5. ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ И ИХ ПРИМЕНЕНИЕ

5.1. Основы комбинаторики.

Для вычисления вероятностей событий широко используются комбинаторные методы, среди которых особое место занимают основные правила комбинаторики — *правило умножения* и *сложения*. При решении задач, использующих понятие классической вероятности, используются следующие комбинаторные объекты: *перестановки, размещения* и *сочетания* без повторов и с повторами элементов.

Основное правило комбинаторики правило умножения формулируется следующим образом.

Если из некоторого конечного множества первый элемент можно выбрать n_1 *способами и после каждого такого выбора второй элемент можно выбрать* n_2 *способами, то оба элемента в указанном порядке можно выбрать* $N = n_1 \cdot n_2$ *способами.*

Данное правило распространяется на случай выбора трех и более элементов.

Пример 5.1. Сколько четырехзначных чисел можно составить из цифр 0, 1, 2, 3, 4, 5 если: а) ни одна цифра не повторяется; б) цифры могут повторяться; в) цифры могут повторяться, но числа должны быть нечетными.

Решение. Рассмотрим множество {0, 1, 2, 3, 4, 5}. В четырехзначном числе $mnpq$ первую цифру m из указанного множества можно выбрать пятью способами, так как 0 выбирать нельзя, т.е. $n_1 = 5$. После этого вторую цифру n можно выбрать, также пятью способами, т.е. $n_2 = 5$. После выбора первых двух цифр, для остальных двух цифр p и q остается, соответственно, 4 и 3 варианта выбора, т.е. $n_3 = 4$ и $n_4 = 3$. В силу комбинаторного правила умножения число четырехзначных чисел требуемого вида будет равно $N = n_1 \cdot n_2 \cdot n_3 \cdot n_4 = 5 \cdot 5 \cdot 4 \cdot 3 = 300$.

Варианты б) и в) предлагаются в качестве упражнений.

Комбинаторное правило суммы формулируется следующим образом**.**

Если некоторый элемент множества можно выбрать n_1 *способом, а другой элемент* n_2 *способами, причем первые и вторые способы взаимно исключают друг друга, то любо первый элемент либо второй можно выбрать* $n_1 + n_2$ *способами.*

Правило суммы применимо для выбора любого конечное число элементов.

Пример 5.2. В группе 15 юношей и 11 девушек. Сколькими способами можно выбрать старосту и его заместителя среди студентов одного пола?

Решение. По правилу умножения двух юношей можно выбрать $15 \cdot 14 = 210$ способами, а двух девушек $11 \cdot 10 = 110$ способами. По правилу суммы старосту и заместителя из двух студентов одного пола можно выбрать $110 + 210 = 320$ способами.

Перестановки элементов конечных множеств. *Перестановками* элементов конечного множество, состоящего из n элементов, называются комбинации, состоящие из всех элементов множества и отличающихся только порядком их расположения. Число P_n всех различных перестановок из n элементов равно произведению чисел от 1 до n (обозначение $n!$), т.е. равно

$$P_n = 1 \cdot 2 \cdot \ldots \cdot n = n!$$

Пример 5.3. Множество M состоит из трех элементов $M = \{a, b, c\}$. Выписать различные перестановок элементов этого множества и определить их число.

Решение. Выпишем вначале перестановки, в которых на первом месте стоит элемент a, затем — перестановки, первым элементом которых является b, далее — перестановки, у которых на первом месте стоит c. В результате получим все перестановки элементов этого множества:

$$a, b, c; \quad a, c, b; \quad b, a, c; \quad b, c, a; \quad c, a, b; \quad c, b, a.$$

Далее, используя формулу для подсчета числа перестановок при $n = 3$, получим $P_3 = 3! = 1 \cdot 2 \cdot 3 = 6$.

Пример 5.4. Сколькими способами могут выстроиться 5 покупателей в очередь в кассу?

Решение. Очевидно, очередь из пяти человек представляет собой перестановку из пяти элементов, следовательно, число способов выстраивания в очередь равно числу различных перестановок из пяти элементов, т.е. равно $P_5 = 5! = 120$.

Размещения. Пусть дано множество из n элементов. *Размещениями* из n элементов по m ($0 < m \le n$) называются комбинации, состоящие из m элементов и различающихся между собой либо элементами, либо порядком их расположения.

Число всевозможных размещений из n элементов по m обозначается A_n^m, и вычисляется по формуле

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

Используя факториалы это число можно записать в следующем виде:

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

Пример 5.5. Студентам нужно сдать 4 экзамена за 8 дней. Сколькими способами можно составить расписание сдачи экзаменов, если в один день можно сдавать только один экзамен?

Решение. Для сдачи всех экзаменов требуется 4 дня, которые можно выбирать из 8 дней. Таким образом, расписанию соответствует размещение из 8 по 4 и, следовательно, число способов составления расписаний равно числу таких размещений, т.е. $A_8^4 = 8 \cdot 7 \cdot 6 \cdot 5 = 1680$.

Пример 5.6. В соревновании участвует 10 команд. Сколькими способами могут распределиться первое, второе и третье место среди этих команд?

Решение. Так как имеются три призовых места, на которые претендуют 10 команд, то число способов распределение призовых мест определяется числом различных комбинаций из 10 по 3, т.е. числом размещений из 10 по 3. Следовательно это число равно $A_{10}^3 = 10 \cdot 9 \cdot 8 = 720$.

Сочетания. Комбинации из n элементов по m ($0 \le m \le n$) называются сочетаниями, если они различаются между собой хотя бы одним элементом. Порядок расположения элементов в сочетаниях не существенен. Число различных сочетаний обозначается C_n^m и вычисляется по формуле

$$C_n^m = \frac{A_n^m}{m!} = \frac{n!}{m! \cdot (n-m)!} = \frac{n \cdot (n-1) \cdot \cdots \cdot (n-m+1)}{1 \cdot 2 \cdot \cdots \cdot m}.$$

Сочетания образуют подмножество в множестве размещений и обладают следующими свойствами:

1) $C_n^m = C_n^{n-m}$,

2) $C_n^m = C_{n-1}^m + C_{n-1}^{m-1}$,

3) $C_n^0 + C_n^1 + C_n^2 + \dots + C_n^m = 2^n$.

Первое свойство используется для вычисления числа сочетаний, если $m > \frac{n}{2}$.

Пример 5.7. Из 22 студентов отбираются 5 студентов для участия в конференции. Сколькими способами можно сформировать группу?

Решение. Выбору из 22 студентов группы, состоящей из 5 студентов, соответствует сочетание, так как порядок расположения студентов в группе не существенен, следовательно, число способов формирований группы равно $C_{22}^5 = \frac{22 \cdot 21 \cdot 20 \cdot 19 \cdot 18}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 26334$.

Перестановки, размещения и сочетания с повторениями. Перестановки, размещения и сочетания, в которых допускается повтор элементов, называются перестановками, размещениями и сочетаниями с повторами.

Если в перестановке, содержащей n элементов, повторяются k различных элементов, при этом 1-й элемент повторяется n_1 раз, 2-ой элемент — n_2 раз, и т.д., k-ый элемент — n_k раз, то число $\overline{P}_n$ таких перестановок с повторениями равно

$$\overline{P}_n = \frac{n!}{n_1! \cdot n_2! \cdot \ldots n_k!},$$

где $n_1 + n_2 + \ldots + n_k = n$.

Формула для подсчета числа $\overline{A}_n^m$ размещений с повторениями из n по m элементов имеет вид

$$\overline{A}_n^m = n^m.$$

Число $\overline{C}_n^m$ сочетаний с повторами из n по m элементов вычисляется по формуле

$$\overline{C}_n^m = C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}.$$

Задачи для самостоятельного решения

1. Сколько различных способов расставить 5 книг на полке?

2. В соревновании участвуют 6 команд. Сколько существует способов распределения мест между ними?

3. В соревновании участвуют 6 команд. Сколькими способами могут быть распределены золотая, серебряная и бронзовая медали ?

4. Сколькими различными способами можно выбрать старосту и заместителя старосты из 20 студентов?

5. Сколько различных способов выбрать из 20 студентов двух для участия в конференции?

6. Сколько есть различных способов рассадить 5 студентов за 10 компьютеров, 5 студентов за 3 компьютера?

7. В вазе стоят 5 красных роз и 8 белых. Сколькими способами можно составить букет из а) 3 цветков; б) из 2 белых роз и одной красной?

5.2. Вероятности случайных событий

Теория вероятностей — раздел высшей математики, изучающий закономерности, которым подчиняются массовые случайные явления. Знание закономерностей, которым подчиняются массовые случайные явления, позволяет предвидеть, как эти явления будут развиваться во времени и принять правильное решение.

Методы теории вероятностей широко применяются в различных отраслях естествознания и техники, теории надежности, теории массового обслуживания, теории ошибок и наблюдений, а также в психологических и социологических исследованиях, психологическом прогнозировании, психодиагностике и т.д.

Особый интерес для психолога и социолога представляют статистические гипотезы, которые выдвигаются по поводу массовых случайных явлений, формулируются на теоретико-вероятностном языке и проверяются математико-статистическими методами. В психологии и социологии объекты изучения обладают всеми нормальными свойствами случайных явлений — случайных событий, случайных величин и функций.

Человек или социальная общность, психические и социологические свойства которых изучаются психологами и социологами формируется и функционирует в стохастической среде, влияние которой тоже подчиняется стохастическим закономерностям. Это приводит к тому, что психологические и социальные гипотезы оказываются в тоже время гипотезами статистическими, которые можно проверять с помощью теоретико-вероятностных и статистических методов.

При проведении некоторых экспериментов их результаты заранее точно предсказать невозможно. Такие эксперименты принято называть случайными. В совокупности всех результатов случайного эксперимента можно выделить взаимно исключающие (которые не могут наступить одновременно). Такие результаты называются *исходами* эксперимента или *элементарными событиями.* Разовое воспроизведение эксперимента называется *испытанием.*

Множество всех исходов (элементарных событий) случайного эксперимента называется *пространством элементарных событий* и обозначается Ω — прописной греческой буквой омега. Произвольное подмножество пространства элементарных событий Ω называется *случайным событием*.

Событие называется *достоверным*, если оно обязательно наступит в результате данного опыта. Событие называется *невозможным*, если оно заведомо не произойдет в результате испытаний, и обозначается $\varnothing$. Событие называется *случайным*, если оно может наступить или не наступить при разовом испытании.

Пример 5.1. Рассмотрим в качестве эксперимент вбрасывания игральной кости и определим следующие события:

A_i — выпало i-очков, $i=1,2,\ldots,6$; B — выпало 10 очков; D — выпало целое число очков; E — четное число очков.

В рассмотренном эксперименте A_i, E — случайные события, B — невозможное событие, а D — достоверное событие. Все перечисленные события взаимно исключающие. Пространством элементарных событий является множество $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Совместные и несовместные события. Два события называются *несовместными*, если при испытании появление одного из них исключает появление другого. В противном случае события называются совместными. События A_1, A_2, ..., A_n называются *попарно несовместными*, если любые два из них несовместны. События называются *равновозможными*, если каждое из них не является более возможным, чем другое.

Операции над событиями. В множестве случайных событий выполняются следующие операции сложения, умножения, вычисление их разности и противоположных событий.

Суммой $A + B$ двух событий A и B называется событие, состоящее в появлении хотя бы одного из них. Суммой n событий A_1, A_2, ... A_n называется событие, состоящее в появлении хотя бы одного из них. Сумма n событий, образующих полную группу, равна Ω.

Произведением AB двух событий называется событие, состоящее в их одновременном появлении. Аналогично определяется произведение n событий. Если события A и B несовместны, то $AB = \varnothing$.

Разностью $A - B$ событий A и B называется событие, при котором событие A происходит, а событие B не происходит.

Противоположным или *дополнительным* к событию A называется событие $\bar{A}$ при котором A не происходит.

Так как случайными событиями являются подмножества пространства элементарных событий, то операции сложения, умножения, вычислении разности и противоположных событий являются аналогами операций объединения, пересечения, вычисления разности и дополнения, выполняемыми над подмножествами. По этой причине свойства операций, выполняемых над случайными событиями, повторяют свойства их аналогов, выполняемых над множествами.

Свойства операций, выполняемых над случайными событиями.

1. $A + B = B + A$;
2. $A + A = A$;
3. $AB = BA$;
4. $AA = A$;
5. $A + (B + C) = (A + B) + C$;
6. $A(BC) = (AB)C$;
7. $A(B+C) = AB + AC$;
8. $(A + B)(A + C) = A + BC$;
9. $A + \Omega = \Omega$;
10. $A\Omega = A$;
11. $A + \bar{A} = \Omega$;
12. $A + \varnothing = A$;
13. $\cdot\varnothing A = \varnothing$;
14. $A + \cdot\, \bar{A} = \Omega$;

15. $B - A = B\bar{A}$.

Пример 5.8. Пусть заданы три произвольных события *A*, *B* и *C*, тогда:

а) ABC — произошли все три события;

б) $A\overline{B}\overline{C}$ — произошло только событие *A*;

в) $A \cdot \overline{B} \cdot \overline{C} + \overline{A} \cdot B \cdot \overline{C} + \overline{A} \cdot \overline{B} \cdot C$ — произошло только одно из этих событий;

г) $\overline{A} \cdot \overline{B} \cdot \overline{C}$ — ни одно из события *A*, *B* и *C* не произошло.

Классическое определение вероятности. Событие *A* называется благоприятствующим событию *B*, если наступление события *A* влечет за собой наступление события *B*.

Вероятностью события A при классическом определении называется отношение числа m элементарных событий, благоприятствующих событию *A*, к общему числу n всех элементарных событий, т.е.

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

Вероятности случайных событий обладают, следующими свойствами::

- вероятность невозможного события равна 0, т.е. $P(\varnothing) = 0$..
- вероятность достоверного события равна 1т.е. $P(\Omega) \ = \ 1$.
- Вероятность случайного события изменяется от 0 до 1, т.е. $0 \le P(A) \le 1$.

Пример 5.9. В урне находятся 11 белых и 5 красных шаров. Какова вероятность того, что наугад вытянутый шар будет красным?

Решение. Пусть *A* — событие, состоящее в том, что вынут красный шар. Число всех элементарных событий равно числу шаров $n = 11 + 5 = 16$. Число элементарных событий, благоприятствующих событию *A*, равно 5, т.е. $m = 5$. Следовательно, вероятность выемки красного шара равна $P(A) = \frac{m}{n} = \frac{5}{16}$.

Частота события. Статистическое определение вероятности. При классическом определении вероятности предполагалось, что все элементарные исходы равновозможные. Но во многих случаях трудно указать основания, позволяющие считать равновозможными все элементарные исходы. В связи с этим появилась необходимость введения ещё одного определения вероятности, называемой статистической.

Относительной частотой события, или *частостью* события, называется отношение числа испытаний, в которых появилось это событие, к общему числу всех произведенных испытаний.

Если обозначить относительную частоту события *A* через $W(A)$, то по определению

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

где m — число испытаний, в которых появлялось событие *A*, а n — число всех произведенных испытаний.

Относительная частота события обладает следующими свойствами:

- частота невозможного события равна 0;
- частота достоверного события равна 1;
- частота случайного события есть число, изменяющееся от 0 до 1;

• частота суммы двух несовместных событий равна сумме их частот.

Наблюдения позволили установить, что относительная частота случайных событий обладает свойствами статистической устойчивости, т.е. она принимает значения, достаточно близкие к вероятности этих событий. В силу этого относительную частоту, являющуюся объективной числовой характеристикой случайного события, считают его вероятностью.

Статистической вероятностью случайного события A называется число, относительно которого колеблется относительная частота события A при достаточно большом числе испытаний.

В случае статистического определения вероятность обладает теми же свойствами, что и частота.

Пример 5.10. Из 400 выбранных наудачу деталей оказалось 12 бракованных. Какова частота события A — выбора бракованной детали?

Решение. Так как в данном случае $m = 12$, $n = 400$, то по определению относительной частоты $W(A) = \frac{m}{n} = \frac{12}{400} = 0{,}03$.

Геометрическое определение вероятности. Классическое определение вероятности предполагает, что число элементарных исходов конечно. На практике встречаются испытания, для которых множество таких исходов бесконечно. В таких ситуациях используется понятие геометрической вероятности.

Рассмотрим на плоскости некоторую область Ω, имеющую площадь S_Ω, и внутри области Ω область D с площадью S_D. В области Ω случайно выбирается точка X. Этот выбор можно интерпретировать как вбрасывание точки X в область Ω. При этом попадание точки в область Ω — достоверное событие, в D — случайное событие (см. рис 5.1).

Предполагается, что все точки области Ω равноправны (все элементарные события равновозможные), т.е. что брошенная точка может попасть в любую точку области Ω и вероятность попасть в область D пропорциональна площади этой области и не зависит от её расположения и формы.

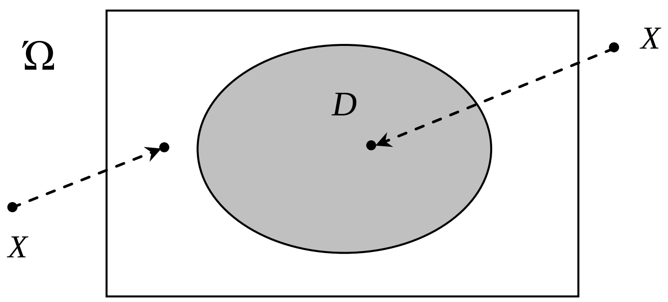

Рис. 5.1.

Пусть событие A — попадание точки X в область D. *Геометрической вероятностью* события A называется отношение площади области D к площади всей области Ω, т.е.

$$P(A) = \frac{S_D}{S_\Omega}.$$

Геометрическое определение вероятности события применимо и в случае, когда области Ω и D обе линейные или объемные. В указанных случаях геометрическая вероятность вычисляется по формулам

$$P(A)=\frac{l_D}{l_\Omega}, \qquad P(A)=\frac{V_D}{V_\Omega}.$$

где l — длина, а V — объем соответствующей области.

Геометрическая вероятность обладает всеми свойствами, присущими классическому определению:

- геометрическая вероятность невозможного события равна 0;
- геометрическая вероятность достоверного события равна 1;
- геометрическая вероятность случайного события есть число, заключенное в пределах от 0 до 1;
- геометрическая вероятность суммы двух несовместных событий равна сумме вероятностей этих событий.

В некоторых случая при проведении испытаний, случайные события удается связать с вбрасыванием случайным образом точки в некоторую часть геометрической области. В таких ситуациях вероятность случайного события полагается равной вероятности попадания точки в часть геометрической области и называется его геометрической вероятностью.

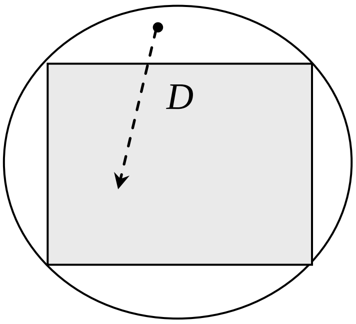

Рис. 5.2.

Пример 5.11. В круг вписан квадрат. В круг наудачу бросается точка. Какова вероятность того, что точка попадет в квадрат?

Решение. Пусть R — радиус круга, a — сторона вписанного квадрата, S — площадь круга, S_1 — площадь квадрата, тогда $S=\pi\cdot R^2$, $a=R\cdot\sqrt{2}$, $S_1=2\cdot R^2$. Пусть A — событие, состоящее в попадании точки в квадрат (см. рис. 5.2), тогда $P(A)=\frac{2\cdot R^2}{\pi\cdot R^2}=\frac{2}{\pi}=0{,}637$. Следовательно, Вероятность попадания точки в квадрат равна 0,637.

5.3. Теоремы сложения и умножения вероятностей.

При нахождении вероятностей сумм и произведений случайных событий используются следующие теоремы.

Теорема 5.1. *Вероятность суммы двух событий равна сумме вероятностей без вероятности их совместного появления, т.е.*

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

Из теоремы 5.1. вытекают следующие утверждения.

- *Если события А и В несовместны, то* $P(A+B) = P(A) + P(B)$.
- *Сумма вероятностей противоположных событий равна* 1, *т.е.* $P(A)+P(\overline{A})=1$.
- *Вероятность противоположного события равна* $P(\overline{A})=1-P(A)$.
- *Вероятность суммы* n *несовместных событий* $A_1, A_2, \ldots A_n$ *равна сумме вероятностей этих событий*

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

• *Если события A_1, A_2, . . . A_n попарно несовместны и образуют полную группу, то*

$$\sum_{i=1}^{n} P(A_i) = 1.$$

Пример 5.12. В урне 27 шаров, среди которых 7 красных, 9 белых и 11 синих. Извлекается один шар. Какова вероятность, что он цветной?

Решение. Цветные шары составляют красные и синие. Извлечение цветного шара означает появление либо красного, либо синего. Всего шаров + 9+ 11 = 27. Пусть событие *A* – извлечение красного шара, тогда *P(A)* = 7/27, событие В – извлечение синего шара, *P(B)* = 11 / 27. Следовательно, вероятность извлечения либо красного, либо синегоо шара равна

P(A+B) = P(A) + P(B) = 7 /27 + 11/ 27 = 18 / 27 = 2 / 3.

Условная вероятность. Независимые события. *Условной вероятностью* события *B* называется его вероятность при условии, что произошло событие *A*. Условная вероятность события *B* обозначается *P*(*B**A*) или $P_A(B)$ и вычисляется по формуле $P(B \setminus A) = \dfrac{P(AB)}{P(A)}$, $P(A) \neq 0.$

Два события *A* и *B* называются *независимыми*, если вероятность появления каждого из них не зависит от того, появилось другое событие или нет. В противном случае события называются *зависимыми.*

Если события *A* и *B* независимы, то *P(B\A) = P(B)* и *P(A \ B) = P(A)*. Вероятность *P(B)* в отличие от условной называется *безусловной вероятностью*.

Условные вероятности обладают всеми свойствами безусловных вероятностей Например, если $\overline{A}$ — противоположное событие к событию *A*, то

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

Пример 5.13. В фирму на реализацию поступило 16 мобильных телефонов, среди которых 7 марки Nokia. Какова вероятность, что второй проданный телефон был данной марки?

Решение. Пусть *A* – событие, состоящее в том, что второй проданный телефон марки Nokia. Вероятность события *A* зависит от того, какой был продан первый телефон. Пусть *B* – событие, состоящее в том, что первый проданный телефон был марки Nokia, *C* – событие, состоящее в том, что первый телефон был другой марки, тогда

$$P(A/B) = \frac{7-1}{16-1} = \frac{6}{15} = \frac{2}{5};\; P(A/C) = \frac{7}{16-1} = \frac{7}{15}.$$

Теорема 5.2. *Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность второго при условии, что первое событие произошло, т.е.*

$$P(AB) = P(A)\, P(B \setminus A).$$

Вероятность произведения двух независимых событий равна произведению их вероятностей, т.е.

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

События независимые в совокупности. События A_1, A_2, ... A_n называются *независимыми в совокупности*, если вероятность наступления любого из них не зависит от того, наступила или нет любая комбинация остальных.

Если события A_1, A_2, ... A_n независимы в совокупности, то вероятность их произведения равна произведению вероятностей этих событий.

$$P(A_1, A_2, ... A_n) = P(A_1)P(A_2) ... P(A_n).$$

Если события A_1, A_2, ... A_n независимы в совокупности то таковыми будут и противоположные им события $\bar{A}_1$, $\bar{A}_2$, … $\bar{A}_n$.

Вероятность наступления события $A = A_1 + A_2 + ... + A_n$, состоящего в появлении хотя бы одного из событий A_1, A_2, ... A_n, независимых в совокупности, находится по формуле:

$$P(A) = 1 - P(\bar{A}_1 \cdot \bar{A}_2 \cdot \ldots \cdot \bar{A}_n) = 1 - P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot \ldots \cdot P(\bar{A}_n)$$

Пример 5.14. Три стрелка попадают в мишень с вероятностями, соответственно, равными 0,85; 0,8 и 0,7. Найти вероятность того, что при одном выстреле хотя бы один из них попадет в мишень.

Решение. Пусть A – событие, состоящее в том, что в цель попал хотя бы один стрелок. Введем в рассмотрение следующие события:

- A_1 — попадание в цель первым стрелком,
- A_2 — попадание в цель вторым стрелком,
- A_3 — попадание в цель третьим стрелком.

По условию задачи $P(A_1) = 0{,}85$; $P(A_2) = 0{,}8$; $P(A_3) = 0{,}7$. Следовательно, вероятности противоположных событий $\bar{A}_1$ — промах первого стрелка, $\bar{A}_2$ — промах второго стрелка, $\bar{A}_3$ – промах третьего стрелка, соответственно, равны:

$$P(\bar{A}_1) = 1 - P(A_1) = 1 - 0{,}85 = 0{,}15; \quad P(\bar{A}_2) = 0{,}2; \quad P(\bar{A}_3) = 0{,}3.$$

Событие, противоположное событию A, состоит в том, что ни один стрелок не попал в цель, т.е. наступлении события $\bar{A}_1 \bar{A}_2 \bar{A}_3$. Таким образом, имеем:

$$P(\bar{A}_1 \bar{A}_2 \bar{A}_3) = P(\bar{A}_1) P(\bar{A}_2) P(\bar{A}_3) = 0{,}15 \cdot 0{,}2 \cdot 0{,}3 = 0{,}009,$$

откуда следует $P(A) = 1 - P(\bar{A}_1 \bar{A}_2 \bar{A}_3) = 1 - 0{,}009 = 0{,}991$.

Формула полной вероятности. Пусть событие A может осуществиться вместе с одним из событий H_1, H_2, ..., H_i, ..., H_n образующих полную группу. Так как заранее неизвестно, с каким из событий H_i может произойти событие A, то событие H_i называются *гипотезами.*

Пусть известны вероятности гипотез $P(H_1)$, $P(H_2)$, ..., $P(H_n)$ и условные вероятности $P(A \setminus H_1)$, $P(A \setminus H_2)$, ..., $P(A \setminus H_n)$ события A. Тогда вероятность события A вычисляется по формуле:

$$P(A) = P(H_1) \cdot P(A \setminus H_1) + P(H_2) \cdot P(A \setminus H_2) + \cdots + P(H_n) \cdot P(A \setminus H_n) =$$

$$= \sum_{i=1}^{n} P(H_i) \cdot P(A \setminus H_i)$$

Пример 5.15. Партия электрических лампочек на 20% изготовлена первым заводом, на 30% — вторым и на 50% — третьим. Вероятности выпуска бракованных лампочек заводами, соответственно, равны 0,01; 0,03 и 0,02. Найти вероятность того, что наудачу выбранная из партии лампочка окажется стандартной.

Решение. Пусть событие A – из партии взята стандартная лампочка. Выбранная стандартная лампочка может быть изготовлена 1-м (гипотеза H_1), 2-м (гипотеза H_2) или 3-м заводом (гипотеза H_3). По условию задачи вероятности гипотез равны:

$$P(H_1) = 20/100 = 0{,}2;\ P(H_2) = 30/100= 0{,}3;\ P(H_3) = 50/100 = 0{,}5,$$

а для условных вероятностей противоположного события $\overline{A}$ справедливы равенства:

$$P(\overline{A}/ H_1) = 0{,}01;\ \ \mathrm{P}(\overline{A}\ / H_2) = 0{,}03;\ \mathrm{P}(\overline{A}\ / H_3) = 0{,}02 = 0{,}98,$$

из которых и формулы $P(A \setminus H_i) = 1 - P(\overline{A} \setminus H_i)$ следует справедливость соотношений:

$P(A / H_1) = 1 – 0{,}01 = 0{,}99$; $P(A / H_2) = 1 – 0{,}03 = 0{,}97$; $P(A / H_3) = 1 – 0{,}02 = 0{,}98$.

Используя формулу полной вероятности, получаем:

$$P(A) = P(H_1)\cdot P(A\setminus H_1) + P(H_2)\cdot P(A\setminus H_2) + P(H_3)\cdot P(A\setminus H_3) =$$
$$= 0{,}2\cdot 0{,}99 + 0{,}3\cdot 0{,}97 + 0{,}5\cdot 0{,}98 = 0.979.$$

Формула Байеса. Если безусловные вероятности гипотез H_i до наступления события A были равны $P(H_i)$, $i = 1, ...n$, то после наступления событие A гипотезы приобретают условные вероятности $P(H_i \setminus A)$, которые не равны их первоначальным вероятностям. Условная вероятность $P(H_i \setminus A)$ i-й гипотезы вычисляется по *формуле Байеса*, которая имеет следующий вид:

$$P(H_i \setminus A) = \frac{P(H_i) \cdot P(A \setminus H_i)}{P(A)}.$$

Пример 5.16. Партия электрических лампочек на 20% изготовлена первым заводом, на 30% — вторым и на 50% — третьим. Вероятности выпуска бракованных лампочек заводами, соответственно, равны 0,01; 0,03 и 0,02. Наудачу взятая из партии лампочка оказалась стандартной. Найти вероятность того, что она изготовлена третьим заводом.

Решение. Введя, как и в решении предыдущего примера 5.15 гипотезы и используя вычисленную полную вероятность события A, с помощью формулы Байеса вычисляем условную вероятность третьей гипотезы:

$$P(H_3 \setminus A) = \frac{P(H_3) \cdot P(A \setminus H_3)}{P(A)} = \frac{0{,}5 \cdot 0{,}98}{0{,}979} = 0{,}5005.$$

Задачи для самостоятельного решения

1. При каждом выстреле, независимо от остальных выстрелов, стрелок попадает в цель с вероятностью 0,8. Ему разрешено стрелять до первого промаха. Какова вероятность, что он произведет ровно три выстрела.

2. В урне 6 синих и 4 красных шара. Из нее подряд извлекают два шара. Какова вероятность, что оба шара синие?

5.4. Повторные независимые испытания.

Схема *повторных испытаний* заключается в проведении серии испытаний при воспроизведении одних и тех же условий, исходы которых принадлежат заданной совокупности случайных событий. Повторные испытания называются *независимыми*, если вероятность наступления любого события при очередном испытании не зависит от результатов предшествующих ему испытаний.

Схема испытаний Бернулли. Простейшей схемой повторных независимых испытаний является *схема испытаний Бернулли.* В данной схеме испытаний возможными являются два исхода: некоторое событие A — «успех» и противоположное ему событие $\overline{A}$ — «неуспех». При этом вероятность успеха, (следовательно, и неуспеха) постоянна и равна p, т.е. $P(A)=p$ и, соответственно, $P(\overline{A})=1-P(A)=q$.

Основными задачами при проведении повторных независимых испытаний являются: а) вычисление вероятности реализации m успехов при проведении n испытаний; б) вычисление вероятности того, что при проведении n испытаний событие A наступит не мене m_1 и не более m_2 раз, указанная вероятность обозначается $P_n(m_1,m_2)$ и равна вероятности реализации неравенства $m_1 \le m \le m_2$, т.е.

$$P_n(m_1,m_2)=P(m_1 \le m \le m_2)$$

Теорема 5.3. (Бернулли) *Вероятность $P_n(m)$ того, что событие A наступит ровно m раз при проведении n испытаний равна*

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m}.$$

где C_n^m — число сочетаний из n элементов по m.

Из теоремы Бернулли и теорем сложения вероятностей вытекает

Следствие 5.1. *Пусть m_1, m_2 — такие целые числа что $0 \le m_1 \le m_2 \le n$, тогда вероятность того, что событие A наступит:*

а) не менее m_1 раз и не более m_2 раз при n испытаниях равна

$$P_n(m_1,m_2)=\sum_{m=m_1}^{m_2} C_n^m \cdot p^m \cdot q^{n-m};$$

б) менее m_2 раз равна $P_n(0 \le m < m_2)=P_n(0)P_n(0)+P_n(1)+...+P_n(m_2-1)$*;*

в) не менее m_1 раз равна $P_n(m_1 \le m < n)=P_n(m_1)+P_n(m_1+1)+...+P_n(n)$.

Пример 5.17. Вероятность, что деталь стандартная равна 0,7. Найти вероятность, что из 7 проверенных деталей ровно 4 детали стандартные.

Решение. Введем событие A — выбрана стандартная деталь. По условию задачи $n=7$, $m=4$, $P(A)=p=0{,}7$, и следовательно $P(\overline{A})=q=1-0{,}7-0{,}3$. Используя формулу Бернулли, вычисляем

$$P_7(4)=C_7^4 \cdot p^4 \cdot q^{7-4}=\frac{7\cdot 6\cdot 5\cdot 4}{1\cdot 2\cdot 3\cdot 4}\cdot(0{,}7)^4\cdot(0{,}3)^3=35\cdot 0{,}2401\cdot 0{,}027\approx 0{,}227.$$

Пример 5.18. Поочередно пять раз подбрасывается монета. Найти вероятность, что при испытаниях выпало более одного герба.

Решение. Введем событие B — при испытаниях выпало более одного герба, и примем за успех событие A — выпал «герб». Тогда число успехов m может принимать значения 0, 1, 2, 3, 4, 5, так как общее число испытаний n = 5. По условию задачи $m>1$, т.е. m может принимать значения $2,3,4,5$, следовательно,

$$P(B) = P(2 \le m \le 5) = P(m=2) + P(m=3) + P(m=4) + P(m=5).$$

Противоположным событием к событию B является событие $\overline{B}$ — при испытаниях выпало менее двух гербов. Следовательно, вычислив в начале вероятность противоположного события (оно вычисляется проще) и, воспользовавшись формулой $P(B) = 1 - P(\overline{B})$, найдем требуемую вероятность.

При реализации события $\overline{B}$ число успехов $m = 0, 1$, следовательно

$$P(\overline{B}) = P(m=0) + P(m=1) = P_5(0) + P_5(1) = C_5^0 \cdot p^0 \cdot q^5 + C_5^1 \cdot p^1 \cdot q^4$$

$$= \left(\frac{1}{2}\right)^5 + 5 \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right)^4 = \frac{1}{32} + \frac{5}{32},$$

так как $p = q = \frac{1}{2}$. Вычислив далее $P(B) = 1 - P(\overline{B}) = 1 - \frac{6}{32} = \frac{13}{16}$, находим искомую вероятность.

Наивероятнейшее число успехов в схеме испытаний Бернулли. Число m_0, которому при заданном n соответствует максимальная вероятность $P_n(m_0)$, называется *наивероятнейшим* числом успеха. Если $p \neq 0$ и $p \neq 1$, то число m_0 можно вычислить, используя двойное неравенство

$$np - q \le m_0 \le np + p,$$

которое равносильно двойному неравенству

$$n(p + 1) - 1 \le m_0 \le n\,(p + 1).$$

При использовании второго неравенства не требуется вычисления $q = 1 - p$.

Формула Пуассона. Формула Бернулли при большом числе испытаний становится неприемлемой ввиду трудности вычисления числа сочетаний и степеней вероятностей p и q. В этом случае используется формула Пуассона, позволяющая приближенно вычислять вероятность $P_n(m)$. Возможность использования формулы Пуассона устанавливает

Теорема 5.4. *Если вероятность успеха* p *при каждом испытании постоянна и достаточно мала, число независимых испытаний* n *достаточно велико и произведение* np *постоянная величина, т.е.* $np = \lambda$, *то вероятность* $P_n(m)$ *наступления* m *успехов при* n *испытаниях приблизительно равна*

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

Функция двух переменных $f(\lambda, m) = \frac{\lambda^m \cdot e^{-\lambda}}{m!}$ называется функцией Пуассона. Эта функция используется для приближенного вычисления вероятности $P_n(m)$ при $p < 0{,}01$, $n > 100$ и $np \le 10$.

Пример 5.19. Завод отправил на базу 500 изделий, вероятность повреждения изделия в пути равна 0,002. Найти вероятность того, что в пути будет повреждено 2 изделия.

Решение. По условию задачи $n = 500$, $p = 0{,}002$, $\lambda = np = 500 \cdot 0{,}002 = 1 < 10$. Применяя формулу Пуассона вычисляем требуемую вероятность

$$P_n(m) = P_{500}(2) = \frac{1^2 \cdot e^{-1}}{2!} = \frac{0{,}368}{2} = 0{,}184.$$

Локальная теорема Муавра-Лапласа. Локальная теорема Муавра-Лапласа вводит еще одну формулу для приближенного вычисления вероятности $P_n(m)$.

Теорема 5.5. *Если вероятность успеха p при каждом испытании постоянна и отлична от нуля и единицы, а число испытаний достаточно велико, то вероятность $P_n(m)$ наступления m успехов при n испытаниях приближенно равна (и тем точнее, чем большее n) значению функции*

$$\frac{1}{\sqrt{npq}}\cdot\varphi(x) \;\; при \;\; x=\frac{m-np}{\sqrt{npq}}, \;\; где \;\; \varphi(x) = \frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}.$$

Функция $y=\varphi(x)$ называется *«малой» функцией Лапласа*, а также *функцией Гаусса*. Таким образом, приближенное значение вероятности наступления m успехов при достаточно большом числе испытаний n испытаниях можно вычислять по формуле

$$P_n(m)\approx\frac{1}{\sqrt{npq}}\cdot\varphi(x), \quad \text{где } x=\frac{m-np}{\sqrt{npq}}. \tag{5.1}$$

Значения функции Гаусса $\varphi(x)$ приведены в специальных таблицах. При вычислении значений этой функции используются ее следующие свойства:

а) функция $\varphi(x)$ четная, т.е. $\varphi(x)=\varphi(-x)$;

б) функция $\varphi(x)$ монотонно убывает, при этом $\varphi(x)\to 0$ при $x\to\pm\infty$;

в) практически, при $x > 4$ или $x < -4$ можно считать, что $\varphi(x)=0$.

Пример 5.20. Вероятность того, что сошедшее с конвейера изделие первого сорта равна 0,9. Какова вероятность того, что из 400 сошедших с конвейера изделий 356 окажутся первого сорта?

Решение. По условию задачи $n = 400$, $m = 356$, $p = 0,9$ и $q = 1–0,9 = 0,1$. Вычислив $npq = 400\cdot0,9\cdot0,1 = 36$, находим значение

$$x=\frac{m-np}{\sqrt{npq}}=\frac{356-400\cdot0,9}{\sqrt{36}}=-0,67.$$

Найдя по таблице значение функции Гаусса $\varphi(-0,67)=\varphi(0,67)=0,3188$ и подставив его и $\sqrt{npq}=\sqrt{36}=6$ в формулу (5.1) получим значение вероятности

$$P_{400}(356)=\frac{1}{\sqrt{npq}}\cdot\varphi(x)=\frac{0,3188}{6}=0,0531.$$

Интегральная теорема Муавра-Лапласа. В данной теореме приводится формула для приближенного вычисления вероятности того, что при n независимых испытаний число успехов m не менее m_1 и не более m_2, т.е. изменяется в пределах от m_1 до m_2.

Теорема 5.6. *Если вероятность p наступления успеха при каждом из n независимых испытаний постоянна и отлична от нуля и единицы, а число испытаний достаточно велико, то вероятность того, что при n испытаниях число успехов m не менее m_1 и не более m_2 приближенно равна значению определенного интеграла*

$$\frac{1}{\sqrt{2\pi}}\int_{x_1}^{x_2} e^{-\frac{x^2}{2}}dx, \text{ где } x_1=\frac{m_1-np}{\sqrt{npq}},\ x_2=\frac{m_2-np}{\sqrt{npq}}.$$

Если ввести в рассмотрение интеграл

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt \tag{5.2}$$

с переменным верхним пределом интегрирования, который является первообразной для малой функции Лапласа

$$y=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$$

то получим следующую формулу для вычисления вероятности

$$P_n(m_1,m_2)=\Phi\left(\frac{m_2-np}{\sqrt{npq}}\right)-\Phi\left(\frac{m_1-np}{\sqrt{npq}}\right). \tag{5.2}$$

Функция (5.2.) называется «*большой*» (или просто) *функцией Лапласа*. Значения этой функции приведены в таблице. При использовании таблицы следует использовать следующие свойства функции $\Phi(x)$:

а) функция $\Phi(x)$ нечетная, т.е. $\Phi(-x)=-\Phi(x)$;

б) предел функции $\Phi(x)$ при $x\to+\infty$ равен 0,5;

в) для значений $x>5$ полагается $\Phi(x)=0{,}5$ и $\Phi(x)=-0{,}5$ при $x<-5$.

С учетом свойств а) — в) в таблице приведены значения функции Лапласа $\Phi(x)$ для $0\le x\le 5$.

Замечание 5.1. В современных учебниках и методических пособиях функция Лапласа полагается равной $\Phi(x)=\frac{2}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$. В силу этого формула (5.2) приобретает следующий вид:

$$P_n(m_1,m_2)=\frac{1}{2}\left(\Phi\left(\frac{m_2-np}{\sqrt{npq}}\right)-\Phi\left(\frac{m_1-np}{\sqrt{npq}}\right)\right).$$

В таблице значений такой функции Лапласа приводятся значения для $0\le x\le 4$, так как практически при $x>4$, значение указанной функции Лапласа равно 1.

Пример 5.21. Вероятность появления события A в каждом из 900 независимых испытаний равна 0,8. Найти вероятность того, что событие A произойдет не менее 710 и не более 740 раз.

Решение. По условию задачи n = 900, m_1 = 710, m_2 = 740, p = 0,8. Вычислив q = 0,2, npq = 900·0,8·0,2 = 12, находим:

$$x_1=\frac{k_1-np}{\sqrt{npq}}=\frac{710-900\cdot 0{,}8}{12}=-0{,}83,\quad x_2=\frac{k_2-np}{\sqrt{npq}}=\frac{740-900\cdot 0{,}8}{12}=1{,}67.$$

Учитывая нечетность находим по таблице значения функции Лапласа:

$\Phi(x_1)=\Phi(-0{,}83)=-\Phi(0{,}83)=-0{,}2967$; $\Phi(x_2)=\Phi(1{,}67)=0{,}4525$

и вычисляем искомую вероятность**:**

$$P_{900}(710, 740) = \Phi(x_2) - \Phi(x_1) = \Phi(1{,}67) - \Phi(-0{,}83) = 0{,}4525 - (-0{,}2967) = 0{,}7492.$$

Задачи для самостоятельного решения

1. Вероятность появления события в каждом из независимых испытаний равна 0,6. Найти вероятность того, что в 15 испытаниях событие наступит ровно 2 раза.

2.Вероятность выигрыша по лотерейному билету равна 0,04. Какова вероятность, что из 12 билетов окажутся 3 выигрышных?

3. Вероятность попадания в мишень при одном выстреле равна 0,7. Найти вероятность, что из 10 выстрелов будет не более 4 попаданий.

4. Вероятность появления события в каждом из независимых испытаний равна 0,8. Найти вероятность того, что в 11 испытаниях событие наступит хотя бы 2 раза.

5. Вероятность появления события в каждом из независимых испытаний равна 0,004. Найти вероятность того, что в 250 испытаниях событие наступит 3 раза.

6. Вероятность опоздания на занятие равна 0, 01. найти вероятность, что из 600 студентов опоздает от 7 до 9 студентов.

7. Телефонная станция обслуживает 400 абонентов. Вероятность ошибки связи за время t равна 0, 02. Найти вероятность, что за время t произойдет менее 3 ошибок.

8. При введении вакцины иммунитет создается с вероятностью 0, 9996. Какова вероятность, что из 25000 детей заболеет более двух? (не более двух).

9. В дисплейном классе имеется 16 компьютеров первого типа и 24 второго типа. Вероятность того, что за время работы на компьютере первого типа не произойдет сбой, равна 0,8, а на компьютере второго типа – 0,7. Найти вероятность, что за время работы на случайно выбранном компьютере сбоя не произошло.

5.5. Дискретные случайные величины и их числовые характеристики

Случайной называется переменная величина, которая при испытаниях принимает значения из заданного числового множества, но какие именно заранее предугадать невозможно.

Примерами случайных величин могут служить число очков при бросании кубика, число выстрелов до первого попадания в цель, курс доллара, расстояние, которое пролетит снаряд при выстреле, время безотказной работы какого-нибудь устройства и т.д.

Практически все количественные характеристики факторов, обуславливающих и влияющих на поведение индивидов социальных систем, функционирование самих систем и отдельных их членов носят вероятностный характер, что является предпосылкой для применения методов теории вероятности и математической статистики при их изучении.

Множество значений, которые может принимать случайная величина может быть конечным, например значения 1, 2, 3, 4, 5, 6 при бросании кубика, так и бесконечным. Например, при стрельбе до первого попадания в цель число выстрелов может совпадать с любым натуральным числом.

Случайная величина, значениями которой являются члены конечной или бесконечной последовательности, называется *дискретной случайной величиной.* В этой ситуации случайная величина принимает отдельные, изолированные возможные значения, которые отделены одно от другого промежутками, в которых нет других возможных значений случайной величины.

Случайные величины обозначаются прописными X, Y, Z, …, а их значения строчными латинскими буквами x, y, z … .

Если X — дискретная случайная величина, то событие $X = x_i$, означает, что случайная величина X принимает значение x_i. Вероятность, с которой это значение принимается, обозначается $P(X = x_i) = p_i$.

Набор пар (x_i, p_i), $i = 1,2,\ldots,n,\ldots$ называется *законом распределения* случайной величины X.

Если случайная величина X конечная, то закон распределения можно задать:

а) таблично (см. таблицу 5.1), упорядочив значения случайной величины по возрастанию;

Таблица 5.1.

X	x_1	x_2	...	x_i	…	x_n
P	p_1	p_2	...	p_i	…	p_n

б) аналитически в виде формулы

$$p_i = f(x_i);$$

в) графически, поставив в соответствие парам (x_i, p_i) точки плоскости $M_i(x_i, p_i)$, $i = 1,2,\ldots,n$, (см. рис. 5.3.)

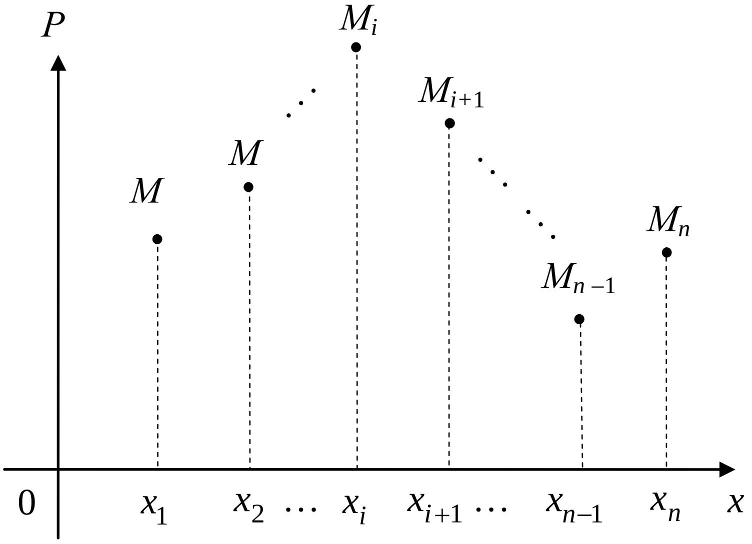

Рис. 5.3.

Так как события $X = x_i$, $i = 1,2,\ldots,n$, несовместны и образуют полную группу, то сумма их вероятностей равна единице, т.е.

$$p_1 + p_2 + \ldots + p_i + \ldots + p_n = \sum_{i=1}^{n} p_i = 1.$$

Если случайная величина бесконечная, то сумма всех вероятностей также равна единице, т.е. справедливо равенство

$$\sum_{i=1}^{+\infty} p_i = 1.$$

Операции над дискретными случайными величинами. Так как случайные величины принимают числовые значения, то их можно складывать, умножать на постоянную величину, находить их разность, произведение и степени.

Произведением дискретной случайной величины X и постоянной величины C называется дискретная случайная величина, обозначаемая CX, возможными значениями которой являются возможные значения X, умноженные на C. Вероятности возможных значений CX равны вероятностям соответствующих им значений X, т.е.

Таблица 5.2.

CX	Cx_1	Cx_2	...	Cx_i	…	Cx_n
P	p_1	p_2	...	p_i	…	p_n

$P(CX = Cx_i) = P(X = x_i) = p_i$. Таким образом, закон распределения CX определяется таблицей 5.2.

Аналогичным образом определяются дискретные случайные величины, являющиеся суммой или разностью случайной величины X и постоянной величины C, которые обозначаются, соответственно, $X \pm C$. Законы распределения указанных величин определяются таблицей 5.3.

Таблица 5.3.

$X \pm C$	$x_1 \pm C$	$x_2 \pm C$	...	$x_i \pm C$	…	$x_n \pm C$
P	p_1	p_2	...	p_i	…	p_n

Пусть Y — дискретная случайная величина, заданная законом распределения, представленным таблицей 5.4, вторая строка которой содержит вероятности возможных значений Y, т.е. $q_j = P(Y = y_j)$, j = 1, 2, …. m. *Суммой* или *разностью* случайных величин X и Y называется случайная величина, обозначаемая соответственно $X \pm Y$, и принимающая возможные значения $x_i \pm y_j$, с вероятностью

Таблица 5.4.

Y	y_1	y_2	...	y_j	…	y_m
P	q_1	q_2	...	q_j	…	q_m

$$P((X = x_i)(Y = y_j)) = p_{ij}, \quad i = 1, 2, \ldots. n, \quad j = 1, 2, \ldots. m, \qquad (5.3)$$

где $(X = x_i)(Y = y_j)$ — произведение событий $X = x_i$ и $Y = y_j$.

Если при вычислении $x_i \pm y_j$ некоторые из них окажутся равными числу z, то вероятность этого возможного значения случайной величины $X \pm Y$, в силу теоремы сложения вероятностей, будет равна сумме вероятностей всех возможных значений $x_i \pm y_j$, равных z. Если X и Y независимые случайные величины, то в силу теоремы произведения вероятностей,

$$P((X = x_i)(Y = y_j)) = p_i q_j, \quad i = 1, 2, \ldots. n, \quad j = 1, 2, \ldots. m. \qquad (5.4)$$

Произведением случайных величин X и Y называется случайная величина, обозначаемая XY. Возможными значениями XY являются произведения

$$x_i y_j, \; i = 1, 2, \ldots, n, \; j =, 2, \ldots, m,$$

вероятности которых, как и для суммы случайных величин, определяются соотношениями (5.3). Если при вычислении произведений $x_i \cdot y_j$ некоторые из них окажутся равными числу z, то вероятность этого возможного значения случайной величины XY также равна сумме вероятностей всех возможных значений $x_i \cdot y_j$, равных z. Если X и Y независимые случайные величины, то вероятность произведения событий $X = x_i$ и $Y = y_j$ определяется формулой (5.4).

Таблица 5.5.

X^k	x_1^k	x_2^k	...	x_i^k	…	x_n^k
P	p_1	p_2	...	p_i	…	p_n

Степенью X^k случайной величины X, где k — целое число большее единицы, называется произведение, содержащее k множителей, каждый из которых — случайная величина X, т.е. $X^k = \underbrace{X \cdot X \ldots X \cdot X}_{k \text{ раз}}$. Возможными значениями степени X^k являются степени x_i^k, $i = 1, 2, \ldots n$, вероятности которых равны вероятностям соответствующих возможных значений X. Следовательно, закон распределения X^k определяется таблицей 5.5.

Пример 5.22. Найти закон распределения произведения независимых случайных величин X и Y, заданных следующими законами распределения:

	x_1	x_2	x_3
X	0	1	2
P	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

	y_1	y_2
Y	1	2
P	$\frac{1}{3}$	$\frac{2}{3}$

Решение. Вычисляем всевозможные произведения $x_i \cdot y_j$, $i = 1, 2, 3$, $j = 1, 2$. и записываем их во вторую строку приведенной ниже таблицы. Затем по формуле (5.4) находим вероятности всех произведений и записываем в третью строку таблицы. Вторая строка показывает, что имеется две пары произведений $x_1 \cdot y_1$, $x_1 \cdot y_2$ и $x_2 \cdot y_2$, $x_3 \cdot y_1$, значения которых равны, соответственно, нулю и двум. Таким образом, для нахождения вероятности возможного значения 0 случайной величины XY, нужно сложить вероятности произведений $x_1 \cdot y_1$, $x_1 \cdot y_2$. В результате получим $P(XY=0) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} = \frac{3}{6}$. По аналогии находим $P(XY=2) = \frac{1}{4} \cdot \frac{2}{3} + \frac{1}{4} \cdot \frac{1}{3} = \frac{3}{12}$.

	$x_1 \cdot y_1$	$x_1 \cdot y_2$	$x_2 \cdot y_1$	$x_2 \cdot y_2$	$x_3 \cdot y_1$	$x_3 \cdot y_2$
X	0	0	1	2	2	4
P	$\frac{1}{2} \cdot \frac{1}{3}$	$\frac{1}{2} \cdot \frac{2}{3}$	$\frac{1}{4} \cdot \frac{1}{3}$	$\frac{1}{4} \cdot \frac{2}{3}$	$\frac{1}{4} \cdot \frac{1}{3}$	$\frac{1}{4} \cdot \frac{2}{3}$

XY	0	1	2	4
P	$\frac{6}{12}$	$\frac{1}{12}$	$\frac{3}{12}$	$\frac{2}{12}$

Приведя к общему знаменателю, равному 12, вычисленные вероятности и вероятности возможных значений 1 и 4 случайной величины XY, получим таблицу, задающую ее закон распределения. Просуммировав вероятности, стоящие в ее второй строке, убеждаемся в справедливости равенства

$$\sum_{i=1}^{4} p_i = \frac{6}{12} + \frac{1}{12} + \frac{3}{12} + \frac{2}{12} = 1.$$

Таким образом, закон распределения случайной величины XY, найден верно. XY

Числовые характеристики дискретных случайных величин. К основным числовым характеристикам случайных величин относятся: математическое ожидание, дисперсия, среднее квадратическое отклонение, мода и медиана, начальные и центральные теоретические моменты.

Математическим ожиданием дискретной случайной величины X называется сумма произведений всех возможных ее значений и вероятностей, с которыми эти значения принимаются.

Математическое ожидание принято обозначать $M(X)$ или MX. Если значениями дискретной случайной величины X являются члены конечной последовательности $x_1, x_2, x_3, \ldots, x_i, \ldots, x_n$, а $p_i = P(X = x_i)$ — вероятность, с которой X принимает значение x_i, $i = 1, 2, \ldots, n$, то

$$M(X) = x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots + x_i p_i + \ldots + x_n \cdot p_n = \sum_{i=1}^{n} x_i p_i .$$

Если дискретная случайная величина принимает значения бесконечной числовой последовательности, т.е. имеет счетное множество возможных значений, то математическое ожидание записывается в виде *числового ряда*

$$M(X) = \sum_{i=1}^{\infty} x_i \cdot p_i$$

и равно сумме этого ряда.

Вероятностный смысл математического ожидания состоит в том, что оно приближенно равно *среднему арифметическому* всех возможных значений дискретной случайной величины, которое обозначается $\overline{X}$ и вычисляется по формуле

$$\overline{X} = \frac{1}{n}(x_1 + x_2 + \ldots + x_i + \ldots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i .$$

По этой причине математическое ожидание случайной величины называют её *средним значением* или *центром распределения*.

Математическое ожидание обладает следующими свойствами:

а) математическое ожидание константы C равно этой константе, т.е. $M(C) = C$.

б) постоянный множитель C можно выносить за знак математического ожидания, т.е.

$$M(CX) = CM(X);$$

в) математическое ожидание суммы (разности) случайных величин равно сумме (разности) их математических ожиданий, т.е.

$$M(X + Y) = M(X) + M(Y), \quad M(X - Y) = M(X) - M(Y);$$

г) математическое ожидание отклонения случайной величины от её математического ожидания равно нулю, т.е.

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

д) математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий, т.е.

$$M(X \cdot Y) = M(X) \cdot M(Y).$$

Приведенные свойства математического ожидания суммы и произведения двух случайных величин переносятся на n независимых случайных величин.

Пример. 5.23. Случайная величина X имеет закон распределения, заданный таблицей. Найти ее математическое ожидание и математическое ожидание случайной величины $3X + 2$.

X	3	4	5
P	0,3	0,4	0,5

Решение. Вычисляя произведение 1-го и 2-го элемента каждого столбца указанной таблицы, и суммируя произведения, находим

$$M(X) = 3 \cdot 0{,}3 + 4 \cdot 0{,}4 + 5 \cdot 0{,}3 = 0{,}9 + 1{,}6 + 1{,}5 = 4.$$

Далее используя свойства математического ожидания, вычисляем

$$M(3X + 2) = M(3X) + M(2) = 3M(X) + 2 = 3 \cdot 4 + 2 = 14.$$

Математическое ожидание $M(X)$ полностью не характеризует случайную величину, а указывает на ее среднее арифметическое $\overline{X}$, вокруг которого группируются ее значения. Нетрудно привести примеры случайных величин, имеющих одно и то же математическое ожидание, но принимающих различные значения. Например, среднегодовая температура может быть равна 0° при изменении температуры в году от –10° до +10° и при ее колебаниях от –30° до +30°. Для оценки рассеивания значений случайной величины X относительно ее математического ожидания $M(X)$ используется числовая характеристика, называемая дисперсией.

Дисперсией случайной величины X называется математическое ожидание квадрата отклонения случайной величины от её математического ожидания. Дисперсия обозначается $D(X)$ и определяется формулой

$$D(X) = M((X - M(X))^2 .$$

В следующем утверждении приводится формула, которой удобно пользоваться, если требуется вычислить обе числовые характеристики $M(X)$ и $D(X)$.

Теорема 5.7. *Дисперсия равна разности математического ожидания квадрата случайной величины X и квадратом ее математического ожидания, т.е. справедлива формула:*

$$D(X) = M(X^2) - (M(X))^2. \tag{5.5}$$

Дисперсия обладает следующими свойствами:

а) дисперсия постоянной величины равна нулю, т.е.

$$D(C) = 0;$$

б) постоянный множитель C значений случайной величины можно выносить за знак дисперсии, возведя его в квадрат, т.е.

$$D(CX) = C^2 D(X);$$

в) дисперсия суммы и разности независимых случайных величин равно сумме их дисперсий т.е.

$$D(X \pm У) = D(X) + D(У);$$

г) Дисперсия случайной величины не изменится, если к этой случайной величине прибавить постоянную C, т.е.

$$D(X + C) = D(X).$$

Дисперсия оценивает величину рассеивания значений случайной величины относительно ее математического ожидания, но обладает одним недостатком. Ее размерностью является размерность квадрата случайной величины. Например, если производятся замеры точности, изготовляемых изделий, измеряемые в миллиметрах (*мм*), то размерность дисперсии будет измеряться в *мм* ^{2}. По этой причине для более точной оценки отклонения значений случайной величины от ее математического ожидания вводится еще одна числовая характеристика.

Средним квадратическим отклонением или *стандартным отклонением* дискретной случайной величины называется квадратный корень из её дисперсии. Данная числовая характеристика обозначается $\sigma(X)$ и вычисляется по формуле:

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

В силу определения размерность среднего квадратического отклонения $\sigma(X)$ совпадает с размерностью случайной величины X.

Пример. 5.24. Найти математическое ожидание, дисперсию и среднее квадратическое отклонение случайной величины X, закон распределения которой задан таблицей.

X	–2	1	3
P	0,1	0,7	0,2

Решение. Так как возможными значениями X^2 являются квадраты значений X и их вероятности совпадают с вероятностями соответствующих значений X, то их математические ожидания равны, следовательно:

$$M(X) = -2\cdot 0{,}1 + 1\cdot 0{,}7 + 3\cdot 0{,}2 = 1{,}1; \quad M(X^2) = (-2)\cdot(-2)\cdot 0{,}1 + 1\cdot 1\cdot 0{,}7 + 9\cdot 0{,}2 = 2{,}9.$$

Далее, используя формулу (5.5), находим дисперсию

$$D(X)= M(X^2) - (M(X))^2 = 2{,}9 - (1{,}1)2 = 1{,}69$$

и, извлекая из нее квадратный корень, определяем среднее квадратическое отклонение $\sigma(X) = \sqrt{D(X)} = \sqrt{1{,}69} = 1{,}3$.

Медианой дискретной случайной величины X называется такое число, обозначаемое $\mathrm{Me}\, X$, которое принадлежит отрезку $[x_l, x_{l+1}]$. Индекс l концов этого отрезка определяются соотношениями

$$\sum_{i=1}^{l} p_i \le 0{,}5; \qquad \sum_{i=1}^{l+1} p_i > 0{,}5.$$

Если число l найдено, то медиана вычисляется по формуле

$$\mathrm{Me}\, X = x_l + \frac{x_{l+1} - x_l}{p_{l+1}}\left(\frac{1}{2} - \sum_{j=1}^{l} p_j\right). \tag{5.6}$$

Модой случайной величины X называется такое ее возможное значение, вероятность которого наибольшая по величине. Мода обозначается $\mathrm{Mo}\, X$ и определяется соотношением

$$P(X = \mathrm{Mo}\, X) = \max\{ p_i \mid i = 1, 2, \ldots, n\}$$

Пример 5.25. Определить моду и вычислить медиану случайной величины X, закон распределения которой задан таблицей.

	x_1	x_2	x_3	x_4·	x_5·	x_6
X	0	0	1	2	2	4
P	0,1	0,15	0,2	0,3	0,1	0,15

Решение. Последовательно выбирая вероятности, начиная с первой и суммируя их, убеждаемся в справедливости соотношений

$$p_1 + p_2 + p_3 = 0{,}1 + 0{,}15 + 0{,}2 = 0{,}45 \le 0{,}5, \quad p_4 + p_5 + p_6 = 0{,}3 + 0{,}1 + 0{,}15 = 0{,}55 > 0{,}5.$$

следовательно, $l = 3$ и медиана принадлежит отрезку $[x_3, x_4] = [1, 2]$. Используя далее формулу (5.6) определяем значение медианы:

$$\mathrm{Me}\, X = 3 + \frac{4-3}{0{,}3}(0{,}5 - 0{,}45) \approx 3{,}17.$$

Наибольшая вероятность возможных значений случайной величины X $p_4 = 0{,}3$, следовательно, мода $\mathrm{Mo}\, X$ равна $x_4 = 2$.

Начальный и центральный теоретические моменты. Данные числовые характеристики используются в математической статистике в *методе моментов* для вычисления приближенных точечных значений неопределенных параметров случайной величины, определяющей количественный или качественный признак случайных событий психологического или социального характера.

Начальным теоретическим моментом k-го порядка случайной величины X, обозначаемым ν_k, называется математическое ожидание ее k-степени, т.е. случайной величины X^k. Таким образом, ν_k вычисляется по формуле

$$\nu_k = M(X^k) = x_1^k p_1 + x_2^k p_2 + \ldots + x_i^k p_i + \ldots + x_n^k p_n = \sum_{i=1}^{n} x_i^k p_i \qquad (5.7)$$

Центральным теоретическим моментом k-го порядка случайной величины X, обозначаемым μ_k, называется математическое ожидание k-степени отклонения X от своего математического ожидания, т.е. — математическое ожидание случайной величины $(X - M(X))^k$. Таким образом, μ_k вычисляется по формуле

$$\mu_k = M((X - M(x))^k) = (x_1 - M(x))^k p_1 + (x_2 - M(x))^k p_2 + \ldots + (x_n - M(x))^k p_n$$
$$= \sum_{i=1}^{n} (x_i - M(x))^k p_i \qquad (5.8)$$

Математическое ожидание $M(X)$ и дисперсия $D(X)$ являются частными случаями, соответственно, начального и центрального теоретического момента, а именно $M(X) = \nu_k$ при $k = 1$, $D(X) = \mu_k$ при $k = 2$.

Используя формулы (5.7) и (5.8) центральные теоретические моменты можно выражать через начальные.

Пример 2.26. Выразить центральные теоретические моменты μ_k, $k = 0, 1$ случайной величины X, закон распределения которой задан таблицей 5.1, через ее начальные теоретические моменты.

Решение. Пусть $k = 0$, тогда все степени x^k и $((x - M(x))^k$ равны нулю и из формул (5.7), (5.8) следует справедливость равенств

$$\mu_0 = \nu_0 = \sum_{i=1}^{n} p_i = 1.$$

При $k = 1$ начальный теоретический момент $\nu_1 = M(X)$. Преобразовав несложным образом правую часть равенств (5.8) при $k = 1$, получим:

$$\sum_{i=1}^{n} (x_i - M(x))p_i = \sum_{i=1}^{n} x_i\, p_i - \sum_{i=1}^{n} M(x)p_i = \sum_{i=1}^{n} x_i\, p_i - M(x)\sum_{i=1}^{n} p_i = M(x) - M(x) = 0,$$

так как сумма всех вероятностей равна единице. Таким образом, $\mu_1 = 0 \cdot \nu_1$.

Функция распределения дискретной случайной величины и ее свойства.

Различные числовые характеристики случайной величины X можно получать с помощью ее функции распределения. Например, вычислять вероятность принадлежности возможных значений X любому конечному и бесконечному промежутку вида $(-\infty, b)$ и $[a, b)$.

Функцией распределения случайной величины X называется функция, обозначаемая $F(x)$, определенная на всей числовой оси, т.е. на интервале $(-\infty, +\infty)$, значение которой равны вероятности того, что в результате испытания значение случайной величины окажется меньше x, т.е. для любого $x \in (-\infty, +\infty)$ значение $F(x)$ определяются формулой

$$F(x) = P(X < x). \tag{5.9}$$

Геометрически равенство (5.9) можно интерпретировать следующим образом: $F(x)$ определяет вероятность того, что точка, соответствующая принятому значению случайной величины, расположена левее точки x на числовой оси. Другими словами, $F(x)$ — вероятность принадлежности принятого случайной величиной значения интервалу $(-\infty, x)$ при испытании.

Функция распределения обладает следующими свойствами:

а) $F(x)$ ограничена и все ее значения принадлежат отрезку $[0, 1]$, т.е. выполняется двойное неравенство $0 \le F(x) \le 1$;

б) $F(x)$ — функция, неубывающая на интервале $(-\infty, +\infty)$, т.е. для любых $x_1 < x_2$ выполняется неравенство $F(x_1) \le F(x_2)$;

в) в любой точке $a \in (-\infty, +\infty)$ функция $F(x)$ — непрерывна слева, т.е. $\lim\limits_{x \to a-0} F(x) = F(x_0)$;

г) если возможные значения случайной величины X принадлежат интервалу (a, b), то $F(x) = 0$ при $x \le a$ и $F(x) = 1$ при $x \ge b$;

д) если возможные значения случайной величины X принадлежат интервалу $(-\infty, +\infty)$, то при неограниченном уменьшении значений аргумента ($x \to -\infty$) значения $F(x)$ стремятся к нулю, а при неограниченном увеличении аргумента ($x \to +\infty$) — к единице, т.е.

$$\lim_{x \to -\infty} F(x) = 0 \text{ и } \lim_{x \to +\infty} F(x) = 1.$$

е) вероятность того, что случайная величина X примет значение из полуинтервала $[a, b)$ равна приращению ее функции распределения на этом полуинтервале, т.е. справедливо равенство

$$P(a \le X < b) = F(b) - F(a).$$

Если закон распределения случайной величины X задан таблицей, в которой ее различные значения упорядочены по возрастанию, то значения ее функции распределения $F(x)$

X	x_1	x_2	...	x_i	x_{i+1}	...	x_n
P	p_1	p_2	...	p_i	p_{i+1}	...	p_n

для любых точек x, таких что $x_i < x \le x_{i+1}$, вычисляются по формуле

$$F(x) = p_1 + p_2 + \ldots + p_j + \ldots + p_i = \sum_{j=1}^{i} p_j. \tag{5.8}$$

Из формулы (5.8) следует, что функция распределения является кусочно-постоянной функцией, принимающей на полуинтервале $(x_i, x_{i+1}]$ постоянное значение, равное

$$s_i = p_1 + p_2 + \ldots + p_j + \ldots + p_i = \sum_{j=1}^{i} p_j.$$

Точками разрыва первого рода функции $F(x)$ являются возможные значения случайной величины x_i, в которых она терпит разрыв, равный p_i.

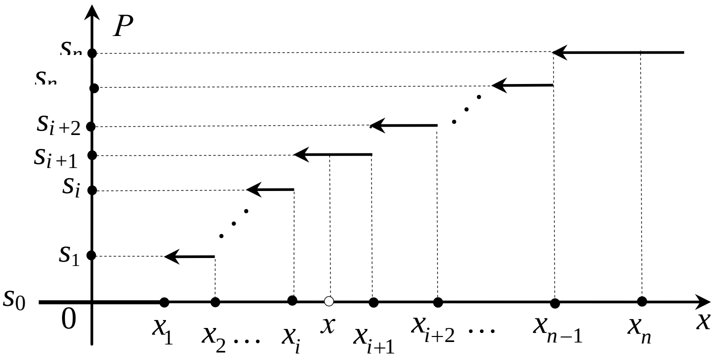

Рис. 5.5.

График функции распределения случайной величины X, закон распределения которой задан таблицей, изображен на рис. 5.4.

Пример 5.26. Случайная величина X задана законом распределения, заданным таблицей. Требуется найти: а) функцию распределения $F(x)$ и изобразить её график; б) вероятность попадания X на $[0,2)$.

	x_1	x_2	x_3
X	–2	1	3
P	0,1	0,7	0,2

Решение. а) Значения данной случайной величины X разбивают интервал $(-\infty,+\infty)$ на полуинтервалы $(-\infty,-2]$, $(-2,1]$, $(1,3]$, $(3,+\infty)$, на которых функция распределения $F(X)$ принимает постоянные значения. Эти значения можно вычислять с помощью формулы (5.8) или для наглядности, используя геометрическую интерпретацию функции распределения. Воспользуемся вторым способом.

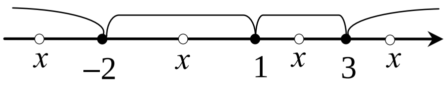

Рис. 5.6.

Для любого x, принадлежащего $(-\infty,-2]$, слева от него нет (см. рис. 5.5) на числовой оси возможных значений случайной величины X. Следовательно, вероятность принадлежности возможных значений X этому интервале равна нулю, т.е.

$$F(x) = P(X < x) = 0$$

для любых $x \in (-\infty,-2]$.

Для любого x, принадлежащего $(-2,1]$, слева от него (см. рис. 5.6) расположено одно возможное значение $x_1 = -2$. Следовательно, вероятность принадлежности возможных значений X этому интервалу равна, вероятности принятия ею значения $x_1 = -2$, т.е.

$$F(x) = P(X < x) = P(X = x_1) = 0{,}1$$

для любых $x \in (-2,1]$.

Для любого x, принадлежащего $(1,3]$, слева от него (см. рис. 5.6) расположено два возможных значения $x_1 = -2$ и $x_2 = 1$. Следовательно, вероятность принадлежности возможных значений X этому интервалу равна, вероятности принятия ею значения $x_1 = -2$ или $x_2 = 1$ и по теореме о сумме вероятностей независимых событий

$$F(x) = P(X < x) = P(X = x_1) + P(X = x_2)) = 0{,}1 + 0{,}7 = 0{,}8$$

для любых $x \in (-2,1]$.

Наконец, для любого x, принадлежащего $(3,+\infty]$, слева от него (см. рис.5.6) расположено все возможных значения $x_1 = -2$, $x_2 = 1$ и $x_3 = 3$. Следовательно, вероятность принадлежности возможных значений X этому интервалу равна, вероятности принятия ею или $x_1 = -2$, или $x_2 = 1$, или $x_3 = 3$ и по теореме о сумме вероятностей несовместных событий

$$F(x) = P(X < x) = P(X = x_1) + P(X = x_2) + P(X = x_3) = 0{,}1 + 0{,}7 + 0{,}2 = 1$$

для любых $x \in (-2{,}1]$.

Записав результаты вычислений в виде формулы, получим аналитическое задание функции распределения следующего вида:

$$F(x) = \begin{cases} 0 & \text{при} \quad -\infty < x \le -2; \\ 0{,}1 & \text{при} \quad -2 < x \le 1; \\ 0{,}8 & \text{при} \quad 1 < x \le 3; \\ 1 & \text{при} \quad 3 < x \le +\infty. \end{cases}$$

б) Для нахождения вероятность попадания X на полуинтервал $[0, 2)$ воспользуемся свойством е) функции распределения

$$P(a \le X < b) = F(b) - F(a).$$

В рассматриваемом случае $a = 0$, $b = 2$. Так как $b \in (1{,}3]$ и $a \in (-2{,}1]$, то используя формулу для вычисления значений функции распределения, получим:

$$F(2) = 0{,}8 \text{ и } F(0) = 0{,}1,$$

следовательно, $P(0 \le X < 2) = F(2) - F(0) = 0{,}8 - 0{,}1 = 0{,}7$.

Законы распределения дискретных случайных величин. При проведении повторных независимых испытаниях по схеме Бернулли число успехов m наступивших при n испытаниях является случайной величиной X, принимающей возможные значения $X = 0, 1, \ldots k, \ldots, m$. В зависимости от способа вычисления вероятностей приведенных значений различают биномиальный закон распределения и закон распределения Пуассона.

Биномиальное распределение. Распределение вероятностей указанной случайной величины X называется биномиальным, если они вычисляются по формуле Бернулли

$$P(X = k) = P_n(k) = C_n^k p^k q^{n-k}, \qquad k = 0{,}1, \ldots, m.$$

Закон назван биномиальным, так как правая часть приведенного равенства совпадает с общим членом разложения бинома Ньютона

$$(p+q)^n = C_n^0 p^0 q^n + C_n^1 p q^{n-1} + \ldots + C_n^k p^k q^{n-k} + \ldots + C_n^n p^n q^0.$$

Основные числовые характеристики случайной величина, распределенной по биномиальному закону вычисляются по формулам:

$$M(X) = np, \qquad D(X) = npq, \qquad \sigma(X) = \sqrt{npq},$$

а сам закон определяется следующей таблицей

P	0	1	2	…	x_k	…	x_n
X	$C_n^0 p^0 q^n$	$C_n^1 p q^{n-1}$	$C_n^2 p^2 q^{n-2}$	…	$C_n^k p^k q^{n-k}$	…	$C_n^n p^n q^0$

Распределение Пуассона. Распределение вероятностей случайной величины X — числа успехов при повторных независимых испытаниях, называется распределением Пуассона, если вероятности ее возможных значений $0, 1, \ldots k, \ldots, m$ вычисляются по формуле Пуассона

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

Данное распределение возникает при проведении при достаточно большого числа испытаний, при которых вероятность успеха p достаточно мала.

Основные числовые характеристики случайной величина, распределенной по биномиальному закону вычисляются по формулам:

$$M(X) = D(X) = \lambda, \quad \sigma(X) = \sqrt{\lambda}\,.$$

Геометрическое распределение. Если по схеме Бернулли проводятся испытания до первой реализации успеха, то X, число проведенных испытаний, является бесконечной случайной величиной, принимающей значения $X = 1, 2, \ldots k, \ldots$.

При $X = k$ наступление первого успеха происходит после реализации $(k-1)$-го неуспеха. Таким образом, событие $X = k$ является произведением $(k-1)$-го неуспеха и одного успеха. Следовательно, вероятность $X = k$ по теореме о произведении независимых событий равна pq^{k-1}, т.е.

$$P(X = k) = pq^{k-1}, \text{ где } k = 1, 2, \ldots .$$

Последовательность произведений pq^{k-1}, $k = 1, 2, \ldots$ является бесконечно убывающей геометрической прогрессией с первым членом p и знаменателем $q < 1$ (тривиальный случай $q = 1$ исключен, так как в этом случае $p = 0$). Следовательно, сумма всех ее членов равна

$$\sum_{k=1}^{+\infty} pq^{k-1} = \frac{p}{1-q} = \frac{p}{1-(1-p)} = 1\,.$$

Основные числовые характеристики случайной величина, имеющей геометрический закон распределения, вычисляются по формулам:

$$M(X) = \frac{1}{p}, \quad D(X) = \frac{q}{p^2}, \quad \sigma(X) = \frac{\sqrt{q}}{p},$$

а сам закон определяется следующей таблицей

P	1	2	3	…	k	…
X	p	pq	pq^2	…	pq^{k-1}	…

Гипергеометрическое распределение. При выборках n элементов из некоторого множества, состоящего из N элементов, в котором имеется M элементов с заданным признаком, среди выбранных элементов может оказаться m элементов с признаком. Возможными значениями m являются числа $0, 1, 2, \ldots , \max \{n, M\}$.

Закон распределения X называется *гипергеометрическим*, если случайная величина X принимает значения $m = 0, 1, 2, \ldots , \max \{n, M\}$ с вероятностями

$$P(X = m) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n},$$

которые вычисляются по формуле классической вероятности с применением комбинаторного правила произведения.

Основные числовые характеристики случайной величина, имеющей гипергеометрический закон распределения, вычисляются по формулам:

$$M(X) = n\frac{M}{N}, \quad D(X) = n\frac{M}{N}\left(1-\frac{M}{N}\right)\left(1-\frac{n-1}{N-1}\right), \; \sigma(X) = \sqrt{M(X)\left(1-\frac{M}{N}\right)\left(1-\frac{n-1}{N-1}\right)}\,.$$

5.6. Непрерывные случайные величины

В случае, когда нельзя отделить одно возможное значение случайной величины от другого промежутком, не содержащим других ее возможных значений, случайная величина называется *непрерывной.*

Из приведенного определения следует, что множеством возможных значений непрерывной случайной величины могут быть только конечные или бесконечные промежутки. Очевидно, что число возможных значений непрерывной случайной величины бесконечно.

Примерами непрерывных случайных величин являются значения временных факторов, влияющих на протекание социальных и психологических и процессов, функционирование и развитие социальных систем, временная реакция индивида на различные стимулы, различные психометрические и социометрические данные.

Непрерывные случайные величины обладают двумя свойствами, отличающими их от дискретных случайных величин:

а) вероятность любого отдельно взятого возможного значения x_0 непрерывной случайной величины X равна нулю, т.е.

$$P(X = x_0) = 0;$$

б) для любой непрерывной случайной величины X выполняются равенства:

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

Для непрерывных случайных величин вводится понятие *функции распределения*, которая называется также *интегральной функцией* и обозначается $F(x)$. Как в дискретном случае, функция распределения $F(x)$ непрерывной случайной величины определена на всем интервале $(-\infty, +\infty)$. Значение $F(x)$ при заданном $x \in (-\infty, +\infty)$ совпадает с вероятностью, того, что X принимает значение, меньшее x, т.е. определяется формулой

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

Интегральная функция обладает всеми свойствами а) — е) функции распределения дискретной случайной величины, которые приведены в подразделе 5.5 на стр. 145.

Введение функции распределения позволяет дать более точное определение непрерывной случайной величины.

Случайная величина X называется непрерывной, если её функция распределения $F(x)$ непрерывна и дифференцируема на всем интервале $(-\infty, +\infty)$, возможно за исключением конечного числа точек.

Плотность распределения. Задание функции распределения является не единственным способом характеризации непрерывной случайной величины. Ее можно задать с помощью функции, которая называется *плотностью распределения* или *дифференциальной функцией*.

Плотностью распределения непрерывной случайной величины называется производная ее функции распределения $F(x)$. Плотность обозначается $p(x)$ и определяется формулой:

$$p(x) = F'(x).$$

График плотности распределения называется *кривой распределения.*

Плотность распределения обладает следующими свойствами:

а) плотность распределения принимает неотрицательные значения на интервале $(-\infty, +\infty)$, т.е. справедливо неравенство

$$p(x) \geq 0 \text{ для любых } x \in (-\infty, +\infty);$$

б) несобственный интеграл с бесконечными пределами интегрирования от плотности распределения равен единице, т.е. справедливо равенство

$$\int_{-\infty}^{+\infty} p(x)\,dx = 1;$$

в) вероятность принятия непрерывной случайной величиной значения из интервала (a, b) равна определенному интегралу с нижним a и верхним b пределами интегрирования, т.е. справедливо равенство

$$P(a < X < b) = \int_{a}^{b} p(x)\,dx;$$

г) функция распределения $F(x)$ равна несобственному интегралу с бесконечным нижним и переменным верхними пределами, т.е. вычисляется по формуле

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

Свойства а) — г) плотности распределения геометрически интерпретируются следующим образом:

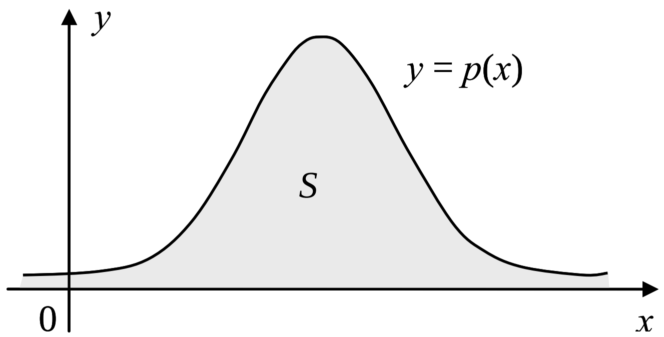

Рис. 5.7.

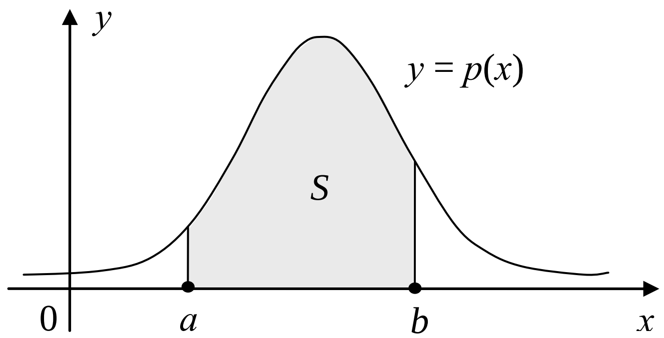

Рис. 5.8.

а) график функции распределения $y = p(x)$ (кривая распределения) расположен выше координатной оси Ox (см. рис. 5.7);

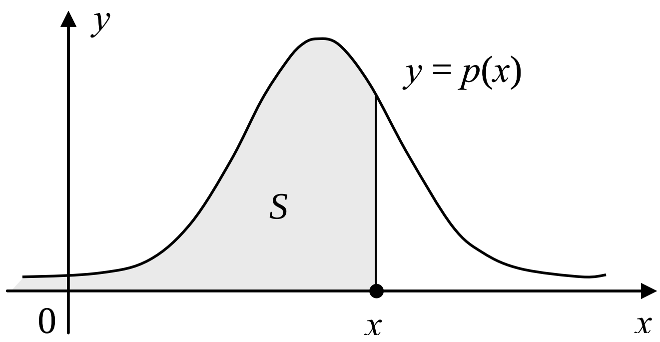

Рис. 5.9.

б) площадь S бесконечной фигуры, ограниченной кривой распределения и координатной осью Ox, равна единице (см. рис. 5.7);

в) вероятность попадания возможного значения непрерывной случайной величины в интервал (a, b) равна площади S плоской фигуры, ограниченной кривой распределения, координатной осью Ox и перпендикулярами, проведенными из концов интервала до пересечения с кривой распределения (см. рис. 5.8);

г) значение функции распределения $F(x)$ в любой точке $x \in (-\infty, +\infty)$ равно площади S бесконечной фигуры, ограниченной кривой распределения, осью Ox и перпендикуляром, проведенным из точки x до пересечения с кривой распределения (см. рис. 5.9).

Числовые характеристики непрерывных случайных величин. К основным числовыми характеристиками непрерывных случайных величин относятся математическое ожидание, дисперсия, среднее квадратическое отклонение, начальный и центральный теоретические моменты, медиана и мода, которые определяются также, как и соответствующие числовые характеристики дискретных случайных величин, и отличаются лишь способом вычисления.

Математическое ожидание $M(X)$ непрерывной случайной величины X, возможные значения которой образуют отрезок $[a, b]$, вычисляется по формуле

$$M(X) = \int_a^b x\,p(x)dx.$$

Если возможными значениями X являются любые точки числовой оси, то математическое ожидание вычисляется с помощью несобственного интеграла с бесконечными пределами интегрирования, т.е.

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

Дисперсией $D(X)$ непрерывной случайной величины X, множеством возможных значений которой является отрезок $[a, b]$, называется определенный интеграл

$$D(X) = \int_a^b (x - M(x))^2 p(x)dx.$$

Если множество возможных значений X совпадает с интервалом $(-\infty, +\infty)$, то дисперсией называется несобственный интеграл

$$D(X) = \int_{-\infty}^{+\infty} (x - M(x))^2 p(x)dx.$$

Математическое ожидание и дисперсия непрерывных случайных величин обладают всеми свойствами аналогичных числовых характеристик дискретных случайных величин. По этой причине приведем лишь часто употребляемые упрощенные формулы для вычисления дисперсии для случайных величин, определенных на конечном отрезке и всей числовой оси:

$$D(X) = \int_a^b x^2 p(x)dx - (M(x))^2, \qquad D(X) = \int_{-\infty}^{+\infty} x^2 p(x)dx - (M(x))^2.$$

Среднее квадратическое отклонение $\sigma(X)$ непрерывной случайной величины X определяется по аналогии с дискретным случаем, т.е.

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

Начальный ν_k *и центральный* μ_k теоретические моменты порядка k непрерывной случайной величины X определяются, как для дискретной, и вычисляются по формулам

$$\nu_k = \int_a^b x^k p(x)dx, \qquad \mu_k = \int_a^b (x - M(x))^k p(x)dx,$$

если X определена на отрезке $[a, b]$. Если X определена на всей числовой оси, то

$$\nu_k = \int_{-\infty}^{+\infty} x^k p(x)dx, \qquad \mu_k = \int_{-\infty}^{+\infty} (x - M(x))^k p(x)dx.$$

Модой непрерывной случайной величины, обозначаемой $\text{Mo}\,X$, называется ее возможное значение, в котором плотность распределения принимает максимальное значение, т.е.

$$p(\text{Mo}\,X) = \max \{ p(x) \mid x \in [a, b] \},$$

если X определена на отрезке $[a, b]$ и

$$p(\text{Mo}\,X) = \max \{ p(x) \mid x \in (-\infty, +\infty) \},$$

если X определена на отрезке всей числовой оси.

Медиана непрерывной случайной величины X определяется, как и в дискретном случае, обозначается $\text{Me}\,X$ и вычисляет по формуле

$$P(X < \text{Me}\,X) = P(X > \text{Me}\,X) = \frac{1}{2}.$$

Так как в силу определения функции распределения $P(X < \text{Me}X) = F(\text{Me}X)$, то для нахождения медианы непрерывной случайной величины X достаточно непосредственно разрешить уравнение $F(\text{Me}X) = \frac{1}{2}$ относительно $\text{Me}\,X$ или, воспользовавшись обратной функцией $F^{-1}(x)$, вычислить

$$\text{Me}X = F^{-1}(F(\text{Me}X)) = F^{-1}\left(\frac{1}{2}\right).$$

Законы распределения непрерывных случайных величин. Непрерывные случайные величины, как и дискретные, имеют различные распределения возможных значений и вероятностей. Плотность распределения непрерывной случайной величины X называется ее *законом распределения*. Закон распределения X полностью определяется ее функцией плотности распределения $p(x)$ (дифференциальной функцией). В зависимости от вида $p(x)$ различают: *равномерный*, *показательный* (*экспоненциальный*) и *нормальный* законы распределения непрерывных случайных величин.

Равномерный закон. Закон распределения вероятностей непрерывной случайной величины X на отрезке $[a, b]$ называется равномерным, если плотность распределения $p(x)$ принимает постоянное значение $\frac{1}{b-a}$ на отрезке, и равна нулю вне него, т.е.

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

Вычислив для данной плотности распределения несобственные интегралы, определяющие функцию распределения $F(x)$, математическое ожидание $M(x)$ и дисперсию $D(x)$, получим следующие формулы для их вычисления:

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

$$M(X) = \frac{a+b}{2}; \qquad D(X) = \frac{(b-a)^2}{12}; \quad \sigma(X) = \frac{|b-a|}{2\sqrt{3}}.$$

Графики функций $p(x)$ и $F(x)$ приведены, соответственно, на рис. 5.10 и 5.11.

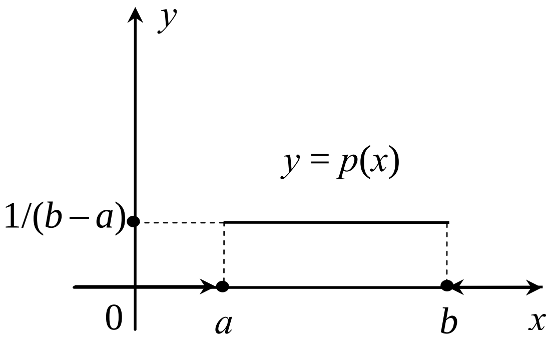

Рис. 5.10.

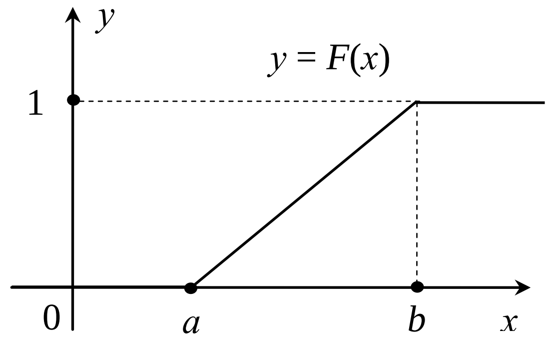

Рис. 5.11.

Показательный (экспоненциальный) закон распределения. Закон распределения непрерывной случайной величины X называется *показательным или экспоненциальным,* если плотность ее распределения определяется функцией

$$p(x) = \begin{cases} \lambda e^{-\lambda x} \ \text{при}\ x \ge 0, \\ 0 \ \text{при}\ x < 0. \end{cases}$$

Вычислив несобственный интеграл $\int\limits_{-\infty}^{x} p(x)\,dx$ для указанной плотности распределения, получим функцию распределения X, следующего вида:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} \ \text{при}\ x \ge 0, \\ 0 \ \text{при}\ x < 0. \end{cases}$$

Вычисление соответствующих несобственных интегралов приводит к формулам для нахождения математического ожидания, дисперсии и среднего квадратического отклонения случайной величины X:

$$M(X) = \sigma(x) = \frac{1}{\lambda}, \quad D(X) = \frac{1}{\lambda^2}.$$

Вероятность попадания значения X в интервал (a, b) вычисляется по формуле

$$P(a < x < b) = e^{-\lambda a} - e^{-\lambda b}.$$

Графики плотности распределения и функции распределения изображены на рис. 5.12 и 5.13.

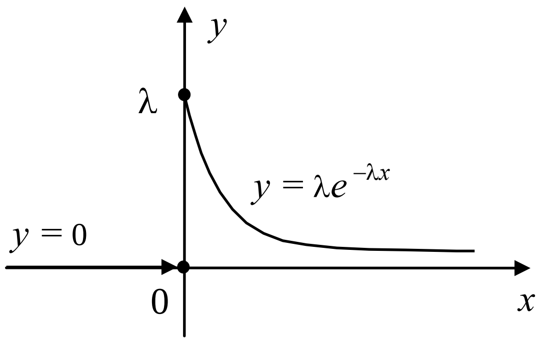

Рис. 5.12.

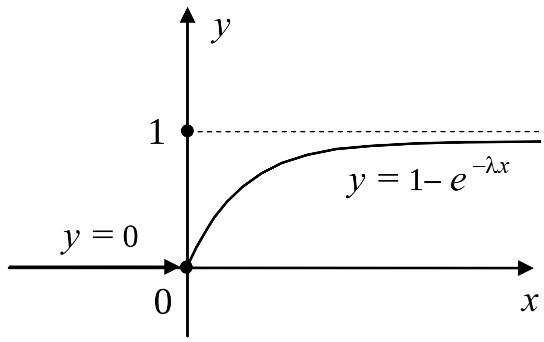

Рис. 5.13.

Нормальный закон распределения. Данный закон распределения имеет универсальный характер. Этому закону подчинены многие случайные явления, процессы, события из разных областей человеческой практики и окружающего мира, таких как экономика, социология, психология, области естественных наук.

Строгое научное обоснование этого феномена дал выдающийся русский математик А.М. Ляпунов. Им было доказано утверждение, названное в честь его теоремой Ляпунова, общая формулировка которой состоит в следующем:

если случайная величина представляет собой сумму очень большого числа взаимно независимых случайных величин, влияние каждой из которых на всю сумму ничтожно мало, то такая случайная величина имеет распределение близкое к нормальному.

В частности нормальное распределение случайной величины всегда возникает при систематических отклонениях ее от своего среднего значения или в тех случаях, когда на нее влияет большое количество случайных факторов.

Закон распределения непрерывной случайной величины X называется *нормальным* (или *законом Гаусса*), если ее плотность распределения $p(x)$ зависит от двух параметров a, σ и имеет следующий вид:

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

Указанная функция является функцией Гаусса специального вида, поэтому ее график называют *нормальной* или *гауссовой кривой*. Эта кривая изображена на рис. 5.14 и обладает следующими свойствами:

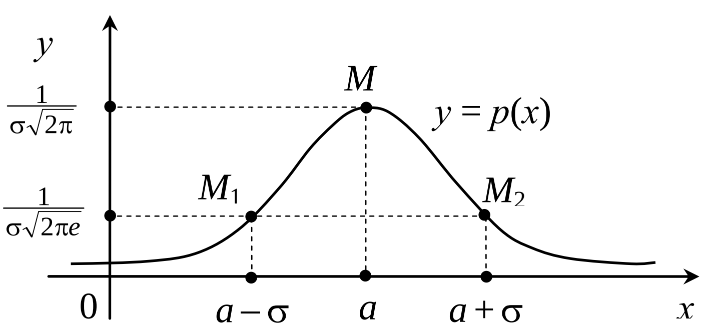

Рис. 5.14.

а) кривая расположена симметрично относительно прямой, заданной уравнением $x = a$;

б) точка $M(a, \frac{1}{\sigma\sqrt{2\pi}})$ — наивысшая точка кривой, т.е. мода $\mathrm{Mo}X = a$ и плотность распределения $p(x)$ достигает максимума $\frac{1}{\sigma\sqrt{2\pi}}$ при при $x = a$;

в) кривая имеет две точки перегиба $M_1(a-\sigma, \frac{1}{\sigma\sqrt{2\pi e}})$ и $M_1(a+\sigma, \frac{1}{\sigma\sqrt{2\pi e}})$.

Математическое ожидание, дисперсия и средне квадратическое отклонение непрерывной случайной величины X определяются равенствами:

$$M(X) = a, \qquad D(X) = \sigma^2, \qquad \sigma(X) = \sigma.$$

Функция распределения $F(x)$ случайной величины X определяется функцией Лапласа и вычисляется по формулам:

$$F(x) = 1 + \Phi\left(\frac{x-a}{\sigma}\right), \text{ если } \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt;$$

$$F(x) = \frac{1}{2} + \frac{1}{2}\Phi\left(\frac{x-a}{\sigma}\right), \text{ если } \Phi(x) = \frac{2}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt.$$

График функции $F(x)$ изображен на рис. 5.15.

Вероятность попадания возможных значений случайной величины X в интервал (α, β) вычисляется по формулам:

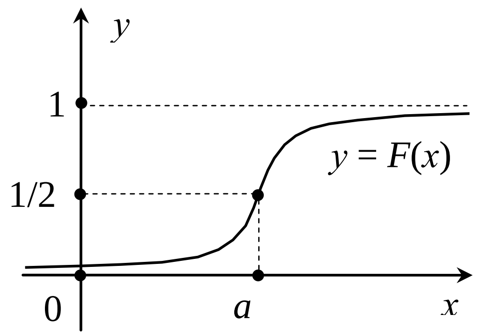

Рис. 5.15.

$$P(\alpha < X < \beta) = \Phi\left(\frac{x-\beta}{\sigma}\right) - \Phi\left(\frac{x-\alpha}{\sigma}\right), \text{ если}$$

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

$$P(\alpha < X < \beta) = \frac{1}{2}\left(\Phi\left(\frac{x-\beta}{\sigma}\right) - \Phi\left(\frac{x-\alpha}{\sigma}\right)\right), \text{ если}$$

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

Важной в практическом и теоретическом отношении является формула, позволяющая вычислять с помощью функции Лапласа вероятность отклонения X от ее математического ожидания $M(x) = a$ на заданную величину ε:

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

Приведенная формула позволяет установить справедливость утверждения, названного «правилом трех сигм», которое определяет интервал, содержащий почти все значения нормально распределенной случайной величины X:

если вероятности случайной величины X распределены по нормальному закону с параметрами a, σ, то практически достоверно, что ее значения заключены в интервале $(a - 3\sigma, a + 3\sigma)$.

Действительно, при $\varepsilon = 3\sigma$ выполняются соотношения

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

Следовательно, с вероятностью, равной 0,9973, неравенство $|X - a| < 3\sigma$ равносильное $a - 3\sigma < X < a + 3\sigma$, можно считать справедливым.

Нормированным или стандартным называется нормальный закон распределения с параметрами $a = 0$ и $\sigma = 1$. Плотностью распределения нормированного закона является функция Гаусса. Для данного закона справедливы все приведенные выше формулы при $a = 0$ и $\sigma = 1$.

Асимметрия и эксцесс являются числовыми характеристиками, которые используются для выявления отличий распределения случайных величин от нормального.

Асимметрией или *коэффициентом асимметрии* называют отношение центрального теоретического момента третьего порядка к кубу среднего квадратического отклонения, т.е. величину

$$A = \frac{\mu_3}{\sigma^3}.$$

При $A = 0$, кривая распределения симметрична относительно прямой, определяемой уравнением $x = \text{Mo}\,X$. При $A > 0$ прямая, заданная уравнением $x = \text{Mo}\,X$, расположена правее прямой, заданной уравнением $x = \text{Mo}\,X$ (см. рис. 5.16). Если $A < 0$, то расположение прямых меняется на противоположное (см. рис. 5.17). На рисунках мода и медиана обозначены, соответственно, x_o и x_e.

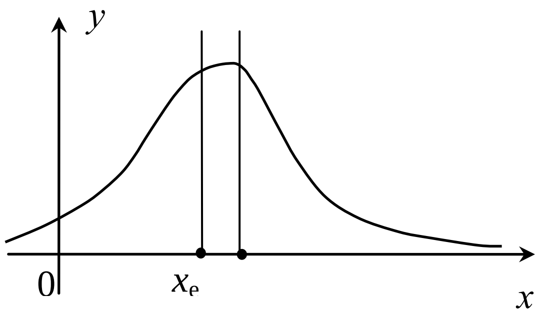

Рис. 5.16.

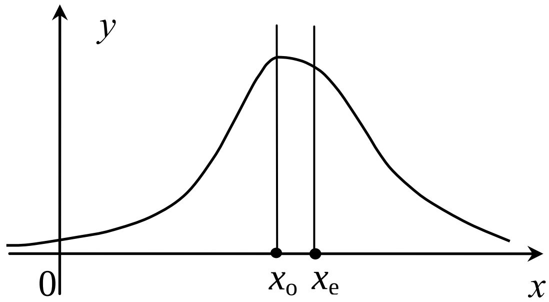

Рис. 5.17.

Эксцессом распределения случайной величины называется числовая характеристика, определяемая равенством

$$E = \frac{\mu_4}{\sigma^4} - 3.$$

При нормальном распределении отношение $\frac{\mu_4}{\sigma^4} = 3$, следовательно, его эксцесс равен нулю. Если эксцесс распределения не равен нулю, то кривая этого распределения отлична от гауссовой кривой при этом:

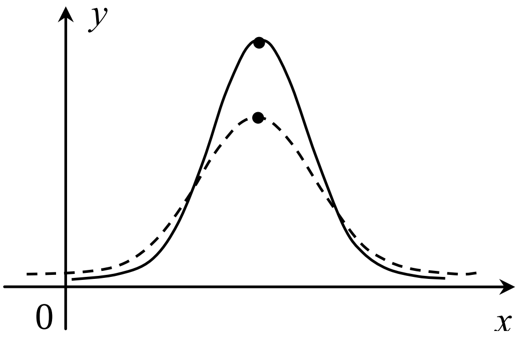

Рис. 5.18. $E > 0$

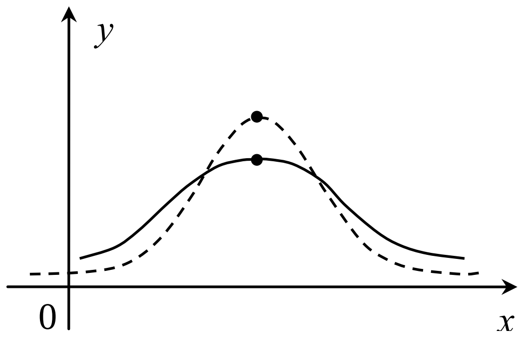

Рис. 5.19. $E < 0$

если эксцесс больше нуля, то «пик» кривой распределения выше и «круче», чем у гауссовой кривой (см. рис. 5.18);

если эксцесс меньше нуля, то пик кривой распределения расположен ниже и более пологий, чем у гауссовой кривой (см. рис. 5.19).

Гауссова кривая на рисунках изображена пунктирной линией.