Нижегородский государственный университет имени Н. И. Лобачевского
Радиофизический факультет

Методическое пособие

# Теория вероятностей

Нижний Новгород, 2018

# Содержание

# Исторические сведения

Возникновение теории вероятностей как науки относят к средним векам, когда появилась возможность и возникла необходимость изучения математическими методами азартных игр (таких как орлянка, кости, рулетка).

Самые ранние работы учёных в области теории вероятностей относятся к XVII веку. Первоначально её основные понятия не имели строго математического описания. Задачи, из которых позже выросла теория вероятностей представляли набор некоторых эмпирических фактов о свойствах реальных событий, которые формулировались с помощью наглядных описаний.

Исследуя прогнозирование выигрыша при бросании костей в письмах друг другу, Блез Паскаль и Пьер Ферма открыли первые вероятностные закономерности. Решением тех же задач занимался и Христиан Гюйгенс. При этом с перепиской Паскаля и Ферма он знаком не был и методику решения изобрёл самостоятельно.

Его статья, в которой он ввёл основные понятия теории вероятностей (понятие вероятности как величину шанса; математическое ожидание для дискретных случаев в виде цены шанса). В своей статье он использует (не сформулированные ещё в явном виде) теоремы сложения и умножения вероятностей. Статья была опубликована в печатном виде на двадцать лет раньше (1657 г.) издания писем Паскаля и Ферма (1679 г.).

Важный вклад в теорию вероятностей внёс Якоб Бернулли, он дал доказательство закона больших чисел в простейшем случае независимых испытаний. В первой половине XIX века теория вероятностей начинает применяться к анализу ошибок наблюдений; Лаплас и Пуассон доказали первые предельные теоремы.

Во второй половине XIX века основной вклад внесли русские учёные П. Л. Чебышёв, А. А. Марков и А. М. Ляпунов. В это время были доказаны закон больших чисел, центральная предельная теорема, а также разработана теория цепей Маркова.

Современный вид теория вероятностей получила благодаря аксиоматике, предложенной Андреем Николаевичем Колмогоровым. В результате теория вероятностей приобрела строгий математический вид и окончательно стала восприниматься как один из разделов математики. Википедия, Статья "Теория вероятностей".

# Часть I
# События и их вероятности

## 1. Элементы комбинаторики. Схемы шансов

В этом параграфе мы подсчитываем число элементарных событий или, проще говоря, исходов, шансев, которые могут возникать в результате эксперимента.

Например, при подбрасывании монеты могут произойти 2 исхода, при подбрасывании игрального кубика могут произойти 6 исходов, при извлечении карты из колоды в 54 листа могут произойти 53 исхода. Такие подсчёты изучают в разделе математики, называемом комбинаторикой.

Пусть $A$ и $B$ — два непересекающихся конечных множества с числом элементов $m$ и $n$ соответственно. Очевидны следующие две леммы.

**Лемма 1.1.** *(о сумме). Число шансов выбрать один элемент либо из $A$ либо из $B$, т.е. из объединения $A \cup B$, равно $m + n$.*

**Лемма 1.2.** *(о произведении). Число шансов выбрать пару элементов, один из $A$, а другой из $B$, равно $mn$, т.е. числу элементов в декартовом произведении $A \times B$.*

Непосредственным обобщением предыдущей леммы является следующая теорема.

**Теорема 1.3.** *Пусть $A_2, A_2, \ldots, A_k$ — конечные непересекающиеся множества, имеющие $n_1, n_2, \ldots, n_k$ элементов соответственно. Выберем из каждого множества по одному элементу. Тогда общее число способов, которыми можно осуществить такой выбор, равно $n_1 n_2 \ldots n_k$.*

*Доказательство.* Ясно, что число способов такого выбора равно числу точек (элементов) в декартовом произведении $A_1 \times A_2 \times \ldots A_k$, т.е. равно $n_1 \cdot n_2 \ldots n_k$. □

### 1.1. Эксперименты выбора шариков

Рассмотрим ящик, содержащий $n$ одинаковых шариков, на которых написаны числа $1, 2, \ldots, n$. Эксперимент состоит в том, что из ящика, не глядя, по одному вынимают $k$ шариков, где $k \leqslant n$. Обозначим через

$$(n_1, n_2, \ldots, n_k)$$

упорядоченный набор чисел, где $n_1$ — номер 1-го вынутого шарика, $n2$ — номер 2-го шарика,…, $n_k$ — номер $k$-го шарика.

Например, из 5 занумерованных шариков выбрали 3 шарика и получился набор (4, 2, 1).

Сколько имеется различных способов вынуть из ящика $k$ шариков? На этот вопрос нельзя дать однозначный ответ, потому что такой эксперимент определён неоднозначно.

Во-первых, не определено, возвращают ли извлеченный шарик обратно в ящик. Во-вторых, не определено, какие наборы номеров считать различными и какие наборы считать одинаковыми.

Рассмотрим следующие возможные условия проведения эксперимента.

1. *Эксперимент с возвращением.* Каждый извлечённый шарик возвращается в ящик. В этом случае в наборе могут появляться одинаковые номера. Например, при выборе трёх шариков из ящика, содержащего пять шариков с номерами 1, 2, 3, 4 и 5, могут появиться наборы (3, 3, 5), (1, 2, 4) и (4, 2, 1).

2. *Эксперимент без возвращений. Извлечённые шарики в ящик не возвращаются.* В этом случае в наборе не могут встречаться одинаковые номера. В рассмотренном выше примере набор (3,3,5) не может появиться, а наборы (1,2,4) и (4,2,1) могут.

Опишем теперь, какие наборы номеров мы будем считать различными. Существуют ровно две возможности.

1. *Эксперимент с учётом порядка.* Два набора номеров считаются различными, если они отличаются либо составом, либо порядком. В рассмотренном выше примере все наборы (3,3,5), (1,2,4) и (4,2,1) считаются различными.

2. *Эксперимент без учёта порядка.* Два набора номеров считаются различными, если они отличаются только составом.

В рассмотренном выше примере наборы (1,2,4) и (4,2,1) доставляют одно и тот же элементарное событие, а набор (3,3,5) — другое.

Подсчитаем теперь, сколько получится различных исходов для каждого из четырёх экспериментов. Заметим, что в литературе такие эксперименты часто называют схемами выбора или схемами шансов. Схема шансов – это условия (с возвратом или без, какие наборы различны и т.д.), при которых проводится эксперимент.

## 1.2. Схема шансов без возвращения и с учетом порядка

**Теорема 1.4.** *В эксперименте без возвращения и с учётом порядка число способов выбрать $k$ элементов из $n$-элементного множества равно*

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

Число $A_n^k$ называется *числом размещений элементов $k$ на $n$ местах.* Читается: «$A$ из $n$ по $k$».

*Доказательство.* При выборе первого шарика имеется $n$ возможностей При выборе первого шарика имеется n возможностей. При выборе второго шарика остаётся $n-1$ возможностей, и т.д. При выборе последнего $k$-го шарика остаётся $n-k+1$ возможностей. По теор. 1.3 общее число наборов равно $n(n-1)\dots(n-k+1)$, что и требовалось доказать. □

**Следствие 1.5.** Число перестановок из $n$ элементов равно $n!$.

*Доказательство.* Очевидно, что перестановка есть результат выбора по схеме без возвращения и с учётом порядка всех $n$ элементов из $n$, т.е. общее число перестановок равно $A_n^2 = n!$. □

## 1.3. Схема шансов без возвращения и без учёта порядка

**Теорема 1.6.** *В эксперименте без возвращения и без учёта порядка число способов извлечь $k$ из $n$-элементного множества равно*

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

*Число $C_n^k$ называется числом сочетаний $k$ элементов из $n$ элементов. Читается: «$C$ из $n$ по $k$»*

*Доказательство.* По следствию 1.5 из $k$ элементов можно образовать $k!$ упорядоченных наборов. Поэтому количество сочетаний (неупорядоченных наборов) в $k!$ раз меньше, чем число размещений. Поделив $A_n^k$ на $k!$, получим требуемый результат. □

## 1.4. Схема шансов с возвращением и с учётом порядка

**Теорема 1.7.** *В эксперименте с возвращением и с учётом порядка число способов извлечь $k$ элементов из $n$-элементного множества равно $n^k$.*

*Доказательство.* При выборе каждого из $k$ шариков имеется $n$ возможностей. По теореме 1.3 общее число наборов равно $n \cdot n \cdot n \cdots n = n^k$. □

## 1.5. Схема шансов с возвращением и без учёта порядка

**Замечание 1.8.** Рассмотрим для примера ящик с двумя шариками 1 и 2, из которого мы вынимаем последовательно два шарика. Без учёта порядка имеется 3 исхода:

$$\{1,1\}, \{1,2\} = \{2,1\}, \{2,2\}$$

.

**Теорема 1.9.** *В эксперименте с возвращением и без учёта порядка число способов извлечь $k$ элементов из $n$-элементного множества равно $C_{n+k-1}^k$.*

*Доказательство.* Т.к. порядок появления шариков не учитывается, то мы учитываем лишь только то, сколько раз в наборе появится $i$-й шарик для каждого $i = 1, 2, \dots, n$. Обозначим через $k_i$ число появлений $i$-го шарика в наборе. Во-первых, $0 \leqslant k_i \leqslant k$, а во-вторых,

$$k_1 + k_2 + \dots + k_n = k.$$

Поставим каждому исходу в соответствие набор чисел $(k_1 1, k_2, \dots, k_n)$. Легко видеть, что это соответствие является взаимно однозначным. Такое соответствие можно рассматривать как способ нумерации наборов. (Например, исходам из замеч. 1.8 ставятся в соответствие следующие номера: $\{1, 1\} \leftrightarrow (2, 0), \{1, 2\} \leftrightarrow (1, 1)$ и $\{2, 2\} \leftrightarrow (0, 2)$. Рассмотрим теперь другой эксперимент. Пусть теперь имеется $n$ урн с номерами $i = 1, 2, \dots, n$, в которых размещаются $k$ неразличимых шариков. Сколько существует способов разложить шарики по урнам? Нас интересует только количество шариков в $i$-й урне для каждого $i$. Обозначим через $k_i$ число шариков в $i$-й урне. Ясно, что $0 \leqslant ki \leqslant k$, и что числа $k_1$ и в этом эксперименте тоже удовлетворяют уравнению

$$k_1 + k_2 + \dots + k_n = k.$$

Исходы этого эксперимента тоже взаимно однозначно описываются наборами чисел
$(k_1, k_2, \dots, k_n)$. Т.о., исходы в эксперименте с урнами и исходы предыдущего эксперимента с ящиком занумерованы одним и тем же набором чисел, поэтому число исходов в обоих экспериментах одно и то же и равно числу решений этого уравнения. Вычислим это число для эксперимента с урнами. Изобразим расположение шариков в урнах с помощью схематичного рисунка. Вертикальными линиями обозначим перегородки между урнами, а кружками — шарики, находящиеся в них. Например,

$$|\bullet\bullet|\bullet\bullet\bullet||\bullet||\bullet\bullet|\bullet|.$$

На рисунке показаны 9 шариков, рассыпанные по 7 урнам: 1-я и 6-я урны содержат по 2 шарика, 2-я урна содержит 3 шарика, 3-я и 5-я урны — пустые и, наконец, 4-я и 7-я урны содержат по одному шарику.

Меняя местами шарики и стенки, можно получить все возможные расположения шариков в урнах. Другими словами, все расположения можно получить, расставляя $k$ шариков и $n - 1$ стенок на $n - 1 + k$ местах. Число $n - 1 + k$ получается следующим образом. Число стенок у $n$ урн равно $n + 1$, и т.к. две крайние стенки двигать нельзя, то число стенок, которые можно двигать равно $n - 1$. Поэтому шарики могут занимать $k$ мест, а стенки урн —

оставшиеся $n-1$ место. По теореме 1.6 число способов расставить $k$ шариков на $n-1+k$ местах и затем расставить стенки на оставшихся $n-1$ местах равно $C^k_{n+k-1}$. Что и требовалось доказать. $\square$

# 2. События, операции над ними и $\sigma$-алгебры событий

Теория вероятностей, как и любая современная математическая теория, начинается с аксиоматических (неопределяемых) понятий. Такими являются следующие понятия.

1. Понятие: *эксперимент = опыт = испытание.* Считается, что все три слова означают одно и то же, т.е. являются синонимами.

2. Понятие: *произойти = возникнуть = появиться.*

3. Понятие: *элементарное событие = элементарный исход = результат = шанс.*

При этом считается, что в результате опыта происходит одно и только одно элементарное событие.

**Определение 2.1.** Множество всех элементарных событий данного эксперимента называется *пространством элементарных событий*, или часто короче *пространством.*

Будем обозначать его через $\Omega$. Ясно, что пространство элементарных событий не пусто, $\Omega \neq \varnothing$.

Как множество пространство $\Omega$ может быть либо конечным[1], либо счётным[2], либо несчётным множеством[3].

**Определение 2.2.** Конечные и счётные пространства элементарных событий называются *дискретными.*

**Примеры**.

1. Эксперимент: подбрасывание монеты. Элементарные события: $o$ – выпадение орла, $p$ – выпадение решётки. Пространство $\Omega = \{o, p\}$ является конечным множеством.

[1]Конечное множество – множество, количество элементов которого конечно, то есть, существует неотрицательное целое число n, равное количеству элементов этого множества. В противном случае множество называется бесконечным.

[2]Счётное множество есть бесконечное множество, элементы которого можно пронумеровать натуральными числами. Более формально: множество $\Omega$ является счётным, если существует биекция $\Omega \leftrightarrow N$, где $N$ обозначает множество всех натуральных чисел.

[3]Множество, не являющееся конечным или счетным, называется несчетным.

2. Эксперимент: одновременное подбрасывание двух монет одного достоинства. Пространство элементарных событий $\Omega = \{(o, o), (o, p), (p, p)\}$ является конечным множеством.

3. Эксперимент: подбрасывание одной монеты до выпадения первого орла. Пространство элементарных событий

$$\Omega = \{o, po, ppo, pppo, ppppo, pppppo, \ldots\}$$

– счётное множество.

4. Эксперимент: на стол $\Omega = I^2 = I \times I$ садится мыльный пузырь и лопается, оставляя под собой точку $(x, y)$. Элементарное событие: появление на плоскости точки$(x, y)$. Пространство элементарных событий $\Omega$ – несчётное множество.

**Определение 2.3.** Если пространство $\Omega$ содержит только одно элементарное событие, то эксперимент называется с детерминированным (определённым) исходом; в противном случае эксперимент называется со случайным исходом.

**Определение 2.4.** Любое подмножество $A \subset \Omega$ пространства элементарных событий называется случайным событием или просто событием. Считается, что событие $A$ произошло, если произошло любое элементарное событие $\Omega$, содержащееся в $A$, см. рис. 1.

**Определение 2.5.** 1) Событие $\Omega$ называется *достоверным.*
2) Событие $\varnothing \subset \Omega$ называется *невозможным.*

Каждое элементарное событие $\omega \in \Omega$ можно рассматривать как одноэлементное подмножество достоверного события $\Omega$, т.е. $\{\omega\} \subset \Omega$. Для изображения событий можно использовать диаграммы Венна[4].

Пусть $A$ и $B$ – события, они показаны на рис. 1.

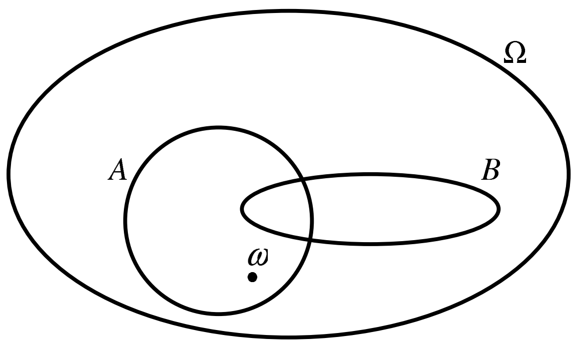


Рис. 1: События в достоверном событии $\Omega$

[4]Джон Венн (John Venn, 1834 – 1923), английский логик.

**Определение 2.6.** 1) Событие $A \cup B$ называется *объединением* событий и состоит в том, что произошло хотя бы одно из событий $A$ или B.

2) Событие $A \cap B$ называется *пересечением* событий $A$ и $B$ и состоит в том, что произошли оба события $A$ и B.

3) Событие $A \setminus B$ называется *разностью* событий $A$ и $B$ и состоит в том, что событие $A$ произошло, а $B$ – нет.

4) Событие $\overline{A} = \Omega \setminus A$ называется *противоположным событию* $A$ и состоит в том, что событие $A$ не произошло. Ясно, что $A = A$ (событие, противоположное к противоположному, является исходным событием). Т.к. $\overline{\varnothing} = \Omega \setminus \varnothing$ и $\overline{\Omega} = \Omega \setminus \Omega = \varnothing$, то невозможное событие $\varnothing$ и достоверное событие $\Omega$ являются взаимно противоположными (друг другу).

5) Говорят, что событие $A$ *влечёт* событие $B$, и пишут $A \subset B$, если при наступлении события $A$ происходит и событие $B$.

**Определение 2.7.** 1) События $A$ и $B$ называются *несовместными*, если $A \cup B$ является невозможным событием, т.е. $A \cup B = \varnothing$.

2) События $A_1, \cdots, A_n$ называются *попарно несовместными*, если для любых $1 \leq i < j \leq n$ события $A_i$ и $A_j$ несовместны.

**Определение 2.8.** Рассмотрим множество $\mathfrak{A}$, элементами которого являются события пространства $\Omega$ (не обязательно все!). Множество $A$ называется алгеброй событий, если достоверное событие $\Omega$ и любые события $A, B \in A$ удовлетворяют аксиомам:

Акс. $\mathcal{A}1$. $\quad \Omega \in \mathfrak{A}$.
Акс. $\mathcal{A}2$. $\quad A \cup B \in \mathfrak{A}$.
Акс. $\mathcal{A}3$. $\quad A \cap B \in \mathfrak{A}$.
Акс. $\mathcal{A}4$. $\quad A \setminus B \in \mathfrak{A}$.

Из аксиомы $\mathcal{A}1$ следует, что алгебра событий не может быть пустой; она всегда содержит достоверное событие $\Omega$. А т.к. $\Omega \setminus \Omega = \varnothing$, то из аксиомы $\mathcal{A}4$ следует, что $\varnothing \in \mathfrak{A}$, т.е. любая алгебра событий $\mathfrak{A}$ содержит вместе с достоверным и невозможное событие.

### Примеры

1) Для любого пространства элементарных событий $\Omega$ набор из двух множеств $\mathfrak{A}_0 = \{\varnothing, \Omega\}$ удовлетворяет аксиомам $\mathcal{A}1$–$\mathcal{A}4$, поэтому $\mathfrak{A}_0 = \{\varnothing, \Omega\}$ является алгеброй событий. Алгебра событий $A_0$ называется тривиальной. Это самая маленькая алгебра событий.

2) В этом примере эксперимент – подбрасывание игральной кости. Пространство элементарных событий есть $\Omega = \{1, 2, 3, 4, 5, 6\}$. Пусть событие $A = \{1, 3, 5\}$ – выпадение нечётного числа очков, а событие $B = \{2, 4, 6\}$ – выпадение чётного числа очков. Множество событий $\mathfrak{A}_1 = \{\varnothing, \Omega, A, B\}$ удовлетворяет аксиомам $\mathcal{A}1$–$\mathcal{A}4$, поэтому является алгеброй событий.

3) Для любого пространства элементарных событий $\Omega$ множество $\mathfrak{B}(\Omega)$ = множество всех подмножеств $\Omega$ удовлетворяет аксиомам $\mathcal{A}1$–$\mathcal{A}4$, поэтому

является алгеброй событий. Эта алгебра событий является самой большой на $\Omega$. Для $\Omega = \{1, 2, 3, 4, 5, 6\}$ эта алгебра событий содержит 64 события.

4) Задайте ещё какую-нибудь алгебру событий на $\Omega = \{1, 2, 3, 4, 5, 6\}$. Сколько различных алгебр событий можно задать на этом пространстве элементарных событий?

**Замечание 2.9.** Из аксиом $\mathcal{A}1$–$\mathcal{A}4$ следует, что если к конечному набору событий из любой алгебры событий применить операции объединения, пересечения и вычитания конечное число раз, то полученное событие тоже содержится в этой алгебре.

**Замечание 2.10.** Если пространство элементарных событий конечно, $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, то любая его алгебра событий тоже конечна. Это следует из того, что множество $\mathfrak{B}(\Omega)$ всех возможных событий, содержащихся в $\Omega$, тоже конечно и содержит $2^n$ событий.

**Определение 2.11.** Алгебра событий $\mathfrak{A}$ называется $\sigma$-алгеброй, если для любого счётного набора событий $A_1, A_2, \dots \in \mathfrak{A}$ выполнена пятая аксиома:

Акс. $\mathcal{A}5$. $\quad \bigcup\limits_{n=1}^{\infty} A_n \in \mathfrak{A}.$

**Замечание 2.12.** Из аксиом $\mathcal{A}1$–$\mathcal{A}4$ не следует, что объединение несчётного количества событий является событием из $\sigma$-алгебры. Рассмотрение несчётных объединений событий приводит к построению т.н. неизмеримых событий, вероятность наступления которых не существует. Первый пример такого неизмеримого события построил Витали[5]. При построении неизмеримого множества Витали используется аксиома теории множеств – аксиома выбора.

**Аксиома выбора.** *Для любого произвольного набора непустых непересекающихся множеств можно составить множество, выбрав в него по одному элементу из каждого множества этого набора.*

**Теорема 2.13** (Теорема Витали – построение неизмеримого множества)**.** *Существуют множества, длина которых не может быть выражена никаким числом.*

*Доказательство.* Для доказательства нам понадобятся лишь следующие очевидные свойства длины:

- длина дуги остается неизменной при повороте окружности вокруг центра;
- длина дуги, которая представляет собой объединение счетного количества попарно непересекающихся дуг, равна сумме длин этих дуг.

[5] Джузеппе Витали (Giuseppe Vitali, 26.08.1875 – 29.02.1932, Italy) – итальянский математик.

Рассмотрим стандартную (единичного радиуса) окружность $S^1$. Она эквивалентна отрезку $[0, 2\pi)$, т.е. её длина равна $2\pi$. На этой окружности центральный угол в радианах равен длине дуги на которую он опирается.

Для любого рационального числа $\frac{p}{q}$ , где $q \neq 0$, рассмотрим дугу длины $\frac{2\pi p}{q}$.

Если отложить её на окружности $S^1$ последовательно $q$ раз, то полученная дуга замкнётся, т.е. начало 1-ой дуги совпадёт с концом $q$-ой.

Для любого иррационального числа $\alpha$ рассмотрим дугу длины $2\pi\alpha$. Ясно, что если отложить эту дугу $n$ раз на окружности $S^1$, то начало и конец дуги $2\pi\alpha n$ не совпадут ни при каком целом n. Это означает, что множество

$$A_0 = \{\ldots, -2\pi\alpha n, \ldots, -4\pi\alpha, -2\pi\alpha, 0, 2\pi\alpha, 4\pi\alpha, \ldots, 2\pi\alpha n, \ldots\}$$

является счётным. Заметим, что для любого целого $n$ поворот на угол $2\pi\alpha n$ переводит множество $A_0$ в себя. Если взять произвольную точку (угол) $\varphi \in [0, 2\pi)$ на окружности, такую чтобы $\varphi \notin A_0$ , то множество

$$A_\varphi = \{\ldots, \varphi-2\pi\alpha n, \ldots, \varphi-4\pi\alpha, \varphi-2\pi\alpha, \varphi, \varphi+2\pi\alpha, \varphi+4\pi\alpha, \ldots, \varphi+2\pi\alpha n, \ldots\}$$

тоже является счётным, и для любого целого $n$ поворот множества $A_\varphi$ на угол $2\pi n\alpha$ тоже переводит его в себя.

Заметим, что т.к. $\varphi \notin A_0$ , то множества $A_0$ и $A_\varphi$ не пересекаются: если бы они имели общую точку, то при $n = \pm1, \pm2, \ldots$ все остальные их точки получились бы прибавлением дуг $2\pi\alpha n$, и тогда бы множества $A_0$ и $A_\varphi$ совпадали.

Заметим, что для любого целого $n$ поворот на угол $2\pi\alpha n$ переводит множество $A_\varphi$ в себя.

Если теперь выбрать угол $\psi \in [0, 2\pi)$, такой чтобы $\psi \notin A_0$ и $\psi \notin A_\varphi$ , то множество

$$A_\psi = \{\ldots, \psi-2\pi\alpha n, \ldots, \psi-4\pi\alpha, \psi-2\pi\alpha, \psi, \psi+2\pi\alpha, \psi+4\pi\alpha, \ldots, \psi+2\pi\alpha n, \ldots\}$$

тоже является счётным, и для любого целого $n$ поворот множества $A_\psi$ на угол $2\pi\alpha n$ тоже переводит его в себя. Заметим, что т.к. $\psi \notin A_0$ и $\psi \notin A_\varphi$, то множество $A_\psi$ не пересекается ни с $A_0$ и ни с $A_\varphi$.

Заметим, что для любого целого $n$ поворот на угол $2\pi\alpha n$ переводит множество $A_\psi$ в себя.

Процесс построения таких непересекающихся множеств можно продолжить неограниченно [6], пока не будут исчерпаны (выбраны) все точки окружности $S^1$ . В каждом таком множестве содержится счетное число точек, и все точки в одном множестве получаются друг из друга поворотами на угол $2\pi\alpha$.

[6] В математике различают два типа бесконечности: потенциальная и актуальная бесконечности.

Потенциальная бесконечность означает, что процесс построения какого-либо объекта может быть продолжен неограниченно. В нашем случае процесс построения множеств $A_\varphi, A_\xi, A_\psi, \ldots$ может быть продолжен неограниченно (если не принимать во внимание, какое время может быть потрачено на это построение). Другой пример потенциальной бесконечности возникает при построении натурального ряда.

Разные множества не пересекаются. Таким образом, окружность $S^1$ является объединением непересекающихся множеств $A_0, A_\varphi, A_\psi, \ldots$ т.е.

$$S^1 = \bigcup_{\lambda \in \{0,\varphi,\psi,\ldots\}} A_\lambda$$

Объединение счётного набора счётных множеств есть счётное множество. А т.к. окружность $S^1$ состоит из несчетного множества точек, то набор построенных множеств $A_0, A_\varphi, A_\psi, \ldots$ является несчётным. Другими словами множество индексов $\{0, \varphi, \psi, \ldots, \}$ – несчётное.

Выберем из каждого множества $A_0, A_\varphi, A_\psi, \ldots$ ровно по одной точке и по аксиоме выбора составим из этих точек множество $B_0$.

Построенное множество $B_0$ называется *множеством Витали.* Для каждого $n = \pm 1, \pm 2, \ldots$ обозначим через $B_n$ множество, которое получается в результате поворота множества $B_0$ на угол $2\pi n\alpha$.

Ясно, что все множества $B_n$ , $n \in \mathbb{Z}$, имеют одну и ту же длину. Т.к. все точки множества $A_\lambda$ , где $\lambda \in \{0, \varphi, \psi, \ldots\}$, можно получить, поворачивая одну из них на углы $2\pi n\alpha$, и т.к. в множестве $B_0$ собраны по одной точке из каждого множества $A_\lambda$ , то объединение всех $B_n$ составляет окружность $S^1$ , т.е.

$$S^1 = \bigcup_{n=-\infty}^{\infty} B_n$$

Оставшуюся часть доказательства проведём методом от противного. Предположим, что длина множества Витали $B_0$ существует, и обозначим её через $\mu(B_0)$. Тогда все множества $B_n$ имеют одну и ту же длину, т.к. получены из $B_0$ поворотом на угол кратный $2\pi\alpha$. Т.к. множества $B_n$ не пересекаются, то сумма их длин равна длине окружности $S^1$ . Поэтому

$$2\pi = \mu(S^1) = \mu\left(\bigcup_{n=-\infty}^{\infty} B_n\right) = \sum_{n=-\infty}^{\infty} \mu(B_n) = \sum_{n=-\infty}^{\infty} \mu(B_0) =$$

$$= \begin{cases} \infty, \text{если } \mu(B_0) > 0 \\ 0, \text{если } \mu(B_0) = 0 \end{cases}$$

---

Если мы выпишем все натуральные числа от 1 до $n$, то ничто не мешает нам написать число $n+1$, и т.д. Потенциальная бесконечность есть бесконечный процесс построения объектов, у которого нет последнего шага. Например, в доказательстве по методу математической индукции.

Под актуальной бесконечностью понимается бесконечная совокупность, построение которой завершено, и все ее элементы наличествуют одновременно. Например, мы будем иметь дело с актуальной бесконечностью, если перечислим весь натуральный ряд полностью и имеем его в законченном виде $N = \{1, 2, \ldots, n, \ldots\}$. Актуальная бесконечность представляет собой весьма сильную идеализацию. В самом деле, она допускает не только возможность построения последующего объекта, если построен предыдущий, но и постулирует, что все возможные объекты уже построены и существуют одновременно. В нашем случае актуальная бесконечность означает, что процесс построения множеств $A_0, A_\varphi, A_\psi, \ldots$ закончен, и мы имеем это множество множеств $\{A_0, A_\varphi, A_\psi, \ldots\}$ налицо.

Полученное противоречие означает, что длины у множества Витали $B_0$ нет вообще никакой (ни нулевой, ни конечной, ни бесконечной). Такие множества и соответствующие им события называются *неизмеримыми*. $\square$

**Замечание 2.14.** 1) В предыдущем примере было доказано существование неизмеримого множества (множества Витали) и предъявлена конструкция его построения. Меняя иррациональное число $\alpha$, можно получить несчётное количество неизмеримых множеств на окружности $S^1 = [0, 2\pi)$.

Если аксиому выбора не признавать, то неизвестно можно ли вообще построить неизмеримые множества. Можно доказать, что аксиома $\mathcal{A}5$ не допускает появление неизмеримых событий.

2) Если $\Omega$ – несчётное множество (отрезок, площадка, объёмное тело), то множество $\mathrm{P}(\Omega)$ всех подмножеств множества $\Omega$ не является $\sigma$-алгеброй, потому что оно содержит неизмеримые события.

**Замечание 2.15.** Все конечные и счётные алгебры событий являются $\sigma$-алгебрами.

**Лемма 2.16.** *Если $A_1, A_2, \ldots$ есть счётный набор событий из $\sigma$-алгебры $A$, то* $\bigcap\limits_{n=1}^{\infty} A_n \in \mathfrak{A}$

*Доказательство.* Пусть $A1, A2, \ldots \in \mathfrak{A}$, Тогда из аксиомы $\mathcal{A}4$ следует, что

$$\overline{A}_1, \overline{A}_2, \ldots \in \mathfrak{A},$$

а из аксиомы $\mathcal{A}5$ следует, что

$$\bigcup_{n=1}^{\infty} \overline{A}_n \in \mathfrak{A}.$$

Тогда по аксиоме 4 дополнение к этому множеству тоже принадлежит $\mathfrak{A}$, т.е.

$$\overline{\bigcup_{n=1}^{\infty} \overline{A}_n} \in \mathfrak{A}$$

по формулам двойственности де Моргана

$$\overline{\bigcup_{n=1}^{\infty} \overline{A}_n} = \bigcap_{n=1}^{\infty} A_n,$$

поэтому

$$\bigcap_{n=1}^{\infty} A_n \in \mathfrak{A}$$

$\square$

**Следствие 2.17.** Лемма 2.16 и аксиома $\mathcal{A}5$ эквивалентны.

# 3. Вероятность и её свойства

**Определение 3.1.** Пусть $\Omega$ — пространство элементарных событий, и $\mathfrak{A}$ — его $\sigma$-алгебра событий. Вероятностью называется функция множества $P$ : $\mathfrak{A} \to \mathbb{R}$, удовлетворяющая следующим аксиомам:

Акс. $\mathcal{P}1$ (*неотрицательность вероятности*). Для любого события $A \in \mathfrak{A}$ выполнено равенство $P(A) \geqslant 0$.

Акс. $\mathcal{P}2$ (*нормированность вероятности*). $P(\Omega) = 1$

Акс. $\mathcal{P}3$ (*счётная аддитивность вероятности как функции множества*).

Для любого счётного набора попарно несовместных событий $A_1, A_2, \dots \in \mathfrak{A}$ имеет место равенство

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

В частности, эта аксиома справедлива и для конечного набора попарно несов местных событий. Акс. $\mathcal{P}1$ (непрерывность вероятности). Для любой убывающей последовательности событий $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$ из $\sigma$-алгебры $\mathfrak{A}$ такой, что $\bigcap_{n=1}^{\infty} A_n = \varnothing$, выполняется равенство

$$\lim_{n\to\infty} P(A_n) = 0$$

**Определение 3.2.** Тройка $(\Omega, (A), P)$ называется *вероятностным пространством.*

Докажем теперь основные свойства вероятности пп. 3.3 – 3.12.

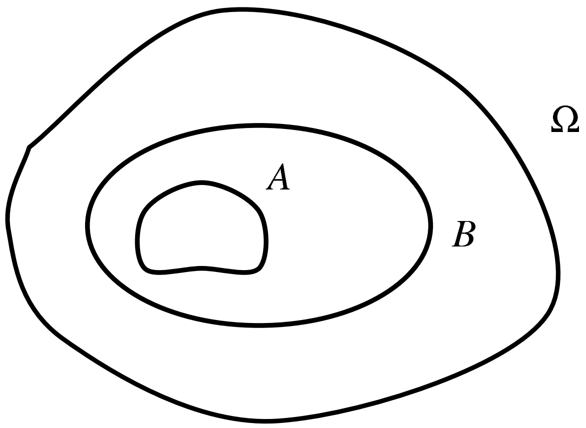


Рис. 2: Если $A \subseteq B$, то $P(B \setminus A = P(B) - P(A))$.

**Лемма 3.3.** *Если событие $A$ влечёт событие $B$, т.е $A \subseteq B$, то $P(B \setminus A) = P(B) - P(A)$ (см. рис. 2)*

*Доказательство.* Т.к. $B = A \cup (B \setminus A)$, и события А и $B \setminus A$ несовместны, то по аксиоме $\mathfrak{P}$ имеем $P(B) = P(A) + P(B \setminus A)$, что и требовалось доказать. □

**Следствие 3.4.** Если $A \subseteq B$, то $P(A) \geqslant P(B)$.

**Лемма 3.5.** $P(\varnothing) = 0$.

*Доказательство.* $P(\varnothing) = P(\Omega \setminus \Omega) = P(\Omega) = P(\Omega) = 1 - 1 = 0$ □

**Лемма 3.6.** *Для любого события $A \in \mathfrak{A}$ выполнено неравенство*

$$0 \leqslant P(A) \leqslant 1$$

*Доказательство.* Т.к. $\varnothing \subseteq A \subseteq \Omega$, то из след. 3.4 следует утверждение леммы. □

**Лемма 3.7.** *Для любого события $A \in (A)$ имеем $P(\bar{A}) = 1 - P(A)$*

*Доказательство.* $P(\bar{A}) = P(\Omega \setminus A) = P(\Omega - P(A) = 1 - P(A)$ □

**Лемма 3.8.** *Для любых событий $A, B \in (A)$ выполнено равенство*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

*Доказательство.* Т.к. $A \cup B = A \cup (B \setminus (A \cup B))$, и события $A$ и $B \setminus (A \cup B)$ несовместны, то применяя сначала Акс. $\mathcal{P}3$, а затем лемму 3.3 получим $P(A \cup B) = P(A) + P(B \setminus (A \cup B)) = P(A) + P(B) - P(A \cup B)$. □

**Следствие 3.9.** Для любых событий $A, B \in \mathfrak{A}$ выполнено равенство

$$P(A \cup B) \geqslant P(A) + P(B).$$

**Следствие 3.10.** Если события А и В несовместны, т.е $A \cap B = \varnothing$, то

$$P(A \cup B) \leqslant P(A) + P(B).$$

*Доказательство.* По лемме 3.8 имеем $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(\varnothing) = P(A) + P(B)$. □

**Следствие 3.11.** $P(A_1 \cup \dots \cup A_n) \leqslant \sum\limits_{i=1}^{n} P(A_i)$.

**Теорема 3.12.**

$$P(A_1 \cup A_2 \cup \dots A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leqslant i < j \leqslant n} P(A_i \cap A_j)$$
$$+ \sum_{1 \leqslant i < j < m \leqslant n} P(A_i \cap A_j \cap A_m) - \dots + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n).$$

*Доказательство.* Применяя метод математической индукции и лемму 3.8 получим требуемый результат. □

# 4. Способы задания и подсчёта вероятности

## 4.1. Экспериментальное нахождение вероятности

В этом пункте описан способ экспериментального определения вероятности наступления так называемых массовых событий.

**Определение 4.1.** Событие A называется *массовым*, если опыт (эксперимент, испытание), при котором событие A может произойти, можно повторить *неограниченное число раз при одних и тех же условиях.*

Методами теории вероятностей изучают (в основном) эксперименты, которые порождают массовые события. Всюду до конца лекций будут рассматриваться только массовые события.

**Пример 4.2.** 1) Опыт: однократное подбрасывание ломаного гроша. События выпадение орла $O$ и выпадение решки $P$ являются массовыми событиями. Заметим, что пространство элементарных событий $\Omega = \{O, P\}$ состоит из конечного числа элементарных исходов.

2) Опыт: подбрасывание ломаного гроша до появления первого орла. События $P, OP, OOP, \ldots, O\ldots OP, \ldots$ являются массовым. Заметим, что в этом примере пространство $\Omega = \{P, OP, OOP, \ldots, O\ldots OP, \ldots\}$ состоит из счётного количества элементарных исходов: положение кольца однозначно определяется положением его центра.

3) Опыт: бросание обручального кольца на клетчатую скатерть (диаметр кольца меньше стороны клетки). Событие $A$: кольцо падает внутрь какой-нибудь клетки, не пересекая её границы. Это событие является массовым. Заметим, что в этом примере пространство состоит из несчётного количества элементарных исходов.

Типичность этих трёх примеров состоит в том, что в теории вероятностей встречаются три типа пространств элементарных событий: конечные, счётные и несчётные.

**Определение 4.3.** Если в результате n испытаний массовое событие $A$ произошло $\mu(A)$ раз, то число $\frac{\mu(A)}{n}$ называется *относительной частотой* появления события $A$.

Для каждого примера 4.2 можно провести $n$ опытов, сосчитать число $\mu(A)$ появлений этого события и подсчитать относительную частоту $\frac{\mu(A)}{n}$.

Массовые события обладают свойством «статистической устойчивости», а именно: *при увеличении числа экспериментов относительная частота* $\frac{\mu(A)}{n}$ *появления события $A$ имеет тенденцию стабилизироваться, стремясь к некоторому числу $P(A)$.*

**Определение 4.4.** Число $P(A) = \lim\limits_{n\to\infty} \frac{\mu(A)}{n}$ называется *статистической вероятностью* события $A$ и находится при больших $n$ по приближённой формуле $P(A) \approx \mu(A)$.

## 4.2. Вероятность на конечном пространстве.

Пусть задано конечное пространство $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, состоящее из $n$ элементарных событий (исходов), и заданы вероятности наступления этих событий $P(\omega_1) = p_1, P(\omega_2) = p_2, \ldots, P(\omega_n) = p_n$ так, что $p_1 + p_2 + \ldots + p_n = 1$. Обозначим через $\omega$ переменную величину, принимающую значения из $\Omega$.

**Определение 4.5.** Такое соответствие записывается в виде таблицы

| $\omega$ | $\omega_1$ | $\omega_2$ | $\ldots$ | $\omega_n$ |
|---|---|---|---|---|
| $P(\xi)$ | $p_1$ | $p_2$ | $\ldots$ | $p_n$ |

которая называется *рядом распределения* случайных исходов или *законом распределения*. Если $n = 2$, то ряд называется *схемой Бернулли*. Если $n \geq 3$, то ряд называется *схемой независимых испытаний с несколькими исходами*.

Ряд распределения задаёт функцию в виде таблицы, которая каждому элементарному исходу ставит в соответствие вероятность его наступления.

Потребуем теперь выполнение аксиомы $\mathcal{P}3$.

**Лемма-определение 4.6** (Формула вероятности на конечном пространстве)**.** Если вероятность, определяемая рядом распределения из опред. 4.5, подчиняется аксиоме $\mathcal{P}3$, то вероятность события $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\} \subset \Omega$ вычисляется по формуле

$$P(A) = p_{i_1} + p_{i_2} + \ldots + p_{i_k},$$

которая называется формулой вероятности на конечном пространстве.

*Доказательство.*

$$P(A) = P(\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}) =$$
$$= P\left(\bigcup_{j=1}^{k} \{\omega_{i_j}\}\right) \overset{\mathcal{P}3}{=} \sum_{j=1}^{k} P(\omega_{i_j}) = \sum_{j=1}^{k} p_{i_j} = p_{i_1} + p_{i_2} + \ldots + p_{i_k}.$$

□

**Пример 4.7.** Простейший ряд распределения имеет эксперимент с детерминированным исходом, $\Omega = \{\omega\}$ (см. опред. 2.3):

| $\xi$ | $\omega$ |
|---|---|
| $P$ | 1 |

Этот тривиальный случай удовлетворяет всем аксиомам и определениям, однако никакого значения в теории вероятностей не имеет. Придётся с этим мириться.

**Определение 4.8.** Пусть в результате опыта могут возникнуть только два события: «успех», который обозначается единицей — 1 и наступает с вероятностью $p$, и «неудача», которая обозначается нулём — 0 и наступает с вероятностью $q = 1 - p$; $\Omega = \{0, 1\}$. Такой опыт называется *схемой Бернулли*[7] и имеет ряд распределения

| $\xi$ | 0 | 1 |
|---|---|---|
| $P$ | $q = 1 - p$ | p |

Схема Бернулли является первым нетривиальным примером вероятности на конечном пространстве.

Схема Бернулли имеет простую интерпретацию. Рассмотрим ломаный грош с вероятностями выпадения орла $p$ и решки $q = 1 - p$. Обозначим появление орла через 1, а его не выпадение через 0, получим схему Бернулли.

**Пример 4.9** (Геометрическая интерпретация схемы независимых испытаний с несколькими исходами.)**.** Рассмотрим произвольный выпуклый многогранника с $n$ гранями, на которых написаны символы $\omega_1, \omega_2, \ldots, \omega_n$.

При этом многогранник должен быть таким, чтобы перпендикуляр, опущенный из его центра тяжести на любую грань, пересекал эту грань в её внутренней точке (чтобы многогранник, падая на эту грань, не перекатывался на другую).

Пусть вероятности выпадения многогранника гранью вниз равны соответственно $p_1, p_2, \ldots, p_n$ , где естественно $p_1 + p_2 + \ldots + p_n = 1$.

Легко видеть, что вероятность $p_i$ пропорциональна телесному углу $\alpha_i$ с вершиной $C$ в центре тяжести многогранника, опирающегося на грань $\omega_i$.

Так как сумма всех телесных углов с вершиной $C$ равна $4\pi$ стерадиан, т.е. $\alpha_1 + \alpha_2 + \ldots + \alpha_n = 4\pi$, то

$$\frac{\alpha_1}{4\pi} + \frac{\alpha_2}{4\pi} + \ldots + \frac{\alpha_n}{4\pi} = 1,$$

поэтому $p_i = 4\pi$.

## 4.3. Классическая вероятность

Классическая вероятность является частным случаем вероятности на конечном пространстве, когда вероятность подчинена принципу равной вероятности. Исторически классическая вероятность применялась в теории азартных игр и появилась раньше вероятности на конечном пространстве.

[7] Якоб Бернулли (Jakob Bernoulli, 1654-1705), швейцарский математик

**Принцип равной вероятности.** *Если на конечном пространстве $\Omega$ вероятность наступления его элементарных исходов $\omega_1, \omega_2, \ldots, \omega_n$ одна и та же,*

$$p_1 = p_2 = \ldots = p_n = p,$$

*то говорят, что вероятность на конечном пространстве удовлетворяет принципу равной вероятности (или равной возможности).*

Выпадение орла и решки при подбрасывании симметричной монеты; выпадение граней при подбрасывании правильного многогранника (тетраэдра, куба, октаэдра, додекаэдра или икосаэдра); появление к.-л. карты из полной колоды карт удовлетворяют принципу равной возможности.

Классическая вероятность характеризуется только числом элементарных исходов $n$ в пространстве $\Omega$. Она имеет ряд распределения

| $\omega$ | $\omega_1$ | $\omega_2$ | $\ldots$ | $\omega_n$ |
|---|---|---|---|---|
| $P(\omega)$ | $p$ | $p$ | $\ldots$ | $p$ |

где $np = 1$, поэтому $p = \frac{1}{n}$.

**Лемма 4.10** (Формула классической вероятности)**.** *Если вероятность $P$ удовлетворяет принципу равной возможности, то вероятность наступления события $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\} \subset \Omega$ определяется по формуле, называемой формулой классической вероятности,*

$$P(A) = \frac{k}{n}$$

*Доказательство.* $P(A) = p_{i_1} + p_{i_2} + \ldots + p_{i_k} = kp = \frac{k}{n}$ □

Если обозначить число $k$ элементов множества $A$ через $\mu(A)$, а число $n$ элементов пространства $\Omega$ через $\mu(\Omega)$, то формулу классической вероятности можно записать в виде

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}$$

Число $\mu(A) = k$ называется числом исходов, *благоприятствующих* наступлению события $A$.

**Замечание 4.11.** 1) Вероятности элементарных исходов в экспериментах с шариками, описанные в теоремах 1.4, 1.6 и 1.7, удовлетворяют принципу равной возможности. Эти вероятности соответственно $1/A_n^k, 1/C_n^k$ и $1/n^k$.

2) Вероятности элементарных исходов в теореме 1.9 (выбор с возвращением и без учёта порядка) не удовлетворяют этому принципу. Например, при $n = 2$ и $k = 2$ вероятности элементарных событий имеют ряд распределения

| $\omega$ | $(1,1)$ | $(1,2)$ | $(2,2)$ |
|---|---|---|---|
| $P(\omega)$ | $1/4$ | $1/2$ | $1/4$ |

3) Следует отметить, что при рассмотрении подобных вопросов ошибались даже такие великие математики, как, например, Д'Аламбер.

Так, однажды у Даламбера спросили, с какой вероятностью монета, брошенная дважды, хотя бы один раз выпадет гербом. Ответ учёного был $\frac{2}{3}$, т.к. он считал, что есть 3 возможных исхода (герб-герб, герб-решка, решка-решка) и среди них 2 благоприятствующих.

Д'Аламбер пренебрегал тем, что эти три возможных исхода не равновозможны.

Правильным ответом является $\frac{3}{4}$ , поскольку из четырёх равновозможных исходов (герб-герб, герб-решка, решкагерб, решка-решка) три благоприятствуют указанному событию.

Точка зрения Д'Аламбера была даже опубликована во Французской энциклопедии в 1754 г. в статье «Герб и решётка» («Croix on pile»).

## 4.4. Вероятность на счётном пространстве

Пусть теперь $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$ – счётное пространство. Пусть $p_1 + p_2 + \ldots + p_n + \ldots = 1$ — сходящийся (к единице) числовой ряд, удовлетворяющий для всех $i \in \mathbb{N}$ условию $0 < p_i < 1$. Последнее условие позволяет трактовать члены этого ряда как вероятности элементарных событий пространства $\Omega$.

**Определение 4.12.** Вероятность элементарных исходов, заданная в виде таблицы

| $\omega$ | $\omega_1$ | $\omega_2$ | $\ldots$ | $\omega_i$ | $\ldots$ |
|---|---|---|---|---|---|
| $P(\omega)$ | $p_1$ | $p_2$ | $\ldots$ | $p_i$ | $\ldots$ |

называется *рядом распределения на счётном пространстве.* Потребуем теперь выполнение аксиомы $\mathcal{P}3$.

**Лемма 4.13** (Формула вероятности на счётном пространстве)**.** *Если события из пространства $\Omega$ подчиняются аксиоме $\mathcal{P}3$, то вероятность события $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}, \ldots\} \subset \Omega$ вычисляется по формуле*

$$P(A) = p_{i_1} + p_{i_2} + \ldots + p_{i_k} + \ldots$$

*Доказательство.*

$$P(A) = P(\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}, \ldots) =$$

$$= P\left(\bigcup_{j=1}^{\infty}\{\omega_{i_j}\}\right) \overset{\mathcal{P}3}{=} \sum_{j=1}^{\infty} P(\omega_{i_j}) = \sum_{j=1}^{\infty} p_{i_j} = p_{i_1} + p_{i_2} + \ldots + p_{i_k} + \ldots$$

□

**Замечание 4.14.** Каждому степенному ряду на той части его интервала сходимости, на которой его члены положительны, можно поставить в соответствие параметрическое семейство распределений со счётным пространством. Рассмотрим, например, экспоненту $e^{\lambda}$ . Её ряд МакЛорена

$$1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^k}{k!} = e^{\lambda}$$

сходится при всех значениях $-\infty < \lambda < \infty$ и имеет положительные члены при $\lambda > 0$. Умножим обе части этого тождества на $e^{-\lambda}$ , получим тождество

$$e^{-\lambda} + \frac{\lambda}{1!}e^{-\lambda} + \frac{\lambda^2}{2!}e^{-\lambda} + \ldots + \frac{\lambda^k}{k!}e^{-\lambda} = 1$$

Составим следующий ряд распределения:

| $k$ | $0$ | $1$ | $2$ | $\ldots$ | $k$ | $\ldots$ |
|---|---|---|---|---|---|---|
| $P_\lambda(k)$ | $e^{-\lambda}$ | $\frac{\lambda}{1!}e^{-\lambda}$ | $\frac{\lambda^2}{2!}e^{-\lambda}$ | $\ldots$ | $\frac{\lambda^k}{k!}e^{-\lambda}$ | $\ldots$ |

**Определение 4.15.** Распределение, реализуемое этим рядом, называется *распределением Пуассона с параметром $\lambda$.*

На рис. 25 показаны функции $P_\lambda(k)$ для для параметра $\lambda = 0,1; 1; 10$, где пунктиром показаны огибающие. В каждом случае сумма длин вертикальных отрезков равна 1.

Распределение Пуассона описывает, например, вероятность $P_\lambda(k) = \frac{\lambda^k e^{-\lambda}}{k!}$ поступления на телефонную станцию $k$ звонков за какой-нибудь фиксированный промежуток времени, где число звонков $k = 0, 1, 2, \ldots$.

.

**Замечание 4.16.** Другим источником построения распределений со счётным пространством являются сходящиеся числовые ряды с положительными членами. Например, известно (Л. Эйлер), что

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{k^2} + \ldots = \frac{\pi^2}{6}.$$

Разделив обе части этого числового тождества на $\frac{\pi^2}{6}$ , можно получить (безымянный) ряд распределения, задаваемый формулой $P(k) = \frac{1}{k^2} \cdot \frac{6}{\pi^2}$ , со счётным пространством $\Omega = \mathbb{N}$. Среди числовых рядов наиболее встречающимися в теории вероятностей являются геометрические прогрессии с положительными членами.

**Определение 4.17.** Если члены ряда $p_1 + p_2 + \ldots + p_n + \ldots = 1$ положительны и являются членами убывающей геометрической прогрессией с первым членом $p_1 = p$ и с знаменателем $q = 1 - p$, то вероятность на счётном пространстве называется *геометрическим распределением.* Геометрическое распределение имеет ряд

| $\tau_1$ | $\omega_1$ | $\omega_2$ | $\ldots$ | $\omega_k$ | $\ldots$ |
|---|---|---|---|---|---|
| $P$ | $q$ | $qp$ | $\ldots$ | $q^{k-1}p$ | $\ldots$ |

**Пример 4.18.** Рассмотрим схему Бернулли (ломаного гроша) с вероятностью «успеха» (выпадения орла – события 1) равной $p$. Она имеет ряд распределения

| $\xi$ | 0 | 1 |
|---|---|---|
| $P$ | $1-p$ | p |

Пусть эксперимент состоит в том, что испытания по схеме Бернулли проводятся неограниченное число раз. С таким экспериментом связаны две случайные величины: $\tau_1$ — номер первого выпавшего орла и $\tau_0$ — число выпавших решек, появившихся до первого орла. Легко подсчитать, что случайные величины $\tau_0$ и $\tau_1$ имеют ряды распределения

| $\tau_0$ | 0 | 1 | $\ldots$ | $k$ | $\ldots$ |
|---|---|---|---|---|---|
| $P$ | $q$ | $qp$ | $\ldots$ | $q^kp$ | $\ldots$ |

и

| $\tau_1$ | 1 | 2 | $\ldots$ | $k$ | $\ldots$ |
|---|---|---|---|---|---|
| $P$ | $q$ | $qp$ | $\ldots$ | $q^{k-1}p$ | $\ldots$ |

В строке вероятностей эти ряды содержат одну и ту же геометрическую прогрессию; они называются соответственно $\tau_0$– и $\tau_1$– геометрическим распределениями. Эти случайные величины связаны очевидной формулой $\tau_1 = \tau_0 + 1$.

Формула вероятности для $\tau_0$ – геометрического распределения определена по формуле

$$P(\tau_0 = k) = q^kp,$$

где $0 < p < 1$ и $q = p - 1$. Формула вероятности для $\tau_1$ – геометрического распределения:

$$P(\tau_1 = k) = q^{k-1}p,$$

где $0 < p < 1$ и $q = p - 1$.

## 4.5. Геометрическая вероятность

Рассмотрим ограниченную измеримую область $\Omega \subset \mathbb{R}^n$ , состоящую из несчётного множества точек. '

Измеримость области означает: на прямой область $\Omega \subset \mathcal{R}^1$ имеет конечную ненулевую длину, на плоскости область $\Omega \subset \mathcal{R}^2$ имеет конечную ненулевую площадь, в пространстве область $\Omega \subset \mathcal{R}^3$ имеет конечный ненулевой объём и т.д.

Пусть на $\Omega$ определена $\sigma$-алгебра $\mathfrak{A}$. Для любого события $A \in \mathfrak{A}$ обозначим через $\mu(A)$ его меру (длину, площадь, объём и т.д. соответственно).

Пусть эксперимент состоит в том, что в область $\Omega$ бросают точку.

**Принцип равномерности**[8]. *Если для любого события $A \in \mathfrak{A}$ его вероятность задаётся по формуле*

$$P(A) = \alpha\mu(A),$$

*где $\alpha$ — постоянное число, независящее от выбора события $A$ (т.е. не зависящее от формы $A$ и его расположения в $\Omega$), то говорят, что вероятность $P(A)$ удовлетворяет принципу равномерности.*

**Лемма 4.19.** *Если вероятность $P(A)$ удовлетворяет принципу равномерности, то*

$$\alpha = \frac{1}{\mu(\Omega)}$$

*Доказательство.* Подставим в формулу $P(A) = \alpha\mu(A)$ достоверное событие $\Omega \in \mathfrak{A}$, получим: $1 = \alpha\mu(\Omega)$. □

**Определение 4.20.** Полученная формула

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}$$

называется формулой *геометрической* вероятности. (Не путать с геометрическим рядом распределения!)

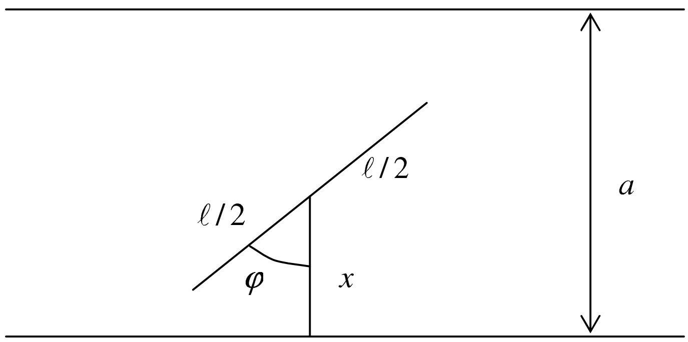


Рис. 3: Задача Бюффона

**Пример 4.21** (Задача Бюффона (1777 г.))**.** На плоскости нарисовано счётное множество параллельных прямых. Расстояние между соседними прямыми равно $a$. На плоскость брошена игла длины $l < a$. Какова вероятность того, что игла пересечёт одну из прямых?

*Решение.* Возможные положения иглы на плоскости полностью определяются двумя координатами: расстоянием $x$ от середины иглы до ближайшей

[8] Вероятности, которые не подчиняются этому принципу, являются главным предметом изучения теории вероятностей. Они будут изучены в гл. 2. «Теория случайных величин».

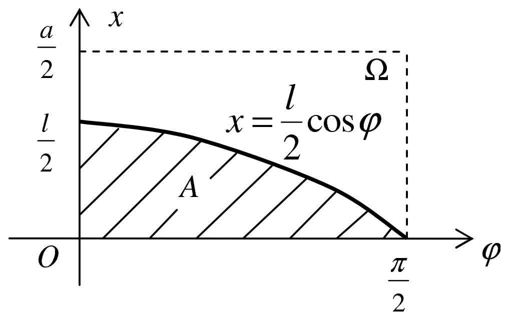


Рис. 4: Пространство элементарных событий omega в задаче Бюффона

прямой и острым углом $\varphi$ между иглой и перпендикуляром к параллельным прямым, см. рис. 3. Ясно, что $x \in [0, \frac{a}{2}]$ и $\varphi \in [0, \frac{\pi}{2}]$. Поэтому множество $\Omega = [0, \frac{\pi}{2}] * [0, \frac{a}{2}]$ есть пространство элементарных исходов этого эксперимента, см. рис. 4, и $\mu(\Omega) = \frac{\pi a}{4}$.

Событие $A =${игла пересечёт одну из прямых} эквивалентно неравенству $A = \{x \leq 2l \cos\varphi\}$, поэтому множество благоприятных исходов располагается в пространстве $\Omega$ под графиком $x = \frac{l}{2}\cos\varphi$. Вычислим

$$\mu(A) = \int_0^{\frac{\pi}{2}} \frac{l}{2}\cos\varphi d\varphi = \frac{l}{2}\sin\varphi\bigg|_0^{\frac{pi}{2}} = \frac{l}{2}$$

Отсюда получаем $P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{2l}{\pi a}$

**Пример 4.22** («Парадокс» Бертрана[9], 1888)**.** В круге наудачу выбирается хорда. Какова вероятность того, что её длина будет больше, чем длина стороны вписанного в круг правильного треугольника?

*Решение.* Обозначим через $A$ событие, состоящее в том, что длина хорды будет больше, чем длина стороны вписанного в круг правильного треугольника. Существует по крайней мере три способа «выбрать наудачу» хорду в круге.

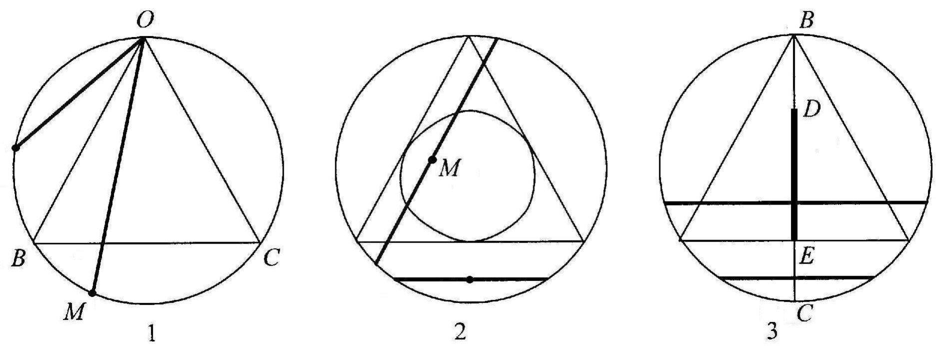


Рис. 5: Парадокс Бертрана

1-й способ. Зафиксируем один конец хорды $O$ на окружности. Положение другого конца хорды $M$ будем считать равномерно распределённым на

[9]Жозеф Луи Франсуа Бертран (Joseph Louis Francois Bertrand, 1822 — 1900), французский математик.

окружности, см. рис. 5.1. Пусть координата конца $M$ хорды есть длина дуги $OM$ окружности, проходимой против часовой стрелки, тогда пространство элементарных событий $\Omega$ есть окружность, т.е. $\mu(\Omega) = 2\pi R$. Благоприятным исходом является положение конца хорды на дуге $BC$. Событие $A$ всех благоприятных исходов есть дуга $BC$, которая составляет третью часть окружности, т.е. $\mu(A) = \frac{2\pi}{3}R$. Поэтому $P(A) = \mu(\Omega) = \frac{1}{3}$.

2-й способ. Для каждой точки $M$ внутри круга (кроме его центра[10]) существует единственная хорда, для которой точка $M$ является её серединой, см. рис. 5.2.

Поэтому, бросая точку $M$ в круг радиуса $R$, можно по ней однозначно восстановить хорду. Середину $M$ хорды будем считать равномерно распределённой в круге. Пространство элементарных событий $\Omega$ есть круг радиуса $R$, его площадь $\mu(\Omega) = \pi R^2$.

Благоприятными событию $A$ являются положения середины $M$ хорды внутри окружности, вписанной в треугольник. Легко подсчитать, что радиус вписанной окружности равен $\frac{R}{2}$ . Поэтому $\mu(A) = \frac{\pi R^2}{4}$ и $P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{1}{4}$.

3-й способ. Можно ограничиться рассмотрением только хорд, которые перпендикулярны какому-нибудь диаметру, например $C$ (остальные положения могут быть получены поворотом), см. рис. 5.3.

Середину хорды будем считать равномерно распределённой на диаметре $BC$. Пространство элементарных событий omega есть диаметр $BC$, которому перпендикулярны хорды, его длина $\mu(\Omega) = 2R$.

Благоприятными событию $A$ являются положения середины хорды на отрезке $DE$, лежащем на диаметре $BC$, так что центр отрезка $DE$ совпадает с центром круга. Можно видеть, что длина отрезка $DE$ есть $\mu(A) = R$. Поэтому $P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{1}{2}$.

Причина разных ответов заключается в том, что условие в круге наудачу выбирается хорда определяет эксперимент не однозначно. В решении мы провели три разных эксперимента по выбору хорды, и поэтому каждом случае был получен правильной ответ.

# 5. Независимые события

**Определение 5.1.** События A и B называются *независимыми*, если выполняется тождество

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

в противном случае они называются *зависимыми*.

**Примеры**.

[10]Т.к. вероятность попадания точки в центр круга равна нулю, то наступление этого события не влияет на вероятность какого-либо события.

1) Рассмотрим колоду карт в 52 листа. Пусть событие $A$ означает вытянуть пику ♠, а $B$ означает вытянуть даму **Д**, Тогда событие $A \cap B$ означает вытянуть пиковую даму **Д**♠. Легко подсчитать вероятности этих событий $P(A \cap B) = 1/52$; $P(A) = 1/4$ и$P(B) = 1/13$. Подставив эти вероятности в равенство опред. 5.1, получим тождество $1/52 = 1/4 \cdot 1/13$; т.е. по опред. 5.1 события $A$ и $B$ независимы (в этой колоде).

2) Рассмотрим полную колоду карт в 54 листа (с двумя шутами). Пусть события $A$ и $B$ — те же как в предыдущем пункте. Легко подсчитать, что в этой колоде $P(A \cap B) = 1/54$, $P(A) = 13/54$ и $P(B) = 2/27$. Таким образом $1/54 \neq 13/54 \cdot 2/27$, и по опред. 5.1 события $A$ и $B$ являются зависимыми в полной колоде. Получилось, что свойство быть или не быть независимыми зависит не от самих событий, а от строения пространства $\Omega$ (52 или 54).

**Лемма 5.2.** *Если события $A$ и $B$ независимы, то пары событий $A$ и $\overline{B}$, $\overline{A}$ и $B$, $\overline{A}$ и $\overline{B}$ тоже являются независимыми.*

*Доказательство.* Докажем, что события $A$ и $\overline{B}$ независимы. Так как $A = (A \cup B) \cup (A \cap \overline{B})$, и события $A \cup B$ и $A \cup \overline{B}$ несовместны, то $P(A) = P(A \cap B) + P(A \cap \overline{B})$. Поэтому $P(A \cap \overline{B}) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\overline{B})$. Независимость пар событий $\overline{A}$ и $B$, $A$ и $\overline{B}$ доказывается аналогично. Доказать самостоятельно. □

**Определение 5.3.** События $A_1, \ldots, A_n$ называются независимыми в совокупности, если для $1 \leqslant i_1 < \cdots < i_k \leqslant n$ выполнено равенство
$P(A_{i1} \cap \cdots \cap A_{ik}) = P(A_{i1}) \cdot \cdots \cdot P(A_{ik})$.

**Замечание 5.4.** Если события $A_1, \ldots, A_n$ независимыми в совокупности, то они попарно независимы. Чтобы это увидеть, достаточно в последнем равенстве положить $k = 2$. Обратное, как показывает следующий пример, не верно.

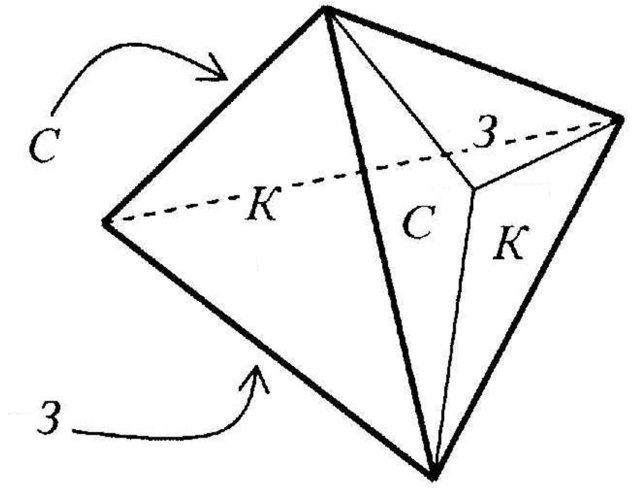


Рис. 6: Тетраэдр Бернштейна

**Пример 5.5** (Тетраэдр Бернштейна[11]). Рассмотрим правильный тетраэдр, три грани которого окрашены в эти синий, зелёный, и красный цвет, см. рис.6, а четвёртая грань разделена на три треугольника, и эти треугольники окрашены в те же три цвета. Обозначим через $B$, $G$ и $R$ события означающие выпадение снизу грани, содержащей соответственно синий (blue), зелёный (green), и красный (red) цвета. Легко видеть, что каждый цвет нарисован на двух из четырёх граней поэтому $P(B) = P(G) = P(R) = 1/2$. Также легко видеть, что появление любой пары из них имеет вероятности $P(B \cap G) = P(G \cap R) = P(B \cap R) = 1/4$. Т.о., для каждой пары из этих событий формула опред. 5.1 выполнена, и следовательно события попарно независимы.

Теперь проверим независимость событий $B, G$ и $R$ в совокупности. Легко видеть, что вероятность $P(B \cap G \cap R)$ выпадения трёхцветной грани равна $1/4$. Это левая часть равенства опред. 5.3. Правая часть равна $P(B) \cdot P(G) \cdot P(R) = 1/8$. Т.е. равенство опред. 5.2 не выполнено, значит события $B, G$ и $R$ зависимы в совокупности.

# 6. Условная вероятность

**Пример**. Игральную кость подбрасывают один раз. Известно, что выпало более трёх очков. Какова при этом вероятность того, что выпало чётное число очков?

*Решение.* $\Omega_1 = 4, 5, 6, A = 4, 6, P(A) = \mu(A)\mu(\Omega) = 3$ .

**Замечание 6.1.** Пусть $\Omega$ – пространство элементарных событий и $B \subset \Omega$ – событие, отличное от невозможного, т.е. $B \neq \varnothing$. Пусть $A \subset \Omega$ – другое событие. Какова вероятность того, что произойдёт событие $A$, при условии, что событие $B$ произошло? Слова *событие $B$ произошло означают*, что новым пространством элементарных событий становится событие $\Omega_1 = B$ и его мера равна $\mu(\Omega_1) = \mu(B)$.

Слова *произойдёт событие $A$, если событие $B$ произошло* означают ту часть события $A$, которая содержится в $B$, т.е. означают *произойдёт событие* $A \cap B$. Ясно, что $A \cap B \subset \Omega$ и $A \cap B \subset B = \Omega_1$.

**Определение 6.2.** Для того, чтобы подчеркнуть, что событие $A \cap B$ есть событие из нового пространства элементарных событий $\Omega_1 = B$ его обозначают $A|B$ и называют *событие $A$ при условии, что событие $B$ произошло.*

Очевидно, что $P(A|B) = \frac{\mu(A|B)}{\mu(B)}$. Вероятность $P(A|B)$ называется *условной вероятностью.*

**Лемма 6.3.** $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

[11]Сергей Натанович Бернштейн (1880 — 1968), советский математик.

*Доказательство.*

$$P(A|B) = \frac{\mu(A|B)}{\mu(B)} = \frac{\mu(A \cap B)}{\mu(B)} = \frac{\frac{\mu(A\cap B)}{\mu(\Omega)}}{\frac{\mu(B)}{\mu(\Omega)}} = \frac{P(A \cap B)}{P(B)}.$$

□

Следующая формула непосредственно следует из леммы 6.3 и традиционно называется *теоремой умножения.*

**Теорема 6.4** (Теорема умножения для двух событий)**.** *Если* $P(B) > 0$, $P(A) > 0$, *то*

$$P(A \cap B) = P(B)P(A|B) = P(A)P(B|A).$$

**Теорема 6.5** (Теорема умножения для n событий)**.**

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A2) \cdot \ldots \cdot P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}),$$

*если все условные вероятности определены.*

*Доказательство.* Доказать методом математической индукции. □

# 7. Формула полной вероятности и формулы Байеса

**Пример**. Три завода производят одну и ту же машину Renault. При этом 1-й завод производит 20%, 2-й — 30%, 3-й — 50% автомобилей. Брак на 1-м заводе составляет 5%, на 2-м — 2%, на 3-м — 4% автомобилей. Автомашины случайным образом поступают в продажу. Найти

а) вероятность купить бракованную машину и

б) условную вероятность того, что машина изготовлена на 1-м заводе. Решение. а) Т.к. сборка машин на заводах не зависима (замеч. 6.1), и

сборка машин на них не совместна (аксиома $\mathcal{P}3$), то вероятность купить бра- кованную машину равна доле бракованных машин во всей продукции, то есть $0,05 \cdot 0,2 + 0,02 \cdot 0,3 + 0,04 \cdot 0,5 = 0,036$.

б) Во этом случае условная вероятность покупки бракованной машины с 1-го завода равна доле брака 1-го завода среди всего количества бракованных машин

$$\frac{0,05 \cdot 0,2}{0,05 \cdot 0,2 + 0,02 \cdot 0,3 + 0,04 \cdot 0,5} = \frac{0,01}{0,036} = 0,2(7).$$

**Определение 7.1.** Если события $H_1, H_2, \dots \subset \Omega$

1) попарно несовместны (т.е. $H_i \cap H_j = \varnothing$ при $i \neq j$),

2) их объединение составляет всё пространство элементарных событий, т.е.

$$\bigcup_{i=1}^{\infty} H_i = \Omega,$$

3)$P(H_i) > 0$ для всех $i \in \mathbb{N}$,

то совокупность $\{H_1, H_2, \dots\}$ называется *полной группой событий*, а события $H_1, H_2, \dots$ называются *гипотезами.*

На 7 показана полная группа событий и некоторое событие A в пространстве $\Omega$.

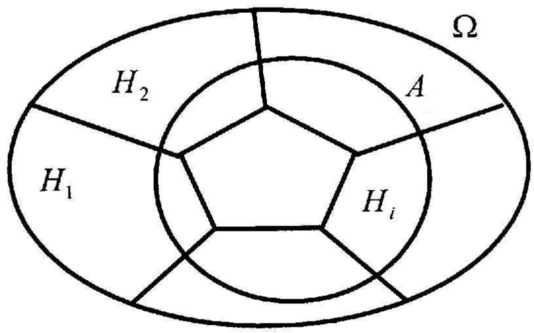


Рис. 7: Полная группа событий и некоторое событие A в пространстве $\Omega$

*Доказательство.* Заметим, что $A = A \cap \Omega = A \cap \left( \bigcup\limits_{i=1}^{\infty} H_i \right) = \bigcup\limits_{i=1}^{\infty} \cap H_i$, где события $A \cap H_1, A \cap H_2, \dots$– попарно несовместны. Используя аддитивность вероятности (аксиома $\mathcal{P}3$), а затем теорему умножения 6.3, получим требуемый результат

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

□

**Теорема 7.2** (формулы Байеса[12])**.** *Пусть $H_1, H_2, \dots$ – полная группа событий и $A$— некоторое событие положительной вероятности. Тогда условная вероятность события $H_k$ при условии, что событие $A$ произошло, для любого $k$ вычисляется по формуле*

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

[12]Томас Байес (Reverend Thomas Bayes, 1702 — 1761), английский математик и пресветерианский священник.

*Доказательство.* По определению условной вероятности имеем

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

где последнее равенство следует из теоремы умножения и формулы полной вероятности $\square$

# 8. Биномиальное распределение

Рассмотрим эксперимент, состоящим в $n$-кратном повторении схемы Бернулли (т.е. в подбрасывании ломаного гроша $n$ раз). Вопрос: «Какова вероятность того, что в результате этих $n$ испытаний «успех» (орёл) выпадет ровно $k$ раз, а остальные $n-k$ раз выпадет «неудача» (решка)?»

Обозначим искомую вероятность (выпадения орла) через $P_n(k)$. Ясно, что этот эксперимент имеет конечное пространство элементарных событий $\Omega = \{0, 1, 2, \ldots, n\}$ – орёл выпал 0 раз, 1 раз, 2 раза, ... , $n$ раз.

**Теорема 8.1.** *Для любого $k \in \Omega$ имеет место формула Бернулли*

$$P_n(k) = C_n^k p^k (1-p)^{n-k} = C_n^k p^k q^{n-k}.$$

*Доказательство.* Рассмотрим один из благоприятных элементарных исходов:

$$(\underbrace{у, у, \ldots, у}_{k \text{ раз}}, \underbrace{н, н, \ldots, н}_{n-k \text{ раз}}) \tag{1}$$

Здесь буквами «у» и «н» обозначены соответственно «успех» и «неудача». Поскольку испытания независимы, вероятность элементарного исхода (первые $k$ испытаний завершились успехом, а остальные неудачей) по опред. 5.1 равна $p^k(1-p)^{n-k}$.

Другие элементарные исходы, благоприятные событию $A$, отличаются от рассмотренного $(\underbrace{у, у, \ldots, у}_{k \text{ раз}}, \underbrace{н, н, \ldots, н}_{n-k \text{ раз}})$ лишь перераспределением $k$ успехов на $n$ местах. По теор. 1.6 существует ровно $C_n^k$ способов расположить $k$ успехов на $n$ местах. Поэтому событие $A$ состоит из $C_n^k$ элементарных исходов, вероятность каждого из них равна $p^k(1-p)^{n-k}$. $\square$

**Определение 8.2.** Эксперимент, имеющий ряд распределения

| $k$ | $P_n(k)$ |
|---|---|
| $0$ | $C_n^0p^0q^n = q^n$ |
| $1$ | $C_n^1p^1q^{n-1} = npq^{n-1}$ |
| $2$ | $C_n^2p^2q^{n-2} = \frac{n(n-1)}{2}p^2q^{n-2}$ |
| $\dots$ | $\dots$ |
| $k$ | $C_n^kp^kq^{n-k} = \frac{n!}{k!(n-k)!}p^kq^{n-k}$ |
| $\dots$ | $\dots$ |
| $n$ | $C_n^np^n = p^n$ |

называется биномиальным распределением.

Название этого распределения связан с тем, что сумма всех вероятностей из второй строки ряда распределения может быть вычислены по формуле бинома Ньютона, т.е.

$$\sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1 \qquad (2)$$

На рис. 8 показан график биномиального распределения $P_{16}(k) = C_16^k \cdot 0.65^k \cdot 0.35^{16-k}$. График является симметричным при $p = q = \frac{1}{2}$.

Если $p > q$, то максимум сдвигается вправо, и наоборот. Возникает естественный вопрос. Какое число успехов при $n$ испытаниях наиболее вероятно? Другими словами, при каком $k$ достигается максимум функции$P_n(k) = C_n^kp^kq^{n-k}$ ? Другими словами, при каком (каких) $k$ вероятность достигает максимума?

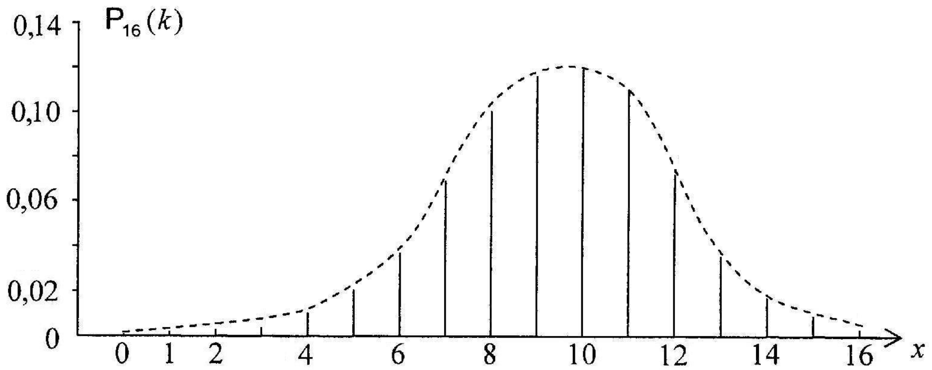


Рис. 8: График биномиального распределения при $n = 16$

**Теорема 8.3.** *Если в схеме Бернулли вероятность появления «успеха» (орла) равна $p$, то в биномиальном распределении с $n$ испытаниями наиболее вероятным числом «успехов» (орлов) является либо*

- *единственное число $[np + p]$, если число $np + p$ не целое, либо*
- *два числа $np + p$ и $np + p + 1$, если число $np + p$ целое.*

*Доказательство.* Сравним отношение чисел $P_n(k)$ и $P_n(k-1)$ с единицей.

$$\frac{P_n(k)}{P_n(k-1)} = \frac{C_n^k p^k q^{n-k}}{C_n^{k-1} p^{k-1} q^{n-k+1}} = \frac{p(n-k+1)}{kq} = 1 + \frac{np+p-k}{kq} \qquad (3)$$

Видно, что

1. $P_n(k) > P_n(k-1)$ при $np+p-k > 0$, т.е. при $k < np+p$;
2. $P_n(k) < P_n(k-1)$ при $np+p-k < 0$, т.е. при $k > np+p$;
3. $P_n(k) = P_n(k-1)$ при $np+p-k = 0$, что возможно лишь, если $np+p$ – целое число.

□

# 9. $k$-номинальное распределение

Полиномиальное распределение возникает в результате повторения $n$ раз схемы Бернулли (см. §8). По аналогии $k$-номиальное распределение возникает в результате повторения $n$ раз схемы независимых испытаний с $k \geqslant 3$ исходами (см. пп. 4.5 и 4.9).

**Пример**. Асимметричный тетраэдр, грани которого обозначены $\omega_1, \omega_2, \omega_3, \omega_4$ подбрасывают 14 раз. Вероятности выпадения этих граней лицом вниз соответственно равны $\frac{1}{2}, \frac{1}{3}, \frac{1}{9}, \frac{1}{18}$. Найти вероятность того, что грани $\omega_1, \omega_2, \omega_3, \omega_4$ выпадут соответственно 5, 3, 4, 2 раза.

*Ответ:* $\frac{14!}{5!3!4!2!}(\frac{1}{2})^5, (\frac{1}{3})^3, (\frac{1}{9})^4, (\frac{1}{18})^2 \simeq 0,00137$

*Замечание* 9.1. Известно, что биномом Ньютона называют формулу

$$(p+q)^n = \sum_{m=0}^{n} = C_n^m p^m q^{n-m},$$

где $C_n^m = \frac{n!}{m!(n-m)!}$

Обобщением бинома Ньютона является следующая формула

$$(p_1 + p_2 + \cdots + p_k)^n = \sum_{m_1+m_2+\cdots+m_k} = \frac{n!}{m_1!m_2!\ldots m_k!} p_1^{m_1} p_2^{m_2} \ldots p_k^{m_k},$$

где $m_1, m_2, \ldots, m_k \geqslant 0$. Эта формула может быть выведена из бинома Ньютона индукцией по k; по аналогии будем назвать её *k-номом Ньютона.*

**Лемма 9.2.** *Сумма k-номиальных коэффициентов равна $k^n$, т.е.*

$$\sum_{m_1+m_2+\cdots+m_k} = \frac{n!}{m_1!m_2!\ldots m_k!} = k^n$$

*Доказательство.* Подставим в $k$-ном Ньютона $p_1 = p_2 = \dots = p_k = 1$, получим требуемый результат. $\square$

**Лемма 9.3.** *Если $p_1 + p_2 + \dots + p_k = 1$, то сумма членов в правой части $k$-нома Ньютона равна единице, т.е.*

$$\sum_{m_1+m_2+\dots+m_k=n} = \frac{n!}{m_1!m_2!\dots m_k!} p_1^{m_1} p_2^{m_2} \dots p_k^{m_k} = 1$$

*Доказательство.* Подставим в $k$-ном Ньютона $p_1 = p_2 = \dots = p_k = 1$, получим требуемый результат. $\square$

**Определение 9.4.** Пусть эксперимент состоит в том, что

1) проводят n независимых испытаний в одинаковых условиях,

2) в каждом испытании появляется одно из k несовместных событий «««< Updated upstream

$$A_1, A_2, \dots, A_k$$

, =======

$$A_1, A_2, \dots, A_k$$

»»»> Stashed changes

3) эти события происходят с вероятностями $p_1, p_2, \dots, p_k$ соответственно,

и 4) $p_1 + p_2 + \dots + p_k = 1$. Тогда такой эксперимент называется схемой независимых испытаний.

**Теорема 9.5.** *Для любой схемы $n$ независимых испытаний и любых $m_1 \geqslant 0, m_2 \geqslant 0, \dots, mk \geqslant 0$, таких что $m_1 + m_2 + \dots + m_k = n$, вероятность $P(m_1, m_2, \dots, m_k)$ того, что события $A_1, A_2, \dots, A_k$ произойдут соответственно $m_1, m_2, \dots, m_k$ раз, определяется по формуле*

$$P(m_1, m_2, \dots, m_k) = \frac{n!}{m_1!m_2!\dots m_k!} p_1^{m_1} p_2^{m_2} \dots p_k^{m_k}.$$

*Доказательство.* Рассмотрим один элементарный исход схемы

$$n$$

независимых испытаний:

$$(\underbrace{A_1, \dots, A_1}_{m_1 \text{ раз}}, \underbrace{A_2, \dots, A_2}_{n-m_2 \text{ раз}}, \dots \underbrace{A_k, \dots, A_k}_{m_k \text{ раз}})$$

Это один из благоприятных исходов: сначала событие $A_1$ произошло $m_1$ раз, затем событие $A_2$ произошло $m_2$ раз, ..., и, наконец, событие $A_k$ произошло $m_k$ раз. Вероятность этого элементарного исхода равна $p_1^{m_1} p_2^{m_2} \dots p_k^{m_k}$.

Все остальные благоприятные исходы отличаются лишь расположением событий из того же набора событий на $n$ местах. Число таких исходов равно числу способов расставить на $n$ местах $m_1$ событий $A_1$, потом $m_2$ событий $A_2, \dots,$ и, наконец, $m_k$ событий $A_k$. По теор. 1.3 и 1.6 это число равно

$$C_n^{m_1} \cdot C_{n-m_2}^{m_2} \cdot C_{n-m_2-m_2}^{m_3} \cdot \dots \cdot C_{n-m_2-\dots m_{k-1}}^{m_k} = [???] = \frac{n!}{m_1!m_2!\dots m_k!}$$

$\square$

*Теперь становится ясно, откуда взялся ответ задачи про асимметричный тетраэдр в начале этого параграфа.*

# 10. Гипергеометрическое распределение

**Задача 10.1.** Из урны, содержащей $n_1$ белых и $n-n_1$ чёрных шаров, вынимают без возвращения $k \le n$ шаров. Найти вероятность события $A$, состоящего в том, что будет вынуто ровно $k_1$ белых и $k-k_1$ чёрных шаров. (Мы полагаем, что $k_1 \le n_1$ и $k - k_1 \le n - n_1$.)

**Решение**. Результатом эксперимента является набор из $k$ шаров. Оказывается, что искомая вероятность не зависит от того, будем ли мы или не будем учитывать порядок следования шаров. Чтобы показать это, подсчитаем её для обоих экспериментов.

1. Эксперимент без учёта порядка. В этом эксперименте общее число $\mu(\Omega)$ элементарных исходов равно числу $k$-элементных подмножеств множества, состоящего из $n$ элементов.

По теореме 1.6 оно равно $\mu(\Omega) = C_n^k$ . Обозначим через $A$ событие, состоящее в появлении набора, содержащего$k_1$ белых шаров и $k - k_1$ чёрных.

По теореме 1.3 число его благоприятных исходов равно произведению числа способов выбрать $k_1$ шаров из $n_1$ белых шаров и числа способов выбрать $k - k_1$ шаров из $n - n_1$ чёрных, т.е $\mu(A) = C_{n_1}^{k_1} \cdot C_{n-n1}^{k-k_1}$.

Получаем, что вероятность появления события $A$ в

$$P(A) = \frac{C_{n_1}^{k_1} \cdot C_{n-n1}^{k-k_1}}{C_n^k} \tag{4}$$

2. Эксперимент с учётом порядка. По теореме 1.4 общее число $\mu(\Omega)$ элементарных исходов равно числу способов разместить $n$ элементов на $k$ местах, т.е. $\mu(\Omega) = A_n^k = n(n-1)\dots(n-k+1)$.

В этом эксперименте при подсчёте числа благоприятных исходов $\mu(A)$ надо учесть как число способов выбрать нужное число $k$ шаров, так и число способов расположить белые и чёрные шары среди выбранных $k$ шаров.

Во-первых, мы подсчитываем число способов расположить $k_1$шаров на $k$ местах. Оно равно $C_k^{k_1}$.

Затем мы подсчитываем число способов расположить $k_1$ белых шаров на $n_1$ местах. Учитывая их порядок, получаем, что это число равно $A_{n_1}^{k_1}$.

И наконец, мы подсчитываем число способов разместить $k - k_1$ чёрных шаров на оставшихся $n - n_1$ местах. Оно равно $A_{n-n_1}^{k-k_1}$.

Перемножая эти числа, по теореме 1.4 получим

$$\mu(A) = C_k^{k_1} \cdot A_{n_1}^{k_1} \cdot A_{n-n1}^{k-k_1} \tag{5}$$

Подсчитывая искомую вероятность, получим

$$P(A) = \frac{C_k^{k_1} \cdot A_{n_1}^{k_1} \cdot A_{n-n1}^{k-k_1}}{A_n^k} = \frac{C_{n_1}^{k_1} \cdot C_{n-n1}^{k-k_1}}{C_n^k} \tag{6}$$

**Определение 10.2.** Ряд распределения, определённый соответствием

$$k_1 \mapsto P(A) = \frac{C_{n_1}^{k_1} \cdot C_{n-n1}^{k-k_1}}{C_n^k}, \tag{7}$$

где $0 \leq k_1 \leq \min(k, n_1)$ и $k - k_1 \leq n - n_1$ называется *гипергеометрическим распределением.*

# Часть II
# Теория случайных величин

## 11. Случайные величины

*На практике элементарные случайные события являются либо числами, либо наборами чисел (случайными векторами). В §§ 11–13 мы изучим одномерные случайные величины, у которых $\Omega \subset \mathbb{R}^1$ — числовое множество.* Начиная с §13 изучим случайные векторы, у которых $\Omega \subset \mathbb{R}$, при $n \geqslant 2$.

Чтобы дать общее определение случайной величины пусть сначала $(\Omega, \mathcal{A}, \mathcal{P})$— произвольное вероятностное пространство.

**Определение 11.1.** Функция $\xi : \Omega \to \mathbb{R}$ называется случайной величиной на $\sigma$-алгебре событий $A$, если для любого числа $x \in \mathbb{R}$ прообраз луча $\xi^{-1}((-\infty, x]) = \{\omega \in \Omega | \xi(\omega) \leqslant x\}$ является событием из $\sigma$-алгебры $A$.

*Замечание* 11.2. 1) Самая простая функция $\xi : \Omega \to \mathbb{R}$ — это постоянная функция, заданная для любого $\omega \in \Omega$ по формуле $\xi(\omega) = c$. Она принимает одно значение и является не случайной, а детерминированной. Она рассматривается в теории вероятностей как частный тривиальный случай.

2)Первая нетривиальная случайная величина (с $\Omega \in \mathbb{R}$) задаётся с помощью тождественной функции $\xi(\omega) = \omega$. В этом случае событие $\xi^{-1}((-\infty, x]) = \{\omega \in \Omega | \xi(\omega) \leqslant x\}$ обозначают короче: $\{\xi \leqslant x\}$.

3) Оказывается, что другие случайные величины ($\xi(\omega) \neq \omega$) могут быть описаны через случайную величину $\xi(\omega) = \omega$. Это будет показано в § 13.

**Определение 11.3.** Функцией распределения случайной величины $\xi$ называется функция $F_\xi(x) : R \to [0, 1]$, определённая по формуле

$$F_\xi(x) = P(\xi \leqslant x)$$

.

Очевидно, что $0 \leqslant F_\xi(x) \leqslant 1$

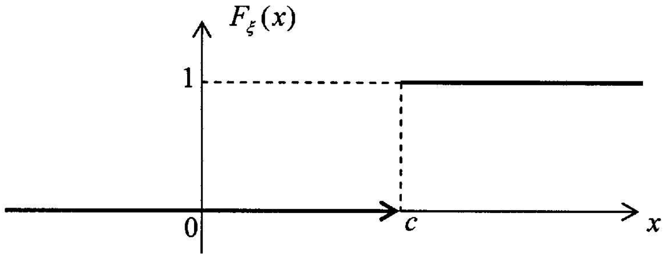


Рис. 9: Функция распределения детерминированной величины.

**Примеры**

1) Детерминированная (вырожденная случайная) величина $\xi : \Omega \to \mathbb{R}$ имеет пространство элементарных событий, состоящее из одного элементарного события, $\Omega = \{\omega\}$. Она принимает только одно значение $\xi(\omega) = c = const \in \mathbb{R}$ с вероятностью равной 1, т.е. имеет ряд распределения

| $\xi$ | $c$ |
|---|---|
| $P$ | 1 |

. Её функция распределения показана на рис. 9.

2) Случайная величина $\xi$, имеющая распределение Бернулли, и принимающая значения 1 (успех) и 0 (неудача) с вероятностями соответственно $p$ и $1 - p$, имеет ряд распределения

| $\xi$ | 0 | 1 |
|---|---|---|
| $P$ | $1 - p$ | $p$ |

и имеет график, показанный на рис. 10.

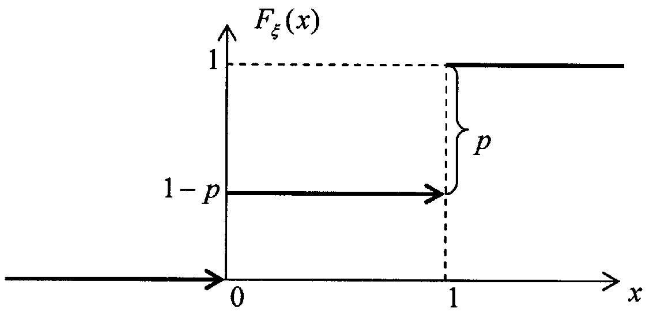


Рис. 10: Функция распределения Бернулли.

3) Случайная величина $\xi$ — номер грани при подбрасывании игральной кости имеет функцию распределения, показанную на рис. 11.

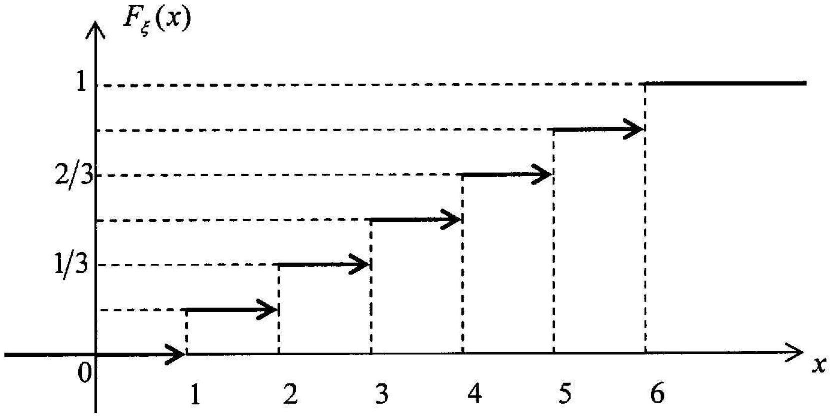


Рис. 11: Функция распределения выпадения числа на игральной кости.

4) Случайная величина $\xi$ — номер появления успеха в геометрическом распределения имеет ряд распределения

| $\xi$ | 1 | 2 | ... | $k$ | ... |
|---|---|---|---|---|---|
| $P$ | $p$ | $pq$ | ... | $pq^{k-1}$ | ... |

и имеет функцию распределения, показанную на рис. 12. В дальнейшем мы будем использовать следующие обозначения для пределов функций слева и справа соответственно:

$$\lim_{\varepsilon\to 0} h(x-\varepsilon) \overset{\text{опр}}{=} h(x-0) \text{ и } \lim_{\varepsilon\to 0} h(x+\varepsilon) \overset{\text{опр}}{=} h(x+0),$$

где всегда $\varepsilon > 0$.

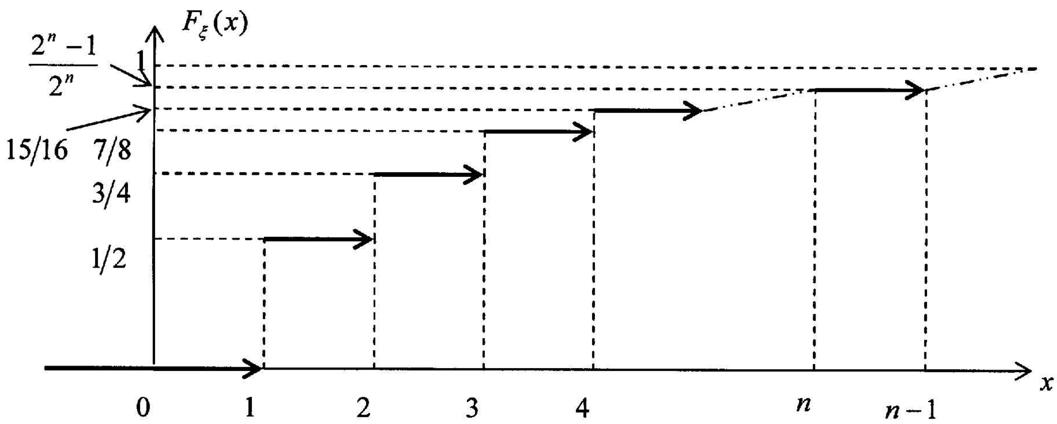


Рис. 12: Функция распределения выпадения числа на игральной кости.

Функция распределения $F_\xi(x)$ обладает следующими свойствами.

**Теорема 11.4.** *1. $F_\xi(x)$ — неубывающая функция, другими словами, если $x_1 < x_2$, то $F_\xi(x_1) \leqslant F_{xi}(x_)$.*

*2. $F_\xi(-\infty) = \lim\limits_{x\to-\infty} F_\xi(x) = 0$ и $F_\xi(+\infty) = \lim\limits_{x\to+\infty} F_\xi(x) = 1$*

*3. $F_\xi(x)$ непрерывна справа: для любой точки $x_0 \in \mathbb{R}$ имеем*

$$F_\xi(x+x_0) = F_\xi(x_0))$$

*Доказательство.* По пп. 11.3, 3.3, 3.1 имеем

1. $F_\xi(x_2) - F_\xi(x_1) = P(\xi \leqslant x_2) - P(\xi \leqslant x_1) = P\{\xi \leqslant x_2\} \setminus P\{\xi \leqslant x_1\} = = P(x_1 < \xi \leqslant x_2) \geqslant 0.$

2. $F_\xi(-\infty) = \lim\limits_{x\to-\infty} F_\xi(x) = \lim\limits_{x\to-\infty} P(\xi \leqslant x) = P(\xi \leqslant -\infty) = 0$ и $F_\xi(+\infty) = \lim\limits_{x\to+\infty} F_\xi(x) = \lim\limits_{x\to+\infty} P(\xi \leqslant x) = P(\xi \leqslant +\infty) = 1.$

3. $F_\xi(x_0 + 0) = \lim\limits_{\varepsilon\to 0} F_\xi(x_0 + \varepsilon) = \lim\limits_{x\to\varepsilon} P(\xi \leqslant x_0 + \varepsilon) = P(\xi \leqslant x_0) = F_\xi(x_0).$

□

**Теорема 11.5.** *Функция распределения имеет не более чем счётное множество скачков.*

*Доказательство.* Разобьём множество H всех скачков на подмножества скачков по убыванию высоты h:

1) $H_1$ – множество скачков высоты $h \in (\frac{1}{2}, 1]$,
2) $H_2$ – множество скачков высоты $h \in (\frac{1}{4}, \frac{1}{2}]$,
3) $H_2$ – множество скачков высоты $h \in (\frac{1}{8}, \frac{1}{4}]$,
...
$n$) – множество скачков высоты $h \in (\frac{1}{2^n}, \frac{1}{2^{n-1}}]$
...

Видно, что множества $H_n$ для всех $n \in \mathbb{Z}$ попарно непересекающиеся, т.е. $H_i \cap H_j = \varnothing$ при $i \neq j$; и $\bigcup\limits_{i=1}^{\infty} = H$. Функция распределения может иметь

Во-первых, не более одного скачка высоты $h \in (\frac{1}{2}, 1]$,т.е. $\mu(H_1) \leqslant 1$;
Во-вторых, не более трёх скачков высоты $h \in (\frac{1}{4}, \frac{1}{2}]$,т.е. $\mu(H_2) \leqslant 3$;
В-третьих, не более семи скачков высоты $h \in (\frac{1}{8}, \frac{1}{4}]$,т.е. $\mu(H_3) \leqslant 7$;
...
В-n-ых, не более $2^n - 1$ скачков высоты $h \in (\frac{1}{2^n}, \frac{1}{2^{n-1}}]$,т.е. $\mu(H_n) \leqslant 2^n - 1$;
...

Можно занумеровать все скачки по убыванию их высоты, нумеруя при этом скачки равной высоты, если такие повстречаются. Эта нумерация ставит множество всех скачков во взаимно однозначное соответствие с множеством $\mathbb{Z}$ натуральных чисел. Что и требовалось доказать.

□

**Теорема 11.6** (О связи вероятности событий-интервалов с функцией распределения.)**.** *Для любых точек $x, a, bin\mathbb{R}$, где $a < b$, имеем*

*1.* $P(\xi < x) = F_\xi(x - 0)$,

*2.* $P(\xi = x) = F_\xi(x) - F_\xi(x - 0)$, *3. если $F_\xi$ непрерывна в точке $x$, то $P(\xi = x) = 0$;*

*4. $P(a < \xi \leqslant b) = F_\xi(b) - F_\xi(a)$,*

*5. $P(a \leqslant \xi \leqslant b) = F_\xi(b) - F_\xi(a-0)$,*

*6. $P(a < \xi < b) = F_\xi(b-0) - F_\xi(a)$*

*7. $P(a \leqslant \xi) = F_\xi(b-0) - F_\xi(a-0)$.*

*Доказательство.* 1. $P(\xi < x) = \lim\limits_{\varepsilon\to 0} P(\xi \leqslant x - \varepsilon) = \lim\limits_{\varepsilon\to 0} F_\xi(x-\varepsilon) = F_\xi(x-0)$.

2. $P(\xi = x) = P(\{\xi \leqslant x\}) \setminus \{\xi < x\}) = P(\xi \leqslant x) - P(\xi < x) = F_\xi(x) - F_\xi(x)$.

3. $P(\xi = x) = F_\xi - F_\xi(x-0) = F_\xi(x) - \lim\limits_{\varepsilon\to 0} F_\xi(x-\varepsilon) = F_\xi(x) - F_\xi(x) = 0$.

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

5. Самостоятельно

6. Самостоятельно

7. Самостоятельно

□

**Следствие 11.7.** Если функция распределения $F_\xi(x)$ непрерывна, то

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

# 12. Абсолютно непрерывные случайные величины

**Определение 12.1.** Случайная величина $\xi$ называется абсолютно непрерывной, если существует такая неотрицательная функция $f_\xi(x)$ (возможно обобщённая, см. §13), что для любого $x \in \mathbb{R}$ функция распределения $F_\xi(x)$ представима в виде[13]

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(t)dt$$

[13] В дальнейшем мы рассматриваем только такие случайные величины, для которых все связанные с ними несобственные интегралы и суммы бесконечных рядов сходятся абсолютно.

При этом функция $f_\xi(x)$ называется плотностью вероятности случайной величины $\xi$

**Замечание 12.2.** Если $F_\xi(x)$ – дифференцируемая функция распределения, то название "плотность вероятности"имеет следующее объяснение. С одной стороны, используя формулу дифференцирования интеграла по верхнему пределу, имеем $\frac{dF_\xi(x)}{dx} = f_\xi(t)|_t = x$. С другой стороны, по определению производной имеем:

$$\frac{dF_\xi(x)}{dx} = \frac{F_\xi(x+dx) - F_\xi(x)}{dx} = \frac{P(\xi \le x+dx) - P(\xi \le x)}{dx} =$$
$$= \frac{P(\{\xi \le x+dx\} \setminus \{\xi \le x\})}{dx} = \frac{P(x < \xi \le x+dx)}{dx}$$

Поэтому

$$f_\xi(x) = \frac{P(x, \xi \le x+dx)}{dx}$$

есть отношение вероятности попадания случайной величины $\xi$ в интервал $(x, x+dx]$ к длине этого интервала $dx$, что физически означает плотность вероятности («массы») случайной величины $\xi$ в точке $x$.

**Теорема 12.3.** *Плотность обладает свойствами:*

*1)* $f_\xi(x) \ge 0$ *для любого* $x$*;*

*2)* $\int\limits_{-\infty}^{\infty} f_\xi(t)dt = 1$

*Доказательство.* 1) По теор. 11.4.1 функция распределения $F_\xi(x)$ – неубывающая, поэтому $\frac{dF_\xi(x)}{dx} = f_\xi(x) \ge 0$

2) $\int\limits_{-\infty}^{\infty} f_\xi(t)dt = lim_{x\to\infty} \int\limits_{-\infty}^{x} f_\xi(t)dt = \lim\limits_{x\to\infty} F_\xi(x) = 1$ по свойству 11.4.2. □

**Определение 12.4.** Пусть область $U \subseteq \mathbb{R}^n$ имеет меру $\mu(U)$. Говорят, что в области $U$ функция $h : U \to \mathbb{R}$ удовлетворяет некоторому свойству $\mathcal{P}$ почти всюду, если подмножество $A \subset U$ , в котором свойство $\mathcal{P}$ не выполняется, имеет меру 0, т.е. $\mu(A) = 0$. Докажем свойства абсолютно непрерывных случайных величин.

**Теорема 12.5.** *Если случайная величина $\xi$ абсолютно непрерывна, то*

*1) её функция распределения $F_\xi(x)$ непрерывна в $\mathbb{R}$ и дифференцируема почти всюду в $\mathbb{R}$, т.е. равенство $f_\xi(x) = \frac{d}{dx}F_\xi(x)$ справедливо почти всюду,*

*2) $P(\xi = x) = 0$ для любого $x \to \mathbb{R}$,*

*3) $P(a \le \xi < b) = P(a < \xi < b) = P(a \le \xi \le b) = P(a < \xi \le b) = \int_a^b f_\xi(t)dt$.*

*Доказательство.* 1) Во-первых, функция $F_\xi(x) = \int_{-\infty}^{x} f_\xi(t)dt$ непрерывна как функция верхнего предела интеграла. Во-вторых, по теор. 11.5 функция $F_\xi(x)$ имеет не более чем счётное множество скачков, поэтому она не дифференцируема не более чем в счётном множестве точек. Любое счётное множество точек на прямой имеет нулевую меру (длину). Поэтому $F_\xi(x)$ дифференцируема почти всюду в $\mathbb{R}$.

2) Это следует из 11.6.3).

3) Из 11.6.4) имеем

$$P(a < \xi \leq b) = F_\xi(b) - F_\xi(a) = \int_{-\infty}^{b} f_\xi(t)dt - \int_{-\infty}^{a} f_\xi(t)dt = \int_{a}^{b} f_\xi(t)dt$$

□

Остальные равенства следуют из 11.7.

### Примеры абсолютно непрерывных случайных величин

**Определение 12.6.** Случайная величина $\xi$ имеет равномерное распределение на отрезке $a, b$, если

$$F_\xi(x) = P(\xi \leq x) = \begin{cases} 0, x < a \\ \dfrac{x-a}{b-a}, a \leq x \leq b \\ 1, x > b \end{cases} \text{ и } f_\xi(x) = \begin{cases} 0, x < a \\ \dfrac{1}{b-a}, a \leq x \leq b \\ 1, x > b \end{cases}$$

Графики функций $F_\xi(x)$ и $f_\xi(x)$ показаны на рис. 13. Заметим, что в точках $a$ и $b$ функция распределения $F_\xi(x)$ не дифференцируема, поэтому значение плотности вероятности в этих точках можно задать как угодно: либо $f_\xi(a) = 0$, либо $f_\xi(a) = \frac{1}{b-a}$, а на другом конце тоже либо $f_\xi(b) = 0$, либо $f_\xi(b) = \frac{1}{b-a}$

**Определение 12.7.** Случайная величина $\xi$ имеет показательное распределение с параметром $\lambda$ (рис. 14), если

$$f_\xi(x) = \begin{cases} 0, x < 0 \\ \lambda e^{-\lambda x}, x \geq 0 \end{cases}$$

**Определение 12.8.** Если для любого $x \in \mathbb{R}$ случайная величина $\xi$ имеет плотность

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

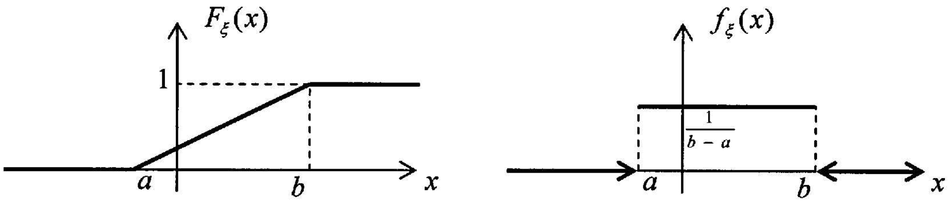


Рис. 13: Равномерное распределение

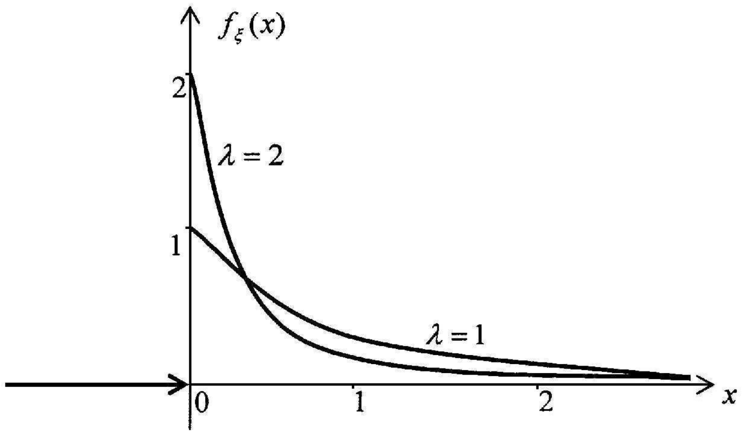


Рис. 14: Показательное распределение

то говорят, что $\xi$ имеет нормальное (или Гаусса[14] ) распределение с математическим ожиданием $a$ и дисперсией $\sigma^2$ , где $a \in \mathbb{R}$ и $\sigma > 0$. (См. рис. 15) Очевидно, что $f_\xi(x) \geq 0$ для любого $x$;

2) Используем табличный интеграл (Пуассона) $\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}$.

$$\int_{-\infty}^{\infty} f_\xi(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \begin{bmatrix} \text{замена переменных} \\ t = \dfrac{x-a}{\sigma}, dx = \sigma dt \end{bmatrix} =$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = 1.$$

**Свойство 12.9.** Легко сосчитать, что расстояние между точками перегиба равно $2\sigma$, поэтому параметр $2\sigma$ является характеристикой ширины графика на уровне точек перегиба.

Заметим, что вероятность попадания нормальной случайной величины в интервал между точками перегиба

$$P(a-\sigma \leq \xi \leq a+\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{a-\sigma}^{a+\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-1}^{1} e^{-\frac{t^2}{2}} dt = 2\Phi(1) \approx 0.6827$$

[14]Карл Фридрих Гаусс (Johann Carl Friedrich Gauß, 1777 — 1855), выдающийся немецкий математик, астроном и физик, считается одним из величайших математиков всех времён.

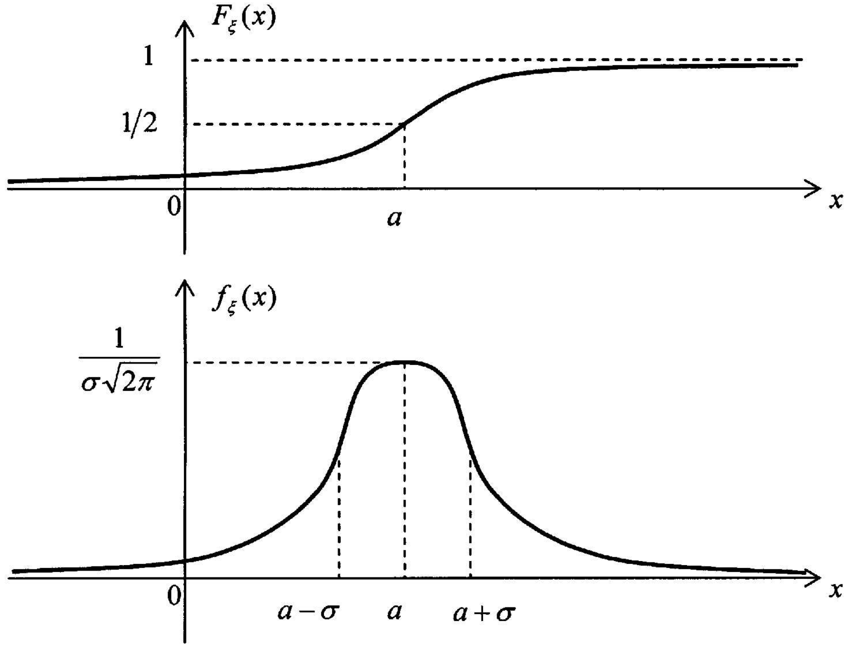


Рис. 15: Нормальное распределение (Гаусса)

где $\Phi(1)$ приближённое значение функции $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ взято из таблицы.

**Свойство 12.10** («Правило трёх сигм»)**.** Если $\xi$ — нормальная случайная величина, то $P(|\xi - a| \leq 3\sigma) = 2\Phi(3) \approx 2 * 0.49865 \approx 0.997$. Запоминать число $0,997$ нет никакого смысла, а вот помнить, что почти вся вероятность («масса») нормального распределения сосредоточена в интервале $[a - 3\sigma, a + 3\sigma]$, всегда полезно.

# 13. Функции Хевисайда и Дирака

**Теорема 13.1.** *Функцией единичного скачка или короче функцией Хевисайда*[15] *называется функция $U : \mathbb{R} \to \mathbb{R}$, заданная по формуле (см. рис. 16)*

$$U(x) = \begin{cases} 0, x < 0 \\ 1, x \geqslant 0 \end{cases}$$

**Пример 13.2.** Функция Хевисайда полезна для записи в строчку формул (многоэтажных) кусочно заданных функций.

[15] Оливер Хевисайд (англ. Oliver Heaviside; 18 мая 1850 — 3 февраля 1925) — английский учёный- самоучка, инженер, математик и физик. Впервые применил комплексные числа для изучения электрических цепей. Переписал уравнения Максвелла из их первоначальной формы, состоявшей из 20 уравнений с 12 переменными, к современной форме, состоящей из 4 дифференциальных уравнений, выраженной в терминах современного векторного анализа. Предложил операционное исчисление (он ввёл обозначение D для дифференциального оператора) и метод решения дифференциальных уравнений с помощью сведения к обыкновенным алгебраическим уравнениям. Ввёл термины: «проводимость», «проницаемость», «индуктивность», «импеданс».

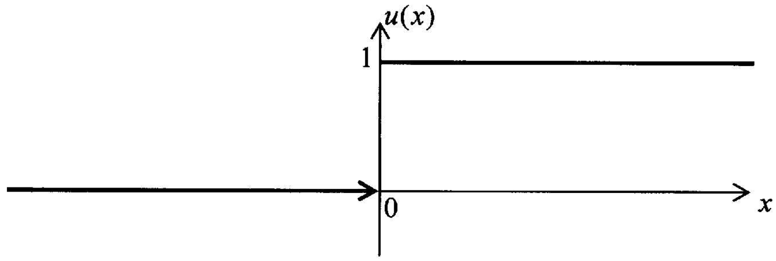


Рис. 16: Функция Хевисайда

1) Функция единичного импульса единичной длины см. рис. ??.

$$U(x) = \begin{cases} 0, x < 0 \\ 1, x \geqslant 0 \end{cases} = u(x) - u(x-1).$$

2) Функция распределения Бернулли, см. рис. 10.

$$F_\xi(x) = \begin{cases} 0, x < c \\ 1-p, 0 \leqslant x < 1 p, 1 \leqslant x \end{cases} = (1-p)u(x) - pu(x-1).$$

3) Функция равномерного распределения, см. рис.

$$F_\xi(x) = \begin{cases} 0, x < a \\ \frac{x-a}{b-a}, a \leqslant x < b 1, x \geqslant b \end{cases} = \frac{x-a}{b-a}u(x-a) - \frac{x-a}{b-a}u(x-b) + u(x-b).$$

4)Интеграл от функции Хевисайда.

$$\int\limits_{-\infty}^{x} f(t)u(t)dt = \int\limits_{0}^{x} f(t)dt$$

в частности,

$$\int\limits_{-\infty}^{x} u(t)dt = xu(x).$$

**Лемма 13.3.** *Кусочная функция*

$$g(x) = \begin{cases} g_0(x), x < x_0 \\ g_1(x), x_0 \leqslant x < x_1 \ldots, \ldots g_n(x), x_{n-1} \leqslant x < x_n g_{n+1}(x), x_n \leqslant x \end{cases}$$

*представима в виде*

$$g(x) = g_0(x)u(x_0 - x) + g_1(x)[u(x-x_0) - u(x-x_1)] + \ldots$$
$$\ldots + g_n(x)[u(x-x_{n-1}) - u(x-x_n)] + g_{n+1}(x)u(x-x_n).$$

*Доказательство.* Функция $g(x)$ является суммой следующих функций

$$\widetilde{g}_0(x) = \begin{cases} g_0(x), x < x_0 0, x_0 \leqslant x \end{cases} = g_0(x)u(x_0 - x),$$

$$\widetilde{g}_1(x) = \begin{cases} 0(x), x < x_0 \text{ или } x_1 \leqslant x \\ g_1(x), x_0 \leqslant x < x_1 \end{cases} = g_1(x)u(x - x_0) - g_1(x)u(x - x_1),$$

$\ldots, \ldots$

$$\widetilde{g}_n(x) = \begin{cases} 0(x), x < x_{n-1} \text{ или } x_n \leqslant x \\ g_n(x), x_{n-1} \leqslant x < x_n \end{cases} = g_n(x)u(x - x_{n-1}) - g_n(x)u(x - x_{n-n}),$$

$$\widetilde{g}_{n+1}(x) = \begin{cases} 0(x), x < x_n \\ g_n(x), x_n \leqslant x < x_n \end{cases} = g_0(x)u(x - x_n).$$

$\square$

**Определение 13.4.** $\delta$-функцией или функцией Дирака называется обобщённая функция $\delta : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$, заданная по формуле

$$\delta(x) = \begin{cases} 0, x \neq 0 \\ \infty, x = 0 \end{cases}$$

и для любого $\varepsilon > 0$ подчинённая условию $\int\limits_{-\varepsilon}^{\varepsilon} \delta(x)dx = 1$. См. рис. 17.

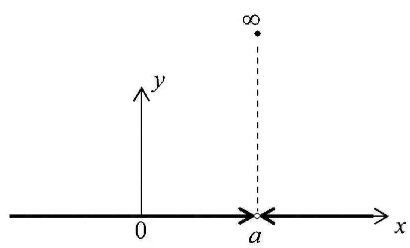


Рис. 17: Функция $y = \delta(x - a)$

**Свойство 13.5.** 1. Для любого $c \in \mathbb{R}$ и любого $\varepsilon > 0$, имеет место формула

$$\int\limits_{c-\varepsilon}^{c+\varepsilon} \delta(x - c)dx = 1$$

2.
$$u(x) = \int\limits_{-\infty}^{x} \delta(t)dt$$

3.
$$u'(x) = \delta(x)$$

4. Для любой непрерывной функции $f(x)$ любого $a \in \mathbb{R}$ имеет место тождество
$$f(x)\delta(x-a) = f(a)\delta(x-a). \qquad (8)$$

5. Для любой непрерывной функции $f(x)$ любого $a \in \mathbb{R}$ имеет место тождество
$$\int\limits_{-\infty}^{x} f(t)\delta(t-a)dt = f(a)u(x-a), \qquad (9)$$
в частности
$$\int\limits_{-\infty}^{x} \delta(t-a)dt = u(x-a) \qquad (10)$$

6. Если функция $f(x)$ непрерывна и все её нули (корни) $x_1, x_2, \ldots, x_k, \ldots$ простые (т.е. имеют кратность 1), то
$$\delta(f(x)) = \sum_k \frac{\delta(x-x_k)}{|f'(x_k)|}. \qquad (11)$$

7. Дельта-функция получается при вычислении преобразования Фурье от константы:
$$\int\limits_{-\infty}^{\infty} e^{ixt}dt = 2\pi\delta(x) \qquad (12)$$

**Пример 13.6.** 1. См.пример 13.2.3.

Дано
$$F_\xi(x) = \frac{x-a}{b-a}[u(x-a) - u(x-b)] + u(x-b). \qquad (13)$$

Найти $f_\xi(x)$.

Решение.

$$f_\xi(x) = F'_\xi(x) =$$

$$\left(\frac{x-a}{b-a}\right)[u(x-a)-u(x-b)] + \frac{x-a}{b-a}[u(x-a)-u(x-b)] + \delta(x-b) =$$
$$= \frac{1}{b-a}[u(x-a)-u(x-b)] + \frac{x-a}{b-a}[\delta(x-a)-\delta(x-b)] + \delta(x-b) =$$
$$= \frac{1}{b-a}[u(x-a)-u(x-b)] + \frac{x-a}{b-a}\delta(x-a) - \frac{x-a}{b-a}\delta(x-b) + \delta(x-b) =$$
$$= \frac{1}{b-a}[u(x-a)-u(x-b)] + \frac{a-a}{b-a}\delta(x-a) - \frac{b-a}{b-a}\delta(x-b) + \delta(x-b) =$$
$$= \frac{1}{b-a}[u(x-a)-u(x-b)].$$

2. Найти плотность вероятности по данной функции распределения (см. рис. 18)

Дано.

$$F_\xi(x) \begin{cases} 0, x < 0 \\ \frac{1}{24}(5x+8), 0 \leqslant x < 2 \\ 1, x \geqslant 2 \end{cases} \quad = \frac{1}{24}(5x+8)[u(x) - x(x-2)] + u(x-2).$$

Решение.

$$f_\xi(x) = F'_\xi(x) =$$
$$= \frac{5}{24}[u(x)-u(x-2)] + \frac{1}{24}(5x+8)[\delta(x)-\delta(x-2)] + \delta(x-2) =$$
$$= \frac{5}{24}[u(x)-u(x-2)] + \frac{1}{3}\delta(x) - \frac{3}{4}\delta(x-2) =$$
$$= \frac{1}{3}\delta(x) + \frac{5}{24}[u(x)-u(x-2)] + \frac{1}{4}\delta(x-2).$$

См. рис. 19

**Определение 13.7.** Функции Хевисайда и Дирака и их композиции относятся к классу так называемых обобщённых функций.

# 14. Функции одной случайной величины

**Замечание 14.1.** Пусть $\xi : \Omega \to \mathbb{R}$ – абсолютно непрерывная случайная величина, имеющая плотность $f_\xi(x)$. Построим с помощью функции $g : \mathbb{R} \to \mathbb{R}$ новую случайную величину по формуле $\eta = g(\xi)$.

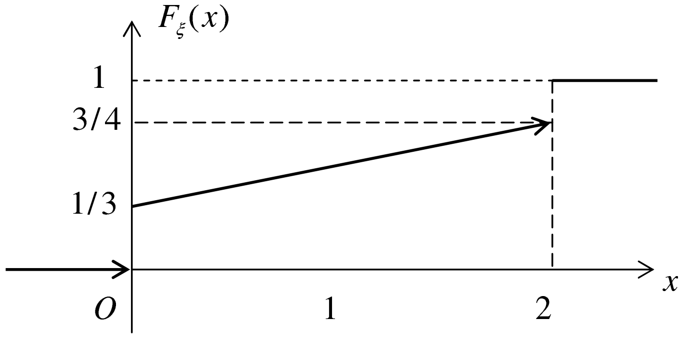


Рис. 18: Функция распределения с двумя скачками

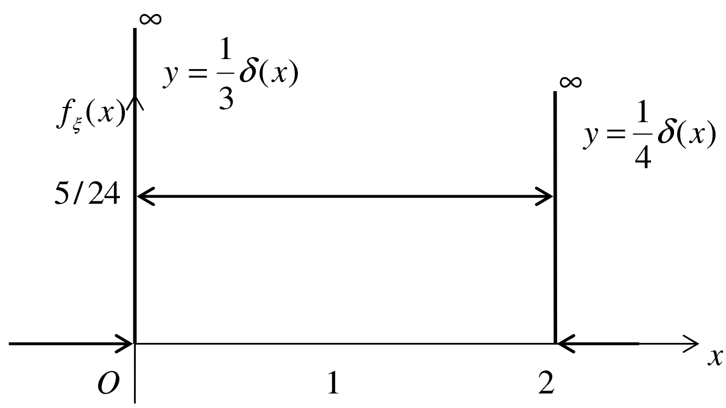


Рис. 19: Площадь под плотностью равна $\frac{1}{3}+\frac{5}{24}\cdot 2+\frac{1}{4}=1$.

Требуется найти функцию распределения и плотность случайной величины $\eta$. Мы решим эту задачу сначала в предположении, что функция $y=g(x)$ дифференцируема и монотонна, т.е. когда во всех точках $x\in\mathbb{R}$ выполнено либо $g'(x)>0$, либо $g'(x)<0$.

**Теорема 14.2.** *Если $\xi$ – абсолютно непрерывная случайная величина, имеющая функцию распределения $F_\xi(x)$ и плотность $f_\xi(x)$, и если $g:\mathbb{R}\to\mathbb{R}$ – дифференцируемая и монотонная функция, то случайная величина $\eta=g(\xi)$ имеет плотность вероятности*

$$f_\eta(y)=f_\xi\left(g^{-1}(y)\right)\left|\frac{\mathrm{d}[g^{-1}(y)]}{\mathrm{d}y}\right|$$

*Доказательство.* Заметим, что если $g:\mathbb{R}\to\mathbb{R}$ – монотонная функция, то существует её функция$g^{-1}:\mathbb{R}\to R$, и выполнено тождество

$$g(g^{-1}(y))\equiv y$$

Дифференцируя его, получим тождество

$$g'(g^{-1}(y))\cdot(g^{-1}(y))'\equiv 1,$$

которое означает, что производные $g'$ и $(g^{-1})'$ – одного знака, т.е. функции $g$ и $g^{-1}$ либо обе возрастающие, либо обе убывающие.

1) Пусть сначала $g$ – возрастающая функция, т.е. $g' > 0$ и $g^{-1} > 0$. Это означает, что неравенство $g(\xi) \le y$ можно записать в виде $\xi \le g^{-1}(y)$.

$$F_\eta(y) = F_{g(\xi)}(y) = P(g(\xi) \le y) = P(\xi \le g^{-1}(y)) = F_\xi(g^{-1}(y)) =$$
$$= \int\limits_{-\infty}^{g^{-1}(y)} f_\xi(t)dt = \begin{bmatrix} t = g^{-1}(\tau), & dt = (g^{-1}(\tau))'d\tau \\ t = -\infty \mapsto \tau = -\infty, & t = g^{-1}(y) \mapsto \tau = y \end{bmatrix} =$$
$$= \int\limits_{-\infty}^{y} (g^{-1}(\tau))' f_\xi(g^{-1}(\tau))d\tau.$$

2) Пусть теперь $g$ – убывающая функция, т.е. $g' < 0$и $g^{-1} < 0$, тогда неравенство $g(\xi) \le y$ можно записать в виде $\xi \ge g^{-1}(y)$.

$$F_\eta(y) = F_{g(\xi)}(y) = P(g(\xi) \le y) = P(\xi \ge g^{-1}(y)) =$$
$$= \int\limits_{g^{-1}(y)}^{\infty} f_\xi(t)dt = \begin{bmatrix} t = g^{-1}(\tau), & dt = (g^{-1}(\tau))'d\tau \\ t = -\infty \mapsto \tau = -\infty, & t = g^{-1}(y) \mapsto \tau = y \end{bmatrix} =$$
$$= \int\limits_{y}^{-\infty} (g^{-1}(\tau))' f_\xi(g^{-1}(\tau))d\tau = \int\limits_{-\infty}^{y} \left|(g^{-1}(\tau))'\right| f_\xi(g^{-1}(\tau))d\tau.$$

Объединяя оба случая в один, получим требуемую формулу. □

**Определение 14.3.** Пусть $A$ – подмножество на прямой $\mathbb{R}$, т.е. $A \subset \mathbb{R}$. Функция $\mathbf{1}_A : \mathbb{R} \to \{0, 1\}$, определённая по формуле

$$\mathbf{1}_A(x) = \begin{cases} 1, & \text{если } x \in A \\ 0, & \text{если } x \notin A \end{cases}$$

называется выделяющей функцией множества $A$ или индикатором $A$. Если $A = [0, \infty)$, то обозначение $\mathbf{1}_A(x)$ сокращают до $\mathbf{1}(x)$. Заметим, что $\mathbf{1}(x) = u(x)$, т.е. совпадает с функцией Хевисайда.

**Замечание 14.4.** Пусть теперь $g : \mathbb{R} \to \mathbb{R}$, $y = g(x)$ – дифференцируемая, кусочно монотонная функция, имеющая интервалы монотонности:

$$\mathcal{D} = \{D1 = (-\infty, a_1], D_2 = (a_1, a_2], \ldots, D_n = (a_{n-1}, \infty)\}$$

Ясно, что все ограничения $g|_{D_i} : D_i \to \mathbb{R}$, определённые по формулам $g|_{D_i}(x) = g(x)$ являются взаимно однозначными функциями и поэтому имеют обратные $\left(g|_{D_i}\right)^{-1}(y) = g|_{D_i}{}^{-1}(y)$ с областями определения $g(D_i)$ соответственно.

**Теорема 14.5** (Без доказательства)**.** *Если $\xi$ – абсолютно непрерывная случайная величина, имеющая функцию распределения $F_\xi(x)$ и плотность $f\xi(x)$, и если $g : \mathbb{R} \to \mathbb{R}$ – кусочно дифференцируемая и кусочно монотонная функция на интервалах $\mathcal{D}$, то случайная величина $\eta = g(\xi)$ имеет плотность вероятности*

$$f_\eta(y) = \sum_{i=1}^{n} f_\xi \left( g|_{D_i}{}^{-1}(y) \right) \cdot \left| \frac{\mathrm{d}\left[ g|_{D_i}{}^{-1}(y) \right]}{\mathrm{d}y} \cdot \mathbf{1}_{g(D_i)}(y) \right|$$

# 15. Случайные векторы и их распределения

# 16. Функции от двух случайных величин

**Замечание 16.1.** Пусть $\xi_1, \xi_2$ – случайные величины с совместной плотностью вероятности $f_{\xi_1}, \xi_2(x_1, x_2)$. Построим с помощью функции $g : \mathbb{R}^2 \to \mathbb{R}$ случайную величину $\eta = g(\xi_1, \xi_2)$. Требуется найти функцию распределения $F_\eta(y)$ и плотность вероятности $f_\eta(y)$ случайной величины $\eta$.

**Лемма 16.2.** *Пусть $y \in \mathbb{R}$, и область $D_y \subseteq \mathbb{R}^2$ состоит из точек $(x_1, x_2)$, удовлетворяющих неравенству $g(x_1, x_2) \leq y$, т.е. $D_y = \{(x_1, x_2) \in \mathbb{R}^2 | g(x_1, x_2) \leq y\}$. Тогда случайная величина $\eta = g(\xi_1, \xi_2)$ имеет функцию распределения*

$$F_\eta(y) = \iint\limits_{D_y} f_{\xi_1}, \xi_2(x_1, x_2) dx_1 dx_2$$

*Доказательство.*

$$F_\eta(y) = \mathsf{P}(g(x_1, x_2) \leq y) = \mathsf{P}((x_1, x_2) \in Dy) = \iint\limits_{D_y} f_{\xi_1}, \xi_2(x_1, x_2) dx_1 dx_2$$

где последнее равенство следует их того, что вероятность попадания случайного вектора $(x_1, x_2)$ в область $D_y$ равна объёму цилиндра, ограниченного сверху графиком плотности вероятности, а снизу – областью $D_y$. □

**Лемма 16.3** (Распределение суммы)**.** *Пусть $\eta = \xi_1 + \xi_2$, тогда*

$$F_\eta(y) = \mathsf{P}(\eta \leq y) = \mathsf{P}(x_1 + x_2 \leq y) = \int\limits_{-\infty}^{\infty} \left( \int\limits_{-\infty}^{y - x_1} f_{\xi_1, \xi_2}(x_1, x_2) dx_2 \right) dx_1$$

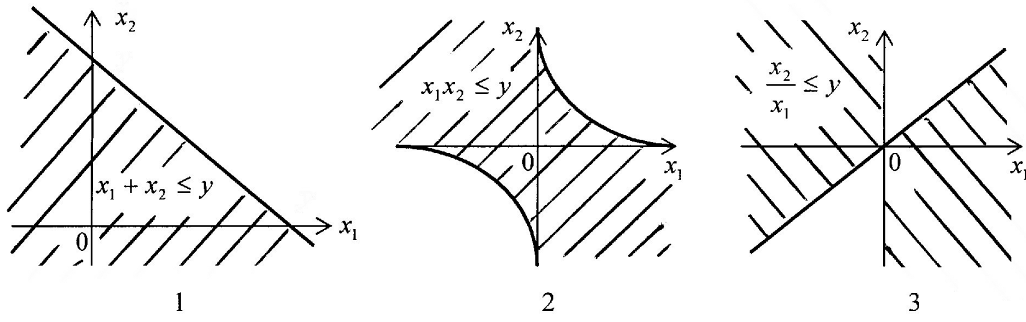


Рис. 20: Случайные функции суммы, произведения и частного

*Доказательство.* Воспользуемся леммой 16.2 для функции $g(x_1, x_2) = x_1 + x_2$.

Заменим двойной интеграл по области $D_y = \{(x_1, x_2) \in \mathbb{R}^2 | x_1 + x_2 \le y\}$ (см. рис. 20.1) на повторный, получим требуемый результат. □

**Лемма 16.4** (Распределение произведения). *Пусть* $\eta = \xi_1\xi_2$ , *тогда*

$$F_\eta(y) = \mathsf{P}(x_1x_2 \le y) = \int\limits_{-\infty}^{0} \left( \int\limits_{y/x_1}^{\infty} f_{\xi_1,\xi_2}(x_1, x_2) dx_2 \right) dx_1 +$$

$$+ \int\limits_{0}^{\infty} \left( \int\limits_{-\infty}^{y/x_1} f_{\xi_1,\xi_2}(x_1, x_2) dx_2 \right) dx_1$$

*Доказательство.* Воспользуемся леммой 16.2 для функции $g(x_1, x_2) = x_1x_2$.

Область $D_y = \{(x_1, x_2) \in \mathbb{R} x_1x_2 \le y\}$ показана на рис. 20.2. Переходя от двойного интеграла к повторному, получим требуемый результат. □

**Лемма 16.5** (Распределение частного). *Пусть* $\eta = \frac{\xi_2}{\xi_1}$, *тогда*

$$F_\eta(y) = \mathsf{P}(\frac{x_2}{x_1} \le y) = \int\limits_{-\infty}^{0} \left( \int\limits_{yx_1}^{\infty} f_{\xi_1,\xi_2}(x_1, x_2) dx_2 \right) dx_1 +$$

$$+ \int\limits_{0}^{\infty} \left( \int\limits_{-\infty}^{yx_1} f_{\xi_1,\xi_2}(x_1, x_2) dx_2 \right) dx_1$$

*Доказательство.* Доказательство. Воспользуемся леммой 16.2 для функции $g(x_1, x_2) = \frac{x_2}{x_1}$.

Область $D_y = \{(x_1, x_2) \in \mathbb{R}^2 \frac{x_2}{x_1} \le y\}$ показана на рис. 20.3 Переходя от двойного интеграла к повторному, получим требуемый результат. □

**Лемма 16.6** (О свёртке). *Если случайные величины $\xi_1$ и $\xi_2$ независимы и имеют плотности вероятности соответственно $f_{\xi_1}(x_1)$ и $f_{\xi_2}(x_2)$, то плотность вероятности суммы $\eta = \xi_1 + \xi_2$ равна свёртке плотностей $f_{\xi_1}$ и $f_{\xi_2}$:*

$$f_\eta(t) = f_{\xi_1+\xi_2}(t) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(u) f_{\xi_2}(t-u)du = \int\limits_{-\infty}^{\infty} f_{\xi_2}(u) f_{\xi_1}(t-u)du$$

*Доказательство.* Воспользуемся леммой 16.3

$$F_\eta(y) = \int\limits_{-\infty}^{\infty} \left( \int\limits_{-\infty}^{y-x_1} f_{\xi_1,\xi_2}(x_1,x_2)dx_2 \right) dx_1 = \int\limits_{-\infty}^{\infty} \left( \int\limits_{-\infty}^{y-x_1} f_{\xi_1}(x_1) f_{\xi_2}(x_2)dx_2 \right) dx_1$$

Сделаем в последнем интеграле замену переменной: $x_2 = t - x_1$ , $dx_2 = dt$. При этом $x_2 \in (-\infty, x - x_1)$ перейдёт в $t \in (-\infty, x)$. В последнем интеграле меняем порядок интегрирования и получаем:

$$\int\limits_{-\infty}^{\infty} \left( \int\limits_{-\infty}^{x} f_{\xi_1}(x_1) f_{\xi_2}(t-x_1)dt \right) dx_1 = \int\limits_{-\infty}^{x} \overbrace{\left( \int\limits_{-\infty}^{\infty} f_{\xi_1}(x_1) f_{\xi_2}(t-x_1)dx_1 \right)}^{f_{\xi_1+\xi_2}(t)} dt =$$

Итак, мы представили функцию распределения $F_\eta(x) = F_{\xi_1+\xi_2}(x)$ в виде $\int\limits_{-\infty}^{x} f_{\xi_1+\xi_2}(t)dt$, где

$$f_{\xi_1+\xi_2}(t) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(x_1) f_{\xi_2}(t-x_1)dx_1 = \int\limits_{-\infty}^{\infty} f_{\xi_1}(u) f_{\xi_2}(t-u)du$$

Второе равенство теоремы получается из первого, если всюду в доказательстве поменять местами индексы 1 и 2. □

**Определение 16.7.** Если две независимые случайные величины имеют одно и то же распределение (возможно, с разными параметрами), и их сумма имеет то же самое распределение, то распределение называется устойчивым относительно суммирования.

**Замечание 16.8.** Следующие леммы могут быть доказаны непосредственно с помощью формулы свёртки, однако более короткие доказательства будут приведены в параграфе 25, используя характеристические функции.

**Лемма 16.9.** *Пусть независимые случайные величины $\xi_n$ и $\xi_m$ имеют биномиальное распределение (см. опред 12.8), с параметрами n и m и вероятностью "успеха"p, тогда случайная величина $\xi_n + \xi_m$ имеет биномиальное распределение $\xi_{n+m}$ . Другими словами, биномиальное распределение является устойчивым относительно суммирования.*

**Лемма 16.10.** *Пусть независимые случайные величины $\xi$ и $\eta$ имеют нормальное распределение (см. опред 12.8), и имеют соответственно математические ожидания $a_1$ и $a_2$ и дисперсии $\sigma_1^2$ и $\sigma_2^2$, тогда случайная величина $\xi+\eta$ имеет нормальное распределение с математическим ожиданием $a_1+a_2$ и дисперсией $\sigma_1^2+\sigma_2^2$. Другими словами, нормальное распределение является устойчивым относительно суммирования.*

**Пример**. Товарняк состоит из 100 вагонов. Масса (измеряемая в тоннах т) произвольного вагона – случайная величина, распределенная по нормальному закону с средним значением 65 тонн и среднеквадратичным отклонением 9 тонн. Локомотив может везти состав массой не более 6600 тонн, в противном случае необходимо прицеплять второй локомотив. Найти вероятность того, что второй локомотив не потребуется.

*Решение.* По лемме 16.10 имеем $a = a_1 + \ldots + a_100 = 65 \ldots 100 = 6500$, $\sigma = \sigma_1^2 + \ldots + \sigma_1 00^2 = 81 \cdot 100 = 8100$.

Вероятность того, что второй локомотив не потребуется, составляет $\mathsf{P}(\xi \leq 6600) = F_\xi(6600) = \frac{1}{2} + \Phi(\frac{x-a}{\sigma}) = \frac{1}{2} + \Phi(\frac{6600-6500}{90}) = \frac{1}{2} + \Phi(\frac{10}{9}) \approx \frac{1}{2} + \Phi(1.11) \approx 0.87$, где $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int\limits_0^z e^{-t^2/2} dt$ – табличный интеграл вероятности нормального распределения.

# 17. Математическое ожидание

# 18. Дисперсия

**Определение 18.1.** 1) Центральным моментом $n$-го порядка случайной величины $\xi$ называется число

$$\mathsf{M}(\xi - \mathsf{M}\xi)^n = \int_{-\infty}^{\infty} (x - \mathsf{M}\xi)^n f_\xi(x) dx$$

2) Дисперсией случайной величины $\xi$ называется центральный момент 2-го порядка, т.е. число

$$\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \int_{-\infty}^{\infty} (x - \mathsf{M}\xi)^2 f_\xi(x) dx.$$

**Замечание 18.2.** Для любой случайной величины $\xi$ её центральный момент 1-го порядка равен нулю, $\mathsf{M}(\xi - \mathsf{M}\xi) = \mathsf{M}\xi - \mathsf{M}\xi = 0$.

**Замечание 18.3.** Если ось $x$ рассматривать как тонкую металлическую проволоку с линейной плотностью массы $f_\xi(x)$, то дисперсия $\mathsf{D}\xi$ есть в точности момент инерции стержня относительно его центра тяжести $\mathsf{M}\xi$.

**Лемма 18.4.** *1) Для любой случайной величины $\xi$ имеем $\mathsf{D}\xi \geq 0$.*

*2) Если $C = const$, то $\mathsf{D}C = 0$.*

*3) Если $C = const$, то для любой случайной величины $\xi$ имеем $\mathsf{D}(C\xi) = C^2\mathsf{D}\xi$.*

*Доказательство.* 1) Т.к. $(x - \mathsf{M}\xi)^2 f_\xi(x)(x) \geq 0$, то интеграл в опред. 18.1.2) не меньше нуля.

2) $\mathsf{D}C = \mathsf{M}(C - \mathsf{M}C)^2 = 0$.

3) $\mathsf{D}(C\xi) = \mathsf{M}[C\xi - \mathsf{M}(C\xi)]^2 = \mathsf{M}[C(\xi - \mathsf{M}\xi)]^2 = C^2\mathsf{M}(\xi - \mathsf{M}\xi)^2 = C^2\mathsf{D}\xi$.

□

**Замечание 18.5.** Дисперсия (по прочтению формулы $\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2$) равна «среднему значению квадрата отклонения случайной величины от её среднего». То есть, дисперсия характеризует отклонение (разброс, кучность) значений случайной величины вокруг её среднего значения $\mathsf{M}\xi$. В таком случае полезно следующее определение.

**Определение 18.6.** Число $\sigma = \sqrt{\mathsf{D}\xi}$ называется средне квадратическим отклонением случайной величины $\xi$.

**Теорема 18.7.** *Для любой случайной величины $\xi$ справедлива формула*

$$\mathsf{D}\xi = \mathsf{M}\xi^2 - (\mathsf{M}\xi)^2.$$

*Доказательство.*

$$\begin{aligned} \mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \mathsf{M}[\xi^2 - 2\xi\mathsf{M}\xi + (\mathsf{M}\xi)^2] = \\ = \mathsf{M}\xi^2 - 2M(\xi\mathsf{M}\xi) + \mathsf{M}[(\mathsf{M}\xi)^2] = \\ = \mathsf{M}\xi^2 - 2(\mathsf{M}\xi)^2 + (\mathsf{M}\xi)^2 = \mathsf{M}\xi^2 - (\mathsf{M}\xi)^2. \end{aligned}$$

□

**Теорема 18.8.** *Если случайные величины $\xi_1$ и $\xi_2$ независимы, то*

$$\mathsf{D}(\xi_1 + \xi_2) = \mathsf{D}\xi_1 + \mathsf{D}\xi_2.$$

*Доказательство.*

$$\begin{aligned} \mathsf{D}(\xi_1 + \xi_2) = \mathsf{M}(\xi_1 + \xi_2)^2 - (\mathsf{M}(\xi_1 + \xi_2))^2 = \\ = \mathsf{M}(\xi_1^2 + 2\xi_1\xi_2 + \xi_2^2) - (\mathsf{M}\xi_1 + \mathsf{M}\xi_2)^2 \overset{17.4.5)}{=} \\ = \mathsf{M}\xi_1^2 + 2\mathsf{M}(\xi_1\xi_2) + \mathsf{M}\xi_2^2 - (\mathsf{M}\xi_1)^2 - 2\mathsf{M}\xi_1\mathsf{M}\xi_2 - (\mathsf{M}\xi_2)^2 \overset{17.4.6)}{=} \\ = \mathsf{M}\xi_1^2 - (\mathsf{M}\xi_1)^2 + \mathsf{M}\xi_2^2 - (\mathsf{M}\xi_2)^2 = \mathsf{D}\xi_1 + \mathsf{D}\xi_2. \end{aligned}$$

□

**Пример 18.9.** 1) Пусть $\xi$ — подчиняется распределению Бернулли, т.е. принимает два значения: 1 с вероятностью $p$ и 0 с вероятностью $q = 1 - p$, тогда

$$\mathsf{M}\xi = 1 \cdot p + 0 \cdot q = p, \quad \mathsf{M}\xi^2 = 1^2 \cdot p + 0^2 \cdot q = p,$$
$$\mathsf{D}\xi = \mathsf{M}\xi^2 - (\mathsf{M}\xi)^2 = p - p^2 = pq.$$

2) Вычислим математическое ожидание и дисперсию случайной величины $\xi$, имеющей биномиальное распределение, т.е. $\xi$ принимает значения $0, 1, 2, \ldots, n$ соответственно с вероятностями $\mathsf{P}(\xi_n = k) = C_n^k p^k (1-p)^{n-k}$. Возьмем $n$ независимых случайных величин $\xi_1, \ldots, \xi_n$ имеющих одно и то же распределение Бернулли, т.е. с $\mathsf{M}\xi_i = p$ и $\mathsf{D}\xi_i = pq$ . Тогда их сумма $\xi = \xi_1 + \ldots + \xi_n$ имеет биномиальное распределение и

$$\mathsf{M}\xi = \sum_{i=1}^{n} \mathsf{M}\xi_i = n\mathsf{M}\xi_1 = np, \quad \mathsf{D}\xi = \mathsf{D}\xi_i = n\mathsf{D}\xi_1 = npq.$$

## 19. Числовые характеристики зависимости случайных величин

# Часть III
# Законы больших чисел

## 20. Неравенство Бьенеме–Чебышёва и неравенство Маркова

Напомним, что мы рассматриваем случайные величины, имеющие конечные начальные и центральные моменты до $n$-го порядка включительно. Метод доказательства неравенств, изучаемых в этом параграфе, принадлежит Чебышёву[16].

**Теорема 20.1** (Неравенство Маркова[17] , 1913 г.). *Для любой случайной величины $\xi$ и для любого $\varepsilon \in (0, \infty)$, имеет место неравенство*

[16]Чебышёв, Пафнутий Львович (распространено неправильное произношение его фамилии с ударением на первый слог) (1821 — 1894), выдающийся русский математик и механик, внёсший большой вклад в теорию вероятностей, теорию приближений, теорию интерполирования функций, интегральное исчисление и картографию. Работая на «оборонку», он улучшил дальнобойность и точность артиллерийской стрельбы, чем оказал большое влияние на развитие русской артиллерии.

[17]Марков, Андрей Андреевич (1856 — 1922), выдающийся русский математик, внёсший большой вклад в теорию вероятностей и математический анализ.

$$\mathsf{P}(|\xi| \geq \varepsilon) \leq \frac{\mathsf{M}|\xi|}{\varepsilon}$$

*Доказательство.* Если $c$ – плотность случайной величины $\xi$, то

$$\mathsf{P}(|\xi| \geq \varepsilon) = \int\limits_{|x| \geq \varepsilon} f_\xi(t)dt$$

Ясно, что для всех $t \in \{|x| \geq \varepsilon\}$ выполнено неравенство $\frac{|t|}{\varepsilon} \geq 1$. Поэтому при замене $f_\xi(t)$ на $\frac{|t|}{\varepsilon} f_\xi(t)$ подынтегральное выражение не уменьшится, т.е.

$$\mathsf{P}(|\xi| \geq \varepsilon) = \int\limits_{|x| \geq \varepsilon} f_\xi(t)dt \leq \int\limits_{|x| \geq \varepsilon} \frac{|t|}{\varepsilon} f_\xi(t)dt = \frac{1}{\varepsilon} \int\limits_{|x| \geq \varepsilon} |t| f_\xi(t)dt.$$

Если теперь мы увеличим область интегрирования с $|x| \geq \varepsilon$ до $(-\infty, \infty)$, то интеграл справа тоже не уменьшится. Окончательно получаем

$$\mathsf{P}(|\xi| \geq \varepsilon) \leq \frac{1}{\varepsilon} \int\limits_{|x| \geq \varepsilon} |t| f_\xi(t)dt \leq \frac{1}{\varepsilon} \int\limits_{-\infty}^{\infty} |t| f_\xi(t)dt = \frac{\mathsf{M}|\xi|}{\varepsilon} \varepsilon$$

□

**Следствие 20.2** (Двойственное неравенство Маркова)**.** Для любой случайной величины $\xi$ и для любого $\varepsilon \in (0, \infty)$ имеет место неравенство

$$\mathsf{P}(|\xi| < \varepsilon) \geq 1 - \frac{\mathsf{M}|\xi|}{\varepsilon}$$

*Доказательство.* По лемме 3.7 имеем $\mathsf{P}(|\xi| < \varepsilon) = 1 - \mathsf{P}(|\xi| \geq \varepsilon)$. Подставим это выражение в неравенство Маркова, получим требуемый результат. □

**Теорема 20.3** (Обобщённое неравенство Маркова)**.** *Для любой случайной величины $\xi$, для любого $\varepsilon \in (0, \infty)$ и любой монотонно возрастающей функции $g : (0, \infty) \to (0, \infty)$ имеет место неравенство* $\mathsf{P}(|\xi| \geq \varepsilon) \leq \frac{\mathsf{M}g(|\xi|)}{g(\varepsilon)}$

*Доказательство.* Поскольку функция $g$ монотонно возрастает, то $\mathsf{P}(|\xi| \geq \varepsilon) = \mathsf{P}(g(|\xi|) \geq g(\varepsilon))$. Оценивая последнюю вероятность согласно неравенству Маркова, получим требуемый результат:

$$\mathsf{P}(g(|\xi|) \geq g(\varepsilon)) \leq \frac{\mathsf{M}g(|\xi|)}{g(\varepsilon)}$$

□

**Следствие 20.4.** Для любой случайной величины $\xi$, для любого $\varepsilon \in (0, \infty)$, и любой монотонно возрастающей функции $g : (0, \infty) \to (0, \infty)$ имеет место неравенство

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geq) \leq \frac{\mathsf{D}\xi}{\varepsilon^2}$$

(двойственное к обобщённому неравенству Маркова). Доказательство аналогично доказательству след. 20.2. В 1853 г. И.-Ж. Бьенеме[18] и в 1866 г. независимо от него П.Л. Чебышёв доказали следующее неравенство.

**Теорема 20.5** (Неравенство Бьенеме – Чебышёва)**.** *Для любой случайной величины $\xi$ и любого $\varepsilon \in (0, \infty)$ имеет место неравенство*

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geq \varepsilon) \leq \frac{\mathsf{D}\xi}{\varepsilon^2}$$

*Доказательство.* Т.к. неравенство Маркова 20.1 справедливо для произвольной случайной величины, то перепишем его для случайной величины $\xi - \mathsf{M}\xi$, получим

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geq \varepsilon) \leq \frac{\mathsf{M}|\xi - \mathsf{M}\xi|^2}{\varepsilon}$$

Зададим функцию $g : (0, \infty) \to (0, \infty)$ по формуле $g(\varepsilon) = \varepsilon^2$. Видно, что она является монотонно возрастающей и поэтому удовлетворяет условиям теор. 20.3, и следовательно, для неё выполнено обобщённое неравенство Маркова

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geq \varepsilon) \leq \frac{\mathsf{M}(\xi - \mathsf{M}\xi)^2}{\varepsilon^2} = \frac{\mathsf{D}\xi}{\varepsilon^2}$$

□

**Следствие 20.6.** Для любой случайной величины $\xi$ и любого $\varepsilon \in (0, \infty)$, имеет место неравенство

$$\mathsf{P}(|\xi - \mathsf{M}\xi| < \varepsilon) \geq 1 - \frac{\mathsf{D}\xi}{\varepsilon^2}$$

(двойственное к неравенству Бьенеме – Чебышёва). Мы будем называть двойственное неравенство тоже неравенством Бьенеме – Чебышёва.

**Следствие 20.7.** Для любой случайной величины $\xi$ вероятность того, что она примет значение, отличающееся от её среднего $\mathsf{M}\xi$ более чем на три корня из её дисперсии, не превосходит $\frac{1}{9}$, т.е.

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geq 3\sqrt{\mathsf{D}\xi}) \leq \frac{1}{9}$$

[18] И.-Ж. Бьенеме (Irenee-Jules Biemayme, 1796 — 1878), французский математик, основные работы по теории вероятностей и математической статистике.

*Доказательство.* Из неравенства Бьенеме-Чебышёва имеем

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geq 3\sqrt{\mathsf{D}\xi}) \leq \frac{\mathsf{D}\xi}{(3\sqrt{\mathsf{D}\xi})^2} = \frac{1}{9}$$

□

# 21. Последовательности случайных величин

# 22. Законы больших чисел

Одним из главных разделов теории вероятностей составляют так называемые «предельные теоремы». Они описывают условия возникновения вероятностных закономерностей при наличии большого числа случайных факторов.

Предельные теоремы можно условно разделить на три типа: 1) закон больших чисел, 2) предельные теоремы биномиального распределения и 3) центральные предельные теоремы.

Этот и следующие параграфы посвящены изучению таких теорем.

**Определение 22.1.** Говорят, что последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ удовлетворяет закону больших чисел (ЗБЧ), если для любого $\varepsilon \in (0, \infty)$ (при некоторых дополнительных условиях) имеют место эквивалентные друг другу равенства

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}\mathsf{M}\xi_k\right| < \varepsilon\right) = 1 \tag{14}$$

и

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}\mathsf{M}\xi_k\right| \geq \varepsilon\right) = 0 \tag{15}$$

Другими словами, среднее арифметическое первых $n$ членов последовательности случайных величин сходится по вероятности к среднему арифметическому математических ожиданий её первых $n$ членов при $n \to \infty$, или в символической форме:

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow{\mathsf{P}} \frac{1}{n}\sum_{k=1}^{n}\mathsf{M}\xi_k \text{ при } n \to \infty$$

**Замечание 22.2.** Закон больших чисел в обобщённом смысле – это совокупность теорем о том, что при некоторых условиях имеют место формулы типа (14) и (15) и их многочисленные обобщения.

Т.к. каждая такая теорема представляет собой конкретный закон больших чисел, то имеет смысл говорить о законах больших чисел.

Широкие условия применимости ЗБЧ были найдены впервые Чебышёвым в 1867 г. Эти условия затем были обобщены А.А. Марковым (старшим).

Окончательное решение проблемы о необходимых и достаточных условиях применимости ЗБЧ было найдено А.Н. Колмогоровым в 1928 г. В этом параграфе мы изучим три наиболее известных таких закона.

**Теорема 22.3** (ЗБЧ в форме Чебышёва). *Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ – последовательность попарно независимых случайных величин, имеющих дисперсии, ограниченные одной и той же константой $C$, т.е.*

$$\mathsf{D}\xi_1 \leq C, \quad \mathsf{D}\xi_2 \leq C, \quad \ldots, \quad \mathsf{D}\xi_n \leq C, \quad \ldots$$

*Тогда для любого $\varepsilon \in (0, \infty)$ при $n \to \infty$*

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k \xrightarrow{\mathsf{P}} \frac{1}{n}\sum_{k=1}^{n} \mathsf{M}\xi_k$$

*Доказательство.* По 18.4.3) и 18.8 имеем

$$\mathsf{D}\left(\frac{1}{n}\sum_{k=1}^{n} \xi_k\right) = \frac{1}{n^2}\mathsf{D}\sum_{k=1}^{n} \xi_k = \frac{1}{n^2}\sum_{k=1}^{n} \mathsf{D}\xi_k \leq \frac{1}{n^2}Cn = \frac{C}{n}$$

По след. 22.6

$$\mathsf{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} \mathsf{M}\xi_k\right| < \varepsilon\right) \geq 1 - \frac{\mathsf{D}\left(\frac{1}{n}\sum\limits_{k=1}^{n} \xi_k\right)}{\varepsilon^2} \geq 1 - \frac{C}{n\varepsilon^2}$$

Переходя к пределу при $n \to \infty$ получим неравенство

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} \mathsf{M}\xi_k\right| < \varepsilon\right) \geq 1$$

откуда по лемме 3.6 получаем требуемый результат. □

**Замечание 22.4.** Из доказательства теоремы 22.3 следует, что двойственное неравенство Бьенеме – Чебышёва можно записать в виде:

$$\mathsf{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} \mathsf{M}\xi_k\right| \geq \varepsilon\right) \leq \frac{\mathsf{D}\left(\frac{1}{n}\sum\limits_{k=1}^{n} \xi_k\right)}{\varepsilon^2} \tag{16}$$

**Теорема 22.5** (ЗБЧ в форме Хинчина[19] (1929)). *Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ – последовательность попарно независимых и одинаково распределённых случайных величин, то есть*

$$\mathsf{M}\xi_1 = \mathsf{M}\xi_2 = \ldots = \mathsf{M}\xi_n = \ldots \quad и \quad \mathsf{D}\xi_1 = \mathsf{D}\xi_2 = \ldots = \mathsf{D}\xi_n = \ldots$$

*Тогда среднее арифметическое $\frac{1}{n}\sum\limits_{k=1}^{n}\xi_k$ сходится по вероятности к $\mathsf{M}\xi_1$ , другими словами*

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow{\mathsf{P}} \mathsf{M}\xi_1$$

*При этом для любого $\varepsilon \in (0, \infty)$ неравенство Бьенеме – Чебышёва можно записать в виде:*

$$\mathsf{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \mathsf{M}\xi_1\right| \geq \varepsilon\right) \leq \frac{\mathsf{D}\xi_1}{n\varepsilon^2}$$

*Доказательство.* Для доказательства достаточно подставить $\frac{1}{n}\sum\limits_{k=1}^{n}\mathsf{M}\xi_k$ в ЗБЧ в форме Чебышёва и в неравенство (16). □

**Теорема 22.6** (ЗБЧ в форме Бернулли (1713)). *Пусть $A$ – событие, которое может произойти в любом из $n$ независимых испытаний с одной и той же вероятностью $p$. Пусть $n(A)$ – число появлений события $A$ в этих $n$ испытаниях. Тогда относительная частота $\frac{n(A)}{n}$ сходится по вероятности к вероятности $p$, т.е.*

$$\frac{n(A)}{n} \xrightarrow{\mathsf{P}} \mathsf{M}\xi_1$$

*При этом для любого $\varepsilon \in (0, \infty)$ неравенство Бьенеме – Чебышёва можно записать в виде:*

$$\mathsf{P}\left(\left|\frac{n(A)}{n} - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2}$$

*Доказательство.* Видно, что ЗБЧ в форме Бернулли удовлетворяет условиям ЗБЧ в форме Хинчина. Среднее арифметическое $\frac{1}{n}\sum\limits_{k=1}^{n}\xi_k$ нулей и единиц, выпавших при $n$ независимых испытаниях, равно частоте появления события $A$ (см. опред 4.2). В примере 18.9.1) мы нашли, что $\mathsf{M}\xi_1 = p\mathsf{D}\xi_1 = p(1-p)$. Подставляя эти величины в ЗБЧ в форме Хинчина, получим ЗБЧ в форме Бернулли. □

[19] Александр Яковлевич Хинчин (1894 – 1959), один из наиболее значимых математиков в советской школе теории вероятностей. Им получены основополагающие результаты в теории функций действительного переменного, теории чисел, теории вероятностей и статистической физике.

**Пример 22.7.** Монету подбрасывают 10 000 раз. Оценить вероятность того, что относительная частота выпадения герба отличается от классической вероятности 1/2 не менее чем на 0,01.

*Решение.* Другими словами, требуется оценить $\mathsf{P}\left(\left|\frac{n(A)}{n}-\frac{1}{2}\right|\geq 0,01\right)$ где $n=10^4$ , $n(A)=\sum\limits_{k=1}^{n}\xi_k$ – число выпадений герба при условии, что $\xi_k$ являются независимыми случайными величинами, имеющими распределение Бернулли с $p=\frac{1}{2}$ , и равные единице, если выпал герб, и нулю – в противном случае.

Применим ЗБЧ в форме Бернулли. Поскольку $\mathsf{D}\xi_1 = p(1-p) = \frac{1}{4}$, то искомая оценка сверху выглядит следующим образом:

$$\mathsf{P}\left(\left|\frac{n(A)}{n}-\frac{1}{2}\right|\geq 0,01\right)\leq\frac{\mathsf{D}\xi_1}{n\cdot 0,01^2}=\frac{1}{4\cdot 10^4\cdot 10^{-4}}=\frac{1}{4}.$$

Другими словами, неравенство Бьенеме – Чебышёва позволяет заключить, что в среднем не более чем в четверти случаев при 10 000 подбрасываниях монеты частота выпадения герба будет отличаться от 1/2 более чем на 0,01. Мы увидим, что эта оценка достаточно грубая, когда изучим так называемую «центральную предельную теорему.»

В заключение параграфа приведём без доказательства ЗБЧ в форме Пуассона.

**Теорема 22.8** (ЗБЧ в форме Пуассона[20] (1837))**.** *Пусть $A$ – событие, которое может произойти в $n$ независимых испытаниях с вероятностями $p_1,p_2,\dots,p_n$ . Пусть $n(A)$ – число появлений события $A$ в этих $n$ испытаниях.*

*Тогда $\frac{n(A)}{n}$ сходится по вероятности к среднему арифметическому вероятностей $p_k$, т.е.*

$$\frac{n(A)}{n}\xrightarrow{\mathsf{P}}\frac{1}{n}\sum_{k=1}^{n}p_k$$

# 23. Предельные теоремы для биномиального распределения

# 24. Характеристические функции

**Замечание 24.1.** Решение многих задач теории вероятностей, особенно тех, которые связаны с суммированием независимых случайных величин, удаётся

[20]Симеон-Дени Пуассон (Simeon-Denis Poisson, 1781 – 1840), знаменитый французский физик и математик. Его труды относятся к разным областям чистой математики, математической физики, теоретической и небесной механики.

получить с помощью т.н. характеристических функций. Связь между преобразованием Фурье и характеристической функцией можно описать следующим образом. Если обозначить преобразование Фурье через $\Phi$, то формулы прямого и обратного преобразования можно записать следующим образом

$$\varphi(t) = \Phi[f] = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-ixt} f(x)dx, f(x) = \Phi^{-1}[\varphi] = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{itx} \varphi(t)dt.$$

Если $\xi$ — случайная величина, имеющая плотность $f_\xi(x)$, то в терминах преобразования Фурье её характеристическая функция есть

$$\Theta_\xi(t) = \sqrt{2\pi}\Phi^{-1}[f_\xi(x)] = \int\limits_{-\infty}^{\infty} e^{itx} dx$$

**Определение 24.2.** Пусть $f_\xi(x)$ – плотность случайной величины $\xi$, тогда функция

$$\Theta_\xi(t) = \mathsf{M} e^{it\xi} = \int\limits_{-\infty}^{\infty} e^{itx} f_\xi(x)dx$$

называется характеристической функцией случайной величины $\xi$.

**Замечание 24.3.** 1) Пусть $\xi$ и $\eta$ – вещесвенные случайные величины. Составим комплексную случайную величину $\xi + i\eta$. Мы распространяем действие знака математического ожидания $\mathsf{M}$ на любую комплексную случайную величину по свойству линейности:

$$\mathsf{M}(\xi + i\eta) = \mathsf{M}\xi + i\mathsf{M}\eta.$$

2) Зная характеристическую функцию $\Theta_\xi(t)$, можно однозначно восстановить функцию распределения, а также плотность вероятности или ряд распределения случайной величины $\xi$. Например, если модуль характеристической функции $\Theta_\xi(t)$ интегрируем на всей прямой, то плотность $f_\xi(x)$ случайной величины $\xi$ находится по формуле

$$f_\xi(x) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{itx} \Theta_\xi(t)dt.$$

**Свойства характеристических функций**

**Теорема 24.4.** *Характеристическая функция $\Theta_\xi(t)$ любой случайной величины $\xi$ равномерно непрерывна. (Без доказательства.)*

**Лемма 24.5.** *1.* $\Theta_\xi(0) = 1$

*2.* $|\Theta_\xi(t)| \le 1$*для* $-\infty < t < \infty$.

*Доказательство.* 1. $\Theta_\xi(0) = \mathsf{M}e^0 = 1$.

2. $|\Theta_\xi(t)| = |\mathsf{M}e^{it\xi}| \le \mathsf{M}|e^{it\xi}| = \mathsf{M}1 = 1$.

□

**Лемма 24.6.** *Если $\eta = a\xi + b$, где $a$ и $b$ — постоянные, то*

$$\Theta_\eta(t) = \Theta_\xi(at)e^{ibt}.$$

*Доказательство.* $\Theta_\eta(t) = \mathsf{M}e^{it\eta} = \mathsf{M}e^{it(a\xi+b)} = e^{ibt}\mathsf{M}e^{iat\xi} = e^{ibt}\Theta_\xi(at)$ □

**Лемма 24.7.** *Если $\xi_1$ и $\xi_2$ – независимые случайные величины, то*

$$\Theta_{\xi_1+\xi_2}(t) = \Theta_{\xi_1}(t) \cdot \Theta_{\xi_2}(t).$$

*Доказательство.* Т.к. $\xi_1$ и $\xi_2$ – независимые величины, то $e^{it\xi_1}$ и $e^{it\xi_2}$ тоже независимы. Тогда

$$\Theta_{\xi_1+\xi_2}(t) = \mathsf{M}e^{it(\xi_1+\xi_2)} = \mathsf{M}e^{it\xi_1} \cdot e^{it\xi_2} \overset{16.5.6)}{=\!=} \mathsf{M}e^{it\xi_1} \cdot \mathsf{M}e^{it\xi_2} = \Theta_{\xi_1}(t) \cdot \Theta_{\xi_2}(t).$$

□

**Следствие 24.8.** Если $\xi_1, \ldots, \xi_n$ — независимые случайные величины, то

$$\Theta_{\xi_1+\ldots+\xi_2}(t) = \Theta_{\xi_1}(t) \cdot \ldots \cdot \Theta_{\xi_n}(t).$$

**Лемма 24.9.** $\Theta_\xi(-t) = \overline{\Theta_\xi(t)}$, *где черта означает комплексное сопряжение.*

*Доказательство.* $\Theta_\xi(-t) = \mathsf{M}e^{-it\xi} = \mathsf{M}\overline{e^{it\xi}} = \overline{\mathsf{M}e^{it\xi}} = \overline{\Theta_\xi(t)}$. □

**Лемма 24.10.** *Если $\xi$ — дискретная случайная величина, заданная рядом распределения (конечным или бесконечным)*

| $\xi$ | $x_1$ | $x_2$ | $\ldots$ | $x_k$ |
|---|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | $\ldots$ | $p_k$ |

*то* $\Theta_\xi(t) = \sum_k e^{itx_k}p_k$

*Доказательство.* Это утверждение непосредственно следует из теор. 17.4.1). □

**Лемма 24.11.** *Пусть существует начальный момент $n$-го порядка, $n = 1, 2, \ldots$, случайной величины $\xi$, т.е. $\mathsf{M}|\xi|^n < \infty$. Тогда её характеристическая функция $\Theta_\xi(t)$ является $n$ раз непрерывно дифференцируемой, и её $n$-я производная в нуле равна*

$$\Theta_\xi^{(n)}(0) = i^n\mathsf{M}\xi^n$$

*Доказательство.* Заметим сначала, что т.к. существует начальный момент $M\xi^n$ , то существуют все моменты $M\xi^k$ при $k < n$.

По лемме 24.5.2) имеем $|\Theta_\xi(t)| \leq 1$, поэтому интеграл в правой части

$$\Theta_\xi(t) = \int\limits_{-\infty}^{\infty} e^{itx} f_\xi(x) dx$$

равномерно сходится по параметру $t$ и следовательно его можно дифференцировать по $t$:

$$\Theta'_\xi(t) = i \int\limits_{-\infty}^{\infty} x e^{itx} f_\xi(x) dx,$$

$$\Theta''_\xi(t) = i^2 \int\limits_{-\infty}^{\infty} x^2 e^{itx} f_\xi(x) dx,$$

$$\dots\dots\dots\dots\dots\dots\dots$$

$$\Theta_\xi^(n)(t) = i^n \int\limits_{-\infty}^{\infty} x^n e^{itx} f_\xi(x) dx.$$

Подставляя в эти равенства $t = 0$, получим требуемый результат: $\Theta'_\xi(0) = i\mathsf{M}\xi, \Theta''_\xi(0) = i^2\mathsf{M}\xi^2, \dots, \Theta_\xi^{(n)}(0) = i^n\mathsf{M}\xi^n$. □

**Лемма 24.12.** *Если существуют начальные моменты* $\mathsf{M}\xi, \mathsf{M}\xi^2, \dots, \mathsf{M}\xi^n$ *случайной величины* $\xi$*, т.е.* $M|\xi|^n < \infty$*, тогда в окрестности нуля её характеристическая функция* $\Theta_\xi(t)$ *разлагается в ряд Тейлора*[21]

$$\Theta_\xi(t) = 1 + \frac{i\mathsf{M}\xi}{1!}t + \frac{i^2\mathsf{M}\xi^2}{2!}t^2 + \dots + \frac{i^n\mathsf{M}\xi^n}{n!}t^n + o(t^n)$$

*Доказательство.* В стандартное разложение функции $\Theta_\xi(t)$ в ряд Тейлора

$$\Theta_\xi(t) = 1 + \frac{\Theta'_\xi(0)}{1!}t + \frac{\Theta''_\xi(0)}{2!}t^2 + \dots + \frac{\Theta_\xi(n)(0)}{n!}t^n + \mathrm{o}(t^n)$$

подставим выражения для $\Theta'_\xi(0), \Theta''_\xi(0), \dots, \Theta_\xi(0)$, доставляемые теор. 24.11, получим требуемый результат. □

В заключение параграфа сформулируем без доказательства теорему Леви, которая устанавливает «непрерывное» соответствие между слабо сходящимися последовательностями случайных величин и поточечно сходящимися

[21] Брук Тэйлор (Brook Taylor, 1685 — 1731), английский математик, именем которого назван ряд (опубликованный им в 1715—1717 гг.), однако этот ряд был известен и применялся ещё в XVII веке Грегори и Ньютоном.

последовательностями характеристических функций. «Непрерывность» этого соответствия состоит в том, что пределу при слабой сходимости соответствует предел при поточечной сходимости и наоборот.

**Теорема 24.13** (Леви[22] о непрерывном соответствии)**.** *Последовательность случайных величин* $\{\xi_n\}_{n=1}^{\infty}$ *слабо сходится к случайной величине* $\xi$ *тогда и только тогда, когда последовательность их характеристических функций* $\{\Theta_{\xi_n}(t)\}_{n=1}^{\infty}$ *поточечно сходится к характеристической функции* $\Theta_{\xi}(t)$.

# 25. Вычисление характеристических функций

# 26. Центральная предельная теорема

# 27. Сферическое, $\xi^2$-распределение и распределение Стьюдента

# 28. Цепи Маркова

[22] Поль Пьер Леви (Paul Pierre Levy, 1886 — 1971), выдающийся французский математик, основные труды по теории вероятностей, функциональному анализу, теории функций и механике.