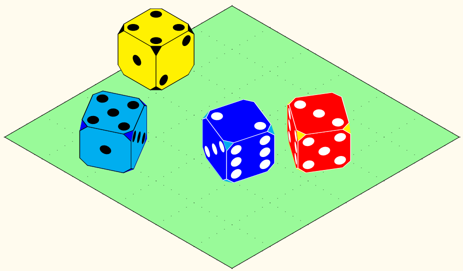
Рязанский Государственный Университет имени С.А. Есенина
Кафедра математики и методики преподавания математических дисциплин

Теория вероятностей, комбинаторика и начала математической статистики

Лектор —
профессор А. Х. Назиев



Содержание

dma@rspu.ryazan.ru

Последняя ревизия: 8 июня 2010 г.

Оглавление

Часть 1. ТЕОРИЯ ВЕРОЯТНОСТЕЙ

1. Предмет теории вероятностей

Теория вероятностей предназначена для изучения закономерностей, присущих случайным массовым явлениям.

Пример 1.1. Имеется мешок с шарами разных цветов. Извлекают шар и регистрируют его цвет. Возвращают шар в мешок, тщательно перемешивают шары, снова извлекают шар, регистрируют его цвет, и так далее.

Цвет извлекаемого шара — случайное массовое явление. Случайное — потому, что заранее невозможно предсказать цвет шара, который будет извлечён. Массовое — потому, что описанные действия могут повторяться неограниченное число раз.

Объясним на этом примере, какие закономерности могут обнаруживаться у случайных массовых явлений. Как мы уже сказали, цвет извлекаемого шара заранее предсказать невозможно, это явление случайное. Но это не означает, что никаких вообще закономерностей в этом явлении обнаружить невозможно. Закономерности есть, но они иного рода. Чтобы их обнаружить, введём понятие относительной частоты.

Обозначим через n общее количество извлечений шаров и через $\mu_n(\mathsf{Б})$ — количество появлений белого шара в этих n извлечениях. Назовём относительной частотой появления белого шара в серии из n испытаний отношение

$$\nu_n(\mathsf{Б}) = \frac{\mu_n(\mathsf{Б})}{n}.$$

В поведении этой величины и обнаруживается закономерность.

Поначалу, при небольших n, её колебания будут довольно значительными, но с увеличением количества извлечений эти колебания станут “затухать”. Относительная частота появления белого шара будет, начиная с некоторого момента, очень мало отличаться от некоторого своего “предельного” значения.

Такое поведение типично для случайных массовых явлений. В общих словах можно сказать, что с увеличением числа наблюдений относительная частота наступления данного события устанавливается на некотором устойчивом “предельном” значении.

Одной из целей теории вероятностей является разработка методов предсказания этой предельной относительной частоты. Покажем на приведённом примере, как могут делаться подобные предсказания.

Пусть всего в мешке N шаров, и M из них — белые. Допустим, что мы произвели 10 000 извлечений шаров. Поскольку нет никаких оснований предпочесть один шар другому, естественно ожидать, что каждый из N шаров появится примерно $\frac{10\,000}{N}$ раз. Имеется M белых шаров. Следовательно за 10 000 извлечений белый шар появится примерно $M \cdot \frac{10\,000}{N}$ раз. Чтобы получить относительную частоту, нужно разделить это число на число извлечений, то есть на 10 000. Полученное число, $\frac{M}{N}$, и есть искомое окончательное устойчивое значение относительной частоты появления белого шара в описанном случайном массовом явлении.

Истолкуем теперь этот результат иначе. Когда мы хотим извлечь один из N одинаковых на ощупь шаров, находящихся в мешке, у нас имеется N возможностей осуществить наше намерение, или, как говорят в теории вероятностей, имеется N возможных исходов. Если мы хотим извлечь именно белый шар, то M шаров “благоприятствуют” нам: имеется M благоприятных исходов. Полученное нами окончательное устойчивое значение относительной частоты появления белого шара оказывается, таки образом, отношением числа благоприятствующих исходов к общему числу возможных исходов. Именно это и взято за основу в классическом определении вероятности.

В классической схеме интересуются вероятностью наступления некоторого события в случайном массовом явлении, которое в каждом отдельном случае (“испытании”) может привести к любому одному из N возможных исходов. При этом предполагается, что нет никаких оснований предпочесть один из этих исходов другому. Коротко говорят: имеется N *равновозможных* исходов испытания. В этих условиях вероятность наступления какого-либо события определяется как отношение числа исходов, благоприятствующих наступлению этого события, к числу всех возможных исходов. Обозначив вероятность события A через $\mathbb{P}(A)$, выразим это определение следующим образом:

$$\mathbb{P}(A) = \frac{\text{число исходов, благоприятствующих событию } A}{\text{число всех равновозможных исходов}}.$$

Пример 1.2. Пусть теперь у нас — два мешка с шарами. В одном N_1 шаров, из которых M_1 белых, в другом N_2 шаров, из которых M_2 белых. Из обоих одновременно извлекается по одному шару. Какова вероятность того, что оба извлечённых шара окажутся белыми?

Поскольку любой шар из первого мешка может быть извлечён с любым шаром из второго мешка, имеется всего $N_1 N_2$ различных возможных исходов, причём нет никаких оснований предпочесть один из этих исходов другому. Таким образом всего имеется ровно $N_1 N_2$ равновозможных исходов. Поскольку любой белый шар из первого мешка может быть извлечён с любым белым шаром из второго мешка, имеется ровно $M_1 M_2$ исходов, в которых оба извлечённых шара окажутся белыми. Иными словами, имеется точно $M_1 M_2$ благоприятствующих исходов. Поэтому искомая вероятность равна

$$\frac{M_1 M_2}{N_1 N_2} = \frac{M_1}{N_1} \cdot \frac{M_2}{N_2}.$$

Но $\frac{M_1}{N_1}$ — это вероятность извлечь белый шар из первого мешка, а $\frac{M_2}{N_2}$ — вероятность извлечь белый шар из второго мешка. Таким образом, вероятность извлечь по белому шару из каждого мешка равна произведению двух вероятностей: вероятности извлечь белый шар из первого мешка и вероятности извлечь белый шар из второго мешка.

Подчеркнём: наш вывод был получен в предположении, что любой шар из первого мешка может быть извлечён с любым шаром из второго мешка, так что извлечение какого-либо шара из одного мешка никак не влияет на то, какой шар будет извлечён из другого мешка. В терминологии теории вероятностей это означает, что событие, состоящее в извлечении какого угодно шара из одного мешка, независимо от события, состоящего в извлечении какого угодно шара из другого мешка. Наш вывод означает, что *если два события независимы, то вероятность наступления того и другого события одновременно равна произведению вероятностей наступления каждого из них отдельно.* Ясно, что этот вывод может быть распространён на любое конечное число независимых событий.

Это — один из основных законов теории вероятностей, называемый законом умножения вероятностей. Далее он будет рассмотрен более подробно.

Рассмотрим одно применение этого закона.

Пример 1.3. Предположим, что вероятность сбить самолёт единичным выстрелом из винтовки равна $0,004$. Допустим, что производится залп из 500 винтовок. Какова вероятность того, что самолёт будет сбит?

Согласно условию, вероятность промаха при единичном выстреле из винтовки равна $0,996$ (почти верный промах). Вероятность того, что самолёт не будет сбит при залпе из 500 винтовок, есть вероятность одновременного наступления 500 событий, состоящих в промахе при выстреле из 1-й, 2-й, $\ldots$, 500-й винтовок. Каждое из этих событий не зависит от других. Значит, в силу закона умножения вероятностей, вероятность того, что самолёт не будет сбит, равна $(0,996)^{500}$. Вычисления показывают, что это число приблизительно равно $0,14$. Таким образом, вероятность того, что самолёт будет сбит, составляет $0,86$. Это означает, что самолёт почти наверняка будет сбит.

На этом примере хорошо видно назначение теории вероятностей. Оно состоит в том, чтобы по известным вероятностям одних событий устанавливать вероятности других событий. Теория вероятностей не скажет нам, какова вероятность поражения самолёта единичным выстрелом из винтовки — это устанавливается практически. Зато она скажет нам, какова вероятность поразить самолёт залпом из данного количества винтовок, *если* известна вероятность успеха в случае единичного выстрела. В этом отношении теория вероятностей нисколько не отличается от других математических дисциплин. Теория групп, например, не скажет нам, ассоциативна или нет данная операция, существует ли для неё нейтральный элемент и каждый ли элемент обладает обратным — это не её задача. Зато она говорит нам, какие ещё факты выполняются, если имеют место те, что берутся в качестве аксиом. Логика не скажет нам, верно или нет, что Иван любит Марью, что Марья счастлива. Зато она скажет нам, верна или нет конъюнкция

Иван любит Марью и Марья счастлива,

если известно, верны или нет её члены. И так далее.

Может показаться, что в первых двух примерах дела обстояли иначе: мы просто вычислили вероятности указанных событий, никаких предположений о вероятностях других событий не делая. На самом же деле и в этих примерах всё происходило точно так же, только оформлено это было немного иначе.

Рассматривая первый пример, мы предполагали, что в мешке находится n шаров и нет никаких оснований предпочесть один шар другому. Тем самым мы неявно предполагали, что для каждого из шаров, находящихся в мешке, вероятность быть извлечённым равна $\frac{1}{N}$.

Точно так же при рассмотрении второго примера неявно предполагалось, что для любой пары шаров, взятых по одному из каждого мешка, вероятность быть извлечённой равна $\frac{1}{N_1N_2}$. И так — в любой задаче, решаемой теорией вероятностей.

Теория вероятностей — это именно *теория*, родственная, скажем, геометрии или теоретической механике. В геометрии не занимаются выяснением "истинной природы" точек и расстояний, а в теоретической механике — "истинной природы" масс и скоростей. Так и в теории вероятностей не занимаются выяснением “истинной природы” событий и вероятностей. Их предполагают данными вместе с их первоначальными свойствами, из которых чисто математически выводят дальнейшие свойства.

2. Основные понятия теории вероятностей

Каждая математическая теория имеет дело с отвлечёнными идеализированными понятиями. Основными понятиями теории вероятностей являются понятия *события* и *вероятности*. С интуитивной точки зрения событие — это любое явление, о котором имеет смысл говорить, что оно происходит, или наступает, а вероятность — это мера осуществимости события.

2.1. События

Различают события *составные* (разложимые) и *простые* (неразложимые).

Неразложимые события называют также *элементарными*, или *атомарными*, или ещё *исходами*.

Множество всех элементарных событий называют *пространством элементарных событий*, или *пространством исходов*. Его чаще всего обозначают через Ω.

Если Ω конечно или счётно (то есть может быть занумеровано натуральными числами), то его называют *дискретным*.

С каждым событием связано множество всех исходов, в которых оно наступает. В современной теории вероятностей принята теоретико-множественная интерпретация событий, заключающаяся в том, что

> *каждое событие <u>отождествляется</u> с множеством тех исходов, в которых оно наступает.*

Благодаря этому оказывается, что

каждое событие является <u>множеством</u> исходов.

Замечание 2.1. Чтобы по достоинству оценить принятие такого решения, следует заметить, что события являются абстрактными объектами. Мы не можем показать событие, мы можем лишь описать его.

Допустим, например, что мы подбросили игральный кубик, и он упал так, как показано на следующем рисунке:

.

Какое *событие* мы видим? Никакого! Мы видим *кубик* с четырьмя точками на верхней грани, одной точкой на передней (обращённой к нам) грани, и пятью точками на правой боковой (если смотреть от нас) грани. События мы связываем с этой картиной мысленно. И этих событий много. Например: «выпало 4 очка», «выпало чётное количество очков», «выпало более 3 очков», «выпало менее пяти очков», «выпало количество очков, десятичное представление квадрата которого оканчивается цифрой 6», и т. д. Показать эти события невозможно. Можно лишь описать их, как это только что и было сделано.

Итак, о событиях мы узнаём через их описания. Это ставит нас перед проблемой: как по двум описаниям событий выяснить, являются ли они описаниями одного и того же события, или же различных.

Снова подбросим игральный кубик и обозначим через X количество выпавших очков. Рассмотрим два события: $\langle X < 1\rangle$, $\langle X > 6\rangle$. Одинаковые это события — или различные?

По приведённым описаниям — совершенно различные (предложение $X < 1$ говорит отнюдь не то же самое, что предложение $X > 6$). А вот множества исходов, отвечающие этим событиям, тождественны (оба пустые). Это даёт нам основание утверждать, что указанные события тождественны: в описанных условиях $\langle X < 1\rangle = \langle X > 6\rangle$. Как видим, *благодаря принятию теоретико-множественной точки зрения на события мы получаем чёткие критерии тождества и различия событий.*

Замечание 2.2. Вообще говоря, не каждое множество исходов является событием (точнее, существуют вероятностные пространства, в которых не каждое множество исходов является событием), но в нашем курсе такие ситуации не встретятся.

Если в исходе a наступает событие A, то говорят, что *исход a принадлежит событию A* и пишут $a \in A$.

Существует событие, не наступающее ни в одном исходе. Его называют *невозможным событием* и обозначают через $\varnothing$.

Существует событие, наступающее в любом исходе. Его называют *достоверным событием* и обозначают через Ω.

Событие, которое не является ни достоверным, ни невозможным, называют *случайным*.

Говорят, что событие A *влечёт* событие B (коротко: $A \subset B$), если всякий раз, когда наступает событие A, наступает и событие B. Иначе говоря: каждый исход, принадлежащий A, принадлежит и B.

События A и B называют *равносильными*, или *тождественными* (коротко: $A = B$), если каждое из них влечёт другое: $A \subset B$ и $B \subset A$.

Любому событию A отвечает событие $\overline{A}$ (читается: не-A), называемое *дополнением события A*, или событием, *противоположным к A*; оно наступает тогда и только тогда, когда не наступает событие A.

Любым двум событиям A и B поставлено в соответствие событие $A + B$, называемое их *суммой*; оно наступает тогда и только тогда, когда наступает хотя бы одно из событий: A или B.

Любым двум событиям A и B поставлено в соответствие событие AB, называемое их *произведением*; оно наступает тогда и только тогда, когда наступают оба события: A и B.

Событие $A\overline{B}$, состоящее в том, что наступает событие A и не наступает событие B, называют *разностью событий A и B* и обозначают через $A-B$. Таким образом, $A-B = A\overline{B}$.

События A и B называют *несовместными*, если событие AB невозможно (то есть, не может быть так, чтобы наступили оба события: A и B).

События A и B несовместны тогда и только тогда, когда каждое из них влечёт дополнение к другому: $A \subset \overline{B}$ и $B \subset \overline{A}$.

События A и B называют *независимыми*, если наступление или ненаступление события A никак не связано с наступлением или ненаступлением события B.

Понятие независимости событий далее будет уточнено.

2.2. Вероятности

Каждому событию A поставлено в соответствие действительное число $\mathbb{P}(A)$, называемое *вероятностью события* A, так, что при этом выполняются следующие условия.

$(P0)$ **(неотрицательность)** для любого события A, $\mathbb{P}(A) \geqslant 0$;

$(P1)$ **(нормированность)** $\mathbb{P}(\Omega) = 1$ (вероятность достоверного события равна единице);

$(P2)$ **(аддитивность)** если события A и B несовместны, то $\mathbb{P}(A+B) = \mathbb{P}(A) + \mathbb{P}(B)$.

В конкретных примерах вероятностей эти свойства доказывают на основе определения вероятности, в общих рассмотрениях — используют в качестве аксиом, из которых выводят остальные свойства. К перечисленным свойствам часто добавляют ещё одно, но оно нам не понадобится, поэтому здесь не приводится.

Пространство исходов с заданной на нём вероятностью называют *вероятностным пространством.*

Теория вероятностей и её термины возникли в XVII веке. В настоящее время все они имеют «переводы» на язык теории множеств. Приведём небольшой словарик.

Теория вероятностей	**Теория множеств**
Элементарное событие, исход	Элемент пространства элементарных событий
Событие	Множество исходов (часть Ω)
Достоверное событие	Множество всех исходов (Ω)
В исходе a наступает событие A	$a \in A$ ($\subset \Omega$)
Невозможное событие	Пустое множество ($\varnothing$)
Сумма событий	Объединение множеств
Произведение событий	Пересечение множеств
Противоположное событие	Дополнение множества
Совместимые события	Множества с непустым пересечением
Несовместимые события	Непересекающиеся множества

3. Примеры вероятностных пространств

Пример 3.1. *Подбрасывание монеты.*

Монета может упасть кверху орлом или решкой, может встать на ребро или куда-нибудь закатиться. Однако мы исключим из рассмотрения эти возможности и ограничимся рассмотрением лишь двух возможных исходов: выпадение орла, выпадение решки. Если монета имеет правильную форму и изготовлена из однородного материала, а подбрасывается над ровной горизонтальной поверхностью, то естественно считать, что эти два исхода равновозможны и других возможных исходов нет. Перечисленные соображения приводят к следующему вероятностному пространству, являющемуся математической моделью описанного реального процесса.

Исходы: $O = \langle$выпал орёл$\rangle$, $P = \langle$выпала решка$\rangle$.

Пространство исходов: $\Omega = \{O,\ P\}$.

События: любые множество исходов.

Элементарная вероятность: $\mathbb{P}(O) = \mathbb{P}(P) = \frac{1}{2}$.

Вероятность: для любого события A полагаем

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega).$$

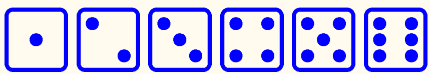
Получается, что, для любого события A,

$$\mathbb{P}(A) = m \cdot \frac{1}{n} = \frac{m}{n},$$

где $n = 2$ — количество элементов пространства Ω („число всех равновозможных случаев“), m — количество элементов множества A („число всех случаев, благоприятствующих событию A“).

Пример 3.2. *Бросание игральной кости.*

Опыт имеет ровно 6 возможных исходов:

.

Если кость изготовлена из однородного материала и симметрична, то все эти исходы равновозможны. Мы приходим к следующей модели описанного процесса.

Исходы: $\omega_i = \langle$выпало i очков$\rangle$, $i = 1,\ \ldots,\ 6$.

Пространство исходов: $\Omega = \{\omega_1,\ \omega_2,\ \ldots,\ \omega_6\}$.

События: любые множества исходов.

Элементарная вероятность:

$$\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = \ldots = \mathbb{P}(\omega_6) = \frac{1}{6}.$$

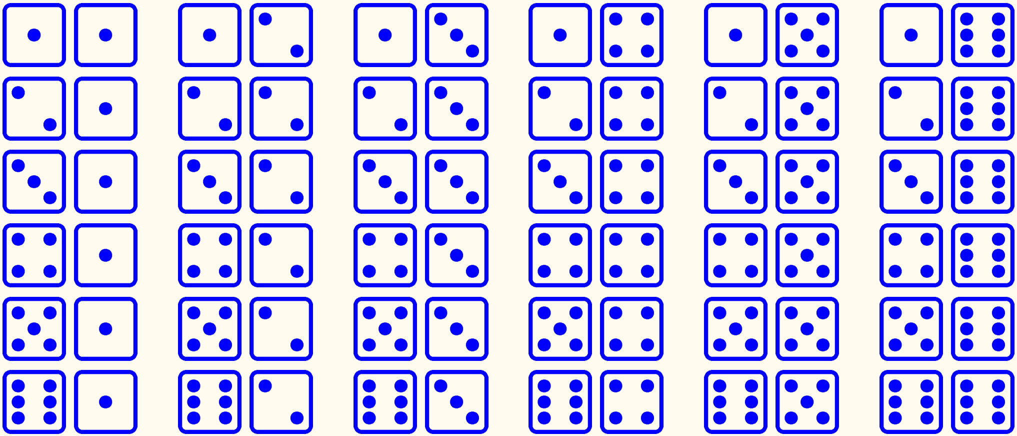
Вероятность: для любого события A полагаем

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \frac{m}{n},$$

где $n = 6$ — количество элементов пространства Ω (“число всех равновозможных случаев”), m — количество элементов множества A (“число всех случаев, благоприятствующих событию A”).

Пример 3.3. *Бросание двух отличимых игральных костей.*

Все возможные исходы представлены на следующем рисунке.



Наглядности эта картина не добавляет. Математическая модель окажется более наглядной. Для получения математической модели назовём одну из костей первой, другую — второй. Каждому исходу отвечает упорядоченная пара $(i,\ j)$, где i — количество очков, выпавшее на первой кости, j — количество очков, выпавшее на второй кости. Множеству всех исходов отвечает множество всех упорядоченных пар $(i,\ j)$ с $i,\ j = 1,\ \ldots,\ 6$, представленных в следующей таблице.

$(1,\ 1)$	$(1,\ 2)$	$(1,\ 3)$	$(1,\ 4)$	$(1,\ 5)$	$(1,\ 6)$
$(2,\ 1)$	$(2,\ 2)$	$(2,\ 3)$	$(2,\ 4)$	$(2,\ 5)$	$(2,\ 6)$
$(3,\ 1)$	$(3,\ 2)$	$(3,\ 3)$	$(3,\ 4)$	$(3,\ 5)$	$(3,\ 6)$
$(4,\ 1)$	$(4,\ 2)$	$(4,\ 3)$	$(4,\ 4)$	$(4,\ 5)$	$(4,\ 6)$
$(5,\ 1)$	$(5,\ 2)$	$(5,\ 3)$	$(5,\ 4)$	$(5,\ 5)$	$(5,\ 6)$
$(6,\ 1)$	$(6,\ 2)$	$(6,\ 3)$	$(6,\ 4)$	$(6,\ 5)$	$(6,\ 6)$

Исходы: $\omega_{ij} = \langle$выпало i очков на первой кости, j очков на второй кости$\rangle$, $i,\ j = 1,\ \ldots,\ 6$. Всего — 36 исходов. Если кости изготовлены из однородного материала и симметричны, то все эти исходы равновозможны.

Пространство исходов: $\Omega = \{\omega_{ij} : i,\ j = 1,\ \ldots,\ 6\}$.

События: любые множества исходов

Элементарная вероятность: $\mathbb{P}(\omega_{ij}) = \frac{1}{36}$.

Вероятность: для любого события A полагаем

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \frac{m}{n},$$

где $n = 36$ — количество элементов пространства Ω (“число всех равновозможных исходов”), m — количество элементов множества A (“число всех исходов, благоприятствующих наступлению события A”).

Пример 3.4. *Бросание двух неотличимых игральных костей.*

Теперь учёт всех возможных исходов без повторения осуществляется иначе. Как и в предыдущем примере, мы можем назвать одну из костей — первой, другую — второй. Однако после броска мы уже будем не в состоянии различить, какая из них — первая, какая — вторая, так что исходы $(i,\ j)$ и $(j,\ i)$ теперь тождественны. Поэтому из пар $(i,\ j)$ и $(j,\ i)$ при любых i, j, нужно оставить только одну. Выбирая ту, в которой первая компонента не больше второй, обнаруживаем, что множеству всех возможных исходов отвечает теперь множество всех упорядоченных пар $(i,\ j)$ с $i \leqslant j$, представленных в следующей таблице.

$$\begin{array}{cccccc}
(1,\ 1) & (1,\ 2) & (1,\ 3) & (1,\ 4) & (1,\ 5) & (1,\ 6) \\
 & (2,\ 2) & (2,\ 3) & (2,\ 4) & (2,\ 5) & (2,\ 6) \\
 & & (3,\ 3) & (3,\ 4) & (3,\ 5) & (3,\ 6) \\
 & & & (4,\ 4) & (4,\ 5) & (4,\ 6) \\
 & & & & (5,\ 5) & (5,\ 6) \\
 & & & & & (6,\ 6)
\end{array}$$

Если считать все эти исходы равновозможными, получится следующее вероятностное пространство.

Исходы: $\omega_{ij} = \langle$выпало i очков на одной кости, j очков на другой кости$\rangle$, $i \leqslant j,\ i,\ j = 1,\ \ldots,\ 6$. Всего — 21 элементарное событие.

Пространство исходов: $\Omega = \{\omega_{ij} : i \leqslant j,\ i,\ j = 1,\ \ldots,\ 6\}$.

События: любые множества исходов.

Вероятность: Полагая $\mathbb{P}(\omega) = \frac{1}{21}$ для любого $\omega \in \Omega$, и

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega)$$

для любого события A, получаем вероятностное пространство, устроенное аналогично предыдущим:

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

где $n = 21$ — количество элементов пространства Ω (“число всех равновозможных случаев”), m — количество элементов множества A (“число всех случаев, благоприятствующих событию A”).

Анализ модели. В полученном вероятностном пространстве вероятности выпадения 6 очков, 7 очков и 8 очков одинаковы и равны $\frac{1}{7}$, ибо каждое из указанных событий распадается на 3 элементарных события. Практика, однако, показывает, что вероятности указанных событий различны: вероятности выпадения 6 очков и выпадения 8 очков равны между собой и немного меньше, чем вероятность выпадения 7 очков.

Исторически подобное расхождение между опытом и априорными (до-опытными) рассуждениями и послужило толчком к возникновению теории вероятностей. Объяснение, данное Паскалем, состояло в следующем.

Кости, хотя и неотличимы, всё же различны: состоят из различных молекул, занимают различное положение в пространстве. Поэтому на самом деле имеют место всё те же 36 исходов, только мы не в состоянии различить их все.

Событие, состоящее в выпадении 7 очков, распадается на 6 элементарных событий из первого пространства, тогда как события, состоящие в выпадении 6 очков и 8 очков — лишь на 5, этим и объясняется различие в вероятностях.

Для получения более адекватной модели описанного реального процесса элементарным событиям из второго пространства вероятности следует приписать иначе, учитывая, что выпадение на обеих костях i очков реализуется при любом $i \in \{1, \ldots, 6\}$ лишь одним способом, выпадение же на одной кости i очков, а на другой j очков при $i \neq j$ — двумя способами.

Это приводит к другому, на этот раз не классическому, вероятностному пространству, в котором пространство исходов то же, что и во втором случае, но элементарная вероятность определена иначе:

$$\mathbb{P}(\omega_{ij}) = \begin{cases} \dfrac{1}{36}, & \text{если } i = j; \\ \dfrac{1}{18}, & \text{если } i \neq j. \end{cases}$$

Теперь событие, состоящее в выпадении 6 очков, и событие, состоящее в выпадении 8 очков, складываются из двух исходов, имеющих вероятность $\frac{1}{18}$, и одного исхода, имеющего вероятность $\frac{1}{36}$ (ω_{15}, ω_{24}, ω_{33} — для первого, ω_{26}, ω_{35}, ω_{44} — для второго) и потому наступают с вероятностью, равной $2 \cdot \frac{1}{18} + \frac{1}{36}$, то есть $\frac{5}{36}$. А вот событие, состоящее в выпадении 7 очков, наступает в исходах ω_{16}, ω_{25}, ω_{34}, каждый из которых имеет вероятность $\frac{1}{18}$, поэтому оно имеет вероятность $3 \cdot \frac{1}{18}$, то есть $\frac{1}{6}$.

Пример 3.5. Размещение трёх отличимых шаров по трём отличимым ящикам.

Назовём ящики первым, вторым и третьим, а шары — a, b и c. Содержимое первого ящика будем указывать на первом месте, второго — на втором, третьего — на третьем. Получим следующий перечень всех возможных случаев.

abc	—	—
—	abc	—
—	—	abc
ab	c	—
ab	—	c
ac	b	—
ac	—	b
bc	a	—
bc	—	a

c	ab	—
—	ab	c
b	ac	—
—	ac	b
a	bc	—
—	bc	a
c	—	ab
—	c	ab
b	—	ac

—	b	ac
a	—	bc
—	a	bc
a	b	c
a	c	b
b	a	c
b	c	a
c	a	b
c	b	a

Исходы: Перечисленные выше размещения шаров по ящикам. Всего — 27 исходов. Естественно предполагать, что все эти исходы равновозможны.

Пространство исходов: $\Omega = \{\omega_i : i = 1, \ldots, 27\}$.

События: любые множества исходов.

Элементарная вероятность: $\mathbb{P}(\omega_i) = \frac{1}{27}$.

Вероятность: для любого события A полагаем

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \frac{m}{n},$$

где $n = 27$ — количество элементов пространства Ω (“число всех равновозможных случаев”), m — количество элементов множества A (“число всех случаев, благоприятствующих событию A”).

Пример 3.6. Размещение трёх неотличимых шаров по трём отличимым ящикам.

Назовём ящики первым, вторым и третьим, шары будем обозначать «горошинкой» ‘∘’. Содержимое первого ящика будем указывать на первом месте, второго — на втором, третьего — на третьем. Получим следующий перечень всех возможных случаев.

ω_1	∘ ∘ ∘	—	—	ω_6	∘	∘ ∘	—
ω_2	—	∘ ∘ ∘	—	ω_7	—	∘ ∘	∘
ω_3	—	—	∘ ∘ ∘	ω_8	∘	—	∘ ∘
ω_4	∘ ∘	∘	—	ω_9	—	∘	∘ ∘
ω_5	∘ ∘	—	∘	ω_{10}	∘	∘	∘

Если предполагать, что все эти исходы равновозможны, придём к следующему — классическому — вероятностному пространству.

Исходы: Перечисленные выше размещения $\omega_1, \ldots, \omega_{10}$ шаров по ящикам. Всего — 10 исходов.

Пространство исходов: $\Omega = \{\omega_i : i = 1, \ldots, 10\}$.

События: любые множества исходов.

Элементарная вероятность: $\mathbb{P}(\omega_i) = \frac{1}{10}$.

Вероятность: для любого события A полагаем

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \frac{m}{n},$$

где $n = 10$ — количество элементов пространства Ω (“число всех равновозможных случаев”), m — количество элементов множества A (“число всех случаев, благоприятствующих событию A”).

Однако шары хотя и неотличимы, всё же различны (состоят из различных молекул, занимают различное положение в пространстве). Поэтому представляется разумным связать с описанным явлением другую модель, основанную на следующем соображении.

Наша неспособность отличить один шар от другого не отражается на сущности физического эксперимента, и на самом деле имеют место всё те же 27 исходов, но мы в состоянии различить только 10 из них. Поэтому событиям, считающимся элементарными в новом пространстве с 10 исходами, следует приписать те же вероятности, которые эти события имеют (не будучи элементарными) в прежнем пространстве с 27 исходами.

Основываясь на этом соображении и используя рассмотренное выше пространство с 27 исходами (обозначим его через S), приходим к следующему определению вероятности элементарных событий для нового пространства (S_1) с 10 исходами.

Каждое из событий ω_1, ω_2, ω_3 наступает лишь в одном исходе из пространства S, поэтому в пространстве S_1 каждому из этих событий следует приписать вероятность $\frac{1}{27}$.

Каждое из событий $\omega_4, \ldots, \omega_9$ наступает в трёх исходах из пространства S, поэтому в пространстве S_1 этим событиям нужно приписать вероятность $3 \cdot \frac{1}{27} = \frac{1}{9}$.

Наконец, событие, состоящее в том, что в каждом ящике оказывается по одному шару, наступает в шести исходах из пространства S, поэтому ему в пространстве S_1 нужно приписать вероятность $6 \cdot \frac{1}{27} = \frac{2}{9}$.

В результате придём к другому вероятностному пространству.

Исходы: Размещения $\omega_1, \ldots, \omega_{10}$ шаров по ящикам. Всего — 10 исходов. Эти исходы НЕ равновозможны.

Пространство исходов: $\Omega = \{\omega_i : i = 1, \ldots, 10\}$.

События: любые множества исходов.

Элементарная вероятность:

ω_i	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
p_i	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{2}{9}$

Вероятность: для любого события A полагаем

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega).$$

Полученное пространство уже не является классическим, ибо в нём вероятность события, вообще говоря, не равна отношению числа благоприятствующих случаев к числу всех равновозможных случаев.

4. Алгебра событий

В следующих теоремах собраны основные свойства операций над событиями. Имеется много других свойств. Они будут формулироваться по мере надобности. Доказательства, как правило, будут оставляться в качестве упражнений.

Теорема 4.1.
Каковы бы ни были события A, B, C:

0° $\varnothing \subset A \subset \Omega$;

1° $AB \subset A \subset A \subset A + B$;

2° *Если* $A \subset B$ *и* $B \subset C$, *то* $A \subset C$;

3° *Если* $A \subset B$ *и* $B \subset A$, *то* $A = B$.

Теорема 4.2 (Свойства сложения)**.**
Каковы бы ни были события A, B, C:

1° $A + \Omega = \Omega$;

2° $A + \varnothing = A$;

3° $A + A = A$;

4° $A + B = B + A$;

5° $A + (B + C) = (A + B) + C$.

Теорема 4.3 (Свойства умножения)**.**
Каковы бы ни были события A, B, C:

1° $A \cdot \Omega = A$;

2° $A \cdot \varnothing = \varnothing$;

3° $A \cdot A = A$;

4° $A \cdot B = B \cdot A$;

5° $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.

Теорема 4.4 (Распределительные законы)**.**
Каковы бы ни были события A, B, C:

1° $A \cdot (B + C) = (A \cdot B) + (A \cdot C)$;

2° $A + (B \cdot C) = (A + B) \cdot (A + C)$.

Теорема 4.5 (Законы поглощения)**.**
Каковы бы ни были события A и B,

$$\text{если } A \subset B, \text{ то } AB = A \text{ и } A + B = B.$$

В частности,

$$A \cdot (A + B) = A \quad \text{и} \quad AB + A = A.$$

Теорема 4.6 (Свойства дополнения)**.**
Каковы бы ни были события A, B, C:

1° $\overline{\varnothing} = \Omega$;

2° $\overline{\Omega} = \varnothing$;

3° $A \cdot \overline{A} = \varnothing$;

4° $A + \overline{A} = \Omega$;

5° $\overline{\overline{A}} = A$;

6° $\overline{A \cdot B} = \overline{A} + \overline{B}$;

7° $\overline{A + B} = \overline{A} \cdot \overline{B}$.

Упражнение 4.1. Предлагаем читателю в качестве весьма полезных и вполне посильных упражнений доказать все пункты приведённых теорем.
Совет. Вспомните основные теоретико-множественные тождества!

5. Простейшие свойства вероятности

Теорема 5.1. *Пусть Ω — пространство исходов, p — вероятность на нём и A, B, A_1, A_2, $\ldots$ — произвольные события. Тогда:*

0° $\mathbb{P}(\varnothing) = 0$;

1° $\mathbb{P}(\Omega) = 1$;

2° *если* $A \subset B$, *то* $\mathbb{P}(A) \leqslant \mathbb{P}(B)$;

3° $0 \leqslant \mathbb{P}(A) \leqslant 1$;

4° *если* $A \cdot B = \varnothing$, *то* $\mathbb{P}(A+B) = \mathbb{P}(A) + \mathbb{P}(B)$; *и вообще,*
если $A_i \cdot A_j = \varnothing$ *при* $i \neq j$, *то* $\mathbb{P}(A_1 + \ldots + A_n) = \mathbb{P}(A_1) + \ldots + \mathbb{P}(A_n)$;

5° $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.

Доказательство. Приведём доказательство для классической схемы определения вероятности.

$\boxed{2}$ Пусть $A \subset B$. Пусть всего исходов n, из них a благоприятствуют событию A, b — событию B. Поскольку $A \subset B$, все исходы, благоприятствующие наступлению события A, содержатся среди тех, которые благоприятствуют B. Значит, $a \leqslant b$, откуда

$$\mathbb{P}(A) = \frac{a}{n} \leqslant \frac{b}{n} = \mathbb{P}(B).$$

$\boxed{3}$ Для любого события A имеем: $\varnothing \subset A \subset \Omega$. Поэтому $\mathbb{P}(\varnothing) \leqslant \mathbb{P}(A) \leqslant \mathbb{P}(\Omega)$. Но $\mathbb{P}(\varnothing) = 0$, а $\mathbb{P}(\Omega) = 1$. Значит, $0 \leqslant \mathbb{P}(A) \leqslant 1$.

$\boxed{5}$ Ясно, что $A + \overline{A} = \Omega$ и $A \cdot \overline{A} = \varnothing$. Поэтому

$$\mathbb{P}(A) + \mathbb{P}(\overline{A}) = \mathbb{P}(\Omega) = 1,$$

откуда и получаем требуемое. □

В силу аддитивности вероятность суммы несовместных событий равна сумме их вероятностей. Рассмотрим теперь вопрос о вероятности суммы *произвольных* событий.

Начнём с двух событий. Итак, пусть A и B — произвольные события, совместимые или нет. Чтобы найти $\mathbb{P}(A+B)$, заметим, что событию $A+B$ благоприятствуют те и только те исходы, которые благоприятствуют: и A, и B; A, но не B; B, но не A. Иначе говоря, $A+B = AB + (A-B) + (B-A)$, причём слагаемые несовместимы.

Пусть всего исходов n, из них событию A благоприятствуют a исходов, событию B — b исходов, событию AB — c исходов. Тогда событию $A-B$ благоприятствуют $a-c$ исходов, событию $B-A$ благоприятствуют $b-c$ исходов. Значит всего событию $A+B$ благоприятствуют $c+(a-c)+(b-c) = a+b-c$ исходов. Поэтому

$$\begin{aligned}\mathbb{P}(A+B) &= \frac{a+b-c}{n}\\ &= \frac{a}{n}+\frac{b}{n}-\frac{c}{n}\\ &= \mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(AB).\end{aligned}$$

Таким образом, **каковы бы ни были события A и B**,

$$\boxed{\mathbb{P}(A+B) = \mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(AB)}.$$

Отсюда без труда выводится аналогичное соотношение для суммы трёх событий:

$$\begin{aligned}\mathbb{P}(A+B+C) &= \mathbb{P}(A+B)+\mathbb{P}(C)-\mathbb{P}((A+B)\cdot C)\\ &= \mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(A\cdot B)+\mathbb{P}(C)-\mathbb{P}((A\cdot C)+(B\cdot C))\\ &= \mathbb{P}(A)+\mathbb{P}(B)+\mathbb{P}(C)-\mathbb{P}(A\cdot B)-(\mathbb{P}(A\cdot C)+\mathbb{P}(B\cdot C)-\mathbb{P}((A\cdot C)\cdot(B\cdot C)))\\ &= \mathbb{P}(A)+\mathbb{P}(B)+\mathbb{P}(C)-\mathbb{P}(A\cdot B)-\mathbb{P}(A\cdot C)-\mathbb{P}(A\cdot C)+\mathbb{P}(A\cdot B\cdot C).\end{aligned}$$

Подобным образом могут быть установлены аналогичные соотношения для четырёх, пяти и большего числа “слагаемых”.

Упражнение 5.1. Проделайте это для четырёх слагаемых.

6. Условные вероятности. Независимость событий

Всё в мире взаимосвязанно. Наступление одних событий влияет на вероятности наступления других событий.

Пример 6.1. Студент Л. Одырь должен был выучить к экзамену ответы на вопросы к 25 билетам, а он выучил лишь к одному. Он идёт сдавать первым. Какова вероятность вытянуть «заветный» билет?

Она равна $\frac{1}{25}$.

Пример 6.2. Но допустим, что Л. Одырь пошёл сдавать вторым, а первый вытянул тот самый единственный билет, выученный Л. Одырем. Какова теперь вероятность для Л. Одыря вытянуть знакомый ему билет?

Теперь она равна 0. Предыдущее событие радикально изменило вероятность наступления желаемого события для Л. Одыря.

Пример 6.3. Теперь допустим, что Л. Одырь пошёл сдавать вторым, а первый вытянул другой билет. Какова в этих условиях вероятность того, что Л. Одырю достанется заветный билет?

В этих условия она равна $\frac{1}{24}$.

Подобные примеры делают естественным рассмотрение понятия условной вероятности — вероятности наступления события B при условии, что наступило событие A. Её обозначают через $p_A(B)$ или $\mathbb{P}(B|A)$. Прежде чем давать общее определение, вычислим эту вероятность для классического случая.

Пример 6.4. Пусть всего имеется n равновозможных исходов, a из них благоприятствуют событию A, b — событию B и c — событию AB. Произведено испытание и в нём наступило событие A. Какова в этих условиях вероятность того, что наступило событие B?

Решение. Поскольку в произведённом испытании наступило событие A, осуществился один из a благоприятствующих событию A исходов. Из этих исходов событию B благоприятствуют ровно c исходов. Значит, в указанных условиях вероятность наступления события B равна отношению c к a:

$$\mathbb{P}(B|A) = \frac{c}{a}.$$

Но

$$\frac{c}{a} = \frac{c}{n} : \frac{a}{n} = \mathbb{P}(AB) : \mathbb{P}(A).$$

Таким образом,

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(AB)}{\mathbb{P}(A)}.$$

Это соотношение принимают в качестве определения условной вероятности в общем случае.

Определение 6.1. Пусть A — событие, вероятность которого отлична от 0, B — произвольное событие. *Вероятностью наступления события B при условии, что наступило событие A*, называется и через $\mathbb{P}(B|A)$ обозначается число, определяемое условием

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(AB)}{\mathbb{P}(A)}.$$

В тех случаях, когда $\mathbb{P}(A) = 0$, вероятность $\mathbb{P}(B|A)$ остаётся не определённой.

Мы подошли к понятию условной вероятности, используя известную вероятность произведения событий. С другой стороны, знание условных вероятностей позволяет вычислять вероятности произведений событий.

Теорема 6.1 (Вероятность произведения)**.** *Каковы бы ни были события A и B,*

$$\mathbb{P}(AB) = \mathbb{P}(A) \cdot \mathbb{P}(B|A).$$

Эта теорема позволяет вычислять вероятности произведения не только двух, но и любого конечного числа событий. Последовательно применяя её, получаем:

$$\begin{aligned}\mathbb{P}(ABC) &= \mathbb{P}(AB) \cdot \mathbb{P}(C|AB)\\ &= \mathbb{P}(A) \cdot \mathbb{P}(B|A) \cdot \mathbb{P}(C|AB);\\ \mathbb{P}(ABCD) &= \mathbb{P}(ABC) \cdot \mathbb{P}(D|ABC)\\ &= \mathbb{P}(A) \cdot \mathbb{P}(B|A) \cdot \mathbb{P}(C|AB) \cdot \mathbb{P}(D|ABC),\end{aligned}$$

и вообще,

$$\mathbb{P}(A_1A_2A_3\ldots A_k) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1A_2) \cdot \ldots \cdot \mathbb{P}(A_k|A_1A_2\ldots A_{k-1}).$$

Упражнение 6.1. В урне находятся шары: 4 белых, 5 красных и 3 синих. Трижды наудачу извлекают по одному шару, не возвращая его обратно. Какова вероятность того, что в первый раз появится белый шар (событие A), во второй — красный (событие B), в третий раз — синий (событие C)?

Определение 6.2. Говорят, что *событие B не зависит от события A*, если $\mathbb{P}(B|A) = \mathbb{P}(B)$.

Теорема 6.2.

1° *Если событие B не зависит от события A, то оно не зависит и от события $\overline{A}$.*

2° *Если событие B не зависит от события A, то и событие A не зависит от события B.*

3° *События A и B независимы тогда и только тогда, когда вероятность произведения этих событий равна произведению их вероятностей:* $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$.

Понятия зависимости и независимости событий представляются интуитивно понятными и не требующими специального определения. Однако это впечатление обманчиво.

Пример 6.5. Из полного набора костей домино наудачу извлекается одна кость. Пусть

$$A = \langle \text{сумма очков на половинках кости равна } 4 \rangle,$$
$$B = \langle \text{сумма очков на половинках кости равна } 6 \rangle,$$
$$C = \langle \text{извлечённая кость — дубль} \rangle.$$

Выясним, зависимы или нет события A и C, B и C. (Подсказывает ли что-нибудь читателю его интуиция?)

Решение. Выложим все кости домино в «треугольник», «покрасив» кости, составляющие событие A, в синий цвет, составляющие событие B — в зелёный (см. картинку далее).

Всего костей 28. Сумма очков равна 4 на трёх костях, ровно одна из которых является дублем. Значит, $\mathbb{P}(C \mid A) = \frac{1}{3}$. Всего дублей 7, значит, $\mathbb{P}(C) = \frac{7}{28} = \frac{1}{4}$. Поскольку $\mathbb{P}(C \mid A) \neq \mathbb{P}(C)$, события A и C зависимы.

Сумма очков равна 6 на четырёх костях, ровно одна из которых является дублем. Поэтому $\mathbb{P}(C \mid B) = \frac{1}{4} = \mathbb{P}(C)$. Значит, события B и C независимы.

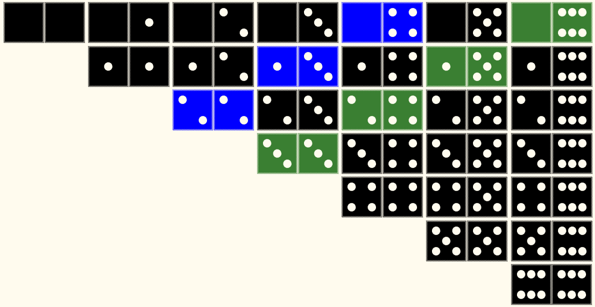


Рис. 1:

Упражнение 6.2. Из полной колоды карт (36 листов) наудачу извлекается одна карта. Пусть событие A состоит в том, что извлечённая карта — дама, событие B — в том, то это — «пика». Предлагается выяснить, зависимы или нет указанные события A и B. (Что говорит Вам на этот счёт Ваша интуиция?)

Предостережение 1. Не следует смешивать понятия независимости и несовместимости событий.

С одной стороны, события могут быть независимыми, но совместимыми. Например, если дважды подбрасывается монета, выпадение герба в первом испытании и выпадение герба во втором испытании — события независимые, но вполне совместимые.

С другой стороны, события могут быть несовместимыми, но зависимыми. Например, если подбрасывается игральная кость, выпадение чётного количества очков (событие A) и выпадение трёх очков (событие B) несовместимы, но зависимы: $\mathbb{P}(B) = 1/6$, а $\mathbb{P}(B|A) = 0$.

Советуем также решить следующее упражнение.

Упражнение 6.3. Докажите, что если события A и B возможны (т. е. $\mathbb{P}(A) > 0$ и $\mathbb{P}(B) > 0$) и независимы, то они совместимы.

7. Формула полной вероятности. Формулы Байеса

Теорема 7.1. *Пусть события $H_1, \ldots, H_n$ попарно несовместны и в сумме покрывают событие A, т. е. $A \subset H_1 + \ldots + H_n$* (см. рис. 2). *Тогда*

$$\mathbb{P}(A) = \mathbb{P}(H_1)\mathbb{P}(A|H_1) + \ldots + \mathbb{P}(H_n)\mathbb{P}(A|H_n).$$

Эта формула называется *формулой полной вероятности.*

Доказательство. По условию $A \subset H_1 + \ldots + H_n$. Значит,

$$\begin{aligned} A &= A \cdot (H_1 + \ldots + H_n) \\ &= A \cdot H_1 + \ldots + A \cdot H_n. \end{aligned}$$

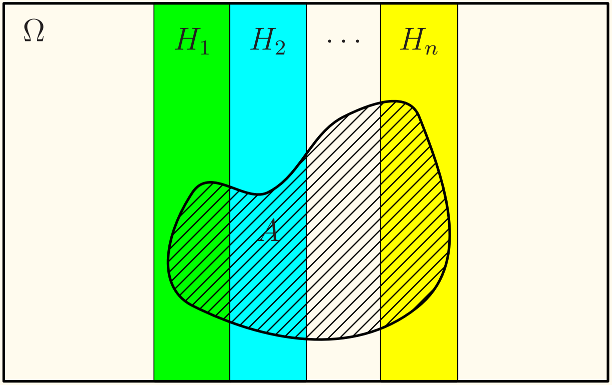


Рис. 2: Иллюстрация к теореме 7.1.

При этом события $H_1, \ldots, H_n$ попарно несовместны. Тем более несовместны события $A \cdot H_1, \ldots, A \cdot H_n$. По формуле произведения вероятностей (теорема 6.1) вероятности этих событий равны, соответственно, $\mathbb{P}(H_1)\mathbb{P}(A|H_1), \ldots, \mathbb{P}(H_n)\mathbb{P}(A|H_n)$. Поэтому

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(A \cdot (H_1 + \ldots + H_n)) \\
&= \mathbb{P}(A \cdot H_1 + \ldots + A \cdot H_n) \\
&= \mathbb{P}(A \cdot H_1) + \ldots + \mathbb{P}(A \cdot H_n) \\
&= \mathbb{P}(H_1)\mathbb{P}(A|H_1) + \ldots + \mathbb{P}(H_n)\mathbb{P}(A|H_n),
\end{aligned}$$

как и утверждалось. □

Пусть выполняются условия предыдущей теоремы. Допустим, что произведён опыт и в результате наступило событие A. Поскольку событие A покрывается событиями H_1, …, H_n, а они попарно несовместны, то вместе с A наступило одно и только одно из событий H_1, …, H_n. Какое именно — сказать невозможно, но можно найти условные вероятности: вероятности, с которыми наступают события H_1, …, H_n *при условии, что наступило событие* A.

Теорема 7.2. *Пусть выполняются условия предыдущей теоремы. Тогда, для любого* $i = 1, \ldots, n$,

$$\mathbb{P}(H_i|A) = \frac{\mathbb{P}(H_i)\mathbb{P}(A|H_i)}{\mathbb{P}(H_1)\mathbb{P}(A|H_1) + \ldots + \mathbb{P}(H_n)\mathbb{P}(A|H_n)}.$$

Эти формулы называются *формулами Байеса*. Заметим, что в знаменателе этих формул стоит выражение для $\mathbb{P}(A)$, полученное по формуле полной вероятности.

Доказательство. Доказательство этих формул чрезвычайно простое. Замечаем, что $A \cdot H_i = H_i \cdot A$ и вычисляем вероятности этих событий по формуле произведения вероятностей:

$$\mathbb{P}(A) \cdot \mathbb{P}(H_i|A) = \mathbb{P}(H_i) \cdot \mathbb{P}(A|H_i).$$

Отсюда

$$\mathbb{P}(H_i|A) = \frac{\mathbb{P}(H_i) \cdot \mathbb{P}(A|H_i)}{\mathbb{P}(A)},$$

а именно это и утверждалось. □

Пример 7.1. Три завода производят электролампы, 1-й — 60% всей продукции, 2-й — 25%, третий — 15%. Первый завод производит 1% брака, второй — 1,5%, третий — 2%.

1) Наудачу выбирается одна лампа. Какова вероятность того, что она — бракованная?

2) Наудачу выбранная лампа оказалась бракованной. Какова вероятность того, что она произведена на первом заводе?

Решение. Пусть

$A = \langle$выбранная лампа оказалась бракованной$\rangle$,

$H_1 = \langle$выбранная лампа произведена на 1-м заводе$\rangle$,

$H_2 = \langle$выбранная лампа произведена на 2-м заводе$\rangle$,

$H_3 = \langle$выбранная лампа произведена на 3-м заводе$\rangle$.

Тогда

$$\begin{array}{ll} \mathbb{P}(H_1) = 0,6, & \mathbb{P}(A \mid H_1) = 0,01, \\ \mathbb{P}(H_2) = 0,25, & \mathbb{P}(A \mid H_2) = 0,015, \\ \mathbb{P}(H_3) = 0,15, & \mathbb{P}(A \mid H_3) = 0,02. \end{array}$$

1) По формуле полной вероятности

$$\mathbb{P}(A) = \mathbb{P}(H_1)\cdot\mathbb{P}(A \mid H_1)+\mathbb{P}(H_2)\cdot\mathbb{P}(A \mid H_2)+\mathbb{P}(H_3)\cdot\mathbb{P}(A \mid H_3) = 0,6\cdot 0,01+0,25\cdot 0,015+0,1\cdot 0,02$$

2) По формуле Байеса

$$\mathbb{P}(H_1 \mid A) = \frac{\mathbb{P}(H_1)\cdot\mathbb{P}(A \mid H_1)}{\mathbb{P}(A)} = \frac{0,6\cdot 0,01}{0,01175} = 0,5106\ldots.$$

Пример 7.2. При обследовании больного имеется подозрение на одно из двух заболеваний: гипотезы H_1 и H_2. Их вероятности оцениваются так: $\mathbb{P}(H_1) = 0,6$, $\mathbb{P}(H_2) = 0,4$. Для уточнения диагноза назначается анализ, результатом которого является положительная или отрицательная реакция. В случае H_1 вероятность положительной реакции равна $0,9$, в случае H_2 — $0,5$.

1) Анализ произвели один раз, реакция была отрицательной. Каковы вероятности гипотез H_1 и H_2 в этих условиях?

2) Анализ произвели ещё раз, и опять реакция была отрицательной. Каковы вероятности гипотез H_1 и H_2 в этих условиях?

Решение. Пусть $A =$⟨результатом анализа явилась положительная реакция⟩. Тогда $\overline{A} =$⟨результатом анализа явилась отрицательная реакция⟩. По условию:

$$\begin{array}{lll} \mathbb{P}(H_1) = 0,6, & \mathbb{P}(A \mid H_1) = 0,9, & \mathbb{P}(\overline{A} \mid H_1) = 0,1; \\ \mathbb{P}(H_2) = 0,4, & \mathbb{P}(A \mid H_2) = 0,5, & \mathbb{P}(\overline{A} \mid H_2) = 0,5. \end{array}$$

1) По формуле полной вероятности

$$\mathbb{P}(\overline{A}) = \mathbb{P}(H_1) \cdot \mathbb{P}(\overline{A} \mid H_1) + \mathbb{P}(H_2) \cdot \mathbb{P}(\overline{A} \mid H_2) = 0,6 \cdot 0,1 + 0,4 \cdot 0,5 = 0,26.$$

Отсюда, по формулам Байеса, находим:

$$\mathbb{P}(H_1 \mid \overline{A}) = \frac{\mathbb{P}(H_1) \cdot \mathbb{P}(\overline{A} \mid H_1)}{\mathbb{P}(\overline{A})} = \frac{0,6 \cdot 0,1}{0,26} = \frac{3}{13};$$

$$\mathbb{P}(H_2 \mid \overline{A}) = \frac{\mathbb{P}(H_2) \cdot \mathbb{P}(\overline{A} \mid H_2)}{\mathbb{P}(\overline{A})} = \frac{0,4 \cdot 0,5}{0,26} = \frac{10}{13}.$$

Если судить по результатам первого анализа, то второе заболевание более вероятно, чем первое, но и у первого она не так уж мала, чтобы можно было его уверенно исключить. Посмотрим, что может дать второй анализ.

2) Результатом второго анализа тоже явилась отрицательная реакция. Можно подумать, что второй анализ ничего не дал. Но нет: после второго анализа у нас не просто два одинаковых результата одного анализа, у нас результат *двукратного* проведения анализа, а это уже нечто новое.

Обозначим через B событие, состоящее в том, что анализ дважды дал отрицательный результат. Поскольку результаты двух анализов независимы,

$$\mathbb{P}(B \mid H_1) = (\mathbb{P}(\overline{A} \mid H_1))^2 = (0,1)^2 = 0,01;$$
$$\mathbb{P}(B \mid H_2) = (\mathbb{P}(\overline{A} \mid H_2))^2 = (0,5)^2 = 0,25.$$

По формуле полной вероятности

$$\mathbb{P}(B) = \mathbb{P}(H_1) \cdot \mathbb{P}(B \mid H_1) + \mathbb{P}(H_2) \cdot \mathbb{P}(B \mid H_2) = 0,6 \cdot 0,01 + 0,4 \cdot 0,25 = 0,106.$$

Отсюда, по формулам Байеса,

$$\mathbb{P}(H_1 \mid B) = \frac{\mathbb{P}(H_1) \cdot \mathbb{P}(B \mid H_1)}{\mathbb{P}(B)} = \frac{0,6 \cdot 0,01}{0,106} = \frac{6}{106};$$
$$\mathbb{P}(H_2 \mid B) = \frac{\mathbb{P}(H_2) \cdot \mathbb{P}(B \mid H_2)}{\mathbb{P}(B)} = \frac{0,4 \cdot 0,25}{0,106} = \frac{100}{106}.$$

Вероятность наличия второго заболевания не просто больше, чем первого, она близка к единице! Наличие первого заболевания маловероятно, наличие второго заболевания весьма вероятно.

В качестве ещё одного, и весьма интересного, приложения формулы Байеса рассмотрим так называемую задачу Монти Холла, или парадокс трёх дверей.

Пример 7.3 (Задача Монти Холла)**.** Предположим, что Вы участвуете в шоу, ведущий которого, Монти Холл, показывает Вам три двери, скажем, A, B и C. Он сообщает Вам, что за одной дверью — автомобиль, за двумя другими — козлы. Монти предлагает Вам выбрать одну из дверей, не открывая её, и получить в подарок то, что скрывается за этой дверью. Он знает, что находится за каждой из дверей, а Вы совершенно определённо не горите желанием выиграть козла. Вы делаете свой выбор, Монти открывает одну из дверей и Вы видите, что за ней находится козёл. Затем Монти великодушно спрашивает Вас, не хотите ли Вы выбрать другую дверь? Перед Вами проблема: менять свой выбор или не менять?

Решение. На первый взгляд кажется, что для каждой из двух оставшихся закрытыми дверей вероятности того, что за ними скрывается приз, равны $\frac{1}{2}$. Вспомним, однако, что у нас есть формула Байеса и попытаемся применить её для получения обоснованного ответа.

Не ограничивая общности предположим, что Вы выбрали дверь A, а Монти Холл открыл дверь B (с козлом позади неё). Обозначим через MB событие $\langle$Монти Холл открыл дверь $B\rangle$, и через $\widetilde{X}$ событие $\langle$автомобиль находится за дверью $X\rangle$, где $X \in \{A,\ B,\ C\}$. Ясно, что события $\widetilde{A}$, $\widetilde{B}$ и $\widetilde{C}$ попарно несовместны. При этом:

- $\mathbb{P}(MB|\widetilde{A}) = \frac{1}{2}$, — потому что что двери B и C для Монти равноправны;
- $\mathbb{P}(MB|\widetilde{B}) = 0$, — потому что Монти ни в коем случае не откроет дверь, за которой скрывается приз;
- $\mathbb{P}(MB|\widetilde{C}) = 1$, — потому что Монти вынужден открыть дверь B, если приз находится за дверью C.

Вычислим теперь вероятности $\mathbb{P}(\widetilde{A}|MB)$ и $\mathbb{P}(\widetilde{C}|MB)$, то есть вероятности того, что приз скрывается за дверью A и за дверью C, при условии, что Монти открыл дверь B.

Имеем:

$$\begin{aligned}\mathbb{P}(\widetilde{A}|MB) &= \frac{\mathbb{P}(MB|\widetilde{A})\cdot\mathbb{P}(\widetilde{A})}{\mathbb{P}(MB|\widetilde{A})\cdot\mathbb{P}(\widetilde{A})+\mathbb{P}(MB|\widetilde{B})\cdot\mathbb{P}(\widetilde{B})+\mathbb{P}(MB|\widetilde{C})\cdot\mathbb{P}(\widetilde{C})}\\ &= \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}+1\cdot\frac{1}{3}}\\ &= \frac{1}{3};\end{aligned}$$

а вот

$$\begin{aligned}\mathbb{P}(\widetilde{C}|MB) &= \frac{\mathbb{P}(MB|\widetilde{C})\cdot\mathbb{P}(\widetilde{C})}{\mathbb{P}(MB|\widetilde{A})\cdot\mathbb{P}(\widetilde{A})+\mathbb{P}(MB|\widetilde{B})\cdot\mathbb{P}(\widetilde{B})+\mathbb{P}(MB|\widetilde{C})\cdot\mathbb{P}(\widetilde{C})}\\ &= \frac{1\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}+1\cdot\frac{1}{3}}\\ &= \frac{2}{3}.\end{aligned}$$

Таким образом, Вы вдвое увеличиваете свои шансы на выигрыш, если меняете свой выбор!

8. Элементы комбинаторики

Пример 8.1. В гостинице 10 свободных мест. В очереди к администратору а) 2 гостя; б) 3 гостя; в) 4 гостя. Сколькими способами эти гости могут быть размещены на свободных местах, если каждое место может быть занято не более чем одним гостем?

Решение. а) Для 1-го гостя существует 10 вариантов размещения. При любом из этих вариантов остаётся 9 свободных мест и, значит, 9 возможностей размещения для 2-го гостя. Всего получается 10 раз по 9, т.е. $10 \cdot 9$ способов размещения двух гостей на 10 свободных местах.

б) Для первых двух гостей имеется $10 \cdot 9$ способов размещения. При любом из этих способов остаётся 8 свободных мест и, значит, 8 вариантов размещения для 3-го гостя. Всего получается $10 \cdot 9$ раз по 8, т.е. $10 \cdot 9 \cdot 8$ способов размещения троих гостей на 10 свободных местах.

в) Для первых трёх гостей существует $10 \cdot 9 \cdot 8$ способов размещения. При любом из этих способов остаётся 7 свободных мест и, значит, 7 способов размещения 4-го гостя. Всего получается $10 \cdot 9 \cdot 8$ раз по 7, т.е. $10 \cdot 9 \cdot 8 \cdot 7$ способов размещения 4-х гостей на 10 свободных местах.

Пример 8.2. В гостинице n свободных мест, в очереди к администратору k гостей. Сколькими способами эти гости могут быть размещены на свободных местах, если на каждое место разрешается помещать не более одного гостя?

Решение. Обозначим искомое количество способов через A_n^k. Для 1-го гостя существует n способов размещения: $A_n^1 = n$. При любом из этих способов остаётся $n-1$ свободное место и, значит, $n-1$ способ размещения 2-го гостя. Всего для 2-х гостей получается A_n^1 раз по $n-1$ способов размещения, так что

$$A_n^2 = A_n^1 \cdot (n-2) = n \cdot (n-1).$$

Как бы ни были размещены два первых гостя, для 3-го гостя остаётся $n-2$ свободных места и, значит, $n-2$ способа размещения. Всего для 3-х гостей получается A_n^2 раз по $(n-2)$ способов размещения, так что

$$A_n^3 = A_n^2 \cdot (n-2) = n \cdot (n-1) \cdot (n-2).$$

. . . .

Когда очередь дойдёт до k-го гостя, $(k-1)$ место из n будет занято, для него останется $n-(k-1)$ свободных мест и, значит, $n-(k-1)$ способов размещения. Всего для k гостей получается A_n^{k-1} раз по $n-(k-1)$ способов, так что

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

Итак,

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

Для A_n^n, числа размещений из n по n, имеется специальное название и специальное обозначение: число n-перестановок, P_n. В этом случае $k=n$, так что $n-(k-1) = n-(n-1) = 1$. Получается произведение всех натуральных чисел от n до 1 (или от 1 до n, если перечислять множители в возрастающем порядке). Это произведение обозначается через $n!$ и называется эн-*факториалом*.

Таким образом,

$$P_n = A_n^n = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1 = n!.$$

Отдельно полагают

$$1! = 1, \quad 0! = 1.$$

Выпишем несколько первых факториалов:

$$0! = 1;\ \ 1! = 1;\ \ 2! = 1 \cdot 2 = 2;$$
$$3! = 1 \cdot 2 \cdot 3 = 2! \cdot 3 = 2 \cdot 3 = 6;$$
$$4! = 1 \cdot 2 \cdot 3 \cdot 4 = 3! \cdot 4 = 6 \cdot 4 = 24;$$
$$5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 4! \cdot 5 = 24 \cdot 5 = 120.$$

Задание 8.1. Доведите вычисления до $10!$.

Пример 8.3. Сколькими способами k человек могут встать в очередь друг за другом?

Решение. Для заполнения первого места имеется k способов. При любом из этих способов для заполнения второго места имеется $k-1$ способ. Всего для заполнения первых двух мест имеется k раз по $k-1$, то есть $k\cdot(k-1)$ способов. При любом из этих способов на третье место остаётся $k-2$ претендента. Получается $k\cdot(k-1)\cdot(k-2)$ способов для заполнения первых трёх мест. И так далее. Для последнего, k-го, остаётся одно место, то есть всего одна возможность. В результате получаем $k\cdot(k-1)\cdot\ldots\cdot 1 = k!$ способов.

Благодаря введению факториалов появляется возможность записать формулу для числа размещений более компактно. Действительно,

$$\begin{aligned} A_n^k &= n\cdot(n-1)\cdot\ldots\cdot(n-(k-1)) \\ &= \frac{n\cdot(n-1)\cdot\ldots\cdot(n-(k-1))\cdot(n-k)\cdot\ldots\cdot 1}{(n-k)\cdot\ldots\cdot 1} \\ &= \frac{n!}{(n-k)!}. \end{aligned}$$

Выписывая крайние члены этой цепочки равенств, получаем:

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

Посмотрим на процесс размещения гостей в гостинице со стороны администратора. У него имеется n свободных мест. Из них он выбирает k мест и распределяет между стоящими в очереди гостями: это — первому, это — второму, …, это — k-му. С математической точки зрения его действия состоят в том, что он выбирает из n свободных мест k мест и упорядочивает эту выборку, то есть располагает в последовательность. Таким образом,

число размещений из n по k равно количеству способов,
которыми могут быть выбраны и упорядочены k *предметов из* n.

Этим и объясняется название „число размещений *из n по k*".

В частности,

число k-перестановок равно количеству способов,
которыми могут быть упорядочены k предметов.

Задание 8.2. Укажите все перестановки чисел 1 и 2; чисел 1, 2, 3; чисел 1, 2, 3, 4.

Пример 8.4. В гостинице n свободных мест, прибыло k гостей. (В очередь к администратору они ещё не встали, но администратор уже прикидывает, какие места он выделит для прибывших.) Сколькими способами администратор из имеющихся в его распоряжении n свободных мест может выбрать k мест для размещения гостей?

Решение. Обозначим искомое число способов через X. Рассмотрим произвольную выборку k мест из n. Распределение этих k мест между гостями будет зависеть от того, в каком порядке гости встанут в очередь друг за другом. Сколькими способами k человек могут встать в очередь друг за другом? Ответ известен: $k!$ способами (см. пример 8.3). Таким образом, каждая выборка может быть распределена между гостями $k!$ способами. Количество выборок равно X, каждая распределяется между гостями (упорядочивается) $k!$ способами. Получается $X \cdot k!$ упорядоченных k-выборок. Но это количество упорядоченных k-выборок есть не что иное, как число размещений из n по k. Таким образом,

$$X \cdot k! = A_n^k,$$

откуда

$$X = \frac{A_n^k}{k!}.$$

Количество способов выбора k предметов из n называют *числом сочетаний из n по k* и обозначают через C_n^k (в России) или $\binom{n}{k}$ (в остальном мире). Как показало решение последнего примера,

$$\boxed{C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!}}.$$

Вычислять числа C_n^k по приведённой формуле — дело весьма утомительное. К счастью, существует более лёгкий способ, найденный Паскалем. Он основан на простейших свойствах чисел C_n^k, отражённых в следующей теореме.

Теорема 8.1. *Для любых неотрицательных целых n и k:*

0° $C_n^0 = C_n^n = 1$;

1° $C_n^1 = n$;

2° $C_n^k = C_n^{n-k}$;

3° $C_n^k + C_n^{k+1} = C_{n+1}^{k+1}$ *(свойство Паскаля).*

Доказательство. $\boxed{3.}$

$$\begin{aligned} C_n^k + C_n^{k+1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-(k+1))!} = \\ &= \frac{n!}{k!(n-(k+1))!}\left(\frac{1}{n-k} + \frac{1}{k+1}\right) = \\ &= \frac{n!}{k!(n-(k+1))!} \cdot \frac{n+1}{(k+1)(n-k)} = \\ &= \frac{(n+1)!}{(k+1)!(n-k)!} = \end{aligned}$$

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

□

Упражнение 8.1. Докажите остальные пункты этой теоремы.

Пользуясь свойствами, перечисленными в теореме 8.1, будем вычислять числа C_n^k по схеме:

$$
\begin{tikzcd}
&&&&& C_0^0 &&&&& \\
&&&& C_1^0 \arrow[dr] && C_1^1 \arrow[dl] &&&& \\
&&& C_2^0 \arrow[dr] && C_2^1 \arrow[dl] \arrow[dr] && C_2^2 \arrow[dl] &&& \\
&& C_3^0 \arrow[dr] && C_3^1 \arrow[dl] \arrow[dr] && C_3^2 \arrow[dl] \arrow[dr] && C_3^3 \arrow[dl] && \\
& C_4^0 \arrow[dr] && C_4^1 \arrow[dl] \arrow[dr] && C_4^2 \arrow[dl] \arrow[dr] && C_4^3 \arrow[dl] \arrow[dr] && C_4^4 \arrow[dl] & \\
C_5^0 && C_5^1 && C_5^2 && C_5^3 && C_5^4 && C_5^5
\end{tikzcd}
$$

. . . .

Начинаем с единицы в вершине и двух единиц в первой строке. В каждой следующей строке по краям ставим единицы, а каждый внутренний член, пользуясь свойством Паскаля 3, получаем как сумму двух чисел, стоящих непосредственно над этим членом (слева и справа) в предыдущей строке.

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
$$

. . . .

Получаемая таким образом схема натуральных чисел называется *треугольником Паскаля.*

Задание 8.3. Добавьте к приведённому треугольнику ещё 10 строк.

9. Применение комбинаторики к подсчёту вероятностей

Пример 9.1 (Задача о статистической выборке)**.** В урне N шаров, из которых K — белые. Из этой урны наудачу извлекается n шаров. Какова вероятность того, что ровно k из них окажутся белыми?

Решение. Количество способов извлечь n шаров из N равно C_N^n. Это — общее количество всех возможных исходов описанного испытания. Теперь подсчитаем количество благоприятствующих исходов.

Количество способов извлечь k белых шаров из K равно C_K^k. Остальные $n-k$ шаров должны быть не белыми. Не белых шаров $N-K$. Количество способов извлечь $n-k$ шаров из $N-K$ равно C_{N-K}^{n-k}. Таким образом, количество благоприятствующих исходов равно $C_K^k \cdot C_{N-K}^{n-k}$, а искомая вероятность равна

$$\frac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n}.$$

Упражнение 9.1. В урне 10 шаров. Из неё наудачу извлекают 2 шара. Вероятность того, что оба окажутся белыми, равна $\frac{2}{15}$. Сколько в урне белых шаров?

10. Последовательности независимых испытаний. Формула Бернулли

Определение 10.1. Если при проведении нескольких испытаний вероятность события A в каждом испытании не зависит от исходов других испытаний, то эти испытания называют *независимыми относительно события* A. Наступление события A в данном испытании называют *успехом* и обозначают букой «У», ненаступление события A называют *неуспехом*, или *неудачей* и обозначают буквой «Н».

Последовательность испытаний называют *испытаниями Бернулли*, если вероятность успеха во всех испытаниях одна и та же. Принято обозначать эту вероятность через p. Понятно, что вероятность неуспеха в испытаниях Бернулли также одна и та же и равна $1-p$. Будем обозначать её через q.

Таким образом, в дальнейшем

$$p = \mathbb{P}(\text{У}), \quad q = 1-p = 1-\mathbb{P}(\text{У}) = \mathbb{P}(\text{Н}).$$

Обозначим через μ_n количество успехов в последовательности из n испытаний Бернулли, и через $P_n(k)$ — вероятность того, что в последовательности из n испытаний Бернулли будет ровно k успехов. Иначе говоря, $P_n(k)$ есть вероятность события $\langle \mu_n = k \rangle$:

$$P_n(k) = \mathbb{P}(\mu_n = k).$$

Мы будем пользоваться обоими этими обозначениями.

Найдём формулу, позволяющую вычислять $P_n(k)$ по данным p, n, k. Сначала рассмотрим частный случай.

Пример 10.1. Производится 5 испытаний Бернулли с вероятностью успеха p. Какова вероятность того, что ровно 2 из этих 5 испытаний завершатся успехом?

Решение. Исходы этих последовательностей испытаний представляются пятичленными последовательностями, каждый член которых есть либо У, либо Н.

Например, одной из таких последовательностей является упорядоченная пятёрка

$$(\text{У}, \text{Н}, \text{Н}, \text{У}, \text{Н}).$$

Каждая такая последовательность представляет произведение событий, состоящих в том, что в первом испытании наступило событие, указанное на первом месте, во втором испытании — событие, указанное на втором месте, ..., в пятом испытании — событие, указанное на пятом месте.

Например, рассмотренной последовательности отвечает произведение событий, состоящих в том, что в первом испытании имел место успех, во втором — неуспех, в третьем — неуспех, в четвёртом — успех, в пятом — неуспех.

Поскольку испытания независимы, вероятность этого произведения событий равна произведению вероятностей сомножителей. Для рассмотренной последовательности это есть произведение

$$pqqpq = p^2q^3.$$

Тот же результат будет получен для любой пятичленной последовательности испытаний, завершившейся ровно двумя успехами.

Сколько имеется таких последовательностей? — Ровно столько, сколькими способами можно выделить два места из пяти. Для этого, как мы знаем, существует ровно C_5^2 способов. Значит, имеется ровно C_5^2 пятичленных последовательностей, в которых успех встречается точно два раза. При этом различные последовательности несовместимы (в каждой серии испытаний реализуется ровно одна из них, так что если появилась одна, то не появилась никакая другая). Поэтому вероятность $P_5(2)$ появления пятичленной последовательности испытаний, в которой успех наступает точно два раза, равна $C_5^2 p^2 q^3$:

$$P_5(2) = C_5^2 p^2 q^3.$$

Теперь рассмотрим общий случай.

Пример 10.2. Производится n испытаний Бернулли с вероятностью успеха p. Какова вероятность того, что в этих n испытаниях будет ровно k успехов?

Рассуждая так же, как в рассмотренном частном случае, получаем:

$$P_n(k) = C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}.$$

Эта формула называется *формулой Бернулли.*

Пример 10.3. Монету подбрасывают 5 раз. Найти вероятности того, что герб выпадет: а) 2 раза, б) менее двух раз, в) не менее двух раз.

Решение. Вероятности выпадения любой из двух сторон монеты одинаковы, т. е. $p = q = 0,5$.

В случае а) $n = 5$, $k = 2$. Отсюда по формуле Бернулли получаем:

$$P_5(2) = C_5^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^4 = \frac{\frac{5!}{2!3!}}{2^6} = \frac{10}{32} = \frac{5}{16}.$$

В случае б) $n = 5$, $k < 2$, т. е. $n = 5$ и $k = 0,\ 1$. Поэтому:

$$P_5(k < 1) = P_5(0) + P_5(1) = C_5^0 \left(\frac{1}{2}\right)^5 + C_5^1 \left(\frac{1}{2}\right)^2 = \frac{1+5}{32} = \frac{3}{16}.$$

В случае в) $n = 5$, $k \geqslant 2$, т. е. $n = 5$ и неверно, что $k < 2$. Поэтому

$$P_5(k \geqslant 2) = 1 - P_5(k < 2) = 1 - \frac{3}{16} = \frac{13}{16}.$$

Пример 10.4. Игральную кость подбрасывают шесть раз. Какова вероятность того, что ровно три раза выпадет более четырёх очков?

Решение. Будем считать успехом выпадение более четырёх очков. Вероятность успеха равна $\frac{1}{3}$ (почему?). Искомая вероятность находится по формуле Бернулли:

$$P_6(3) = C_6^3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^3 = 20 \cdot \frac{1}{27} \cdot \frac{8}{27} = \frac{160}{729}.$$

Пример 10.5. Игральную кость подбрасывают шесть раз. Какова вероятность того, что более трёх раз выпадет более четырёх очков?

Решение. И здесь будем считать успехом выпадение более четырёх очков. Вероятность успеха равна $\frac{1}{3}$. Искомая вероятность есть вероятность события $\langle \mu_6 > 3 \rangle$, которое распадается на три попарно несовместимых события: $\langle \mu_6 = 4 \rangle$, $\langle \mu_6 = 5 \rangle$, $\langle \mu_6 = 6 \rangle$. Вероятность каждого из этих событий находится по формуле Бернулли. В итоге получаем:

$$\mathbb{P}(\mu_6 > 3) = P_6(4)+P_6(5)+P_6(6) = C_6^4\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2+C_6^5\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^1+C_6^6\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^0 =$$

$$= \left(\frac{1}{3}\right)^6 \cdot (15\cdot 2^2 + 5\cdot 2^1 + 1\cdot 2^0) = \frac{71}{729}.$$

Пример 10.6. Студент решает задачу по теории вероятностей. Вероятность того, что он получит правильный ответ, равна $p = 0,4$. Вокруг него сидят 5 однокурсников, каждый из которых способен получить правильный ответ с той же вероятностью. Как лучше поступить студенту: решать задачу самому или положиться на мнение большинства из этих пяти однокурсников?

Решение. Вычислим вероятность того, что большинство из пяти однокурсников дадут правильный ответ. Большинство из пяти — это 3, 4 или 5. Поэтому:

$$\begin{aligned} P(k \geqslant 3) &= P_5(3) + P_5(4) + P_5(5) = \\ &= C_5^3(0,4)^3(0,6)^2 + C_5^4(0,4)^4(0,6)^1 + C_5^5(0,4)^5(0,6)^0 = \\ &= 10\cdot 0,064\cdot 0,36 + 5\cdot 0,256\cdot 0,6 + 1\cdot 0,01024\cdot 1 = \\ &= 0,31744. \end{aligned}$$

Ответ. Вероятность того, что большинство из пяти однокурсников дадут правильный ответ, меньше вероятности получить правильный ответ самому. Лучше решать задачу самостоятельно.

Упражнение 10.1. Игральную кость бросают 5 раз. а) Найдите вероятность того, что более трёх раз выпадет более четырёх очков. б) Самостоятельно составьте и решите ещё 5 подобных задач.

Упражнение 10.2. Игральную кость бросают 5 раз. а) Что более вероятно: двукратное выпадение более 4-х очков — или трёхкратное выпадение более 2-х очков? б) Самостоятельно составьте и решите ещё 5 подобных задач.

Упражнение 10.3. Испытание состоит в пятикратном подбрасывании монеты и повторяется 5 раз. а) Найдите вероятность того, что более трёх раз выпадет более трёх гербов. б) Самостоятельно составьте и решите ещё 5 подобных задач.

Упражнение 10.4. Испытание состоит в пятикратном подбрасывании монеты и повторяется 5 раз. а) Что более вероятно: трёхкратное выпадение более двух гербов — или двукратное выпадение более трёх гербов? б) Самостоятельно составьте и решите ещё 5 подобных задач.

Упражнение 10.5 (Задача де Мерэ)**.** Сколько раз нужно подбросить пару игральных костей, чтобы вероятность появления хотя бы один раз двух шестёрок была не менее $\frac{1}{2}$?

11. Наиболее вероятное число успехов

Пример 11.1. В родильном доме за 1 сутки появилось на свет 10 детей. Пусть вероятность рождения мальчика равна $0,5$. Найдём вероятности того, что родилось 0, 1, $\ldots$, 10 мальчиков.

Решение. Производим вычисления по формуле Бернулли с $n = 10$, $p = q = \frac{1}{2}$ и $k = 0$, 1, 2, $\ldots$ 10. Замечаем, что, в силу равенства $C_n^k = C_n^{n-k}$ и предположения $p = q$ имеют место равенства $P_n(k) = P_n(n-k)$, и что, при любом k,

$$p^k q^{n-k} = p^k p^{n-k} = p^n = \left(\frac{1}{2}\right)^{10} = \frac{1}{2^{10}} = \frac{1}{1024}.$$

Получаем:

$$P_{10}(0) = P_{10}(10) = C_{10}^0 \cdot \frac{1}{1024} = \frac{1}{1024};$$

$$P_{10}(1) = P_{10}(9) = C_{10}^{1} \cdot \frac{1}{1024} = \frac{10}{1024};$$
$$P_{10}(2) = P_{10}(8) = C_{10}^{2} \cdot \frac{1}{1024} = \frac{45}{1024};$$
$$P_{10}(3) = P_{10}(7) = C_{10}^{3} \cdot \frac{1}{1024} = \frac{120}{1024};$$
$$P_{10}(4) = P_{10}(6) = C_{10}^{4} \cdot \frac{1}{1024} = \frac{210}{1024};$$
$$P_{10}(5) = C_{10}^{5} \cdot \frac{1}{1024} = \frac{252}{1024}.$$

Как видим, в рассмотренном примере вероятности $P_n(k)$ сначала возрастают, достигают наибольшего значения, а затем убывают. Наибольшую величину имеет вероятность $P_{10}(5)$. Иначе говоря, наиболее вероятно появление пяти мальчиков. Этого можно было ожидать. Но вот что неожиданно, так это значение этой наибольшей вероятности. Оно равно отнюдь не $0,5$, как можно было бы подумать, а лишь $\frac{252}{1024}$, то есть примерно вдвое меньше.

Теперь рассмотрим общим общий случай. Оказывается, и в общем случае при фиксированном n имеет место похожая картина: при увеличении k вероятности $P_n(k)$ сначала возрастают, достигают наибольшего значения, а затем убывают.

Чтобы убедиться в этом и найти наибольшее значение, фиксируем n и выясним, при каких k имеет место неравенство $P_n(k) < P_n(k+1)$. Имеем:

$$\begin{aligned} P_n(k) < P_n(k+1) &\Leftrightarrow C_n^k p^k q^{n-k} < C_n^{k+1} p^k q^{n-(k+1)} \\ &\Leftrightarrow \frac{n! p^k q^{n-k}}{k!(n-k)!} < \frac{n! p^{k+1} q^{n-(k+1)}}{(k+1)!(n-(k+1))!} \\ &\Leftrightarrow \frac{q}{n-k} < \frac{p}{k+1} \end{aligned}$$

$$\begin{aligned}&\Leftrightarrow (k+1)q < (n-k)p\\&\Leftrightarrow k(p+q) < np-q\\&\Leftrightarrow k < np-q.\end{aligned}$$

Итак,

$$P_n(k) < P_n(k+1) \Leftrightarrow k < np-q.$$

Совершенно аналогично устанавливается, что

$$P_n(k) > P_n(k+1) \Leftrightarrow k > np-q$$
$$P_n(k) = P_n(k+1) \Leftrightarrow k = np-q.$$

Попытаемся понять, что это означает.

Начнём с последнего соотношения. В силу случайного соотношения между n, p и q может оказаться, что $np-q$ — целое число. Тогда при $k_0 = np-q$ окажется, что $P_n(k_0) = P_n(k_0+1)$. Для $1 \leqslant k \leqslant k_0-1$ будет выполняться неравенство $k < np-q$ и потому — неравенства

$$P_n(1) < P_n(2) < \ldots < P_n(k_0-1) < P_n(k_0),$$

для $k_0+1 \leqslant k \leqslant n$ будет выполняться неравенство $k > np-q$ и потому — неравенства

$$P_n(k_0+1) > P_n(k_0+2) > \ldots > P_n(n-1) > P_n(n).$$

В целом будет выполняться цепочка неравенств

$$P_n(1) < P_n(2) < \ldots < P_n(k_0) = P_n(k_0+1) > P_n(k_0+2) > \ldots > P_n(n),$$

показывающая, что (если $np-q$ — целое число!) наибольшее значение величина $P_n(k)$ принимает при двух значениях k: $k = np-q$ и $k = np-q+1 = np+p$.

Если же $np - q$ — не целое, то существует ровно одно целое число k_0, заключённое между $np - q$ и $np - q + 1 = np + p$. Тогда получаем следующую цепочку неравенств:

$$P_n(1) < P_n(2) < \ldots < P_n(k_0) > P_n(k_0 + 1) > P_n(k_0 + 2) > \ldots > P_n(n).$$

В этом случае наибольшее значение величина $P_n(k)$ принимает при (единственном) k_0, удовлетворяющем неравенствам

$$np - q < k_0 < np + p.$$

Подведём итог. Наибольшее значение величина $P_n(k)$ принимает при целых k_0, удовлетворяющих неравенствам

$$np - q \leqslant k_0 \leqslant np + p.$$

Если число $np - q$ — целое, таких k_0 два: $k_0 = np - q$ и $k_0 = np + p$. Если же число $np - q$ — не целое (как чаще всего и бывает), то наибольшее значение величина $P_n(k)$ принимает при единственном k_0, удовлетворяющем указанным неравенствам.

Определение 11.1. Значение $k = k_0$, при котором величина $P_n(k)$ достигает наибольшего значения, называют *наиболее вероятным числом наступлений события A в серии из n испытаний.*

Вывод: чтобы найти наиболее вероятное число наступлений события A в серии из n испытаний Бернулли, в каждом из которых событие A наступает с вероятностью p, нужно найти все целые числа k_0, удовлетворяющие системе неравенств

$$np - q \leqslant k_0 \leqslant np + p.$$

Таких чисел может быть либо два (если число $np - q$ — целое), либо одно (если число $np - q$ — не целое). В частности, если число np — целое, то именно оно и является наиболее вероятным числом наступлений события A в серии из n испытаний.

Пример 11.2. Куплено а) 8, б) 9 семян, каждое из которых с вероятностью 0,8 даст всходы. Какое наиболее вероятное число k_0 из этих семян даст всходы?

Решение. В обоих случаях $p = 0,8$, $q = 0,2$. В случае а) $n = 8$ и мы имеем

$$8 \cdot 0,8 - 0,2 \leqslant k_0 \leqslant 8 \cdot 0,8 + 0,8,$$

откуда $k_0 = 7$.

В случае б)

$$9 \cdot 0,8 - 0,2 \leqslant k_0 \leqslant 9 \cdot 0,8 + 0,8,$$

то есть

$$7 \leqslant k_0 \leqslant 8,$$

откуда $k_0 = 8$ или 9.

Упражнение 11.1. Монета подбрасывается 16 раз. Каково наиболее вероятное количество случаев выпадения герба? Какова эта наибольшая вероятность? А если монета подбрасывается 17 раз? 210 раз? 211 раз? 4000 раз? 4001 раз?

Упражнение 11.2. В родильном доме за день появилось на свет 10 детей. Каково наиболее вероятное количество мальчиков среди них? Какова эта наибольшая вероятность? А если детей 11? 12? 13? 14? 15? (Вероятность рождения мальчика считать равной 0,5.)

Упражнение 11.3. Игральная кость подбрасывается 15 раз. Каково наиболее вероятное количество испытаний, в которых выпадет 6 очков? 5 очков? более четырёх очков? А если кость подбрасывается 30 раз? 31 раз? 32 раза? 33 раза? 34 раза? 35 раз? 36 раз?

Упражнение 11.4. Испытание состоит в пятикратном подбрасывании монеты и повторяется 35 раз. Каково наиболее вероятное количество испытаний, в которых выпадет три герба? четыре герба? пять гербов? более трёх гербов? А если испытание повторяется 50 раз? 75 раз? 100 раз? 150 раз?

Упражнение 11.5. Испытание состоит в двукратном подбрасывании игральной кости и повторяется 15 раз. Каково наиболее вероятное количество испытаний, в которых сумма выпавших очков окажется равна 9? 10? 11? 12? более десяти? более девяти? более восьми? Каковы эти наибольшие вероятности? А если испытание повторяется 16 раз? 17 раз? 18 раз? 19 раз? 20 раз? 21 раз? 50 раз? 75 раз? 100 раз? 150 раз?

Упражнение 11.6. Производится 99 испытаний Бернулли. Наиболее вероятное количество успехов равно 60. Какова вероятность успеха?

12. Локальная приближённая формула Муавра-Лапласа

Вычисление вероятностей $P_n(k)$ при больших n и k требует большой вычислительной работы. В таких случаях применяют приближённую формулу, указанную в приведённой далее локальной теореме Лапласа. Прежде чем формулировать эту теорему, введём обозначение:

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

Так определённая функция называется *функцией Гаусса*, а её график — *кривой Гаусса*.

Упражнение 12.1. Исследуйте функцию Гаусса и постройте её график.

Имеются таблицы значений функции φ, по которым можно с достаточно высокой степенью точности найти практически любое значение этой функции, а значит, и вычислить вероятность, о которой идёт речь в следующей теореме. Функция φ — чётная, поэтому таблицах даются её значения только для положительных значений x.

Теорема 12.1 (Локальная приближённая формула Муавра-Лапласа)**.** *Пусть производится серия из n испытаний Бернулли с вероятностью успеха p, причем $0 < p < 1$. Тогда при достаточно больших n*

$$P_n(k) \approx \frac{1}{\sqrt{npq}}\varphi(x),$$

где

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}, \quad x = \frac{k - np}{\sqrt{npq}}.$$

Слова «при достаточно больших n» означают, что при постоянном p точность указанной формулы возрастает с увеличением n. Более определённо: точность улучшается с ростом произведения npq. На практике эту формулу применяют в случае, если $npq \geqslant 10$.

Пример 12.1. Вероятность выпуска бракованного изделия равна 0,3. Найти вероятность того, что среди 100 выпущенных изделий будет ровно 60 изделий без брака.

Решение. Будем считать испытанием проверку отдельно взятого изделия на наличие или отсутствие брака, успехом — отсутствие брака. По условию количество испытаний $n = 100$, $p = 0,7$, $q = 0,3$, $k = 60$. Поскольку $npq = 100 \cdot 0,7 \cdot 0,3 = 21 > 10$, применима локальная формула Муавра-Лапласа.

Последовательно вычисляем:

$$\sqrt{npq} = \sqrt{100 \cdot 0,7 \cdot 0,3} = 4,58,$$

$$x = \frac{k - np}{\sqrt{npq}} = \frac{10}{4,58} = 2,18.$$

Теперь для найденного аргумента x по таблице значений функции φ находим соответствующее значение $\varphi(x)$; оно равно 0,0371. Подстановка этого числа в локальную формулу Муавра-Лапласа дает приближенное значение искомой вероятности:

$$P_{100}(60) \approx \frac{0,0371}{4,58} = 0,008.$$

13. Интегральная приближённая формула Муавра-Лапласа

Теорема 13.1 (Интегральная приближённая формула Муавра-Лапласа)**.** *Пусть производится серия из n испытаний Бернулли с вероятностью успеха p, причем $0 < p < 1$. Тогда при достаточно больших n*

$$P_n(k_1 \leqslant k \leqslant k_2) \approx \Phi(x_2) - \Phi(x_1),$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{1}{2}t^2} dt, \quad x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}.$$

И эту формулу формулу следует применять лишь в тех случаях, когда $npq \geqslant 10$.

Функцию Φ называют *функцией Лапласа*. Для неё также составлены таблицы значений. Функция $\Phi(x)$ — нечётная, поэтому в таблицах даются её значения только для положительных значений x.

Пример 13.1. Всхожесть семян равна 0,8. Куплено 900 семян. Какова вероятность того, что от 710 до 740 из них дадут всходы?

Решение. В этом примере $n = 900$, $p = 0,8$, $q = 0,2$, $k_1 = 710$, $k_2 = 740$. Требуется найти $P_{900}(710 \leqslant k \leqslant 740)$. Непосредственное применение формулы Бернулли в этом случае даёт точное значение искомой вероятности:

$$P_{900}(710 \leqslant k \leqslant 740) = P_{900}(710) + P_{900}(711) + \ldots + P_{900}(740).$$

Однако нахождение этого точного значения сопряжено с большими вычислительными трудностями. Поэтому будем искать приближённое значение искомой величины. Воспользуемся интегральной формулой Муавра-Лапласа. Поскольку $npq = 900 \cdot 0,8 \cdot 0,2 = 144 > 10$, эта формула применима.

Вычисляем последовательно:

$$\sqrt{npq} = \sqrt{900 \cdot 0,8 \cdot 0,2} = \sqrt{144} = 12,$$

$$x_1 = \frac{k_1 - np}{\sqrt{npq}} = \frac{710 - 720}{12} = \frac{-10}{12} \approx -0,83,$$

$$x_2 = \frac{k_2 - np}{\sqrt{npq}} = \frac{740 - 720}{12} = \frac{20}{12} \approx 1,67.$$

По таблице значений функции Лапласа находим:

$$\Phi(-0,83) = -\Phi(0,83) \approx -0,2967, \quad \Phi(1,67) \approx 0,4525.$$

Отсюда

$$P_{900}(710 \leqslant k \leqslant 740) \approx 0,4525 - (-0,2967) = 0,7492.$$

14. Предельная теорема и приближённая формула Пуассона

Как уже отмечалось, точность приближённых формул Муавра-Лапласа возрастает с увеличением произведения npq. Отсюда видно, что чем ближе одно из чисел, p или q, к нулю (а другое, следовательно, к единице), тем большим следует брать n. Поэтому в случае близости одной из вероятностей p или q к нулю приближёнными формулами Муавра-Лапласа обычно не пользуются. В указанном случае значительно более точной является приведённая далее в этом пункте приближённая формула Пуассона. Но сначала —

Теорема 14.1 (Предельная теорема Пуассона). *Пусть производится последовательность серий испытаний, состоящих из 1, 2, ..., n, ... испытаний, причём вероятность успеха в каждой серии постоянна и равна* $p_n = \frac{\lambda}{n}$, *где* λ — *постоянная величина, не зависящая от* n. *Тогда для любого* k

$$P_n(k) \to \frac{\lambda^k}{k!} e^{-\lambda} \text{ при } n \to \infty.$$

Доказательство. По формуле Бернулли

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}.$$

В нашем случае $p = \frac{\lambda}{n}$, поэтому

$$\begin{aligned} P_n(k) &= \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \frac{n(n-1)\ldots(n-k+1)}{n^k \left(1 - \frac{\lambda}{n}\right)^k} = \\ &= \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \frac{1 \cdot \left(1 - \frac{1}{n}\right) \ldots \left(1 - \frac{k-1}{n}\right)}{\left(1 - \frac{\lambda}{n}\right)^k}. \end{aligned}$$

Но, как известно

$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}, \quad \text{а} \quad \lim_{n\to\infty} \frac{1 \cdot \left(1 - \frac{1}{n}\right) \ldots \left(1 - \frac{k-1}{n}\right)}{\left(1 - \frac{\lambda}{n}\right)^k} = 1.$$

Отсюда и получаем требуемое предельное соотношение. □

Из предельной теоремы Пуассона вытекает

Следствие 14.1 (Приближённая формула Пуассона). *При достаточно больших n и малых p*

$$P_n(k) \approx \frac{\lambda^k}{k!} e^{-\lambda}.$$

Точность этой формулы определяется числом np^2. А именно, для любого k

$$\left| P_n(k) - \frac{\lambda^k}{k!} e^{-\lambda} \right| < np^2.$$

Более того, для любой последовательности попарно различных чисел k_1, k_2, …, k_s, …,

$$\left| P_n(k) - \left(\frac{\lambda^{k_1}}{k_1!} + \frac{\lambda^{k_2}}{k_2!} + \ldots \right) e^{-\lambda} \right| < np^2.$$

Пример 14.1. С базы в магазин отправлено 400 бутылок пива. Вероятность того, что бутылка разобьётся в пути, равна 0,5%. Какова вероятность того, что в пути разобьются: 3 бутылки? от 5 до 7 бутылок?

Решение. В этом примере $n = 400$, $p = 0,005$, $\lambda = np = 2$. По приближённой формуле Пуассона находим

$$P_{400}(3) \approx \frac{2^3}{3!} e^{-2} \approx 0,1804;$$

$$P_{400}(5 \leqslant k \leqslant 7) \approx e^{-2} \left(\frac{2^5}{5!} + \frac{2^6}{6!} + \frac{2^7}{7!} \right) \approx 0,05.$$

15. Случайные величины

Вернёмся к рассмотренному выше примеру с бросанием двух отличимых игральных костей. Сопоставим каждому исходу ω_{jk}, помимо вероятности, сумму $j+k$ очков, выпавших на костях. Получим функцию $X\colon \Omega \to \mathbb{R}$, определённую условием:

$$X(\omega_{jk}) = j + k.$$

Такие функции называют случайными величинами.

Определение 15.1. Пусть $(\Omega,\ p)$ — вероятностное пространство. *Вещественная случайная величина на* Ω — это функция $X\colon \Omega \to \mathbb{R}$.

Случайная величина называется *дискретной*, если множество всех её значений конечно или счётно.

Выпишем явно все значения случайной величины X, с рассмотрения которой мы начали этот параграф.

$$\begin{aligned}
&X(\omega_{11}) = 2;\\
&X(\omega_{12}) = X(\omega_{21}) = 3;\\
&X(\omega_{13}) = X(\omega_{22}) = X(\omega_{31}) = 4;\\
&X(\omega_{14}) = X(\omega_{23}) = X(\omega_{32}) = X(\omega_{41}) = 5;\\
&X(\omega_{15}) = X(\omega_{24}) = X(\omega_{33}) = X(\omega_{42}) = X(\omega_{51}) = 6;\\
&X(\omega_{16}) = X(\omega_{25}) = X(\omega_{34}) = X(\omega_{43}) = X(\omega_{52}) = X(\omega_{61}) = 7;\\
&X(\omega_{26}) = X(\omega_{35}) = X(\omega_{44}) = X(\omega_{53}) = X(\omega_{62}) = 8;\\
&X(\omega_{36}) = X(\omega_{45}) = X(\omega_{54}) = X(\omega_{63}) = 9;\\
&X(\omega_{46}) = X(\omega_{55}) = X(\omega_{64}) = 10;\\
&X(\omega_{56}) = X(\omega_{65}) = 11;
\end{aligned}$$

$$X(\omega_{66}) = 12.$$

Будем обозначать через $\langle X = x \rangle$ событие, состоящее в том, что случайная величина X приняла значение x, и через $\mathbb{P}(X = x)$ — вероятность этого события. Тогда:

$$\begin{aligned}
\langle X = 2 \rangle &= \{\omega_{11}\}; & \mathbb{P}(X = 2) &= 1/36;\\
\langle X = 3 \rangle &= \{\omega_{12},\ \omega_{21}\}; & \mathbb{P}(X = 3) &= 2/36 = 1/18;\\
\langle X = 4 \rangle &= \{\omega_{13},\ \omega_{22},\ \omega_{31}\}; & \mathbb{P}(X = 4) &= 3/36 = 1/12;\\
\langle X = 5 \rangle &= \{\omega_{14},\ \omega_{23},\ \omega_{32},\ \omega_{41}\}; & \mathbb{P}(X = 5) &= 4/36 = 1/9;\\
\langle X = 6 \rangle &= \{\omega_{15},\ \omega_{24},\ \omega_{33},\ \omega_{42},\ \omega_{51}\}; & \mathbb{P}(X = 6) &= 5/36;\\
\langle X = 7 \rangle &= \{\omega_{16},\ \omega_{25},\ \omega_{34},\ \omega_{43},\ \omega_{52},\ \omega_{61}\}; & \mathbb{P}(X = 7) &= 6/36 = 1/6;\\
\langle X = 8 \rangle &= \{\omega_{26},\ \omega_{35},\ \omega_{44},\ \omega_{53},\ \omega_{62}\}; & \mathbb{P}(X = 8) &= 5/36;\\
\langle X = 9 \rangle &= \{\omega_{36},\ \omega_{45},\ \omega_{54},\ \omega_{63}\}; & \mathbb{P}(X = 9) &= 4/36 = 1/9;\\
\langle X = 10 \rangle &= \{\omega_{46},\ \omega_{55},\ \omega_{64}\}; & \mathbb{P}(X = 10) &= 3/36 = 1/12;\\
\langle X = 11 \rangle &= \{\omega_{56},\ \omega_{65}\}; & \mathbb{P}(X = 11) &= 2/36 = 1/18;\\
\langle X = 12 \rangle &= \{\omega_{66}\}; & \mathbb{P}(X = 12) &= 1/36.
\end{aligned}$$

В большинстве случаев бывает нужна информация, представленная в правом столбце, а не сведения о том, какие значения принимает случайная величина на отдельных исходах. Удобно представлять эту информацию таблицей вида

$$\begin{array}{c|c|c|c|c}
X & x_1 & x_2 & \ldots & x_n \\
\hline
\mathbb{P} & p_1 & p_2 & \ldots & p_n
\end{array},$$

где $x_1,\ x_2,\ \ldots,\ x_n$ — все (попарно различные) значения, принимаемые случайной величиной X, а $p_1,\ p_2,\ \ldots,\ p_n$ — вероятности, с которыми величина X принимает эти значения.

Получаемая таблица называется *законом распределения вероятностей* случайной величины X. (Говорят также «*таблица распределения вероятностей*».)

В рассмотренном примере закон распределения вероятностей задаётся таблицей:

X	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Заметим, что поскольку значения x_i попарно различны, события $X = x_i$ попарно несовместны, а поскольку $x_1, \dots, x_n$ — это все значения случайной величины, в каждом исходе какое-нибудь из них обязательно принимается. Значит, сумма всех событий $\langle X = x_i \rangle$ является достоверным событием и, тем самым, сумма вероятностей этих событий равна 1:

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

Пример 15.1. Куплено 5 семян, каждое из которых даст всходы с вероятностью 0,9. Составьте таблицу распределения вероятностей для числа семян, которые дадут всходы.

Решение. Обозначим через X случайную величину, равную количеству семян, которые дадут всходы. Её значениями являются 0, 1, 2, 3, 4, 5 и только эти числа. Вероятности, с которыми принимаются эти значения, определяются по формуле Бернулли с $n = 5$, $p = 0,9$, $q = 1 - p = 0,1$: $P_5(k) = C_5^k (0,9)^k (0,1)^{5-k}$. Таким образом, рассматриваемая случайная величина задаётся таблицей:

X	0	1	2	3	4	5
p_i	$P_5(0)$	$P_5(1)$	$P_5(2)$	$P_5(3)$	$P_5(4)$	$P_5(5)$

Задание 15.1. Вычислить величины $P_5(0), P_5(1), \ldots, P_5(5)$ и заполнить ими нижнюю строчку таблицы.

Упражнение 15.1. Из полного комплекта костей домино наудачу извлекается одна кость. Составьте таблицу распределения вероятностей для случайной величины, равной сумме очков, указанных на половинках кости.

Важными примерами случайных величин являются индикаторы событий.

Пример 15.2. Пусть A — произвольное событие. *Индикатором события* A называется случайная величина I_A, определяемая условием:

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

Как видно из определения, эта случайная величина принимает всего два значения: 1 и 0. При этом $I_A(\omega) = 1$ тогда и только тогда, когда $\omega \in A$, $I_A(\omega) = 0$ тогда и только тогда, когда $\omega \notin A$. Иначе говоря, $[I_A = 1] = A$, $[I_A = 0] = \overline{A}$. Таким образом, эта случайная величина имеет следующий закон распределения вероятностей:

I_A	0	1
$\mathbb{P}$	$1 - \mathbb{P}(A)$	$\mathbb{P}(A)$

16. Арифметические операции над случайными величинами

Арифметические операции над случайными величинами производятся так же, как над любыми функциями.

Определение 16.1. Пусть X, Y — случайные величины на вероятностном пространстве $(\Omega,\ \mathbb{P})$, $\alpha \in \mathbb{R}$. Для любого $\omega \in \Omega$ полагаем по определению:

$$(X+Y)(\omega) = X(\omega) + Y(\omega);$$
$$(\alpha \cdot X)(\omega) = \alpha \cdot X(\omega);$$
$$(X \cdot Y)(\omega) = X(\omega) \cdot Y(\omega),$$

и так далее.

Как видим, всё просто и естественно. Однако составление законов распределения вероятностей для так определённых случайных величин требует внимания.

Пример 16.1. Производится два испытания Бернулли с вероятностью успеха p, B_1 — количество успехов в 1-м испытании, B_2 — во втором. Составим законы распределения вероятностей для случайных величин B_1, B_2, $2B_1$, $B_1 + B_1$, $B_1 + B_2$.

Решение. Законы распределения вероятностей для случайных величин B_1 и B_2 одинаковы и представляются следующими таблицами:

B_1	0	1
	q	p

,

B_2	0	1
	q	p

.

Для $2B_1$ возможными значениями являются 0, 2 и только эти числа. При этом:

$$2B_1 = 0 \Leftrightarrow B_1 = 0, \quad 2B_1 = 2 \Leftrightarrow B_1 = 1,$$

поэтому

$$\mathbb{P}(2B_1 = 0) = \mathbb{P}(B_1 = 0) = q, \quad \mathbb{P}(2B_1 = 2) = \mathbb{P}(B_1 = 1) = p.$$

Таким образом, распределение вероятностей случайной величины $2B_1$ задаётся таблицей

$2B_1$	0	2
	q	p

.

Для величины $X = B_1 + B_1$ логически возможны значения $0+0$, $0+1$, $1+0$ и $1+1$, то есть 0, 1 и 2. Вычислим вероятности, с которыми принимаются эти значения. Ясно, что

$$X = 0 \Leftrightarrow B_1 + B_1 = 0 \Leftrightarrow B_1 = 0, \quad B_1 + B_1 = 2 \Leftrightarrow B_1 = 1.$$

А вот

$$X = 1 \Leftrightarrow (B_1 = 0 \text{ и } B_1 = 1) \text{ или } (B_1 = 1 \text{ и } B_1 = 0).$$

Поскольку случайная величина не может принимать сразу два значения, событие $\langle X = 1\rangle$ на самом деле невозможно, то есть случайная величина $X = B_1 + B_1$ не может принимать значения 1.

Таким образом, распределение вероятностей для X определяется таблицей

$B_1 + B_1$	0	2
	q	p

,

из которой видно, что, как и следовало ожидать, $B_1 + B_1 = 2B_1$.

Наконец, рассмотрим случайную величину $Y = B_1 + B_2$. Для неё логически возможны значения $0+0$, $0+1$, $1+0$ и $1+1$, то есть 0, 1 и 2. Вычислим вероятности, с которыми принимаются эти значения. Ясно, что

$$B_1 + B_2 = 0 \Leftrightarrow B_1 = 0 \text{ и } B_2 = 0, \quad B_1 + B_2 = 2 \Leftrightarrow B_1 = 1 \text{ и } B_2 = 1.$$

Значит,

$$\mathbb{P}(Y = 0) = \mathbb{P}(\langle B_1 = 0\rangle \cdot \langle B_2 = 0\rangle) = \mathbb{P}(B_1 = 0) \cdot \mathbb{P}(B_2 = 0) = q \cdot q = q^2$$

и аналогично

$$\mathbb{P}(Y = 2) = p^2.$$

Наконец

$$Y = 1 \Leftrightarrow B_1 + B_2 = 1 \Leftrightarrow (B_1 = 1 \text{ и } B_2 = 0) \text{ или } (B_1 = 0 \text{ и } B_2 = 1).$$

Значит,

$$\mathbb{P}(Y = 1) = pq + qp = 2pq.$$

Всё это означает, что распределение вероятностей случайной величины $B_1 + B_2$ даётся таблицей:

$B_1 + B_2$	0	1	2
	q^2	$2pq$	p^2

Это и понятно: $B_1 + B_2 = \mu_2$, а распределение для μ_2 именно такое.

Упражнение 16.1. Постройте таблицу распределения вероятностей для $B_1 + B_2 + B_3$ и убедитесь, что она такая же, как и для μ_3.

Упражнение 16.2. Дважды бросается игральная кость. Пусть X_1, соотв., X_2 — количество очков, выпавших при первом, соотв., втором броске. Постройте таблицы распределения вероятностей для случайных величин $X_1 + X_2$, $X_1 - X_2$, $X_1 \cdot X_2$, X_1^2, $X_1^2 + X_2^2$, $X_1^2 - X_2^2$, $(X_1 + X_2)^2$, $(X_1 - X_2)^2$.

17. Математическое ожидание

Пример 17.1. Пусть на предприятии работают N человек и при этом

$$\begin{gathered}
k_1 \text{ человек получают зарплату } x_1, \\
k_2 \text{ человек получают зарплату } x_2, \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots, \\
k_n \text{ человек получают зарплату } x_n,
\end{gathered}$$

где все x_j попарно различны и $k_1 + k_2 + \ldots + k_n = N$. Вычислим среднюю зарплату M на этом предприятии. Имеем:

$$\begin{aligned}
M &= \frac{x_1 k_1 + x_2 k_2 + \ldots + x_n k_n}{k_1 + k_2 + \ldots + k_n} \\
&= \frac{x_1 k_1 + x_2 k_2 + \ldots + x_n k_n}{N} \\
&= x_1 \frac{k_1}{N} + x_2 \frac{k_2}{N} + \ldots + x_n \frac{k_n}{N}.
\end{aligned}$$

Переведём рассмотренный пример на язык теории вероятностей. Для каждого работника ω этого предприятия обозначим через $X(\omega)$ его зарплату. Получим пространство элементарных событий $\Omega = \{\omega_1,\ \omega_2,\ \ldots,\ \omega_N\}$ и случайную величину X на нём, определённую условием:

$$X(\omega) = \{\text{зарплата работника } \omega\}.$$

Вероятность того, что определённая нами случайная величина X принимает значение x_j, есть вероятность того, что выбранный наугад работник получает зарплату x_j, и равна отношению количества k_j работников, получающих зарплату x_j, к общему числу N всех работников:

$$\mathbb{P}(X = x_j) = \frac{k_j}{N}.$$

Переходя к этим обозначениям, обнаруживаем, что вычисленная выше средняя зарплата равна

$$\begin{aligned} M &= x_1\frac{k_1}{N} + x_2\frac{k_2}{N} + \ldots + x_n\frac{k_n}{N} \\ &= x_1\mathbb{P}(X = x_1) + x_2\mathbb{P}(X = x_2) + \ldots + x_n\mathbb{P}(X = x_n) \\ &= \sum_{j=1}^{n} x_j p_j. \end{aligned}$$

То, что у нас получилось, называют средним значением, или математическим ожиданием случайной величины.

Определение 17.1. *Математическим ожиданием* случайной величины X называют число

$$\mathbb{M}(X) = \sum_{i=1}^{n} x_j p_j,$$

где $x_1, \ldots, x_n$ — полный перечень значений величины X, $p_1, \ldots, p_n$ — вероятности, с которыми величина X принимает эти значения.

Пример 17.2. Предприниматель размышляет над тем, куда лучше вложить деньги: в киоск для торговли мороженым или в палатку для торговли хлебом.

В первом случае он с вероятностью 0,5 обеспечит себе годовую прибыль в 5000 у.е., с вероятностью 0,2 — в 10000 у.е. и с вероятностью 0,3 — в 3000 у.е.

Во втором случае прогноз таков: 5500 с вероятностью 0,6, 5000 с вероятностью 0,3 и 6500 с вероятностью 0,1.

В каком случае математическое ожидание дохода больше?

Решение. Для каждого из двух возможных решений годовая прибыль является случайной величиной. Обозначим эти величины через X и Y. Построим их таблицы распределения.

X	3000	5000	10000
p	0,3	0,5	0,2

Y	5000	5500	6500
p	0,3	0,6	0,1

Найдём математические ожидания:

$$\mathbb{M}(X) = 3000 \cdot 0,3 + 5000 \cdot 0,5 + 10000 \cdot 0,2 = 5400 \text{ долл.},$$
$$\mathbb{M}(Y) = 5000 \cdot 0,3 + 5500 \cdot 0,6 + 6500 \cdot 0,1 = 5450 \text{ долл.},$$

Получается, что $\mathbb{M}(X) < \mathbb{M}(Y)$. Таким образом, математическое ожидание для палатки больше.

Определение 17.2 (Для дискретных случайных величин)**.** Случайные величины X и Y называют *независимыми*, если они принимают свои значения независимо одна от другой, т. е. для любого значения x величины X и любого значения y величины Y события $\langle X = x\rangle$ и $\langle Y = y\rangle$ независимы.

Теорема 17.1 (Свойства математического ожидания)**.**

1° $\mathbb{M}(\mathbf{0}) = 0$, $\mathbb{M}(\mathbf{1}) = 1$ *и, вообще,* $\mathbb{M}(\alpha \cdot \mathbf{1}) = \alpha$ *для любого* $\alpha \in \mathbb{R}$.

2° *Для любого события* A, $\mathbb{M}(I_A) = \mathbb{P}(A)$.

3° $\mathbb{M}(X + Y) = \mathbb{M}(X) + \mathbb{M}(Y)$.

4° $\mathbb{M}(\alpha \cdot X) = \alpha \cdot \mathbb{M}(X)$.

5° $\mathbb{M}(\alpha_1 X_1 + \ldots + \alpha_s X_s) = \alpha_1 \mathbb{M}(X_1) + \ldots + \alpha_s \mathbb{M}(X_s)$.

6° $\mathbb{M}(X - \mathbb{M}(X)) = 0$.

7° *Если* $X \geqslant 0$, *то* $\mathbb{M}(X) \geqslant 0$.

8° *Если* $X \leqslant Y$, *то* $\mathbb{M}(X) \leqslant \mathbb{M}(Y)$.

9° *Если* $X \geqslant 0$, *и* $\mathbb{M}(X) = 0$, *то* $\mathbb{P}(X = 0) = 1$.

10° *Если случайные величины* X *и* Y *независимы, то математическое ожидание произведения этих величин равно произведению их математических ожиданий:*

$$\mathbb{M}(X \cdot Y) = \mathbb{M}(X) \cdot \mathbb{M}(Y).$$

Пример 17.3. Рассмотрим последовательность испытаний Бернулли с вероятностью успеха p. Пусть B_j — случайная величина, равная количеству успехов в испытании номер j, то есть равная 1, если испытание номер j завершается успехом, и 0, если неуспехом. Найдём $\mathbb{M}(B_j)$.

Решение. $\mathbb{M}(B_j) = 1 \cdot p + 0 \cdot (1 - p) = p$.

Пример 17.4. Обозначим через μ_n количество успехов в серии из n испытаний Бернулли с вероятностью успеха p. Найдём $\mathbb{M}(\mu_n)$.

Решение. Ясно, что

$$\mu_n = B_1 + B_2 + \ldots + B_n = \sum_{j=1}^{n} B_j.$$

Применяя свойство 5°, получаем:

$$\begin{aligned}\mathbb{M}(\mu_n) &= \mathbb{M}(B_1) + \mathbb{M}(B_2) + \ldots + \mathbb{M}(B_n) = \\ &= \underbrace{p + p + \ldots + p}_{n \text{ слагаемых}} = \\ &= np.\end{aligned}$$

Таким образом,

$$\mathbb{M}(\mu_n) = np.$$

18. Дисперсия

Для общего представления о распределении случайной величины важно знать не только её среднее значение, но и то, насколько отклоняется величина от этого среднего значения.

Рассмотрим, например, случайные величины X и Y со следующими законами распределения:

X	101	99
p	0,5	0,5

Y	0	200
p	0,5	0,5

Первая имеет место, например, если Вы с вероятностью 0,5 можете заработать 101 у.е. и с вероятностью 0,5 — 99 у.е., вторая, если Вы с вероятностью 0,5 можете заработать 200 у.е., и с вероятностью 0,5 — ничего.

Математические ожидания обеих величин одинаковы:

$$\mathbb{M}(X) = \mathbb{M}(Y) = 100.$$

Но для величины X её отклонение от $\mathbb{M}(X)$ незначительно, а вот для величины Y — весьма заметно. Это отклонение является показателем предсказуемости: в первом случае — стабильная, предсказуемая прибыль около 100 у. е., во втором случае — риск. Что́ Вы предпочтёте — это уже Ваше дело, но, принимая решение, Вы должны знать, насколько Вы рискуете.

Для измерения уклонения случайной величины от её математического ожидания используются различные показатели. Одним из таких показателей является дисперсия.

Определение 18.1. *Дисперсией* случайной величины X называется и через $\mathbb{D}(X)$ обозначается математическое ожидание квадрата отклонения случайной величины от её математического ожидания:

$$\mathbb{D}(X) = \mathbb{M}((X - \mathbb{M}(X))^2).$$

Объясним, как она получается. Пусть X — произвольная величина и $M(X)$ — её математическое ожидание. Вычитая из X число $M(X)$ получаем новую случайную величину

$$X - \mathbb{M}(X).$$

Квадрат этой случайной величины также является случайной величиной:

$$(X - \mathbb{M}(X))^2.$$

Математическое ожидание этой случайной величины и есть дисперсия случайной величины X.

Теорема 18.1 (Свойства дисперсии)**.**

1° $\mathbb{D}(X) \geqslant 0$.

2° $\mathbb{D}(X) = \mathbb{M}(X^2) - (\mathbb{M}(X))^2$.

3° *Для любого* $\alpha \in \mathbb{R}$

$$\mathbb{D}(\alpha X) = \alpha^2 \mathbb{D}(X), \quad \mathbb{D}(X + \alpha) = \mathbb{D}(X).$$

4° *Если случайные величины* $X_1, X_2, \ldots, X_n$ *независимы, то*

$$\mathbb{D}(X_1 + X_2 + \ldots + X_n) = \mathbb{D}(X_1) + \mathbb{D}(X_2) + \ldots + \mathbb{D}(X_n).$$

Пример 18.1. Рассмотрим опять последовательность испытаний Бернулли с вероятностью успеха p. Обозначим через B_j количество успехов в испытании номер j, $j \in \mathbb{N}$, а через μ_n — количество успехов в серии из n испытаний (так что

$$\mu_n = B_1 + B_2 + \ldots + B_n).$$

Найдём $\mathbb{D}(B_j)$ и $\mathbb{D}(\mu_n)$.

Решение. Для любого j,

$$\mathbb{D}(B_j) = \mathbb{M}(B_j^2) - (\mathbb{M}(B_j))^2 = \mathbb{M}(B_j) - (\mathbb{M}(B_j))^2 = p - p^2 = \mathbb{P}(1-p) = pq,$$

а поскольку испытания независимы, то

$$\mathbb{D}(\mu_n) = \mathbb{D}(B_1 + \ldots + B_n) = \mathbb{D}(B_1) + \ldots + \mathbb{D}(B_n) = n \cdot pq.$$

19. Основные виды законов распределения дискретных случайных величин

1. Распределение вероятностей для случайной величины X, принимающей n значений $x_1, \ldots, x_n$ с одинаковыми вероятностями, задаётся формулой

$$p_k = P_n(X = x_k) = \frac{1}{n}.$$

Такое распределение называется *конечным дискретно-равномерным.*

2. Распределение вероятностей для числа μ_n успехов в последовательности из n испытаний Бернулли задаётся формулой

$$p_k = P(\mu_n = k) = P_n(k) = C_n^k p^k q^{n-k}, \quad k = 0, \ \ldots, \ n.$$

Такое распределение вероятностей называется *биномиальным.* (Название объясняется тем, что значения вероятности вычисляются по той же формуле, что и члены разложения в формуле бинома Ньютона).

3. Пусть производится последовательность испытаний Бернулли, X — номер первого испытания, завершившегося успехом. Распределение вероятностей для этой случайной величины задаётся формулой

$$p_k = P(X = k) = pq^{m-1}, \quad k = 0, \ 1, \ 2, \ \ldots \ .$$

Такое распределение вероятностей называется *геометрическим.* (Название объясняется тем, что значения вероятности составляют геометрическую прогрессию).

4. *Распределение Пуассона* задаётся формулой

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

20. Непрерывные случайные величины

Напомним, что случайную величину называют непрерывной, если множество её значений является промежутком. Задание такой случайной величины таблицей распределения вероятностей оказывается невозможным. Этому — две причины. Первая состоит в том, что элементы промежутка — значения случайной велчины — не могут быть перечислены (как говорят, промежуток несчётен). Вторая причина состоит в том, что непрерывная случайная величина каждое отдельное значение принимает с вероятностью, равной нулю: для любого x

$$\mathbb{P}(X = x) = 0.$$

Поэтому для задания непрерывных случайных величин используют другие средства.

При рассмотрении непрерывной случайной величины чаще всего требуется знать, какова вероятность того, что её значения принадлежат тому или иному промежутку. В соответствии с этим для задания непрерывной случайной величины нужно указать все вероятности вида $\mathbb{P}(\alpha < X < \beta)$, где $\alpha,\ \beta \in \mathbb{R}$ или $\alpha = -\infty$ или $\beta = +\infty$. Но, оказывается, достаточно указать лишь вероятности вида $\mathbb{P}(X < \beta)$ с $\beta \in \mathbb{R}$ — тогда и все вероятности вида $\mathbb{P}(\alpha < X < \beta)$ определятся.

Действительно, для любых $\alpha,\ \beta \in \mathbb{R}$ при $\alpha < \beta$ имеем:

$$\mathbb{P}(X < \beta) = \mathbb{P}(X < \alpha) + \mathbb{P}(X = \alpha) + \mathbb{P}(\alpha < X < \beta) = \mathbb{P}(X < \alpha) + \mathbb{P}(\alpha < X < \beta),$$

откуда

$$\mathbb{P}(\alpha < X < \beta) = \mathbb{P}(X < \beta) - \mathbb{P}(X < \alpha).$$

Определение 20.1. Функцию $F_X\colon \mathbb{R} \to \mathbb{R}$, связанную со случайной величиной X условием

$$F_X(x) = \mathbb{P}(X < x),$$

называют *функцией распределения случайной величины* X.

В дискретном случае имеется возможность задать случайную величину таблицей распределения и построить по этой таблице функцию распределения. В непрерывном случае такой возможности нет. Здесь задание функции распределения является способом задания случайной величины.

Из непрерывных случайных величин чаще всего рассматривают такие, функции распределения которых обладают производной во всех точках, за исключением, может быть, конечного их числа.

Определение 20.2. Пусть функция распределения F_X случайной величины X обладает производной во всех точках за исключением, может быть, конечного их числа. Тогда производную функции распределения называют *плотностью распределения вероятностей случайной величины* X и обозначают через p_X или f_X. Часто говорят короче: плотность распределения случайной величины X, или ещё короче: плотность случайной величины X. Индекс 'X' часто опускают (когда это не ведёт к недоразумениям).

Таким образом, плотность p_X (или f_X) распределения случайной величины X связана с функцией распределения F_X соотношением:

$$p_X = f_X = F'_X.$$

С другой стороны, функция распределения F непрерывной случайной величины X может быть определена по заданной плотности f формулой

$$F(x) = \int_{-\infty}^{x} f(t)\,dt.$$

Благодаря этому вероятность попадания непрерывной случайной величины X в заданный интервал $(\alpha,\ \beta)$ определяется по формуле

$$\mathbb{P}(\alpha < X < \beta) = F(\beta) - F(\alpha) = \int\limits_{\alpha}^{\beta} f(x)\, dx,$$

где $F(x)$ — функция распределения, $f(x)$ — плотность вероятности. Такой же будет вероятность попадания на отрезок $[\alpha, \beta]$ и на любой из полуотрезков $(\alpha, \beta]$ и $[\alpha, \beta)$.

Предложение 20.1 (Основные свойства функций распределения)**.**

1° *Функция распределения является неубывающей на всей числовой прямой.*

2° *Функция распределения непрерывна слева в каждой точке числовой прямой.*

3° $F(-\infty) = 0, \quad F(+\infty) = 1$ (*т. е.* $\lim\limits_{x\to-\infty} F(x) = 0,\ \lim\limits_{x\to+\infty} F(x) = 1$).

Математическое ожидание $\mathbb{M}(X)$ и дисперсия $\mathbb{D}(X)$ непрерывной случайной величины X определяются формулами

$$\mathbb{M}(X) = \int\limits_{-\infty}^{\infty} x f(x)\, dx, \quad \mathbb{D}(X) = \int\limits_{-\infty}^{\infty} [x - \mathbb{M}(X)]^2 f(x)\, dx,$$

где $f(x)$ — плотность вероятности. Как и в дискретном случае, $\mathbb{D}(X)$ можно также вычислять по формуле

$$\mathbb{D}(X) = \mathbb{M}(X^2) - \mathbb{M}^2(X),$$

где $\mathbb{M}(X^2) = \int\limits_{-\infty}^{\infty} x^2 f(x)\, dx.$

Определение 20.3. Случайную величину X называют *стандартной*, если её математическое ожидание равно нулю, а дисперсия — единице:

$$\mathbb{M}(X) = 0, \quad \sigma^2(X) = 1.$$

Теорема 20.1 (Теорема стандартизации)**.** *Пусть X — случайная величина, a — её математическое ожидание, σ^2 — дисперсия (т.е. $\mathbb{M}(X) = a$, $\mathbb{D}(X) = \sigma^2$). Тогда:*

1° *случайная величина*

$$\mathrm{St}(X) = \frac{X - a}{\sigma}$$

стандартна;

2° *для любых α, $\beta \in \mathbb{R}$*

$$\mathbb{P}(X < \beta) = \mathbb{P}\left(\mathrm{St}(X) < \frac{\beta - a}{\sigma}\right),$$

$$\mathbb{P}(X > \alpha) = \mathbb{P}\left(\mathrm{St}(X) > \frac{\alpha - a}{\sigma}\right),$$

$$\mathbb{P}(\alpha < X < \beta) = \mathbb{P}\left(\frac{\alpha - a}{\sigma} < \mathrm{St}(X) < \frac{\beta - a}{\sigma}\right).$$

Доказательство. Это тотчас же следует из свойств математического ожидания и дисперсии. □

Упражнение 20.1. Проведите доказательство во всех деталях.

Пример 20.1. Случайная величина X задана функцией распределения

$$F(x) = \begin{cases} 0, & \text{если } x \leqslant -1, \\ \dfrac{1}{3}x + \dfrac{1}{3}, & \text{если } -1 < x \leqslant 2, \\ 1, & \text{если } x > 2. \end{cases}$$

Найти вероятность того, что в результате испытания она примет значение

а) меньшее -2,
б) меньшее 1,
в) меньшее 1 и большее 0,
г) не меньшее 1,
д) не меньшее 3.

Решение.

а) $$\mathbb{P}(-\infty < X < -2) = F(-2) - F(-\infty) = 0 - 0 = 0.$$

б) $$\mathbb{P}(-\infty < X < 1) = F(1) - F(-\infty) = \left(\frac{1}{3}x + \frac{1}{3}\right)\Big|_{x=1} - 0 = \frac{2}{3}.$$

в) $$\mathbb{P}(0 < X < 1) = F(1) - F(0) = \left(\frac{1}{3}x + \frac{1}{3}\right)\Big|_{x=1} - \left(\frac{1}{3}x + \frac{1}{3}\right)\Big|_{x=0} =$$
$$= \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.$$

г) $$\mathbb{P}(X \geqslant 1) = P(1 \leqslant X < \infty) = F(\infty) - F(1) = 1 - \frac{2}{3} = \frac{1}{3}.$$

д) $$\mathbb{P}(X \geqslant 3) = \mathbb{P}(3 \leqslant X < \infty) = F(\infty) - F(3) = 1 - 1 = 0.$$

Пример 20.2. Случайная величина X задана плотностью

$$f(x) = \begin{cases} 0, & \text{если } x \leqslant 0 \\ \dfrac{1}{2}\sin x, & \text{если } 0 < x \leqslant \pi \\ 0, & \text{если } x > \pi. \end{cases}$$

Найти

а) функцию распределения $F(x)$,

б) вероятность того, что в результате испытания случайная величина X примет значение, заключенное в интервале $(0,\ \pi/3)$.

Решение. а) $F(x) = \int\limits_{-\infty}^{x} f(t)\,dt.$

При $x \leqslant 0$

$$F(x) = \int\limits_{-\infty}^{x} 0\,dt = C\Big|_{-\infty}^{x} = C - C = 0.$$

При $0 < x \leqslant \pi$

$$\begin{aligned} F(x) &= \int\limits_{-\infty}^{x} f(t)\,dt = \int\limits_{-\infty}^{0} 0\,dt + \int\limits_{0}^{x} \frac{1}{2}\sin t\,dt = \\ &= 0 + \frac{1}{2}(-\cos t)\Big|_{0}^{x} = \frac{1}{2}\cos t\Big|_{x}^{0} = \frac{1}{2}(\cos 0 - \cos x) = \frac{1}{2}(1 - \cos x). \end{aligned}$$

При $x > \pi$

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{0} 0\,dt + \int_{0}^{\pi} \frac{1}{2}\sin t\,dt + \int_{\pi}^{x} 0\,dt = \\ &= 0 + \frac{1}{2}\int_{0}^{\pi} \sin t\,dt + 0 = \frac{1}{2}(-\cos t)\Big|_0^{\pi} = \\ &= \frac{1}{2}\cos t\Big|_{\pi}^{0} = \frac{1}{2}(\cos 0 - \cos\pi) = \frac{1}{2}\cdot 2 = 1. \end{aligned}$$

Таким образом,

$$F(x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ \frac{1}{2}(1-\cos x), & \text{если } 0 < x \leqslant \pi, \\ 1, & \text{если } x > \pi. \end{cases}$$

$$\begin{aligned} \text{б) } \mathbb{P}\left(0 < X < \frac{\pi}{3}\right) &= \int_{0}^{\pi/3} \frac{1}{2}\sin t\,dt = \frac{1}{2}(-\cos t)\Big|_0^{\pi/3} = \\ &= \frac{1}{2}\cos t\Big|_{\pi/3}^{0} = \frac{1}{2}\left(\cos 0 - \cos\frac{\pi}{3}\right) = \frac{1}{2}\left(1-\frac{1}{2}\right) = \frac{1}{4}. \end{aligned}$$

Пример 20.3. Случайная величина X задана функцией распределения

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

Найти $\mathbb{M}(X)$ и $\mathbb{D}(X)$.

Решение.

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

$$\begin{aligned} \mathbb{M}(X) &= \int\limits_{-\infty}^{\infty} x f(x)\, dx = \int\limits_{-\infty}^{0} 0 \cdot x\, dx + \int\limits_{0}^{1} x \cdot 2x\, dx + \int\limits_{1}^{\infty} x \cdot 0\, dx = \\ &= c\Big|_{-\infty}^{0} + 2\int\limits_{0}^{1} x^2\, dx + c\Big|_{1}^{\infty} = 0 + 2\frac{x^3}{3}\Big|_0^1 + 0 = \\ &= \frac{2}{3}(1^3 - 0^3) = \frac{2}{3}, \quad M^2(X) = \frac{4}{9}, \end{aligned}$$

$$\begin{aligned} \mathbb{M}(X^2) &= \int\limits_{-\infty}^{\infty} x^2 f(x)\, dx = \int\limits_{0}^{1} x^2 \cdot 2x\, dx = \\ &= 2\int\limits_{0}^{1} x^3\, dx = 2\frac{x^4}{4}\Big|_0^1 = \frac{x^4}{2}\Big|_0^1 = \frac{1}{2}, \end{aligned}$$

$$\mathbb{D}(X) \;=\; \mathbb{M}(X^2) - \mathbb{M}^2(X) = \frac{1}{2} - \frac{4}{9} = \frac{1}{18} \approx 0{,}06.$$

Упражнение 20.2. Случайная величина X задана функцией распределения

$$F(x) = \begin{cases} 0, & \text{если } x \leqslant 2 \\ \dfrac{1}{2}x - 1, & \text{если } 2 < x \leqslant 4 \\ 1, & \text{если } x > 4. \end{cases}$$

Найти вероятность того, что в результате испытания она примет значение

а) меньшее 1,

б) меньшее 3,

в) меньшее 3,5 и большее 2,5,

г) не меньшее 3,

д) не меньшее 5.

Ответ: а) 0, б) 1/2, в) 1/2, г) 1/2, д) 0.

Упражнение 20.3. Случайная величина X задана плотностью

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

Найти

а) функцию распределения $F(x)$,

б) вероятность того, что в результате испытания случайная величина X примет значение в интервале $(0, \pi/6)$.

Ответ:

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

б) $\mathbb{P}(0 < X < \pi/6) = 1/4 = 0{,}25.$

Упражнение 20.4. Случайная величина X задана функцией распределения

$$F(x) = \begin{cases} 0, & \text{если } x \leqslant 0 \\ \frac{1}{2}(x^2 + x^3), & \text{если } 0 < x \leqslant 1 \\ 1, & \text{если } x > 1. \end{cases}$$

Найти $\mathbb{M}(X)$ и $\mathbb{D}(X)$.

Ответ: $\mathbb{M}(X) = 17/24 \approx 0{,}71, \quad \mathbb{D}(X) \approx 0{,}05.$

21. Нормальное распределение

Непрерывная случайная величина X называется *нормально распределенной с параметрами a и σ*, если её плотность вероятности $f(x)$ имеет вид

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

Можно показать, что в таком случае

$$\mathbb{M}(X) = a, \quad \mathbb{D}(X) = \sigma^2.$$

Иначе говоря, для нормально распределённой случайной величины X с параметрами a и σ число a является её математическим ожиданием, число σ — стандартным отклонением.

Полагая $a = 0$, $\sigma = 1$ в определении нормально распределённой случайной величины, получаем, что *плотностью распределения стандартной нормально распределённой случайной величины является функция Гаусса*

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

а функцией распределения —

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} \mathrm{d}t = \frac{1}{2} + \Phi(x),$$

где $\Phi(x)$ — *функция Лапласа.*

Стандартную нормально распределённую случайную величину будем обозначать через U. Повторим, что плотностью распределения для U является функция Гаусса φ, функцией распределения — функция $\frac{1}{2} + \Phi(x)$.

Предложение 21.1. *Вероятность попадания нормальной случайной величины X с математическим ожиданием a и дисперсией σ^2 в заданный промежуток $(\alpha,\ \beta)$ определяется по формуле*

$$\mathbb{P}(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right).$$

Доказательство. Действительно,

$$\mathbb{P}(\alpha < X < \beta) = \mathbb{P}\left(\frac{\alpha - a}{\sigma} < \frac{X - a}{\sigma} < \frac{\beta - a}{\sigma}\right) = \mathbb{P}\left(\frac{\alpha - a}{\sigma} < U < \frac{\beta - a}{\sigma}\right) =$$

$$= \Psi\left(\frac{\beta - a}{\sigma}\right) - \Psi\left(\frac{\alpha - a}{\sigma}\right) = \left(\frac{1}{2} + \Phi\left(\frac{\beta - a}{\sigma}\right)\right) - \left(\frac{1}{2} + \Phi\left(\frac{\alpha - a}{\sigma}\right)\right) =$$
$$= \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right).$$

□

Предложение 21.2. *Вероятность отклонения нормально распределенной случайной величины X от её математического ожидания a не более чем на δ определяется по формуле*

$$\mathbb{P}(|X - a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right).$$

В частности,

$$\mathbb{P}(|X - a| < \sigma) = 2\Phi(1) \approx 0,6826,$$
$$\mathbb{P}(|X - a| < 2\sigma) = 2\Phi(2) \approx 0,9544,$$
$$\mathbb{P}(|X - a| < 3\sigma) = 2\Phi(3) \approx 0,9973.$$

Доказательство. Действительно,

$$\mathbb{P}(|X - a| < \delta) = \mathbb{P}(-\delta < X - a < \delta) = \mathbb{P}\left(\frac{-\delta}{\sigma} < \frac{X - a}{\sigma} < \frac{\delta}{\sigma}\right) =$$
$$= \mathbb{P}\left(\frac{-\delta}{\sigma} < U < \frac{\delta}{\sigma}\right) = \Phi\left(\frac{\delta}{\sigma}\right) - \Phi\left(-\frac{\delta}{\sigma}\right) = 2\Phi\left(\frac{\delta}{\sigma}\right).$$

□

В силу последнего утверждения доказанного предложения, случайное событие, состоящее в том, что нормально распределённая случайная величина X отклоняется от своего математического ожидания менее чем на 3σ, можно считать практически достоверным, т. к. его вероятность близка к 1. Этот факт называют «*правилом* 3σ».

Пример 21.1. Случайная величина X распределена нормально. Ее математическое ожидание равно 6, дисперсия равна 9. Найти вероятность того, что в результате испытания X примет значение, заключенное в интервале $(3, 12)$.

Решение. $a = 6, \quad \sigma^2 = 9$, следовательно, $\sigma = 3$.

$$P(3 < X < 12) = \Phi\left(\frac{12-6}{3}\right) - \Phi\left(\frac{3-6}{3}\right) = \Phi(2) - \Phi(-1) =$$

$$= \Phi(2) + \Phi(1) \approx 0{,}4772 + 0{,}3413 = 0{,}8185.$$

Пример 21.2. Случайная величина X распределена нормально. Её дисперсия равна 4. Найти вероятность того, что абсолютная величина отклонения случайной величины X от её математического ожидания a не превосходит 3.

Решение. Численное значение a не задано. Однако, оно и не требуется для решения задачи. По условию $\sigma^2 = 4, \quad \sigma = 2$. Значит

$$\mathbb{P}(|X-a| < 3) = \mathbb{P}(-3 < X-a < 3) = \mathbb{P}\left(\frac{-3}{\sigma} < \frac{X-a}{\sigma} < \frac{3}{\sigma}\right) = 2\Phi\left(\frac{3}{2}\right) = 2\Phi(1{,}5) \approx 2\cdot 0{,}4332$$

Пример 21.3. Случайная величина X распределена нормально, $a = 6$, $\sigma = 3$. Найти интервал $(\alpha, \ \beta)$ с серединой в точке a, в который с вероятностью 0,9973 попадает X в результате испытания.

Решение.

$$P(\alpha < X < \beta) = 0{,}9973.$$

По «правилу 3σ»

$$P(|X-a| < 3\sigma) = 0{,}9973$$

или

$$P(a - 3\sigma < X < a + 3\sigma) \approx 0{,}9973.$$

Откуда

$$\alpha = a - 3\sigma = 6 - 3 \cdot 3 = -3,$$
$$\beta = a + 3\sigma = 6 + 3 \cdot 3 = 15.$$

Т. о. искомый интервал — $(-3,\ 15)$.

Пример 21.4. Автомат фасует рис в пачки так, что вес пачек с рисом является нормально распределенной случайной величиной с математическим ожиданием $a = 1$ кг и стандартным отклонением $\sigma = 10$ г.

1) Какой процент пачек имеет вес: а) менее 990 г; б) от 990 г до 1010 г; в) более 1010 г?

2) В каком интервале с серединой в точке a заключён вес произвольно выбранной пачки с вероятностью: а) 0,9974; б) 0,9; в) 0,95; г) 0,99?

Решение. Обозначим через X случайную величину, выражающую вес произвольно выбранной пачки с рисом. По условию $\mathbb{M}(X) = a = 1000$ г, $\sigma(X) = \sigma = 10$ г.

1а) Процент пачек, имеющих вес менее 990 г, — это вероятность того, что вес выбранной наугад пачки окажется менее 990 г и не менее 0, то есть $\mathbb{P}(0 \leqslant X < 990) = \mathbb{P}(0 < X < 990)$. При этом

$$\mathbb{P}(0 < X < 990) = \mathbb{P}\left(\frac{0-a}{\sigma} < \frac{X-a}{\sigma} < \frac{990-a}{\sigma}\right) = \mathbb{P}\left(\frac{-1000}{10} < U < \frac{990-1000}{10}\right) =$$
$$= \mathbb{P}(-100 < U < -1) = \Phi(-1) - \Phi(-100) = \Phi(100) - \Phi(1) \approx \frac{1}{2} - \Phi(1).$$

По таблице значений функции Лапласа находим её значение в точке 1: $\Phi(1) = 0,3413$.

Таким образом, $\mathbb{P}(0 < X < 990) = 0,5 - 0,3413 = 0,1587$. Это означает, что искомый процент равен 15,87%.

1б), 1в) Рассуждая аналогично предыдущему, получаем:

$$\mathbb{P}(990 < X < 1010) = \mathbb{P}(-1 < U < 1) = \Phi(1) - \Phi(-1) = 2\Phi(1) = 0,6826;$$
$$\mathbb{P}(X > 1010) = \mathbb{P}(U > 1) = 1 - \Psi(1) = 1 - (0,5 + \Phi(1)) = 0,5 - \Phi(1) = 0,1587.$$

Иначе говоря, искомые проценты равны, соответственно, 68,26% и 15,87%.

2а) По правилу трёх сигм 0,9974 — это $\mathbb{P}(|X-a| < 3\sigma)$. В данном случае $\mathbb{P}(|X-a| < 3\sigma) = \mathbb{P}(|X-a| < 30)$. Значит, с вероятностью 0,9974 вес пачек с рисом, расфасованных этим автоматом, заключён между 970 г и 1030 г.

2б) В силу предложения 21.2

$$\mathbb{P}(|X-a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right).$$

В таблице значений функции Лапласа находим значение $\Phi(x)$, для которого $2\Phi(x) = 0,9$, то есть значение $\Phi(x) = 0,45$ (или ближайшее к нему). Ближайшим является 0,4505. Оно соответствует аргументу $x = 1,65$. Это означает, что

$$\frac{\delta}{\sigma} = 1,65, \text{ то есть } \frac{\delta}{10} = 1,65,$$

откуда $\delta = 16,5$. Таким образом, с вероятностью 0,9 вес пачек с рисом, расфасованных этим автоматом, заключён между 983,5 г и 1016,5 г.

2в, г) Эти случаи рассматриваются аналогично. Предлагаем читателю разобрать их самостоятельно.

Упражнение 21.1. Случайная величина X распределена нормально. Ее математическое ожидание равно 10, дисперсия равна 16. Найти вероятность того, что в результате испытания X примет значение, заключенное в интервале $(5,\ 15)$.

Ответ. $P(5 < X < 15) \approx 0{,}7888$.

Упражнение 21.2. Случайная величина X распределена нормально. Ее дисперсия равна 25. Найти вероятность того, что отклонение X от ее математического ожидания по абсолютной величине не превосходит 4.

Ответ. $P(|X - a| < 4) \approx 0{,}5762$.

Упражнение 21.3. Случайная величина X распределена нормально, $a = 5$, $\sigma = 4$. Найти интервал $(\alpha,\ \beta)$, в который с вероятностью 0,9973 попадает X в результате испытания, если a является серединой этого интервала.

Ответ. $(-7,\ 17)$.

22. Законы больших чисел

Теорема 22.1 (Неравенство Чебышёва). *Пусть X — неотрицательная случайная величина на вероятностном пространстве $(\Omega,\ p)$. Тогда, для любого $\varepsilon > 0$,*

$$\mathbb{P}(X \geqslant \varepsilon) \leqslant \frac{\mathbb{M}(X)}{\varepsilon}.$$

Доказательство. Заметим, что, для любого исхода $\omega \in \Omega$, либо $X(\omega) \geqslant \varepsilon$, либо $X(\omega) < \varepsilon$. Иначе говоря,

$$\langle X \geqslant \varepsilon \rangle + \langle X < \varepsilon \rangle = \Omega,$$

причём слагаемые несовместны. На языке индикаторов это означает, что

$$I_{\langle X \geqslant \varepsilon \rangle} + I_{\langle X < \varepsilon \rangle} = I_\Omega = \mathbf{1}.$$

Поэтому

$$X = X \cdot \mathbf{1} = X \cdot \left(I_{\langle X \geqslant \varepsilon \rangle} + I_{\langle X < \varepsilon \rangle}\right) = X \cdot I_{\langle X \geqslant \varepsilon \rangle} + X \cdot I_{\langle X < \varepsilon \rangle} \geqslant X \cdot I_{\langle X \geqslant \varepsilon \rangle} \geqslant \varepsilon \cdot I_{\langle X \geqslant \varepsilon \rangle}.$$

Отсюда по свойствам математических ожиданий получаем:

$$\mathbb{M}(X) \geqslant M\big(\varepsilon \cdot I_{\langle X \geqslant \varepsilon \rangle}\big) \geqslant \varepsilon M\big(I_{\langle X \geqslant \varepsilon \rangle}\big) = \varepsilon \cdot p\,\langle X \geqslant \varepsilon \rangle,$$

откуда и вытекает требуемое. □

Следствие 22.1 (Следствия неравенства Чебышёва)**.** *Пусть* X *— произвольная случайная величина на вероятностном пространстве* $(\Omega,\ p)$*. Тогда, для любого* $\varepsilon > 0$*:*

1°

$$\mathbb{P}\big(|X| \geqslant \varepsilon\big) \leqslant \frac{\mathbb{M}(|X|)}{\varepsilon};$$

2°

$$\mathbb{P}\big(|X| \geqslant \varepsilon\big) = \mathbb{P}\big(X^2 \geqslant \varepsilon^2\big) \leqslant \frac{\mathbb{M}(X^2)}{\varepsilon^2};$$

3°

$$\mathbb{P}\big(|X - \mathbb{M}(X)| \geqslant \varepsilon\big) \leqslant \frac{\mathbb{D}(|X|)}{\varepsilon^2}.$$

Теорема 22.2 (Закон больших чисел в форме Чебышёва)**.** *Пусть* $X_1,\ X_2,\ \ldots,\ X_n,\ \ldots$ *— попарно независимые случайные величины с ограниченными в совокупности дисперсиями (то есть такие, что существует константа* c *такая, что, при всех* i*,* $D[X_i] \leqslant c$*). Тогда, для любого* $\varepsilon > 0$*,*

$$\mathbb{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \frac{\mathbb{M}(X_1) + \ldots + \mathbb{M}(X_n)}{n}\right| \geqslant \varepsilon\right) \to 0 \ \textit{при}\ n \to \infty.$$

Доказательство. Определим случайную величину σ_n условием

$$\sigma_n = \frac{1}{n} \cdot (X_1 + \ldots + X_n),$$

$n = 1,\ 2,\ \ldots$ Тогда

$$\mathbb{M}(\sigma_n) = \frac{1}{n} \cdot (\mathbb{M}(X_1) + \ldots + \mathbb{M}(X_n)).$$

Значит

$$\mathbb{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \frac{\mathbb{M}(X_1) + \ldots + \mathbb{M}(X_n)}{n}\right| \geqslant \varepsilon\right) = \mathbb{P}\left(|\sigma_n - \mathbb{M}(\sigma_n)| \geqslant \varepsilon\right)$$

$$\leqslant \frac{D(\sigma_n)}{\varepsilon^2} = \frac{D(X_1) + \ldots + D(X_n)}{n^2\varepsilon^2} \leqslant \frac{nc}{n^2\varepsilon^2} = \frac{1}{n} \cdot \frac{c}{\varepsilon^2} \to 0 \text{ при } n \to \infty,$$

что и требовалось доказать. □

Следствие 22.2. *Пусть, вдобавок к условиям предыдущей теоремы,* $\mathbb{M}(X_1) = \mathbb{M}(X_2) = \ldots = a$. *Тогда, для любого* $\varepsilon \geqslant 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - a\right| \geqslant \varepsilon\right) \to 0 \textit{ при } n \to \infty.$$

И в заключение — закон больших чисел в форме Бернулли.

Теорема 22.3 (Закон больших чисел в форме Бернулли)**.** *Пусть вероятность наступления события* A *в последовательности из* n *независимых испытаний постоянна и равна* p. *Тогда, для любого* $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right) \to 1 \textit{ при } n \to \infty.$$

Доказательство. Используя установленные ранее соотношения $\mathbb{M}(\mu_n) = np$ и $D(\mu_n) = npq$, получаем:

$$\mathbb{P}\left(\left|\frac{\mu_n}{n} - p\right| \geqslant \varepsilon\right) = \mathbb{P}\left(|\mu_n - np| \geqslant n\varepsilon\right)$$

$$
\begin{aligned}
&= \mathbb{P}\left(|\mu_n - \mathbb{M}(\mu_n)| \geqslant n\varepsilon\right)\\
&\leqslant \frac{D(\mu_n)}{n^2\varepsilon^2} = \frac{npq}{n^2\varepsilon^2}\\
&= \frac{1}{n}\cdot\frac{pq}{\varepsilon^2} \to 0 \text{ при } n\to\infty.
\end{aligned}
$$

Значит, вероятность события $\left|\frac{\mu_n}{n} - p\right| < \varepsilon$, противоположного событию $\left|\frac{\mu_n}{n} - p\right| \geqslant \varepsilon$, стремится к 1 при $n\to\infty$, — как и утверждалось. □

Этот закон является оправданием статистического подхода к определению вероятности, с которого мы начали наше знакомство с теорией вероятностей. А именно, он говорит, что, как бы мало ни было положительное число ε, относительная частота $\frac{\mu_n}{n}$ наступления события A при достаточно больших n с вероятностью, сколь угодно близкой к единице, будет отличаться от вероятности p события A меньше чем на ε.

Отметим неравенство, полученное в процессе последнего доказательства:

$$\mathbb{P}\left(\left|\frac{\mu_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{1}{n}\cdot\frac{pq}{\varepsilon^2},$$

и его следствие

$$\mathbb{P}\left(\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right) \geqslant 1 - \frac{1}{n}\cdot\frac{pq}{\varepsilon^2}.$$

Эти неравенства часто используются для оценки вероятностей, производимой по результатам многократных испытаний.

Пример 22.1. Изготовлена партия стержней. Средняя длина стержней равна 50 см, среднее квадратическое отклонение равно 0,2 см. Оценить снизу вероятность того, что длина взятого наугад стержня окажется не менее 49,5 см и не более 50,5 см.

Решение. В качестве случайной величины X берём длину стержня. По условию $\mathbb{M}(X) = 50$, $\mathbb{D}(X) = \sigma^2 = (0,2)^2 = 0,04$. Нужно оценить снизу вероятность события $\langle 49,5 \leqslant X \leqslant 50,5 \rangle$. Имеем:

$$49,5 \leqslant X \leqslant 50,5 \iff -0,5 \leqslant X - 50 \leqslant 0,5 \iff |X - 50| \leqslant 0,5.$$

Оценку снизу для этой вероятности получаем с помощью следствия 3° из неравенства Чебышёва. В силу этого следствия (с $\mathbb{D}(X) = 0,04$, $\varepsilon = 0,5$)

$$\mathbb{P}(|X - 50| \geqslant 0,5) \leqslant \frac{0,04}{0,25},$$

значит

$$\mathbb{P}(|X - 50| \leqslant 0,5) \geqslant 1 - \frac{0,04}{0,25} = 0,84.$$

Пример 22.2. Игральный кубик подбрасывается 350 раз. Оценить снизу вероятность того, что среднее арифметическое числа выпавших очков отклонится от математического ожидания не более чем на 0,2.

Решение. Сначала найдём математическое ожидание и дисперсию числа очков, выпадающих при одном подбрасывании (проделайте это!). Они равны, соответственно, 3,5 и $\frac{35}{12}$. Обозначим через X_i ($i = 1, \ldots, 350$) количество очков, выпавших при i-м подбрасывании. Эти случайные величины 1) попарно независимы, 2) имеют ограниченные в совокупности дисперсии и 3) имеют одно и то же математическое ожидание. Все условия следствия из закона больших чисел в форме Чебышёва выполнены. Применяя это следствие с $n = 50$, $\mathbb{D}(X) = \frac{35}{12}$, $a = 3,5$, $\varepsilon = 0,2$, получаем:

$$\mathbb{P}\left(\left|\frac{1}{350}\sum_{i=1}^{350} X_i - 3,5\right| \leqslant 0,2\right) \geqslant 1 - \frac{35}{12 \cdot 350 \cdot 0,04} \approx 0,792.$$

Пример 22.3. Игральную кость подбросили 320 раз. Оценим вероятность того, что относительная частота выпадения пяти очков отклонится от вероятности этого события (в ту или иную сторону) не более чем на 0,03.

Решение. Используя второе из отмеченных выше неравенств с $p = \frac{1}{6}$, $q = \frac{5}{6}$, $\varepsilon = 0,03$, получаем:

$$\mathbb{P}\left(\left|\frac{\mu_n}{n} - \frac{1}{6}\right| < 0,03\right) \geqslant 1 - \frac{1}{320} \cdot \frac{\frac{1}{6} \cdot \frac{5}{6}}{(0,03)^2} \approx 0,518.$$

23. Понятие о центральной предельной теореме

В рассмотренных выше законах больших чисел устанавливалось, что при определённых условиях *значения* случайных величин сходятся по вероятности к некоторым *числам*. В них ничего не утверждается о законах распределения случайных величин. Предельные законы распределения составляют предмет другой группы теорем, объединённых под названием центральной предельной теоремы.

Смысл центральной предельной теоремы состоит в том, что *если случайная величина может быть представлена в виде суммы достаточно большого числа независимых (или слабо зависимых) случайных величин, каждая из которых в отдельности сравнительно мало влияет на сумму, то эта случайная величина распределена приблизительно по нормальному закону*. Так как условия, при которых имеет место центральная предельная теорема, весьма часто выполняются, нормальный закон распределения является наиболее распространённым из законов распределения, наиболее часто встречающимся при изучении случайных явлений природы или общества. Отсюда и его название.

Часть 2. НАЧАЛА МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

24. Предмет и задачи математической статистики

Математическая статистика изучает методы обработки и осмысления результатов многократно повторяемых случайных событий.

Понятие случайного события определяется в теории вероятностей, обработка результатов также производится при помощи теоретически разработанных вероятностных методов.

Задачей математической статистики является а) построение и б) оценка адекватности идеальных вероятностных моделей реальных процессов.

В качестве иллюстрации рассмотрим следующий пример. В теории вероятностей обычно рассматривается модель, в которой при бросании игрального кубика вероятность выпадения любого числа очков от 1 до 6 одинакова и равна $1/6$. Однако для „реального" кубика может случиться так, что при бросании его, например, 1000 раз шестёрка выпадет в 300 случаях. В принципе такое может произойти и в рамках модели с вероятностью $1/6$, однако здравый смысл подсказывает, что скорее всего у кубика смещён центр тяжести. Это означает, что в дальнейшем имеет смысл использовать другую гипотезу относительно вероятности выпадения шестерки при бросании данного кубика. Например, довольно логично предположить, что эта вероятность близка к 0,3.

Для процесса построения и применения моделей характерно следующее обстоятельство: чем больше данных, тем точнее, адекватнее модель. В полной мере это относится к статистическим моделям.

Поскольку мы имеем дело со случайными событиями, то рекомендации, полученные на основе статистических соображений, всегда носят вероятностный характер. Однако это не снижает их ценности. Напротив, вероятностный характер модели является показателем близости к описываемой реальной ситуации, которая зачастую слишком сложна для детерминированного описания.

Известно, какое значение в экономике, сельском хозяйстве, биологии, медицине и т. д. играют статистические методы изучения случайных явлений. Обычно к ним прибегают в тех случаях, когда требуется изучить распределение большой совокупности предметов (явлений, индивидуумов) по некоторому признаку. Например, можно интересоваться распределением множества людей по возрасту, множества животных данного вида по весу, распределением пахотных земель по урожайности, изделий определенного наименования по сортности, распределением больных гриппом по их реакции на данное лекарство.

Так как практически любой признак допускает количественную оценку, то, вместо того чтобы говорить о распределении предметов по признаку, можно говорить о распределении некоторой случайной величины X; опыт, с которым связана величина X, заключается в выборе наугад одного представителя данной совокупности, а значение, принимаемое X, есть значение признака для этого представителя.

Понятно, что исчерпывающее описание такого распределения можно получить, выяснив значения признака для *всех без исключения* представителей данной совокупности. В отдельных ситуациях так и поступают: например, данные о распределении жителей той или иной страны по полу, возрасту и т. д. получают при всеобщих переписях населения, производимых раз в несколько десятилетий. Однако такой способ «поголовного» обследования всей изучаемой совокупности связан с рядом трудностей. Одна из них — это (как правило) *большой объем* совокупности. В некоторых случаях имеется еще и трудность принципиального характера, заключающаяся в том, что рассматриваемая нами совокупность не существует в готовом виде, а является лишь *воображаемой.* Например, если нас интересует распределение ошибки, допускаемой измерительным прибором, то изучаемая совокупность представляет собой совокупность *всех мыслимых измерений*, которые можно произвести с помощью данного прибора. Ясно, что обследовать все элементы такой совокупности невозможно.

Чтобы обойти указанные трудности, обследование всей совокупности заменяют обследованием небольшой, выбранной наугад, её части. Такую часть обычно называют *выборкой*; в противоположность ей вся совокупность называется *генеральной совокупностью.* Разумеется, при этом желательно, чтобы результаты обследования выборки отражали характерные, основные черты изучаемого признака; для этого объем выборки не должен быть чрезмерно мал. Например, о распределении жителей Москвы по размерам носимой ими одежды нельзя судить по результатам обследования одной квартиры; в этом смысле данные, относящиеся к целому дому или группе домов, более показательны. Разработка методов, позволяющих *по результатам обследования выборки делать обоснованные заключения о распределении признака по всей совокупности*, и есть одна из важнейших задач математической статистики.

25. Первичная обработка данных

Обработка данных и получение на ее основе каких-либо рекомендаций относительно принятия того или иного управленческого решения — процесс, вообще говоря, многоэтапный.

Обычно полученные в результате наблюдений данные представляют собой набор чисел. Просматривая этот набор, как правило, трудно выявить какую-либо закономерность. Поэтому данные подвергают некоторой первичной обработке, целью которой является упрощение дальнейшего анализа. Рассмотрим подробно один из возможных способов.

Обычно данные, полученные в результате регистрации значений некоторой случайной величины, представляют собой набор чисел:

$$x_1,\ x_2,\ \dots,\ x_n$$

(некоторые значения могут совпадать). Этот набор чисел называется *выборкой*.

Дальнейшие действия зависят от того, как много в выборке различных чисел. Если мы имеем дело с дискретной случайной величиной, то различных чисел немного; если с непрерывной случайной величиной, то могут и все числа оказаться различными. Поэтому далее рассмотрим два этих случая по отдельности.

25.1. Дискретный случай

Первый этап обработки выборки — это составление вариационного ряда. Его получают так: среди всех чисел x_i отбирают различные и располагают в порядке возрастания —

$$\alpha_1,\ \alpha_2,\ \dots,\ \alpha_m,$$

где $\alpha_1 < \alpha_2 < \ldots < \alpha_m$.

Следующий этап обработки выборки — составление дискретной таблицы частот:

α_1	α_2	$\dots$	α_m
k_1	k_2	$\dots$	k_m
$n_1 = k_1/n$	$n_2 = k_2/n$	$\dots$	$n_m = k_m/n$

Здесь n — число всех измерений, k_i — число измерений, в которых наблюдалось значение α_i. Величины k_i называются *частотами*, а величины $n_i = k_i/n$ — *относительными частотами*. Графической иллюстрацией дискретной таблицы частот является *столбиковая диаграмма* (рис. 3).

Замечание 25.1. Частоты и относительные частоты пропорциональны, поэтому при построении столбиковой диаграммы на вертикальной оси можно указывать значения либо относительных частот, либо частот — визуальное восприятие от этого не зависит.

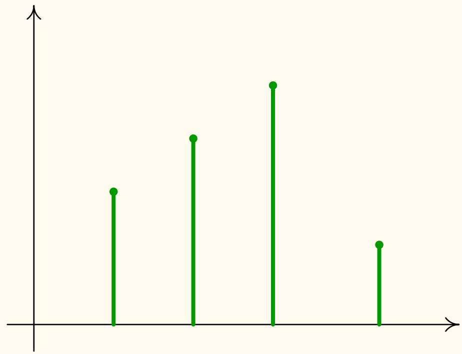


Рис. 3: Примерный вид столбиковой диаграммы

Пример 25.1. Пусть нашей задачей является выявление картины успеваемости студентов, сдавших экзамен по дисциплине «Теория вероятностей». На курсе 56 человек. Полученные студентами отметки представляют собой (в порядке алфавитного списка) следующий набор чисел:

3, 4, 5, 4, 3, 3, 5, 4, 3, 5, 5, 2, 3, 5, 3, 5, 3, 5, 4, 4, 3, 3, 4, 3, 4, 3, 3, 5, 3, 3, 4, 3, 4, 3, 5, 3, 4, 4, 3, 5, 3, 3, 5, 4, 2, 5, 3, 4, 2, 3, 5, 4, 3, 5, 3, 5.

Это и есть исходные данные — выборка. Здесь случайная величина — отметка на экзамене, а числа, составляющие выборку, представляют собой результаты измерения этой случайной величины.

Составляем вариационный ряд:

$$2,\ 3,\ 4,\ 5.$$

Теперь надо подсчитать, сколько раз встречается каждая из оценок. Это можно сделать непосредственно, однако существует и другой способ. Выписывают значения 2, 3, 4, 5 — по одному на каждой строке. Затем просматривают выборку, одно число за другим, и для каждого значения ставят «палочку» (вертикальный отрезок) в соответствующей строке. (Обычно эту работу проделывают вдвоём: один просматривает выборку и называет числа, другой ставит «палочки».) После этого подсчитывают число отрезков в каждой строке.

В данном случае имеем:

2: ||| — 3 значения,
3: |||||||||||||||||||||||| — 24 значения,
4: |||||||||||||| — 14 значений,
5: ||||||||||||||| — 15 значений.

Составляем таблицу частот и относительных частот:

2	3	4	5
3	24	14	15
$3/56$	$24/56$	$14/56$	$15/56$

(В последней строке представлены *относительные* частоты, получаемые делением частот на число измерений $n = 56$.)

И, наконец, строим столбиковую диаграмму (рис. 4).

25.2. Непрерывный случай

Если число различных значений в выборке велико, вычислять частоту каждого из них не имеет большого смысла. Например, если все значения в выборке различны, то дискретная таблица частот имеет вид

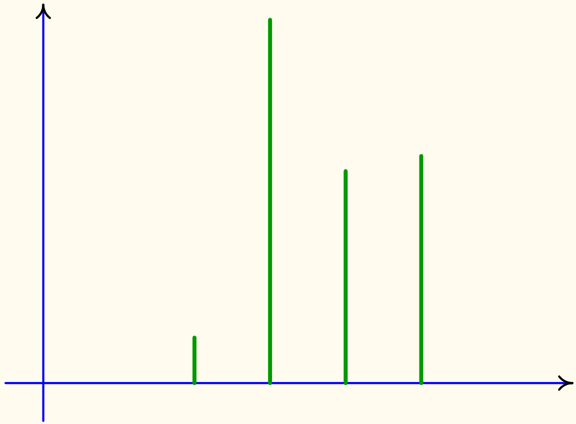


Рис. 4: Столбиковая диаграмма, иллюстрирующая полученную таблицу

α_1	α_2	$\dots$	α_m
1	1	$\dots$	1
$1/n$	$1/n$	$\dots$	$1/n$

Понятно, что такая таблица даёт весьма малую информацию о распределении случайной величины. Поэтому поступают иначе.

Весь промежуток изменения значений выборки, от минимального до максимального, разбивают на интервалы. После этого для каждого интервала подсчитывают интервальную частоту — число значений из выборки, попавших в данный интервал, затем — относительную интервальную частоту. В результате получают интервальную таблицу частот:

$[\mu_1,\ \mu_2]$	$(\mu_2,\ \mu_3]$	$\dots$	$(\mu_m,\ \mu_{m+1}]$
k_1	k_2	$\dots$	k_m
$n_1 = k_1/n$	$n_2 = k_2/n$	$\dots$	$n_m = k_m/n$

Здесь

- n — число всех измерений,
- m — число интервалов,
- k_i — количество чисел, приходящихся на i-й интервал (i-я частота),
- $n_i = k_i/n$ — относительная частота попадания в i-й интервал (i-я относительная частота).

Интервалы обычно берут одинаковой длины, хотя это и не обязательно.

Графической иллюстрацией интервальной таблицы частот является *гистограмма* — ступенчатая фигура, у которой, для любого i, основанием i-й ступеньки является интервал $(\mu_i;\ \mu_{i+1}]$, а площадь этой ступеньки равна n_i $(i = 1,\ \dots,\ m)$ — см. рис. 5.

Число интервалов m выбирают из соображений наглядности получающейся гистограммы. Обычно m лежит в пределах от 5 до 15.

Понятно, что если интервалы $(\mu_i;\ \mu_{i+1}]$ выбраны одинаковой длины, то площади ступенек гистограммы пропорциональны их высотам; в этом случае можно отмечать на оси ординат просто частоты k_i.

Пример 25.2. Предположим, что студенты некоторой группы, состоящей из 25 человек, написали контрольную работу. Каждый студент получил определенное количество баллов. Приведем эти баллы (в порядке алфавитного списка группы):

$$75,\ 145,\ 150,\ 180,\ 125,\ 150,\ 150,\ 165,\ 95,\ 135,\ 130,\ 70,\ 130,\ 105,\ 135,\ 135,\ 100,\ 160,\ 60,\ 85,\ 120,\ 60,\ 145,\ 150,\ 135.$$

Требуется построить интервальную таблицу частот и гистограмму.

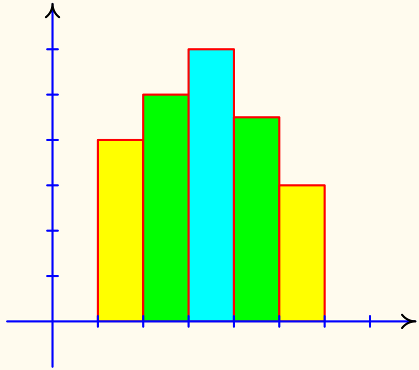


Рис. 5: Примерный вид гистограммы

Найдём среди приведённых чисел минимальное и максимальное — это числа 60 и 180. Таким образом, все значения лежат на отрезке $[60;\ 180]$.

Разобьём этот отрезок, например, на $m = 6$ равных частей.

Подсчитаем число значений, попавших в каждый интервал (воспользуемся процедурой, описанной в примере 1):

$$\begin{array}{ll}
[60;\ 80]: & |||| \text{ — 4 значения,}\\
(80;\ 100]: & ||| \text{ — 3 значения,}\\
(100;\ 120): & || \text{ — 2 значения,}\\
(120;\ 140]: & ||||||| \text{ — 7 значений,}\\
(140;\ 160]: & ||||||| \text{ — 7 значений,}\\
(160;\ 180]: & || \text{ — 2 значения.}
\end{array}$$

Составляем интервальную таблицу частот:

$[60;\ 80]$	$(80;\ 100]$	$(100;\ 120]$	$(120;\ 140]$	$(140;\ 160]$	$(160;\ 180]$
4	3	2	7	7	2
$\frac{4}{25} = 0,16$	$\frac{3}{25} = 0,12$	$\frac{2}{25} = 0,08$	$\frac{7}{25} = 0,28$	$\frac{7}{25} = 0,28$	$\frac{2}{25} = 0,08$

И, наконец, строим гистограмму (рис. 6).

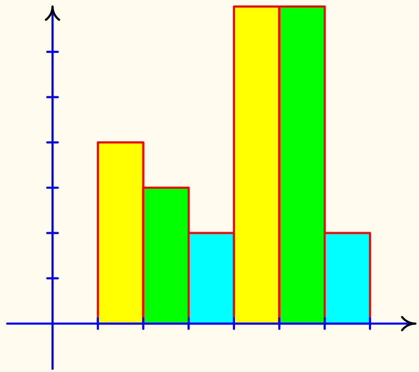


Рис. 6: Гистограмма к примеру 25.2

26. Оценки параметров распределения

По виду таблицы частот или гистограммы можно строить гипотезы об истинном характере распределения случайной величины X. На практике, однако, редко встречается такое положение, когда изучаемый закон распределения неизвестен *полностью*. Чаще всего *вид* закона распределения бывает ясен заранее (из каких-либо теоретических соображений), а требуется найти только некоторые *параметры*, от которых он зависит. Например, если заранее известно, что закон распределения случайной величины нормальный, то задача сводится к нахождению значений двух параметров, a и σ. Впрочем, в некоторых задачах и сам вид закона распределения несуществен, а требуется найти только его числовые характеристики. Во всех подобных случаях можно обойтись сравнительно небольшим числом наблюдений — порядка одного или нескольких десятков.

26.1. Требования, предъявляемые к оценкам параметров

Итак, допустим, что закон распределения случайной величины X содержит некоторый параметр θ. Численное значение этого параметра не указано (хотя оно и является вполне определенным числом). В связи с этим возникает такая задача: исходя из набора значений величины X, полученного в результате n независимых опытов, оценить значение параметра θ.

Любая оценка для θ — обозначим ее $\widetilde{\theta}$ — будет представлять собой, естественно, некоторое выражение, составленное из x_1, x_2,..., x_n:

$$\widetilde{\theta} = \widetilde{\theta}(x_1,\ x_2,\ \ldots,\ x_n).$$

Тем самым $\widetilde{\theta}$ будет случайной величиной (принимающей свои значения в результате n опытов над X). Ее закон распределения будет зависеть от закона распределения случайной величины X (последнему подчинена каждая из величин x_1, x_2, ..., x_n) и от числа опытов n.

Естественно предъявить к оценке $\widetilde{\theta}$ ряд требований.

1. Желательно, чтобы, пользуясь величиной $\widetilde{\theta}$ вместо θ, мы не делали систематических ошибок ни в сторону занижения, ни в сторону завышения, т. е. чтобы было

$$\mathbb{M}(\widetilde{\theta}) = \theta.$$

Оценка, удовлетворяющая такому условию, называется *несмещённой*. Требование несмещённости особенно важно при малом числе опытов.

2. Желательно, чтобы с увеличением числа n опытов значения случайной величины $\widetilde{\theta}$ концентрировались около θ все более тесно, т. е. чтобы было

$$\mathbb{D}(\widetilde{\theta}) \to 0 \text{ при } n \to \infty.$$

Оценку, обладающую таким свойством, условимся называют *состоятельной*.

На практике не всегда удается удовлетворить перечисленным требованиям, так как при прочих условиях желательно, чтобы выражение для функции $\widetilde{\theta}(x_1, x_2, \ldots, x_n)$ было не слишком сложным. В некоторых случаях, например, применяют незначительно смещенные оценки.

26.2. Оценка для математического ожидания

Рассмотрение конкретных оценок мы начнем с наиболее важного случая — оценки для математического ожидания $a = \mathbb{M}(X)$. В качестве такой оценки естественно принять так называемое *эмпирическое среднее*, т. е. среднее арифметическое $\overline{X}$ всех полученных значений величины X:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Чтобы подчеркнуть случайный характер величины $\overline{X}$, перепишем это равенство в виде

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

где через X_i мы обозначаем значение случайной величины X, полученное в i-м опыте (в прежней записи x_i).

Случайные величины X_1, X_2, $\ldots$, X_n имеют один и тот же закон распределения (он совпадает с законом распределения величины X), поэтому

$$\mathbb{M}(\overline{X}) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{M}(X_i) = \frac{1}{n}n\mathbb{M}(X) = a. \tag{1}$$

Таким образом, оценка $\overline{X}$ для математического ожидания является несмещённой. Дисперсия этой оценки

$$\mathbb{D}(\overline{X}) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{D}(X_i) = \frac{1}{n^2}nD = \frac{D}{n}. \tag{2}$$

где D — дисперсия случайной величины X. Отсюда вытекает состоятельность оценки $\overline{X}$.

26.3. Оценки для дисперсии

Так как по самому определению дисперсия D есть математическое ожидание случайной величины $(X-a)^2$, то естественной оценкой для D представляется выражение

$$\overline{D} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

— среднее арифметическое квадратов отклонений от эмпирического среднего $\overline{X}$; эту оценку называют *эмпирической дисперсией*.

Можно показать, что $\overline{D}$ является состоятельной оценкой. Весьма неожиданным, однако, является то, что $\overline{D}$ является *смещённой оценкой для* D. Подсчет, который мы здесь не приводим, показывает, что математическое ожидание величины $\overline{D}$ не совпадает с числом D, а несколько меньше последнего:

$$\mathbb{M}(\overline{D}) = \frac{n-1}{n} D.$$

Отсюда видно, что несмещённой оценкой дисперсии является величина

$$\overline{\overline{D}} = \frac{n}{n-1} D = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2.$$

Действительно,

$$\mathbb{M}(\overline{\overline{D}}) = \frac{n}{n-1} \mathbb{M}(\overline{D}) = D.$$

Оценку $\overline{\overline{D}}$ называют *несмещённой*, или *исправленной эмпирической дисперсией*. Так как $\overline{\overline{D}}$ отличается от $\overline{D}$ множителем $\frac{n}{n-1}$, стремящимся к 1 при $n \to \infty$, то при больших значениях n практически безразлично, какой из двух оценок мы пользуемся.

С чисто технической точки зрения вычисление оценки $\overline{D}$ удобно производить с помощью формулы

$$\overline{D} = \overline{X^2} - (\overline{X})^2 \qquad (3)$$

являющейся аналогом формулы

$$\mathbb{D}(X) = \mathbb{M}(X^2) - M^2(X).$$

(В правой части формулы (3) стоит разность между эмпирическим средним величины X^2 и квадратом эмпирического среднего величины X.)

Пример 26.1. По данным четырёх вступительных экзаменов составлена таблица:

Сумма баллов	12	13	14	15	16	17	18	19	20
Количество абитуриентов	1	3	7	15	21	30	12	8	3

Найти эмпирическое среднее и эмпирические дисперсии (смещенную и несмещенную) для величины X — суммы баллов.

Решение. В данном случае число всех наблюдений

$$n = 1 + 3 + 7 + 15 + 21 + 30 + 12 + 8 + 3 = 100.$$

Отсюда находим:

$$\overline{X} = \frac{1}{100}(12 \cdot 1 + 13 \cdot 3 + 14 \cdot 7 + 15 \cdot 15 + 16 \cdot 21 + 17 \cdot 30 + +18 \cdot 12 + 19 \cdot 8 + 20 \cdot 3) = \frac{1648}{100} =$$

$$\overline{X^2} = \frac{1}{100}(12^2 \cdot 1 + 13^2 \cdot 3 + 14^2 \cdot 7 + 15^2 \cdot 15 + 16^2 \cdot 21 + 17^2 \cdot 30 + +18^2 \cdot 12 + 19^2 \cdot 8 + 20^2 \cdot 3)$$

$$\overline{D} = 272,08 - (16,48)^2 = 0,4896,$$

$$\overline{\overline{D}} = \frac{100}{99} 0,4896 = 0,4945\ldots.$$

Итак, искомыми оценками являются: $\overline{X} = 16,48$, $\overline{D} = 0,4896$, $\overline{\overline{D}} = 0,4945\ldots$.

27. Доверительные оценки

27.1. Доверительные вероятности и доверительные интервалы

До сих пор мы ставили своей задачей оценить неизвестный параметр θ одним числом $\widetilde{\theta}$. Такая оценка называется *точечной*. При большом числе опытов точечная оценка, как правило, близка к неизвестному параметру. Однако если число наблюдений мало, то случайный характер величины $\widetilde{\theta}$ может привести к значительному расхождению между $\widetilde{\theta}$ и θ. Тогда возникает задача о приближении параметра θ не одним числом, а целым *интервалом*

$$(\widetilde{\theta}_1,\ \widetilde{\theta}_2)$$

так, чтобы вероятность поглощения этим интервалом параметра θ, т. е. вероятность двойного неравенства

$$\widetilde{\theta}_1(x_1,\ \ldots,\ x_n) < \theta < \widetilde{\theta}_2(x_1,\ \ldots,\ x_n) \tag{4}$$

была не меньше заданного числа α.

Подчеркнем, что $\widetilde{\theta}_1$, и $\widetilde{\theta}_2$ суть случайные величины, в то время как θ — некоторое вполне определенное (хотя и неизвестное нам) число; поэтому событие (4) является случайным событием, что дает право говорить о вероятности его наступления.

Если число α выбрать достаточно большим, например 0,95 или 0,99, то событие (4) можно считать практически достоверным; следовательно, получив опытные значения $(x_1, \ \ldots, \ x_n)$ случайной величины X и построив по ним интервал $(\widetilde{\theta}_1, \ \widetilde{\theta}_2)$, можно быть практически уверенными в том, что неизвестный параметр θ окажется заключенным в этом интервале.

Вероятность α называется *доверительной вероятностью*, а соответствующий интервал $(\widetilde{\theta}_1, \ \widetilde{\theta}_2)$ — *доверительным интервалом* (отвечающим доверительной вероятности α).

27.2. Построение доверительного интервала для математического ожидания нормального распределения

Перейдем к вопросу о построении доверительного интервала. При этом ограничимся тем случаем, когда величина X имеет нормальное распределение с параметрами a (математическое ожидание) и σ (среднее квадратичное отклонение). Для параметра a требуется на основе опытных данных построить доверительный интервал, отвечающий доверительной вероятности α.

Эта задача имеет большое практическое значение, особенно при обработке результатов измерений. В самом деле, допустим, что производится серия независимых измерений для определения некоторой физической величины a. На результат измерения оказывает влияние огромное количество случайных факторов, таких, как колебание атмосферных условий, сотрясения измерительного прибора, усталость наблюдателя и т. п. Каждый из этих факторов порождает незначительную ошибку в измерении случайной величины. Результирующая ошибка будет суммой большого числа малых случайных величин и, в силу центральной предельной теоремы, окажется распределённой по нормальному закону. Следовательно, и результат измерения

$$X = a + \text{ошибка}$$

имеет нормальное распределение. Если при этом отсутствует систематическая ошибка, то

$$\mathbb{M}(X) = a.$$

Поэтому основная задача обработки результатов измерений — *оценка истинного значения измеряемой величины* — математически формулируется как задача оценки математического ожидания (или, как его ещё называют, центра) нормального распределения.

Частичное решение этой задачи даёт эмпирическое среднее. Однако если число n измерений невелико, то значительно больший интерес представляет доверительная оценка, т. е. такой интервал $(a_1,\ a_2)$, который с заданной доверительной вероятностью (или, как говорят, с заданной *надёжностью*) накрывает число a.

Задачу построения доверительного интервала для a обычно ставят в двух вариантах:

1) когда σ известно;

2) когда σ неизвестно.

Мы рассмотрим только первую.

27.3. Доверительный интервал для a при известном σ

Пусть σ известно. Примем во внимание, что случайная величина $\sum_{i=1}^{n} X_i$, от которой a отличается лишь постоянным множителем $\frac{1}{n}$, подчиняется нормальному закону (этот факт вытекает из так называемой центральной предельной теоремы: сумма независимых случайных величин, распределенных каждая по нормальному закону, сама имеет нормальное распределение). Следовательно, величина $\overline{X}$ тоже распределена нормально с математическим ожиданием a и средним квадратичным отклонением $\frac{\sigma}{\sqrt{n}}$ (см. формулы (1) и (2) предыдущего параграфа). Рассмотрим случайную величину

$$U = \frac{\overline{X} - a}{\frac{\sigma}{\sqrt{n}}}.$$

Её распределение тоже является нормальным, причём с математическим ожиданием 0 и дисперсией 1. Пользуясь этим, можно по данному α найти такое число t_α, чтобы было

$$\mathbb{P}(-t_\alpha < U < t_\alpha) = \alpha. \tag{5}$$

В самом деле, вероятность события $-t_\alpha < U < t_\alpha$ равна $\Phi(t_\alpha) - \Phi(-t_\alpha)$, т. е. $2\Phi(t_\alpha)$ (где $\Phi(x)$ — функция Лапласа). Значит, для нахождения искомого числа t_α достаточно решить уравнение $2\Phi(t_\alpha) = \alpha$, то есть $\Phi(t_\alpha) = \frac{\alpha}{2}$. Это делается с помощью таблицы значений функции Лапласа. Получив t_α, мы можем утверждать, что вероятность события $-t_\alpha < U < t_\alpha$, или, в более подробной записи, события

$$-t_\alpha < \frac{\overline{X} - a}{\frac{\sigma}{\sqrt{n}}} < t_\alpha$$

равна α. Поскольку это эквивалентно

$$\overline{X} - t_\alpha \frac{\sigma}{\sqrt{n}} < m < \overline{X} + t_\alpha \frac{\sigma}{\sqrt{n}},$$

то можно, следовательно, утверждать, что вероятность события $-t_\alpha < U < t_\alpha$ равна α. Это означает, что интервал

$$\left(\overline{X} - t_\alpha \frac{\sigma}{\sqrt{n}},\ \overline{X} + t_\alpha \frac{\sigma}{\sqrt{n}}\right)$$

будет доверительным интервалом для математического ожидания a, отвечающим доверительной вероятности α. Заметим, что длина $2t_\alpha \frac{\sigma}{\sqrt{n}}$ этого интервала оказалась постоянной (не зависящей от опытных данных), хотя, разумеется, она и зависит от взятого α; центр интервала находится в случайной точке $\bar{X}$.

Пример 27.1. Автомат, фасующий чай в пачки, работает со стандартным отклонением $\sigma = 5$ г. Произведена случайная выборка объёмом $n = 30$ пачек. Средний вес пачки чая в выборке оказался равен $\overline{X} = 101$ г. Найти доверительный интервал для среднего веса пая в пачке для генеральной совокупности с доверительной вероятностью $\alpha = 0,95$.

Решение. По условию $\alpha = 0,95$, значит, $\frac{\alpha}{2} = 0,475$. По таблице значений функции Лапласа находим *среди значений* число 0,475 или ближайшее к нему число. В таблице имеется значение 0,475. Соответствующее ему значение аргумента равно 1,96. Таким образом, в нашем случае $t_\alpha = 1,96$. Значит, искомым доверительным интервалом является интервал

$$\left(101 - 1,96 \cdot \frac{5}{\sqrt{30}},\ 101 + 1,96 \cdot \frac{5}{\sqrt{30}}\right),$$

то есть интервал (99,21, 102,79).

Упражнение 27.1. Автомат, фасующий чай в пачки, работает со стандартным отклонением $\sigma = 3$ г. Произведена случайная выборка объёмом 40 пачек. Средний вес чая в пачке из выборки оказался равен $\overline{X} = 49$ г. Найти доверительный интервал для среднего веса чая в пачке в генеральной совокупности с доверительной вероятностью $\alpha = 0,99$.