МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

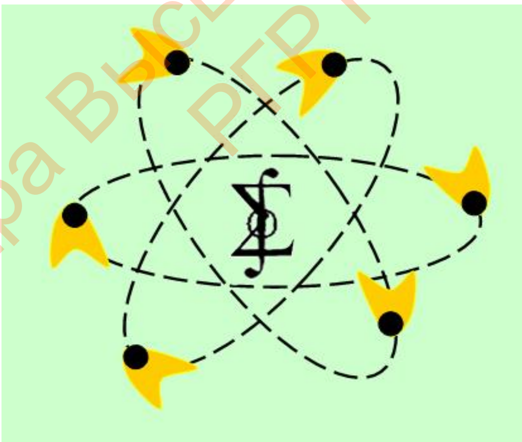
РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

К.В. БУХЕНСКИЙ,

Н.В. ЕЛКИНА, Н.Н. МАСЛОВА

КРАТКИЙ КУРС МАТЕМАТИКИ

Часть 4



Рязань 2014

Министерство образования и науки Российской Федерации

Рязанский государственный радиотехнический университет

К.В. БУХЕНСКИЙ,

Н.В. ЕЛКИНА, Н.Н. МАСЛОВА

КРАТКИЙ КУРС МАТЕМАТИКИ

Часть 4

Учебное пособие

Рязань 2014

УДК 517.38

Краткий курс математики. Часть 4: учеб. пособие / К.В. Бухенский, Н.В. Елкина, Н.Н. Маслова; Рязан. гос. радиотехн. ун-т. – Рязань, 2014. – 140 с.

Содержит основные сведения курса высшей математики в соответствии с Федеральным государственным образовательным стандартом высшего профессионального образования.

Предназначено для студентов направления 240100 «Химическая технология», а также для студентов всех направлений и специальностей, изучающих теорию вероятностей и математическую статистику.

Табл. 29. Ил. 38. Библиогр.: 10 назв.

Событие, случайная величина, закон распределения, математическая статистика

Печатается по решению редакционно-издательского совета Рязанского государственного радиотехнического университета.

Рецензент: кафедра высшей математики Рязанского государственного радиотехнического университета (доц. кафедры канд. физ.-мат. наук А.Б. Дюбуа)

ГЛАВА 16. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ВЕРОЯТНОСТЕЙ

1. Случайные события

Теория вероятностей изучает свойства массовых случайных событий, способных многократно повторяться при воспроизведении определенного комплекса условий. Основное свойство любого случайного события, независимо от его природы, – мера, или вероятность его осуществления.

1.1. Серии опытов со случайными исходами. Частота. Свойства частот

Наблюдаемые нами события или явления можно разделить на три вида: достоверные, невозможные, случайные. Будем говорить, что произведено испытание или опыт, если будет осуществлена определенная совокупность условий.

Достоверным называют событие, которое обязательно произойдет в результате проведения испытания.

Пример. В сосуде содержится вода при нормальном давлении и температуре +20 ^{0}С (атмосферное давление и температура составляют в данном случае совокупность условий). Событие, заключающееся в том, что в результате испытания «вода в сосуде находится в жидком состоянии», – достоверное событие.

Достоверное событие обозначается символом Ω.

Невозможным называют событие, которое заведомо не произойдет, если проведено испытание (опыт). Например, если осуществлены условия предыдущего примера, то событие «вода в сосуде находится в твердом состоянии» – заведомо не произойдет. Невозможное событие обозначается символом $\varnothing$.

Случайное событие – событие, которое в результате испытания может либо произойти, либо не произойти. Например, если брошена вверх монета, то она может упасть так, что сверху будет либо герб, либо решка.

Теория вероятностей не ставит перед собой вопрос: произойдет ли единичное событие или нет. Здесь трудно обнаружить какие-то закономерности, и поэтому мало оснований строить какую-то теорию. Однако если обратить внимание на последовательность большого числа одинаковых экспериментов, то обнаружится интересное явление. Если индивидуальные результаты опытов ведут себя очень «неправильно», то средние результаты обнаруживают некоторые закономерности.

Пусть n_A – число выпадений герба после проведения n испытаний. Рассмотрим отношение $\frac{n_A}{n}$, называемое относительной частотой выпадения герба. Обозначим ее $P^*(A)$:

$$P^*(A) = \frac{n_A}{n}.$$

Многократно проводились опыты, их результаты приведены в табл.16.1.

Таблица 16.1

Число бросаний n (экспериментатор)	n_A	$\frac{n_A}{n}$
4040 (Бюффон, 18 век)	2048	0,5069
12000 (Пирсон)	6019	0,5016
24000 (Пирсон)	12012	0,5005

Видим, что с ростом n частота $\frac{n_A}{n}$ стремится к 0,5. Это свойство называют *статистической устойчивостью*, а сам предел $\lim\limits_{n\to\infty} P^*(A) = \lim\limits_{n\to\infty} \frac{n_A}{n}$ – *статистической вероятностью*. Массовые случайные явления, как правило, обладают этим свойством. Относительная частота события $P^*(A)$ обладает также следующими свойствами:

1. $0 \le P^*(A) \le 1$.
2. $P^*(\varnothing) = 0$.
3. $P^*(\Omega) = 1$.
4. $P^*(A+B) = P^*(A) + P^*(B)$

в том случае, если события A и B не могут произойти одновременно.

1.2. Пространство элементарных событий. Случайные события, операции над событиями и отношения между ними

При построении любой математической теории прежде всего следует договориться об определении некоторых понятий и о принятии некоторых аксиом. Для математического описания испытаний нам потребуется понятие пространства элементарных событий, которое обозначается буквой греческого алфавита Ω (омега). Заметим, что иногда вместо словосочетания «элементарные события» используется термин «элементарные исходы».

Определение. Пространство элементарных событий есть множество, элементами которого являются исходы, наблюдаемые в данном эксперименте (опыте).

Элементарные исходы обозначаются буквой ω, т.е. $\Omega = \{\omega\}$. Рассмотрим несколько примеров.

Пример 1 (подбрасывание монеты). Элементарными событиями являются: выпадение «герба» ($\omega_1 = \text{Г}$) и выпадение «решки» ($\omega_2 = \text{Р}$). Поэтому $\Omega = \{\text{Г}, \text{Р}\}$.

Пример 2 (монета подбрасывается два раза).

В этом случае $\Omega = \{\text{ГГ}, \text{ГР}, \text{РГ}, \text{РР}\}$.

Пример 3 (подбрасывание игральной кости).

Пусть $\omega_i = \{$на верхней грани выпало i очков$\}$, $i = \overline{1, 6}$. Тогда $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$.

Описав понятие пространства элементарных событий (исходов), зададимся вопросом: что понимать под событием A? Пусть, например, событие A={выпадение на верхней грани четного числа очков} (см. пример 3). Ясно, что каждое элементарное событие $\omega_2, \omega_4, \omega_6$ дает нам полную информацию о том, произошло событие A или не произошло. Множество $\{\omega_2, \omega_4, \omega_6\}$ является подмножеством Ω. При этом, если ω принадлежит этому множеству, то можно утверждать, что событие A произошло. Если же ω не принадлежит множеству $\{\omega_2, \omega_4, \omega_6\}$, то событие A не произошло. Элементы множества $\{\omega_2, \omega_4, \omega_6\}$ называются элементарными исходами, благоприятствующими событию A.

Определение. *Случайным событием* называется любое подмножество пространства элементарных событий Ω.

Принято все пространство элементарных исходов Ω тоже включать в число событий, наблюдаемых в данном эксперименте. Так как Ω обязательно происходит в данном эксперименте, то оно формально отождествляется с достоверным событием. В свою очередь невозможное событие трактуется как пустое множество, не содержащее ни одного из элементарных исходов, и поэтому обозначается символом $\varnothing$.

Если из наступления события A необходимо следует наступление события B, то A называется частным случаем B (обозначается $A \subset B$). Событие A есть частный случай события B тогда и только тогда, когда каждый исход, благоприятствующий событию A, будет таковым же для события B.

Пример 4. Рассматривается стрельба по мишени. Пусть события: A={попадание в малый круг}; B={попадание в большой круг}. Тогда $A \subset B$.

По определению полагают, что для любого события A имеет место: $\varnothing \subset A$, $A \subset \Omega$. Если же одновременно $A \subset B$ и $B \subset A$, то $A = B$. Другими словами, события A и B называются равными, если осуществление события A влечет за собой осуществление события B, и наоборот.

Определение. *Суммой (объединением)* событий A и B называется такое событие C, которое означает наступление хотя бы одного из событий A или B (обозначение $C = A + B$). Событие C состоит из тех элементарных исходов ω, которые благоприятствуют хотя бы одному из событий A или B.

Пример 5. Рассмотрим участок электрической цепи MN, содержащий два последовательно включенных элемента $Э_1$, $Э_2$ (рис. 16.1).

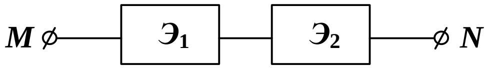


Рис. 16.1

Пусть событие A_i={не работает i-й элемент $Э_i$}; A={не работает участок MN}. Тогда $A = A_1 + A_2$.

Определение. *Произведением (пересечением)* событий A и B называется такое событие C, которое означает одновременное наступление событий A и B в одном опыте (обозначение: $C = A \cdot B$). Событие C состоит из тех элементарных исходов, которые благоприятствуют сразу обоим событиям A и B.

Так, если в примере 5 элементы $Э_1$, $Э_2$ включены параллельно, то $A = A_1 \cdot A_2$ (рис. 16.2).

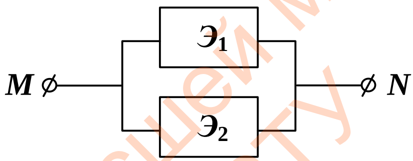


Рис. 16.2

Определение. События A и B называются *несовместными,* если наступление одного из них в данном эксперименте исключает наступление другого, т.е. $A \cdot B = \varnothing$.

Пример 6. При стрельбе по мишени события: A={попадание в мишень}; B={промах} являются несовместными.

В примере 5 события A_1 и A_2 – совместные.

Определение. Несколько событий образуют *полную группу событий,* если они попарно несовместны и в результате испытания происходит только одно из них.

Определение. *Разностью* событий A и B называется событие $C = A - B$, происходящее тогда и только тогда, когда происходит событие A, но не происходит событие B.

Определение. Если A – случайное событие, то противоположным событием (его обозначают $\overline{A}$) называется событие, состоящее в том, что A не происходит. Событие $\overline{A}$

состоит из всех тех элементов Ω, которые не благоприятствуют A.

Пример 7. В условиях примера 3 обозначим A={выпадение на верней грани четного числа очков}, $\overline{A}$={выпадение на верхней грани нечетного числа очков}.
$A=\{\omega_2,\omega_4,\omega_6\}$, $\overline{A}=\{\omega_1,\omega_3,\omega_5\}$, $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}$.

Справедливы следующие равенства, характеризующие свойства операций над событиями:

1. Коммутативность:
$$A+B=B+A,\quad A\cdot B=B\cdot A.$$

2. Ассоциативность:
$$A+(B+C)=(A+B)+C,\quad A\cdot(B\cdot C)=(A\cdot B)\cdot C.$$

3. Дистрибутивность:
$$A+(B\cdot C)=(A+B)\cdot(A+C),\quad A\cdot(B+C)=(A\cdot B)+(A\cdot C).$$

4. Идемпотентность:
$$A+A=A,\quad A\cdot A=A.$$

5. Законы де Моргана:
$$\overline{A+B}=\overline{A}\cdot\overline{B},\quad \overline{A\cdot B}=\overline{A}+\overline{B}.$$

6. Закон двойного отрицания:
$$\overline{\overline{A}}=A.$$

7. Закон исключенного третьего:
$$A+\overline{A}=\Omega.$$

8. Закон противоречия:
$$A\cdot\overline{A}=\varnothing.$$

9. Свойства достоверного и невозможного событий:
$$A+\Omega=\Omega,\quad A\cdot\Omega=A,\quad A+\varnothing=A,\quad A\cdot\varnothing=\varnothing,$$
$$\overline{\Omega}=\varnothing,\quad \overline{\varnothing}=\Omega.$$

10. Свойства разности событий:
$$A-B=A\cdot\overline{B},\quad \Omega-A=\overline{A},\quad A-\Omega=\varnothing,$$
$$A-\varnothing=A,\quad \varnothing-A=\varnothing,\quad A-A=\varnothing.$$

С математической точки зрения пространство (совокупность) всех элементарных событий, возможных в опыте, – это некоторое множество, а элементарные события – го элементы. Однако в теории вероятностей для обозначения используемых понятий по традиции применяются свои термины, отличающиеся от терминов теории множеств. В табл. 16.2 установлено соответствие между терминологическими рядами этих двух математических дисциплин.

Таблица 16.2

Теория вероятностей	**Теория множеств**
Пространство элементарных событий	Множество
Элементарное событие	Элемент этого множества
Событие	Подмножество
Достоверное событие	Подмножество, совпадающее с множеством
Невозможное событие	Пустое подмножество $\varnothing$
Сумма $A+B$ событий A и B	Объединение $A \bigcup B$
Произведение $A \cdot B$ событий A и B	Пересечение $A \bigcap B$
Событие, противоположное событию A	Дополнение $\overline{A}$
События A и B несовместны	$A \bigcap B = \varnothing$ пусто
События A и B совместны	$A \bigcap B \neq \varnothing$ не пусто

Эти два параллельных терминологических ряда сложились потому, что основные понятия теории вероятностей и её терминология сформировались в XVII-XVIII вв., а теория множеств возникла в конце XIX в. независимо от теории вероятностей и получила распространение в XX в.

1.3. Классическая схема и определение вероятности. Геометрические вероятности

Предположим, что из некоторых соображений, основанных на экспериментальных данных, или из соображений симметрии известно, что исходы опыта равновозможны. Предположим также, что пространство элементарных событий содержит конечное число элементов

$$\Omega = \{\omega_1, \omega_2, ..., \omega_n\}.$$

Пусть A – случайное событие, которому благоприятствуют m элементарных исходов (т.е. m есть число элементов в A).

Определение. Вероятностью $P(A)$ события A называется отношение числа m элементарных исходов, благоприятствующих этому событию, к общему числу n элементарных исходов:

$$P(A) = \frac{m}{n}. \tag{16.1}$$

Из приведенного определения следуют свойства:

1) для любого события A верно $0 \le P(A) \le 1$;
2) $P(\varnothing) = 0$ (вероятность невозможного события равна 0);
3) $P(\Omega) = 1$ (вероятность достоверного события равна 1);
4) для несовместных событий $P(A + B) = P(A) + P(B)$.

В простых ситуациях, когда число исходов опыта невелико, применение формулы (16.1) для решения не представляет никаких затруднений.

Пример 8. Монета подбрасывается дважды. Найти вероятность того, что герб выпадет хотя бы один раз.

Решение. Пространство Ω состоит из четырех равновозможных исходов: $\Omega = \{\text{ГГ, ГР, РГ, РР}\}$.

Событие A={герб выпадает хотя бы один раз} состоит из трех исходов (элементов): $A = \{\text{ГГ, ГР, РГ, РР}\}$. Поэтому

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

Пример 9. Из колоды в 36 карт вынимается одна карта. Какова вероятность появления карты пиковой масти?

Решение. В данном случае пространство Ω состоит из 36 равновозможных исходов, т.е. $n = 36$. Число благоприятствующих исходов событию A={появляется карта пиковой масти} равно 9, т.е. $m = 9$. Следовательно,

$$P(A) = \frac{9}{36} = \frac{1}{4}.$$

Пусть число опытов бесконечно, т.е. Ω – непрерывное пространство элементарных событий. Предположим, что пространством элементарных событий служит некоторое множество точек области Ω на плоскости, имеющих площадь соответственно. Элементарное событие представляется точкой (x, y), попадающей в область Ω. Будем считать, что случайная точка (x, y) равновозможно попадает в любую точку Ω; вероятность её попадания в любую часть $A \subset \Omega$ пропорциональна площади $S(A)$ и не зависит от ее формы и расположения. Тогда вероятность события A определяется формулой:

$$P(A) = \frac{S(A)}{S(\Omega)}. \tag{16.2а}$$

Пример 10. В круг радиусом R вписан правильный треугольник со стороной a. Найти вероятность того, что точка, брошенная в этот круг, попадет в данный треугольник (рис.16.3).

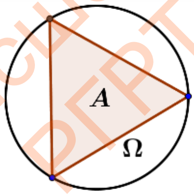


Рис. 16.3

Решение.

$$P(A) = \frac{S(A)}{S(\Omega)} = \left| \begin{array}{l} a = R\sqrt{3} \\ S(A) = \dfrac{\sqrt{3}a^2}{4} \end{array} \right| = \frac{3\sqrt{3}R^2}{4\pi R^2} \approx 0{,}41.$$

Геометрическое определение вероятности события применимо и в том случае, когда области Ω и A обе линейные или объемные. В первом случае

$$P(A) = \frac{\ell(A)}{\ell(\Omega)}, \tag{16.2б}$$

во втором

$$P(A)=\frac{V(A)}{V(\Omega)}. \qquad (16.2в)$$

Все три формулы (16.2а), (16.2б), (16.2в) можно записать в виде

$$P(A)=\frac{mes(A)}{mes(\Omega)}, \qquad (16.2)$$

где через mes обозначена мера множества, т.е. длина, или площадь, или объем.

Пример 11. Известно, что через данную станцию метро каждые 5 минут проходит электропоезд. Пассажир, не знающий расписания движения электропоездов, приходит на станцию. Найти вероятность того, что он будет ожидать электропоезд не менее 4 минут.

Решение. Пространство элементарных событий Ω представляет собой отрезок $[0;5]$. Пусть x – случайное время прихода пассажира; $0\le x\le 5$. Событие $A=${пассажир ожидает не менее 4 минут} есть множество точек из Ω, удовлетворяющих неравенству $5-x\ge 4$, откуда $0\le x\le 1$.

Следовательно, $S(A)=\ell[0;1]=1$, $S(\Omega)=\ell[0;5]=5$, поэтому $P(A)=\frac{1}{5}=0{,}2$.

1.4. Теоремы умножения и сложения вероятностей

Говорить о вероятности события $P(A)$ как о мере наступления (осуществления) случайного события A имеет смысл только при осуществлении определенного комплекса условий. При изменении этих условий изменится и вероятность события A. Так, если к комплексу условий, при котором изучалась вероятность $P(A)$, добавить новое условие, состоящее в появлении события B, то получим другое значение вероятности $P(A/B)$ – *условную вероятность события* A при условии, что произойдет событие B. Вероятность $P(A)$ в отличие от условной называют *безусловной*.

Пример. Дважды бросается игральная кость. Пусть событие A={при первом бросании выпадает 1}, событие B={сумма очков на верхней грани меньше 4}. Найти $P(A/B)$ и $P(B/A)$.

Решение. Событие B произошло, если произошел один из трех следующих элементарных исходов: $B=\{(1,1);(1,2);(2,1)\}$. Из них событию A благоприятствуют только два исхода: $(1,1)$ и $(1,2)$. Следовательно, $P(A/B)=\dfrac{2}{3}$.

Подсчитаем теперь условную вероятность $P(B/A)$. Событие A произошло, если произошел один из шести следующих элементарных исходов.

$$A=\{(1,1);(1,2);(1,3);(1,4);(1,5);(1,6)\}.$$

Среди них лишь два исхода – $(1,1)$ и $(1,2)$ благоприятствуют событию B. Поэтому $P(B/A)=\dfrac{2}{6}=\dfrac{1}{3}$.

Видим, что в конкретных случаях условная вероятность $P(A/B)$ вычисляется теми же способами, что и безусловная, с той разницей, что при вычислении условной вероятности пространством элементарных событий следует считать множество элементарных исходов, благоприятствующих событию B.

Теорема (умножения вероятностей). Вероятность произведения двух событий равна вероятности одного из них, умноженной на условную вероятность второго события при условии, что произошло первое:

$$P(A/B)=P(B)\cdot P(A/B)=P(A)\cdot P(B/A). \qquad (16.3)$$

Доказательство. Выведем формулу для вычисления условной вероятности. Пусть событиям A и B благоприятствуют соответственно m и k элементарных исходов из n. Тогда по формуле классической вероятности: $P(A)=\dfrac{m}{n}$ и $P(B)=\dfrac{k}{n}$.

Пусть событию A при условии, что событие B произошло, благоприятствуют r элементарных исходов (рис. 16.4). Тогда условная вероятность события A равна

$$P(A/B)=\frac{r}{k}.$$

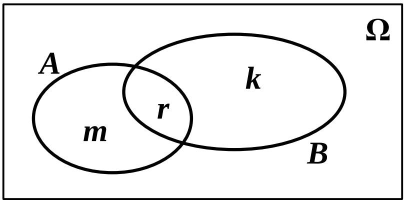


Рис. 16.4

Поделив числитель и знаменатель правой части формулы на n, получим

$$P(A/B)=\frac{r/n}{k/n}=\frac{P(A\cdot B)}{P(B)}.$$

Дело в том, что событию $A\cdot B$ благоприятствует r исходов и, следовательно, $\frac{r}{n}$ – его безусловная вероятность.

Из последней формулы для условной вероятности следует формула для вероятности произведения событий:

$$P(A\cdot B)=P(B)\cdot P(A/B)=P(A)\cdot P(B/A). \qquad (16.4)$$

Пример. Дважды бросается игральная кость. Пусть событие $A=\{$при первом бросании выпадает 1$\}$, событие $B=\{$сумма очков на верхней грани меньше 4$\}$. Найти $P(A\cdot B)$.

Решение. $P(A\cdot B)=P(B)\cdot P(A/B)=\frac{3}{36}\cdot\frac{2}{3}=\frac{1}{18}$;

$P(A\cdot B)=P(A)\cdot P(B/A)=\frac{1}{6}\cdot\frac{1}{3}=\frac{1}{18}$.

Определение. События A и B называются *независимыми*, если вероятность появления события A не зависит от того, произошло событие B или нет, и наоборот.

Таким образом, $P(A/B)=P(A)$ и $P(B/A)=P(B)$.

Для независимых событий, как следует из (16.4),

$$P(A\cdot B)=P(A)\cdot P(B). \qquad (16.4a)$$

Для того чтобы обобщить теорему умножения на несколько событий, вводят понятие *независимости событий в совокупности.*

Определение. Несколько событий A_1, A_2, …, A_n называются *независимыми в совокупности*, если каждое из них и любая комбинация остальных (содержащая либо все остальные события, либо часть из них) есть независимые события.

Для событий A_1, …, A_n, независимых в совокупности, справедлива формула

$$P(A_1 \cdot A_2 \cdot \ldots \cdot A_n) = P(A_1) \cdot P(A_2) \cdot \ldots \cdot P(A_n). \qquad (16.4б)$$

Рассмотрим вопрос о нахождении вероятности $P(A+B)$ суммы событий $A+B$. Предположим, что A и B – совместные события.

Теорема (сложения вероятностей). Вероятность суммы двух совместных событий равна сумме их вероятностей без вероятности их произведения:

$$P(A+B) = P(A) + P(B) - P(A \cdot B). \qquad (16.5)$$

Доказательство. Докажем справедливость этой формулы, используя понятие геометрической вероятности.

Введем в рассмотрение событие A_1 (заштриховано на рис. 16.5). При этом $A+B = A_1 + B$. Так как события A_1 и B несовместны ($A_1 \cdot B = \varnothing$), то для них по 4-му свойству вероятностей находим: $P(A+B) = P(A_1) + P(B)$. Далее $A = A_1 + A \cdot B$ и $A_1 \cdot (AB) = \varnothing$. Следовательно, $P(A) = P(A_1) + P(A \cdot B)$, откуда:

$$P(A_1) = P(A) - P(A \cdot B)$$

и $P(A+B) = P(A) + P(B) - P(A \cdot B)$.

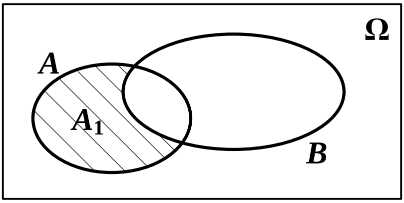


Рис. 16.5

Если же события A и B являются несовместными, то следует использовать формулу

$$P(A+B)=P(A)+P(B). \qquad (16.5а)$$

Замечание. Если имеется n попарно несовместных событий A_1, A_2, …, A_n, то

$$P(A_1+A_2+\ldots+A_n)=P(A_1)+P(A_2)+\ldots+P(A_n). \qquad (16.5б)$$

Пример. Электрическая цепь между точками M и N составлена по следующей схеме (рис. 16.6).

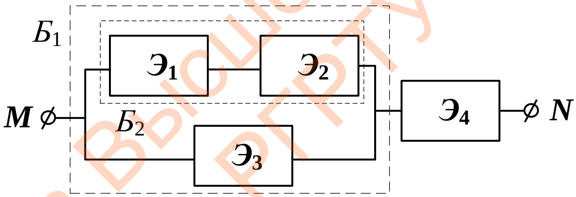


Рис. 16.6

Выход из строя за время T различных элементов цепи – независимые в совокупности события, имеющие вероятности p_1, p_2, p_3, p_4 соответственно. Отказ какого-то элемента приводит к разрыву той ветви цепи, в которой он установлен. Найти вероятность разрыва цепи за установленное время.

Решение. Введем в рассмотрение следующие события:

A={разрыв цепи за установленное время};

B_i ={ выход из строя блока $Б_i$, i=1, 2};

A_i ={выход из строя элемента $Э_i$, $i=\overline{1,4}$}, причем $P(A_i)=p_i, i=\overline{1,4}$.

Так как блок $Б_1$ и элемент $Э_4$ соединены последовательно, то $A = B_1 + A_4$ и по теореме сложения вероятностей

$$P(A) = P(B_1) + P(A_4) - P(B_1 \cdot A_4) =$$
$$= P(B_1) + p_4 - P(B_1) \cdot p_4 = (1 - p_4) \cdot P(B_1) + p_4 . \qquad (16.6)$$

В блоке $Б_1$ элемент $Э_3$ и блок $Б_2$ соединены параллельно. Следовательно, $B_1 = B_2 \cdot A_3$ и в силу независимости событий B_2 и A_3 имеем:

$$P(B_1) = P(B_2) \cdot P(A_3) = P(B_2) \cdot p_3 . \qquad (16.7)$$

Аналогично для блока $Б_2$ получаем $B_2 = A_1 \cdot A_2$ и

$$P(B_2) = P(A_1) + P(A_2) - P(A_1) \cdot P(A_2) =$$
$$= p_1 + p_2 - p_1 \cdot p_2 . \qquad (16.8)$$

Подставляя (16.8) в (16.7), а затем (16.7) в (16.6), находим

$$P(B_1) = (p_1 + p_2 - p_1 p_2) \cdot p_3 ,$$
$$P(A) = (1 - p_4) \cdot (p_1 + p_2 - p_1 p_2) \cdot p_3 + p_4 .$$

1.5. Теорема о полной вероятности. Формула Байеса

Предположим, что событие A наступает при условии появления одного из попарно несовместных событий H_1, H_2, …, H_n. Эти события образуют полную группу событий, т.е. $H_i \cdot H_j = \varnothing$ при $\forall i \neq j$ и $H_1 + H_2 + \ldots + H_n = \Omega$.

Обычно события H_1, H_2, …, H_n называются *гипотезами.* Если известны вероятности гипотез $P(H_1)$, $P(H_2)$, …, и условные вероятности события A при каждой гипотезе $P(A/H_1)$, $P(A/H_2)$, …, $P(A/H_n)$, то справедлива следующая теорема.

Теорема (о полной вероятности). При сделанных выше предположениях относительно события A и гипотез H_1, H_2, …, H_n полная вероятность события A равна

$$P(A)=P(H_1)\cdot P(A/H_1)+P(H_2)\cdot P(A/H_2)+\ldots+P(H_n)\cdot P(A/H_n)$$

$$\text{или } P(A)=\sum_{i=1}^{n}P(H_i)\cdot P(A/H_i). \qquad (16.9)$$

Формула (16.9) называется *формулой полной вероятности*.

Доказательство. Так как событие A может наступить при условии появления одного из попарно несовместных событий H_1, H_2, …, H_n, то исходы, благоприятствующие событию A, содержатся в попарно несовместных событиях $A\cdot H_1$, $A\cdot H_2$, …, $A\cdot H_n$, так как

$$A=A\cdot\Omega=A\cdot(H_1+\ldots+H_n)=A\cdot H_1+\ldots+A\cdot H_n.$$

Воспользовавшись теоремой сложения вероятностей для несовместных событий, найдем:

$$P(A)=P(A\cdot H_1+\ldots+A\cdot H_n)=P(A\cdot H_1)+\ldots+P(A\cdot H_n).$$

По теореме умножения вероятностей имеем для $\forall i=\overline{1,n}$ $P(A\cdot H_i)=P(H_i)\cdot P(A/H_i)$. Отсюда получим:

$$P(A)=P(H_1)\cdot P(A/H_1)+\ldots+P(H_n)\cdot P(A/H_n)=\sum_{i=1}^{n}P(H_i)\cdot P(A/H_i).$$

Пример. Два стрелка произвели по одному выстрелу по цели. Вероятность попадания в цель первым стрелком равна 0,8; а вторым – 0,4. Найти вероятность того, что цель поражена одной пулей.

Решение. Относительно результатов опыта (стрельба по цели) можно высказать следующие гипотезы: H_1={оба стрелка не попали}, H_2={оба попали}; H_3={первый попал; второй – не попал}; H_4={первый не попал; второй – попал}.

Тогда вероятности этих гипотез равны:

$$P(H_1)=0{,}2\cdot 0{,}6=0{,}12;\; P(H_2)=0{,}8\cdot 0{,}4=0{,}32;$$

$$P(H_3)=0{,}8\cdot 0{,}6=0{,}48;\; P(H_4)=0{,}2\cdot 0{,}4=0{,}08;$$

$$P(H_1)+P(H_2)+P(H_3)+P(H_4)=0{,}12+0{,}32+0{,}48+0{,}08=1.$$

По условию задачи находим условные вероятности события A={цель поражена одной пулей}:

$$P(A/H_1)=0;\ P(A/H_2)=0;\ P(A/H_3)=1;\ P(A/H_4)=1.$$

Следовательно, полная вероятность события A равна:

$$P(A)=0{,}12\cdot 0+0{,}32\cdot 0+0{,}48\cdot 1+0{,}08\cdot 1=0{,}56.$$

Рассмотрим теперь такую ситуацию. Пусть по-прежнему известны *априорные* (доопытные) вероятности $P(H_1)$ и условные вероятности $P(A/H_i)$. Затем производится опыт, в результате которого установлено, что событие A произошло. Как этот факт влияет на оценку вероятностей гипотез? Таким образом, приходим к постановке следующей задачи.

Задача *(о переоценке гипотез).* Найти *апостериорные* (послеопытные) вероятности $P(H_1/A)$, $P(H_2/A)$, …, $P(H_n/A)$.

Решение. Действительно,

$$P(H_i/A)=\frac{P(A\cdot H_i)}{P(A)}=\frac{P(H_i)\cdot P(A/H_i)}{P(A)} \quad \text{или}$$

$$P(H_i/A)=\frac{P(H_i)\cdot P(A/H_i)}{\sum\limits_{i=1}^{n}P(H_i)\cdot P(A/H_i)}. \tag{16.10}$$

Формула (16.10) называется *формулой Байеса*. Имеют место соотношения:

$$P(H_1)+P(H_2)+P(H_3)+P(H_4)=1;$$

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

Первое из них является следствием того, что гипотезы H_i попарно несовместны и их сумма есть достоверное событие Ω. Второе следует из (16.10), так как

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

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

Пример. Пусть в условиях предыдущего примера произошло событие A. Найти апостериорные вероятности гипотез.

Решение. С учетом того, что $P(A/H_1)=P(A/H_2)=0$ и $P(A/H_3)=P(A/H_4)=1$, используя формулу (16.10), находим:

$$P(H_1/A)=0;\ P(H_2/A)=0;$$

$$P(H_3/A)=\frac{P(H_3)\cdot P(A/H_3)}{P(A)}=\frac{0{,}48}{0{,}56}=\frac{6}{7}\approx 0{,}86;$$

$$P(H_4/A)=\frac{P(H_4)\cdot P(A/H_4)}{P(A)}=\frac{0{,}08}{0{,}56}=\frac{1}{7}\approx 0{,}14.$$

Напомним, что априорные вероятности гипотез были следующими: $P(H_1)=0{,}12$; $P(H_2)=0{,}32$; $P(H_3)=0{,}48$; $P(H_4)=0{,}08$.

1.6. Последовательность независимых испытаний

1.6.1. Схема и формула Бернулли

Предположим, что проводится серия из n независимых одинаковых испытаний. Испытания называются независимыми в том случае, если результаты любой группы испытаний не влияют на вероятности любых событий в других испытаниях.

В общем случае в результате испытания может наступить один из r исходов ($r \geq 2$). Будем считать, что в каждом испытании (опыте) этой серии наблюдается наступление (или ненаступление) одного и того же события A. Если в некотором опыте событие A произошло, то говорят, что в этом опыте имел место «успех»; в противном случае говорят, что была «неудача». Пусть вероятность успеха в *каждом* опыте равна p, а соответственно вероятность неудачи $q=1-p$. Такие условия испытаний называют *схемой Бернулли*.

Задача. Найти вероятность того, что в серии из n испытаний Бернулли произойдет *ровно* m *успехов.*

Эта вероятность обозначается обычно через $P_n(m)$.

Элементарный исход в данном эксперименте (который состоит в проведении одной серии из n независимых испытаний) удобно обозначить набором из n символов:

$$(x_1, x_2, \dots, x_n), \tag{16.11}$$

каждый из которых может быть нулем или единицей. При этом если на каком-то месте стоит единица, то это означает, что в испытании с этим номером был отмечен успех, а если стоит нуль, то неудача. Например, если серия состоит из трех испытаний, то исход (101) означает, что в первом и третьем испытаниях был успех, а во втором – неудача.

Ввиду независимости испытаний вероятность каждого исхода типа (16.11) равна произведению вероятностей

$$\underbrace{p \cdot p \cdot \ldots \cdot p}_{m \text{ раз}} \cdot \underbrace{q \cdot q \cdot \ldots \cdot q}_{n \text{ раз}} = p^m \cdot q^n, \qquad (16.12)$$

где m – число единиц в данном исходе. Событию {произошло ровно m успехов} благоприятствуют все те исходы типа (16.11), которые содержат ровно m единиц. Вероятность каждого такого исхода вычисляется по формуле (16.12), а их число равно числу сочетаний из n символов (общее число позиций в исходах) по m символов (те позиции, на которых стоят единицы). Поэтому искомая вероятность вычисляется по формуле

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m}, \qquad (16.13)$$

которая называется *формулой Бернулли*.

Замечание. Если воспользоваться формулой бинома Ньютона $(p+q)^n = \sum_{m=0}^{n} C_n^m p^m \cdot q^{n-m}$, то $\sum_{m=0}^{n} P_n(m) = (p+q)^n = 1^n = 1$.

Поэтому совокупность вероятностей $P_n(m)$ называют *биномиальным законом* распределения вероятностей.

Пример. Что вероятнее выиграть у равносильного противника: 3 партии из 5 или 4 из 7?

Решение. По формуле Бернулли находим

$$P_5(3) = C_5^3 \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^{5-3} = \frac{5!}{3! \cdot 2!} \cdot \frac{1}{2^5} = \frac{4 \cdot 5}{2} \cdot \frac{1}{32} = \frac{10}{32} = \frac{40}{128},$$

$$P_7(4) = C_7^4 \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^{7-4} = \frac{7!}{4! \cdot 3!} \cdot \frac{1}{2^7} = \frac{5 \cdot 6 \cdot 7}{3!} \cdot \frac{1}{128} = \frac{35}{128}.$$

Отсюда видно, что $P_5(3) > P_7(4)$.

1.6.2. Предельные теоремы в схеме Бернулли: локальная и интегральная теоремы Муавра – Лапласа, теорема Пуассона

Если число испытаний n велико, то использовать формулу Бернулли становится трудно: приходится сталкиваться с вычислением факториалов больших чисел. Как вычислить искомую вероятность, не прибегая к формуле Бернулли? Ответ на этот вопрос дает следующая теорема.

Теорема (локальная теорема Муавра – Лапласа).

Если вероятность успеха p в каждом испытании постоянна и больше нуля, то

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x), \tag{16.14}$$

где $x = \dfrac{m - np}{\sqrt{npq}}$, $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ – нормированная функция Гаусса.

Замечания. 1. Функция $\varphi(x)$ – четная, т.е. $\varphi(x) = \varphi(-x)$. Её значения находят по таблицам (см. приложение 1).

2. При $x > 3$ функция $\varphi(x)$ практически равна нулю.

3. Формула (16.14) справедлива при $npq > 10$.

Пример. Найти вероятность попадания в цель ровно 80 раз при 400 выстрелах, если вероятность попадания при одном выстреле равна 0,2.

Решение. В данном случае $n = 400$, $m = 80$, $p = 0{,}2$; $q = 1 - p = 0{,}8$; $npq = 400 \cdot 0{,}2 \cdot 0{,}8 = 64$, $\sqrt{npq} = 8$. Находим:

$$x = \frac{80 - 400 \cdot 0{,}2}{8} = 0, \quad \varphi(x) = 0{,}3989,$$

$$P_{400}(80) \approx \frac{1}{8} \cdot 0{,}3989 = 0{,}04980.$$

Непосредственный подсчет по формуле Бернулли дает следующий результат: $P_{400}(80) = 0{,}04980$.

Для вычисления при больших n вероятности $P_n(m_1 \le m \le m_2)$ того, что число успехов m в n испытаниях Бернулли находится между m_1 и m_2, используется интегральная теорема Муавра – Лапласа.

Теорема (интегральная теорема Муавра – Лапласа). Если вероятность успеха p в каждом испытании постоянна и больше нуля, то

$$P_n(m_1 \le m \le m_2) = \Phi(x'') - \Phi(x'), \tag{16.15}$$

где $x' = \dfrac{m_1 - np}{\sqrt{npq}}$, $x'' = \dfrac{m_2 - np}{\sqrt{npq}}$, $\Phi(x)$ – функция Лапласа (интеграл вероятностей):

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

Замечание. Функция $\Phi(x)$ – нечетная, т.е. $\Phi(-x) = -\Phi(x)$. При $x > 5$ можно считать, что $\Phi(x) \approx 0{,}5$. Функция Лапласа $\Phi(x)$ табулирована (см. приложение 2).

Графики функций $\varphi(x)$, $\Phi(x)$ приведены на рис. 16.7, 16.8.

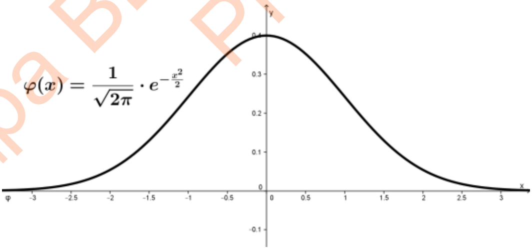


Рис. 16.7

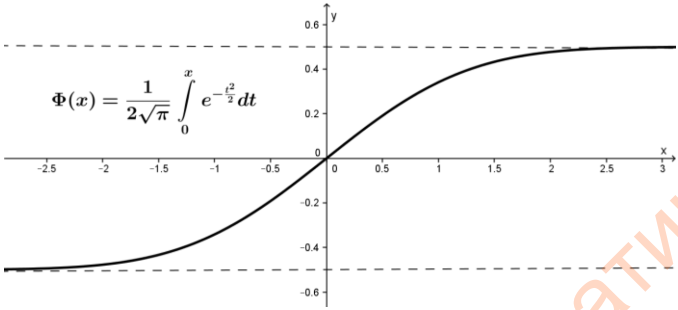


Рис. 16.8

Пример. По каналу связи передано 100 знаков. Вероятность искажения знака помехами равна 0,1. Действие помех на каждый знак происходит независимо. Какова вероятность того, что при передаче будет не более 15 искажений.

Решение. Искомую вероятность находим по формуле (16.15). По условию задачи: $n = 100$, $m_1 = 0$, $m_2 = 15$, $p = 0{,}1$, $q = 0{,}9$. Далее вычисляем:

$$\sqrt{npq} = \sqrt{100 \cdot 0{,}1 \cdot 0{,}9} = \sqrt{9} = 3, \quad \frac{1}{\sqrt{npq}} = \frac{1}{3},$$

$$x' = \frac{0 - 10}{\sqrt{npq}} = -\frac{10}{3} \approx -3{,}333, \quad x'' = \frac{15 - 10}{3} = \frac{5}{3} \approx 1{,}667.$$

Находим по таблице:

$$\Phi(x') = \Phi(-3{,}333) = -\Phi(3{,}333) = -0{,}4995;$$

$$\Phi(x'') = \Phi(1{,}667) = 0{,}4525.$$

В итоге $P_{100}(0 \le m \le 15) = 0{,}4525 - (-0{,}4995) = 0{,}952$.

Предположим теперь, что число испытаний n велико, а вероятность p появления успеха в одном испытании мала. Уточним понятие такого редкого события. В схеме Бернулли мы заранее фиксировали число испытаний n и вероятность p наступления события в одном испытании.

Рассмотрим теперь серии с различными n. Теперь при различных n вероятность p может быть разной. Поэтому вместо p будем писать p_n, вместо q пишем q_n. Положим далее $\lambda_n = p_n \cdot n$. При этом, исходя из определения вероятности события, получаем смысл величины λ_n – среднее число появления события в серии из n испытаний.

Будем в дальнейшем рассматривать серии испытаний, в которых λ_n стремится к λ при $n\to\infty$, т.е. $\lim\limits_{n\to\infty}\lambda_n = \lambda$. Это означает, что с ростом длины n серии испытаний среднее число наступления события (успеха) в серии стабилизируется, стремится к некоторому пределу λ.

Теорема (Пуассона). Пусть p_n – вероятность наступления события в каждом испытании серии из n испытаний Бернулли. Если $\lambda_n = p_n \cdot n \to \lambda$ при $n\to\infty$ и $0<\lambda<\infty$, то

$$P_n(m) \approx \frac{\lambda^m}{m!}e^{-\lambda}. \qquad (16.16)$$

Замечание. На практике число испытаний n конечно, и трудно проверить, выполняется ли условие $p_n \cdot n \to \lambda$ при $n\to\infty$. Однако из этого условия следует, что $p_n = \dfrac{\lambda}{n} \to 0$ при $n\to\infty$. Поэтому, если в схеме Бернулли n – велико, а p – мало, то в качестве значения параметра λ берут $\lambda = n\cdot p$.

Доказательство. Рассмотрим

$$\lim_{n\to\infty}P_n(m) = \lim_{n\to\infty}C_n^m p^m(1-p)^{n-m} = \lim_{n\to\infty}\frac{A_n^m}{P_m}p^m(1-p)^{n-m} =$$

$$= \lim_{n\to\infty}\frac{n(n-1)\ldots[n-(m-1)]}{1\cdot 2 \ldots \cdot m}\left(\frac{\lambda}{n}\right)^m\left(1-\frac{\lambda}{n}\right)^{n-m} =$$

$$= \frac{\lambda^m}{m!}\lim_{n\to\infty}\frac{n(n-1)\ldots[n-(m-1)]}{\underbrace{n\cdot n\cdot \ldots \cdot n}_{m\text{ раз}}}\cdot\left(1-\frac{\lambda}{n}\right)^{n-m} =$$

$$= \frac{\lambda^m}{m!}\lim_{n\to\infty}1\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\ldots\frac{[n-(m-1)]}{n}\cdot\left(1-\frac{\lambda}{n}\right)^{n-m} =$$

$$=\frac{\lambda^m}{m!}\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{m-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-m}=$$

$$=\frac{\lambda^m}{m!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n-m}=\frac{\lambda^m}{m!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n}\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-m}=\frac{\lambda^m}{m!}\cdot e^{-\lambda}.$$

Пример. Коммутатор обслуживает 100 абонентов. Вероятность того, что в течение 1 минуты абонент позвонит на коммутатор, равна 0,01. Найти вероятность того, что в течение 1 минуты позвонят 2 абонента.

Решение. В нашем случае $n=100$, $m=2$, $p=0{,}01$; $\lambda=p\cdot n=1$. По формуле (16.16) находим:

$$P_{100}(2)=\frac{1^2}{2!}\cdot e^{-1}=\frac{1}{2e}\approx 0{,}1839\,.$$

Замечания. 1.Функция, используемая в теореме Пуассона,

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

табулирована и представляет собой таблицу с двумя входами (см. приложение 3).

2. Теорему Пуассона целесообразно использовать при $npq\le 10$.

2. Случайные величины

Понятие случайной величины – одно из основных в теории вероятностей. Случайная величина является числовой характеристикой результата эксперимента, т.е. числовой функцией, определенной на пространстве элементарных событий.

2.1. Определение случайной величины. Функция распределения случайной величины

Будем фиксировать число выпавших «гербов» при подбрасывании двух монет. В зависимости от исхода опыта это число может оказаться равным 0, 1 и 2. Построим математическую модель данного опыта. Выберем в качестве элементарных следующие события: ω_1=(Г,Г)={на первой монете выпал «герб», на второй – также «герб»}; ω_2=(Р,Г)={на первой монете выпала «решка», на второй – «герб»}; ω_3=(Г,Р)={на первой монете выпал «герб», на второй – «решка»}; ω_4=(Р,Р)={на обеих монетах выпала «решка»}. Число выпавших «гербов» X является функцией от элементарного события:

$$X(\omega_1)=X(\Gamma,\Gamma)=2;\quad X(\omega_2)=1;\quad X(\omega_3)=1;\quad X(\omega_4)=0,$$

т.е. каждому событию ставится в соответствие число $X(\omega)$.

Исход опыта является случайным, поэтому значение X является случайным. Функция $X=X(\omega_i)$ – случайная величина, определенная на множестве элементарных событий $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$. Можно сформулировать следующее определение.

***Определение* 1.** *Случайной величиной* (СВ) называется функция $X=X(\omega)$, определенная на множестве элементарных событий.

Существует равносильное определение случайной величины.

***Определение* 2.** *Случайной величиной* (СВ) называют величину, которая в результате опыта принимает одно и только одно значение, наперед неизвестное и зависящее от ряда случайных причин, которые заранее не могут быть учтены.

Случайные величины обозначаются прописными латинскими $X, Y, Z, \ldots$ или греческими $\xi, \eta, \zeta, \ldots$ буквами. Значения СВ обозначаются строчными латинским буквами $x, y, z, \ldots$ Случайные величины делятся на дискретные и непрерывные.

Определение. Случайная величина называется *дискретной* (ДСВ), если она принимает значения из некоторого дискретного конечного $x_1, x_2, \ldots, x_n$ или счетного множества чисел $x_1, x_2, \ldots$ При этом $P(X=x_i)=p_i\geq 0;\quad p_1+p_2+\ldots+p_n=1$.

Примеры: 1) число студентов, присутствующих на данном занятии; 2) число успехов в n испытаниях Бернулли; 3) число попаданий в самолет, достаточное для вывода его из строя (возможные значения 1, 2, 3, … n, …); 4) число заявок на АТС.

Определение. Случайная величина называется *непрерывной* (НСВ), если ее возможные значения не могут быть заранее перечислены и непрерывно заполняют некоторый промежуток.

Примеры: 1) время безотказной работы устройства; 2) ошибка измерителя высоты; 3) расстояние от точки попадания до центра мишени при выстреле.

Важной характеристикой СВ является *закон распределения* СВ. Под ним понимается всякое соответствие, устанавливающее связь между возможными значениями СВ и соответствующими им вероятностями. Про СВ будем говорить, что она *подчинена* данному закону распределения.

Закон распределения может быть задан таблично (ряд распределения); графически (многоугольник распределения), аналитически (с помощью формул).

Простейшей формой закона распределения для ДСВ является таблица, в которой перечислены возможные значения СВ и соответствующие им вероятности. Это *ряд распределения* (см. табл.16.3).

Таблица 16.3

$X = x_i$	x_1	x_2	$\cdots$	x_n
$P(X = x_i) = p_i$	p_1	p_2	$\cdots$	p_n

Чтобы придать ряду распределения более наглядный вид, прибегают к его графическому изображению (рис. 16.9). Это многоугольник распределения.

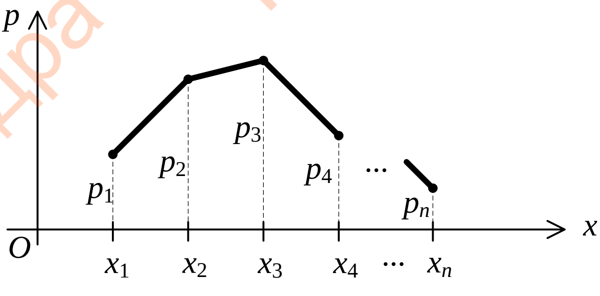


Рис. 16.9

Рассмотрим пример аналитического задания СВ. Пусть $X =$ {число успехов в n независимых испытаниях}:

$$P(X = m) = \frac{\lambda^m}{m!} e^{-\lambda}; \quad \lambda > 0; \quad m = 0, 1, 2, \ldots, n\,.$$

В данном случае закон распределения СВ описывается законом Пуассона.

Ряд распределения СВ является исчерпывающей характеристикой ДСВ. Однако в общем случае эта характеристика не является универсальной. Нетрудно убедиться, что для НСВ такой характеристики построить нельзя. Составить таблицу (ряд распределения), в которой были бы перечислены все возможные значения НСВ, невозможно. Поэтому для количественной характеристики СВ удобно воспользоваться не вероятностью события $(X = x_i)$, а вероятностью события $(X < x)$, где x – текущая переменная. Вероятность этого события, очевидно, зависит от x, т.е. является функцией от x. Эта функция называется *функцией распределения СВ* X и обозначается как $F(x)$. Итак, по определению

$$F(x) = P(X < x). \tag{16.17}$$

Геометрически (рис. 16.10) значение $F(x)$ равно вероятности попадания случайной точки X в промежуток $(-\infty; x)$.

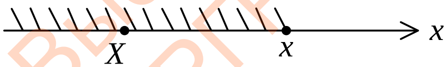


Рис. 16.10

Функция распределения существует для всех СВ: как дискретных, так и непрерывных, и характеризует СВ с вероятностной точки зрения, т.е. является одной из форм закона распределения.

2.2. Дискретная случайная величина

Пример 1. Построим функцию распределения СВ, равной числу выпадений «герба» при бросании двух монет. Ее ряд распределения имеет вид (табл.16.4):

Таблица 16.4

$X = x_i$	0	1	2
p_i	0,25	0,5	0,25

Будем строить $F(x)$, исходя из определения.

1. Пусть $x \le 0$. Например, $x = -0{,}5$, тогда $F(-0{,}5) = P(X < -0{,}5) = 0$, так как событие $(X < -0{,}5)$ – невозможное. Аналогично $F(0) = P(X < 0) = 0$ (рис. 16.11а).

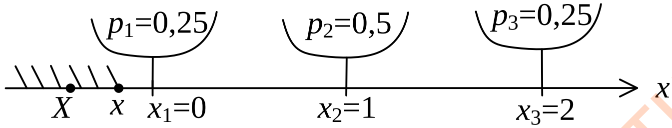


Рис. 16.11а

2. Пусть $0 < x \le 1$. Так, для $x = 0{,}5$ получим $F(0{,}5) = P(X < 0{,}5) = P(X = 0) = 0{,}25$ (рис. 16.11б).

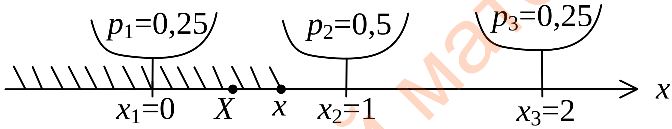


Рис. 16.11б

Легко убедиться, что $F(x) = 0{,}25$ в данном случае.

3. Пусть $1 < x \le 2$. Аналогичными рассуждениями находим $F(x) = 0{,}75$ (рис. 16.11в). Например,

$$F(1{,}5) = P(X < 1{,}5) = P(X = 0) + P(X = 1) = 0{,}25 + 0{,}5 = 0{,}75.$$

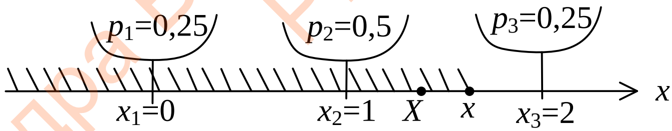


Рис. 16.11в

4. При $x > 2$ имеем $F(x) = 1$, так как в этом случае $F(x) = P(X < x) = 1$ как достоверное событие (рис. 16.11г).

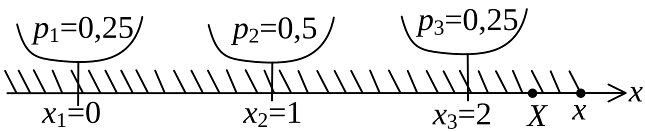


Рис. 16.11г

Таким образом,

$$F(x)=\begin{cases}0, & x\le 0;\\ 0{,}25, & 0<x\le 1;\\ 0{,}75, & 1<x\le 2;\\ 1, & x>2.\end{cases} \tag{16.18}$$

График функции $F(x)$ изображен на рис. 16.12.

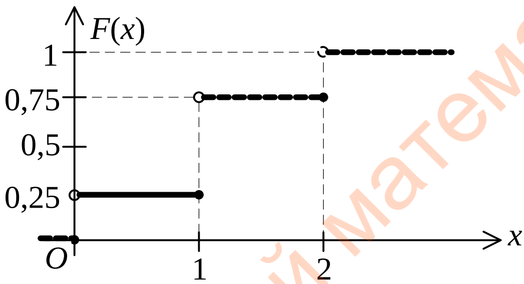


Рис. 16.12

Можно сделать выводы:

1) $F(x)$ постоянна на промежутках $(-\infty, x_1]$, $(x_1, x_2]$, …, $(x_n, +\infty)$;

2) в точках x_1, x_2, …, x_n функция $F(x)$ имеет скачок, равный вероятности того, что СВ примет это значение, т.е. p_i (см. рис. 16.12).

Свойства функции распределения

1. $0\le F(x)\le 1$.

2. $F(x_2)\ge F(x_1)$, при $x_2>x_1$, т.е. $F(x)$ – неубывающая функция (см. рис. 16.12).

Доказательство.

В самом деле, событие $(X<x_2)$ есть сумма двух несовместных событий $(X<x_1)$ и $(x_1\le X<x_2)$. По теореме сложения вероятностей $P(X<x_2)=P(X<x_1)+P(x_1\le X<x_2)$, или $F(x_2)=F(x_1)+P(x_1\le X<x_2)$, или

$$P(x_1 \le X < x_2) = F(x_2) - F(x_1). \qquad (16.19)$$

Так как $P(x_1 \le X < x_2) \ge 0$, то $F(x_2) \ge F(x_1)$.

3. $F(x) \to 1$ при $x \to +\infty$.

Доказательство. Рассмотрим последовательность $\{x_n\}$ (рис.16.13) такую, что $\lim\limits_{n \to +\infty} x_n = +\infty$,

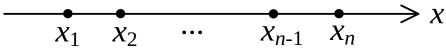


Рис. 16.13

и попарно несовместные события $A_1, ..., A_n$:

$$A_1 = \{X < x_1\}, \;\; A_2 = \{x_1 \le X < x_2\}, ..., \;\; A_n = \{x_{n-1} \le X < x_n\}.$$

Для события $A = \{X < x_n\}$ можно записать

$$A = A_1 + A_2 + ... + A_n \text{ и}$$

$$P(A) = F(x_n) = P(A_1) + P(A_2) + ... + P(A_n).$$

Переходя к пределу при $n \to \infty$ и учитывая, что $\sum\limits_{n=1}^{\infty} A_n = \Omega$, получаем

$$\lim_{n \to \infty} F(x) = P(\Omega) = 1.$$

4. $F(x) \to 0$ при $x \to -\infty$. Доказательство аналогично п.3.

5. Функция $F(x)$ непрерывна слева, т.е. $\lim\limits_{x \to x_i - 0} F(x) = F(x_i)$.

Следствия. 1. $P(X \ge x) = 1 - P(X < x) = 1 - F(x)$.

2. Для дискретной СВ $F(x) = \sum\limits_{x_i < x} p_i$.

3. Вероятность попадания СВ в заданный промежуток вычисляется по формуле (16.19):

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

Рассмотрим примеры дискретных распределений.

Пример 1. Биномиальное распределение.

Определение. Случайная величина X называется распределенной *по биномиальному закону*, если ряд распределения имеет вид табл.16.5.

Таблица 16.5

$X = x_m$	0	1	…	m	…	n
p_m	q^n	npq^{n-1}	…	$C_n^m p^m q^{n-m}$	…	p^n

При этом СВ X принимает целочисленные значения 0, 1, 2, …, $m, \ldots, n$ с вероятностями

$$p_m = P(X = m) = C_n^m p^m q^{n-m},$$

где $q = 1 - p$, $0 \le p \le 1$.

C_n^m – биномиальный коэффициент, отсюда название «биномиальное распределение». Это одно из основных распределений вероятностей, связанных с последовательностью независимых испытаний. Например, распределение СВ, равной числу успехов m в n испытаниях Бернулли. Каждую случайную величину X,имеющую биномиальное распределение с параметрами n, p, можно представить в виде суммы n независимых СВ $X_1, \ldots, X_n$, имеющих биномиальное распределение с параметрами $n = 1$ и p.

Пример 2. Распределение Пуассона.

Определение. *Распределение Пуассона* (пуассоновское распределение) – это распределение СВ X с целочисленными неотрицательными значениями 0, 1, 2, …, m, …, заданное формулой

$$P(X = m) = \frac{\lambda^m}{m!} e^{-\lambda}; \ \lambda > 0 \text{ – параметр.}$$

Ряд распределения имеет вид табл.16.6.

Таблица 16.6

$X = x_m$	0	1	2	…	m	…
p_m	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2!} e^{-\lambda}$	…	$\frac{\lambda^m}{m!} e^{-\lambda}$	…

Таким образом, СВ X принимает счетное число возможных значений. При этом

$$\sum_{m=0}^{\infty} p_m = e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda^2}{2!} e^{-\lambda} + \ldots =$$

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

Распределение Пуассона применяется для описания редких отклонений от нормы в хорошо отлаженном процессе (ошибки, сбои, опечатки, редкие болезни, редкие признаки). Такое распределение хорошо описывает процесс радиоактивного распада и многие другие физические явления.

Пример 3. Геометрическое распределение.

Определение. Случайная величина X называется распределенной *по геометрическому закону*, если она принимает значения 0, 1, 2, 3, … , m, ... с вероятностями $p_m = P(X = x_m) = q^m p$, т.е. ряд распределения имеет вид, представленный в табл.16.7.

Таблица 16.7

$X = x_m$	0	1	2	…	m	…
p_m	p	qp	$q^2 p$	…	$q^{m-1} p$	…

Замечание. Вероятности p_m образуют бесконечно убывающую геометрическую прогрессию $p,\ qp,\ q^2 p,\ \ldots$ со знаменателем q.

Покажем, что $\sum_{m=1}^{\infty} p_m = 1$:

$$\sum_{m=1}^{\infty} p_m = \sum_{m=1}^{\infty} pq^{m-1} = p\sum_{m=1}^{\infty} q^{m-1} = p \cdot \frac{1}{1-q} = \frac{p}{p} = 1.$$

По геометрическому закону распределена СВ X – число опытов в схеме Бернулли, проведенных до первого успеха.

Например, число выстрелов до первого попадания, число испытаний прибора до первого отказа.

2.3. Непрерывная случайная величина. Плотность распределения вероятностей

Функция распределения любой дискретной случайной величины – всегда разрывная ступенчатая функция, скачки которой происходят в точках, соответствующих возможным значениям случайной величины, и равны вероятностям этих значений. Сумма всех скачков функции $F(x)$ равна единице. По мере увеличения числа возможных значений случайной величины и уменьшения интервалов между ними число скачков становится больше, а сами скачки – меньше. Случайная величина постепенно приближается к непрерывной величине, а ее функция распределения – к непрерывной функции.

Пусть имеется непрерывная случайная величина X с функцией распределения $F(x)$, которую предположим непрерывной и дифференцируемой. Тогда

$$P(x \le X < x + \Delta x) = F(x + \Delta x) - F(x),$$

т.е. вероятность попадания случайной величины на промежуток $(x;\, x + \Delta x)$ есть приращение функции $F(x)$. Рассмотрим отношение этой вероятности к длине участка, т.е. среднюю вероятность, приходящуюся на единицу длины на этом участке, и будем приближать Δx к нулю. В пределе получим $F'(x)$. Введем обозначение

$$p(x) = F'(x). \tag{16.20}$$

Функция $p(x)$ характеризует как бы плотность, с которой распределяются значения случайной величины в окрестности данной точки. Эта функция называется *плотностью распределения* непрерывной случайной величины. Кривая, изображающая плотность распределения, называется *кривой распределения*.

Плотность распределения, так же как и функция распределения, есть одна из форм закона распределения. В противоположность функции распределения эта форма не является универсальной: она существует только для непрерывных случайных величин. Из формулы (16.20) следует

$$F(x)=\int_{-\infty}^{x} p(t)dt\,. \tag{16.21}$$

Поэтому можно сформулировать следующее определение НСВ.

Определение. СВ называется *непрерывной* СВ, если существует такая неотрицательная функция $p(x)$, что для $\forall x$ функция распределения равна $F(x)=\int_{-\infty}^{x} p(t)\,dt$.

Свойства плотности распределения:

1) $p(x)\geq 0$, так как функция $F(x)$ неубывающая;

2) $\int_{-\infty}^{+\infty} p(x)dx=1$, так как $F(+\infty)=1$.

Следствие. Если $X\in(a,b)$, то $\int_{a}^{b} p(x)dx=1$.

Геометрически свойства 1 и 2 означают, что:

1) кривая распределения лежит не ниже оси абсцисс;
2) полная площадь, ограниченная кривой распределения и осью абсцисс, равна единице;
3) вероятностный смысл плотности распределения:

$$F(x+\Delta x)-F(x)\approx dF(x)=F'(x)dx=p(x)\cdot\Delta x$$

или

$$P(x<X<x+\Delta x)\approx p(x)\cdot\Delta x\,.$$

Чем больше значение $p(x)$ в точке x, тем больше вероятность того, что случайная величина попадет в заданный небольшой интервал $(x;\,x+\Delta x)$.

Замечание. Функция распределения, как всякая вероятность, есть величина безразмерная. Размерность $p(x)$ обратна размерности случайной величины.

Пример 1. Функция распределения непрерывной случайной величины задана выражением

$$F(x)=\begin{cases}0, & x\le 0;\\ ax^2, & 0<x\le 4;\\ 1, & x>4.\end{cases}$$

Найти: 1) коэффициент a; 2) плотность распределения $p(x)$; 3) вероятность попадания X в промежуток $(1;3)$.

Решение. 1. Так как функция распределения непрерывна, то при $x=4$ имеем $F(x-0)=F(x+0)$ или $ax^2=1$, откуда $a=\dfrac{1}{16}$.

2. По определению плоности распределения получим:

$$p(x)=F'(x)=\begin{cases}0, & x\le 0;\\ \frac{1}{8}x, & 0<x<4;\\ 0, & x\ge 4.\end{cases}$$

Если объединить первую и последнюю строки, то получим:

$$p(x)=\begin{cases}\frac{1}{8}x, & x\in(0;4);\\ 0, & x\notin(0;4).\end{cases}$$

3. По формуле (16.19) находим

$$P(1\le X<3)=F(3)-F(1)=\frac{1}{16}(9-1)=0{,}5\,.$$

Замечание. Для непрерывной случайной величины вероятность любого отдельного значения равна нулю, так как

$$P(X=a)=\lim_{\Delta x\to 0}P(a\le X<a+\Delta x)=\lim_{\Delta x\to 0}\left[F(a+\Delta x)-F(a)\right]=0$$

в силу непрерывности функции $F(x)$.

Рассмотрим еще один пример, иллюстрирующий нахождение $F(x)$ по известной $p(x)$.

Пример 2. Плотность распределения непрерывной случайной величины имеет вид

$$p(x)=\begin{cases}0, & x\le 0;\\ 1, & 0<x\le 1;\\ 0, & x>1.\end{cases}$$

Найти функцию распределения $F(x)$.

Решение. Так как $p(x)$ имеет различное аналитическое выражение на промежутках $(-\infty;0]$, $(0;1]$ и $(1;+\infty)$, то рассмотрим три случая.

1. Пусть $x\le 0$. По формуле (16.21) находим

$$F(x)=\int_{-\infty}^{x}p(t)dt=\int_{-\infty}^{x}0\cdot dt=0.$$

2. Пусть $0<x\le 1$. Тогда

$$F(x)=\int_{-\infty}^{x}p(t)dt=\int_{-\infty}^{0}p(t)dt+\int_{0}^{x}p(t)dt=\int_{-\infty}^{x}0\cdot dt+\int_{0}^{x}1\cdot dt=x.$$

3. Пусть $x>1$. В этом случае имеем

$$F(x)=\int_{-\infty}^{x}p(t)dt=\int_{-\infty}^{0}p(t)dt+\int_{0}^{1}p(t)dt+\int_{1}^{x}p(t)dt=$$

$$=\int_{-\infty}^{0}0\cdot dt+\int_{0}^{1}1\cdot dt+\int_{1}^{x}0\cdot dt=1.$$

Обобщая все три случая, получаем:

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

Графики функций $p(x)$ и $F(x)$ представлены на рис. 16.14.

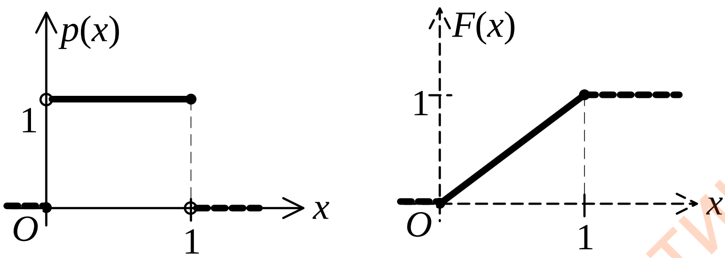


Рис. 16.14

2.3.1. Вероятность попадания непрерывной случайной величины в заданный промежуток

Ранее было получено, что

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

С учетом формулы (16.21) и сделанного выше замечания $(P(X = a) = 0)$ получаем:

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

$$= F(b) - F(a) = \int_{-\infty}^{b} p(t)dt - \int_{-\infty}^{a} p(t)dt = \int_{a}^{b} p(t)dt = \int_{a}^{b} p(x)dx\,. \quad (16.22)$$

Так, искомая вероятность в примере 1 может быть подсчитана следующим образом:

$$P(1 < X < 3) = \int_{1}^{3} \frac{1}{8} x dx = \frac{1}{8} \cdot \frac{x^2}{2}\Bigg|_{1}^{3} = \frac{1}{16}(9 - 1) = 0{,}5\,.$$

Геометрически (16.22) можно истолковать так: вероятность того, что непрерывная случайная величина примет значение, принадлежащее интервалу $(a;b)$ или другим промежуткам: $[a;b]$, $(a;b]$, $[a;b)$, равна площади криволинейной трапеции, ограниченной осью Ox, кривой распределения $p(x)$ и прямыми $x = a$ и $x = b$ (рис. 16.15).

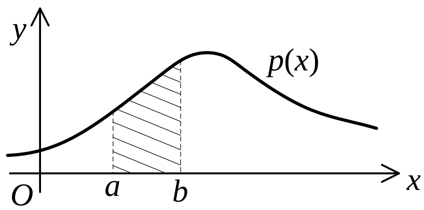


Рис. 16.15

2.3.2. Примеры непрерывных распределений: равномерное, показательное, нормальное

Пример 1. Равномерное распределение.

В некоторых задачах практики встречаются непрерывные случайные величины, о которых известно, что их возможные значения лежат в пределах промежутка $[a;b]$, все значения одинаково вероятны. Последнее означает, что все значения случайной величины обладают одной и той же плотностью вероятности. Говорят, что такие случайные величины имеют *равномерное* распределение.

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

График $p(x)$ представлен на рис. 16.16.

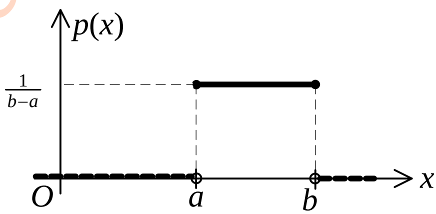


Рис. 16.16

Функция распределения для равномерного закона имеет вид рис. 16.17.

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

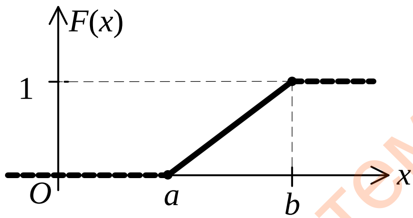


Рис. 16.17

Приведем пример равномерно распределенной случайной величины. Шкала измерительного прибора (вольтметр, амперметр, манометр и т.д.) проградуирована в некоторых единицах. Ошибку при округлении отсчета до ближайшего целого деления можно рассматривать как случайную величину, которая может с постоянной плотностью принимать любое значение между двумя соседними целыми делениями. Другие примеры: ошибка округления при взвешивании; время ожидания пассажира на платформе, когда поезда идут через фиксированный интервал времени.

Пример 2. Показательное распределение.

Показательным (экспоненциальным) называют распределение вероятностей, описываемое плотностью (рис. 16.18):

$$p(x)=\begin{cases}0, & x<0;\\ \lambda e^{-\lambda x}, & x\ge 0, \quad \lambda>0.\end{cases}$$

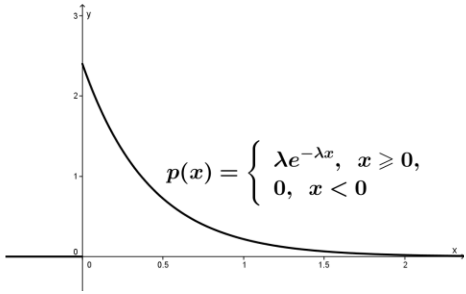


Рис. 16.18

Функция распределения для показательного закона имеет вид

$$F(x) = \begin{cases} 0, & x < 0; \\ 1 - e^{-\lambda x} & x \geq 0. \end{cases}$$

График $F(x)$ приведен на рис. 16.19. Показательное распределение определяется одним параметром λ .

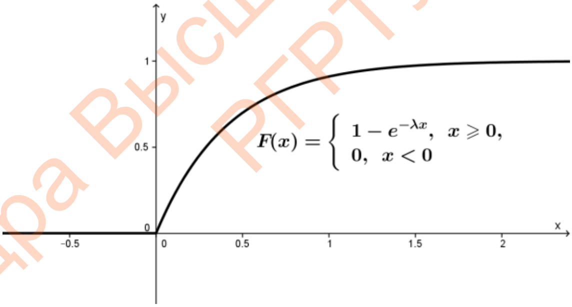


Рис.16.19

Пример 3. Нормальное распределение.

Нормальным называют распределение вероятностей непрерывной случайной величины, которое описывается плотностью

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

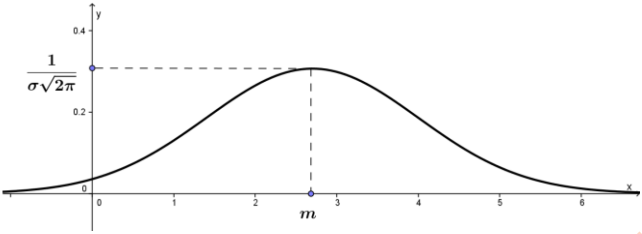


Рис.16.20

График распределения приведен на рис. 16.20. Нормальное распределение определяется двумя параметрами: m и σ. Тот факт, что непрерывная случайная величина X распределена нормально, коротко записывают так: $X \sim N(m,\sigma)$. Это распределение играет особую роль в теории вероятностей. Это связано с тем, что при выполнении определенных условий сумма большого числа случайных величин имеет распределение, близкое к нормальному.

Найдем функцию распределения для нормированного нормального закона ($m=0$ и $\sigma=1$):

$$F(x)=\int_{-\infty}^{x} p(t)dt=\int_{-\infty}^{x}\varphi(t)dt=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt=$$

$$=0{,}5+\frac{1}{\sqrt{2\pi}}\int_{0}^{x} e^{-\frac{t^2}{2}}dt=0{,}5+\Phi(x),$$

где $\varphi(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$ – нормированная функция Гаусса,

$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{0}^{x} e^{-\frac{t^2}{2}}dt$ – функция Лапласа.

График функции распределения нормального закона приведен на рис. 16.21.

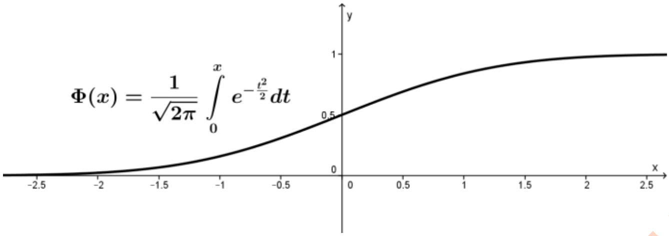


Рис. 16.21

Если $X \sim N(m, \sigma)$, то

$$P(a < X < b) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right),$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt$ – функция Лапласа. В частности,

$$P(|X-m| < \delta) = P(m-\delta < X < m+\delta) =$$

$$= \Phi\left(\frac{\delta}{\sigma}\right) - \Phi\left(\frac{-\delta}{\sigma}\right) = 2\Phi\left(\frac{\delta}{\sigma}\right).$$

Замечание. Кроме рассмотренных примеров непрерывных распределений для описания явлений, существуют другие распределения: Коши, Максвелла и т.д. Из распределений часто получают некоторые правила (как следствия), используемые на практике. Например, из нормального распределения выводят правило «трех сигм»:

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

2.3.3. Распределения Пирсона (хи-квадрат), Стьюдента и Фишера

С помощью нормального распределения определяются три распределения, которые в настоящее время часто используются при статистической обработке данных.

Распределение Пирсона χ^2 (хи-квадрат) – распределение случайной величины

$$X = X_1^2 + X_2^2 + \ldots + X_n^2,$$

где случайные величины $X_1, X_2, ..., X_n$ независимы и имеют одно и то же распределение $N(0,1)$. При этом число слагаемых, т.е. n, называется «числом степени свободы» распределения хи-квадрат.

Распределение хи-квадрат используют при оценивании дисперсии (с помощью доверительного интервала), при проверке гипотез согласия, однородности, независимости переменных, принимающих конечное число значений, и во многих других задачах статистического анализа данных.

Распределение t Стьюдента – это распределение случайной величины

$$T = \frac{U\sqrt{n}}{\sqrt{X}},$$

где случайные величины U и X независимы, U имеет стандартное нормальное распределение $N(0,1)$, а X – распределение хи-квадрат с n степенями свободы. При этом n называется «числом степеней свободы» распределения Стьюдента.

Распределение Стьюдента введено в 1908 г. английским статистиком В.Госсетом, работавшим на фабрике, выпускающей пиво. Вероятностно-статистические методы использовались для принятия экономических и технических решений на этой фабрике. Поэтому ее руководство запрещало В.Госсету публиковать научные статьи под своим именем. Таким способом охранялась коммерческая тайна, «ноу-хау» в виде вероятностно-статистических методов, разработанных В.Госсетом. Однако он имел возможность публиковаться под псевдонимом «Стьюдент».

В настоящее время распределение Стьюдента – одно из наиболее известных распределений среди используемых при анализе реальных данных. Его применяют при оценивании математического ожидания, прогнозного значения и других характеристик с помощью доверительных интервалов, при проверке гипотез о значениях математических ожиданий, коэффициентов регрессионной зависимости, гипотез однородности выборок и т.д.

Распределение Фишера – это распределение случайной величины

$$F = \frac{X_1/k_1}{X_2/k_2},$$

где случайные величины X_1 и X_2 независимы и имеют распределения хи-квадрат с числом степеней свободы k_1 и k_2 соответственно. При этом пара (k_1, k_2) – пара «чисел степеней свободы» распределения Фишера, а именно: k_1 – число степеней свободы числителя, а k_2 – число степеней свободы знаменателя.

Распределение случайной величины F названо в честь великого английского статистика Р. Фишера (1890-1962), активно использовавшего его в своих работах. Это распределение используют при проверке гипотез об адекватности модели в регрессионном анализе, о равенстве дисперсий и в других задачах прикладной статистики.

Выражения для функций распределения хи-квадрат, Стьюдента и Фишера, их плотностей и характеристик, а также таблицы, необходимые для их практического использования, можно найти в специальной литературе.

2.4. Числовые характеристики случайных величин

Функция распределения дает полную информацию о законе распределения случайной величины. Однако эта информация является труднообозримой. Часто встречаются ситуации, когда закон распределения случайной величины неизвестен и приходится ограничиваться меньшими сведениями о случайной величине, например некоторыми числами – числовыми характеристиками. Числовые характеристики можно условно разделить на три группы: характеристики положения (математическое ожидание, мода, медиана); характеристики рассеивания (дисперсия; среднее квадратическое отклонение); характеристики формы (коэффициент асимметрии, коэффициент эксцесса).

2.4.1. Математическое ожидание случайной величины

2.4.1.1. Математическое ожидание дискретной случайной величины

Рассмотрим дискретную случайную величину X и ее ряд распределения (табл.16.8).

Таблица 16.8

X	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

Определение. *Математическим ожиданием* дискретной случайной величины называется сумма произведений всех возможных значений СВ на вероятности этих значений (обозначение: MX, M_X, $M(X)$, $M[X]$, m_X, m).

Таким образом,

$$M(X) = x_1 \cdot p_1 + \ldots + x_n \cdot p_n = \sum_{i=1}^{n} x_i \cdot p_i \,. \qquad (16.23)$$

Для того чтобы уяснить смысл математического ожидания и происхождение этого термина, рассмотрим следующую задачу.

Рассматривается лотерея любого вида (денежная, книжная, вещевая и т.п.), содержащая N билетов. При этом на N_1 билетов лотереи выпадает выигрыш объемом a_1 (в частности, может быть, что $a_1 = 0$): на N_2 билетов выпадает выигрыш объемом a_2 и т.д. На N_n билетов выпадает выигрыш объемом a_n. Некто приобрел один билет и оценивает *ожидаемый выигрыш.*

Введем в рассмотрение случайную величину $X =${выигрыш в лотерее на один приобретенный билет}. Очевидно, что это дискретная случайная величина. Составим ряд распределения (табл.16.9).

Таблица 16.9

X	a_1	a_2	…	a_n

p	$\dfrac{N_1}{N}$	$\dfrac{N_2}{N}$	…	$\dfrac{N_n}{N}$

Находим

$$M(X)=a_1\cdot\frac{N_1}{N}+a_2\cdot\frac{N_2}{N}+\ldots+a_n\cdot\frac{N_n}{N}=\frac{a_1N_1+a_2N_2+\ldots+a_nN_n}{N}=$$

$$=\frac{\textit{выигрышный фонд лотереи}}{\textit{общее}\text{ количество билетов лотереи}}=$$

$$=\left[\textit{средний}\text{ выигрыш на один билет}\right].$$

Таким образом, математическое ожидание случайной величины X равно среднему ожидаемому выигрышу на один приобретенный билет.

Математическое ожидание иногда называют просто средним значением случайной величины. Это среднее значение есть некоторое число, являющееся как бы «представителем» случайной величины и заменяющее ее при грубо ориентированных расчетах.

Замечание. Пусть $x_{\min}$, $x_{\max}$ – минимальные и максимальные возможные значения случайной величины X. Тогда выполняется условие $x_{\min}<M(X)<x_{\max}$.

Примеры. 1. Математическое ожидание дискретной случайной величины, распределенной по биномиальному закону.

Пусть случайная величина $\mu=\{$число наступления события в n независимых испытаниях$\}$; p – вероятность наступления события в одном испытании; $q=1-p$. Находим M_μ.

По определению имеем

$$M_\mu=\sum_{m=0}^{n}m\cdot P_n(m)=\sum_{m=0}^{n}m\cdot C_n^m\cdot p^m\cdot q^{n-m}.$$

Для подсчета последней суммы воспользуемся биномом Ньютона:

$$(q+p)^n=\sum_{m=0}^{n}C_n^m\cdot p^m\cdot q^{n-m}. \qquad (16.24)$$

Продифференцируем (16.24) по p:

$$n\cdot(q+p)^{n-1}=\sum_{m=0}^{n}C_n^m m\cdot p^{m-1}\cdot q^{n-m}.$$

Умножая последнюю формулу на p, получаем

$$p \cdot n(q+p)^{n-1} = \sum_{m=0}^{n} C_n^m \cdot m p^m q^{n-m}. \tag{16.25}$$

С учетом $q+p=1$ из (16.25) получаем, что

$$\sum_{m=0}^{n} C_n^m \cdot m p^m q^{n-m} = np.$$

Следовательно, для биномиального закона получили

$$M_\mu = n \cdot p. \tag{16.26}$$

2. Математическое ожидание дискретной случайной величины, распределенной по закону Пуассона:

$$M(X) = \sum_{m=1}^{\infty} m \cdot p_m = \sum_{m=1}^{\infty} m \cdot \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=1}^{\infty} m \cdot \frac{\lambda^m}{m!} =$$

$$= e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!} = \lambda e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda \ .$$

3. Математическое ожидание дискретной случайной величины, распределенной по геометрическому закону:

$$M(X) = \sum_{m=1}^{\infty} m\, p_m = \sum_{m=1}^{\infty} m p q^{m-1} = p \sum_{m=1}^{\infty} m \cdot q^{m-1} = p \sum_{m=1}^{\infty} \frac{d(q^m)}{dq} =$$

$$= p \frac{d}{dq} \left(\sum_{m=1}^{\infty} q^m \right) = p \frac{d}{dq} \left(\frac{q}{1-q} \right) = p \frac{1}{(1-q)^2} = \frac{1}{p} \ .$$

2.4.1.2. Математическое ожидание непрерывной случайной величины

Рассмотрим непрерывную случайную величину X с плотностью $p(x)$ и возможными значениями из отрезка $[a; b]$. Разобьем отрезок произвольным образом на n частей точками $x_0 = a$, x_1, x_2, …, x_{k-1}, x_k, …, $x_n = b$ (рис. 16.22).

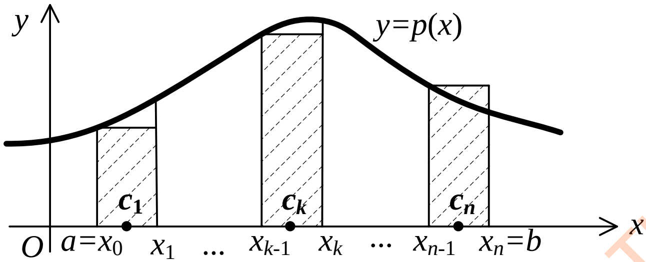


Рис. 16.22

Определим длины отрезков $\Delta x_k = x_k - x_{k-1}$, $k = \overline{1,n}$, и $\lambda = \max\limits_k \Delta x_k$. Выберем на отрезках случайным образом точки c_1, c_2, …, c_n и рассмотрим случайную величину C (дискретную), заданную рядом распределения:

C	c_1	c_2	…	c_k	…	c_n
p	p_1	p_2	…	p_k	…	p_n

При этом учтем, что

$$P(C = c_1) = p_1 = P(x_0 \le X \le x_1) \approx p(x_0)\Delta x_1;$$

$$P(C = c_2) = p_2 = P(x_1 \le X \le x_2) \approx p(x_1)\Delta x_2; \ \ldots;$$

$$P(C = c_k) = p_k = P(x_{k-1} \le X \le x_k) \approx p(x_{k-1})\Delta x_k; \ \ldots$$

$$P(C = c_n) = p_n = P(x_{n-1} \le X \le x_n) \approx p(x_{n-1})\Delta x_n.$$

Теперь $\quad M(C) = \sum\limits_{k=1}^{n} c_k p_k \approx \sum\limits_{k=1}^{n} c_k p(x_{k-1})\Delta x_k \approx \sum\limits_{k=1}^{n} c_k \cdot p(c_k)\Delta x_k$.

Последняя сумма есть интегральная сумма для функции $x \cdot p(x)$. Поэтому при $\lambda \to 0$, исходя из определения определенного интеграла, получим $M(X) = M(C)$ и

$$M(X) = \int_a^b x \cdot p(x)dx \quad \text{или} \quad M(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx,$$

если возможные значения непрерывной случайной величины X принадлежат всей числовой оси.

Определение. *Математическим ожиданием* непрерывной случайной величины называется несобственный интеграл

$$M(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx \qquad (16.27)$$

от произведения x и плотности распределения вероятностей $p(x)$.

Замечания. 1. Если $p(x) = p(-x)$ (функция $p(x)$ – четная), то $M(X) = 0$.

2. $M(X - M(X)) = 0$, так как

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

$$= \int_{-\infty}^{+\infty} x \cdot p(x)dx - m_X \int_{-\infty}^{+\infty} p(x)dx = m_X - m_X \cdot 1 = 0.$$

3. Есть случайные величины, для которых математическое ожидание не существует. Например, случайная величина, имеющая распределение Коши: $p(x) = \dfrac{1}{\pi(1 + x^2)}$.

Примеры. 1. Найдем математическое ожидание случайной величины, распределенной по нормальному закону.

Решение. Пусть $X \sim N(m, \sigma)$.

Тогда $M_X = \dfrac{1}{\sigma\sqrt{2\pi}} \displaystyle\int_{-\infty}^{+\infty} xe^{-\frac{(x-m)^2}{2\sigma^2}} dx$. Сделаем замену в интеграле $\dfrac{x - m}{\sigma} = t$; $x = m + \sigma t$; $dx = \sigma\, dt$; $-\infty < t < +\infty$. Теперь

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

$$= m \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t \cdot e^{-\frac{t^2}{2}} dt. \qquad (16.28)$$

В (16.28) второй интеграл равен нулю как интеграл от нечетной функции в симметричных пределах. Заметим далее, что $\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ есть плотность распределения случайной величины $X \sim N(0;1)$. Поэтому $\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = 1$ и из (16.28) следует:

$$M_X = m. \tag{16.29}$$

2. Найдем математическое ожидание случайной величины, распределенной по равномерному закону:

$$p(x) = \begin{cases} 0, & x \notin [a;b], \\ \frac{1}{b-a}, & x \in [a;b]. \end{cases}$$

Решение. Воспользуемся формулой (16.27):

$$M(X) = \int\limits_{-\infty}^{+\infty} x \cdot p(x) dx = \int\limits_{-\infty}^{a} xp(x) dx + \int\limits_{a}^{b} xp(x) dx + \int\limits_{b}^{+\infty} xp(x) dx =$$

$$= \int\limits_{-\infty}^{a} x \cdot 0 \cdot dx + \int\limits_{a}^{b} x \cdot \frac{1}{b-a} dx + \int\limits_{b}^{+\infty} x \cdot 0 \cdot dx =$$

$$= \frac{1}{b-a} \cdot \frac{x^2}{2}\Bigg|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}.$$

3. Найдем математическое ожидание случайной величины, распределенной по показательному закону:

$$p(x) = \begin{cases} 0, & x < 0; \\ \lambda e^{-\lambda x}, & x \ge 0, \quad \lambda > 0. \end{cases}$$

Решение. Также по формуле (16.27):

$$M(X) = \int\limits_{-\infty}^{+\infty} x \cdot p(x) dx = \int\limits_{-\infty}^{0} x \cdot p(x) dx + \int\limits_{0}^{+\infty} x \cdot p(x) dx =$$

$$= \int_{-\infty}^{0} x \cdot 0 dx + \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} dx = 0 + \lambda \lim_{b \to +\infty} \int_{0}^{b} x \cdot e^{-\lambda x} dx =$$

$$= \begin{vmatrix} u = x & du = dx \\ dv = e^{-\lambda x} dx & v = -\dfrac{1}{\lambda} e^{-\lambda x} \end{vmatrix} =$$

$$= \lambda \lim_{b \to +\infty} \left(-\frac{1}{\lambda} e^{-\lambda x} x \Big|_0^b + \frac{1}{\lambda} \int_0^b e^{-\lambda x} dx \right) = \lim_{b \to +\infty} \left(-e^{-\lambda x} x \Big|_0^b + \int_0^b e^{-\lambda x} dx \right) =$$

$$= \lim_{b \to +\infty} \left(-e^{-\lambda x} x \Big|_0^b - \frac{1}{\lambda} e^{-\lambda x} \Big|_0^b \right) = \frac{1}{\lambda} = M(X).$$

2.4.2. Основные свойства математического ожидания

1. Математическое ожидание постоянной величины равно этой постоянной:

$$M(C) = C.$$

Постоянную C можно рассматривать как случайную величину, имеющую следующий ряд распределения (табл.16.10).

Таблица 16.10

C	C
p	1

Поэтому $M(C) = C \cdot 1 = C$.

2. Постоянную можно выносить за знак математического ожидания:

$$M(C \cdot X) = C \cdot M(X),\ C = const.$$

Доказательство аналогично доказательству первого свойства.

3. Если X и Y – независимые случайные величины, то:

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

Для иллюстрации доказательства рассмотрим случайные величины X и Y (табл. 16.11).

Таблица 16.11

X	x_1	x_2

Y	y_1	y_2

p	p_1	p_2

p	g_1	g_2

При этом $p_1 + p_2 = 1$ и $g_1 + g_2 = 1$.

Возможные значения случайной величины $X \cdot Y$: $x_1 y_1$, $x_1 y_2$, $x_2 y_1$, $x_2 y_2$. Будем считать для простоты, что все они различны.

$P\left(X \cdot Y = x_i y_j\right) = P\left(X = x_i\right) \cdot P\left(Y = y_j\right) = p_i g_j$. Ряд распределения случайной величины $X \cdot Y$ приведен в табл. 16.12.

Таблица 16.12

$X \cdot Y$	$x_1 y_1$	$x_1 y_2$	$x_2 y_1$	$x_2 y_2$
P	$p_1 g_1$	$p_1 g_2$	$p_2 g_1$	$p_2 g_2$

Теперь

$$M\left(X \cdot Y\right) = x_1 y_1 \cdot p_1 g_1 + x_1 y_2 \cdot p_1 g_2 + x_2 y_1 \cdot p_2 g_1 + x_2 y_2 \cdot p_2 g_2 =$$

$$= x_1 p_1 \left(y_1 g_1 + y_2 g_2\right) + x_2 p_2 \left(y_1 g_1 + y_2 g_2\right) =$$

$$= \left(x_1 p_1 + x_2 p_2\right)\left(y_1 g_1 + y_2 g_2\right) = M\left(X\right) \cdot M\left(Y\right).$$

4. $M\left(X + Y\right) = M\left(X\right) + M\left(Y\right)$.

Доказательство. Рассмотрим, как и раньше, случайные величины (табл. 16.10) и будем считать, что возможные значения случайной величины $X + Y$ различны между собой (табл. 16.13).

Таблица 16.13

$X + Y$	$x_1 + y_1$	$x_1 + y_2$	$x_2 + y_1$	$x_2 + y_2$
P	p_{11}	p_{12}	p_{21}	p_{22}

Здесь $p_{ij} = P\left(X + Y = x_i + y_j\right)$. Так как не предполагалось независимости X и Y, то в общем случае $p_{ij} \neq p_i \cdot g_j$. По определению получим:

$$M\left(X + Y\right) = \left(x_1 + y_1\right) p_{11} + \left(x_1 + y_2\right) p_{12} + \left(x_2 + y_1\right) p_{21} + \left(x_2 + y_2\right) p_{22} =$$

$$= x_1\left(p_{11} + p_{12}\right) + x_2\left(p_{21} + p_{22}\right) + y_1\left(p_{11} + p_{21}\right) + y_2\left(p_{12} + p_{22}\right).$$

Покажем, что $p_{11}+p_{12}=p_1$. Для этого рассмотрим два события: $A=\{X=x_1\}$ и $B=\{X+Y=x_1+y_1$ или $X+Y=x_1+y_2\}$. Появление события A влечет появление события B и наоборот. Следовательно, $A=B$. Но $P(A)=p_1$, $P(B)=p_{11}+p_{12}$ по теореме сложения вероятностей. Аналогично, показывается, что $p_{21}+p_{22}=p_2$, $p_{11}+p_{21}=g_1$ и $p_{12}+p_{22}=g_2$. Тогда

$$M(X+Y)=x_1\cdot p_1+x_2\cdot p_2+y_2g_1+y_2g_2=M(X)+M(Y).$$

Следствие. Математическое ожидание суммы нескольких случайных величин $X_1,\ldots,X_n$ равно сумме математических ожиданий слагаемых

$$M(X_1+\ldots+X_2)=M(X_1)+\ldots+M(X_n).$$

Доказательство проводится методом математической индукции.

Пример. С помощью сформулированного следствия легко получить выражение для математического ожидания числа появления события в n независимых испытаниях. Будем считать, что $\mu=\mu_1+\mu_2+\ldots+\mu_n$, где μ_i – случайная величина, равная числу появления события в i-м испытании, т.е. ее ряд распределения имеет вид табл. 16.14.

Таблица 16.14

μ_i	0	1
P	q	p

Тогда $M_{\mu_i}=0\cdot q+1\cdot p=p$ и

$$M_\mu=M(\mu_1+\mu_2+\ldots+\mu_n)=\sum_{i=1}^{n}M_{\mu_i}=p\cdot n.$$

2.4.3. Дисперсия случайной величины

Кроме характеристик положения, употребляются другие характеристики, описывающие различные свойства случайной величины. Нетрудно указать случайные величины, которые имеют одинаковые математические ожидания, но возможные значения их существенно различны. Например,

X	-0,001	0,001
P	0,5	0,5

Y	-500	500
P	0,5	0,5

Здесь $M(X)=M(Y)=0$. Но случайная величина X имеет возможные значения, близкие к ее математическому ожиданию, а Y имеет возможные значения, далекие от своего математического ожидания. Таким образом, математическое ожидание полностью случайную величину не характеризует. Для решения ряда практических задач требуется оценить рассеивание (разброс) возможных значений случайной величины вокруг ее математического ожидания.

На первый взгляд кажется, что для случайной величины X, имеющей ряд распределения табл. 16.15, для оценки рассеивания достаточно вычислить $x_1 - M_X$, …, $x_n - M_X$, т.е. все возможные значения отклонения случайной величины от математического ожидания, а затем найти их среднее значение, т.е. математическое ожидание $M(x_n - M_X)$.

Таблица 16.15

X	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

Однако

$$M(x_n - M_X) = M_X - M(M_X) = M_X - M_X = 0.$$

Это означает, что возможные положительные отклонения погашаются возможными отрицательными, а среднее отклонение равно нулю.

Чтобы избежать такого эффекта взаимного погашения, можно взять модули возможных отклонений $|x_i - M_X|$ или квадраты возможных отклонений $(x_i - M_X)^2$. Второй путь считается предпочтительным.

Определение. *Дисперсией* случайной величины X [обозначение DX или $D(X)$ или D_X] называется математическое ожидание квадрата отклонения случайной величины от ее математического ожидания:

$$D(X)=M\left((X-M_X)^2\right). \qquad (16.30)$$

Замечания. 1. Из определения следует, что дисперсия случайной величины есть величина неслучайная.

2. Дисперсия есть величина неотрицательная, так как $(X-M_X)^2\geq 0$ и $M\left((X-M_X)^2\right)\geq 0$.

3. Дисперсия случайной величины есть характеристика рассеивания, разбросанности возможных значений случайной величины около ее математического ожидания. В этом состоит ее вероятностный смысл.

Из (16.30) для вычисления дисперсии можно получить более удобную формулу:

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

т.е. дисперсия есть разность между математическим ожиданием квадрата случайной величины и квадратом ее математического ожидания:

$$D(X)=M\left(X^2\right)-M_X^2. \qquad (16.31)$$

Для дискретной случайной величины (ДСВ) формулы (16.30) и (16.31) принимают вид:

$$D(X)=\sum_{i=1}^{n}(x_i-M_X)^2 p_i, \quad D(X)=\sum_{i=1}^{n}x_i^2 p_i-(M_X)^2. \qquad (16.32)$$

Для непрерывной случайной величины (НСВ) формулы (16.30) и (16.31) имеют вид:

$$D(X)=\int_{-\infty}^{+\infty}(x-M_X)^2 p(x)dx, \quad D(X)=\int_{-\infty}^{+\infty}x^2 p(x)dx-(M_X)^2. \qquad 16.33)$$

Пример 1. Функция распределения случайной величины X задана формулой

$$F(x)=\begin{cases}0, & x\leq 0,\\ x, & 0<x\leq 1,\\ 0, & x>1.\end{cases}$$

Найти $M(X)$ и $D(X)$.

Решение. Находим функцию $p(x)$, где $p(x)=F'(x)$:

$$p(x)=\begin{cases}0, & x\le 0,\\ 1, & 0<x\le 1,\\ 0, & x>1,\end{cases}=\begin{cases}1, & x\in[0;1],\\ 0, & x\notin[0;1].\end{cases}$$

Таким образом, имеем дело с равномерным законом распределения. Далее находим:

$$M(X)=\int_{-\infty}^{+\infty}x\cdot p(x)dx=\int_0^1 x\cdot 1dx=\left.\frac{x^2}{2}\right|_0^1=\frac{1}{2},$$

$$M\left(X^2\right)=\int_0^1 x^2dx=\left.\frac{x^3}{3}\right|_0^1=\frac{1}{3},$$

$$D(X)=M\left(X^2\right)-M^2(X)=\frac{1}{3}-\left(\frac{1}{2}\right)^2=\frac{1}{3}-\frac{1}{4}=\frac{1}{12}.$$

Пример 2. Если случайная величина X распределена равномерно в интервале $(a;b)$ или на отрезке $[a;b]$, то $D_X=\dfrac{(b-a)^2}{12}$. Доказать самостоятельно.

Дисперсия случайной величины имеет размерность квадрата случайной величины, что не всегда удобно для качественной характеристики случайной величины. Поэтому для характеристики рассеивания случайной величины удобнее пользоваться величиной, размерность которой совпадает с размерностью случайной величины. Для этого из дисперсии извлекают квадратный корень. Полученная величина называется *средним квадратическим отклонением* (обозначается σ_X). Итак, $\sigma_X=\sqrt{D_X}$.

Замечание. Когда не возникает сомнения, к какой случайной величине относится характеристика рассеивания, допускается опускать символ X в D_X и σ_X, а записывать просто D и σ соответственно.

Выделим основные свойства дисперсии.

1. Дисперсия постоянной величины C равна 0.

Доказательство. Согласно (16.30) и свойствам математического ожидания

$$DX = M\left(\left(C - M_C\right)^2\right) = M\left(\left(C - C\right)^2\right) = M(0) = 0.$$

Это свойство очевидно, так как постоянная величина сохраняет одно и то же значение и, следовательно, рассеивания около математического ожидания не имеет.

2. Постоянный множитель можно выносить за знак дисперсии, возводя его в квадрат:

$$D(C \cdot X) = C^2 \cdot D(X).$$

Доказательство.

$$D(C \cdot X) = M\left[C \cdot X - M(C \cdot X)\right]^2 = M\left[C \cdot X - C \cdot M(X)\right]^2 =$$
$$= M\left[C \cdot \left(X - M(X)\right)\right]^2 = M\left[C^2 \cdot \left(X - M(X)\right)^2\right] =$$
$$= C^2 \cdot M\left[\left(X - M(X)\right)^2\right] = C^2 \cdot D(X).$$

3. Если X и Y – независимые случайные величины, то $D(X + Y) = D(X) + D(Y)$.

Доказательство.

$$D(X + Y) = M\left((X + Y)^2\right) - \left(M(X + Y)\right)^2 =$$
$$= M\left(X^2 + 2XY + Y^2\right) - \left(M(X) + M(Y)\right)^2 =$$
$$= M\left(X^2\right) + 2M(XY) + M\left(Y^2\right) - \left(M_X\right)^2 - 2M_X \cdot M_Y - \left(M_Y\right)^2.$$

Так как X и Y – независимые случайные величины, то $M(X \cdot Y) = M_X \cdot M_Y$. Тогда

$$D(X + Y) = M\left(X^2\right) + M\left(Y^2\right) - \left(M_X\right)^2 - \left(M_Y\right)^2 =$$
$$= \left(M\left(X^2\right) - \left(M_X\right)^2\right) + \left(M\left(Y^2\right) - \left(M_Y\right)^2\right) = D(X) + D(Y).$$

Следствие. $D(C + X) = D(X)$ в силу независимости величин C и X. Это следует также из того, что величины $C + X$ и C отличаются лишь началом отсчета. Поэтому они одинаково рассеяны относительно своих математических ожиданий $M(X) + C$ и C.

Аналогично можно доказать, что дисперсия суммы нескольких взаимно независимых случайных величин равна сумме дисперсий этих величин.

4. Среднее квадратическое отклонение среднего арифметического n независимых случайных величин X_1, …, X_n с одинаковой дисперсией σ^2 равно $\dfrac{\sigma}{\sqrt{n}}$.

Доказательство. Найдем среднее арифметическое $\overline{X}_n$ случайных величин X_1, …, X_n:

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i .$$

Тогда дисперсия среднего арифметического равна:

$$D(\overline{X}_n) = D\left(\frac{1}{n}\sum_i X_i\right) = \frac{1}{n^2} D\left(\sum_i X_i\right) = \frac{1}{n^2}\sum_i D(X_i) =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}.$$

В итоге получим среднее квадратическое отклонение:

$$\sigma_{\overline{X}_n} = \sqrt{D(\overline{X}_n)} = \frac{\sigma}{\sqrt{n}}.$$

5. Дисперсия разности двух независимых случайных величин X и Y равна сумме их дисперсий, т.е. $D(X - Y) = D(X) + D(Y)$.

Доказательство.

$$D(X - Y) = D(X + (-1)\cdot Y) = D(X) + D((-1)\cdot Y) =$$

$$= D(X) + (-1)^2 \cdot D(Y) = D(X) + D(Y).$$

Примеры. 1. Дисперсия дискретной случайной величины, распределенной по биномиальному закону.

Рассматривается последовательность n независимых испытаний Бернулли. Пусть случайная величина μ_i ={число появления события A в i-м испытании}, μ ={число появления события A в n испытаниях}. Очевидно, $\mu = \mu_1 + \mu_2 + \ldots + \mu_n$. Найдем $D(\mu)$. Так как величины μ_1, …, μ_n взаимно независимы, то $D(\mu) = \sum_{i=1}^{n} D_{\mu_i}$. При этом

μ_i	0	1
P	q	p

μ_i^2	0	1
P	q	p

$$M_{\mu_i} = 0 \cdot q + 1 \cdot p = p, \; M_{\mu_i^2} = 0 \cdot q + 1 \cdot p = p ,$$

$$D(\mu_i) = M(\mu_i^2) - (M_{\mu_i})^2 = p - p^2 = p(1-p) = pq .$$

Следовательно, $D(\mu) = npq$ и $\sigma_\mu = \sqrt{npq}$.

2. Можно показать, что дисперсия дискретной случайной величины, распределенной по закону Пуассона, равна $D_X = \lambda$.

3. Можно показать, что дисперсия дискретной случайной величины, распределенной по геометрическому закону, равна $D_X = \dfrac{q}{p^2}$.

4. Дисперсия случайной величины, распределенной по нормальному закону. Пусть $X \sim N(m, \sigma)$. Найдем D_X:

$$D_X = M(X^2) - M^2(X) = \int_{-\infty}^{+\infty} x^2 p(x) dx - m^2 ,$$

где $p(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$. Рассмотрим $M(X^2) = \int_{-\infty}^{+\infty} x^2 p(x) dx$.

$$M(X^2) = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \left| \begin{array}{l} \dfrac{x-m}{\sigma} = t \\ x = m + \sigma t \\ dx = \sigma\, dt \end{array} \right| =$$

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

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

$$=\frac{m^2}{\sqrt{2\pi}}Y_1+\frac{2m\sigma}{\sqrt{2\pi}}Y_2+\frac{\sigma^2}{\sqrt{2\pi}}Y_3.$$

Так как $\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$ – плотность случайной величины $X\sim N(0;1)$, то по свойству плотности распределения:

$$\frac{m^2}{\sqrt{2\pi}}Y_1=m^2\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=m^2\cdot 1=m^2.$$

Интеграл Y_2 равен 0 как интеграл от нечетной функции в симметричных пределах. Для вычисления Y_3 воспользуемся интегрированием по частям:

$$Y_3=\int_{-\infty}^{+\infty}t^2e^{-\frac{t^2}{2}}dt=-\int_{-\infty}^{+\infty}td\left(e^{-\frac{t^2}{2}}\right)=-\left(te^{-\frac{t^2}{2}}\Bigg|_{-\infty}^{+\infty}-\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt\right)=\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt,$$

поэтому $\frac{\sigma^2}{\sqrt{2\pi}}\cdot Y_3=\sigma^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=\sigma^2$. Окончательно имеем:

$$M\left(X^2\right)=m^2+0+\sigma^2=m^2+\sigma^2,$$

тогда $D_X=M\left(X^2\right)-M^2(X)=m^2+\sigma^2-m^2=\sigma^2$.

5. Дисперсия случайной величины, распределенной по равномерному закону:

$$DX=\int_{-\infty}^{+\infty}x^2p(x)dx-M^2X=\int_a^b x^2\cdot\frac{1}{b-a}dx-\left(\frac{b+a}{2}\right)^2=$$

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

6. Можно показать, что дисперсия случайной величины, распределенной по показательному закону, равна $D_X = \dfrac{1}{\lambda^2}$.

Замечание. К характеристикам рассеивания, кроме дисперсии и среднего квадратического отклонения, относится *коэффициент вариации* – это отношение среднего квадратического отклонения к математическому ожиданию:

$$v = \frac{\sigma}{M(X)}.$$

Коэффициент вариации применяется при $M(X) > 0$ и измеряет разброс в относительных единицах, в то время как среднее квадратическое отклонение – в абсолютных. Например, для равномерно распределенной случайной величины X среднее квадратическое отклонение равно

$$\sigma = \frac{b-a}{2\sqrt{3}},$$

а коэффициент вариации таков:

$$v = \frac{b-a}{\sqrt{3}(a+b)}.$$

2.4.4. Нормированная случайная величина

В математической модели случайная величина описывает те или иные параметры изучаемого случайного явления. Числовые значения исходных параметров, а также математическое ожидание, дисперсия, среднее квадратическое отклонение и т.д. зависят от выбора масштаба измерения и положения математического ожидания. Чтобы избежать такой зависимости, случайную величину приводят к некоторому стандартному виду.

Определение. Случайную величину X^0 называют *нормированной,* если $M(X^0) = 0$ и $D(X^0) = 1$.

Для случайной величины X выполним преобразование $X^0 = \dfrac{X - M(X)}{\sigma_X}$, называемое нормированием случайной величины X. При этом случайная величина X^0 является нормированной. В самом деле,

$$M\left(X^0\right) = M\left(\frac{X - M_X}{\sigma_X}\right) = \frac{1}{\sigma_X} M\left(X - M_X\right) = \frac{1}{\sigma_X}\left(M_X - M_X\right) = 0,$$

$$D\left(X^0\right) = D\left(\frac{X - M_X}{\sigma_X}\right) = \frac{1}{\sigma_X^2} D\left(X - M_X\right) = \frac{D_X}{\sigma_X{}^2} = 1.$$

Приводя случайную величину X к нормированной случайной величине X^0, мы как бы меняем начало отсчета и масштаб измерения исходной случайной величины. Теперь единицей измерения становится σ_{X^0}. При этом сама случайная величина X^0 является безразмерной и не зависит от выбора масштаба измерения X.

Примеры. 1. Для ДСВ, распределенной по биномиальному закону, нормированной СВ является $X^0 = \dfrac{m - np}{\sqrt{npq}}$ (докажите!).

2. Для НСВ, распределенной по нормальному закону, нормированной СВ является $X^0 = \dfrac{X - m}{\sigma}$ (докажите!).

2.4.5. Мода и медиана. Квантили

Кроме математического ожидания характеристиками положения СВ служат мода и медиана.

Определение. *Модой* ДСВ X называется ее значение, принимаемое с наибольшей вероятностью по сравнению с двумя соседними значениями. Обозначается $Mo(X)$.

Для НСВ $Mo(X)$ – точка локального максимума плотности распределения $p(x)$ (рис. 16.23). Если мода единственна, то закон распределения СВ называется *унимодальным*, в противном случае – *полимодальным.*

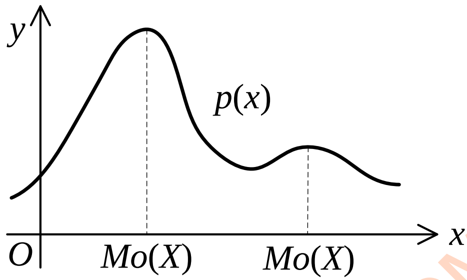


Рис. 16.23

В геометрии есть понятие «медиана» – отрезок, соединяющий вершину треугольника с серединой противоположной стороны. В математической статистике медиана делит пополам не сторону треугольника, а распределение случайной величины.

Определение. *Медианой* НСВ X [обозначение $Me(X)$] называется такое ее значение, для которого

$$P(X < Me(X)) = P(X \geq Me(X)) = \frac{1}{2}, \tag{16.34}$$

т.е. одинаково вероятно, окажется ли СВ меньше $Me(X)$ или больше $Me(X)$ (рис. 16.24).

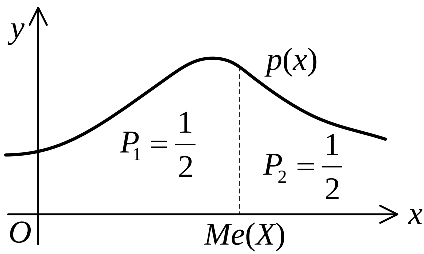


Рис. 16.24

С помощью функции распределения $F(x)$ равенство (16.34) можно записать в виде $F(Me_X)=1-F(Me_X)$. Отсюда $F(Me_X)=\dfrac{1}{2}$.

Замечания. 1.Для ДСВ медиана обычно не определяется.

2. С точки зрения одной из современных концепций – теории устойчивых статистических процедур – медиана более подходящая характеристика случайной величины, чем математическое ожидание. Рассмотрим пример.

Предположим, что в одной комнате оказалось 19 бедняков и один миллиардер. Каждый кладёт на стол деньги – бедняки из кармана, а миллиардер – из чемодана. По \$5 кладёт каждый бедняк, а миллиардер – \$1 млрд (10^9). В сумме получается \$1 000 000 095. Если мы разделим деньги равными долями на 20 человек, то получим \$50 000 004,75. Это будет среднее арифметическое значение суммы наличных, которая была у всех 20 человек в этой комнате.

Медиана в этом случае будет равна \$5 (полусумма десятого и одиннадцатого, *срединных* значений ранжированного ряда). Можно интерпретировать это следующим образом. Разделив нашу компанию на две равные группы по 10 человек, мы можем утверждать, что в первой группе каждый положил на стол не больше \$5, во второй же не меньше \$5. В общем случае можно сказать, что медиана – это то, сколько принёс с собой *средний* человек. Наоборот, среднее арифметическое – неподходящая характеристика, так как оно значительно превышает сумму наличных, имеющуюся у среднего человека.

3. Каждая из трех характеристик – математическое ожидание, медиана, мода – описывает «центр» распределения вероятностей. Понятие «центр» можно определять разными способами – отсюда три разные характеристики. Однако для важного класса распределений – симметричных унимодальных – все три характеристики совпадают.

Примеры. 1. Для СВ, распределенной по нормальному закону, $M(X)=Mo(X)=Me(X)=m$.

2. Для СВ, распределенной по показательному закону, $MX=\dfrac{1}{\lambda}$, $Mo(X)=0$, $Me(X)=\dfrac{\ln 2}{\lambda}$.

Определение. *Квантилью* x_p *порядка* p СВ X называется решение уравнения

$$F(x_p)=p, \tag{16.35}$$

где $0<p<1$.

Квантиль порядка p – значение случайной величины, для которой функция распределения принимает значение p. Может случиться, что это условие выполняется для всех значений x, принадлежащих этому интервалу (т.е. функция распределения постоянна на этом интервале и равна p). Тогда каждое такое значение называется «квантилью порядка p». Для непрерывных функций распределения, как правило, существует единственная квантиль x_p порядка p (см. рис. 16.25).

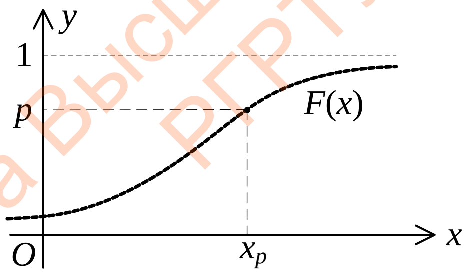


Рис. 16.25

Пример. Найдем квантиль x_p порядка p для функции $F(x)$ для равномерного закона распределения:

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

Решение.

При $0 < p < 1$ квантиль x_p находится из уравнения

$$\frac{x-a}{b-a} = p,$$

т.е. $x_p = a + p(b-a) = a(1-p) + bp$.

При $p = 0$ любое $x \le a$ является квантилью порядка $p = 0$. Квантилью порядка $p = 1$ является любое число $x \ge b$.

Замечание. Медиана является квантилью уровня 0,5, т.е. $Me(X) = x_{0,5}$. Таким образом, равенство $F(x_{0,5}) = 0{,}5$ означает, что вероятность попасть левее $x_{0,5}$ и вероятность попасть правее $x_{0,5}$ (или непосредственно в $x_{0,5}$) равны между собой и равны $\frac{1}{2}$, т.е. $P(X < x_{0,5}) = P(X \ge x_{0,5}) = \frac{1}{2}$.

2.4.6. Моменты случайных величин

Рассмотренные числовые характеристики случайной величины – математическое ожидание и дисперсия – являются частными случаями понятия момента случайной величины. На практике применяются моменты двух видов: начальные и центральные.

Определение. *Начальным моментом* k*-го порядка* случайной величины X называется математическое ожидание k-й степени этой случайной величины (обозначается ν_k):

$$\nu_k = M(X^k). \tag{16.36}$$

Определение. *Центральным моментом* k*-го порядка* случайной величины X называется математическое ожидание k-й степени соответствующей центрированной случайной величины (обозначение μ_k):

$$\mu_k = M\left((X - M_X)^k\right). \tag{16.37}$$

Таким образом, $\nu_1 = M(X)$, $\nu_2 = M(X^2)$, $\mu_1 = 0$, $\mu_2 = D(X)$. Центральные моменты выражаются через начальные: $\mu_2 = \nu_2 - \nu_1^2$, $\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3$ и т.д.

Центральный момент μ_3 служит для характеристики асимметрии (или «скошенности») распределения. Для получения безразмерной характеристики асимметрии вводят *коэффициент асимметрии* $\gamma_1 = \dfrac{\mu_3}{\sigma^3}$.

Если $\gamma_1 < 0$, то кривая распределения более полога слева от $Mo(X)$; если $\gamma_1 > 0$, то кривая распределения более полога справа от $Mo(X)$ (рис. 16.26).

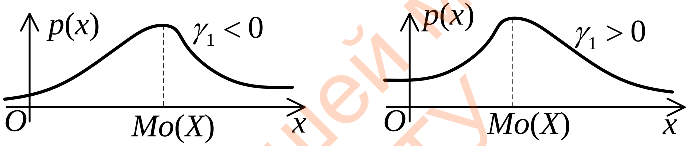


Рис. 16.26

Четвертый центральный момент μ_4 служит для характеристики «крутости», т.е. островершинности или плосковершинности распределения. Эти свойства распределения описываются с помощью так называемого *коэффициента эксцесса* $\gamma_2 = \dfrac{\mu_4}{\sigma^4} - 3$ (число -3 в конце формулы введено для того, чтобы коэффициент эксцесса нормального распределения был равен нулю).

Величина γ_2 характеризует островершинность или плосковершинность распределения. Для нормального закона распределения $\gamma_1 = 0$ и $\gamma_2 = 0$, остальные распределения сравниваются с нормальным: если $\gamma_2 > 0$ – более островершинные, а плосковершинные распределения имеют $\gamma_2 < 0$ (рис. 16.27).

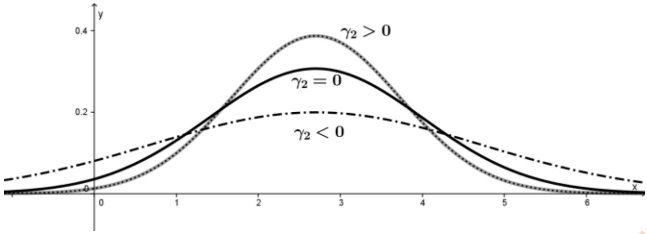


Рис. 16.27

3. Двумерная случайная величина

На одном и том же пространстве элементарных событий Ω может быть определена не одна, а несколько случайных величин. Необходимость в этом возникает, например, при моделировании ситуации, когда объект характеризуется несколькими случайными параметрами. Например, точка попадания снаряда определяется (описывается) не одной случайной величиной, а двумя: абсциссой и ординатой и может быть рассмотрена как комплекс или система двух случайных величин. Осколок, образовавшийся при разрыве снаряда, характеризуется рядом случайных величин: массой, размерами, начальной скоростью, направлением полета и т.д.

Свойства системы нескольких случайных величин не исчерпываются свойствами отдельных величин ее составляющих: помимо этого они включают также взаимные связи между случайными величинами.

3.1. Случайные векторы. Двумерная величина и функция ее распределения

При рассмотрении вопросов, связанных с системами случайных величин, удобно пользоваться геометрической интерпретацией системы. Например, систему двух случайных величин (X, Y) можно изображать случайной точкой на плоскости Oxy. Часто вместо образа случайной точки для геометрической интерпретации системы случайных величин пользуются образом случайного вектора. Так, систему двух случайных величин при этом рассматривают как случайный радиус-вектор на плоскости Oxy с координатами $(x,\ y)$.

Определение. Функцией распределения двумерной случайной величины (X, Y) называется вероятность совместного выполнения двух событий $(X < x)$ и $(Y < y)$:

$$F(x, y) = P(X < x,\ Y < y). \tag{16.38}$$

Если воспользоваться геометрической интерпретацией, то функция распределения $F(x, y)$ есть вероятность попадания случайной точки (X, Y) в бесконечный квадрант с вершиной $(x,\ y)$, лежащий левее и ниже этой вершины (рис. 16.28).

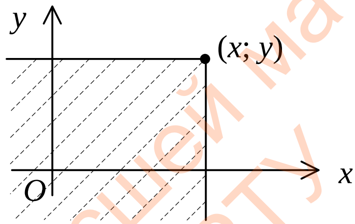


Рис. 16.28

Отметим некоторые свойства функции распределения:

1) $0 \le F(x, y) \le 1$;

2) $F(x, y)$ – неубывающая функция по каждому аргументу, т.е.

$$\forall x_1, x_2 \in \mathbf{R}: \quad x_1 < x_2 \Rightarrow F(x_1, y) \le F(x_2, y),$$

$$\forall y_1, y_2 \in \mathbf{R}: \quad y_1 < y_2 \Rightarrow F(x, y_2) \le F(x, y_2);$$

3) при $x \to -\infty$ верны условия $F(x, y) \to 0$ и $F(-\infty, y) = 0$. Также при $y \to -\infty$ верно $F(x, y) \to 0$ и $F(x, -\infty) = 0$. Обобщением этого будет $F(-\infty, -\infty) = 0$.

В этом свойстве можно наглядно убедиться, неограниченно отодвигая влево правую границу квадранта ($x \to -\infty$), вниз его верхнюю границу ($y \to -\infty$) или делая это одновременно с обеими границами (рис. 16.28). При этом вероятность попадания в квадрант стремится к нулю;

4) $F(+\infty, +\infty) = 1$;

5) $F(x,+\infty)=F(x),\ F(+\infty,y)=F(y).$

Функции распределения $F(x),\ F(y)$ в этом случае называются *частными*, или *маргинальными.*

Так, имеем

$$F(x,\ +\infty)=P(X<x,\ Y<+\infty)=P(X<x)\cdot P(Y<+\infty)=$$
$$=P(X<x)\cdot 1=F(x).$$

3.2. Дискретные и непрерывные двумерные случайные величины

Рассмотрим дискретную двумерную случайную величину (ДДСВ). Пусть возможные значения составляющих таковы:

$$X\to\{x_1,\,x_2,\dots,\,x_m\},\ Y\to\{y_1,y_2,\dots,y_n\}.$$

Обычно ДДСВ задают в виде *матрицы распределения.* При этом возможные значения X и Y располагают в порядке возрастания. Таким образом, матрица распределения $(p_{i\,j})$ имеет вид табл. 16.16.

Таблица 16.16

	y_1	…	y_j	…	y_n	p_i
x_1	p_{11}	…	p_{1j}	…	p_{1n}	p_1
…	…	…	…	…	…	…
x_i	p_{i1}	…	$p_{i\,j}$	…	$p_{i\,n}$	$\sum\limits_j p_{i\,j}=p_i$
…	…	…	…	…	…	…
x_m	p_{m1}	…	$p_{m\,j}$	…	$p_{m\,n}$	p_m
p_j	p_1	…	p_j	…	p_n	$\sum\limits_i\sum\limits_j p_{i\,j}=1$

При этом $p_{i\,j}=P(X=x_i,Y=y_j)$. В последнем столбце помещены суммы элементов по строкам:

$$p_i=\sum_{j=1}^{n}p_{ij}=P(X=x_i).$$

Аналогично элементы в последней строке равны сумме элементов по соответствующим столбцам:

$$p_j = \sum_{i=1}^{m} p_{ij} = P(Y = y_j).$$

Для функции распределения ДДСВ находим

$$F(x, y) = \sum_{x_i < x} \left(\sum_{y_j < y} p_{ij} \right).$$

Пример 1. ДДСВ задана матрицей распределения (табл.16.17).

Таблица 16.17

	0	0,1	0,2	0,3	p_i
5	0,2	0,1	0,05	0,05	0,4
6	0	0,15	0,15	0,1	0,4
7	0	0	0,1	0,1	0,2
p_j	0,2	0,25	0,03	0,25	1

Найти $F(5{,}5; 0{,}25)$.

Решение.

$$F(5{,}5; 0{,}25) = \sum_{x_i < 5{,}5} \left(\sum_{y_j < 0{,}25} p_{ij} \right) = p_{11} + p_{12} + p_{13} =$$

$$= 0{,}2 + 0{,}1 + 0{,}05 = 0{,}35,$$

так как

$$P(X < 5{,}5; Y < 0{,}25) =$$

$$= P(X = 5, Y = 0) + P(X = 5, Y = 0{,}1) + P(X = 5, Y = 0{,}2).$$

По матрице распределения можно найти частные, или маргинальные, распределения (табл. 16.18). Так, из последнего столбца находим вероятности p_i, с которыми случайная величина X принимает значения 5, 6, 7:

$$P(X = x_i) = P(X = x_i, Y = y_1) + P(X = x_i, Y = y_2) +$$

$$+ \ldots + P(X = x_i, Y = y_n).$$

Таблица 16.18

X	5	6	7
p	0,4	0,4	0,2

Y	0	0,1	0,2	0,3
p	0,2	0,25	0,3	0,25

Для непрерывной двумерной случайной величины (НДСВ) одной из основных характеристик, наряду с функцией распределения $F(x,y)$, является *совместная плотность распределения* $p(x,y)$. При этом

$$p(x,y)=\frac{\partial^2 F(x,y)}{\partial x \partial y}. \qquad (16.39)$$

Так как графиком функции двух переменных является некоторая поверхность, то геометрически функцию $p(x,y)$ можно истолковать как *поверхность распределения.* Отметим некоторые свойства плотности распределения.

1. Неотрицательность $p(x,y)\geq 0$.
2. Условие нормировки: $\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} p(x,y)dxdy=1$.
3. Выражение маргинальных функций распределения:

$$F(x,y)=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} p(u,v)dudv;$$

$$F(x)=F(x,+\infty)=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{+\infty} p(u,v)dudv;$$

$$F(y)=F(+\infty,y)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{y} p(u,v)dudv.$$

Дифференцируя по x и y соответственно, находим:

$$p(x)=F'(x)=\int\limits_{-\infty}^{+\infty} p(x,y)dy; \quad p(y)=F'(y)=\int\limits_{-\infty}^{+\infty} p(x,y)dx.$$

3.3. Вероятность попадания двумерной случайной величины в заданную область

Рассмотрим для простоты прямоугольную область (рис. 16.29)

$D = \{a_1 < X < a_2, \ \ b_1 < Y < b_2\}$.

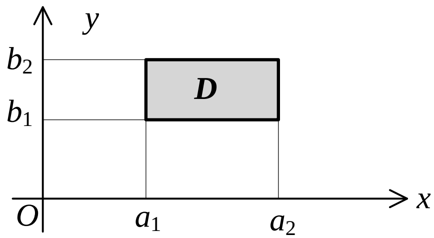


Рис. 16.29

Согласно геометрическому смыслу функции распределения $F(a_2,b_2)-F(a_1,b_2)$ есть вероятность попадания в полосу с правым верхним углом (a_2,b_2). Разность $F(a_2,b_1)-F(a_1,b_1)$ есть вероятность попадания в полосу с правым верхним углом (a_2,b_1). Следовательно, разность этих вероятностей дает вероятность попадания в область D:

$$P((x,y)\in D) = F(a_2,b_2) - F(a_1,b_2) - F(a_2,b_1) + F(a_1,b_1) . \qquad (16.40)$$

В общем случае вероятность попадания двумерной случайной величины в заданную область равна

$$P((x,y)\in D) = \iint_D p(x,y)dxdy . \qquad (16.41)$$

Пример 2. Пусть двумерная случайная величина (X,Y) имеет следующую функцию распределения:

$$F(x,y) = \begin{cases} 1 - e^{-ax} - e^{-by} + e^{-ax-by}, & x \ge 0, \ y \ge 0, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Найти:

а) вероятность попадания в область $D = \{0 < x < 1, \ \ 0 < y < 1\}$;

б) $p(x,y)$; в) $F(x)$, $F(y)$; г) $p(x)$, $p(y)$.

Решение.

а) по формуле (16.40)

$$P((x,y)\in D) = F(1,1) - F(0,1) - F(1,0) + F(0,0) =$$
$$= (1 - e^{-a} - e^{-b} + e^{-a-b}) - (-e^{-b} + e^{-b}) - (1 - e^{-a} - 1 + e^{-a}) + 0 =$$

$=1-e^{-a}-e^{-b}+e^{-a-b}$;

б) при использовании формулы (16.39) получим

$$p(x,y)=\begin{cases} abe^{-ax-by}, & x\ge 0,\ y\ge 0, \\ 0 & \text{в остальных случаях;} \end{cases} \qquad (16.42)$$

в) $F(x)=F(x,+\infty)=\begin{cases} 1-e^{-ax}, & x\ge 0, \\ 0, & x<0. \end{cases}$

$$F(y)=F(+\infty,y)=\begin{cases} 1-e^{-by}, & y\ge 0, \\ 0, & y<0; \end{cases}$$

г) при $x<0$ имеем $p(x,y)=0$, поэтому $p(x)=0$. При $x\ge 0$ получим:

$$p(x)=\int_{-\infty}^{+\infty}p(x,y)dy=\int_{0}^{+\infty}abe^{-ax-by}dy=abe^{-ax}\cdot\left.\frac{e^{-by}}{-b}\right|_{0}^{+\infty}=ae^{-bx}.$$

$$p(x)=\begin{cases} ae^{-ax}, & x\ge 0, \\ 0, & x<0; \end{cases} \quad p(y)=\begin{cases} be^{-by}, & y\ge 0, \\ 0, & y<0. \end{cases} \qquad (16.43)$$

Замечание. При изучении систем случайных величин следует обращать внимание на степень и характер их зависимости. Случайная величина Y называется независимой от случайной величины X, если закон распределения Y не зависит от того, какое значение приняла величина X. Зависимость или независимость случайных величин всегда взаимны: если величина Y не зависит от X, то и величина X не зависит от Y.

Необходимое и достаточное условие независимости случайных величин X и Y состоит в следующем:

$$p(x,y)=p(x)\cdot p(y). \qquad (16.44)$$

Например, из (16.42), (16.43) следует, что $p(x,y)=p(x)\cdot p(y)$. Следовательно, X и Y в примере 2 – независимые случайные величины.

3.4. Числовые характеристики двумерной случайной величины

Изучая случайную величину, мы ввели в рассмотрение числовые характеристики этой случайной величины. Часто бывает достаточно знать одну или несколько числовых характеристик случайной величины, которые давали бы менее полное, но достаточно наглядное представление. Аналогичные числовые характеристики можно ввести и для системы двух случайных величин. Основными числовыми характеристиками двумерной случайной величины (X, Y) являются: математические ожидания $M(X)$, $M(Y)$; дисперсии $D(X)$, $D(Y)$, корреляционный момент или ковариация $\operatorname{cov}(X, Y)$.

3.4.1. Математическое ожидание и дисперсия двумерной случайной величины. Свойства

Дискретная двумерная случайная величина (ДДСВ) задается матрицей распределения (p_{ij}), где $(p_{ij}) = P(X = x_i, Y = y_j)$.

Y X	y_i	…	y_j	…	y_n	p_i
x_i	p_{1i}	…	p_{1j}	…	p_{1n}	p_1
….	…	…	…	…	…	…
x_i	p_{i1}	…	p_{ij}	…	p_{in}	p_i
...	…	…	...	…	…	…
x_m	p_{m1}	…	p_{mj}	…	p_{mn}	p_m
p_j	p_1	…	p_j	…	p_n	

В последнем столбце матрицы помещены суммы элементов по строкам $p_i = \sum_j p_{ij} = P(X = x_i)$. В последней строке элементы равны сумме элементов по соответствующим столбцам $p_j = \sum_i p_{ij} = P(Y = y_j)$. Далее:

$$M(X) = \sum_i x_i p_i = \sum_i x_i P(X = x_i) = \sum_i \left(\sum_j x_i p_{ij} \right).$$

Аналогично для случайной величины Y:

$$M(Y)=\sum_j y_j p_j=\sum_j y_j P(Y=y_j)=\sum_j\left(\sum_i y_j p_{ij}\right).$$

Для примера 1 находим

$$M(X)=5\cdot 0{,}4+6\cdot 0{,}4+7\cdot 0{,}2=5{,}8;$$

$$M(Y)=0\cdot 0{,}2+0{,}1\cdot 0{,}25+0{,}2\cdot 0{,}3+0{,}3\cdot 0{,}25=0{,}16.$$

Совокупность $M(X)$ и $M(Y)$ представляет собой *характеристику положения* системы. Геометрически это координаты средней точки на плоскости, вокруг которой происходит рассеивание случайной точки (X,Y).

Для непрерывной двумерной случайной величины (НДСВ) с известной совместной плотностью распределения $p(x,y)$ имеем $M(X)=\int\limits_{-\infty}^{+\infty} x\,p(x)dx$. Поскольку $p(x)=\int\limits_{-\infty}^{+\infty} p(x,y)dy$, то

$$M(X)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} x\,p(x,y)dxdy. \tag{16.45}$$

Аналогично

$$M(Y)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} y\,p(x,y)dxdy. \tag{16.46}$$

Для примера 2 находим

$$M(X)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} x\cdot abe^{ax-by}dxdy=ab\int\limits_{-\infty}^{+\infty} xe^{-ax}dx\cdot\int\limits_{-\infty}^{+\infty} e^{-by}dy.$$

При этом $\int\limits_{0}^{+\infty} xe^{-ax}dx=\dfrac{1}{a^2}$ и $\int\limits_{0}^{+\infty} e^{-by}dy=\dfrac{1}{b}$. Следовательно, $M(X)=ab\cdot\dfrac{1}{a^2}\cdot\dfrac{1}{b}=\dfrac{1}{a}$. Аналогично $M(Y)=\dfrac{1}{b}$.

Для дисперсии ДДСВ имеем

$$D(X)=\sum_i (x_i-M_X)^2 p_i=\sum_i (x_i-M_X)^2 P(X=x_i)=$$

$$=\sum_i (x_i - M_X)^2 \sum_j p_{ij} = \sum_i \sum_j (x_i - M_X)^2 p_{ij}.$$

$$D(Y) = \sum_j (y_j - M_Y)^2 p_j = \sum_j (y_j - M_Y)^2 P(Y = y_i) =$$

$$=\sum_j (y_j - M_Y)^2 \sum_i p_{ij} = \sum_i \sum_j (y_j - M_Y)^2 p_{ij}.$$

Так же, как и для одномерной случайной величины, справедливы формулы

$$D(X) = M(X^2) - M^2(X),\ D(Y) = M(Y^2) - M^2(Y).$$

Так, для примера 1 находим

$$M(X^2) = 0{,}4 \cdot 25 + 0{,}4 \cdot 36 + 0{,}2 \cdot 49 = 34{,}2,$$

$$D(X) = 34{,}2 - (5{,}8)^2 = 0{,}56.$$

$$M(Y^2) = 0{,}01 \cdot 0{,}25 + 0{,}04 \cdot 0{,}3 + 0{,}09 \cdot 0{,}25 = 0{,}037,$$

$$D(Y) = 0{,}037 - (0{,}16)^2 = 0{,}0114.$$

Для НДСВ получаем

$$D(X) = \int_{-\infty}^{+\infty} (x - M_X)^2 p(x)dx = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - M_X)^2 p(x, y)dxdy$$

или $D(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2 p(x, y)dxdy - (M_X)^2$.

Аналогичные формулы можно получить и для $D(Y)$.

Дисперсии $D(X)$ и $D(Y)$ характеризуют рассеивание случайной точки (X, Y) в направлении осей Ox и Oy. Наряду с дисперсиями $D(X), D(Y)$ рассматривают средние квадратические отклонения σ_X, σ_Y.

Для примера 2 имеем

$$M(X^2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2 abe^{-ax-by} dxdy = ab \cdot \int_0^{+\infty} x^2 e^{-ax} dx \cdot \int_0^{+\infty} e^{-by} dy.$$

При этом $\int_0^{+\infty} e^{-by} dy = \frac{1}{b}$, а также

$$\int_{-\infty}^{+\infty} x^2 e^{-ax} dx = -\frac{1}{a}\int_0^{+\infty} x^2 d\left(e^{-ax}\right) = -\frac{1}{a}\left(x^2 e^{-ax}\Big|_0^{+\infty} - \int_0^{+\infty} e^{-ax}\cdot 2x dx \right) =$$

$$= \frac{2}{a}\int_{-\infty}^{+\infty} x e^{-ax} dx = \frac{2}{a}\cdot\frac{1}{a^2} = \frac{2}{a^3},$$

тогда $M\left(X^2\right) = ab\cdot\frac{2}{a^3}\cdot\frac{1}{b} = \frac{2}{a^2}$, $D(X) = \frac{2}{a^2} - \left(\frac{1}{a}\right)^2 = \frac{1}{a^2}$.

Аналогично $D(Y) = \frac{1}{b^2}$; средние квадратические отклонения: $\sigma_X = \frac{1}{a}$, $\sigma_Y = \frac{1}{b}$.

3.4.2. Ковариация случайных величин. Коэффициент корреляции и его свойства

Определение. *Ковариацией* случайных величин X и Y [обозначение $\operatorname{cov}(X,Y)$ или σ_{XY}] называется математическое ожидание произведения их отклонений

$$\operatorname{cov}(X,Y) = M\left[\left(X - M_X\right)\cdot\left(Y - M_Y\right)\right]. \qquad (16.47)$$

Замечания. 1. Иногда вместо термина «ковариация» употребляют выражение *корреляционный момент*. 2. Ковариация происходит от латинского co – совместно; vario – изменяю.

Преобразуем формулу (16.47):

$$\begin{aligned}\operatorname{cov}(X,Y) &= M[X\cdot Y - X\cdot M_Y - Y\cdot M_X + M_X\cdot M_Y] = \\ &= M(XY) - 2M_X\cdot M_Y + M_X\cdot M_Y = M(XY) - M_X\cdot M_Y.\end{aligned} \qquad (16.48)$$

Для ДДСВ формулы (16.47) и (16.48) для подсчета ковариации принимают вид

$$\operatorname{cov}(X,Y) = \sum_i\sum_j\left(x_i - M_X\right)\left(y_j - M_Y\right),$$

$$\operatorname{cov}(X,Y) = \sum_i\sum_j x_j y_j\cdot p_{ij} - M_X\cdot M_Y.$$

Для данных примера 1

$$M(XY) = \sum_{i=1}^{3}\sum_{j=1}^{4} x_i y_j p_{ij} = 5\cdot 0{,}02 + 5\cdot 0{,}1\cdot 0{,}1 + 5\cdot 0{,}2\cdot 0{,}05 +$$

$$+5\cdot 0{,}3\cdot 0{,}05+6\cdot 0{,}1+0{,}15+6\cdot 0{,}2\cdot 0{,}15+6\cdot 0{,}1\cdot 0{,}3+$$
$$+7\cdot 0{,}2\cdot 0{,}1+7\cdot 0{,}3\cdot 0{,}1=0{,}975.$$
$$\operatorname{cov}(X,Y)=0{,}975-5{,}8\cdot 0{,}16=0{,}047.$$

Некоторые свойства ковариации

1. $\operatorname{cov}(X,Y)=\operatorname{cov}(Y,X)$ (симметрия).
2. $\operatorname{cov}(X,X)=D(X)$.
3. $\operatorname{cov}(X+c_1,Y+c_2)=\operatorname{cov}(X,Y)$, где c_1,c_2 – постоянные.
4. $D(X+Y)=D(X)+D(Y)+2\operatorname{cov}(X,Y)$.
5. $\operatorname{cov}(X,Y)$ – величина размерная, $[\operatorname{cov}(X,Y)]=[X]\times[Y]$.
6. Если X и Y – независимые случайные величины, то $M(X\cdot Y)=M(X)\cdot M(Y)$ и, следовательно, $\operatorname{cov}(X,Y)=0$. Обратное неверно, т.е. из некоррелированности случайных величин не следует их независимость.

Таким образом, условия независимости более жесткие, чем условия некоррелированности.

Следствие. Если $\operatorname{cov}(X,Y)\neq 0$, то X и Y – зависимые случайные величины. В самом деле, если предположить, что X и Y – независимые случайные величины, то в силу 6-го свойства $\operatorname{cov}(X,Y)=0$, что противоречит исходному предположению.

Для двумерной случайной величины или двумерного случайного вектора (X,Y) вводится понятие *ковариационной матрицы* (обозначается $\mathbf{\Sigma}$):

$$\mathbf{\Sigma}=\begin{pmatrix}\operatorname{cov}(X,X) & \operatorname{cov}(X,Y)\\ \operatorname{cov}(Y,X) & \operatorname{cov}(Y,Y)\end{pmatrix}. \qquad (16.49)$$

Ковариационная матрица обладает следующими свойствами:

1. $\mathbf{\Sigma}$ – симметричная матрица.

2. На главной диагонали находятся дисперсии компонент случайного вектора, т.е. $\operatorname{cov}(X,X)=D(X)$ и $\operatorname{cov}(Y,Y)=D(Y)$.

Определитель матрицы $\mathbf{\Sigma}$, равный $|\mathbf{\Sigma}| = D(X)\cdot D(Y) - \operatorname{cov}^2(X,Y)$, называется *обобщенной дисперсией.* Это есть мера рассеивания двумерной случайной величины.

Для примера 1:

$$|\mathbf{\Sigma}| = D(X)\cdot D(Y) - \operatorname{cov}^2(X,Y) = \begin{vmatrix} 0{,}56 & 0{,}047 \\ 0{,}047 & 0{,}0114 \end{vmatrix} = -0{,}001975 .$$

В качестве количественной характеристики зависимости случайных величин X,Y используют безразмерную величину – *коэффициент корреляции* r_{XY}:

$$r_{XY} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \cdot \sigma_Y} . \tag{16.50}$$

При этом $|r_{XY}| \le 1$.

Для примера 1:

$$r_{XY} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{0{,}047}{\sqrt{0{,}56\sqrt{0{,}0114}}} \approx 0{,}587 .$$

Свойства коэффициента корреляции

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

2. $r_{XY} = 0$, если случайные величины X и Y независимы.

3. Если $|r_{XY}| = 1$, то между случайными величинами X и Y существует линейная функциональная зависимость.

Из независимости двух случайных величин следует их *некоррелированность*, т.е. равенство $r_{XY} = 0$. Однако некоррелированность двух случайных величин еще не означает их независимость.

4. Закон больших чисел. Центральная предельная теорема

В теории вероятностей и ее приложениях часто рассматриваются случайные величины, являющиеся, в свою очередь, суммами большого числа случайных величин. Непосредственное вычисление распределения вероятности суммы большого числа случайных величин связано со значительными трудностями.

Теоремы, относящиеся к закону больших чисел, устанавливают условия, при которых среднее арифметическое случайных величин обладает свойством устойчивости.

4.1. Закон больших чисел

Под ***законом больших чисел*** в широком смысле понимается *общий принцип, согласно которому*, по формулировке академика А.Н. Колмогорова, *совокупное действие большого числа случайных факторов приводит (при некоторых весьма общих условиях) к результату, почти не зависящему от случая*. Другими словами, при большом числе случайных величин их средний результат перестает быть случайным и может быть предсказан с большой степенью определенности.

Неравенство Маркова (лемма Чебышева). Если случайная величина X принимает только неотрицательные значения и имеет математическое ожидание, то для $A>0$ верны неравенства

$$P(X > A) \le \frac{M(X)}{A}; \qquad (16.51)$$

$$P(X \le A) \ge 1 - \frac{M(X)}{A}. \qquad (16.52)$$

Неравенство Чебышева. Для любой случайной величины, имеющей математическое ожидание и дисперсию, верны неравенства

$$P(|X-a| > \varepsilon) \le \frac{D(X)}{\varepsilon^2}; \qquad (16.53)$$

$$P(|X-a| \le \varepsilon) \ge 1 - \frac{D(X)}{\varepsilon^2}. \qquad (16.54)$$

Теорема Чебышева. Если дисперсии n независимых случайных величин X_1, X_2, …, X_n ограничены одной и той же постоянной, то при неограниченном увеличении числа n средняя арифметическая случайных величин сходится по вероятности к средней арифметической их математических ожиданий a_1, a_2, …, a_n, т.е.

$$\lim_{n\to\infty} P\left(\left|\frac{x_1+x_2+\ldots+x_n}{n}-\frac{a_1+a_2+\ldots+a_n}{n}\right|\le\varepsilon\right)=1 \qquad (16.55)$$

$$\text{или} \quad \frac{\sum_{i=1}^{n} x_i}{n} \xrightarrow{p} \frac{\sum_{i=1}^{n} a_i}{n}. \qquad (16.55а)$$

Теорема Бернулли. Частость события в n повторных независимых испытаниях, в каждом из которых оно может произойти с одной и той же вероятностью p, при неограниченном увеличении числа n сходится по вероятности к вероятности p этого события в отдельном испытании, т.е.

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)=1 \qquad (16.56)$$

$$\text{или} \quad \frac{m}{n} \xrightarrow{p} p. \qquad (16.56а)$$

Теоремы Чебышева и Бернулли (следствие теоремы Чебышева) при $n\to\infty$ устанавливают факт приближения с вероятностью, как угодно близкой к единице (*сходится по вероятности*), случайных величин $\frac{1}{n}\sum_{i=1}^{n} X_i$ и $\frac{m}{n}$ к неслучайным детерминированным величинам соответственно $\frac{1}{n}\sum_{i=1}^{n} a_i$ и p.

Так, например, при $n\to\infty$ средняя $\frac{1}{n}\sum_{i=1}^{n} X_i$ результатов измерения X_i $(i=1,2,\ldots,n)$ постоянной величины a при отсутствии систематических погрешностей сходится по вероятности к истинному значению a, а частость $\frac{m}{n}$ рождения мальчика в n независимых испытаниях – к вероятности этого события $p\approx 0{,}515$ (по многолетним данным).

4.2. Центральная предельная теорема

Центральная предельная теорема представляет группу теорем, посвященных установлению условий, при которых возникает нормальный закон распределения. Среди этих теорем важнейшее место принадлежит теореме Ляпунова.

Теорема Ляпунова. Если X_1, X_2, …, X_n – независимые случайные величины, у каждой из которых существует математическое ожидание $M(X_i) = a$, дисперсия $D(X_i) = \sigma^2$, абслютный центральный момент третьего порядка $M\left(|X_i - a_i|^3\right) = m_i$ и

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} m_i}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{\frac{3}{2}}} = 0, \tag{16.57}$$

то закон распределения суммы $\sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n$ при $n \to \infty$ неограниченно приближается к нормальному с математическим ожиданием $\sum_{i=1}^{n} a_i$ и дисперсией $\sum_{i=1}^{n} \sigma_i^2$.

Это означает, что при суммировании n случайных величин, среди которых нет резко выделяющихся по своему влиянию на рассеяние их суммы $\sum_{i=1}^{n} X_i$, закон ее распределения при $n \to \infty$ будет неограниченно приближаться к нормальному закону.

Например, если X_1, X_2, …, X_n – случайные величины, характеризующие потребление электроэнергии в квартирах многоквартирного дома, то суммарное потребление электроэнергии всего дома $\sum_{i=1}^{n} X_i$ будет иметь нормальный закон распределения.

В частности, если X_1, X_2, …, X_n *одинаково распределены*, то закон распределения их суммы $\sum_{i=1}^{n} X_i$ при $n \to \infty$ неограниченно приближается к нормальному.

Напомним, что закон нормального распределения имеет большое теоретическое значение: с его помощью выведен целый ряд других важных распределений; построены различные статистические критерии (χ^2, t и F – распределения и опирающиеся на них критерии).

ГЛАВА 17. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

В теории вероятностей выводятся правила, которые позволяют по вероятностям одних случайных событий вычислять вероятности других, с ними связанных, по числовым характеристикам и функциям распределения одних случайных величин подсчитывать функции распределения и числовые характеристики других. Но возникает вопрос: как найти эти исходные вероятности, функции распределения и числовые характеристики? Как оценить хотя бы приближенные их значения? Это является предметом исследования другой науки о массовых случайных явлениях, которая получила наименование «математическая статистика».

Математическая статистика – это раздел прикладной математики, в котором рассматриваются методы выявления и количественного описания, т.е. построения математических моделей, реально существующих массовых случайных явлений, на основании результатов повторных наблюдений.

Как наука математическая статистика возникла только в 20-м веке. Однако отдельные задачи возникали и рассматривались даже в 17-м веке.

Термин «статистика» происходит от латинского слова «статус» – состояние. Первоначально, в 18-м веке, когда статистика начала оформляться в научную дисциплину, термин «статистика» связывался с системой описания фактов, характеризующих состояние государства.

В настоящее время статистика состоит из следующих трех разделов:

1) сбор статистических сведений, т.е. сведений, характеризующих отдельные единицы каких-либо массовых совокупностей;

2) статистическое исследование полученных данных, заключающееся в выяснении тех закономерностей, которые могут быть установлены на основе данных массового наблюдения;

3) разработка приемов статистического наблюдения и анализа статистических данных. Последний раздел и составляет содержание математической статистики.

1. Основные задачи математической статистики

Математическая статистика непосредственно занимается вопросами практического использования математических моделей и методов теории вероятностей при изучении реальных явлений.

Все задачи математической статистики касаются вопросов обработки результатов наблюдений над массовыми явлениями, но в зависимости от характера изучаемой величины и целей исследования могут принимать ту или иную форму. Можно, однако, выделить наиболее типичные, основные задачи математической статистики.

1. *Оценка неизвестного закона распределения случайной величины на основании результатов измерений.*

Пусть на основании измерений случайной величины ее значения $x_1, x_2, \ldots, x_n$. По ним надо приближенно построить или оценить закон распределения случайной величины X, т.е. приближенно найти ее функцию распределения $F(x)$ или плотность распределения $p(x)$.

2. *Оценка неизвестных параметров распределения.*

Пусть из теоретических соображений известны $F(x)$ или $p(x)$ случайной величины X, но они содержат некоторые неизвестные параметры [например, $M(X)$ или $D(X)$]. На основании экспериментальных данных необходимо вычислить приближенные значения этих параметров или, как говорят, оценить эти параметры.

3. *Статистическая проверка гипотез.*

Пусть при измерении случайной величины X получены ее значения $x_1, x_2, \ldots, x_n$. На основании этого (или по другим соображениям) предполагают, что функция распределения случайной величины имеет вид $F(x)$. Необходимо выяснить, согласуются ли результаты эксперимента с гипотезой о том, что случайная величина имеет закон распределения $F(x)$.

2. Выборки и их характеристики

2.1. Генеральная совокупность и выборка

Пусть требуется изучить совокупность одинаковых (однородных) объектов относительно некоторого количественного или качественного признака. Как это возможно осуществить? Существуют два пути:

1) исследуют каждый из объектов (сплошное обследование);

2) случайно отбирают из всей совокупности ограниченное число объектов и подвергают их изучению (выборочное обследование).

Определение. *Генеральной совокупностью* (ГС) называют совокупность всех подлежащих изучению объектов интересующего нас типа или возможных результатов измерений, проводимых в неизменных условиях над одним объектом.

Из определения следует, что непрерывные ГС (состоящие из наблюдений признаков непрерывной природы) всегда бесконечны. Дискретные ГС могут быть как бесконечными, так и конечными.

Пусть N – число объектов ГС или объем ГС.

Сплошное обследование ГС бывает либо слишком трудоемко (в случае больших N), либо принципиально невозможно (в случае бесконечных ГС).

Определение. *Выборочной совокупностью* или *выборкой* называется совокупность объектов, случайно отобранных из ГС.

Каждая выборка характеризуется числом объектов n, называемым *объемом выборки.* Если занумеровать объекты выборки, то ее можно записать как $x_1, x_2, \ldots, x_n$, где x_i – элемент выборки или *варианта*; n – объем выборки.

Выборка из данной ГС – это результат ограниченного ряда наблюдений $x_1, x_2, \ldots, x_n$ случайной величины X. Таким образом, выборку можно рассматривать как некий эмпирический аналог ГС. При этом $n << N$.

Для того чтобы по данным выборки можно было делать достоверные выводы, необходимо, чтобы объекты выборки правильно ее представили. Выборка должна правильно представлять пропорции ГС, т.е. выборка должна быть *представитель-*

ной (*репрезентативной*). По закону больших чисел этого можно достигнуть, если:

– каждый объект выборки выбран случайно из ГС;

– все объекты имеют одинаковую вероятность попасть в выборку.

Вся совокупность наблюдаемых значений случайной величины X представляет собой первичный статистический материал, который подлежит дальнейшей обработке. Прежде всего этот материал следует упорядочить.

2.2. Статистическое распределение. Эмпирическая функция распределения

Если записать элементы выборки в порядке возрастания, то, таким образом, упорядоченная выборка называется *вариационным рядом.* Если несколько элементов выборки имеют одинаковое значение, то они между собой располагаются в произвольном порядке, т.е.

$$x^{(1)} \le x^{(2)} \le \ldots \le x^{(k)}.$$

Одной из характеристик вариационного ряда является *статистическое распределение* выборки или *статистический ряд.* Под ним понимается соответствие между наблюдаемыми вариантами и их частотами n_i или относительными частотами p_i^*, при этом $p_i^* = \frac{n_i}{n}$, где n – объем выборки. Статистическое распределение задается в виде табл. 17.1.

Таблица 17.1

X	x_1	$\ldots$	x_i	$\ldots$	x_k
$n_i\left(p_i^*\right)$	$n_1\left(p_1^*\right)$	$\ldots$	$n_i\left(p_i^*\right)$	$\ldots$	$n_k\left(p_k^*\right)$

Статистическое распределение изображается графически в виде *полигона* частот (относительных частот)

Определение. *Полигоном частот* называют ломаную, отрезки которой соединяют точки с координатами (x_1, n_1), (x_2, n_2), …, (x_k, n_k).

Определение. *Полигоном относительных частот* (*частостей*) называют ломаную, отрезки которой соединяют точки с координатами (x_1, p_1^*), (x_2, p_2^*), …, (x_k, p_k^*).

Пусть $x_1, x_2, \ldots, x_n$ – выборка, x – произвольное число. Обозначим через $\nu_n(x)$ – число элементов выборки, которые меньше X .

Определение. Функция

$$F_n^*(x) = p^*(X < x) = \frac{\nu_n(x)}{n} \tag{17.1}$$

называется *статистической* (*эмпирической*) *функцией распределения.* Ее график представляет собой ступенчатую кривую, которая в точках $x = x_i$ имеет скачки. Левее точки x_1 эмпирическая функция распределения $F_n^*(x)$ равна 0 , а правее точки x_n функция $F_n^*(x)$ равна 1. В точках разрыва она непрерывна слева. График $F_n^*(x)$ дает визуальное представление о графике функции распределения $F(x)$ генеральной совокупности. Для этого требуется «сгладить» ступенчатую кривую.

Пример 1. Наблюдаемые значения случайной величины X представлены в виде статистического ряда (табл. 17.2).

Таблица 17.2

X	1	3	4	6	7
n_i	2	4	6	5	3
p_i^*	$2/20$	$4/20$	$6/20$	$5/20$	$3/20$

Построить полигон частот. Составить эмпирическую функцию распределения и построить ее график.

Решение.

На плоскости xOy отмечаем точки с координатами (x_i, n_i) и соединяем их отрезками прямых, получаем полигон частот, который представлен на рис. 17.1.

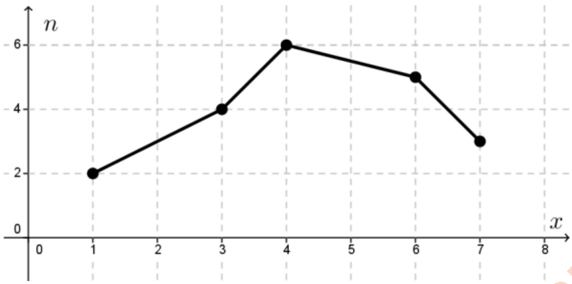


Рис. 17.1

Составим эмпирическую функцию распределения в соответствии с правилом $F_n^*(x) = p^*(X < x)$.

$$F_n^*(x) = \begin{cases} 0 & \text{при} \quad x \le 1; \\ 0{,}1 & \text{при} \quad 1 < x \le 3; \\ 0{,}3 & \text{при} \quad 3 < x \le 4; \\ 0{,}6 & \text{при} \quad 4 < x \le 6; \\ 0{,}85 & \text{при} \quad 6 < x \le 7; \\ 1 & \text{при} \quad 7 < x. \end{cases}$$

График эмпирической функции распределения приведен на рис. 17.2.

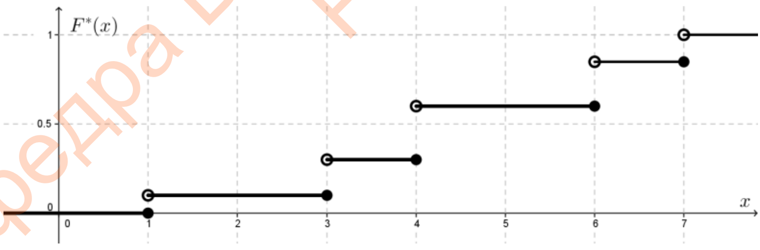


Рис. 17.2

2.3. Интервальный статистический ряд. Гистограмма

Если объем выборки велик ($n > 50$), то с точки зрения упрощения дальнейшей статистической обработки результатов наблюдений удобно перейти к так называемым «группированным» выборочным данным.

По имеющейся выборке строится *интервальный статистический ряд*. Для этого весь обследованный диапазон наблюдаемых значений $[x_{\min}; x_{\max}]$ разбивается на определенное число S равных непересекающихся интервалов. Чтобы интервальный ряд не был громоздким и отражал основные свойства распределения, следует придерживаться следующих рекомендаций.

Выбор количества интервалов существенно зависит от объема выборки n. Для примерной ориентации в выборе S можно пользоваться приближенной формулой

$$S \approx [\log_2 n] + 1 \approx 1 + 3{,}332 \lg n,$$

которую следует воспринимать как оценку снизу для S (особенно при больших n).

Длина интервала h находится по формуле Стерджесса:

$$h = \frac{x_{\max} - x_{\min}}{1 + 3{,}332 \lg n}, \qquad (17.2)$$

где $R = x_{\max} - x_{\min}$ – *размах выборки*.

Таким образом, длина интервала h и число интервалов S связаны соотношением

$$h = \frac{x_{\max} - x_{\min}}{S}.$$

При этом значение h можно округлять до удобного для вычислений значения.

Затем составляют интервальный статистический ряд распределения частот или относительных частот.

Интервалы изменения	$x_{\min}; x_{\min} + h$	$x_{\min} + h; x_{\min} + 2h$	…	$x_{\max} - h; x_{\max}$
n_i				
p_i^*				

После того как интервалы выбраны, определяют количество элементов выборки, попавших в каждый интервал. При этом придерживаются такого правила: элемент, попавший на правую границу данного интервала, относят к следующему правому интервалу.

Замечание. Иногда прибавляют по $\frac{h}{2}$ к левому и правому концам диапазона значений выборки. Таким образом, за правый конец первого интервала берут $x_{\min} + \frac{h}{2}$. При этом первый интервал обозначается как «до $x_{\min} + \frac{h}{2}$», последний – «свыше $x_{\max} - l$», $0 < l < h$.

В результате получают интервальный статистический ряд распределения частот или относительных частот (табл. 17.3).

Таблица 17.3

Интервалы изменения	До $x_{\min} + \frac{h}{2}$	$x_{\min} + \frac{h}{2}$, $x_{\min} + \frac{3}{2}h$	…	Свыше $x_{\max} - l$
n_i				
p_i^*				

Определение. *Гистограммой частот (относительных частот)* называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длины h, а высоты равны отношению $\frac{n_i}{h}$ (плотность частоты) или $\frac{p_i^*}{h}$ (плотность относительной частоты).

Сумма площадей всех прямоугольников для гистограммы частот равна объему выборки, а для гистограммы относительных частот – 1. Гистограмма – это графическое изображение интервального статистического ряда распределения, являющееся аналогом плотности распределения случайной величины X. Для большей наглядности полученную ступенчатую фигуру «сглаживают», соединяя плавной линией примерно середины соседних площадок гистограммы.

Пример 2. Имеются данные о значении ресурса до списания (в тыс. ч) химического оборудования, представленные в виде выборки объемом $n = 55$ (табл.17.4).

Таблица 17.4

14	16	20	15	18	12	13	10	17	15	12
17	11	17	13	11	17	16	12	16	13	16
12	19	18	9	7	18	15	11	11	14	13
10	16	15	17	21	13	17	16	13	15	15
18	14	16	14	10	14	8	15	16	16	14

Построить гистограмму.

Решение. По имеющейся выборке строим интервальный статистический ряд распределения. Для этого упорядочим данные в порядке возрастания (табл.17.4а).

Таблица 17.4а

7	8	9	10	10	10	11	11	11	11	12
12	12	12	13	13	13	13	13	13	14	14
14	14	14	14	15	15	15	15	15	15	15
16	16	16	16	16	16	16	16	16	17	17
17	17	17	17	18	18	18	18	19	20	21

При этом размах выборки равен

$$R = x_{\max} - x_{\min} = 21 - 7 = 14.$$

По формуле Стерджесса $h = \dfrac{x_{\max} - x_{\min}}{1 + 3{,}332 \lg n}$ находим:

$$h = \frac{14}{1 + 3{,}332 \lg 55} = \frac{14}{1 + 3{,}332 \cdot 1{,}74} \approx 2{,}06.$$

Округляем h до удобного для вычисления значения $h = 2{,}0$. Тогда число интервалов равно $S = R : h = 14 : 2 = 7$.

Составляем интервальный статистический ряд распределения частот n_i – количество данных на определенном интервале и относительных частот $p_i^* = \dfrac{n_i}{n}$. Для построения гистограммы вычисляем плотность частоты и относительной частоты в каждом интервале (табл. 17.5).

Таблица 17.5

Номер интервала	Границы интервала	Середина интервала	Частота интервала n_i	Плотность частоты $\frac{n_i}{h}$	Относительная частота $p_i^* = \frac{n_i}{n}$	Плотность относит. частоты $\frac{p_i^*}{h}$
1	[7;9)	8	2	1	0,036	0,018
2	[9;11)	10	4	2	0,073	0,036
3	[11;13)	12	8	4	0,145	0,073
4	[13;15)	14	12	6	0,218	0,109
5	[15;17)	16	16	8	0,291	0,145
6	[17;19)	18	10	5	0,182	0,091
7	[19;21]	20	3	1,5	0,055	0,027

Гистограмма частот приведена на рис. 17.3, гистограмма относительных частот – на рис. 17.4.

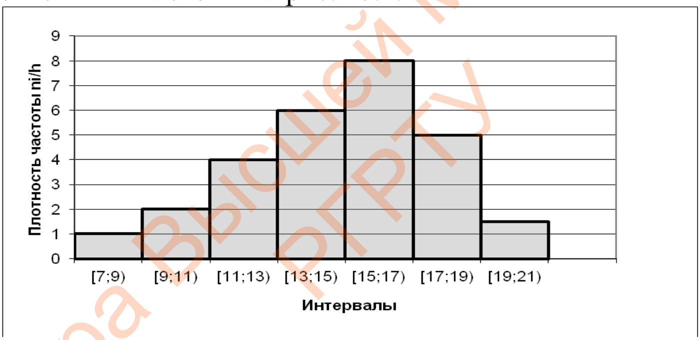


Рис. 17.3

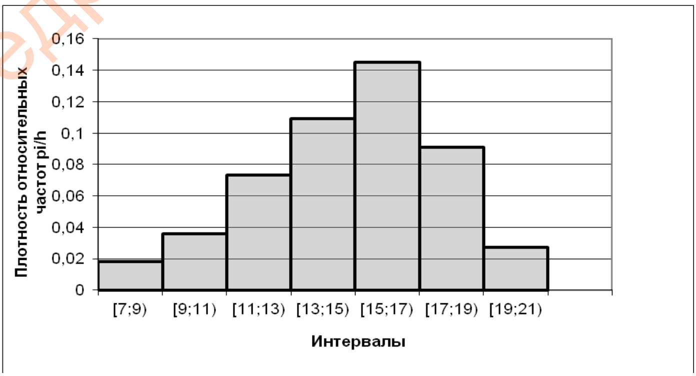


Рис. 17.4

«Сглаженная» ступенчатая фигура позволяет высказать гипотезу о том, что случайная величина X распределена по нормальному закону.

2.4. Числовые характеристики статистического распределения

Пусть в нашем распоряжении имеются данные выборки $x_1, x_2, \ldots, x_n$ случайной величины X. Можно ли по ним оценить параметры распределения X (математическое ожидание, дисперсию и т.д.)?

Определим некоторые числовые характеристики статистического распределения (табл.17.6), позволяющие найти лишь точку, около которой находится неизвестный оцениваемый параметр.

Таблица 17.6

X	x_1	x_2	…	x_k
n_i	n_1	n_2	…	n_k

Определение. *Выборочным средним* $\bar{x}$ называется среднее арифметическое всех значений выборки

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot n_i \,. \qquad (17.3)$$

Выборочное среднее можно также найти по формуле

$$\bar{x} = \sum_{i=1}^{k} x_i \cdot p_i^* \,. \qquad (17.3а)$$

Замечания. 1. Выборочное среднее является аналогом математического ожидания случайной величины.

2. В случае интервального статистического ряда в качестве x_i берут середины интервалов, а n_i – соответствующие им частоты.

Определение. *Выборочной дисперсией* D_x называется среднее арифметическое квадратов отклонений значений выборки от выборочного среднего $\bar{x}$:

$$D_x = \frac{1}{n}\sum_{i=1}^{k} n_i\left(x_i - \bar{x}\right)^2 . \qquad (17.4)$$

Вычисление выборочной дисперсии можно упростить, используя следующую формулу:

$$D_x = \overline{x^2} - \left(\bar{x}\right)^2, \qquad (17.4а)$$

где $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} x_i^2 \cdot n_i$ – средний квадрат значений выборки.

Кроме выборочного среднего и выборочной дисперсии применяются другие характеристики статистического ряда: $\sigma_x = \sqrt{D_x}$ – *выборочное среднее квадратическое отклонение;*

$S_x^2 = \frac{n}{n-1} D_x$ – *исправленная выборочная дисперсия;*

$S_x = \sqrt{\frac{n}{n-1} D_x}$ – *исправленное выборочное среднее квадратическое отклонение*, или $S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{k} n_i\left(x_i - \bar{x}\right)^2}$.

Кроме этих числовых характеристик, используются также:

– *мода* Mo – варианта, имеющая наибольшую частоту;

– *медиана* Me – варианта, делящая вариационный ряд на две равные части;

– *коэффициент вариации* $\upsilon = \frac{\sigma_x}{\bar{x}} \cdot 100\%$ (служит для сравнения величин рассеивания двух вариационных рядов: тот из рядов имеет большее рассеивание, у которого коэффициент вариации больше);

– *среднее абсолютное отклонение* $\theta_{абс} = \frac{\sum n_i \left|x_i - x\right|}{n}$ и т.д.

Пример 1. Для случайной величины X, представленной в виде статистического ряда (табл. 17.2), найдем выборочное среднее и выборочную дисперсию.

Решение. Сначала найдем выборочное среднее:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot n_i = \frac{1}{20}\sum_{i=1}^{k} x_i \cdot n_i =$$

$$= \frac{1}{20}(1\cdot 2 + 3\cdot 4 + 4\cdot 6 + 6\cdot 5 + 7\cdot 3) = 4{,}45.$$

Далее для вычисления выборочной дисперсии $D_x = \overline{x^2} - (\bar{x})^2$ предварительно найдем средний квадрат:

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} x_i^2 \cdot n_i = \frac{1}{20}\left(1^2\cdot 2 + 3^2\cdot 4 + 4^2\cdot 6 + 6^2\cdot 5 + 7^2\cdot 3\right) =$$

$$= \frac{1}{20}(2 + 36 + 144 + 180 + 147) = 25{,}45;$$

$$D_x = \overline{x^2} - (\bar{x})^2 = 25{,}45 - (4{,}45)^2 = 25{,}45 - 19{,}8025 \approx 5{,}65.$$

Найдем исправленную выборочную дисперсию:

$$S_x^2 = \frac{n}{n-1} D_x = \frac{20}{19}\cdot 5{,}65 \approx 5{,}95.$$

Ответ: $\bar{x} = 4{,}45$, $D_x \approx 5{,}65$.

Пример 2. Для случайной величины X, представленной в виде интервального ряда (табл. 17.5), найдем выборочное среднее и выборочную дисперсию.

Решение. Оформим решение в виде табл. 17.7, где в качестве x_i берем середины интервалов, а n_i – соответствующие им частоты.

Таблица 17.7

i	x_i	n_i	$x_i \cdot n_i$	x_i^2	$x_i^2 \cdot n_i$
1	8	2	16	64	128
2	10	4	40	100	400
3	12	8	96	144	1152
4	14	12	168	196	2352
5	16	16	256	256	4096
6	18	10	180	324	3240
7	20	3	60	400	1200
Сумма		55	816	1484	12568

Тогда выборочное среднее:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot n_i = \frac{1}{55}\sum_{i=1}^{7} x_i \cdot n_i = \frac{816}{55} \approx 14{,}836.$$

Далее находим средний квадрат:

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} x_i^2 \cdot n_i = \frac{1}{55}\sum_{i=1}^{7} x_i^2 \cdot n_i = \frac{12568}{55} \approx 228{,}509.$$

Поэтому выборочная дисперсия будет равна:

$$D_x = \overline{x^2} - (\bar{x})^2 \approx 228{,}509 - (14{,}836)^2 \approx 8{,}402.$$

Ответ: $\bar{x} \approx 14{,}836$, $D_x \approx 8{,}402$.

3. Оценки неизвестных параметров распределения случайных величин

Чтобы найти закон распределения в форме функции распределения $F(x)$ или плотности распределения вероятностей $p(x)$ случайной величины X из опыта, надо располагать достаточно обширным статистическим материалом, порядка нескольких сотен наблюдений.

Однако на практике часто приходится иметь дело со статистическим материалом весьма ограниченного объема – с двумя-тремя десятками наблюдений и даже меньше. Такого ограниченного материала явно недостаточно для того, чтобы найти заранее неизвестный закон распределения. Однако он позволяет хотя бы ориентировочно определить важнейшие числовые характеристики случайной величины: математическое ожидание, дисперсию и т.д.

Часто вид закона распределения известен заранее и требуется найти только его некоторые числовые параметры, от которых он зависит.

3.1. Точечные оценки неизвестных параметров распределения случайной величины по выборке

Предположим, что имеется случайная величина X, закон распределения которой содержит неизвестный параметр θ (это может быть математическое ожидание, дисперсия и т.д.). Желая найти этот параметр или вычислить его приближенное значение, производят n опытов, в результате которых получают n значений (наблюдений, измерений) случайной величины X.

Существуют два способа оценки параметров: точечный и интервальный.

Точечные методы указывают лишь точку, около которой находится неизвестный оцениваемый параметр. С помощью интервальных методов можно найти интервал, в котором с некоторой, как правило большой (выбираемой самим исследователем), вероятностью находится неизвестное значение параметра.

Определение. Любая функция $\gamma(x_1, x_2, ..., x_n)$ от результатов наблюдений случайной величины X называется *статистикой.*

Статистика $\hat{\theta}$, используемая в качестве приближенного значения неизвестного параметра θ, называется *статистической оценкой.* Таким образом, задача статистического оценивания неизвестного параметра θ заключается в построении такой функции $\hat{\theta}(x_1, x_2, ..., x_n)$ от имеющейся в распоряжении исследователя выборки, которая давала бы в определенном смысле наиболее точные приближенные значения для истинного (неизвестного нам) значения параметра θ.

Например, в качестве статистической оценки $\hat{\theta}$ неизвестного математического ожидания генеральной совокупности можно использовать

$$1)\ \frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \bar{x} \text{ или } 2)\ \frac{x_{\min} + x_{\max}}{2}.$$

Заметим, что все статистические оценки являются случайными величинами: при переходе от одной выборки к другой конкретные значения статистической оценки, подсчитанные по одной и той же формуле, будут подвержены некоторому разбросу. Однако значения статистической оценки $\hat{\theta}$, подсчитанные по разным выборкам, должны концентрироваться около истинного неизвестного значения параметра θ.

Возникает вопрос о требованиях, которые следует предъявить к статистическим оценкам, чтобы эти оценки в каком-то определенном смысле были надежными. Эти требования формулируются обычно с помощью следующих трех свойств оценок: *несмещенности, эффективности, состоятельности*.

3.1.1. Несмещенность

Предположим, что $\hat{\theta}$ – статистическая оценка неизвестного параметра θ. С этой целью подсчитываем значения $\hat{\theta}_i$, являющиеся результатом подстановки в $\hat{\theta}(x_1, x_2, ..., x_n)$ i-й по порядку выборки объемом n. Полученные значения $\hat{\theta}_1$, $\hat{\theta}_2$,..., $\hat{\theta}_i$ можно рассматривать как возможные значения случайной величины $\hat{\theta}$.

Найдем математическое ожидание $M(\hat{\theta})$ случайной величины $\hat{\theta}$ как результат усреднения по всем выборкам данного объема n. При этом можем получить: $M(\hat{\theta}) < \theta$, $M(\hat{\theta}) > \theta$, $M(\hat{\theta}) = \theta$. В последнем случае статистическую оценку называют *несмещенной.* Удовлетворение этому требованию устраняет систематическую погрешность оценивания параметра θ.

Определение. Оценка $\hat{\theta}$ неизвестного параметра θ называется *несмещенной*, если

$$M(\hat{\theta}) = \theta. \qquad (17.5)$$

Покажем, что выборочное среднее $\bar{x}$ есть несмещенная оценка неизвестного математического ожидания $M(X)$ генеральной совокупности. Не ограничивая общности, будем считать, что все элементы выборки различны. Рассматриваем x_1, x_2, …, x_n как значения независимых, одинаково распределенных случайных величин X_1, X_2, …, X_n, для которых $M(X_1) = M(X_2) = ... = M(X_n) = M(X)$.

Находим:

$$M(\bar{x}) = M\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n} M\left(\sum_{i=}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} M(x_i) =$$

$$= \frac{1}{n}\sum_{i=1}^{n} M\left(X_i\right) = \frac{1}{n}\cdot n \cdot M\left(X\right) = M\left(X\right).$$

Можно показать, что выборочная дисперсия

$$D_x \frac{1}{n}\sum_{i}(x_{\mathrm{i}} - \overline{x})^2$$

является смещенной оценкой дисперсии генеральной совокупности D. При этом $M\left(D_x\right) = \dfrac{n-1}{n} D\left(X\right)$.

Если смещение оценки удалось выяснить, то оно легко устраняется. Так, для устранения смещения выборочной дисперсии достаточно перейти к оценке S_x^2 (исправленная выборочная дисперсия)

$$S_x^2 = \frac{n}{n-1} D_x, \tag{17.6}$$

которая уже будет несмещенной. В самом деле,

$$M\left(S_x^2\right) = M\left(\frac{n}{n-1} D_x\right) = \frac{n}{n-1} M\left(D_x\right) = \frac{n}{n-1}\cdot\frac{n-1}{n}\cdot D\left(X\right) = D\left(X\right).$$

Из (17.6) следует, что требование несмещенности особенно существенно при малом количестве наблюдений. При $n \to \infty$ имеем $\dfrac{n}{n-1} \to 1$. На практике при $n < 30$ пользуются исправленной выборочной дисперсией.

3.1.2. Эффективность

Предположим, что мы имеем две несмещенные оценки $\hat{\theta}^{(1)}$ и $\hat{\theta}^{(2)}$ неизвестного параметра θ. Несмотря на то, что обе оценки являются несмещенными, значения $\hat{\theta}_i^{(1)}$ и $\hat{\theta}_i^{(2)}$, где $i = \overline{1, k}$, могут быть по-разному разбросаны около истинного значения θ.

Более тесная концентрация оценок, полученных одним из способов, около истинного значения θ, очевидно, приводит нас к мысли о большей эффективности данной оценки по сравнению с другой.

Как измеряется степень случайного разброса значений оценки θ? В качестве такой меры берется $D(\hat{\theta})$, так как в случае несмещенных оценок имеем:

$$D(\hat{\theta}) = M\left(\left(\hat{\theta} - M_{\theta}\right)^2\right) = M\left(\left(\hat{\theta} - \theta\right)^2\right).$$

Определение. Статистическая оценка $\hat{\theta}$ параметра θ называется *эффективной,* если она среди всех прочих оценок того же самого параметра имеет наименьшую дисперсию:

$$D\left(\hat{\theta}(x_1, x_2, \ldots, x_n)\right) = D_{\min}. \tag{17.7}$$

3.1.3. Состоятельность

Определение. Оценка $\hat{\theta}$ неизвестного параметра θ называется *состоятельной,* если по мере роста числа наблюдений n (т.е. при $n \to \infty$) она стремится по вероятности к оцениваемому значению θ, т.е. если для любого сколь угодно малого $\varepsilon > 0$ вероятность события $\left(\left|\theta - \hat{\theta}\right| < \varepsilon\right)$ стремится к нулю при $n \to \infty$:

$$\lim_{n \to \infty} P\left(\left|\theta - \hat{\theta}\right| < \varepsilon\right) = 1. \tag{17.8}$$

Требование состоятельности представляется необходимым для того, чтобы оценка имела практический смысл, так как в противном случае увеличение объема исходной информации не будет «приближать нас к истине». Свойство состоятельности может проявляться лишь при очень больших объемах выборки. Очевидно, если дисперсия несмещенной оценки стремится к нулю при $n \to \infty$, то такая оценка оказывается и состоятельной.

Замечание. Выборочное среднее $\bar{x}$ и исправленная дисперсия S_x^2 являются эффективными и состоятельными оценками соответственно для математического ожидания и дисперсии исследуемой случайной величины X.

3.2. Интервальные оценки параметров распределения

Ранее были рассмотрены способы получения точечных оценок неизвестных параметров по выборочным данным. Как далеко может находиться значение случайной величины $\hat{\theta}$ от неизвестного числа θ, оценкой которого оно является?

Мерой точности оценки $\hat{\theta}$ неизвестного параметра θ служит величина $\left|\theta-\hat{\theta}\right|$ – расстояние от случайной величины $\hat{\theta}$ до неизвестного параметра. В частности, нельзя ли указать такую величину δ, которая с заданной вероятностью p, близкой к единице, гарантировала бы выполнение неравенства $\left|\theta-\hat{\theta}\right|<\delta$?

Определение. Случайный интервал $\left(\hat{\theta}-\delta;\hat{\theta}+\delta\right)$, границы которого полностью определяются выборкой и для которого выполняется равенство

$$P\left(\left|\theta-\hat{\theta}\right|<\delta\right)=p \tag{17.9}$$

или $$P\left(\hat{\theta}-\delta<\theta<\hat{\theta}+\delta\right)=p, \tag{17.9a}$$

называется *доверительным интервалом* для параметра θ, соответствующим *доверительной вероятности* p.

Доверительную вероятность p называют также *надежностью,* а величину $\alpha=1-p$ *уровнем значимости*. Обычно в качестве p используют значения 0,90; 0,95; 0,99; 0,999. Ширина доверительного интервала зависит:

1) от объема выборки n (уменьшается с ростом n);
2) от p (увеличивается при $p\to 1$).

3.2.1. Доверительные интервалы для математического ожидания и дисперсии в случае нормального распределения

3.2.1.1. Доверительный интервал для математического ожидания при известной дисперсии σ^2

Будем рассматривать значения выборки $x_1,x_2,\ldots,x_n$ как реализации одинаково распределенных случайных величин $X_1,X_2,\ldots,X_n$. При этом $M(X_i)=m$, $D(X_i)=\sigma^2$, $X_i\sim N\left(m,\sigma^2\right)$. Известно, что линейная функция от нормальных случайных величин является нормальной. При этом для $\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ имеем $M(\overline{x})=m$ и $D(\overline{x})=\frac{\sigma^2}{n}$.

Замечание. Докажем, что $D(\overline{x})=\frac{\sigma^2}{n}$.

Действительно, имеем $D(x_i)=\sigma^2$. Тогда

$$D(\overline{x})=D\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)=\frac{1}{n^2}\cdot D\left(\sum_{i=1}^{n}x_i\right)=\frac{nD(x_i)}{n^2}=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}.$$

Считаем оценкой математического ожидания СВ выборочное среднее, т.е. $\hat{\theta}=\overline{x}$ и $\theta=m$.

Доверительный интервал будем искать из условия (17.9):

$$P(|\overline{x}-m|<\delta)=p.$$

С другой стороны, для нормального распределения

$$P(|\overline{x}-m|<\delta)=2\Phi\left(\frac{\delta}{\sigma(\overline{x})}\right). \tag{17.10}$$

Подставляя $\sigma(\overline{x})=\frac{\sigma}{\sqrt{n}}$ в (17.10), получаем

$$P(|\overline{x}-m|<\delta)=2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right).$$

Обозначая $t=\frac{\delta\sqrt{n}}{\sigma}$, приходим к выражению $\delta=\frac{t\sigma}{\sqrt{n}}$, поэтому

$$P\left(|\overline{x}-m|<\frac{t\sigma}{\sqrt{n}}\right)=2\Phi(t)=p \text{ или}$$

$$P\left(\overline{x}-\frac{t\sigma}{\sqrt{n}}<m<\overline{x}+\frac{t\sigma}{\sqrt{n}}\right)=2\Phi(t)=p,$$

где $\Phi(x)$ – нормированная функция Лапласа (приложение 2):

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

Следовательно, доверительный интервал в рассматриваемом случае имеет вид $\left(\bar{x}-\frac{t\sigma}{\sqrt{n}};\bar{x}+\frac{t\sigma}{\sqrt{n}}\right)$. Очевидно, что точность оценки равна $\frac{t\sigma}{\sqrt{n}}$.

Таким образом, можно утверждать, что с вероятностью p доверительный интервал $\left(\bar{x}-\frac{t\sigma}{\sqrt{n}};\bar{x}+\frac{t\sigma}{\sqrt{n}}\right)$ накрывает неизвестное значение математического ожидания.

Для нахождения доверительного интервала:

1) по заданному значению p из равенства $\Phi(t)=\frac{p}{2}$ по таблицам функции $\Phi(x)$ (приложение 2) находим аргумент t;

2) вычисляем точность оценки $\frac{t\sigma}{\sqrt{n}}$;

3) определяем доверительный интервал:

$$\left(\bar{x}-\frac{t\sigma}{\sqrt{n}};\bar{x}+\frac{t\sigma}{\sqrt{n}}\right). \tag{17.11}$$

Пример. По выборке

10,5; 10,8; 11,2; 10,9; 10,4; 10,6; 10,9; 11,0; 10,3; 10,8

из нормально распределенной генеральной совокупности с $\sigma^2=0{,}06$ найти доверительный интервал для математического ожидания с доверительной вероятностью $0{,}9$.

Решение. Так как $p = 0{,}9$ и $n = 10$ (количество данных), то из равенства $\Phi(t) = \frac{p}{2} = 0{,}45$ по таблице приложения 2 находим $t = 1{,}645$. Далее, $\frac{t\sigma}{\sqrt{n}} = 1{,}645 \cdot \frac{\sqrt{0{,}06}}{\sqrt{10}} = 0{,}13$.

Выборочная средняя равна

$$\bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = \frac{107{,}4}{10} = 10{,}74.$$

Границы доверительного интервала:

$$\bar{x} - \frac{t\sigma}{\sqrt{n}} = 10{,}74 - 0{,}13 = 10{,}61; \quad \bar{x} + \frac{t\sigma}{\sqrt{n}} = 10{,}74 + 0{,}13 = 10{,}87.$$

Таким образом, $10{,}61 < m < 10{,}87$.

Доверительная вероятность $0{,}9$ указывает на то, что если произведено достаточно большое число выборок одинакового объема $n = 10$, то $90\,\%$ из них определяет такие доверительные интервалы, в которых математическое ожидание m действительно заключено.

3.2.1.2. Доверительный интервал для математического ожидания при неизвестной дисперсии

В случае когда σ^2 неизвестна, доверительный интервал имеет вид:

$$\left(\bar{x} - \frac{t_p S_x}{\sqrt{n}};\ \bar{x} + \frac{t_p S_x}{\sqrt{n}}\right), \qquad (17.12)$$

где значение t_p находится из таблицы квантилей распределения Стьюдента (приложение 5), $S_x = \sqrt{\frac{n}{n-1} D_x}$ – исправленное выборочное среднее квадратическое отклонение, или $S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{k} n_i (x_i - \bar{x})^2}$.

3.2.1.3. Доверительный интервал для дисперсии с известным математическим ожиданием

$$\frac{\tau D_x}{\chi^2_\beta(\tau)} < \sigma^2 < \frac{\tau D_x}{\chi^2_{1-\beta}(\tau)}, \qquad (17.13)$$

где D_x – выборочная дисперсия, $\chi^2_\beta(\tau)$ и $\chi^2_{1-\beta}(\tau)$ – квантили порядка β и $1-\beta$ соответственно распределения хи-квадрат с τ степенями свободы; $\beta = \frac{1+p}{2} = 1 - \frac{\alpha}{2}$.

Замечание. Число степеней свободы определяется по формуле $\tau = s - r - 1$, где s – число частичных интервалов наблюдения; r – число параметров предполагаемого распределения, которые оценены по данным выборки.

3.2.1.4. Доверительный интервал для дисперсии с неизвестным математическим ожиданием

$$\frac{(\tau-1)S_x^2}{\chi^2_\beta(\tau-1)} < \sigma^2 < \frac{(\tau-1)S_x^2}{\chi^2_{1-\beta}(\tau-1)}, \qquad (17.14)$$

где S_x^2 – исправленная выборочная дисперсия, $\chi^2_\beta(\tau)$ и $\chi^2_{1-\beta}(\tau)$ – квантили порядка β и $1-\beta$ соответственно распределения хи-квадрат с τ степенями свободы; $\beta = \frac{1+p}{2} = 1 - \frac{\alpha}{2}$.

4. Статистическая проверка гипотез

На разных стадиях статистического исследования возникает необходимость в формулировке и экспериментальной проверке некоторых предположительных утверждений (гипотез) относительно природы или величины неизвестных параметров. По своему прикладному содержанию высказываемые в ходе статистической обработки данных гипотезы можно подразделить на два основных типа.

1. Гипотезы о типе закона распределения исследуемой случайной величины X.

2. Гипотезы о числовых значениях параметров исследуемой генеральной совокупности.

Существуют и другие типы гипотез.

Определение. Процедура обоснованного сопоставления высказанной гипотезы с имеющимися в нашем распоряжении выборочными данными $x_1, x_2, ..., x_n$, осуществляемая с помощью того или иного *статистического критерия*, называется *статистической проверкой гипотез*.

4.1. Постановка задачи

Пусть в результате эксперимента получена выборка $x_1, x_2, ..., x_n$ значений случайной величины X. Будем обозначать в дальнейшем высказанную нами гипотезу через H_0. Ее называют *нулевой (основной)* гипотезой. Наряду с выдвинутой гипотезой рассматривают и противоречащую ей гипотезу H_1, называемую *конкурирующей (альтернативной)*. Наша цель – проверить, не противоречит ли высказанная нами гипотеза H_0 имеющимся выборочным данным.

Выбирая окончательно в качестве рабочей одну из гипотез – нулевую или конкурирующую, мы используем следующую логическую схему (алгоритм) (рис. 17.5).

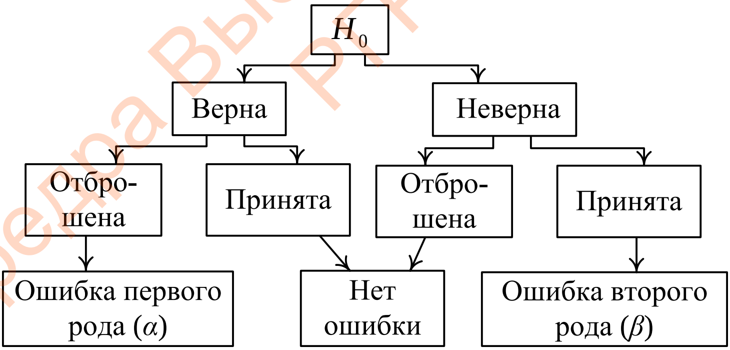


Рис. 17.5

В первом приближении можно считать, что нам одинаково «вредны» ошибки как первого, так и второго рода. Более актуальным является вопрос – а как их избежать или хотя бы снизить вероятность их появления?

Содержательный смысл статистического критерия заключается в том, что с помощью него определяется мера расхождения выборочных данных с гипотезой H_0.

По своему назначению и характеру решаемых задач статистические критерии чрезвычайно разнообразны. Однако их объединяет общность логической схемы, по которой они строятся.

4.2. Общая логическая схема статистического критерия

1. Выдвигается гипотеза H_0.

2. Задаются величиной так называемого *уровня значимости* критерия α, который есть вероятность ошибочного отклонения гипотезы H_0.

В какой-то небольшой доле случаев α гипотеза H_0 может оказаться отвергнутой, в то время как на самом деле она является справедливой, или, наоборот, в какой-то небольшой доле случаев β мы можем принять нашу гипотезу, в то время как на самом деле она является ошибочной, а справедлива гипотеза H_1. При фиксированном объеме выборки обычно задаются величиной α вероятности ошибочного отклонения проверяемой гипотезы H_0 (нулевой). Эту вероятность ошибочного отклонения гипотезы H_0 принято называть *ошибкой первого рода* или уровнем значимости. Как правило, пользуются некоторыми стандартными значениями уровня значимости, к ним можно отнести величины $\alpha = 0{,}1;\ 0{,}05;\ 0{,}025;\ 0{,}01;\ 0{,}005;\ 0{,}001$.

Особенно распространенной является величина уровня значимости, равная $0{,}05$. Она означает, что в среднем в пяти случаях из ста мы будем ошибочно отвергать высказанную гипотезу.

3. Задаются некоторой функцией от результатов наблюдения – *критической статистикой* $\gamma^{(n)} = \gamma(x_1, \ldots, x_n)$. Критическая статистика, как и всякая функция от результатов наблюдения, является случайной величиной и в предположении справедливости гипотезы H_0 подчинена некоторому хорошо изу-

ченному (затабулированному) закону распределения с плотностью $p_\gamma(u)$.

Общий содержательный смысл критической статистики таков: как правило, им определяется мера расхождения имеющихся в нашем распоряжении выборочных данных $x_1, x_2, \ldots, x_n$ с высказанной гипотезой H_0.

4. Из таблиц распределения $p_\gamma(u)$ при заданном α находятся две точки $\gamma_{\frac{\alpha}{2}}^{\min}$ и $\gamma_{\frac{\alpha}{2}}^{\max}$, разделяющие всю область значений случайной величины $\gamma^{(n)}$ на три части: область неправдоподобно малых (I) (вероятность попадания в эту область равна $\frac{\alpha}{2}$); неправдоподобно больших (III) (вероятность попадания в эту область равна $1-\frac{\alpha}{2}$); естественных или правдоподобных (II) (в условиях справедливости гипотезы H_0) значений случайной величины (рис.17.6).

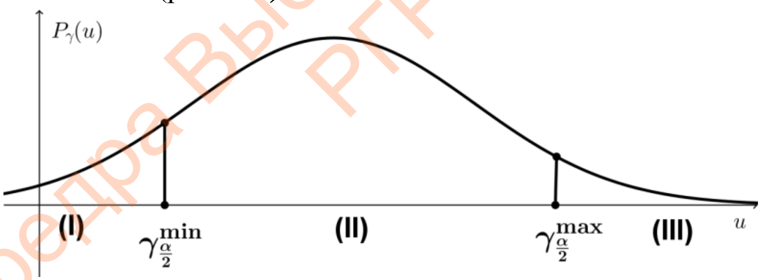


Рис. 17.6

5. В функцию $\gamma^{(n)}$ подставляются выборочные данные $x_1, x_2, \ldots, x_n$. Если вычисленное значение принадлежит области правдоподобных значений $\gamma^{(n)}$, то гипотеза H_0 считается *не противоречащей* выборочным данным. В противном случае делается вывод, что $\gamma^{(n)}$ на самом деле не подчиняется закону

$p_\gamma(u)$. Этот вывод, как легко понять, сопровождается вероятностью ошибки, равной α (ошибка первого рода).

4.3. Критерий согласия Пирсона (хи-квадрат)

Пусть в результате наблюдения (эксперимента) получена выборка $x_1, x_2, ..., x_n$ значений случайной величины X. Необходимо проверить гипотезу о том, что выборка произведена из генеральной совокупности с функцией распределения $F(x)$.

Критерием согласия называют статистический критерий проверки гипотезы о предполагаемом законе неизвестного распределения.

Ограничимся описанием применения критерия Пирсона к проверке гипотезы о нормальном распределении генеральной совокупности. Заметим, что критерий Пирсона применяется и для других распределений.

1. Весь интервал наблюдения делим на промежутки

$$(-\infty; x_1],\ (x_1; x_2],\ (x_2; x_3],\ \ldots,\ (x_{s-1}; +\infty).$$

2. Находим эмпирические частоты n_i, выборочное среднее $\bar{x}$, исправленное выборочное среднее квадратическое отклонение S_x.

3. Вычисляем теоретические частоты $n_i^{\text{т}}$ по формуле

$$n_i^{\text{т}} = n \cdot p_i = n \cdot \left[\Phi\left(\frac{x_i - \bar{x}}{S_x} \right) - \Phi\left(\frac{x_{i-1} - \bar{x}}{S_x} \right) \right],$$

где $i = \overline{1, s}$, $\Phi(x)$ – функция Лапласа (нормированная).

4. Вычисляем значение критической статистики $\gamma^{(n)}$ по формуле

$$\gamma^{(n)} = \sum_{i=1}^{s} \frac{\left(n_i - n_i^{\text{т}}\right)^2}{n_i^{\text{т}}}. \qquad (17.15)$$

Случайная величина $\gamma^{(n)}$, определенная по формуле (17.15), при $n \to \infty$ имеет распределение хи-квадрат χ^2 с τ степенями свободы и плотностью распределения $p_\tau\left(\chi^2\right)$. Таким образом,

$$\gamma^{(n)} \equiv \chi^2 .$$

Поэтому обозначим

$$\chi^2_{набл} = \sum_{i=1}^{s} \frac{\left(n_i - n_i^{\text{T}}\right)^2}{n_i^{\text{T}}} . \qquad (17.16)$$

Для оценивания согласия опытных данных с принятой гипотезой о законе распределения $F(x)$ назначают доверительную вероятность p или уровень значимости $\alpha = 1 - p$. Далее из уравнения $\int\limits_0^{\chi^2_{крит}} p_\tau\left(\chi^2\right)dz = p$ или $\int\limits_{\chi^2_{крит}}^{+\infty} p_\tau\left(\chi^2\right)dz = \alpha$ находят величину $\chi^2_{крит}$, называемую квантилью порядка p распределения хи-квадрат. С вероятностью p, численно равной площади заштрихованной части фигуры на рис. 17.7, случайная величина не превысит $\chi^2_{крит}$.

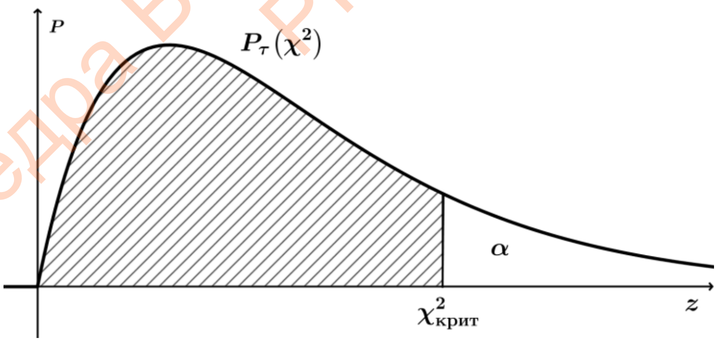


Рис. 17.7

5. Определяем число степеней свободы $\tau = s - r - 1$, где s – число частных интервалов наблюдения; r – число параметров предполагаемого распределения, которые оценены по данным

выборки. В нашем случае оцениваются два параметра (математическое ожидание и дисперсия). Следовательно, $r = 2$ и $\tau = s - 3$.

6. По заданному значению α (уровень значимости) или по заданной доверительной вероятности $p = 1 - \alpha$ по таблице приложения 4 находим значение $\chi^2_{\alpha,\tau} = \chi^2_{крит}$.

7. Вывод: если $\chi^2_{набл} \geq \chi^2_{крит}$, то отклоняем принятую гипотезу относительно вида распределения. В противном случае, т.е. при $\chi^2_{набл} < \chi^2_{крит}$ считаем гипотезу не противоречащей экспериментальным данным.

Пример. Интервалы наблюдения и эмпирические частоты представлены в табл.17.8.

Таблица 17.8

i	1	2	3	4	5	6
x	(-∞;2,62]	(2,62;3,16]	(3,16;3,7]	(3,7;4,24]	(4,24;4,78]	(4,78;+∞]
n_i	1	1	5	5	5	3

При этом $\bar{x} = 3{,}978$; S_x=0,739; n=20.

Приняв в качестве нулевой гипотезу H_0: генеральная совокупность, из которой извлечена выборка, имеет нормальное распределение, проверить ее, пользуясь критерием Пирсона при уровне значимости α=0,01.

Решение

Вычисляем теоретические частоты $n_i^{\text{т}}$, учитывая, что $\Phi(-x) = -\Phi(x)$:

$$1)\ n_1^{\text{т}} = n \cdot p_1 = n \cdot \left[\Phi\left(\frac{x_1 - \bar{x}}{S_x}\right) - \Phi\left(\frac{x_0 - \bar{x}}{S_x}\right)\right] =$$

$$= 20 \cdot \left[\Phi\left(\frac{2{,}62 - 3{,}978}{0{,}739}\right) - \Phi(-\infty)\right] = 20 \cdot [\Phi(-1{,}838) - \Phi(-\infty)] =$$

$$= 20 \cdot [-0{,}467 + 0{,}5] = 20 \cdot 0{,}033 = 0{,}66;$$

$$2)\ n_2^{\text{т}} = n \cdot p_2 = n \cdot \left[\Phi\left(\frac{x_2 - \bar{x}}{S_x}\right) - \Phi\left(\frac{x_1 - \bar{x}}{S_x}\right)\right] =$$

$$= 20 \cdot \left[\Phi\left(\frac{3{,}16 - 3{,}978}{0{,}739} \right) - \Phi(-1{,}838) \right] = 20 \cdot [\Phi(-1{,}107) - \Phi(-1{,}838)] =$$

$= 20 \cdot [-0{,}366 + 0{,}467] = 20 \cdot 0{,}101 = 2{,}02$ и т.д.

Эти расчеты удобно свести в таблицу.

i	x_i	x_{i+1}	$x_i - \bar{x}$	$x_{i+1} - \bar{x}$	$z_i = \frac{x_i - \bar{x}}{S_x}$	$z_{i+1} = \frac{x_{i+1} - \bar{x}}{S_x}$
1	$-\infty$	2,62		-1,358	$-\infty$	-1,838
2	2,62	3,16	-1,358	-0,818	-1,838	-1,107
3	3,16	3,7	-0,818	-0,278	-1,107	-0,376
4	3,7	4,24	-0,278	0,262	-0,376	0,355
5	4,24	4,78	0,262	0,802	0,355	1,085
6	4,78	$+\infty$	0,802		1,085	$+\infty$

Далее находим теоретические вероятности $p_i = \Phi(z_{i+1}) - \Phi(z_i)$ и теоретические частоты $n_i^{\text{т}} = n \cdot p_i$, где $n = 20$, а значения $\Phi(x)$ находим по таблице приложения 2.

z_i	z_{i+1}	$\Phi(z_i)$	$\Phi(z_{i+1})$	p_i	$n_i^{\text{т}}$
$-\infty$	-1,838	-0,5	-0,467	0,033	0,661
-1,838	-1,107	-0,467	-0,366	0,101	2,022
-1,107	-0,376	-0,366	-0,147	0,219	4,384
-0,376	0,355	-0,147	0,139	0,285	5,703
0,355	1,085	0,139	0,361	0,223	4,451
1,085	$+\infty$	0,361	0,5	0,139	2,778
			Σ	1	20

Вычисляем $\chi^2_{набл} = \sum_{i=1}^{s} \frac{(n_i - n_i^{\text{т}})^2}{n_i^{\text{т}}}$, для чего составим расчетную таблицу.

i	n_i	$n_i^{\text{т}}$	$n_i - n_i^{\text{т}}$	$(n_i - n_i^{\text{т}})^2$	$\frac{(n_i - n_i^{\text{т}})^2}{n_i^{\text{т}}}$	n_i^2	$\frac{n_i^2}{n_i^{\text{т}}}$
1	1	0,661	0,339	0,1148	0,174	1	1,512
2	1	2,022	-1,022	1,0449	0,517	1	0,495
3	5	4,384	0,616	0,3789	0,086	25	5,702
4	5	5,703	-0,703	0,4939	0,087	25	4,384

5	5	4,451	0,549	0,3011	0,068	25	5,616
6	3	2,778	0,222	0,0492	0,018	9	3,24
Σ	20	20		$\chi^2_{набл}$	0,950		20,95

Для контроля вычислений формулу (17.16) преобразуют к виду $\chi^2_{набл} = \sum_{i=1}^{s} \frac{n_i^2}{n_i^{\mathrm{T}}} - n$, в которой в данном случае имеем $\chi^2_{набл} = 20{,}95 - 20 = 0{,}95$.

Задан уровень значимости $\alpha = 0{,}01$. Найдем число степеней свободы τ, учитывая, что число интервалов наблюдений $s = 6$, поэтому $\tau = s - 3 = 6 - 3 = 3$. По таблице χ^2 для $\alpha = 0{,}01$ и $\tau = 3$ находим критическое значение $\chi^2_{крит} = \chi^2_{0,99}(3) = 11{,}3$ (см. приложение 4).

Так как $\chi^2_{набл} = 0{,}942 < \chi^2_{крит} = 11{,}3$, то считаем гипотезу о нормальном распределении генеральной совокупности не противоречащей экспериментальным данным.

5. Корреляция и регрессия

5.1. Основные понятия

Отметим факт существования различных зависимостей между переменными. В случае функциональной зависимости Y от X каждому значению переменной x по некоторому вполне определенному закону ставится в соответствие единственное значение переменной y.

При статистической зависимости каждому значению x случайной величины X отвечает множество возможных значений y случайной величины Y с известным законом условного распределения $p_x(y)$. Поэтому говорить здесь о функциональной зависимости Y от X в общепринятом смысле не приходится.

При исследовании статистической зависимости возникают следующие основные вопросы:

1. Имеется ли вообще какая-либо связь между исследуемыми переменными и как измерить тесноту этой связи?

2. Каков общий математический вид связи между Y и X, т.е. как определяется общая структура соответствующей математической модели?

3. Как, отправляясь от выбранной общей структуры модели, проводя обработку исходных данных, получить оценку $\hat{\theta}_0, \hat{\theta}_1, \ldots$ неизвестных параметров модели?

Корреляционный и регрессионный анализы являются смежными разделами математической статистики и предназначены для изучения по выборочным данным статистической зависимости ряда величин, где некоторые являются случайными. Понятия корреляции и регрессии появились в середине XIX века в работах Гальтона и Пирсона. Корреляция – от лат. «correlation» – соотношение, взаимосвязь. Термин «регрессия» (от лат. «regression» – движение назад) ввел Гальтон. Он изучал связь между ростом детей и ростом их родителей. При этом обнаружил явление «регрессия к среднему», т.е. рост детей очень высоких родителей имел тенденцию быть ближе к средней величине.

Если X, Y – случайные величины, то при изучении статистической зависимости между Y и X, говорят о корреляции между Y и X, о корреляционной зависимости Y и X. Примерами корреляционной зависимости является статистическая взаимосвязь между отдельными частями человеческого тела (длиной руки и длиной ноги, весом и ростом и т.д.), обусловленная их взаимосвязью и влиянием определенных первичных факторов, связанных, прежде всего, с наследственностью.

Если же X – неслучайная величина, значения которой задаются перед началом наблюдений, Y – зависимая случайная величина, то говорят о регрессионной зависимости между Y и X, о регрессии между Y и X. Примером регрессионной зависимости служит зависимость между урожайностью определенной с/х культуры и влияющими на нее природными и экономическими факторами. Здесь из внестатистических соображений ясно, что дожди и удобрения влияют на урожай, а не наоборот. Следовательно, надо изучать зависимость урожайности от количества дождей и других природно-экономических факторов. В регрессионном анализе изучают функциональную зависимость условной средней $M_x(y)$ случайной величины Y от X, т.е.

$$M_x(y) = f(x, \theta_0, \theta_1, \theta_2, \ldots), \tag{17.17}$$

где $\theta_0, \theta_1, \theta_2, \ldots$ – неизвестные параметры (коэффициенты регрессии), оценки которых нужно найти по результатам наблюдений. При этом вид функции предполагается известным.

5.2. Вычисление коэффициента корреляции случайных величин по выборкам

Рассмотрим для примера парную корреляцию, занимающуюся изучением характеристик двух случайных величин Y и X. По выборке (x_1, y_1), (x_2, y_2), …, (x_n, y_n) общую картину их взаимной изменчивости можно получить, изобразив на координатной плоскости все выборочные точки. Это изображение называется *корреляционным полем*. Если заранее известно, что между исследуемыми величинами имеется *линейная связь* [функция f в (17.17) – линейная], а их совместное распределение нормально, то характеристикой тесноты связи является *коэффициент корреляции*, определяемый соотношением:

$$r_{XY} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{M[(X - M_X)\cdot(Y - M_Y)]}{\sigma_X \cdot \sigma_Y}. \tag{17.18}$$

По данным выборки можно получить оценку $r = \hat{r}_{XY}$ коэффициента корреляции r_{XY} (выборочный коэффициент корреляции):

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}. \tag{17.19}$$

Формулу (17.19) удобно для расчетов представить в виде

$$r = \frac{\sum_i x_i y_i - \frac{\sum x_i \cdot \sum y_i}{n}}{\sqrt{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}} \cdot \sqrt{\sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}}}. \qquad (17.20)$$

Отметим, что $|r| \le 1$. Если Y и X независимы, то $r = 0$. При $|r| = 1$ имеем функциональную зависимость между Y и X.

Если $r > 0$, то говорят о *положительной корреляции*, т.е. большие значения одной случайной величины встречаются обычно с большими значениями другой, а меньшие значения величин также сочетаются друг с другом.

При $r < 0$ имеем *отрицательную корреляцию*, обладающую противоположными свойствами. Чем больше $|r|$, тем теснее связь между переменными и с тем большим основанием найденная взаимосвязь может быть использована. Для качественной оценки тесноты корреляционной связи можно использовать таблицу Чеддока (табл. 17.9).

Таблица 17.9

Диапазон	$0{,}1 \div 0{,}3$	$0{,}3 \div 0{,}5$	$0{,}5 \div 0{,}7$	$0{,}7 \div 0{,}9$	$0{,}9 \div 0{,}99$
Характеристика тесноты связи	Слабая	Умеренная	Заметная	Высокая	Весьма высокая

5.3. Линейное уравнение регрессии и определение его коэффициентов по методу наименьших квадратов

В случае парной линейной регрессии имеется одна неслучайная величина X (она называется также фактором) и одна зависимая случайная величина Y. Функция регрессии $f(x) = M_x(y)$, описывающая поведение условных средних случайной величины Y при $X = x$ в зависимости от изменения x, является линейной, т.е.

$$f(x) = \theta_0 + \theta_1 x, \tag{17.21}$$

где θ_0, θ_1 – некоторые константы (неизвестные параметры регрессии).

Предполагается, что имеется выборка значений y_i случайной величины Y и отвечающих ей значений x_i детерминированной величины X, т.е. набор n пар чисел (x_1, y_1), (x_2, y_2), …, (x_n, y_n). Требуется по этому набору получить оценки параметров θ_0, θ_1 линейной регрессии. Подставив их в (17.21), получим *выборочное уравнение регрессии*:

$$y = \hat{\theta}_0 + \hat{\theta}_1 x, \tag{17.22}$$

которое является оценкой взаимосвязи $f(x) = \theta_0 + \theta_1 x$. Как это сделать?

Статистические оценки $\hat{\theta}_0, \hat{\theta}_1$ параметров регрессии θ_0, θ_1 выбираются таким образом, чтобы эмпирические значения $\hat{y}_i = \hat{\theta}_0 + \hat{\theta}_1 x_i$ как можно ближе приближались к фактическим значениям y_i. В качестве меры близости обычно выбирают сумму квадратов отклонений

$$Q(\theta_0, \theta_1) = \sum_{i=1}^{n} \left[y_i - (\theta_0 + \theta_1 x_i) \right]^2, \tag{17.23}$$

являющуюся функцией двух переменных θ_0, θ_1. Оценки $\hat{\theta}_0, \hat{\theta}_1$, обладающие наилучшими свойствами, определяются из условия минимума функции $Q(\theta_0, \theta_1)$.

Применим метод наименьших квадратов (МНК). Запишем необходимые условия минимума функции $Q(\theta_0, \theta_1)$, т.е. находим частные производные и приравниваем их к нулю. Имеем

$$\begin{cases} \dfrac{\partial Q}{\partial \theta_0} = -2\sum_i (y_i - \theta_0 - \theta_1 x_i) = 0, \\ \dfrac{\partial Q}{\partial \theta_1} = -2\sum_i (y_i - \theta_0 - \theta_1 x_i) x_i = 0, \end{cases}$$

откуда получаем систему линейных алгебраических уравнений (СЛАУ) для определения оценок $\hat{\theta}_0, \hat{\theta}_1$, которая называется *системой нормальных уравнений*:

$$\begin{cases} \hat{\theta}_0 n + \hat{\theta}_1 \sum x_i = \sum y_i, \\ \hat{\theta}_0 \sum x_i + \hat{\theta}_1 \sum x_i^2 = \sum x_i y_i. \end{cases} \qquad (17.24)$$

Разделив первое уравнение системы (17.24) на n, получим

$$\hat{\theta}_0 + \hat{\theta}_1 \bar{x} = \bar{y}. \qquad (17.25)$$

Подставляя $\hat{\theta}_0 = \bar{y} - \hat{\theta}_1 \bar{x}$ во второе уравнение системы (17.24) и решая его, получаем

$$\hat{\theta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}. \qquad (17.26)$$

Определяя $\hat{\theta}_1$ по (17.26), находим из (17.25) значение $\hat{\theta}_0$.

Выборочное уравнение регрессии (17.22) можно представить в другой форме. Имеем: $\hat{y} = \hat{\theta}_0 + \hat{\theta}_1 x$ и из (17.25) $\bar{y} = \hat{\theta}_0 + \hat{\theta}_1 \bar{x}$. Вычитая, получаем

$$\hat{y} - \bar{y} = \hat{\theta}_1 (x - \bar{x}), \qquad (17.27)$$

т.е. эмпирическая линия регрессии проходит через точку с координатами $(\bar{x}, \bar{y})$.

Между выборочным коэффициентом корреляции r и оценкой $\hat{\theta}_1$ существует следующее соотношение:

$$r = \hat{\theta}_1 \frac{\sigma_x}{\sigma_y} \text{ или } \hat{\theta}_1 = r \frac{\sigma_y}{\sigma_x},$$

где $\sigma_x = \sqrt{\dfrac{\sum x_i^2}{n} - \left(\dfrac{\sum x_i}{n}\right)^2}$ и $\sigma_y = \sqrt{\dfrac{\sum y_i^2}{n} - \left(\dfrac{\sum y_i}{n}\right)^2}$.

Подставляя найденное выражение для $\hat{\theta}_1$ в уравнение (17.27), получаем выборочное уравнение регрессии в более симметричной форме:

$$\hat{y} - \bar{y} = r\frac{\sigma_y}{\sigma_x}(x - \bar{x}) \quad \Leftrightarrow \quad \frac{\hat{y} - \bar{y}}{\sigma_y} = r\frac{x - \bar{x}}{\sigma_x}. \qquad (17.28)$$

Пример. Получены данные (табл. 17.10) о зависимости выходного напряжения Y (В) от входного напряжения X (В) технического устройства. Построить корреляционное поле (диаграмму рассеяния), найти выборочный коэффициент корреляции.

Таблица 17.10

x_i	2	8	6	3	5	2	12	10	5	14
y_i	4	5	3	3	4	3	6	5	3	6

Решение. Строим корреляционное поле, нанося на координатную плоскость XOY точки (x_i, y_i), $i = \overline{1, 10}$ (рис. 17.8).

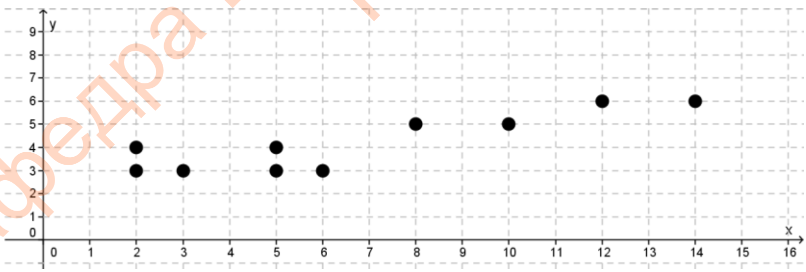


Рис. 17.8

Для расчета r удобно составить табл. 17.11.

Таблица 17.11

i	x_i	x_i^2	y_i	y_i^2	$x_i y_i$	$x_i + y_i$	$(x_i + y_i)^2$
1	2	4	4	16	8	6	36

Окончание таблицы 17.11

2	8	64	5	25	40	13	169
3	6	36	3	9	18	9	81
4	3	9	3	9	9	6	36
5	5	25	4	16	20	9	81
6	2	4	3	9	6	5	25
7	12	144	6	36	72	18	324
8	10	100	5	25	50	15	225
9	5	25	3	9	15	8	64
10	14	196	6	36	84	20	400
$\sum$	67	607	42	190	322	109	1441

По данным измерений имеем: $n = 10$, $\sum x_i = 67$, $\sum y_i = 42$, $\bar{x} = 6{,}7$, $\bar{y} = 4{,}2$, $\sum x_i^2 = 607$, $\sum y_i^2 = 190$, $\sum x_i y_i = 322$.

Для контроля вычислений находим: $\sum (x_i + y_i)^2 = 1441$. С другой стороны,

$$\sum (x_i + y_i)^2 = \sum x_i^2 + 2\sum x_i y_i + \sum y_i^2 =$$
$$= 607 + 2 \cdot 322 + 190 = 1441,$$

следовательно, вычисления верны.

По формуле (17.20) найдем выборочный коэффициент корреляции

$$r = \frac{\sum\limits_i x_i y_i - \frac{\sum x_i \cdot \sum y_i}{n}}{\sqrt{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}} \cdot \sqrt{\sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}}} =$$

$$= \frac{322 - \frac{67 \cdot 42}{10}}{\sqrt{607 - \frac{(67)^2}{10}} \cdot \sqrt{190 - \frac{(42)^2}{10}}} \approx 0{,}88.$$

Так как значение r близко к 1, то характеристика тесноты связи высокая.

Далее вычисляем:

$$\sigma_x = \sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2} = \sqrt{\frac{607}{10} - \left(\frac{67}{10}\right)^2} \approx 3{,}976,$$

$$\sigma_y = \sqrt{\frac{\sum y_i^2}{n} - \left(\frac{\sum y_i}{n}\right)^2} = \sqrt{\frac{190}{10} - \left(\frac{42}{10}\right)^2} \approx 1{,}166,$$

$$r\frac{\sigma_y}{\sigma_x} \approx 0{,}88 \cdot \frac{1{,}166}{3{,}976} \approx 0{,}258.$$

Т.е. выборочное уравнение регрессии (17.28) примет вид:

$$\hat{y} - 4{,}2 = 0{,}258(x - 6{,}7) \Leftrightarrow \hat{y} = 0{,}258x + 2{,}47.$$

График линии регрессии приведен на рис. 17.9.

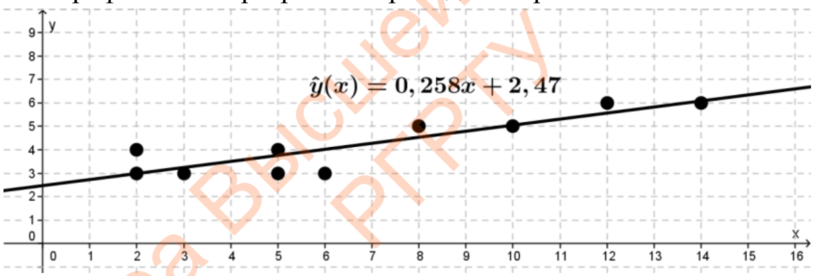


Рис. 17.9

Исследование взаимозависимых величин приводит к теории корреляции как к разделу теории вероятностей и корреляционному анализу как к разделу математической статистики. Исследование зависимости случайной величины от ряда неслучайных и случайных величин приводит к моделям регрессии и регрессионному анализу на базе выборочных данных. Теория вероятностей и математическая статистика представляют лишь инструмент для изучения статистической зависимости, но не ставят целью установление причинной связи. Представления и гипотезы о причинной связи должны быть привнесены из некоторой другой теории, которая позволяет содержательно объяснить изучаемое явление.

ПРИЛОЖЕНИЯ

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

x	Сотые доли x									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3725	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
x	Десятые доли x									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
2,	0540	0440	0355	0283	0224	0175	0136	0104	0079	0060
3,	0044	0033	0024	0017	0012	0009	0006	0004	0003	0002
4,	0001									

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

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0,0000	4000	0080	0120	0160	0199	0239	0279	0319	0359
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0753
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2257	2291	2324	2357	2389	2422	2454	2486	2517	2549
0,7	2580	2611	2642	2673	2704	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1,0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1,6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1,9	0,4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
x	Десятые доли x									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
2,	4773	4851	4861	4893	4918	4938	4953	4965	4974	4981
3,	4987	4990	4993	4995	4997	4998	4998	4999	4999	5000

Приложение 3. Значения функции $\dfrac{\lambda^m}{m!}e^{-\lambda}$

m \ a	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**
0	0,90484	0,81873	0,74082	0,67032	0,60653
1	0,09048	0,16375	0,22225	0,26813	0,30327
2	0,00452	0,01637	0,03334	0,05363	0,07582
3	0,00015	0,00109	0,00333	0,00715	0,01264
4		0,00005	0,00025	0,00072	0,00158
5			0,00002	0,00006	0,00016

m \ a	**0,6**	**0,7**	**0,8**	**0,9**
0	0,548812	0,496585	0,449329	0,40657
1	0,329287	0,34761	0,359463	0,365913
2	0,098786	0,121663	0,143785	0,164661
3	0,019757	0,028388	0,038343	0,049398
4	0,002964	0,004968	0,007669	0,011115
5	0,000356	0,000696	0,001227	0,002001
6	0,00036	0,00008	0,000164	0,0003

m \ a	**1,0**	**2,0**	**3,0**	**4,0**	**5,0**
0	0,367879	0,135335	0,049787	0,018316	0,006738
1	0,367879	0,270671	0,149361	0,073263	0,03369
2	0,18394	0,270671	0,224042	0,146525	0,084224
3	0,061313	0,180447	0,224042	0,195367	0,140374
4	0,015328	0,090224	0,168031	0,195367	0,175467
5	0,003066	0,036089	0,100819	0,156293	0,175467
6	0,000511	0,01203	0,050409	0,104196	0,146223
7	0,00007	0,003437	0,021604	0,05954	0,104445
8	0,00001	0,000859	0,008102	0,02977	0,065278
9		0,000191	0,002701	0,013231	0,036266
10		0,00004	0,00081	0,005292	0,018133
11		0,00001	0,000221	0,001925	0,008242

Приложение 4. Квантили $\chi^2_{\alpha,\tau}$ распределения χ^2_τ (τ – число степеней свободы)

τ	Уровень значимости α					
	0,01	**0,025**	**0,05**	**0,95**	**0,975**	**0,99**
1	6,63	5,02	3,84	0,0039	0,00098	0,0002
2	9,21	7,38	5,99	0,103	0,051	0,020
3	11,34	9,35	7,81	0,352	0,216	0,115
4	13,28	11,14	9,49	0,711	0,484	0,297
5	15,09	12,83	11,07	1,15	0,831	0,554
6	16,81	14,45	12,59	1,64	1,24	0,872
7	18,48	16,01	14,07	2,17	1,69	1,24
8	20,09	17,53	15,51	2,73	2,18	1,65
9	21,67	19,02	16,92	3,33	2,70	2,09
10	23,21	20,48	18,31	3,94	3,25	2,56
11	24,72	21,92	19,68	4,57	3,82	3,05
12	26,22	23,34	21,03	5,23	4,40	3,57
13	27,69	24,74	22,36	5,89	5,01	4,11
14	29,14	26,12	23,68	6,57	5,63	4,66
15	30,58	27,49	25,00	7,26	6,26	5,23
16	32,00	28,85	26,30	7,96	6,91	5,81
17	33,41	30,19	27,59	8,67	7,56	6,41
18	34,81	31,53	28,87	9,39	8,23	7,01
19	36,19	32,85	30,14	10,12	8,91	7,63
20	37,57	34,17	31,41	10,85	9,59	8,26
21	38,93	35,48	32,67	11,59	10,28	8,90
22	40,29	36,78	33,92	12,34	10,98	9,54
23	41,64	38,08	35,17	13,09	11,69	10,20
24	42,98	39,36	36,42	13,85	12,40	10,86
25	44,31	40,65	37,65	14,61	13,12	11,52
26	45,64	41,92	38,89	15,38	13,84	12,20
27	46,96	43,19	40,11	16,15	14,57	12,88
28	48,28	44,46	41,34	16,93	15,31	13,56
29	49,59	45,72	42,56	17,71	16,05	14,26
30	50,89	46,98	43,77	18,49	16,79	14,95

Приложение 5. Квантили t-распределения Стьюдента

τ	Уровень значимости α					
	0,1	**0,05**	**0,02**	**0,01**	**0,002**	**0,001**
1	6,31	12,71	31,82	63,66	318,31	636,62
2	2,92	4,30	6,96	9,92	22,33	31,60
3	2,35	3,18	4,54	5,84	10,21	12,92
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,02	2,57	3,36	4,03	5,89	6,87
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,41
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,02	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,97
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,72	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,50	3,79
23	1,71	2,07	2,50	2,81	3,48	3,77
24	1,71	2,06	2,49	2,80	3,47	3,75
25	1,71	2,06	2,49	2,79	3,45	3,73
26	1,71	2,06	2,48	2,78	3,43	3,71
27	1,70	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,47	2,76	3,41	3,67
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46

Приложение 6. Краткие теоретические сведения по теории вероятностей.

1. Случайные события

1.1. Классическое определение вероятности события A:

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

где m – число исходов, благоприятных A; n – число всех исходов.

1.2. Геометрическая вероятность события A:

$$P(A)=\frac{mes(A)}{mes(\Omega)},$$

где $mes(A)$ – мера множества A, $mes(\Omega)$ – мера ПЭС Ω.

1.3. Вероятность суммы событий:

а) для совместных $P(A+B)=P(A)+P(B)-P(A\cdot B)$;

б) для несовместных $P(A+B)=P(A)+P(B)$.

1.4. Условная вероятность:

$$P(A/B)=P(B)\cdot P(A/B)=P(A)\cdot P(B/A).$$

Вероятность произведения:

$$P(A\cdot B)=P(B)\cdot P(A/B)=P(A)\cdot P(B/A).$$

Условие независимости событий:

$$P(A\cdot B)=P(A)\cdot P(B).$$

1.5. Формула полной вероятности: $P(A)=\sum\limits_{i=1}^{n}P(H_i)\cdot P(A/H_i)$,

где $\bigcup\limits_{i=1}^{n}H_i=\Omega$, $H_i\cdot H_j=\varnothing$.

Формула Байеса: $P(H_i/A)=\dfrac{P(H_i)\cdot P(A/H_i)}{\sum\limits_{i=1}^{n}P(H_i)\cdot P(A/H_i)}$.

1.6. Формула Бернулли: $P_n(m)=C_n^m\cdot p^m\cdot q^{n-m}$, $p+q=1$.

2. Случайные величины

<table>
<tr><td>Дискретные случайные величины (ДСВ)</td><td>Непрерывные случайные величины (НСВ)</td></tr>
<tr><td colspan="2">2.1. Законы распределения
Функция распределения:
$$F(x)=P(X<x),$$
$F(-\infty)=0;\ F(+\infty)=1;\ F(x)\geq 0;$ неубывающая.</td></tr>
<tr><td>$$F(x)=\sum_i p_i,\quad x_i<x;$$
$F(x)$ – ступенчатая.
Ряд распределения
<table><tr><td>X</td><td>x_i</td><td>…</td><td>x_n</td></tr><tr><td>P</td><td>p_i</td><td>…</td><td>p_n</td></tr></table>
$$p_i>0.$$</td><td>$$F(x)=\int_{-\infty}^{x} f(x)dx;$$
$F(x)$ – непрерывная.
Плотность распределения
$$f(x)=F'(x);$$
$$f(x)\geq 0.$$</td></tr>
<tr><td colspan="2">Условие нормировки:</td></tr>
<tr><td>$$\sum_{i=1}^{n} p_i=1.$$</td><td>$$\int_{-\infty}^{+\infty} f(x)dx=1.$$</td></tr>
<tr><td colspan="2">2.2. Числовые характеристики
Математическое ожидание:</td></tr>
<tr><td>$$M(X)=\sum_{i=1}^{n} x_i p_i;$$</td><td>$$M(X)=\int_{-\infty}^{+\infty} x\cdot f(x)dx;$$</td></tr>
<tr><td colspan="2">Дисперсия:
$$D(X)=M\left([X-M(X)]^2\right)=M(X^2)-M^2(X);$$</td></tr>
<tr><td>$$D(X)=\sum_{i=1}^{n} x_i^2\cdot p_i-M^2(X)=$$
$$=\sum_{i=1}^{n}(x_i-M(X))^2\cdot p_i.$$</td><td>$$D(X)=\int_{-\infty}^{+\infty} x^2\cdot f(x)dx-M^2(X)=$$
$$=\int_{-\infty}^{+\infty}(x-M(X))^2\cdot f(x)dx.$$</td></tr>
<tr><td colspan="2">Среднее квадратическое отклонение: $\sigma(X)=\sqrt{D(X)}$</td></tr>
</table>

2.3. Свойства числовых характеристик	
Математическое ожидание: $x_{\min} \le M(X) \le x_{\max}$; $M(C) = C$, $C = const$; $M(C \cdot X) = C \cdot M(X)$; $M(aX + b) = aM(X) + b$; $M(X + Y) = M(X) + M(Y)$; $M(X \cdot Y) = M(X) \cdot M(Y) +$ $+ \operatorname{cov}(X, Y)$	Дисперсия: $D(X) \ge 0$; $D(C) = 0$, $C = const$; $D(C \cdot X) = C^2 \cdot D(X)$; $D(aX + b) = a^2 D(X)$; $D(X + Y) = D(X) + D(Y) +$ $+ 2\operatorname{cov}(X, Y)$; $D(X - Y) = D(X) + D(Y)$ $D(X) = \operatorname{cov}(X, X)$
Ковариация: $\operatorname{cov}(aX + b; cY + d) = ac \cdot \operatorname{cov}(X, Y)$; $\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X)$; $\operatorname{cov}(X, X) = D(X)$; $\lvert\operatorname{cov}(X, Y)\rvert \le \sqrt{D(X) \cdot D(Y)}$; $\operatorname{cov}(X, Y) = 0$ для независимых величин X и Y; $\operatorname{cov}(X, Y) \ne 0$ для зависимых величин X и Y	Коэффициент корреляции: $\lvert r_{XY}\rvert \le 1$; $r_{XY} = r_{YX}$; $r[aX + b; cY + d] = \dfrac{ac}{\lvert ac\rvert} \cdot r_{XY}$; $r[X; aX + b] = \dfrac{a}{\lvert a\rvert} = \pm 1$; $r_{XY} = 0$, если величины X и Y независимы; $\lvert r_{XY}\rvert = 1$, если между X и Y существует линейная функциональная зависимость

3. Случайные векторы

Дискретный случайный вектор	Непрерывный случайный вектор

3.1. Законы распределения

Функция распределения:

$$F(x,y)=P(X<x,\ Y<y);$$

$$F(-\infty,y)=F(x,-\infty)=F(-\infty,-\infty)=0;$$

$$F(+\infty,+\infty)=1;$$

$$F(x,+\infty)=F(x);\ F(+\infty,y)=F(y);$$

$F(x,y)$ – неубывающая по x и y.

Дискретный случайный вектор

$$F(x,y)=\sum_i\sum_j P_{ij},\ \ x_i<x, y_j<y.$$

Таблица распределения

$X \backslash Y$	y_1	…	y_j	…	y_m
x_1	p_{11}	…	p_{1j}	…	p_{1m}
…	…	…	…	…	…
x_i	p_{i1}	…	p_{ij}	…	p_{im}
…	…	…	…	…	…
x_n	p_{n1}	…	p_{nj}	…	p_{nm}

$$p_{ij}\geq 0;$$

$$\sum_{j=1}^{m}p_{ij}=P_i(X=x_i);$$

$$\sum_{i=1}^{n}p_{ij}=P_j(Y=y_j).$$

Непрерывный случайный вектор

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

Плотность распределения:

$$f(x,y)=\frac{\partial^2F(x,y)}{\partial x\,\partial y};$$

$$F(x,y)\geq 0;$$

$$\int_{-\infty}^{+\infty}f(x,y)dy=f_x(x);$$

$$\int_{-\infty}^{+\infty}f(x,y)dx=f_y(y).$$

Условие нормировки:

$$\sum_{i=1}^{n}\sum_{j=1}^{m}p_{ij}=1$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1$$

Вероятность попадания в область D

$$P((x,y)\in D)=\iint_D p(x,y)dxdy$$

3.2. Числовые характеристики

Математическое ожидание:

$$M(X)=\sum_i\left(\sum_j x_i p_{ij}\right);\qquad M(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\,p(x,y)dxdy;$$

$$M(Y)=\sum_j\left(\sum_i y_j p_{ij}\right).\qquad M(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\,p(x,y)dxdy.$$

Дисперсия:

$$D(X)=\sum_i\sum_j (x_i-M_X)^2 p_{ij};\qquad D(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2 p(x,y)dxdy-M_X^2;$$

$$D(Y)=\sum_i\sum_j (y_j-M_Y)^2 p_{ij}.\qquad D(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y^2 p(x,y)dxdy-M_Y^2.$$

Ковариация:

$$\operatorname{cov}(X,Y)=M[(X-M_X)\cdot(Y-M_Y)]=M(X\cdot Y)-M(X)\cdot M(Y)$$

Коэффициент корреляции: $r_{XY}=\dfrac{\operatorname{cov}(X,Y)}{\sigma_X\cdot\sigma_Y}$

3.3. Независимые случайные величины

Условие независимости: $F(x,y)=F(x)\cdot F(y)$;

$$p_{ij}=p_i\cdot p_j.\qquad f(x,y)=f(x)\cdot f(y).$$

Свойства числовых характеристик:

$$M(X\cdot Y)=M(X)\cdot M(Y);\; D(X+Y)=D(X)+D(Y);$$

$$\operatorname{cov}(X,Y)=0;\; r_{XY}=0.$$

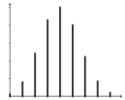
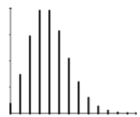
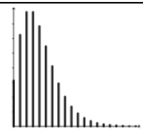
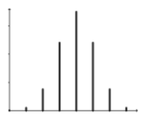
3.4. Зависимые случайные величины

Условные законы распределения:

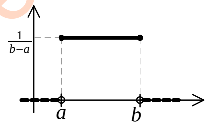
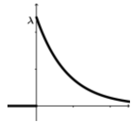
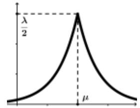
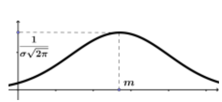
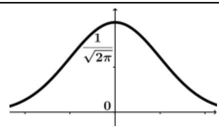
$$P(Y=y_j/X=x_i)=\frac{p_{ij}}{p_i};\qquad f(Y/X)=\frac{f(x,y)}{f_x(x)};$$

$$P(X=x_i/Y=y_j)=\frac{p_{ij}}{p_j}\qquad f(X/Y)=\frac{f(x,y)}{f_y(y)}$$

Приложение 7. Дискретные случайные величины

Закон распределения	Обозначение	Аналитическое выражение	Параметры	Эскиз закона распределения	$M(X)$	$D(X)$
Биномиальный (Бернулли)	$B(n,p)$	$p_m = C_n^m p^m q^{n-m}$	$n,\ p$		np	npq
Пуассона	$P(\lambda)$	$p_m = \dfrac{\lambda^m}{m!} e^{-\lambda};\ \lambda > 0$	λ		λ	λ
Геометрический (Фарри)	$Geom(p)$	$p_m = q^m p$	p		$\dfrac{1}{p}$	$\dfrac{q}{p^2}$
Гипергеометрический	$HG(N,M,n)$	$p_m = \dfrac{C_M^m C_{N-M}^{n-m}}{C_N^n}$	N, M, n		$n \cdot \dfrac{M}{N}$	$\dfrac{M(N-M)}{N^2} \times \dfrac{n(N-n)}{(N-1)}$

Приложение 8. Непрерывные случайные величины

Закон распределения	Обозна-чение	Аналитическое выражение плотности распределения $p(x)$	Парамет-ры	График плотности распределения	*M(X)*	*D(X)*
Равномерный (прямоугольный)	$R[a,b]$	$\begin{cases} 0, & x \notin [a,b]; \\ \frac{1}{b-a}, & x \in [a,b] \end{cases}$	a, b		$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
Показательный (экспоненциальный)	$E(\lambda)$	$\begin{cases} 0, & x < 0; \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases}$	λ		$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
Лапласа (двусторонний экспоненциальный)	$E(\lambda, \mu)$	$\dfrac{\lambda}{2} e^{-\lambda \lvert x-\mu \rvert}$	λ, μ		μ	$\dfrac{2}{\lambda^2}$
Нормальный	$N(m, \sigma)$	$\dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$	m, σ		m	σ^2
Нормальный стандартный	$N(0,1)$	$\dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	$m = 0,$ $\sigma = 1$		0	1

Библиографический список

1. Галушкина Ю.И. Конспект лекций по дискретной математике / Ю.И. Галушкина, А.Н. Марьямов. – М.: Айрис-пресс, 2007.
2. Соболева Т.С. Дискретная математика: учебник для студ. вузов /Т.С. Соболева, А.В. Чечкин; под ред. А.В. Чечкина. – М.: Издательский центр «Академия», 2006.
3. Курош А.Г. Курс высшей алгебры. – СПб.: Лань, 2003.
4. Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Высшая школа, 2004.
5. Письменный Д.Т. Конспект лекций по теории вероятностей, математической статистике и случайным процессам. –М.: Айрис.-пресс, 2007.
6. Вентцель Е.С. Теория вероятностей и ее инженерные приложения. – М.: Академия, 2003.
7. Орлов А.И. Вероятность и прикладная статистика: основные факты. – М.:КноРус, 2010.
8. Тарасов В.В., Елкина Н.В. Дискретная математика: учеб. пособие. – Рязань: РГРТУ, 2009.
9. Красин В.П. и др. Математическая статистика: метод. указания к пр.зан. – Рязань: РГРТУ, 2000.
10. Курашин В.Н. Элементы математической статистики: учеб. пособие. – Рязань: РФВУС, 2004.

ОГЛАВЛЕНИЕ

Б у х е н с к и й Кирилл Валентинович
Е л к и н а Наталья Викторовна
М а с л о в а   Наталия Николаевна

Краткий курс математики. Часть 4

Редактор Р.К. Мангутова
Корректор С.В. Макушина
Подписано в печать 25.06.14. Формат бумаги 60×84 1/16.
Бумага газетная. Печать трафаретная. Усл. печ. л. 8,75.
Тираж 60 экз. Заказ
Рязанский государственный радиотехнический университет.
390005, Рязань, ул. Гагарина, 59/1.
Редакционно-издательский центр РГРТУ.