ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Механико-математический факультет
Кафедра теории вероятностей и математической статистики

Н. И. Чернова

Теория вероятностей

Учебное пособие

Новосибирск
2007

УДК 519.21
ББК В17я73-2
Ч493

Чернова Н. И. Теория вероятностей: Учеб. пособие / Новосиб. гос. ун-т. Новосибирск, 2007. 160 с.

ISBN 978-5-94356-506-9

Учебное пособие содержит полный курс лекций по теории вероятностей, более 10 лет читаемый автором на первом курсе отделения экономики экономического факультета НГУ. Подбор материала является традиционным для курса теории вероятностей, излагаемого студентам экономических специальностей университетов, и включает основы теории меры, основы комбинаторики, элементарную и аналитическую теорию вероятностей, предельные теоремы теории вероятностей.

Пособие соответствует требованиям Государственного образовательного стандарта к профессиональным образовательным программам по специальности 061800 «Математические методы в экономике».

Предназначено для студентов экономических специальностей.

Рецензент
канд. физ.-мат. наук А. П. Ковалевский

Издание подготовлено в рамках выполнения инновационно-образовательной программы «Инновационные образовательные программы и технологии, реализуемые на принципах партнерства классического университета, науки, бизнеса и государства» национального проекта «Образование».

ISBN 978-5-94356-506-9

ОГЛАВЛЕНИЕ

— Знаете что, милый Арамис? — сказал д'Артаньян, ненавидевший стихи почти так же сильно, как латынь. — Добавьте к достоинству трудности достоинство краткости, и вы сможете быть уверены в том, что ваша поэма будет иметь никак не менее двух достоинств.

А. Дюма. Три мушкетёра

ПРЕДИСЛОВИЕ

Учебное пособие содержит курс лекций по теории вероятностей, читаемый автором студентам 1-го курса отделения экономики экономического факультета НГУ. Курс теории вероятностей опирается на курсы математического анализа и алгебры и служит основой для следующего за ним курса математической статистики, которая, в свою очередь, является основой для одной из главных дисциплин специализации экономистов-математиков — эконометрики.

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

Курс разбит на 12 глав. Первые две главы знакомят читателя с основными понятиями и аксиомами теории вероятностей. Третья глава посвящена условной вероятности и её следствиям. Здесь же обсуждается понятие независимости событий. Схема Бернулли рассмотрена в четвёртой главе. Здесь же встречаются первые предельные теоремы.

Пятая глава знакомит читателя с понятием случайной величины и её распределения. Здесь же дан перечень стандартных семейств распределений и изучаются их свойства. Глава шестая посвящена изучению преобразований случайных величин.

В седьмой 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 восьмая посвящена математическим ожиданиям, дисперсиям и другим важным числовым характеристикам распределений. Математические ожидания вводятся отдельно для дискретного и для абсолютно непрерывного распределения. Коэффициент корреляции обсуждается в главе девятой.

Глава десятая знакомит читателя с видами сходимостей последовательностей случайных величин и законами больших чисел. В одиннадцатой главе излагаются свойства слабой сходимости и центральная предельная теорема для независимых и одинаково распределённых слагаемых. В последней главе рассмотрены характеристические функции и их применение в теории вероятностей.

В конце курса приведён 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.

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

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

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 $\square$.

Автор искренне признателен своим коллегам по кафедре теории вероятностей и математической статистики ММФ НГУ за то влияние, которое они оказали на содержание данного курса и на методические воззрения автора.

ГЛАВА I

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ

…Я употребил обыкновенную хитрость романистов: начал повесть эффектными сценами, вырванными из средины или конца её, прикрыл их туманом. Ты, публика, добра, очень добра, а потому ты неразборчива и недогадлива. На тебя нельзя положиться, что ты с первых страниц можешь различить, будет ли содержание повести стоить того, чтобы прочесть её, у тебя плохое чутьё, оно нуждается в пособии, а пособий этих два: или имя автора, или эффектность манеры.

Н. Г. Чернышевский. Что делать?

§ 1. Элементы комбинаторики

Научимся подсчитывать число «шансов». О числе шансов говорят, когда возможно несколько результатов какого-либо действия (выбор карты из колоды, подбрасывание кубика или монетки). Формулы комбинаторики позволяют посчитать число способов проделать действие или число его возможных результатов.

Теорема о перемножении шансов. 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 1. *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 или с любым другим из k элементов множества A. Таким образом, всего возможно km пар, в которых первый элемент выбран из множества A, а второй — из множества B. □

Упражнение. С помощью теоремы 1 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 $9{\cdot}9{\cdot}8$.

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

На этот вопрос нельзя дать однозначный ответ, пока мы не знаем:

а) как организован выбор;

б) 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 (какие наборы 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 различными наборами.

2. Выбор *без учёта порядка*: два набора номеров шаров считаются различными, если они отличаются составом. Так, наборы $(1,5,2)$ и $(2,5,3)$ различны, а наборы $(1,5,2)$ и $(2,5,1)$ не различаются.

Подсчитаем, сколько возможно различных результатов для каждой из четырёх схем выбора: с возвращением или без возвращения, и в каждом из этих случаев — с учётом порядка или без учёта.

Теорема 2. *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.

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 1, получаем, что общее число возможных наборов из k шаров равно произведению k сомножителей $n(n-1)\cdot\ldots\cdot(n-k+1)$. Здесь последний сомножитель $n-k+1$ есть число способов выбрать k-й шар из оставшихся в урне шаров. □

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 количество 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 восьми разных книг;

в) из алфавита выбирают три разные буквы и составляют слово;

г) из различных ненулевых цифр составляют трёхзначное число;

д) 36 карт в колоде перемешивают и выкладывают на стол в ряд.

Теорема 3. *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. Упорядочить k различных номеров шаров можно $k!$ 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 количество различных результатов в следующих экспериментах:

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

б) из двадцати учеников класса выбирают троих дежурных.

Теорема 4. *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 $n \cdot n \cdot \ldots \cdot n = n^k$. $\square$

Упражнение. Найти количество различных результатов в следующих экспериментах:

а) монету подбрасывают пять раз;

б) пятизначное число составляют из одних нечётных цифр.

в) обезьяна напечатала на машинке слово из десяти букв;

г) составляют слово длиной в 10 символов из нулей и единиц;

д) игральную кость подбрасывают четырежды.

Выбор с возвращением и без учёта порядка. Рассмотрим урну с двумя пронумерованными шарами и перечислим результаты выбора двух шариков из этой урны при выборе с возвращением. Если учитывать порядок, то исходов получится четыре:

$$(1,\ 1),\ (2,\ 2),\ (1,\ 2),\ (2,\ 1).$$

Если порядок не учитывать, то следует объявить два последних исхода одним и тем же результатом эксперимента. Исходов окажется три:

дважды вынут 1-й шар, дважды вынут 2-й шар, вынуты разные шары.

Видим, что в схеме выбора без учёта порядка получилось три различных результата, тогда как при выборе с учётом порядка различных результатов было четыре. Никаким делением на «число каких-нибудь перестановок», которое помогло избавиться от учёта порядка при выборе без возвращения, число 3 из числа 4 получить не удастся.

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

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

Упражнение. Проверить, что при $n = 2$ и $k = 2$ получается ровно три исхода.

Доказательство. Рассмотрим, чем отличаются друг от друга два разных результата такой схемы выбора. Нам не важен порядок следования номеров, т. е. мы учитываем только, *сколько раз* в нашем наборе из k номеров шаров появился каждый номер. Поэтому результат выбора можно представить набором чисел $k_1, \ \ldots, \ k_n$, в котором $k_i \geqslant 0$ — число появлений шара номер i в наборе, $k_1 + \ldots + k_n = k$. Два результата выбора с возвращением и без учёта порядка различаются, если соответствующие им упорядоченные наборы $k_1, \ \ldots, \ k_n$ не совпадают.

Представим себе другой эксперимент, имеющий точно такие же результаты, и посчитаем их количество. Есть n ящиков, в которых размещаются k шаров. Нас интересует только число шаров в каждом ящике. Результатом эксперимента снова является набор чисел $k_1, \ \ldots, \ k_n$, где $k_i \geqslant 0$ равно числу шаров в ящике с номером i, $k_1 + \ldots + k_n = k$.

А теперь изобразим результат такого размещения в виде схемы, в которой вертикальные линии обозначают перегородки между ящиками, а точки — находящиеся в ящиках шары:

$$| \bullet \bullet \bullet | \, | \bullet | \bullet \bullet | \bullet \bullet | \, | \bullet |$$

Мы видим результат размещения девяти шаров по семи ящикам. Первый ящик содержит три шара, второй и шестой ящики пусты, третий ящик содержит один шар, в четвёртом и пятом ящиках лежит по два шара. Переложим один шар из первого ящика во второй и изобразим таким же образом ещё два результата размещения:

$$| \bullet \bullet | \bullet | \bullet | \bullet \bullet | \bullet \bullet | \, | \bullet |$$

$$| \, | \, | \, | \, | \, | \, | \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet |$$

Видим, что все размещения можно получить, меняя между собой шары и перегородки или расставляя k шаров на $n-1+k$ местах. Число $n-1+k$ получается так: у n ящиков есть ровно $n+1$ перегородка, считая крайние, но из них перемещать можно лишь $n-1$ внутреннюю перегородку. Таким образом, имеется $n-1+k$ мест, которые можно занять шарами либо внутренними перегородками. Перебрав все возможные способы расставить k шаров на этих $n-1+k$ местах, переберём и все нужные размещения. Осталось заметить, что по теореме 3 существует $C_{n-1+k}^{k} = C_{n+k-1}^{n-1}$ способов выбрать места для k шаров на $n-1+k$ местах. □

Упражнение. Найти:

а) количество способов разложить число $k \in \mathbb{N}$ в сумму n целых неотрицательных слагаемых, если важен порядок следования слагаемых;

б) число различных производных порядка k функции n переменных;

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

§ 2. События и операции над ними

Пространство элементарных исходов. Основным понятием теории вероятностей является множество всех возможных результатов данного случайного эксперимента.

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* непустое 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.

Замечание. Вообще говоря, можно называть событиями не *любые* подмножества множества Ω, а лишь элементы некоторого набора подмножеств. О смысле такого ограничения мы поговорим позднее.

Итак, элементарный исход — это мельчайший неделимый результат эксперимента, а событие может состоять из одного или 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:

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

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 $D = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$ — на костях выпадет одинаковое число очков.

П р и м е р 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:

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

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 — множество всех точек стола. Оно бесконечно и несчётно. Событием можно назвать, например, попадание центра монеты на лист бумаги, лежащий на столе, в левую или правую половину стола.

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

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

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 $\Omega_1 = \{б,\ ч\}$ — 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-нибудь иначе). Тогда множество $\Omega_2 = \{б_1,\ б_2,\ ч\}$ будет состоять из трёх равновозможных элементарных исходов.

Пример 7. В коробке лежат один чёрный и два белых шара. Из коробки достают наугад два шара. Порядок следования шаров нам безразличен. Занумеруем шары, чтобы элементарные исходы были равновозможными (это может оказаться удобным). Пространство элементарных исходов состоит из трёх элементов:

$$\Omega = \{(б_1, б_2),\ (б_1, ч),\ (б_2, ч)\}.$$

Событие «вынуты два белых шара» включает один исход $\omega_1 = (б_1, б_2)$, а событие «вынуты разноцветные шары» состоит из двух исходов: $\omega_2 = (б_1, ч)$, $\omega_3 = (б_2, ч)$.

Можно, как в примере 6, рассмотреть пространство элементарных исходов, состоящее из двух элементов: $\Omega_1 = \{(б, б),\ (б, ч)\}$ — 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 (рис. 1).

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.

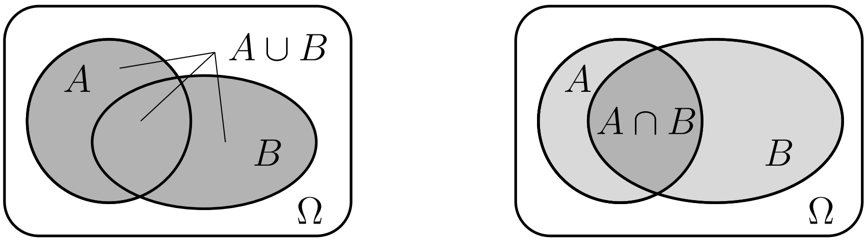

Рис. 1. Объединение и пересечение событий

Дополнением $A \setminus B$ события B до A называется событие, состоящее в том, что произошло 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 A не произошло. Событие $\overline{A}$ состоит из элементарных исходов, не входящих в множество A (рис. 2).

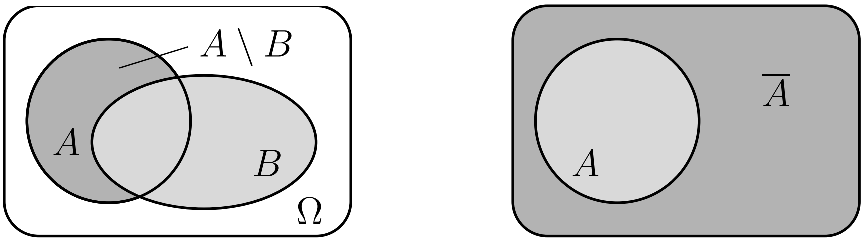

Рис. 2. Дополнение и противоположное событие

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. Кроме того, объединение и пересечение событий связаны очень важными соотношениями двойственности:

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

Пример 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* первые две 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: $C = A_1 \cup A_2 \cup A_3$ и $C = \overline{B}$.

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 в множество B, т. е. A содержится в B (рис. 3).

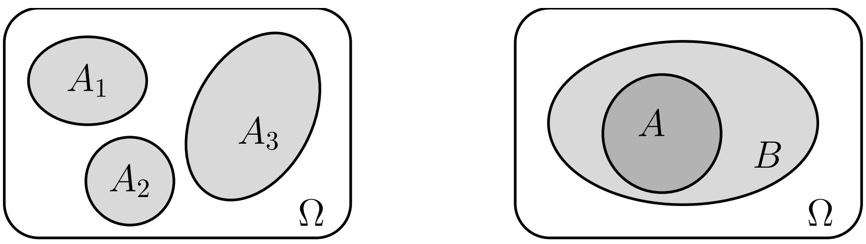

Рис. 3. Попарно несовместные и вложенные события

П р и м е р 9. При бросании двух игральных костей события «сумма очков равна четырём» и «на первой кости выпало шесть очков» несовместны: они не могут случиться одновременно.

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».

§ 3. Дискретное пространство элементарных исходов

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

Так, эксперименты из примеров 1, 2, 3, 5, 6 и 7 (но не 4) приводят к дискретным пространствам элементарных исходов.

З а м е ч а н и е. Множество счётно, если существует взаимно-однозначное соответствие между этим множеством и множеством всех натуральных чисел. Счётными множествами являются множество $\mathbb{N}$ натуральных чисел, множество $\mathbb{Z}$ целых чисел (*доказать*!), множество $\mathbb{Q}$ рациональных чисел (*доказать*!), множество чётных чисел и т. д. Множество конечно, если оно состоит из конечного числа элементов.

Чтобы определить вероятность любого события на таком пространстве, присвоим вероятность каждому элементарному исходу в отдельности, т. е. снабдим вероятностями мельчайшие «кирпичики» — элементарные исходы, из которых составляется любое событие. Вероятность каждого события найдём как сумму вероятностей входящих в него исходов.

О п р е д е л е н и е 3. Сопоставим каждому элементарному исходу ω_i 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 $\mathsf{P}(A) = 0$.

З а м е ч а н и е. Позднее, познакомившись с аксиоматикой теории вероятностей, мы зададим вероятности событий непосредственно, а не через вероятности элементарных исходов. Ведь сложением вероятностей элементарных исходов можно получить лишь вероятность события, состоящего не более чем из счётного числа элементарных исходов (иначе само понятие суммирования не определено). Но на дискретном пространстве элементарных исходов всегда возможно определить вероятности событий согласно определению 3.

П р и м е р 10. В эксперименте из примера 5 монета подбрасывается до первого выпадения герба. Присвоим элементарным исходам следующие вероятности:

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

Проверим, что сумма вероятностей элементарных исходов равна единице: 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:

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

П р и м е р 11. На том же самом множестве $\Omega = \mathbb{N}$ зададим вероятности так: $p_1 = \ldots = p_{100} = 0{,}01$, $p_i = 0$ для $i > 100$.

П р и м е р 12. Пусть теперь $\Omega = \mathbb{N} \cup \{0\}$ — множество целых неотрицательных чисел. Положим

$$p_i = \frac{7^i}{i!}\, e^{-7} \quad \text{для} \quad i = 0, 1, 2, \ldots$$

Проверим, равна ли единице сумма вероятностей всех элементарных исходов. Собрав разложенную в ряд Тейлора экспоненту, получим

$$\sum_{\omega_i \in \Omega} p_i = e^{-7} \cdot \sum_{i=0}^{\infty} \frac{7^i}{i!} = e^{-7} e^{7} = 1.$$

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.

Классическое определение вероятности. 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 из-за симметрии в эксперименте (симметричная монета, хорошо перемешанная колода карт, правильная игральная кость, отсутствие оснований предпочесть один результат эксперимента 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 $\dfrac{1}{N}$. Если событие $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ состоит из k элементарных исходов, то вероятность этого события равна

$$\mathsf{P}(A) = p_{i_1} + \ldots + p_{i_k} = k \cdot \frac{1}{N} = \frac{|A|}{|\Omega|}. \tag{2}$$

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* этому событию, к общему числу равновозможных исходов».

Полезно сравнить это определение с формулировкой автора определения, Я. Бернулли[1]: «Вероятность есть степень достоверности и отличается от неё как часть от целого» (*Искусство предположений*, 1713 г.)

Итак, вычисление вероятности в классической схеме сводится к подсчёту общего числа исходов (шансов) и числа исходов, благоприятствующих событию. Число шансов вычисляют с помощью формул комбинаторики.

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.

[1] Jacob Bernoulli (27.12.1654—16.08.1705, Basel, Switzerland).

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

$$(\textit{герб, герб}),\ (\textit{решка, решка}),\ (\textit{решка, герб}),\ (\textit{герб, решка}).$$

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

$$(\textit{два герба}),\ (\textit{две решки}),\ (\textit{один герб и одна решка}).$$

Первые два исхода имеют вероятности по $\frac{1}{4}$, а вероятность последнего равна $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. 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 $\frac{1}{21}$?

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

Пример 14. Из урны, в которой K белых и $N-K$ чёрных шаров, наудачу и без возвращения вынимают n шаров, где $n \leqslant N$ (рис. 4). Термин «наудачу» означает, что появление любого набора из n шаров равновозможно. Найти вероятность того, что будет выбрано k белых и $n-k$ чёрных шаров.

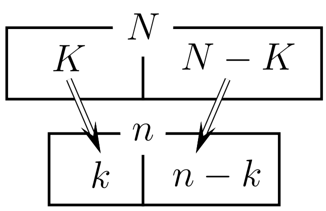

Рис. 4. Выбор n шаров из N

Решение. Результат эксперимента — набор из n шаров. Можно не учитывать порядок следования шаров в наборе. Общее число элементарных исходов по теореме 3 равно $|\Omega| = C_N^n$. 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 $\mathsf{P}(A_k) = 0$.

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. По теореме о перемножении шансов число благоприятных исходов равно $|A_k| = C_K^k\, C_{N-K}^{n-k}$,

$$\mathsf{P}(A_k) = \frac{|A_k|}{|\,\Omega\,|} = \frac{C_K^k\, C_{N-K}^{n-k}}{C_N^n}\,. \tag{3}$$

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 (3) 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 вещественной прямой.

Пример 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.

Решение. Всего возможно $|\Omega| = 5!$ перестановок карточек. Заметим, что перестановка двух карточек с буквой А не меняет слова. Поэтому есть два благоприятных исхода: ЛА_1МПА_2 и ЛА_2МПА_1. Вероятность получить нужное слово равна $\dfrac{2}{5!} = \dfrac{1}{60}$.

Пример 16. Игральная кость подбрасывается трижды. Найти вероятность получить в сумме четыре очка.

Решение. Общее число равновозможных элементарных исходов есть $|\Omega| = 6^3$. Сумма очков равна четырём, если на двух костях выпали единицы, и на одной — двойка. Этому событию благоприятствуют три элементарных исхода: $(1, 1, 2)$, $(1, 2, 1)$, $(2, 1, 1)$. Поэтому искомая вероятность равна $\dfrac{3}{6^3} = \dfrac{1}{72}$.

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 один из способов задания вероятностей на таком пространстве исходов.

§ 4. Геометрическая вероятность

Рассмотрим какую-нибудь область Ω в $\mathbb{R}^k$ (на прямой, на плоскости, в пространстве). Предположим, что «мера» Ω (длина, площадь, объём соответственно) конечна. Пусть случайный эксперимент состоит в том, что мы наудачу бросаем в эту область точку. Термин «наудачу» означает, что вероятность попадания точки в любую часть $A \subseteq \Omega$ не зависит от формы или расположения A внутри Ω. а зависит лишь от «меры» области A.

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

$$\mathsf{P}(A) = \frac{\mu(A)}{\mu(\Omega)}\,. \tag{4}$$

Если для точки, брошенной в область Ω, выполнены условия геометрического определения вероятности, то говорят, что точка *равномерно распределена* в области Ω.

П р и м е р 17. Точка наудачу бросается на отрезок $[0,\,1]$. Вероятность ей попасть в точку $0{,}5$ равна нулю, так как равна нулю мера множества, состоящего из одной точки («длина точки»). Но попадание в точку $0{,}5$ не является невозможным событием — это один из элементарных исходов.

П р и м е р 18. 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 (4), 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.

П р и м е р 19 (з а д а ч а о в с т р е ч е). 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 $[0, 1]$. Обозначим через $\xi \in [0, 1]$ и $\eta \in [0, 1]$ моменты прихода X и Y в течение этого часа (рис. 5). Результатами эксперимента являются всевозможные пары точек (ξ, η) 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:

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

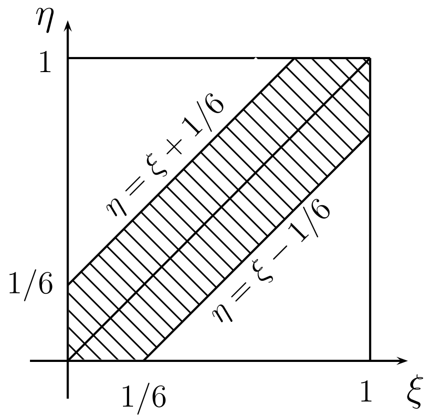

Рис. 5. Задача о встрече

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

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

Существование неизмеримых множеств. Заканчивая обсуждение понятия геометрической вероятности, отметим следующее неприятное обстоятельство.

Если даже эксперимент удовлетворяет геометрическому определению вероятности, далеко не для всех множеств $A \subset \Omega$ вероятность может быть вычислена как отношение меры A к мере Ω. Причиной этого является существование так называемых «неизмеримых» множеств, т. е. множеств, мера которых не существует.

П р и м е р 20 (м н о ж е с т в о В и т а л и[2]). В этом примере мы построим множество на отрезке, «длина» которого не существует. Нам понадобятся лишь следующие очевидные свойства «длины» множества: длина множества остаётся неизменной при сдвиге всех точек этого множества; длина множества, составленного из счётного объединения попарно непересекающихся множеств, равняется сумме длин этих множеств.

Рассмотрим окружность с радиусом 1 (то же, что отрезок $[0, 2\pi]$). Возьмём любое иррациональное число α. Поскольку оно иррационально, число $n\alpha$ не является целым ни при каком целом $n \neq 0$.

---

[2]Giuseppe Vitali (26.08.1875—29.02.1932, Italy).

Поэтому если взять произвольную точку (угол) $x \in [0, 2\pi]$ на окружности и перечислить все точки, которые получаются поворотом x на угол $2\pi n\alpha$, где $n = \pm 1, \pm 2, \ldots$, то мы ни разу не вернёмся в точку x. Точек, получившихся из x такими поворотами, счётное число. Объединим их в один класс. С любой другой точкой окружности можно тоже связать класс точек, получающихся из неё поворотами на $2\pi n\alpha$ при целых n. Эти классы либо совпадают, либо не имеют общих точек. Таким образом, вся окружность разбивается на классы точек. В каждом классе счётное число точек, и все точки в одном классе получаются друг из друга такими поворотами. Разные классы не пересекаются.

Искомое множество A_0 определим так: возьмём из каждого такого класса ровно по одной точке. Пусть множество A_n получается поворотом всех точек множества A_0 на угол $2\pi n\alpha$, $n = \pm 1, \pm 2, \ldots$

Так как все точки одного класса можно получить, поворачивая любую из них на угол $2\pi n\alpha$, $n = \pm 1, \pm 2, \ldots$, а в множестве A_0 собрано по одной точке из каждого класса, то, поворачивая это множество, получим все точки окружности.

Предположим, что «длина» $l(A_0)$ множества A_0 существует. Тогда все множества A_n имеют ту же длину, так как получены из A_0 поворотом. Но все эти множества не пересекаются, поэтому «длина» их объединения равна сумме их длин и равна длине отрезка $[0, 2\pi]$:

$$2\pi = l\left(\bigcup_{n=-\infty}^{\infty} A_n\right) = \sum_{n=-\infty}^{\infty} l(A_n) = \sum_{n=-\infty}^{\infty} l(A_0) = \begin{cases} \infty, & \text{если } l(A_0) > 0, \\ 0, & \text{если } l(A_0) = 0. \end{cases}$$

Полученное противоречие означает, что длина множества A_0 просто *не существует*.

Итак, мы построили множество на отрезке, длина которого не существует (неизмеримое множество). Пользуясь геометрическим определением вероятности, мы *не можем* определить вероятность попадания точки в такое неизмеримое множество. Но если не для любого $A \subseteq \Omega$ мы можем определить вероятность, следует сузить класс множеств, называемых «событиями», оставив в этом классе только те множества, вероятность которых определена.

В следующей главе мы займёмся, следуя Колмогорову[3], построением аксиоматики теории вероятностей: познакомимся с понятиями σ-алгебры (или поля) событий, вероятностной меры и вероятностного пространства.

[3]Андрей Николаевич Колмогоров (25.04.1903—20.10.1987).

ГЛАВА II

АКСИОМАТИКА ТЕОРИИ ВЕРОЯТНОСТЕЙ

Математик должен знать меру, норму и предел

Фольклор ММФ НГУ

§ 1. Алгебра и сигма-алгебра событий

Алгебра событий. Пусть Ω — пространство элементарных исходов некоторого случайного эксперимента (т. е. непустое множество произвольной природы). Мы собираемся определить набор подмножеств Ω, которые будут называться событиями, и затем задать вероятность как функцию, определённую *только* на множестве событий.

Итак, событиями мы будем называть не любые подмножества Ω, а лишь элементы некоторого выделенного набора подмножеств множества Ω. При этом необходимо позаботиться, чтобы этот набор подмножеств был *замкнут* относительно обычных операций над событиями, т. е. чтобы объединение, пересечение, дополнение событий снова давало событие. Сначала введём понятие алгебры множеств.

Определение 4. Множество $\mathcal{A}$, элементами которого являются подмножества множества Ω (не обязательно все) называется *алгеброй*, если оно удовлетворяет следующим условиям:

(A1) $\Omega \in \mathcal{A}$ (алгебра содержит достоверное событие);

(A2) если $A \in \mathcal{A}$, то $\overline{A} \in \mathcal{A}$ (вместе с любым множеством алгебра содержит противоположное к нему);

(A3) если $A \in \mathcal{A}$ и $B \in \mathcal{A}$, то $A \cup B \in \mathcal{A}$ (вместе с любыми двумя множествами алгебра содержит их объединение).

Из (A1) и (A2) следует, что пустое множество $\varnothing = \overline{\Omega}$ также содержится в $\mathcal{A}$, т. е. алгебра содержит и невозможное событие.

Из условия (A3) следует, что вместе с любым *конечным* набором множеств алгебра содержит их объединение: для любого $n \in \mathbb{N}$, для любых $A_1, \ldots, A_n \in \mathcal{A}$ выполнено $A_1 \cup \ldots \cup A_n \in \mathcal{A}$.

Вместо замкнутости относительно объединения можно требовать замкнутость относительно пересечения.

Свойство 1. *В определении 4 можно заменить* (A3) *на* (A4):

(A4) *если* $A \in \mathcal{A}$ *и* $B \in \mathcal{A}$, *то* $A \cap B \in \mathcal{A}$.

Доказательство. Докажем, что при выполнении (A1) и (A2) из (A3) следует (A4). Если A, $B \in \mathcal{A}$, то $\overline{A} \in \mathcal{A}$ и $\overline{B} \in \mathcal{A}$ по свойству (A2). Тогда из (A3) следует, что $\overline{A} \cup \overline{B} \in \mathcal{A}$. Вновь пользуясь (A2), получим, что дополнение $\overline{\overline{A} \cup \overline{B}}$ к этому множеству также принадлежит алгебре $\mathcal{A}$. В силу формул двойственности, дополнение к объединению равно пересечению дополнений:

$$A \cap B = \overline{\overline{A} \cup \overline{B}} \in \mathcal{A}.$$

Аналогично доказывается, что при выполнении (A1) и (A2) из (A4) следует (A3), т. е. эти два свойства в определении взаимозаменяемы. $\square$

Пример 21. Пусть $\Omega = \{\spadesuit, \clubsuit, \diamondsuit, \heartsuit\}$ — пространство элементарных исходов. Следующие наборы подмножеств Ω являются алгебрами (*проверьте это по определению*!):

а) $\mathcal{A} = \{\Omega, \varnothing\} = \{\{\spadesuit, \clubsuit, \diamondsuit, \heartsuit\}, \varnothing\}$ — *тривиальная* алгебра;

б) $\mathcal{A} = \{\{\spadesuit, \clubsuit, \diamondsuit, \heartsuit\}, \varnothing, \{\diamondsuit\}, \{\spadesuit, \clubsuit, \heartsuit\}\}$;

в) $\mathcal{A} = 2^{\Omega}$ — множество всех подмножеств Ω.

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

Сигма-алгебра событий. В теории вероятностей часто возникает необходимость объединять счётные наборы событий и считать событием результат такого объединения. При этом свойства (A3) алгебры оказывается недостаточно: из него не вытекает, что объединение счётной последовательности множеств из алгебры снова принадлежит алгебре. Поэтому разумно наложить более суровые ограничения на класс событий.

Определение 5. Множество $\mathfrak{F}$, элементами которого являются подмножества множества Ω (не обязательно все) называется σ-*алгеброй* (σ-алгеброй событий), если выполнены следующие условия:

(S1) $\Omega \in \mathfrak{F}$ (σ-алгебра событий содержит достоверное событие);

(S2) если $A \in \mathfrak{F}$, то $\overline{A} \in \mathfrak{F}$ (вместе с любым событием σ-алгебра содержит противоположное событие);

(S3) если $A_1, A_2, \ldots \in \mathfrak{F}$, то $A_1 \cup A_2 \cup \ldots \in \mathfrak{F}$ (вместе с любым *счётным* набором событий σ-алгебра содержит их объединение).

Упражнение. Доказать, что вместо (S1) достаточно предположить непустоту множества $\mathcal{F}$. Вывести из (S1) и (S2), что $\varnothing \in \mathcal{F}$.

Этого набора аксиом достаточно для замкнутости множества $\mathcal{F}$ относительно счётного числа любых других операций над событиями. В частности, аналогично свойству 1 проверяется следующее утверждение.

Свойство 2. *В определении 5 можно заменить* (S3) *на* (S4)*:*
(S4) *если* $A_1,\ A_2,\ \ldots \in \mathcal{F}$, *то* $A_1 \cap A_2 \cap \ldots \in \mathcal{F}$.

Свойство 3. *Всякая* σ*-алгебра является алгеброй.*

Доказательство. Пусть $\mathcal{F}$ — σ-алгебра. Нужно проверить, что она удовлетворяет свойству (A3), т. е. для любых $A \in \mathcal{F}$ и $B \in \mathcal{F}$ выполняется $A \cup B \in \mathcal{F}$.

Превратим пару A, B в счётную последовательность событий так: A, B, B, B, $\ldots$, т. е. положим $A_1 = A$, $A_i = B$ при всех $i \geqslant 2$. Объединение $A \cup B$ совпадает с объединением всех множеств A_i из этой бесконечной последовательности. А так как $\mathcal{F}$ — σ-алгебра, то

$$A \cup B = \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}. \qquad \square$$

Итак, всякая σ-алгебра автоматически является алгеброй, но не наоборот. Приведём пример алгебры, не являющейся σ-алгеброй.

Пример 22. Пусть $\Omega = \mathbb{R}$, и пусть $\mathcal{A}$ — множество, содержащее любые конечные подмножества $\mathbb{R}$ (т. е. состоящие из конечного числа точек, в том числе пустое) и их дополнения. Так, множество $\{0,\ 2,\ \pi\}$ принадлежит $\mathcal{A}$, множество $(-\infty,\ -7{,}2) \cup (-7{,}2,\ 5) \cup (5,\ \infty)$ принадлежит $\mathcal{A}$.

Легко проверить, что множество $\mathcal{A}$ является алгеброй. Действительно, пустое множество и само $\Omega = \mathbb{R}$ там содержатся, дополнение к любому конечному подмножеству множества вещественных чисел содержится в $\mathcal{A}$ по определению, дополнение к множеству вида $\mathbb{R} \setminus A$ для конечных A совпадает с A и также принадлежит $\mathcal{A}$ по определению. Свойство (A3) проверяется непосредственно: объединение любых конечных множеств снова конечно и поэтому принадлежит $\mathcal{A}$. Объединение конечного множества с множеством вида $\mathbb{R} \setminus A$, где A конечно, есть снова множество вида $\mathbb{R} \setminus B$, где B конечно (или пусто) и т. д.

Однако алгебра $\mathcal{A}$ не содержит ни одного счётного множества точек. Действительно, объединяя конечные множества в конечном числе, мы можем получить только конечное множество. Например, натуральный ряд $\mathbb{N}$ не принадлежит $\mathcal{A}$. Поэтому $\mathcal{A}$ не является σ-алгеброй: для бесконечной, но счётной последовательности одноточечных множеств $A_i = \{i\}$ из $\mathcal{A}$ их объединение $\mathbb{N} = A_1 \cup A_2 \cup \ldots$ не принадлежит $\mathcal{A}$.

Все алгебры из примера 21 являются σ-алгебрами, поскольку содержат лишь конечное число элементов. Вообще, на конечном множестве Ω понятия алгебры и σ-алгебры совпадают. Множество всех подмножеств Ω является σ-алгеброй для любого Ω (*проверить*!).

Борелевская[4] σ-алгебра. Приведём пример σ-алгебры, которая нам будет необходима в дальнейшем,— σ-алгебры *бо́релевских* множеств на вещественной прямой.

Борелевской сигма-алгеброй в $\mathbb{R}$ называется *самая маленькая* среди всех возможных σ-алгебр, содержащих *любые* интервалы на прямой. Разумеется, σ-алгебры, содержащие все интервалы, существуют. Например, множество всех подмножеств $\mathbb{R}$ — это σ-алгебра, и она содержит все интервалы. Что же такое «самая маленькая σ-алгебра» из нескольких данных? Обратимся к примерам.

Пусть $\Omega = \mathbb{R}$ — вещественная прямая. Рассмотрим некоторые наборы множеств, не являющиеся σ-алгебрами, и увидим, как их можно дополнить до σ-алгебр.

П р и м е р 23. Множество $\mathfrak{A} = \{\mathbb{R}, \varnothing, [0, 1], \{0\}\}$ не является σ-алгеброй, так как, например, $\overline{\{0\}} = \mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty) \notin \mathfrak{A}$. Самый маленький набор множеств, содержащий $\mathfrak{A}$ и являющийся σ-алгеброй (*минимальная* σ-алгебра), получится, если включить в него всевозможные объединения, пересечения и дополнения множеств из $\mathfrak{A}$: $\mathcal{F} = \{\mathbb{R}, \varnothing, [0, 1], \{0\}, (-\infty, 0) \cup (1, \infty), (0, 1], (-\infty, 0] \cup (1, \infty), (-\infty, 0) \cup (0, \infty)\}$.

О п р е д е л е н и е 6. Минимальной σ-алгеброй, содержащей набор множеств $\mathfrak{A}$, называется пересечение всех σ-алгебр, содержащих $\mathfrak{A}$.

Ещё раз напомним, что пересекать в определении 6 есть что: хотя бы одна σ-алгебра, содержащая данный набор множеств, всегда найдётся — это σ-алгебра всех подмножеств Ω.

У п р а ж н е н и е. Доказать, что пересечение *двух* σ-алгебр, содержащих набор множеств $\mathfrak{A}$, снова является σ-алгеброй, содержащей $\mathfrak{A}$.

У п р а ж н е н и е. Найти минимальную σ-алгебру, содержащую следующий набор подмножеств $\mathbb{R}$: $\mathfrak{A} = \{\mathbb{R}, \varnothing, [0, 1], \{3\}\}$.

Дадим определение борелевской сигма-алгебры. Пусть по-прежнему $\Omega = \mathbb{R}$, а множество $\mathfrak{A}$ состоит из *всевозможных* открытых интерва-

[4] Félix Edouard Justin Emile Borel (7.01.1871—3.02.1956, France).

лов $(a,\, b)$, где $a < b$: $\mathfrak{A} = \{(a,\, b) \mid -\infty < a < b < \infty\}$. Это множество всех интервалов не является ни алгеброй, ни σ-алгеброй.

Определение 7. Минимальная σ-алгебра, содержащая множество $\mathfrak{A}$ всех интервалов на вещественной прямой, называется *борелевской* σ-алгеброй в $\mathbb{R}$ и обозначается $\mathfrak{B}(\mathbb{R})$.

Перечислим некоторые множества на прямой, содержащиеся в $\mathfrak{B}(\mathbb{R})$ по определению. Таковы все привычные нам множества. Чтобы получить множество, не содержащееся в $\mathfrak{B}(\mathbb{R})$, требуются специальные построения. Итак, мы знаем, что все интервалы на прямой принадлежат $\mathfrak{B}(\mathbb{R})$, и $\mathfrak{B}(\mathbb{R})$ — σ-алгебра. Отсюда сразу следует, что $\mathfrak{B}(\mathbb{R})$ содержит любое множество, которое можно получить из интервалов с помощью счётного числа операций объединения или пересечения, а также взятием дополнения.

В частности, $\mathbb{R} \in \mathfrak{B}(\mathbb{R})$ по свойству (S1). Далее, все одноточечные множества $\{x\}$, где $x \in \mathbb{R}$, принадлежат $\mathfrak{B}(\mathbb{R})$. Действительно, интервалы $\left(x - \dfrac{1}{n},\, x + \dfrac{1}{n}\right)$ принадлежат $\mathfrak{B}(\mathbb{R})$, по определению, при любом n. Их счётное пересечение также принадлежит $\mathfrak{B}(\mathbb{R})$ по свойству (S4):

$$\{x\} = \bigcap_{n=1}^{\infty} \left(x - \frac{1}{n},\, x + \frac{1}{n}\right) \in \mathfrak{B}(\mathbb{R}).$$

Далее, любой интервал вида $(a,\, b\,]$ (или $[a,\, b)$, или $[a,\, b\,]$), где $a < b$, принадлежит $\mathfrak{B}(\mathbb{R})$ как объединение открытого интервала и точки (или двух точек): $(a,\, b\,] = (a,\, b) \cup \{b\}$.

Упражнение. Докажите, что множество натуральных чисел $\mathbb{N}$ и множество рациональных чисел $\mathbb{Q}$ принадлежат $\mathfrak{B}(\mathbb{R})$.

Борелевская σ-алгебра в $\mathbb{R}^n$ строится совершенно так же, как в $\mathbb{R}$. Это должна быть минимальная σ-алгебра, содержащая все множества вида $(a_1,\, b_1) \times \ldots \times (a_n,\, b_n)$ — уже не интервалы, как в $\mathbb{R}$, а прямоугольники в $\mathbb{R}^2$, параллелепипеды в $\mathbb{R}^3$ и т. д. Вместе с ними $\mathfrak{B}(\mathbb{R}^n)$ содержит любые множества, являющиеся «предельными» для объединений измельчающихся прямоугольников. Например, круг в $\mathbb{R}^2$ является борелевским множеством — можно изнутри или снаружи приблизить его объединениями прямоугольников.

Итак, мы определили специальный класс $\mathcal{F}$ подмножеств Ω, названный σ-алгеброй событий. Применение счётного числа любых операций к множествам из $\mathcal{F}$ снова дает множество из $\mathcal{F}$, т. е. не выводит за рамки этого класса. *Событиями* будем называть только множества $A \in \mathcal{F}$.

Определим теперь понятие *вероятности* как функции, определённой на множестве событий (функции, которая каждому событию ставит в соответствие число — вероятность этого события).

А чтобы читателю сразу стало понятно, о чём пойдёт речь, добавим: вероятность мы определим как *неотрицательную нормированную меру*, заданную на σ-алгебре $\mathcal{F}$ подмножеств Ω. Следующий параграф познакомит нас с понятиями меры и вероятностной меры.

§ 2. Мера и вероятностная мера

Определение 8. Пусть Ω — некоторое непустое множество, $\mathcal{F}$ — σ-алгебра его подмножеств. Функция $\mu : \mathcal{F} \to \mathbb{R} \cup \{+\infty\}$ называется *мерой* на $(\Omega, \mathcal{F})$, если она удовлетворяет условиям:

(μ1) $\mu(A) \geqslant 0$ для любого множества $A \in \mathcal{F}$;

(μ2) для любого счётного набора попарно непересекающихся множеств $A_1, A_2, A_3, \ldots \in \mathcal{F}$ (т. е. такого, что $A_i \cap A_j = \varnothing$ при всех $i \neq j$) мера их объединения равна сумме их мер:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

(«счётная аддитивность» или «σ-аддитивность» меры).

Упражнение. Зачем в свойстве (μ2) требуется, чтобы события не пересекались? Может ли какая-нибудь функция $\mu : \mathcal{F} \to \mathbb{R}$ удовлетворять свойству $\mu(A \cup B) = \mu(A) + \mu(B)$ при любых событиях A и B? Привести пример такой функции и доказать, что других не существует.

Упражнение. Указать область определения и область значений функции μ. Для каких $A \subset \Omega$ определено значение $\mu(A)$?

Пример 24. Пусть $\Omega = \{a, b, c\}$, $\mathcal{F} = 2^{\Omega}$ — множество всех подмножеств Ω. Зададим меру μ на $\mathcal{F}$ так: $\mu\{a\} = 3$, $\mu\{b\} = 17$, $\mu\{c\} = 1$, $\mu\{a, b\} = 20$, $\mu\{a, c\} = 4$, $\mu\{b, c\} = 18$, $\mu\{a, b, c\} = 21$, $\mu(\varnothing) = 0$. Для краткости записи мы вместо $\mu(\{a\})$ писали всюду $\mu\{a\}$.

Пример 25. Пусть $\Omega = \mathbb{N}$, $\mathcal{F} = 2^{\mathbb{N}}$ — множество всех подмножеств натурального ряда. Зададим меру μ на $\mathcal{F}$ так: $\mu(A) = |A|$ — число элементов в множестве A (бесконечность, если множество A бесконечно).

Пример 26 (мера Лебега[5]). Когда мы говорили о геометрической вероятности, мы использовали термин «мера области A в $\mathbb{R}^m$», имея в виду «длину» на прямой, «площадь» на плоскости, «объём» в трёхмер-

[5] Henri Léon Lebesgue (28.06.1875—26.07.1941, France).

ном пространстве. Являются ли все эти «длины-площади-объёмы» настоящими мерами в смысле определения 8? Мы решим этот вопрос для прямой, оставляя плоскость и пространство большей размерности читателю.

Рассмотрим вещественную прямую с σ-алгеброй борелевских множеств. Эта σ-алгебра, по определению, есть наименьшая σ-алгебра, содержащая все интервалы. Для каждого интервала (a, b) число $b-a$ назовём *длиной* интервала (a, b).

Мы не станем доказывать следующее утверждение:

Теорема 6. *Существует единственная мера* λ *на* $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, *значение которой на любом интервале равно его длине:* $\lambda(a, b) = b - a$. *Эта мера называется мерой Лебега*.

Нам пригодится свойство, которым обладает любая мера. Это свойство *непрерывности меры* иногда называют *аксиомой непрерывности*, имея в виду, что ею можно заменить $(\mu 2)$ в определении 8.

Теорема 7 (свойство непрерывности меры). *Пусть дана убывающая последовательность* $B_1 \supseteq B_2 \supseteq B_3 \supseteq \ldots$ *множеств из* $\mathfrak{F}$, *причём* $\mu(B_1) < \infty$. *Пусть* $B = \bigcap\limits_{n=1}^{\infty} B_n$. *Тогда* $\mu(B) = \lim\limits_{n\to\infty} \mu(B_n)$.

Доказательство. Обозначим через C_n кольца: $C_n = B_n \setminus B_{n+1}$. Множества B, C_1, C_2, $\ldots$ попарно не пересекаются. Тогда из представлений

$$B_1 = B \cup \left(\bigcup_{i=1}^{\infty} C_i\right), \quad B_n = B \cup \left(\bigcup_{i=n}^{\infty} C_i\right)$$

вытекают, в силу аксиомы $(\mu 2)$, соответствующие равенства и для мер:

$$\mu(B_1) = \mu(B) + \sum_{i=1}^{\infty} \mu(C_i), \qquad \mu(B_n) = \mu(B) + \sum_{i=n}^{\infty} \mu(C_i).$$

Первая сумма $\sum\limits_{i=1}^{\infty} \mu(C_i)$ в силу условия $\mu(B_1) < \infty$ есть сумма абсолютно сходящегося ряда (составленного из неотрицательных слагаемых). Из сходимости этого ряда следует, что «хвост» ряда, равный $\sum\limits_{i=n}^{\infty} \mu(C_i)$, стремится к нулю при $n \to \infty$. Поэтому

$$\mu(B_n) = \mu(B) + \sum_{i=n}^{\infty} \mu(C_i) \underset{n\to\infty}{\longrightarrow} \mu(B) + 0 = \mu(B). \qquad \square$$

В полезности этого свойства легко убедиться упражнениями.

Упражнение. Используя аксиому непрерывности меры для убывающей последовательности множеств $B_n = (x - 1/n,\, x + 1/n)$, доказать, что мера Лебега одноточечного подмножества $\{x\}$ вещественной прямой равна нулю: $\lambda\{x\} = 0$. Используя этот факт, доказать, что $\lambda(\mathbb{N}) = 0$, $\lambda(\mathbb{Z}) = 0$, $\lambda(\mathbb{Q}) = 0$, $\lambda(a,\, b) = \lambda[a,\, b]$.

Замечание. В отсутствие предположения $\mu(B_1) < \infty$ свойство $\mu(B) = \lim\limits_{n\to\infty} \mu(B_n)$ может не выполняться.

Например, зададим меру на $\mathfrak{B}(\mathbb{R})$ так: $\mu(B) = 0$, если B не более чем счётно, иначе $\mu(B) = \infty$. Тогда для множеств $B_n = (x - 1/n,\, x + 1/n)$ имеем:

$$B = \bigcap_{n=1}^{\infty} B_n = \{x\}, \quad \mu(B_n) = \infty \not\to \mu(B) = 0.$$

Наконец, мы в состоянии определить понятие *вероятности* как нормированной меры.

Определение 9. Пусть Ω — непустое множество, $\mathcal{F}$ — σ-алгебра его подмножеств. Мера $\mu : \mathcal{F} \to \mathbb{R}$ называется *нормированной*, если $\mu(\Omega) = 1$. Другое название нормированной меры — *вероятность*.

То же самое ещё раз и подробно:

Определение 10. Пусть Ω — пространство элементарных исходов, $\mathcal{F}$ — σ-алгебра его подмножеств (событий). *Вероятностью* или *вероятностной мерой* на $(\Omega,\, \mathcal{F})$ называется функция $\mathsf{P} : \mathcal{F} \to \mathbb{R}$, обладающая свойствами:

(P1) $\mathsf{P}(A) \geqslant 0$ для любого события $A \in \mathcal{F}$;

(P2) для любого счётного набора попарно несовместных событий $A_1,\, A_2,\, A_3,\, \ldots \in \mathcal{F}$ имеет место равенство

$$\mathsf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathsf{P}(A_i);$$

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

Свойства (P1) — (P3) называют *аксиомами вероятности*.

Определение 11. Тройка $\langle \Omega, \mathcal{F}, \mathsf{P} \rangle$, в которой Ω — пространство элементарных исходов, $\mathcal{F}$ — σ-алгебра его подмножеств и P — вероятностная мера на $\mathcal{F}$, называется *вероятностным пространством*.

Докажем свойства вероятности, вытекающие из аксиом. Ниже мы не будем всякий раз оговаривать, что имеем дело только с событиями.

Теорема 8. *Вероятность обладает следующими свойствами.*

1. $\mathsf{P}(\varnothing) = 0$.

2. *Для любого* ***конечного*** *набора попарно несовместных событий* $A_1, \ldots, A_n \in \mathcal{F}$ *имеет место равенство:*

$$\mathsf{P}(A_1 \cup \ldots \cup A_n) = \mathsf{P}(A_1) + \ldots + \mathsf{P}(A_n).$$

3. $\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A)$.

4. *Если* $A \subseteq B$, *то* $\mathsf{P}(B \setminus A) = \mathsf{P}(B) - \mathsf{P}(A)$.

5. *Если* $A \subseteq B$, *то* $\mathsf{P}(A) \leqslant \mathsf{P}(B)$ (*монотонность вероятности*).

6. $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$.

7. $\mathsf{P}(A_1 \cup \ldots \cup A_n) \leqslant \sum\limits_{i=1}^{n} \mathsf{P}(A_i)$.

8. *Формула включения-исключения*:

$$\mathsf{P}(A_1 \cup \ldots \cup A_n) = \sum_{i=1}^{n} \mathsf{P}(A_i) - \sum_{i<j} \mathsf{P}(A_i A_j) \; + \\ + \sum_{i<j<m} \mathsf{P}(A_i A_j A_m) - \ldots + (-1)^{n-1} \mathsf{P}(A_1 A_2 \ldots A_n). \qquad (5)$$

Доказательство. 1. События $A_1 = \Omega$, $A_i = \varnothing$, где $i \geqslant 2$, попарно несовместны, и их объединение есть Ω. По аксиоме (P2),

$$1 = \mathsf{P}(\Omega) = \sum_{i=1}^{\infty} \mathsf{P}(A_i) = 1 + \sum_{i=2}^{\infty} \mathsf{P}(\varnothing).$$

Это возможно только в случае $\mathsf{P}(\varnothing) = 0$.

2. Положим $A_i = \varnothing$ при любом $i > n$. События $A_1, \ldots, A_n, \varnothing, \varnothing, \ldots$ попарно несовместны, и по аксиоме (P2),

$$\mathsf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \mathsf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathsf{P}(A_i) = \sum_{i=1}^{n} \mathsf{P}(A_i).$$

3. Достоверное событие $\Omega = A \cup \overline{A}$ есть объединение двух несовместных событий A и $\overline{A}$. По свойству 2 получим $1 = \mathsf{P}(\Omega) = \mathsf{P}(A) + \mathsf{P}(\overline{A})$.

4 и 5. Событие B равно объединению двух несовместных событий: $B = A \cup (B \setminus A)$. Согласно свойству 2, $\mathsf{P}(B) = \mathsf{P}(A) + \mathsf{P}(B \setminus A) \geqslant \mathsf{P}(A)$.

6. Событие $A \cup B$ можно разложить в объединение двух несовместных событий $A \cup B = A \cup (B \setminus AB)$, причём $AB \subseteq B$. По свойствам 2 и 4 получим $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B \setminus AB) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB)$.

7. При $n = 2$ неравенство вытекает из свойства 6:

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

Упражнение. Докажите свойство 7 и формулу (5) с помощью математической индукции. □

Приведём пример задачи, в которой использование формулы включения-исключения — самый простой путь решения.

Пример 27 (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 предназначенный ему конверт.

Решение. 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 используем формулу (5). По классическому определению вероятности вычислим вероятности всех событий A_i и их пересечений. Элементарными исходами будут всевозможные перестановки n писем по n конвертам. Их общее число есть $|\Omega| = n!$, и событию A_i благоприятны $(n-1)!$ из них, а именно перестановки всех писем, кроме i-го, лежащего в своём конверте. Поэтому $\mathsf{P}(A_i) = \dfrac{(n-1)!}{n!} = \dfrac{1}{n}$ — одна и та же для всех i. Точно так же

$$\mathsf{P}(A_iA_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}, \quad \mathsf{P}(A_iA_jA_m) = \frac{1}{n(n-1)(n-2)} \quad \text{и т. д.}$$

Вычислим количество слагаемых в каждой сумме в формуле (5). 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 (5), получаем:

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

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

ГЛАВА III

УСЛОВНАЯ ВЕРОЯТНОСТЬ И НЕЗАВИСИМОСТЬ

...Здесь является вопрос ... относительно влияния прошлого на вероятность будущего.

П. Лаплас. Опыт философии теории вероятностей

§ 1. Условная вероятность

Пример 28. 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 $\mathsf{P}(A \mid B)$.

Определение 12. Условной вероятностью события A при условии, что произошло событие B, называется число

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

Условная вероятность определена только в случае, когда $\mathsf{P}(B) > 0$.

Следует отличать условную вероятность одного события при осуществлении другого от вероятности им одновременно произойти.

Это определение бывает полезно использовать не для вычисления условной вероятности, а для последовательного вычисления вероятности нескольким событиям случиться одновременно, если известны соответствующие условные вероятности. Справедливы следующие «теоремы умножения вероятностей».

Теорема 9. *Если* $\mathsf{P}(B) > 0$ *и* $\mathsf{P}(A) > 0$, *то*

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

Теорема 10. *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.

Упражнение. Доказать теорему 10 методом математической индукции. Доказать, что все условные вероятности в теореме 10 определены тогда и только тогда, когда $\mathsf{P}(A_1 \ldots A_{n-1}) > 0$.

§ 2. Независимость событий

Определение 13. События A и B называются *независимыми*, если $\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$.

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

Решение. Вероятность вытянуть туза равна $\mathsf{P}(A) = \dfrac{4}{36} = \dfrac{1}{9}$. Вероятность вытянуть пиковую карту равна $\mathsf{P}(B) = \dfrac{1}{4}$. Пересечение этих событий означает появление туза пик и имеет вероятность $\mathsf{P}(AB) = \dfrac{1}{36}$. События A и B независимы, так как $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$.

Естественно считать события A и B независимыми, когда условная вероятность A при условии, что B произошло, остаётся такой же, как и безусловная. Убедимся, что этим свойством обладают события, независимые согласно определению 13.

Свойство 4. *Пусть* $\mathsf{P}(B) > 0$. *Тогда события* A *и* B *независимы тогда и только тогда, когда* $\mathsf{P}(A \mid B) = \mathsf{P}(A)$.

Упражнение. Доказать по определению условной вероятности.

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.

Свойство 5. *Пусть события* A *и* B *несовместны. Тогда независимыми они будут только в том случае, если* $\mathsf{P}(A) = 0$ *или* $\mathsf{P}(B) = 0$.

Это свойство (*а вы его доказали?*) означает, что в невырожденном случае (когда вероятности событий положительны) несовместные события не могут быть независимыми. Зависимость между ними — просто причинно-следственная: если $A \cap B = \varnothing$, то $A \subseteq \overline{B}$, т. е. при выполнении A событие B *не происходит*.

Упражнение. Доказать с помощью свойства монотонности вероятности, что событие A, вероятность которого равна нулю или единице, не зависит ни от какого события B, в том числе и от самого себя.

Свойство 6. *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 $A = (A \cap B) \cup (A \cap \overline{B})$, и события $A \cap B$ и $A \cap \overline{B}$ несовместны, то $\mathsf{P}(A) = \mathsf{P}(A \cap B) + \mathsf{P}(A \cap \overline{B})$. Поэтому $\mathsf{P}(A \cap \overline{B}) = \mathsf{P}(A) - \mathsf{P}(A \cap B) = \mathsf{P}(A) - \mathsf{P}(A)\mathsf{P}(B) = \mathsf{P}(A)(1 - \mathsf{P}(B)) = \mathsf{P}(A)\mathsf{P}(\overline{B})$. Остальные утверждения вытекают из первого. □

Если у нас не два, а большее число событий, выполнение только одного равенства $\mathsf{P}(A_1 \cap \ldots \cap A_n) = \mathsf{P}(A_1) \cdot \ldots \cdot \mathsf{P}(A_n)$ вовсе не означает независимости этих событий. Например, при таком равенстве события A_1 и A_2 вполне могут оказаться зависимыми.

Пример 30. Пусть $0 < \mathsf{P}(A) < 1$. События $A_1 = A_2 = A$, $A_3 = \emptyset$ обладают свойством

$$0 = \mathsf{P}(A_1 \cap A_2 \cap A_3) = \mathsf{P}(A_1) \cdot \mathsf{P}(A_2) \cdot \mathsf{P}(A_3) = 0,$$

что не мешает событиям A_1 и A_2 быть зависимыми:

$$\mathsf{P}(A_1 \cap A_2) = \mathsf{P}(A) \neq \mathsf{P}(A_1) \cdot \mathsf{P}(A_2) = \big(\mathsf{P}(A)\big)^2.$$

Хотелось бы независимостью нескольких событий считать такое свойство, при котором любые комбинации этих событий будут независимы между собой: например, независимы $A_1 \cap A_2$ и $A_3 \cup A_4 \cup A_5$.

Определение 14. События $A_1, \ldots, A_n$ называются независимыми *в совокупности*, если для любого $1 \leqslant k \leqslant n$ и любого набора различных меж собой индексов $1 \leqslant i_1, \ldots, i_k \leqslant n$ имеет место равенство

$$\mathsf{P}(A_{i_1} \cap \ldots \cap A_{i_k}) = \mathsf{P}(A_{i_1}) \cdot \ldots \cdot \mathsf{P}(A_{i_k}). \tag{6}$$

Замечание. Если события $A_1, \ldots, A_n$ независимы в совокупности, то они попарно независимы, т. е. любые два события A_i и A_j независимы. Достаточно в равенстве (6) взять $k = 2$. Обратное, как показывает следующий пример, неверно: из попарной независимости не вытекает независимость в совокупности.

Пример 31 (пример Бернштейна[6]). 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 (соответственно 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.

Заметьте, что равенство (6) выполнено при $k = 2$, но не при $k = 3$.

§ 3. Формула полной вероятности

Пример 32. Есть три завода, производящих одну и ту же продукцию. Первый завод производит 25 %, второй завод — 35, третий — 40 % всей производимой 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.

Первая вероятность равна доле бракованных изделий в объеме всей продукции, т. е. $0{,}05 \cdot 0{,}25 + 0{,}03 \cdot 0{,}35 + 0{,}04 \cdot 0{,}4$. Вторая вероятность равна доле брака первого завода среди всего брака, т. е.

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

Определение 15. Конечный или счётный набор попарно несовместных событий $H_1, H_2, \ldots$ таких, что $\mathsf{P}(H_i) > 0$ для всех i и $H_1 \cup H_2 \cup \ldots = \Omega$, называется *полной группой событий* или *разбиением* пространства Ω.

События $H_1, H_2, \ldots$, образующие полную группу событий, часто называют *гипотезами.* При подходящем выборе гипотез для любого события A могут быть сравнительно просто вычислены $\mathsf{P}(A \mid H_i)$ и собственно $\mathsf{P}(H_i)$. Как, используя эти данные, посчитать вероятность события A?

[6]Сергей Натанович Бернштейн (5.03.1880—26.10.1968).

Т е о р е м а 11 (ф о р м у л а п о л н о й в е р о я т н о с т и). *Пусть дана полная группа событий* $H_1,\ H_2,\ \ldots$ *Тогда вероятность любого события* A *может быть вычислена по формуле*

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

Д о к а з а т е л ь с т в о. Заметим, что

$$A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^{\infty} H_i \right) = \bigcup_{i=1}^{\infty} (A \cap H_i),$$

и события $A \cap H_1,\ A \cap H_2,\ \ldots$ попарно несовместны. Поэтому

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

Во втором равенстве мы использовали σ-аддитивность вероятностной меры (*а что это*?), а в третьем — теорему 9 умножения вероятностей. □

§ 4. Формула Байеса[7]

Т е о р е м а 12 (ф о р м у л а Б а й е с а). *Пусть* $H_1, H_2, \ldots$ — *полная группа событий, и* A — *некоторое событие, вероятность которого положительна. Тогда условная вероятность того, что имело место событие* H_k, *если в результате эксперимента наблюдалось событие* A, *может быть вычислена по формуле*

$$\mathsf{P}(H_k\,|\,A) = \frac{\mathsf{P}(H_k)\,\mathsf{P}(A\,|\,H_k)}{\sum\limits_{i=1}^{\infty} \mathsf{P}(H_i)\,\mathsf{P}(A\,|\,H_i)}.$$

Д о к а з а т е л ь с т в о. По определению условной вероятности,

$$\mathsf{P}(H_k\,|\,A) = \frac{\mathsf{P}(H_k \cap A)}{\mathsf{P}(A)} = \frac{\mathsf{P}(H_k)\,\mathsf{P}(A\,|\,H_k)}{\sum\limits_{i=1}^{\infty} \mathsf{P}(H_i)\,\mathsf{P}(A\,|\,H_i)}. \qquad \square$$

П р и м е р 33. Вернёмся к примеру 32. Изделие выбирается наудачу из всей произведённой продукции. Рассмотрим три гипотезы: $H_i = \{\text{изделие изготовлено } i\text{-м заводом}\}$, $i = 1, 2, 3$. Вероятности этих событий даны: $\mathsf{P}(H_1) = 0{,}25$, $\mathsf{P}(H_2) = 0{,}35$, $\mathsf{P}(H_3) = 0{,}4$.

Пусть $A = \{\text{изделие оказалось бракованным}\}$. Даны также условные вероятности $P(A\,|\,H_1) = 0{,}05$, $P(A\,|\,H_2) = 0{,}03$, $P(A\,|\,H_3) = 0{,}04$.

[7] Thomas Bayes (1702—17.04.1761, England).

Убедитесь, что полученные нами в примере 32 вероятности совпадают с вероятностями, вычисленными по формуле полной вероятности и по формуле Байеса.

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.

П р и м е р 34. 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

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

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.

ГЛАВА IV

СХЕМА БЕРНУЛЛИ

На дне глубокого сосуда лежат спокойно n шаров.
Поочерёдно их оттуда таскают двое дураков.
Сие занятье им приятно: они таскают t минут,
И, вынув шар, его обратно в сосуд немедленно кладут.
Ввиду условия такого, сколь вероятность велика,
Что первый был глупей второго, когда шаров он вынул k?

В. П. Скитович

§ 1. Распределение числа успехов в n испытаниях

Определение 16. *Схемой Бернулли* называется последовательность независимых в совокупности испытаний, в каждом из которых возможны лишь два исхода — «успех» и «неудача», при этом успех в одном испытании происходит с вероятностью $p \in (0,1)$, а неудача — с вероятностью $q = 1 - p$.

Под независимостью в совокупности *испытаний* понимается независимость в совокупности любых событий, относящихся к разным испытаниям. В испытаниях схемы Бернулли, когда с одним испытанием можно связать только два взаимоисключающих события, независимость в совокупности испытаний означает, что при любом n независимы в совокупности события $A_1 = \{$успех в первом испытании$\}$, $\ldots$, $A_n = \{$успех в n-м испытании$\}$. Эти события принадлежат одному и тому же пространству элементарных исходов, полученному декартовым произведением бесконечного числа двухэлементных множеств $\{у, н\}$:

$$\Omega = \{(a_1, a_2, \ldots) \mid a_i \in \{у, н\}\}.$$

Здесь буквами «$у$» и «$н$» обозначены успешный и неудачный результаты испытаний соответственно.

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

Теорема 13 (формула Бернулли). *При любом* $k = 0, 1, \ldots, n$ *имеет место равенство:*

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

Доказательство. Событие $A = \{\nu_n = k\}$ означает, что в n испытаниях схемы Бернулли произошло ровно k успехов. Рассмотрим один из благоприятствующих событию A элементарных исходов:

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

когда первые k испытаний завершились успехом, остальные неудачей. Поскольку испытания независимы, вероятность такого элементарного исхода равна $p^k q^{n-k}$. Другие благоприятствующие событию A элементарные исходы 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 также равна $p^k q^{n-k}$. □

Определение 17. Набор чисел $\left\{C_n^k p^k q^{n-k}, \, k = 0, 1, \ldots, n\right\}$ называется *биномиальным* распределением вероятностей.

§ 2. Номер первого успешного испытания

Рассмотрим схему Бернулли с вероятностью успеха p в одном испытании. Введём величину τ со значениями $1, 2, 3, \ldots$, равную *номеру первого успешного испытания*.

Теорема 14. *Вероятность того, что первый успех произойдёт в испытании с номером* $k \in \mathbb{N} = \{1, 2, 3, \ldots\}$, *равна* $\mathsf{P}(\tau = k) = pq^{k-1}$.

Доказательство. Вероятность первым $k - 1$ испытаниям завершиться неудачей, а последнему — успехом, равна

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

Определение 18. Набор чисел $\{pq^{k-1}, \ \ k = 1, 2, 3, \ldots\}$ называется *геометрическим* распределением вероятностей.

Геометрическое распределение вероятностей обладает интересным свойством, которое можно назвать свойством «нестарения».

Теорема 15. *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 (измеряемым целым числом часов) некоторого устройства, то данному равенству можно придать следующее звучание: *вероятность работающему устройству проработать ещё сколько-то часов не зависит от того момента, когда мы начали отсчёт времени, или от того, сколько уже работает устройство*. Общепринятое название этого свойства — свойство *отсутствия последействия.*

Доказательство. По определению условной вероятности,

$$\mathsf{P}(\tau > n + k \,|\, \tau > n) = \frac{\mathsf{P}(\tau > n + k,\, \tau > n)}{\mathsf{P}(\tau > n)} = \frac{\mathsf{P}(\tau > n + k)}{\mathsf{P}(\tau > n)}. \tag{7}$$

Последнее равенство следует из того, что событие $\{\tau > n + k\}$ влечёт событие $\{\tau > n\}$, поэтому пересечение этих событий есть $\{\tau > n + k\}$. Найдём для целого $m \geqslant 0$ вероятность $\mathsf{P}(\tau > m)$: событие $\{\tau > m\}$ означает в точности, что в схеме Бернулли первые m испытаний завершились неудачами, т. е. его вероятность равна q^m. Возвращаясь к (7), получим

$$\mathsf{P}(\tau > n + k \,|\, \tau > n) = \frac{\mathsf{P}(\tau > n + k)}{\mathsf{P}(\tau > n)} = \frac{q^{n+k}}{q^n} = q^k = \mathsf{P}(\tau > k).$$

Теорема 15 доказана. □

§ 3. Независимые испытания с несколькими исходами

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

Пример 35. Игральная кость подбрасывается 15 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.

Теорема 16. *Для любого n и любых неотрицательных целых чисел $n_1, \ldots, n_m$, сумма которых равна n, верна формула*

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

Доказательство. 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 35 и выписать ответ: вероятность получить десять троек, три единицы и ещё два других очка равна

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

так как вероятности выпадения тройки и единицы равны по $1/6$, а вероятность третьего исхода (выпала любая другая грань) равна $4/6$.

§ 4. Теорема Пуассона[8] для схемы Бернулли

Предположим, нам нужна вероятность получить не менее семи успехов в тысяче испытаний схемы Бернулли с вероятностью успеха 0,003. Вероятность этого события равна любому из следующих выражений, вычислить которые довольно сложно:

$$\sum_{k=7}^{1\,000} C_{1\,000}^{k}\, (0{,}003)^k\, (0{,}997)^{1\,000-k} = 1 - \sum_{k=0}^{6} C_{1\,000}^{k}\, (0{,}003)^k\, (0{,}997)^{1\,000-k}.$$

Сформулируем теорему о приближённом вычислении вероятности иметь k успехов в большом числе испытаний Бернулли с маленькой вероятностью успеха p. Термин «большое число» должен означать $n \to \infty$. Если при этом p остаётся неизменной, то вероятность получить любое

[8] Siméon Denis Poisson (21.06.1781—25.04.1840, France).

заданное число успехов уменьшается до нуля. Необходимо чтобы вероятность успеха $p = p_n$ уменьшалась одновременно с ростом числа испытаний. Но от испытания к испытанию вероятность успеха меняться не может (см. определение схемы Бернулли). Поэтому нам придётся рассмотреть так называемую «схему серий»: если испытание одно, то вероятность успеха в нём равна p_1, если испытаний два, то вероятность успеха в каждом равна p_2 и т. д. Вероятность успеха меняется не внутри одной серии испытаний, а от серии к серии, когда меняется общее число испытаний.

Теорема 17 (теорема Пуассона). *Пусть $n \to \infty$ и $p_n \to 0$ так, что $np_n \to \lambda > 0$. Тогда для любого $k \geqslant 0$ вероятность получить k успехов в n испытаниях схемы Бернулли с вероятностью успеха p_n стремится к величине $\lambda^k\, e^{-\lambda} / k!$:*

$$\mathsf{P}(\nu_n = k) = C_n^k\, p_n^k\, (1 - p_n)^{n-k} \;\to\; \frac{\lambda^k}{k!}\, e^{-\lambda}.$$

Доказательство. Положим $\lambda_n = np_n$. По условию $\lambda_n \to \lambda > 0$. Подставим $p_n = \lambda_n / n$ в формулу Бернулли:

$$\begin{aligned} C_n^k\, p_n^k\, (1 - p_n)^{n-k} &= C_n^k \frac{\lambda_n^k}{n^k} \left(1 - \frac{\lambda_n}{n}\right)^{n-k} = \\ &= \underbrace{\frac{n(n-1)\dots(n-k+1)}{n^k}}_{\substack{\downarrow \\ 1}} \frac{\lambda_n^k}{k!} \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^{n}}_{\substack{\downarrow \\ e^{-\lambda}}} \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^{-k}}_{\substack{\downarrow \\ 1}} \longrightarrow \frac{\lambda^k}{k!}\, e^{-\lambda}. \end{aligned} \tag{8}$$

В соотношении (8) мы воспользовались тем, что $\lambda_n^k \to \lambda^k$ и замечательным пределом $(1 - \lambda_n / n)^n \to e^{-\lambda}$. □

Определение 19. Набор чисел $\left\{ \dfrac{\lambda^k}{k!}\, e^{-\lambda},\ k = 0, 1, 2, \dots \right\}$ называется *распределением Пуассона* с параметром $\lambda > 0$.

По теореме 17 можно приближённо посчитать вероятность получить не менее семи успехов в тысяче испытаний схемы Бернулли с вероятностью успеха 0,003, с вычисления которой мы начали. Поскольку $n = 1\,000$ «велико», а $p_n = 0{,}003$ «мало», то, взяв $\lambda = np_n = 3$, можно записать приближённое равенство

$$1 - \sum_{k=0}^{6} C_{1\,000}^k\, (0{,}003)^k\, (0{,}997)^{1\,000-k} \approx 1 - \sum_{k=0}^{6} \frac{3^k}{k!}\, e^{-3} \approx 0{,}034. \tag{9}$$

Осталось решить, а достаточно ли $n = 10^3$ велико, а $p_n = 0{,}003$ мало, чтобы заменить точную вероятность на её приближённое значение. Для этого нужно уметь оценивать разницу между этими вероятностями.

Следующую очень полезную теорему мы, исключительно из экономии времени, доказывать не станем.

Теорема 18 (уточнённая теорема Пуассона). *Пусть A — произвольное множество целых неотрицательных чисел, ν_n — число успехов в n испытаниях схемы Бернулли с вероятностью успеха p, $\lambda = np$, Справедливо неравенство*

$$\left|\mathsf{P}(\nu_n \in A) - \sum_{k\in A} \frac{\lambda^k}{k!}\, e^{-\lambda}\right| = \left|\sum_{k\in A} C_n^k p^k (1-p)^{n-k} - \sum_{k\in A} \frac{\lambda^k}{k!}\, e^{-\lambda}\right| \leqslant \min(p,\, np^2).$$

Таким образом, теорема 18 предоставляет нам возможность самим решать, достаточно ли n велико, а p мало, руководствуясь полученной величиной погрешности. Какова же погрешность в формуле (9)? Взяв $A = \{0, 1, \ldots, 6\}$, имеем

$$\bigl|\,\mathsf{P}(\nu_{1000} \geqslant 7) - 0{,}034\,\bigr| = \left|\Bigl(1 - \mathsf{P}(\nu_{1000} \leqslant 6)\Bigr) - \left(1 - \sum_{k=0}^{6} \frac{3^k}{k!}\, e^{-3}\right)\right| =$$

$$= \left|\mathsf{P}(\nu_{1000} \leqslant 6) - \sum_{k=0}^{6} \frac{3^k}{k!}\, e^{-3}\right| \leqslant \min(p,\, np^2) = 0{,}003.$$

Таким образом, можно утверждать, что искомая вероятность заключена в границах $(0{,}034 - 0{,}003,\ 0{,}034 + 0{,}003) = (0{,}031,\ 0{,}037)$.

ГЛАВА V
СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И ИХ РАСПРЕДЕЛЕНИЯ

То, что остаётся после всех этих абстракций, не следует ли… считать тем реальным и неизменным содержанием, которое навязывается существам всех видов с одинаковой необходимостью, потому что оно не зависит ни от индивида, ни от момента времени, ни от точки зрения?..

А. Рей. Современная философия. 1908

§ 1. Случайные величины

Мы уже видели, что для многих экспериментов нет никаких различий в подсчёте *вероятностей* событий, тогда как элементарные исходы в этих экспериментах очень различаются. Но нас и должны интересовать именно вероятности событий, а не структура пространства элементарных исходов. Поэтому пора во всех таких «похожих» экспериментах вместо самых разных исходов использовать, например, числа. Иначе говоря, каждому элементарному исходу поставить в соответствие некоторое вещественное число, и работать только с числами.

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

Определение 20. Функция $\xi : \Omega \to \mathbb{R}$ называется *случайной величиной*, если для любого борелевского множества $B \in \mathfrak{B}(\mathbb{R})$ множество $\xi^{-1}(B)$ является событием, т. е. принадлежит σ-алгебре $\mathcal{F}$.

Множество $\xi^{-1}(B) = \{\omega \mid \xi(\omega) \in B\}$, состоящее из тех элементарных исходов ω, для которых $\xi(\omega)$ принадлежит B, называется полным прообразом множества B.

Замечание. Вообще, пусть функция f действует из множества X в множество Y, и заданы σ-алгебры $\mathcal{F}$ и $\mathcal{G}$ подмножеств X и Y соответственно. Функция f называется *измеримой*, если для любого множества $B \in \mathcal{G}$ его полный прообраз $f^{-1}(B)$ принадлежит $\mathcal{F}$.

Замечание. Читатель, не желающий забивать себе голову абстракциями, связанными с σ-алгебрами событий и с измеримостью, может смело считать, что любое множество элементарных исходов есть событие, и, следовательно, случайная величина есть произвольная функция из Ω в $\mathbb{R}$. Неприятностей на практике это не влечёт, так что всё дальнейшее в этом параграфе можно пропустить.

Теперь, избавившись от нелюбопытных читателей, попробуем понять, зачем случайной величине нужна измеримость.

Если задана случайная величина ξ, нам может потребоваться вычислить вероятности вида $\mathsf{P}(\xi = 5) = \mathsf{P}\{\omega \,|\, \xi(\omega) = 5\}$, $\mathsf{P}(\xi \in [-3, 7])$, $\mathsf{P}(\xi \geqslant 3{,}2)$, $\mathsf{P}(\xi < 0)$ и вообще самые разные вероятности попадания в борелевские множества на прямой. Это возможно лишь если множества, стоящие под знаком вероятности, являются событиями — ведь вероятность есть функция, определённая только на σ-алгебре событий. Требование измеримости равносильно тому, что для любого борелевского множества B определена вероятность $\mathsf{P}(\xi \in B)$.

Можно потребовать в определении 20 чего-нибудь другого. Например, чтобы событием было попадание в любой интервал: $\{\omega \mid \xi(\omega) \in (a, b)\} \in \mathcal{F}$, или в любой полуинтервал: $\{\omega \mid \xi(\omega) < x\} \in \mathcal{F}$.

О п р е д е л е н и е 21. Функция $\xi : \Omega \to \mathbb{R}$ называется случайной величиной, если для любых вещественных $a < b$ множество $\{\omega \,|\, \xi(\omega) \in (a,\, b)\}$ принадлежит σ-алгебре $\mathcal{F}$.

Д о к а з а т е л ь с т в о. Докажем эквивалентность определений 20 и 21. Если ξ — случайная величина в смысле определения 20, то она будет случайной величиной и в смысле определения 21, поскольку любой интервал $(a,\, b)$ является борелевским множеством.

Докажем, что верно и обратное. Пусть для любого интервала $B = (a,\, b)$ выполнено $\xi^{-1}(B) \in \mathcal{F}$. Мы должны доказать, что то же самое верно для любых борелевских множеств. Соберём в множестве $\mathcal{A} = \{B \subseteq \mathbb{R} \mid \xi^{-1}(B) \in \mathcal{F}\}$ все такие подмножества B вещественной прямой, что их прообразы являются событиями. По определению, $B \in \mathcal{A}$ тогда и только тогда, когда множество $\xi^{-1}(B)$ принадлежит $\mathcal{F}$.

Множество $\mathcal{A}$ уже содержит все интервалы $(a,\, b)$. Покажем, что множество $\mathcal{A}$ является σ-алгеброй.

1. Убедимся, что $\mathbb{R} \in \mathcal{A}$. Но $\xi^{-1}(\mathbb{R}) = \Omega \in \mathcal{F}$ и, следовательно, $\mathbb{R} \in \mathcal{A}$.

2. Убедимся, что $\overline{B} \in \mathcal{A}$ для любого $B \in \mathcal{A}$. Пусть $\xi^{-1}(B) \in \mathcal{F}$. Тогда $\xi^{-1}(\overline{B}) = \{\omega \mid \xi(\omega) \notin B\} = \Omega \setminus \xi^{-1}(B) \in \mathcal{F}$, так как $\mathcal{F}$ — σ-алгебра.

3. Убедимся, что $B_1 \cup B_2 \cup \ldots \in \mathcal{A}$ для любых $B_1, B_2, \ldots \in \mathcal{A}$. Пусть $\xi^{-1}(B_i) \in \mathfrak{F}$ для всех $i \geqslant 1$. Но $\mathfrak{F}$ — σ-алгебра, поэтому

$$\xi^{-1}\Big(\bigcup_{i=1}^{\infty} B_i\Big) = \Big\{\omega \ \Big|\ \xi(\omega) \in \bigcup_{i=1}^{\infty} B_i\Big\} = \bigcup_{i=1}^{\infty} \xi^{-1}(B_i) \in \mathfrak{F}.$$

Мы доказали, что $\mathcal{A}$ — σ-алгебра и содержит все интервалы на прямой. Но $\mathfrak{B}(\mathbb{R})$ — наименьшая из σ-алгебр, содержащих все интервалы на прямой. Следовательно, $\mathfrak{B}(\mathbb{R}) \subseteq \mathcal{A}$. □

Приведём примеры измеримых и неизмеримых функций.

Пример 36. Подбрасываем кубик. Пусть $\Omega = \{1, 2, 3, 4, 5, 6\}$ и две функции из Ω в $\mathbb{R}$ заданы так: $\xi(\omega) = \omega$, $\eta(\omega) = \omega^2$.

Пока не задана σ-алгебра $\mathfrak{F}$, нельзя говорить об измеримости. Функция, измеримая относительно какой-то σ-алгебры $\mathfrak{F}$, может не быть таковой для другой $\mathfrak{F}$.

1. Если $\mathfrak{F}$ есть множество *всех* подмножеств Ω, то ξ и η являются случайными величинами, поскольку любое множество элементарных исходов принадлежит $\mathfrak{F}$, в том числе и $\{\omega \,|\, \xi(\omega) \in B\}$ или $\{\omega \,|\, \eta(\omega) \in B\}$. Можно записать соответствие между значениями случайных величин ξ и η и вероятностями принимать эти значения в виде *таблицы распределения вероятностей* или, короче, *таблицы распределения*:

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

η	1	4	9	16	25	36
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Здесь $\mathsf{P}(\xi = 1) = \ldots = \mathsf{P}(\xi = 6) = \mathsf{P}(\eta = 1) = \ldots = \mathsf{P}(\eta = 36) = \frac{1}{6}$.

2. Пусть σ-алгебра событий $\mathfrak{F}$ состоит из четырёх множеств:

$$\mathfrak{F} = \big\{\Omega,\ \varnothing,\ \{1, 3, 5\},\ \{2, 4, 6\}\big\}, \tag{10}$$

т. е. событием является, кроме достоверного и невозможного событий, выпадение чётного или нечётного числа очков. Убедимся, что при такой сравнительно бедной σ-алгебре ни ξ, ни η не являются случайными величинами. Возьмём, скажем, $B = \{4\}$. Видим, что $\{\omega \,|\, \xi(\omega) = 4\} = \{4\} \notin \mathfrak{F}$ и $\{\omega \,|\, \eta(\omega) = 4\} = \{2\} \notin \mathfrak{F}$.

Упражнение: а) описать все функции, измеримые относительно σ-алгебры $\mathfrak{F}$ из (10);

б) доказать, что ξ и η не являются случайными величинами, если $\mathfrak{F} = \{\Omega, \varnothing\}$ — тривиальная σ-алгебра;

в) доказать, что относительно тривиальной σ-алгебры измеримы только функции вида $\xi(\omega) = c$ (постоянные).

Пример 37. Пусть $\Omega = [0,\, 2\pi]$, $\mathcal{F} = \mathfrak{B}(\mathbb{R}) \cap [0,\, 2\pi]$ — сигма-алгебра борелевских подмножеств отрезка $[0,\, 2\pi]$, $\mathsf{P}(B) = \lambda(B)/2\pi$ — геометрическая вероятность на $\mathcal{F}$ и A — неизмеримое множество Витали, построенное нами в примере 20 (с. 25). Функция

$$\xi(\omega) = I_A(\omega) = \begin{cases} 1, & \text{если } \omega \in A, \\ 0, & \text{если } \omega \notin A \end{cases}$$

не является случайной величиной, поскольку, например, прообраз единицы $\{\omega \,|\, \xi(\omega) = 1\} = A$ не принадлежит $\mathcal{F}$. И вероятность для ξ попасть в единицу $\mathsf{P}(\xi = 1) = \lambda(A)/2\pi$ просто не существует.

Познакомимся с важным понятием — «распределение» случайной величины и опишем различные типы распределений случайных величин.

§ 2. Распределения случайных величин

Определение 22. *Распределением* случайной величины ξ называется вероятностная мера $\mu(B) = \mathsf{P}(\xi \in B)$ на множестве борелевских подмножеств $\mathbb{R}$.

Можно представлять себе распределение случайной величины ξ как соответствие между множествами $B \in \mathfrak{B}(\mathbb{R})$ и вероятностями $\mathsf{P}(\xi \in B)$.

Распределения случайных величин суть основные объекты изучения в теории вероятностей. Мы не будем, как правило, интересоваться тем, из какого множества действует функция и каким именно элементарным исходам сопоставляет свои возможные значения. Нас будет интересовать лишь, *с какой вероятностью* эти значения принимаются. Приведём несколько примеров совершенно разных случайных величин, имеющих одно и то же распределение (одинаково распределённых).

Пример 38. Один раз бросается правильная монета. Пространство Ω состоит из двух элементарных исходов — *герб* и *решка*. В качестве σ-алгебры рассмотрим множество всех подмножеств Ω, вероятность зададим как в классической схеме. Положим

$\xi(\omega) = 1$, если $\omega =$ *герб*, и $\xi(\omega) = 0$, если $\omega =$ *решка*;

$\eta(\omega) = 0$, если $\omega =$ *герб*, и $\eta(\omega) = 1$, если $\omega =$ *решка*.

Очевидно, что для любого множества $B \subseteq \mathbb{R}$ вероятности принадлежать B для ξ и для η одинаковы. Тем не менее ни для одного элементарного исхода ω значения $\xi(\omega)$ и $\eta(\omega)$ не совпадают. Иными словами, ξ и η одинаково распределены, но не одинаковы как функции.

Пример 39. Точка наудачу бросается на отрезок $[0,\, 1]$. В этом случае Ω есть отрезок $[0,\, 1]$ с σ-алгеброй борелевских подмножеств $[0,\, 1]$

и мерой Лебега в качестве вероятности. Читатель убедится, что две совершенно разные функции: $\xi(\omega) = \omega$ и $\eta(\omega) = 1 - \omega$ (расстояния до упавшей точки от левого и от правого концов отрезка соответственно) обладают одинаковыми вероятностями принимать значения внутри любых борелевских множеств B. Вероятности эти равны мере Лебега пересечения множеств B и $[0, 1]$. Эти случайные величины одинаково распределены, но не одинаковы: их значения совпадают лишь при одном элементарном исходе $\omega = 0{,}5$ (*нарисовать графики функций* $\xi(\omega)$ *и* $\eta(\omega)$).

Пример 40. На том же отрезке $\Omega = [0, 1]$ построим две функции: $\xi(\omega) = 0$ при всех ω; $\eta(\omega) = 0$ при всех ω, кроме $\omega = 0{,}5$, а в точке $\omega = 0{,}5$ положим $\eta(\omega) = -17$.

Поскольку мера Лебега точки (она же — вероятность) равна нулю, распределения величин ξ и η одинаковы. Теперь $\xi(\omega)$ и $\eta(\omega)$ снова не совпадают как функции, но отличаются их значения лишь на множестве нулевой вероятности — только в точке $\omega = 0{,}5$. В этом случае говорят, что ξ и η совпадают «почти наверное»: $\mathsf{P}(\xi = \eta) = 1$.

Опишем различные типы распределений случайных величин. Вся вероятностная масса может быть сосредоточена в нескольких точках прямой, а может быть «размазана» по некоторому интервалу или по всей прямой. В зависимости от типа множества, на котором сосредоточена вся единичная вероятностная масса, распределения делят на дискретные, абсолютно непрерывные, сингулярные и их смеси.

Определение 23. Случайная величина ξ имеет *дискретное* распределение, если существует конечный или счётный набор чисел $a_1, a_2, \ldots$ такой, что

$$\mathsf{P}(\xi = a_i) > 0 \ \text{ для всех } i, \qquad \sum_{i=1}^{\infty} \mathsf{P}(\xi = a_i) = 1.$$

Итак, случайная величина ξ имеет дискретное распределение, если она принимает не более чем счётное число значений. Значения эти иначе называют *атомами*: ξ имеет атом в точке x, если $\mathsf{P}(\xi = x) > 0$.

Если случайная величина ξ имеет дискретное распределение, то для любого $B \subseteq \mathbb{R}$

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

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	$\ldots$

Определение 24. Случайная величина ξ имеет *абсолютно непрерывное* распределение, если существует неотрицательная функция $f_\xi(x)$ такая, что для любого борелевского множества B имеет место равенство:

$$\mathsf{P}(\xi \in B) = \int\limits_B f_\xi(x)\,dx.$$

Функцию $f_\xi(x)$ называют плотностью распределения величины ξ.

Замечание. Интеграл выше есть интеграл Лебега, а не Римана. Вполне достаточно, если читатель, не знакомый с интегралом Лебега, будет представлять его себе просто как площадь под графиком подынтегральной функции над множеством B. При этом площадь над множеством B, имеющим нулевую меру Лебега, равна нулю. Заметим, что любая функция, отличающаяся от функции $f_\xi(x)$ лишь в конечном или счётном числе точек (или на множестве нулевой меры Лебега), будет являться плотностью того же распределения, так как интеграл не изменится от изменения подынтегральной функции на множестве меры нуль.

Теорема 19. *Плотность распределения обладает свойствами:*

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

Доказательство. Свойство (f1) выполнено по определению плотности, свойство (f2) также следует из определения 24. Действительно, если в качестве борелевского множества B взять всю числовую прямую, получим:

$$\mathsf{P}(\xi \in \mathbb{R}) = 1 = \int\limits_{\mathbb{R}} f_\xi(x)\;dx. \qquad \square$$

Эти два свойства полностью характеризуют класс плотностей.

Теорема 20. *Если функция f обладает свойствами* (f1) *и* (f2), *то существует вероятностное пространство и случайная величина ξ на нём, для которой f является плотностью распределения.*

Доказательство. Пусть Ω — область, заключенная между осью абсцисс и графиком функции f. Площадь области Ω равна единице по свойству (f2). Возьмём в качестве $\mathcal{F}$ множество борелевских подмножеств Ω, а в качестве вероятности P — меру Лебега (площадь) на множествах из $\mathcal{F}$. И пусть случайная величина ξ — абсцисса точки, наудачу брошенной в эту область. Тогда для любого $B \in \mathfrak{B}(\mathbb{R})$ выполнено

$$\mathsf{P}(\xi \in B) = \mathsf{P}(\text{точка попала в } D_B) = \frac{\text{площадь } D_B}{\text{площадь } \Omega} = \int\limits_B f(x)\,dx. \qquad (11)$$

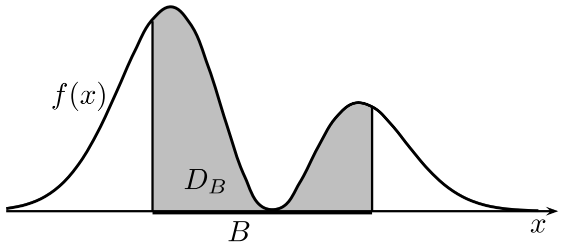

Рис. 6. Область $D_B = \{\xi \in B\}$

Здесь область D_B есть криволинейная трапеция с основанием B под графиком плотности (рис. 6). По определению, равенство (11) означает, что функция f является плотностью распределения величины ξ. □

Отметим полезное свойство абсолютно непрерывных распределений.

Свойство 7. *Если случайная величина* ξ *имеет абсолютно непрерывное распределение, то* $\mathsf{P}(\xi = x) = 0$ *для любого* $x \in \mathbb{R}$.

Доказательство сразу следует из определения 24 и следующего за ним замечания: интеграл по области интегрирования, состоящей из одной точки, равен нулю. □

Можно выделить ещё один особый класс распределений, сосредоточенных, в отличие от абсолютно непрерывных распределений, на множестве нулевой меры Лебега, но не имеющих, в отличие от дискретных, атома ни в одной точке этого множества.

Определение 25. Случайная величина ξ имеет *сингулярное* распределение, если существует борелевское множество B с нулевой лебеговой мерой $\lambda(B) = 0$ такое, что $\mathsf{P}(\xi \in B) = 1$, но при этом $\mathsf{P}(\xi = x) = 0$ для любой точки $x \in B$.

Можно отметить следующее свойство сингулярных распределений. Множество B, на котором сосредоточено всё распределение, не может состоять из конечного или счётного числа точек. Действительно, если B конечно или счётно, то $\mathsf{P}(\xi \in B) = \sum \mathsf{P}(\xi = x_i)$, где суммирование ведётся по всем $x_i \in B$. Последняя сумма равна нулю как сумма счётного числа нулей, что противоречит предположению $\mathsf{P}(\xi \in B) = 1$.

Таким образом, любое сингулярное распределение сосредоточено на несчётном множестве с нулевой мерой Лебега. Примером такого множества может служить канторовское совершенное множество, а примером такого распределения — лестница Кантора[9] (*выяснить, что это*).

Наконец, распределение может быть выпуклой линейной комбинацией дискретного, абсолютно непрерывного и сингулярного распределений.

Определение 26. Случайная величина ξ имеет *смешанное* распределение, если найдутся такие случайные величины ξ_1, ξ_2 и ξ_3 — с дискретным, абсолютно непрерывным и сингулярным распределениями соответственно (или такие три распределения), и числа p_1, p_2, $p_3 \in [0, 1)$, $p_1 + p_2 + p_3 = 1$, что для любого $B \in \mathfrak{B}(\mathbb{R})$ имеет место равенство

$$\mathsf{P}(\xi \in B) = p_1\,\mathsf{P}(\xi_1 \in B) + p_2\,\mathsf{P}(\xi_2 \in B) + p_3\,\mathsf{P}(\xi_3 \in B).$$

По заданным на одном вероятностном пространстве случайным величинам ξ_1, ξ_2, ξ_3 и числам $p_1 + p_2 + p_3 = 1$ можно построить случайную величину со смешанным распределением так: пусть φ — случайная величина на том же вероятностном пространстве с дискретным распределением $\mathsf{P}(\varphi = k) = p_k$ для $k = 1, 2, 3$, и пусть при любом k и любом $B \in \mathfrak{B}(\mathbb{R})$ события $\{\varphi = k\}$ и $\{\xi_k \in B\}$ независимы.

Построим случайную величину ξ так: если $\varphi(\omega) = k$, то положим $\xi(\omega) = \xi_k(\omega)$. Её распределение найдём по формуле полной вероятности:

$$\mathsf{P}(\xi \in B) = \mathsf{P}(\xi_1 \in B,\, \varphi = 1) + \mathsf{P}(\xi_2 \in B,\, \varphi = 2) + \mathsf{P}(\xi_3 \in B,\, \varphi = 3).$$

В силу независимости событий под знаком каждой из вероятностей,

$$\mathsf{P}(\xi \in B) = p_1\,\mathsf{P}(\xi_1 \in B) + p_2\,\mathsf{P}(\xi_2 \in B) + p_3\,\mathsf{P}(\xi_3 \in B).$$

Никаких других видов распределений, кроме перечисленных выше, не существует (*доказано Лебегом*).

§ 3. Функция распределения

Описание распределения набором вероятностей $\mathsf{P}(\xi \in B)$ не очень удобно: слишком много существует борелевских множеств. Мы описали дискретные распределения таблицей распределения, абсолютно непрерывные — плотностью распределения. Попробуем поискать какой-нибудь универсальный способ описать любое возможное распределение.

Можно поставить вопрос иначе: распределение есть набор вероятностей попадания в любые борелевские множества на прямой. Нельзя ли обойтись знанием вероятностей попадания в какой-нибудь меньший набор множеств на прямой? Борелевская σ-алгебра $\mathfrak{B}(\mathbb{R})$ порождается интервалами (равно как и лучами $(-\infty,\, x)$), поэтому можно ограничиться только

[9] Georg Ferdinand Ludwig Philipp Cantor (3.03.1845, Russia — 6.01.1918, Germany).

вероятностями попадания в такие лучи для всех $x \in \mathbb{R}$. А уже с их помощью можно будет определить и вероятность попасть в любое борелевское множество.

З а м е ч а н и е. Можно с таким же успехом ограничиться набором вероятностей попадания в интервалы $(-\infty,\, x]$, или в $(x,\, \infty)$, или в $[x,\, \infty)$.

О п р е д е л е н и е 27. *Функцией распределения* случайной величины ξ называется функция $F_\xi : \mathbb{R} \to [0,\, 1]$, при каждом $x \in \mathbb{R}$ равная вероятности случайной величине ξ принимать значения, меньшие x:

$$F_\xi(x) = \mathsf{P}(\xi < x) = \mathsf{P}\{\omega : \xi(\omega) < x\}.$$

Перечислим основные дискретные и абсолютно непрерывные распределения и найдём их функции распределения.

§ 4. Примеры дискретных распределений

Вырожденное распределение. Говорят, что случайная величина ξ имеет вырожденное распределение в точке $c \in \mathbb{R}$, и пишут: $\xi \in \mathrm{I}_c$, если ξ принимает единственное значение c с вероятностью 1, т. е. $\mathsf{P}(\xi = c) = 1$. Функция распределения ξ имеет вид

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

Распределение Бернулли. Говорят, что случайная величина ξ имеет распределение Бернулли с параметром p, и пишут: $\xi \in \mathrm{B}_p$, если ξ принимает значения 1 и 0 с вероятностями p и $1 - p$ соответственно. Случайная величина ξ с таким распределением равна *числу успехов* в одном испытании схемы Бернулли с вероятностью успеха p: ни одного успеха или один успех.

Функция распределения случайной величины ξ такова:

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

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

ξ	0	1	$\ldots$	k	$\ldots$	n
P	$(1-p)^n$	$np(1-p)^{n-1}$	$\ldots$	$C_n^k p^k (1-p)^{n-k}$	$\ldots$	p^n

Например, количество выпавших шестёрок при двадцати подбрасываниях правильной игральной кости имеет биномиальное распределение $\mathrm{B}_{20,\,\frac{1}{6}}$. Распределение Бернулли совпадает с распределением $\mathrm{B}_{1,\,p}$.

Геометрическое распределение. Говорят, что случайная величина τ имеет геометрическое распределение с параметром $p \in (0,\,1)$, и пишут $\tau \in \mathrm{G}_p$, если τ принимает значения $k = 1,\,2,\,3,\,\ldots$ с вероятностями $\mathsf{P}(\tau = k) = p(1-p)^{k-1}$. Случайная величина с таким распределением имеет смысл *номера первого успешного испытания* в схеме Бернулли с вероятностью успеха p. Таблица распределения τ имеет вид

τ	1	2	$\ldots$	k	$\ldots$
P	p	$p(1-p)$	$\ldots$	$p(1-p)^{k-1}$	$\ldots$

Распределение Пуассона. Говорят, что случайная величина ξ имеет распределение Пуассона с параметром $\lambda > 0$, и пишут: $\xi \in \Pi_\lambda$, если ξ принимает значения $k = 0,\,1,\,2,\,\ldots$ с вероятностями $\mathsf{P}(\xi = k) = \dfrac{\lambda^k}{k!}\,e^{-\lambda}$.

Распределение Пуассона возникло в теореме Пуассона (с. 47) как предельное распределение для числа успехов в n испытаниях схемы Бернулли, когда число испытаний n увеличивается, а вероятность успеха уменьшается обратно пропорционально n. Поэтому распределение Пуассона называют иначе *распределением числа редких событий.*

Гипергеометрическое распределение. Говорят, что случайная величина ξ имеет гипергеометрическое распределение с параметрами n, N и K, где $K \leqslant N$, $n \leqslant N$, если ξ принимает целые значения k такие, что $0 \leqslant k \leqslant K$, $0 \leqslant n-k \leqslant N-K$, с вероятностями $\mathsf{P}(\xi = k) = C_K^k C_{N-K}^{n-k} \,/\, C_N^n$. Случайная величина с таким распределением имеет смысл числа белых шаров среди n шаров, выбранных наудачу и без возвращения из урны, содержащей K белых шаров и $N-K$ не белых.

Упражнение. Построить графики функций распределения для вырожденного распределения, распределений Бернулли и Пуассона, биномиального и геометрического распределений.

§ 5. Примеры абсолютно непрерывных распределений

Равномерное распределение. Говорят, что ξ имеет равномерное распределение на отрезке $[a, b]$, и пишут: $\xi \in \mathrm{U}_{a,b}$, если плотность распределения ξ постоянна на отрезке $[a, b]$ и равна нулю вне него:

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

Площадь под графиком этой функции равна единице, $f_\xi(x) \geqslant 0$. Поэтому $f_\xi(x)$ является плотностью распределения.

Случайная величина $\xi \in \mathrm{U}_{a,b}$ имеет смысл *координаты точки, выбранной наудачу на отрезке* $[a, b]$. Вычислим по определению 24 функцию распределения случайной величины ξ:

$$F_\xi(x) = \mathsf{P}(\xi < x) = \int\limits_{-\infty}^{x} f_\xi(t)\,dt = \begin{cases} \int\limits_{-\infty}^{x} 0\,dt, & x < a, \\ \int\limits_{-\infty}^{a} 0\,dt + \int\limits_{a}^{x} \dfrac{1}{b-a}\,dt, & a \leqslant x \leqslant b, \\ \int\limits_{-\infty}^{a} 0\,dt + \int\limits_{a}^{b} \dfrac{1}{b-a}\,dt + \int\limits_{b}^{x} 0\,dt, & x > b. \end{cases}$$

Получим следующую непрерывную функцию распределения:

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

Графики плотности и функции распределения равномерного распределения на отрезке $[a, b]$ изображены на рис. 7.

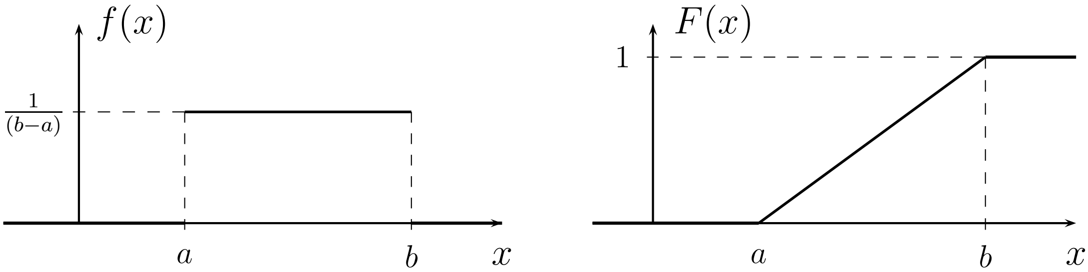

Рис. 7. Плотность и функция распределения $\mathrm{U}_{a,b}$

Показательное распределение. Говорят, что ξ имеет показательное (экспоненциальное) распределение с параметром $\alpha > 0$, и пишут:

$\xi \Subset \mathrm{E}_\alpha$, если ξ имеет следующую плотность распределения:

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

Функция распределения случайной величины ξ непрерывна:

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

Графики плотности и функции распределения показательного распределения с параметром α приведены на рис. 8.

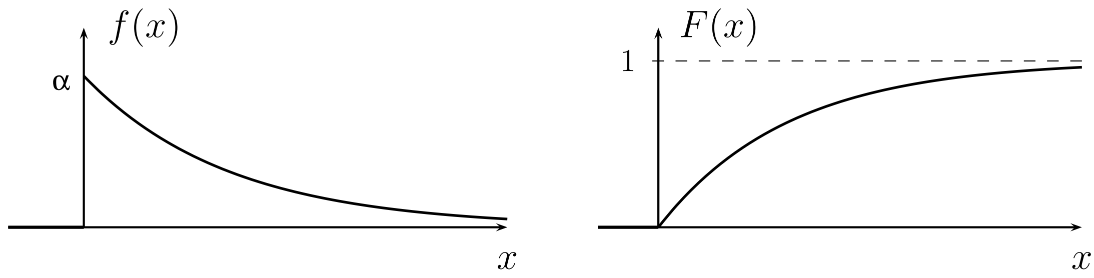

Рис. 8. Плотность и функция распределения E_α

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 является единственным абсолютно непрерывным распределением, для которого выполнено свойство «нестарения» (и в этом смысле оно является непрерывным аналогом дискретного геометрического распределения).

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

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

Упражнение. Доказать теорему 21. Доказать далее, что если неотрицательная случайная величина ξ с абсолютно непрерывным распределением обладает свойством (12) при любых $x, y > 0$, то она имеет показательное распределение с некоторым параметром α.

Нормальное распределение. Говорят, что ξ имеет нормальное (гауссовское[10]) распределение с параметрами a и σ^2, где $a \in \mathbb{R}$, $\sigma > 0$, и пишут: $\xi \in\; \mathrm{N}_{a,\,\sigma^2}$, если ξ имеет следующую плотность распределения:

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

На рис. 9 приведены графики плотностей нормальных распределений с одним и тем же параметром a и разными значениями параметра σ.

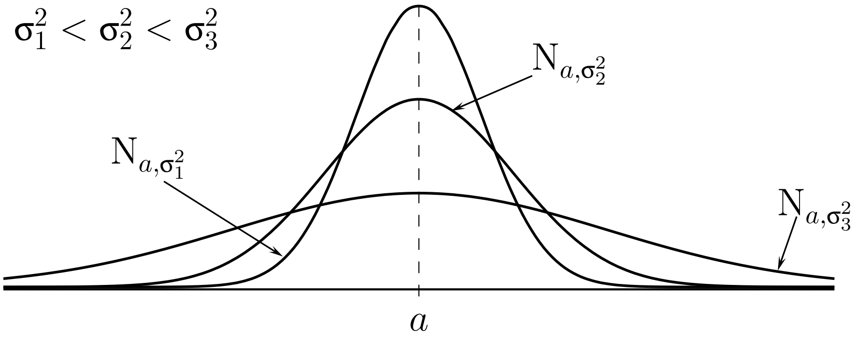

Рис. 9. Плотности нормальных распределений

Убедимся, что $f_\xi(x)$ является плотностью распределения. Так как $f_\xi(x) > 0$ для всех $x \in \mathbb{R}$, то свойство (f1) выполнено. Проверим (f2):

$$\int\limits_{-\infty}^{\infty} f_\xi(x)\,dx \;=\; \int\limits_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx = \left[\begin{array}{c}\text{замена переменных}\\ t = \dfrac{x-a}{\sigma},\; dx = \sigma\,dt\end{array}\right] =$$

$$= \int\limits_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\; e^{-t^2/2}\sigma\,dt = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-t^2/2}\,dt \;=\; \frac{I}{\sqrt{2\pi}} \;=1,$$

где через I обозначен табличный интеграл (интеграл Пуассона[11])

$$I = \int\limits_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi}.$$

[10]Johann Carl Friedrich Gauss (30.04.1777—23.02.1855, Germany).

[11] Этот интеграл вычисляется так:

$$I^2 = \int\limits_{-\infty}^{\infty} e^{-x^2/2}\,dx \int\limits_{-\infty}^{\infty} e^{-y^2/2}\,dy \;=\; \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\,dx\,dy.$$

Далее полярная замена переменных: $x = r\cos\varphi$, $y = r\sin\varphi$, $dx\,dy = r\,dr\,d\varphi$, $x^2 + y^2 = r^2$:

$$I^2 = \int\limits_0^{2\pi}\int\limits_0^{\infty} re^{-r^2/2}\,dr\,d\varphi = \int\limits_0^{2\pi}\int\limits_0^{\infty} e^{-r^2/2}\,d(r^2/2)\,d\varphi = 2\pi, \quad I = \sqrt{2\pi}.$$

Нормальное распределение $\mathrm{N}_{0,1}$ с параметрами $a = 0$ и $\sigma^2 = 1$ называется *стандартным нормальным* распределением. Плотность стандартного нормального распределения равна $f_\xi(x) = \dfrac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$.

Мы будем использовать специальное обозначение $\Phi_{a,\sigma^2}(x)$ для функции распределения нормального закона N_{a,σ^2} (рис. 10). Первообразная функции e^{-x^2} не может быть выражена через элементарные функции. Поэтому функцию $\Phi_{a,\sigma^2}(x)$ можно записать лишь в виде интеграла

$$\Phi_{a,\sigma^2}(x) = \int\limits_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(t-a)^2}{2\sigma^2}}\, dt, \qquad \Phi_{0,1}(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t^2}{2}}\, dt.$$

Функция $\Phi_{0,1}(x)$ табулирована, т. е. её значения при различных вещественных x вычислены. Их можно найти в соответствующих таблицах.

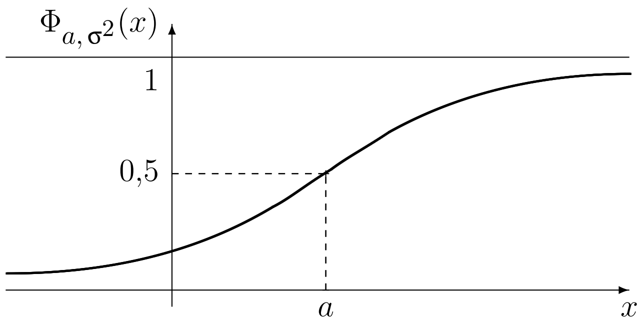

Рис. 10. Функция распределения нормального распределения N_{a,σ^2}

Гамма-распределение. Говорят, что ξ имеет гамма-распределение с параметрами $\alpha > 0$, $\lambda > 0$, и пишут: $\xi \in \Gamma_{\alpha,\lambda}$, если ξ имеет следующую плотность распределения:

$$f_\xi(x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ c \cdot x^{\lambda-1} e^{-\alpha x}, & \text{если } x > 0, \end{cases}$$

где постоянная c вычисляется из свойства (f2) плотности так:

$$1 = \int\limits_{-\infty}^{\infty} f_\xi(x)\, dx = c \int\limits_{0}^{\infty} x^{\lambda-1} e^{-\alpha x}\, dx = \frac{c}{\alpha^\lambda} \int\limits_{0}^{\infty} (\alpha x)^{\lambda-1} e^{-\alpha x}\, d(\alpha x) = \frac{c}{\alpha^\lambda}\, \Gamma(\lambda),$$

откуда $c = \alpha^\lambda / \Gamma(\lambda)$. Здесь через $\Gamma(\lambda)$ обозначен интеграл

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

называемый гамма-функцией Эйлера[12]; $\Gamma(k) = (k-1)!$ при целых положительных k, $\Gamma(1) = 1$. Замена в интеграле Пуассона даст $\Gamma(1/2) = \sqrt{\pi}$.

Полезно отметить, что показательное распределение есть частный случай гамма-распределения: $\mathrm{E}_\alpha = \Gamma_{\alpha,1}$.

Упражнение. 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.

Функцию распределения гамма-распределения можно записать, вообще говоря, только в виде интеграла:

$$F_\xi(x) = \frac{\alpha^\lambda}{\Gamma(\lambda)} \int\limits_0^x t^{\lambda-1} e^{-\alpha t}\, dt.$$

Но при целых значениях параметра λ интегрированием по частям этот интеграл можно превратить в сумму:

$$F_\xi(x) = 1 - \sum_{k=0}^{\lambda-1} \frac{(\alpha x)^k}{k!}\, e^{-\alpha x} = \sum_{k=\lambda}^{\infty} \frac{(\alpha x)^k}{k!}\, e^{-\alpha x}. \qquad (13)$$

Упражнение. Доказать первое из равенств (13). Доказать следующее забавное равенство: $F_\xi(x) = \mathsf{P}(\eta \geqslant \lambda)$, где $\eta \in \Pi_{\alpha x}$.

Распределение Коши. Говорят, что ξ имеет распределение Коши[13] с параметрами $a \in \mathbb{R}$, $\sigma > 0$, и пишут: $\xi \in \mathrm{C}_{a,\sigma}$, если ξ имеет следующую плотность распределения:

$$f_\xi(x) = \frac{1}{\pi}\, \frac{\sigma}{\sigma^2 + (x-a)^2} \quad \text{для любого} \quad x \in \mathbb{R}.$$

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 ξ с распределением Коши равна $F_\xi(x) = \dfrac{1}{2} + \dfrac{1}{\pi} \operatorname{arctg}\left(\dfrac{x-a}{\sigma}\right)$ при всех x.

[12] Leonhard Euler (15.04.1707, Switzerland — 18.09.1783, Russia).
[13] Augustin Louis Cauchy (21.08.1789—23.05.1857, France).

Распределение Парето. Говорят, что ξ имеет распределение Парето[14] с параметром $\alpha > 0$, если ξ имеет следующие плотность и функцию распределения:

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

Часто рассматривают более широкий класс распределений Парето, сосредоточенных не на $[1, \infty)$, а на $[c, \infty)$ при $c > 0$.

С другими абсолютно непрерывными распределениями (хи-квадрат Пирсона, распределениями Стью́дента, Фишера, Колмогорова, Лапласа) мы познакомимся при изучении математической статистики. С распределениями Вейбулла, логарифмически нормальным и некоторыми другими читатель познакомится в дальнейших курсах.

§ 6. Свойства функций распределения

Общие свойства функций распределения. Функцией распределения случайной величины ξ мы назвали функцию $F_\xi(x) = \mathsf{P}(\xi < x)$.

Т е о р е м а 22. *Любая функция распределения обладает свойствами:*

(F1) *она не убывает: если $x_1 < x_2$, то $F_\xi(x_1) \leqslant F_\xi(x_2)$;*

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

(F3) *она в любой точке непрерывна слева:*

$$F_\xi(x_0 - 0) = \lim_{x\to x_0-0} F_\xi(x) = F_\xi(x_0).$$

Д о к а з а т е л ь с т в о с в о й с т в а (F1). Для любых чисел $x_1 < x_2$ событие $\{\xi < x_1\}$ влечёт событие $\{\xi < x_2\}$, т. е. $\{\xi < x_1\} \subseteq \{\xi < x_2\}$. Но вероятность — монотонная функция событий, поэтому

$$F_\xi(x_1) = \mathsf{P}\{\xi < x_1\} \leqslant \mathsf{P}\{\xi < x_2\} = F_\xi(x_2). \qquad \square$$

Для доказательства остальных свойств нам понадобится свойство непрерывности вероятностной меры (с. 33, теорема 7).

Д о к а з а т е л ь с т в о с в о й с т в а (F2). Заметим сначала, что существование пределов в свойствах (F2), (F3) вытекает из монотонности и ограниченности функции $F_\xi(x)$. Остаётся лишь доказать равенства $\lim\limits_{x\to-\infty} F_\xi(x) = 0$, $\lim\limits_{x\to+\infty} F_\xi(x) = 1$ и $\lim\limits_{x\to x_0-0} F_\xi(x) = F_\xi(x_0)$. Для этого в каждом случае достаточно найти предел по какой-нибудь подпоследова-

[14] Vilfredo Pareto (15.07.1848—20.08.1923, France, Italy, Switzerland).

тельности $\{x_n\}$, так как существование предела влечёт совпадение всех частичных пределов.

Докажем, что $F_\xi(-n) \to 0$ при $n \to \infty$. Рассмотрим вложенную убывающую последовательность событий $B_n = \{\xi < -n\}$:

$$B_{n+1} = \big\{\xi < -(n+1)\big\} \subseteq B_n = \big\{\xi < -n\big\} \text{ для любых } n \geqslant 1.$$

Пересечение B всех этих событий состоит из тех и только тех ω, для которых $\xi(\omega)$ меньше любого вещественного числа. Но для любого элементарного исхода ω значение $\xi(\omega)$ вещественно, и не может быть меньше всех вещественных чисел. Иначе говоря, пересечение событий B_n не содержит элементарных исходов, т. е. $B = \bigcap B_n = \varnothing$. По свойству непрерывности меры, $F_\xi(-n) = \mathsf{P}(B_n) \to \mathsf{P}(B) = 0$ при $n \to \infty$.

Точно так же докажем остальные свойства.

Покажем, что $F_\xi(n) \to 1$ при $n \to \infty$, т. е. $1 - F_\xi(n) = \mathsf{P}(\xi \geqslant n) \to 0$. Обозначим через B_n событие $B_n = \{\xi \geqslant n\}$. События B_n вложены:

$$B_{n+1} = \big\{\xi \geqslant (n+1)\big\} \subseteq B_n = \big\{\xi \geqslant n\big\} \text{ для любых } n \geqslant 1,$$

а пересечение B этих событий снова пусто: оно означает, что ξ больше любого вещественного числа. По свойству непрерывности меры,

$$1 - F_\xi(n) = \mathsf{P}(B_n) \to \mathsf{P}(B) = 0 \text{ при } n \to \infty. \qquad \square$$

Д о к а з а т е л ь с т в о с в о й с т в а (F3). Достаточно доказать, что $F_\xi(x_0 - 1/n) \to F_\xi(x_0)$ при $n \to \infty$. Иначе говоря, доказать сходимость к нулю следующей разности:

$$F_\xi(x_0) - F_\xi\Big(x_0 - \frac{1}{n}\Big) = \mathsf{P}(\xi < x_0) - \mathsf{P}\Big(\xi < x_0 - \frac{1}{n}\Big) = \mathsf{P}\Big(x_0 - \frac{1}{n} \leqslant \xi < x_0\Big).$$

У п р а ж н е н и е. Обозначьте событие $\{x_0 - 1/n \leqslant \xi < x_0\}$ через B_n, и попробуйте снова воспользоваться свойством непрерывности меры. $\square$

Следующая теорема говорит о том, что три доказанных свойства полностью описывают класс функций распределения. То, что любая функция распределения ими обладает, мы с вами доказали, а теорема утверждает, что любая функция с такими свойствами есть функция распределения.

Т е о р е м а 23. *Если функция $F : \mathbb{R} \to [0, 1]$ удовлетворяет свойствам* (F1)–(F3), *то F есть функция распределения некоторой случайной величины ξ, т. е. найдётся вероятностное пространство $\langle \Omega, \mathcal{F}, \mathsf{P} \rangle$ и случайная величина ξ на нём такая, что $F(x) \equiv F_\xi(x)$.*

Эту теорему мы доказывать не станем. Хотя её можно попробовать доказать конструктивно — предъявив то вероятностное пространство (проще всего отрезок $\Omega = [0, 1]$ с σ-алгеброй борелевских множеств и мерой Лебега) и ту случайную величину, о существовании которых идёт речь.

Упражнение. Непременно попробуйте сделать это! Например, можно проверить, не подойдёт ли $\xi(\omega) = \sup\{x \mid F(x) < \omega\}$.

Помимо отмеченных в теореме 22, функции распределения обладают следующими свойствами:

Свойство 8. *В любой точке x_0 разница $F_\xi(x_0+0) - F_\xi(x_0)$ равна $\mathsf{P}(\xi = x_0)$. Иначе говоря, $F_\xi(x_0+0) = F_\xi(x_0) + \mathsf{P}(\xi = x_0) = \mathsf{P}(\xi \leqslant x_0)$.*

Упражнение. Докажите (так же, как мы доказывали (F2) и (F3)).

Разница $F_\xi(x_0+0) - F_\xi(x_0)$ между пределом при стремлении к x_0 справа и значением в точке x_0 есть величина скачка функции распределения. Эта величина равна нулю, если функция распределения непрерывна (справа) в точке x_0. Слева функция распределения непрерывна всегда.

Замечание. Очень часто функцией распределения называют $\mathsf{P}(\xi \leqslant x)$. Эта функция отличается от определённой выше лишь тем, что она непрерывна *справа*, а не слева. И вероятность $\mathsf{P}(\xi = x_0)$ для неё равна величине скачка слева, а не справа.

Свойство 9. *Для любой случайной величины ξ*

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

Доказательство. Разобьём событие $\{\xi < b\}$ в объединение несовместных событий: $\{\xi < a\} \cup \{a \leqslant \xi < b\} = \{\xi < b\}$. По свойству аддитивности вероятности, $\mathsf{P}\{\xi < a\} + \mathsf{P}\{a \leqslant \xi < b\} = \mathsf{P}\{\xi < b\}$, или $F_\xi(a) + \mathsf{P}\{a \leqslant \xi < b\} = F_\xi(b)$, что и требовалось доказать. □

Функция распределения дискретного распределения. Согласно определению дискретного распределения, его функция распределения может быть найдена по таблице распределения так:

$$F_\xi(x) = \mathsf{P}(\xi < x) = \sum_{k\,:\,a_k < x} \mathsf{P}(\xi = a_k).$$

Из свойств 8 и 9 вытекает следующее свойство.

Свойство 10. *Случайная величина ξ имеет дискретное распределение тогда и только тогда, когда функция распределения $F_\xi(x)$ имеет в точках a_i скачки с величиной $p_i = \mathsf{P}(\xi = a_i) = F_\xi(a_i + 0) - F_\xi(a_i)$, и растёт только за счёт скачков.*

Упражнение. Доказать, что любая функция распределения имеет не более чем счётное число точек разрыва (или «скачков»). Сколько скачков с величиной более $1/2$ может иметь функция распределения?А скачков с величиной более $1/3$? Более $1/4$?

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

$$F_\xi(x) = \mathsf{P}(\xi < x) = \mathsf{P}(\xi \in (-\infty,\, x)) = \int\limits_{-\infty}^{x} f_\xi(t)\, dt. \tag{14}$$

Поскольку функция распределения однозначно определяет распределение случайной величины, можно считать возможность представить функцию распределения интегралом (14) от неотрицательной функции определением абсолютно непрерывного распределения.

Свойство 11. *Если случайная величина ξ имеет абсолютно непрерывное распределение, то её функция распределения всюду непрерывна.*

Доказательство. Этот факт следует из свойств 7 и 8. Заметим, что оно есть также следствие представления (14) и непрерывности интеграла как функции верхнего предела. □

Свойство 12. *Если ξ имеет абсолютно непрерывное распределение, то её функция распределения дифференцируема почти всюду,*

$$f_\xi(x) = F'_\xi(x) = \frac{d}{dx}\, F_\xi(x) \quad \text{для почти всех } x.$$

Замечание. Термин для «почти всех» означает «для всех, кроме (возможно) x из некоторого множества нулевой меры Лебега».

Заметим, что *любая* функция распределения дифференцируема почти всюду. Например, функции распределения равномерного распределения и распределения Бернулли дифференцируемы всюду, кроме двух точек. Но у равномерного распределения плотность существует, а у распределения Бернулли — нет. Поэтому возможность дифференцировать функцию распределения никакого отношения к существованию плотности не имеет. Даже если мы дополнительно потребуем непрерывности функции распределения, этого не будет достаточно для абсолютной непрерывности распределения. Например, далее мы увидим, что функция распределения сингулярного распределения непрерывна и дифференцируема почти всюду, однако плотности у этого распределения нет, так как производная функции распределения почти всюду равна нулю.

Опираясь на свойство 12 и формулу (14), можно сформулировать такой критерий абсолютной непрерывности распределения: распределение с функцией распределения $F_\xi(x)$ абсолютно непрерывно, если при всех x

$$F_\xi(x) = \int\limits_{-\infty}^{x} F'_\xi(t)\,dt.$$

Наконец, очевидным следствием определения абсолютно непрерывного распределения и свойства 11 является следующее свойство.

С в о й с т в о 13. *Если случайная величина ξ имеет абсолютно непрерывное распределение, то для любых $a < b$ имеют место равенства:*

$$\mathsf{P}(a < \xi < b) = \mathsf{P}(a \leqslant \xi < b) = \mathsf{P}(a < \xi \leqslant b) = \mathsf{P}(a \leqslant \xi \leqslant b) = \int\limits_a^b f_\xi(t)\,dt.$$

Функция распределения сингулярного распределения. Для полноты картины посмотрим, какие свойства имеет функция распределения сингулярного распределения. По определению 25, случайная величина ξ с сингулярным распределением принимает с единичной вероятностью лишь значения из некоторого борелевского множества B с нулевой лебеговой мерой. Поэтому $\mathsf{P}(\xi \in \mathbb{R} \setminus B) = 0$. Но по свойству (9), если $\mathsf{P}(\xi \in [a, b)) = 0$, то $F_\xi(b) = F_\xi(a)$. Это означает, что расти функция распределения может лишь в точках множества B. На всём остальном множестве $\mathbb{R} \setminus B$ функция распределения имеет почти всюду нулевую производную. Тем не менее, $F_\xi(x)$ всюду непрерывна, поскольку $\mathsf{P}(\xi = x) = 0$ для любой точки $x \in \mathbb{R}$. Примером такой функции распределения служит лестница Кантора (рис. 11).

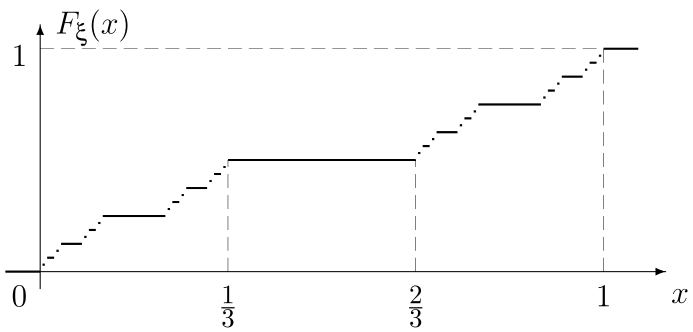

Рис. 11. Лестница Кантора

Функция распределения смешанного распределения. Функция распределения смешанного распределения есть линейная комбинация функций распределения дискретного, абсолютно непрерывного и сингулярного распределений. Если смешивать дискретное и абсолютно непрерывное распределения, то функция распределения будет иметь скачки в точках значений дискретного распределения и участки непрерывного роста, приращение на которых восстанавливается по её производной.

§ 7. Свойства нормального распределения

Рассмотрим отдельно свойства самого главного распределения. Сначала установим связь между функциями $\Phi_{a,\sigma^2}(x)$ и $\Phi_{0,1}(x)$.

Свойство 14. *Для любого* $x \in \mathbb{R}$ *справедливо соотношение:*

$$\Phi_{a,\sigma^2}(x) = \Phi_{0,1}\Big(\frac{x-a}{\sigma}\Big).$$

Доказательство. Действительно,

$$\Phi_{a,\sigma^2}(x) = \int\limits_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(t-a)^2/\,2\sigma^2} dt = \int\limits_{-\infty}^{\frac{x-a}{\sigma}} \frac{1}{\sqrt{2\pi}}\, e^{-y^2/\,2} dy = \Phi_{0,1}\Big(\frac{x-a}{\sigma}\Big).$$

Мы сделали замену переменных $y = \dfrac{t-a}{\sigma}$, $dt = \sigma\, dy$, верхняя граница интегрирования $t = x$ при такой замене перешла в $y = \dfrac{x-a}{\sigma}$. □

То же самое для случайных величин можно сформулировать так:

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

Доказательство. Убедимся, что случайная величина η имеет функцию распределения $\Phi_{0,1}(x)$:

$$\begin{aligned} F_\eta(x) &= \mathsf{P}(\eta < x) = \mathsf{P}\Big(\frac{\xi - a}{\sigma} < x\Big) = \mathsf{P}(\xi < \sigma x + a) = \\ &= \Phi_{a,\sigma^2}(\sigma x + a) = \Phi_{0,1}\Big(\frac{\sigma x + a - a}{\sigma}\Big) = \Phi_{0,1}(x). \quad \square \end{aligned}$$

Следствие 3. *Если* $\xi \in \mathrm{N}_{a,\sigma^2}$, *то*

$$\mathsf{P}(x_1 < \xi < x_2) = \Phi_{a,\sigma^2}(x_2) - \Phi_{a,\sigma^2}(x_1) = \Phi_{0,1}\Big(\frac{x_2 - a}{\sigma}\Big) - \Phi_{0,1}\Big(\frac{x_1 - a}{\sigma}\Big).$$

Видим, что вычисление любых вероятностей для нормально распределённой случайной величины сводится к вычислению функции распределения $\Phi_{0,1}(x)$. Она обладает следующими свойствами (*нарисуйте их на графике плотности стандартного нормального распределения*).

Свойство 15. $\Phi_{0,1}(0) = 0{,}5$, $\quad \Phi_{0,1}(-x) = 1 - \Phi_{0,1}(x)$.

Свойство 16. *Если* $\xi \in \mathrm{N}_{0,1}$, *то для любого* $x > 0$

$$\mathsf{P}(|\xi| < x) = 1 - 2\Phi_{0,1}(-x) = 2\Phi_{0,1}(x) - 1.$$

Доказательство. При $x > 0$ имеем

$$\begin{aligned}\mathsf{P}(|\xi| < x) &= \mathsf{P}(-x < \xi < x) = \Phi_{0,1}(x) - \Phi_{0,1}(-x) = \\ &= 1 - 2\Phi_{0,1}(-x) = 2\Phi_{0,1}(x) - 1. \qquad \square\end{aligned}$$

Свойство 17 (правило трёх сигм). *Если* $\xi \in \mathrm{N}_{a,\sigma^2}$, *то*

$$\mathsf{P}(|\xi - a| \geqslant 3\sigma) = 0{,}0027 \quad (\textit{совсем мало}).$$

Доказательство. Перейдём к противоположному событию:

$$\mathsf{P}\big(|\xi - a| \geqslant 3\sigma\big) = 1 - \mathsf{P}\big(|\xi - a| < 3\sigma\big) = 1 - \mathsf{P}\left(\left|\frac{\xi - a}{\sigma}\right| < 3\right).$$

Но величина $\eta = \dfrac{\xi - a}{\sigma}$ имеет стандартное нормальное распределение и можно использовать свойство 16:

$$1 - \mathsf{P}(|\eta| < 3) = 2\Phi_{0,1}(-3) = 2 \cdot 0{,}00135 = 0{,}0027.$$

Число $\Phi_{0,1}(-3) = 0{,}00135$ полезно отыскать в таблице. $\square$

Большого смысла в запоминании числа 0,0027 нет, но полезно помнить, что почти вся масса нормального распределения сосредоточена в границах от $a - 3\sigma$ до $a + 3\sigma$.

ГЛАВА VI

ПРЕОБРАЗОВАНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

…По данному соотношению между флюентами определить соотношение между флюксиями.

И. Ньютон. Метод флюксий и бесконечных рядов…

§ 1. Измеримость функций от случайных величин

Пусть на вероятностном пространстве $\langle\Omega, \mathcal{F}, \mathsf{P}\rangle$ задана случайная величина ξ. Если $g(\xi)$ — случайная величина, то полезно уметь находить распределение $g(\xi)$ 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. Если же нам необходимы значения показательно распределённой случайной величины, нужно найти преобразование, которое из равномерного распределения сделает показательное.

Вопрос об измеримости $g(\xi)$ решает следующая теорема.

Теорема 24. *Пусть ξ — случайная величина, а $g : \mathbb{R} \to \mathbb{R}$ — борелевская (измеримая по Борелю) функция, т. е. такая, что для всякого борелевского множества B его прообраз $g^{-1}(B)$ есть снова борелевское множество. Тогда $g(\xi)$ — случайная величина.*

Доказательство. Проверим, что прообраз любого борелевского множества при отображении $g(\xi) : \Omega \to \mathbb{R}$ является событием. Возьмём произвольное $B \in \mathfrak{B}(\mathbb{R})$ и положим $B_1 = g^{-1}(B)$. Множество B_1 борелевское, так как функция g измерима по Борелю. Но тогда $(g(\xi))^{-1}(B) = \xi^{-1}(B_1)$. Это множество принадлежит $\mathcal{F}$, поскольку B_1 — борелевское множество и ξ — случайная величина. □

Борелевскими являются все привычные нам функции. Функцией, неизмеримой по Борелю, будет, например, индикаторная функция неизмеримого множества Витали. Вообще говоря, неизмеримые функции суть объекты экзотические, в обычной жизни не встречающиеся.

§ 2. Распределения функций от случайных величин

Линейные и монотонные преобразования. Если ξ имеет дискретное распределение, то для любой $g(\xi)$ также имеет дискретное распределение, и таблица её распределения находится просто по определению. Поэтому мы будем рассматривать преобразования случайных величин с абсолютно непрерывными распределениями.

Пусть случайная величина ξ имеет функцию распределения $F_\xi(x)$ и плотность распределения $f_\xi(x)$. Построим с помощью борелевской функции $g : \mathbb{R} \to \mathbb{R}$ случайную величину $\eta = g(\xi)$. Требуется найти плотность распределения величины η (если таковая существует).

Замечание. 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.

Упражнение. Привести пример плотности распределения $f_\xi(x)$ и *непрерывной* функции g таких, что $\eta = g(\xi)$ имеет: а) дискретное распределение; б) невырожденное дискретное распределение.

Для отыскания плотности распределения мы не можем просто продифференцировать функцию распределения, поскольку не знаем, существует ли плотность. Следует доказать, что распределение абсолютно непрерывно. Но доказывая это, мы попутно найдём и плотность распределения. Действительно, у нас есть следующий путь доказательства абсолютной непрерывности распределения. Если для любого x мы представим функцию распределения величины η в виде

$$F_\eta(x) = \int\limits_{-\infty}^{x} h(y)\,dy, \qquad \text{где } h(y) \geqslant 0,$$

то в качестве плотности распределения величины η можно взять подынтегральную функцию: $f_\eta(x) = h(x)$. Другой путь — продифференцировать функцию распределения и убедиться, что производная является плотностью распределения, т. е. обладает свойствами (f1) и (f2).

Теорема 25. *Пусть ξ имеет функцию распределения $F_\xi(x)$ и плотность распределения $f_\xi(x)$, и постоянная a отлична от нуля. Тогда случайная величина $\eta = a\xi + b$ имеет плотность распределения*

$$f_\eta(x) = \frac{1}{|a|}\, f_\xi\Big(\frac{x-b}{a}\Big).$$

Доказательство. Пусть сначала $a > 0$.

$$F_\eta(x) = \mathsf{P}(a\xi + b < x) = \mathsf{P}\Big(\xi < \frac{x-b}{a}\Big) = F_\xi\Big(\frac{x-b}{a}\Big) = \int\limits_{-\infty}^{(x-b)/a} f_\xi(t)\,dt.$$

Сделаем замену переменной в последнем интеграле. Переменную t заменим на новую переменную y так: $t = \dfrac{y-b}{a}$. Тогда $dt = \dfrac{dy}{a}$, нижняя граница области интегрирования $t = -\infty$ перейдёт в $y = -\infty$, верхняя граница $t = \dfrac{x-b}{a}$ перейдёт в $y = x$. Получим

$$F_\eta(x) = \int\limits_{-\infty}^{x} \frac{1}{a}\, f_\xi\Big(\frac{y-b}{a}\Big)\,dy.$$

Функция под интегралом есть искомая плотность распределения $f_\eta(y)$ случайной величины $\eta = a\xi + b$ при $a > 0$.

Пусть теперь $a < 0$.

$$F_\eta(x) = \mathsf{P}(a\xi + b < x) = \mathsf{P}\Big(\xi > \frac{x-b}{a}\Big) = \int\limits_{(x-b)/a}^{+\infty} f_\xi(t)\,dt.$$

Сделаем ту же замену переменной. Но теперь граница интегрирования $t = +\infty$ перейдёт в $y = -\infty$, поскольку $a < 0$. Получим

$$F_\eta(x) = \int\limits_{x}^{-\infty} \frac{1}{a}\, f_\xi\Big(\frac{y-b}{a}\Big)\,dy = \int\limits_{-\infty}^{x} \frac{1}{|a|}\, f_\xi\Big(\frac{y-b}{a}\Big)\,dy.$$

Функция под интегралом — плотность распределения $f_\eta(y)$ случайной величины $\eta = a\xi + b$ при $a < 0$. $\square$

Из теоремы 25 следуют уже знакомые нам утверждения.

Следствие 4. *Если* $\xi \in \mathrm{N}_{0,1}$, *то* $\eta = \sigma\xi + a \in \mathrm{N}_{a,\sigma^2}$.

Доказательство. Действительно,

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

Следствие 5. *Если* $\xi \in \mathrm{N}_{a,\sigma^2}$, *то* $\dfrac{\xi - a}{\sigma} \in \mathrm{N}_{0,1}$.

Следствие 6. *Если* $\xi \in \mathrm{U}_{0,1}$, *то* $a\xi + b \in \mathrm{U}_{b,\,a+b}$ *при* $a > 0$.

Следствие 7. *Если* $\xi \in \mathrm{E}_\alpha$, *то* $\alpha\xi \in \mathrm{E}_1$.

Квантильное преобразование. Полезно уметь строить случайные величины с заданным распределением по равномерно распределённой случайной величине (например, по результату датчика случайных чисел).

Теорема 26. *Пусть функция распределения $F(x) = F_\xi(x)$ непрерывна. Тогда случайная величина $\eta = F(\xi)$ имеет равномерное на отрезке $[0, 1]$ распределение.*

Доказательство. Найдём функцию распределения случайной величины η. Заметим, что всегда $0 \leqslant \eta \leqslant 1$. Предположим сначала, что функция F всюду возрастает. Тогда она обратима, и поэтому

$$F_\eta(x) = \mathsf{P}(F(\xi) < x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ \mathsf{P}(\xi < F^{-1}(x)), & \text{если } x \in (0, 1), \\ 1, & \text{если } x \geqslant 1. \end{cases} \tag{15}$$

Но $\mathsf{P}(\xi < F^{-1}(x)) = F\left(F^{-1}(x)\right) = x$, т. е. $\eta \in \mathrm{U}_{0,1}$.

Если функция F не является всюду возрастающей, то у неё есть участки постоянства. В этом случае просто обозначим через $F^{-1}(x)$ самую левую точку из замкнутого множества $\{t \mid F(t) = x\}$. При таком понимании «обратной» функции равенства (15) остаются справедливыми вместе с равенством $\mathsf{P}(\xi < F^{-1}(x)) = F\left(F^{-1}(x)\right) = x$ для любого $x \in (0, 1)$. □

Обратим теорему 26. Следующее утверждение верно не только для непрерывных, но для любых функций распределения F. Обозначим через $F^{-1}(x)$ точную нижнюю грань множества тех t, для которых $F(t) \geqslant x$:

$$F^{-1}(x) = \inf\{t \mid F(t) \geqslant x\}.$$

Для непрерывной функции F это определение «обратной функции» совпадает с введённым в доказательстве теоремы 26.

Теорема 27. *Пусть $\eta \in \mathrm{U}_{0,1}$, а F — произвольная функция распределения. Тогда случайная величина $\xi = F^{-1}(\eta)$ («квантильное преобразование» над η) имеет функцию распределения F.*

Следствие 8. *Пусть $\eta \in \mathrm{U}_{0,1}$. Верны соотношения:*

$$-\frac{1}{\alpha}\ln(1-\eta) \in \mathrm{E}_\alpha, \quad a + \sigma\,\mathrm{tg}(\pi\eta - \pi/2) \in C_{a,\sigma}, \quad \Phi_{0,1}^{-1}(\eta) \in \mathrm{N}_{0,1}.$$

Упражнение. Доказать теорему 27 и следствие 8, а также продолжить список соотношений. Как получить случайную величину с распределением Парето? А с нормальным распределением? (Указание: так её никто не получает).

ГЛАВА VII

МНОГОМЕРНЫЕ РАСПРЕДЕЛЕНИЯ

Не следует множить сущности сверх необходимости.

Принцип «бритва Оккама»

§ 1. Совместное распределение

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

Определение 28. Функция

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

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. Для простоты обозначений ограничимся вектором (ξ_1, ξ_2) из двух величин.

(F0) Для любых x_1, x_2 верно неравенство: $0 \leqslant F_{\xi_1, \xi_2}(x_1, x_2) \leqslant 1$.

(F1) $F_{\xi_1, \xi_2}(x_1, x_2)$ не убывает по каждой координате вектора (x_1, x_2).

(F2) Для любого $i = 1, 2$ существует $\lim\limits_{x_i \to -\infty} F_{\xi_1, \xi_2}(x_1, x_2) = 0$. Существует двойной предел $\lim\limits_{x_1 \to +\infty} \lim\limits_{x_2 \to +\infty} F_{\xi_1, \xi_2}(x_1, x_2) = 1$.

(F3) Функция $F_{\xi_1, \xi_2}(x_1, x_2)$ по каждой координате вектора (x_1, x_2) непрерывна слева.

(F4) Чтобы по функции совместного распределения восстановить функции распределения ξ_1 и ξ_2 в отдельности, следует устремить мешающую переменную к $+\infty$:

$$\lim_{x_1 \to +\infty} F_{\xi_1, \xi_2}(x_1, x_2) = F_{\xi_2}(x_2), \quad \lim_{x_2 \to +\infty} F_{\xi_1, \xi_2}(x_1, x_2) = F_{\xi_1}(x_1). \tag{16}$$

Доказательство всех этих свойств совершенно аналогично одномерному случаю. Но теперь свойств (F0)—(F3) не хватает для описания класса функций совместного распределения. Иначе говоря, выполнение этих свойств для некоторой функции $F : \mathbb{R}^2 \to \mathbb{R}$ не гарантирует, что эта функция является функцией распределения некоторого случайного вектора.

Упражнение. Доказать, что функция

$$F(x_1, x_2) = \begin{cases} 0, & \text{если } x_1 \leqslant 0 \text{ или } x_2 \leqslant 0 \text{ или } x_1 + x_2 \leqslant 1, \\ 1, & \text{если одновременно } x_1 > 0,\, x_2 > 0,\, x_1 + x_2 > 1 \end{cases}$$

удовлетворяет всем свойствам (F0)—(F3), но не является функцией распределения никакого вектора (ξ_1, ξ_2) хотя бы потому, что, найдись такой вектор, найдётся и прямоугольник $[a_1, b_1) \times [a_2, b_2)$, «вероятность» попасть в который (вычисленная с помощью этой якобы «функции распределения») отрицательна: $\mathsf{P}(a_1 \leqslant \xi_1 < b_1,\ a_2 \leqslant \xi_2 < b_2) < 0$.

Легко убедиться (*убедиться, что легко*), что вероятность вектору (ξ_1, ξ_2) попасть в прямоугольник $[a_1, b_1) \times [a_2, b_2)$ по функции распределения этого вектора вычисляется так: $\mathsf{P}(a_1 \leqslant \xi_1 < b_1, a_2 \leqslant \xi_2 < b_2) = F(b_1, b_2) + F(a_1, a_2) - F(a_1, b_2) - F(b_1, a_2)$.

Дополнительно к свойствам (F0)—(F3) от функции F требуют неотрицательности этого выражения (при любых $a_1 < b_1,\ a_2 < b_2$).

§ 2. Типы многомерных распределений

Ограничимся рассмотрением двух типичных случаев: когда совместное распределение координат случайного вектора (ξ_1, ξ_2) либо дискретно, либо абсолютно непрерывно. Заметим, что сингулярные совместные распределения тоже не являются редкостью, в отличие от одномерного случая: стоит бросить точку наудачу на отрезок на плоскости, и мы получим сингулярное совместное распределение (*доказать*).

Определение 29. Случайные величины $\xi_1,\ \xi_2$ имеют *дискретное* совместное распределение, если существует конечный или счётный набор пар чисел $\{a_i, b_j\}$ такой, что

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mathsf{P}(\xi_1 = a_i,\ \xi_2 = b_j) = 1.$$

Таблицу, на пересечении i-й строки и j-го столбца которой стоит вероятность $\mathsf{P}(\xi_1 = a_i,\ \xi_2 = b_j)$, называют таблицей совместного распределения случайных величин ξ_1 и ξ_2.

Таблицы распределения каждой из случайных величин $\xi_1,\ \xi_2$ в отдельности (таблицы частных, или *маргинальных* распределений) восстанавливаются по таблице совместного распределения с помощью формул

$$\mathsf{P}(\xi_1 = a_i) = \sum_{j=1}^{\infty} \mathsf{P}(\xi_1 = a_i,\, \xi_2 = b_j), \quad \mathsf{P}(\xi_2 = b_j) = \sum_{i=1}^{\infty} \mathsf{P}(\xi_1 = a_i,\, \xi_2 = b_j).$$

Так, первое равенство следует из того, что набор $\{\xi_2 = b_1\}$, $\{\xi_2 = b_2\}$, $\ldots$ есть полная группа событий, и поэтому событие $\{\xi_1 = a_i\}$ раскладывается в объединение попарно несовместных событий:

$$\{\xi_1 = a_i\} = \bigcup_{j=1}^{\infty} \{\xi_1 = a_i,\, \xi_2 = b_j\}.$$

Определение 30. Случайные величины ξ_1, ξ_2 имеют *абсолютно непрерывное* совместное распределение, если существует неотрицательная функция $f_{\xi_1,\xi_2}(x,\, y)$ такая, что для любого множества $B \in \mathfrak{B}(\mathbb{R}^2)$ имеет место равенство

$$\mathsf{P}((\xi_1,\, \xi_2) \in B) = \iint\limits_B f_{\xi_1,\xi_2}(x,\, y)\, dx\, dy.$$

Если такая функция $f_{\xi_1,\xi_2}(x,\, y)$ существует, она называется *плотностью совместного распределения* случайных величин ξ_1, ξ_2.

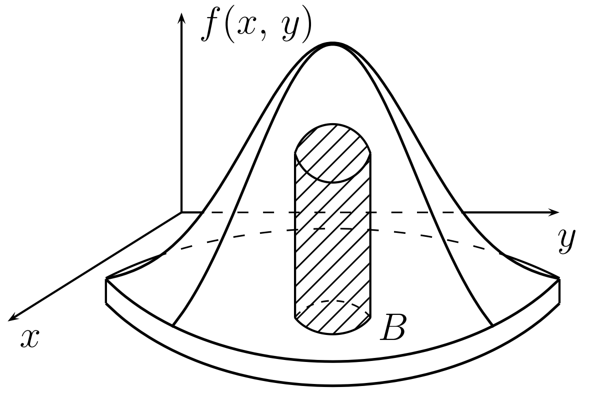

Рис. 12. Плотность в $\mathbb{R}^2$

Достаточно, если двойной интеграл по множеству B читатель будет понимать как объём области под графиком функции $f_{\xi_1,\xi_2}(x,\, y)$ над множеством B в плоскости переменных $(x,\, y)$, как показано на рис. 12.

Плотность совместного распределения обладает такими же свойствами, как и плотность распределения одной случайной величины:

(f1) неотрицательность: $f_{\xi_1,\xi_2}(x,\, y) \geqslant 0$ для любых $x,\, y \in \mathbb{R}$;

(f2) нормированность: $\displaystyle\iint\limits_{\mathbb{R}^2} f_{\xi_1,\,\xi_2}(x,\,y)\,dx\,dy = 1.$

Справедливо и обратное: любая функция, обладающая этими свойствами, является плотностью некоторого совместного распределения. Доказательство этого факта ничем не отличается от одномерного случая.

Если случайные величины $\xi_1,\ \xi_2$ имеют абсолютно непрерывное совместное распределение, то для любых $x_1,\,x_2$ имеет место равенство

$$F_{\xi_1,\,\xi_2}(x_1,\,x_2) = \mathsf{P}(\xi_1 < x_1,\,\xi_2 < x_2) = \int\limits_{-\infty}^{x_1}\left(\int\limits_{-\infty}^{x_2} f_{\xi_1,\,\xi_2}(x,\,y)\;dy\right)dx.$$

По функции совместного распределения его плотность находится как смешанная частная производная: $f_{\xi_1,\,\xi_2}(x,\,y) = \dfrac{\partial^2}{\partial x\partial y}\,F_{\xi_1,\,\xi_2}(x,\,y)$ для почти всех $(x,\,y)$.

Из существования плотностей ξ_1 и ξ_2 не следует абсолютная непрерывность совместного распределения этих случайных величин. Например, вектор $(\xi,\,\xi)$ принимает значения только на диагонали в $\mathbb{R}^2$ и уже поэтому не имеет плотности распределения (его распределение сингулярно). Обратное же свойство, как показывает следующая теорема, всегда верно: если совместное распределение абсолютно непрерывно, то и частные распределения тоже таковы.

Т е о р е м а 28. *Если случайные величины ξ_1 и ξ_2 имеют абсолютно непрерывное совместное распределение с плотностью $f(x,\,y)$, то ξ_1 и ξ_2 в отдельности также имеют абсолютно непрерывное распределение с плотностями:*

$$f_{\xi_1}(x) = \int\limits_{-\infty}^{\infty} f(x,\,y)\,dy; \quad f_{\xi_2}(y) = \int\limits_{-\infty}^{\infty} f(x,\,y)\,dx.$$

Для $n > 2$ плотности случайных величин $\xi_1,\,\ldots,\,\xi_n$ по плотности их совместного распределения $f(x_1,\,\ldots,\,x_n)$ находятся интегрированием функции f по всем «лишним» координатам.

Д о к а з а т е л ь с т в о. Например, в силу равенств (16),

$$F_{\xi_1}(x_1) = \lim_{x_2\to+\infty} F_{\xi_1,\,\xi_2}(x_1,\,x_2) = \int\limits_{-\infty}^{x_1}\left(\int\limits_{-\infty}^{\infty} f(x,\,y)\,dy\right)dx = \int\limits_{-\infty}^{x_1} f_{\xi_1}(x)\,dx.$$

Аналогично устанавливается и справедливость второго равенства. □

§ 3. Примеры многомерных распределений

Приведём два наиболее употребительных примера абсолютно непрерывных многомерных распределений.

Равномерное распределение. Пусть $S \subset \mathbb{R}^n$ — борелевское множество с конечной лебеговой мерой $\lambda(S)$. Говорят, что вектор $(\xi_1, \ldots, \xi_n)$ имеет равномерное распределение в области S, если плотность совместного распределения $f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)$ постоянна в области S и равна нулю вне этой области:

$$f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \begin{cases} \dfrac{1}{\lambda(S)}, & \text{если } (x_1, \ldots, x_n) \in S, \\ 0, & \text{если } (x_1, \ldots, x_n) \notin S. \end{cases} \qquad (17)$$

Убедимся, что эта функция является плотностью распределения:

$$\int\limits_{\mathbb{R}^n} f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = \frac{1}{\lambda(S)} \int\limits_S dx_1 \ldots dx_n = \frac{1}{\lambda(S)}\, \lambda(S) = 1.$$

Как и в одномерном случае, вектор $(\xi_1, \ldots, \xi_n)$ с равномерным распределением в области S есть просто вектор координат точки, брошенной наудачу в область S.

Многомерное нормальное распределение. Пусть $\Sigma > 0$ — положительно определённая симметричная матрица $(n \times n)$, матрица Σ^{-1} — обратная к Σ, и $\vec{a} \in \mathbb{R}^n$ — n-мерный вектор-столбец. Транспонированный вектор 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 ковариаций Σ, если плотность совместного распределения $f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)$ равна

$$f_{\vec{\xi}}(\vec{x}) = \frac{1}{\sqrt{\det\Sigma}\,\left(\sqrt{2\pi}\right)^n} \exp\left\{-\frac{1}{2}\,(\vec{x}-\vec{a})^T \cdot \Sigma^{-1} \cdot (\vec{x}-\vec{a})\right\}.$$

Мы не будем проверять, что эта функция является плотностью совместного распределения, поскольку для этого требуется умение заменять переменные в многомерном интеграле. Выражение $(\vec{x}-\vec{a})^T\Sigma^{-1}(\vec{x}-\vec{a})$ в показателе экспоненты является квадратичной формой от переменных $(x_i - a_i)$. Действительно, для матрицы $B = \Sigma^{-1}$ с элементами b_{ij} имеем

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

Подробно с многомерным нормальным распределением мы познакомимся в курсе математической статистики, и там же выясним, что означают слова «с вектором средних $\vec{a}$ и матрицей ковариаций Σ».

В частном случае, когда Σ — диагональная матрица с элементами $\sigma_1^2, \ldots, \sigma_n^2$ на диагонали, совместная плотность превращается в произведение плотностей нормальных случайных величин:

$$\begin{aligned} f_{\vec{\xi}}(\vec{x}) &= \frac{1}{\sigma_1 \ldots \sigma_n \left(\sqrt{2\pi}\right)^n} \exp\left\{ -\frac{1}{2} \sum_{i=1}^{n} \frac{1}{\sigma_i^2}(x_i - a_i)^2 \right\} = \\ &= \frac{1}{\sigma_1\sqrt{2\pi}}\, e^{-\frac{(x_1-a)^2}{2\sigma_1^2}} \cdot \ldots \cdot \frac{1}{\sigma_n\sqrt{2\pi}}\, e^{-\frac{(x_n-a)^2}{2\sigma_n^2}}. \end{aligned}$$

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. Заметим (но не будем доказывать), что применение к набору случайных величин многих привычных нам функций не выводит нас из класса случайных величин. Интересующийся читатель может попробовать доказать, например, что сумма двух случайных величин есть снова случайная величина.

Следующие два примера показывают, что знания только частных распределений двух случайных величин недостаточно для отыскания распределения, например, суммы этих величин. 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 41. Рассмотрим две случайные величины ξ и η с одним и тем же распределением Бернулли с параметром $p = 1/2$ и следующей таблицей совместного распределения: для $0 \leqslant r \leqslant 1/2$ положим

$$\mathsf{P}(\xi = 0,\, \eta = 0) = r, \qquad \mathsf{P}(\xi = 0,\, \eta = 1) = \frac{1}{2} - r,$$
$$\mathsf{P}(\xi = 1,\, \eta = 0) = \frac{1}{2} - r, \qquad \mathsf{P}(\xi = 1,\, \eta = 1) = r,$$

Если $r = 0$, 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 $r = 1/2$, то $\mathsf{P}(\xi + \eta = 0) = \mathsf{P}(\xi + \eta = 2) = 1/2$, т. е. $\xi + \eta$ имеет невырожденное дискретное распределение, принимая значения 0 и 2 с равными вероятностями.

Взяв $r = 1/4$, получим $\mathsf{P}(\xi + \eta = 0) = 1/4$, $\mathsf{P}(\xi + \eta = 2) = 1/4$ и $\mathsf{P}(\xi + \eta = 1) = 1/2$, т. е. $\xi + \eta \in \mathrm{B}_{2, \frac{1}{2}}$.

Если взять $r = 1/3$, получим $\mathsf{P}(\xi + \eta = 0) = 1/3$, $\mathsf{P}(\xi + \eta = 1) = 1/3$ и $\mathsf{P}(\xi + \eta = 2) = 1/3$, т. е. $\xi + \eta$ принимает значения 1, 2 и 3 с равными вероятностями (это *не* биномиальное распределение).

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.

Пример 42. Пусть случайная величина 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*).

Распределение функции нескольких случайных величин может определяться их частными распределениями только если совместное распределение этих случайных величин определяется их частными распределениями. Так бывает для независимых случайных величин.

§ 5. Независимость случайных величин

Определение 31. Случайные величины $\xi_1, \ldots, \xi_n$ называют *независимыми* (в совокупности), если *для любого* набора борелевских множеств $B_1, \ldots, B_n \in \mathfrak{B}(\mathbb{R})$ имеет место равенство

$$\mathsf{P}(\xi_1 \in B_1, \ldots, \xi_n \in B_n) = \mathsf{P}(\xi_1 \in B_1) \cdot \ldots \cdot \mathsf{P}(\xi_n \in B_n).$$

Определение 32. Случайные величины $\xi_1, \ldots, \xi_n$ называют *попарно независимыми*, если независимы любые две из них.

Оба этих определения годятся не только для конечного набора случайных величин, но и для их бесконечной последовательности.

Замечание. Независимость случайных величин в совокупности влечёт попарную независимость. Достаточно в определении независимости в качестве «лишних» борелевских множеств взять $\mathbb{R}$.

П р и м е р 43. Вспомним пример Бернштейна 31. Свяжем с событиями A, B и C случайные величины ξ_1, ξ_2 и ξ_3 — индикаторы этих событий. Например, $\xi_1 = 1$, если A произошло, и $\xi_1 = 0$, если A не произошло. Случайные величины ξ_1, ξ_2 и ξ_3 независимы попарно (проверить), но зависимы в совокупности:

$$\mathsf{P}(\xi_1 = 1, \xi_2 = 1, \xi_3 = 1) = \mathsf{P}(A \cap B \cap C) = \frac{1}{4},$$
$$\mathsf{P}(\xi_1 = 1)\,\mathsf{P}(\xi_2 = 1)\,\mathsf{P}(\xi_3 = 1) = \mathsf{P}(A)\,\mathsf{P}(B)\,\mathsf{P}(C) = \frac{1}{8}.$$

Попарная независимость случайных величин встречается редко. Поэтому всюду, где мы будем употреблять термин «независимы», будет подразумеваться независимость в совокупности.

Определение независимости можно сформулировать в терминах функций распределения.

О п р е д е л е н и е 33. Случайные величины $\xi_1, \ldots, \xi_n$ независимы (в совокупности), если *для любых* $x_1, \ldots, x_n$ имеет место равенство

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n).$$

Описать независимость случайных величин с дискретным распределением можно с помощью таблицы их совместного распределения.

О п р е д е л е н и е 34. Случайные величины $\xi_1, \ldots, \xi_n$ с дискретным распределением независимы (в совокупности), если *для любых* чисел $a_1, \ldots, a_n$ имеет место равенство

$$\mathsf{P}(\xi_1 = a_1, \ldots, \xi_n = a_n) = \mathsf{P}(\xi_1 = a_1) \cdot \ldots \cdot \mathsf{P}(\xi_n = a_n).$$

У п р а ж н е н и е. Доказать, что из независимости в смысле определения 31 следует независимость в смысле определения 33.

У п р а ж н е н и е. Доказать, что для случайных величин с дискретным распределением определения 31 и 34 эквивалентны.

Для случайных величин с абсолютно непрерывными распределениями справедливо утверждение.

Т е о р е м а 29. *Случайные величины $\xi_1, \ldots, \xi_n$ с абсолютно непрерывными распределениями независимы (в совокупности) тогда и только тогда, когда плотность их совместного распределения существует и равна произведению плотностей, т. е. для любых $x_1, \ldots, x_n$ имеет место равенство: $f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = f_{\xi_1}(x_1) \cdot \ldots \cdot f_{\xi_n}(x_n)$.*

З а м е ч а н и е. Плотность распределения определяется с точностью до её значений на множестве нулевой лебеговой меры (распределение не меняется от изменения плотности на множестве нулевой меры). Поэтому равенство плотности совместного распределения и произведения плотностей можно понимать тоже как равенство «почти всюду».

Доказательство. Пусть случайные величины $\xi_1, \ldots, \xi_n$ независимы, т. е. для любых $x_1, \ldots, x_n$

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n).$$

Но произведение функций распределения записывается произведением интегралов, или одним n-мерным интегралом:

$$F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n) = \int\limits_{-\infty}^{x_1} f_{\xi_1}(s_1)\, ds_1 \cdot \ldots \cdot \int\limits_{-\infty}^{x_n} f_{\xi_n}(s_n)\, ds_n =$$

$$= \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} f_{\xi_1}(s_1) \ldots f_{\xi_n}(s_n)\, ds_1 \ldots ds_n = F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n).$$

Мы представили функцию совместного распределения в виде интеграла от плотности совместного распределения, которая оказалась равной произведению плотностей частных распределений.

Пусть теперь известно, что плотность совместного распределения существует и распадается в произведение плотностей. В таком случае функция совместного распределения распадается в произведение функций распределения:

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} f_{\xi_1}(s_1) \ldots f_{\xi_n}(s_n)\, ds_1 \ldots ds_n =$$

$$= F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n),$$

т. е. случайные величины независимы согласно определению 33. □

§ 6. Функции от двух случайных величин

Пусть ξ_1 и ξ_2 — случайные величины с плотностью совместного распределения $f_{\xi_1,\xi_2}(x_1, x_2)$, и задана борелевская функция $g : \mathbb{R}^2 \to \mathbb{R}$. Требуется найти функцию (а если существует, то и плотность) распределения случайной величины $\eta = g(\xi_1, \xi_2)$.

Пользуясь тем, что вероятность случайному вектору попасть в некоторую область можно вычислить как объем под графиком плотности распределения вектора над этой областью, сформулируем утверждение.

Теорема 30. *Пусть $x \in \mathbb{R}$, и область $D_x \subseteq \mathbb{R}^2$ состоит из точек $(u,\, v)$ таких, что $g(u,\, v) < x$. Тогда случайная величина $\eta = g(\xi_1,\, \xi_2)$ имеет функцию распределения*

$$F_\eta(x) = \mathsf{P}\big(g(\xi_1,\, \xi_2) < x\big) = \mathsf{P}\big((\xi_1,\, \xi_2) \in D_x\big) = \iint\limits_{D_x} f_{\xi_1,\xi_2}(u,\, v)\, du\, dv.$$

Далее в этой главе предполагается, что случайные величины ξ_1 и ξ_2 независимы, т. е. $f_{\xi_1,\xi_2}(u,\, v) \equiv f_{\xi_1}(u)\, f_{\xi_2}(v)$. В этом случае распределение величины $g(\xi_1,\, \xi_2)$ полностью определяется частными распределениями величин ξ_1 и ξ_2.

Следствие 9 (формула свёртки). *Если случайные величины ξ_1 и ξ_2 независимы и имеют абсолютно непрерывные распределения с плотностями $f_{\xi_1}(u)$ и $f_{\xi_2}(v)$, то плотность распределения суммы $\xi_1 + \xi_2$ существует и равна «свёртке» плотностей f_{ξ_1} и f_{ξ_2}:*

$$f_{\xi_1+\xi_2}(t) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(u)\; f_{\xi_2}(t-u)\; du = \int\limits_{-\infty}^{\infty} f_{\xi_2}(u)\; f_{\xi_1}(t-u)\; du. \tag{18}$$

Доказательство. Воспользуемся утверждением теоремы 30 для борелевской функции $g(u,v) = u+v$. Интегрирование по двумерной области $D_x = \{(u,\, v) \mid u+v < x\}$ можно заменить последовательным вычислением двух интегралов: наружного — по переменной u, меняющейся в пределах от $-\infty$ до $+\infty$, и внутреннего — по переменной v, которая при каждом u должна быть меньше, чем $x-u$. Поэтому

$$F_{\xi_1+\xi_2}(x) = \iint\limits_{D_x} f_{\xi_1}(u)\; f_{\xi_2}(v)\, dv\, du = \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{x-u} f_{\xi_1}(u) f_{\xi_2}(v)\, dv \right) du.$$

Сделаем в последнем интеграле замену переменной v на t так: $v = t-u$. При этом $v \in (-\infty,\, x-u)$ перейдёт в $t \in (-\infty,\, x)$, $dv = dt$. В полученном интеграле меняем порядок интегрирования:

$$F_{\xi_1+\xi_2}(x) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{x} f_{\xi_1}(u)\; f_{\xi_2}(t-u)\; dt\; du = \int\limits_{-\infty}^{x} \left(\int\limits_{-\infty}^{\infty} f_{\xi_1}(u)\; f_{\xi_2}(t-u)\, du \right) dt.$$

Итак, мы представили функцию распределения $F_{\xi_1+\xi_2}(x)$ в виде интеграла от $-\infty$ до x от плотности распределения $f_{\xi_1+\xi_2}(t)$ из формулы свёртки (18). □

Следствие 9 не только предлагает формулу для вычисления плотности распределения суммы, но и утверждает, что сумма двух независимых случайных величин с абсолютно непрерывными распределениями также имеет абсолютно непрерывное распределение.

Упражнение. Для тех, кто уже ничему не удивляется: привести пример двух случайных величин с абсолютно непрерывными распределениями таких, что их сумма имеет вырожденное распределение.

Если даже одна из двух независимых случайных величин имеет дискретное, а вторая — абсолютно непрерывное распределение, то их сумма тоже имеет абсолютно непрерывное распределение:

Упражнение. Пусть величина ξ имеет таблицу распределения $\mathsf{P}(\xi = a_i) = p_i$, а η имеет плотность распределения $f_\eta(x)$, и эти величины независимы. Доказать, что $\xi + \eta$ имеет плотность распределения $f_{\xi+\eta}(x) = \sum p_i f_\eta(x - a_i)$. Для вычисления функции распределения суммы использовать формулу полной вероятности.

§ 7. Примеры использования формулы свёртки

Пример 44. Пусть независимые случайные величины ξ и η имеют стандартное нормальное распределение. Докажем, что их сумма имеет нормальное распределение с параметрами $a = 0$ и $\sigma^2 = 2$.

Доказательство. По формуле свёртки, плотность суммы равна

$$f_{\xi+\eta}(x) = \int\limits_{-\infty}^{\infty} \frac{1}{2\pi}\, e^{-u^2/2}\, e^{-(x-u)^2/2}\, du = \int\limits_{-\infty}^{\infty} \frac{1}{2\pi}\, e^{-\left(u^2 + \frac{x^2}{2} - xu\right)}\, du =$$

$$= e^{-x^2/4} \int\limits_{-\infty}^{\infty} \frac{1}{2\pi}\, e^{-\left(u - \frac{x}{2}\right)^2}\, du = \frac{1}{2\sqrt{\pi}}\, e^{-x^2/4} \int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}}\, e^{-v^2}\, dv = \frac{e^{-x^2/4}}{2\sqrt{\pi}}.$$

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 распределения суммы есть плотность нормального распределения с параметрами 0 и 2. $\square$

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 следующих утверждениях перечислены практически все устойчивые распределения.

Лемма 1. *Пусть случайные величины* $\xi \in \Pi_\lambda$ *и* $\eta \in \Pi_\mu$ *независимы. Тогда* $\xi + \eta \in \Pi_{\lambda+\mu}$.

Доказательство. Найдём таблицу распределения суммы. Для любого целого $k \geqslant 0$

$$\begin{aligned}\mathsf{P}(\xi+\eta=k) &= \sum_{i=0}^{k} \mathsf{P}(\xi=i,\, \eta=k-i) = \\ &= \sum_{i=0}^{k} \mathsf{P}(\xi=i)\cdot\mathsf{P}(\eta=k-i) = \sum_{i=0}^{k} \frac{\lambda^i}{i!}\, e^{-\lambda}\cdot\frac{\mu^{k-i}}{(k-i)!}\, e^{-\mu} = \\ &= e^{-(\lambda+\mu)}\frac{1}{k!}\,\sum_{i=0}^{k}\frac{k!}{i!\,(k-i)!}\,\lambda^i\mu^{k-i} = e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^k}{k!}.\end{aligned}$$

В последнем равенстве мы воспользовались биномом Ньютона. □

Лемма 2. *Пусть случайные величины* $\xi \in \mathrm{B}_{n,p}$ *и* $\eta \in \mathrm{B}_{m,p}$ *независимы. Тогда* $\xi + \eta \in \mathrm{B}_{n+m,p}$.

Смысл леммы 2 совершенно понятен: складывая количество успехов в первых n и в следующих m независимых испытаниях одной и той же схемы Бернулли, получаем количество успехов в $n+m$ испытаниях. Полезно доказать это утверждение аналогично тому, как мы доказали лемму 1.

Лемма 3. *Пусть случайные величины* $\xi \in \mathrm{N}_{a_1,\sigma_1^2}$ *и* $\eta \in \mathrm{N}_{a_2,\sigma_2^2}$ *независимы. Тогда* $\xi + \eta \in \mathrm{N}_{a_1+a_2,\,\sigma_1^2+\sigma_2^2}$.

Лемма 4. *Пусть случайные величины* $\xi \in \Gamma_{\alpha,\lambda_1}$ *и* $\eta \in \Gamma_{\alpha,\lambda_2}$ *независимы. Тогда* $\xi + \eta \in \Gamma_{\alpha,\lambda_1+\lambda_2}$.

Эти утверждения мы докажем позднее, используя аппарат характеристических функций, хотя при некотором терпении можно попробовать доказать их напрямую с помощью формулы свёртки.

Показательное распределение не устойчиво по суммированию, однако оно является частным случаем гамма-распределения, которое уже устойчиво относительно суммирования. Докажем частный случай леммы 4.

Лемма 5. *Пусть независимые случайные величины* $\xi_1, \ldots, \xi_n$ *имеют показательное распределение* E_α. *Тогда* $\xi_1 + \ldots + \xi_n \in \Gamma_{\alpha,n}$.

Доказательство. Докажем утверждение по индукции. При $n=1$ оно верно в силу равенства $\mathrm{E}_\alpha = \Gamma_{\alpha,1}$. Пусть утверждение леммы справедливо для $n=k-1$. Докажем, что оно верно и для $n=k$. По предположению индукции, $S_{k-1} = \xi_1 + \ldots + \xi_{k-1}$ имеет распределение $\Gamma_{\alpha,k-1}$, т. е. плотность распределения величины S_{k-1} равна

$$f_{S_{k-1}}(x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ \dfrac{\alpha^{k-1}}{(k-2)!}\, x^{k-2} e^{-\alpha x}, & \text{если } x > 0. \end{cases}$$

Тогда по формуле свёртки плотность суммы $S_k = \xi_1 + \ldots + \xi_k$ равна

$$f_{S_k}(x) = \int\limits_{-\infty}^{\infty} f_{S_{k-1}}(u) f_{\xi_k}(x-u)\, du = \int\limits_{0}^{\infty} \frac{\alpha^{k-1}}{(k-2)!}\, u^{k-2} e^{-\alpha u} f_{\xi_k}(x-u)\, du.$$

Так как $f_{\xi_k}(x-u) = 0$ при $x - u < 0$, т. е. при $u > x$, то плотность под интегралом отлична от нуля, только если переменная интегрирования изменяется в пределах $0 \leqslant u \leqslant x$ при $x > 0$. При $x \leqslant 0$ подынтегральная функция равна нулю. При $x > 0$ имеем

$$f_{S_k}(x) = \int\limits_{0}^{x} \frac{\alpha^{k-1}}{(k-2)!}\, u^{k-2} e^{-\alpha u}\, \alpha\, e^{-\alpha(x-u)}\, du = \frac{\alpha^k}{(k-1)!}\, x^{k-1} e^{-\alpha x}.$$

Поэтому $S_k \in \Gamma_{\alpha, k}$, что и требовалось доказать. □

П р и м е р 45. 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 $\xi, \eta \in \mathrm{U}_{0,1}$ — независимые случайные величины. Случайные величины ξ и η можно считать координатами точки, брошенной наудачу в единичный квадрат.

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. 13 треугольник (при $0 < t \leqslant 1$) либо пятиугольник (при $1 < t \leqslant 2$). Получим функцию распределения и плотность распределения суммы двух независимых равномерно распределённых на отрезке $[0,\, 1]$ случайных величин:

$$F_{\xi+\eta}(t) = \begin{cases} 0, & t \leqslant 0, \\ \frac{1}{2}\, t^2, & 0 < t \leqslant 1, \\ 1 - \frac{1}{2}\,(2-t)^2, & 1 < t \leqslant 2, \\ 1, & t > 2; \end{cases} \qquad f_{\xi+\eta}(t) = \begin{cases} 0, & t \not\in (0,\, 2), \\ t, & 0 < t \leqslant 1, \\ 2-t, & 1 < t \leqslant 2. \end{cases}$$

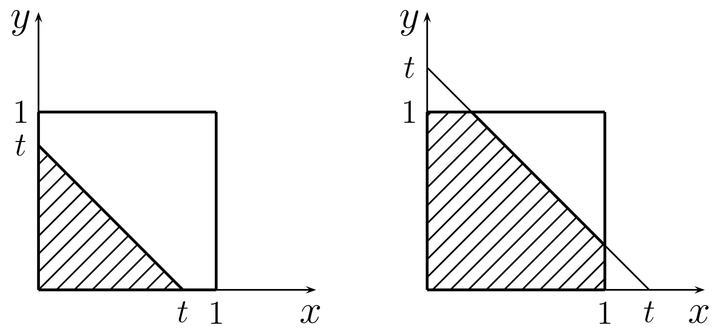

Рис. 13. Область $\{\xi+\eta<t\}$ в зависимости от t

Полученное распределение называется «треугольным распределением» Симпсона. Видим, что распределение суммы независимых случайных величин с равномерным распределением не является равномерным.

П р и м е р 46. Найдём функцию и плотность распределения частного двух независимых случайных величин ξ и η, имеющих показательное распределение с параметром 1.

При $x>0$ по теореме 30 имеем

$$\mathsf{P}\left(\frac{\eta}{\xi}<x\right)=\iint\limits_{D_x} f_\xi(u)f_\eta(v)\,du\,dv,$$

где область D_x есть множество точек $(u,\,v)$ таких, что $\dfrac{v}{u}<x$. При этом достаточно ограничиться положительными значениями u и v: показательно распределённые случайные величины могут принимать отрицательные значения лишь с нулевой вероятностью.

Вычислим интеграл по области $D_x=\{(u,v)\,|\,0<u<\infty,\;0<v<ux\}$:

$$\mathsf{P}\left(\frac{\eta}{\xi}<x\right)=\int\limits_0^\infty\left(\int\limits_0^{ux}e^{-u}e^{-v}\,dv\right)du=1-\frac{1}{x+1}.$$

У п р а ж н е н и е. Провести вычисления и получить ответ.

Таким образом, функция и плотность распределения частного имеют вид

$$F(x)=\begin{cases}1-\dfrac{1}{x+1}, & x>0,\\ 0, & x\leqslant 0;\end{cases}\qquad f(x)=\begin{cases}\dfrac{1}{(x+1)^2}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

ГЛАВА VIII

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ РАСПРЕДЕЛЕНИЙ

Если я имею одинаковые шансы на получение a или b, то цена моему ожиданию равна $(a+b)/2$.

Христиан Гюйгенс. О расчётах в азартной игре

§ 1. Математическое ожидание случайной величины

Определение 35. *Математическим ожиданием* $\mathsf{E}\,\xi$ случайной величины ξ с дискретным распределением называется *число*

$$\mathsf{E}\,\xi = \sum_k a_k p_k = \sum_k a_k \mathsf{P}(\xi = a_k),$$

если данный ряд абсолютно сходится, т. е. если $\sum |a_i| p_i < \infty$. В противном случае говорят, что математическое ожидание *не существует*.

Определение 36. *Математическим ожиданием* $\mathsf{E}\,\xi$ случайной величины ξ с абсолютно непрерывным распределением с плотностью распределения $f_\xi(x)$ называется *число*

$$\mathsf{E}\,\xi = \int\limits_{-\infty}^{\infty} x\, f_\xi(x)\, dx,$$

если этот интеграл абсолютно сходится, т. е. если

$$\int\limits_{-\infty}^{\infty} |x|\, f_\xi(x)\, dx < \infty.$$

Математическое ожидание (иначе называемое средним значением или первым моментом) имеет простой физический смысл: если на прямой разместить единичную массу, поместив в точки a_i массу p_i (для дискретного распределения) или «размазав» её с плотностью $f_\xi(x)$ (для абсолютно непрерывного распределения), то точка $\mathsf{E}\,\xi$ будет координатой «центра тяжести» прямой.

Пример 47. 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.

Пример 48. Пусть случайная величина ξ — координата точки, брошенной наудачу на отрезок $[a,\, b]$. Тогда

$$\mathsf{E}\,\xi = \int\limits_a^b x \cdot \frac{1}{b-a}\,dx = \frac{x^2}{2(b-a)}\bigg|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

Центр тяжести равномерного распределения есть середина отрезка.

§ 2. Свойства математического ожидания

Во всех свойствах предполагается, что рассматриваемые математические ожидания существуют.

(E1) Для произвольной борелевской функции $g : \mathbb{R} \to \mathbb{R}$

$$\mathsf{E}\,g(\xi) = \begin{cases} \sum\limits_k g(a_k)\mathsf{P}(\xi = a_k), & \text{если распределение } \xi \text{ дискретно и ряд абсолютно сходится;} \\ \int\limits_{-\infty}^{\infty} g(x) f_\xi(x)\,dx, & \text{если распределение } \xi \text{ абсолютно непрерывно и интеграл абсолютно сходится.} \end{cases}$$

Доказательство. Мы докажем это свойство (как и почти все дальнейшие) только для дискретного распределения. Пусть $g(\xi)$ принимает значения $c_1,\, c_2,\, \ldots$ с вероятностями

$$\mathsf{P}(g(\xi) = c_m) = \sum_{k:\, g(a_k)=c_m} \mathsf{P}(\xi = a_k)\,.$$

Тогда

$$\mathsf{E}\,g(\xi) = \sum_m c_m \mathsf{P}(g(\xi) = c_m) = \sum_m c_m \sum_{k:\, g(a_k)=c_m} \mathsf{P}(\xi = a_k) =$$
$$= \sum_m \sum_{k:\, g(a_k)=c_m} g(a_k)\,\mathsf{P}(\xi = a_k) = \sum_k g(a_k)\,\mathsf{P}(\xi = a_k)\,. \qquad \square$$

Следствие 10. *Математическое ожидание* ξ *существует тогда и только тогда, когда* $\mathsf{E}\,|\xi| < \infty$.

Доказательство. Условием существование математического ожидания является абсолютная сходимость ряда или интеграла в определениях 35 и 36. Это в точности есть условие $\mathsf{E}\,g(\xi) < \infty$ при $g(x) = |x|$. □

(E2) Математическое ожидание постоянной равно ей самой: $\mathsf{E}\,c = c$.

(E3) Постоянную можно вынести за знак математического ожидания:

$$\mathsf{E}\,(c\,\xi) = c\,\mathsf{E}\,\xi.$$

Доказательство. Следует из свойства (E1) при $g(x) = c\,x$. □

(E4) Математическое ожидание суммы *любых* случайных величин равно сумме их математических ожиданий, если только эти математические ожидания существуют:

$$\mathsf{E}\,(\xi + \eta) = \mathsf{E}\,\xi + \mathsf{E}\,\eta.$$

Доказательство. Пусть случайные величины ξ и η имеют дискретные распределения со значениями x_k и y_n соответственно. Для борелевской функции $g : \mathbb{R}^2 \to \mathbb{R}$ можно доказать свойство, аналогичное (E1) (*сделать это*). Воспользуемся этим свойством для $g(x, y) = x + y$:

$$\begin{aligned}
\mathsf{E}\,(\xi + \eta) &= \sum_{k,\,n} (x_k + y_n)\mathsf{P}(\xi = x_k,\, \eta = y_n) = \\
&= \sum_k x_k \sum_n \mathsf{P}(\xi = x_k,\, \eta = y_n) + \sum_n y_n \sum_k \mathsf{P}(\xi = x_k,\, \eta = y_n) = \\
&= \sum_k x_k\,\mathsf{P}(\xi = x_k) + \sum_n y_n\,\mathsf{P}(\eta = y_n) = \mathsf{E}\,\xi + \mathsf{E}\,\eta. \qquad \square
\end{aligned}$$

(E5) Если $\xi \geqslant 0$ п. н., т. е. если $\mathsf{P}(\xi \geqslant 0) = 1$, то $\mathsf{E}\,\xi \geqslant 0$.

Упражнение. Доказать для дискретного и для абсолютно непрерывного распределений.

Замечание. Сокращение «п. н.» читается как «почти наверное» и означает «с вероятностью 1». По определению, математическое ожидание — это числовая характеристика *распределения*. Распределение же не изменится от изменения случайной величины на множестве нулевой вероятности. Поэтому, например, даже если $\xi(\omega) \geqslant 0$ не при всех ω, а на множестве единичной вероятности, математическое ожидание ξ всё равно неотрицательно.

(E6) Если $\xi \geqslant 0$ п. н. и при этом $\mathsf{E}\,\xi = 0$, то $\xi = 0$ п. н.

Доказательство. Это свойство мы докажем, заранее предполагая, что ξ имеет дискретное распределение с неотрицательными значениями $a_k \geqslant 0$. Равенство $\mathsf{E}\xi = \sum a_k p_k = 0$ означает, что все слагаемые в этой сумме равны нулю, т. е. все вероятности p_k нулевые, кроме вероятности, соответствующей значению $a_k = 0$. □

Из свойств (E5) и (E6) следуют полезные утверждения.

Следствие 11. *Если* $\xi \leqslant \eta$ *п. н., то* $\mathsf{E}\xi \leqslant \mathsf{E}\eta$.

Следствие 12. *Если* $a \leqslant \xi \leqslant b$ *п. н., то* $a \leqslant \mathsf{E}\xi \leqslant b$.

(E7) Если ξ и η независимы и их математические ожидания существуют, то $\mathsf{E}(\xi\eta) = \mathsf{E}\xi\,\mathsf{E}\eta$.

Доказательство. В дискретном случае

$$\begin{aligned} \mathsf{E}(\xi\eta) &= \sum_{k,\,n} (x_k\,y_n)\,\mathsf{P}(\xi = x_k,\, \eta = y_n) = \\ &= \sum_k x_k\,\mathsf{P}(\xi = x_k)\,\sum_n y_n\,\mathsf{P}(\eta = y_n) = \mathsf{E}\xi\,\mathsf{E}\eta. \quad \square \end{aligned}$$

Замечание. 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 49. Пусть ξ принимает значения 0 и ± 1 с вероятностями по $1/3$ каждое, 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 $\mathsf{E}(\xi\eta) = \mathsf{E}\xi\,\mathsf{E}\eta$.

Пример 50. Пусть $\varphi \in \mathrm{U}_{0,\,2\pi}$, и пусть $\xi = \cos\varphi$ и $\eta = \sin\varphi$ — заведомо зависимые случайные величины. Например:

$$\mathsf{P}\Big(\xi > \frac{1}{\sqrt{2}},\, \eta > \frac{1}{\sqrt{2}}\Big) = 0 \neq \mathsf{P}\Big(\xi > \frac{1}{\sqrt{2}}\Big)\mathsf{P}\Big(\eta > \frac{1}{\sqrt{2}}\Big) > 0.$$

Но математическое ожидание их произведения равно произведению их математических ожиданий из-за симметричности распределений ξ, η и $\xi\eta$ относительно нуля. Действительно, по свойству (E1) имеем

$$\mathsf{E}\xi = \int_0^{2\pi} \frac{1}{2\pi}\,\cos x\,dx = 0, \quad \mathsf{E}\eta = \int_0^{2\pi} \frac{1}{2\pi}\,\sin x\,dx = 0,$$

$$\mathsf{E}\xi\eta = \int_0^{2\pi} \frac{1}{2\pi}\,\cos x\,\sin x\,dx = 0 = \mathsf{E}\xi\,\mathsf{E}\eta.$$

§ 3. Дисперсия и моменты старших порядков

Определение 37. Пусть $\mathsf{E}\,|\xi|^k < \infty$. Число $\mathsf{E}\,\xi^k$ называется моментом порядка k или k-м моментом случайной величины ξ, число $\mathsf{E}\,|\xi|^k$ называется абсолютным k-м моментом, $\mathsf{E}\,(\xi - \mathsf{E}\,\xi)^k$ называется центральным k-м моментом, и $\mathsf{E}\,|\xi - \mathsf{E}\,\xi|^k$ — 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 51. 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:

$$\begin{aligned}
\mathsf{E}\,\xi &= 0 \cdot 0{,}99999 + 100 \cdot 0{,}00001 = 0{,}001,\\
\mathsf{E}\,\xi^2 &= 0^2 \cdot 0{,}99999 + 100^2 \cdot 0{,}00001 = 0{,}1,\\
\mathsf{E}\,\xi^4 &= 0^4 \cdot 0{,}99999 + 100^4 \cdot 0{,}00001 = 1\,000,\\
\mathsf{E}\,\xi^6 &= 0^6 \cdot 0{,}99999 + 100^6 \cdot 0{,}00001 = 10\,000\,000.
\end{aligned}$$

Пример 52. Дисперсия $\mathsf{D}\,\xi = \mathsf{E}\,(\xi - \mathsf{E}\,\xi)^2$ есть «среднее значение квадрата отклонения случайной величины ξ от своего 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 38. Число $\sigma = \sqrt{\mathsf{D}\,\xi}$ называют *среднеквадратическим отклонением* случайной величины ξ.

Чтобы прояснить связь моментов разных порядков, докажем несколько неравенств. Во-первых, получим очевидное утверждение, обеспечивающее существование моментов меньших порядков, если существуют моменты более высокого порядка.

Теорема 31. *Если существует момент порядка $t > 0$ случайной величины ξ, то существуют и её моменты порядка s при $0 < s < t$.*

Доказательство. Для любого числа x верно неравенство

$$|x|^s \leqslant \max\{\,|x|^t,\ 1\} \leqslant |x|^t + 1.$$

Действительно, $|x|^s \leqslant |x|^t$ при $|x| > 1$, и $|x|^s \leqslant 1$ при $|x| \leqslant 1$.

Из этого неравенства следует, что $|\xi(\omega)|^s \leqslant |\xi(\omega)|^t + 1$ для всех ω. Но следствие 11 позволяет из неравенства для случайных величин получить такое же неравенство для их математических ожиданий:

$$\mathsf{E}\,|\xi|^s \leqslant \mathsf{E}\,|\xi|^t + 1.$$

Момент порядка t существует, т. е. $\mathsf{E}\,|\xi|^t < \infty$. Поэтому и $\mathsf{E}\,|\xi|^s < \infty$. □

Докажем ещё одно чрезвычайно полезное неравенство.

Теорема 32 (неравенство Йенсена[15]). *Пусть вещественнозначная функция g выпукла («выпукла вниз», т. е. её надграфик есть выпуклое множество). Тогда для любой случайной величины ξ с конечным первым моментом верно неравенство:* $\mathsf{E}\,g(\xi) \geqslant g(\mathsf{E}\,\xi)$. *Для вогнутых функций знак неравенства меняется на противоположный.*

Доказательство. Нам понадобится следующее свойство.

Лемма 6. *Пусть функция g выпукла. Тогда для всякого x_0 найдётся число $c(x_0)$ такое, что при всех x*

$$g(x) \geqslant g(x_0) + c(x_0)(x - x_0).$$

Это свойство очевидно и означает, что график выпуклой функции лежит полностью выше любой из касательных к этому графику.

Возьмём в условиях леммы $x_0 = \mathsf{E}\,\xi$, $x = \xi$. Тогда

$$g(\xi) \geqslant g(\mathsf{E}\,\xi) + c(\mathsf{E}\,\xi)(\xi - \mathsf{E}\,\xi).$$

Вычислим математическое ожидание обеих частей неравенства. Так как $\mathsf{E}\,(\xi - \mathsf{E}\,\xi) = 0$, и неравенство между математическими ожиданиями сохраняется по следствию 11, то $\mathsf{E}\,g(\xi) \geqslant g(\mathsf{E}\,\xi)$. □

Следующее неравенство связывает моменты разных порядков.

Следствие 13. *Если* $\mathsf{E}\,|\xi|^t < \infty$, *то для любого* $0 < s < t$

$$\sqrt[s]{\mathsf{E}\,|\xi|^s} \leqslant \sqrt[t]{\mathsf{E}\,|\xi|^t}$$

Доказательство. Поскольку $0 < s < t$, то $g(x) = |x|^{t/s}$ — выпуклая функция. По неравенству Йенсена для $\eta = |\xi|^s$,

$$(\mathsf{E}\,|\xi|^s)^{t/s} = (\mathsf{E}\,\eta)^{t/s} = g(\mathsf{E}\,\eta) \leqslant \mathsf{E}\,g(\eta) = \mathsf{E}\,|\eta|^{t/s} = \mathsf{E}\,|\xi|^{s\cdot t/s} = \mathsf{E}\,|\xi|^t.$$

Осталось извлечь из обеих частей корень степени t. □

[15] Johan Ludwig William Valdemar Jensen (8.05.1859—5.03.1925, Denmark).

Из неравенства Йенсена вытекают, например, неравенства:

$$\mathsf{E}\, e^{\xi} \geqslant e^{\mathsf{E}\,\xi}, \qquad \mathsf{E}\,\xi^2 \geqslant (\mathsf{E}\,\xi)^2, \qquad \mathsf{E}\,|\xi| \geqslant |\mathsf{E}\,\xi|,$$
$$\mathsf{E}\,\ln\xi \leqslant \ln(\mathsf{E}\,\xi), \qquad \mathsf{E}\,\frac{1}{\xi} \geqslant \frac{1}{\mathsf{E}\,\xi}, \qquad \mathsf{E}\,\sqrt{\xi} \leqslant \sqrt{\mathsf{E}\,\xi}.$$

Последние три неравенства верны для положительных ξ.

§ 4. Свойства дисперсии

Свойства дисперсии следуют из соответствующих свойств математического ожидания. Заметим, что из существования второго момента следует существование математического ожидания случайной величины и конечность дисперсии. Во всех свойствах ниже предполагается существование вторых моментов случайных величин.

(D1) Дисперсия может быть вычислена по формуле: $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2$.

Д о к а з а т е л ь с т в о. Положим 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: $\mathsf{D}\,(c\xi) = c^2\,\mathsf{D}\,\xi$.

У п р а ж н е н и е. Доказать.

(D3) Дисперсия всегда неотрицательна: $\mathsf{D}\,\xi \geqslant 0$. Дисперсия обращается в нуль лишь для вырожденного распределения: если $\mathsf{D}\,\xi = 0$, то $\xi = const$ п. н. и наоборот.

Д о к а з а т е л ь с т в о. Дисперсия есть математическое ожидание почти наверное неотрицательной случайной величины $(\xi - \mathsf{E}\,\xi)^2$, и неотрицательность дисперсии следует из свойства (E5). Далее, по свойству (E6) из равенства дисперсии нулю вытекает $(\xi - \mathsf{E}\,\xi)^2 = 0$ п. н., т. е. $\xi = \mathsf{E}\,\xi$ п. н. И наоборот, если $\xi = c$ п. н., то $\mathsf{D}\,\xi = \mathsf{E}\,(c - \mathsf{E}\,c)^2 = 0$. □

(D4) Дисперсия не зависит от сдвига случайной величины на постоянную: $\mathsf{D}\,(\xi + c) = \mathsf{D}\,\xi$.

У п р а ж н е н и е. Доказать.

(D5) Если ξ и η независимы, то $\mathsf{D}\,(\xi + \eta) = \mathsf{D}\,\xi + \mathsf{D}\,\eta$.

Д о к а з а т е л ь с т в о. Действительно,

$$\mathsf{D}\,(\xi + \eta) = \mathsf{E}\,(\xi + \eta)^2 - (\mathsf{E}\,(\xi + \eta))^2 =$$
$$= \mathsf{E}\,\xi^2 + \mathsf{E}\,\eta^2 + 2\mathsf{E}\,(\xi\eta) - (\mathsf{E}\,\xi)^2 - (\mathsf{E}\,\eta)^2 - 2\mathsf{E}\,\xi\mathsf{E}\,\eta = \mathsf{D}\,\xi + \mathsf{D}\,\eta,$$

так как математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий. □

Замечание. См. замечание 2.

Следствие 14. *Если* ξ *и* η *независимы, то*

$$\mathsf{D}\,(\xi-\eta)=\mathsf{D}\,(\xi+\eta)=\mathsf{D}\,\xi+\mathsf{D}\,\eta.$$

Доказательство. Из свойств (D5) 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 15. *Для произвольных случайных величин* ξ *и* η *с конечными вторыми моментами имеет место равенство*

$$\mathsf{D}\,(\xi+\eta)=\mathsf{D}\,\xi+\mathsf{D}\,\eta+2\big(\mathsf{E}\,(\xi\eta)-\mathsf{E}\,\xi\,\mathsf{E}\,\eta\big).$$

(D6) Минимум среднеквадратического отклонения случайной величины ξ от точек числовой прямой есть среднеквадратическое отклонение ξ от её математического ожидания: $\mathsf{D}\,\xi=\mathsf{E}\,(\xi-\mathsf{E}\,\xi)^2=\min\limits_a \mathsf{E}\,(\xi-a)^2$.

Доказательство. Сравним величину $\mathsf{E}\,(\xi-a)^2$ с дисперсией:

$$\mathsf{E}\,(\xi-a)^2=\mathsf{E}\,\big((\xi-\mathsf{E}\,\xi)+(\mathsf{E}\,\xi-a)\big)^2=$$
$$=\mathsf{D}\,\xi+\big(\mathsf{E}\,\xi-a\big)^2+2(\mathsf{E}\,\xi-\mathsf{E}\,\xi)\,\big(\mathsf{E}\,\xi-a\big)=\mathsf{D}\,\xi+\big(\mathsf{E}\,\xi-a\big)^2\geqslant\mathsf{D}\,\xi,$$

и последнее неравенство превращается в равенство лишь при $a=\mathsf{E}\,\xi$. $\square$

§ 5. Математические ожидания и дисперсии стандартных распределений

Пример 53 (вырожденное распределение I_c). Математическое ожидание и дисперсию этого распределения мы знаем из свойств (E2) и (D3): $\mathsf{E}\,c=c$, $\mathsf{D}\,c=0$.

Пример 54 (распределение Бернулли B_p). Вычислим два момента и дисперсию: $\mathsf{E}\,\xi=1\cdot p+0\cdot q=p$; $\mathsf{E}\,\xi^2=1^2\cdot p+0^2\cdot q=p$; $\mathsf{D}\,\xi=\mathsf{E}\,\xi^2-(\mathsf{E}\,\xi)^2=p-p^2=pq$.

Пример 55 (биномиальное распределение $\mathrm{B}_{n,p}$). Используем свойство устойчивости биномиального распределения относительно суммирования — лемму 2 (с. 86). Возьмём на каком-нибудь вероятностном пространстве n независимых случайных величин $\xi_1,\ldots,\xi_n$ 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) получаем

$$\mathsf{E}\,S_n=\sum_{i=1}^n\mathsf{E}\,\xi_i=n\mathsf{E}\,\xi_1=np.$$

А 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 $\xi \Subset \mathrm{B}_{n,p}$.

Пример 56 (геометрическое распределение G_p). Вычислим математическое ожидание ξ:

$$\begin{aligned}\mathsf{E}\,\xi &= \sum_{k=1}^{\infty} k\,p\,q^{k-1} = p\sum_{k=1}^{\infty} k\,q^{k-1} = p\sum_{k=1}^{\infty} \frac{dq^k}{dq} = \\ &= p\,\frac{d}{dq}\left(\sum_{k=1}^{\infty} q^k\right) = p\,\frac{d}{dq}\left(\frac{q}{1-q}\right) = p\,\frac{1}{(1-q)^2} = \frac{1}{p}.\end{aligned}$$

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:

$$\mathsf{D}\,\xi = \mathsf{E}\,\xi(\xi-1) + \mathsf{E}\,\xi - (\mathsf{E}\,\xi)^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q-1+p}{p^2} = \frac{q}{p^2}.$$

Пример 57 (распределение Пуассона Π_λ). Вычислим математическое ожидание ξ:

$$\begin{aligned}\mathsf{E}\,\xi &= \sum_{k=0}^{\infty} k\,\frac{\lambda^k}{k!}\,e^{-\lambda} = e^{-\lambda}\sum_{k=1}^{\infty} k\,\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \\ &= \lambda e^{-\lambda}\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda e^{-\lambda}e^{\lambda} = \lambda.\end{aligned}$$

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 $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2 = \lambda$.

Пример 58 (равномерное распределение $\mathrm{U}_{a,b}$). Математическое ожидание $\mathsf{E}\,\xi = \dfrac{a+b}{2}$ найдено в примере 48. Вычислим второй момент:

$$\mathsf{E}\,\xi^2 = \int\limits_{-\infty}^{\infty} x^2 f_\xi(x)\,dx = \int\limits_a^b x^2\,\frac{1}{b-a}\,dx = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}.$$

Дисперсия равна $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2 = (b-a)^2/\,12$.

Пример 59 (стандартное нормальное распределение $\mathrm{N}_{0,1}$). Математическое ожидание этого распределения существует, поскольку

$$\mathsf{E}\,|\xi| = \frac{2}{\sqrt{2\pi}}\int\limits_0^{\infty} xe^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}}\int\limits_0^{\infty} e^{-x^2/2}\,d(x^2/2) = \frac{2}{\sqrt{2\pi}} < \infty.$$

Математическое ожидание ξ равно нулю:

$$\mathsf{E}\,\xi = \int\limits_{-\infty}^{\infty} x f_\xi(x)\,dx = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} x\,e^{-x^2/2}\,dx = 0,$$

так как под сходящимся интегралом стоит нечётная функция. Далее,

$$\mathsf{E}\,\xi^2 = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} x^2\,e^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}}\int\limits_0^{\infty} x^2\,e^{-x^2/2}\,dx = -\frac{2}{\sqrt{2\pi}}\int\limits_0^{\infty} x\,de^{-x^2/2} =$$

$$= -\frac{2x}{\sqrt{2\pi}}\,e^{-x^2/2}\Big|_0^{\infty} + 2\int\limits_0^{\infty}\frac{1}{\sqrt{2\pi}}\,e^{-x^2/2}\,dx = 0 + \int\limits_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\,e^{-x^2/2}\,dx = 1.$$

Поэтому $\mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2 = 1 - 0 = 1$.

Пример 60 (нормальное распределение N_{a,σ^2}). Мы знаем, что если $\xi \in \mathrm{N}_{a,\sigma^2}$, то $\eta = \dfrac{\xi - a}{\sigma} \in \mathrm{N}_{0,1}$. Математическое ожидание $\mathsf{E}\,\eta = 0$ и дисперсия $\mathsf{D}\,\eta = 1$ стандартного нормального распределения вычислены выше. Тогда

$$\mathsf{E}\,\xi = \mathsf{E}\,(\sigma\eta + a) = \sigma\mathsf{E}\,\eta + a = a; \quad \mathsf{D}\,\xi = \mathsf{D}\,(\sigma\eta + a) = \sigma^2\mathsf{D}\,\eta = \sigma^2.$$

Итак, параметры a и σ^2 нормального распределения суть его математическое ожидание и дисперсия.

П р и м е р 61 (показательное распределение E_α). Найдём для произвольного $k \in \mathbb{N}$ 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:

$$\Gamma(k+1) = \int\limits_0^{\infty} u^k\,e^{-u}\,du = k!$$

Из формулы для момента порядка k находим

$$\mathsf{E}\,\xi = \frac{1}{\alpha}, \quad \mathsf{E}\,\xi^2 = \frac{2}{\alpha^2}, \quad \mathsf{D}\,\xi = \mathsf{E}\,\xi^2 - (\mathsf{E}\,\xi)^2 = \frac{1}{\alpha^2}.$$

П р и м е р 62 (стандартное распределение Коши $\mathrm{C}_{0,1}$). Математическое ожидание распределения Коши не существует, так как 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 потому, что подынтегральная функция ведёт себя на бесконечности как $1/x$. Поэтому не существуют ни дисперсия, ни моменты более высоких порядков этого распределения. То же самое можно сказать про распределение Коши $\mathrm{C}_{a,\sigma}$.

У п р а ж н е н и е. Доказать, что ни один момент порядка $k \leqslant -1$ распределения Коши также не существует. Найти все такие $\alpha \in \mathbb{R}$, при которых существует момент порядка α.

П р и м е р 63 (распределение Парето). У распределения Парето существуют только моменты порядка $t < \alpha$, поскольку

$$\mathsf{E}\,|\xi|^t = \int\limits_1^{\infty} x^t\,\alpha\,\frac{1}{x^{\alpha+1}}\,dx = \int\limits_1^{\infty} \alpha\,\frac{1}{x^{\alpha-t+1}}\,dx$$

сходится при $t < \alpha$, когда подынтегральная функция на бесконечности ведёт себя как $1\,/\,x^{s+1}$, где $s = \alpha - t > 0$.

У п р а ж н е н и е. Вычислить момент порядка $t < \alpha$ распределения Парето. При каких α у этого распределения существует дисперсия? А 2 317-й момент?

§ 6. Другие числовые характеристики распределений

Распределения можно характеризовать и многими другими показателями, большинство из которых находит основное применение в статистике. Здесь мы только кратко познакомимся 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. Так, 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 для всех $\mu \in [1,\ 2]$

$$\mathsf{P}(\xi \leqslant \mu) \geqslant \frac{1}{2}, \quad \mathsf{P}(\xi \geqslant \mu) \geqslant \frac{1}{2}.$$

Часто в таких случаях в качестве μ 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» (рис. 14). Если распределение имеет плотность f, то площади каждой из областей под графиком плотности слева и справа от точки μ одинаковы.

Медиана является одной из *квантилей* распределения. Пусть для простоты функция распределения F непрерывна и строго монотонна. Тогда *квантилью* уровня $\delta \in (0,\ 1)$ называется решение уравнения $F(x_\delta) = \delta$.

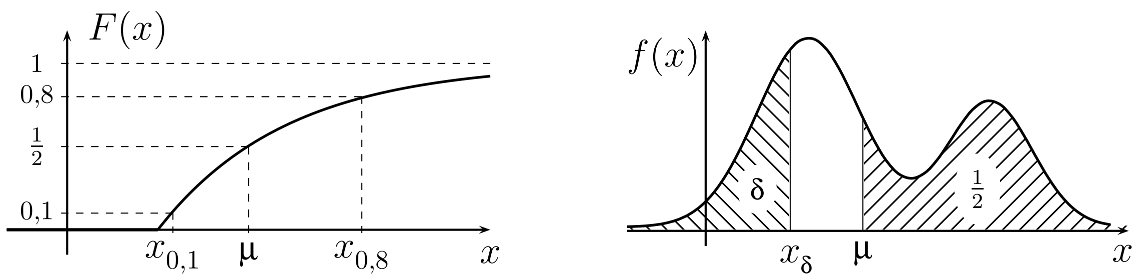

Рис. 14. Медиана и квантили на графике функции распределения и плотности

Квантиль x_δ уровня δ отрезает от области под графиком плотности область с площадью δ слева от себя, и с площадью $1-\delta$ — справа. Медиана является квантилью уровня $\delta = \frac{1}{2}$.

Квантили уровней, кратных 0,01, в прикладной статистике называют *процентилями,* квантили уровней, кратных 0,1, — *децилями,* уровней, кратных 0,25, — *квартилями.*

Модой абсолютно непрерывного распределения называют любую точку локального максимума плотности распределения. Для дискретных распределений модой считают любое значение a_i, вероятность которого больше, чем вероятности соседних значений (соседнего, если таковое одно).

Для нормального распределения N_{a,σ^2} медиана, математическое ожидание и мода равны a. Распределение, обладающее единственной модой, называют *унимодальным.* Идеальным примером унимодального распределения является нормальное распределение. Плотность произвольного унимодального распределения может быть как более плоской (равномерное распределение), так и более «островершинной» (показательное распределение) по сравнению с плотностью нормального распределения, может быть симметричной либо наклонённой в одну сторону. Для описания таких свойств плотности используют *коэффициент эксцесса* и *коэффициент асимметрии*.

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. Если $\beta_1 > 0$, то график плотности распределения имеет более крутой наклон 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. Четвёртый момент этого распределения равен трём: $\mathsf{E}\,\eta^4 = 3$ (*вычислить аналогично второму моменту в примере* 59). Поэтому $\beta_2 = 0$.

При $\beta_2 > 0$ плотность распределения имеет более острую вершину, чем у нормального распределения, при $\beta_2 < 0$, наоборот, более плоскую.

ГЛАВА IX
ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ЗАВИСИМОСТИ

Кажется, нельзя сомневаться ни в истине того, что всё в мире может быть представлено числами; ни в справедливости того, что всякая в нём перемена и отношение выражается аналитической функцией. Между тем обширный взгляд теории допускает существование зависимости только в том смысле, чтобы числа, одни с другими в связи, принимать как бы данными вместе.

Н. И. Лобачевский. Об исчезании тригонометрических строк

§ 1. 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

$$\mathsf{D}\,(\xi+\eta)=\mathsf{D}\,\xi+\mathsf{D}\,\eta+2\big(\mathsf{E}\,(\xi\eta)-\mathsf{E}\,\xi\,\mathsf{E}\,\eta\,\big). \tag{19}$$

Величина $\mathsf{E}\,(\xi\eta)-\mathsf{E}\,\xi\,\mathsf{E}\,\eta$ равняется нулю, если случайные величины ξ и η независимы (свойство (E7) математического ожидания). С другой стороны, из равенства её нулю вовсе не следует независимость, как показывают примеры 49 и 50 (с. 92). Эту величину используют как «индикатор наличия зависимости» между двумя случайными величинами.

Определение 39. *Ковариацией* $\operatorname{cov}(\xi,\,\eta)$ случайных величин ξ и η называется число $\operatorname{cov}(\xi,\,\eta)=\mathsf{E}\,\big((\xi-\mathsf{E}\,\xi)(\eta-\mathsf{E}\,\eta)\big)$.

Свойство 18. *Справедливы равенства:* $\operatorname{cov}(\xi,\,\eta)=\mathsf{E}\,(\xi\eta)-\mathsf{E}\,\xi\,\mathsf{E}\,\eta$; $\operatorname{cov}(\xi,\,\xi)=\mathsf{D}\,\xi$; $\operatorname{cov}(\xi,\,\eta)=\operatorname{cov}(\eta,\,\xi)$; $\operatorname{cov}(c\cdot\xi,\,\eta)=c\cdot\operatorname{cov}(\xi,\,\eta)$.

Упражнение. Доказать свойство 18.

Упражнение. Доказать следующее свойство 19, пользуясь равенствами

$$(a+b)^2=a^2+b^2+ab+ba=a^2+b^2+2ab=aa+bb+ab+ba$$

и получив аналогичные равенства для квадрата суммы n слагаемых.

Свойство 19. *Дисперсия суммы нескольких случайных величин вычисляется по любой из следующих формул:*

$$\begin{aligned}\mathsf{D}\,(\xi_1 + \ldots + \xi_n) &= \sum_{i=1}^{n} \mathsf{D}\,\xi_i + \sum_{i \neq j} \operatorname{cov}(\xi_i, \xi_j) = \\ &= \sum_{i=1}^{n} \mathsf{D}\,\xi_i + 2\sum_{i<j} \operatorname{cov}(\xi_i, \xi_j) = \sum_{i,j} \operatorname{cov}(\xi_i, \xi_j).\end{aligned}$$

Обсудим достоинства и недостатки ковариации, как величины, характеризующей зависимость двух случайных величин.

Если ковариация $\operatorname{cov}(\xi, \eta)$ отлична от нуля, то величины ξ и η зависимы. Чтобы судить о наличии зависимости согласно любому из определений независимости, требуется знать совместное распределение пары ξ и η. Но найти совместное распределение часто бывает сложнее, чем посчитать математическое ожидание произведения ξ и η. Если нам повезёт, и математическое ожидание $\xi\eta$ не будет равняться произведению их математических ожиданий, мы установим зависимость ξ и η *не находя* их совместного распределения. Это очень хорошо.

Пример 64. 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:

$$\mathsf{E}\,\bigl(\xi(\xi+\eta)\bigr) = \mathsf{E}\,\xi^2 + \mathsf{E}\,\xi\,\mathsf{E}\,\eta, \qquad \mathsf{E}\,\xi\,\mathsf{E}\,(\xi+\eta) = (\mathsf{E}\,\xi)^2 + \mathsf{E}\,\xi\,\mathsf{E}\,\eta.$$

Вычитая одно из другого, получим $\operatorname{cov}(\xi, \xi+\eta) = \mathsf{D}\,\xi > 0$. Следовательно, ξ и $\xi + \eta$ зависимы.

Упражнение. Доказать, что ξ и $\xi + \eta$ независимы, если $\mathsf{D}\,\xi = 0$.

Величина $\operatorname{cov}(\xi, \eta)$ не является «безразмерной»: если ξ — объем газа в сосуде, а η — давление этого газа, то ковариация измеряется в м$^3 \times$ Па. Иначе говоря, при умножении ξ или η на 100 ковариация тоже увеличится в 100 раз. Но от умножения на 100 величины не стали «более зависимыми», так что большое значение ковариации не означает более сильной зависимости. Это очень плохо.

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.

З а м е ч а н и е. Говоря о «силе» зависимости между случайными величинами, мы имеем в виду следующее. Самая сильная зависимость — функциональная, а из функциональных — линейная зависимость, когда $\xi = a\eta + b$ п. н. Бывают гораздо более слабые зависимости. Так, если по последовательности независимых случайных величин $\xi_1, \xi_2, \ldots$ построить величины $\xi = \xi_1 + \ldots + \xi_{24} + \xi_{25}$ и $\eta = \xi_{25} + \xi_{26} + \ldots + \xi_{90}$, то эти величины зависимы, но очень «слабо»: через единственное общее слагаемое ξ_{25}. Сильно ли зависимы число гербов в первых 25 подбрасываниях монеты и число гербов в испытаниях с 25-го по 90-е?

Итак, следующая величина есть всего лишь ковариация, нормированная нужным образом.

§ 2. Коэффициент корреляции

О п р е д е л е н и е 40. *Коэффициентом корреляции* $\rho(\xi, \eta)$ случайных величин ξ и η, дисперсии которых существуют и отличны от нуля, называется число

$$\rho(\xi, \eta) = \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{\mathsf{D}\,\xi}\sqrt{\mathsf{D}\,\eta}}.$$

З а м е ч а н и е. Чтобы разглядеть «устройство» коэффициента корреляции, распишем по определению числитель и знаменатель:

$$\rho(\xi, \eta) = \frac{\mathsf{E}\left((\xi - \mathsf{E}\,\xi)(\eta - \mathsf{E}\,\eta)\right)}{\sqrt{\mathsf{E}\,(\xi - \mathsf{E}\,\xi)^2}\sqrt{\mathsf{E}\,(\eta - \mathsf{E}\,\eta)^2}}.$$

Перед нами — «косинус угла» между двумя элементами $\xi - \mathsf{E}\,\xi$ и $\eta - \mathsf{E}\,\eta$ гильбертова пространства, образованного случайными величинами с нулевым математическим ожиданием и конечным вторым моментом, снабженного скалярным произведением $\operatorname{cov}(\xi, \eta)$ и «нормой», равной корню из дисперсии, или корню из скалярного произведения $\operatorname{cov}(\xi, \xi)$.

П р и м е р 65. Рассмотрим продолжение примера 64, но пусть ξ и η будут не только независимыми, но и одинаково распределёнными случайными величинами, и их дисперсия отлична от нуля. 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 величин ξ и $\xi + \eta$ равен косинусу угла 45°, образованного «векторами» ξ и $\xi + \eta$, когда ξ и η «ортогональны» и их «длина» одинакова.

У п р а ж н е н и е. Чтобы аналогия не заходила слишком далеко, и у читателя не возникло искушения любые случайные величины рисовать стрелочками на плоскости и вместо подсчёта математических ожиданий измерять углы, полезно убедиться, например, что коэффициент корреляции величин ξ и ξ^2 равен:

а) нулю, если ξ имеет нормальное распределение с нулевым средним;

б) $2/\sqrt{5}$, если ξ имеет показательное распределение.

Т е о р е м а 33. *Коэффициент корреляции обладает свойствами:*

1) *если* ξ *и* η *независимы, то* $\rho(\xi,\eta)=0$;

2) *всегда* $|\rho(\xi,\eta)|\leqslant 1$;

3) $|\rho(\xi,\eta)|=1$ *тогда и только тогда, когда* ξ *и* η *п. н. линейно связаны, т. е. существуют числа* $a\neq 0$ *и* b *такие, что* $\mathsf{P}(\eta=a\xi+b)=1$.

Д о к а з а т е л ь с т в о. Свойство (1) мы уже много раз (*сколько*?) упоминали и один раз доказали. Более того, при рассмотрении свойств математического ожидания мы привели примеры 49 и 50 — два из многих возможных примеров того, что свойство (1) в обратную сторону неверно.

Докажем свойство (2). 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:

$$\mathsf{E}\,\widehat{\xi}=\mathsf{E}\,\frac{\xi-\mathsf{E}\,\xi}{\sqrt{\mathsf{D}\,\xi}}=\frac{\mathsf{E}\,\xi-\mathsf{E}\,\xi}{\sqrt{\mathsf{D}\,\xi}}=0;\qquad \mathsf{E}\,\widehat{\xi}^{\,2}=\mathsf{D}\,\widehat{\xi}=\mathsf{D}\,\frac{\xi-\mathsf{E}\,\xi}{\sqrt{\mathsf{D}\,\xi}}=\frac{\mathsf{D}\,(\xi-\mathsf{E}\,\xi)}{\mathsf{D}\,\xi}=1.$$

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 (2):

$$-1=-\frac{1}{2}\,\mathsf{E}\left(\widehat{\xi}^{\,2}+\widehat{\eta}^{\,2}\right)\;\leqslant\;\rho(\xi,\,\eta)=\mathsf{E}\left(\widehat{\xi}\cdot\widehat{\eta}\right)\;\leqslant\;\frac{1}{2}\,\mathsf{E}\left(\widehat{\xi}^{\,2}+\widehat{\eta}^{\,2}\right)=1.\qquad(20)$$

Докажем свойство (3). 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 (3): 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 $\rho(\xi, \eta) = \rho\big(\widehat{\xi}, \widehat{\eta}\big) = 1$. Тогда второе неравенство в формуле (20) превращается в равенство:

$$\mathsf{E}\left(\widehat{\xi} \cdot \widehat{\eta}\right) = \frac{1}{2}\mathsf{E}\left(\widehat{\xi}^2 + \widehat{\eta}^2\right), \quad \text{т. е.} \quad \mathsf{E}\left(\widehat{\xi} - \widehat{\eta}\right)^2 = 0.$$

Если математическое ожидание неотрицательной случайной величины $\big(\widehat{\xi} - \widehat{\eta}\big)^2$ равно нулю, то $\big(\widehat{\xi} - \widehat{\eta}\big)^2 = 0$, п. н. Поэтому с единичной вероятностью

$$\frac{\eta - \mathsf{E}\eta}{\sqrt{\mathsf{D}\eta}} = \frac{\xi - \mathsf{E}\xi}{\sqrt{\mathsf{D}\xi}}, \qquad \eta = \frac{\sqrt{\mathsf{D}\eta}}{\sqrt{\mathsf{D}\xi}}\,\xi + \mathsf{E}\eta - \frac{\sqrt{\mathsf{D}\eta}}{\sqrt{\mathsf{D}\xi}}\,\mathsf{E}\xi = a\xi + b.$$

В случае $\rho(\xi, \eta) = -1$ нужно рассмотреть первое неравенство в формуле (20) и повторить рассуждения. Тем самым теорема 33 доказана. $\square$

Полезно знать следующие часто употребляемые термины.

Определение 41. 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 $\eta = a\xi + b$ п. н. Так, $\rho(\xi, \eta) = 1$ означает, что чем больше ξ, тем больше и η. Напротив, $\rho(\xi, \eta) = -1$ означает, что чем больше ξ, тем меньше η. Похожим образом можно трактовать знак коэффициента корреляции и в случае, когда $|\rho(\xi, \eta)| < 1$, помня при этом, что зависимость между ξ и η теперь уже не линейная и, возможно, даже не функциональная.

Так, величины ξ и $\xi + \eta$ в примерах 64 и 65 положительно коррелированы, но их зависимость не функциональная.

Следующее свойство показывает, что модуль коэффициента корреляции не меняется при линейных преобразованиях случайных величин.

Свойство 20. *Для любых случайных величин ξ и η с конечной и ненулевой дисперсией при любых постоянных $a \neq 0$ и b имеет место равенство $\rho(a\xi + b, \eta) = \operatorname{sgn}(a) \cdot \rho(\xi, \eta)$, где $\operatorname{sgn}(a) = \dfrac{a}{|a|}$ — знак a.*

Доказательство. Запишем $\rho(a\xi + b, \eta)$, не забывая про свойства дисперсии:

$$\rho(a\xi + b, \eta) = \frac{\operatorname{cov}(a\xi + b, \eta)}{\sqrt{\mathsf{D}\,(a\xi + b)}\,\sqrt{\mathsf{D}\eta}} = \frac{a\operatorname{cov}(\xi, \eta)}{\sqrt{a^2\mathsf{D}\xi}\,\sqrt{\mathsf{D}\eta}} = \frac{a}{|a|} \cdot \rho(\xi, \eta). \qquad \square$$

П р и м е р 66. Если ξ и η суть координаты точки, брошенной наудачу в треугольник D с вершинами $(2,0)$, $(0,0)$ и $(0,1)$, то их коэффициент корреляции $\rho(\xi,\eta)$ отрицателен. Это можно объяснить так: чем больше ξ, тем меньше у η возможностей быть большой.

Полезно убедиться в этом, проверив справедливость следующих высказываний. Во-первых,

$$f_\xi(x)=\begin{cases}1-\dfrac{x}{2}, & 0\leqslant x\leqslant 2,\\ 0, & \text{иначе};\end{cases}\qquad f_\eta(y)=\begin{cases}2-2y, & 0\leqslant y\leqslant 1,\\ 0, & \text{иначе}\end{cases}$$

и вычисленные по этим плотностям средние (*вычислить*) равны соответственно $\mathsf{E}\,\xi=2/3$ и $\mathsf{E}\,\eta=1/3$.

Во-вторых, по определению многомерного равномерного распределения в области D,

$$\mathsf{E}\,(\xi\eta)=\iint\limits_D x\cdot y\cdot 1\;dx\,dy=\int\limits_0^2\int\limits_0^{1-x/2} x\,y\;dy\;dx=\frac{1}{6}.$$

Ковариация (а с ней и коэффициент корреляции) отрицательна.

У п р а ж н е н и е. Почему коэффициент корреляции в примере 66 существует? Какие свойства случайных величин гарантируют конечность второго момента? А из их ограниченности следует существование моментов? По какому из свойств математического ожидания это так?

П р и м е р 67. Найдём коэффициент корреляции между числом выпадений единицы и числом выпадений шестерки при n подбрасываниях правильной игральной кости.

Обозначим для $i\in\{1,\ldots,6\}$ через ξ_i случайную величину, равную числу выпадений грани с i очками при n подбрасываниях кубика. Посчитаем $\mathrm{cov}(\xi_1,\xi_6)$. Каждая из случайных величин ξ_i имеет биномиальное распределение с параметрами n и $\dfrac{1}{6}$, поэтому $\mathsf{E}\,\xi_i=\dfrac{n}{6}$, $\mathsf{D}\,\xi_i=\dfrac{5n}{36}$.

Далее заметим, что $\xi_1+\ldots+\xi_6=n$. Из-за симметрии кубика математические ожидания $\mathsf{E}\,\xi_1\xi_2$, $\mathsf{E}\,\xi_1\xi_3$, ..., $\mathsf{E}\,\xi_1\xi_6$ одинаковы, но отличаются от $\mathsf{E}\,\xi_1\xi_1=\mathsf{E}\,\xi_1^2=\mathsf{D}\,\xi_1+(\mathsf{E}\,\xi_1)^2=\dfrac{5n}{36}+\dfrac{n^2}{36}$. Посчитаем $\mathsf{E}\,\xi_1(\xi_1+\cdots+\xi_6)$. С одной стороны, это число равно

$$\mathsf{E}\,\xi_1(\xi_1+\ldots+\xi_6)=\mathsf{E}\,\xi_1\cdot n=\frac{n^2}{6}.$$

С другой стороны,

$$\mathsf{E}\,\xi_1(\xi_1+\ldots+\xi_6)=\mathsf{E}\,\xi_1^2+5\mathsf{E}\,\xi_1\xi_6=\frac{5n}{36}+\frac{n^2}{36}+5\mathsf{E}\,\xi_1\xi_6.$$

Отсюда $5\mathsf{E}\,\xi_1\xi_6=\dfrac{n^2}{6}-\dfrac{5n}{36}-\dfrac{n^2}{36}$, т. е. $\mathsf{E}\,\xi_1\xi_6=\dfrac{n^2-n}{36}$. Следовательно, искомый коэффициент корреляции равен

$$\rho(\xi_1,\,\xi_6)=\frac{\mathsf{E}\,\xi_1\xi_6-\mathsf{E}\,\xi_1\mathsf{E}\,\xi_6}{\sqrt{\mathsf{D}\,\xi_1\mathsf{D}\,\xi_6}}=\frac{(n^2-n)/36-n^2/36}{5n/36}=-\frac{1}{5}.$$

Интересно, что полученный коэффициент корреляции не зависит от n.

Упражнение. Объяснить знак величины $\rho(\xi_1,\,\xi_6)$. Вычислить коэффициент корреляции числа единиц и числа двоек при n подбрасываниях правильной игральной кости.

Пример 68. Вычислим математическое ожидание и дисперсию гипергеометрического распределения. Мы не могли сделать это раньше, так как очень не хотели вычислять следующие суммы:

$$\mathsf{E}\,\xi=\sum_k k\,\frac{C_K^k C_{N-K}^{n-k}}{C_N^n},\qquad \mathsf{E}\,\xi^2=\sum_k k^2\,\frac{C_K^k C_{N-K}^{n-k}}{C_N^n},$$

где, напомним (чтобы читатель окончательно отказался от мысли вычислить эти суммы напрямую), суммирование ведётся по целым k таким, что $0\leqslant k\leqslant K$ и $0\leqslant n-k\leqslant N-K$.

Рассмотрим урну, содержащую K белых шаров и $N-K$ не белых, и пусть из неё наудачу и без возвращения выбирают по одному n шаров. Свяжем случайную величину ξ, равную числу белых шаров среди n выбранных, с результатами отдельных извлечений шаров.

Обозначим через ξ_i, где $i=1,\,\ldots,\,n$, «индикатор» того, что i-й по счёту вынутый шар оказался белым: $\xi_i=1$, если при i-м извлечении появился белый шар, иначе $\xi_i=0$. Тогда $\xi=\xi_1+\ldots+\xi_n$ — число появившихся белых шаров, и математическое ожидание считается просто:

$$\mathsf{E}\,\xi=\mathsf{E}\,(\xi_1+\ldots+\xi_n)=\mathsf{E}\,\xi_1+\ldots+\mathsf{E}\,\xi_n.$$

Убедимся, что случайные величины $\xi_1,\,\ldots,\,\xi_n$ имеют одно и то же распределение Бернулли B_p, где $p=K\,/\,N$.

Пронумеруем шары: белые — номерами от одного до K, остальные — номерами от $K+1$ до N. Элементарным исходом опыта является набор из n номеров шаров в схеме выбора n элементов из N без возвращения и с учётом порядка. Общее число исходов равно $|\Omega|=A_N^n$.

Вычислим вероятность события $A_i = \{\xi_i = 1\}$. Событие A_i включает в себя элементарные исходы (наборы), в которых на i-м месте стоит любой из номеров белых шаров, а остальные $n-1$ место занимают любые из оставшихся $N-1$ номеров. По теореме 1 о перемножении шансов число благоприятных событию A_i исходов есть произведение K и A_{N-1}^{n-1}. Здесь K есть число способов поставить на i-е место один из номеров белых шаров, A_{N-1}^{n-1} — число способов после этого разместить на оставшихся $n-1$ местах остальные $N-1$ номеров шаров. Но тогда

$$p = \mathsf{P}(\xi_i = 1) = \mathsf{P}(A_i) = \frac{|A_i|}{|\,\Omega\,|} = \frac{K\,A_{N-1}^{n-1}}{A_N^n} = \frac{K}{N}\,,$$

что совершенно очевидно: вероятность 20-му шару быть белым, если мы ничего не знаем про первые 19, точно такая же, как вероятность первому шару быть белым и равна отношению числа белых шаров к числу всех.

Вернёмся к математическому ожиданию:

$$\mathsf{E}\,\xi = \mathsf{E}\,\xi_1 + \ldots + \mathsf{E}\,\xi_n = n\mathsf{E}\,\xi_1 = np = \frac{nK}{N}\,.$$

Вычислим дисперсию ξ. До сих пор мы не интересовались совместным распределением $\xi_1, \ldots, \xi_n$: для вычисления математического ожидания их суммы нам было достаточно знания маргинальных распределений этих величин. Но дисперсия суммы уже не всегда равна сумме дисперсий. Зависимость величин $\xi_1, \ldots, \xi_n$ очевидна: если, скажем, случилось событие $A_1 = \{\xi_1 = 1\}$, то вероятность второму шару быть белым уже не равна отношению $K\,/\,N$:

$$\mathsf{P}(\xi_2 = 1 \mid \xi_1 = 1) = \frac{K-1}{N-1} \neq \frac{K}{N} = \mathsf{P}(\xi_2 = 1).$$

Поэтому при вычислении дисперсии будем пользоваться свойством 19. Вычислим ковариацию величин ξ_i и ξ_j, $i \neq j$. Для этого сначала посчитаем $\mathsf{E}\,(\xi_i\xi_j)$. Произведение $\xi_i\xi_j$ снова имеет распределение Бернулли: $\xi_i\xi_j = 1$, если при i-м и j-м извлечениях появились белые шары. Вероятность этого события равна

$$\mathsf{P}(\xi_i\xi_j = 1) = \mathsf{P}(A_i \cap A_j) = \frac{|A_i \cap A_j|}{|\,\Omega\,|} = \frac{K(K-1)A_{N-2}^{n-2}}{A_N^n} = \frac{K(K-1)}{N(N-1)}.$$

Тогда

$$\operatorname{cov}(\xi_i,\, \xi_j) = \mathsf{E}\,(\xi_i\xi_j) - \mathsf{E}\,\xi_i\mathsf{E}\,\xi_j = \frac{K(K-1)}{N(N-1)} - \frac{K}{N}\,\frac{K}{N} = -\frac{K(N-K)}{N^2(N-1)}.$$

Подставляя одинаковые дисперсии $\mathsf{D}\,\xi_i = p(1-p)$ и эти не зависящие от i и j ковариации в формулу дисперсии суммы, получаем

$$\begin{aligned}\mathsf{D}\,\xi \;&=\; \mathsf{D}\,(\xi_1+\ldots+\xi_n) = \sum_{i=1}^{n}\mathsf{D}\,\xi_i + \sum_{i\neq j}\operatorname{cov}(\xi_i,\xi_j) = \\ &=\; np(1-p) + n(n-1)\operatorname{cov}(\xi_1,\,\xi_2) = \\ &=\; n\,\frac{K}{N}\Bigl(1-\frac{K}{N}\Bigr) - n(n-1)\frac{K(N-K)}{N^2(N-1)} = n\,\frac{K}{N}\Bigl(1-\frac{K}{N}\Bigr)\Bigl(1-\frac{n-1}{N-1}\Bigr).\end{aligned}$$

Заметим любопытнейшую вещь: если вынимать шары *с возвращением*, то испытания станут независимыми испытаниями в схеме Бернулли; ставшие независимыми величины ξ_i в сумме дадут число белых шаров, имеющее биномиальное распределение с параметрами n и $p = \dfrac{K}{N}$ и точно такое же математическое ожидание $np = \dfrac{nK}{N}$, как и у числа белых шаров при выборе *без возвращения.*

Дисперсия же у числа белых шаров при выборе без возвращения меньше, чем при выборе с возвращением — за счёт отрицательной коррелированности слагаемых ξ_i и ξ_j при $i \neq j$.

ГЛАВА X

СХОДИМОСТЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ СЛУЧАЙНЫХ ВЕЛИЧИН

> Откуда, наконец, вытекает то удивительное, по-видимому, следствие, что, если бы наблюдения над всеми событиями продолжать всю вечность, причём вероятность, наконец, перешла бы в полную достоверность, то было бы замечено, что в мире всё управляется точными отношениями и постоянным законом изменений, так что даже в вещах, в высшей степени случайных, мы принуждены были бы признать как бы некоторую необходимость и, скажу я, рок.
>
> *Якоб Бернулли. Искусство предположений*

§ 1. Сходимости «почти наверное» и «по вероятности»

Напомним, что случайная величина есть (измеримая) функция из некоторого непустого множества Ω в множество действительных чисел. Последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ есть тем самым последовательность функций, определённых на одном и том же множестве Ω. Существуют разные виды сходимости последовательности *функций.* Давать определение любой сходимости мы будем, опираясь на сходимость *числовых* последовательностей, как на уже известное основное понятие.

В частности, при каждом новом $\omega \in \Omega$ мы имеем новую *числовую* последовательность $\xi_1(\omega), \xi_2(\omega), \xi_3(\omega), \ldots$ Поэтому можно говорить о сходимости последовательности значений функций в данной точке ω, а также во всех остальных точках $\omega \in \Omega$. В теории вероятностей можно не обращать внимание на неприятности, происходящие с нулевой вероятностью. Поэтому вместо сходимости «всюду» принято рассматривать сходимость «почти всюду», или «почти наверное».

Определение 42. Говорят, что последовательность $\{\xi_n\}$ сходится *почти наверное* к случайной величине ξ при $n \to \infty$, и пишут: $\xi_n \to \xi$ п. н., если $\mathsf{P}\{\omega \mid \xi_n(\omega) \to \xi(\omega) \text{ при } n \to \infty\} = 1$. Иначе говоря, если $\xi_n(\omega) \to \xi(\omega)$ при $n \to \infty$ для всех $\omega \in \Omega$, кроме, возможно, $\omega \in A$, где A — событие, имеющее нулевую вероятность.

Заметим сразу: определение сходимости «почти наверное» требует знания того, как устроены отображения $\omega \mapsto \xi_n(\omega)$. В задачах же теории вероятностей, как правило, известны не сами случайные величины, а лишь их *распределения.*

Можем ли мы, обладая только информацией о распределениях, говорить о какой-либо сходимости последовательности случайных величин $\{\xi_n\}$ к случайной величине ξ?

Можно, скажем, потребовать, чтобы вероятность тех элементарных исходов ω, для которых $\xi_n(\omega)$ не попадает в «ε-окрестность» числа $\xi(\omega)$, уменьшалась до нуля с ростом n. Такая сходимость в функциональном анализе называется сходимостью «по мере», а в теории вероятностей — сходимостью «по вероятности».

Определение 43. Говорят, что последовательность случайных величин $\{\xi_n\}$ сходится *по вероятности* к случайной величине ξ при $n \to \infty$, и пишут $\xi_n \xrightarrow{\text{p}} \xi$, если для любого $\varepsilon > 0$

$$\mathsf{P}\left(|\xi_n - \xi| \geqslant \varepsilon\right) \to 0 \text{ при } n \to \infty \quad (\text{или } \mathsf{P}\left(|\xi_n - \xi| < \varepsilon\right) \to 1 \text{ при } n \to \infty).$$

Пример 69. Рассмотрим последовательность $\xi_1, \xi_2, \ldots,$ в которой все величины имеют *разные* распределения: величина ξ_n принимает значения 0 и n^7 с вероятностями $\mathsf{P}\big(\xi_n = n^7\big) = 1/n = 1 - \mathsf{P}(\xi_n = 0)$. Докажем, что эта последовательность сходится по вероятности к нулю.

Зафиксируем произвольное $\varepsilon > 0$. Для всех n начиная с некоторого n_0 такого, что $n_0^7 > \varepsilon$, верно равенство $\mathsf{P}(\xi_n \geqslant \varepsilon) = \mathsf{P}(\xi_n = n^7) = 1/n$. Поэтому

$$\mathsf{P}\big(|\xi_n - 0| \geqslant \varepsilon\big) = \mathsf{P}\big(\xi_n \geqslant \varepsilon\big) = \mathsf{P}\big(\xi_n = n^7\big) = \frac{1}{n} \ \to \ 0 \ \text{ при } \ n \to \infty.$$

Итак, случайные величины ξ_n с ростом n могут принимать всё бо́льшие и бо́льшие значения, но со всё меньшей и меньшей вероятностью.

Например, последовательность $\{\xi_n\}$ можно задать на вероятностном пространстве $\langle \Omega, \mathcal{F}, \mathsf{P} \rangle = \langle [0, 1], \mathfrak{B}([0, 1]), \lambda \rangle$ так: положим $\xi_n(\omega) = 0$ для $\omega \in [0, 1 - 1/n]$ и $\xi_n(\omega) = n^7$ для $\omega \in (1 - 1/n, 1]$.

Заметим, что сходимость по вероятности имеет место совершенно независимо от того, как именно заданы случайные величины на Ω, поскольку определяется лишь их распределениями.

Замечание. Иное дело — сходимость «почти наверное». Если, скажем, задать случайные величины как указано выше, то сходимость «почти наверное» будет иметь место. Действительно, для всякого $\omega \in [0, 1)$ найдётся такое n_0, что $\omega \in [0, 1 - 1/n_0]$, и поэтому для всех $n \geqslant n_0$ все $\xi_n(\omega)$ равны нулю.

Можно попробовать задать случайные величины ξ_n на отрезке $[0, 1]$ как-нибудь иначе, чтобы не было сходимости почти наверное. Для этого нужно заставить отрезок длины $1/n$, на котором $\xi_n(\omega) = n^7$, «бегать» по отрезку $[0, 1]$, чтобы любая точка $\omega \in [0, 1]$ попадала внутрь этого отрезка бесконечное число раз, и, тем самым, для любого ω существовала подпоследовательность $\xi_{n_k}(\omega) \to \infty$.

Сходимость по вероятности не обязательно сопровождается сходимостью математических ожиданий или моментов других порядков: из $\xi_n \xrightarrow{\text{p}} \xi$ *не следует*, что $\mathsf{E}\,\xi_n \to \mathsf{E}\,\xi$. Действительно, в примере 69 имеет место сходимость $\xi_n \xrightarrow{\text{p}} \xi = 0$, но $\mathsf{E}\,\xi_n = n^6 \not\to \mathsf{E}\,\xi = 0$. При этом вообще последовательность $\mathsf{E}\,\xi_n$ неограниченно возрастает.

А если вместо значения n^7 взять n (с той же вероятностью $1/n$), то получим $\mathsf{E}\,\xi_n = 1 \not\to \mathsf{E}\,\xi = 0$. Но теперь хотя бы предел у последовательности математических ожиданий конечен.

Если же ξ_n принимает значения 0 и $\sqrt{n}$ с вероятностями из примера 69, то $\mathsf{E}\,\xi_n = 1/\sqrt{n} \to \mathsf{E}\,\xi = 0$, но уже вторые моменты сходиться ко второму моменту ξ не будут: $\mathsf{E}\,\xi_n^2 = 1 \not\to \mathsf{E}\,\xi^2 = 0$.

Однако сходимость математических ожиданий и других моментов сходящихся последовательностей бывает чрезвычайно важна в различных задачах статистики. Существуют условия, при выполнении которых сходимость по вероятности $\xi_n \xrightarrow{\text{p}} \xi$ влечёт сходимость математических ожиданий $\mathsf{E}\,\xi_n \to \mathsf{E}\,\xi$.

Сформулируем без доказательства следующее утверждение.

Теорема 34. *Пусть* $\xi_n \xrightarrow{\text{p}} \xi$ *при* $n \to \infty$. *Тогда для сходимости* $\mathsf{E}\,\xi_n \to \mathsf{E}\,\xi$ *достаточно выполнения любого из следующих условий*:

1. Все члены последовательности ограничены одной и той же постоянной: $|\xi_n| \leqslant C = const$.

2. Все члены последовательности ограничены одной и той же случайной величиной с конечным первым моментом: $|\xi_n| \leqslant \eta$, $\mathsf{E}\,\eta < \infty$.

3. Существует $\alpha > 1$ такое, что $\mathsf{E}\,|\xi_n|^\alpha \leqslant C = const$ для любого n.

Самым слабым в этом списке является третье условие, наиболее ограничительным — первое. Ни одно из этих условий не является необходимым для сходимости математических ожиданий (*найти контрпример*).

Сходимость по вероятности, так же как и любая другая сходимость, не портится под действием непрерывной функции.

Свойство 21. *Пусть функция g действует из $\mathbb{R}$ в $\mathbb{R}$.*

1. *Если $\xi_n \xrightarrow{\mathrm{p}} \xi$ и функция $g(x)$ непрерывна, то $g(\xi_n) \xrightarrow{\mathrm{p}} g(\xi)$.*
2. *Если $\xi_n \xrightarrow{\mathrm{p}} c$ и $g(x)$ непрерывна в точке c, то $g(\xi_n) \xrightarrow{\mathrm{p}} g(c)$.*

Доказательство. Простое доказательство первого утверждения можно предложить в двух случаях, которыми мы и ограничимся: если $\xi = c = \mathrm{const}$ (и тогда достаточно, чтобы g была непрерывна в точке c) или если функция g равномерно непрерывна (*а что это значит*?).

И в том и в другом случае для любого $\varepsilon > 0$ найдётся такое $\delta > 0$, что для любого ω, удовлетворяющего условию $|\xi_n(\omega) - \xi(\omega)| < \delta$, выполняется неравенство $|g(\xi_n(\omega)) - g(\xi(\omega))| < \varepsilon$.

Другими словами, событие $\bigl\{|\xi_n - \xi| < \delta\bigr\}$ влечёт за собой событие $\bigl\{|g(\xi_n) - g(\xi)| < \varepsilon\bigr\}$. Следовательно, вероятность первого не больше вероятности второго. Но, какое бы ни было $\delta > 0$, вероятность первого события стремится к единице по определению сходимости по вероятности:

$$1 \longleftarrow \mathsf{P}\bigl(|\xi_n - \xi| < \delta\bigr) \leqslant \mathsf{P}\bigl(|g(\xi_n) - g(\xi)| < \varepsilon\bigr) \leqslant 1.$$

Тогда вероятность второго события также стремится к единице. □

То же самое можно утверждать и для непрерывной функции многих переменных, применённой к нескольким сходящимся последовательностям.

Свойство 22. *Пусть функция g отображает $\mathbb{R}^2$ в $\mathbb{R}$.*

1. *Если $\xi_n \xrightarrow{\mathrm{p}} \xi$, $\eta_n \xrightarrow{\mathrm{p}} \eta$ при $n \to \infty$, функция $g(x, y)$ всюду непрерывна, то $g(\xi_n, \eta_n) \xrightarrow{\mathrm{p}} g(\xi, \eta)$.*
2. *Если $\xi_n \xrightarrow{\mathrm{p}} c_1$, $\eta_n \xrightarrow{\mathrm{p}} c_2$ при $n \to \infty$, функция $g(x)$ непрерывна в точке (c_1, c_2), то $g(\xi_n, \eta_n) \xrightarrow{\mathrm{p}} g(c_1, c_2)$.*

Доказательство. Докажем опять только второе свойство. Воспользуемся определением непрерывности функции двух переменных: для любого $\varepsilon > 0$ найдётся такое $\delta > 0$, что для любого ω, принадлежащего одновременно двум событиям

$$A_n = \bigl\{|\xi_n(\omega) - c_1| < \delta\bigr\}, \qquad B_n = \bigl\{|\eta_n(\omega) - c_2| < \delta\bigr\},$$

выполняется неравенство

$$|g(\xi_n(\omega), \eta_n(\omega)) - g(c_1, c_2)| < \varepsilon.$$

Тогда событие $A_n \cap B_n$ влечёт событие $C = \bigl\{|g(\xi_n, \eta_n) - g(c_1, c_2)| < \varepsilon\bigr\}$, поэтому вероятность первого не больше вероятности второго. Но вероятность пересечения двух событий, вероятности которых стремятся к единице, также стремится к единице:

$$\mathsf{P}(A_n \cap B_n) = 1 - \mathsf{P}\big(\overline{A_n} \cup \overline{B_n}\big) \geqslant 1 - \mathsf{P}\big(\overline{A_n}\big) - \mathsf{P}\big(\overline{B_n}\big) \to 1.$$

Поэтому $\mathsf{P}(C) \geqslant \mathsf{P}(A_n \cap B_n) \to 1$ при $n \to \infty$. □

Из свойства 22 вытекают обычные свойства пределов, хорошо знакомые нам по числовым последовательностям. Например, функции $g(x, y) = x + y$ и $g(x, y) = xy$ непрерывны в $\mathbb{R}^2$, поэтому предел суммы (произведения) сходящихся по вероятности последовательностей равен сумме (произведению) пределов.

Свойство 23. *Если* $\xi_n \xrightarrow{\text{p}} \xi$ *и* $\eta_n \xrightarrow{\text{p}} \eta$, *то* $\xi_n + \eta_n \xrightarrow{\text{p}} \xi + \eta$ *и* $\xi_n \cdot \eta_n \xrightarrow{\text{p}} \xi \cdot \eta$.

Сходимость «почти наверное» сильнее сходимости по вероятности.

Свойство 24. *Если* $\xi_n \to \xi$ *п. н., то* $\xi_n \xrightarrow{\text{p}} \xi$.

Доказательство. Ограничимся для простоты случаем, когда $\xi_n(\omega) \to \xi(\omega)$ *для любого* ω. Зафиксируем $\omega \in \Omega$. По определению предела, $\xi_n(\omega) \to \xi(\omega)$ при $n \to \infty$, если для всякого $\varepsilon > 0$ найдётся $N = N(\omega, \varepsilon) \geqslant 0$ такое, что для всех $n > N$ выполняется неравенство $|\xi_n(\omega) - \xi(\omega)| < \varepsilon$.

Событие $A = \{n > N(\omega, \varepsilon)\}$ влечёт событие $B = \{|\xi_n(\omega) - \xi(\omega)| < \varepsilon\}$. Тогда

$$1 \geqslant \mathsf{P}(B) \geqslant \mathsf{P}(A) = \mathsf{P}\big(N(\omega, \varepsilon) < n\big) = F_{N(\varepsilon,\omega)}(n) \to 1 \text{ при } n \to \infty$$

по свойству (F2) функций распределения. Мы получили, что $\mathsf{P}(B) \to 1$, т. е. $\xi_n \xrightarrow{\text{p}} \xi$. □

§ 2. Неравенства Чебышёва

Чтобы доказывать сходимость по вероятности, требуется уметь вычислять $\mathsf{P}\left(|\xi_n - \xi| \geqslant \varepsilon\right)$ при больших n. Но для этого нужно знать распределение ξ_n, что не всегда возможно.

Полезно иметь неравенства, позволяющие оценивать вероятность $\mathsf{P}\left(|\xi_n - \xi| \geqslant \varepsilon\right)$ сверху. Тогда для доказательства сходимости по вероятности было бы достаточно устремить к нулю эту оценку.

Все неравенства в этом параграфе принято относить к одному классу неравенств Чебышёва[16].

[16] Пафнутий Львович Чебышёв (16.05.1821—8.12.1894).

Теорема 35 (неравенство Маркова[17]). *Если* $\mathsf{E}\,|\xi| < \infty$, *то для любого* $x > 0$

$$\mathsf{P}\big(|\xi| \geqslant x\big) \leqslant \frac{\mathsf{E}\,|\xi|}{x}\,.$$

Доказательство. Нам потребуется следующее понятие.

Определение 44. 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 $p = \mathsf{P}(I(A) = 1) = \mathsf{P}(A)$ 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

$$|\xi| = |\xi| \cdot I(|\xi| < x) + |\xi| \cdot I(|\xi| \geqslant x) \geqslant |\xi| \cdot I(|\xi| \geqslant x) \geqslant x \cdot I(|\xi| \geqslant x).$$

Тогда $\mathsf{E}\,|\xi| \geqslant \mathsf{E}\,\big(x \cdot I(|\xi| \geqslant x)\big) = x \cdot \mathsf{P}\big(|\xi| \geqslant x\big)$. Осталось разделить обе части этого неравенства на положительное число x. □

Следствие 16 (обобщённое неравенство Чебышёва). *Пусть функция* g *не убывает и неотрицательна на* $\mathbb{R}$. *Если* $\mathsf{E}\,g(\xi) < \infty$, *то для любого* $x \in \mathbb{R}$

$$\mathsf{P}\big(\xi \geqslant x\big) \leqslant \frac{\mathsf{E}\,g(\xi)}{g(x)}.$$

Доказательство. Заметим, что $\mathsf{P}\big(\xi \geqslant x\big) \leqslant \mathsf{P}\big(g(\xi) \geqslant g(x)\big)$, поскольку функция g не убывает. Оценим последнюю вероятность по неравенству Маркова, которое можно применять в силу неотрицательности g:

$$\mathsf{P}\big(g(\xi) \geqslant g(x)\big) \leqslant \frac{\mathsf{E}\,g(\xi)}{g(x)}. \qquad \square$$

Упражнение. Записать предыдущее неравенство для функции $g(x) = e^x$ и получить *экспоненциальное неравенство Чебышёва.*

Следствие 17 (неравенство Чебышёва). *Если* $\mathsf{D}\,\xi$ *существует, то для любого* $x > 0$

$$\mathsf{P}\big(|\xi - \mathsf{E}\,\xi| \geqslant x\big) \leqslant \frac{\mathsf{D}\,\xi}{x^2}.$$

Доказательство. Для $x > 0$ неравенство $|\xi - \mathsf{E}\,\xi| \geqslant x$ равносильно неравенству $(\xi - \mathsf{E}\,\xi)^2 \geqslant x^2$, поэтому

$$\mathsf{P}\big(|\xi - \mathsf{E}\,\xi| \geqslant x\big) = \mathsf{P}\big((\xi - \mathsf{E}\,\xi)^2 \geqslant x^2\big) \leqslant \frac{\mathsf{E}\,(\xi - \mathsf{E}\,\xi)^2}{x^2} = \frac{\mathsf{D}\,\xi}{x^2}. \quad \square$$

[17] Андрей Андреевич Марков (14.06.1856—20.07.1922).

Неравенство Чебышёва позволяет, помимо всего прочего, получать абсолютные оценки для вероятности того, что стандартизованная случайная величина превзойдёт некоторое значение: для любого $x > 0$

$$\mathsf{P}\left(\left|\frac{\xi - \mathsf{E}\,\xi}{\sqrt{\mathsf{D}\,\xi}}\right| \geqslant x\right) = \mathsf{P}\left(|\xi - \mathsf{E}\,\xi| \geqslant x\sqrt{\mathsf{D}\,\xi}\right) \leqslant \frac{\mathsf{D}\,\xi}{x^2\mathsf{D}\,\xi} = \frac{1}{x^2}.$$

Например, при $x = 10$ эта вероятность не превышает 0,01.

§ 3. Законы больших чисел

Определение 45. Говорят, что последовательность случайных величин $\xi_1, \xi_2, \ldots$ с конечными первыми моментами *удовлетворяет закону больших чисел* (ЗБЧ), если

$$\frac{\xi_1 + \ldots + \xi_n}{n} - \frac{\mathsf{E}\,\xi_1 + \ldots + \mathsf{E}\,\xi_n}{n} \xrightarrow{\text{p}} 0 \quad \text{при} \quad n \to \infty. \tag{21}$$

Законами больших чисел принято называть утверждения о том, при каких условиях последовательность случайных величин удовлетворяет закону больших чисел.

Выясним сначала, когда выполнен ЗБЧ для последовательности независимых и одинаково распределённых случайных величин.

Теорема 36 (ЗБЧ Чебышёва). *Для любой последовательности $\xi_1, \xi_2, \ldots$ попарно независимых и одинаково распределённых случайных величин с конечным вторым моментом $\mathsf{E}\,\xi_1^2 < \infty$ имеет место сходимость*

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{\text{p}} \mathsf{E}\,\xi_1. \tag{22}$$

Заметим, что если величины одинаково распределены, то их математические ожидания одинаковы (и равны, например, $\mathsf{E}\,\xi_1$), поэтому свойство (21) можно записать в виде (22).

ЗБЧ утверждает, что среднее арифметическое большого числа случайных слагаемых «стабилизируется» с ростом этого числа. Как бы сильно каждая случайная величина ни отклонялась от своего среднего значения, при суммировании эти отклонения «взаимно гасятся», так что среднее арифметическое приближается к постоянной величине.

В дальнейшем мы увидим, что требование конечности второго момента (или дисперсии) связано исключительно со способом доказательства, и что 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

$$\mathsf{E}\left(\frac{S_n}{n}\right) = \frac{\mathsf{E}\,\xi_1 + \ldots + \mathsf{E}\,\xi_n}{n} = \frac{n\,\mathsf{E}\,\xi_1}{n} = \mathsf{E}\,\xi_1.$$

Пусть $\varepsilon > 0$. Воспользуемся неравенством Чебышёва (следствие 17):

$$\mathsf{P}\left(\left|\frac{S_n}{n} - \mathsf{E}\left(\frac{S_n}{n}\right)\right| \geqslant \varepsilon\right) \leqslant \frac{\mathsf{D}\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{\mathsf{D}\,S_n}{n^2\varepsilon^2} = \frac{\mathsf{D}\,\xi_1 + \ldots + \mathsf{D}\,\xi_n}{n^2\varepsilon^2} =$$
$$= \frac{n\,\mathsf{D}\,\xi_1}{n^2\varepsilon^2} = \frac{\mathsf{D}\,\xi_1}{n\varepsilon^2} \to 0 \quad \text{при} \quad n \to \infty, \tag{23}$$

так как $\mathsf{D}\,\xi_1 < \infty$. Дисперсия суммы превратилась в сумму дисперсий в силу попарной независимости слагаемых, из-за которой все ковариации $\mathrm{cov}(\xi_i, \xi_j)$ в свойстве 19 (с. 103) обратились в нуль при $i \neq j$. □

Замечание. Мы не только доказали сходимость, но и получили оценку для вероятности среднему арифметическому любого числа попарно независимых и одинаково распределённых величин отличаться от $\mathsf{E}\,\xi_1$ более чем на заданное ε:

$$\mathsf{P}\left(\left|\frac{\xi_1 + \ldots + \xi_n}{n} - \mathsf{E}\,\xi_1\right| \geqslant \varepsilon\right) \leqslant \frac{\mathsf{D}\,\xi_1}{n\varepsilon^2}. \tag{24}$$

Попарную независимость слагаемых в ЗБЧ Чебышёва можно заменить их попарной некоррелированностью, ничего не меняя в доказательстве. ЗБЧ может выполняться и для последовательности зависимых и разнораспределённых слагаемых. Из неравенства Чебышёва сразу вытекает следующее достаточное условие выполнения ЗБЧ для последовательности произвольных случайных величин.

Теорема 37 (ЗБЧ Маркова). *Последовательность случайных величин $\xi_1, \xi_2, \ldots$ с конечными вторыми моментами удовлетворяет ЗБЧ, если $\mathsf{D}\,S_n = o(n^2)$, т. е. если $\frac{\mathsf{D}\,S_n}{n^2} \to 0$ при $n \to \infty$.*

Теорема Маркова утверждает, что ЗБЧ выполнен, если дисперсия суммы n слагаемых растёт не слишком быстро с ростом n.

Сильная зависимость слагаемых приводит обычно к невыполнению ЗБЧ. Если, например, $\mathsf{D}\,\xi_1 \neq 0$ и $\xi_n \equiv \xi_1$, то $S_n = n\xi_1$, и свойство (22) не выполнено (*убедиться в этом*!). В этом случае не выполнено и достаточное условие для ЗБЧ: $\mathsf{D}\,S_n = \mathsf{D}\,(n\xi_1) = cn^2$. Для одинаково распределённых слагаемых дисперсия суммы ещё быстрее расти уже не может.

Следующее утверждение мы докажем чуть позже. Сравните его условия с условиями ЗБЧ Чебышёва.

Т е о р е м а 38 (З Б Ч Х и н ч и н а[18]). *Для любой последовательности* $\xi_1, \xi_2, \ldots$ *независимых (в совокупности) и одинаково распределённых случайных величин с конечным первым моментом* $\mathsf{E}\,|\xi_1| < \infty$ *имеет место сходимость:*

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{\text{p}} \mathsf{E}\,\xi_1.$$

Итак, чтобы последовательность независимых и одинаково распределённых случайных величин удовлетворяла ЗБЧ, достаточно существования *первого* момента слагаемых. Более того, в условиях теоремы 38 имеет место и сходимость п. н. последовательности $(\xi_1 + \ldots + \xi_n)/n$ к $\mathsf{E}\,\xi_1$. Это утверждение называется усиленным законом больших чисел (УЗБЧ) Колмогорова, и его мы доказывать не будем.

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.

Т е о р е м а 39 (З Б Ч Б е р н у л л и). *Пусть событие A может произойти в любом из n независимых испытаний с одной и той же вероятностью p, и пусть $\nu_n(A)$ — число осуществлений события A в n испытаниях. Тогда* $\dfrac{\nu_n(A)}{n} \xrightarrow{\text{p}} p$. *При этом для любого* $\varepsilon > 0$

$$\mathsf{P}\left(\left|\frac{\nu_n(A)}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{p(1-p)}{n\varepsilon^2}.$$

Д о к а з а т е л ь с т в о. Заметим, что $\nu_n(A)$ есть сумма независимых, одинаково распределённых случайных величин, имеющих распределение Бернулли с параметром $p = \mathsf{P}(A)$ (индикаторов того, что в соответствующем испытании произошло A): $\nu_n(A) = \xi_1 + \ldots + \xi_n$, где

$$\xi_i = \begin{cases} 1, & \text{если } A \text{ произошло в } i\text{-м испытании;} \\ 0, & \text{если } A \text{ не произошло в } i\text{-м испытании;} \end{cases}$$

и $\mathsf{E}\,\xi_1 = \mathsf{P}(A) = p$, $\mathsf{D}\,\xi_1 = p(1-p)$. Осталось воспользоваться ЗБЧ в форме Чебышёва и неравенством (24). □

П р и м е р 70. Монета подбрасывается 10^4 раз. Оценим вероятность того, что частота выпадения герба отличается от $\frac{1}{2}$ на 0,01 или более.

[18]Александр Яковлевич Хинчин (19.07.1894—18.11.1959).

Пусть $\xi_1, \ldots, \xi_n$ — независимые случайные величины, каждая из которых имеет распределение Бернулли с параметром $p = 1/2$ и равна единице, если при соответствующем подбрасывании выпал герб, и нулю иначе. Нужно оценить $\mathsf{P}\left(\left|\frac{\nu_n}{n} - \frac{1}{2}\right| \geqslant 0{,}01\right)$, где $n = 10^4$, а $\nu_n = \sum\limits_{i=1}^{n} \xi_i$ — число выпадений герба. Поскольку $\mathsf{D}\,\xi_1 = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$, искомая оценка сверху выглядит так:

$$\mathsf{P}\left(\left|\frac{\nu_n}{n} - \frac{1}{2}\right| \geqslant 0{,}01\right) \leqslant \frac{\mathsf{D}\,\xi_1}{n \cdot 0{,}01^2} = \frac{1}{4 \cdot 10^4 \cdot 10^{-4}} = \frac{1}{4}.$$

Итак, 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. Мы увидим, насколько это грубая оценка, когда познакомимся с *центральной предельной теоремой.*

ГЛАВА XI

ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

> Из этой первой лекции по теории вероятностей я запомнил только полузнакомый термин «математическое ожидание». Незнакомец употреблял этот термин неоднократно, и каждый раз я представлял себе большое помещение, вроде зала ожидания, с кафельным полом, где сидят люди с портфелями и бюварами и, подбрасывая время от времени к потолку монетки и бутерброды, сосредоточенно чего-то ожидают. До сих пор я часто вижу это во сне. Но тут незнакомец оглушил меня звонким термином «предельная теорема Муавра — Лапласа» и сказал, что всё это к делу не относится.
>
> *Аркадий и Борис Стругацкие. Стажёры*

§ 1. Как быстро среднее арифметическое сходится к математическому ожиданию?

Пусть, как в законе больших чисел Чебышёва, $S_n = \xi_1 + \ldots + \xi_n$ — сумма n независимых и одинаково распределённых величин с конечной дисперсией. Тогда по ЗБЧ $\dfrac{S_n}{n} \xrightarrow{\text{p}} \mathsf{E}\,\xi_1$ с ростом n. Или, после приведения к общему знаменателю,

$$\frac{S_n - n\,\mathsf{E}\,\xi_1}{n} \xrightarrow{\text{p}} 0.$$

Если при делении на n мы получили в пределе нуль (в смысле некоторой, всё равно какой, сходимости), резонно задать себе вопрос: а не слишком ли на большую величину мы поделили? Нельзя ли поделить на что-нибудь, растущее к бесконечности медленнее, чем n, чтобы получить в пределе не нуль (но и не бесконечность)?

Можно поставить тот же вопрос иначе. Есть последовательность, стремящаяся к нулю. Можно ли её домножить на что-либо растущее, чтобы «погасить» это стремление к нулю и получить, тем самым, что-нибудь конечное и ненулевое в пределе?

Оказывается, что уже последовательность случайных величин
$$\frac{S_n - n\,\mathsf{E}\,\xi_1}{\sqrt{n}} = \sqrt{n}\cdot\frac{S_n - n\,\mathsf{E}\,\xi_1}{n}$$
не сходится к нулю. Распределение членов этой последовательности становится всё более похожим на нормальное распределение! Можно считать, что такая последовательность сходится к случайной величине, имеющей нормальное распределение, но сходится никак не по вероятности, а только в смысле сходимости распределений, или «слабой сходимости».

§ 2. Слабая сходимость

Пусть задана последовательность случайных величин $\xi_1, \xi_2, \ldots$, задано некоторое распределение $\mathcal{F}$ с функцией распределения F_ξ и пусть ξ — произвольная случайная величина, имеющая распределение $\mathcal{F}$.

Определение 46. Говорят, что последовательность случайных величин $\xi_1, \xi_2, \ldots$ сходится *слабо* или *по распределению* к случайной величине ξ и пишут: $\xi_n \Rightarrow \xi$, если для любого x такого, что функция распределения F_ξ непрерывна в точке x, имеет место сходимость $F_{\xi_n}(x) \to F_\xi(x)$ при $n \to \infty$.

Итак, слабая сходимость — это сходимость функций распределения во всех точках непрерывности предельной функции распределения.

Замечание. Заметим, что сходимость $\xi_n \Rightarrow \xi$ есть сходимость распределений, а не случайных величин: если «предельную» величину ξ заменить на другую величину η с тем же распределением, ничего не изменится: в том же смысле $\xi_n \Rightarrow \eta$.

Следующее свойство очевидно. Если нет — нужно вернуться к определению и свойствам функций распределения.

Свойство 25. *Если* $\xi_n \Rightarrow \xi$, *и функция распределения* F_ξ *непрерывна в точках* a *и* b, *то* $\mathsf{P}(\xi_n \in (a, b)) \to \mathsf{P}(\xi \in (a, b))$. *Если во всех точках* a *и* b *непрерывности функции распределения* F_ξ *имеет место сходимость* $\mathsf{P}(\xi_n \in (a, b)) \to \mathsf{P}(\xi \in (a, b))$, *то* $\xi_n \Rightarrow \xi$.

Свойство 26. 1. *Если* $\xi_n \xrightarrow{\text{p}} \xi$, *то* $\xi_n \Rightarrow \xi$.

2. *Если* $\xi_n \Rightarrow c = const$, *то* $\xi_n \xrightarrow{\text{p}} c$.

Итак, сходимость по вероятности влечёт слабую сходимость. Обратное утверждение в общем случае смысла не имеет (см. замечание выше). Однако из слабой сходимости *к постоянной* вытекает сходимость по вероятности.

Доказательство. Первое утверждение мы докажем чуть позже.

Докажем, что слабая сходимость к постояннной влечёт сходимость по вероятности. Пусть $\xi_n \Rightarrow c$, т. е.

$$F_{\xi_n}(x) \to F_c(x) = \begin{cases} 0, & x \leqslant c; \\ 1, & x > c \end{cases}$$

при любом x, являющемся точкой непрерывности предельной функции $F_c(x)$, т. е. при всех $x \neq c$.

Возьмём произвольное $\varepsilon > 0$ и докажем, что $\mathsf{P}(|\xi_n - c| < \varepsilon) \to 1$:

$$\mathsf{P}(-\varepsilon < \xi_n - c < \varepsilon) = \mathsf{P}(c - \varepsilon < \xi_n < c + \varepsilon) \geqslant \mathsf{P}(c - \varepsilon/2 \leqslant \xi_n < c + \varepsilon) =$$

$$= F_{\xi_n}(c + \varepsilon) - F_{\xi_n}(c - \varepsilon/2) \ \to \ F_c(c + \varepsilon) - F_c(c - \varepsilon/2) = 1 - 0 = 1,$$

поскольку в точках $c+\varepsilon$ и $c-\varepsilon/2$ функция F_c непрерывна, и, следовательно, имеет место сходимость последовательностей $F_{\xi_n}(c+\varepsilon)$ к $F_c(c+\varepsilon) = 1$ и $F_{\xi_n}(c - \varepsilon/2)$ к $F_c(c - \varepsilon/2) = 0$.

Осталось заметить, что $\mathsf{P}(|\xi_n - c| < \varepsilon)$ не бывает больше 1, так что по свойству предела зажатой последовательности $\mathsf{P}(|\xi_n - c| < \varepsilon) \to 1$. □

Следующее свойство приводит пример операций, которые можно применять к слабо сходящимся последовательностям — домножать их на последовательности, сходящиеся по вероятности к постоянным величинам.

Замечание. Свойство «предел суммы равен сумме пределов» для слабой сходимости просто бессмысленно: сходимости $\xi_n \Rightarrow \xi$, $\eta_n \Rightarrow \eta$ означают, что нам известны предельные распределения этих последовательностей. Но предельное распределение их суммы может быть различным в зависимости от совместного распределения ξ_n и η_n. Иное дело, когда одно из предельных распределений вырождено. Тогда предельная функция распределения суммы или произведения определена однозначно.

Свойство 27. 1. *Если* $\xi_n \xrightarrow{\text{p}} c = const$ *и* $\eta_n \Rightarrow \eta$, *то* $\xi_n \cdot \eta_n \Rightarrow c\eta$.

2. *Если* $\xi_n \xrightarrow{\text{p}} c = const$ *и* $\eta_n \Rightarrow \eta$, *то* $\xi_n + \eta_n \Rightarrow c + \eta$.

Доказательство. Нелюбознательный читатель может пропустить это доказательство, вернувшись к нему при втором прочтении.

Заметим вначале, что если $\eta_n \Rightarrow \eta$, то $c\eta_n \Rightarrow c\eta$ и $c + \eta_n \Rightarrow c + \eta$ (доказать). Поэтому достаточно доказать первое утверждение свойства 27 при $c = 1$, а второе утверждение — при $c = 0$.

Рассмотрим второе утверждение, оставив первое любознательному читателю. Пусть $\xi_n \xrightarrow{\text{p}} 0$ и $\eta_n \Rightarrow \eta$. Докажем, что тогда $\xi_n + \eta_n \Rightarrow \eta$.

Пусть x_0 — точка непрерывности функции распределения $F_\eta(x)$. Требуется доказать, что имеет место сходимость $F_{\xi_n+\eta_n}(x_0)$ к $F_\eta(x_0)$. Зафиксируем $\varepsilon > 0$ такое, что $F_\eta(x)$ непрерывна в точках $x_0 \pm \varepsilon$.

События $H_1 = \{|\xi_n| \geqslant \varepsilon\}$ и $H_2 = \{|\xi_n| < \varepsilon\}$ образуют полную группу, поэтому

$$F_{\xi_n+\eta_n}(x_0) = \mathsf{P}(\xi_n + \eta_n < x_0,\, H_1) + \mathsf{P}(\xi_n + \eta_n < x_0,\, H_2) = P_1 + P_2.$$

Оценим $P_1 + P_2$ сверху и снизу. Для P_1 имеем

$$0 \leqslant P_1 = \mathsf{P}(\xi_n + \eta_n < x_0,\, H_1) \leqslant \mathsf{P}(H_1) = \mathsf{P}(|\xi_n| \geqslant \varepsilon),$$

и последняя вероятность может быть выбором n сделана сколь угодно малой. Для P_2, с одной стороны,

$$P_2 = \mathsf{P}(\xi_n + \eta_n < x_0,\; -\varepsilon < \xi_n < \varepsilon) \leqslant \mathsf{P}(-\varepsilon + \eta_n < x_0) = F_{\eta_n}(x_0 + \varepsilon).$$

Выше мы воспользовались тем, что если $-\varepsilon < \xi_n$ и $\xi_n + \eta_n < x_0$, то тем более $-\varepsilon + \eta_n < x_0$. С другой стороны,

$$\begin{aligned} P_2 &= \mathsf{P}(\xi_n + \eta_n < x_0,\, -\varepsilon < \xi_n < \varepsilon) \geqslant \mathsf{P}(\varepsilon + \eta_n < x_0,\, -\varepsilon < \xi_n < \varepsilon) \geqslant \\ &\geqslant \mathsf{P}(\varepsilon + \eta_n < x_0) - \mathsf{P}(|\xi_n| \geqslant \varepsilon) = F_{\eta_n}(x_0 - \varepsilon) - \mathsf{P}(|\xi_n| \geqslant \varepsilon). \end{aligned}$$

Здесь первое неравенство объясняется включением

$$\{\varepsilon + \eta_n < x_0,\, -\varepsilon < \xi_n < \varepsilon\} \subseteq \{\xi_n + \eta_n < x_0,\, -\varepsilon < \xi_n < \varepsilon\},$$

которое получилось заменой в событии $\{\varepsilon + \eta_n < x_0\}$ числа ε на меньшую величину ξ_n, $\xi_n < \varepsilon$. Второе неравенство следует из свойств:

$$\mathsf{P}(A\overline{B}) \leqslant \mathsf{P}(\overline{B}),\ \text{поэтому}\ \mathsf{P}(AB) = \mathsf{P}(A) - \mathsf{P}(A\overline{B}) \geqslant \mathsf{P}(A) - \mathsf{P}(\overline{B}).$$

Мы получили оценки снизу и сверху для $P_1 + P_2$, т. е. для $F_{\xi_n+\eta_n}(x_0)$:

$$F_{\eta_n}(x_0 - \varepsilon) - \mathsf{P}(|\xi_n| \geqslant \varepsilon) \leqslant F_{\xi_n+\eta_n}(x_0) \leqslant \mathsf{P}(|\xi_n| \geqslant \varepsilon) + F_{\eta_n}(x_0 + \varepsilon).$$

Устремляя n к бесконечности и вспоминая, что $x_0 \pm \varepsilon$ — точки непрерывности функции распределения F_η, получаем

$$F_\eta(x_0 - \varepsilon) \leqslant \underline{\lim}\, F_{\xi_n+\eta_n}(x_0) \leqslant \overline{\lim}\, F_{\xi_n+\eta_n}(x_0) \leqslant F_\eta(x_0 + \varepsilon). \tag{25}$$

У любой функции распределения не более чем счётное множество точек разрыва. Поэтому можно выбрать такую уменьшающуюся до нуля последовательность ε, что в точках $x_0 \pm \varepsilon$ функция распределения F_η будет непрерывной и, следовательно, останутся верны неравенства (25). Переходя к пределу по такой последовательности $\varepsilon \to 0$ и помня, что x_0 — точка непрерывности функции F_η, получаем, что нижний и верхний пределы $F_{\xi_n+\eta_n}(x_0)$ при $n \to \infty$ совпадают и равны $F_\eta(x_0)$. □

Доказательство утверждения (1) из свойства 26. В качестве простого следствия из только что доказанного второго утверждения свойства 27 покажем, что сходимость $\xi_n \xrightarrow{\text{p}} \xi$ по вероятности влечёт слабую сходимость $\xi_n \Rightarrow \xi$.

Представим ξ_n в виде суммы $\xi_n = (\xi_n - \xi) + \xi$. Здесь последовательность $\xi_n - \xi$ по вероятности стремится к нулю, а «последовательность» ξ слабо сходится к ξ. Поэтому их сумма слабо сходится к ξ. $\square$

Получим ещё одно следствие из свойства 27. Для удобства ссылок назовём следующее утверждение «теоремой о двойном пределе».

Теорема 40. *Пусть $\xi_n \Rightarrow \xi$, причём функция распределения случайной величины ξ непрерывна всюду, и пусть $x_n \to x_0 \in [-\infty, \infty]$ — числовая последовательность. Тогда $F_{\xi_n}(x_n) \to F_\xi(x_0)$.*

В формулировке теоремы мы для краткости использовали запись $F_\xi(\pm\infty)$, которую следует понимать так: $F_\xi(-\infty) = 0$, $F_\xi(+\infty) = 1$.

Доказательство. Если $x_0 \in \mathbb{R}$, то утверждение теоремы следует из свойства 27. Действительно, из $x_n \to x_0$ следует, что $x_n \xrightarrow{\text{p}} x_0$. Тогда $\xi_n - x_n \Rightarrow \xi - x_0$ по свойству 27. Функция распределения $F_{\xi - x_0}(x)$ отличается от $F_\xi(x)$ лишь сдвигом и тоже непрерывна всюду, поэтому имеет место сходимость функций распределения в любой точке. В частности, в точке $x = 0$ имеет место сходимость при $n \to \infty$

$$F_{\xi_n}(x_n) = \mathsf{P}(\xi_n - x_n < 0) = F_{\xi_n - x_n}(0) \to F_{\xi - x_0}(0) = F_\xi(x_0).$$

Пусть теперь $x_0 = -\infty$. Нужно доказать, что $F_{\xi_n}(x_n) \to F_\xi(-\infty) = 0$.

По определению, $x_n \to -\infty$ с ростом n, если для любого $M > 0$ существует N такое, что при $n \geqslant N$ выполнено неравенство: $x_n \leqslant -M$. В силу монотонности функций распределения, $0 \leqslant F_{\xi_n}(x_n) \leqslant F_{\xi_n}(-M)$. В точке $-M$, как и в любой иной точке, имеет место сходимость функций распределения $F_{\xi_n}(-M) \to F_\xi(-M)$. Выбором M величина $F_\xi(-M)$ может быть сделана сколь угодно близкой к нулю. Тем самым верхний предел последовательности $F_{\xi_n}(x_n)$ оказывается зажат между нулём и сколь угодно малой величиной, т. е. равняется нулю.

Случай $x_0 = +\infty$ проверяется аналогично. $\square$

Основной источник слабо сходящихся последовательностей и необычайно мощное и универсальное средство для асимптотического анализа *распределений сумм* независимых и одинаково распределённых случайных величин предоставляет нам центральная предельная теорема.

§ 3. Центральная предельная теорема

Мы будем называть следующее утверждение «ЦПТ Ляпунова[19]», но сформулируем и докажем теорему Ляпунова только в частном случае — для последовательности независимых и одинаково распределённых случайных величин. Как и ранее, через S_n обозначена сумма первых n случайных величин в последовательности: $S_n = \xi_1 + \ldots + \xi_n$.

Теорема 41 (ЦПТ Ляпунова). *Пусть* $\xi_1, \xi_2, \ldots$ — *независимые и одинаково распределённые случайные величины с конечной и ненулевой дисперсией:* $0 < \mathsf{D}\,\xi_1 < \infty$. *Тогда имеет место слабая сходимость*

$$\frac{S_n - n\,\mathsf{E}\,\xi_1}{\sqrt{n\,\mathsf{D}\,\xi_1}} \Rightarrow \mathrm{N}_{0,1}$$

последовательности центрированных и нормированных сумм случайных величин к стандартному нормальному распределению.

Пользуясь определением и свойствами слабой сходимости и заметив, что функция распределения $\Phi_{a,\sigma^2}(x)$ любого нормального закона непрерывна всюду на $\mathbb{R}$ (*почему*?), утверждение ЦПТ можно сформулировать любым из следующих способов.

Следствие 18. *Пусть* $\xi_1, \xi_2, \ldots$ — *независимые и одинаково распределённые случайные величины с конечной и ненулевой дисперсией. Тогда выполнены утверждения*:

а) *для любых вещественных* $x < y$ *при* $n \to \infty$ *имеет место сходимость*

$$\mathsf{P}\left(x \leqslant \frac{S_n - n\,\mathsf{E}\,\xi_1}{\sqrt{n\,\mathsf{D}\,\xi_1}} \leqslant y\right) \to \Phi_{0,1}(y) - \Phi_{0,1}(x) = \int\limits_x^y \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt;$$

б) *если* η — *произвольная случайная величина со стандартным нормальным распределением, то*

$$\sqrt{n}\,\left(\frac{S_n}{n} - \mathsf{E}\,\xi_1\right) = \frac{S_n - n\,\mathsf{E}\,\xi_1}{\sqrt{n}} \Rightarrow \sqrt{\mathsf{D}\,\xi_1}\cdot\eta \in \mathrm{N}_{0,\mathsf{D}\,\xi_1}.$$

Мы докажем центральную предельную теорему и закон больших чисел в форме Хинчина в следующей главе. Нам потребуется для этого познакомиться с мощным математическим инструментом, который в математике

[19] Александр Михайлович Ляпунов (6.06.1857—3.11.1918).

обычно называют преобразованиями Фурье, а в теории вероятностей — характеристическими функциями.

§ 4. Предельная теорема Муавра — Лапласа

Получим в качестве следствия из ЦПТ Ляпунова предельную теорему Муавра[20] и Лапласа[21]. Подобно ЗБЧ Бернулли, это утверждение годится только для схемы Бернулли.

Теорема 42 (предельная теорема Муавра — Лапласа). *Пусть событие A может произойти в любом из n независимых испытаний с одной и той же вероятностью p и пусть $\nu_n(A)$ — число осуществлений события A в n испытаниях. Тогда*

$$\frac{\nu_n(A) - np}{\sqrt{np(1-p)}} \Rightarrow \mathrm{N}_{0,1} \ \text{ при } \ n \to \infty,$$

т. е. для любых вещественных $x < y$ имеет место сходимость

$$\mathsf{P}\left(x \leqslant \frac{\nu_n(A) - np}{\sqrt{np(1-p)}} \leqslant y\right) \to \Phi_{0,1}(y) - \Phi_{0,1}(x) = \int\limits_x^y \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt;$$

Доказательство. Величина $\nu_n(A)$ есть сумма независимых, одинаково распределённых случайных величин, имеющих распределение Бернулли с параметром, равным вероятности успеха p: $\nu_n(A) = \xi_1 + \ldots + \xi_n$, где $\mathsf{E}\,\xi_1 = p$, $\mathsf{D}\,\xi_1 = p(1-p)$. Осталось воспользоваться ЦПТ. □

Пример 71. Задача из примера 70 (с. 119). Требуется найти

$$\mathsf{P}\left(\left|\frac{\nu_n}{n} - \frac{1}{2}\right| \geqslant 0{,}01\right),$$

где $n = 10\,000$, ν_n — число выпадений герба. Вычислим вероятность дополнительного события. Домножим обе части неравенства под знаком вероятности на $\sqrt{n} = 100$ и поделим на $\sqrt{p(1-p)} = 1/2$.

$$\mathsf{P}\left(\left|\frac{\nu_n}{n} - \frac{1}{2}\right| < 0{,}01\right) = \mathsf{P}\left(\frac{\sqrt{n}}{\sqrt{p(1-p)}}\left|\frac{\nu_n}{n} - p\right| < 0{,}01\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) =$$

$$= \mathsf{P}\left(\frac{\sqrt{n}}{\sqrt{p(1-p)}}\left|\frac{\nu_n}{n} - p\right| < 2\right) = \mathsf{P}\left(-2 < \frac{\nu_n - np}{\sqrt{np(1-p)}} < 2\right) \approx$$

$$\approx \Phi_{0,1}(2) - \Phi_{0,1}(-2) = 1 - 2\Phi_{0,1}(-2) = 1 - 2 \cdot 0{,}0228 = 1 - 0{,}0456.$$

[20] Abraham de Moivre (26.05.1667—27.11.1754, France, England).

[21] Pierre-Simon Laplace (23.03.1749—5.03.1827, France).

Искомая вероятность примерно равна $0{,}0456$:
$$\mathsf{P}\left(\left|\frac{\nu_n}{n}-\frac{1}{2}\right|\geqslant 0{,}01\right)=1-\mathsf{P}\left(\left|\frac{\nu_n}{n}-\frac{1}{2}\right|<0{,}01\right)\approx 0{,}0456.$$

Центральной предельной теоремой пользуются для приближённого вычисления вероятностей, связанных с суммами большого числа независимых и одинаково распределённых величин. При этом распределение центрированной и нормированной суммы заменяют на стандартное нормальное распределение. Насколько велика ошибка при такой замене (погрешность приближения)?

Упражнение. Какие ещё предельные теоремы для схемы Бернулли вы знаете? Что такое теорема Пуассона? Найти её. Какова погрешность пуассоновского приближения? Вычислить её. Объяснить, почему теорема Пуассона не применима в задаче из примера 71.

В примере 71 мы вычислили вероятность приближённо. Следующий результат позволяет оценить погрешность приближения в ЦПТ.

Теорема 43 (неравенство Берри—Эссéена). *В условиях ЦПТ для любого $x\in\mathbb{R}$ и для любого распределения ξ_1 с конечным третьим моментом*
$$\left|\;\mathsf{P}\left(\frac{S_n-n\,\mathsf{E}\,\xi_1}{\sqrt{n\,\mathsf{D}\,\xi_1}}<x\right)\;-\;\Phi_{0,1}(x)\;\right|\;\leqslant\;C\cdot\frac{\mathsf{E}\,|\xi_1-\mathsf{E}\,\xi_1|^3}{\sqrt{n}\big(\sqrt{\mathsf{D}\,\xi_1}\,\big)^3}.$$

Замечание. В качестве постоянной C можно брать число $0{,}4$.

Продолжение примера 71. Проверьте, что для случайной величины ξ_1 с распределением Бернулли
$$\mathsf{E}\,|\xi_1-\mathsf{E}\,\xi_1|^3=|0-p\,|^3\,\mathsf{P}(\xi_1=0)+|1-p\,|^3\,\mathsf{P}(\xi_1=1)=pq(p^2+q^2).$$

Поэтому разница между левой и правой частями приближённого равенства «$\approx$» в примере 71 при $n=10^4$ и $p=q=\dfrac{1}{2}$ не превышает величины
$$C\cdot\frac{pq(p^2+q^2)}{\sqrt{npq}(\sqrt{pq})^3}=C\cdot\frac{p^2+q^2}{\sqrt{n}\,\sqrt{pq}}\leqslant 0{,}4\cdot\frac{1}{100}=0{,}004,$$

т. е. искомая вероятность $\mathsf{P}\left(\left|\dfrac{\nu_n}{n}-\dfrac{1}{2}\right|>0{,}01\right)$ не больше, чем $0{,}0456+$ $+0{,}004$. Уместно сравнить этот ответ с оценкой $\dfrac{1}{4}$, полученной с помощью ЗБЧ в примере 70.

ГЛАВА XII

ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ

> Я напрямик спросил, какую пользу можно извлечь от изучения его работ о покере. «Примерно такую же, как от чтения персидской поэзии», — ответил фон Нейман.
>
> *Д. Мак-Дональд. Игра называется бизнес*

§ 1. Определение и примеры

В этой главе $i = \sqrt{-1}$ — мнимая единица, t — вещественная переменная, $e^{it} = \cos t + i \sin t$ — формула Эйлера, $\mathsf{E}\,(\eta + i\zeta) = \mathsf{E}\,\eta + i\,\mathsf{E}\,\zeta$ — способ вычисления математического ожидания комплекснозначной случайной величины $\eta + i\zeta$, если математические ожидания её действительной (η) и мнимой (ζ) частей существуют.

Как всегда, модулем комплексного числа $z = x + iy$ называется положительное число $|z| = \sqrt{x^2 + y^2}$, так что $\left|e^{it}\right| = 1$.

Определение 47. Функция $\varphi_\xi(t) = \mathsf{E}\,e^{it\xi}$ вещественной переменной t называется *характеристической функцией* случайной величины ξ.

Пример 72. Пусть случайная величина ξ имеет распределение Бернулли с параметром p. Её характеристическая функция равна

$$\varphi_\xi(t) = \mathsf{E}\,e^{it\xi} = e^{it\cdot 0}\,\mathsf{P}(\xi = 0) + e^{it\cdot 1}\,\mathsf{P}(\xi = 1) = 1 - p + pe^{it}.$$

Пример 73. Пусть случайная величина ξ имеет биномиальное распределение с параметрами n и p. Её характеристическая функция равна

$$\begin{aligned}\varphi_\xi(t) &= \mathsf{E}\,e^{it\xi} = \sum_{k=0}^{n} e^{it\cdot k}\,\mathsf{P}(\xi = k) = \sum_{k=0}^{n} e^{it\cdot k}\,C_n^k\,p^k\,(1-p)^{n-k} = \\ &= \sum_{k=0}^{n} C_n^k\left(pe^{it}\right)^k (1-p)^{n-k} = \left(1 - p + pe^{it}\right)^n.\end{aligned}$$

Последнее равенство есть бином Ньютона.

Пример 74. Пусть случайная величина ξ имеет распределение Пуассона с параметром λ. Её характеристическая функция равна

$$\begin{aligned}\varphi_\xi(t) &= \mathsf{E}\, e^{it\xi} = \sum_{k=0}^{\infty} e^{it\cdot k}\,\mathsf{P}(\xi = k) = \sum_{k=0}^{\infty} e^{it\cdot k}\,\frac{\lambda^k}{k!}\,e^{-\lambda} = \\ &= e^{-\lambda}\sum_{k=0}^{\infty}\frac{\left(\lambda e^{it}\right)^k}{k!} = e^{-\lambda}e^{\lambda e^{it}} = \exp\{\lambda\left(e^{it}-1\right)\}.\end{aligned}$$

Пример 75. Пусть случайная величина ξ имеет гамма-распределение с параметрами α и λ. Её характеристическая функция равна

$$\begin{aligned}\varphi_\xi(t) &= \mathsf{E}\, e^{it\xi} = \int\limits_0^{\infty} e^{it\cdot x}\, f_\xi(x)\,dx = \int\limits_0^{\infty} e^{itx}\,\frac{\alpha^\lambda}{\Gamma(\lambda)}\,x^{\lambda-1}\,e^{-\alpha x}\,dx = \\ &= \frac{\alpha^\lambda}{\Gamma(\lambda)}\int\limits_0^{\infty} x^{\lambda-1}\,e^{-x(\alpha-it)}\,dx = \left(\frac{\alpha}{\alpha-it}\right)^{\lambda} = \left(1-\frac{it}{\alpha}\right)^{-\lambda}.\end{aligned}$$

Интеграл мы вычислили с помощью гамма-функции: замена $y = x(\alpha-it)$ даёт

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

В качестве следствия получим, что для случайной величины ξ с показательным распределением $\mathrm{E}_\alpha = \Gamma_{\alpha,1}$ характеристическая функция равна $\varphi_\xi(t) = \dfrac{\alpha}{\alpha-it}$.

Пример 76. Пусть случайная величина ξ имеет стандартное нормальное распределение. Её характеристическая функция равна

$$\begin{aligned}\varphi_\xi(t) &= \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{itx}\,e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-t^2/2}\,e^{-(x-it)^2/2}\,dx = \\ &= e^{-t^2/2}\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-(x-it)^2/2}\,d(x-it) = e^{-t^2/2}.\end{aligned}$$

При интегрировании мы выделили полный квадрат в показателе экспоненты и вспомнили, чему равен интеграл по $\mathbb{R}$ от функции $\dfrac{1}{\sqrt{2\pi}}\,e^{-u^2/2}$ (*а чему он равен*?).

Самое время остановиться и спросить: «Ну и что? Зачем нам эти функции и какой от них прок?» Давайте познакомимся с замечательными свойствами характеристических функций.

§ 2. Свойства характеристических функций

(Ф1). Характеристическая функция всегда существует:

$$|\varphi_\xi(t)| = \left|\mathsf{E}\, e^{it\xi}\right| \leqslant 1.$$

Доказательство. Воспользуемся свойством $\mathsf{D}\eta \geqslant 0$, равносильным неравенству $\left(\mathsf{E}\,\eta\right)^2 \leqslant \mathsf{E}\eta^2$:

$$\begin{aligned} |\varphi_\xi(t)|^2 &= \left|\mathsf{E}\,\cos(t\xi) + i\mathsf{E}\,\sin(t\xi)\right|^2 = \left(\mathsf{E}\,\cos(t\xi)\right)^2 + \left(\mathsf{E}\,\sin(t\xi)\right)^2 \leqslant \\ &\leqslant \mathsf{E}\,\cos^2(t\xi) + \mathsf{E}\,\sin^2(t\xi) = \mathsf{E}\left(\cos^2(t\xi) + \sin^2(t\xi)\right) = \mathsf{E}\,1 = 1. \quad \square \end{aligned}$$

(Ф2). По характеристической функции однозначно восстанавливается распределение (функция распределения, плотность или таблица распределения). Другими словами, если две случайные величины имеют одинаковые характеристические функции, то и распределения этих величин совпадают.

Формулы, с помощью которых по характеристической функции восстанавливается распределение, в анализе называют формулами «обратного преобразования Фурье». Например, если модуль характеристической функции интегрируем на всей прямой, то у случайной величины есть плотность распределения и она находится по формуле

$$f_\xi(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-itx}\,\varphi_\xi(t)\,dt.$$

Ни одна из формул обратного преобразования Фурье нам не понадобится.

(Ф3). Характеристическая функция случайной величины $a + b\xi$ связана с характеристической функцией случайной величины ξ равенством

$$\varphi_{a+b\xi}(t) = \mathsf{E}\,e^{it(a+b\xi)} = e^{ita}\mathsf{E}\,e^{i(tb)\xi} = e^{ita}\,\varphi_\xi(tb).$$

Пример 77. Вычислим характеристическую функцию случайной величины ξ, имеющей нормальное распределение с параметрами a и σ^2. Мы знаем, что у «стандартизованной» случайной величины $\eta = (\xi - a)/\sigma$ характеристическая функция равна $\varphi_\eta(t) = e^{-t^2/2}$. Тогда характеристическая функция величины $\xi = a + \sigma\eta$ равна

$$\varphi_\xi(t) = \varphi_{a+\sigma\eta}(t) = e^{ita}\,\varphi_\eta(t\sigma) = e^{ita}\,e^{-(t\sigma)^2/2}.$$

(Ф4). Характеристическая функция суммы независимых случайных величин равна произведению характеристических функций слагаемых: если случайные величины ξ и η независимы, то, по свойству (E7) математических ожиданий,

$$\varphi_{\xi+\eta}(t) = \mathsf{E}\, e^{it(\xi+\eta)} = \mathsf{E}\, e^{it\xi}\, \mathsf{E}\, e^{it\eta} = \varphi_\xi(t)\, \varphi_\eta(t).$$

З а м е ч а н и е. Чтобы характеристическая функция суммы n случайных величин распадалась в произведение их характеристических функций, попарной независимости слагаемых не хватит. То же самое можно сказать про свойство (E7) математических ожиданий.

Замечательным свойством (Ф4) мы сразу же воспользуемся, как обещали, для доказательства леммы 3 (с. 86), утверждающей устойчивость нормального распределения относительно суммирования.

Д о к а з а т е л ь с т в о л е м м ы 3. Пусть $\xi \in \mathrm{N}_{a_1, \sigma_1^2}$ и $\eta \in \mathrm{N}_{a_2, \sigma_2^2}$ независимы. Характеристическая функция суммы $\xi + \eta$ равна

$$\varphi_{\xi+\eta}(t) = \varphi_\xi(t)\, \varphi_\eta(t) = e^{ita_1} e^{-t^2\sigma_1^2/2} e^{ita_2} e^{-t^2\sigma_2^2/2} = e^{it(a_1+a_2)} e^{-t^2(\sigma_1^2+\sigma_2^2)/2}.$$

Видим, что характеристическая функция суммы есть характеристическая функция нормального распределения с параметрами $a_1 + a_2$ и $\sigma_1^2 + \sigma_2^2$. Следовательно, $\xi + \eta \in \mathrm{N}_{a_1+a_2, \sigma_1^2+\sigma_2^2}$ по свойству (Ф2). □

Д о к а з а т е л ь с т в о л е м м 1, 2, 4 (с. 86). Докажем свойства устойчивости по суммированию биномиального распределения, распределения Пуассона и гамма-распределения, используя характеристические функции из примеров 72— 75.

Для независимых случайных величин с распределениями Пуассона Π_λ и Π_μ характеристическая функция суммы

$$\varphi_{\xi+\eta}(t) = \exp\left\{\lambda\left(e^{it} - 1\right)\right\} \exp\left\{\mu\left(e^{it} - 1\right)\right\} = \exp\left\{(\lambda+\mu)\left(e^{it} - 1\right)\right\}$$

равна характеристической функции распределения Пуассона $\Pi_{\lambda+\mu}$.

Для независимых случайных величин с биномиальными распределениями $\mathrm{B}_{n,p}$ и $\mathrm{B}_{m,p}$ характеристическая функция суммы

$$\varphi_{\xi+\eta}(t) = \varphi_\xi(t)\, \varphi_\eta(t) = \left(1 - p + pe^{it}\right)^n \left(1 - p + pe^{it}\right)^m = \left(1 - p + pe^{it}\right)^{n+m}$$

равна характеристической функции биномиального распределения с параметрами $n + m$ и p.

Для независимых случайных величин с гамма-распределениями $\Gamma_{\alpha,\lambda_1}$ и $\Gamma_{\alpha,\lambda_2}$ характеристическая функция суммы

$$\varphi_{\xi+\eta}(t) = \left(1 - \frac{it}{\alpha}\right)^{-\lambda_1}\left(1 - \frac{it}{\alpha}\right)^{-\lambda_2} = \left(1 - \frac{it}{\alpha}\right)^{-(\lambda_1+\lambda_2)}$$

равна характеристической функции гамма-распределения $\Gamma_{\alpha,\lambda_1+\lambda_2}$. □

(Ф5.) Пусть существует момент порядка $k \in \mathbb{N}$ случайной величины ξ, т. е. $\mathsf{E}\,|\xi|^k < \infty$. Тогда характеристическая функция $\varphi_\xi(t)$ непрерывно дифференцируема k раз и её k-я производная *в нуле* связана с моментом порядка k равенством

$$\varphi_\xi^{(k)}(0) = \left(\frac{d^k}{d\,t^k}\,\mathsf{E}\,e^{it\xi}\right)\Bigg|_{t=0} = \left(\mathsf{E}\,i^k\,\xi^k\,e^{it\xi}\right)\Bigg|_{t=0} = i^k\,\mathsf{E}\,\xi^k.$$

Существование и непрерывность k-й производной, равно как и законность переноса производной под знак математического ожидания, мы доказывать не будем.

Упражнение. Доказать, что для случайной величины ξ со стандартным нормальным распределением момент чётного порядка $2k$ равен

$$\mathsf{E}\,\xi^{2k} = (2k-1)!! \; = \; (2k-1)\cdot(2k-3)\cdot\ldots\cdot 3\cdot 1.$$

Доказать по определению, что все моменты нечётных порядков стандартного нормального распределения существуют и равны нулю.

Как только появились производные высших порядков, самое время разложить функцию в ряд Тейлора[22].

(Ф6). Пусть существует момент порядка $k \in \mathbb{N}$ случайной величины ξ, т. е. $\mathsf{E}\,|\xi|^k < \infty$. Тогда характеристическая функция $\varphi_\xi(t)$ в окрестности точки $t = 0$ разлагается в ряд Тейлора

$$\begin{aligned}\varphi_\xi(t) \; &= \; \varphi_\xi(0) + \sum_{j=1}^{k}\frac{t^j}{j!}\,\varphi_\xi^{(j)}(0) + o(|t^k|) = 1 + \sum_{j=1}^{k}\frac{i^j t^j}{j!}\,\mathsf{E}\,\xi^j + o(|t^k|) = \\ &= \; 1 + it\,\mathsf{E}\,\xi - \frac{t^2}{2}\,\mathsf{E}\,\xi^2 + \ldots + \frac{i^k t^k}{k!}\,\mathsf{E}\,\xi^k + o(|t^k|).\end{aligned}$$

Ряды Тейлора бывают особенно полезны в теории пределов. Следующее основное свойство характеристических функций потребуется нам для доказательства предельных теорем, и это свойство — последняя теорема, оставленная нами без доказательства.

Теорема 44 (теорема о непрерывном соответствии[23]). *Случайные величины ξ_n слабо сходятся к случайной величине ξ тогда*

[22] Brook Taylor (18.08.1685—29.12.1731, England).

и только тогда, когда для любого t характеристические функции $\varphi_{\xi_n}(t)$ *сходятся к характеристической функции* $\varphi_\xi(t)$.

Сформулированная теорема устанавливает непрерывное соответствие между классами $\langle F_\xi, \Rightarrow\rangle$ функций распределения со слабой сходимостью и $\langle \varphi_\xi, \to\rangle$ характеристических функций со сходимостью в каждой точке. «Непрерывность» этого соответствия — в том, что пределу в одном классе относительно заданной в этом классе сходимости соответствует предел в другом классе относительно сходимости, заданной в этом другом классе.

Осталось воспользоваться теоремой о непрерывном соответствии и доказать ЗБЧ в форме Хинчина и ЦПТ.

§ 3. Доказательство ЗБЧ Хинчина

Пусть $\xi_1, \xi_2, \ldots$ — последовательность независимых в совокупности и одинаково распределённых случайных величин с конечным *первым* моментом $\mathsf{E}\,|\xi_1| < \infty$. Обозначим через a математическое ожидание $\mathsf{E}\,\xi_1$. Требуется доказать, что

$$\frac{S_n}{n} = \frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{\text{p}} a.$$

По свойству 26 (с. 122) сходимость по вероятности *к постоянной* эквивалентна слабой сходимости. Так как a — постоянная, достаточно доказать слабую сходимость $\dfrac{S_n}{n}$ к a. По теореме о непрерывном соответствии, эта сходимость имеет место тогда и только тогда, когда для любого $t \in \mathbb{R}$ сходятся характеристические функции

$$\varphi_{S_n/n}(t) \to \varphi_a(t) = \mathsf{E}\,e^{ita} = e^{ita}.$$

Найдём характеристическую функцию случайной величины $\dfrac{S_n}{n}$. Пользуясь свойствами (Ф3) и (Ф4), получаем

$$\varphi_{S_n/n}(t) = \varphi_{S_n}\left(\frac{t}{n}\right) = \left(\varphi_{\xi_1}\left(\frac{t}{n}\right)\right)^n.$$

Вспомним, что первый момент ξ_1 существует, поэтому свойство (Ф6) позволяет разложить $\varphi_{\xi_1}(t)$ в ряд Тейлора в окрестности нуля:

$$\varphi_{\xi_1}(t) \;=\; 1 + it\,\mathsf{E}\,\xi_1 + o(|t|) = 1 + ita + o(|t|).$$

В точке t/n соответственно

$$\varphi_{\xi_1}\left(\frac{t}{n}\right) \;=\; 1 + \frac{ita}{n} + o\left(\left|\frac{t}{n}\right|\right),$$

[23] Paul Pierre Lévy (15.09.1886—15.12.1971, France).

$$\varphi_{S_n/n}(t) \;=\; \left(\varphi_{\xi_1}\left(\frac{t}{n}\right)\right)^n = \left(1 + \frac{ita}{n} + o\left(\left|\frac{t}{n}\right|\right)\right)^n.$$

При $n \to \infty$, пользуясь «замечательным пределом» $\left(1+\dfrac{x}{n}\right)^n \to e^x$, получаем

$$\varphi_{S_n/n}(t) \;=\; \left(1 + \frac{ita}{n} + o\left(\left|\frac{t}{n}\right|\right)\right)^n \;\to\; e^{ita},$$

что и требовалось доказать. □

§ 4. Доказательство центральной предельной теоремы

Пусть $\xi_1, \xi_2, \ldots$ — последовательность независимых в совокупности и одинаково распределённых случайных величин с конечной и ненулевой дисперсией. Обозначим через a математическое ожидание $\mathsf{E}\,\xi_1$ и через σ^2 — дисперсию $\mathsf{D}\,\xi_1$. Требуется доказать, что

$$\frac{S_n - na}{\sigma\sqrt{n}} = \frac{\xi_1 + \ldots + \xi_n - na}{\sigma\sqrt{n}} \;\Rightarrow\; \mathrm{N}_{0,1}.$$

Введём «стандартизованные» случайные величины $\zeta_i = (\xi_i - a)/\sigma$ — независимые случайные величины с нулевыми математическими ожиданиями и единичными дисперсиями. Пусть Z_n есть их сумма:

$$Z_n = \zeta_1 + \ldots + \zeta_n = \frac{S_n - na}{\sigma}.$$

Требуется доказать, что последовательность $\dfrac{Z_n}{\sqrt{n}}$ слабо сходится к стандартному нормальному распределению. Характеристическая функция величины $\dfrac{Z_n}{\sqrt{n}}$ равна

$$\varphi_{Z_n/\sqrt{n}}(t) = \varphi_{Z_n}\left(\frac{t}{\sqrt{n}}\right) = \left(\varphi_{\zeta_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n. \qquad (26)$$

Характеристическую функцию случайной величины ζ_1 можно разложить в ряд Тейлора, в коэффициентах которого использовать известные моменты $\mathsf{E}\,\zeta_1 = 0$, $\mathsf{E}\,\zeta_1^2 = \mathsf{D}\,\zeta_1 = 1$:

$$\varphi_{\zeta_1}(t) = 1 + it\,\mathsf{E}\,\zeta_1 - \frac{t^2}{2}\,\mathsf{E}\,\zeta_1^2 + o(t^2) = 1 - \frac{t^2}{2} + o(t^2).$$

Подставим это разложение, взятое в точке $t/\sqrt{n}$, в равенство (26) и устремим n к бесконечности. Ещё раз воспользуемся замечательным пределом:

$$\varphi_{Z_n/\sqrt{n}}(t) = \left(\varphi_{\zeta_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n \to e^{-t^2/2} \text{ при } n \to \infty.$$

В пределе получили характеристическую функцию стандартного нормального распределения. По теореме о непрерывном соответствии можно сделать вывод о слабой сходимости

$$\frac{Z_n}{\sqrt{n}} = \frac{S_n - na}{\sigma\sqrt{n}} \Rightarrow \mathrm{N}_{0,1}. \qquad \square$$

Попробуйте теперь сами.

Упражнение. Пусть при любом $\lambda > 0$ случайная величина ξ_λ имеет распределение Пуассона с параметром λ. Используя теорему о непрерывном соответствии, доказать, что случайные величины $(\xi_\lambda - \lambda)/\sqrt{\lambda}$ слабо сходятся к стандартному нормальному распределению при $\lambda \to \infty$.

Характеристическая функция случайной величины ξ_λ вычислена в примере 74.

КОНТРОЛЬНЫЕ ВОПРОСЫ

Вопросы по главам I–IV

1. Что такое пространство элементарных исходов?

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

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

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

5. В урне лежат два шарика: белый и чёрный. Наудачу вытаскивают один, возвращают обратно и снова вытаскивают один. Описать пространство элементарных исходов.

6. В урне два белых и три чёрных шара. Вытаскивают наудачу один шар. Выписать все равновозможные элементарные исходы опыта.

7. В урне два белых шара и один чёрный. Наугад берут два шара. Выписать все равновозможные элементарные исходы опыта.

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: а) произошло хотя бы одно; б) *не* произошло хотя бы одно; в) не произошло ни одно; г) случились все три события A, B, C одновременно; д) событие A произошло, а события B и C не произошли.

11. В коробке девять деталей. Событие $A = \{$*в коробке все детали дефектные*$\}$. Описать событие $\overline{A}$.

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

13. Что дают в объединении событие и противоположное к нему? В пересечении? Чему равно дополнение к объединению событий? К пересечению событий?

14. В каком случае два события несовместны?

15. Будут ли несовместными события «на первой кости выпало чётное

число очков» и «на второй кости выпало нечётное число очков» при бросании двух игральных костей?

16. Что такое попарная несовместность событий?

17. Чему равно пересечение трёх попарно несовместных событий?

18. Что больше: объединение или пересечение событий?

19. Объединение двух событий влечёт их пересечение или наоборот?

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

21. Нарисовать графически, что событие A влечёт событие B.

22. Бросают три игральных кости. Как соотносятся события: «на 1-й и 2-й костях выпали единицы» и «на всех костях выпали единицы»?

23. Как соотносятся события $A = A_1 \cap A_2 \cap A_3$ и $B = A_1 \cap A_2$?

24. Сформулировать определение вероятности на дискретном пространстве элементарных исходов.

25. Задать какую-нибудь вероятность на $\Omega = \mathbb{N}$ как на дискретном пространстве элементарных исходов.

26. Задать какую-нибудь вероятность на $\Omega = \mathbb{Z}$ как на дискретном пространстве элементарных исходов.

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

28. Сформулировать классическое определение вероятности.

29. В урне 22 белых и три чёрных шара. Вытаскивают наудачу один шар. С какой вероятностью он белый?

30. В урне пять шаров. Из урны 100 000 раз вытаскивали наудачу один шар, возвращая его обратно. Белый шар был вынут 40 035 раз. Как вы думаете, сколько белых шаров в урне?

31. Какова вероятность ровно один раз выбросить герб при двух подбрасываниях правильной монеты?

32. Какова вероятность хотя бы один раз выбросить герб при двух подбрасываниях правильной монеты?

33. Бросают два раза игральную кость. Какова вероятность, что оба раза выпадет шесть очков?

34. Каково число элементарных исходов при выборе без возвращения, с учётом порядка?

35. Есть пять различных шариков. Сколькими способами их можно разместить в ряд?

36. Что вычисляет число C_n^k при выборе шаров из урны?

37. Что вычисляет число A_n^k при выборе шаров из урны?

38. В урне пять шаров. Выбирают два шара без возвращения и без учёта порядка. Найти $|\Omega|$.

39. Что такое гипергеометрическое распределение вероятностей?

40. Как вычисляется $\mathsf{P}(A)$ согласно геометрическому определению вероятности?

41. Две точки наудачу и независимо друг от друга бросаются на отрезок. Какова вероятность их координатам совпасть?

42. Привести пример $A \neq \varnothing$ такого, что $\mathsf{P}(A) = 0$.

43. Привести пример $A \neq \Omega$ такого, что $\mathsf{P}(A) = 1$.

44. Равносильны ли свойства: $\mathsf{P}(A \cap B) = 0$ и $A \cap B = \varnothing$? Если «нет», что из чего вытекает?

45. Равносильны ли свойства: $\mathsf{P}(A \cup B) = 1$ и $A \cup B = \Omega$? Если «нет», что из чего вытекает?

46. Определение алгебры подмножеств Ω.

47. Задано пространство $\Omega = \{1, 2, 3, 4\}$. Является ли алгеброй множество $\mathcal{F} = \{\varnothing, \{1, 2, 3, 4\}, \{1, 2\}, \{2, 3, 4\}, \{2\}\}$?

48. Задать какую-нибудь алгебру на множестве $\Omega = \{0, 1, \ldots, 10\}$.

49. Записать 2^{Ω}, если $\Omega = \{$*герб*, *решка*$\}$. Является ли 2^{Ω} алгеброй?

50. Записать 2^{Ω}, если $\Omega = \{\diamondsuit, \clubsuit\}$. Является ли 2^{Ω} алгеброй?

51. Записать 2^{Ω}, если $\Omega = \{a, b, c\}$. Является ли 2^{Ω} алгеброй?

52. Записать 2^{Ω}, если $\Omega = \{1,\ 2,\ 3\}$. Является ли 2^{Ω} алгеброй?

53. Найти $\left|2^{\Omega}\right|$, если $\Omega = \{1,\ 2,\ 3,\ 4,\ 5\}$.

54. Пусть $\mathcal{F}$ — алгебра подмножеств Ω. Докажите, что $A \setminus B \in \mathcal{F}$, если $A,\ B \in \mathcal{F}$.

55. Пусть $\mathcal{F}$ — алгебра подмножеств Ω. Верно ли, что $A \cup B \cup C \in \mathcal{F}$, если $A,\ B,\ C \in \mathcal{F}$?

56. Сформулировать определение сигма-алгебры событий.

57. Что такое событие? Какие подмножества Ω являются событиями, а какие не являются?

58. Пространство элементарных исходов Ω состоит из четырёх точек: $\Omega = \{\diamondsuit, \clubsuit, \heartsuit, \spadesuit\}$. Привести пример σ-алгебры $\mathcal{F}$ событий, состоящей более чем из двух событий.

59. Является ли сигма-алгеброй множество всех подмножеств Ω?

60. Привести пример алгебры, не являющейся σ-алгеброй.

61. Пусть $\mathcal{F}$ — алгебра подмножеств Ω и $A_1,\ A_2,\ \ldots \in \mathcal{F}$. Докажите, что $\bigcup\limits_{i=1}^{n} A_i \in \mathcal{F}$ для любого $n \in \mathbb{N}$. Следует ли отсюда, что $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$?

62. Всякая ли алгебра является σ-алгеброй? Всякая ли σ-алгебра яв-

ляется алгеброй?

63. Может ли σ-алгебра событий состоять из одного события? Из двух? Из трёх? Из четырёх?

64. Сформулировать определение меры.

65. Какова область определения и область значений меры? Может ли мера принимать бесконечные значения?

66. Чему равна площадь всей плоскости $\mathbb{R}^2$? Длина прямой $\mathbb{R}$?

67. Является ли функция μ такая, что $\mu(A) = 0$ для всех A, мерой на $(\Omega, \mathcal{F})$, где $\Omega = \{a, b, c\}$, $\mathcal{F} = 2^\Omega$?

68. Пусть $\Omega = \{0,\, 1,\, 2\}$, $\mathcal{F} = 2^\Omega$, $\mu(B) = 5$, если $1 \in B$, и $\mu(B) = 0$ иначе. Выписать $\mu(B)$ для всех $B \in \mathcal{F}$.

69. Сформулировать определение вероятностной меры.

70. Какова область определения и область возможных значений вероятностной меры?

71. Пусть задано вероятностное пространство $\langle \Omega, \mathcal{F}, \mathsf{P} \rangle$. Для каких множеств $A \subseteq \Omega$ определена вероятность $\mathsf{P}(A)$, а для каких нет?

72. Каких значений не может принимать вероятность?

73. Чему равна вероятность достоверного события? Невозможного?

74. Что такое счётная аддитивность вероятностной меры?

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

76. Чему равна вероятность объединения счётного числа попарно несовместных событий?

77. Что такое вероятностное пространство?

78. Пусть $\Omega = \{a,\, b,\, c\}$. Построить какое-нибудь вероятностное пространство на Ω.

79. Пусть $\Omega = \{a,\, b,\, c\}$, $\mathcal{F} = 2^\Omega$. Задана вероятность P такая, что $\mathsf{P}\{a,\, c\} = 5/8$ и $\mathsf{P}\{b,\, c\} = 7/8$. Найти вероятности элементарных исходов $\mathsf{P}\{a\}$, $\mathsf{P}\{b\}$, $\mathsf{P}\{c\}$.

80. Пусть $\Omega = \mathbb{N}$. Задать какое-нибудь вероятностное пространство.

81. Пусть $\Omega = \mathbb{R}$. Задать какое-нибудь вероятностное пространство.

82. Доказать, исходя из определения вероятностной меры, что $\mathsf{P}(\varnothing) = 0$ и $\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A)$.

83. Как связаны вероятности прямого и противоположного событий?

84. Что такое монотонность вероятности?

85. Пусть событие A влечёт событие B. Что можно сказать про их вероятности?

86. Что больше: $\mathsf{P}(A \cap B)$ или $\mathsf{P}(A)$?

87. Что больше: вероятность объединения или вероятность пересечения двух событий?

88. Чему равна вероятность объединения двух событий? Когда вероятность объединения равна сумме вероятностей?

89. Может ли вероятность объединения двух совместных событий равняться сумме их вероятностей? Привести пример.

90. Пусть событие B влечёт событие A. Всегда ли верно, что $\mathsf{P}(A \backslash B) = \mathsf{P}(A) - \mathsf{P}(B)$? Всегда ли верно, что $\mathsf{P}(B \backslash A) = \mathsf{P}(B) - \mathsf{P}(A)$?

91. Записать формулу включения-исключения.

92. Сформулировать свойство непрерывности меры.

93. Зачем в свойстве непрерывности меры требуется конечность меры множества B_1?

94. Что такое сигма-алгебра, порожденная набором множеств $\mathfrak{A}$?

95. Определение борелевской σ-алгебры $\mathfrak{B}(\mathbb{R})$.

96. Является ли интервал $(1,\,5)$ борелевским множеством?

97. Доказать по определению, что $[0,1)$, $[1,\,2]$, $\{4\}$ являются борелевскими множествами.

98. Является ли множество $(0,\,1)\cup(2,\,3)$ борелевским?

99. На борелевской σ-алгебре в $\mathbb{R}$ задана функция: $\mu(A) = 1$ для любого борелевского множества A. Является ли μ вероятностной мерой?

100. Сформулировать определение меры Лебега в $\mathbb{R}$.

101. Чему равна мера Лебега отрезка $[0,\,1]$? Множества $\{0,\ 1\}$? Множества $\mathbb{Z}$? Луча $(0,\,+\infty)$?

102. Сформулировать определение условной вероятности.

103. Может ли условная вероятность равняться безусловной?

104. Может ли условная вероятность равняться единице, нулю?

105. Чему равна вероятность пересечения двух произвольных событий? Двух независимых событий?

106. Привести теорему умножения для n событий. Когда она верна?

107. Как вычислять $\mathsf{P}(ABC)$, если эта вероятность ненулевая?

108. Что такое полная группа событий? Чему равна сумма вероятностей событий из полной группы?

109. Дважды бросается монета. Образуют ли события «герб при первом броске» и «герб при втором броске» полную группу?

110. Записать формулу полной вероятности.

111. Записать формулу Байеса. При каких условиях она верна?

112. Сформулировать определение независимости двух событий.

113. Из колоды карт выбирают наугад одну. Независимы ли события

«выбрана пика» и «выбран туз»? Независимы ли события «выбрана пика» и «выбрана бубна»?

114. Дважды бросают правильную монету. Независимы ли события «при первом броске выпал герб» и «при втором броске выпала решка»? Независимы ли события «при первом броске выпал герб» и «при первом броске выпала решка»?

115. Могут ли несовместные события быть независимы?

116. Могут ли два независимых события образовать полную группу?

117. Всегда ли событие зависит от самого себя?

118. Зависит ли невозможное событие от самого себя? Достоверное?

119. Привести пример события, не являющегося невозможным или достоверным, но не зависящего от самого себя.

120. Независимы ли события, противоположные к независимым?

121. События A и B независимы. Чему равна $\mathsf{P}(A \cap \overline{B})$?

122. Выразить вероятность объединения двух независимых событий через вероятности этих событий.

123. Дать определение независимости n событий в совокупности.

124. Выписать все условия, при которых события A, B, C, D независимы в совокупности.

125. Следует ли из равенства $\mathsf{P}(A \cap B \cap C) = \mathsf{P}(A)\mathsf{P}(B)\mathsf{P}(C)$ независимость A, B, C в совокупности?

126. Достаточно ли попарной независимости событий для независимости в совокупности?

127. Проводится пять независимых испытаний с вероятностью успеха p в каждом из них. Какова вероятность, что сначала произойдут два успеха, потом три неудачи?

128. Есть симметричная монета. Чему равна вероятность получить при 20-м броске герб, если перед этим 19 раз выпадали решки?

129. Бросают три раза правильную монету. Какова вероятность, что в первый раз выпадет герб, а в остальные два — решки?

130. Записать формулу Бернулли.

131. Какова вероятность получить ровно один успех в n испытаниях схемы Бернулли с вероятностью успеха p?

132. Какова вероятность получить три герба после пяти подбрасываний правильной монеты?

133. Какова вероятность не получить ни одного успеха в пяти испытаниях схемы Бернулли с вероятностью успеха $1/4$?

134. Какова вероятность получить четыре успеха в 10 испытаниях схе-

мы Бернулли?

135. Какова вероятность получить не более четырёх успехов в 10 испытаниях схемы Бернулли?

136. Какова вероятность выбросить шесть очков не менее 75 раз при 200 подбрасываниях правильной игральной кости?

137. Какова вероятность впервые выбросить шесть очков при восьмом подбрасывании игральной кости?

138. Какова вероятность первому успеху в схеме Бернулли случиться в пятом испытании, в 10-м испытании?

Вопросы по главам V–IX

1. Сформулировать определение случайной величины.

2. Привести пример функции, не являющейся случайной величиной.

3. Дать определение распределения случайной величины. Какие виды распределений возможны? Чем они отличаются друг от друга?

4. Сформулировать определения дискретного, абсолютно непрерывного и сингулярного распределений.

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

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

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

8. Бросается один раз правильная монета. Построить две различные случайные величины с одним и тем же распределением $B_{1/2}$.

9. Перечислить основные дискретные распределения. Записать таблицу распределения и функцию распределения каждого. Построить графики всех функций распределения.

10. Перечислить основные абсолютно непрерывные распределения. Записать плотность распределения и функцию распределения каждого и построить их графики.

11. Привести пять примеров распределений таких, что $0 \leqslant \xi \leqslant 3$ п. н.

12. Существует ли плотность у распределения Пуассона? Если «да», какова она?

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

14. Как вычислить $\mathsf{P}(\xi \in [-2, 5])$ для $\xi \Subset \Pi_\lambda$?

15. Как вычислить $\mathsf{P}(\xi \in [2, 4])$, $\mathsf{P}(\xi < 3)$, $\mathsf{P}(\xi > 3)$ для случайной величины с абсолютно непрерывным распределением?

16. На графике плотности распределения $\mathrm{N}_{0,1}$ указать вероятности $\mathsf{P}(0 \leqslant \xi \leqslant 2)$, $\mathsf{P}(\xi < -2)$ и $\mathsf{P}(\xi > 3)$.

17. Может ли плотность распределения равняться нулю при всех значениях аргумента, единице, двойке?

18. Необходимое и достаточное условие того, что f является плотностью распределения.

19. Пусть f и g — плотности распределений. Являются ли плотностями распределения функции $2f$, $f+g$, $\dfrac{f+g}{2}$, $2f+2g$, $f-g$, $\dfrac{2}{3}f+\dfrac{1}{3}g$?

20. Сформулировать определение и перечислить основные свойства функции распределения.

21. Для каждого свойства функций распределения нарисовать график любой функции, не обладающей этим свойством.

22. Сформулировать необходимое и достаточное условие того, что функция F является функцией распределения.

23. Нарисовать график функции распределения случайной величины ξ, если $\mathsf{P}(\xi = 5) = 1$.

24. Может ли такая функция F являться функцией распределения:

а) F — чётная функция;

б) $F(x) = \dfrac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$ для любого x;

в) $F(-99) = 0$, $F(0) = 1/4$, $F(101) = 1$?

25. Верно ли, что $F_\xi(1 - 1/n) \to F_\xi(1)$ при $n \to \infty$? Объяснить.

26. Всегда ли $\mathsf{P}(\xi < 1/n) \to \mathsf{P}(\xi < 0)$ при $n \to \infty$? Если нет, привести соответствующий пример.

27. Всегда ли $\mathsf{P}(\xi < -1/n) \to \mathsf{P}(\xi < 0)$ при $n \to \infty$? Почему?

28. Найти пределы $\lim\limits_{n\to\infty} F_\xi(1 + 1/n)$, $\lim\limits_{n\to\infty} F_\xi(-n)$ и $\lim\limits_{n\to\infty} F_\xi(n)$.

29. Как выглядит график функции распределения дискретного распределения? Как по таблице дискретного распределения нарисовать график функции распределения и наоборот?

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

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

32. Дана функция распределения: $F(x) = 0$ при $x < 0$; $F(x) = x/2$ при $x \in [0,\ 1]$ и $F(x) = 1$ при $x > 1$. Существует ли плотность этого распределения?

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

34. Нарисовать график любой функции распределения $F_\xi(x)$ такой, что $\mathsf{P}(\xi = 1) = 1/2$.

35. Как найти плотность распределения по функции распределения?

36. Закончить высказывание: $\mathsf{P}(\xi < 2) = \mathsf{P}(\xi \leqslant 2)$ тогда и только тогда, когда ...

37. Закончить высказывание: $\mathsf{P}(\xi > 0) = 1 - F_\xi(0)$ тогда и только тогда, когда ...

38. Как вычислять вероятность $\mathsf{P}(x_1 < \xi < x_2)$, если $\xi \in \mathrm{N}_{a,\sigma^2}$?

39. На графике функции распределения показательного распределения указать вероятность $\mathsf{P}(1 \leqslant \xi \leqslant 2)$. Ту же вероятность указать на графике плотности этого распределения.

40. Чему равна $\mathsf{P}(\xi < 0)$ для $\xi \in \mathrm{N}_{0,1}$? Что можно сказать про x, если $\Phi_{0,1}(x) < 1/2$? Чему равна $\mathsf{P}(\xi < a)$ для $\xi \in \mathrm{N}_{a,\sigma^2}$?

41. Найти по таблице $\mathsf{P}(\xi < -3)$, $\mathsf{P}(\xi < -1{,}96)$, $\mathsf{P}(\xi < -1{,}6)$, $\mathsf{P}(\xi < 1{,}6)$, $\mathsf{P}(\xi < 1{,}96)$ и $\mathsf{P}(\xi < 3)$ для $\xi \in \mathrm{N}_{0,1}$.

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

43. Как по плотности распределения величины ξ найти плотности распределения величин $-\xi$, 2ξ, $\xi + 2$?

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

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

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

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

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

49. Каким преобразованием можно случайную величину $\xi \in \mathrm{N}_{5,9}$ превратить в $\eta \in \mathrm{N}_{0,1}$ и наоборот?

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

51. Что такое функция распределения случайного вектора?

52. Как по функции распределения вектора находят функции распределения его координат?

53. Что такое таблица совместного распределения?

54. Как по таблице совместного распределения двух случайных величин находят их частные распределения?

55. Какими свойствами обладает плотность распределения случайного вектора?

56. Как по плотности совместного распределения двух случайных величин находят их частные плотности?

57. Можно ли найти совместное распределение по частным распределениям?

58. Привести пример того, что при одних и тех же частных распределениях возможны разные совместные.

59. Что такое многомерное нормальное распределение?

60. Сформулировать определение независимости в совокупности n случайных величин.

61. Как из независимости в совокупности n случайных величин вытекает их попарная независимость?

62. Для каких-то 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 случайных величин ξ и η?

63. Дать определение зависимости случайных величин ξ и η.

64. Верно ли, что если $\mathsf{P}(\xi < 0,\ \eta < 0) = \mathsf{P}(\xi < 0) \cdot \mathsf{P}(\eta < 0)$, то ξ и η независимы?

65. Верно ли равенство $\mathsf{P}(\xi \in \mathbb{R},\ \eta \in \mathbb{R}) = \mathsf{P}(\xi \in \mathbb{R}) \times \mathsf{P}(\eta \in \mathbb{R})$? Можно ли отсюда сделать вывод, что ξ и η независимы? Почему?

66. Привести пример зависимых случайных величин ξ и η таких, что для любого x верно равенство $\mathsf{P}(\xi < x,\ \eta < x) = \mathsf{P}(\xi < x) \cdot \mathsf{P}(\eta < x)$.

67. Дать определение независимости двух случайных величин с дискретными распределениями.

68. Дать определение независимости двух случайных величин с абсолютно непрерывными распределениями.

69. Случайные величины $\xi_1, \ldots, \xi_n$ независимы в совокупности и имеют стандартное нормальное распределение. Выписать плотность совместного распределения величин $\xi_1, \ldots, \xi_n$.

70. Пусть $\xi \Subset \mathrm{B}_{1/2}$, $\eta = \xi$. Проверить, зависимы ли ξ и η.

71. Пусть $\xi \Subset \Pi_1$, $\eta = \xi$. Проверить, зависимы ли ξ и η.

72. В каком случае случайная величина ξ не зависит от себя самой?

73. Как вычислить плотность распределения суммы двух независимых случайных величин, зная плотность распределения каждой?

74. Сформулировать теорему об устойчивости распределения Пуассо-

на по суммированию.

75. Привести пример случайных величин $\xi \in \Pi_\lambda$, $\eta \in \Pi_\mu$ таких, что распределение $\xi + \eta$ не является пуассоновским.

76. Сформулировать теорему об устойчивости биномиального распределения по суммированию.

77. Привести пример случайных величин $\xi \in \mathrm{B}_{n,p}$, $\eta \in \mathrm{B}_{m,p}$ таких, что распределение $\xi + \eta$ не является биномиальным.

78. Привести пример, когда сумма двух одинаково распределённых случайных величин с распределением B_p имеет распределение, отличное от $B_{2,p}$.

79. Сформулировать теорему об устойчивости нормального распределения по суммированию.

80. Привести пример случайных величин $\xi \in \mathrm{N}_{0,1}$, $\eta \in \mathrm{N}_{0,1}$ таких, что $\xi + \eta \notin \mathrm{N}_{0,2}$.

81. Привести пример случайных величин $\xi \in \mathrm{N}_{0,1}$, $\eta \in \mathrm{N}_{0,1}$ таких, что распределение $\xi + \eta$ не является нормальным.

82. Пусть $\xi \in \mathrm{N}_{1,9}$ и $\eta \in \mathrm{N}_{1,1}$ — независимые случайные величины. Какое распределение имеет $\xi - \eta$?

83. Сформулировать теорему об устойчивости гамма-распределения относительно суммирования.

84. Имеет ли сумма независимых и равномерно распределённых слагаемых равномерное распределение?

85. Дать определение математического ожидания для дискретного распределения. Когда существует математическое ожидание случайной величины с дискретным распределением?

86. Дать определение математического ожидания для абсолютно непрерывного распределения. Когда существует математическое ожидание случайной величины с абсолютно непрерывным распределением?

87. Одинаковы ли математические ожидания у двух разных случайных величин с одним и тем же распределением?

88. Какой физический смысл имеет математическое ожидание?

89. Всегда ли математическое ожидание существует?

90. Привести пример распределения, математическое ожидание которого не существует.

91. Привести пример распределения случайной величины с математическим ожиданием -3.

92. Перечислить математические ожидания и дисперсии всех основных распределений.

93. Пользуясь свойствами математического ожидания, вычислить $\mathsf{E}\,(3\xi)$ и $\mathsf{E}\,(\xi+1)$ для $\xi \in \mathrm{N}_{a,\sigma^2}$.

94. Всегда ли математическое ожидание суммы равно сумме математических ожиданий?

95. Всегда ли математическое ожидание произведения равно произведению математических ожиданий?

96. Как вычислять второй момент для показательного распределения, четвёртый?

97. Записать формулу для вычисления $\mathsf{E}(\xi^2 e^{\xi})$, если $\xi \in B_{n,p}$.

98. Записать формулу для вычисления $\mathsf{E}(2^{\xi}\cos\xi)$, если $\xi \in \Pi_{\lambda}$.

99. Записать формулу для вычисления $\mathsf{E}\,\xi e^{-\xi}$ для $\xi \in \mathrm{E}_{\alpha}$.

100. Записать формулу для вычисления $\mathsf{E}\sqrt{\xi}$ для $\xi \in \mathrm{E}_{\alpha}$.

101. Известно, что $\mathsf{P}(\xi \in (-5,\,5)) = 1$. Что можно сказать про $\mathsf{E}\,\xi$?

102. Когда возможно равенство $\mathsf{E}\,|\xi| = 0$? Почему?

103. Сформулировать неравенство Йенсена.

104. Сравнить $\mathsf{E}\,(e^{\xi})$ и $e^{\mathsf{E}\,\xi}$, $\mathsf{E}\,\ln\xi$ и $\ln(\mathsf{E}\,\xi)$.

105. Дать определение и привести основные свойства дисперсии.

106. Как изменится дисперсия при изменении случайной величины вдвое?

107. Можно ли привести пример распределения с дисперсией -1?

108. Что можно сказать про случайную величину ξ, если $\mathsf{D}\,\xi = 0$?

109. Всегда ли дисперсия суммы равна сумме дисперсий?

110. Пусть $\xi \in \Pi_{\lambda}$. Чему равна дисперсия $\mathsf{D}(2-3\xi)$?

111. Найти $\mathsf{D}S_n$, где $S_n = \xi_1 + \ldots + \xi_n$ — сумма независимых и одинаково распределённых случайных величин с конечной дисперсией σ^2.

112. Пусть $\xi_1, \xi_2 \in \mathrm{E}_{\alpha}$ независимы. Вычислить $\mathsf{D}(\xi_1 - \xi_2)$.

113. Сравнить $\mathsf{E}\xi^2$ и $(\mathsf{E}\xi)^2$. Когда эти величины сопадают?

114. Чему равна $\mathsf{D}(4\xi - 3\eta)$ для произвольных случайных величин ξ и η с конечными вторыми моментами?

115. Пусть $\xi_1, \xi_2, \xi_3 \in \mathrm{N}_{0,1}$ независимы. Сравнить $\mathsf{D}(\xi_1 + \xi_2 + \xi_3)$ и $\mathsf{D}(3\xi_1)$.

116. Записать определение и свойства коэффициента корреляции.

117. Если $\xi = 2\eta$, чему равен их коэффициент корреляции?

118. Что можно сказать про случайные величины, если их коэффициент корреляции равен -1?

119. Если обе случайные величины увеличить вдвое, как изменится их коэффициент корреляции?

120. Если обе случайные величины увеличить на два, как изменится

их коэффициент корреляции?

121. Чему равен коэффициент корреляции независимых случайных величин? Может ли коэффициент корреляции двух зависимых случайных величин равняться нулю?

122. Для того чтобы $\mathsf{D}(\xi+\eta)=\mathsf{D}\xi+\mathsf{D}\eta$, необходимо и достаточно, чтобы ξ и η были независимы или некоррелированы? Обосновать.

Вопросы по главам X–XII

1. Сформулировать определение сходимости «почти наверное».
2. Сформулировать определение сходимости по вероятности.
3. Дать определение предела числовой последовательности.
4. Привести пример сходящейся по вероятности последовательности.
5. Сходится ли по вероятности сходящаяся числовая последовательность? Сходится ли она почти наверное?
6. Известно, что $\mathsf{P}(|\xi_n-\xi|\geqslant 0,001)\to 0$ при $n\to\infty$. Верно ли, что тогда $\xi_n \xrightarrow{p} \xi$?
7. Куда сходится по вероятности последовательность $\xi_n \in \mathrm{B}_{1-\frac{1}{n}}$?
8. Какими свойствами обладает сходимость по вероятности?
9. Какая из сходимостей сильнее: почти наверное или по вероятности?
10. Пусть $\mathsf{E}\,|\xi|=1$. Оценить с помощью неравенства Маркова вероятность $\mathsf{P}(|\xi|>3)$.
11. Сформулировать неравенства Маркова и Чебышёва.
12. Какие вероятности позволяет оценивать неравенство Чебышёва?
13. 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?
14. Чем можно оценить вероятность случайной величине отличаться от своего математического ожидания более чем на три корня из дисперсии, более чем на четыре, на пять?
15. Что означают слова «последовательность удовлетворяет ЗБЧ»?
16. Каков смысл закона больших чисел?
17. Куда сходятся средние арифметические независимых и одинаково распределённых случайных величин с конечной дисперсией?
18. Как себя ведёт отношение числа успехов в схеме Бернулли к числу испытаний с ростом последнего?
19. 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?
20. Сформулировать ЗБЧ в формах Чебышёва, Маркова, Хинчина.

21. Пусть $\xi_1, \xi_2, \ldots$ — независимые и одинаково распределённые случайные величины. При каком условии существует и чему равен предел последовательности $\dfrac{\xi_1^3 + \ldots + \xi_n^3}{n}$ при $n \to \infty$?

22. Сформулировать усиленный ЗБЧ Колмогорова.

23. Может ли последовательность зависимых случайных величин удовлетворять ЗБЧ? Привести пример.

24. Может ли последовательность разнораспределённых случайных величин удовлетворять ЗБЧ? Привести пример.

25. Определение слабой сходимости.

26. Расшифровать по определению запись $\xi_n \Rightarrow 0$.

27. Доказать по определению, что последовательность $\xi_n = -\dfrac{1}{n}$ слабо сходится к нулю.

28. Доказать по определению, что последовательность $\xi_n = \dfrac{1}{n}$ слабо сходится к нулю.

29. Как связаны слабая сходимость и сходимость по вероятности?

30. Сформулировать теорему о двойном пределе.

31. Перечислить свойства слабой сходимости.

32. Сформулировать ЦПТ.

33. К какому распределению в условиях ЦПТ приближается распределение случайной величины $\dfrac{S_n - \mathsf{E}\, S_n}{\sqrt{\mathsf{D}\, S_n}}$ при $n \to \infty$?

34. Чему равно математическое ожидание и дисперсия случайной величины $\dfrac{S_n - \mathsf{E}\, S_n}{\sqrt{\mathsf{D}\, S_n}}$?

35. В условиях ЦПТ как себя ведут при $n \to \infty$

$$\mathsf{P}\left(\frac{S_n - n\mathsf{E}\xi_1}{\sqrt{n\mathsf{D}\xi_1}} < 1/n\right) \quad \text{и} \quad \mathsf{P}\left(\frac{S_n - n\mathsf{E}\xi_1}{\sqrt{n\mathsf{D}\xi_1}} < n\right)?$$

36. Выполнено ли утверждение ЦПТ для независимых случайных величин $\xi_i \in \Pi_\lambda$?

37. Привести пример задачи, для решения которой необходима ЦПТ.

38. В условиях ЦПТ каков предел при $n \to \infty$ последовательности вероятностей $\mathsf{P}\Big(\dfrac{S_n}{n} < \mathsf{E}\,\xi_1\Big)$?

ПРИЛОЖЕНИЕ

Таблица 1

Основные дискретные распределения

Название, 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 $p \in (0,\, 1)$	$k = 1,\, 2, \ldots$	$p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
Гипергеомет-рическое $n, K, N \in \mathbb{N}$ $0 \leqslant n, K \leqslant N$	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

Название, обозначение, параметры	Плотность распределения	$\mathsf{E}\xi$	$\mathsf{D}\xi$	Асимметрия	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) N_{a,σ^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 распределения стандартного нормального закона

$$\Phi_{0,1}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

x	$\Phi_{0,1}(x)$	x	$\Phi_{0,1}(x)$	x	$\Phi_{0,1}(x)$	x	$\Phi_{0,1}(x)$
-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	$0,0013$	$-2,34$	$0,0096$	$-1,84$	$0,0329$	$-1,51$	$0,0655$
$-2,98$	$0,0014$	$-2,32$	$0,0102$	$-1,83$	$0,0336$	$-1,5$	$0,0668$
$-2,96$	$0,0015$	$-2,3$	$0,0107$	$-1,82$	$0,0344$	$-1,49$	$0,0681$
$-2,94$	$0,0016$	$-2,28$	$0,0113$	$-1,81$	$0,0351$	$-1,48$	$0,0694$
$-2,92$	$0,0018$	$-2,26$	$0,0119$	$-1,8$	$0,0359$	$-1,47$	$0,0708$
$-2,9$	$0,0019$	$-2,24$	$0,0125$	$-1,79$	$0,0367$	$-1,46$	$0,0721$
$-2,88$	$0,0020$	$-2,22$	$0,0132$	$-1,78$	$0,0375$	$-1,45$	$0,0735$
$-2,86$	$0,0021$	$-2,2$	$0,0139$	$-1,77$	$0,0384$	$-1,44$	$0,0749$
$-2,84$	$0,0023$	$-2,18$	$0,0146$	$-1,76$	$0,0392$	$-1,43$	$0,0764$
$-2,82$	$0,0024$	$-2,16$	$0,0154$	$-1,75$	$0,0401$	$-1,42$	$0,0778$
$-2,8$	$0,0026$	$-2,14$	$0,0162$	$-1,74$	$0,0409$	$-1,41$	$0,0793$
$-2,78$	$0,0027$	$-2,12$	$0,0170$	$-1,73$	$0,0418$	$-1,4$	$0,0808$
$-2,76$	$0,0029$	$-2,1$	$0,0179$	$-1,72$	$0,0427$	$-1,39$	$0,0823$
$-2,74$	$0,0031$	$-2,08$	$0,0188$	$-1,71$	$0,0436$	$-1,38$	$0,0838$
$-2,72$	$0,0033$	$-2,06$	$0,0197$	$-1,7$	$0,0446$	$-1,37$	$0,0853$
$-2,7$	$0,0035$	$-2,04$	$0,0207$	$-1,69$	$0,0455$	$-1,36$	$0,0869$
$-2,68$	$0,0037$	$-2,02$	$0,0217$	$-1,68$	$0,0465$	$-1,35$	$0,0885$
$-2,66$	$0,0039$	-2	$0,0228$	$-1,67$	$0,0475$	$-1,34$	$0,0901$
$-2,64$	$0,0041$	$-1,99$	$0,0233$	$-1,66$	$0,0485$	$-1,33$	$0,0918$
$-2,62$	$0,0044$	$-1,98$	$0,0239$	$-1,65$	$0,0495$	$-1,32$	$0,0934$
$-2,6$	$0,0047$	$-1,97$	$0,0244$	$-1,64$	$0,0505$	$-1,31$	$0,0951$
$-2,58$	$0,0049$	$-1,96$	$0,0250$	$-1,63$	$0,0516$	$-1,3$	$0,0968$
$-2,56$	$0,0052$	$-1,95$	$0,0256$	$-1,62$	$0,0526$	$-1,29$	$0,0985$
$-2,54$	$0,0055$	$-1,94$	$0,0262$	$-1,61$	$0,0537$	$-1,28$	$0,1003$
$-2,52$	$0,0059$	$-1,93$	$0,0268$	$-1,6$	$0,0548$	$-1,27$	$0,1020$
$-2,5$	$0,0062$	$-1,92$	$0,0274$	$-1,59$	$0,0559$	$-1,26$	$0,1038$

Окончание табл. 3

x	$\Phi_{0,1}(x)$	x	$\Phi_{0,1}(x)$	x	$\Phi_{0,1}(x)$	x	$\Phi_{0,1}(x)$
$-1,25$	$0,1056$	$-0,93$	$0,1762$	$-0,61$	$0,2709$	$-0,29$	$0,3859$
$-1,24$	$0,1075$	$-0,92$	$0,1788$	$-0,6$	$0,2743$	$-0,28$	$0,3897$
$-1,23$	$0,1093$	$-0,91$	$0,1814$	$-0,59$	$0,2776$	$-0,27$	$0,3936$
$-1,22$	$0,1112$	$-0,9$	$0,1841$	$-0,58$	$0,2810$	$-0,26$	$0,3974$
$-1,21$	$0,1131$	$-0,89$	$0,1867$	$-0,57$	$0,2843$	$-0,25$	$0,4013$
$-1,2$	$0,1151$	$-0,88$	$0,1894$	$-0,56$	$0,2877$	$-0,24$	$0,4052$
$-1,19$	$0,1170$	$-0,87$	$0,1922$	$-0,55$	$0,2912$	$-0,23$	$0,4090$
$-1,18$	$0,1190$	$-0,86$	$0,1949$	$-0,54$	$0,2946$	$-0,22$	$0,4129$
$-1,17$	$0,1210$	$-0,85$	$0,1977$	$-0,53$	$0,2981$	$-0,21$	$0,4168$
$-1,16$	$0,1230$	$-0,84$	$0,2005$	$-0,52$	$0,3015$	$-0,2$	$0,4207$
$-1,15$	$0,1251$	$-0,83$	$0,2033$	$-0,51$	$0,3050$	$-0,19$	$0,4247$
$-1,14$	$0,1271$	$-0,82$	$0,2061$	$-0,5$	$0,3085$	$-0,18$	$0,4286$
$-1,13$	$0,1292$	$-0,81$	$0,2090$	$-0,49$	$0,3121$	$-0,17$	$0,4325$
$-1,12$	$0,1314$	$-0,8$	$0,2119$	$-0,48$	$0,3156$	$-0,16$	$0,4364$
$-1,11$	$0,1335$	$-0,79$	$0,2148$	$-0,47$	$0,3192$	$-0,15$	$0,4404$
$-1,1$	$0,1357$	$-0,78$	$0,2177$	$-0,46$	$0,3228$	$-0,14$	$0,4443$
$-1,09$	$0,1379$	$-0,77$	$0,2206$	$-0,45$	$0,3264$	$-0,13$	$0,4483$
$-1,08$	$0,1401$	$-0,76$	$0,2236$	$-0,44$	$0,3300$	$-0,12$	$0,4522$
$-1,07$	$0,1423$	$-0,75$	$0,2266$	$-0,43$	$0,3336$	$-0,11$	$0,4562$
$-1,06$	$0,1446$	$-0,74$	$0,2296$	$-0,42$	$0,3372$	$-0,1$	$0,4602$
$-1,05$	$0,1469$	$-0,73$	$0,2327$	$-0,41$	$0,3409$	$-0,09$	$0,4641$
$-1,04$	$0,1492$	$-0,72$	$0,2358$	$-0,4$	$0,3446$	$-0,08$	$0,4681$
$-1,03$	$0,1515$	$-0,71$	$0,2389$	$-0,39$	$0,3483$	$-0,07$	$0,4721$
$-1,02$	$0,1539$	$-0,7$	$0,2420$	$-0,38$	$0,3520$	$-0,06$	$0,4761$
$-1,01$	$0,1562$	$-0,69$	$0,2451$	$-0,37$	$0,3557$	$-0,05$	$0,4801$
-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 $\Phi_{0,1}(x)$ находят по такому правилу:

$$\Phi_{0,1}(x) = 1 - \Phi_{0,1}(-x).$$

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

СПИСОК ЛИТЕРАТУРЫ

Гнеденко Б. В. Курс теории вероятностей. М.: Наука, 1988. 406 с.

Боровков А. А. Теория вероятностей. М.: Наука, 1986. 432 с.

Колемаев В. А., Калинина В. Н. Теория вероятностей и математическая статистика. М.: Инфра-М, 1997. 302 с.

Бочаров П. П., Печинкин А. В. Теория вероятностей. Математическая статистика. М.: Гардарика, 1998. 328 с.

Пугачев В. С. Теория вероятностей и математическая статистика. М.: Физматлит, 2002. 496 с.

Коршунов Д. А., Фосс С. Г. Сборник задач и упражнений по теории вероятностей. Новосибирск: НГУ, 2003. 119 с.

Вентцель Е. С., Овчаров Л. А. Задачи и упражнения по теории вероятностей. М.: Высш. шк., 2000. 366 с.

Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике. М.: Высш. образование, 2006. 404 с.

Учебное издание

Чернова Наталья Исааковна

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Учебное пособие

Редактор С. Д. Андреева

Подписано в печать 10.05.2007 г.
Формат $60 \times 84\ ^1/_{16}$. Офсетная печать.
Уч.-изд. л. 10. Усл. печ. л. 9,3. Тираж 200 экз.
Заказ №

Редакционно-издательский центр НГУ.
630090, Новосибирск-90, ул. Пирогова, 2.