Федеральное агентство морского и речного транспорта
Федеральное государственное бюджетное образовательное учреждение высшего образования

ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ МОРСКОГО И РЕЧНОГО ФЛОТА имени адмирала С. О. МАКАРОВА

Институт ВОДНОГО ТРАНСПОРТА
Кафедра математики

М. Ю. Ястребов

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Учебное пособие

Рекомендовано к изданию Редакционно-издательским советом ГУМРФ имени адмирала С. О. Макарова в качестве учебного пособия

Санкт-Петербург
Издательство ГУМРФ имени адмирала С. О. Макарова
2018

ББК 22.171
УДК 519.21
Я85

Ястребов, М. Ю. Теория вероятностей: учеб. пособие / М. Ю. Ястребов. — СПб.: Изд-во ГУМРФ им. адм. С. О. Макарова, 2018. — 96 с.

Я85

Предназначено для студентов второго курса технических и экономических направлений. Содержание соответствует рабочим программам дисциплины «Математика».

Рекомендовано к изданию Редакционно-издательским советом ГУМРФ имени адмирала С. О. Макарова в качестве учебного пособия. Протокол № 2 от 12.04.2018.

Рецензент:

Васин А. В., д-р тех. наук, проф. (ФГБОУ ВО «ГУМРФ им. адм. С. О. Макарова»).

Введение

Предметом изучения в теории вероятностей являются количественные закономерности массовых случайных явлений, событий.

Массовость означает практическую или мыслимую возможность многократно производить какое-либо испытание в неизменных условиях.

Случайность означает отсутствие однозначной связи между воспроизведением условий испытания и наступлением или ненаступлением какого-либо заранее оговорённого события. Случайность связана как с неизбежным влиянием неконтролируемых условий, так и с погрешностью обеспечения контролируемых условий.

Вероятность выступает как числовая характеристика возможности появления в результате испытания заранее оговоренного события.

Традиционными сферами приложения теории вероятностей являются страховое дело, военное дело, социометрия, анализ ошибок измерения, производство массовой однородной продукции, системы массового обслуживания (службы скорой помощи, телефонные сети и т. д.).

1. КОМБИНАТОРИКА

1.1. Факториалы

Определение
Числом $n!$ (читается «*n-факториал*») называется произведение всех натуральных чисел от 1 до n:

$$1!=1;\ \ 2!=1\cdot 2=2;\ \text{далее при } n\geq 3\ \ n!=1\cdot 2\cdot ...\cdot n.$$

Дополнительно полагают $0!=1$.

Замечание. Данное определение для нуль-факториала позволяет сохранить единообразие многих формул, в которых участвуют факториалы.

Таким образом, $3!=6;\ \ 4!=24;\ \ 5!=120;\ \ 6!=720;\ \ 7!=5040$.

Факториалы быстро растут; так, например, $10!=3628800$.

Факториалы соседних натуральных чисел связаны рекуррентным соотношением $(n+1)!=n!(n+1)$.

Сокращение факториалов: $\dfrac{7!}{10!}=\dfrac{7!}{7!\cdot 8\cdot 9\cdot 10}=\dfrac{1}{8\cdot 9\cdot 10}=\dfrac{1}{720}$.

1.2. Принцип умножения

Принцип умножения (принцип произведения) задаёт правило для подсчёта количества различных наборов из r элементов в случае, когда последние выбираются, соответственно, по одному из r конечных множеств. Благодаря этому принципу подсчёт количества вариантов во многих случаях приводит к большим числам.

1. Если для пары (a,b) первый член может быть выбран из k элементов $\{a_1,a_2,...,a_k\}$, а второй — из n элементов $\{b_1,b_2,...,b_n\}$, то общее количество таких пар равно произведению kn.

Действительно, каждый фиксированный элемент a_i $(1\leq i\leq k)$ порождает n пар: $(a_i,b_1);\ (a_i,b_2);\ ...;(a_i,b_n)$. Всего, таким образом, получается k раз по n пар.

2. Если для тройки (a,b,c) первый элемент может быть выбран из k элементов $\{a_1,a_2,...,a_k\}$, второй — из n элементов $\{b_1,b_2,...,b_n\}$, а третий — из l элементов $\{c_1,c_2,...,c_l\}$, то общее количество троек равно произведению knl.

Действительно, каждая фиксированная пара (a_i, b_j) порождает l троек: $(a_i, b_j, c_1); (a_i, b_j, c_2);\ ...;\ (a_i, b_j, c_l)$. Всего, таким образом, получается kn раз по l троек (рис. 1).

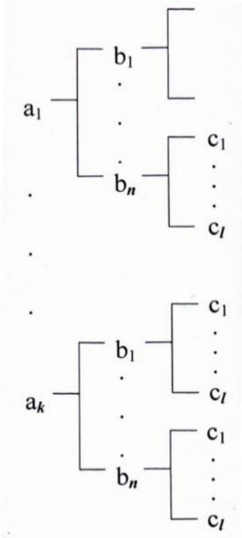

Рис. 1

3. В общем случае (доказывается по индукции), если строится набор из r элементов, причем первый член может быть выбран k_1 способами, второй — k_2 способами и т. д., наконец, последний — k_r способами, то общее количество r-членных наборов равно произведению $k_1 k_2 ... k_r$.

Для случая пар принцип умножения иллюстрируется (по аналогии с матрицами) прямоугольной таблицей, в которой пара (a_i, b_j) стоит на пересечении i-й строки и j-го столбца. Для случая троек (а также наборов из большего числа элементов) принцип умножения иллюстрируется деревом вариантов, которое приведено на рис. 1.

Примеры

1. Количество четырехзначных натуральных чисел, у которых все цифры разные и первая цифра отлична от нуля, по принципу умножения равно $9 \cdot 9 \cdot 8 \cdot 7 = 4536$.

2. Количество пятизначных натуральных чисел, которые могут быть записаны с помощью цифр 2, 4, 6, равно $3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 3^5$.

3. Каждое подмножество множества $\{a_1, a_2, ..., a_k\}$ из k элементов взаимно однозначно характеризуется набором из k нулей и единиц: если элемент a_i входит в подмножество, то на i-м месте набора стоит единица, в противном случае — нуль. Например, пустому множеству соответствует набор из нулей, самому множеству — набор из единиц, одноэлементному подмножеству $\{a_i\}$ — набор, в котором на i-м месте стоит единица, а на остальных местах нули. Количество таких наборов, а значит, и количество всех подмножеств, равно 2^n.

1.3. Перестановки

> **Определение**
> *Перестановкой из n элементов* $a_1, a_2, ..., a_n$ называется их расположение в определённом порядке $(a_{i_1}, a_{i_2}, ..., a_{i_n})$.

Две перестановки считаются различными, если хотя бы один элемент занимает в них разные позиции.

Пример. Все перестановки из трех элементов — цифр 1, 2, 3:

(1,2,3); (1,3,2); (2,1,3); (2,3,1); (3,1,2); (3,2,1).

Число перестановок из n элементов принято обозначать через P_n. Последний пример показывает, что $P_3 = 6$.

Теорема. *Для числа перестановок справедлива формула*

$$P_n = n! \qquad (1)$$

Доказательство. При построении перестановки из n элементов первый член может быть выбран n способами. Для выбора второго члена перестановки остаётся $n-1$ способов, поскольку один элемент уже использован. Для выбора третьего члена остаётся $n-2$ способов, и т. д. Для предпоследнего члена перестановки остаётся два способа, и, наконец, для последнего — один способ, поскольку остался только один неиспользованный элемент.

По принципу умножения общее число различных перестановок равно произведению $P_n = n(n-1)(n-2)...2 \cdot 1 = n!$ ■

Примеры

1. Очередь у причала из четырех судов под грузовую обработку может быть сформирована $4! = 24$ способами.

2. Колода из 36 карт может быть растасована $36!$ способами.

1.4. Размещения

Определение
Размещением из n *различных элементов по* k *элементов* $(0 \le k \le n)$ называется упорядоченный набор каких-либо k из этих элементов.

Два размещения из n по k считаются различными, если они различаются составом и / или порядком следования входящих в них элементов.

Пример. Все размещения из трех элементов — цифр 1, 2, 3 по два:

(1,2); (2,1); (1,3); (3,1); (2,3); (3,2).

Число размещений из n по k принято обозначать через A_n^k. Последний пример показывает, что $A_3^2 = 6$.

Теорема. *Для числа размещений справедлива формула*

$$A_n^k = \frac{n!}{(n-k)!} = n(n-1)(n-2)...(n-k+1) \ . \tag{2}$$

Доказательство. При построении размещения из n по k первый член может быть выбран n способами. Для выбора второго члена размещения остаётся $n-1$ способов, поскольку один элемент уже использован. Для выбора третьего члена остаётся $n-2$ способов, и т. д. Для последнего, k-го члена размещения остаётся $n-k+1$ способов. По принципу умножения общее число различных размещений равно произведению

$$A_n^k = n(n-1)(n-2) \cdot ... \cdot (n-k+1) = \frac{n!}{(n-k)!} \, . \ \blacksquare$$

Пример. $A_5^3 = \frac{5!}{2!} = 5 \cdot 4 \cdot 3 = 60$.

Замечание. Если $k = n$, то размещение из n по n является перестановкой из n элементов. Обе формулы — (1) и (2) — дают в этом случае одинаковый результат $A_n^n = \frac{n!}{0!} = \frac{n!}{1} = P_n$.

Пример. Количество m четырехзначных чисел, все цифры которых различны и не равны нулю, есть количество размещений из девяти элементов — цифр 1, 2, 3, 4, 5, 6, 7, 8, 9 по четыре элемента $m = A_9^4 = 9 \cdot 8 \cdot 7 \cdot 6 = 3024$.

1.5. Сочетания

Определение
Сочетанием из n *различных элементов по* k *элементов* $(0 \le k \le n)$ называется набор каких-либо k из этих элементов без учёта порядка их следования.

Два сочетания из n по k считаются различными, если они различаются составом входящих в них элементов.

Пример. Все сочетания из пяти элементов — цифр 1, 2, 3, 4, 5 по два

(1,2); (1,3); (1,4); (1,5); (2,3); (2,4); (2,5); (3,4); (3,5); (4,5).

Число сочетаний из n по k принято обозначать через C_n^k. Последний пример показывает, что $C_5^2 = 10$.

Теорема. *Для числа сочетаний справедлива формула*

$$C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!} \ . \tag{3}$$

Доказательство. Размещения из n по k с одинаковым составом элементов, но разным порядком их следования, образуют перестановки из k элементов, число которых, согласно формуле (1), равно $k!$

Все эти перестановки приводят к одному сочетанию с тем же составом элементов. Поэтому число сочетаний из n по k в $k!$ раз меньше, чем число размещений из n по k: $C_n^k = \dfrac{A_n^k}{k!}$. Второе равенство в формуле (3) следует теперь из (2). ■

Примеры

1. $C_7^3 = \dfrac{7!}{3!\cdot 4!} = 35$.

2. $C_4^0 = 1;\ \ C_4^1 = 4;\ \ C_4^2 = 6;\ \ C_4^3 = 4;\ \ C_4^4 = 1$.

3. $C_5^0 = 1;\ \ C_5^1 = 5;\ \ C_5^2 = 10;\ \ C_5^3 = 10;\ \ C_5^4 = 5;\ \ C_5^5 = 1$.

4. $C_n^0 = C_n^n = 1;\ \ \ C_n^1 = C_n^{n-1} = n$.

Числа C_n^k называют *биномиальными коэффициентами*, поскольку они участвуют в разложении бинома (двучлена) $(1+x)^n$ по степеням x (бином Ньютона):

$$(x+y)^n = \sum_{i=0}^{n} C_n^i \cdot x^i y^{n-i} = y^n + C_n^1 \cdot xy^{n-1} + C_n^2 \cdot x^2 y^{n-2} + \ldots + C_n^n \cdot x^n \ .$$

Частными случаями разложения бинома являются хорошо известные формулы для квадрата и куба суммы.

Примеры

1. Количество m способов, которыми можно заполнить карточку спортлото «5 из 36», есть число способов, которыми из 36 имеющихся чисел можно выбрать пять для зачеркивания. Поскольку порядок зачеркивания выбранных чисел не играет роли, мы имеем дело с сочетаниями из 36 по пять. Число способов выражается числом сочетаний

$$m = C_{36}^{5} = \frac{36!}{5!\cdot 31!} = \frac{32\cdot 33\cdot 34\cdot 35\cdot 36}{2\cdot 3\cdot 4\cdot 5} = 376992\,.$$

2. Количество l способов, которыми можно заполнить карточку спортлото «5 из 36», угадав ровно три номера, найдем с помощью принципа умножения. Три выигрышных («угаданных») номера должны быть выбраны из шести выигрышных номеров; это можно сделать $C_5^3 = 10$ способами. Два оставшихся номера должны быть выбраны из 31 невыигрышного номера; это можно сделать $C_{31}^2 = 465$ способами. Поэтому общее число способов $l = 10\cdot 465 = 4650$.

3. Количество l способов, которыми в серии испытаний из n бросаний монет может ровно k раз выпасть герб — это число способов, которыми из n номеров испытаний $1, 2, \ldots, n$ можно выбрать k номеров для успешных испытаний. Поскольку порядок, в котором назначаются номера успешных испытаний, не играет роли, мы имеем дело с сочетаниями. Поэтому $l = C_n^k$.

2. СЛУЧАЙНЫЕ СОБЫТИЯ

2.1. Классификация случайных событий

Исходным понятием технико-экономических приложений теории вероятностей является *испытание*, под которым понимают воспроизведение неизменного комплекса условий. В результате испытания может наступить или не наступить заранее оговоренный исход. Как уже отмечалось, отсутствие однозначной связи между проведением испытания и наступлением или ненаступлением исхода (события) обусловлено, во-первых, влиянием неконтролируемых условий, а во-вторых, неизбежной погрешностью, с которой воспроизводятся контролируемые условия.

Предполагается (по крайней мере, теоретически), что испытание можно проводить неограниченное количество раз. Однократное испытание и связанные с ним исходы не являются предметом изучения теории вероятностей.

> **Определение**
> *Случайным событием* (или просто событием) называется заранее оговорённый исход испытания.

Объектом изучения теории вероятностей являются не любые испытания и связанные с ними случайные события, а только такие, для которых при большом числе испытаний проявляются статистические (количественные) закономерности. Крайними полюсами неопределённости исхода являются достоверность и невозможность.

> **Определение**
> Случайное событие называется *достоверным*, если оно заведомо наступает при каждой реализации испытания.

Обозначение достоверного события: Ω .

> **Определение**
> Случайное событие называется *невозможным*, если оно заведомо не может наступить ни при какой реализации испытания.

Обозначение невозможного события: $\varnothing$.

Удобно геометрически иллюстрировать достоверное событие Ω квадратом со стороной (а значит, и площадью) 1; невозможное событие при этом, в соответствии с обозначением, трактуется как пустое множество, а всякое другое случайное событие A , связанное с данным испытанием, изображается фигурой (множеством точек) внутри единичного квадрата (рис. 2).

Определение
Событие A влечёт за собой событие B, если всякий раз при наступлении A наступает и B.

Обозначение: $A \subset B$.

Геометрически это иллюстрируется включением фигуры A в фигуру B (рис. 3).

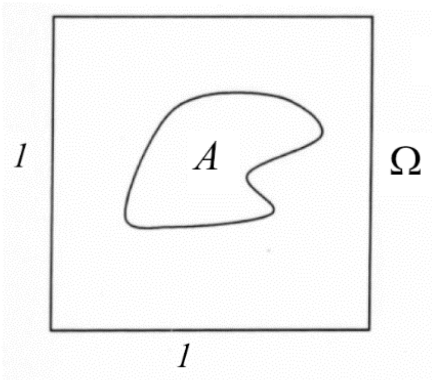

Рис. 2

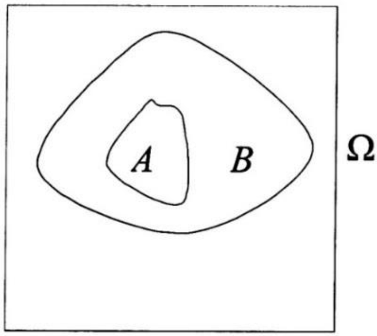

Рис. 3

Иллюстрация, как и обозначение, становятся особенно наглядными, если считать испытанием выстрел по мишени, событиями A и B — попадание в фигуры A и B соответственно.

Любое событие A влечёт за собой достоверное событие $A \subset \Omega$.

Удобно считать, что невозможное событие влечёт за собой любое событие: $\varnothing \subset A$ (аналогично тому, как пустое множество считается содержащемся во всяком множестве).

2.2. Операции над событиями

1. Сумма событий

Определение
Событие C называется суммой событий A и B, если оно заключатся в том, что в результате испытания наступает, по крайней мере, одно из событий A или B.

Обозначение: $C = A + B$.

Геометрически сумма событий иллюстрируется объединением фигур A и B (рис. 4).

Описание суммы событий связано с употреблением союза ИЛИ, причем «нестрогого» или «неальтернативного» ИЛИ, когда допускаются обе возможности.

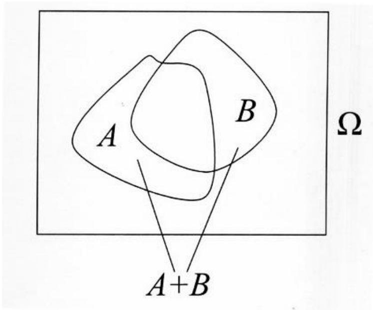

Рис. 4

В примере со стрельбой по мишени наступление суммы событий $A+B$ означает попадание в объединение фигур A и B.

Определение

Событие C называется суммой бесконечного числа событий

$$C = A_1 + A_2 + \ldots + A_n + \ldots = \sum_{i=1}^{\infty} A_i ,$$

если оно заключается в том, что в результате испытания наступает, по крайней мере, одно из событий $A_1, A_2, \ldots, A_n, \ldots$.

Пример. Испытание — стрельба по мишени до первого попадания. Пусть событие A_i означает, что первое попадание происходит при i-м выстреле. Тогда событие

$$C = A_1 + A_2 + \ldots + A_n + \ldots$$

означает, что мишень вообще когда-нибудь (после конечного числа выстрелов) будет поражена.

2. Произведение событий

Определение

Событие C называется произведением событий A и B, если оно заключатся в том, что в результате испытания наступают оба события A и B.

Обозначение: $C = AB$.

Геометрически произведение событий иллюстрируется пересечением фигур A и B (рис. 5).

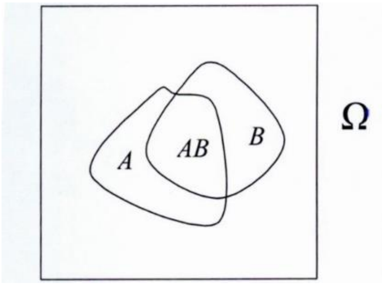

Рис. 5

Описание произведения событий связано с употреблением союза «И».

В примере со стрельбой по мишени наступление произведения событий AB означает попадание в пересечение фигур A и B.

Определения

1. События A и B называются *несовместными*, если невозможно их совместное наступление при одном испытании, т. е. если $AB = \varnothing$.

2. События $A_1, A_2, \ldots, A_n, \ldots$, образующие конечную или бесконечную совокупность, называются *попарно несовместными*, если любые два из них несовместны: $A_i A_j = \varnothing$ при $i \neq j$.

Геометрически несовместность событий A и B иллюстрируется тем, что фигуры A и B не пересекаются (рис. 6).

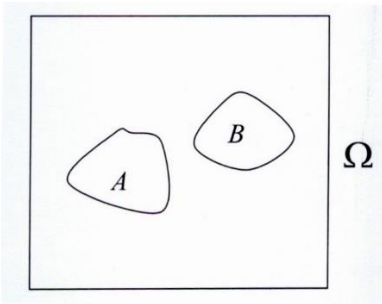

Рис. 6

3. Противоположное событие

Определение

Событие C называется *противоположным* событию A, если оно заключатся в том, что в результате испытания событие A не наступает.

Обозначение: $C = \overline{A}$.

Геометрически противоположное событие $\overline{A}$ иллюстрируется дополнением фигуры A до Ω (рис. 7).

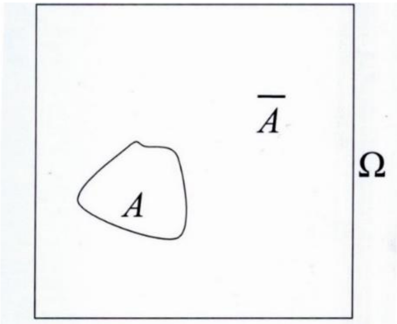

Рис. 7

Описание противоположного события связано с употреблением отрицательной частицы НЕ.

В примере со стрельбой по мишени наступление противоположного события $\overline{A}$ означает непопадание в фигуру A.

Порядок выполнения операций, как и в случае чисел, регулируется скобками. По умолчанию принят следующий порядок действий: сначала выполняются операции перехода к противоположному событию, затем операции умножения, и последними — операции сложения.

Пример. Запись $AB + \overline{C}$ равносильна записи с использованием скобок $(AB) + (\overline{C})$.

2.3. Свойства операций над событиями

Из определения операций над событиями вытекает ряд свойств этих операций, которые удобно разбить на несколько групп.

1. Операции с участием достоверного и невозможного событий:

$$\overline{\Omega} = \varnothing; \quad \overline{\varnothing} = \Omega.$$

Для всякого события A:

$$A + \Omega = \Omega; \quad A + \varnothing = A; \quad A\Omega = A; \quad A\varnothing = \varnothing.$$

2. Операции с одним событием:

$$A + A = A; \quad AA = A; \quad \overline{(\overline{A})} = A; \quad A + \overline{A} = \Omega; \quad A\overline{A} = \varnothing.$$

3. Алгебраические свойства. Для любых двух событий A и B имеют место:

– коммутативность (переместительный закон):

$$A+B=B+A; \quad AB=BA;$$

– ассоциативность (сочетательный закон):

$$(A+B)+C=A+(B+C); \quad (AB)C=A(BC);$$

– дистрибутивность (распределительный закон):

$$A(B+C)=AB+AC;$$

$$A+(BC)=(A+B)(A+C).$$

Замечание. Свойства этой группы, за исключением последнего, имеют привычный вид, поскольку они справедливы для операций сложения и умножения чисел. Последнее свойство имеет аналогию в свойстве операций объединения и пересечения множеств, а также в свойстве логических операций над высказываниями — дизъюнкции (логического сложения, операции ИЛИ) и конъюнкции (логического умножения, операции И) [11].

4. Связь сложения и умножения событий с переходом к противоположному событию выражается *законом двойственности*

$$\overline{A+B}=\overline{A}\cdot\overline{B}; \quad \overline{AB}=\overline{A}+\overline{B}.$$

5. Если событие A влечёт за собой событие B,

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

2.4. Относительная частота события

Определение

Пусть при проведении n испытаний событие A наступило $k(A)$ раз. *Относительной частотой события* A в данной серии испытаний называется отношение

$$w_n(A)=\frac{k(A)}{n}. \tag{4}$$

Пример. Испытание — однократное бросание монеты. Событие A — выпадение герба. Если после семи бросаний оказалось, что герб выпал пять раз, то $n=7$; $k(A)=5$; $w_7(A)=\dfrac{5}{7}$.

Замечания

1. Для вычисления относительной частоты необходимо фактически осуществить серию испытаний.

2. В разных сериях испытаний относительная частота может принимать различные значения.

3. Как показывает практика, для некоторых испытаний и связанных с ними случайных событий A относительная частота проявляет свойство устойчивости: при большом числе испытаний n значения $w(A)$ колеблются в небольшом диапазоне. Именно такие испытания и события рассматриваются в теории вероятностей.

Свойства относительной частоты

1. Относительная частота случайного события неотрицательна

$$w_n(A) \geq 0 .$$

Действительно, в формуле (4) $n > 0;\ \ k(A) \geq 0 \Rightarrow \dfrac{k(A)}{n} \geq 0$.

2. Для достоверного события $w_n(\Omega) = 1$.

Действительно, достоверное событие наступает при каждой реализации испытания, так что $k(A) = n$.

3. Для невозможного события $w_n(\varnothing) = 0$.

Действительно, невозможное событие не наступает ни при одной реализации испытания, так что $k(A) = 0$.

4. Если события A и B несовместны, то

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

Действительно, в случае несовместности A и B событие $A+B$ наступает либо только если наступает A, либо только если наступает B. Поэтому

$$k(A+B) = k(A) + k(B) ,$$

так что

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

3. СВОЙСТВА ВЕРОЯТНОСТИ

3.1. Аксиоматическое введение вероятностей

Исходными объектами при аксиоматическом введении вероятностей являются случайные события, связанные с определённым испытанием. Основанием для введения аксиом в приводимом ниже виде являются соответствующие свойства относительной частоты.

Предполагается, что на совокупности $\{A\}$ случайных событий A задана числовая функция $P(A)$, выражающая вероятность события A и удовлетворяющая следующим *аксиомам*.

Аксиома 1. $P(A) \geq 0$ — вероятность случайного события неотрицательна.

Аксиома 2. $P(\Omega) = 1$ — вероятность достоверного события равна единице.

Аксиома 3. Для конечной или бесконечной совокупности попарно несовместных событий $A_1, A_2, \ldots$ справедливо равенство

$$P(A_1 + A_2 + \ldots) = P(A_1) + P(A_2) + \ldots \text{ —}$$

вероятность суммы попарно несовместных событий равна сумме их вероятностей.

Замечание. Если в аксиоме 3 число событий-слагаемых бесконечно, то сумма в правой части понимается как сумма сходящегося ряда.

При геометрической интерпретации случайных событий (см. п. 2.1) в качестве их вероятностей выступают площади соответствующих фигур. При этом площадь объемлющего квадрата, представляющего достоверное событие, равна единице.

3.2. Непосредственные следствия из аксиом

Теорема 1. *Вероятность невозможного события равна нулю*

$$P(\varnothing) = 0 .$$

Доказательство. По свойству операции сложения событий (п. 2.3) имеем: $\Omega = \Omega + \varnothing$. Поскольку события Ω и $\varnothing$ несовместны, то по аксиоме 3: $P(\Omega) = P(\Omega) + P(\varnothing) \;\Rightarrow\; P(\varnothing) = 0$. ■

Теорема 2. *Если событие* A *влечёт за собой событие* B*, то* $P(A) \leq P(B)$: $A \subset B \;\Rightarrow\; P(A) \leq P(B)$.

Доказательство. По свойствам операций над событиями имеем

$$B = \Omega B = (A + \overline{A})B = AB + \overline{A}B = A + \overline{A}B .$$

Поскольку события A и $\overline{A}B$ несовместны, то $P(B) = P(A) + P(\overline{A}B) \geq P(A)$, так как $P(\overline{A}B) \geq 0$. ■

Теорема 3. *Вероятность случайного события не превосходит единицы*

$$P(A) \leq 1 .$$

Доказательство. Поскольку $A \subset \Omega$, то по предыдущей теореме $P(A) \leq P(\Omega) = 1$. ■

Теорема 4. *Для вероятности противоположного события*

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

Доказательство. По свойствам операций над событиями (п. 2.3):

$$\Omega = A + \overline{A} \ \Rightarrow \ 1 = P(\Omega) = P(A) + P(\overline{A}) \ \Rightarrow P(\overline{A}) = 1 - P(A) . \ ■$$

Теорема 5 (теорема сложения). *Для любых двух событий A и B*

$$P(A + B) = P(A) + P(B) - P(AB) . \qquad (5)$$

Замечание. Геометрически это означает (рис. 8), что для вычисления площади объединения фигур A и B нужно из суммы их площадей вычесть площадь общей части AB, поскольку при сложении она учтена дважды.

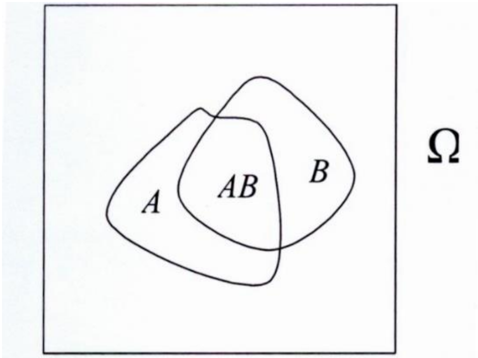

Рис. 8

Доказательство. По свойствам операций над событиями имеем

$$A + B = A\Omega + \Omega B = A(B + \overline{B}) + (A + \overline{A})B = AB + A\overline{B} + AB + \overline{A}B = AB + A\overline{B} + \overline{A}B$$

— сумма попарно несовместных событий (геометрически это означает, что объединение фигур A и B состоит из трех частей: их общей части, той части A, которая не пересекается с B, и той части B, которая не пересекается с A).

Поэтому (аксиома 3)

$$P(A+B)=P(AB)+P(A\overline{B})+P(\overline{A}B). \qquad (6)$$

Далее $A=A\Omega=A(B+\overline{B})=AB+A\overline{B}$ — сумма несовместных событий, так что $P(A)=P(AB)+P(A\overline{B})$,

и
$$P(A\overline{B})=P(A)-P(AB). \qquad (7)$$

Аналогично

$$P(\overline{A}B)=P(B)-P(AB). \qquad (8)$$

Подставляя (7) и (8) в (6), получаем

$$P(A+B)=P(A)+P(B)-P(AB). \blacksquare$$

Замечание. Формула (5) вместе с третьей аксиомой дают две формы теоремы сложения:

– для произвольных событий

$$P(A+B)=P(A)+P(B)-P(AB);$$

– для несовместных событий

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

3.3. Схема равновозможных исходов

Основным содержанием теории вероятностей как науки, ориентированной на приложения, является определение вероятностей одних, более сложных случайных событий по вероятностям других, более простых событий. Поэтому большое практическое значение имеют методы вычисления (или обоснованного назначения) этих исходных вероятностей.

Одним из возможных способов является проведение большого числа n испытаний, после чего вероятность события $P(A)$ полагается равной его относительной частоте $w(A)$ («статистическое определение вероятности»)

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

где k — число тех испытаний, при которых событие A наступило.

Данный способ имеет ряд недостатков:

1) такой подход применим только в случае, когда относительная частота $w(A)$ с ростом n проявляет устойчивость, колеблясь в небольшом диапазоне;

2) заранее неизвестно, будет ли относительная частота с увеличением числа испытаний n проявлять устойчивость. Неизвестно также, какой

должна быть нижняя граница для n, чтобы подобная устойчивость начала проявляться;

3) осуществление отдельного испытания может быть связано с большими материальными затратами и техническими трудностями;

4) значение вероятности $P(A)$ не является однозначно определённым (даже при наличии устойчивости относительной частоты она все же может принимать в разных сериях из n испытаний различные значения).

Альтернативой этому (статистическому) определению вероятности случайного события является её вычисление по «схеме равновозможных исходов». Указанный метод применяется в ситуациях, когда условия испытания обладают по отношению к исходам известной симметрией. Последнее имеет, например, место при контроле качества массовой однородной продукции, проведении лотерей, жеребьевок, а также в других ситуациях «случайного», «справедливого» выбора.

Схема равновозможных исходов (или «классическое определение вероятности») основана на следующем утверждении: если событию A благоприятствуют k из n равновозможных (равновероятных) исходов, попарно несовместных и в совокупности исчерпывающих все возможные исходы испытания, то для вероятности $P(A)$ справедлива формула

$$P(A)=\frac{k}{n}\ . \qquad (9)$$

Таким образом, вероятность случайного события равна отношению числа благоприятных исходов к общему числу исходов. Более точно это формулируется следующим образом.

Теорема. *Пусть события $H_1, H_2, \ldots, H_n$ удовлетворяют трем условиям:*

1) события попарно несовместны: $H_i H_j = \varnothing$ при $i \neq j$;

2) события равновероятны: $P(H_1) = P(H_2) = \ldots = P(H_n)$;

3) в результате испытания одно из событий обязательно наступает

$$H_1 + H_2 + \ldots + H_n = \Omega .$$

Тогда для вероятности события

$$A = H_{i_1} + H_{i_2} + \ldots + H_{i_k}\ (1 \le i_1 < i_2 < \ldots < i_k \le n\ )$$

справедлива формула

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

Доказательство. Убедимся сначала, что для любого $i = 1, 2, \ldots, n$ имеет место равенство: $P(H_i) = \frac{1}{n}$. Действительно, ввиду попарной несовместности событий H_i

$$1 = P(\Omega) = P(H_1 + H_2 + \ldots + H_n) = P(H_1) + P(H_2) + \ldots + P(H_n) .$$

В силу условия 2) в последней сумме все слагаемые одинаковы, так что

$$P(H_1) = P(H_2) = \ldots = P(H_n) = \frac{1}{n} .$$

Поэтому $P(A) = P(H_{i_1}) + P(H_{i_2}) + \ldots + P(H_{i_k}) = \frac{k}{n}$ (сумма k слагаемых, каждое из которых равно $\frac{1}{n}$). ■

3.4. Алгоритм реализации схемы равновозможных исходов

Практическое использование схемы равновозможных исходов опирается на алгоритм, который удобно сформулировать в виде последовательности вопросов. В зависимости от конкретной задачи ответы на некоторые из них могут быть тривиальны и прямо вытекать из условия задачи, из описания испытания. Однако в целом они необходимы для правильного применения последней теоремы.

1. В чем заключается произвольный исход испытания?
2. Являются ли исходы попарно несовместными?
3. Являются ли исходы равновероятными? (Положительный ответ может опираться на симметрию условий испытания по отношению к его исходам).
4. Каково общее число n всех возможных исходов?
5. Чем характеризуется благоприятный исход (при котором наступает событие A)?
6. Каково число k благоприятных исходов?
7. Вычисление вероятности по формуле (9).

Рассмотрим примеры использования алгоритма при решении задач на определение вероятностей случайных событий.

Пример 1. Из урны, содержащей пять белых и семь чёрных (перемешанных) шаров, наугад извлекается один шар. Найти вероятность того, что извлечённый шар окажется белым.

Решение. Событие A — извлечённый шар является белым. Отвечаем последовательно на вопросы алгоритма.

1. Отдельный исход испытания заключается в извлечении одного из имеющихся в урне шаров.

2. Поскольку извлекается ровно один шар, то разные исходы попарно несовместны.

3. Поскольку шары одинаковы по форме, то все исходы равновероятны.

4. Общее число возможных исходов равно количеству шаров в урне $n = 12$.

5. Исход испытания благоприятствует событию A, когда извлечён один из пяти белых шаров.

6. Число благоприятных исходов $k = 5$.

7. $P(A) = \dfrac{5}{12}$.

Пример 2. Два раза бросается монета. Найти вероятность того, что герб выпадет хотя бы один раз.

Решение. Событие A — в паре выпавших сторон хотя бы один раз встречается герб. Отвечаем последовательно на вопросы алгоритма.

1. При однократном бросании монеты могут быть два равновероятных (в силу симметрии монеты) исхода: выпадение герба (Г) и выпадение решетки (Р). Поэтому для испытания, состоящего из двух последовательных бросаний, отдельный исход можно обозначить парой из двух символов, каждый из которых может принимать значения «Г» или «Р»: (Г Г); (Р Р); (Г Р); (Р Г).

2. Разные исходы попарно несовместны, поскольку различаются результатом, по крайней мере, одного бросания.

3. Поскольку монета предполагается симметричной, и результаты отдельных бросаний друг на друга не влияют, то все исходы равновероятны.

4. Общее число возможных исходов равно количеству пар: $n = 4$.

5. Благоприятный исход имеет место, если в паре символов хотя бы один раз встречается «Г»: (Г Г); (Г Р); (Р Г).

6. Число благоприятных исходов (число таких пар) $k = 3$.

7. $P(A) = \dfrac{3}{4}$.

Пример 3. После семи бросаний монеты герб выпал ровно четыре раза. Найти вероятность того, что герб выпал при первых четырех бросаниях.

Решение. Номера успешных испытаний с выпадением герба по условию принимают четыре каких-либо значения из совокупности $\{1, 2, 3, 4, 5, 6, 7\}$. Выбрать четыре номера из семи можно C_7^4 способами, так что общее число исходов $n = C_7^4 = 35$. Лишь один исход является благоприятным, — когда успешны бросания с номерами $1, 2, 3, 4$; таким образом, $k = 1$, и искомая вероятность равна $\frac{1}{35}$.

3.5. Эмпирический закон больших чисел

Если вероятность некоторого случайного события A может быть вычислена по схеме равновозможных исходов или каким-либо другим способом, основанным на аксиомах теории вероятностей, то её называют теоретической вероятностью. Как показывает практика, при большом числе испытаний относительная частота $w_n(A)$ устойчиво колеблется вокруг теоретической вероятности $P(A)$. Это почерпнутое из практики положение носит название *эмпирического закона больших чисел*. Оно является в некотором смысле таким же законом природы, как, например, закон всемирного тяготения, хотя, в отличие от последнего, не выражается в виде формулы, является «качественным».

В частности, события, имеющие малую теоретическую вероятность, и на самом деле происходят редко (каждый, кто пытался выиграть в лотерее автомобиль, может вспомнить, как часто это ему удавалось).

Практическое следствие из задачи, решённой в последнем примере, таково: следует ожидать, что при большом числе испытаний относительная частота окажется близкой к числу $\frac{3}{4}$, т. е. примерно в 75 % случаев в результате двух бросаний монеты хотя бы один раз выпадет герб.

3.6. Условная вероятность

Определение
Пусть при проведении n испытаний число наступлений события AB (т. е. события A вместе с событием B) равно $k(AB)$, причем B наступило $k(B)$ раз. *Условной относительной частотой* события A при условии наступления события B называется отношение $\frac{k(AB)}{k(B)}$.

Обозначение условной относительной частоты: $w_n(A/B)$.

Таким образом, $w_n(A/B) = \frac{k(AB)}{k(B)}$.

Смысл данного определения заключается в следующем. Если учитывать только те испытания, при которых наступало событие B, то общее число испытаний уменьшается с n до $k(B)$, а число успешных испытаний уменьшается по сравнению с $k(A)$ и оказывается равным числу одновременных наступлений A и B.

Пример. Пусть после проведения 10 испытаний событие A наступило пять раз, событие B — 6 раз, а события A и B одновременно наступили четыре раза (рис. 9). Тогда обычная, безусловная относительная частота $w_{10}(A) = \dfrac{5}{10}$, а условная относительная частота $w_{10}(A/B) = \dfrac{4}{6}$.

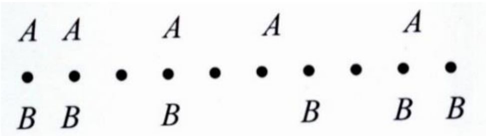

Рис. 9

Теорема. *Для условной относительной частоты справедлива формула*

$$w_n(A/B) = \frac{w_n(AB)}{w_n(B)}. \tag{10}$$

Доказательство. Деля числитель и знаменатель на общее число проведённых испытаний n, получаем

$$w_n(A/B) = \frac{k(AB)}{k(B)} = \frac{\left(\dfrac{k(AB)}{n}\right)}{\left(\dfrac{k(B)}{n}\right)} = \frac{w_n(AB)}{w_n(B)}. \blacksquare$$

Поскольку (в соответствии с эмпирическим законом больших чисел) относительные частоты при большом числе испытаний колеблются вокруг теоретических вероятностей, формула (10) для условной относительной частоты является основанием для следующего определения условной вероятности:

Определение

Пусть $P(B) \neq 0$. *Условной вероятностью* события A при условии наступления события B называется отношение

$$P(A/B) = \frac{P(AB)}{P(B)}. \tag{11}$$

3.7. Теорема умножения

Теорема (умножения). *Если* $P(B) \neq 0$ *(так что существует условная вероятность* $P(A/B)$*), то для вероятности произведения событий справедлива формула*

$$P(AB) = P(B)P(A/B) \ . \qquad (12)$$

Доказательство. Достаточно в формуле (11) обе части равенства умножить на $P(B)$. ■

Замечания

1. Формально, в соответствии с определением (11) условной вероятности, для вычисления $P(A/B)$ следует предварительно и независимо от неё найти $P(AB)$ и $P(B)$. На практике, однако, условную вероятность $P(A/B)$ находят, исходя из анализа испытания, по схеме равновозможных исходов, а формулу (12) используют для вычисления вероятности произведения.

2. Если отличны от нуля обе вероятности $P(A)$ и $P(B)$, то определены обе условные вероятности $P(A/B)$ и $P(B/A)$. Тогда теорема умножения принимает вид

$$P(AB) = P(A)P(B/A) = P(B)P(A/B) \, . \qquad (13)$$

3. Для произведения трех событий теорема умножения даёт

$$P(ABC) = P(A(BC)) = P(A)P(BC/A) \, ,$$

или $\quad P(ABC) = P((AB)C)) = P(AB)P(C/AB) = P(A)P(B/A)P(C/AB)$.

Пример. Испытание: из урны, содержащей 3 белых и 6 чёрных шаров, наугад извлекается шар; после этого извлекается еще один шар. Найти вероятность события C, которое заключается в том, что первый извлечённый шар окажется белым, а второй — чёрным.

Решение. Введём события: A — первый извлечённый шар окажется белым; B — второй извлечённый шар окажется чёрным. Тогда $C = AB$. По схеме равновозможных исходов $P(A) = \frac{3}{9}$. Если событие A уже наступило, то в урне 2 белых и 6 чёрных шаров, так что $P_A(B) = \frac{6}{8}$. По теореме умножения

$$P(C) = P(AB) = P(A)P(B/A) = \frac{3}{9} \cdot \frac{6}{8} = \frac{1}{4} \, .$$

3.8. Независимость событий

I. Независимость двух событий

Опираясь на статистическое определение вероятности (п. 3.3), события A и B следует считать независимыми, если при большом числе испытаний наступления события B не влияют на частоту наступления события A

$$w_n(A/B) = w_n(A),$$

или, в соответствии с определением условной относительной частоты (п. 3.6)

$$\frac{w_n(AB)}{w_n(B)} = w_n(A) \;\Rightarrow\; w_n(AB) = w_n(A)w_n(B).$$

Поскольку, согласно эмпирическому закону больших чисел, относительные частоты при большом числе испытаний колеблются вокруг теоретических вероятностей, последнее равенство является основанием для следующего определения независимости событий:

> **Определение**
> События A и B называются *независимыми*, если вероятность произведения этих событий равна произведению их вероятностей
> $$P(AB) = P(A)P(B). \qquad (14)$$

Формально, в соответствии с данным определением, для решения вопроса о независимости событий необходимо предварительно вычислить все три вероятности $P(AB),\ P(A),\ P(B)$, после чего проверить, выполняется ли равенство (13). На практике, однако, независимость событий A и B устанавливают путем содержательного их анализа, а формулу (13) используют для отыскания вероятности произведения событий.

Таким образом, имеются *две формы теоремы умножения:*

1) для произвольных событий

$$P(AB) = P(A)P(B/A) = P(B)P(A/B);$$

2) для независимых событий

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

Пример. Испытание: одновременно бросаются две монеты. Найти вероятность события C, состоящего в том, что на обеих монетах выпал герб.

Решение. Введём события: A — на первой монете выпал герб, B — на второй монете выпал герб. Тогда $C = AB$. События A и B явно не влияют друг на друга, их следует считать независимыми.

Тогда $$P(C) = P(AB) = P(A)P(B) = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}.$$

Теорема (независимость для противоположных событий). *Если события A и B независимы, то независимы также пары событий*

$$A \text{ и } \overline{B}, \ \overline{A} \text{ и } B, \ \overline{A} \text{ и } \overline{B}.$$

Доказательство. Докажем, например независимость событий A и $\overline{B}$. По условию $P(AB) = P(A)P(B)$; кроме того, $P(\overline{B}) = 1 - P(B)$.

По свойствам операций над событиями (п. 2.3) имеем

$A = A\Omega = A(B + \overline{B}) = AB + A\overline{B}$ — сумма несовместных событий; отсюда

$$P(A\overline{B}) = P(A) - P(AB) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\overline{B}). \ \blacksquare$$

Теорема (критерий независимости двух событий). *Пусть $P(B) \neq 0$. Для того, чтобы события A и B были независимы, необходимо и достаточно, чтобы условная вероятность события A совпадала с его безусловной вероятностью $P(A/B) = P(A)$.*

Доказательство

1. *Необходимость*. Если A и B независимы, то выполняется равенство $P(AB) = P(A)P(B)$. С другой стороны, по теореме умножения $P(AB) = P(B)P(A/B)$. Отсюда $P(A)P(B) = P(B)P(A/B)$. Поскольку $P(B) \neq 0$, получаем $P(A/B) = P(A)$.

2. *Достаточность*. Пусть $P(A/B) = P(A)$. Тогда, применяя теорему умножения, получаем $P(AB) = P(B)P(A/B) = P(B)P(A)$, т. е., согласно определению, события A и B независимы. $\blacksquare$

Теорема (о независимости от Ω и $\varnothing$). *Любое событие A не зависит от достоверного события и от невозможного события*.

Доказательство

1. $P(A\Omega) = P(A) = P(A) \cdot 1 = P(A)P(\Omega)$, так что A и Ω независимы.

2. $P(A\varnothing) = P(\varnothing) = 0 = P(A) \cdot 0 = P(A)P(\varnothing)$, так что A и $\varnothing$ независимы.

II. Независимость событий в совокупности

Для трех и более событий $A_1, A_2, A_3, \ldots, A_n$ их взаимная независимость («независимость в совокупности») означает не только то, что любые два из них не влияют друг на друга (попарная независимость)

$$P(A_i A_j) = P(A_i)P(A_j), \ (i \neq j), \tag{15}$$

но и что для любого подмножества из трех, четырех и т. д. событий этой совокупности вероятность произведения событий равна произведению их вероятностей:

$$P(A_i A_j A_k) = P(A_i)P(A_j)P(A_k), \ (i \neq j \neq k); \tag{16}$$

$$P(A_iA_jA_kA_l) = P(A_i)P(A_j)P(A_k)P(A_l)\text{, } (\, i \neq j \neq k \neq l \,) \tag{17}$$

и т. д. вплоть до условия

$$P(A_1A_2...A_n) = P(A_1)P(A_2)...P(A_n)\ . \tag{18}$$

Недостаточность попарных соотношений (15) для справедливости совокупности равенств (16) – (18) показывает пример С. Н. Бернштейна.

Пример С. Н. Бернштейна. Наугад бросается игральная кость, имеющая форму правильного тетраэдра, четыре грани которого имеют, соответственно, белую, синюю, красную и тройную бело-сине-красную (полосатую) окраску.

Решение. B — на выпавшей грани присутствует белый цвет, S — на выпавшей грани присутствует синий цвет, K — на выпавшей грани присутствует красный цвет. По схеме равновозможных исходов легко убедиться, что $P(B) = P(S) = P(K) = \frac{2}{4} = \frac{1}{2}$. Далее, произведение любых двух из них означает выпадение полосатой грани, так что $P(BS) = P(SK) =$ $= P(BK) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$. Значит, условие (15) выполняется. В то же время, $P(BSK) \ = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$, и условие (16) не выполняется.

Пример. Три игрока поочередно бросают шестигранную игральную кость. Найти вероятность события D заключающегося в том, что все три раза выпадет шестерка.

Решение. Введём события A, B, C — выпадение шестерки, соответственно, у первого, второго, и третьего игрока. Тогда $D = ABC$ — произведение независимых событий. Поэтому

$$P(D) = P(A)P(B)P(C) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216} \approx 0.005\ .$$

3.9. Формула полной вероятности

Определение

События $H_1, ..., H_n$ образуют *полную группу*, если выполняются два условия:

1) в результате испытания одно из них обязательно наступает, т. е. их сумма есть достоверное событие: $H_1 + ... + H_n = \Omega$;

2) события попарно несовместны, т. е. $H_iH_j = \varnothing$ при $i \neq j$.

Теорема. *Пусть выполняются два условия:*

1) *события* $H_1,\dots,H_n$ *(«гипотезы») образуют полную группу;*

2) *события* $H_1,\dots,H_n$ *имеют ненулевые вероятности:* $P(H_i)\neq 0$ $(i=1,\dots,n)$.

Тогда для всякого события A *справедлива формула*

$$P(A)=P(H_1)P(A/H_1)+\dots+P(H_n)P(A/H_n),$$

или в краткой записи

$$P(A)=\sum_{i=1}^{n}P(H_i)P(A/H_i)\ . \qquad (19)$$

Замечания

1. Формула (19) носит название *формулы полной вероятности.*

2. Третье условие теоремы обеспечивает существование условных вероятностей $P(A/H_i)$ (см. п. 3.6).

Доказательство. По свойствам операций над событиями имеем $A=A\Omega=A(H_1+\dots+H_n)=AH_1+\dots+AH_n$ — сумма попарно несовместных событий. Поэтому, применяя последовательно аксиому сложения и теорему умножения, получаем

$$P(A)=P(AH_1)+\dots+P(AH_n)=P(H_1)P(A/H_1)+\dots+P(H_n)P(A/H_n)\ . \blacksquare$$

Решение задач с применением формулы полной вероятности рекомендуется проводить по следующему плану.

1. Введение гипотез H_i.

2. Проверка полноты и попарной несовместности гипотез.

3. Вычисление вероятностей гипотез $P(H_i)$ (например, по схеме равновозможных исходов).

4. Вычисление условных вероятностей $P(A/H_i)$.

5. Применение формулы (19).

Пример. В первой урне содержится 5 белых и 7 чёрных шаров, во второй — 3 белых и 4 чёрных шара; из первой урны наугад (не глядя) перекладывается один шар во вторую урну, после чего из второй урны наугад извлекается шар. Найти вероятность того, что будет извлечён белый шар.

Решение. Пусть событие A состоит в том, что из второй урны извлечён белый шар. Введём события (гипотезы):

H_1 — из первой урны переложен белый шар;

H_2 — из первой урны переложен чёрный шар.

Эти гипотезы образуют полную группу и несовместны. Далее, по схеме равновозможных исходов: $P(H_1)=\frac{5}{12}$, $P(H_2)=\frac{7}{12}$. Найдем условные вероятности. Если имела место гипотеза H_1, т. е. был переложен белый шар, то во второй урне стало 4 белых и 4 чёрных шара; поэтому $P(A/H_1)=\frac{4}{8}$. Если имела место гипотеза H_2, т. е. был переложен чёрный шар, то во второй урне стало 3 белых и 5 чёрных шаров; поэтому $P(A/H_2)=\frac{3}{8}$. По формуле полной вероятности получаем

$$P(A)=\frac{5}{12}\cdot\frac{4}{8}+\frac{7}{12}\cdot\frac{3}{8}=\frac{41}{96}.$$

3.10. Формулы Байеса

Рассмотрим следующую ситуацию. Имеются две внешне неразличимые урны. В урне № 1 находятся 10^6 (миллион) белых шаров и один чёрный шар; в урне № 2, наоборот, находятся один белый шар и 10^6 чёрных. Из схемы равновозможных исходов следует, что вероятность выбрать наугад урну № 1 (гипотеза H_1), как и вероятность выбрать наугад урну № 2 (гипотеза H_2), обе равны $\frac{1}{2}$. Эти вероятности называют *априорными* (латинское «a priori» означает «до опыта»).

Предположим теперь, что из выбранной наугад урны также наугад извлекли один шар, и он оказался белым. Понятно, что, скорее всего, т. е. с вероятностью, близкой к единице, это шар из урны № 1 (это *апостериорная* вероятность гипотезы H_1; латинское «a posteriori» означает «после опыта»). Наоборот, «почти невозможно», т. е. с апостериорной вероятностью, близкой к нулю, это шар из урны № 2.

Таким образом, информация о наступлении некоторого случайного события изменяет наши представления о вероятностях гипотез. Формулы Байеса дают точное количественное выражение для апостериорных вероятностей гипотез.

Теорема. *Пусть для событий $H_1,\ldots,H_n$ («гипотез») и события A выполняются три условия:*

1) *гипотезы $H_1,\ldots,H_n$ образуют полную группу;*

2) *гипотезы* $H_1,\dots,H_n$ *имеют ненулевые вероятности* $P(H_i)\neq 0$ $(i=1,\dots,n)$*;*

3) $P(A)\neq 0$.

Тогда при $j=1,2,\dots,n$ *справедлива формула*

$$P(H_j/A)=\frac{P(H_j)P(A/H_j)}{P(H_1)P(A/H_1)+\dots+P(H_n)P(A/H_n)},$$

или в краткой записи

$$P(H_j/A)=\frac{P(H_j)P(A/H_j)}{\sum\limits_{i=1}^{n}P(H_i)P(A/H_i)}. \qquad (20)$$

Замечания

1. Формула (20) носит название *формулы Байеса.*

2. Второе условие теоремы обеспечивает существование условных вероятностей $P(A/H_i)$ (см. п. 3.6), а третье — существование условных вероятностей $P(H_i/A)$ (см. п. 3.6).

Доказательство. По теореме умножения имеем (см. формулу (13) в п. 3.7)

$$P(H_jA)=P(H_j)P(A/H_j)=P(A)P(H_j/A),$$

откуда $P(H_j/A)=\dfrac{P(H_j)P(A/H_j)}{P(A)}$. Заменяя в знаменателе число $P(A)$ его выражением по формуле полной вероятности (19), приходим к (20). ■

Решение задач с применением формулы Байеса рекомендуется проводить по следующему плану.

1. Введение гипотез H_i.

2. Проверка полноты и попарной несовместности гипотез.

3. Вычисление вероятностей гипотез $P(H_i)$ (например, по схеме равновозможных исходов).

4. Вычисление условных вероятностей $P(A/H_i)$.

5. Применение формулы (20).

Пример. В двух урнах первого типа находится по два белых и три чёрных шара, а в пяти урнах второго типа – соответственно по четыре белых и одному чёрному. Наугад выбирается урна, и из нее извлекается шар. Этот шар оказался белым. Найти вероятность того, что он извлечён из урны первого типа.

Решение. Пусть событие A заключается в том, что извлечённый шар оказался белым. Введём гипотезы:

H_1 – выбрана урна первого типа;

H_2 – выбрана урна второго типа.

По схеме равновозможных исходов $P(H_1)=\frac{2}{7}$, $P(H_2)=\frac{5}{7}$. Далее, $P(A/H_1)=\frac{2}{5}$, $P(A/H_2)=\frac{4}{5}$. По формуле (20) получаем

$$P(H_1/A)=\frac{\frac{2}{7}\cdot\frac{2}{5}}{\frac{2}{7}\cdot\frac{2}{5}+\frac{5}{7}\cdot\frac{4}{5}}=\frac{1}{6}.$$

3.11. Схема независимых испытаний Бернулли

> **Определение**
> Испытания $И_1, И_2, \ldots, И_n$ образуют относительно исхода A последовательность независимых испытаний по схеме Бернулли, если выполняются два условия:
> 1) исходы испытаний независимы в совокупности;
> 2) вероятность исхода A во всех испытаниях одинакова и равна p.

Терминология: A — успех, $P(A)=p$ — вероятность успеха, $\overline{A}$ — неудача, $q=P(\overline{A})=1-p$ — вероятность неудачи.

Пример. Многократные бросания монеты образуют по отношению к исходу A, заключающемуся в выпадении герба, последовательность независимых испытаний по схеме Бернулли с вероятностью успеха $p=\frac{1}{2}$.

Введём обозначение: $P_n(k)$ — вероятность того, что при проведении n испытаний по схеме Бернулли успех будет иметь место ровно k раз.

Теорема (о вероятности числа успехов). *Справедлива формула*

$$P_n(k)=C_n^k p^k q^{n-k} \tag{21}$$

(здесь $C_n^k=\frac{n!}{k!(n-k)!}$ — число сочетаний из n по k; см. п. 1.5).

Доказательство. Пусть D — событие, которое заключается в том, что при проведении n испытаний успех будет иметь место ровно k раз, так что $P_n(k)=P(D)$. Тогда D является суммой попарно несовместных событий

D_j, каждое из которых имеет вид произведения (одновременного наступления) событий

$$D_j = B_1 B_2 ... B_n, \tag{22}$$

причем среди множителей B_i k раз встречается успех A и $n-k$ раз — неудача $\overline{A}$:

$$D = \sum_j D_j$$

и

$$P(D) = \sum_j P(D_j). \tag{23}$$

В серии из n испытаний можно выбрать k номеров успешных испытаний C_n^k различными способами. Поэтому число слагаемых в сумме (23) равно C_n^k (см. последние примеры в пп. 1.5 и 3.4).

Вероятность отдельного слагаемого вычисляется по теореме умножения для независимых событий (исходы разных испытаний, по предположению, не влияют друг на друга)

$$P(D_j) = P(B_1)P(B_2)...P(B_n) = p^k q^{n-k},$$

поскольку в произведении (22) содержится k множителей-успехов, имеющих вероятность p, и $n-k$ множителей-неудач, вероятность которых равна q.

В результате в (23) имеем C_n^k одинаковых слагаемых, равных $p^k q^{n-k}$. Поэтому $P(D) = C_n^k p^k q^{n-k}$. ■

Пример. Из урны, содержащей два белых и три чёрных шара, наугад извлекается один шар (после чего возвращается обратно). Успех A — извлечённый шар оказался белым. Найти вероятность того, что из семи испытаний ровно четыре будут успешными.

Решение. По схеме равновозможных исходов находим вероятность успеха в одном испытании $p = P(A) = \frac{2}{5}$. Далее $q = 1 - p = \frac{3}{5}$, $n = 7$, $k = 4$. Применяя формулу (21), получаем

$$P_7(4) = C_7^4 \cdot \left(\frac{2}{5}\right)^4 \cdot \left(\frac{3}{5}\right)^3 = \frac{7! \cdot 2^4 \cdot 3^3}{4! \cdot 3! \cdot 5^7}.$$

3.12. Локальная теорема Лапласа

Последний пример показывает, что даже при относительно небольших значениях n вычисление вероятностей $P_n(k)$ по точной формуле (19) с использованием факториалов и степеней затруднительно. Поэтому имеется потребность в такой приближённой формуле для вероятности $P_n(k)$, которая не приводила бы к слишком громоздким вычислениям с непомерно большими по записи промежуточными результатами. Такую формулу доставляет локальная теорема Лапласа.

I. Дифференциальная функция Лапласа

> **Определение**
> Дифференциальной функцией Лапласа называется функция $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.

График дифференциальной функции Лапласа («колокол») приведен на рис. 10.

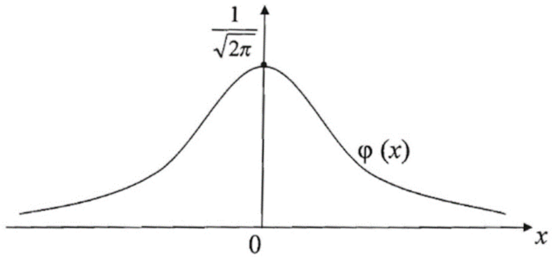

Рис. 10

Свойства функции $\varphi(x)$:

1) $\varphi(x)>0$ при всех x;

2) $\varphi(x)$ — чётная функция, т. е. $\varphi(-x)=\varphi(x)$. График функции симметричен относительно оси ординат;

3) $\lim\limits_{x\to\pm\infty}\varphi(x)=0$.

Стремление к нулю в последнем пределе достаточно быстрое. Так, с точностью до четырех знаков после запятой $\varphi(3{,}9)=0{,}0001$.

Для отыскания значений функции $\varphi(x)$ имеются таблицы и стандартные компьютерные программы.

II. Предельное равенство

Введём для n испытаний по схеме Бернулли с вероятностью успеха p обозначения: $x=\dfrac{k-np}{\sqrt{npq}}$, $\alpha(k,n)=\dfrac{1}{\sqrt{npq}}\phi(x)$, где k — количество успехов ($0\le k\le n$), $q=1-p$.

Теорема. *Пусть вероятность успеха* p *в серии независимых испытаний по схеме Бернулли удовлетворяет условию* $0<p<1$.

Тогда

$$\lim_{n\to\infty}\frac{P_n(k)}{\alpha(k,n)}=1. \qquad (24)$$

Замечания

1. Предельное равенство (24) означает, что при больших n имеет место приближённое равенство

$$P_n(k)\approx\alpha(k,n) .$$

2. Точность приближённого равенства повышается с ростом n.

Пример. Вероятность успеха в отдельном испытании $p=0{,}2$. Найти приближённое значение вероятности 80 успехов в 400 испытаниях.

Решение. Имеем $n=400,\ \ k=80,\ \ p=0{,}2$. Далее вычисляем:

$$q=0{,}8;\ \sqrt{npq}=8;\ \ \frac{1}{\sqrt{npq}}=\frac{1}{8}=0{,}125;\ x=0 .$$

По таблице для дифференциальной функции Лапласа находим $\varphi(0)\approx0{,}3989$. Окончательно $P_{400}(80)\approx0{,}125\cdot0{,}3989\approx0{,}0500$. Таким образом, следует ожидать, что при большом числе серий, по 400 испытаний в каждой, относительная частота тех серий, при которых успех имел место ровно 80 раз, составит около 5 %.

3.13. Интегральная теорема Лапласа

I. Интегральная функция Лапласа

Определение

Интегральной функцией Лапласа называется интеграл с переменным верхним пределом

$$\Phi(x)=\int_0^x\varphi(t)\,dt=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}\,dt.$$

Для отыскания значений функции $\Phi(x)$ имеются таблицы и стандартные компьютерные программы. График интегральной функции Лапласа приведен на рис. 11. Геометрически $\Phi(x)$ выражает площадь заштрихованной части криволинейной трапеции на рис. 12.

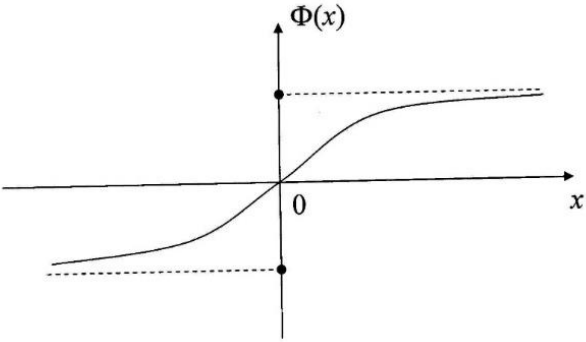

Рис. 11

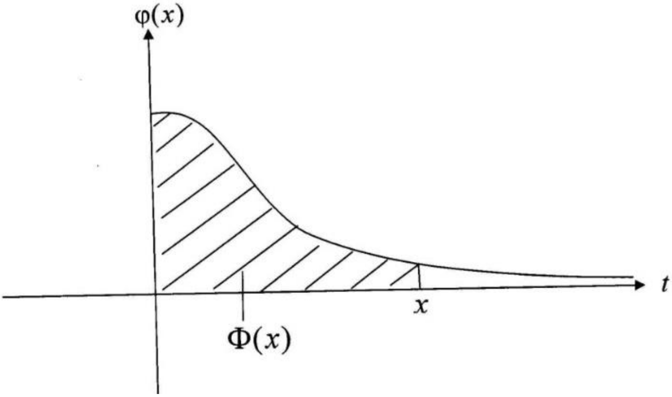

Рис. 12

Свойства функции $\Phi(x)$:

1) $-\dfrac{1}{2} < \Phi(x) < \dfrac{1}{2}$ при всех x;

2) $\Phi(x)$ — нечётная функция, т. е. $\Phi(-x) = -\Phi(x)$. График функции симметричен относительно начала координат;

3) $\lim\limits_{x\to-\infty}\Phi(x)=-\dfrac{1}{2}$; $\lim\limits_{x\to+\infty}\Phi(x)=\dfrac{1}{2}$. Значения $\Phi(x)$ достаточно быстро приближаются к своим предельным значениям. Так, с точностью до шести знаков после запятой $\Phi(4{,}5)=0{,}499997$;

4) $\varphi(x)$ является производной для $\Phi(x)$: $\Phi'(x)=\varphi(x)$. Это следует из теоремы о производной интеграла с переменным верхним пределом. Следовательно, $\Phi(x)$ — одна из первообразных функции $\varphi(x)$;

5) функция $\Phi(x)$ строго возрастает.

Действительно, её производная $\varphi(x)$ положительна при всех x.

II. Предельное равенство

Введём обозначение: $P_n(k_1,k_2)$ — вероятность того, что в серии из n испытаний по схеме Бернулли число успехов k лежит в пределах: $k_1\le k\le k_2$. Пусть, как и в п. 3.12, n — количество испытаний по схеме Бернулли, p — вероятность успеха, $q=1-p$, $x_i=\dfrac{k_i-np}{\sqrt{npq}}$ $(i=1,2)$.

Теорема. *Пусть для вероятности успеха p в серии независимых испытаний по схеме Бернулли выполняется условие $0<p<1$. Тогда для вероятности $P_n(k_1,k_2)$*

$$\lim_{n\to\infty}\left(P_n(k_1,k_2)-\int_{x_1}^{x_2}\varphi(t)\,dt\right)=0\ , \tag{25}$$

или, учитывая определение функции φ,

$$\lim_{n\to\infty}\left(P_n(k_1,k_2)-\frac{1}{\sqrt{2\pi}}\int_{x_1}^{x_2}e^{-\frac{t^2}{2}}\,dt\right)=0\ . \tag{26}$$

Замечания

1. Поскольку $\Phi(x)$ является первообразной для $\varphi(x)$, то интеграл в формулах (25) и (26), согласно формуле Ньютона–Лейбница, равен разности значений первообразной

$$\int_{x_1}^{x_2}\varphi(t)\,dt=\Phi(x_2)-\Phi(x_1)\,.$$

2. Предельное равенство (25) означает, что при больших n имеет место приближённое равенство

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

3. Точность приближённого равенства повышается с ростом n.

Пример. Пусть вероятность успеха $p=0{,}8$, число испытаний $n=100$, границы для числа успехов $k_1=75$, $k_2=90$. Найти $P_{100}(75{,}90)$.

Решение. Имеем: $q=0{,}2$; $\sqrt{npq}=4$; $x_1=-1{,}25$; $x_2=2{,}5$.
Поэтому

$$P_{100}(75{,}90)\approx\Phi(2{,}5)-\Phi(-1{,}25)=\Phi(2{,}5)+\Phi(1{,}25)\approx$$
$$\approx 0{,}4938+0{,}3944=0{,}8882.$$

Таким образом, следует ожидать, что при большом числе серий, из 100 испытаний каждая относительная частота тех серий, в которых число успехов k окажется в пределах $75\le k\le 90$, составит около 89 %.

4. ДИСКРЕТНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

4.1. Понятие случайной величины

Определение
Случайной величиной называется числовая величина, удовлетворяющая двум условиям:

1) в результате испытания она принимает неизвестные наперёд значения;

2) принятие ею значения, лежащего в каком-либо наперёд заданном промежутке, является случайным событием, имеющим определённую вероятность.

Замечание. Второе условие означает, что в теории вероятностей рассматриваются в качестве специальных объектов (случайных величин) не любые величины с неизвестными наперед значениями, а только такие, которые:

– во-первых, связаны с многократно воспроизводимыми (в неизменных контролируемых условиях) испытаниями;

– во-вторых, принимают значения из любого наперед заданного диапазона с относительной частотой, которая проявляет при большом числе испытаний свойство устойчивости (п. 2.4).

Будем обозначать случайные величины прописными латинскими буквами: X, Y, Z и т. п.

4.2. Функция распределения случайной величины

Определение
Функцией распределения случайной величины X называется функция $F(x)$, заданная при каждом $x \in (-\infty, +\infty)$ формулой

$$F(x) = P(X < x) \ .$$

Таким образом, $F(x)$ выражает вероятность случайного события, которое заключается в том, что в результате испытания случайная величина X примет значение, меньшее аргумента x.

Свойства функции распределения:

1) значения функции распределения принадлежат отрезку $[0,1]$: $0 \le F(x) \le 1$. Действительно, $F(x)$ есть вероятность некоторого случайного события, а вероятность всегда лежит в указанных пределах (п. 3.2); ■

2) $P(a \le X < b) = F(b) - F(a)$. *Доказательство*. Введём случайные события: $A = (X < a)$; $B = (X < b)$; $C = (a \le X < b)$, так что $F(a) = P(A)$, $F(b) = P(B)$. Тогда $B = A + C$, и правая часть есть сумма несовместных событий (рис. 13).

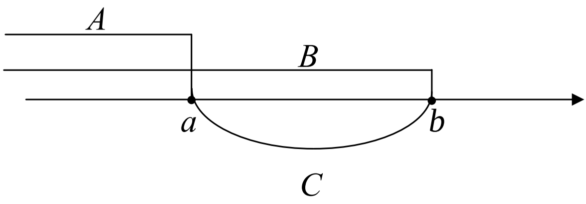

Рис. 13

Отсюда $P(B) = P(A) + P(C)$, и $P(C) = P(B) - P(A)$; ■

3) функция распределения $F(x)$ является неубывающей $F(a) \le F(b)$ при $a < b$. *Доказательство.* В предыдущих обозначениях событие A влечёт за собой событие B: $A \subset B$. Следовательно, $P(A) \le P(B)$, т. е. $F(a) \le F(b)$; ■

4) если все возможные значения случайной величины принадлежат отрезку $[a, b]$, то:

$$F(x) = 0 \text{ при } x < a;$$
$$F(x) = 1 \text{ при } x > b.$$

Доказательство. При $x \le a$ событие $(X < x)$ является невозможным, поэтому $F(x) = P(X < x) = P(\varnothing) = 0$. При $x > b$ событие $(X < x)$ является достоверным, поэтому $F(x) = P(X < x) = P(\Omega) = 1$. ■

Укажем без доказательства следующие свойства:

5) поведение на бесконечности:

$$\lim_{x \to -\infty} F(x) = 0; \quad \lim_{x \to +\infty} F(x) = 1$$

или в другой записи $\qquad F(-\infty) = 0; \ F(+\infty) = 1;$

6) функция распределения непрерывна слева, т. е. для каждого x_0 левосторонний предел функции в точке x_0 равен значению функции в этой точке

$$\lim_{x \to x_0 - 0} F(x) = F(x_0)$$

или в другой записи $\qquad F(x_0 - 0) = F(x_0)$.

4.3. Закон распределения дискретной случайной величины

Определение

Случайная величина X называется *дискретной*, если все её возможные значения можно представить в виде конечной или бесконечной последовательности: $x_1, x_2, ..., x_n, ...$.

Определение

Законом распределения дискретной случайной величины X называется таблица (конечная или бесконечная), содержащая все её возможные значения x_i и вероятности принятия этих значений $p_i = P(X = x_i)$.

$X:$	x_1	x_2	$\dots$	x_n	$\dots$
$P:$	p_1	p_2	$\dots$	p_n	$\dots$

Будем обозначать закон распределения также в виде $(x_1, x_2, ..., x_n, ...; p_1, p_2, ..., p_n, ...)$ или $(x_i; p_i)$.

Примеры

1. Бросание игральной кости. Случайная величина X — количество выпавших очков. Возможные значения — числа $1, 2, 3, 4, 5, 6$. Для каждого из этих значений схема равновозможных исходов даёт вероятность 1/6. Закон распределения имеет вид

$X:$	1	2	3	4	5	6
$P:$	1/6	1/6	1/6	1/6	1/6	1/6

2. Три раза бросается монета. Случайная величина X — количество выпадений герба. Возможные значения — числа $0, 1, 2, 3$. Вероятности значений находятся по схеме Бернулли как вероятности числа успехов (п. 3.11), где вероятность успеха в отдельном испытании $p = 1/2$, вероятность неудачи $q = 1 - p = \frac{1}{2}$:

$$P(X = 0) = P_3(0) = C_3^0 \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^3 = \frac{1}{8};$$

$$P(X = 1) = P_3(1) = \frac{3}{8};$$

$$P(X = 2) = P_3(2) = \frac{3}{8};$$

$$P(X = 3) = P_3(3) = \frac{1}{8}.$$

Закон распределения имеет вид

X:	0	1	2	3
P:	1/8	3/8	3/8	1/8

3. Стрельба по мишени до первого попадания. Вероятность попадания при отдельном выстреле равна $0{,}6$. Случайная величина X — количество произведенных выстрелов. Возможные значения: $1, 2, \ldots, n, \ldots$; множество значений бесконечно. Найдем вероятности этих значений, считая результаты отдельных выстрелов независимыми событиями.

Если первое попадание произошло при n-м выстреле, то первые $n-1$ выстрелов были промахами, а последний, n-й, — попаданием. Введём события A_n — попадание при n-м выстреле ($n = 1, 2, \ldots$), так что

$$P(A_n) = 0{,}6; \;\; P(\overline{A_n}) = 1 - 0{,}6 = 0{,}4.$$

Тогда по теореме умножения для независимых событий (п. 3.8)

$$P(X = n) = P(\overline{A}_1 \overline{A}_2 \ldots \overline{A}_{n-1} A_n) = 0{,}4^{n-1} \cdot 0{,}6.$$

Теорема. *Сумма вероятностей закона распределения дискретной случайной величины равна* 1

$$\sum_i P(X = x_i) = 1 \;. \qquad (27)$$

(*если множество возможных значений бесконечно, то сумма в (27) понимается как сумма ряда, т. е. как предел частичных сумм*).

Доказательство. События $(X = x_i)$ при разных i попарно несовместны. Поскольку одно из возможных значений обязательно реализуется в результате испытания, то $\sum_i (X = x_i) = \Omega$.

Тогда

$$\sum_i P(X = x_i) = P\left(\sum_i (X = x_i)\right) = P(\Omega) = 1 \,. \;\blacksquare$$

4.4. Функция распределения дискретной случайной величины

Рассмотрим задачу нахождения функции распределения $F(x)$ дискретной случайной величины X на примере следующего закона распределения

X:	–1	1	2	4
P:	0,2	0,3	0,4	0,1

Четыре возможных значения случайной величины разбивают числовую ось на пять промежутков. Рассмотрим поочередно возможные случаи расположения аргумента x.

1. Если $x \le -1$, то событие $(X < x)$ является невозможным, в том числе и для крайней точки промежутка $x = -1$. Поэтому $F(x) = P(X < x) = P(\varnothing) = 0$.

2. Если $-1 < x \le 1$, то событие $(X < x)$ совпадает с событием $(X = -1)$, поскольку левее x имеется лишь одно возможное значение случайной величины, а именно, -1.
Значит, $F(x) = P(X < x) = P(X = -1) = 0{,}2.$

3. Если $1 < x \le 2$, то событие $(X < x)$ означает, что $(X = -1)$ или $(X = 1)$, поскольку левее x имеются только два возможных значения случайной величины X. Таким образом, имеем для этого случая: $(X < x) = (X = -1) + (X = 1)$ — сумма несовместных событий.
Значит, $F(x) = P(X < x) = P(X = -1) + P(X = 1) = 0{,}2 + 0{,}3 = 0{,}5.$

4. Если $2 < x \le 4$, то событие $(X < x)$ означает, что $(X = -1)$, или $(X = 1)$, или $(X = 2)$, поскольку левее x имеются три указанных значения. Таким образом, в этом случае $(X < x) = (X = -1) + (X = 1) + (X = 2)$ — сумма попарно несовместных событий.
Значит, $F(x) = P(X < x) = P(X = -1) + P(X = 1) + P(X = 2) = 0{,}2 + 0{,}3 + 0{,}4 = 0{,}9.$

5. Наконец, при $x > 4$ событие $(X < x)$ является достоверным, поскольку все возможные значения случайной величины лежат левее x. Следовательно,

$$F(x) = P(X < x) = P(\Omega) = 1.$$

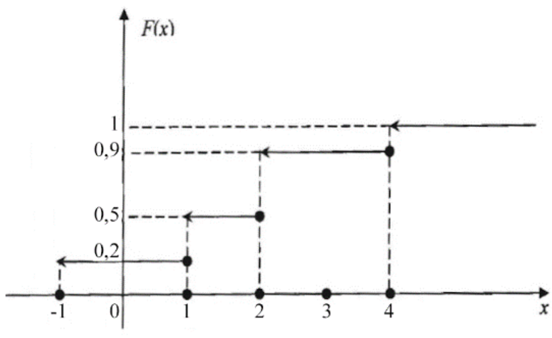

Рис. 14

Итак, график функции распределения имеет ступенчатый характер (рис. 14). Точка на правом конце каждой «ступеньки» означает, что именно на ней находится точка графика с данной абсциссой. Наоборот, стрелка на левом конце указывает, что крайняя левая точка «выколота». Таким образом, значения функции вблизи слева от аргумента x совпадают со значением функции в самой точке x, и, значит, имеют место непрерывность слева и разрыв первого рода (скачок) справа.

График на рис. 14 наглядно иллюстрирует общие свойства функции распределения, описанные в п. 4.2.

4.5. Математическое ожидание дискретной случайной величины

Пусть дискретная случайная величина X принимает значения $x_1, x_2, \ldots, x_n$ с вероятностями, соответственно, $p_1, p_2, \ldots, p_n$. Если при проведении N испытаний значение x_i появилось, k_i раз (так что $k_1+k_2+\ldots+k_n=N$), то среднее арифметическое $\overline{X}$ реализованных значений представляется в виде

$$\overline{X}=\frac{k_1x_1+k_2x_2+\ldots+k_nx_n}{N}=\frac{k_1}{N}x_1+\frac{k_2}{N}x_2+\ldots+\frac{k_n}{N}x_n .$$

Отношение $\frac{k_i}{N}=w_i$ является относительной частотой события $(X=x_i)$. При большом числе испытаний относительные частоты w_i колеблются вокруг соответствующих теоретических вероятностей p_i. Поэтому $\overline{X}$ колеблется вокруг значения $p_1x_1+p_2x_2+\ldots+p_nx_n$.

Это является основанием для следующего определения.

Определения

1. *Математическим ожиданием* дискретной случайной величины X с конечным множеством значений $x_1, x_2, \ldots, x_n$ и с вероятностями этих значений, соответственно, $p_1, p_2, \ldots, p_n$ называется число

$$M(X) = p_1x_1 + p_2x_2 + \ldots + p_nx_n \tag{28}$$

или в краткой записи $M(X)=\sum_{i=1}^{n}p_ix_i$.

2. *Математическим ожиданием* дискретной случайной величины X с бесконечным множеством значений $x_1, x_2, \ldots, x_n, \ldots$ и с вероятностями, $p_1, p_2, \ldots, p_n, \ldots$ называется сумма ряда

$$M(X)=\sum_{i=1}^{\infty}p_ix_i . \tag{29}$$

При этом предполагается, что ряд является *абсолютно сходящимся* (т. е. существует предел последовательности частичных сумм, составленных из модулей его членов). Если абсолютной сходимости нет, то считают, что математическое ожидание не существует.

Статистический смысл математического ожидания: $M(X)$ — *это число, вокруг которого колеблется среднее арифметическое реализованных значений случайной величины при большом числе испытаний*.

Пример. Пусть дискретная случайная величина имеет закон распределения

X:	$X:$	–1	1	2	4
P:	$P:$	0,2	0,3	0,4	0,1

Тогда $M(X)=0,2\cdot(-1)+0,3\cdot 1+0,4\cdot 2+0,1\cdot 4=1,3$.

В соответствии с эмпирическим законом больших чисел, следует ожидать, что при большом числе испытаний среднее арифметическое реализованных значений случайной величины окажется близким к числу 1,3.

4.6. Свойства математического ожидания

1. Если случайная величина X является постоянной («неслучайной») величиной $X=a$, т. е. имеет закон распределения

$$\begin{array}{ll} X: & a \\ P: & 1 \end{array}$$

то $M(X)=a$, или $M(a)=a$.

Доказательство. По определению $M(X)=1\cdot a=a$. ■

2. Постоянный множитель можно выносить за знак математического ожидания: $M(kX)=kM(X)$.

Доказательство. Ограничимся случаем конечного множества значений. Пусть закон распределения имеет вид

$$(x_1, x_2, \ldots, x_n;\ p_1, p_2, \ldots, p_n).$$

Рассмотрим два возможных случая:

1) $k = 0$. Тогда kX имеет закон распределения

$$\begin{array}{ll} kX: & 0 \\ P: & 1, \end{array}$$

так что $M(kX)=M(0)=0$, и также $kM(X)=0\cdot M(X)=0$;

2) $k\neq 0$. Тогда условия $kX=kx_i$ и $X=x_i$ равносильны, и поэтому $(kX=kx_i)=(X=x_i)$; $P(kX=kx_i)=P(X=x_i)=p_i$.

Следовательно, закон распределения случайной величины kX имеет вид

$$(kx_1, kx_2, ..., kx_n;\ p_1, p_2, ..., p_n).$$

Поэтому
$$M(kX) = \sum_{i=1}^{n} p_i k x_i = k \sum_{i=1}^{n} p_i x_i = kM(X). \blacksquare$$

3. **Теорема сложения для математического ожидания**: *пусть случайные величины X и Y имеют математические ожидания $M(X)$ и $M(Y)$. Тогда* $M(X+Y) = M(X) + M(Y)$.

Доказательство. Ограничимся случаем конечного множества значений. Пусть законы распределения для X и Y имеют вид:

$$(x_1, x_2, ..., x_n;\ p_1, p_2, ..., p_n);$$

$$(y_1, y_2, ..., y_r;\ q_1, q_2, ..., q_r).$$

Случайная величина $X+Y$ принимает nr значений вида $x_i + y_j$, где $i = 1, 2, ..., n;\quad j = 1, 2, ..., r$. Ограничимся для простоты случаем, когда все эти значения различны. Пусть $p_{ij} = P(X+Y = x_i + y_j)$ — вероятности указанных значений.

Тогда
$$M(X+Y) = \sum_{i=1}^{n}\sum_{j=1}^{r} p_{ij}(x_i + y_j) = \sum_{i=1}^{n}\sum_{j=1}^{r} p_{ij} x_i + \sum_{i=1}^{n}\sum_{j=1}^{r} p_{ij} y_j.$$

Производя в каждой из двойных сумм группировку слагаемых и вынося за знак внутренней суммы общий множитель (x_i и y_j соответственно), получаем

$$M(X+Y) = \sum_{i=1}^{n}\left(x_i \sum_{j=1}^{r} p_{ij}\right) + \sum_{j=1}^{r}\left(y_j \sum_{i=1}^{n} p_{ij}\right).$$

Убедимся теперь, что $\sum_{j=1}^{r} p_{ij} = p_i$. Для этого введём события:

$$A = (X = x_i);\ H_1 = (Y = y_1);\ H_2 = (Y = y_2);\ \ldots;\ H_r = (Y = y_r).$$

События H_j $(j = 1, ..., r)$, а значит, и AH_j попарно несовместны, поэтому (сумма вероятностей событий равна вероятности их суммы):

$$\sum_{j=1}^{r} p_{ij} = \sum P\big((X = x_i)(Y = y_j)\big) = \sum_{j=1}^{r} P(AH_j) =$$

$$= P(AH_1 + AH_2 + ... + AH_n) = P\big(A(H_1 + H_2 + ... + H_n)\big) =$$

$$= P(A\Omega) = P(A) = p_i.$$

Аналогично $\sum_{i=1}^{n} p_{ij} = q_j$.

Тогда $$M(X+Y) = \sum_{i=1}^{n} x_i p_i + \sum y_j q_j = M(X) + M(Y) \ . \blacksquare$$

4. **Теорема умножения для математического ожидания**

Определения

1) случайные величины X_1, X_2, …, X_n называются *независимыми*, если для любых промежутков $\langle a_1, b_1\rangle$, $\langle a_2, b_2\rangle$, …, $\langle a_n, b_n\rangle$ независимы в совокупности события $(X_1 \in \langle a_1, b_1\rangle)$, $(X_2 \in \langle a_2, b_2\rangle)$, …, $(X_n \in \langle a_n, b_n\rangle)$;

2) случайные величины $X_1, X_2, ..., X_n, ...$, образующие бесконечную последовательность, называются *независимыми*, если для любого натурального n любые n случайных величин этой последовательности независимы.

Замечание. В частности, если X и Y независимы, то для любых чисел a и b

$$P\big((X=a)(Y=b)\big) = P\big((x \in [a,a])(Y \in [b,b])\big) =$$
$$= P(X \in [a,a])P(Y \in [b,b]) = P(X=a)P(Y=b) \ .$$

Теорема. *Пусть случайные величины X и Y независимы и имеют математические ожидания $M(X)$ и $M(Y)$.*

Тогда $$M(XY) = M(X)M(Y) \ .$$

Доказательство. Ограничимся случаем конечного множества значений. Пусть законы распределения для X и Y имеют вид:

$$(x_1, x_2, ..., x_n;\ p_1, p_2, ..., p_n) \ ;$$
$$(y_1, y_2, ..., y_r;\ q_1, q_2, ..., q_r) \ .$$

Случайная величина XY принимает nr значений вида $x_i y_j$, где $i = 1, 2, ..., n; \quad j = 1, 2, ..., r$. Ограничимся для простоты случаем, когда все эти произведения различны. Обозначим через p_{ij} вероятности этих значений. Тогда, поскольку вероятность произведения независимых событий равна произведению их вероятностей,

$$p_{ij} = P\big((X = x_i)(Y = y_j)\big) = P(X = x_i)P(Y = y_j) = p_i q_j \ .$$

Далее $$M(XY) = \sum_{i=1}^{n}\sum_{j=1}^{r} p_{ij} x_i y_j = \sum_{i=1}^{n}\sum_{j=1}^{r} p_i q_j x_i y_j \ .$$

Производя в двойной сумме группировку слагаемых и вынося за знак внутренней суммы общий множитель $p_i x_i$, получаем

$$M(XY)=\sum_{i=1}^{n}\left(p_i x_i \sum_{j=1}^{r} q_j y_j\right)=\sum_{i=1}^{n} p_i x_i M(Y)=M(Y)\sum_{i=1}^{n} p_i x_i=M(Y)M(X). \blacksquare$$

4.7. Дисперсия дискретной случайной величины

Рассмотрим следующий пример. Пусть случайные величины X и Y имеют следующие законы распределения:

X:	–2	–1	0	1	2
P:	0,2	0,2	0,2	0,2	0,2

Y:	–200	–100	0	100	200
P:	0,2	0,2	0,2	0,2	0,2

Нетрудно проверить, что они имеют равные математические ожидания: $M(X)=M(Y)=0$. В то же время рассеивание значений вокруг математического ожидания у Y явно больше, чем у X. Это рассеивание выражается отклонениями реализованных значений от математического ожидания, т. е. разностями $x_i - M(X)$ и $y_i - M(Y)$, соответственно.

Определение
Пусть у случайной величины X существует математическое ожидание $M(X)$. Случайная величина $X - M(X)$ со значениями $x_i - M(X)$ называется *отклонением* (от математического ожидания).

Если при большом числе реализаций случайной величины просуммировать полученные отклонения, то их значения разных знаков в значительной степени погашают друг друга, и такая сумма не может служить мерой рассеивания значений случайной величины вокруг математического ожидания. Для того чтобы избежать подобного взаимного погашения, отклонения перед суммированием возводят в квадрат.

Определение
Пусть у случайной величины X существует математическое ожидание $M(X)$. Её *дисперсией* называется число

$$D(X)=M\left(\left(X-M(X)\right)^2\right), \qquad (30)$$

т. е. математическое ожидание квадрата отклонения.

Статистический смысл дисперсии: дисперсия служит мерой рассеивания значений случайной величины вокруг математического ожидания.

Общее определение дисперсии принимает применительно к дискретной случайной величине следующий вид:

1) если X является дискретной случайной величиной с конечным множеством значений и законом распределения

$$(x_1, x_2, ..., x_n;\ p_1, p_2, ..., p_n),$$

то отклонение $X - M(X)$ имеет закон распределения

$$(x_1 - M(X), x_2 - M(X), ..., x_n - M(X);\ \ p_1, p_2, ..., p_n),$$

и, в соответствии с определением математического ожидания,

$$D(X) = \sum_{i=1}^{n} p_i \left(x_i - M(X)\right)^2 \ ; \qquad (31)$$

2) если X является дискретной случайной величиной с бесконечным множеством значений и законом распределения

$$(x_1, x_2, ..., x_n, ...;\ \ p_1, p_2, ..., p_n, ...),$$

то

$$D(X) = \sum_{i=1}^{\infty} p_i \left(x_i - M(X)\right)^2 . \qquad (32)$$

Если ряд в правой части (32) расходится, то считают, что дисперсия не существует.

Замечание. Если у случайной величины не существует математического ожидания, то понятие дисперсии для нее не вводится.

Примеры

1. Пусть закон распределения имеет вид

X:	–2	0	1
P:	0,1	0,5	0,4

Вычисление дисперсии предполагает предварительное вычисление математического ожидания $M(X) = 0,1 \cdot (-2) + 0,5 \cdot 0 + 0,4 \cdot 1 = 0,2$.

Далее, $D(X) = 0,1 \cdot \left(-2 - 0,2\right)^2 + 0,5 \cdot \left(0 - 0,2\right)^2 + 0,4 \cdot \left(1 - 0,2\right)^2 = 1,08$.

2. Для приведенных в начале этого параграфа законов распределения получаем:

$$M(X) = M(Y) = 0;$$

$$D(X) = 0,2 \cdot \left((-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2\right) = 2;$$

$$D(Y) = 0,2 \cdot \left((-200)^2 + (-100)^2 + 0^2 + 100^2 + 200^2\right) = 20000.$$

Среднее квадратическое отклонение. В случае, когда случайная величина X имеет размерность (метры, килограммы и т. п.), размерность дисперсии $D(X)$ равна квадрату размерности X. Поэтому наряду с дисперсией в качестве меры рассеивания значений случайной величины вокруг математического ожидания применяют также арифметический квадратный корень из дисперсии. Последний уже имеет размерность, совпадающую с размерностью X.

Определение
Средним квадратическим отклонением случайной величины X называется число

$$\sigma(X) = \sqrt{D(X)}.$$

4.8. Свойства дисперсии

1. Для дисперсии справедлива формула

$$D(X) = M(X^2) - (M(X))^2 . \qquad (33)$$

Доказательство. По свойствам математического ожидания

$$D(X) = M\left((X - M(X))^2\right) = M\left(X^2 - 2X \cdot M(X) + (M(X))^2\right) =$$

$$= M\left(X^2\right) + M(-2X \cdot M(X)) + M\left((M(X))^2\right).$$

Во втором слагаемом постоянный множитель $-2M(X)$ вынесем за знак математического ожидания; в третьем слагаемом математическое ожидание константы $(M(X))^2$ равно самой этой константе. В результате получаем

$$D(X) = M\left(X^2\right) - 2M(X)M(X) + (M(X))^2 = M\left(X^2\right) - (M(X))^2. \blacksquare$$

2. $D(X) \geq 0$. Действительно, формулы (31) и (32) приводят к суммам неотрицательных слагаемых. $\blacksquare$

3. $D(kX) = k^2 D(X)$.

Доказательство. Применим формулу (33)

$$D(kX) = M\left((kX)^2\right) - (M(kX))^2 = M\left(k^2 X^2\right) - (kM(X))^2 =$$

$$= k^2 M\left(X^2\right) - k^2 (M(X))^2 = k^2\left(M\left(X^2\right) - (M(X))^2\right) =$$

$$= k^2 D(X). \blacksquare$$

4. Если случайная величина X является постоянной («неслучайной») величиной: $X = a = \text{const}$, т. е. имеет закон распределения

$$X: \quad a$$
$$P: \quad 1,$$

то $D(X)=0$.

Доказательство. По свойству математического ожидания $M(X)=a,\ M\left(X^2\right)=a^2$. Поэтому в формуле (33) для дисперсии

$$D(X)=M(a^2)-(M(a))^2=a^2-a^2=0. \blacksquare$$

5. Обратно, если $D(X)=0$, то случайная величина X является постоянной, $X=a=\text{const}$.

Доказательство. Пусть, X — дискретная случайная величина, и $D(X)=0$. Если бы она с ненулевыми вероятностями принимала, по крайней мере, два разных значения, т. е. имела закон распределения $(x_1,x_2,...;\ p_1,p_2,...)$,

то $$D(X)=p_1(x_1-M(X))^2+p_2(x_2-M(X))^2+...>0,$$

поскольку, по крайней мере, одно из неотрицательных слагаемых строго больше нуля. $\blacksquare$

6. **Теорема сложения для дисперсии.** *Если случайные величины X и Y независимы, то* $D(X\pm Y)=D(X)+D(Y)$.

Доказательство. По теореме умножения для математических ожиданий $M(XY)=M(X)M(Y)$. Применим к $D(X\pm Y)$ формулу (7):

$$D(X\pm Y)=M\left((X\pm Y)^2\right)-(M(X\pm Y))^2=$$
$$=M\left(X^2\pm 2XY+Y^2\right)-(M(X)\pm M(Y))^2=$$
$$=M(X^2)\pm 2M(XY)+M(Y^2)-(M(X))^2\mp 2M(X)M(Y)-(M(Y))^2=$$
$$=\left[M(X^2)-(M(X))^2\right]+\left[M(Y^2)-(M(Y))^2\right]+$$
$$\pm[2M(X)M(Y)-2M(X)M(Y)]=$$
$$=D(X)+D(Y)+0. \blacksquare$$

Замечание. Аналогичное утверждение имеет место и для суммы/разности нескольких независимых случайных величин.
Например,
$$D(X+Y+Z)=D((X+Y)+Z)=D(X+Y)+D(Z)=D(X)+D(Y)+D(Z).$$

Лемма (о независимости константы). *Если случайная величина X является постоянной:* $X=a=\text{const}$, *т. е. с вероятностью* 1 *принимает значение* a, *то для всякой случайной величины Y имеет место независимость X и Y.*

Доказательство. Пусть $<c_1, d_1>$ и $<c_2, d_2>$ произвольные промежутки. Рассмотрим два возможных случая:

1) $a \in <c_1, d_1>$. Тогда событие $(X \in <c_1, d_1>) = \Omega$ является достоверным, и $P(X \in <c_1, d_1>) = 1$.

Далее

$$P\big((X \in <c_1, d_1>) \cdot (Y \in <c_2, d_2>)\big) =$$
$$= P\big(\Omega \cdot (Y \in <c_2, d_2>)\big) = P\big(Y \in <c_2, d_2>\big) =$$
$$= 1 \cdot P\big(Y \in <c_2, d_2>\big) = P\big((X \in <c_1, d_1>) \cdot P(Y \in <c_2, d_2>\big),$$

так что условие независимости выполнено;

2) $a \notin <c_1, d_1>$. Тогда событие $(X \in <c_1, d_1>) = \varnothing$ является невозможным, и $P(X \in <c_1, d_1>) = 0$.

Далее

$$P\big((X \in <c_1, d_1>) \cdot (Y \in <c_2, d_2>)\big) = P\big(\varnothing \cdot (Y \in <c_2, d_2>)\big) = P(\varnothing) = 0 =$$
$$= 0 \cdot P\big(Y \in <c_2, d_2>\big) = P\big((X \in <c_1, d_1>) \cdot P(Y \in <c_2, d_2>\big). \blacksquare$$

Следствие. *Если случайная величина X является постоянной, $X = a = \text{const}$, то для всякой случайной величины Y имеет место равенство $D(a + Y) = D(Y)$.*

Доказательство. Поскольку случайные величины $X = a$ и Y независимы, то

$$D(a + Y) = D(a) + D(Y) = 0 + D(Y) = D(Y). \blacksquare$$

7. **Математическое ожидание и дисперсия среднего арифметического**

Теорема. *Если случайные величины $X_1, X_2, ..., X_n$ независимы и имеют одинаковое математическое ожидание, равное m, и одинаковую дисперсию, равную d, то для их среднего арифметического $Y = \frac{1}{n}(X_1 + X_2 + ... + X_n)$ справедливы формулы*:

$$M(Y) = m; \quad D(Y) = \frac{d}{n}. \tag{34}$$

Доказательство. Прежде всего, среднее арифметическое Y имеет такое же математическое ожидание, как и X

$$M\left(\frac{1}{n}(X_1 + X_2 + ... + X_n)\right) = \frac{1}{n}\big(M(X_1) + M(X_2) + ... + M(X_n)\big) = \frac{1}{n} nm = m.$$

Далее, используя уже доказанные свойства дисперсии, получаем

$$D(Y) = D\left(\frac{1}{n}(X_1 + X_2 + ... + X_n)\right) = = \frac{1}{n^2}(D(X_1 + X_2 + ... + X_n)) =$$

$$= \frac{1}{n^2}(D(X_1) + D(X_2) + ... + D(X_n)) = \frac{1}{n^2} nd = \frac{d}{n}. \blacksquare$$

Замечание. Из формулы (34) следует, что дисперсия среднего арифметического в n раз меньше исходной дисперсии отдельного слагаемого. Иными словами, *среднее арифметическое имеет меньшее рассеивание вокруг математического ожидания* m. Это связано с тем, что в среднем арифметическом при суммировании отклонения разных знаков в значительной степени погашают друг друга.

Последнее находит применение в практике измерений. Так, например, в навигации принято производить измерения по приборам трижды и в качестве результата брать среднее арифметическое полученных значений.

4.9. Биномиальное распределение

Определение

Дискретная случайная величина X имеет *биномиальное распределение* (распределена по *биномиальному закону*), если её возможные значения $0, 1, 2, ..., k, ..., n-1, n$ выражают число успехов k в n испытаниях по схеме Бернулли с вероятностью успеха p, а соответствующие вероятности равны вероятностям числа успехов $P(X = k) =$ $= P_n(k) = C_n^k p^k q^{n-k}$. Таким образом, закон биномиального распределения имеет вид

$$(0, 1, ..., n; \;\; P_n(0); P_n(1), ..., P_n(n)).$$

Для отыскания числовых характеристик биномиального распределения (математического ожидания и дисперсии) введём вспомогательные случайные величины: Z_k — *индикатор* k-го испытания ($k = 1, 2, ..., n$):

$Z_k = 0$, если в k-м испытании имела место неудача;

$Z_k = 1$, если в k-м испытании имел место успех.

Случайные величины Z_k независимы, поскольку связаны с исходами независимых испытаний, и имеют одинаковые законы распределения $(0, 1; \;\; p, q)$. Найдем их числовые характеристики:

$$M(Z_k) = q \cdot 0 + p \cdot 1 = p;$$

$$D(Z_k) = q(0 - p)^2 + p(1 - p)^2 = qp^2 + pq^2 = pq(p + q) = pq.$$

Исходная случайная величина (число успехов) X равна сумме индикаторов $X=\sum_{k=1}^{n} Z_k$ (в сумме справа столько единиц, сколько раз в n испытаниях имел место успех, а остальные слагаемые равны нулю).

По свойствам математического ожидания и дисперсии:

$$M(X)=M\left(\sum_{k=1}^{n} Z_k\right)=\sum_{k=1}^{n} M(Z_k)=np\,;$$

$$D(X)=D\left(\sum_{k=1}^{n} Z_k\right)=\sum_{k=1}^{n} D(Z_k)=npq\,.$$

Итак, *для случайной величины X, имеющей биномиальное распределение,*

$$M(X)=np;\;\; D(X)=npq\;.$$

4.10. Распределение Пуассона

Определение

Дискретная случайная величина X с бесконечным множеством значений распределена по *закону Пуассона с параметром* $\lambda>0$, если она принимает значения $k=0,1,2,...,n,...$ с вероятностями

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}\,.$$

Распределение Пуассона имеют случайные величины, описывающие, например, работу АТС (пример системы массового обслуживания), катодную эмиссию электронов.

Найдем математическое ожидание и дисперсию распределения Пуассона, имея в виду, что

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=1+\lambda+\frac{\lambda^2}{2!}+...+\frac{\lambda^k}{k!}+...=e^{\lambda}\;\text{—}$$

ряд Маклорена для показательной функции.

1. $M(X)=\sum_{k=0}^{\infty} kP(X=k)=\sum_{k=0}^{\infty} k\frac{\lambda^k e^{-\lambda}}{k!}$. Первое слагаемое ряда равно нулю из-за множителя $k=0$; далее, при $k\ge 1$ после сокращения в каждом слагаемом ряда получаем: $\frac{k}{k!}=\frac{1}{(k-1)!}$. Поэтому, вынося постоянный множитель $\lambda e^{-\lambda}$ за знак ряда, получаем

$$M(X)=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}=\lambda e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\ldots\right)=\lambda e^{-\lambda}e^{\lambda}=\lambda .$$

2. Используем для вычисления дисперсии формулу (33) и уже известное значение математического ожидания

$$D(X)=M\left(X^2\right)-(M(X))^2=M\left(X^2\right)-\lambda^2 .$$

Случайная величина X^2 имеет закон распределения

$$\left(0^2,1^2,2^2,\ldots,k^2,\ldots;\ e^{-\lambda},\frac{\lambda e^{-\lambda}}{1!},\frac{\lambda^2 e^{-\lambda}}{2!},\ldots,\frac{\lambda^k e^{-\lambda}}{k!},\ldots\right).$$

Поэтому

$$M\left(X^2\right)=\sum_{k=0}^{\infty}k^2\frac{\lambda^k e^{-\lambda}}{k!}=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{k\lambda^{k-1}}{(k-1)!}=\lambda e^{-\lambda}\left(1+\frac{2\lambda}{1!}+\frac{3\lambda^2}{2!}+\ldots+\frac{(k+1)\lambda^k}{k!}+\ldots\right).$$

Каждую дробь в скобках представляем в виде суммы двух слагаемых:

$$\frac{2\lambda}{1!}=\frac{\lambda}{1!}+\frac{\lambda}{1!}=\frac{\lambda}{1!}+\lambda\ ;\ \frac{3\lambda^2}{2!}=\frac{\lambda^2}{2!}+\frac{2\lambda^2}{2!}=\frac{\lambda^2}{2!}+\frac{\lambda^2}{1!}\ ;\ \ldots;$$

$$\frac{(k+1)\lambda^k}{k!}=\frac{\lambda^k}{k!}+\frac{k\lambda^k}{k!}=\frac{\lambda^k}{k!}+\frac{\lambda^k}{(k-1)!}\ \text{ и т. д.}$$

Группируя сначала все первые слагаемые, а затем из суммы вторых слагаемых вынося общий множитель λ, получаем:

$$M\left(X^2\right)=\lambda e^{-\lambda}\left[\left(1+\lambda+\frac{\lambda^2}{2!}+\ldots+\frac{\lambda^k}{k!}+\ldots\right)+\right.$$

$$\left.+\lambda\left(1+\lambda+\frac{\lambda^2}{2!}+\ldots+\frac{\lambda^k}{k!}+\ldots\right)\right]=\lambda e^{-\lambda}[e^{\lambda}+\lambda e^{\lambda}]=\lambda(1+\lambda) .$$

Отсюда по формуле (7)

$$D(X)=(\lambda+\lambda^2)-\lambda^2=\lambda .$$

Итак, *для случайной величины* X, *имеющей распределение Пуассона с параметром* λ:

$$M(X)=\lambda;\ \ D(X)=\lambda.$$

5. НЕПРЕРЫВНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

5.1. Плотность непрерывной случайной величины

Напомним, что функция распределения $F(X)$ случайной величины X определяется формулой: $F(X) = P(X < x)$.

Определение
Случайная величина X называется *непрерывной*, если существует неотрицательная кусочно-непрерывная функция $p(x)$, интегрируемая на $(-\infty, +\infty)$ и такая, что функция распределения представима в виде интеграла с переменным верхним пределом

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

Функция $p(x)$ называется в этом случае *плотностью распределения* случайной величины.

Свойства плотности распределения

1. В точках непрерывности плотность является производной функции распределения

$$F'(X) = p(x)\,. \tag{35}$$

В случае, когда плотность $p(x)$ непрерывна на всей числовой оси, это следует из теоремы о производной интеграла с переменным верхним пределом [12]. ■

Следствие. В точках непрерывности плотности $p(x)$ функция распределения $F(X)$ дифференцируема (а значит, и непрерывна) и является первообразной для плотности.

2.
$$\int_{-\infty}^{+\infty} p(x)\,dx = 1\,. \tag{36}$$

Доказательство. По определению несобственного интеграла

$$\int_{-\infty}^{+\infty} p(x)\,dx = \int_{-\infty}^{0} p(x)\,dx + \int_{0}^{+\infty} p(x)\,dx =$$
$$= \big(F(0) - F(-\infty)\big) + \big(F(+\infty) - F(0)\big) =$$
$$= \big(F(0) - 0\big) + \big(1 - F(0)\big) = 1\,. \ ■$$

3. Для непрерывной случайной величины X вероятность принять значение из полуоткрытого промежутка $[a, b)$ («вероятность попадания в промежуток») равна интегралу от плотности по этому промежутку

$$P(a \le X < b) = \int_a^b p(x)\,dx\,. \tag{37}$$

Доказательство. По свойству функции распределения

$$P(a \le X < b) = F(b) - F(a) = \int_{-\infty}^{b} p(x)\,dx - \int_{-\infty}^{a} p(x)\,dx = \int_a^b p(x)\,dx\,. \blacksquare$$

Замечание. Поскольку определённый интеграл в формуле (37) равен площади криволинейной трапеции, то при одинаковой длине промежутков больше вероятность попадания в тот из них, у которого больше площадь соответствующей криволинейной трапеции. Так, для плотности, график которой изображён на рис. 15, вероятность попадания в промежуток $[1, 2)$ больше, чем вероятность попадания в промежуток $[3, 4)$.

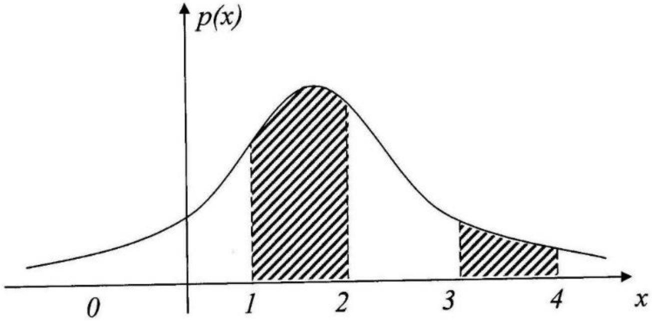

Рис. 15

5.2. Особенность непрерывной случайной величины

Теорема. *Для непрерывной случайной величины X вероятность принять любое наперёд заданное значение равна нулю*

$$\forall a:\ P(X = a) = 0$$

(иными словами, «вероятность попадания в точку» равна нулю).

Замечание. Таким образом, непрерывные случайные величины дают примеры случайных событий, а именно $(X = a)$, которые, не будучи невозможными, имеют, тем не менее, вероятность, равную нулю.

Доказательство. Для любого $h>0$ событие $(X=a)$ влечёт за собой событие $(a\le X<a+h)$, поэтому в силу непрерывности функции распределения

$$0\le P(X=a)\le P(a\le X<a+h)=F(a+h)-F(a)\underset{h\to+0}{\to}0.$$

Отсюда $P(X=a)=0$. ■

Следствие. Для непрерывной случайной величины вероятность попадания в промежуток не зависит от вида промежутка

$$P(a<X<b)=P(a\le X\le b)=P(a\le X<b)=P(a<X\le b)$$

(все эти вероятности равны $F(b)-F(a)$).

Доказательство. Докажем, например, первое равенство.
Имеем $(a\le X\le b)=(a<X<b)+(X=a)+(X=b)$ — сумма попарно несовместных событий.
Отсюда:

$$P(a\le X\le b)=P(a<X<b)+P(X=a)+P(X=b)=$$
$$=P(a<X<b)+0+0=P(a<X<b). ■$$

5.3. Вероятностный смысл плотности распределения

Если случайная величина X непрерывна, то для её плотности $p(x)$ в точках дифференцируемости имеем

$$p(x)=F'(x)=\lim_{\Delta x\to0}\frac{F(x+\Delta x)-F(x)}{\Delta x}.$$

Поэтому при малых $\Delta x>0$ имеет место приближённое равенство

$$d(F(x))=p(x)\Delta x\approx F(x+\Delta x)-F(x)=P(X\in[x,x+\Delta x]).$$

Итак, *при малых $\Delta x>0$ вероятность того, что непрерывная случайная величина примет значение из отрезка $[x,x+\Delta x]$, приближённо равна произведению плотности в точке x на длину отрезка*

$$P(X\in[x,x+\Delta x])\approx p(x)\Delta x.$$

5.4. Математическое ожидание непрерывной случайной величины

I. Наводящее рассуждение

Для дискретной случайной величины с законом распределения $(x_1,x_2,...,x_n;\ p_1,p_2,...,p_n)$ математическое ожидание вычисляется, согласно п. 1.5, по формуле

$$M(X)=p_1x_1+p_2x_2+...+p_nx_n\ .$$

Если отдельные значения x_i можно окружить непересекающимися малыми промежутками с длинами Δx_i, то вероятности p_i этих значений являются одновременно вероятностями попадания в соответствующие отрезки

$$p_i=P(X\in[x_i,x_i+\Delta x_i]),$$

и формула для математического ожидания принимает вид

$$M(X)=\sum_{i=1}^{n}x_i\cdot P(X\in[x_i,x_i+\Delta x_i]).\qquad(38)$$

Для непрерывной случайной величины X имеем

$$P(X\in[x_i,x_i+\Delta x_i])\approx p(x_i)\Delta x_i,$$

так что сумма, аналогичная (38), имеет вид

$$\sum_{i=1}^{n}x_i\,p(x_i)\Delta x_i$$

и является интегральной суммой для несобственного интеграла $\int\limits_{-\infty}^{+\infty}xp(x)\,dx$.

II. Определение математического ожидания

Определение

Математическим ожиданием непрерывной случайной величины X с плотностью $p(x)$ называется несобственный интеграл

$$M(X)=\int\limits_{-\infty}^{+\infty}xp(x)\,dx,\qquad(39)$$

причем предполагается, что этот интеграл сходится абсолютно, т. е. сходится интеграл $\int\limits_{-\infty}^{+\infty}|x|\,p(x)\,dx$.

Если абсолютной сходимости нет, то для такой непрерывной случайной величины математическое ожидание не определено.

III. Математическое ожидание функции случайного аргумента

Пусть X — непрерывная случайная величина с плотностью $p(x)$, и $f(u)$ — функция числового аргумента, которая непрерывна на $(-\infty,+\infty)$. Тогда $f(X)$ принимает вместе со случайным аргументом X случайные значения и является случайной величиной.

Можно доказать, что *случайная величина* $f(X)$ *также является непрерывной, и для её математического ожидания справедлива формула*

$$M(f(X)) = \int\limits_{-\infty}^{+\infty} f(x)p(x)\,dx. \qquad (40)$$

IV. Свойства математического ожидания непрерывной случайной величины

Эти свойства аналогичны свойствам математического ожидания в случае дискретной случайной величины (п.1.6); перечислим их заново.

1. Постоянный множитель можно выносить за знак математического ожидания $M(kX) = kM(X)$.

2. Пусть случайные величины X и Y имеют математические ожидания $M(X)$ и $M(Y)$. Тогда $M(X+Y) = M(X) + M(Y)$.

3. Пусть случайные величины X и Y независимы и имеют математические ожидания $M(X)$ и $M(Y)$. Тогда $M(XY) = M(X)M(Y)$.

5.5. Дисперсия непрерывной случайной величины

Напомним (п. 1.7), что дисперсия служит мерой рассеивания значений случайной величины вокруг математического ожидания и определяется как математическое ожидание квадрата отклонения

$$D(X) = M\left((X - M(X))^2\right).$$

Квадрат отклонения $(X - M(X))^2$ является частным случаем функции случайного аргумента X, а именно, когда

$$f(u) = (u - M(X))^2.$$

Поэтому, в соответствии с общей формулой (40), для дисперсии непрерывной случайной величины X с плотностью $p(x)$ получаем формулу

$$D(X) = \int\limits_{-\infty}^{+\infty} (x - M(X))^2\, p(x)dx. \qquad (41)$$

Если несобственный интеграл в формуле (41) расходится, то считают, что дисперсия не существует.

Свойства дисперсии непрерывной случайной величины также аналогичны свойствам дисперсии в дискретном случае (п. 1.6); перечислим их заново.

1. $D(X) = M\left(X^2\right) - (M(X))^2 = \int\limits_{-\infty}^{+\infty} x^2 p(x)dx - (M(X))^2.$

2. $D(X) \geq 0$.

3. $D(kX) = k^2 D(X)$.

4. Если случайные величины X и Y независимы, то $D(X \pm Y) = D(X) + D(Y)$.

Для непрерывной случайной величины сохраняется определение среднего квадратического отклонения $\sigma(X) = \sqrt{D(X)}$.

5.6. Нормальное распределение

Определение
Непрерывная случайная величина X имеет *нормальное распределение (распределение Гаусса) с параметрами* a *и* $\sigma > 0$, если её плотность имеет вид

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}. \qquad (42)$$

График плотности нормального распределения, изображённый на рис. 16, называется *нормальной кривой* или *кривой Гаусса*.

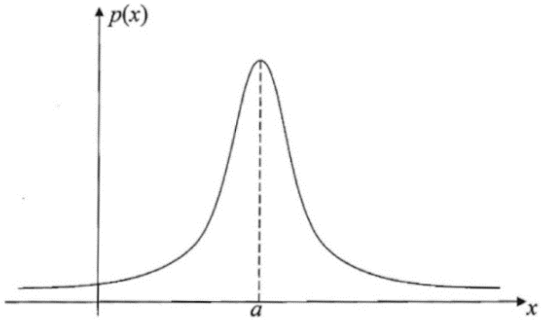

Рис. 16

Из графика видно, что из интервалов значений одинаковой длины более близкие к a имеют бóльшую вероятность, попадание в них происходит чаще. При удалении от a вероятность попадания в интервал уменьшается.

Такое поведение вероятностей, а значит, и относительных частот, характерно для многих случайных величин. Например, если a — средний рост, а X — рост произвольно выбранного человека, то люди, чей рост близок к среднему, встречаются часто, а «великаны» и «карлики» — крайне редко.

Замечание. При $a = 0$, $\sigma = 1$ плотность нормального распределения является дифференциальной функцией Лапласа

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

Функция распределения F_N в этом случае задаётся выражением

$$F_N(x)=\int_{-\infty}^{x}\varphi(t)dt=\int_{-\infty}^{0}\varphi(t)dt+\int_{0}^{x}\varphi(t)dt=\frac{1}{2}+\Phi(x)\,,$$

где $\Phi(x)$ — интегральная функция Лапласа. Итак,

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

Зависимость нормальной кривой от параметра a (при $a_1<a_2$ и постоянном σ) изображена на рис. 17. Вертикальная прямая $x=a$ является осью симметрии нормальной кривой.

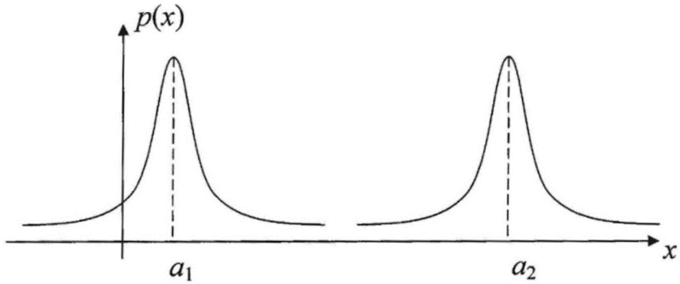

Рис. 17

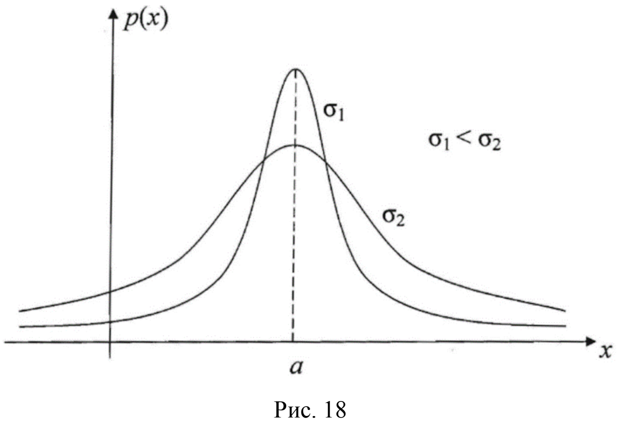

Рис. 18

Зависимость нормальной кривой от параметра σ (при постоянном a) изображена на рис. 18. При увеличении σ кривая становится более пологой, так что далекие от a значения случайной величины X приобретают большую вероятность, реализуются чаще; разброс значений вокруг a при этом увеличивается.

Эти свойства нормальной кривой проясняет следующая теорема.

Теорема (о вероятностном смысле параметров a и σ). *Для случайной величины* X*, имеющей нормальное распределение с параметрами* a *и* σ

$$M(X) = a; \quad D(X) = \sigma^2.$$

Доказательство. При вычислении интегралов, выражающих $M(X)$ и $D(X)$, воспользуемся заменой переменной. В соответствии с формулой (13)

$$M(X) = \int_{-\infty}^{+\infty} xp(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} xe^{-\frac{(x-a)^2}{2\sigma^2}}\,dx =$$

$$= \left\{ \begin{array}{c} \dfrac{x-a}{\sigma} = t;\ x = \sigma t + a;\ dx = \sigma dt \\ x = -\infty \ \Rightarrow\ t = -\infty \\ x = +\infty \ \Rightarrow\ t = +\infty \end{array} \right\} =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma t + a) e^{-\frac{t^2}{2}} \sigma\,dt =$$

$$= \frac{\sigma^2}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} te^{-\frac{t^2}{2}}\,dt + \frac{a\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt\ .$$

В полученном выражении первое слагаемое является сходящимся интегралом от нечётной функции $te^{-\frac{t^2}{2}}$ по симметричному промежутку и поэтому равно нулю. Второе слагаемое представляет собою умноженный на $\frac{a}{\sqrt{2\pi}}$ *интеграл Пуассона*, про который известно, что он равен $\sqrt{2\pi}$

$$\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}$$

(в этом смысл множителя $\frac{1}{\sqrt{2\pi}}$ в формуле для плотности — он обеспечивает выполнение равенства (36)). В результате получаем $M(X) = a$.

Аналогичными вычислениями устанавливается равенство $D(X) = \sigma^2$. ■

Получим выражение для функции распределения $F_{a,\sigma}(x)$ произвольного нормального закона с параметрами a и σ:

$$F_{a,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt =$$

(проводим замену переменной)

$$= \left\{ \begin{array}{ll} \dfrac{t-a}{\sigma} = u; & t = -\infty \Rightarrow u = -\infty; \\ t = \sigma u + a & t = x \Rightarrow u = \dfrac{x-a}{\sigma}; \\ dt = \sigma du & \end{array} \right\} =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{u^2}{2}} du = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right).$$

Нормальное распределение играет исключительно важную роль при математическом описании многих процессов, имеющих вероятностную, случайную (говорят также *стохастическую*) природу.

5.7. Показательное распределение

Определение

Непрерывная случайная величина X имеет *показательное (экспоненциальное) распределение с параметром* $\lambda > 0$, если её плотность имеет вид

$$p(x) = \begin{cases} 0, \text{ если } x < 0; \\ \lambda e^{-\lambda x}, \text{ если } x \geq 0. \end{cases} \qquad (43)$$

График плотности показательного распределения изображён на рис. 19.

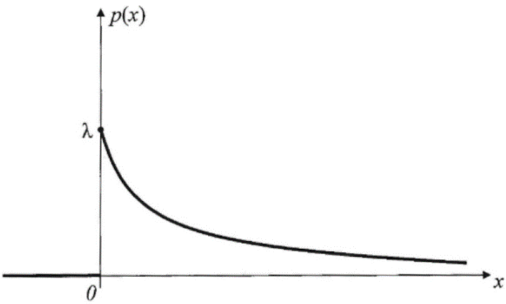

Рис. 19

Теорема. *Если непрерывная случайная величина имеет показательное распределение с параметром* λ, *то для её математического ожидания и дисперсии справедливы формулы:*

$$M(X)=\frac{1}{\lambda};\ \ D(X)=\frac{1}{\lambda^2}.$$

Доказательство. Заметим, что по правилу Лопиталя

$$\lim_{x\to+\infty} xe^{-\lambda x}=\lim_{x\to+\infty}\frac{x}{e^{\lambda x}}=\left(\frac{\infty}{\infty}\right)=\lim_{x\to+\infty}\frac{(x)'}{\left(e^{\lambda x}\right)'}=\lim_{x\to\infty}\frac{1}{\lambda e^{\lambda x}}=0.$$

Интегрируя по частям, получим

$$M(X)=\int_{-\infty}^{+\infty} xp(x)\,dx=\int_{-\infty}^{0} x\cdot 0\,dx+\int_{0}^{+\infty} x\lambda e^{-\lambda x}\,dx=$$

$$=0+\int_{0}^{+\infty}\lambda e^{-\lambda x}x\,dx=\left\{\begin{array}{l|l} u=x & du=dx \\ dv=\lambda e^{-\lambda x}dx & v=-e^{-\lambda x}\end{array}\right\}=$$

$$=-xe^{-\lambda x}\Big|_{0}^{+\infty}+\int_{0}^{+\infty} e^{-\lambda x}dx=$$

$$=-\left(\lim_{x\to+\infty} xe^{-\lambda x}-0\right)+\frac{1}{-\lambda}\int_{0}^{+\infty} e^{-\lambda x}d(-\lambda x)=$$

$$-(0-0)+\frac{1}{-\lambda}\lim_{x\to+\infty}\left(e^{-\lambda x}-1\right)=\frac{1}{-\lambda}(0-1)=\frac{1}{\lambda}.$$

Далее

$$D(X)=M(X^2)-\left(M(X)\right)^2=M(X^2)-\frac{1}{\lambda^2}.$$

Согласно формуле (40)

$$M(X^2)=\int_{-\infty}^{+\infty} p(x)x^2\,dx=\lambda\int_{0}^{+\infty} e^{-\lambda x}x^2dx.$$

Аналогично предыдущему, интегрируя по частям дважды, получим $M(X^2)=\dfrac{2}{\lambda^2}$, так что $D(X)=\dfrac{2}{\lambda^2}-\dfrac{1}{\lambda^2}=\dfrac{1}{\lambda^2}$. ■

Аналогичными вычислениями получается выражение для функции распределения показательного распределения

$$F(x)=\int_{-\infty}^{x} p(t)dt=\begin{cases}0, & x<0;\\ 1-e^{-\lambda x}, & x\ge 0.\end{cases}$$

5.8. Равномерное распределение

Определение
Непрерывная случайная величина X имеет *равномерное распределение на отрезке* $[a,b]$, если её плотность имеет вид:

$$p(x)=\begin{cases} 0, \text{ если } x\notin[a,b]; \\ c, \text{ если } x\in[a,b]. \end{cases} \qquad (44)$$

График плотности равномерного распределения изображён на рис. 20.

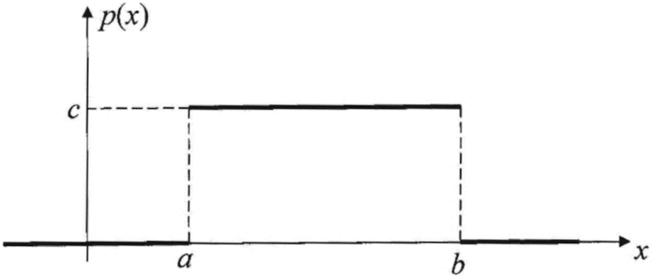

Рис. 20

Теорема. *Если непрерывная случайная величина имеет равномерное распределение на отрезке* $[a,b]$*, то:*

$$c=\frac{1}{b-a};\ M(X)=\frac{a+b}{2};\ D(X)=\frac{(b-a)^2}{12}.$$

Замечание. Число $\frac{a+b}{2}$ — середина отрезка $[a,b]$; число $(b-a)$ — длина отрезка $[a,b]$;

Доказательство. Для определения параметра c воспользуемся свойством плотности (36)

$$1=\int_{-\infty}^{+\infty}p(x)\,dx=\int_a^b c\,dx=c(b-a),$$

откуда

$$c=\frac{1}{b-a}.$$

Далее

$$M(X)=\int_{-\infty}^{+\infty}p(x)x\,dx=\frac{1}{b-a}\int_a^b x\,dx=\frac{1}{b-a}\cdot\frac{x^2}{2}\Big|_a^b=$$

$$=\frac{1}{b-a}\cdot\frac{b^2-a^2}{2}=\frac{a+b}{2}.$$

$$D(X)=M(X^2)-\left(M(X)\right)^2=\frac{1}{b-a}\int_a^b x^2dx-\left(\frac{a+b}{2}\right)^2=$$

$$=\frac{1}{b-a}\cdot\frac{x^3}{3}\Bigg|_a^b-\frac{a^2+2ab+b^2}{4}=\frac{1}{b-a}\cdot\frac{b^3-a^3}{3}-\frac{a^2+2ab+b^2}{4}=$$

$$=\frac{4(b^2+ab+a^2)-3(a^2+2ab+b^2)}{12}=\frac{b^2-2ab+a^2}{12}=\frac{(b-a)^2}{12}.\ \blacksquare$$

Замечание. Полученные значения математического ожидания и дисперсии равномерного распределения хорошо иллюстрируют их статистический смысл. Так, в силу симметрии графика плотности относительно середины отрезка, при большом числе реализаций случайной величины одинаково часто будут встречаться значения случайной величины с обеих сторон от этой середины. Поэтому среднее арифметическое должно оказаться близким к ней. Чем больше длина отрезка, т. е. число $b-a$, тем на большем промежутке «размазаны» возможные значения, тем больше должна быть дисперсия, которая как раз и пропорциональна квадрату длины отрезка $[a,b]$.

Аналогичными вычислениями получается выражение для функции распределения равномерного распределения

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

5.9. Преобразование случайных величин

I. Линейное преобразование нормального закона

Теорема. *Если случайная величина X распределена по нормальному закону с параметрами a и σ, то при $k\neq 0$ случайная величина $Y=kX+b$, полученная из X линейным преобразованием вида $y=kx+b$, также имеет нормальное распределение с параметрами $a_1=ka+b$ и $\sigma_1=|k|\sigma$.*

Доказательство. Рассмотрим случай $k>0$. Тогда функция $y=kx+b$ строго возрастает, и имеет место равносильность неравенств:

$$Y<x\Leftrightarrow kX+b<x\Leftrightarrow X<\frac{x-b}{k}.$$

Поэтому также имеет место равенство событий

$$(Y < x) = \left(X < \frac{x-b}{k} \right).$$

Найдем функцию распределения $F_Y(x)$ случайной величины Y и убедимся, что она соответствует нормальному распределению.

$$F_Y(x) = P(Y < x) = P\left(X < \frac{x-b}{k} \right) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{x-b}{k}} e^{-\frac{(t-a)^2}{2\sigma^2}} dt =$$

$$= \left\{ \begin{array}{l} u = kt + b;\ t = \dfrac{u-b}{k};\ dt = \dfrac{du}{k} \\ t = -\infty \ \Rightarrow \ u = -\infty \\ t = \dfrac{x-b}{k} \ \Rightarrow \ u = x \end{array} \right\} =$$

$$= \frac{1}{(k\sigma)\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2\sigma^2}\left(\frac{u-b}{k} - a\right)^2} du =$$

$$= \frac{1}{(k\sigma)\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(u-(ka+b))^2}{2(k\sigma)^2}} du =$$

$$= \frac{1}{\sigma_1\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(u-a_1)^2}{2\sigma_1^2}} du.$$

Пришли к функции распределения нормального закона с параметрами $a_1 = ka + b$ и $\sigma_1 = |k|\sigma$. Аналогично рассматривается и случай $k < 0$. ■

II. Общий случай преобразования случайной величины

Пусть теперь $y = f(x)$ — не линейная, а произвольная непрерывная строго монотонная функция. Аналогичными выкладками можно при заданных функции распределения $F_X(x)$ и плотности $p_X(x)$ случайной величины X найти функцию распределения $F_Y(x)$ и плотность $p_Y(x)$ случайной величины $Y = f(X)$. Дело сводится к замене переменной в соответствующем интеграле.

Пример. Пусть $Y = X^3$, т. е. $f(x) = x^3$. Из равносильности неравенств $Y < x \Leftrightarrow X < \sqrt[3]{x}$ вытекает равенство событий $(Y < x) = \left(X < \sqrt[3]{x}\right)$.

Поэтому $$F_Y(x) = P(Y < x) = P(X < \sqrt[3]{x}) = \int_{-\infty}^{\sqrt[3]{x}} p_X(t)\,dt =$$

$$= \left\{ \begin{array}{l} t = \sqrt[3]{u};\ \ u = t^3;\ \ dt = \dfrac{1}{3}u^{-\frac{2}{3}} \\ t = -\infty \Rightarrow u = -\infty \\ t = \sqrt[3]{x} \Rightarrow u = x \end{array} \right\} =$$

$$= \int_{-\infty}^{x} \frac{1}{3}u^{-\frac{2}{3}} p_X(\sqrt[3]{u})\,du\ .$$

Подынтегральная функция в последнем выражении является плотностью распределения случайной величины Y

$$p_Y(x) = \frac{1}{3}u^{-\frac{2}{3}} p_X(\sqrt[3]{u})\ .$$

5.10. Вероятность попадания в промежуток для нормального распределения

I. Вероятность попадания в произвольный промежуток

Теорема. *Пусть непрерывная случайная величина* X *имеет нормальное распределение с параметрами* a *и* σ. *Тогда для всякого промежутка* $\langle A, B\rangle$ *вероятность попадания значения* X *в этот промежуток задаётся формулой*

$$P(X \in \langle A, B\rangle) = \Phi\left(\frac{B-a}{\sigma}\right) - \Phi\left(\frac{A-a}{\sigma}\right)\ . \qquad (45)$$

Доказательство. Поскольку для непрерывной случайной величины вероятность попадания в промежуток не зависит от типа промежутка (п. 2.2), докажем формулу (45) для интервала (A, B). Введём случайную величину

$$Y = \frac{X-a}{\sigma} = \frac{1}{\sigma}X + \left(-\frac{a}{\sigma}\right).$$

Она получена из X линейным преобразованием. По предыдущей теореме Y имеет нормальное распределение с параметрами $a_1 = \dfrac{1}{\sigma}a + \left(-\dfrac{a}{\sigma}\right) = 0$ и $\sigma_1 = \left|-\dfrac{1}{\sigma}\right|\sigma = 1$. Её плотностью является дифференциальная функция Лапласа $\varphi(x)$, одной из первообразных которой является интегральная функция Лапласа $\Phi(x)$.

Ввиду равносильности неравенств

$$A < X < B \Leftrightarrow \frac{A-a}{\sigma} < \frac{X-a}{\sigma} < \frac{B-a}{\sigma} \Leftrightarrow \frac{A-a}{\sigma} < Y < \frac{B-a}{\sigma},$$

получаем для вероятностей

$$P(A < X < B) = P\left(\frac{A-a}{\sigma} < Y < \frac{B-a}{\sigma}\right) = \int_{\frac{A-a}{\sigma}}^{\frac{B-a}{\sigma}} \varphi(x)\,dx .$$

Применяя к последнему интегралу формулу Ньютона-Лейбница с первообразной Φ, получаем окончательно

$$P(A < X < B) = \Phi\left(\frac{B-a}{\sigma}\right) - \Phi\left(\frac{A-a}{\sigma}\right). \blacksquare$$

Пример. Пусть X имеет нормальное распределение с параметрами a = 30 и σ = 10. Найдем вероятность попадания в отрезок [10, 40].

Здесь $A = 10,\ B = 40,\ \frac{A-a}{\sigma} = -2,\ \frac{B-a}{\sigma} = 1$. Учитывая нечётность функции Ф, получаем

$$P(X \in [10, 40]) = \Phi(1) - \Phi(-2) = \Phi(1) + \Phi(2) \approx \approx 0{,}3413 + 0{,}4772 = 0{,}8185.$$

В соответствии с эмпирическим законом больших чисел, следует ожидать, что при большом числе испытаний относительная частота попадания реализованного значения случайной величины в отрезок [10, 40] окажется близкой к 82 %.

II. Вероятность отклонения от математического ожидания

Теорема. *Пусть непрерывная случайная величина X имеет нормальное распределение с параметрами* a *и* σ. *Тогда для всякого* $\delta > 0$ *вероятность отклонения значения X от математического ожидания* a *по модулю меньше чем на* δ, *задаётся формулой*

$$P(|X-a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right). \tag{46}$$

Доказательство. Применим формулу (45) при $A = a - \delta,\ B = a + \delta$, так что $\frac{B-a}{\sigma} = \frac{\delta}{\sigma},\ \frac{A-a}{\sigma} = -\frac{\delta}{\sigma}$. Поскольку функция Φ является нечётной, то

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

Пример. Пусть X имеет нормальное распределение, и $\sigma = 2$. Найдем при $\delta = 1{,}2$ вероятность отклонения от математического ожидания

$$P(|X-a|<1{,}2)=2\Phi\left(\frac{1{,}2}{2}\right)=2\Phi(0{,}6)\approx 2\cdot 0{,}2257=0{,}4514.$$

III. Правило «трёх сигм»

Применим последнюю теорему и формулу (46) к отклонению $\delta=3\sigma$. При этом

$$P(|X-a|<3\sigma)=2\Phi\left(\frac{3\sigma}{\sigma}\right)=2\Phi(3)\approx 2\cdot 0{,}49865=0{,}9973.$$

Итак, для нормально распределённой случайной величины с параметрами a и σ вероятность отклонения реализованного значения от математического ожидания менее чем на 3σ приближённо равна 0,9973. Во многих практических ситуациях случайное событие с такой вероятностью принято считать *практически достоверным*.

Поэтому полагают, что *практически все реализуемые значения нормально распределённой случайной величины с параметрами* a *и* σ *попадают в интервал* $(a-3\sigma, a+3\sigma)$. В этом и заключается «правило трех сигм».

5.11. Корреляция случайных величин

1. Нормированные случайные величины

Определение

Случайная величина X называется *центрированной*, если она имеет математическое ожидание, равное нулю: $M(X)=0$.

Пример. Случайная величина X, распределённая по нормальному закону с параметрами $a=0$ и σ, является центрированной, поскольку $M(X)=a=0$.

Напомним, что для случайной величины X, имеющей математическое ожидание $M(X)=m$, случайная величина $\overset{\circ}{X}=X-m$ называется *отклонением* (отклонением X от математического ожидания).

Теорема. *Отклонение* $\overset{\circ}{X}$ *является центрированной случайной величиной.*

Доказательство. По свойствам математического ожидания

$$M(\overset{\circ}{X})=M(X-m)=M(X)-M(m)=m-m=0. \blacksquare$$

Определение

Случайная величина X называется *нормированной*, если она имеет математическое ожидание, равное нулю, и дисперсию, равную единице: $M(X)=0,\ \ D(X)=1$.

Теорема. *Для случайной величины* X, *у которой* $M(X)=m$, $D(X)=\sigma^2>0$ *(так что* $\sigma=\sigma(X)$ *— среднеквадратическое отклонение), случайная величина*

$$X'=\frac{X-m}{\sigma}=\frac{\overset{\circ}{X}}{\sigma} \tag{47}$$

является нормированной.

Доказательство. По свойствам математического ожидания и дисперсии:

$$M(X')=\frac{1}{\sigma}M(\overset{\circ}{X})=\frac{1}{\sigma}\cdot 0=0\,;$$

$$D(X')=\frac{1}{\sigma^2}\big(D(X)+D(-m)\big)=\frac{1}{\sigma^2}\left(\sigma^2+0\right)=1\,. \blacksquare$$

Теорема. *Для нормированной случайной величины* X' *справедлива формула*

$$D(X')=M(X'^2)\,. \tag{48}$$

Доказательство. По формуле (33)

$$D(X')=M(X'^2)-\big(M(X')\big)^2=M(X'^2)-0^2=M(X'^2)\,. \blacksquare$$

2. Корреляционный момент

Определение

Пусть случайные величины X и Y имеют математические ожидания $M(X)$ и $M(Y)$. Их корреляционным моментом μ_{XY} называется математическое ожидание произведения отклонений

$$\mu_{XY}=M(\overset{\circ}{X}\overset{\circ}{Y})=M\big((X-M(X))\cdot(Y-M(Y))\big)\,.$$

Определения

1. Случайные величины X и Y называются *коррелированными*, если их корреляционный момент не равен нулю: $\mu_{XY}\neq 0$.

2. Случайные величины X и Y называются *некоррелированными*, если их корреляционный момент равен нулю: $\mu_{XY}=0$.

Теорема. *Если случайные величины* X *и* Y *независимы, то их корреляционный момент равен нулю:* $\mu_{XY}=0$.

Доказательство. По свойствам математического ожидания:

$$\mu_{XY}=M\big((X-M(X))\cdot(Y-M(Y))\big)=$$

$$=M\big(XY-M(X)Y-M(Y)X+M(X)M(Y)\big)=$$

$$= M(XY) - M(X)M(Y) - M(X)M(Y) + M(X)M(Y) =$$

$$= M(XY) - M(X)M(Y) = 0$$

(последнее равенство имеет место по теореме умножения для математических ожиданий независимых случайных величин). ■

Следствие. Если случайные величины X и Y являются коррелированными, то они зависимы.

3. Коэффициент корреляции

Определение

Коэффициентом корреляции случайных величин X и Y, имеющих корреляционный момент μ_{XY} и средние квадратические отклонения $\sigma_X > 0,\ \sigma_Y > 0$, называется число

$$r_{XY} = \frac{\mu_{XY}}{\sigma_X \sigma_Y}.$$

В то время как корреляционный момент μ_{XY} является размерной величиной, значение которой зависит от выбора единиц измерения X и Y, коэффициент корреляции r_{XY} является безразмерной величиной.

Теорема (об оценке коэффициента корреляции). *Справедливо неравенство*

$$|r_{XY}| \le 1.$$

Доказательство. Пусть X' и Y' — соответствующие нормированные случайные величины, полученные по формуле (21). Тогда, внося постоянные множители под знак математического ожидания, имеем

$$r_{XY} = \frac{M(\overset{\circ}{X}\overset{\circ}{Y})}{\sigma_X \sigma_Y} = M\left(\frac{\overset{\circ}{X}}{\sigma_X}\frac{\overset{\circ}{Y}}{\sigma_Y}\right) = M(X'Y'). \tag{49}$$

Применим к дисперсии формулу (33)

$$0 \le D(X' \pm Y') = M\left((X' \pm Y')^2\right) - \left(M(X' \pm Y')\right)^2 =$$

$$= \left[M(X'^2) \pm 2M(X'Y') + M(Y'^2)\right] - (0 \pm 0)^2 =$$

(применим формулу (48) к первому и третьему слагаемым, формулу (49) — ко второму)

$$= D(X') \pm 2r_{XY} + D(Y') = 1 \pm 2r_{XY} + 1 = 2 \pm 2r_{XY}.$$

Итак, $$0 \le 2 \pm 2r_{XY} \Rightarrow |r_{XY}| \le 1. \;■$$

Замечание. В ходе доказательства для нормированных случайных величин установлено равенство
$$D(X' \pm Y') = 2(1 \pm r_{XY}). \qquad (50)$$

Теорема (необходимое условие независимости). *Если случайные величины X и Y независимы, то $r_{XY} = 0$.*

Доказательство. Поскольку X и Y независимы, то
$$\mu_{XY} = 0 \Rightarrow r_{XY} = 0. \blacksquare$$

Теорема (критерий линейной связи). *Для того чтобы случайные величины X и Y были связаны функциональной линейной зависимостью вида $Y = kX + b$ $(k \neq 0)$, необходимо и достаточно выполнение условия $|r_{XY}| = 1$.*

Доказательство.

1. *Необходимость*. Пусть $Y = kX + b$; по свойствам математического ожидания и дисперсии
$$M(Y) = kM(X) + b;$$
$$D(Y) = D(kX + b) = k^2 D(X) + 0 = k^2 D(X) \Rightarrow$$
$$\Rightarrow \sigma_Y = \sigma(Y) = \sqrt{D(Y)} = \sqrt{k^2 D(X)} = |k|\sigma_X .$$
Теперь
$$r_{XY} = \frac{M\left((X - M(X)) \cdot (Y - M(Y))\right)}{\sigma_X \sigma_Y} =$$
$$= \frac{M\left((X - M(X))(kX + b - kM(X) - b)\right)}{\sigma_X |k| \sigma_X} =$$
$$= \frac{kM\left((X - M(X))(X - M(X))\right)}{\sigma_X |k| \sigma_X} = \frac{kD(X)}{\sigma_X |k| \sigma_X} = \frac{k}{|k|} = \pm 1.$$

2. *Достаточность*. Пусть $|r_{XY}| = 1$, т. е. $r_{XY} = \pm 1$. Если, например, $r_{XY} = 1$, то с учётом (50)
$$0 \leq D(X' - Y') = 2(1 - r_{XY}) = 2(1 - 1) = 0,$$
так что $D(X' - Y') = 0$. Тогда, по свойству дисперсии $X' - Y' = a = \text{const}$, т. е.
$$\frac{X - M(X)}{\sigma_X} - \frac{Y - M(Y)}{\sigma_Y} = a \Rightarrow$$
$$\Rightarrow Y = \frac{\sigma_Y}{\sigma_X} X + \left(M(Y) - \frac{\sigma_Y}{\sigma_X} M(X) - \sigma_Y a \right).$$

Остаётся положить:
$$k = \frac{\sigma_Y}{\sigma_X}; \; b = \left(M(Y) - \frac{\sigma_Y}{\sigma_X} M(X) - \sigma_Y a \right). \blacksquare$$

6. ЗАКОН БОЛЬШИХ ЧИСЕЛ

6.1. Первое неравенство Чебышева

Теорема. *Если случайная величина X имеет математическое ожидание $M(X)$ и принимает только неотрицательные значения, то*

$$P(X \geq 1) \leq M(X). \tag{51}$$

Доказательство. Рассмотрим случай, когда X — непрерывная случайная величина с плотностью $p(x)$. Тогда, поскольку при $x \in (1, +\infty)$ выполняется неравенство $p(x) \leq xp(x)$:

$$P(X \geq 1) = \int_1^{+\infty} p(x)\,dx \leq \int_1^{+\infty} xp(x)\,dx \leq$$

(добавляем неотрицательное слагаемое — интеграл по отрезку $[0,1]$ от неотрицательной функции $xp(x)$)

$$\leq \int_1^{+\infty} xp(x)\,dx + \int_0^1 xp(x)\,dx =$$

(добавляем нулевое слагаемое — интеграл по промежутку $(-\infty, 0]$, на котором $p(x) = 0$ в силу условия на X)

$$= \int_1^{+\infty} xp(x)\,dx + \int_0^1 xp(x)\,dx + \int_{-\infty}^0 xp(x)\,dx =$$

$$= \int_{-\infty}^{+\infty} xp(x)\,dx = M(X). \blacksquare$$

6.2. Второе неравенство Чебышева

Теорема. *Если случайная величина X имеет математическое ожидание $M(X) = m$ и дисперсию $D(X) = d$, то для всякого $\varepsilon > 0$ выполняется неравенство*

$$P(|X - m| < \varepsilon) \geq 1 - \frac{d}{\varepsilon^2}. \tag{52}$$

Замечание. В левой части неравенства (52) стоит вероятность того, что реализованное значение случайной величины X отклонится от математического ожидания меньше чем на ε.

Доказательство. Имеет место равносильность неравенств

$$|X-m|<\varepsilon \Leftrightarrow \frac{|X-m|}{\varepsilon}<1 \Leftrightarrow \frac{(X-m)^2}{\varepsilon^2}<1.$$

Поэтому $$P(|X-m|<\varepsilon)=P\left(\frac{(X-m)^2}{\varepsilon^2}<1\right).$$

Поскольку события $\left(\frac{(X-m)^2}{\varepsilon^2}<1\right)$ и $\left(\frac{(X-m)^2}{\varepsilon^2}\geq 1\right)$ противоположны,

то $$P(|X-m|<\varepsilon)=1-P\left(\frac{(X-m)^2}{\varepsilon^2}\geq 1\right).$$

Применим первое неравенство Чебышева к неотрицательной случайной величине $\frac{(X-m)^2}{\varepsilon^2}$:

$$P\left(\frac{(X-m)^2}{\varepsilon^2}\geq 1\right)\leq M\left(\frac{(X-m)^2}{\varepsilon^2}\right)=\frac{1}{\varepsilon^2}M((X-m)^2)=\frac{d}{\varepsilon^2}$$

(последнее равенство — в силу определения дисперсии как математического ожидания квадрата отклонения). Теперь

$$P(|X-m|<\varepsilon)=1-P\left(\frac{(X-m)^2}{\varepsilon^2}\geq 1\right)\geq 1-\frac{d}{\varepsilon^2}. \blacksquare$$

6.3. Сходимость по вероятности

Определение. Число a называется *пределом по вероятности* последовательности случайных величин X_n, если для любого $\varepsilon>0$

$$\lim_{n\to\infty} P(|X_n-a|<\varepsilon)=1. \qquad (53)$$

Говорят также, что X_n *сходится по вероятности к числу* a.

Обозначение: $X_n \xrightarrow{(P)} a$.

Замечания

1. Равенство (53) означает, что с увеличением номера n вероятность события $(|X_n-a|<\varepsilon)$ неограниченно приближается к единице, и, значит, событие становится все более достоверным, происходит всё чаще.

2. Ввиду равносильности неравенств

$$|X_n-a|<\varepsilon \Leftrightarrow |(X_n-a)-0|<\varepsilon,$$

имеет место равносильность двух условий сходимости по вероятности

$$X_n \xrightarrow{(P)} a \Leftrightarrow (X_n-a)\xrightarrow{(P)} 0.$$

6.4. Общий закон больших чисел в форме Чебышева

Термином «закон больших чисел» объединяют круг теорем, утверждающих, что с вероятностью близкой к единице произойдёт некоторое случайное событие A, зависящее от неограниченно увеличивающегося числа других случайных событий, каждое из которых оказывает лишь незначительное влияние на A.

Теорема. *Пусть последовательность случайных величин* X_n *удовлетворяет трем условиям:*

1) случайные величины X_n *независимы (п. 4.6);*

2) они имеют математические ожидания $M(X_n)$ *и дисперсии* $D(X_n)$;

3) дисперсии ограничены в совокупности, т. е. при всех n $D(X_n) \le C = \text{const}$.

Тогда

$$\frac{X_1 + X_2 + ... + X_n}{n} - \frac{M(X_1) + M(X_2) + ... + M(X_n)}{n} \xrightarrow{(P)} 0 . \tag{54}$$

Замечание. Левая часть формулы (54) есть разность между средним арифметическим первых n случайных величин и средним арифметическим их математических ожиданий. Таким образом, теорема означает, что с увеличением числа слагаемых n среднее арифметическое реализованных значений практически всегда оказывается близким к среднему арифметическому математических ожиданий.

Доказательство. Положим $Y_n = \dfrac{X_1 + X_2 + ... + X_n}{n}$. По свойствам математического ожидания

$$M(Y_n) = M\left(\frac{X_1 + X_2 + ... + X_n}{n}\right) =$$

$$= \frac{1}{n}\big(M(X_1) + M(X_2) + ... + M(X_n)\big) .$$

Поэтому равенство (54), которое нужно доказать, принимает вид

$$Y_n - M(Y_n) \xrightarrow{(P)} 0 .$$

Зададим произвольное $\varepsilon > 0$. Применим к Y_n второе неравенство Чебышева

$$1 \ge P(|Y_n - M(Y_n)| < \varepsilon) \ge 1 - \frac{D(Y_n)}{\varepsilon^2} . \tag{*}$$

По свойствам дисперсии

$$0 \le D(Y_n) = D\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n} D(Y_i) \le$$

(пользуемся условием $D(Y_i) \le C$)

$$\le \frac{nC}{n^2} = \frac{C}{n} \underset{n\to\infty}{\to} 0 .$$

Тогда по принципу сжатой переменной, примененному к последовательности $D(Y_n)$, получаем

$$\lim_{n\to\infty} D(Y_n) = 0 \Rightarrow \lim_{n\to\infty}\left(1 - \frac{D(Y_n)}{\varepsilon^2}\right) = 1 .$$

Теперь, применяя тот же принцип сжатой переменной к последовательности $P(|Y_n - M(Y_n)| < \varepsilon)$, получаем из двойного неравенства (*), что

$$\lim_{n\to\infty} P(|Y_n - M(Y_n)| < \varepsilon) = 1 . \blacksquare$$

6.5. Частный закон больших чисел в форме Чебышева

Теорема. *Пусть случайные величины* X_n *независимы, и имеют одинаковые математические ожидания* $M(X_n) = m$ *и одинаковые дисперсии* $D(X_n) = d$ *. Тогда имеет место сходимость по вероятности*

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{(P)} m . \qquad (55)$$

Замечание. Условия теоремы выполняются, в частности, если независимые случайные величины имеют одинаковое распределение (одинаковую функцию распределения), и у них существуют математическое ожидание и дисперсия.

Доказательство. Выполняются все условия предыдущей теоремы; при этом

$$\frac{M(X_1) + M(X_2) + \ldots + M(X_n)}{n} = \frac{nm}{n} = m .$$

Согласно общему закону больших чисел, имеем

$$\frac{X_1 + X_2 + \ldots + X_n}{n} - m \xrightarrow{(P)} 0 \Rightarrow \frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{(P)} m . \blacksquare$$

Замечание. Формулы (54) и (55) являются математическим выражением давно установленного эмпирически факта устойчивости среднего арифметического большого числа независимых случайных величин — устойчивости, при которой существенные отклонения среднего арифметического реализованных значений от общего математического ожидания отдельных слагаемых являются редкими событиями. Это существенно дополняет свойство дисперсии среднего арифметического (п. 4.8).

Причиной такой устойчивости, как уже отмечалось в п. 1.8, является взаимное погашение отклонений разных знаков отдельных слагаемых при их суммировании в среднем арифметическом.

6.6. Закон больших чисел в форме Я. Бернулли

Теорема. *Пусть случайная величина* $X_n = w(A) = \dfrac{k(A)}{n}$ *— относительная частота события* A *в* n *независимых испытаниях по схеме Бернулли с вероятностью успеха* $P(A) = p$. *Тогда имеет место сходимость по вероятности*

$$w(A) \xrightarrow{(P)} p\,. \tag{56}$$

Замечание. Закон больших чисел в форме Бернулли является математическим выражением эмпирического закона больших чисел, в соответствии с которым при большом числе испытаний относительная частота $w(A)$ колеблется вблизи теоретической вероятности p (п. 3.5).

Доказательство. Введём вспомогательные случайные величины Z_k — индикатор k-го испытания (п. 4.9):

$Z_k = 0$, если в k-м испытании имела место неудача;

$Z_k = 1$, если в k-м испытании имел место успех.

Случайные величины Z_k независимы, поскольку связаны с исходами независимых испытаний, Они имеют одинаковый закон распределения $(0,1;\ \ p,q)$, одинаковые математические ожидания и одинаковые дисперсии: $M(Z_k) = p$, $D(Z_k) = pq$.

Случайная величина $k(A)$ (число успехов) есть сумма индикаторов: $k(A) = \sum\limits_{k=1}^{n} Z_k$ (в сумме справа столько единиц, сколько раз в n испытаниях имел место успех, а остальные слагаемые равны нулю). Относительная частота $w(A)$ есть среднее арифметическое индикаторов

$$w(A) = \frac{k(A)}{n} = \frac{\sum\limits_{k=1}^{n} Z_k}{n}\,.$$

Согласно частному закону больших чисел в форме Чебышева $w(A) \xrightarrow{P} p$. ■

6.7. Центральная предельная теорема

Эмпирически замечен факт: результат совместного влияния большого числа независимых случайных величин, каждая из которых в отдельности влияет на общую сумму лишь незначительно, приводит к нормальному распределению.

Термином «центральная предельная теорема» (ЦПТ) объединяют круг теорем, которые в различных формах описывают математически это стремление к нормальному закону. Приведём без доказательства формулировку одной из теорем круга ЦПТ.

Теорема (ЦПТ в форме Леви). *Пусть последовательность случайных величин* X_n *удовлетворяет трем условиям:*

1) все X_n *независимы;*

2) все X_n *имеют одинаковую функцию распределения: при всех* n *выполняется равенство* $F_{X_n}(x) = F(x)$*;*

3) все X_n *имеют математические ожидания и дисперсии:* $M(X_n) = m,\ \ D(X_n) = \sigma^2$ *(эти характеристики одинаковы для всех случайных величин в силу второго условия).*

Тогда для функций распределения F_{Y_n} *случайных величин* $Y_n = \dfrac{X_1 + X_2 + \ldots + X_n - nm}{\sigma\sqrt{n}}$ *при каждом* x *выполняется равенство* $\lim\limits_{n\to\infty} F_{Y_n}(x) = F_N(x)$*, где* $F_N(x) = \dfrac{1}{2} + \Phi(x)$ *— функция распределения нормального закона с параметрами* $a = 0$ *и* $\sigma = 1$ *(п. 5.6).*

Замечание. По свойствам математического ожидания и дисперсии случайные величины Y_n имеют те же характеристики, что и предельное распределение $F_N(x)$: $M(Y_n) = 0;\ \ D(Y_n) = 1$.

7. ДВУМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

7.1. Функция распределения двумерной случайной величины

Определение
Двумерной случайной величиной (случайным вектором, системой двух случайных величин) Z называется упорядоченная пара случайных величин X и Y: $Z=(X,Y)$. При этом случайные величины X и Y называются *составляющими* Z.

Аналогично определяется n-мерный случайный вектор

$$Z=(X_1, X_2, ..., X_n).$$

Геометрически двумерная случайная величина Z изображается случайной точкой плоскости со случайными координатами X и Y (рис. 21).

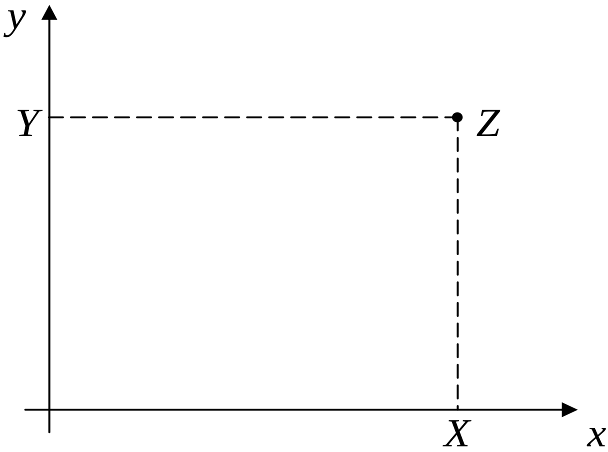

Рис. 21

Определение
Функцией распределения двумерной случайной величины $Z=(X,Y)$ называется функция двух переменных $F(x,y)$, задаваемая формулой

$$F(x,y)=P(X<x,Y<y).$$

$F(x,y)$ выражает вероятность попадания случайной точки (X,Y) в область D плоскости Oxy (рис. 22).

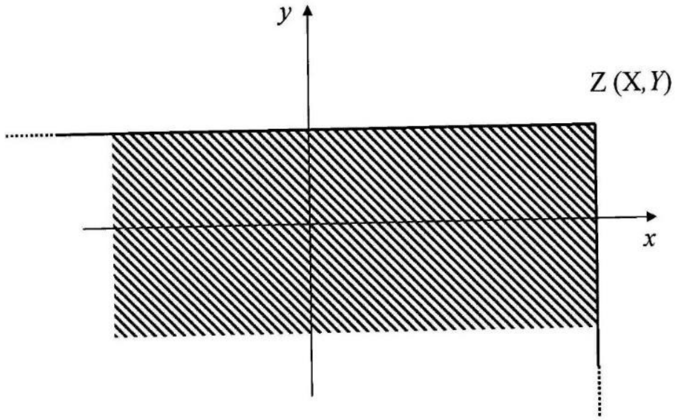

Рис. 22

Свойства функции распределения

1. Для любых x и y выполняется неравенство

$$0 \le F(x,y) \le 1.$$

Это следует из общих свойств вероятности случайного события (п. 3.2).

2. Функция $F(x,y)$ является неубывающей по каждому аргументу:

при $x_1 < x_2$: $F(x_1,y) \le F(x_2,y)$;

при $y_1 < y_2$: $F(x,y_1) \le F(x,y_2)$.

Доказательство. Пусть, например, $x_1 < x_2$.

Тогда $(X < x_2, Y < y) = (X < x_1, Y < y) + (x_1 \le X < x_2, Y < y)$ — сумма несовместных событий. Следовательно

$$P(X < x_2, Y < y) = P(X < x_1, Y < y) + P(x_1 \le X < x_2, Y < y), \qquad (57)$$

откуда $P(X < x_2, Y < y) \ge P(X < x_1, Y < y)$, так как последнее слагаемое в (57), будучи вероятностью случайного события, неотрицательно. ■

3. Поведение функции распределения на бесконечности:

$$F(-\infty, y) = \lim_{x \to -\infty} F(x,y) = 0;$$

$$F(x, -\infty) = \lim_{y \to -\infty} F(x,y) = 0;$$

$$F(-\infty, -\infty) = \lim_{x,y \to -\infty} F(x,y) = 0;$$

$$F(+\infty, +\infty) = \lim_{x,y \to +\infty} F(x,y) = 1$$

(без доказательства).

4. Если $F(x,y)$ – функция распределения двумерной случайной величины (X,Y), а $F_X(x)$, $F_y(y)$ — функции распределения составляющих X и Y, то $F(+\infty,y)=F_Y(y)$, $F(x,+\infty)=F_X(x)$ (без доказательства).

Определение
Функции распределения $F_X(x)$ и $F_Y(y)$ составляющих двумерной случайной величины называются *частными распределениями*.

5. Вероятность попадания случайной точки в полосу (рис. 23):

$$P(x_1 \le X < x_2, Y < y) = F(x_2, y) - F(x_1, y)\ ; \qquad (58)$$

$$P(x < X,\ y_1 \le Y < y_2) = F(x, y_2) - F(x, y_1)\ . \qquad (59)$$

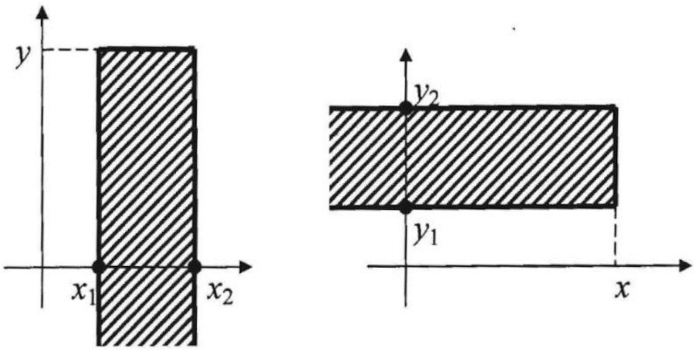

Рис. 23

Доказательство. Докажем, например, первое равенство. Событие $(X < x_2, Y < y) = (X < x_1, Y < y) + (x_1 \le X < x_2, Y < y)$ — сумма попарно несовместных событий. Переходя к вероятностям, получаем

$$P(X < x_2, Y < y) =$$
$$= P(X < x_1, Y < y) + P(x_1 \le X < x_2, Y < y) \Rightarrow$$
$$\Rightarrow F(x_2, y) = F(x_1, y) + P(x_1 \le X < x_2)\ ,$$

откуда следует нужное равенство. ■

6. Вероятность попадания случайной точки в прямоугольник (рис. 24)

$$P(x_1 \le X < x_2,\ y_1 \le Y < y) =$$
$$= [F(x_2, y_2) - F(x_1, y_2)] - [F(x_2, y_1) - F(x_1, y_1)]. \qquad (60)$$

Доказательство. Имеет место равенство случайных событий (рис. 25)
$(x_1 \le X < x_2, Y < y_2) = (x_1 \le X < x_2, Y < y_1) + (x_1 \le X < x_2,\ y_1 \le Y < y_2)$ —

сумма несовместных событий. Переходя к вероятностям, получаем

$$P(x_1 \le X < x_2, y_1 \le Y < y_2) =$$
$$P(x_1 \le X < x_2, Y < y_2) - P(x_1 \le X < x_2, Y < y_1).$$

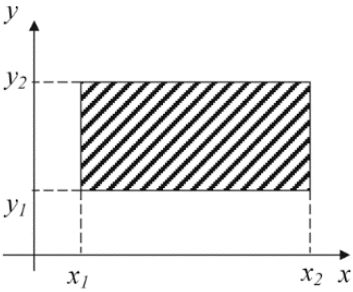

Рис. 24

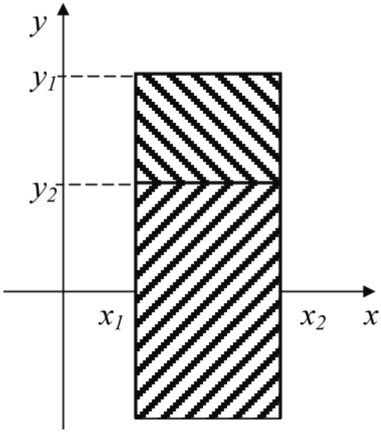

Рис. 25

Обе вероятности в правой части — это вероятности попадания в соответствующие полосы. Применяя к ним формулу (58) при $y = y_2$ и $y = y_1$ соответственно, получаем требуемое равенство (60). ■

7.2. Дискретные двумерные случайные величины

Определение
Двумерная случайная величина (X,Y) называется *дискретной*, если дискретными являются обе её составляющих X и Y.

Пусть законы распределения составляющих имеют вид $(x_i; p_i)$ и $(y_j; q_j)$. Обозначим через p_{ij} вероятность

$$p_{ij} = P\big((X = x_i)(Y = y_j)\big) = P\big((X,Y) = (x_i, y_j)\big).$$

Закон распределения двумерной случайной величины имеет вид $\big((x_i, y_j), p_{ij}\big)$, причем, как и в одномерном случае, выполняется равенство: $\sum\limits_{i,j} p_{ij} = P(\Omega) = 1$.

Если составляющие X и Y независимы, то по теореме умножения

$$p_{ij} = P\big((X = x_i)(Y = y_j)\big) = P(X = x_i)P(Y = y_j) = p_i q_j.$$

7.3. Непрерывные двумерные случайные величины

Определение

Двумерная случайная величина (X, Y) называется *непрерывной*, если существует неотрицательная интегрируемая в R^2 функция $p(x,y)$ такая, что функция распределения $F(x,y)$ представима в виде:

$$F(x,y) = \iint\limits_D p(u,v)\,dudv = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} p(u,v)\,dudv \qquad (61)$$

(область интегрирования D изображена на рис. 26).

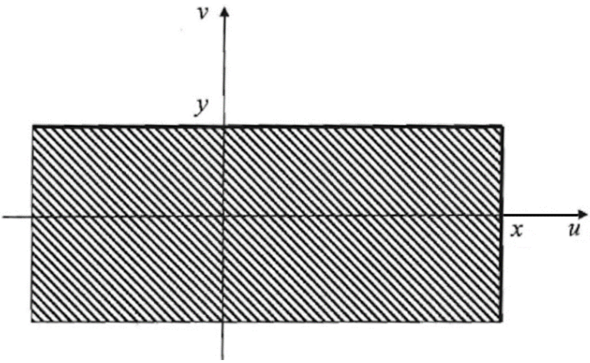

Рис. 26

Функция $p(x,y)$ называется *плотностью распределения*.

Замечание. Для двойного интеграла $\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} p(u,v)\,dudv$ с переменными верхними пределами справедлива теорема, аналогичная теореме о производной определённого интеграла с переменным верхним пределом [12]

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

Вероятностный смысл плотности

Пусть двумерная случайная величина (X,Y) непрерывна и имеет непрерывную плотность $p(x,y)$. Обозначим через P_{ABCD} вероятность попадания случайной точки $(X;Y)$ в прямоугольник $ABCD$ с площадью S_{ABCD} (рис. 27).

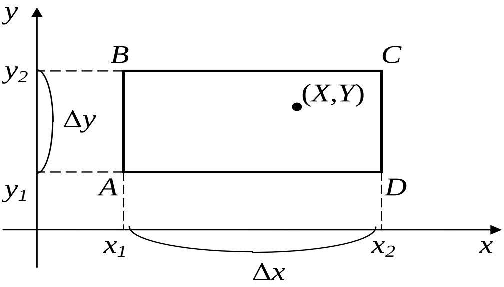

Рис. 27

По формулам (60) и (61)

$$P_{ABCD} = \left[\int\limits_{-\infty}^{x_2} \int\limits_{-\infty}^{y_2} p(x,y)\,dxdy - \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{y_2} p(x,y)\,dxdy \right] -$$

$$- \left[\int\limits_{-\infty}^{x_2} \int\limits_{-\infty}^{y_1} p(x,y)\,dxdy - \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{y_1} p(x,y)\,dxdy \right] =$$

(применим свойство аддитивности к внешнему интегралу)

$$= \int\limits_{x_1}^{x_2} \int\limits_{-\infty}^{y_2} p(x,y)\,dxdy - \int\limits_{x_1}^{x_2} \int\limits_{-\infty}^{y_1} p(x,y)\,dxdy =$$

$$= \int\limits_{x_1}^{x_2} \left[\int\limits_{-\infty}^{y_2} p(x,y)dy - \int\limits_{-\infty}^{y_1} p(x,y)dy \right] dx =$$

(применим свойство аддитивности к внутреннему интегралу)

$$= \int\limits_{x_1}^{x_2} \int\limits_{y_1}^{y_2} p(x,y)\,dxdy\,.$$

Итак,

$$P_{ABCD} = \int\limits_{x_1}^{x_2} \int\limits_{y_1}^{y_2} p(x,y)\,dxdy\,. \tag{62}$$

Вывод: *вероятность попадания непрерывно распределённой случайной точки в прямоугольник равна двойному интегралу от плотности по этому прямоугольнику*.

Следствие: Применяя к двойному интегралу в (62) теорему о среднем, получаем для вероятности попадания в прямоугольник *ABCD* со сторонами $\Delta x,\ \Delta y > 0$:

$$P_{ABCD} = p(\xi,\eta) S_{ABCD} = p(\xi,\eta)\Delta x \Delta y\,, \tag{63}$$

где (ξ,η) — некоторая точка прямоугольника.

Отсюда $p(\xi,\eta)=\frac{P_{ABCD}}{\Delta x\Delta y}$. Если теперь $\Delta x\to 0,\ \Delta y\to 0$, то точка (ξ,η) неограниченно приближается к точке (x,y), так что

$$p(x,y)=\lim_{\Delta x,\,\Delta y\to 0}\frac{P_{ABCD}}{S_{ABCD}}. \tag{64}$$

Итак, *плотность распределения непрерывной двумерной случайной величины при малых* $\Delta x,\ \Delta y>0$ *приближённо равна отношению вероятности попадания в малый прямоугольник к площади этого прямоугольника*.

7.4. Вероятность попадания случайной точки в заданную область

Теорема. *Пусть двумерная случайная величина (X, Y) непрерывна с плотностью* $p(x,y)$, *D — произвольная область плоскости Oxy. Тогда*

$$P\big((X,Y)\in D\big)=\iint_D p(x,y)\,dxdy. \tag{65}$$

Поясним теорему. Разобьем область D вертикальными и горизонтальными прямыми на непересекающиеся частичные области D_i (почти все они являются прямоугольниками).

По формуле (63) для этих прямоугольников

$$P\big((X,Y)\in D_i\big)=p(\xi_i,\eta_i)\Delta x_i\Delta y_i,$$

откуда, суммируя по всем прямоугольникам

$$\sum_i P\big((X,Y)\in D_i\big)=\sum_i p(\xi_i,\eta_i)\Delta x_i\Delta y_i.$$

В левой части последнего равенства — сумма вероятностей попарно несовместных событий. По аксиоме сложения, примененной «в обратном направлении», она равна вероятности суммы этих событий; эта сумма событий означает попадание в объединение частичных областей

$$P\left((X,Y)\in\bigcup_i D_i\right)=\sum_i p(\xi_i,\eta_i)\Delta x_i\Delta y_i.$$

Если ранг разбиения стремится к нулю, то объединение частичных прямоугольных областей стремится покрыть всю область D, а интегральная сумма в правой части стремится к соответствующему двойному интегралу. В пределе получаем

$$P\big((X,Y)\in D\big)=\iint_D p(x,y)\,dxdy.\ \blacksquare$$

7.5. Свойства плотности непрерывной двумерной случайной величины

1. $\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} p(x,y)\,dxdy = 1$.

Доказательство. Беря в формуле (65) в качестве области D всю плоскость Oxy, получаем

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} p(x,y)\,dxdy = P(\Omega) = 1. \blacksquare$$

2. Составляющие X и Y непрерывной случайной величины (X, Y) также являются непрерывными случайными величинами, и для их плотностей справедливы формулы:

$$P_X(x) = \int\limits_{-\infty}^{+\infty} p(x,y)\,dy;\ P_Y(y) = \int\limits_{-\infty}^{+\infty} p(x,y)\,dx.$$

Доказательство. Проведём доказательство для составляющей X. По свойству 4 функции распределения (п. 7.1) имеем

$$F_X(x) = F(x,+\infty) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{+\infty} p(x,y)\,dxdy = \int\limits_{-\infty}^{x}\left(\int\limits_{-\infty}^{+\infty} p(x,y)\,dy\right)dx = \int\limits_{-\infty}^{x} p_X(x)\,dx,$$

где $p_X(x) = \int\limits_{-\infty}^{+\infty} p(x,y)\,dy$, так что составляющая X непрерывна, и её плотностью является функция $p_X(x)$. $\blacksquare$

7.6. Условные законы распределения составляющих

I. Случай дискретной двумерной случайной величины

Пусть для дискретной двумерной случайной величины (X,Y) законы распределения составляющих X и Y имеют вид: $(x_i; p_i)$ и $(y_j; q_j)$. Обозначим, как и ранее, через p_{ij} вероятность

$$p_{ij} = P\big((X = x_i)(Y = y_j)\big) = P\big((X,Y) = (x_i, y_j)\big).$$

Зафиксируем событие $(Y = y_j)$. Если составляющие не предполагаются независимыми, то тогда зависимы события $A = (X = x_i)$ и $B = (Y = y_j)$, и можно говорить об условной вероятности

$$P_B(A) = P(X = x_i | Y = y_j) = \frac{P\big((X = x_i)(Y = y_j)\big)}{P(Y = y_j)} = \frac{p_{ij}}{q_j}. \qquad (66)$$

Определение

Условным законом распределения составляющей X при $Y = y_j$ называется таблица

$X \mid Y = y_j:$	…	x_i	…
$P:$	…	$\dfrac{p_{ij}}{q_j}$	…

Аналогично для фиксированного события ($X = x_i$) *условным законом распределения составляющей Y при* $X = x_i$ называется таблица

$Y \mid X = x_i:$	…	y_j	…
$P:$	…	$\dfrac{p_{ij}}{p_i}$	…

Замечание. Если составляющие X и Y независимы, то по теореме умножения:

$$p_{ij} = p_i q_j \Rightarrow \frac{p_{ij}}{q_j} = p_i; \quad \frac{p_{ij}}{p_i} = q_j,$$

и условные законы распределения составляющих совпадают с безусловными.

II. Случай непрерывной двумерной случайной величины

Пусть непрерывная двумерная случайная величина (X, Y) имеет плотность распределения $p(x, y)$, а плотности распределения составляющих – функции $p_X(x)$ и $p_Y(y)$ — не принимают нулевых значений. Обобщая формулу (66) на непрерывный случай, получаем следующее.

Определение

Условной плотностью распределения составляющей X при условии $Y = y$ называется функция $\varphi(x|y)$ аргумента x

$$\varphi(x|y) = \frac{p(x, y)}{p_Y(y)}.$$

Аналогично *условной плотностью распределения составляющей* Y при условии $X = x$ называется функция $\psi(y|x)$ аргумента y

$$\psi(y / x) = \frac{p(x, y)}{p_X(x)}.$$

7.7. Критерии независимости составляющих

Напомним, что случайные величины X и Y независимы, если для любых промежутков $<a,b>$ и $<c,d>$ являются независимыми события $(X \in <a,b>)$ и $(Y \in <c,d>)$, так что вероятность их произведения равна произведению соответствующих вероятностей

$$P\big((X \in <a,b>)(Y \in <c,d>)\big) = P(X \in <a,b>)P(Y \in <c,d>) .$$

Теорема. *Для того чтобы составляющие X и Y двумерной случайной величины (X,Y) были независимы, необходимо и достаточно, чтобы её функция распределения $F(x,y)$ была равна произведению функций распределения составляющих*:

$$F(x,y) = F_X(x)F_Y(y) \ . \tag{67}$$

Доказательство.

1. *Необходимость*. Пусть составляющие X и Y независимы. Тогда

$$F(x,y) = P((X<x)(Y<y)) = P(X<x)P(Y<y) = F_X(x)F_Y(y) .$$

2. Не приводя исчерпывающего доказательства, дадим пояснения, относящиеся к *достаточности*. Пусть $F(x,y) = F_X(x)F_Y(y)$. По определению функции распределения в двумерном и одномерном случае, это означает, что

$$P\big((X<x)(Y<y)\big) = P(X<x)P(Y<y) ,$$

т. е. события $(X<x)$ и $(Y<y)$ независимы. Можно доказать, что тогда и для любых промежутков $<a,b>$ и $<c,d>$ события $(X \in <a,b>)$ и $(Y \in <c,d>)$ также независимы. Последнее и означает независимость случайных величин X и Y . ■

Следствие. Для того чтобы составляющие X и Y непрерывной двумерной случайной величины (X,Y) были независимы, необходимо и достаточно, чтобы выполнялось равенство для плотностей

$$p(x,y) = p_X(x)p_Y(y) .$$

Доказательство.

1. *Необходимость*. Пусть составляющие X и Y независимы. Тогда справедливо равенство (41). Дифференцируя его дважды по x и по y , получаем

$$\frac{\partial^2 F}{\partial x \partial y} = F'_x(x)F'_y(y) \text{, или } p(x,y) = p_X(x)p_Y(y) .$$

2. *Достаточность*. Пусть $p(x,y)=p_X(x)p_Y(y)$. Интегрируя это равенство по множеству $D=\{(u,v)/-\infty<u<x;\ -\infty<v<y\}$ (см. рис. 26), получаем

$$\int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)\,dudv=\int_{-\infty}^{x}\int_{-\infty}^{y} p_X(u)p_Y(y)\,dxdy=$$

(выносим за знак внутреннего интеграла множители, не зависящие от y)

$$=\int_{-\infty}^{x}\left(p_X(u)\,du\cdot\int_{-\infty}^{y} p_Y(v)\,dv\right)=\int_{-\infty}^{x} p_X(u)\,du\cdot\int_{-\infty}^{y} p_Y(v)\,dv\ .$$

Применяя к левой части формулу (61), а к правой — определение функции распределения в одномерном случае (п.5.1), получаем

$$F(x,y)=F_X(x)F_Y(y)\ .\ \blacksquare$$

Библиографический список

1. Вентцель Е. С. Теория вероятностей: учебник для вузов / Е. С. Вентцель. — М.: Высш. шк., 2001. — 575 с.

2. Гмурман В. Е. Теория вероятностей и математическая статистика / В. Е. Гмурман. — М.: Высш. шк., 2002. — 478 с.

3. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман. — М.: Высш. шк., 1979. — 400 с.

4. Гнеденко Б. В. Элементарное введение в теорию вероятностей / Б. В. Гнеденко, А. Я. Хинчин. — М.: УРСС, 2003. — 205 с.

5. Гнеденко Б. В. Курс теории вероятностей / Б. В. Гнеденко. — М.: УРСС, 2001. — 318 с.

6. Данко П. Е. Высшая математика в упражнениях и задачах. Часть II / П. Е. Данко, А. Г. Попов, Т. Я. Кожевникова. — М.: Высш. шк., 1997. — 416 с.

7. Коваленко И. Н. Теория вероятностей и математическая статистика / И. Н. Коваленко, А. А. Филиппова. — М.: Высш. шк., 1982. — 256 с.

8. Шкадова А. Р. Теория вероятностей / А. Р. Шкадова, А. П. Нырков. — СПб.: Изд-во СПГУВК, 2003. — 198 с.

9. Ястребов М. Ю. Схема равновозможных исходов. СПб.: Изд-во СПГУВК, 1994. — 13 с.

10. Ястребов М. Ю. Производная и исследование функций / М. Ю. Ястребов. — СПб.: Изд-во СПГУВК, 2003. — 45 с.

11. Ястребов М. Ю. Введение в математическую логику / М. Ю. Ястребов. — СПб.: Изд-во СПГУВК, 2003. — 71 с.

12. Ястребов М. Ю. Математика. Неопределённый и определённый интегралы / М. Ю. Ястребов. — СПб.: Изд-во СПГУВК. 2004. — 55 с.

Оглавление

	Введение…………...	3
1.	КОМБИНАТОРИКА…………………………………………………	5
1.1.	Факториалы ………………………………………………………	5
1.2.	Принцип умножения………………………………........................	5
1.3.	Перестановки………………………………………………………	7
1.4.	Размещения…………………………………………………………	8
1.5.	Сочетания………………………………………………………...	9
2.	СЛУЧАЙНЫЕ СОБЫТИЯ………………………………………	11
2.1.	Классификация случайных событий……………………………	11
2.2.	Операции над событиями ………………………………………	12
2.3.	Свойства операций над событиями………………………………	15
2.4.	Относительная частота события………………………………..	16
3.	СВОЙСТВА ВЕРОЯТНОСТИ ……………...……………………	18
3.1.	Аксиоматическое введение вероятностей………………………..	18
3.2.	Непосредственные следствия из аксиом ………………………	18
3.3.	Схема равновозможных исходов ……………...............................	20
3.4.	Алгоритм реализации схемы равновозможных исходов……………..	22
3.5.	Эмпирический закон больших чисел……………………………..	24
3.6.	Условная вероятность………………………………………………..	24
3.7.	Теорема умножения …………………………………………………	26
3.8.	Независимость событий …………………………………………..	27
3.9.	Формула полной вероятности …………………….......................	29
3.10.	Формулы Байеса…………………………………………………	31
3.11.	Схема независимых испытаний Бернулли………………………..	33
3.12.	Локальная теорема Лапласа………………………………………..	35
3.13.	Интегральная теорема Лапласа ………………………………....	36
4.	ДИСКРЕТНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ…………………………	40
4.1.	Понятие случайной величины………………………………………	40
4.2.	Функция распределения случайной величины………………………..	40
4.3.	Закон распределения дискретной случайной величины………………	42
4.4.	Функция распределения дискретной случайной величины…………..	43
4.5.	Математическое ожидание дискретной случайной величины……….	45
4.6.	Свойства математического ожидания………………........................	46
4.7.	Дисперсия дискретной случайной величины……………………………	49
4.8.	Свойства дисперсии………………………………………………………..	51
4.9.	Биномиальное распределение………………………………………..	54
4.10.	Распределение Пуассона………………………………………………..	55
5.	НЕПРЕРЫВНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ……………………	57
5.1.	Плотность непрерывной случайной величины………………………..	57
5.2.	Особенность непрерывной случайной величины……………………..	58

5.3. Вероятностный смысл плотности распределения…………………….. 59
5.4. Математическое ожидание непрерывной случайной величины…….... 59
5.5. Дисперсия непрерывной случайной величины……………………... 61
5.6. Нормальное распределение…………………………………………... 62
5.7 Показательное распределение……………………………………….. 65
5.8 Равномерное распределение…………………………………………. 67
5.9 Преобразование случайных величин………………………………... 68
5.10 Вероятность попадания в промежуток для нормального распределения………………………………………………………… 70
5.11 Корреляция случайных величин……………………………………... 72

6. ЗАКОН БОЛЬШИХ ЧИСЕЛ…………………………………………. 76
6.1. Первое неравенство Чебышева……………………………………… 76
6.2. Второе неравенство Чебышева……………………………………… 76
6.3. Сходимость по вероятности ………... 77
6.4. Общий закон больших чисел в форме Чебышева…………………... 78
6.5. Частный закон больших чисел в форме Чебышева…………………. 79
6.6. Закон больших чисел в форме Я. Бернулли………………………… 80
6.7. Центральная предельная теорема…………………………………… 81

7. ДВУМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ………………………….. 82
7.1. Функция распределения двумерной случайной величины…………… 82
7.2. Дискретные двумерные случайные величины……………………….. 85
7.3. Непрерывные двумерные случайные величины……………………… 86
7.4. Вероятность попадания случайной точки в заданную область………. 88
7.5. Свойства плотности непрерывной двумерной случайной величины... 89
7.6. Условные законы распределения составляющих……………………... 90
7.7. Критерии независимости составляющих……………………………… 91

Библиографический список………..…………………........................... 93

Учебное издание

Ястребов Михаил Юрьевич, канд. экон. наук, доц.

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Учебное пособие

Издательство ГУМРФ им. адм. С. О. Макарова

198035, Санкт-Петербург, Межевой канал, 2
Тел.: (812) 748-97-19, 748-97-23
e-mail: izdat@gumrf.ru

Ответственный за выпуск	Сатикова Т. Ф.
Редактирование	Середова Т. В.
Компьютерная верстка	Тюленева Е. И.

Подписано в печать 14.05.2018
Формат 60×90/16. Бумага офсетная. Гарнитура Times New Roman
Усл. печ. л. 6,0. Тираж 100 экз. Заказ № 96/18