf{P}(4\,900 \leqslant \nu_n \leqslant 5\,100) \;\approx\; \Phi\left(\frac{5\,100 - 5\,000}{\sqrt{2\,500}}\right) - \Phi\left(\frac{4\,900 - 5\,000}{\sqrt{2\,500}}\right) =$$
$$= \Phi(2) - \Phi(-2) = 0{,}97725 - 0{,}02275 = 0{,}9545.$$

Мы получили странный результат: с вероятностью более 0,95 число гербов после десяти тысяч бросков монеты лежит в границах $5\,000 \pm 100$. Иными словами, есть только 5% шансов, что количество выпавших гербов будет отличаться от 5 000 более чем на 100.

Диапазоном, в котором возможное число гербов лежит с вероятностью единица (т. е. обязательно), будет отрезок от 0 до $10\,000$. Чтобы сузить этот диапазон, придётся пожертвовать достоверностью нашего предсказания. Интересно, в каких границах лежит число гербов с вероятностью *чуть* меньшей единицы? Например, с вероятностью 0,999? Будут ли это границы от 10 до 9 990? Или от 1 000 до 9 000? А может быть, от 2 000 до 8 000?

Воспользуемся теоремой Муавра — Лапласа. Требуется найти x такое, что

$$\mathsf{P}(5\,000 - x \leqslant \nu_n \leqslant 5\,000 + x) \approx \Phi\left(\frac{x}{\sqrt{2\,500}}\right) - \Phi\left(-\frac{x}{\sqrt{2\,500}}\right) = \\ = \Phi(0{,}02\,x) - \Phi(-0{,}02\,x) = 0{,}999.$$

Функция Φ обладает следующим свойством: $\Phi(-x) = 1 - \Phi(x)$. Поэтому

$$0{,}999 = \Phi(0{,}02\,x) - \Phi(-0{,}02\,x) = 1 - 2\,\Phi(-0{,}02\,x),$$

откуда $\Phi(-0{,}02\,x) = 0{,}0005$. По таблице находим $0{,}02\,x = 3{,}29$, $x = 164{,}5$. Итак, с вероятностью 0,999 число гербов после десяти тысяч бросков монеты лежит в границах от $5\,000 - 165$ до $5\,000 + 165$.

Мы получили диапазон длиной всего 330. Тем не менее, имеется лишь одна десятая процента шансов обнаружить после 10 000 бросков монеты число гербов за пределами этого диапазона.

Сформулируем без доказательства ещё одно утверждение, позволяющее вычислять локальную вероятность $\mathsf{P}(\nu_n = k)$ при большом числе испытаний n и не слишком малой вероятности успеха p.

Т е о р е м а 16 (локальная теорема Муавра — Лапласа). *Вероятность получить k успехов в n испытаниях схемы Бернулли можно вычислить по следующей приближённой формуле:*

$$\mathsf{P}(\nu_n = k) \approx \frac{1}{\sqrt{npq}}\,\varphi\left(\frac{k - np}{\sqrt{npq}}\right),$$

где $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}\,e^{-x^2/2}$.

Значения функции $\varphi(x)$ либо вычисляют вручную, либо находят по соответствующим таблицам.

П р и м е р 38. Вычислим вероятность получить ровно 5 000 гербов после десяти тысяч подбрасываний правильной монеты.

Р е ш е н и е. По теореме 16,

$$\mathsf{P}(\nu_n = 5\,000) \approx \frac{1}{\sqrt{2\,500}}\,\varphi(0) = 0{,}02 \cdot \frac{1}{\sqrt{2\pi}} \approx 0{,}00798.$$

При таком большом числе испытаний приближённая формула даёт практически точный ответ: все указанные значащие цифры ответа верны.

§ 5. Независимые испытания с несколькими исходами

Рассмотрим схему независимых испытаний уже не с двумя, а с бóльшим количеством возможных результатов в каждом испытании.

П р и м е р 39. Игральная кость подбрасывается пятнадцать раз. Найти вероятность того, что выпадет ровно десять троек и три единицы.

Здесь каждое испытание имеет три, а не два исхода: выпадение тройки, выпадение единицы, выпадение любой другой грани. Поэтому воспользоваться формулой для числа успехов в схеме Бернулли не удастся.

Попробуем вывести подходящую формулу. Пусть в одном испытании возможны m исходов: $1, 2, \ldots, m$, и i-й исход в одном испытании случается с вероятностью p_i, где $p_1 + \ldots + p_m = 1$.

Обозначим через $P(n_1, \ldots, n_m)$ вероятность того, что в n независимых испытаниях первый исход случится n_1 раз, второй исход — n_2 раз, и т. д., наконец, m-й исход — n_m раз.

Т е о р е м а 17. *Для любого n и любых неотрицательных целых чисел $n_1, \ldots, n_m$, сумма которых равна n, верна формула:*

$$P(n_1, \ldots, n_m) = \frac{n!}{n_1! \ldots n_m!}\, p_1^{n_1} \cdot \ldots \cdot p_m^{n_m}.$$

Д о к а з а т е л ь с т в о. Рассмотрим один элементарный исход, благоприятствующий выпадению n_1 единиц, n_2 двоек, и т. д.:

$$(\underbrace{1, \ldots, 1}_{n_1}, \underbrace{2, \ldots, 2}_{n_2}, \ldots, \underbrace{m, \ldots, m}_{n_m}).$$

Это результат n экспериментов, когда все нужные исходы появились в некотором заранее заданном порядке. Вероятность такого результата равна произведению вероятностей $p_1^{n_1} \cdot \ldots \cdot p_m^{n_m}$. Остальные благоприятные исходы отличаются лишь расположением чисел $1, 2, \ldots, m$ на n местах. Число таких исходов равно числу способов расположить на n местах n_1 единиц, n_2 двоек, и т. д. Это число равно

$$C_n^{n_1} \cdot C_{n-n_1}^{n_2} \cdot C_{n-n_1-n_2}^{n_3} \cdot \ldots \cdot C_{n-n_1-\ldots-n_{m-1}}^{n_m} = \frac{n!}{n_1! \ldots n_m!}. \quad \square$$

Теперь мы можем вернуться к примеру 39 и выписать ответ: так как вероятности выпадения тройки и единицы равны по $1/6$, а вероятность третьего исхода (выпала любая другая грань) равна $4/6$, то вероятность получить десять троек, три единицы и ещё два других очка равна

$$P(10, 3, 2) = \frac{15!}{10!\, 3!\, 2!} \cdot \left(\frac{1}{6}\right)^{10} \cdot \left(\frac{1}{6}\right)^{3} \cdot \left(\frac{4}{6}\right)^{2}.$$

§ 6. Контрольные вопросы

1. Что такое схема Бернулли?

2. Какова вероятность получить пять успехов в пяти испытаниях схемы Бернулли? Пять неудач?

3. Какова вероятность получить пять успехов в десяти испытаниях схемы Бернулли?

4. Какова вероятность получить сначала пять успехов, а затем пять неудач в десяти испытаниях схемы Бернулли?

5. Почему в формуле Бернулли присутствует число сочетаний?

6. Какова вероятность первому успеху случиться в первом испытании? В пятом?

7. Какова вероятность номеру первого успеха быть больше пяти?

8. Что такое биномиальное распределение?

9. Чему равна сумма по всем k вероятностей в формуле Бернулли?

10. Что такое геометрическое распределение?

11. Чему равна сумма вероятностей в геометрическом распределении?

12. Что такое отсутствие последействия?

13. Для чего нужна теорема Пуассона?

14. Можно ли пользоваться теоремой Пуассона при $n = 100$, $p = 0{,}5$?

15. Чем по теореме Пуассона можно приблизить вероятность иметь хотя бы один успех в $1\,000$ испытаний с вероятностью успеха $0{,}001$?

16. Для чего нужна локальная теорема Муавра — Лапласа?

17. Для чего нужна интегральная теорема Муавра — Лапласа?

18. Можно ли пользоваться теоремами Муавра — Лапласа при $n = 100$ и $p = 0{,}5$? А при $n = 100$ и $p = 0{,}01$?

19. Как вычислять значения функции $\Phi(x)$?

20. Приведите пример задачи, для решения которой необходима интегральная теорема Муавра — Лапласа.

21. Приведите пример задачи, для решения которой необходима локальная теорема Муавра — Лапласа.

22. В какой точке функция $\varphi(x)$ достигает наибольшего значения?

23. При каком k максимальна вероятность получить k успехов в большом числе испытаний схемы Бернулли?

24. По какой формуле можно вычислить вероятность при шести бросаниях игральной кости выбросить по разу каждое число очков на кости?

ГЛАВА V
СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И ИХ РАСПРЕДЕЛЕНИЯ

Мы уже видели, что для многих экспериментов нет никаких различий в подсчёте *вероятностей* событий, тогда как элементарные исходы в этих экспериментах очень различаются. С вероятностью 0,5 при броске монеты выпадает герб, на игральной кости — чётное число очков, точка падает на левую половину отрезка, вынутая из колоды карта оказывается красной и т. д. Пора во всех таких «похожих» экспериментах для обозначения элементарных исходов использовать, например, числа. Иначе говоря, пора каждый элементарный исход заменить действительным числом, не обязательно уникальным, и работать только с числами.

§ 1. Случайные величины

Пусть задано вероятностное пространство $\langle \Omega, \mathcal{F}, \mathsf{P} \rangle$.

Определение 12. *Случайной величиной* называется функция, которая каждому элементарному исходу $\omega \in \Omega$ ставит в соответствие действительное число, т. е. функция, действующая из Ω в $\mathbb{R}$.

Случайные величины мы будем обозначать греческими буквами: ξ (кси), η (эта), χ (хи), ν (ню), τ (тау), ψ (пси) и др.

Корректности ради от случайных величин требуют так называемой *измеримости*. А именно, требуют, чтобы для любого действительного числа x множество элементарных исходов ω, для которых $\xi(\omega) < x$, принадлежало полю событий $\mathcal{F}$. Но функции, не обладающие этим свойством, являются исключительно редкими и экзотическими объектами, поэтому *нигде больше* мы измеримость упоминать не будем.

Пример 40. Подбрасывают один раз правильную игральную кость. Рассмотрим $\Omega = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$ и определим следующие случайные величины: 1) $\xi(\omega) = \omega$: эта случайная величина равна числу выпавших на кости очков, 2) $\eta(\omega) = 1$, если $\omega = 2,\ 4,\ 6$, иначе $\eta(\omega) = 0$: эта случайная величина служит индикатором того, выпало ли на кости чётное или нечётное число очков. В первом случае она становится равна единице, во втором — нулю.

Случайная величина ξ принимает значения $1, 2, 3, 4, 5, 6$ с вероятностями по одной шестой каждое. Например, $\xi = 1$, когда на кости выпало одно очко, т. е. с вероятностью $\frac{1}{6}$. Можно записать соответствие между значениями случайной величины ξ и вероятностями принимать эти значения в виде *таблицы распределения вероятностей* или, коротко, *таблицы распределения*:

ξ	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

В первой строке таблицы перечислены *значения* случайной величины ξ, во второй строке — *вероятности,* с которыми она принимает эти значения:

$$\mathsf{P}(\xi = 1) = \mathsf{P}(\xi = 2) = \ldots = \mathsf{P}(\xi = 6) = \frac{1}{6}.$$

Для случайной величины η таблица распределения выглядит так:

$$\mathsf{P}(\eta = 0) = \mathsf{P}(\text{выпало 1, 3 или 5 очков}) = 0{,}5;$$
$$\mathsf{P}(\eta = 1) = \mathsf{P}(\text{выпало 2, 4 или 6 очков}) = 0{,}5.$$

Пример 41. Стержень длиной 5 см ломается на две части в наудачу выбранном месте. Длину каждого из полученных обломков можно считать случайной величиной, принимающей любые значения из отрезка $[0, 5]$. Если ξ — длина левой части стержня, то $\eta = 5 - \xi$ — длина правой.

Пользуясь геометрической вероятностью, мы можем вычислить вероятности различных событий, связанных со случайными величинами ξ и η, как это делалось в примере 20 на стр. 31.

Например, $\mathsf{P}(1 \leqslant \xi \leqslant 2) = 0{,}2$ — вероятность того, что длина левой части окажется от 1 до 2 см, $\mathsf{P}(\eta \geqslant 3) = 0{,}4$ — вероятность правому обломку быть длиннее 3 см, $\mathsf{P}(\xi = 2) = 0$ — вероятность поделить стержень на части *ровно* в 2 и 3 см длиной, $\mathsf{P}(\eta \geqslant 6) = 0$ — вероятность невозможного события.

Пример 42. Правильная монета подбрасывается 10 раз. Случайная величина ν_{10} из примера 34 на стр. 44 равна количеству выпавших гербов.

Эта случайная величина принимает целые значения от 0 до 10. В примере 34 вычислены некоторые вероятности, связанные с ней:

$$\mathsf{P}(\nu_{10} = 4) = \mathsf{P}(\nu_{10} = 6) \approx 0{,}205, \quad \mathsf{P}(\nu_{10} = 5) \approx 0{,}246.$$

Можно записать все вероятности по формуле Бернулли: для $k = 0, 1, \ldots, 10$

$$\mathsf{P}(\nu_n = k) = C_{10}^k \cdot 0{,}5^{10}.$$

Из-за симметрии монеты число выпавших *решек* имеет такое же распределение с такой же таблицей распределения.

§ 2. Распределения случайных величин

Опишем различные типы распределений случайных величин. Случайная величина может принимать отдельные значения, т. е. вся вероятностная масса может быть сосредоточена в нескольких точках прямой. Случайная величина может принимать любые значения из некоторого интервала или отрезка, т. е. вероятностная масса может быть «размазана» по некоторому интервалу или по всей прямой. В зависимости от этого распределения делят на дискретные, абсолютно непрерывные и их смеси.

Мы ограничимся в основном первыми двумя видами распределений.

Определение 13. Случайная величина ξ имеет *дискретное* распределение, если существует конечный или счётный набор чисел $a_1, a_2, \ldots$ и набор вероятностей $p_1 = \mathsf{P}(\xi = a_1),\ p_2 = \mathsf{P}(\xi = a_2), \ldots$ таких, что

$$p_1 + p_2 + \ldots = 1.$$

Итак, случайная величина ξ имеет дискретное распределение, если множество её значений конечно или счётно. Если случайная величина ξ имеет дискретное распределение, то для любого подмножества прямой $B \subseteq \mathbb{R}$

$$\mathsf{P}(\xi \in B) = \sum_{a_i \in B} \mathsf{P}(\xi = a_i).$$

Дискретное распределение удобно задавать следующей таблицей, в которой $p_i = \mathsf{P}(\xi = a_i)$:

ξ	a_1	a_2	a_3	$\ldots$
P	p_1	p_2	p_3	$\ldots$

Случайные величины из примеров 40 и 42 имеют дискретные распределения, в отличие от случайных величин из примера 41. Они имеют абсолютно непрерывные распределения.

Определение 14. Случайная величина ξ имеет *абсолютно непрерывное* распределение, если существует неотрицательная функция $f(x)$ такая, что для любого интервала $(a,\ b)$ имеет место равенство:

$$\mathsf{P}(a < \xi < b) = \int\limits_a^b f(x)\,dx.$$

Функция $f(x)$ называется *плотностью распределения.*

Если нам будет нужно различать плотности распределения разных случайных величин, мы будем использовать обозначения $f_\xi(x),\ f_\eta(x)$ и т. д.

Мы уже встречались в § 3 главы II с плотностями распределения. Определение 14 напоминает, что для случайной величины ξ с плотностью распределения $f(x)$ вероятность принимать значения внутри любого интервала вычисляется как площадь под графиком $f(x)$ над этим интервалом.

Теорема 18. *Любая плотность распределения обладает свойствами*:

$$\text{(f1)}\quad f(x) \geqslant 0 \quad \textit{для любого } x; \qquad \text{(f2)}\quad \int\limits_{-\infty}^{\infty} f(x)\,dx = 1.$$

Доказательство. Свойство (f1) выполнено по определению 14, свойство (f2) также следует из него:

$$1 = \mathsf{P}(-\infty < \xi < \infty) = \int\limits_{-\infty}^{\infty} f(x)\,dx. \qquad \square$$

Свойство (f2) означает, что для любой плотности распределения площадь всей области под графиком плотности распределения выше оси абсцисс равна единице. Заметим, что *любая функция, удовлетворяющая* (f1) *и* (f2), *является плотностью распределения некоторой случайной величины.*

§ 3. Функция распределения и её свойства

Определение и примеры. Мы описали дискретные распределения таблицей распределения, абсолютно непрерывные — плотностью распределения. Существует универсальный способ описать любое распределение. Это можно сделать с помощью функции распределения.

Определение 15. *Функцией распределения* случайной величины ξ называется функция $F(x)$, при каждом x равная вероятности случайной величине ξ принимать значения, меньшие x:

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

Смысл функции распределения весьма прост. Представим себе, что на числовой прямой размещена единичная масса. Причём размещена она там в соответствии с распределением ξ: если оно дискретное, частички p_i этой массы помещены в точки с координатами a_i. Если же распределение абсолютно непрерывно, масса «размазана» по прямой с плотностью $f(x)$, т. е. на участок $[x, x + dx]$ приходится масса $f(x)\,dx$. Тогда функция распределения $F(x)$ при каждом x равна массе, расположенной слева от точки x.

С ростом x наблюдается увеличение массы, находящейся слева от x. Это увличение происходит либо непрерывно (если масса распределена по всем точкам прямой), либо скачками (если масса помещена в отдельные точки прямой). Если x неограниченно передвигать влево, масса слева от x будет уменьшаться до нуля. Если x так же неограниченно передвигать вправо, масса слева от x будет расти до единицы — до общей массы на всей прямой.

Рассмотрим примеры того, как находить функции распределения.

П р и м е р 43. Дважды подбрасывают правильную монету. Найти функцию распределения числа выпавших гербов.

Р е ш е н и е. Число выпавших гербов — это случайная величина ξ, которая может принимать значения 0, 1, 2 с вероятностями p_0, p_1, p_2. Вероятности можно найти или по формуле Бернулли, или напрямую:

$$p_0 = \mathsf{P}(\xi = 0) = \mathsf{P}(\textit{две решки}) = 0{,}5 \cdot 0{,}5 = 0{,}25,$$
$$p_1 = \mathsf{P}(\xi = 1) = \mathsf{P}(\textit{герб, решка} \text{ или } \textit{решка, герб}) = 2 \cdot 0{,}5 \cdot 0{,}5 = 0{,}5,$$
$$p_2 = \mathsf{P}(\xi = 2) = \mathsf{P}(\textit{два герба}) = 0{,}25 = 1 - p_0 - p_1.$$

Функция распределения $F(x)$ равна вероятности события $\{\xi < x\}$ и её значения зависят от x следующим образом:

а) при $x \leqslant 0$ событие $\{\xi < x\}$ невозможно, его вероятность нулевая;

б) если $0 < x \leqslant 1$, то событие $\{\xi < x\}$ происходит в том и только в том случае, когда $\xi = 0$, т. е. с вероятностью $p_0 = 0{,}25$;

в) если $1 < x \leqslant 2$, то событие $\{\xi < x\}$ происходит, если $\xi = 0$ или $\xi = 1$, т. е. с вероятностью $p_0 + p_1 = 0{,}75$;

г) при $x > 2$ событие $\{\xi < x\}$ обязательно происходит, и его вероятность равна единице.

$$F(x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ 0{,}25, & \text{если } 0 < x \leqslant 1, \\ 0{,}75, & \text{если } 1 < x \leqslant 2, \\ 1, & \text{если } x > 2. \end{cases}$$

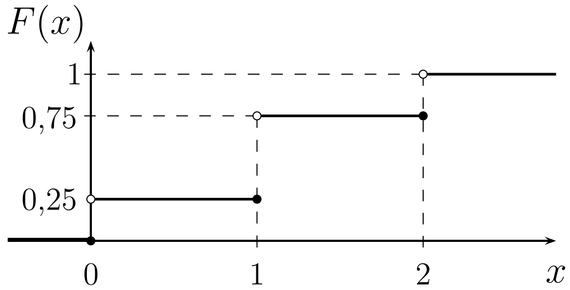

Рис. 10.

Закрашенные и незакрашенные точки на графике указывают на значения функции в точке разрыва. Например, закрашенная точкой (0, 0) и «пустая» точка (0, 0,25) означают, что $F_\xi(0) = 0$, а не 0,25.

Заметим, что функция распределения имеет разрывы в точках 0, 1 и 2 — именно в тех точках, где ξ принимает значения. Приращения функции $F_\xi(x)$ в этих точках равны соответственно $p_0 = 0{,}25$, $p_1 = 0{,}5$ и $p_2 = 0{,}25$.

Аналогично строится график функции распределения для любого дискретного распределения со значениями a_i и вероятностями p_i.

С в о й с т в о 2. *Функция распределения дискретного распределения растёт только в точках разрыва: в каждой точке a_i она имеет скачок величиной p_i. За счёт скачков функция $F(x)$ растёт от* 0 *до* 1.

П р и м е р 44. Точка наудачу бросается в круг с радиусом R. Найти функцию распределения расстояния от этой точки до центра круга.

Р е ш е н и е. Случайная величина ξ равна расстоянию от наудачу выбранной в круге точки до центра круга и принимает любые значения от 0 до R. Поэтому при $x \leqslant 0$ событие $\{\xi < x\}$ невозможно, и $F_\xi(x) = 0$. А при $x > R$ событие $\{\xi < x\}$ происходит обязательно, и $F_\xi(x) = 1$ для таких x.

Если же $0 < x \leqslant R$, то событие $\{\xi < x\}$ происходит при попадании точки в круг с радиусом x и тем же центром. Вероятность этого события, по геометрическому определению вероятности, равна отношению площадей кругов: $F_\xi(x) = \dfrac{\pi x^2}{\pi R^2} = \dfrac{x^2}{R^2}$. Окончательно получаем:

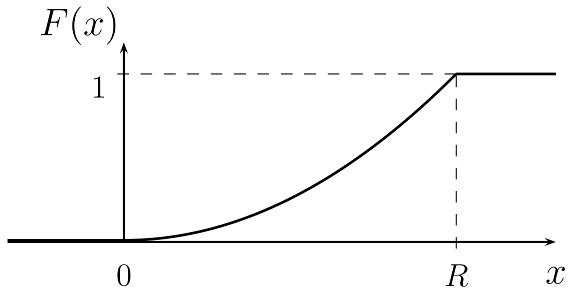

Рис. 11. График функции $F(x)$

$$F(x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ \dfrac{x^2}{R^2}, & \text{если } 0 < x \leqslant R, \\ 1, & \text{если } x > 2. \end{cases}$$

Здесь, в отличие от предыдущего примера, функция $F(x)$ непрерывна.

Сделаем важное замечание: две разные случайные величины могут иметь одну и ту же функцию распределения, одну и ту же таблицу или плотность распределения. В этом случае говорят, что данные случайные величины *одинаково распределены*. Иными словами, функция распределения характеризует не столько саму случайную величину, сколько её распределение.

П р и м е р 45. По результату одного подбрасывания правильной монеты построим две случайные величины:

$$\xi = \begin{cases} 1, & \text{если выпал герб}, \\ 0, & \text{если выпала решка}, \end{cases} \qquad \eta = \begin{cases} 0, & \text{если выпал герб}, \\ 1, & \text{если выпала решка}. \end{cases}$$

Эти случайные величины никогда не равны одна другой: если монета выпала гербом, то $\xi = 1$, а $\eta = 0$. Если решкой — наоборот. Однако распределения этих случайных величин одинаковы:

ξ	0	1
P	0,5	0,5

η	0	1
P	0,5	0,5

Одинаковыми будут и функции распределения этих случайных величин: $F_\xi(x) = \mathsf{P}(\xi < x) = \mathsf{P}(\eta < x) = F_\eta(x)$ для всех x.

Общие свойства функций распределения. Основные свойства функций распределения заключены в теореме 19. Верно и обратное к ней утверждение: любая функция, обладающая перечисленными в теореме свойствами, является функцией распределения некоторой случайной величины.

Т е о р е м а 19. *Любая функция распределения обладает свойствами:*

(F1) $F(x)$ *не убывает: если* $x_1 < x_2$, *то* $F(x_1) \leqslant F(x_2)$;

(F2) $F(x)$ *имеет пределы* $\lim\limits_{x\to-\infty} F(x) = 0$ *и* $\lim\limits_{x\to+\infty} F(x) = 1$;

(F3) $F(x)$ *в любой точке непрерывна слева:* $F(x-0) = F(x)$.

В свойстве (F3) через $F(x-0)$ обозначен предел значений функции F при стремлении её аргумента к точке x слева (т. е. по области $x_n < x$):

$$F(x-0) = \lim_{x_n\to x-0} F(x_n).$$

Д о к а з а т е л ь с т в о. Докажем свойство (F1). Рассмотрим произвольные числа $x_1 < x_2$. Если $\xi < x_1$, то тем более $\xi < x_2$. Это означает, что первое событие влечёт второе, и их вероятности тоже упорядочены:

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

Ограничимся, не приводя строгого доказательства свойств (F2) и (F3), следующими соображениями.

Вероятность события $\{\xi < x\}$ при $x \to -\infty$ стремится к вероятности невозможного события $\{\xi = -\infty\}$, т. е. к нулю. Вероятность события $\{\xi < x\}$ при $x \to +\infty$ стремится к вероятности достоверного события $\{\xi < +\infty\}$, т. е. к единице. Если же x_n устремить к x слева, то вероятность события $\{\xi < x_n\}$ стремится к вероятности события $\{\xi < x\}$, т. е. к $F(x)$. $\square$

Функции распределения обладают также следующим полезным свойством:

(F4) *Для любой случайной величины* ξ *имеет место равенство:*

$$\mathsf{P}(a \leqslant \xi < b) = F(b) - F(a). \tag{10}$$

Д о к а з а т е л ь с т в о. Заметим, что $\{\xi < a\} \cup \{a \leqslant \xi < b\} = \{\xi < b\}$, и первые два события несовместны. Тогда из аксиомы (P2) следует, что

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

Это как раз и означает, что

$$F(a) + \mathsf{P}(a \leqslant \xi < b) = F(b). \qquad \square$$

Если функция распределения непрерывна в точках a и b, то следующие четыре вероятности одинаковы, и любая из них вычисляется по формуле (10):

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

Свойства абсолютно непрерывного распределения. Пусть случайная величина ξ имеет абсолютно непрерывное распределение с плотностью $f(x)$. Тогда функция распределения в любой точке x может быть найдена по плотности распределения так:

$$F(x) = \mathsf{P}(\xi < x) = \mathsf{P}(\xi \in (-\infty, x)) = \int\limits_{-\infty}^{x} f(t)\,dt. \tag{11}$$

Интеграл (11) непрерывен по x, поэтому функция распределения случайной величины с абсолютно непрерывным распределением *всюду непрерывна*.

Из равенства (11) следует также, что плотность абсолютно непрерывного распределения равна производной от функции распределения: $f(x) = F'(x)$.

Это равенство выполняется для всех x, кроме, разве что, конечного или счётного множества точек, в которых производная функции $F(x)$ может вообще не существовать. В таких точках плотность можно задавать как угодно.

Сформулируем очень важный критерий абсолютной непрерывности.

Т е о р е м а 20. *Если непрерывная функция распределения $F(x)$ дифференцируема всюду, за исключением, возможно, конечного или счётного множества точек, то распределение является абсолютно непрерывным.*

Так, в примере 44 функция распределения $F(x)$ везде непрерывна и дифференцируема всюду, кроме точек 0 и R. Плотность этого распределения равна $f(x) = F'(x)$ (*продифференцировать!*).

§ 4. Примеры дискретных распределений

Вырожденное распределение I_c. Случайная величина ξ имеет вырожденное распределение в точке c, если ξ принимает лишь одно значение c, т. е. $\mathsf{P}(\xi = c) = 1$. Для краткости мы будем употреблять запись $\xi \sim \mathrm{I}_c$.

Распределение Бернулли B_p. Случайная величина ξ имеет распределение Бернулли с параметром $p \in (0, 1)$, если ξ принимает значение 1 с вероятностью p и значение 0 с вероятностью $q = 1 - p$. Случайная величина ξ с таким распределением равна *числу успехов в одном испытании* схемы Бернулли с вероятностью успеха p: либо ни одного успеха, либо один успех. Таблица распределения случайной величины $\xi \sim \mathrm{B}_p$ имеет вид:

ξ	0	1
P	q	p

Например, случайные величины ξ и η из примера 45 имеют одно и то же распределение Бернулли с параметром $p = \dfrac{1}{2}$.

Биномиальное распределение $\mathrm{B}_{n,p}$. Случайная величина ξ имеет биномиальное распределение с параметрами $n \in \mathbb{N}$ и $p \in (0,\,1)$, если ξ принимает значения $k = 0,\,1,\,\ldots,\,n$ с вероятностями $\mathsf{P}(\xi = k) = C_n^k p^k q^{n-k}$. Случайная величина с таким распределением равна *числу успехов в n испытаниях* схемы Бернулли с вероятностью успеха p. Таблица распределения случайной величины $\xi \sim \mathrm{B}_{n,p}$ имеет вид:

ξ	0	1	$\ldots$	k	$\ldots$	n
P	q^n	npq^{n-1}	$\ldots$	$C_n^k p^k q^{n-k}$	$\ldots$	p^n

Распределение Бернулли B_p совпадает с распределением $\mathrm{B}_{1,\,p}$.

Например, случайная величина ν_{10} из примера 42 имеет биномиальное распределение $\mathrm{B}_{10,\,\frac{1}{2}}$ с параметрами $n = 10$, $p = \dfrac{1}{2}$. Количество выпавших шестёрок при двадцати подбрасываниях правильной игральной кости имеет биномиальное распределение $\mathrm{B}_{20,\,\frac{1}{6}}$.

Геометрическое распределение G_p. Случайная величина ξ имеет геометрическое распределение с параметром $p \in (0,\,1)$, если ξ принимает значения $k = 1,\,2,\,3,\,\ldots$ с вероятностями $\mathsf{P}(\xi = k) = pq^{k-1}$. Случайная величина с таким распределением равна *номеру первого успешного испытания* в схеме Бернулли. Таблица распределения случайной величины $\xi \sim \mathrm{G}_p$ имеет вид:

ξ	1	2	$\ldots$	k	$\ldots$
P	p	pq	$\ldots$	pq^{k-1}	$\ldots$

Так, количество проведённых испытаний в примере 11 на стр. 21 имеет геометрическое распределение с параметром $p = \dfrac{1}{2}$.

Распределение Пуассона Π_λ. Случайная величина ξ имеет распределение Пуассона с параметром $\lambda > 0$, если ξ принимает целые неотрицательные значения $k = 0,\,1,\,2,\,\ldots$ с вероятностями $\mathsf{P}(\xi = k) = \dfrac{\lambda^k}{k!}\,e^{-\lambda}$. Таблица распределения случайной величины $\xi \sim \Pi_\lambda$ имеет вид:

τ	0	1	2	$\ldots$	k	$\ldots$
P	$e^{-\lambda}$	$\lambda\,e^{-\lambda}$	$\dfrac{\lambda^2}{2}\,e^{-\lambda}$	$\ldots$	$\dfrac{\lambda^k}{k!}\,e^{-\lambda}$	$\ldots$

Распределение Пуассона возникло в теореме Пуассона (стр. 46) как предельное распределение для числа успехов в n испытаниях схемы Бернулли, когда число испытаний n увеличивается, а вероятность успеха $p \sim \dfrac{\lambda}{n}$ уменьшается обратно пропорционально n. Поэтому распределение Пуассона называют иногда *распределением числа редких событий*.

Гипергеометрическое распределение. Пусть случайная величина ξ равна числу белых шаров среди n шаров, выбранных наудачу и без возвращения из урны с K белыми и $N-K$ чёрными шарами. Распределение этой случайной величины называется *гипергеометрическим* распределением. Случайная величина ξ принимает целые значения $k = 0, 1, \ldots, n$ с вероятностями

$$\mathsf{P}(\xi = k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}.$$

Упражнение. Построить графики функций распределения для всех перечисленных выше распределений, как это сделано в примере 43.

§ 5. Примеры абсолютно непрерывных распределений

Равномерное распределение $\mathrm{U}_{a,b}$. Случайная величина ξ имеет равномерное распределение на отрезке $[a, b]$, если её плотность распределения постоянна на отрезке $[a, b]$ и равна нулю вне него:

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

Случайная величина $\xi \sim \mathrm{U}_{a,b}$ имеет смысл *координаты точки, выбранной наудачу* на отрезке $[a, b]$. Функция распределения случайной величины ξ всюду непрерывна и имеет вид:

$$F(x) = \begin{cases} 0, & \text{если } x < a, \\ \dfrac{x-a}{b-a}, & \text{если } a \leqslant x \leqslant b, \\ 1, & \text{если } x > b. \end{cases}$$

Мы много раз встречались с равномерным распределением. Так, длина каждого из двух обломков стержня в примере 42 имеет равномерное распределение на отрезке $[0, 5]$.

Графики плотности и функции распределения равномерного распределения на отрезке $[a, b]$ изображены на рис. 12.

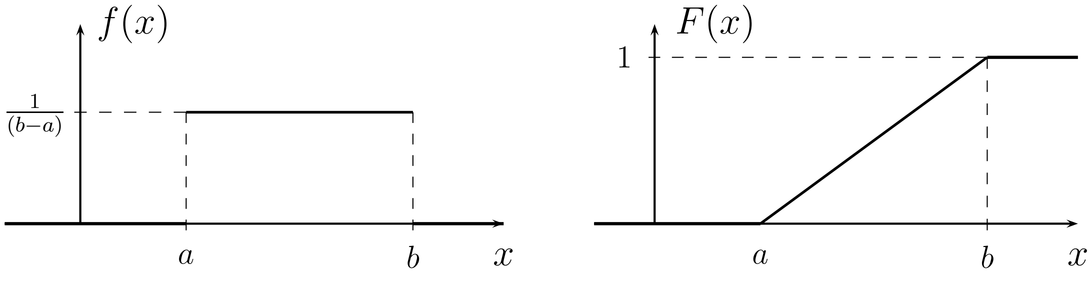

Рис. 12. Плотность и функция распределения $\mathrm{U}_{a,b}$

Показательное распределение E_α. Случайная величина ξ имеет показательное (экспоненциальное) распределение с параметром $\alpha > 0$, если ξ имеет следующие плотность и функцию распределения:

$$f(x) = \begin{cases} 0, & \text{если } x < 0, \\ \alpha e^{-\alpha x}, & \text{если } x \geqslant 0; \end{cases} \qquad F(x) = \begin{cases} 0, & \text{если } x < 0, \\ 1 - e^{-\alpha x}, & \text{если } x \geqslant 0. \end{cases}$$

Графики плотности и функции распределения показательного распределения с параметром α изображены на рис. 13.

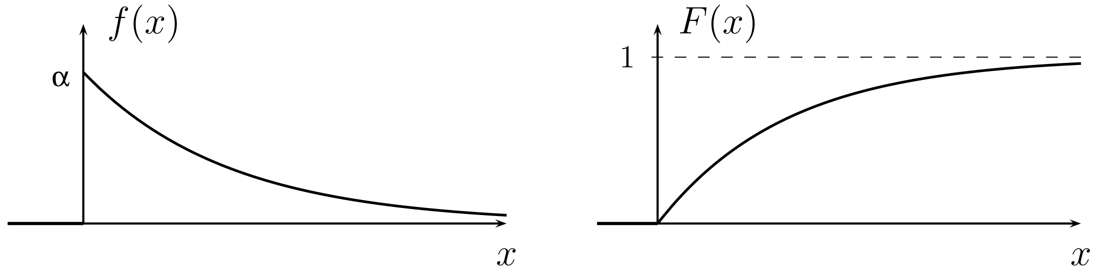

Рис. 13. Плотность и функция распределения E_α

Плотность показательного распределения равна нулю на отрицательной полуоси, поэтому вероятность события $\{\xi < 0\}$ нулевая — случайная величина с показательным распределением не может быть отрицательна. К тому же плотность отлична от нуля на всей положительной полуоси, поэтому случайная величина с показательным распределением может принимать сколь угодно большие положительные значения: для всякого x вероятность события $\{\xi > x\}$ не равна нулю.

Показательное распределение обладает свойством *отсутствия последействия*. Мы встречались с этим свойством у геометрического распределения.

Теорема 21. *Пусть $\xi \sim \mathrm{E}_\alpha$. Тогда для любых x, $y > 0$*

$$\mathsf{P}(\xi \geqslant x + y \mid \xi \geqslant x) = \mathsf{P}(\xi \geqslant y). \tag{12}$$

Доказательство. Заметим, что для любого $t > 0$

$$\mathsf{P}(\xi \geqslant t) = 1 - \mathsf{P}(\xi < t) = 1 - F(t) = e^{-\alpha t}.$$

Поэтому

$$\mathsf{P}(\xi \geqslant x + y \mid \xi \geqslant x) = \frac{\mathsf{P}(\xi \geqslant x + y,\ \xi \geqslant x)}{\mathsf{P}(\xi \geqslant x)} = \frac{\mathsf{P}(\xi \geqslant x + y)}{\mathsf{P}(\xi \geqslant x)} = $$
$$= \frac{e^{-\alpha(x+y)}}{e^{-\alpha x}} = e^{-\alpha y} = \mathsf{P}(\xi \geqslant y). \qquad \square$$

Нормальное распределение N_{a,σ^2}. Случайная величина ξ имеет нормальное (или гауссовское) распределение с параметрами a и σ^2, где $a \in \mathbb{R}$, $\sigma > 0$, если ξ имеет следующую плотность распределения:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

На рис. 14 показаны графики плотностей нормальных распределений с одним и тем же значением параметра a и разными значениями параметра σ.

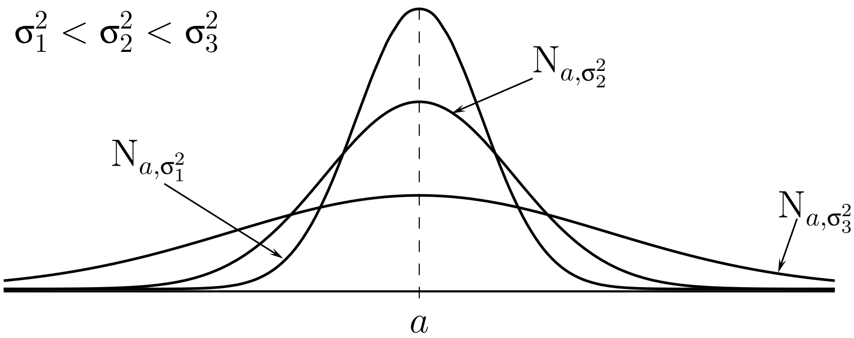

Рис. 14. Плотности нормальных распределений

Нормальное распределение $N_{0,1}$ с параметрами $a = 0$ и $\sigma^2 = 1$ называется *стандартным нормальным* распределением. Плотность стандартного нормального распределения равна $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Она является чётной функцией, и её график (кривая Гаусса) симметричен относительно прямой $x = 0$. Читатель вспомнит, что эта плотность уже возникала в локальной теореме Муавра — Лапласа (стр. 48).

Мы будем использовать специальное обозначение $\Phi(x)$ для функции распределения *стандартного нормального* распределения.

Первообразная функции e^{-x^2} не может быть выражена через элементарные функции. Поэтому функцию $\Phi(x)$ (так же как и функцию распределения произвольного нормального распределения) можно записать лишь в виде интеграла:

$$\Phi(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t^2}{2}}\, dt.$$

Функция $\Phi(x)$ встречалась нам в интегральной теореме Муавра — Лапласа (стр. 47). Она *табулирована,* т. е. её значения при различных действительных x вычислены (см. таблицу на стр. 121).

Познакомимся со свойствами нормального распределения. Прежде всего установим связь между функцией распределения произвольного нормального закона распределения и $\Phi(x)$.

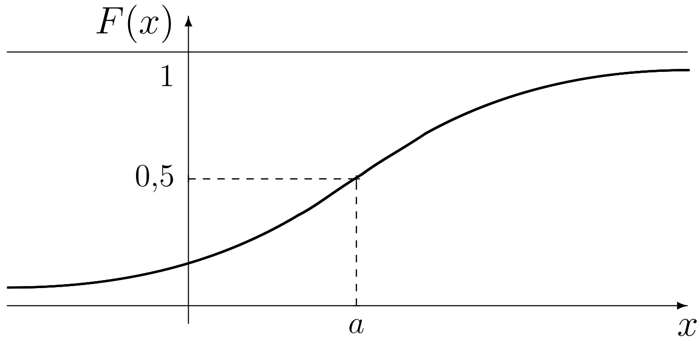

Рис. 15. Функция распределения нормального распределения N_{a,σ^2}

С в о й с т в о 3. *Пусть случайная величина* ξ *имеет нормальное распределение* N_{a,σ^2}. *Тогда*

$$F_\xi(x) = \Phi\left(\frac{x-a}{\sigma}\right), \qquad \frac{\xi - a}{\sigma} \sim \mathrm{N}_{0,1}.$$

Д о к а з а т е л ь с т в о. Докажем первое равенство. Функцию распределения $F_\xi(x) = \mathsf{P}(\xi < x)$ запишем в виде интеграла от плотности и сделаем в этом интеграле замену переменных $\dfrac{t-a}{\sigma} = y$, $dt = \sigma\, dy$:

$$F_\xi(x) = \int\limits_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(t-a)^2}{2\sigma^2}}\, dt \;=\; \int\limits_{-\infty}^{\frac{x-a}{\sigma}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{y^2}{2}}\, dy \;=\; \Phi\left(\frac{x-a}{\sigma}\right).$$

Читатель обратил внимание, что после такой замены область интегрирования $t \in (-\infty, x)$ заменилась на область интегрирования $y \in \left(-\infty, \dfrac{x-a}{\sigma}\right)$.

Докажем второе свойство. Вычислим, используя первое равенство, функцию распределения величины $\eta = \dfrac{\xi - a}{\sigma}$:

$$\begin{aligned} F_\eta(x) \;&=\; \mathsf{P}(\eta < x) = \mathsf{P}\left(\frac{\xi - a}{\sigma} < x\right) = \mathsf{P}(\xi < \sigma x + a) \;= \\ &=\; F_\xi(\sigma x + a) = \Phi\left(\frac{\sigma x + a - a}{\sigma}\right) = \Phi(x). \end{aligned}$$

Функция распределения случайной величины η равна $\Phi(x)$, поэтому η имеет стандартное нормальное распределение. □

Из свойства 3 и свойства (F4) функций распределения следует полезное равенство для вычисления вероятностей попадания в интервалы.

С в о й с т в о 4. *Если* $\xi \sim \mathrm{N}_{a,\sigma^2}$, *то*

$$\mathsf{P}(x_1 < \xi < x_2) \;=\; \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right).$$

Мы показали, что вычисление любых вероятностей для нормально распределённой случайной величины сводится к вычислению функции распределения $\Phi(x)$. Она обладает следующими свойствами.

Свойство 5. $\Phi(0) = 0{,}5, \quad \Phi(-x) = 1 - \Phi(x).$

Доказательство. Величина $\Phi(0)$ равна площади, заключённой под графиком кривой Гаусса левее оси ординат. Из-за симметрии плотности относительно прямой $x = 0$ эта площадь равна половине всей площади под графиком плотности, т. е. 0,5.

Второе свойство полезно нарисовать и проверить на графике *плотности* стандартного нормального распределения. □

Следующее свойство стандартного нормального распределения показано на рис. 16. Площадь центральной области равна единице за вычетом площадей двух симметричных «хвостов».

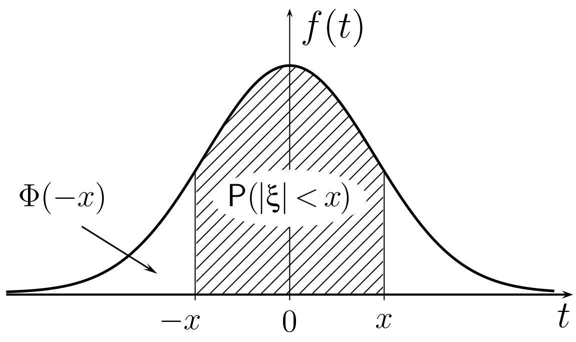

Рис. 16.

Свойство 6. *Если* $\xi \sim \mathrm{N}_{0,1}$, *то для любого* $x > 0$

$$\mathsf{P}(|\xi| < x) = 1 - 2\Phi(-x).$$

Доказательство. При $x > 0$ имеем:

$$\mathsf{P}(|\xi| < x) = \mathsf{P}(-x < \xi < x) = \Phi(x) - \Phi(-x) = 1 - 2\Phi(-x). \quad \square$$

Свойство 7 (правило трёх сигм). *Если* $\xi \sim \mathrm{N}_{a,\sigma^2}$, *то*

$$\mathsf{P}(|\xi - a| < 3\sigma) = 0{,}9973 \quad (\textit{т. е. почти единица}).$$

Доказательство. $\mathsf{P}\big(|\xi - a| < 3\sigma\big) = \mathsf{P}\left(\left|\frac{\xi - a}{\sigma}\right| < 3\right)$. Случайная величина $\eta = \frac{\xi - a}{\sigma}$ имеет распределение $\mathrm{N}_{0,1}$. По свойству 6,

$$\mathsf{P}(|\eta| < 3) = (1 - 2\Phi(-3)) = 1 - 2 \cdot 0{,}00135 = 0{,}9973. \quad \square$$

Большого смысла в запоминании числа 0,9973 нет, но полезно помнить, что почти вся масса (более 99%) нормального распределения сосредоточена в границах от $a - 3\sigma$ до $a + 3\sigma$.

П р и м е р 46. Станок-автомат изготовляет детали, длина которых по стандарту может отклоняться от 125 мм не более, чем на 1 мм. Среди продукции станка 7% нестандартных деталей. Считая, что длины деталей имеют нормальное распределение $\mathrm{N}_{125,\sigma^2}$, найти параметр σ^2.

Р е ш е н и е. Обозначим через ξ длину типичной детали. Деталь нестандартна, если $|\xi-125|>1$. Это событие происходит с вероятностью 0,07 — именно так следует понимать фразу про семь процентов нестандартных деталей. Раскроем модуль и используем свойство 4, а затем 5:

$$\begin{aligned}0{,}07 &= \mathsf{P}(|\xi-125|>1) = 1-\mathsf{P}(|\xi-125|\leqslant 1) = \\ &= 1-\mathsf{P}(124\leqslant\xi\leqslant 126) = 1-\left(\Phi\left(\frac{126-125}{\sigma}\right)-\Phi\left(\frac{124-125}{\sigma}\right)\right) = \\ &= 1-\left(\Phi\left(\sigma^{-1}\right)-\Phi\left(-\sigma^{-1}\right)\right) = 2\Phi\left(-\sigma^{-1}\right).\end{aligned}$$

По таблице на стр. 121, зная $\Phi(-\sigma^{-1})=0{,}035$, найдём $\sigma^{-1}=1{,}81$. Отсюда $\sigma^2=0,305$ мм^2.

Гамма-распределение $\Gamma_{\alpha,\lambda}$. Случайная величина ξ имеет гамма-распределение с параметрами $\alpha>0$, $\lambda>0$, если ξ имеет следующую плотность распределения:

$$f(x)=\begin{cases}0, & \text{если } x\leqslant 0,\\ \dfrac{\alpha^\lambda}{\Gamma(\lambda)}\cdot x^{\lambda-1}e^{-\alpha x}, & \text{если } x>0.\end{cases}$$

В выражении для плотности участвует *гамма-функция* Эйлера:

$$\Gamma(\lambda)=\int\limits_0^\infty x^{\lambda-1}e^{-x}\,dx.$$

Следующие свойства гамма-функции бывают полезны для вычисления определённых интегралов, содержащих произведение степенной и показательной функций: $\Gamma(\lambda)=(\lambda-1)\Gamma(\lambda-1)$, $\Gamma(n)=(n-1)!$ при целых положительных n, $\Gamma(1)=1$, $\Gamma(0{,}5)=\sqrt{\pi}$.

Отметим, что показательное распределение является частным случаем гамма-распределения: $\mathrm{E}_\alpha=\Gamma_{\alpha,1}$. Гамма-распределение с параметрами $\alpha=\frac{1}{2}$ и $\lambda=\frac{n}{2}$ при целом положительном n называется *хи-квадрат* (χ^2) *распределением с n степенями свободы.* С этим распределением мы познакомимся в курсе математической статистики. Гамма-распределение с целым положительным $\lambda=n$ называется *распределением Эрланга.*

У п р а ж н е н и е. Нарисовать график плотности распределения $\Gamma_{\alpha,\lambda}$ при $\lambda<1$, при $\lambda=1$ и при $\lambda>1$, отметить на этом графике точки экстремума, точки перегиба и иные особенности графика.

Распределение Коши $\mathrm{C}_{a,\sigma}$. Случайная величина ξ имеет распределение Коши с параметрами $a \in (-\infty, \infty)$, $\sigma > 0$, если ξ имеет плотность распределения

$$f(x) = \frac{1}{\pi}\,\frac{\sigma}{\sigma^2 + (x-a)^2}.$$

Плотность распределения Коши симметрична относительно прямой $x = a$ и похожа на плотность нормального распределения, но имеет более толстые «хвосты» на $\pm\infty$. Функция распределения случайной величины $\xi \sim \mathrm{C}_{a,\sigma}$ равна

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}\left(\frac{x-a}{\sigma}\right).$$

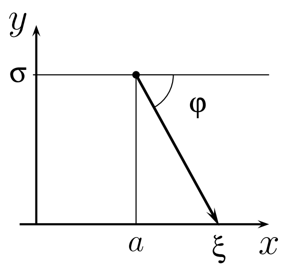

Рис. 17.

Распределение Коши имеет простой физический смысл (см. рис. 17). Если из точки плоскости с координатами (a, σ) испустить луч в направлении оси OX под случайным углом φ, выбранным наудачу из интервала $(0, \pi)$, то абсцисса точки пересечения луча с осью OX будет иметь распределение Коши $\mathrm{C}_{a,\sigma}$.

Действительно, $\xi = \sigma \operatorname{ctg}\varphi + a = \sigma \operatorname{tg}\left(\frac{\pi}{2} - \varphi\right) + a$,

$$\mathsf{P}(\xi < x) = \mathsf{P}\left(\frac{\pi}{2} - \varphi < \operatorname{arctg}\left(\frac{x-a}{\sigma}\right)\right) = \mathsf{P}\left(\varphi > \frac{\pi}{2} - \operatorname{arctg}\left(\frac{x-a}{\sigma}\right)\right).$$

Поскольку φ выбирается наудачу на интервале $(0, \pi)$, вероятность есть отношение длин соответствующих интервалов:

$$\mathsf{P}\left(\varphi > \frac{\pi}{2} - \operatorname{arctg}\left(\frac{x-a}{\sigma}\right)\right) = \frac{\pi - \left(\frac{\pi}{2} - \operatorname{arctg}\left(\frac{x-a}{\sigma}\right)\right)}{\pi} = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}\left(\frac{x-a}{\sigma}\right).$$

Распределение Парето. Случайная величина ξ имеет распределение Парето с параметром $\alpha > 0$, если ξ имеет следующие плотность и функцию распределения:

$$f(x) = \begin{cases} \dfrac{\alpha}{x^{\alpha+1}}, & \text{если } x \geqslant 1, \\ 0, & \text{если } x < 1; \end{cases} \qquad F(x) = \begin{cases} 1 - \dfrac{1}{x^{\alpha}}, & \text{если } x \geqslant 1, \\ 0, & \text{если } x < 1. \end{cases}$$

С другими абсолютно непрерывными распределениями (Стью́дента, Фишера, Колмогорова, Лапласа) мы познакомимся при изучении математической статистики. В различных задачах читателю, возможно, встретятся распределения Вейбулла, логарифмически нормальное и некоторые другие интересные абсолютно непрерывные распределения. Плотности всех этих распределений можно найти в таблице на стр. 120.

§ 6. Преобразования случайных величин

Если ξ — случайная величина, а $g(x)$ — действительная функция действительного аргумента, то $g(\xi)$ — новая случайная величина, и нужно уметь находить её распределение по распределению ξ. Эта проблема возникает, например, при моделировании случайных величин с заданным распределением. Датчик случайных чисел может генерировать лишь значения случайных величин с равномерным распределением. А если нам необходимы значения показательно распределённой величины, нужно знать, какое преобразование применить, чтобы из равномерного распределения получить показательное.

Преобразования дискретных распределений. Если ξ имеет дискретное распределение, то величина $g(\xi)$ также имеет дискретное распределение, и таблица её распределения находится просто по определению, как показано в следующем примере.

П р и м е р 47. Случайная величина ξ имеет таблицу распределения

ξ	-2	-1	0	1	2
P	0,1	0,3	0,1	0,2	0,3

Найти таблицу распределения случайной величины $\eta = \xi^2$.

Р е ш е н и е. Возводя значения случайной величины ξ в квадрат, увидим, что η принимает только значения 0, 1, 4. При этом $\eta = 0$ только когда $\xi = 0$, т. е. с вероятностью 0,1, а каждое из значений 1 и 4 величина η принимает в двух случаях:

$$\{\eta = 1\} = \{\xi = -1\} \cup \{\xi = 1\}, \quad \{\eta = 4\} = \{\xi = -2\} \cup \{\xi = 2\}.$$

По аксиоме вероятности (P2), вероятность объединения несовместных событий равна сумме их вероятностей:

$$\begin{aligned} \mathsf{P}(\eta = 1) &= \mathsf{P}(\xi = -1) + \mathsf{P}(\xi = 1) = 0{,}3 + 0{,}2 = 0{,}5, \\ \mathsf{P}(\eta = 4) &= \mathsf{P}(\xi = -2) + \mathsf{P}(\xi = 2) = 0{,}1 + 0{,}3 = 0{,}4. \end{aligned}$$

Получаем следующую таблицу распределения:

η	0	1	4
P	0,1	0,5	0,4

Преобразования абсолютно непрерывных распределений. Пусть случайная величина ξ имеет функцию распределения $F(x)$ и плотность распределения $f(x)$. Построим с помощью функции $g(x)$ случайную величину $\eta = g(\xi)$. Требуется найти функцию распределения и, если существует, плотность распределения величины η.

Плотность распределения случайной величины $\eta = g(\xi)$ существует далеко не при любых функциях g. Так, если функция g кусочно-постоянна, то η имеет дискретное распределение, и плотность её распределения не существует. Например, если функция $g(x) = 1$ при всех x, то величина $\eta = g(\xi) = 1$ имеет вырожденное распределение $\mathbf{I}_1$.

Теорема 22. *Пусть случайная величина ξ имеет плотность распределения $f(x)$, и постоянная a отлична от нуля. Тогда $\eta = a\xi + b$ имеет плотность распределения*

$$f_\eta(x) = \frac{1}{|a|}\, f\left(\frac{x-b}{a}\right).$$

Доказательство. Пусть сначала $a > 0$.

$$F_\eta(x) = \mathsf{P}(a\xi + b < x) = \mathsf{P}\left(\xi < \frac{x-b}{a}\right) = F\left(\frac{x-b}{a}\right).$$

Продифференцируем это выражение по x как сложную функцию:

$$f_\eta(x) = (F_\eta(x))' = \frac{d}{dx} F\left(\frac{x-b}{a}\right) = \left(\frac{x-b}{a}\right)' \cdot f\left(\frac{x-b}{a}\right) = \frac{1}{a} f\left(\frac{x-b}{a}\right).$$

Пусть теперь $a < 0$.

$$F_\eta(x) = \mathsf{P}(a\xi + b < x) = \mathsf{P}\left(\xi > \frac{x-b}{a}\right) = 1 - F\left(\frac{x-b}{a}\right).$$

Продифференцируем это выражение по x :

$$f_\eta(x) = (F_\eta(x))' = \frac{d}{dx}\left(1 - F\left(\frac{x-b}{a}\right)\right) = -\frac{1}{a} f\left(\frac{x-b}{a}\right) = \frac{1}{|a|} f\left(\frac{x-b}{a}\right).$$

Убедиться в том, что производная $f_\eta(x)$ действительно является плотностью распределения, можно, например, проверив для неё свойство (f2) плотности.

Из теоремы 22 следуют (*проверить!*) уже знакомые нам утверждения:

Следствие 2. *Если $\xi \sim \mathrm{N}_{0,1}$, то $\eta = \sigma\xi + a \sim \mathrm{N}_{a,\sigma^2}$.*

Доказательство. Действительно,

$$f_\eta(x) = \frac{1}{\sigma} \cdot f_\xi\left(\frac{x-a}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-a)^2}{2\sigma^2}}. \qquad \square$$

Следствие 3. *Если $\xi \sim \mathrm{N}_{a,\sigma^2}$, то её линейное преобразование $c\xi + d$ при $c \neq 0$ снова имеет нормальное распределение.*

Этот факт следует запомнить: линейное преобразование оставляет распределение нормальным. Полезно также убедиться, что $c\xi + d \sim \mathrm{N}_{ca+d,\, c^2\sigma^2}$.

Следствие 4. *Если $\xi \sim \mathrm{U}_{a,b}$, то её линейное преобразование $c\xi + d$ при $c \neq 0$ снова имеет равномерное распределение.*

Границы интервала, на котором сосредоточено распределение $c\xi + d$, получаются линейным преобразованием границ исходного интервала (a, b). Например, $c\xi + d \sim \mathrm{U}_{ca+d,\, cb+d}$ при $c > 0$.

Следствие 5. *Если* $\xi \sim \mathrm{E}_\alpha$, *то* $\alpha\xi \sim \mathrm{E}_1$.

Пример 48. Случайная величина ξ имеет стандартное нормальное распределение. Найти плотность распределения величины ξ^2.

Решение. Для $x \leqslant 0$ функция распределения величины $\eta = \xi^2$ равна

$$F_\eta(x) = \mathsf{P}(\xi^2 < x) = 0, \text{ поэтому } f_\eta(x) = 0.$$

Для $x > 0$ имеем: $F_\eta(x) = \mathsf{P}(\xi^2 < x) = \mathsf{P}\left(|\xi| < \sqrt{x}\right) = 1 - 2\Phi\left(-\sqrt{x}\right)$. Производная функции $\Phi(x)$ равна плотности стандартного нормального распределения $\varphi(x)$, поэтому при $x > 0$

$$f_\eta(x) = (F_\eta(x))' = -2\varphi\left(-\sqrt{x}\right) \cdot \left(-\sqrt{x}\right)' = \frac{1}{\sqrt{x}}\,\varphi\left(-\sqrt{x}\right) = \frac{1}{\sqrt{2\pi x}}\, e^{-x/2}.$$

Мы получили следующую плотность:

$$f_\eta(x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ \dfrac{1}{\sqrt{2\pi x}}\, e^{-x/2}, & \text{если } x > 0. \end{cases}$$

Это плотность гамма-распределения с параметрами $\alpha = \dfrac{1}{2}$, $\lambda = \dfrac{1}{2}$. Тем самым мы доказали, что $\eta \sim \Gamma_{1/2,\,1/2}$. Другое название этого распределения — распределение хи-квадрат с одной степенью свободы.

Пример 49. Случайная величина ξ имеет равномерное распределение на отрезке $[0, 1]$. Найти распределение величины $-\ln\xi$.

Решение. Функция распределения величины $\eta = -\ln\xi$ равна

$$F_\eta(x) = \mathsf{P}(-\ln\xi < x) = \mathsf{P}(\xi > e^{-x}) = 1 - \mathsf{P}(\xi \leqslant e^{-x}) = 1 - F_\xi(e^{-x}).$$

Подставим $e^{-x} > 0$ вместо x в функцию распределения $\mathrm{U}_{0,\,1}$:

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

Окончательно получаем:

$$F_\eta(x) = 1 - F_\xi(e^{-x}) = \begin{cases} 1 - e^{-x}, & \text{если } x \geqslant 0, \\ 0, & \text{если } x < 0, \end{cases}$$

Итак, $\eta \sim \mathrm{E}_1$. Мы получили замечательный факт: применение функции $g(x) = -\ln x$ к равномерно распределённой на отрезке $[0, 1]$ случайной величине приводит к показательному распределению с параметром 1. Если же мы хотим получить показательное распределение с произвольным параметром α, следует воспользоваться следствием 5: $\zeta = -\dfrac{1}{\alpha}\ln\xi \sim \mathrm{E}_\alpha$.

§ 7. Контрольные вопросы

1. Для чего нужны случайные величины?

2. Что такое случайная величина?

3. Привести примеры случайных величин.

4. Что называют таблицей распределения вероятностей? У каких случайных величин есть таблица распределения вероятностей?

5. Сколько значений может принимать случайная величина с дискретным распределением?

6. Как вычислить $\mathsf{P}(\xi \in [2, 4])$ для случайной величины с дискретным распределением?

7. Могут ли две разные случайные величины иметь одинаковые таблицы распределения?

8. Совпадают ли количества очков при первом и при втором броске игральной кости? Одинаковы ли распределения этих случайных величин?

9. Совпадают ли результаты первого и второго бросаний одной и той же монеты? Одинаковы ли распределения соответствующих случайных величин?

10. Сколько значений может принимать случайная величина с абсолютно непрерывным распределением?

11. Что такое плотность распределения?

12. Как вычислить $\mathsf{P}(\xi \in [2, 4])$ для случайной величины с абсолютно непрерывным распределением?

13. Как вычислить $\mathsf{P}(\xi < 3)$ для случайной величины с абсолютно непрерывным распределением?

14. Как вычислить $\mathsf{P}(\xi > 3)$ для случайной величины с абсолютно непрерывным распределением?

15. Чему равен интеграл от плотности распределения по всей прямой?

16. Может ли плотность распределения принимать отрицательные значения?

17. Может ли плотность распределения равняться нулю при всех значениях аргумента? Единице? Двойке? Чему равно значение интеграла по всей прямой от каждой из этих функций?

18. Что такое функция распределения случайной величины?

19. Перечислите и объясните свойства функции распределения.

20. Для каждого свойства функций распределения нарисуйте график любой функции, не обладающей этим свойством.

21. Как выглядит функция распределения дискретного распределения? Чему равны величины её скачков?

22. Как по таблице дискретного распределения нарисовать график функции распределения?

23. Как по графику функции распределения дискретного закона восстановить таблицу распределения?

24. Как по функции распределения вычислить вероятность $\mathsf{P}(2 \leqslant \xi < 3)$?

25. Как по функции распределения вычислить вероятность $\mathsf{P}(\xi \geqslant 3)$?

26. Может ли функция распределения абсолютно непрерывного распределения иметь разрывы?

27. Чему для любого x равна $\mathsf{P}(\xi = x)$, если ξ имеет абсолютно непрерывное распределение?

28. Как плотность распределения находится по функции распределения?

29. Перечислите основные дискретные распределения. Запишите таблицу распределения каждого.

30. Перечислите основные абсолютно непрерывные распределения. Запишите плотность и функцию распределения каждого. Постройте графики всех плотностей и функций распределения.

31. Как связаны функция распределения нормального закона и функция $\Phi(x)$?

32. Как вычислять вероятность $\mathsf{P}(x_1 < \xi < x_2)$, если ξ имеет нормальное распределение с параметрами a и σ^2?

33. Как из нормально распределённой случайной величины сделать величину со стандартным нормальным распределением?

34. Гауссовская кривая — это график плотности или функции распределения? Какому распределению этот график отвечает?

35. Чему равна $\mathsf{P}(\xi < 0)$ для $\xi \sim \mathrm{N}_{0,1}$?

36. Чему равна $\mathsf{P}(\xi < a)$ для $\xi \sim \mathrm{N}_{a,\sigma^2}$?

37. Найти $\mathsf{P}(\xi < -3)$, $\mathsf{P}(\xi < -2)$, $\mathsf{P}(\xi < -1{,}6)$, $\mathsf{P}(\xi < 1{,}6)$, $\mathsf{P}(\xi < 2)$ и $\mathsf{P}(\xi < 3)$ для $\xi \sim \mathrm{N}_{0,1}$.

38. Как связаны плотности распределения величин ξ и $a\xi + b$?

39. Как по плотности распределения величины ξ найти плотность распределения величины $-\xi$? 2ξ? $\xi + 2$?

40. Если случайная величина ξ имеет нормальное распределение, каким будет распределение случайной величины $-\xi$?

41. Если случайная величина ξ имеет нормальное распределение, каким будет распределение случайной величины $5\xi + 7$?

42. Каким преобразованием можно случайную величину $\xi \sim \mathrm{U}_{0,5}$ превратить в $\eta \sim \mathrm{U}_{0,1}$?

43. Каким преобразованием можно случайную величину $\xi \sim \mathrm{U}_{0,1}$ превратить в $\eta \sim \mathrm{U}_{0,5}$? А в $\eta \sim \mathrm{E}_1$?

44. Каким преобразованием можно случайную величину $\xi \sim \mathrm{E}_5$ превратить в $\eta \sim \mathrm{E}_1$?

45. Каким преобразованием можно случайную величину $\xi \sim \mathrm{E}_1$ превратить в $\eta \sim \mathrm{E}_5$?

46. Каким преобразованием можно случайную величину $\xi \sim \mathrm{N}_{5,9}$ превратить в $\eta \sim \mathrm{N}_{0,1}$?

47. Каким преобразованием можно случайную величину $\xi \sim \mathrm{N}_{0,1}$ превратить в $\eta \sim \mathrm{N}_{5,9}$? А в $\eta \sim \mathrm{N}_{-5,9}$?

48. Если случайная величина ξ имеет стандартное нормальное распределение, каким будет распределение случайной величины ξ^2?

ГЛАВА VI

МНОГОМЕРНЫЕ РАСПРЕДЕЛЕНИЯ

В практических задачах приходится иметь дело не с одной, но с несколькими взаимосвязанными случайными величинами. Примером может служить набор экономических или производственных показателей, характеризующих какой-то процесс. Набор $\xi_1, \ldots, \xi_n$ из нескольких случайных величин называют случайным вектором или многомерной случайной величиной. Нет никаких принципиальных отличий между тем, как описать распределение одномерной и многомерной случайных величин.

§ 1. Совместное распределение

Пусть случайные величины $\xi_1, \ldots, \xi_n$ заданы на одном вероятностном пространстве $\langle\Omega, \mathcal{F}, \mathsf{P}\rangle$.

Определение 16. Функция

$$F(x_1, \ldots, x_n) = \mathsf{P}(\xi_1 < x_1, \ldots, \xi_n < x_n)$$

называется функцией распределения вектора $(\xi_1, \ldots, \xi_n)$ или функцией *совместного* распределения случайных величин $\xi_1, \ldots, \xi_n$.

Перечислим свойства функции совместного распределения, ограничившись, для простоты, вектором (ξ, η) из двух величин. Функция распределения $F(x, y) = \mathsf{P}(\xi < x, \eta < y)$ не убывает по x и по y; стремится к нулю, если любую из переменных устремить к $-\infty$; стремится к единице, если *обе* переменные устремить к $+\infty$; непрерывна слева по каждой переменной.

Следующее свойство позволяет по функции совместного распределения пары (ξ, η) находить функции распределения величин ξ и η в отдельности.

(F4) Чтобы по функции совместного распределения восстановить функции распределения ξ и η в отдельности, следует устремить «лишнюю» переменную к $+\infty$:

$$\lim_{x\to+\infty} F(x, y) = F_\eta(y), \qquad \lim_{y\to+\infty} F(x, y) = F_\xi(x).$$

Действительно, с ростом x событие $\{\xi < x\}$ становится всё более достоверным, поэтому $\mathsf{P}(\xi < x, \eta < y) \to \mathsf{P}(\eta < y)$.

§ 2. Виды многомерных распределений

Рассмотрим два типичных случая: когда *совместное* распределение координат случайного вектора либо дискретно, либо абсолютно непрерывно.

Дискретное совместное распределение. Совместное расределение величин ξ и η *дискретно,* если каждая из них имеет дискретное распределение. Если ξ принимает значения $a_1, a_2, \ldots$, а η принимает значения $b_1, b_2, \ldots$, то пара (ξ, η) принимает всевозможные значения (a_i, b_j). Таблицу, на пересечении i-й строки и j-го столбца которой стоит вероятность $p_{i,j} = \mathsf{P}(\xi = a_i,\ \eta = b_j)$, называют *таблицей совместного распределения* случайных величин ξ и η. Вероятности $p_{i,j}$ в сумме дают единицу:

$$\sum_i \sum_j p_{i,j} = 1.$$

Таблицы распределения каждой из случайных величин ξ, η в отдельности (таблицы *частных,* или *маргинальных* распределений) восстанавливаются по таблице совместного распределения суммированием по строке или столбцу:

$$\mathsf{P}(\xi = a_i) = \sum_j p_{i,j}\,, \qquad \mathsf{P}(\eta = b_j) = \sum_i p_{i,j}\,. \tag{13}$$

Так, первое равенство в 13 следует из того, что набор $\{\eta = b_1\}$, $\{\eta = b_2\}$, $\ldots$ является полной группой событий, поэтому событие $\{\xi = a_i\}$ раскладывается в объединение попарно несовместных событий:

$$\{\xi = a_i\} = \bigcup_j \{\xi = a_i,\ \eta = b_j\}, \quad \mathsf{P}(\xi = a_i) = \sum_j \mathsf{P}(\xi = a_i,\ \eta = b_j) = \sum_j p_{i,j}\,.$$

Второе равенство в 13 получится, если разложить по полной группе событий событие $\{\eta = b_j\}$.

П р и м е р 50. Правильную монету подбрасывают трижды. Случайная величина ξ равна числу гербов, выпавших в первых двух испытаниях, случайная величина η равна числу гербов, выпавших в трёх испытаниях. Построить таблицу совместного распределения ξ и η.

Р е ш е н и е. Возможны значения $\xi = 0, 1, 2$ и $\eta = 0, 1, 2, 3$. Но не все пары значений имеют положительную вероятность.

$$\mathsf{P}(\xi = 0,\ \eta = 0) = \mathsf{P}(ppp) = \frac{1}{8}, \quad \mathsf{P}(\xi = 1,\ \eta = 1) = \mathsf{P}(pгp) + \mathsf{P}(гpp) = \frac{1}{4},$$

$$\mathsf{P}(\xi = 0,\ \eta = 1) = \mathsf{P}(ppг) = \frac{1}{8}, \quad \mathsf{P}(\xi = 1,\ \eta = 2) = \mathsf{P}(pгг) + \mathsf{P}(гpг) = \frac{1}{4},$$

$$\mathsf{P}(\xi = 2,\ \eta = 2) = \mathsf{P}(ггp) = \frac{1}{8}, \quad \mathsf{P}(\xi = 2,\ \eta = 3) = \mathsf{P}(ггг) = \frac{1}{8}.$$

Остальные значения пар (ξ, η) невозможны. Составим таблицу:

	$\eta = 0$	$\eta = 1$	$\eta = 2$	$\eta = 3$	$\mathsf{P}(\xi = a_i)$
$\xi = 0$	$1/8$	$1/8$	0	0	$1/4$
$\xi = 1$	0	$1/4$	$1/4$	0	$1/2$
$\xi = 2$	0	0	$1/8$	$1/8$	$1/4$
$\mathsf{P}(\eta = b_j)$	$1/8$	$3/8$	$3/8$	$1/8$	$\boldsymbol{\Sigma = 1}$

Нижняя строка состоит из сумм вероятностей по каждому столбцу — это таблица распределения η. Аналогично правый столбец состоит из сумм по строкам — это таблица распределения ξ. Для проверки правильности вычислений заметим, что по этим таблицам $\xi \sim \mathrm{B}_{2,\,0{,}5}$, $\eta \sim \mathrm{B}_{3,\,0{,}5}$, как и должно быть.

Абсолютно непрерывное совместное распределение. Дадим определение.

Определение 17. Говорят, что случайные величины ξ и η имеют *абсолютно непрерывное совместное распределение*, если существует неотрицательная функция $f(x, y)$ такая, что для любых $a < b$, $c < d$ имеет место равенство

$$\mathsf{P}(a < \xi < b,\ c < \eta < d) = \int\limits_a^b dx \int\limits_c^d f(x, y)\, dy.$$

Если такая функция $f(x, y)$ существует, она называется плотностью совместного распределения случайных величин ξ и η.

Если случайные величины ξ, η имеют абсолютно непрерывное совместное распределение, то для любых x, y имеет место равенство:

$$F(x, y) = \mathsf{P}(\xi < x, \eta < y) = \int\limits_{-\infty}^{x} du \int\limits_{-\infty}^{y} f(u, v)\, dv. \tag{14}$$

Плотность совместного распределения имеет те же свойства, что и плотность распределения одной случайной величины: неотрицательность и нормированность:

$$f(x, y) \geqslant 0, \qquad \int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{\infty} f(x, y)\, dy \;=\; 1.$$

По функции совместного распределения его плотность находится как смешанная частная производная (в точках, где она существует):

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y).$$

Т е о р е м а 23. *Если случайные величины* ξ *и* η *имеют абсолютно непрерывное совместное распределение с плотностью* $f(x, y)$, *то* ξ *и* η *в отдельности тоже имеют абсолютно непрерывные распределения с плотностями:*

$$f_\xi(x) = \int\limits_{-\infty}^{\infty} f(x, y)\,dy\,; \qquad f_\eta(y) = \int\limits_{-\infty}^{\infty} f(x, y)\,dx.$$

Итак, чтобы найти плотности координат вектора, нужно проинтегрировать плотность совместного распределения по всем «лишним» координатам. Это свойство сразу следует из свойства (F4) функции совместного распределения, достаточно в равенстве (14) устремить к бесконечности переменную x или переменную y.

П р и м е р 51. Привести пример случайных величин ξ и η с абсолютно непрерывными распределениями таких, что пара (ξ, η) не имеет плотности распределения.

Р е ш е н и е. Пусть $\xi \sim \mathrm{U}_{0,1}$ — координата точки, брошенной наудачу на отрезок $[0, 1]$, и пусть $\eta = \xi$. Тогда пара (ξ, η) принимает значения только на диагонали единичного квадрата, поэтому плотности распределения у этой пары не существует. Чтобы это показать, найдём функцию совместного распределения:

$$F(x, y) = \mathsf{P}(\xi < x, \eta < y) = \mathsf{P}(\xi < x, \xi < y) = \begin{cases} F_\xi(x), & \text{если } x \leqslant y, \\ F_\xi(y), & \text{если } y < x. \end{cases}$$

Вторая смешанная производная этой функции по x и по y существует везде внутри квадрата, кроме диагонали, и равна нулю. Но функция $f(x, y) = 0$ не является плотностью распределения, поскольку объём области под графиком этой функции не равен единице (*а чему он равен?*).

П р и м е р 52. Возможно ли, зная распределения случайных величин ξ и η в отдельности, найти их совместное распределение?

Р е ш е н и е. Ответ на этот вопрос, разумеется, отрицательный.

Пусть, как в предыдущем примере, $\xi \sim \mathrm{U}_{0,1}$ — координата точки, брошенной наудачу на отрезок $[0, 1]$, и пусть $\eta = 1 - \xi$. Заметим, что $\eta \sim \mathrm{U}_{0,1}$.

Найдём функцию совместного распределения величин ξ и η:

$$F(x, y) = \mathsf{P}(\xi < x, 1 - \xi < y) = \mathsf{P}(1 - y < \xi < x) = F_\xi(x) - F_\xi(1 - y),$$

если $1 - y < x$, иначе $F(x,y) = 0$.

В этом и в предыдущем примерах $\xi, \eta \sim \mathrm{U}_{0,1}$ имеют одни и те же частные распределения, но совместные распределения ξ и η в этих примерах оказались различными.

§ 3. Примеры многомерных распределений

Приведём два наиболее употребительных примера абсолютно непрерывных многомерных распределений. Первое из них — распределение координаты точки, брошенной наудачу в область S на плоскости. С этим распределением мы давно знакомы по геометрическому определению вероятности.

Равномерное распределение. Пусть $S \subset \mathbb{R}^2$ — множество с конечной площадью $\mu(S)$. Пара (ξ, η) имеет равномерное распределение в области S, если плотность совместного распределения $f(x, y)$ постоянна в области S и равна нулю вне этой области:

$$f(x, y) = \begin{cases} \dfrac{1}{\mu(S)}, & \text{если } (x, y) \in S, \\ 0, & \text{если } (x, y) \notin S. \end{cases} \tag{15}$$

Эта функция является плотностью распределения, поскольку объём, заключённый под графиком этой функции, есть объём цилиндра с основанием S и с высотой $1/\mu(S)$, и равен единице.

Пример 53. Точка с координатами (ξ, η) наудачу выбирается в треугольнике с вершинами $(0, 0)$, $(2, 0)$, $(0, 1)$. Найти плотность совместного распределения и частные плотности распределения координат точки.

Решение. Совместное распределение по условию является равномерным распределением в данном треугольнике S с единичной площадью. Плотность этого распределения равна

$$f(x, y) = \begin{cases} 1, & \text{если } (x, y) \in S, \\ 0, & \text{если } (x, y) \notin S. \end{cases}$$

Плотности распределения величин ξ и η вычислим по теореме 23. Направление и границы интегрирования показаны на рис. 18.

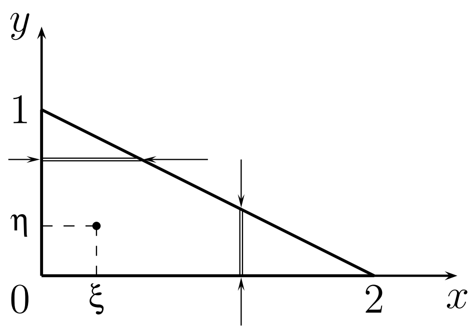

Рис. 18.

Интегрирование $f(x, y)$ по y при фиксированном $x \in (0, 2)$ показано на рис. 18 вертикальным отрезком, интегрирование по x при фиксированном

$y \in (0,\,1)$ — горизонтальным:

$$f_\xi(x) = \int\limits_0^{1-\frac{x}{2}} 1\;dy = 1 - \frac{x}{2}\,, \qquad f_\eta(y) = \int\limits_0^{2-2y} 1\;dx = 2 - 2y.$$

При остальных значениях x и y обе плотности равны нулю. Распределение каждой координаты в отдельности уже не является равномерным распределением. Это и понятно: для абсциссы наудачу выбранной в данном треугольнике точки более вероятны значения вблизи нуля, чем вблизи двойки; и для ординаты значения около нуля вероятнее значений вблизи единицы.

Плотности распределений компонент можно найти и с помощью геометрического определения вероятности: вероятность величине ξ принимать значения из интервала $(x,\,x+dx)$ равна отношению площади узкой вертикальной полосы с основанием $(x,\,x+dx)$ к площади всего треугольника (в данном случае — к единице): $\mathsf{P}(\xi \in (x,\,x+dx)) = \left(1 - \frac{x}{2}\right)dx$. С другой стороны, эта вероятность пропорциональна плотности: $\mathsf{P}(\xi \in (x,\,x+dx)) = f_\xi(x)\,dx$. Поэтому плотность распределения случайной величины ξ равна высоте этой полосы $1 - \frac{x}{2}$. Точно так же плотность распределения величины η равна длине горизонтальной полосы с основанием $(y,\,y+dy)$, т. е. $2-2y$.

Следующее многомерное распределение чрезвычайно важно и часто встречается в экономических приложениях.

Многомерное нормальное распределение. Пусть $\Sigma > 0$ — положительно определённая[2] симметричная матрица $(n \times n)$, матрица Σ^{-1} — обратная к Σ, и $\vec{a} \in \mathbb{R}^n$ — n-мерный вектор-столбец. Транспонированный вектор, т. е. вектор-строку, мы будем обозначать так: $\vec{a}^{\,T} = (a_1,\,\ldots,\,a_n)$.

Говорят, что вектор $(\xi_1,\,\ldots,\,\xi_n)$ имеет *многомерное нормальное* распределение $\mathrm{N}_{\vec{a},\,\Sigma}$ с вектором средних $\vec{a}$ и матрицей коварииаций Σ, если плотность совместного распределения координат этого вектора равна

$$f(x_1,\,\ldots,\,x_n) = \frac{1}{\sqrt{\det\Sigma}\,\left(\sqrt{2\pi}\right)^n}\;\exp\left\{-\frac{1}{2}\,(\vec{x}-\vec{a})^T\cdot\Sigma^{-1}\cdot(\vec{x}-\vec{a})\right\}.$$

Выражение $(\vec{x}-\vec{a})^T\,\Sigma^{-1}\,(\vec{x}-\vec{a})$ в показателе экспоненты является квадратичной формой от переменных $(x_i - a_i)$: для матрицы $B = \Sigma^{-1}$ с элементами b_{ij} получим

$$(\vec{x}-\vec{a})^T\,B\,(\vec{x}-\vec{a}) = \sum_{i=1}^n\sum_{j=1}^n b_{ij}(x_i-a_i)(x_j-a_j).$$

[2]Матрица положительно определена, если все её собственные значения положительны.

Плотность, например, двумерного нормального распределения имеет вид колоколообразной «шапочки», похожей на плотность на рис. 8, стр. 32. Линией уровня этой шапочки, получающейся в сечении плоскостью $f(x,\, y) = \mathrm{const}$, будет эллипс $(\vec{x}-\vec{a})^T B(\vec{x}-\vec{a}) = c$. Этот эллипс называется *эллипсом рассеивания* (в $\mathbb{R}^n$ — эллипсоидом).

По теореме 23 можно вычислить плотности координат вектора $(\xi_1,\, \ldots,\, \xi_n)$ с многомерным нормальным распределением. Величина ξ_i имеет нормальное распределение $\mathrm{N}_{a_i,\sigma_i^2}$, где $\sigma_i^2 = \Sigma_{ii}$ — i-й диагональный элемент матрицы Σ.

Подробно с многомерным нормальным распределением мы познакомимся в курсе математической статистики. А смысл заклинаний «с вектором средних $\vec{a}$ и матрицей ковариаций Σ» выяснится уже скоро.

В частном случае, когда матрица Σ диагональна с элементами $\sigma_1^2,\, \ldots,\, \sigma_n^2$ на главной диагонали, плотность совместного распределения превращается в произведение плотностей нормальных случайных величин:

$$\begin{aligned} f(x_1,\ldots,x_n) &= \frac{1}{\sigma_1\ldots\sigma_n\left(\sqrt{2\pi}\right)^n}\,\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}(x_i-a_i)^2\right\} = \\ &= \frac{1}{\sigma_1\sqrt{2\pi}}\,e^{-\frac{(x_1-a)^2}{2\sigma_1^2}}\cdot\ldots\cdot\frac{1}{\sigma_n\sqrt{2\pi}}\,e^{-\frac{(x_n-a)^2}{2\sigma_n^2}}. \end{aligned} \tag{16}$$

Скоро мы увидим, что это равенство означает *независимость* случайных величин $\xi_1,\, \ldots,\, \xi_n$.

§ 4. Роль совместного распределения

Если нам известно совместное распределение двух или нескольких случайных величин, становится возможным отыскать распределение суммы, разности, произведения, частного, иных функций от этих случайных величин.

Следующие три простых примера показывают, что знания только частных распределений двух случайных величин недостаточно для отыскания, например, распределения их суммы. Для этого необходимо знать их совместное распределение. Распределение суммы (и любой иной функции) *не определяется,* вообще говоря, распределениями слагаемых: при одних и тех же распределениях слагаемых распределение суммы может быть разным в зависимости от *совместного распределения* слагаемых.

П р и м е р 54. В примере 51 сумма $\xi+\eta=\xi+\xi=2\xi$ имеет равномерное распределение $\mathrm{U}_{0,\,2}$, а в примере 52 сумма $\xi+\eta=\xi+1-\xi=1$ имеет вырожденное распределение I_1, хотя распределение ξ и η в этих примерах одно и то же — равномерное на отрезке $[0,\, 1]$. Но совместные распределения различаются, и распределение суммы это почувствовало.

Пример 55. Пусть ξ имеет стандартное нормальное распределение.

Возьмём $\eta = -\xi$. Тогда η тоже имеет стандартное нормальное распределение, а сумма $\xi + \eta = 0$ имеет вырожденное распределение.

Возьмём теперь $\eta = \xi$. Тогда сумма $\xi + \eta = 2\xi$ имеет уже не вырожденное, а нормальное распределение $\mathrm{N}_{0,4}$ (*проверить!*).

Пример 56. Рассмотрим две случайные величины ξ и η с одним и тем же распределением Бернулли с параметром $p = 0{,}5$ и следующей таблицей совместного распределения: для $0 \leqslant r \leqslant 0{,}5$ положим

	$\eta = 0$	$\eta = 1$	$\mathsf{P}(\xi = a_i)$
$\xi = 0$	r	$0{,}5 - r$	$0{,}5$
$\xi = 1$	$0{,}5 - r$	r	$0{,}5$
$\mathsf{P}(\eta = b_j)$	$0{,}5$	$0{,}5$	$\mathbf{\Sigma = 1}$

Если взять $r = 0$, то $\mathsf{P}(\xi + \eta = 1) = \mathsf{P}(\xi = 0,\, \eta = 1) + \mathsf{P}(\xi = 1,\, \eta = 0) = 1$, т. е. распределение $\xi + \eta$ вырождено в точке 1.

Если $r = 0{,}5$, то $\mathsf{P}(\xi + \eta = 0) = \mathsf{P}(\xi + \eta = 2) = 0{,}5$, т. е. $\xi + \eta$ принимает значения 0 и 2 с равными вероятностями.

Если $r = 0{,}25$, то $\mathsf{P}(\xi + \eta = 0) = 0{,}25 = \mathsf{P}(\xi + \eta = 2)$, $\mathsf{P}(\xi + \eta = 1) = 0{,}5$, т. е. $\xi + \eta$ имеет биномиальное распределение с параметрами 2 и 0,5.

Ещё раз отметим, что частные распределения ξ и η от r не зависят. Распределение суммы меняется вместе с совместным распределением ξ и η при неизменных частных распределениях величин ξ и η.

Распределение функции от нескольких случайных величин определяется их частными распределениями, только если ими определяется совместное распределение. Для этого достаточно, например, потребовать *независимости* этих случайных величин — когда совместное распределение оказывается равным произведению распределений координат.

§ 5. Независимость случайных величин

Определение 18. Случайные величины $\xi_1, \ldots, \xi_n$ называют независимыми, если для любых множеств $B_1, \ldots, B_n$ имеет место равенство:

$$\mathsf{P}(\xi_1 \in B_1, \ldots, \xi_n \in B_n) = \mathsf{P}(\xi_1 \in B_1) \cdot \ldots \cdot \mathsf{P}(\xi_n \in B_n).$$

Можно сформулировать более удобное равносильное определение.

Определение 19. Случайные величины $\xi_1, \ldots, \xi_n$ независимы, если функция совместного распределения распадается в произведение частных функций распределения, т. е. для любых $x_1, \ldots, x_n$ имеет место равенство:

$$F(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n).$$

Для случайных величин с дискретными или с абсолютно непрерывными распределениями эквивалентные определения независимости выглядят так.

Определение 20. Случайные величины $\xi_1, \ldots, \xi_n$ с дискретным распределением независимы, если для любых чисел $a_1, \ldots, a_n$ имеет место равенство: $\mathsf{P}(\xi_1 = a_1, \ldots, \xi_n = a_n) = \mathsf{P}(\xi_1 = a_1) \cdot \ldots \cdot \mathsf{P}(\xi_n = a_n)$.

Определение 21. Случайные величины $\xi_1, \ldots, \xi_n$ с абсолютно непрерывным совместным распределением независимы, если плотность совместного распределения распадается в произведение плотностей, т.е. для любых $x_1, \ldots, x_n$ имеет место равенство: $f(x_1, \ldots, x_n) = f_{\xi_1}(x_1) \cdot \ldots \cdot f_{\xi_n}(x_n)$.

§ 6. Функции от двух случайных величин

Пусть ξ и η — случайные величины с плотностью совместного распределения $f(x, y)$, и задана функция $g(x, y)$ со значениями в $\mathbb{R}$. Требуется найти функцию распределения (а если существует, то и плотность распределения) случайной величины $\psi = g(\xi, \eta)$.

Мы знаем, что вероятность случайному вектору попасть в некоторую область B на плоскости можно вычислить как объём под графиком плотности совместного распределения над этой областью.

Теорема 24. *Пусть $t \in \mathbb{R}$, и область B_t состоит из точек (x, y) таких, что $g(x, y) < t$. Тогда случайная величина $\psi = g(\xi, \eta)$ имеет функцию распределения*

$$F_\psi(t) = \mathsf{P}\big(g(\xi, \eta) < t\big) = \mathsf{P}\big((\xi, \eta) \in B_t\big) = \iint\limits_{B_t} f(x, y)\, dx\, dy.$$

Далее в этой главе будем рассматривать только независимые случайные величины ξ и η, т. е. такие, что $f(x, y) \equiv f_\xi(x)\, f_\eta(y)$. В этом случае распределение величины $g(\xi, \eta)$ полностью определяется распределениями ξ и η.

Следствие 6 (формула свёртки). *Если случайные величины ξ и η независимы и имеют абсолютно непрерывные распределения с плотностями распределений $f_\xi(x)$ и $f_\eta(y)$, то плотность распределения суммы $\xi + \eta$ равна «свёртке» плотностей f_ξ и f_η:*

$$f_{\xi+\eta}(z) = \int\limits_{-\infty}^{\infty} f_\xi(x)\ f_\eta(z - x)\ dx = \int\limits_{-\infty}^{\infty} f_\eta(y)\ f_\xi(z - y)\ dy. \tag{17}$$

Доказательство. Применим теорему 24 с функцией $g(x, y) = x + y$. Интегрирование по области $B_t = \{(x, y) \mid x + y < t\}$ можно заменить последовательным вычислением двух интегралов: наружного — по переменной x, меняющейся в пределах от $-\infty$ до $+\infty$, и внутреннего — по переменной y, которая при каждом x должна быть меньше, чем $t-x$.

$$F_{\xi+\eta}(t) = \iint\limits_{B_t} f_\xi(x)\, f_\eta(y)\, dx\, dy = \int\limits_{-\infty}^{\infty} f_\xi(x)\; dx \int\limits_{-\infty}^{t-x} f_\eta(y)\, dy.$$

Сделаем в последнем интеграле замену переменной y на z так: $z=y+x$. При этом $y\in(-\infty,\, t-x)$ перейдёт в $z\in(-\infty,\, t)$, $dz=dy$. В полученном интеграле меняем порядок интегрирования:

$$F_{\xi+\eta}(t) = \int\limits_{-\infty}^{\infty} f_\xi(x)\; dx \int\limits_{-\infty}^{t} f_\eta(z-x)\; dz = \int\limits_{-\infty}^{t} dz \int\limits_{-\infty}^{\infty} f_\xi(x)\, f_\eta(z-x)\, dx.$$

Итак, мы представили функцию распределения $F_{\xi+\eta}(t)$ в виде интеграла от плотности, задаваемой формулой свёртки. $\square$

§ 7. Устойчивость распределений по суммированию

П р и м е р 57. Пусть независимые случайные величины ξ и η имеют стандартное нормальное распределение. Докажем, что их сумма снова имеет нормальное распределение, но уже с параметрами $a=0$ и $\sigma^2=2$.

Д о к а з а т е л ь с т в о. По формуле свёртки, плотность суммы равна

$$f_{\xi+\eta}(z) = \int\limits_{-\infty}^{\infty} \frac{1}{2\pi}\, e^{-x^2/2}\, e^{-(z-x)^2/2}\; dx = \int\limits_{-\infty}^{\infty} \frac{1}{2\pi}\; e^{-\left(x^2+\frac{z^2}{2}-zx\right)}\; dx =$$

$$= e^{-z^2/4} \int\limits_{-\infty}^{\infty} \frac{1}{2\pi}\, e^{-\left(x-\frac{z}{2}\right)^2}\; dx = \frac{1}{2\sqrt{\pi}}\, e^{-z^2/4} \int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}}\, e^{-v^2}\; dv = \frac{e^{-z^2/4}}{2\sqrt{\pi}}.$$

Последний интеграл равен единице, поскольку под интегралом стоит плотность нормального распределения с параметрами $a=0$ и $\sigma^2=\dfrac{1}{2}$.

Итак, мы получили, что плотность суммы есть плотность нормального распределения с параметрами 0 и 2. $\square$

Верно и более общее утверждение (*доказать!*).

С в о й с т в о 8. *Пусть случайные величины* $\xi \sim \mathrm{N}_{a_1,\sigma_1^2}$ *и* $\eta \sim \mathrm{N}_{a_2,\sigma_2^2}$ *независимы. Тогда* $\xi+\eta \sim \mathrm{N}_{a_1+a_2,\,\sigma_1^2+\sigma_2^2}$.

Если сумма двух независимых случайных величин из одного и того же распределения (возможно, с разными параметрами) имеет такое же распределение, говорят, что это распределение *устойчиво* относительно суммирования.

В утверждениях этого параграфа перечислены почти все устойчивые по суммированию распределения. Часть утверждений мы докажем.

Свойство 9. *Пусть случайные величины* $\xi \sim \mathrm{B}_{n,p}$ *и* $\eta \sim \mathrm{B}_{m,p}$ *независимы. Тогда* $\xi + \eta \sim \mathrm{B}_{n+m,p}$.

Это утверждение имеет простую интерпретацию: если сложить число успехов в первых n и в следующих m независимых испытаниях схемы Бернулли, получится число успехов в $n+m$ испытаниях.

Из свойства 9, в частности, следует, что сумма n независимых случайных величин с одним и тем же распределением Бернулли $\mathrm{B}_p = \mathrm{B}_{1,p}$ имеет биномиальное распределение $\mathrm{B}_{n,p}$. Этим фактом мы скоро воспользуемся.

Свойство 10. *Пусть случайные величины* $\xi \sim \Pi_\lambda$ *и* $\eta \sim \Pi_\mu$ *независимы. Тогда* $\xi + \eta \sim \Pi_{\lambda+\mu}$.

Доказательство. Найдём таблицу распределения суммы. Для любого целого $k \geqslant 0$

$$\mathsf{P}(\xi+\eta=k) = \mathsf{P}(\xi=0,\eta=k) + \mathsf{P}(\xi=1,\eta=k-1) + \ldots + \mathsf{P}(\xi=k,\eta=0) =$$
$$= \sum_{i=0}^{k} \mathsf{P}(\xi=i)\cdot\mathsf{P}(\eta=k-i) = \sum_{i=0}^{k} \frac{\lambda^i}{i!}e^{-\lambda}\cdot\frac{\mu^{k-i}}{(k-i)!}e^{-\mu} =$$
$$= e^{-(\lambda+\mu)}\frac{1}{k!}\sum_{i=0}^{k}\frac{k!}{i!\,(k-i)!}\lambda^i\mu^{k-i} = e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^k}{k!}.$$

В последнем равенстве мы воспользовались биномом Ньютона. □

Свойство 11. *Пусть случайные величины* $\xi \sim \Gamma_{\alpha,\lambda_1}$ *и* $\eta \sim \Gamma_{\alpha,\lambda_2}$ *независимы. Тогда* $\xi+\eta \sim \Gamma_{\alpha,\lambda_1+\lambda_2}$.

Доказательство. Плотности распределений величин ξ и η равны соответственно

$$f_\xi(x) = \begin{cases} 0, & x \leqslant 0,\\ \dfrac{\alpha^{\lambda_1}}{\Gamma(\lambda_1)}x^{\lambda_1-1}e^{-\alpha x}, & x>0,\end{cases} \qquad f_\eta(y) = \begin{cases} 0, & y \leqslant 0,\\ \dfrac{\alpha^{\lambda_2}}{\Gamma(\lambda_2)}y^{\lambda_2-1}e^{-\alpha y}, & y>0.\end{cases}$$

Тогда по формуле свёртки плотность суммы равна

$$f_{\xi+\eta}(z) = \int\limits_0^\infty \frac{\alpha^{\lambda_1}}{\Gamma(\lambda_1)}x^{\lambda_1-1}e^{-\alpha x}f_\eta(z-x)\,dx.$$

Но $f_\eta(z-x)=0$ при $x>z$. Поэтому при $z>0$ плотность под интегралом отлична от нуля, только если переменная интегрирования изменяется в пределах $0<x<z$. При $z\leqslant 0$ подынтегральная функция всегда равна нулю.

При $z > 0$ имеем:

$$f_{\xi+\eta}(z) = \int\limits_0^z \frac{\alpha^{\lambda_1}}{\Gamma(\lambda_1)}\, x^{\lambda_1-1} e^{-\alpha x} \frac{\alpha^{\lambda_2}}{\Gamma(\lambda_2)}\, (z-x)^{\lambda_2-1} e^{-\alpha(z-x)}\, dx = \\ = \frac{\alpha^{\lambda_1+\lambda_2}}{\Gamma(\lambda_1)\Gamma(\lambda_2)}\, e^{-\alpha z} \int\limits_0^z x^{\lambda_1-1}(z-x)^{\lambda_2-1} dx.$$

Последний интеграл заменой $x = tz$ сводится к *бета-функции*

$$B(a,b) = \int\limits_0^1 t^{a-1}(1-t)^{b-1} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad a > 0,\ b > 0.$$

После замены получаем:

$$f_{\xi+\eta}(z) = \frac{\alpha^{\lambda_1+\lambda_2}}{\Gamma(\lambda_1)\Gamma(\lambda_2)}\, e^{-\alpha z} z^{\lambda_1+\lambda_2-1} \int\limits_0^1 t^{\lambda_1-1}(1-t)^{\lambda_2-1} dt = \frac{\alpha^{\lambda_1+\lambda_2}}{\Gamma(\lambda_1+\lambda_2)}\, z^{\lambda_1+\lambda_2-1} e^{-\alpha z}.$$

Поэтому $\xi + \eta \sim \Gamma_{\alpha,\,\lambda_1+\lambda_2}$, что и требовалось доказать. □

Показательное распределение не устойчиво по суммированию: сумма независимых и одинаково показательно распределённых величин имеет не показательное, но гамма-распределение.

С в о й с т в о 12. *Пусть независимые случайные величины* $\xi, \ldots, \xi_n$ *имеют показательное распределение* $\mathrm{E}_\alpha = \Gamma_{\alpha,1}$. *Тогда* $\xi + \ldots + \xi_n \sim \Gamma_{\alpha,\,n}$.

П р и м е р 58. Равномерное распределение не является устойчивым относительно суммирования. Найдём функцию и плотность распределения суммы двух независимых случайных величин с одинаковым равномерным на отрезке $[0,\ 1]$ распределением, но не по формуле свёртки, а используя геометрическую вероятность.

Пусть $\xi,\ \eta \sim \mathrm{U}_{0,1}$ — независимые случайные величины. Пару $(\xi,\ \eta)$ можно считать координатой точки, брошенной наудачу в единичный квадрат.

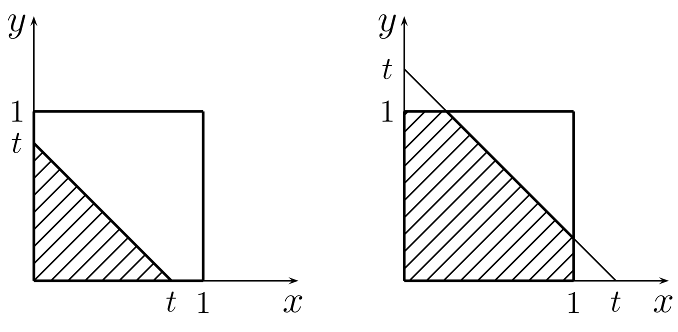

Рис. 19.

Тогда $F_{\xi+\eta}(t) = \mathsf{P}(\xi+\eta < t)$ равна площади области внутри квадрата под прямой $y = t - x$. Эта область — заштрихованный на рис. 19 треугольник при $0 < t \leqslant 1$ или пятиугольник при $1 < t \leqslant 2$.

Получим следующие функцию распределения и плотность распределения:

$$F_{\xi+\eta}(t) = \begin{cases} 0, & t \leqslant 0, \\ \frac{1}{2}t^2, & 0 < t \leqslant 1, \\ 1-\frac{1}{2}(2-t)^2, & 1 < t \leqslant 2, \\ 1, & t > 2; \end{cases} \qquad f_{\xi+\eta}(t) = \begin{cases} 0, & t \notin (0,\, 2), \\ t, & 0 < t \leqslant 1, \\ 2-t, & 1 < t \leqslant 2. \end{cases}$$

Это плотность так называемого «треугольного» распределения Симпсона. Распределение суммы независимых случайных величин с равномерным распределением не является равномерным.

§ 8. Контрольные вопросы

1. Что такое функция распределения случайного вектора?
2. Как по функции распределения вектора находят функции распределения его координат?
3. Что такое таблица совместного распределения?
4. Как по таблице совместного распределения двух случайных величин находят их частные распределения?
5. Какими свойствами обладает плотность совместного распределения?
6. Как по плотности совместного распределения двух случайных величин находят их частные плотности?
7. Если плотность распределения случайного вектора, состоящего из n случайных величин, проинтегрировать по одной из переменных в границах от $-\infty$ до $+\infty$, что за плотность получится? Будет ли это плотность вектора из $n-1$ случайной величины?
8. Можно ли найти совместное распределение по частным распределениям?
9. Привести пример того, что при одних и тех же частных распределениях возможны разные совместные.
10. Что такое многомерное равномерное распределение?
11. Что такое многомерное нормальное распределение?
12. Дать определение независимости двух случайных величин.
13. Пусть для каких-то двух множеств B_1 и B_2 оказалось верно равенство $\mathsf{P}(\xi \in B_1,\, \eta \in B_2) = \mathsf{P}(\xi \in B_1) \cdot \mathsf{P}(\eta \in B_2)$. Следует ли отсюда независимость величин ξ и η?

14. Дать определение независимости двух случайных величин с дискретными распределениями.

15. Пусть для случайных величин ξ и η с дискретными распределениями оказалось, что $\mathsf{P}(\xi = 0, \eta = 0) = \mathsf{P}(\xi = 0) \cdot \mathsf{P}(\eta = 0)$. Следует ли отсюда независимость величин ξ и η?

16. Дать определение зависимости двух случайных величин.

17. Как записать плотность совместного распределения двух независимых случайных величин, зная плотность распределения каждой?

18. Как вычислить плотность распределения суммы двух независимых случайных величин, зная плотность распределения каждой?

19. Что такое устойчивость распределения по суммированию?

20. Перечислить устойчивые по суммированию распределения.

21. Зачем в свойствах устойчивости требуется независимость слагаемых?

22. Имеет ли сумма независимых и равномерно распределённых слагаемых равномерное распределение?

23. Для случайной величины $\xi \sim \Pi_\lambda$ сумма $\xi+\xi = 2\xi$ принимает только чётные значения, и поэтому её распределение не является распределением Пуассона. Не противоречит ли это свойству 10?

24. Для случайных величин $\xi \sim \mathrm{N}_{0,1}$ и $-\xi \sim \mathrm{N}_{0,1}$ распределение случайной величины $\xi + (-\xi) = 0$ вырождено, и поэтому не является нормальным. Не противоречит ли это свойству 8?

25. Объясните, откуда следует факт: для любых независимых случайных величин ξ и η с нормальными распределениями любая линейная комбинация $a\xi + b\eta + c$, где $ab \neq 0$, снова имеет нормальное распределение.

ГЛАВА VII

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ РАСПРЕДЕЛЕНИЙ

Знание закона распределения случайной величины даёт нам полную информацию о её вероятностном поведении. Существуют, однако, и более грубые числовые характеристики, позволяющие, даже не зная распределения, многое сказать о случайной величине. К таким характеристикам относятся центр тяжести распределения, или его математическое ожидание; момент инерции распределения, или его дисперсия; другие характеристики положения и рассеивания.

§ 1. Математическое ожидание случайной величины

Определение 22. *Математическим ожиданием* (*средним значением*, *первым моментом*) случайной величины ξ, имеющей дискретное распределение со значениями a_1, a_2, ..., называется *число*

$$\mathsf{E}\,\xi = \sum_i a_i p_i = \sum_i a_i \mathsf{P}(\xi = a_i),$$

если данный ряд абсолютно сходится, т. е. если $\sum |a_i| p_i < \infty$. В противном случае говорят, что математическое ожидание не существует.

Определение 23. *Математическим ожиданием* случайной величины ξ, имеющей абсолютно непрерывное распределение с плотностью распределения $f(x)$, называется *число*

$$\mathsf{E}\,\xi = \int\limits_{-\infty}^{\infty} x f(x) dx,$$

если этот интеграл абсолютно сходится, т.е. если $\int\limits_{-\infty}^{\infty} |x| f(x)\, dx < \infty$. Иначе математическое ожидание не существует.

Математическое ожидание имеет простой физический смысл: если на прямой, как на невесомом стержне, распределить единичную массу, поместив в точки a_i массу p_i или «размазав» её с плотностью $f(x)$, то точка $\mathsf{E}\,\xi$ будет координатой «центра тяжести» прямой. Если математическое ожидание существует, то стержень, подвешенный в этой точке, будет находиться в состоянии равновесия.

П р и м е р 59. Пусть случайная величина ξ равна числу очков, выпадающих при одном подбрасывании кубика. Тогда

$$\mathsf{E}\,\xi = \sum_{k=1}^{6} k \cdot \frac{1}{6} = \frac{1}{6}\,(1+2+3+4+5+6) = 3{,}5.$$

В среднем при одном подбрасывании кубика выпадает $3{,}5$ очка. Это несмотря на то, что ни при одном подбрасывании $3{,}5$ очка выпасть не может.

П р и м е р 60. Пусть случайная величина ξ — координата точки, брошенной наудачу на отрезок $[a,\, b]$. Тогда

$$\mathsf{E}\,\xi = \int\limits_a^b x \cdot \frac{1}{b-a}\,dx = \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}.$$

Центр тяжести равномерного распределения на отрезке — середина отрезка.

В § 3 вычислены математические ожидания основных распределений. Там же можно увидеть примеры распределений, математические ожидания которых не существуют (примеры 73 и 74).

Свойства математического ожидания. Во всех свойствах ниже предполагается, что рассматриваемые математические ожидания существуют. Часть свойств мы оставим без доказательства, а остальные докажем только для абсолютно непрерывных распределений. Первое свойство позволяет вычислять математические ожидания функций от случайных величин по исходному распределению этих величин.

(E1) Для произвольной функции $g(x)$ со значениями в $\mathbb{R}$

$$\mathsf{E}\,g(\xi) = \begin{cases} \sum\limits_k g(a_k)\mathsf{P}(\xi = a_k), & \text{если распределение } \xi \text{ дискретно;} \\ \int\limits_{-\infty}^{\infty} g(x) f_\xi(x)\,dx, & \text{если распределение } \xi \text{ абсолютно непрерывно.} \end{cases}$$

Такое же свойство верно и для числовых функций нескольких аргументов $g(x_1, \ldots, x_n)$, если ξ — вектор из n случайных величин, а в сумме и в интеграле участвует их совместное распределение. Например, для $g(x,y) = x+y$ и для случайных величин ξ и η с плотностью совместного распределения $f(x,\, y)$ верно:

$$\mathsf{E}\,(\xi+\eta) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x+y) f(x,\, y)\,dx\,dy. \qquad (18)$$

(E2) Математическое ожидание постоянной равно ей самой: $\mathsf{E}\,c = c$.

(E3) Постоянный множитель можно вынести за знак математического ожидания: $\mathsf{E}\,(c\,\xi) = c\,\mathsf{E}\,\xi$.

Это свойство следует из свойства (E1) при $g(x) = c\,x$.

(E4) Математическое ожидание суммы *любых* случайных величин равно сумме их математических ожиданий: $\mathsf{E}\,(\xi+\eta) = \mathsf{E}\,\xi + \mathsf{E}\,\eta$.

Доказательство. Воспользуемся равенством (18) и теоремой 23:

$$\begin{aligned}\mathsf{E}\,(\xi+\eta) &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x+y)f(x,\,y)\,dx\,dy = \\ &= \int\limits_{-\infty}^{\infty} x\,dx\int\limits_{-\infty}^{\infty} f(x,\,y)\,dy \;+\; \int\limits_{-\infty}^{\infty} y\,dy\int\limits_{-\infty}^{\infty} f(x,\,y)\,dx = \\ &= \int\limits_{-\infty}^{\infty} x\,f_\xi(x)\,dx \;+\; \int\limits_{-\infty}^{\infty} y\,f_\eta(y)\,dy \;=\; \mathsf{E}\,\xi + \mathsf{E}\,\eta. \qquad \square\end{aligned}$$

(E5) Если $\xi \geqslant 0$, то $\mathsf{E}\,\xi \geqslant 0$.

Доказательство. Неотрицательность ξ означает, что $a_i \geqslant 0$ при всех i в случае дискретного распределения, либо $f_\xi(x) = 0$ при $x < 0$ — для абсолютно непрерывного распределения. И в том, и в другом случае имеем:

$$\mathsf{E}\,\xi = \sum a_i p_i \geqslant 0 \quad \text{или} \quad \mathsf{E}\,\xi = \int\limits_{0}^{\infty} x\,f(x)\,dx \geqslant 0. \qquad \square$$

Из свойства (E5) вытекает множество полезных утверждений, например:

Следствие 7. *Если* $\xi \leqslant \eta$, *то* $\mathsf{E}\,\xi \leqslant \mathsf{E}\,\eta$.

Следствие 8. *Если* $a \leqslant \xi \leqslant b$, *то* $a \leqslant \mathsf{E}\,\xi \leqslant b$.

Второе особенно очевидно: центр тяжести стержня не может находиться вне отрезка, если вся масса сосредоточена на этом отрезке.

(E7) Математическое ожидание произведения *независимых* случайных величин равно произведению их математических ожиданий.

Доказательство. В равенстве (18) заменим сложение умножением и плотность совместного распределения произведением плотностей:

$$\begin{aligned}\mathsf{E}\,(\xi\eta) &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} xy\,f_\xi(x)\,f_\eta(y)\,dx\,dy = \\ &= \int\limits_{-\infty}^{\infty} x\,f_\xi(x)\,dx\int\limits_{-\infty}^{\infty} y\,f_\eta(y)\,dy \;=\; \mathsf{E}\,\xi\,\mathsf{E}\,\eta. \qquad \square\end{aligned}$$

Замечание 1. Обратное утверждение к свойству (E7) неверно: из равенства $\mathsf{E}(\xi\eta) = \mathsf{E}\,\xi\,\mathsf{E}\,\eta$ *не следует* независимость величин ξ и η.

Пример 61. Пусть ξ принимает значения 0 и ± 1 с одинаковыми вероятностями, и $\eta = \xi^2$. Это зависимые случайные величины:

$$\mathsf{P}(\xi = 1, \eta = 0) = \mathsf{P}(\xi = 1, \xi^2 = 0) = 0 \neq \frac{1}{3} \cdot \frac{1}{3} = \mathsf{P}(\xi = 1)\,\mathsf{P}(\eta = 0).$$

Однако $\mathsf{E}\,\xi = 0$ и $\mathsf{E}(\xi\eta) = \mathsf{E}(\xi^3) = 0$, поэтому $\mathsf{E}(\xi\eta) = \mathsf{E}\,\xi\,\mathsf{E}\,\eta$.

§ 2. Дисперсия и моменты старших порядков

Определение 24. Пусть $\mathsf{E}\,|\xi|^k < \infty$. Число $\mathsf{E}\,\xi^k$ называется моментом порядка k или k-м моментом случайной величины ξ, число $\mathsf{E}\,|\xi|^k$ — абсолютным k-м моментом, число $\mathsf{E}(\xi - \mathsf{E}\,\xi)^k$ — центральным k-м моментом, и число $\mathsf{E}\,|\xi - \mathsf{E}\,\xi|^k$ — абсолютным центральным k-м моментом случайной величины ξ. Число $\mathsf{D}\,\xi = \mathsf{E}(\xi - \mathsf{E}\,\xi)^2$ (центральный момент второго порядка) называется *дисперсией* случайной величины ξ.

Пример 62. Пусть, скажем, случайная величина ξ принимает значение 0 с вероятностью 0,99999, и значение 100 с вероятностью 0,00001. Посмотрим, как моменты разных порядков реагируют на большие, но маловероятные значения случайной величины:

$$\begin{aligned}
\mathsf{E}\,\xi &= 0 \cdot 0{,}99999 + 100 \cdot 0{,}00001 = 0{,}001,\\
\mathsf{E}\,\xi^2 &= 0^2 \cdot 0{,}99999 + 100^2 \cdot 0{,}00001 = 0{,}1,\\
\mathsf{E}\,\xi^6 &= 0^6 \cdot 0{,}99999 + 100^6 \cdot 0{,}00001 = 10\,000\,000.
\end{aligned}$$

Пример 63. Дисперсия $\mathsf{D}\,\xi = \mathsf{E}(\xi - \mathsf{E}\,\xi)^2$ есть *среднее значение квадрата отклонения случайной величины ξ от её среднего*. Посмотрим, за что эта величина отвечает.

Пусть случайная величина ξ принимает значения ± 1 с равными вероятностями, а случайная величина η — значения ± 10 с равными вероятностями. Тогда $\mathsf{E}\,\xi = \mathsf{E}\,\eta = 0$, поэтому $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 = 1$, $\mathsf{D}\,\eta = \mathsf{E}\,\eta^2 = 100$.

Говорят, что дисперсия характеризует *степень разброса* значений случайной величины вокруг её математического ожидания.

Определение 25. Число $\sigma = \sqrt{\mathsf{D}\,\xi}$ называют *среднеквадратическим отклонением* случайной величины ξ.

Сравнение моментов. Моменты меньших порядков существуют, если если существуют моменты более высокого порядка.

Теорема 25. *Если существует момент порядка $t > 0$ случайной величины ξ, то существует и её момент порядка s, где $0 < s < t$.*

Доказательство. Заметим, что $|\xi|^s \leqslant |\xi|^t + 1$. Но следствие 7 позволяет из неравенства для случайных величин получить такое же неравенство для их математических ожиданий: $\mathsf{E}\,|\xi|^s \leqslant \mathsf{E}\,|\xi|^t + 1 < \infty$. □

Сформулируем без доказательства очень полезное неравенство.

Теорема 26 (неравенство Йенсена). *Пусть функция $g(x)$ на своей области определения выпукла, т. е. область над графиком этой функции есть выпуклое множество. Тогда для любой случайной величины ξ верно неравенство:* $\mathsf{E}\,g(\xi) \geqslant g(\mathsf{E}\,\xi)$. *Для вогнутых функций знак неравенства меняется на противоположный.*

В частности,

$$\mathsf{E}\,e^{\xi} \geqslant e^{\mathsf{E}\,\xi}, \qquad \mathsf{E}\,\xi^2 \geqslant (\mathsf{E}\,\xi)^2, \qquad \mathsf{E}\,|\xi| \geqslant |\mathsf{E}\,\xi|,$$

$$\mathsf{E}\,\ln\xi \leqslant \ln(\mathsf{E}\,\xi), \qquad \mathsf{E}\,\frac{1}{\xi} \geqslant \frac{1}{\mathsf{E}\,\xi}, \qquad \mathsf{E}\,\sqrt{\xi} \leqslant \sqrt{\mathsf{E}\,\xi}.$$

Последние три неравенства верны для положительных ξ.

Свойства дисперсии. Свойства дисперсии следуют из соответствующих свойств математического ожидания. Во всех свойствах предполагается существование вторых моментов случайных величин. Тогда матожидания существуют по теореме 25.

(D1) Дисперсия может быть вычислена по формуле: $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2$.

Доказательство. Обозначим для удобства $a = \mathsf{E}\,\xi$. Тогда

$$\mathsf{D}\,\xi = \mathsf{E}\,(\xi - a)^2 = \mathsf{E}\,(\xi^2 - 2a\xi + a^2) = \mathsf{E}\,\xi^2 - 2a\mathsf{E}\,\xi + a^2 = \mathsf{E}\,\xi^2 - a^2.$$

(D2) При умножении случайной величины на постоянную c дисперсия увеличивается в c^2 раз: $\mathsf{D}\,(c\xi) = c^2\,\mathsf{D}\,\xi$ (*доказать!*).

(D3) Дисперсия всегда неотрицательна: $\mathsf{D}\,\xi \geqslant 0$.

Доказательство. Пусть $a = \mathsf{E}\,\xi$. Дисперсия есть математическое ожидание неотрицательной случайной величины $(\xi - a)^2$, и неотрицательность дисперсии следует из свойства (E5). □

(D4) Дисперсия обращается в нуль лишь для вырожденного распределения: если $\mathsf{D}\,\xi = 0$, то $\xi = \mathrm{const}$, и наоборот.

Доказательство. Равенство $\mathsf{D}\,\xi = 0$ означает, грубо говоря, что центр тяжести массы, распределённой справа от нуля, оказался равен нулю. Это возможно только если вся масса сосредоточена в нуле, т. е. $(\xi - a)^2 = 0$, $\xi = a = \mathrm{const}$. И наоборот: если $\xi = c$, то $\mathsf{D}\,\xi = \mathsf{E}\,(c - \mathsf{E}\,c)^2 = \mathsf{E}\,0 = 0$. □

(D5) Дисперсия не зависит от сдвига случайной величины на постоянную: $\mathsf{D}\,(\xi + c) = \mathsf{D}\,\xi$ (*доказать!*).

(D6) Если ξ и η независимы, то $\mathsf{D}\,(\xi + \eta) = \mathsf{D}\,\xi + \mathsf{D}\,\eta$.

Доказательство. Действительно, применяя свойство (E7), получим:

$$\mathsf{D}(\xi+\eta) = \mathsf{E}(\xi+\eta)^2 - (\mathsf{E}(\xi+\eta))^2 =$$
$$= \mathsf{E}\,\xi^2 + \mathsf{E}\,\eta^2 + 2\mathsf{E}(\xi\eta) - (\mathsf{E}\,\xi)^2 - (\mathsf{E}\,\eta)^2 - 2\mathsf{E}\,\xi\mathsf{E}\,\eta = \mathsf{D}\,\xi + \mathsf{D}\,\eta. \quad \square$$

Замечание 2. См. замечание 1.

Следствие 9. *Если* ξ *и* η *независимы, то* $\mathsf{D}(\xi-\eta) = \mathsf{D}\,\xi + \mathsf{D}\,\eta$.

Доказательство. Из свойств (D6) и (D2) получим:

$$\mathsf{D}(\xi-\eta) = \mathsf{D}(\xi+(-\eta)) = \mathsf{D}\,\xi + \mathsf{D}(-\eta) = \mathsf{D}\,\xi + (-1)^2\mathsf{D}\,\eta = \mathsf{D}\,\xi + \mathsf{D}\,\eta. \quad \square$$

Следствие 10. *Для произвольных случайных величин* ξ *и* η *имеет место равенство:*

$$\mathsf{D}(\xi\pm\eta) = \mathsf{D}\,\xi + \mathsf{D}\,\eta \pm 2\big(\mathsf{E}(\xi\eta) - \mathsf{E}\,\xi\,\mathsf{E}\,\eta\big).$$

Это равенство мы получили в доказательстве свойства (D6). Разность $\mathsf{E}(\xi\eta) - \mathsf{E}\,\xi\,\mathsf{E}\,\eta$ обращается в нуль для независимых ξ и η. Мы изучим подробно эту разность в следующей главе.

§ 3. Средние и дисперсии стандартных распределений

Пример 64 (вырожденное распределение I_c). Математическое ожидание и дисперсию этого распределения мы знаем из свойств (E2) и (D3):

$$\mathsf{E}\,c = c, \qquad \mathsf{D}\,c = 0.$$

Пример 65 (распределение Бернулли B_p). Вычислим первый и второй моменты: $\mathsf{E}\,\xi = 1\cdot p + 0\cdot q = p$; $\mathsf{E}\,\xi^2 = 1^2\cdot p + 0^2\cdot q = p$. По ним найдём дисперсию: $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2 = p - p^2 = pq$.

Пример 66 (биномиальное распределение $\mathrm{B}_{n,p}$). Используем устойчивость биномиального распределения относительно суммирования — свойство 9 (стр. 83). Пусть $\xi_1, \ldots, \xi_n$ — независимые случайные величины с распределением Бернулли $\mathrm{B}_p = \mathrm{B}_{1,p}$. Тогда их сумма $S_n = \xi_1 + \ldots + \xi_n$ имеет распределение $\mathrm{B}_{n,p}$, и по свойству (E4),

$$\mathsf{E}\,S_n = \sum_{i=1}^{n} \mathsf{E}\,\xi_i = n\mathsf{E}\,\xi_1 = np.$$

Поскольку ξ_i независимы, и дисперсия каждой равна pq, то

$$\mathsf{D}\,S_n = \sum_{i=1}^{n} \mathsf{D}\,\xi_i = n\mathsf{D}\,\xi_1 = npq.$$

Итак, $\mathsf{E}\,\xi = np$, $\mathsf{D}\,\xi = npq$ для $\xi \sim \mathrm{B}_{n,p}$.

П р и м е р 67 (геометрическое распределение G_p). Вычислим $\mathsf{E}\,\xi$:

$$\begin{aligned}\mathsf{E}\,\xi &= \sum_{k=1}^{\infty} k\,p\,q^{k-1} = p\sum_{k=1}^{\infty} k\,q^{k-1} = p\sum_{k=1}^{\infty}\frac{dq^k}{dq} = \\ &= p\,\frac{d}{dq}\left(\sum_{k=1}^{\infty} q^k\right) = p\,\frac{d}{dq}\left(\frac{q}{1-q}\right) = p\,\frac{1}{(1-q)^2} = \frac{1}{p}.\end{aligned}$$

Вычислим так называемый «второй факториальный момент» ξ:

$$\begin{aligned}\mathsf{E}\,\xi(\xi-1) &= \sum_{k=1}^{\infty} k(k-1)\,p\,q^{k-1} = p\,q\sum_{k=0}^{\infty}\frac{d^2q^k}{dq^2} = p\,q\,\frac{d^2}{dq^2}\left(\sum_{k=0}^{\infty} q^k\right) = \\ &= p\,q\,\frac{d^2}{dq^2}\left(\frac{1}{1-q}\right) = p\,q\,\frac{2}{(1-q)^3} = \frac{2q}{p^2}.\end{aligned}$$

Найдём дисперсию через второй факториальный момент:

$$\mathsf{D}\,\xi = \mathsf{E}\,\xi(\xi-1) + \mathsf{E}\,\xi - (\mathsf{E}\,\xi)^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q-1+p}{p^2} = \frac{q}{p^2}.$$

П р и м е р 68 (распределение Пуассона Π_λ). Вычислим $\mathsf{E}\,\xi$:

$$\begin{aligned}\mathsf{E}\,\xi &= \sum_{k=0}^{\infty} k\,\frac{\lambda^k}{k!}\,e^{-\lambda} = e^{-\lambda}\sum_{k=1}^{\infty} k\,\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!} = \\ &= \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!} = \lambda e^{-\lambda}e^{\lambda} = \lambda.\end{aligned}$$

Моменты более высоких порядков легко находятся через факториальные моменты $\mathsf{E}\,\xi^{[m]} = \mathsf{E}\,\xi(\xi-1)\ldots(\xi-m+1)$ порядка m. Так, второй факториальный момент ξ равен

$$\mathsf{E}\,\xi(\xi-1) = \sum_{k=0}^{\infty} k(k-1)\,\frac{\lambda^k}{k!}\,e^{-\lambda} = \lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2 e^{-\lambda}e^{\lambda} = \lambda^2.$$

Поэтому $\mathsf{E}\,\xi^2 = \mathsf{E}\,\xi(\xi-1) + \mathsf{E}\,\xi = \lambda^2 + \lambda$ и $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2 = \lambda$.

П р и м е р 69 (равномерное распределение $\mathrm{U}_{a,b}$). Математическое ожидание $\mathsf{E}\,\xi = \dfrac{a+b}{2}$ мы вычислили в примере 60. Второй момент равен

$$\mathsf{E}\,\xi^2 = \int_{-\infty}^{\infty} x^2 f_\xi(x)\,dx = \int_a^b x^2\,\frac{1}{b-a}\,dx = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}.$$

Дисперсия равна $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2 = \dfrac{(b-a)^2}{12}$.

П р и м е р 70 (стандартное нормальное распределение $\mathrm{N}_{0,1}$). Математическое ожидание этого распределения существует, поскольку $\mathsf{E}\,|\xi| < \infty$:

$$\mathsf{E}\,|\xi| = \frac{2}{\sqrt{2\pi}} \int\limits_0^\infty x e^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}} \int\limits_0^\infty e^{-x^2/2}\,d(x^2/2) = \frac{2}{\sqrt{2\pi}} < \infty.$$

Математическое ожидание ξ равно

$$\mathsf{E}\,\xi = \int\limits_{-\infty}^{\infty} x f_\xi(x)\,dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} x\,e^{-x^2/2}\,dx = 0,$$

т. к. интеграл сходится, а подынтегральная функция нечётна. Далее,

$$\mathsf{E}\,\xi^2 = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} x^2\,e^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}} \int\limits_0^\infty x^2\,e^{-x^2/2}\,dx = -\frac{2}{\sqrt{2\pi}} \int\limits_0^\infty x\,de^{-x^2/2} =$$

$$= -\frac{2x}{\sqrt{2\pi}}\,e^{-x^2/2}\Bigg|_0^\infty + 2\int\limits_0^\infty \frac{1}{\sqrt{2\pi}}\,e^{-x^2/2}\,dx = 0 + \int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\,e^{-x^2/2}\,dx = 1.$$

Поэтому $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2 = 1 - 0 = 1$.

П р и м е р 71 (нормальное распределение N_{a,σ^2}). Мы знаем, что если $\xi \sim \mathrm{N}_{a,\sigma^2}$, то $\eta = \dfrac{\xi - a}{\sigma}$ имеет стандартное нормальное распределение. Мы только что вычислили $\mathsf{E}\,\eta = 0$, $\mathsf{D}\,\eta = 1$. Тогда (*над каждым равенством подписать, каким свойствам оно обязано!*)

$$\mathsf{E}\,\xi = \mathsf{E}\,(\sigma\eta + a) = \sigma\mathsf{E}\,\eta + a = a; \quad \mathsf{D}\,\xi = \mathsf{D}\,(\sigma\eta + a) = \sigma^2\mathsf{D}\,\eta = \sigma^2.$$

Итак, параметры a и σ^2 нормального распределения суть его математическое ожидание и дисперсия.

П р и м е р 72 (показательное распределение E_α). Найдём для произвольного натурального числа k момент порядка k:

$$\mathsf{E}\,\xi^k = \int\limits_{-\infty}^{\infty} x^k f_\xi(x)\,dx = \int\limits_0^\infty x^k\,\alpha\,e^{-\alpha x}\,dx = \frac{1}{\alpha^k} \int\limits_0^\infty (\alpha x)^k\,e^{-\alpha x}\,d(\alpha x) = \frac{k!}{\alpha^k}.$$

В последнем равенстве мы воспользовались гамма-функцией Эйлера:

$$\Gamma(k+1) = \int\limits_0^\infty u^k\,e^{-u}\,du = k!$$

Тогда $\mathsf{E}\,\xi = \dfrac{1}{\alpha}$, $\mathsf{E}\,\xi^2 = \dfrac{2}{\alpha^2}$, $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2 = \dfrac{1}{\alpha^2}$.

П р и м е р 73 (стандартное распределение Коши $C_{0,1}$). Математическое ожидание распределения Коши *не существует*, т. к. расходится интеграл

$$\mathsf{E}\,|\xi| = \int\limits_{-\infty}^{\infty} |x|\,\frac{1}{\pi(1+x^2)}\,dx = \int\limits_{0}^{\infty} \frac{1}{\pi(1+x^2)}\,dx^2 = \lim_{x\to+\infty}\frac{1}{\pi}\ln(1+x^2) = +\infty.$$

Расходится он из-за того, что подынтегральная функция ведёт себя на бесконечности как $1/x$. Поэтому не существуют ни дисперсия, ни моменты более высоких порядков этого распределения (так же как и распределения $C_{a,\sigma}$).

П р и м е р 74 (распределение Парето). У распределения Парето существуют только моменты порядка $t < \alpha$, поскольку

$$\mathsf{E}\,|\xi|^t = \int\limits_{1}^{\infty} x^t\,\alpha\,\frac{1}{x^{\alpha+1}}\,dx = \alpha\int\limits_{1}^{\infty}\frac{1}{x^{(\alpha-t)+1}}\,dx$$

сходится при $\alpha - t > 0$ и расходится при $\alpha - t \leqslant 0$.

У п р а ж н е н и е. Посчитать момент порядка $t < \alpha$ распределения Парето. При каких α у этого распределения существует дисперсия? А две тысячи триста семнадцатый момент?

§ 4. Производящая функция моментов

Рассмотрим случайную величину ξ, у которой конечны моменты всех положительных порядков. Например, такова случайная величина с показательным распределением. Составим при некотором $t \in \mathbb{R}$ ряд

$$1 + t\mathsf{E}\,\xi + \frac{t^2}{2!}\,\mathsf{E}\,\xi^2 + \frac{t^3}{3!}\,\mathsf{E}\,\xi^3 + \ldots \tag{19}$$

и предположим, что этот ряд сходится, и что его сумма равна

$$\mathsf{E}\left(1 + t\xi + \frac{t^2\xi^2}{2!} + \frac{t^3\xi^3}{3!} + \ldots\right) = \mathsf{E}\,e^{t\xi}.$$

О п р е д е л е н и е 26. *Производящей функцией моментов* называется функция $M_\xi(t)$ действительного аргумента t, равная

$$M_\xi(t) = \mathsf{E}\,e^{t\xi}.$$

Эта функция определена, если указанное математическое ожидание существует для всех t из некоторого отрезка, включающего $t = 0$.

Вычислять производящую функцию моментов удобно по свойству (E1) математического ожидания. Смысл введения этой дополнительной функции кроется в её великолепных свойствах, главное из которых: *разным распределениям отвечают разные производящие функции моментов.* Название производящей функции моментов указывает ещё на одно полезное свойство: разлагая её в ряд Тейлора по степеням t, получаем коэффициенты разложения, равные моментам случайной величины. Соответственно, k-я производная этой функции в нуле равна моменту k-го порядка:

$$\mathsf{E}\,\xi = M'_\xi(0), \quad \mathsf{E}\,\xi^2 = M''_\xi(0), \quad \mathsf{E}\,\xi^k = M^{(k)}_\xi(0).$$

Полезное свойство производящей функции моментов — производящая функция моментов суммы *независимых* случайных величин равна произведению их производящих функций:

$$M_{\xi+\eta}(t) = M_\xi(t) \cdot M_\eta(t).$$

Пример 75. Пусть $\xi \sim \mathrm{E}_\alpha$. Вычислим $\mathsf{E}\,\xi^3$ по производящей функции моментов. При $t < \alpha$

$$M_\xi(t) = \mathsf{E}\,e^{t\xi} = \int\limits_0^\infty e^{tx}\alpha\,e^{-\alpha x}dx = \alpha\int\limits_0^\infty e^{x(t-\alpha)}dx = \frac{\alpha}{\alpha - t} = \left(1 - \frac{t}{\alpha}\right)^{-1}.$$

При $\frac{t}{\alpha} < 1$ функцию $M_\xi(t) = \left(1 - \frac{t}{\alpha}\right)^{-1}$ можно разложить в ряд:

$$M_\xi(t) = 1 + \frac{t}{\alpha} + \frac{t^2}{\alpha^2} + \frac{t^3}{\alpha^3} + \ldots = 1 + t\cdot\frac{1}{\alpha} + \frac{t^2}{2!}\cdot\frac{2!}{\alpha^2} + \frac{t^3}{3!}\cdot\frac{3!}{\alpha^3} + \ldots$$

Сравнивая этот ряд с разложением (19), находим $\mathsf{E}\,\xi^3 = \frac{3!}{\alpha^3}$.

Пример 76. Пусть $\xi \sim \mathrm{N}_{0,1}$. Вычислим $\mathsf{E}\,\xi^4$ по производящей функции моментов:

$$M_\xi(t) = \mathsf{E}\,e^{t\xi} = \int\limits_{-\infty}^{\infty} e^{tx}\frac{1}{\sqrt{2\pi}}\,e^{-x^2/2}dx = e^{t^2/2}\int\limits_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\,e^{-(x-t)^2/2}dx = e^{t^2/2}.$$

Последний интеграл равен единице, поскольку это интеграл по всей прямой от плотности нормального распределения с параметрами t и 1.

При любом t функцию $M_\xi(t) = e^{t^2/2}$ можно разложить в ряд:

$$M_\xi(t) = 1 + \frac{t^2}{2} + \frac{t^4}{2^2\cdot 2!} + \frac{t^6}{2^3\cdot 3!} + \ldots = 1 + \frac{t^2}{2!} + \frac{t^4}{4!}\cdot 3 + \frac{t^6}{6!}\cdot 3\cdot 5 + \ldots$$

Сравнивая этот ряд с разложением (19), находим $\mathsf{E}\,\xi^4 = 3$. Видно, что все моменты нечётных порядков стандартного нормального распределения равны нулю, а моменты чётного порядка $2k$ равны произведению всех нечётных чисел от 1 до $2k-1$ включительно: $\mathsf{E}\,\xi^{2k} = 1\cdot 3\cdot 5\cdot\ldots\cdot(2k-1)$.

§ 5. Другие числовые характеристики распределений

Распределения можно характеризовать и многими другими показателями, большинство из которых находит основное применение в статистике. Здесь мы только кратко познакомим читателя с их определениями.

Медианой распределения случайной величины ξ называется любое из *чисел* μ таких, что

$$\mathsf{P}(\xi \leqslant \mu) \geqslant \frac{1}{2}, \quad \mathsf{P}(\xi \geqslant \mu) \geqslant \frac{1}{2}.$$

Медиана распределения всегда существует, но может быть не единственна. Например, у биномиального распределения с параметрами 3 и $\frac{1}{2}$ медианой будет любое число из отрезка $[1, 2]$. Действительно, ξ принимает значения 0, 1, 2 и 3 с вероятностями соответственно $\frac{1}{8}$, $\frac{3}{8}$, $\frac{3}{8}$ и $\frac{1}{8}$. Поэтому

$$\mathsf{P}(\xi \leqslant 1) = \frac{1}{2}, \quad \mathsf{P}(\xi \geqslant 2) = \frac{1}{2}.$$

Часто в таких случаях в качестве медианы берут середину «отрезка медиан».

Для распределений с непрерывной и строго монотонной функцией распределения F медиана является единственным решением уравнения $F(\mu) = \frac{1}{2}$. Это точка, левее и правее которой на числовой прямой сосредоточено ровно по половине всей вероятностной «массы». Если распределение имеет плотность f, то площади каждой из областей под графиком плотности слева и справа от точки μ равны друг другу и равны $\frac{1}{2}$.

Медиана является одной из *квантилей* распределения. Пусть, для простоты, функция распределения F непрерывна и строго монотонна. Тогда *квантилью* уровня δ, где $\delta \in (0, 1)$, называется решение x_δ уравнения $F(x_\delta) = \delta$. Квантиль уровня δ отрезает от области под графиком плотности область с площадью δ слева от себя. Справа от x_δ площадь области под графиком плотности равна $1 - \delta$. Медиана является квантилью уровня $\frac{1}{2}$.

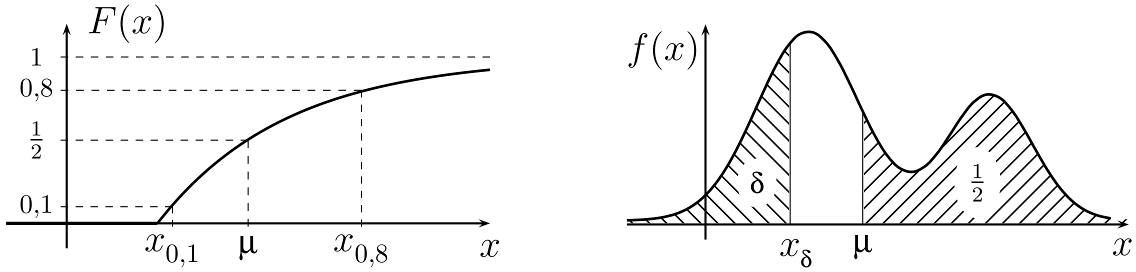

Рис. 20. Медиана и квантили на графике функции распределения и плотности

Квантили уровней, кратных 0,01, называют *процентилями,* квантили уровней, кратных 0,1, — *децилями,* уровней, кратных 0,25, — *квартилями.*

Для нормального распределения N_{a,σ^2} и медиана, и математическое ожидание равны a. Для несимметричных распределений медиана и математическое ожидание обычно не совпадают.

П р и м е р 77. Найдём по таблице значений функции $\Phi(x)$ (стр. 121) значения некоторых часто употребляющихся квантилей стандартного нормального распределения:

$$x_{0{,}005} = -x_{0{,}995} = -2{,}576, \quad x_{0{,}025} = -x_{0{,}975} = -1{,}96,$$

$$x_{0{,}05} = -x_{0{,}95} = -1{,}645, \quad x_{0{,}1} = -x_{0{,}9} = -1{,}282, \quad x_{0{,}5} = 0.$$

Модой абсолютно непрерывного распределения называют любую точку локального максимума плотности распределения. Для дискретных распределений модой считают любое значение a_i, вероятность которого p_i больше, чем вероятности соседних значений.

Распределение, обладающее единственной модой, называют *унимодальным.* Идеальным примером унимодального распределения является стандартное нормальное распределение. Плотность произвольного унимодального распределения может быть как более плоской, так и более «островершинной» по сравнению с кривой Гаусса, симметричной либо наклонённой в одну сторону. Для описания таких свойств плотности используют *коэффициент эксцесса* и *коэффициент асимметрии* распределения.

Коэффициентом асимметрии распределения с конечным третьим моментом называется число

$$\beta_1 = \mathsf{E}\left(\frac{\xi - a}{\sigma}\right)^3,$$

где $a = \mathsf{E}\,\xi$, $\sigma = \sqrt{\mathsf{D}\,\xi}$.

Для симметричных распределений коэффициент асимметрии равен нулю. При $\beta_1 > 0$ график плотности распределения более крутой слева и более пологий справа, при $\beta_1 < 0$ — наоборот.

Коэффициентом эксцесса распределения с конечным четвёртым моментом называется число

$$\beta_2 = \mathsf{E}\left(\frac{\xi - a}{\sigma}\right)^4 - 3.$$

Для всех нормальных распределений коэффициент эксцесса равен нулю. Действительно, для $\xi \sim \mathrm{N}_{a,\sigma^2}$ величина $\eta = \dfrac{\xi - a}{\sigma}$ имеет стандартное нормальное распределение по свойству 3 (стр. 63). Четвёртый момент $\mathsf{E}\,\eta^4 = 3$ этого распределения мы вычислили в примере 76. Поэтому $\beta_2 = 0$.

При $\beta_2 > 0$ плотность распределения имеет более острую вершину, чем кривая Гаусса, при $\beta_2 < 0$ — наоборот, более плоскую.

§ 6. Числовые характеристики зависимости случайных величин

Зависимость двух случайных величин можно полностью описать их совместным распределением. Однако нахождение совместного распределения часто невозможно либо из-за его сложности, либо по причине отсутствия полной информации. Хотелось бы иметь более простые характеристики, позволяющие судить о наличии и силе зависимости случайных величин без вычисления их совместного распределения.

Ковариация двух случайных величин. Мы знаем, что для независимых случайных величин с конечными вторыми моментами дисперсия их суммы равна сумме их дисперсий. В общем случае дисперсия суммы равна

$$\mathsf{D}\,(\xi+\eta) = \mathsf{D}\,\xi + \mathsf{D}\,\eta + 2\bigl(\mathsf{E}\,(\xi\eta) - \mathsf{E}\,\xi\,\mathsf{E}\,\eta\bigr). \tag{20}$$

По свойству (E7) математического ожидания, величина $\mathsf{E}\,(\xi\eta) - \mathsf{E}\,\xi\,\mathsf{E}\,\eta$ равна нулю, если случайные величины ξ и η независимы, но из равенства её нулю независимость отнюдь не следует: величины могут быть даже функционально зависимы, но иметь нулевую ковариацию, как в примере 61 на стр. 90. Тем не менее эту величину часто используют как некоторый «индикатор наличия зависимости» между двумя случайными величинами.

Определение 27. *Ковариацией* $\mathrm{cov}(\xi,\,\eta)$ случайных величин ξ и η называется число $\mathrm{cov}(\xi,\,\eta) = \mathsf{E}\,\bigl((\xi-\mathsf{E}\,\xi)(\eta-\mathsf{E}\,\eta)\bigr) = \mathsf{E}\,(\xi\eta) - \mathsf{E}\,\xi\,\mathsf{E}\,\eta$.

Свойство 13. *Ковариация обладает следующими свойствами:*

$$\mathrm{cov}(\xi,\,\xi) = \mathsf{D}\,\xi, \quad \mathrm{cov}(\xi,\,\eta) = \mathrm{cov}(\eta,\,\xi), \quad \mathrm{cov}(c\,\xi,\,\eta) = c\,\mathrm{cov}(\xi,\,\eta).$$

Упражнение. Доказать свойство 13.

Свойство 14. *Дисперсия суммы нескольких случайных величин вычисляется по любой из следующих формул:*

$$\mathsf{D}\,(\xi_1+\ldots+\xi_n) = \sum_{i=1}^{n}\mathsf{D}\,\xi_i + \sum_{i\neq j}\mathrm{cov}(\xi_i,\,\xi_j) = \sum_{i=1}^{n}\mathsf{D}\,\xi_i + 2\sum_{i<j}\mathrm{cov}(\xi_i,\,\xi_j).$$

Упражнение. Доказать свойство 14, пользуясь равенствами

$$(a_1+\ldots+a_n)^2 = \sum_{i=1}^{n}a_i^2 + \sum_{i\neq j}a_ia_j = \sum_{i=1}^{n}a_i^2 + 2\sum_{i<j}a_ia_j.$$

Говоря о многомерном нормальном распределении в § 3 главы VI (стр. 78), мы произносили слова «матрица ковариаций». Дадим им определение.

Матрица, составленная из ковариаций случайных величин ξ_i и ξ_j, называется *матрицей ковариаций* случайного вектора $(\xi_1,\ldots,\xi_n)$. Эта матрица имеет следующий вид:

$$\Sigma = \begin{pmatrix} \mathsf{D}\,\xi_1 & \mathrm{cov}(\xi_1,\,\xi_2) & \ldots & \mathrm{cov}(\xi_1,\,\xi_n) \\ \mathrm{cov}(\xi_1,\,\xi_2) & \mathsf{D}\,\xi_2 & \ldots & \mathrm{cov}(\xi_2,\,\xi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(\xi_1,\,\xi_n) & \mathrm{cov}(\xi_2,\,\xi_n) & \ldots & \mathsf{D}\,\xi_n \end{pmatrix}.$$

П р и м е р 78. Пусть случайный вектор $(\xi_1, \ldots, \xi_n)$ имеет многомерное нормальное распределение и *диагональную* матрицу ковариаций Σ: все ковариации ξ_i и ξ_j при $i \neq j$ равны нулю.

Тогда диагональной будет и матрица, обратная к Σ. Поэтому справедливо равенство (16) на стр. 79: плотность совместного распределения распадается в произведение плотностей, т. е. случайные величины независимы. Итак, для вектора, имеющего *многомерное нормальное распределение,* равенство нулю всех ковариаций равносильно независимости координат вектора.

П р и м е р 79. Покажем, что с помощью ковариации можно судить о зависимости даже тогда, когда для вычисления совместного распределения недостаточно данных. Пусть ξ и η — независимые случайные величины, и дисперсия ξ отлична от нуля (*что это значит?*). Покажем, что ξ и $\xi+\eta$ зависимы:

$$\mathsf{E}\left(\xi(\xi+\eta)\right) = \mathsf{E}\,\xi^2 + \mathsf{E}\,\xi\,\mathsf{E}\,\eta, \qquad \mathsf{E}\,\xi \cdot \mathsf{E}\,(\xi+\eta) = (\mathsf{E}\,\xi)^2 + \mathsf{E}\,\xi\,\mathsf{E}\,\eta,$$

поэтому $\mathrm{cov}(\xi, \xi+\eta) = \mathsf{E}\,\xi^2 + \mathsf{E}\,\xi\,\mathsf{E}\,\eta - \left((\mathsf{E}\,\xi)^2 + \mathsf{E}\,\xi\,\mathsf{E}\,\eta\right) = \mathsf{D}\,\xi > 0$. Следовательно, ξ и $\xi+\eta$ зависимы.

Но величина $\mathrm{cov}(\xi,\,\eta)$ не отражает силу зависимости: скажем, при увеличении ξ и η в десять раз ковариация увеличится в сто раз, но величины не станут «более зависимыми». Нужно как-то нормировать ковариацию, получив из неё «безразмерную» величину, абсолютное значение которой:

а) не менялось бы при умножении случайных величин на число;

б) свидетельствовало бы о «силе зависимости» случайных величин.

Следующая величина есть всего лишь нормированная ковариация.

Коэффициент корреляции и его свойства. Дадим определение.

О п р е д е л е н и е 28. *Коэффициентом корреляции* $\rho(\xi, \eta)$ случайных величин ξ и η, дисперсии которых существуют и отличны от нуля, называется число

$$\rho(\xi, \eta) = \frac{\mathrm{cov}(\xi,\,\eta)}{\sqrt{\mathsf{D}\,\xi}\sqrt{\mathsf{D}\,\eta}}.$$

П р и м е р 80. Рассмотрим продолжение примера 79, но пусть ξ и η не только независимы, но и имеют одинаковую дисперсию, не равную нулю.

Найдём коэффициент корреляции величин ξ и $\xi+\eta$:

$$\rho(\xi,\xi+\eta)=\frac{\operatorname{cov}(\xi,\xi+\eta)}{\sqrt{\mathsf{D}\,\xi}\sqrt{\mathsf{D}\,(\xi+\eta)}}=\frac{\mathsf{D}\,\xi}{\sqrt{\mathsf{D}\,\xi}\sqrt{\mathsf{D}\,\xi+\mathsf{D}\,\eta}}=\frac{\mathsf{D}\,\xi}{\sqrt{\mathsf{D}\,\xi}\sqrt{2\mathsf{D}\,\xi}}=\frac{1}{\sqrt{2}}.$$

Коэффициент корреляции оказался равен косинусу угла в 45°, образованного «векторами» ξ и $\xi+\eta$, когда ξ и η «ортогональны» (независимы) и их «длина» (дисперсия) одинакова.

Теорема 27. *Коэффициент корреляции обладает свойствами:*

(R1) *если* ξ *и* η *независимы, то* $\rho(\xi,\eta)=0$;

(R2) $|\rho(\xi,\eta)|\leqslant 1$;

(R3) $|\rho(\xi,\eta)|=1$ *тогда и только тогда, когда* ξ *и* η *линейно связаны.*

Доказательство. Свойство (R1) мы уже много раз (*сколько?*) упоминали и один раз доказали, и даже привели пример того, что свойство (R1) в обратную сторону неверно (пример 61 на стр. 90).

Докажем (R2). Рассмотрим преобразование $\widehat{\xi}=\dfrac{\xi-\mathsf{E}\,\xi}{\sqrt{\mathsf{D}\,\xi}}$ случайной величины, называемое *стандартизацией*. Случайная величина $\widehat{\xi}$ имеет нулевое математическое ожидание и единичную дисперсию:

$$\mathsf{E}\,\widehat{\xi}=\mathsf{E}\,\frac{\xi-\mathsf{E}\,\xi}{\sqrt{\mathsf{D}\,\xi}}=\frac{\mathsf{E}\,\xi-\mathsf{E}\,\xi}{\sqrt{\mathsf{D}\,\xi}}=0;\quad \mathsf{E}\,\widehat{\xi}^2=\mathsf{D}\,\widehat{\xi}=\mathsf{D}\,\frac{\xi-\mathsf{E}\,\xi}{\sqrt{\mathsf{D}\,\xi}}=\frac{\mathsf{D}\,(\xi-\mathsf{E}\,\xi)}{\mathsf{D}\,\xi}=1.$$

Коэффициент корреляции теперь запишется проще: $\rho(\xi,\,\eta)=\mathsf{E}\left(\widehat{\xi}\cdot\widehat{\eta}\right)$.

Далее, неравенство $(x\pm y)^2\geqslant 0$ равносильно неравенству

$$-\frac{1}{2}\,(x^2+y^2)\leqslant xy\leqslant\frac{1}{2}\,(x^2+y^2).$$

Подставив в него $\widehat{\xi}$ вместо x, $\widehat{\eta}$ вместо y и взяв математические ожидания всех частей неравенства, получим свойство (R2):

$$-1=-\frac{1}{2}\,\mathsf{E}\left(\widehat{\xi}^2+\widehat{\eta}^2\right)\ \leqslant\ \rho(\xi,\,\eta)=\mathsf{E}\left(\widehat{\xi}\cdot\widehat{\eta}\right)\ \leqslant\ \frac{1}{2}\,\mathsf{E}\left(\widehat{\xi}^2+\widehat{\eta}^2\right)=1. \tag{21}$$

Докажем свойство (R3). В одну сторону утверждение проверяется непосредственно: если $\eta=a\xi+b$, то

$$\rho(\xi,\,a\xi+b)=\frac{\mathsf{E}\,(\xi(a\xi+b))-\mathsf{E}\,\xi\cdot\mathsf{E}\,(a\xi+b)}{\sqrt{\mathsf{D}\,\xi}\sqrt{\mathsf{D}\,(a\xi+b)}}=\frac{a\mathsf{D}\,\xi}{\sqrt{\mathsf{D}\,\xi}\sqrt{a^2\mathsf{D}\,\xi}}=\begin{cases}1, & a>0;\\ -1, & a<0.\end{cases}$$

Докажем вторую часть свойства (R3): если $|\rho(\xi,\,\eta)|=1$, то существуют числа $a\neq 0$ и b такие, что $\mathsf{P}(\eta=a\xi+b)=1$.

Рассмотрим сначала случай $\rho(\xi,\,\eta)=\rho\left(\widehat{\xi},\,\widehat{\eta}\right)=1$. Это возможно только когда второе неравенство в формуле (21) превращается в равенство:

$$\mathsf{E}\left(\widehat{\xi}\cdot\widehat{\eta}\right)=\frac{1}{2}\,\mathsf{E}\left(\widehat{\xi}^2+\widehat{\eta}^2\right),\quad\text{т. е.}\quad \mathsf{E}\left(\widehat{\xi}-\widehat{\eta}\right)^2=0.$$

Математическое ожидание положительной случайной величины $\left(\widehat{\xi}-\widehat{\eta}\right)^2$ может быть равно нулю, если только $\left(\widehat{\xi}-\widehat{\eta}\right)^2=0$, т. е. если

$$\frac{\eta-\mathsf{E}\,\eta}{\sqrt{\mathsf{D}\,\eta}}=\frac{\xi-\mathsf{E}\,\xi}{\sqrt{\mathsf{D}\,\xi}}, \qquad \eta=\frac{\sqrt{\mathsf{D}\,\eta}}{\sqrt{\mathsf{D}\,\xi}}\,\xi+\mathsf{E}\,\eta-\frac{\sqrt{\mathsf{D}\,\eta}}{\sqrt{\mathsf{D}\,\xi}}\,\mathsf{E}\,\xi=a\xi+b.$$

В случае $\rho(\xi,\eta)=-1$ нужно рассмотреть первое неравенство в формуле (21) и повторить рассуждения. Тем самым теорема 27 доказана. $\square$

Полезно знать следующие часто употребляемые термины. Говорят, что ξ и η отрицательно коррелированы, если $\rho(\xi,\eta)<0$; положительно коррелированы, если $\rho(\xi,\eta)>0$; некоррелированы, если $\rho(\xi,\eta)=0$.

Смысл знака $\rho(\xi,\eta)$ хорошо виден в случае $\rho(\xi,\eta)=\pm 1$. Тогда знак ρ равен знаку a в равенстве $\eta=a\xi+b$. Так, $\rho(\xi,\eta)=1$ означает, что чем больше ξ, тем больше η. Напротив, $\rho(\xi,\eta)=-1$ означает, что чем больше ξ, тем меньше η.

Знак коэффициента корреляции в случае, когда $|\rho(\xi,\eta)|<1$, можно трактовать как тенденцию «в среднем»: чем больше ξ, тем *в среднем* больше (меньше) η. В этом случае зависимость между ξ и η уже не линейная и, возможно, даже не функциональная.

Следующее свойство показывает, что модуль коэффициента корреляции не меняется при линейных преобразованиях случайных величин (а вот знак изменится на противоположный, если $a<0$).

Свойство 15. *При любых постоянных $a\neq 0$ и b имеет место равенство:* $|\rho(a\xi+b,\eta)|=|\rho(\xi,\eta)|$.

Упражнение. Доказать свойство 15.

Пример 81. Найдём коэффициент корреляции случайных величин из примера 50 на стр. 74.

По таблице совместного распределения вычислим $\mathsf{E}\,(\xi\eta)$:

$$\mathsf{E}\,(\xi\eta)=\sum_{i,j}a_i\,b_j\,p_{i,j}=1\cdot 1\cdot\frac{1}{4}+1\cdot 2\cdot\frac{1}{4}+2\cdot 2\cdot\frac{1}{8}+2\cdot 3\cdot\frac{1}{8}=2,$$

Зная распределения ξ и η, найдём их среднее и дисперсию:

$$\mathsf{E}\,\xi=2\cdot\frac{1}{2}=1,\ \ \mathsf{D}\,\xi=2\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{2},\ \ \mathsf{E}\,\eta=3\cdot\frac{1}{2}=\frac{3}{2},\ \ \mathsf{D}\,\eta=3\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{3}{4}.$$

Вычислим коэффициент корреляции:

$$\operatorname{cov}(\xi,\eta)=2-1\cdot\frac{3}{2}=\frac{1}{2},\quad \rho(\xi,\eta)=\frac{1/2}{\sqrt{1/2}\cdot\sqrt{3/4}}=\frac{\sqrt{2}}{\sqrt{3}}.$$

§ 7. Контрольные вопросы

1. Дать определение математического ожидания случайной величины с дискретным распределением.

2. Дать определение математического ожидания случайной величины с абсолютно непрерывным распределением.

3. Математическое ожидание случайной величины — число или функция?

4. Одинаковы ли математические ожидания у двух разных случайных величин с одним и тем же распределением?

5. Какой физический смысл имеет математическое ожидание?

6. Всегда ли математическое ожидание существует?

7. Привести пример распределения, математическое ожидание которого не существует.

8. Сколько в среднем очков выпадает при бросании игральной кости?

9. Как вычислять математическое ожидание функции от случайной величины с дискретным распределением? С абсолютно непрерывным распределением?

10. Пользуясь свойствами математического ожидания, вычислить $\mathsf{E}\,1$, $\mathsf{E}\,3\xi$ и $\mathsf{E}\,\xi+1$ для $\xi \sim \mathrm{N}_{a,\,\sigma^2}$.

11. Всегда ли математическое ожидание суммы равно сумме математических ожиданий?

12. Всегда ли математическое ожидание произведения равно произведению математических ожиданий?

13. Какой физический смысл имеет дисперсия?

14. Что такое среднеквадратическое отклонение?

15. Что позволяет сравнивать неравенство Йенсена?

16. Как меняется дисперсия при изменении случайной величины вдвое?

17. Можно ли привести пример распределения с дисперсией -1?

18. Что можно сказать про случайную величину, дисперсия которой нулевая?

19. Всегда ли дисперсия суммы равна сумме дисперсий?

20. Перечислить математические ожидания и дисперсии всех основных распределений.

21. Чему равен шестой момент стандартного нормального распределения?

22. Что такое производящая функция моментов и как она «производит» моменты?

23. Чем медиана отличается от моды?

24. Какую площадь под графиком плотности отрезает медиана слева? Справа?

25. Что такое квантиль? Какую площадь под графиком плотности отрезает квантиль уровня 0,05 слева? Справа?

26. Найти все децили нормального стандартного распределения и построить по ним график функции $\Phi(x)$.

27. Для чего служат коэффициенты эксцесса и асимметрии?

28. Чему равны коэффициенты эксцесса и асимметрии стандартного нормального распределения? Любого нормального распределения?

29. Что показывает коэффициент корреляции?

30. Если $\xi = 2\eta$, чему равен их коэффициент корреляции?

31. Что можно сказать про случайные величины, если их коэффициент корреляции равен -1?

32. Чему равен коэффициент корреляции числа гербов и числа решек при ста бросках монеты?

33. Если каждую из двух случайных величин увеличить вдвое, как изменится их коэффициент корреляции?

34. Если каждую из двух случайных величин увеличить на два, как изменится их коэффициент корреляции?

35. Чему равен коэффициент корреляции независимых случайных величин?

36. Чему равен коэффициент корреляции ξ и ξ^2, где $\xi \sim \mathrm{N}_{0,1}$? Можно ли отсюда сделать вывод, что эти величины независимы?

ГЛАВА VIII

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Мы уже знакомы с некоторыми предельными теоремами теории вероятностей и знаем, что они служат для приближённого вычисления вероятностей событий, связанных с очень длинными сериями независимых испытаний. Все предельные теоремы, которые мы рассмотрим в настоящей главе, касаются поведения последовательности сумм (или средних арифметических) большого числа независимых случайных величин. Эти теоремы можно разбить на две категории: грубые свойства, называемые *законами больших чисел*, утверждают, что средние арифметические большого числа слагаемых приближаются в некотором смысле к постоянной, неслучайной величине (стабилизируются). Так, доля успехов в схеме Бернулли приближается к вероятности успеха в одном испытании. Более тонкие свойства, называемые собственно *предельными теоремами*, касаются скорости сближения среднего арифметического с постоянной и позволяют заменять распределение суммы подходящим нормальным распределением.

§ 1. Сходимость по вероятности

Пусть $\xi_1, \xi_2, \ldots$ — бесконечная последовательность случайных величин, заданных на одном и том же вероятностном пространстве. Можно ли говорить о сходимости этой последовательности в том же смысле, в каком сходятся числовые последовательности?

Напомним определение: числовая последовательность $\{b_n\}$ сходится к числу b при $n \to \infty$, если для любого $\varepsilon > 0$ найдётся число $N = N(\varepsilon)$ такое, что для любого $n \geqslant N$ выполнено неравенство $|b_n - b| < \varepsilon$.

Для *случайных величин* такое определение сходимости ξ_n к числу b было бы слишком жёстким: оно требует, чтобы для n, начиная с некоторого, все значения случайных величин ξ_n попадали в интервал $(b - \varepsilon,\, b + \varepsilon)$. Например, бросая n раз правильную монету, мы с положительной (но маленькой) вероятностью можем получить n гербов. В этом случае доля гербов среди всех испытаний равна единице и от ожидаемой в пределе $\frac{1}{2}$ отличается существенно. Но вероятность столь неприятного события — получить одни лишь гербы — уменьшается с ростом числа испытаний.

Итак, потребуем, чтобы вероятность события $\{|\xi_n - b| \geqslant \varepsilon\}$, состоящего из тех элементарных исходов ω, для которых $\xi_n(\omega)$ не попадает в «ε-окрестность» числа b, уменьшалась до нуля с ростом n. Такая сходимость называется сходимостью «по вероятности».

Определение 29. Говорят, что последовательность случайных величин $\{\xi_n\}$ сходится *по вероятности* к числу b при $n \to \infty$, и пишут: $\xi_n \xrightarrow{\text{p}} b$, если для любого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\,(|\xi_n - b| \geqslant \varepsilon) = 0 \quad \text{(или, равносильно,} \quad \lim_{n\to\infty} \mathsf{P}\,(|\xi_n - b| < \varepsilon) = 1).$$

Сходимость по вероятности обладает обычными свойствами пределов числовых последовательностей: *предел суммы* (*произведения, частного*) *сходящихся по вероятности последовательностей равен сумме* (*произведению, частному*) *пределов.*

Сходимость по вероятности не портится под действием непрерывной функции: *если* $\xi_n \xrightarrow{\text{p}} b$ *и* $g(x)$ *непрерывна в точке* b, *то* $g(\xi_n) \xrightarrow{\text{p}} g(b)$.

Чтобы устанавливать сходимость по вероятности, нужно уметь доказывать сходимость вероятностей $\mathsf{P}\,(|\xi_n - b| \geqslant \varepsilon)$ к нулю. Прямое вычисление этих вероятностей не всегда возможно, особенно если ξ_n — сумма большого числа случайных величин. Получим оценки сверху для таких вероятностей.

§ 2. Вероятностные неравенства

Теорема 28 (неравенство Маркова). *Для любого* $x > 0$

$$\mathsf{P}\big(|\xi| \geqslant x\big) \leqslant \frac{\mathsf{E}\,|\xi|}{x}.$$

Доказательство. Если $\mathsf{E}\,|\xi| = \infty$, неравенство становится тривиальным. Поэтому будем считать, что $\mathsf{E}\,\xi$ существует.

Нам потребуется следующее понятие. Назовём *индикатором* события A случайную величину $I(A)$, равную единице, если событие A произошло, и нулю, если A не произошло.

По определению, величина $I(A)$ имеет распределение Бернулли с параметром $p = \mathsf{P}(A)$, и её математическое ожидание равно вероятности успеха $p = \mathsf{P}(A)$. Индикаторы прямого и противоположного событий связаны равенством $I(A) + I(\overline{A}) = 1$. Поэтому для $A = \{|\xi| \geqslant x\}$ имеем:

$$|\xi| = |\xi| \cdot \big(I(A) + I(\overline{A})\big) \geqslant |\xi| \cdot I(A) \geqslant x \cdot I(A).$$

Тогда

$$\mathsf{E}\,|\xi| \geqslant \mathsf{E}\,\big(x \cdot I(A)\big) = x \cdot \mathsf{P}\big(A\big) = x \cdot \mathsf{P}\big(|\xi| \geqslant x\big). \tag{22}$$

Осталось разделить обе части неравенства (22) на положительное x. □

С л е д с т в и е 11 (неравенство Чебышёва). *Пусть* $\mathsf{D}\,\xi$ *существует. Тогда для любого* $x > 0$

$$\mathsf{P}\big(|\xi - \mathsf{E}\,\xi| \geqslant x\big) \leqslant \frac{\mathsf{D}\,\xi}{x^2}.$$

Д о к а з а т е л ь с т в о. Для $x > 0$ неравенство $|\xi - \mathsf{E}\,\xi| \geqslant x$ равносильно неравенству $(\xi - \mathsf{E}\,\xi)^2 \geqslant x^2$, поэтому

$$\mathsf{P}\big(|\xi - \mathsf{E}\,\xi| \geqslant x\big) = \mathsf{P}\big((\xi - \mathsf{E}\,\xi)^2 \geqslant x^2\big) \leqslant \frac{\mathsf{E}\,(\xi - \mathsf{E}\,\xi)^2}{x^2} = \frac{\mathsf{D}\,\xi}{x^2}. \quad \square$$

В качестве примера использования неравенства Чебышёва оценим вероятность случайной величине отличаться от своего математического ожидания более чем на три корня из дисперсии. Разумеется, для каждого распределения величина этой вероятности своя: для нормального распределения, например, 0,0027 — см. свойство 7. Мы получим верную для всех распределений с конечной дисперсией оценку сверху для вероятности случайной величине отличаться от своего математического ожидания более чем на три корня из дисперсии.

С л е д с т в и е 12. *Если* $\mathsf{D}\,\xi < \infty$, *то* $\mathsf{P}\big(|\xi - \mathsf{E}\,\xi| \geqslant 3\sqrt{\mathsf{D}\,\xi}\big) \leqslant \dfrac{1}{9}$.

Д о к а з а т е л ь с т в о. Согласно следствию 11,

$$\mathsf{P}\big(|\xi - \mathsf{E}\,\xi| \geqslant 3\sqrt{\mathsf{D}\,\xi}\big) \leqslant \frac{\mathsf{D}\,\xi}{\big(3\sqrt{\mathsf{D}\,\xi}\big)^2} = \frac{1}{9}. \qquad \square$$

§ 3. Законы больших чисел

Законами больших чисел (ЗБЧ) принято называть утверждения о том, при каких условиях среднее арифметическое случайных величин «стабилизируется» с ростом числа слагаемых. Всюду, где будут встречаться одинаково распределённые случайные величины ξ_i, число $a = \mathsf{E}\,\xi_1$ будет обозначать их (у всех одинаковое) математическое ожидание, σ^2 — их дисперсию.

Т е о р е м а 29 (ЗБЧ Чебышёва). *Для любой последовательности* $\xi_1, \xi_2, \ldots$ *независимых и одинаково распределённых случайных величин с конечной дисперсией имеет место сходимость:*

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{\text{p}} a. \tag{23}$$

ЗБЧ утверждает, что среднее арифметическое большого числа одинаково устроенных случайных слагаемых сближается со средним значением одного слагаемого. Как бы сильно каждая случайная величина не отклонялась от своего математического ожидания, при суммировании эти отклонения «взаимно гасятся», так что среднее арифметическое приближается к постоянной величине.

Требование конечности дисперсии связано исключительно со способом доказательства: утверждение останется верным, если требовать существования только первого момента.

Доказательство. Обозначим через $S_n = \xi_1 + \ldots + \xi_n$ сумму первых n случайных величин. Из линейности математического ожидания получим:

$$\mathsf{E}\left(\frac{S_n}{n}\right) = \frac{\mathsf{E}\,\xi_1 + \ldots + \mathsf{E}\,\xi_n}{n} = \frac{na}{n} = a.$$

Пусть $\varepsilon > 0$. Воспользуемся неравенством Чебышёва (следствие 11) и тем, что дисперсия суммы независимых слагаемых равна сумме их дисперсий:

$$\mathsf{P}\left(\left|\frac{S_n}{n} - a\right| \geqslant \varepsilon\right) \leqslant \frac{\mathsf{D}\,(S_n/n)}{\varepsilon^2} = \frac{\mathsf{D}\,S_n}{n^2\varepsilon^2} = \frac{\mathsf{D}\,\xi_1 + \ldots + \mathsf{D}\,\xi_n}{n^2\varepsilon^2} =$$
$$= \frac{n\,\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0 \text{ при } n \to \infty. \qquad \square$$

Замечание 3. Мы не только доказали сходимость, но и получили оценку для вероятности среднему арифметическому любого числа независимых и одинаково распределённых величин отличаться от $a = \mathsf{E}\,\xi_1$ более чем на заданное ε:

$$\mathsf{P}\left(\left|\frac{\xi_1 + \ldots + \xi_n}{n} - a\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}. \tag{24}$$

Получим в качестве следствия из ЗБЧ Чебышёва закон больших чисел Бернулли. В отличие от ЗБЧ Чебышёва, описывающего предельное поведение среднего арифметического случайных величин с *произвольными* распределениями, ЗБЧ Бернулли — утверждение только для схемы Бернулли.

Теорема 30 (ЗБЧ Бернулли). *Пусть* ν_n *— число успехов в* n *испытаниях схемы Бернулли с вероятностью успеха* p. *Тогда* $\frac{\nu_n}{n} \xrightarrow{\text{p}} p$. *При этом для любого* $\varepsilon > 0$

$$\mathsf{P}\left(\left|\frac{\nu_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{pq}{n\varepsilon^2}.$$

Доказательство. Величина ν_n — сумма независимых, одинаково распределённых случайных величин, имеющих распределение Бернулли с параметром p (индикаторов того, что в соответствующем испытании случился успех): $\nu_n = \xi_1 + \ldots + \xi_n$, где

$$\xi_i = \begin{cases} 1, & \text{если в } i\text{-м испытании успех;} \\ 0, & \text{если в } i\text{-м испытании неудача;} \end{cases}$$

$a = \mathsf{E}\,\xi_1 = p$, $\sigma^2 = \mathsf{D}\,\xi_1 = pq$. Осталось воспользоваться ЗБЧ в форме Чебышёва и неравенством (24). $\square$

П р и м е р 82. Правильная монета подбрасывается 10 000 раз. Оценим вероятность того, что число гербов отличается от 5 000 менее, чем на 100.

Пусть ν_n — число гербов, выпавших в $n = 10\,000$ испытаниях. Нужно оценить вероятность события

$$\{4\,900 < \nu_n < 5\,100\} = \left\{0{,}49 < \frac{\nu_n}{n} < 0{,}51\right\} = \left\{\left|\frac{\nu_n}{n} - \frac{1}{2}\right| < 0{,}01\right\}.$$

Поскольку $p = q = \frac{1}{2}$, искомая оценка выглядит так:

$$\mathsf{P}(4\,900 < \nu_n < 5\,100) = 1 - \mathsf{P}\left(\left|\frac{\nu_n}{n} - \frac{1}{2}\right| \geqslant 0{,}01\right) \geqslant 1 - \frac{pq}{n \cdot 0{,}01^2} = \frac{3}{4}.$$

Итак, ЗБЧ Бернулли (а на самом деле, неравенство Чебышёва) позволяет заключить, что в среднем не более чем в четверти случаев при 10 000 подбрасываниях монеты частота выпадения герба будет отличаться от $\frac{1}{2}$ на одну сотую или больше. На самом деле искомую вероятность мы вычислили в примере 37 на стр. 47. Она равна примерно 0,9545. Разница этого ответа с оценкой $\frac{3}{4}$ из неравенства Чебышёва показывает, насколько это неравенство является грубым. Для вычисления вероятностей следует пользоваться более тонкими предельными теоремами.

§ 4. Центральная предельная теорема

Одинаково распределённые слагаемые. Если в условиях ЗБЧ среднее арифметическое $\frac{S_n}{n} = \frac{\xi_1 + \ldots + \xi_n}{n}$ сходится по вероятности к a, то разность $\frac{S_n}{n} - a$ сходится к нулю. Оказывается, что эта разность ведёт себя подобно величине $\frac{\sigma}{\sqrt{n}}\eta$, где η — случайная величина со стандартным нормальным распределением. Иначе говоря, с ростом n *распределение* величины

$$\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - a\right) = \frac{S_n - na}{\sigma\sqrt{n}}$$

становится всё более похожим на *стандартное нормальное распределение.* Это утверждение и называется «центральной предельной теоремой».

Т е о р е м а 31 (центральная предельная теорема). *Пусть $\xi_1, \xi_2, \ldots$ — независимые и одинаково распределённые случайные величины с конечной и ненулевой дисперсией σ^2 и математическим ожиданием a.*

Тогда для любых $x < y$ имеет место сходимость при $n \to \infty$

$$\mathsf{P}\left(x < \frac{S_n - na}{\sigma\sqrt{n}} < y\right) \to \Phi(y) - \Phi(x) = \frac{1}{\sqrt{2\pi}}\int\limits_x^y e^{-t^2/2}dt.$$

Заметим, что дробь $\dfrac{S_n - na}{\sigma\sqrt{n}}$ получилась *стандртизацией* величины S_n: мы просто вычли из S_n её математическое ожидание $\mathsf{E}\,S_n = na$ и поделили эту разность на корень из дисперсии $\sqrt{\mathsf{D}\,S_n} = \sqrt{n\sigma^2} = \sigma\sqrt{n}$.

Можно записать следующие приближённые равенства:

$$\mathsf{P}(x<S_n<y) = \mathsf{P}\left(\frac{x-na}{\sigma\sqrt{n}}<\frac{S_n-na}{\sigma\sqrt{n}}<\frac{y-na}{\sigma\sqrt{n}}\right) \approx \Phi\left(\frac{y-na}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{x-na}{\sigma\sqrt{n}}\right); \quad (25)$$

$$\mathsf{P}\left(x < \frac{S_n}{n} < y\right) = \mathsf{P}(nx < S_n < ny) \approx \Phi\left(\sqrt{n}\,\frac{y-a}{\sigma}\right) - \Phi\left(\sqrt{n}\,\frac{x-a}{\sigma}\right); \quad (26)$$

З а м е ч а н и е 4. Ещё раз напомним, что функция распределения стандартного нормального закона ищется либо по таблице (стр. 121), либо с помощью какого-либо программного обеспечения, но никак не путем нахождения первообразной от плотности нормального стандартного распределения.

П р и м е р 83. При составлении отчёта требуется сложить 10 000 чисел, каждое из которых округлено с точностью до 10^{-5}. Предположим, что ошибки, возникающие при округлении чисел, независимы и равномерно распределены в интервале $(-0{,}5\cdot 10^{-5},\ 0{,}5\cdot 10^{-5})$. Найти границы, в которых с вероятностью 0,99 будет лежать суммарная ошибка.

Р е ш е н и е. Заметим, что ошибки округления симметричны относительно нуля, и суммарная ошибка должна лежать в симметричных относительно нуля границах. Требуется найти такое $x > 0$, что $\mathsf{P}(-x < S_n < x) \approx 0{,}99$, где $n = 10\,000$, $S_n = \xi_1+\ldots+\xi_n$, а слагаемые ξ_i суть независимые случайные величины с равномерным распределением в интервале $(-0{,}5\cdot 10^{-5},\ 0{,}5\cdot 10^{-5})$. Вычислим математическое ожидание и дисперсию слагаемых. Помним, что для равномерного распределения математическое ожидание равно середине отрезка, дисперсия равна одной двенадцатой квадрата длины отрезка:

$$a = \mathsf{E}\,\xi_1 = 0, \quad \sigma^2 = \mathsf{D}\,\xi_1 = \frac{10^{-10}}{12}, \quad \sigma = \frac{1}{2\sqrt{3}\cdot 10^5}.$$

По формуле (25),

$$\mathsf{P}(-x < S_n < x) \approx \Phi\left(\frac{x}{\sigma\sqrt{n}}\right) - \Phi\left(-\frac{x}{\sigma\sqrt{n}}\right) = \Phi\left(2\sqrt{3}\cdot 10^3 x\right) - \Phi\left(-2\sqrt{3}\cdot 10^3 x\right).$$

Чтобы найти x, найдём по таблице функции $\Phi(x)$ или возьмём в примере 77 на стр. 98 такое число $u = 2\sqrt{3}\cdot 10^3 x$, что

$$\Phi(u) - \Phi(-u) = 1 - 2\Phi(-u) = 0{,}99, \quad \text{т. е. } \Phi(-u) = 0{,}005.$$

Число $-u$ равно квантили уровня 0,005 стандартного нормального распределения, поэтому $u = 2{,}576$, $x = \dfrac{2{,}576}{2\sqrt{3}}\cdot 10^{-3} \approx 74\cdot 10^{-5}$.

Итак, суммарная ошибка округления десяти тысяч чисел с вероятностью не менее 0,99 лежит в границах $(-74\cdot 10^{-5},\ 74\cdot 10^{-5})$.

ЦПТ в схеме Бернулли. Пусть случайные величины ξ_i имеют распределение Бернулли и связаны с результатами отдельных испытаний схемы Бернулли, $a = p$, $\sigma = pq$, а величина $S_n = \nu_n$ есть число успехов в n испытаниях схемы Бернулли. Тогда центральная предельная теорема и её следствие (25) превращаются в уже знакомую нам интегральную теорему Муавра — Лапласа (стр. 47).

Теорема 32 (интегральная теорема Муавра — Лапласа). *Пусть ν_n — число успехов в n испытаниях схемы Бернулли с вероятностью успеха p. Тогда для любых $x < y$ имеет место сходимость при $n \to \infty$*

$$\mathsf{P}\left(x < \frac{\nu_n - np}{\sqrt{npq}} < y\right) \to \Phi(y) - \Phi(x),$$

т.е. для любых $k_1 < k_2$ число успехов заключено в границах от k_1 до k_2 с вероятностью

$$\mathsf{P}(k_1 \leqslant \nu_n \leqslant k_2) \;\approx\; \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right).$$

Разнораспределённые слагаемые. Центральной предельной теоремой пользуются также для приближённого вычисления вероятностей, связанных с суммами большого числа независимых, но не обязательно одинаково распределённых величин. Однако слагаемые должны быть не слишком «разными» — они должны вносить примерно одинаковый вклад в сумму, ни одно из слагаемых не должно существенно выделяться. Сформулируем один из возможных вариантов ЦПТ для разнораспределённых случайных величин.

Пусть $\xi_1, \xi_2, \ldots$ — независимые случайные величины с конечными и ненулевыми дисперсиями $\mathsf{D}\,\xi_i = \sigma_i^2$ и математическими ожиданиями $\mathsf{E}\,\xi_i = a_i$.

Говорят, что последовательность $\xi_1, \xi_2, \ldots$ удовлетворяет *условию Ляпунова*, если следующая *дробь Ляпунова*

$$\frac{\mathsf{E}\,|\xi_1 - a_1|^3 + \ldots + \mathsf{E}\,|\xi_n - a_n|^3}{\left(\sqrt{\sigma_1^2 + \ldots + \sigma_n^2}\right)^3}$$

стремится к нулю при $n \to \infty$.

Теорема 33 (ЦПТ Ляпунова). *Пусть $\xi_1, \xi_2, \ldots$ — независимые случайные величины, удовлетворяющие условию Ляпунова. Тогда для любых $x < y$ имеет место сходимость при $n \to \infty$*

$$\mathsf{P}\left(x < \frac{S_n - \mathsf{E}\,S_n}{\sqrt{\mathsf{D}\,S_n}} < y\right) = \mathsf{P}\left(x < \frac{S_n - (a_1 + \ldots + a_n)}{\sqrt{\sigma_1^2 + \ldots + \sigma_n^2}} < y\right) \to \Phi(y) - \Phi(x).$$

Все теоремы, сформулированные в этом разделе, мы оставим без доказательства.

§ 5. Контрольные вопросы

1. Дать определение сходимости по вероятности.

2. Какими свойствами обладает сходимость по вероятности?

3. Пусть $\mathsf{E}\,|\xi| = 1$. Оценить с помощью неравенства Маркова вероятность $\mathsf{P}(|\xi| > 3)$.

4. Какие вероятности позволяет оценивать неравенство Чебышёва?

5. Как по неравенству Чебышёва оценить вероятность $\mathsf{P}(|\xi - \mathsf{E}\,\xi| < x)$, если $x > 0$ и $\mathsf{D}\,\xi$ существует? Будет ли это оценка сверху или снизу?

6. Чем можно оценить вероятность случайной величине отличаться от своего математического ожидания более чем на три корня из дисперсии? Более, чем на четыре? На пять?

7. Каков смысл закона больших чисел?

8. Куда сходятся средние арифметические независимых и одинаково распределённых случайных величин с конечной дисперсией?

9. Как себя ведёт отношение числа успехов в схеме Бернулли к числу испытаний с ростом последнего?

10. Можно ли при каком-нибудь большом числе бросаний правильной монеты гарантировать, что частота выпадения орла отклонится от 0,5 не более, чем на 0,05?

11. Сформулировать ЦПТ. Какие вероятности она позволяет вычислять?

12. К какому распределению в условиях ЦПТ приближается распределение величины $\dfrac{S_n - \mathsf{E}\,S_n}{\sqrt{\mathsf{D}\,S_n}}$?

13. Чему равно математическое ожидание и дисперсия величины $\dfrac{S_n - \mathsf{E}\,S_n}{\sqrt{\mathsf{D}\,S_n}}$?

14. Для каких случайных величин верна ЦПТ Ляпунова?

15. Привести пример задачи, для решения которой необходима ЦПТ.

ГЛАВА IX

ЭЛЕМЕНТЫ ТЕОРИИ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

В этой главе мы рассмотрим последовательности, состоящие, как правило, из зависимых между собой случайных величин с конечным или счётным числом значений. Такую последовательность можно считать процессом, описывающим изменения в состоянии какого-то объекта, если этот объект может находиться в одном из нескольких состояний и за один шаг менять случайным образом своё состояние. Мы рассмотрим здесь только последовательности, обладающие *марковским свойством* — состояние объекта в любой следующий момент времени зависит только от его состояния в настоящий момент времени, но не зависит от того, в каких состояниях объект находился до этого момента. Это свойство читают так: *будущее при известном настоящем не зависит от прошлого.*

§ 1. Цепи Маркова

Пусть мы наблюдаем в отдельные моменты времени какой-то объект, который может находиться в одном из состояний E_k, где $k \in \mathbb{Z}$. Свяжем с этим объектом последовательность целочисленных случайных величин $\xi_0, \xi_1, \xi_2, \ldots$ так: случайная величина ξ_n равна *номеру* состояния наблюдаемого объекта в момент времени n (или в n-ом испытании). Мы переходим от самих состояний к их номерам просто для удобства, и будем говорить «состояние k» вместо «состояние E_k».

П р и м е р 84. В студенческой столовой в качестве первого блюда в разные дни подают щи, борщ, солянку, рассольник, уху, окрошку, харчо и куриный суп. Можно рассматривать объект «первое блюдо в меню». Каждый день этот объект находится в одном из восьми возможных состояний. Можно перейти к случайным величинам ξ_n со значениями $1, 2, \ldots, 8$ так: например, $\xi_3 = 5$, если третьего сентября в столовой уха.

П р и м е р 85. Число клиентов, посетивших банк вчера, сегодня, завтра, послезавтра и т. д. — последовательность целочисленных случайных величин.

П р и м е р 86. Численность (в тысячах человек) населения города в этом году, в предыдущем, в следующем и т. д. — последовательность целочисленных случайных величин.

Описывать динамику изменения состояний объекта будем с помощью вероятностей. Во-первых, мы должны иметь распределение вероятностей для состояния ξ_0 в начальный момент времени: $p_k = \mathsf{P}(\xi_0 = k)$, $\quad \sum p_k = 1$.

Предположим теперь, что мы знаем все состояния $k_0, \ldots, k_{n-1}, i$ объекта в моменты времени $0, \ldots, n-1, n$. В каком состоянии окажется объект в момент времени $n+1$? Это будущее состояние не определяется по предыстории объекта однозначно, а является случайным. Зависимость его от предыстории содержится в вероятностях: при данной предыстории объекта он в момент времени $n+1$ окажется в состоянии j с вероятностью

$$\mathsf{P}(\xi_{n+1} = j \mid \xi_0 = k_0, \ldots, \xi_{n-1} = k_{n-1}, \xi_n = i) = p(k_0, \ldots, k_{n-1}, i, j).$$

Эта вероятность зависит от $n+2$ переменных. Если и далее вероятности переходов зависят от всей предыстории объекта, то описывать поведение объекта становится всё сложнее.

Поэтому рассматривают последовательности $\{\xi_n\}$, для которых динамику задавать проще. Например, такие последовательности, для которых заданные выше вероятности зависят лишь от j (следующее состояние) и от i (состояние в настоящий момент), но не зависят от тех состояний, которые посетил объект до того, как попасть в i. Дадим точное определение.

О п р е д е л е н и е 30. Последовательность целочисленных случайных величин $\xi_0, \xi_1, \xi_2, \ldots$ называется *цепью Маркова* (ЦМ), если для любого момента времени n и для любых состояний $k_0, \ldots, k_{n-1}, i, j$

$$\begin{aligned}\mathsf{P}(\xi_{n+1} = j \mid \xi_0 = k_0, \ldots, \xi_{n-1} = k_{n-1}, \xi_n = i) &= \\ = \mathsf{P}(\xi_{n+1} = j \mid \xi_n = i) &= p_{i,j}(n+1).\end{aligned}$$

Цепь Маркова называется *однородной*, если вероятность $p_{i,j}(n+1)$ перейти за один шаг из состояния i в состояние j на $(n+1)$-ом шаге *не зависит от номера шага* $n+1$: $p_{i,j}(n+1) = p_{i,j}$. Мы будем рассматривать только однородные цепи Маркова.

Распределение величины ξ_0 называют *начальным распределением* цепи Маркова. Динамика любой однородной ЦМ полностью описывается заданием начального распределения $p_k = \mathsf{P}(\xi_0 = k)$ и *вероятностей перехода* $p_{i,j}$ за один шаг из состояния i в состояние j.

Матрица P с элементами $p_{i,j}$ называется *матрицей переходных вероятностей* ЦМ. Эта квадратная матрица не обязательно имеет конечные размеры — её размеры совпадают с числом состояний цепи. Матрица P является *стохастической*, т. е. состоит из неотрицательных чисел, сумма которых в каждой строке равна единице. Действительно, сумма по i-й строке вероятностей перехода из состояния i за один шаг в одно из возможных состояний равна вероятности перейти за шаг хоть куда-нибудь, т. е. единице:

$$P = \|p_{i,j}\|, \quad p_{i,j} \geqslant 0, \quad \sum_j p_{i,j} = \sum_j \mathsf{P}(\xi_1 = j \mid \xi_0 = i) = 1.$$

Обозначим через $p_{i,j}^{(n)}$ вероятность *за n шагов* перейти из состояния i в j:

$$p_{i,j}^{(n)} = \mathsf{P}(\xi_n = i \mid \xi_0 = j) = \mathsf{P}(\xi_{n+m} = i \mid \xi_m = j), \qquad P^{(n)} = \|p_{i,j}^{(n)}\|.$$

Вычислим эту вероятность. Выйдя из состояния i, попасть в j за n шагов можно, если за $(n-1)$ шаг попасть в любое состояние k, и уже из него за один шаг попасть в j:

$$p_{i,j}^{(n)} = \sum_k \mathsf{P}(\xi_{n-1} = k \mid \xi_0 = i) \cdot \mathsf{P}(\xi_n = j \mid \xi_{n-1} = k) = \sum_k p_{i,k}^{(n-1)} p_{k,j}.$$

Мы показали, что (i,j)-й элемент матрицы $P^{(n)}$ получается умножением строки матрицы $P^{(n-1)}$ на столбец матрицы P. Поэтому $P^{(n)} = P^{(n-1)} \cdot P = P^{(n-2)} \cdot P^2 = \ldots = P^n$. Итак, *матрица вероятностей перехода за n шагов есть n-я степень матрицы вероятностей перехода за один шаг.*

Чтобы найти безусловные вероятности цепи Маркова находиться после n шагов в состояниях $j \in \mathbb{Z}$ (распределение величины ξ_n), воспользуемся формулой полной вероятности:

$$\mathsf{P}(\xi_n = j) = \sum_i \mathsf{P}(\xi_0 = i) \cdot \mathsf{P}(\xi_n = j \mid \xi_0 = i) = \sum_i p_i\, p_{i,j}^{(n)}. \tag{27}$$

Пример 87. Лягушка прыгает по точкам 1, 2, 3. В каждой точке она может задержаться на единицу времени с вероятностью $\frac{1}{2}$, либо куда-то перепрыгнуть. Из точек 1 и 3 она обязательно прыгает в точку 2, из точки 2 она может перепрыгнуть в любую из соседних с равными вероятностями.

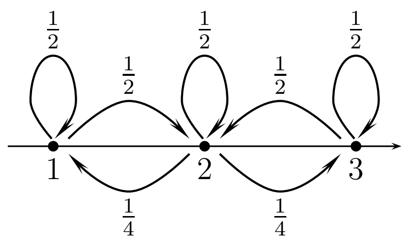

Рис. 21. Блуждание лягушки

Запишем матрицу P вероятностей перехода этой однородной ЦМ и найдём матрицу вероятностей перехода за два шага:

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,1} & p_{3,2} & p_{3,3} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{pmatrix}, \quad P^{(2)} = P^2 = \begin{pmatrix} 3/8 & 1/2 & 1/8 \\ 1/4 & 1/2 & 1/4 \\ 1/8 & 1/2 & 3/8 \end{pmatrix}.$$

Можно вычислить вероятности перехода за 3, 4 и более шагов. Возникает естественный вопрос: пусть лягушка прыгает уже очень давно. Каковы шансы в произвольный момент времени обнаружить её в точке 1? В точке 2?

§ 2. Эргодичность конечной цепи Маркова

Рассмотрим поведение вероятностей $p_{i,j}^{(n)}$ перейти из состояния i в состояние j за n шагов при увеличении n. Предположим, что $p_{i,j}^{(n)}$ при $n \to \infty$ имеют предел π_j, не зависящий от i. Это означает, что вероятность попасть в состояние j через большой интервал времени не зависит от того, с какого состояния цепь стартовала. ЦМ «забыла» своё начальное положение. Величины π_j можно трактовать как вероятности цепи оказаться в состоянии j через очень большой интервал времени. Распределение $\mathsf{P}(\xi = j) = \pi_j$ называется *стационарным распределением* цепи Маркова, а такое поведение вероятностей перехода за n шагов, когда $p_{i,j}^{(n)} \to \pi_j$ при $n \to \infty$, называется *эргодичностью* ЦМ.

Следующее утверждение приводит пример условий, при которых цепь Маркова будет эргодичной. Ограничимся для простоты случаем ЦМ с конечным числом состояний.

Т е о р е м а 34. *Пусть ЦМ $\{\xi_n,\ n \geqslant 0\}$ имеет конечное число состояний и для некоторого n_0 все элементы матрицы $P^{(n_0)}$ (матрицы вероятностей перехода за n_0 шагов) отличны от нуля. Тогда при любых i и j существуют пределы*

$$\lim_{n\to\infty} p_{i,j}^{(n)} = \pi_j > 0.$$

Вероятности π_j являются единственным решением системы уравнений

$$\sum_j \pi_j = 1, \quad \pi_j = \sum_i \pi_i\, p_{i,j}. \tag{28}$$

Система уравнений (28) получается из формулы полной вероятности: вероятность π_j цепи, забывшей о своём начальном распределении, находиться в состоянии j равна сумме вероятностей цепи на предыдущем шаге находиться в одном из состояний i и за один шаг перейти из i в j.

П р и м е р 88. Существует ли в примере 87 стационарное распределение? Условиям теоремы удовлетворяет $n_0 = 2$: вычисленная нами матрица P^2 состоит из ненулевых элементов. Найдём стационарное распределение из системы (28):

$$\begin{cases} \pi_1 + \pi_2 + \pi_3 = 1, & \pi_1 = \frac{1}{2}\pi_1 + \frac{1}{4}\pi_2, \\ \pi_2 = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2 + \frac{1}{2}\pi_3, & \pi_3 = \frac{1}{4}\pi_2 + \frac{1}{2}\pi_3, \end{cases} \qquad \pi_1 = \frac{1}{4}, \ \pi_2 = \frac{1}{2}, \ \pi_3 = \frac{1}{4}.$$

Итак, давно прыгающая лягушка обнаруживается в точках 1 и 3 с вероятностями по $\frac{1}{4}$, и в точке 2 с вероятностью $\frac{1}{2}$.

Система уравнений (28) означает также следующее: если в качестве начального распределения ЦМ взять стационарное: $p_k = \pi_k$, то распределение ЦМ и дальше останется тем же самым стационарным. Действительно, по формуле (27) получим:

$$\mathsf{P}(\xi_1 = j) = \sum_i \mathsf{P}(\xi_0 = i) \cdot \mathsf{P}(\xi_1 = j \mid \xi_0 = i) = \sum_i \pi_i \, p_{i,j} = \pi_j.$$

То есть ЦМ, стартовавшая со стационарным начальным распределением, на каждом шаге имеет то же самое распределение. В этом случае говорят, что цепь работает в стационарном режиме.

Рассмотрим ещё один пример.

Пример 89. Если студент накануне опоздал на занятие, сегодня он имеет 80% шансов прийти вовремя. Если же он вчера пришёл вовремя, то в 40% случаев опоздает сегодня. Как часто студент опаздывает на занятия?

Решение. Имеется цепь Маркова с двумя состояниями: E_1 — студент сегодня опоздал на занятия и E_2 — студент сегодня пришёл вовремя. Матрица вероятностей перехода этой цепи равна

$$P = \begin{pmatrix} 0,2 & 0,8 \\ 0,4 & 0,6 \end{pmatrix}.$$

Все вероятности в этой матрице отличны от нуля, поэтому цепь эргодична. Решая для этой матрицы систему уравнений (28), получим:

$$\pi_1 + \pi_2 = 1, \quad \pi_1 = 0,2\pi_1 + 0,4\pi_2; \qquad \pi_1 = 1/3, \quad \pi_2 = 2/3.$$

Итак, функционирующая в стационарном режиме цепь Маркова с вероятностью $1/3$ пребывает в состоянии 1. Возвращаясь к студенту, находим, что с вероятностью $1/3$ он опаздывает на занятия. Иначе говоря, в среднем он опаздывает на треть занятий.

§ 3. Контрольные вопросы

1. Что такое марковское свойство?
2. Дать определение цепи Маркова.
3. Что означает однородность цепи Маркова?
4. Будет ли цепью Маркова последовательность независимых и одинаково распределённых целочисленных случайных величин? Будет ли она однородной цепью Маркова?
5. Будет ли цепью Маркова последовательность сумм независимых и одинаково распределённых целочисленных случайных величин? Будет ли она однородной цепью Маркова?

6. Будет ли цепью Маркова последовательность произведений независимых и одинаково распределённых целочисленных случайных величин? Будет ли она однородной цепью Маркова?

7. Как по начальному распределению и переходным вероятностям найти распределение однородной ЦМ на любом шаге? Что изменится для неоднородной цепи Маркова?

8. Каким свойством обладают стохастические матрицы?

9. Как ведёт себя ЦМ с единичной матрицей переходных вероятностей?

10. Сколько переходных вероятностей будет у ЦМ с восемью состояниями?

11. Как найти вероятности перехода ЦМ за два шага? За три шага?

12. Что такое эргодичность цепи Маркова?

13. Что называют стационарным распределением ЦМ? Каков его смысл?

14. При каких условиях ЦМ с конечным числом состояний эргодична?

15. Как искать стационарное распределение?

16. Какое распределение имеет ЦМ на каждом шаге, если она стартовала со стационарного начального распределения?

ПРИЛОЖЕНИЕ

Таблица 1. Основные дискретные распределения

Распределение, обозначение, параметры	Возможные значения k	$\mathsf{P}(\xi = k)$	$\mathsf{E}\,\xi$	$\mathsf{D}\,\xi$
Вырожденное I_c, $c \in \mathbb{R}$	c	$\mathsf{P}(\xi = c) = 1$	c	0
Бернулли B_p $p \in (0,\, 1)$	$k = 0,\, 1$	$\mathsf{P}(\xi = 0) = 1-p$, $\mathsf{P}(\xi = 1) = p$	p	$p(1-p)$
Биномиальное $\mathrm{B}_{n,\,p}$ $p \in (0,\, 1)$ $n = 1, 2, \ldots$	$k = 0, \ldots, n$	$C_n^k p^k (1-p)^{n-k}$	np	$np(1-p)$
Пуассона Π_λ $\lambda > 0$	$k = 0,\, 1,\, 2, \ldots$	$\dfrac{\lambda^k}{k!}\, e^{-\lambda}$	λ	λ
Геометрическое G_p $p \in (0,\, 1)$	$k = 1,\, 2, \ldots$	$p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
Гипергеометрическое $n, K, N \in \mathbb{N}$ $0 \leqslant n, K \leqslant N$	целые от $\max(0,\, n+K-N)$ до $\min(n,\, K)$	$\dfrac{C_K^k C_{N_K}^{n-k}}{C_N^n}$	$n\dfrac{K}{N}$	$n\dfrac{K}{N}\left(1-\dfrac{K}{N}\right)\dfrac{N-n}{N-1}$

Таблица 2. Основные абсолютно непрерывные распределения

Распределение, обозначение, параметры	Плотность распределения	$\mathsf{E}\,\xi$	$\mathsf{D}\,\xi$	Асим-метрия	Эксцесс
Равномерное на отрезке $[a,\,b]$, $\mathrm{U}_{a,\,b}$, $a<b$	$\begin{cases}\dfrac{1}{b-a}, & x\in[a,\,b],\\ 0, & x\notin[a,\,b]\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	0	$-1{,}2$
Показательное (экспоненциальное) $\mathrm{E}_\alpha=\Gamma_{\alpha,\,1}$, $\alpha>0$	$\begin{cases}\alpha\, e^{-\alpha x}, & x>0,\\ 0, & x\leqslant 0\end{cases}$	$\dfrac{1}{\alpha}$	$\dfrac{1}{\alpha^2}$	2	6
Нормальное (гауссовское) $\mathrm{N}_{a,\,\sigma^2}$, $a\in\mathbb{R}$, $\sigma>0$	$\dfrac{1}{\sigma\sqrt{2\pi}}\,e^{-(x-a)^2/2\sigma^2}$, $-\infty<x<\infty$	a	σ^2	0	0
Коши $\mathrm{C}_{a,\,\sigma}$, $a\in\mathbb{R}$, $\sigma>0$	$\dfrac{1}{\pi}\,\dfrac{\sigma}{\sigma^2+(x-a)^2}$, $-\infty<x<\infty$	—	—	—	—
Гамма $\Gamma_{\alpha,\,\lambda}$, $\alpha>0,\ \lambda>0$	$\begin{cases}\dfrac{\alpha^\lambda}{\Gamma(\lambda)}\,x^{\lambda-1}e^{-\alpha x}, & x>0,\\ 0, & x\leqslant 0\end{cases}$	$\dfrac{\lambda}{\alpha}$	$\dfrac{\lambda}{\alpha^2}$	$\dfrac{2}{\sqrt{\lambda}}$	$\dfrac{6}{\lambda}$
Лапласа $\mathrm{L}_{\alpha,\,\mu}$, $\alpha>0,\ \mu\in\mathbb{R}$	$\dfrac{\alpha}{2}\,e^{-\alpha\lvert x-\mu\rvert}$, $-\infty<x<\infty$	μ	$\dfrac{2}{\alpha^2}$	0	3
Парето, $\alpha>0$	$\begin{cases}\dfrac{\alpha}{x^{\alpha+1}}, & x\geqslant 1,\\ 0, & x<1\end{cases}$	$\dfrac{\alpha}{\alpha-1}$ $(\alpha>1)$	$\dfrac{\alpha}{(\alpha-1)^2(\alpha-2)}$, $\alpha>2$	$\sqrt{\dfrac{\alpha-2}{\alpha}}\,\dfrac{2(\alpha+1)}{\alpha-3}$, $\alpha>3$	$\dfrac{6(\alpha^3+\alpha^2-6\alpha-2)}{\alpha(\alpha-3)(\alpha-4)}$, $\alpha>4$

Таблица 3. Функция $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt$

(функция распределения стандартного нормального закона)

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
-5	$0,0000003$	$-2,48$	$0,0066$	$-1,91$	$0,0281$	$-1,58$	$0,0571$
$-4,5$	$0,0000034$	$-2,46$	$0,0069$	$-1,9$	$0,0287$	$-1,57$	$0,0582$
-4	$0,0000317$	$-2,44$	$0,0073$	$-1,89$	$0,0294$	$-1,56$	$0,0594$
$-3,8$	$0,0000724$	$-2,42$	$0,0078$	$-1,88$	$0,0301$	$-1,55$	$0,0606$
$-3,6$	$0,0001591$	$-2,4$	$0,0082$	$-1,87$	$0,0307$	$-1,54$	$0,0618$
$-3,4$	$0,0003370$	$-2,38$	$0,0087$	$-1,86$	$0,0314$	$-1,53$	$0,0630$
$-3,2$	$0,0006872$	$-2,36$	$0,0091$	$-1,85$	$0,0322$	$-1,52$	$0,0643$
-3	$0,00135$	$-2,34$	$0,0096$	$-1,84$	$0,0329$	$-1,51$	$0,0655$
$-2,98$	$0,00144$	$-2,32$	$0,0102$	$-1,83$	$0,0336$	$-1,5$	$0,0668$
$-2,96$	$0,00154$	$-2,3$	$0,0107$	$-1,82$	$0,0344$	$-1,49$	$0,0681$
$-2,94$	$0,00164$	$-2,28$	$0,0113$	$-1,81$	$0,0351$	$-1,48$	$0,0694$
$-2,92$	$0,00175$	$-2,26$	$0,0119$	$-1,8$	$0,0359$	$-1,47$	$0,0708$
$-2,9$	$0,00187$	$-2,24$	$0,0125$	$-1,79$	$0,0367$	$-1,46$	$0,0721$
$-2,88$	$0,00199$	$-2,22$	$0,0132$	$-1,78$	$0,0375$	$-1,45$	$0,0735$
$-2,86$	$0,00212$	$-2,2$	$0,0139$	$-1,77$	$0,0384$	$-1,44$	$0,0749$
$-2,84$	$0,00226$	$-2,18$	$0,0146$	$-1,76$	$0,0392$	$-1,43$	$0,0764$
$-2,82$	$0,00240$	$-2,16$	$0,0154$	$-1,75$	$0,0401$	$-1,42$	$0,0778$
$-2,8$	$0,00256$	$-2,14$	$0,0162$	$-1,74$	$0,0409$	$-1,41$	$0,0793$
$-2,78$	$0,00272$	$-2,12$	$0,0170$	$-1,73$	$0,0418$	$-1,4$	$0,0808$
$-2,76$	$0,00289$	$-2,1$	$0,0179$	$-1,72$	$0,0427$	$-1,39$	$0,0823$
$-2,74$	$0,00307$	$-2,08$	$0,0188$	$-1,71$	$0,0436$	$-1,38$	$0,0838$
$-2,72$	$0,00326$	$-2,06$	$0,0197$	$-1,7$	$0,0446$	$-1,37$	$0,0853$
$-2,7$	$0,00347$	$-2,04$	$0,0207$	$-1,69$	$0,0455$	$-1,36$	$0,0869$
$-2,68$	$0,00368$	$-2,02$	$0,0217$	$-1,68$	$0,0465$	$-1,35$	$0,0885$
$-2,66$	$0,00390$	-2	$0,0228$	$-1,67$	$0,0475$	$-1,34$	$0,0901$
$-2,64$	$0,00415$	$-1,99$	$0,0233$	$-1,66$	$0,0485$	$-1,33$	$0,0918$
$-2,62$	$0,00440$	$-1,98$	$0,0239$	$-1,65$	$0,0495$	$-1,32$	$0,0934$
$-2,6$	$0,00466$	$-1,97$	$0,0244$	$-1,64$	$0,0505$	$-1,31$	$0,0951$
$-2,58$	$0,00494$	$-1,96$	$0,0250$	$-1,63$	$0,0516$	$-1,3$	$0,0968$
$-2,56$	$0,00523$	$-1,95$	$0,0256$	$-1,62$	$0,0526$	$-1,29$	$0,0985$
$-2,54$	$0,00554$	$-1,94$	$0,0262$	$-1,61$	$0,0537$	$-1,28$	$0,1003$
$-2,52$	$0,00590$	$-1,93$	$0,0268$	$-1,6$	$0,0548$	$-1,27$	$0,1020$
$-2,5$	$0,00621$	$-1,92$	$0,0274$	$-1,59$	$0,0559$	$-1,26$	$0,1038$

Окончание табл. 3

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
$-1,25$	$0,1056$	$-0,93$	$0,1762$	$-0,61$	$0,2709$	$-0,29$	$0,3859$
$-1,24$	$0,1075$	$-0,92$	$0,1788$	$-0,6$	$0,2743$	$-0,28$	$0,3897$
$-1,23$	$0,1093$	$-0,91$	$0,1814$	$-0,59$	$0,2776$	$-0,27$	$0,3936$
$-1,22$	$0,1112$	$-0,9$	$0,1841$	$-0,58$	$0,2810$	$-0,26$	$0,3974$
$-1,21$	$0,1131$	$-0,89$	$0,1867$	$-0,57$	$0,2843$	$-0,25$	$0,4013$
$-1,2$	$0,1151$	$-0,88$	$0,1894$	$-0,56$	$0,2877$	$-0,24$	$0,4052$
$-1,19$	$0,1170$	$-0,87$	$0,1922$	$-0,55$	$0,2912$	$-0,23$	$0,4090$
$-1,18$	$0,1190$	$-0,86$	$0,1949$	$-0,54$	$0,2946$	$-0,22$	$0,4129$
$-1,17$	$0,1210$	$-0,85$	$0,1977$	$-0,53$	$0,2981$	$-0,21$	$0,4168$
$-1,16$	$0,1230$	$-0,84$	$0,2005$	$-0,52$	$0,3015$	$-0,2$	$0,4207$
$-1,15$	$0,1251$	$-0,83$	$0,2033$	$-0,51$	$0,3050$	$-0,19$	$0,4247$
$-1,14$	$0,1271$	$-0,82$	$0,2061$	$-0,5$	$0,3085$	$-0,18$	$0,4286$
$-1,13$	$0,1292$	$-0,81$	$0,2090$	$-0,49$	$0,3121$	$-0,17$	$0,4325$
$-1,12$	$0,1314$	$-0,8$	$0,2119$	$-0,48$	$0,3156$	$-0,16$	$0,4364$
$-1,11$	$0,1335$	$-0,79$	$0,2148$	$-0,47$	$0,3192$	$-0,15$	$0,4404$
$-1,1$	$0,1357$	$-0,78$	$0,2177$	$-0,46$	$0,3228$	$-0,14$	$0,4443$
$-1,09$	$0,1379$	$-0,77$	$0,2206$	$-0,45$	$0,3264$	$-0,13$	$0,4483$
$-1,08$	$0,1401$	$-0,76$	$0,2236$	$-0,44$	$0,3300$	$-0,12$	$0,4522$
$-1,07$	$0,1423$	$-0,75$	$0,2266$	$-0,43$	$0,3336$	$-0,11$	$0,4562$
$-1,06$	$0,1446$	$-0,74$	$0,2296$	$-0,42$	$0,3372$	$-0,1$	$0,4602$
$-1,05$	$0,1469$	$-0,73$	$0,2327$	$-0,41$	$0,3409$	$-0,09$	$0,4641$
$-1,04$	$0,1492$	$-0,72$	$0,2358$	$-0,4$	$0,3446$	$-0,08$	$0,4681$
$-1,03$	$0,1515$	$-0,71$	$0,2389$	$-0,39$	$0,3483$	$-0,07$	$0,4721$
$-1,02$	$0,1539$	$-0,7$	$0,2420$	$-0,38$	$0,3520$	$-0,06$	$0,4761$
$-1,01$	$0,1562$	$-0,69$	$0,2451$	$-0,37$	$0,3557$	$-0,05$	$0,4801$
-1	$0,1587$	$-0,68$	$0,2483$	$-0,36$	$0,3594$	$-0,04$	$0,4840$
$-0,99$	$0,1611$	$-0,67$	$0,2514$	$-0,35$	$0,3632$	$-0,03$	$0,4880$
$-0,98$	$0,1635$	$-0,66$	$0,2546$	$-0,34$	$0,3669$	$-0,02$	$0,4920$
$-0,97$	$0,1660$	$-0,65$	$0,2578$	$-0,33$	$0,3707$	$-0,01$	$0,4960$
$-0,96$	$0,1685$	$-0,64$	$0,2611$	$-0,32$	$0,3745$	0	$0,5000$
$-0,95$	$0,1711$	$-0,63$	$0,2643$	$-0,31$	$0,3783$		
$-0,94$	$0,1736$	$-0,62$	$0,2676$	$-0,3$	$0,3821$		

При $x > 0$ значения $\Phi(x)$ находят по правилу: $\Phi(x) = 1 - \Phi(-x)$.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

СПИСОК ЛИТЕРАТУРЫ

1. *Бочаров П. П., Печинкин А. В.* Теория вероятностей. Математическая статистика. М.: Гардарика, 1998, 328 с.
2. *Вентцель Е. С.* Теория вероятностей. М.: Высшая школа, 1999, 576 с.
3. *Гнеденко Б. В.* Курс теории вероятностей. М.: Наука, 1988, 406 с.
4. *Колемаев В. А., Калинина В. Н.* Теория вероятностей и математическая статистика. М.: ИНФРА-М, 1997, 302 с.
5. *Пугачев В. С.* Теория вероятностей и математическая статистика. М.: ФИЗМАТЛИТ, 2002, 496 с.
6. *Чистяков В. П.* Курс теории вероятностей. М.: Агар, 2000, 255 с.
7. *Гмурман В. Е.* Руководство к решению задач по теории вероятностей и математической статистике. М.: Высшее образование, 2006, 404 с.
8. Сборник задач по теории вероятностей, математической статистике и теории случайных функций / Под редакцией А. А. Свешникова. М.: Наука, 1970, 656 с.
9. *Вентцель Е. С., Овчаров Л. А.* Задачи и упражнения по теории вероятностей. М.: Высшая школа, 2000, 366 с.

Наталья Исааковна Чернова

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Учебное пособие

Редактор: О. А. Игнатова
Корректор:

Подписано в печать
Формат бумаги $62 \times 84/16$, отпечатано на ризографе, шрифт №10,
изд. л. , зак. № , тир. – экз., СибГУТИ
630102, Новосибирск, ул. Кирова, 86.