Министерство науки и высшего образования Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ
ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ

И.В. Зайцева, Ю.Б. Ржонсницкая

МАТЕМАТИКА И СТАТИСТИКА

Санкт-Петербург
РГГМУ
2021

УДК [519.6:512.64](075.8)
ББК 22.17+22.143я73

Рецензент:

Тарасенко Елена Олеговна, к.ф.-м.н., доцент, доцент кафедры вычислительной математики и кибернетики ФГАОУ ВО «Северо-Кавказский федеральный университет»

Зайцева И.В., Ржонсницкая Ю.Б.
З-17 Математика и статистика: учебное пособие. – СПб: РГГМУ, 2021. – 184 с.

В учебном пособии изложен материал по курсу математики и статистики. В учебном пособии рассматриваются такие важные понятия матрицы, определителя и ранга матрицы, систем линейных уравнений, дифференциального исчисления, теории вероятностей и математической статистики. Авторы излагают те теоретические вопросы, знание которых является необходимым минимумом для усвоения материала, рассматриваемого в курсе «Математика и статистика». Наряду со сведениями теоретического характера в учебном пособии разбирается значительное число примеров и задач, цель которых – уяснение основных понятий. Учебное пособие предназначено для студентов высших учебных заведений. Также может быть полезно инженерам и научным работникам разных специальностей, изучающим или использующим методы линейной алгебры.

Содержание

ВВЕДЕНИЕ

В данном учебном пособии излагаются основные факты курса «Математика и статистика». Учебное пособие охватывает основные разделы курса линейной алгебры, математического анализа, теории вероятностей и математической статистики.

Пособие включает две главы, которые разбиваются на параграфы. В учебном пособии рассматриваются такие важные понятия матрицы, определителя и ранга матрицы, систем линейных уравнений, дифференциального исчисления, теории вероятностей и математической статистики. Авторы излагают те теоретические вопросы, знание которых является необходимым минимумом для усвоения материала, рассматриваемого в курсе «Математика и статистика». Наряду со сведениями теоретического характера в учебном пособии разбирается значительное число примеров и задач, цель которых – уяснение основных понятий.

Учебное пособие предназначено для студентов высших учебных заведений. Основным назначением пособия является помощь студентам в изучении и активном усвоении сложного материала. В результате изучения данного пособия студенты должны уметь выполнять действия с матрицами, знать свойства действий над матрицами, находить определители квадратных матриц и использовать свойства определителей, вычислять обратную матрицу, освоить методы решения систем линейных алгебраических уравнений, находить производную, вычислять числовые характеристики случайных величин.

Также учебное пособие может быть полезно инженерам и научным работникам разных специальностей, изучающим или использующим методы линейной алгебры.

Дополнительные теоретические сведения для более глубокого изучения того или иного параграфа можно получить из книг, приведенных в списке литературы.

ГЛАВА I. МАТЕМАТИКА

§1. Матрицы и операции над ними

1.1. Матрицы

Матрицей размерности $m \times n$ называется прямоугольная таблица чисел, содержащая m строк и n столбцов. Матрицы обычно обозначаются прописными буквами латинского алфавита A, B, C, …

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{или} \quad \underset{m\times n}{A} = \left(a_{ij}\right) \text{ - матрица}$$

размерности $m \times n$. Здесь a_{ij} - элемент матрицы A. i – номер строки ($i = 1, 2, 3, \dots, m$), j – номер столбца ($j = 1, 2, 3, \dots, n$). $m \times n$ – размерность матрицы.

Например, матрицу A размерности 3×2 можно записать следующим образом: $\underset{3\times 2}{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$ или $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$. Наряду с круглыми скобками используют и другие скобки: $[\ ]$ или $\|\ \|$. Например $\underset{2\times 3}{B} = \begin{Vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{Vmatrix}$.

Две матрицы А и В равны тогда и только тогда, когда они имеют одинаковую размерность и равны их соответствующие элементы.

$$\underset{m\times n}{A} = \underset{m\times n}{B} \Leftrightarrow a_{ij} = b_{ij},\ i = 1, 2, 3, \ldots, m,\ j = 1, 2, 3, \ldots, n.$$

Матрица А называется квадратной, если число строк равно числу столбцов, т.е. имеет размерность n × n. В этом случае говорят, что матрица А имеет порядок n (или n-го порядка). Например, $A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ - матрица второго порядка.

Элементы квадратной матрицы $a_{11}, a_{22}, a_{33}, \ldots, a_{nn}$ образуют главную диагональ матрицы. Например, в матрице $A = \begin{pmatrix} 2 & 1 & 8 \\ 3 & 6 & 9 \\ 5 & 7 & 11 \end{pmatrix}$ элементы 2, 6, 11 образуют главную диагональ.

Квадратная матрица, у которой все элементы, кроме элементов главной диагонали равны нулю, называется диагональной матрицей. Например, $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ - диагональная матрица третьего порядка.

Диагональная матрица, у которой все элементы главной диагонали равны единице, называется единичной матрицей. Единичная матрица обычно обозначается буквой Е или I. Например, $E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ - единичная матрица второго порядка; $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ - единичная матрица третьего порядка.

Матрица любой размерности, у которой все ее элементы равны нулю, называется нулевой матрицей или нуль матрицей. Нуль матрицу обозначают О. Например,

$$\underset{m\times n}{O} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \cdots & \dots & \cdots & \cdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$ - нуль матрица размерности m × n.

Матрица, содержащая один столбец или одну строку, называется матрица-столбец (вектор-столбец) или матрица-строка (вектор-строка). Например, $\underset{m\times 1}{A} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$ - матрица-столбец, $\underset{1\times n}{C} = \begin{pmatrix} c_1 & c_2 & \dots & c_n \end{pmatrix}$ - матрица-строка.

1.2. Операции над матрицами.

1. Умножение матрицы на число. Результатом умножения матрицы на число $\alpha \neq 0$ является матрица,

каждый элемент которой умножен на это число. $\alpha \cdot A = B$, где $b_{ij} = \alpha \cdot a_{ij}$. Например, $2 \cdot \begin{pmatrix} 1 & 4 \\ -5 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 8 \\ -10 & 6 \end{pmatrix}$.

Из определения следует, что $\alpha \cdot A = A \cdot \alpha$.

2. Сложение двух матриц. Складывать можно только матрицы одной размерности. Результатом сложения двух матриц будет матрица, каждый элемент которой есть сумма соответствующих элементов слагаемых. $\underset{m\times n}{A} + \underset{m\times n}{B} = \underset{m\times n}{C}$, где $c_{ij} = a_{ij} + b_{ij}$.

Свойства.

$A + B = B + A$ - переместительное свойство.

$(A + B) + C = A + (B + C)$ - сочетательное свойство.

Пример. $\begin{pmatrix} 1 & -3 & 0 \\ 4 & 5 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 2 & 1 \\ 3 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 \\ 7 & 5 & 0 \end{pmatrix}$.

3. Вычитание матриц. $\underset{m\times n}{A} - \underset{m\times n}{B} = \underset{m\times n}{C}$, где $c_{ij} = a_{ij} - b_{ij}$.

Пример. $A=\begin{pmatrix}1 & 7\\ 5 & 1\\ 6 & 3\end{pmatrix}$, $B=\begin{pmatrix}2 & 3\\ 7 & 2\\ 1 & 1\end{pmatrix}$. $A-B=\begin{pmatrix}1 & 7\\ 5 & 1\\ 6 & 3\end{pmatrix}-$

$$\begin{pmatrix}2 & 3\\ 7 & 2\\ 1 & 1\end{pmatrix}=\begin{pmatrix}1 & 7\\ 5 & 1\\ 6 & 3\end{pmatrix}+\begin{pmatrix}-2 & -3\\ -7 & -2\\ -1 & -1\end{pmatrix}=\begin{pmatrix}-1 & 4\\ -2 & -1\\ 5 & 2\end{pmatrix}.$$

4. Умножение матриц. Произведением матрицы $\underset{m\times p}{A}=\left(a_{ij}\right)$ на матрицу $\underset{p\times n}{B}=\left(b_{ij}\right)$ называется матрица $\underset{m\times n}{C}=\left(c_{ij}\right)$ ($C=A\cdot B$), каждый элемент которой c_{ij} равен сумме произведений элементов i-й строки матрицы A на соответствующие элементы j-го столбца матрицы B (скалярному произведению i-й строки матрицы A на j-й столбец матрицы B).

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+a_{i3}b_{3j}+\ldots+a_{ip}b_{pj}=\sum_{k=1}^{p}a_{ik}b_{kj},$$

$$i=1, 2, 3,\ldots, m, \quad j=1, 2, 3, \ldots, n.$$

Матрицы можно перемножать, если число столбцов первой матрицы равно числу строк второй матрицы. У произведения строк будет столько, сколько строк у первого сомножителя, а столбцов – сколько столбцов у второго сомножителя. Например,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{pmatrix}$$, где $c_{11} = a_{11}b_{11} + a_{12}b_{21}$, $c_{12} = a_{11}b_{12} + a_{12}b_{22}$, $c_{21} = a_{21}b_{11} + a_{22}b_{21}$, $c_{22} = a_{21}b_{12} + a_{22}b_{22}$, $c_{31} = a_{31}b_{11} + a_{32}b_{21}$, $c_{32} = a_{31}b_{12} + a_{32}b_{22}$.

Пример 1. Вычислить $A \cdot B$, если $A = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 4 & 5 \\ 7 & 3 & 8 \end{pmatrix}$, $B = \begin{pmatrix} 4 & 9 \\ 2 & 7 \\ 10 & 6 \end{pmatrix}$.

Решение. Размерность матрицы произведения: $\underset{3\times 3}{A} \cdot \underset{3\times 2}{B} = \underset{3\times 2}{C}$. Элементы матрицы С:

$c_{11} = 2 \cdot 4 + 1 \cdot 2 + 3 \cdot 10 = 8 + 2 + 30 = 40$,

$c_{12} = 2 \cdot 9 + 1 \cdot 7 + 3 \cdot 6 = 18 + 7 + 18 = 43$,

$c_{21} = 0 \cdot 4 + 4 \cdot 2 + 5 \cdot 10 = 0 + 8 + 50 = 58$,

$c_{22} = 0 \cdot 9 + 4 \cdot 7 + 5 \cdot 6 = 0 + 28 + 30 = 58$,

$c_{31} = 7 \cdot 4 + 3 \cdot 2 + 8 \cdot 10 = 28 + 6 + 80 = 114$,

$c_{32} = 7 \cdot 9 + 3 \cdot 7 + 8 \cdot 6 = 63 + 21 + 48 = 132$.

Таким образом, $C=\begin{pmatrix}40 & 43\\ 58 & 58\\ 114 & 138\end{pmatrix}$.

Пример 2. Вычислить $A\cdot B$, если $A=\begin{pmatrix}3 & 0\\ 1 & 1\\ 5 & 2\end{pmatrix}$, $B=\begin{pmatrix}4 & 7\\ 6 & 8\end{pmatrix}$.

Решение. $\begin{pmatrix}3 & 0\\ 1 & 1\\ 5 & 2\end{pmatrix}\cdot\begin{pmatrix}4 & 7\\ 6 & 8\end{pmatrix}=\begin{pmatrix}12 & 21\\ 10 & 15\\ 32 & 51\end{pmatrix}$.

Пример 3. Вычислить $A\cdot B$ и $B\cdot A$, если $A=\begin{pmatrix}3 & 4 & 1 & 5\end{pmatrix}$, $B=\begin{pmatrix}1\\ 2\\ 0\\ 1\end{pmatrix}$.

Решение. $\underset{1\times 4}{A}\cdot \underset{4\times 1}{B}=\underset{1\times 1}{C}$. $C=\begin{pmatrix}3 & 4 & 1 & 5\end{pmatrix}\cdot\begin{pmatrix}1\\ 2\\ 0\\ 1\end{pmatrix}=\begin{pmatrix}16\end{pmatrix}$.

$\underset{4\times 1}{B}\cdot\underset{1\times 4}{A}=\underset{4\times 4}{C}$. $C=\begin{pmatrix}1\\2\\0\\1\end{pmatrix}\cdot\begin{pmatrix}3 & 4 & 1 & 5\end{pmatrix}=\begin{pmatrix}3 & 4 & 1 & 5\\6 & 8 & 2 & 10\\0 & 0 & 0 & 0\\3 & 4 & 1 & 5\end{pmatrix}$.

Из рассмотренного примера видно, что $A\cdot B\neq B\cdot A$.

Пример 4. Вычислить $A\cdot B$ и $B\cdot A$, если $A=\begin{pmatrix}2 & 1\\3 & 4\end{pmatrix}$, $B=\begin{pmatrix}3\\1\end{pmatrix}$.

Решение. $\underset{2\times 2}{A}\cdot\underset{2\times 1}{B}=\underset{2\times 1}{C}$. $C=\begin{pmatrix}2 & 1\\3 & 4\end{pmatrix}\cdot\begin{pmatrix}3\\1\end{pmatrix}=\begin{pmatrix}7\\13\end{pmatrix}$.
Произведение $B\cdot A$ не определено, так как число строк первой матрицы не равно числу столбцов второй.

Приведем некоторые свойства приведенных операций. Некоторые из этих свойств совпадают со свойствами операций над числами.

1) $\alpha(A+B)=\alpha A+\alpha B$. 2) $A(B+C)=AB+AC$. 3) $(A+B)C=AC+BC$. 4) $\alpha(AB)=(\alpha A)B=A(\alpha B)$. 5) $A(BC)=(AB)C$. 6) $O\cdot A=A\cdot O=O$.

7) $IA=AI=A$.

Как уже отмечали, что, вообще говоря, $AB \neq BA$. В том случае, когда равенство есть $AB = BA$, например, 6) и 7) свойства, матрицы называются перестановочными. Например, $A = \begin{pmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & -1 \end{pmatrix}$, $B = \begin{pmatrix} -8 & 29 & -11 \\ -5 & 18 & -7 \\ 1 & -3 & 1 \end{pmatrix}$.

$A \cdot B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $B \cdot A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Матрицы A и B - перестановочные.

5. Возведение в степень. Целой положительной степенью A^m квадратной матрицы A называется произведение m матриц, равных A, т.е. $A^m = \underbrace{A \cdot A \cdot \ldots \cdot A}_{m \text{ раз}}$.

По определению $A^0 = I,\ A^1 = A$. Следовательно, выполняются свойства: $A^m \cdot A^n = A^{m+n}$, $\left(A^m\right)^n = A^{mn}$.

6. Транспонирование матрицы. Если в прямоугольной матрице $\underset{m \times n}{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$ строки заменить столбцами, а столбцы строками, сохраняя порядок следования элементов, то получим матрицу, которая

называется транспонированной. Ее обозначают A^T.

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}.$$

Транспонированная матрица будет иметь размерность $n \times m$. Переход от матрицы A к матрице A^T называется транспонированием матрицы A.

Пример 1. $A = \begin{pmatrix} 1 & 3 & 5 \\ 0 & 6 & 3 \\ -1 & 5 & 4 \end{pmatrix}$. Тогда $A^T = \begin{pmatrix} 1 & 0 & -1 \\ 3 & 6 & 5 \\ 5 & 3 & 4 \end{pmatrix}$.

Пример 2. $C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$. $C^T = \begin{pmatrix} c_1 & c_2 & \dots & c_m \end{pmatrix}$.

Свойства транспонирования: 1) $\left(A^T\right)^T = A$. 2) $(\alpha A)^T = \alpha \cdot A^T$. 3) $(A+B)^T = A^T + B^T$. 4) $(AB)^T = B^T \cdot A^T$.

§2. Определитель квадратной матрицы.

Квадратной матрице n-го порядка можно поставить в соответствие число, которое называется определителем (детерминантом). Это число обозначается $|A|$, или Δ, или $\det A$.

1.Определитель матрицы первого порядка. Пусть матрица A – матрица первого порядка, т.е. $A=\left(a_{11}\right)$. Тогда $\det A=|A|=a_{11}$.

2. Определитель матрицы второго порядка. Определителем матрицы второго порядка $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & {}_{22} \end{pmatrix}$ называется число, которое вычисляется по формуле: $|A|=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & {}_{22} \end{vmatrix}=a_{11}a_{22}-a_{21}a_{12}$. Например, Определитель матрицы $A=\begin{pmatrix} 1 & 6 \\ 5 & 7 \end{pmatrix}$. $|A|=\begin{vmatrix} 1 & 6 \\ 5 & 7 \end{vmatrix}=1\cdot 7-5\cdot 6=7-30=-23$.

3. Определитель матрицы третьего порядка $A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Определителем этой матрицы называется число, которое вычисляется по формуле:

$$|A|=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{21}a_{32}a_{13}-a_{31}a_{22}a_{13}-a_{21}a_{12}a_{33}-a_{32}a_{23}a_{11}$$

. Эту формулу легко запомнить, используя схему, которая называется правилом треугольников (или Сарроса):

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

произведения элементов, лежащих на главной диагонали и произведения элементов, лежащих в вершинах треугольников, одна сторона которых параллельна главной диагонали, берутся со знаком плюс. Таким образом, имеем три слагаемых: $a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13}$.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

произведения элементов, лежащих на дополнительной диагонали и произведения элементов, лежащих в вершинах треугольников, одна сторона которых параллельна дополнительной диагонали, берутся со знаком минус. Три слагаемых со знаком минус: $-a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{32}a_{23}a_{11}$.

Пример. Вычислить определитель матрицы

$$A = \begin{pmatrix} -2 & 0 & -2 \\ 4 & 3 & 2 \\ 1 & 3 & -5 \end{pmatrix}.$$

Решение.

$$\Delta = \begin{vmatrix} -2 & 0 & -2 \\ 4 & 3 & 2 \\ 1 & 3 & -5 \end{vmatrix} =$$

$$= -2 \cdot 3 \cdot (-5) + 0 \cdot 2 \cdot 1 + 4 \cdot 3 \cdot (-2) - 1 \cdot 3 \cdot (-2) - 4 \cdot 0 \cdot (-5) - 3 \cdot 2 \cdot (-2) =$$

= 30 + 0 - 24 + 6 - 0 + 12 = 24.

Определитель третьего порядка можно вычислить, используя другое правило. Допишем под определителем его первые две строки. Тогда произведения элементов, лежащих на диагоналях, идущих сверху вниз и слева направо берутся со знаком плюс. Произведения элементов, лежащих на диагоналях, идущих снизу вверх и слева направо берутся со знаком минус.

$$\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}.$$

Пример. Вычислить определитель матрицы $A = \begin{pmatrix} 1 & -3 & -1 \\ 2 & -1 & 4 \\ 7 & 7 & 1 \end{pmatrix}$.

Решение. $\det A = \begin{vmatrix} 1 & -3 & -1 \\ 2 & -1 & 4 \\ 7 & 7 & 1 \end{vmatrix} = \begin{array}{ccc} 1 & -3 & -1 \\ 2 & -1 & 4 \\ 7 & 7 & 1 \\ 1 & -3 & -1 \\ 2 & -1 & 4 \end{array} =$

$= 1\cdot(-1)\cdot 1 + 2\cdot 7\cdot(-1) + 7\cdot(-3)\cdot 4 - 7\cdot(-1)\cdot(-1) - 1\cdot 7\cdot 4 - 2\cdot(-3)\cdot 1 =$

-1 – 14 – 84 – 7 – 28 + 6 = -128.

Вычисление определителей более высокого порядка основано на вычислении определителей низших порядков. Методы вычисления рассмотрим ниже, после рассмотрения свойств определителей, доказательство которых в общем случае мы опустим. Для определителей второго и третьего порядков справедливость свойств легко проверить, пользуясь правилами вычисления определителей.

2.1. Свойства определителей

1. Определитель не изменится, если его строки заменить столбцами. $\left(|A| = |A^T|\right)$.

2. При перестановке двух строк или двух столбцов знак определителя изменится на противоположный.

3. Общий множитель какой-либо строки или какого-либо столбца можно вынести за знак определителя.

4. Если все элементы какого-либо столбца или, какой либо строки равны нулю, то определитель равен нулю.

5. Если определитель имеет два одинаковых столбца или две одинаковых строки, то он равен нулю.

6. Если элементы двух строк или двух столбцов пропорциональны, то определитель равен нулю.

7. Определитель не изменится, если к элементам некоторой строки (столбца) прибавить соответствующие элементы другой строки (столбца), умноженные на некоторое число $\alpha \neq 0$.

2.2. Миноры и алгебраические дополнения квадратной матрицы.

Рассмотрим матрицу n-го порядка

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Минором μ_{ij} элемента a_{ij} матрицы A называется определитель матрицы n – 1-го порядка, полученный из матрицы A вычеркиванием i-й строки и j-го столбца. Всего у матрицы n-го порядка будет n^2 миноров. Например, у матрицы третьего порядка $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, будет 9 миноров: $\mu_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{32}a_{23}$,

$$\mu_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = a_{21}a_{33} - a_{31}a_{23},$$

$$\mu_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = a_{21}a_{32} - a_{31}a_{22},$$

$$\mu_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = a_{12}a_{33} - a_{31}a_{13},$$

$$\mu_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = a_{11}a_{33} - a_{31}a_{13},$$

$$\mu_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} = a_{11}a_{32} - a_{31}a_{12},$$

$$\mu_{31} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = a_{12}a_{23} - a_{22}a_{13},$$

$$\mu_{32} = \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} = a_{11}a_{13} - a_{21}a_{13},$$

$$\mu_{33} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Пример. Вычислить минор μ_{23} матрицы $A = \begin{pmatrix} 1 & -3 & -1 \\ 2 & -1 & 4 \\ 7 & 7 & 1 \end{pmatrix}$.

Решение. $\mu_{23} = \begin{vmatrix} 1 & -3 \\ 7 & 7 \end{vmatrix} = 1 \cdot 7 - 7 \cdot (-3) = 7 + 21 = 28$.

Алгебраическим дополнением элемента a_{ij} матрицы A называется его минор, взятый со знаком плюс, если сумма $i + j$ - четное число, и со знаком минус, если эта сумма нечетное число. Алгебраическое дополнение элемента a_{ij} обозначается A_{ij}. Таким образом, $A_{ij} = (-1)^{i+j}\mu_{ij}$.

Например, в матрице $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$$A_{22} = (-1)^{2+2}\mu_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = a_{11}a_{33} - a_{31}a_{13},$$

$$A_{12} = (-1)^{1+2}\mu_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = -(a_{21}a_{33} - a_{31}a_{23}).$$

Свойство 8. Определитель n-го порядка равен сумме произведений элементов любой строки (любого столбца) на их алгебраические дополнения, т.е.

$$\det A = \sum_{i=1}^{n} a_{ij} A_{ij} = a_{1j} A_{1j} + a_{2j} A_{2j} + \ldots a_{nj} A_{nj}, \qquad j = 1, 2, \ldots, n$$

(разложение определителя по элементам j-го столбца) или

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij} = a_{i1} A_{i1} + a_{i2} A_{i2} + \ldots a_{in} A_{in}, \qquad i = 1, 2, \ldots, n$$

(разложение определителя по элементам i-й строки). Таким образом, вычисление определителя n-го порядка сводится к вычислению n определителей (n-1)-го порядка

Проиллюстрируем это свойство на примере определителя третьего порядка.

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} A_{11} + a_{12} A_{12} + a_{13} A_{13}$$

$$= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} =$$

$$= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) =$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} - a_{13}a_{31}a_{22}.$$

Для разложения определителя обычно выбирают ту строку или тот столбец, где больше нулевых элементов, так как соответствующие им слагаемые в разложении будут равны нулю.

Пример 1. Вычислить определитель $\begin{vmatrix} 3 & 5 & 7 & 8 \\ 0 & 22 & 21 & 25 \\ 0 & 5 & 3 & 2 \\ 1 & -1 & 7 & 4 \end{vmatrix}$.

Решение. Разложим определитель по элементам первого столбца. Имеем:

$$\begin{vmatrix} 3 & 5 & 7 & 8 \\ 0 & 22 & 21 & 25 \\ 0 & 5 & 3 & 2 \\ 1 & -1 & 7 & 4 \end{vmatrix} = 3 \cdot A_{11} + 0 \cdot A_{21} + 0 \cdot A_{31} + 1 \cdot A_{41} =$$

$$3\begin{vmatrix} 22 & 21 & 25 \\ 5 & 3 & 2 \\ -1 & 7 & 4 \end{vmatrix} - \begin{vmatrix} 5 & 7 & 8 \\ 22 & 21 & 25 \\ 5 & 3 & 2 \end{vmatrix} = 122.$$

Пример 2. Вычислить определитель $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$.

Решение. В данном определителя в первой строке на первом месте стоит 1. С помощью первой строки, применяя свойство 7, получим нули в первом столбце. Операции над строками запишем справа от определителя.

$$\begin{vmatrix}1&2&3&4\\2&3&4&1\\3&4&1&2\\4&1&2&3\end{vmatrix}\begin{matrix}\\l_2-2l_1\\l_3-3l_1\\l_4-4l_1\end{matrix}=\begin{vmatrix}1&2&3&4\\0&-1&-2&-7\\0&-2&-8&-10\\0&-7&-10&-13\end{vmatrix}$$. Далее разложим полученный определитель по элементам первого столбца.

$$\begin{vmatrix}1&2&3&4\\0&-1&-2&-7\\0&-2&-8&-10\\0&-7&-10&-13\end{vmatrix}=1\cdot\begin{vmatrix}-1&-2&-7\\-2&-8&-10\\-7&-10&-13\end{vmatrix}=160.$$

Пример 3. Вычислить определитель матрицы

$$A=\begin{pmatrix}1&1&1&3\\2&3&2&7\\3&1&1&5\\5&-1&-1&3\end{pmatrix}.$$

Решение.

$$|A|=\begin{vmatrix}1&1&1&3\\2&3&2&7\\3&1&1&5\\5&-1&-1&3\end{vmatrix}\begin{matrix}\\l_2-2l\\l_3-3l_1\\l_4-5l_1\end{matrix}=\begin{vmatrix}1&1&1&3\\0&1&0&1\\0&-2&-2&-4\\0&-6&-6&-12\end{vmatrix}.$$ В полученном определителе элементы третьей и четвертой строк пропорциональны, следовательно, определитель равен нулю (свойство 4). $|A|=0$.

Свойство 9. Сумма произведений элементов какой-либо строки (столбца) на алгебраические дополнения соответствующих элементов другой строки (столбца) равна нулю. $\sum_{j=1}^{n} a_{ij} A_{kj} = a_{i1} A_{k1} + a_{i2} A_{k2} + \ldots a_{in} A_{kn} = 0$, $i \neq k$ или $\sum_{i=1}^{n} a_{ij} A_{ik} = a_{1j} A_{1k} + a_{2j} A_{2k} + \ldots a_{nj} A_{nk} = 0$, $j \neq k$.

Для определителя третьего порядка $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ имеем $a_{11}A_{12} + a_{21}A_{22} + a_{31}A_{32} = 0$, $a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23} = 0$, $\ldots$ (Проверить самостоятельно).

§3. Ранг матрицы

Рассмотрим матрицу $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$.
Выделим в матрице *A* какие-либо *k* строк и *k* столбцов $(k \le \min(m,n))$. Определитель $|A_k|$, составленный из элементов матрицы *A*, стоящих на пересечении выделенных строк и столбцов называется минором *k*-го порядка данной матрицы *A*.

Определение. Рангом матрицы A называется наивысший порядок отличных от нуля миноров. Обозначается $r(A)$ или $rang(A)$.

Пусть A прямоугольная матрица размерности (m x n).Из определения следует:

1) $r(A)=0 \Leftrightarrow A=O$; 2) $r(A)\le \min(m,n)$.

Пример. Найти ранг матрицы $A=\begin{pmatrix} 1 & 3 & 5 & -1 \\ 2 & -1 & -3 & 4 \\ 5 & 1 & -1 & 7 \\ 7 & 7 & 9 & 1 \end{pmatrix}$.

Решение. $1\le r(A)\le 4$, так как у матрицы A есть элементы отличные от нуля и, порядок матрицы равен 4. Выберем минор второго порядка, например, расположенный в верхнем левом углу матрицы A. $\begin{vmatrix} 1 & 3 \\ 2 & -1 \end{vmatrix}=-7\ne 0$. Следовательно, $r(A)\ge 2$. Будем теперь последовательно вычислять миноры матрицы A третьего порядка, окаймляющие отличный от нуля минор второго порядка. $\begin{vmatrix} 1 & 3 & 5 \\ 2 & -1 & -3 \\ 5 & 1 & -1 \end{vmatrix}=1-45+10-1+45-10=0$;

$\begin{vmatrix} 1 & 3 & -1 \\ 2 & -1 & 4 \\ 5 & 1 & 7 \end{vmatrix}=0$; $\begin{vmatrix} 1 & 3 & -1 \\ 2 & -1 & 4 \\ 7 & 7 & 1 \end{vmatrix}=28\ne 0$. Следовательно, $r(A)\ge 3$. Вычислим единственный минор четвертого порядка $\begin{vmatrix} 1 & 3 & 5 & -1 \\ 2 & -1 & -3 & 4 \\ 5 & 1 & -1 & 7 \\ 7 & 7 & 9 & 1 \end{vmatrix}\begin{matrix} \\ l_2-2l_1 \\ l_3-5l_1 \\ l_4-7l_1 \end{matrix}=\begin{vmatrix} 1 & 3 & 5 & -1 \\ 0 & -7 & -13 & 6 \\ 0 & -14 & -26 & 12 \\ 0 & -14 & -26 & 8 \end{vmatrix}=0$,

так как вторая и третья строки пропорциональны.

Ответ: $r(A)=3$.

Такой метод нахождения ранга матрицы называется методом окаймляющих миноров. Он достаточно трудоемкий. Для облегчения этой задачи используются преобразования матрицы, сохраняющие ее ранг.

Элементарными преобразованиями матрицы, не изменяющими ее ранг, являются: 1) Отбрасывание нулевой строки (столбца). 2) Умножение всех элементов некоторой строки (столбца) на число, не равное нулю. 3) Перестановка двух строк (столбцов). 4) Прибавление к каждому элементу некоторой строки (столбца) соответствующих элементов другой строки (столбца), умноженных на любое число. 5) Транспонирование матрицы.

С помощью элементарных преобразований можно привести матрицу к треугольному или трапециидальному виду. Тогда ранг матрицы будет равен числу ненулевых строк преобразованной матрицы.

Пример. Найти ранг матрицы $A=\begin{pmatrix} 1 & 3 & 5 & -1 \\ 2 & -1 & -3 & 4 \\ 5 & 1 & -1 & 7 \\ 7 & 7 & 9 & 1 \end{pmatrix}$ с помощью элементарных преобразований.

Решение.

$$\begin{pmatrix} 1 & 3 & 5 & -1 \\ 2 & -1 & -3 & 4 \\ 5 & 1 & -1 & 7 \\ 7 & 7 & 9 & 1 \end{pmatrix}\begin{array}{l} \\ l_2-2l_1 \\ l_3-5l_1 \\ l_4-7l \end{array} \rightarrow \begin{pmatrix} 1 & 3 & 5 & -1 \\ 0 & -7 & -13 & 6 \\ 0 & -14 & -26 & 12 \\ 0 & -14 & -26 & 8 \end{pmatrix}\begin{array}{l} \\ \\ l_3-2l_2 \\ l_4-2l_2 \end{array} \rightarrow$$

$$\to \begin{pmatrix} 1 & 3 & 5 & -1 \\ 0 & -7 & -13 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 5 & -1 \\ 0 & -7 & -13 & 6 \\ 0 & 0 & 0 & -4 \end{pmatrix}.$$ Имеем три ненулевых строки. Ранг матрицы A равен 3.

Ответ: $r(A)=3$.

§4. Обратная матрица

Пусть дана квадратная матрица n-го порядка

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Определение. Квадратная матрица называется невырожденной, если ее определитель $|A|$ не равен нулю.

Если $|A|=0$, то матрица называется вырожденной.

Определение. Матрица $\tilde{A}$ называется присоединенной (союзной) матрицей к матрице А, если она составлена из алгебраических дополнений матрицы A^T.

Таким образом, $\widetilde{A} = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$, где, напомним, $A_{ij} = (-1)^{i+j}\mu_{ij}$, μ_{ij} -минор элемента a_{ij} матрицы А, т.е. это определитель (n – 1)-го порядка, полученный из матрицы А вычеркиванием i-й строки и j-го столбца.

Пример. Найти матрицу $\widetilde{A}$, присоединенную к матрице $A = \begin{pmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & -1 \end{pmatrix}$.

Решение. Ищем миноры и алгебраические дополнения элементов матрицы А. $\mu_{11} = \begin{vmatrix} -3 & 1 \\ -5 & -1 \end{vmatrix} = 3 + 5 = 8$, $A_{11} = (-1)^{1+1} \cdot 8 = 8$, $\mu_{12} = \begin{vmatrix} 2 & 1 \\ 3 & -1 \end{vmatrix} = -2 - 3 = -5$, $A_{12} = (-1)^{1+2} \cdot (-5) = 5$, $\mu_{13} = \begin{vmatrix} 2 & -3 \\ 3 & -5 \end{vmatrix} = -1$, $A_{13} = -1$, $\mu_{21} = \begin{vmatrix} -4 & 5 \\ -5 & -1 \end{vmatrix} = 29$, $A_{21} = -29$, $\mu_{22} = \begin{vmatrix} 3 & 5 \\ 3 & -1 \end{vmatrix} = -18$, $A_{22} = -18$, $\mu_{23} = \begin{vmatrix} 3 & -4 \\ 3 & -5 \end{vmatrix} = -3$, $A_{23} = 3$, $\mu_{31} = \begin{vmatrix} -4 & 5 \\ -3 & 1 \end{vmatrix} = 11$,

$A_{31}=11$, $\mu_{32}=\begin{vmatrix}3 & 5\\ 2 & 1\end{vmatrix}=-7$, $A_{32}=7$, $\mu_{33}=\begin{vmatrix}3 & -4\\ 2 & -3\end{vmatrix}=-1$, $A_{33}=-1$. Записываем присоединенную матрицу

$$\tilde{A}=\begin{pmatrix}A_{11} & A_{21} & A_{31}\\ A_{12} & A_{22} & A_{32}\\ A_{13} & A_{23} & A_{33}\end{pmatrix}=\begin{pmatrix}8 & -29 & 11\\ 5 & -18 & 7\\ -1 & 3 & -1\end{pmatrix}.$$

Определение. Матрица A^{-1} называется обратной матрицей к матрице А, если $A\cdot A^{-1}=A^{-1}\cdot A=I$.

Теорема. Любая невырожденная матрица имеет обратную. Обратная матрица вычисляется по формуле

$$A^{-1}=\frac{1}{|A|}\cdot\tilde{A}.$$

Доказательство. Найдем произведения $A\cdot\tilde{A}$ и $\tilde{A}\cdot A$. Имеем:

$$A\cdot\tilde{A}=\begin{pmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ a_{n1} & a_{n2} & \dots & a_{nn}\end{pmatrix}\cdot\begin{pmatrix}A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \dots & \dots & \dots & \dots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{pmatrix}=$$

$$=\begin{pmatrix}a_{11}A_{11}+a_{12}A_{12}+\ldots+a_{1n}A_{1n} & a_{11}A_{21}+a_{12}A_{22}+\ldots+a_{1n}A_{2n} & \cdots & a_{11}A_{n1}+a_{12}A_{n2}+\ldots a_{1n}A_{nn}\\ a_{21}A_{12}+a_{22}A_{12}+\ldots+a_{2n}A_{1n} & a_{21}A_{21}+a_{22}A_{22}+\ldots+a_{2n}A_{2n} & \cdots & a_{21}A_{n1}+a_{22}A_{n2}+\ldots+a_{2n}A_{nn}\\ \dots & \dots & \dots & \dots\\ a_{n1}A_{11}+a_{n2}A_{12}+\ldots+a_{nn}A_{1n} & a_{n1}A_{21}+a_{n2}A_{22}+\ldots+a_{nn}A_{2n} & \cdots & a_{n1}A_{n1}+a_{n2}A_{n2}+\ldots+a_{nn}A_{nn}\end{pmatrix}$$

$=$

$$=\begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & |A| \end{pmatrix}$$. Аналогично проверяется, что $\widetilde{A} \cdot A$

$$=\begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & |A| \end{pmatrix}$$. Следовательно, имеем,

$$A \cdot A^{-1} = A^{-1} \cdot A = A \cdot \frac{\widetilde{A}}{|A|} = \frac{\widetilde{A}}{|A|} \cdot A = \frac{1}{|A|} \begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & |A| \end{pmatrix} =$$

$$=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I$$. Что и требовалось доказать.

4.1. Свойства обратной матрицы

1) $\left|A^{-1}\right| = \dfrac{1}{|A|}$. 2) $\left(A^{-1}\right)^{-1} = A$. 3) $\left(A^{m}\right)^{-1} = \left(A^{-1}\right)^{m} = A^{-m}$. 4) $(AB)^{-1} = B^{-1}A^{-1}$. 5) $\left(A^{-1}\right)^{T} = \left(A^{T}\right)^{-1}$.

Пример 1. Найти A^{-1}, если $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Решение. 1) Ищем определитель матрицы А. $|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2 \neq 0$. 2) Ищем алгебраические дополнения элементов матрицы А. $\mu_{11} = 4$, $A_{11} = 4$, $\mu_{12} = 3$, $A_{12} = -3$, $\mu_{21} = 2$, $A_{21} = -2$ $\mu_{22} = 1$, $A_{22} = 1$. 3) Составляем присоединенную матрицу. $\tilde{A} = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$. 4) Выписываем обратную матрицу. $A^{-1} = -\frac{1}{2}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$.

Проверка. $A \cdot A^{-1} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Ответ. $A^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$.

Пример 2. Найти A^{-1}, если $A = \begin{pmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & -1 \end{pmatrix}$.

Решение. 1)

$$|A| = \begin{vmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & -1 \end{vmatrix} = 9 - 50 - 12 - (-45 + 8 - 15) = -53 + 52 = -1.$$

2) $\mu_{11} = \begin{vmatrix} -3 & 1 \\ -5 & -1 \end{vmatrix} = 3 + 5 = 8$, $A_{11} = (-1)^{1+1} \cdot 8 = 8$,

$\mu_{12} = \begin{vmatrix} 2 & 1 \\ 3 & -1 \end{vmatrix} = -2 - 3 = -5$, $A_{12} = (-1)^{1+2} \cdot (-5) = 5$,

$\mu_{13} = \begin{vmatrix} 2 & -3 \\ 3 & -5 \end{vmatrix} = -1$, $A_{13} = -1$, $\mu_{21} = \begin{vmatrix} -4 & 5 \\ -5 & -1 \end{vmatrix} = 29$, $A_{21} = -29$,

$\mu_{22} = \begin{vmatrix} 3 & 5 \\ 3 & -1 \end{vmatrix} = -18$, $A_{22} = -18$, $\mu_{23} = \begin{vmatrix} 3 & -4 \\ 3 & -5 \end{vmatrix} = -3$, $A_{23} = 3$,

$\mu_{31} = \begin{vmatrix} -4 & 5 \\ -3 & 1 \end{vmatrix} = 11$, $A_{31} = 11$, $\mu_{32} = \begin{vmatrix} 3 & 5 \\ 2 & 1 \end{vmatrix} = -7$, $A_{32} = 7$,

$\mu_{33} = \begin{vmatrix} 3 & -4 \\ 2 & -3 \end{vmatrix} = -1$, $A_{33} = -1$.

3) $$\tilde{A} = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = \begin{pmatrix} 8 & -29 & 11 \\ 5 & -18 & 7 \\ -1 & 3 & -1 \end{pmatrix}.$$

4) $$A^{-1} = \frac{1}{-1} \begin{pmatrix} 8 & -29 & 11 \\ 5 & -18 & 7 \\ -1 & 3 & -1 \end{pmatrix} = \begin{pmatrix} -8 & 29 & -11 \\ -5 & 18 & -7 \\ 1 & -3 & 1 \end{pmatrix}.$$

Проверка.

$$A \cdot A^{-1} = \begin{pmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & -1 \end{pmatrix} \cdot \begin{pmatrix} -8 & 29 & -11 \\ -5 & 18 & -7 \\ 1 & -3 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ответ.

$$A^{-1} = \begin{pmatrix} -8 & 29 & -11 \\ -5 & 18 & -7 \\ 1 & -3 & 1 \end{pmatrix}.$$

4.2. Нахождение обратной матрицы с помощью элементарных преобразований (метод Гаусса)

Элементарные преобразования матрицы – это умножение данной матрицы на некоторую матрицу. Продемонстрируем это на примере матрицы третьего порядка.

1) Перестановка двух строк (2-й и 3-й):
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}.$$

2) Умножение элементов некоторой строки на число α (второй строки):
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \alpha \cdot a_{21} & \alpha \cdot a_{22} & \alpha \cdot a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$
.

3) Сложение двух строк (вторую строку заменим суммой второй и третьей строк):
$$\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}+a_{31}&a_{22}+a_{32}&a_{23}+a_{33}\\a_{31}&a_{32}&a_{33}\end{pmatrix}.$$

На использовании этих преобразований основан метод Гаусса вычисления обратной матрицы.

Присоединяем к матрице A справа единичную матрицу I того же порядка. С помощью элементарных преобразований над строками расширенной матрицы, преобразуем матрицу A в единичную матрицу. Тогда справа единичная матрица преобразуется в A^{-1}. Т.е. $A|I \xrightarrow{\text{элементарные преобразования}} I\,|\,A^{-1}$. Фактически мы умножили слева матрицы A и I на матрицу A^{-1}, т.е. $A^{-1}\cdot A|\,A^{-1}\cdot I = I\,|\,A^{-1}$.

Пример. Найти методом Гаусса A^{-1}, если
$$A=\begin{pmatrix}2&2&3\\1&-1&0\\-1&2&1\end{pmatrix}.$$

Решение. Проверяем, что матрица A невырожденная. $|A| = -2+0+6-3-2-0=-1$. Присоединяем к матрице A справа единичную матрицу и выполняем элементарные преобразования.
$$\left(\begin{array}{ccc|ccc}2&2&3&1&0&0\\1&-1&0&0&1&0\\-1&2&1&0&0&1\end{array}\right)\xrightarrow{\text{переставляем первую и вторую строки}}$$

$$\to\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 2 & 2 & 3 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 & 0 & 1 \end{array}\right)\begin{array}{l} \\ l_2-2l_1 \\ l_3+l_1 \end{array}\to$$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 4 & 3 & 1 & -2 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{array}\right)\xrightarrow{\text{переставляем вторую и третью строки}}$$

$$\to\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 4 & 3 & 1 & -2 & 0 \end{array}\right)\begin{array}{l} l_1+l_2 \\ \\ l_3-4l_2 \end{array}\to$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 & -6 & -4 \end{array}\right)\begin{array}{l} \\ \\ l_3\cdot(-1) \end{array}\to$$

$$\to\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 6 & 4 \end{array}\right)\begin{array}{l} l_1-l_3 \\ l_2-l_3 \\ \\ \end{array}\to\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -4 & -3 \\ 0 & 1 & 0 & 1 & -5 & -3 \\ 0 & 0 & 1 & -1 & 6 & 4 \end{array}\right).$$

Проверка. $\begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}\cdot$

$$\begin{pmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ответ. $A^{-1}=\begin{pmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{pmatrix}.$

Упражнения.

Найти A^{-1}, если:

1\. $A=\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix}$. 2. $A=\begin{pmatrix}4 & -1\\ 1 & 1\end{pmatrix}$. 3. $A=\begin{pmatrix}2 & 3\\ 3 & -2\end{pmatrix}$.

4\. $A=\begin{pmatrix}5 & 3\\ 2 & -3\end{pmatrix}$.

5\. $A=\begin{pmatrix}2 & -1 & -1\\ 3 & 4 & -2\\ 3 & -2 & 4\end{pmatrix}$. 6. $A=\begin{pmatrix}1 & 2 & 5\\ 1 & -1 & 3\\ 3 & -6 & -1\end{pmatrix}$. 7. $A=\begin{pmatrix}1 & 2 & 3\\ 2 & -1 & -2\\ 3 & 2 & -1\end{pmatrix}$.

8\. $A=\begin{pmatrix}1 & 1 & 2\\ 2 & -1 & -1\\ -2 & 1 & -1\end{pmatrix}$.

Найти значение выражения.

9\. $g(A)=2A^2-4A^{-1}+3A$, если $A=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$.

10\. $g(A)=A+12A^{-1}-3A^2$, если $A=\begin{pmatrix}4 & -3\\ 4 & 6\end{pmatrix}$.

§5. Элементы векторной алгебры и аналитической геометрии

5.1. Векторы. Линейные операции над векторами

Скалярная величина – это величина, которая характеризуется только числом. ***Векторная величина*** – это величина, которая для своей характеристики кроме числа, требует еще указать направление (сила, скорость и др.).

Вектор – это направленный отрезок $(\mathbf{AB}, \mathbf{a})$, один из концов которого называется началом (A), а другой – концом (B). ***Длина вектора*** – скаляр, характеризующий вектор $(|\mathbf{AB}| = \mathbf{a})$. Направление вектора характеризуется заданием начала и конца или углами, которые он образует с осями координат $(\alpha; \beta; \gamma)$. Косинусы этих углов – ***направляющие косинусы*** – обладают свойством $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1$.

Над векторами можно производить алгебраические операции: умножать на число, складывать и вычитать. Свойства этих операций над векторами в основном совпадают со свойствами аналогичных операций в алгебре.

Вектор $\lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \ldots + \lambda_n \mathbf{a}_n$, полученный в результате проведения нескольких линейных операций, называется ***линейной комбинацией*** векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. Числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ называются коэффициентами этой линейной комбинации.

Понятия, которые нужно запомнить!

1. ***Орт*** – вектор, длина которого равна 1 (единичный вектор).
2. ***Нулевой вектор*** – вектор, длина которого равна 0, а направление не определено.
3. ***Векторы коллинеарны***, если они лежат на параллельных прямых линиях $(\mathbf{a} \parallel \mathbf{b})$. Условие коллинеарности $\lambda_1 \mathbf{a} + \lambda_2 \mathbf{b} = 0$ или $\mathbf{b} = \lambda \mathbf{a}$.
4. ***Векторы перпендикулярны***, если они образуют угол 90° $(\mathbf{a} \perp \mathbf{b})$.

5. ***Векторы равны*** $(\mathbf{a}=\mathbf{b})$, если они коллинеарны, имеют равные длины $(|\mathbf{a}|=|\mathbf{b}|)$ и одинаково направлены (сонаправлены $\mathbf{a}\uparrow\uparrow\mathbf{b}$).
6. ***Векторы противоположны***, если $|\mathbf{a}|=|\mathbf{b}|$, они коллинеарны и противоположно направлены $(\mathbf{a}\uparrow\downarrow\mathbf{b})$.
7. ***Векторы компланарны***, если они лежат в одной плоскости или в параллельных плоскостях.
8. ***Базисные орты*** – орты $\mathbf{i}, \mathbf{j}, \mathbf{k}$, направленные по осям координат.

Система векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ называется ***линейно-зависимой***, если существуют числа $\lambda_1, \lambda_2, \ldots, \lambda_n$, из которых хотя бы одно отлично от нуля, такие, что $\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_n\mathbf{a}_n=0$.

Если для системы векторов $\mathbf{a}_i$ $(1\le i\le n)$ это равенство верно только в случае, когда числа $\lambda_i=0$, то эта система называется ***линейно независимой***.

Пару векторов будем называть «упорядоченной парой», если указано, какой вектор пары считается первым, а какой вторым.

Упорядоченная пара линейно независимых (неколлинеарных) векторов $\mathbf{e}_1$, $\mathbf{e}_2$ называется ***базисной системой векторов*** (***базисом***) на плоскости, определяемой заданными векторами.

Теорема разложения 1. *Всякий вектор на плоскости может быть представлен как линейная комбинация базисной системы векторов, это представление* (*разложение по базисной системе*) *единственно* $\mathbf{a}=a_x\mathbf{e}_1+a_y\mathbf{e}_2$, *где* a_x, a_y *координаты вектора* $\mathbf{a}$ *в базисе* $\mathbf{e}_1$, $\mathbf{e}_2$.

Упорядоченная тройка (если указано, какой вектор тройки считается первым, какой вторым и какой третьим) линейно независимых (некомпланарных) векторов $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ называется ***базисной системой векторов*** (***базисом***) в пространстве.

Теорема разложения 2. *Всякий вектор может быть представлен как линейная комбинация базисной системы векторов; это представление (разложение по базисной системе) единственно* $\mathbf{a} = a_x \mathbf{e}_1 + a_y \mathbf{e}_2 + a_z \mathbf{e}_3$, *где* a_x, a_y, a_z координаты вектора $\mathbf{a}$ *в базисе* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

Базис называется ***ортонормированным***, если его векторы взаимно перпендикулярны и имеют единичную длину. Обозначают такой базис $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$.

5.2. Скалярное произведение

Скалярным произведением двух векторов $\mathbf{a}$ и $\mathbf{b}$ называется ***число***, которое обозначается $\mathbf{a}\cdot\mathbf{b}$ и равно произведению модулей данных векторов на косинус угла между ними

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}|\cdot|\mathbf{b}|\cos(\mathbf{a},^\wedge \mathbf{b}). \qquad (2.1)$$

Если $\mathbf{a} = a_x \mathbf{i} + a_y \mathbf{j} + a_z \mathbf{k}$, $\mathbf{b} = b_x \mathbf{i} + b_y \mathbf{j} + b_z \mathbf{k}$, то

$$\mathbf{a}\cdot\mathbf{b} = a_x b_x + a_y b_y + a_z b_z, \qquad (2.2)$$

$$|\mathbf{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}, \qquad (2.3)$$

$$\cos(\mathbf{a},^\wedge \mathbf{b}) = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{a_x b_x + a_y b_y + a_z b_z}{\sqrt{a_x^2 + a_y^2 + a_z^2}\sqrt{b_x^2 + b_y^2 + b_z^2}}. \qquad (2.4)$$

5.3. Основные задачи на прямую линию на плоскости

Любому уравнению первой степени $Ax + By + C = 0$ соответствует прямая линия на плоскости Oxy и наоборот. В зависимости от условия задачи уравнение $Ax + By + C = 0$ может быть записано по-разному.

Общее уравнение прямой линии

$$Ax + By + C = 0. \qquad (2.5)$$

Уравнение прямой линии с угловым коэффициентом

$$y = kx + b, \qquad (2.6)$$

$k = \mathrm{tg}\,\alpha$ – угловой коэффициент прямой.

Угол наклона прямой линии (α) измеряется от положительного направления оси Ox до прямой линии (против часовой стрелки). Начальная ордината b – отрезок, отсекаемый прямой от оси Oy .

Уравнение прямой линии в отрезках

$$\frac{x}{a} + \frac{y}{b} = 1, \qquad (2.7)$$

a, b – отрезки, отсекаемые прямой линией на осях Ox, Oy соответственно.

Нормальное уравнение прямой линии

$$x\cos\alpha + y\sin\alpha - p = 0. \qquad (2.8)$$

α – угол наклона нормали прямой, p – расстояние от начала координат до прямой линии.

Если даны две прямые $A_1x + B_1y + C_1 = 0,\ A_2x + B_2y + C_2 = 0$, то координаты точки пересечения находят как решение системы $\begin{cases} A_1x + B_1y + C_1 = 0, \\ A_2x + B_2y + C_2 = 0. \end{cases}$

Угол между прямыми линиями определяется формулой

$$\mathrm{tg}\,\varphi = \left|\frac{k_2 - k_1}{1 + k_2k_1}\right|, \qquad (2.9)$$

где k_1 и k_2 – угловые коэффициенты сторон угла, $0 \le \varphi < \frac{\pi}{2}$ (φ – наименьший угол между двумя прямыми).

Если ***прямые линии***
$y = k_1x + b_1,\ y = k_2x + b_2\ (A_1x + B_1y + C_1 = 0,\ A_1x + B_1y + C_1 = 0)$
параллельны, то $k_1 = k_2\ \left(\frac{A_1}{A_2} = \frac{B_1}{B_2}\right)$,
если *перпендикулярны*, то $k_1 \cdot k_2 = -1\ (A_1A_2 + B_1B_2 = 0)$.

Уравнение прямой линии, проходящей через данную точку $M_0(x_0;\, y_0)$ ***в данном направлении*** $(\mathrm{tg}\,\alpha = k)$

$$y - y_0 = k(x - x_0)\,. \qquad (2.10)$$

Если k не задано (произвольно), то это же уравнение – ***уравнение пучка прямых линий.***

Уравнение прямой, проходящей через две точки $M_1(x_1;\, y_1), M_2(x_2;\, y_2)$,

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}\,. \qquad (2.11)$$

Расстояние от точки до прямой линии. Чтобы найти расстояние от точки $(x_0;\, y_0)$ до прямой $Ax + By + C = 0$, нужно уравнение прямой привести к нормальному виду и в него вместо текущих координат подставить координаты точки $(x_0;\, y_0)$:

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}\,. \qquad (2.12)$$

5.4. Определение основных элементов в треугольнике

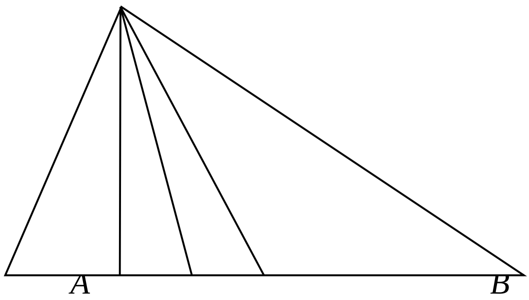
1) Определение вершин треугольника по уравнениям сторон. Приведем решение задачи в общем виде. Даны уравнения сторон: (AB), (AC), (CB). Найти вершины A, B, C.

Из схематического чертежа, очевидно, что каждая из вершин – точка пересечения соответствующих сторон. Для определения вершины A нужно решить систему $\begin{cases}(AB)\\(AC)\end{cases}$, вершины B – систему $\begin{cases}(AB)\\(BC)\end{cases}$, вершины C – систему $\begin{cases}(AC)\\(CB)\end{cases}$.

2) Вычисление длин сторон, медиан, биссектрис и высот по координатам вершин.

а) Дано: $A(x_A;\, y_A),\ B(x_B;\, y_B),\ C(x_C;\, y_C)$. Найти длины сторон.

C



Длина каждой из сторон – расстояние между соответствующими вершинами. Это одна из основных задач на метод координат и можно воспользоваться готовой формулой

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\,. \qquad (2.13)$$

Имеем $AB = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}$.

Аналогично находятся длины остальных сторон

б) Дано: $A(x_A; y_A),\ B(x_B; y_B),\ C(x_C; y_C)$. Найти длины медиан.

Длина медианы CM – расстояние CM. Можно воспользоваться формулой (2.13), но сначала нужно найти координаты точки M по формулам деления отрезка пополам:

$$x_M = \frac{x_A + x_B}{2};\ y_M = \frac{y_A + y_B}{2}\,. \qquad (2.14)$$

Аналогично находятся длины остальных медиан.

в) Дано: $A(x_A; y_A),\ B(x_B; y_B),\ C(x_C; y_C)$. Найти длины биссектрис.

Длина биссектрисы CN – расстояние между двумя точками, координаты точки N неизвестны. По свойству биссектрис точка N делит сторону AB в отношении пропорциональном длинам прилежащих сторон

$$x_N = \frac{x_A + \lambda x_B}{1+\lambda},\ y_N = \frac{y_A + \lambda y_B}{1+\lambda}\,, \qquad (2.15)$$

где $\lambda = \frac{AN}{NB} = \frac{AC}{CB}$. После определения координат точки N нужно воспользоваться формулой (2.13).

г) Дано: $A(x_A; y_A),\ B(x_B; y_B),\ C(x_C; y_C)$. Найти длины высот.

Длина высоты – расстояние от точки до прямой линии. Если уравнение AB имеет вид $A_1x + B_1y + C_1 = 0$, то по формуле (2.12)

$$AB = \frac{\left|A_1x_C + B_1y_C + C_1\right|}{\sqrt{A_1^2 + B_1^2}},$$

где $C(x_C; y_C)$ – вершина, из которой опущена высота.

Определение уравнений сторон, медиан, биссектрис и высот по координатам вершин.

Дано: $A(x_A; y_A),\ B(x_B; y_B),\ C(x_C; y_C)$. Найти $(CB), (CM), (CN), (CH)$.

а) Уравнение стороны CB – уравнение прямой, проходящей через точки C и B определяем по формуле (2.11) $\frac{y - y_C}{y_B - y_C} = \frac{x - x_C}{x_B - x_C}$.

б) Уравнение медианы CM определяется уравнением (2.11), но координаты точки M еще нужно найти по формулам деления отрезка пополам (2.14).

в) Уравнение биссектрисы CN также определяется уравнением (2.11), координаты точки N можно найти по формулам (2.15) деления отрезка в отношении $\lambda = \frac{AC}{CB}$.

г) Уравнение высоты CH определяется как уравнение прямой, проходящей через вершину C перпендикулярно стороне AB

$$y - y_C = k_{CH}(x - x_C),\ k_{CH} = -\frac{1}{k_{AB}}.$$

§6. Системы линейных уравнений и методы их решения

6.1. Системы линейных уравнений

Системой линейных алгебраических уравнений, содержащей m уравнение и n неизвестных, называется выражение вида:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases}$$

(1)

Здесь числа a_{ij} - коэффициенты системы, i – номер уравнения, i = 1, 2, …, m,
j – номер неизвестного (x_j), j = 1, 2, …, n, b_i, i = 1, 2, …, m – свободные члены.

Систему уравнений (1) можно записать в матричной форме. Для этого введем обозначения:
$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$ - матрица, составленная из коэффициентов системы. Называется матрицей коэффициентов системы. $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ - матрица-столбец

неизвестных. $B=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ - матрица-столбец свободных членов. Тогда система уравнений (1) запишется в матричном виде $A\cdot X=B$.

Если все свободные члены в системе уравнений равны нулю, то такая система называется однородной. В этом случае $A\cdot X=O$, где O – матрица-столбец, состоящая из нулей.

Матрица, составленная из коэффициентов системы и присоединенного к ней столбца свободных членов $A|B=\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$, называется расширенной матрицей коэффициентов системы.

Решением системы называется вектор $C=\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$, где $x_1=c_1,\ \ x_2=c_2,\ \ \ldots,\ \ x_n=c_n$ при подстановке координат которого в систему (1), все уравнения обращаются в тождества.

Система уравнений называется совместной, если она имеет хотя бы одно решение.

Система уравнений называется несовместной, если она не имеет решений.

Система уравнений называется определенной, если она имеет единственное решение.

Система уравнений называется неопределенной, если она имеет больше одного решения.

Две системы называются эквивалентными (равносильными), если они имеют одно и то же множество решений. Эквивалентные системы получаются при элементарных преобразованиях над строками расширенной матрицы коэффициентов системы.

6.2. Решение системы линейных уравнений методом обратной матрицы

Рассмотрим систему n линейных уравнений с n неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n. \end{cases}$$

Матрица коэффициентов системы $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$. Определитель $|A| = \Delta$ называется главным определителем системы. В матричном виде система записывается $A \cdot X = B$.

Найдем решение данной системы, в случае $|A| = \Delta \neq 0$.

Умножим обе части уравнения $A \cdot X = B$ слева на A^{-1}. Получим $A^{-1} \cdot A \cdot X = A^{-1} \cdot B$. Так как $A^{-1} \cdot A = I$ и $I \cdot X = X$, то $X = A^{-1} \cdot B$. Таким образом, задача свелась к нахождению обратной матрицы коэффициентов системы.

Метод обратной матрицы можно применять для решения матричных уравнений, когда *A, B, X, Y* –матрицы соответствующей размерности: $AX = B,\ X = A^{-1}B$; $YA = B,\ Y = BA^{-1}$.

Пример. Решить систему уравнений методом обратной матрицы. $\begin{cases} 3x_1 - 3x_2 + x_3 = 7, \\ -3x_1 + 5x_2 - 2x_3 = -10, \\ x_1 - 2x_2 + x_3 = 4. \end{cases}$

Решение. Введем обозначения $A = \begin{pmatrix} 3 & -3 & 1 \\ -3 & 5 & -2 \\ 1 & -2 & 1 \end{pmatrix}$ - матрица коэффициентов системы. $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ - матрица-столбец неизвестных. $B = \begin{pmatrix} 7 \\ -10 \\ 4 \end{pmatrix}$ - матрица-столбец свободных членов. Исходная система в матричной форме будет иметь вид $AX = B$. Ее решение $X = A^{-1} \cdot B$. Будем искать A^{-1}.

$$|A| = \Delta = \begin{vmatrix} 3 & -3 & 1 \\ -3 & 5 & -2 \\ 1 & -2 & 1 \end{vmatrix} = 15 + 6 + 6 - 5 - 9 - 12 = 1.$$

Матрица A невырожденная. Следовательно, обратная матрица A^{-1} существует. Найдем A^{-1} методом Гаусса.

$$\left(\begin{array}{ccc|ccc} 3 & -3 & 1 & 1 & 0 & 0 \\ -3 & 5 & -2 & 0 & 1 & 0 \\ 1 & -2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{меняем местами первую и третью строки}}$$

$$\left(\begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 0 & 1 \\ -3 & 5 & -2 & 0 & 1 & 0 \\ 3 & -3 & 1 & 1 & 0 & 0 \end{array}\right) \begin{array}{l} \\ l_2+3l_1 \to \\ l_3-3l_1 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 1 & 3 \\ 0 & 3 & -2 & 1 & 0 & -3 \end{array}\right) l_2\cdot(-1) \to$$

$$\left(\begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & -1 & -3 \\ 0 & 3 & -2 & 1 & 0 & -3 \end{array}\right) \begin{array}{l} l_1+2l_2 \\ \qquad\quad \to \\ l_3-3l_2 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & -2 & -5 \\ 0 & 1 & -1 & 0 & -1 & -3 \\ 0 & 0 & 1 & 1 & 3 & 6 \end{array}\right) \begin{array}{l} l_1+l_3 \\ l_2+l_3 \to \\ \end{array} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 & 3 & 6 \end{array}\right).$$

$$A^{-1}=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}. \qquad X=A^{-1}\cdot B=$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}\cdot\begin{pmatrix} 7 \\ -10 \\ 4 \end{pmatrix}=\begin{pmatrix} 7-10+4 \\ 7-20+12 \\ 3-30+24 \end{pmatrix}=\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Проверка. 3 +3 + 1 = 7, 7 = 7 – верно.

Ответ. (1, -1, 1).

6.3. Формулы Крамера

Дана система n линейных уравнений с n неизвестными $AX=B$. Ее решение $X=A^{-1}\cdot B=\frac{1}{\Delta}\cdot\tilde{A}\cdot B$.

Распишем эту формулу поэлементно:

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} A_{11}b_1 + A_{21}b_2 + \ldots + A_{n1}b_n \\ A_{12}b_1 + A_{22}b_2 + \ldots + A_{n2}b_n \\ \cdots \\ A_{1n}b_1 + A_{2n}b_2 + \ldots + A_{nn}b_n \end{pmatrix}.$$

Обозначим:

$$\Delta_1 = \begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix} = b_1 A_{11} + b_2 A_{21} + ... + b_n A_{n1},$$

$$\Delta_2 = \begin{vmatrix} a_{11} & b_1 & \cdots & a_{1n} \\ a_{21} & b_2 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & b_n & \cdots & a_{nn} \end{vmatrix} = b_1 A_{12} + b_2 A_{22} + ... + b_n A_{n2},$$

....

$$\Delta_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & b_1 \\ a_{21} & a_{22} & \cdots & b_2 \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & b_n \end{vmatrix} = b_1 A_{1n} + b_2 A_{2n} + ... + b_n A_{nn}$$ -

дополнительные определители. Определитель Δ_i получаются из главного определителя Δ заменой i-го столбца на столбец свободных членов B.

Тогда $\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} \Delta_1 \\ \Delta_2 \\ \cdots \\ \Delta_n \end{pmatrix} = \begin{pmatrix} \frac{\Delta_1}{\Delta} \\ \frac{\Delta_2}{\Delta} \\ \cdots \\ \frac{\Delta_n}{\Delta} \end{pmatrix}$. Таким образом, получаем формулы Крамера:

$$x_1 = \frac{\Delta_1}{\Delta},\ x_2 = \frac{\Delta_2}{\Delta},\ \ldots,\ x_i = \frac{\Delta_i}{\Delta},\ \ldots,\ x_n = \frac{\Delta_n}{\Delta}.$$

6.4. Исследование системы уравнений, с использованием определителей.

1. Если $\Delta \neq 0$, то система имеет единственное решение.

2. Если $\Delta = 0$ и хотя бы один из дополнительных определителей $\Delta_i \neq 0$, то система несовместна.

3. Если $\Delta = 0$ и все $\Delta_i = 0,\ i = 1, 2, \ldots, n$, то система имеет бесконечно много решений. Смысл выражения «бесконечно много решений» поймем после рассмотрения метода Гаусса.

Пример 1. Решить систему $\begin{cases} 2x_1 + 3x_2 = 1, \\ 3x_1 + 7x_2 = -1. \end{cases}$

Решение. Вычисляем главный определитель $\Delta = \begin{vmatrix} 2 & 3 \\ 3 & 7 \end{vmatrix} = 14 - 9 = 5$ и дополнительные определители: $\Delta_1 = \begin{vmatrix} 1 & 3 \\ -1 & 7 \end{vmatrix} = 7 + 3 = 10$, $\Delta_2 = \begin{vmatrix} 2 & 1 \\ 3 & -1 \end{vmatrix} = -2 - 3 = -5$. По формулам Крамера имеем: $x_1 = \frac{\Delta_1}{\Delta} = \frac{10}{5} = 2$, $x_2 = \frac{\Delta_2}{\Delta} = \frac{-5}{5} = -1$.

Ответ. (2, -1).

Пример 2. Решить систему $\begin{cases} 3x_1 - x_2 - x_3 = -2, \\ 2x_1 + 3x_2 - x_3 = -5, \\ x_1 + 2x_2 + 3x_3 = -3. \end{cases}$

Решение. Вычисляем определители:

$$\Delta = \begin{vmatrix} 3 & -1 & -1 \\ 2 & 3 & -1 \\ 1 & 2 & 3 \end{vmatrix} = 27 + 1 - 4 + 3 + 6 + 6 = 39,$$

$$\Delta_1 = \begin{vmatrix} -2 & -1 & -1 \\ -5 & 3 & -1 \\ -3 & 2 & 3 \end{vmatrix} = -18 - 3 + 10 - 9 - 15 - 4 = -39,$$

$$\Delta_2 = \begin{vmatrix} 3 & -2 & -1 \\ 2 & -5 & -1 \\ 1 & -3 & 3 \end{vmatrix} = -45 + 2 + 6 - 5 + 12 - 9 = -39,$$

$$\Delta_3 = \begin{vmatrix} 3 & -1 & -2 \\ 2 & 3 & -5 \\ 1 & 2 & -3 \end{vmatrix} = -27 + 5 - 8 + 6 - 6 + 30 = 0.$$

$$x_1 = \frac{-39}{39} = -1,\ x_2 = -1,\ x_3 = 0.$$

Ответ. (-1, -1, 0).

6.5. Метод Гаусса

Дана система m линейных уравнений с n неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases}$$

В матричном виде эта система запишется $A \cdot X = B$.

Расширенная матрица системы

$$A\,|\,B = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

Метод Гаусса – это метод последовательного исключения неизвестных, используя свойства уравнений: перестановка уравнений; умножение левой и правой частей уравнения на число, не равное нулю; прибавление к обеим частям одного из уравнений системы соответствующих частей другого уравнения, умноженных на любое число. В результате таких преобразований получается система эквивалентная исходной.

Могут возникнуть следующие строки:

1. Одно из уравнений примет вид: $0x_1 + 0x_2 + \ldots + 0x_n = b \ (b \neq 0)$. Это противоречивое уравнение. В этом случае система не имеет решений.
2. Одно из уравнений примет вид: $0x_1 + 0x_2 + \ldots + 0x_n = 0$. Его можно отбросить и продолжить процесс преобразований.

Обычно преобразовывают расширенную матрицу коэффициентов системы, приводя ее к треугольному или диагональному виду.

Если в результате преобразований, на месте матрицы коэффициентов, мы получаем единичную матрицу, то на месте свободных членов получается решение исходной системы уравнений. Т.е. выполняется $A|B \xrightarrow{\text{элементарные преобразования}} I\,|\,A^{-1}B$. В этом случае решение системы есть матрица-столбец $X = A^{-1}B$.

Если в результате преобразований возникнет строка, где на месте коэффициентов стоят нули, а на месте свободного члена – число, отличное от нуля, то система не имеет решений.

В противном случае расширенная матрица коэффициентов преобразуется к виду:

$A_1\,|\,A_2\,|\,B_1$, где A_1 - невырожденная матрица порядка r (после преобразований расширенной матрицы коэффициентов осталось r ненулевых строк, $r<n$), A_2 - матрица размерности $r\times(n-r)$ и B_1 - преобразованный столбец свободных членов. Система уравнений в матричной записи примет вид: $A_1X_1 + A_2X_2 = B_1$. Здесь X_1 - матрица-столбец из r неизвестных, соответствующих матрице A_1, X_2 - матрица-столбец из n-r оставшихся неизвестных. Назовем X_1 базисными неизвестными, а X_2 - свободными. Перенесем A_2X_2 в правую часть уравнения. Тогда имеем $A_1X_1 = -A_2X_2 + B_1$. Отсюда $X_1 = -A_1^{-1}A_2X_2 + A_1^{-1}B_1$. Давая свободным неизвестным X_2 произвольные значения, получим общее решение исходной системы уравнений.

Пример 1. Решить систему уравнений

$$\begin{cases} x_1 + x_2 - x_3 = 1, \\ 2x_1 - 3x_2 + x_3 = 0, \\ 3x_1 + x_2 + x_3 = 5. \end{cases}$$

Решение. $\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 2 & -3 & 1 & 0 \\ 3 & 1 & 1 & 5 \end{array}\right)\begin{array}{l} \\ l_2 - 2l_1 \rightarrow \\ l_3 - 3l_1 \end{array}$

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -5 & 3 & -2 \\ 0 & -2 & 4 & 2 \end{array}\right)\begin{array}{l} \\ \\ l_3 : (-2) \end{array} \rightarrow$$

$$\rightarrow \left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -5 & 3 & -2 \\ 0 & 1 & -2 & -1 \end{array}\right) \xrightarrow{\text{меняем местами вторую и третью строки}}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & -2 & -1 \\ 0 & -5 & 3 & -2 \end{array}\right)\begin{array}{l} l_1 - l_2 \\ \\ l_3 + 5l_2 \end{array} \rightarrow \quad \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & -7 & -7 \end{array}\right)\begin{array}{l} \\ \\ l_3 : (-7) \end{array} \rightarrow$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right)\begin{array}{l} l_1 - l_3 \\ l_2 + 2l_3 \\ \\ \end{array} \rightarrow \quad \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$ Таким образом, $x_1 = 1, x_2 = 2, x_3 = 3.$

Ответ. (1, 1, 1).

Пример 2. Решить систему уравнений $\begin{cases} x_1 + x_2 - x_3 = 1, \\ 2x_1 - 3x_2 + x_3 = 0, \\ 3x_1 - 2x_2 = 2. \end{cases}$

Решение. $\begin{pmatrix} 1 & 1 & -1 & \bigg| & 1 \\ 2 & -3 & 1 & \bigg| & 0 \\ 3 & -2 & 0 & \bigg| & 2 \end{pmatrix} \begin{matrix} \\ l_2 - 2l_1 \\ l_3 - 3l_1 \end{matrix} \to$

$\begin{pmatrix} 1 & 1 & -1 & \bigg| & 1 \\ 0 & -5 & 3 & \bigg| & -2 \\ 0 & -5 & 3 & \bigg| & -1 \end{pmatrix} \begin{matrix} \\ \\ l_3 - l_2 \end{matrix} \to \begin{pmatrix} 1 & 1 & -1 & \bigg| & 1 \\ 0 & -5 & 3 & \bigg| & -2 \\ 0 & 0 & 0 & \bigg| & 1 \end{pmatrix}$. Последняя строка на месте коэффициентов содержит нули, а на месте свободного члена – единицу.

Ответ. Система несовместна.

Пример 3. Решить систему уравнений
$$\begin{cases} 2x_1 + 7x_2 + 3x_3 + x_4 = 6, \\ 3x_1 + 5x_2 + 2x_3 + 2x_4 = 4, \\ 9x_1 + 4x_2 + x_3 + 7x_4 = 2, \\ 6x_1 - x_2 - x_3 + 5x_4 = -2. \end{cases}$$

Решение. $\begin{pmatrix} 2 & 7 & 3 & 1 & \bigg| & 6 \\ 3 & 5 & 2 & 2 & \bigg| & 4 \\ 9 & 4 & 1 & 7 & \bigg| & 2 \\ 6 & -1 & -1 & 5 & \bigg| & -2 \end{pmatrix} \begin{matrix} \\ l_2 - l_1 \\ \\ \\ \end{matrix} \to$

$$\to \begin{pmatrix} 2 & 7 & 3 & 1 & \bigg| & 6 \\ 1 & -2 & -1 & 1 & \bigg| & -2 \\ 9 & 4 & 1 & 7 & \bigg| & 2 \\ 6 & -1 & -1 & 5 & \bigg| & -2 \end{pmatrix} \xrightarrow{\textit{меняем местами первую ивторую строки}}$$

$$\begin{pmatrix} 1 & -2 & -1 & 1 & \bigg| & -2 \\ 2 & 7 & 3 & 2 & \bigg| & 6 \\ 9 & 4 & 1 & 7 & \bigg| & 2 \\ 6 & -1 & -1 & 5 & \bigg| & -2 \end{pmatrix} \begin{matrix} \\ l_2 - 2l_1 \\ l_3 - 9l_1 \\ l_4 - 6l_1 \end{matrix} \to$$

$$\to\left(\begin{array}{cccc|c}1 & -2 & -1 & 1 & -2\\ 0 & 11 & 5 & -1 & 10\\ 0 & 22 & 10 & -2 & 20\\ 0 & 11 & 5 & -1 & 10\end{array}\right)\begin{array}{l}\\ \\ l_3-2l_2\\ l_4-l_2\end{array}\to$$

$$\left(\begin{array}{cccc|c}1 & -2 & -1 & 1 & -2\\ 0 & 11 & 5 & -1 & 10\\ 0 & 22 & 10 & -2 & 20\\ 0 & 11 & 5 & -1 & 10\end{array}\right)\begin{array}{l}\\ \\ l_3-2l_2\\ l_4-l_2\end{array}\to$$

$$\to\left(\begin{array}{cccc|c}1 & -2 & -1 & 1 & -2\\ 0 & 11 & 5 & -1 & 10\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{array}\right)\to$$

$$\left(\begin{array}{cccc|c}1 & -2 & -1 & 1 & -2\\ 0 & 11 & 5 & -1 & 10\end{array}\right)\begin{array}{l}\\ l_2\cdot(-1)\end{array}\to$$

$$\to\left(\begin{array}{cccc|c}1 & -2 & -1 & 1 & -2\\ 0 & -11 & -5 & 1 & -10\end{array}\right)\begin{array}{l}l_1-l_2\\ \\ \end{array}\to\left(\begin{array}{cccc|c}1 & 9 & 4 & 0 & 8\\ 0 & -11 & -5 & 1 & -10\end{array}\right).$$

Пусть $A_1=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, тогда $A_2=\begin{pmatrix}9 & 4\\ -11 & -5\end{pmatrix}$. Имеем $A_1\begin{pmatrix}x_1\\ x_4\end{pmatrix}+A_2\begin{pmatrix}x_2\\ x_3\end{pmatrix}=\begin{pmatrix}8\\ -10\end{pmatrix}$. Отсюда $\begin{pmatrix}x_1\\ x_4\end{pmatrix}=-\begin{pmatrix}9 & 4\\ -11 & -5\end{pmatrix}\begin{pmatrix}x_2\\ x_3\end{pmatrix}+\begin{pmatrix}8\\ -10\end{pmatrix}$. Давая свободным неизвестным x_2 и x_3 произвольные значения $x_2=c_1,\ x_3=c_2,\ c_1,c_2\in R$, получаем: $x_1=-9c_1-4c_2+8$, $x_4=11x_2+5x_3-10$.

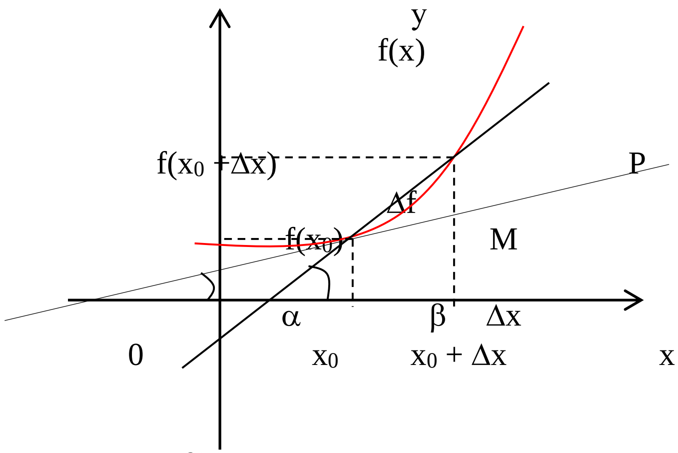
Ответ. $(-9c_1 - 4c_2 + 8; c_1; c_2; 11c_1 + 5c_2 - 10)$, $c_1, c_2 \in R$.

§7. Дифференциальное исчисление функций одной переменной

7.1. Производная функции, ее геометрический и физический смысл

Определение. Производной функции f(x) в точке x = x0 называется предел отношения приращения функции в этой точке к приращению аргумента, если он существует.

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$



Пусть *f(x)* определена на некотором промежутке (a, b). Тогда $tg\beta = \frac{\Delta f}{\Delta x}$ – тангенс угла наклона секущей MP к графику функции.

$$\lim_{\Delta x\to 0} tg\beta = \lim_{\Delta x\to 0}\frac{\Delta f}{\Delta x} = f'(x_0) = tg\alpha,$$

где α - угол наклона касательной к графику функции *f(x)* в точке (x_0, $f(x_0)$).

Угол между кривыми может быть определен как угол между касательными, проведенными к этим кривым в какой- либо точке.

Уравнение касательной к кривой: $y - y_0 = f'(x_0)(x - x_0)$

Уравнение нормали к кривой: $y - y_0 = -\frac{1}{f'(x_0)}(x - x_0)$.

Фактически производная функции показывает как бы скорость изменения функции, как изменяется функция при изменении переменной.

Физический смысл производной функции f(t), где t- время, а f(t)- закон движения (изменения координат) – мгновенная скорость движения.

Соответственно, вторая производная функции- скорость изменения скорости, т.е. ускорение.

Односторонние производные функции в точке.

Определение. Правой (левой) производной функции *f(x)* в точке x = x_0 называется правое (левое) значение предела отношения $\frac{\Delta f}{\Delta x}$ при условии, что это отношение существует.

$$f'_{+}(x_0) = \lim_{\Delta x \to 0+} \frac{\Delta f}{\Delta x} \qquad f'_{-}(x_0) = \lim_{\Delta x \to 0-} \frac{\Delta f}{\Delta x}$$

Если функция *f(x)* имеет производную в некоторой точке x = x_0, то она имеет в этой точке односторонние производные. Однако, обратное утверждение неверно. Во-первых функция может иметь разрыв в точке x_0, а во-вторых, даже если функция непрерывна в точке x_0, она может быть в ней не дифференцируема.

Например: f(x) = $|x|$- имеет в точке x = 0 и левую и правую производную, непрерывна в этой точке, однако, не имеет в ней производной.

Теорема. (Необходимое условие существования производной) *Если функция f(x) имеет производную в точке x_0, то она непрерывна в этой точке.*

Понятно, что это условие не является достаточным.

7.2. Основные правила дифференцирования.

Обозначим *f(x)* = *u*, *g(x)* = *v*- функции, дифференцируемые в точке x.

1) $(u \pm v)' = u' \pm v'$

2) $(u \cdot v)' = u \cdot v' + u' \cdot v$

3) $\left(\frac{u}{v}\right)' = \frac{u'v - v'u}{v^2}$, если $v \neq 0$

Эти правила могут быть легко доказаны на основе теорем о пределах.

Производные основных элементарных функций.

1) $C' = 0$;

2) $(x^m)' = mx^{m-1}$;

3) $\left(\sqrt{x}\right)' = \dfrac{1}{2\sqrt{x}}$

4) $\left(\dfrac{1}{x}\right)' = -\dfrac{1}{x^2}$

5) $\left(e^x\right)' = e^x$

6) $\left(a^x\right)' = a^x \ln a$

7) $(\ln x)' = \dfrac{1}{x}$

8) $(\log_a x)' = \dfrac{1}{x \ln a}$

9) $(\sin x)' = \cos x$

10) $(\cos x)' = -\sin x$

11) $(tgx)' = \dfrac{1}{\cos^2 x}$

12) $(ctgx)' = -\dfrac{1}{\sin^2 x}$

13) $(\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}$

14) $(\arccos x)' = -\dfrac{1}{\sqrt{1-x^2}}$

15) $(arctgx)' = \dfrac{1}{1+x^2}$

16) $(arcctgx)' = -\dfrac{1}{1+x^2}$

7.3. Производная сложной функции.

Теорема. *Пусть y = f(x); u = g(x), причем область значений функции u входит в область определения функции f.*

Тогда $y' = f'(u) \cdot u'$

Доказательство.

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x}$$

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta u \to 0} \frac{\Delta y}{\Delta u} \cdot \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x}$$

(с учетом того, что если $\Delta x \to 0$, то $\Delta u \to 0$, т.к. u = g(x) – непрерывная функция)

Тогда $\dfrac{dy}{dx}=\dfrac{dy}{du}\cdot\dfrac{du}{dx}$

Теорема доказана.

7.4. Логарифмическое дифференцирование

Рассмотрим функцию $y=\ln|x|=\begin{cases}\ln x, & при \quad x>0\\ \ln(-x), & при \quad x<0\end{cases}$.

Тогда (ln|x|)′= $\dfrac{1}{x}$, т.к. $(\ln x)'=\dfrac{1}{x};\quad (\ln(-x))'=\dfrac{(-x)'}{x}=\dfrac{1}{x}$.

Учитывая полученный результат, можно записать $\left(\ln|f(x)|\right)'=\dfrac{f'(x)}{f(x)}$.

Отношение $\dfrac{f'(x)}{f(x)}$ называется **логарифмической производной** функции f(x).

Способ **логарифмического дифференцирования** состоит в том, что сначала находят логарифмическую производную функции, а затем производную самой функции по формуле

$f'(x)=(\ln|f(x)|)'\cdot f(x)$

Способ логарифмического дифференцирования удобно применять для нахождения производных сложных, особенно показательных функций, для которых непосредственное вычисление производной с использованием правил дифференцирования представляется трудоемким.

7.5. Производная показательно- степенной функции

Функция называется показательной, если независимая переменная входит в показатель степени, и степенной, если переменная является основанием. Если же и основание и показатель степени зависят от переменной, то такая функция будет показательно – степенной.

Пусть u = f(x) и v = g(x) – функции, имеющие производные в точке x, f(x)>0.

Найдем производную функции y = u^v. Логарифмируя, получим:

lny = vlnu

$$\frac{y'}{y} = v'\ln u + v\frac{u'}{u}$$

$$y' = u^v\left(v\frac{u'}{u} + v'\ln u\right)$$

$$\left(u^v\right)' = vu^{v-1}u' + u^v v'\ln u$$

Пример. Найти производную функции $f(x) = (x^2 + 3x)^{x\cos x}$.

По полученной выше формуле получаем: $u = x^2 + 3x; \qquad v = x\cos x;$

Производные этих функций: $u' = 2x + 3; \quad v' = \cos x - x\sin x;$

Окончательно:

$$f'(x) = x\cos x\cdot(x^2+3x)^{x\cos x-1}\cdot(2x+3) + (x^2+3x)^{x\cos x}(\cos x - x\sin x)\ln(x^2+3x)$$

7.6. Производная обратных функций

Пусть требуется найти производную функции *y = f(x)* при условии, что обратная ей функция *x = g(y)* имеет производную, отличную от нуля в соответствующей точке.

Для решения этой задачи дифференцируем функцию x = g(y) по x:

$$y' = \frac{1}{g'(y)}$$

$1 = g'(y)y'$, т.к. $g'(y) \neq 0$

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}$$

т.е. производная обратной функции обратна по величине производной данной функции.

Пример. Найти формулу для производной функции arctg.

Функция arctg является функцией, обратной функции tg, т.е. ее производная может быть найдена следующим образом:

$y = tgx; \qquad x = arctgy;$

Известно, что $y' = (tgx)' = \dfrac{1}{\cos^2 x};$

По приведенной выше формуле получаем:

$$y' = \frac{1}{d(arctgy)/dx}; \qquad \frac{d(arctgy)}{dy} = \frac{1}{1/\cos^2 x}$$

Т.к. $\dfrac{1}{\cos^2 x} = 1 + tg^2 x = 1 + y^2;$ то можно записать окончательную формулу для производной арктангенса:

$$(arctgy)' = \frac{1}{1+y^2};$$

Таким образом получены все формулы для производных арксинуса, арккосинуса и других обратных функций, приведенных в таблице производных.

7.7. Дифференциал функции

Пусть функция y = f(x) имеет производную в точке x:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x)$$

Тогда можно записать: $\dfrac{\Delta y}{\Delta x} = f'(x) + \alpha$, где $\alpha \to 0$, при $\Delta x \to 0$.

Следовательно: $\Delta y = f'(x) \cdot \Delta x + \alpha \cdot \Delta x$.

Величина $\alpha\Delta x$- бесконечно малая более высокого порядка, чем $f'(x)\Delta x$, т.е. $f'(x)\Delta x$- главная часть приращения Δy.

Определение. Дифференциалом функции f(x) в точке x называется главня линейная часть приращения функции.

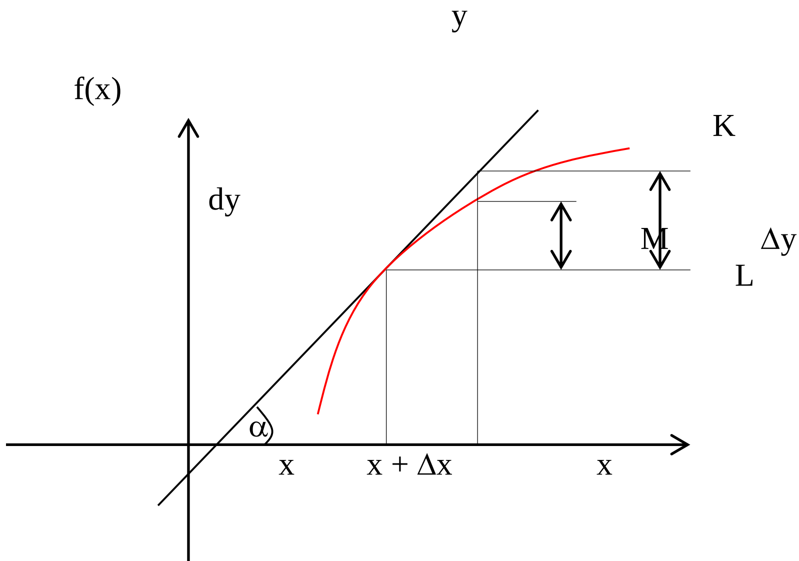
Обозначается dy или df(x).

Из определения следует, что $dy = f'(x)\Delta x$ или

$dy = f'(x)dx$.

Можно также записать: $$f'(x) = \frac{dy}{dx}$$

Геометрический смысл дифференциала.



Из треугольника ΔMKL: KL = $dy = tg\alpha \cdot \Delta x = y' \cdot \Delta x$

Таким образом, дифференциал функции *f(x)* в точке x равен приращению ординаты касательной к графику этой функции в рассматриваемой точке.

Свойства дифференциала

Если $u = f(x)$ и $v = g(x)$- функции, дифференцируемые в точке x, то непосредственно из определения дифференциала следуют следующие свойства:

1) $d(u \pm v) = (u \pm v)'dx = u'dx \pm v'dx = du \pm dv$

2) $d(uv) = (uv)'dx = (u'v + v'u)dx = vdu + udv$

3) $d(Cu) = Cdu$

4) $d\left(\frac{u}{v}\right) = \frac{vdu - udv}{v^2}$

Дифференциал сложной функции

Инвариантная форма записи дифференциала.

Пусть $y = f(x)$, $x = g(t)$, т.е y- сложная функция.

Тогда $dy = f'(x)g'(t)dt = f'(x)dx$.

Видно, что форма записи дифференциала dy не зависит от того, будет ли x независимой переменной или функцией какой- то другой переменной, в связи с чем эта форма записи называется инвариантной формой записи дифференциала.

Однако, если x- независимая переменная, то

$dx = \Delta x$, но

если x зависит от t, то $\Delta x \neq dx$.

Таким образом форма записи $dy = f'(x)\Delta x$ не является инвариантной.

Пример. Найти производную функции $y = x\cos x\sin x + \frac{1}{2}\cos^2 x$.

Сначала преобразуем данную функцию: $y = \frac{1}{2}\sin 2x + \frac{1}{2}\cos^2 x$

$$y' = \frac{1}{2}\sin 2x + \frac{1}{2}x2\cos 2x + \frac{1}{2}2\cos x(-\sin x) = \frac{1}{2}\sin 2x + x\cos 2x - \sin x\cos x = x\cos 2x.$$

Пример. Найти производную функции $y = \frac{x^2 e^{x^2}}{x^2 + 1}$.

$$y' = \frac{(2xe^{x^2} + x^2 2xe^{x^2})(x^2+1) - (2x)x^2 e^{x^2}}{(x^2+1)^2} = \frac{2x^3 e^{x^2} + 2x^5 e^{x^2} + 2xe^{x^2} + 2x^3 e^{x^2} - 2x^3 e^{x^2}}{(x^2+1)^2} =$$

$$= \frac{2xe^{x^2}(x^4+1+x^2)}{(x^2+1)^2}$$

Пример. Найти производную функции $y = \ln tg\frac{x}{2} - \frac{x}{\sin x}$

$$y' = \frac{1}{tg\frac{x}{2}} \cdot \frac{1}{\cos^2\frac{x}{2}} \cdot \frac{1}{2} - \frac{\sin x - x\cos x}{\sin^2 x} = \frac{1}{2\sin\frac{x}{2}\cos\frac{x}{2}} - \frac{\sin x - x\cos x}{\sin^2 x} = \frac{\sin x - \sin x + x\cos x}{\sin^2 x} =$$

$$= \frac{x\cos x}{\sin^2 x}$$

Пример. Найти производную функции $y = arctg\frac{2x^4}{1-x^8}$

$$y' = \frac{1}{\left(1 + \frac{4x^8}{(1-x^8)^2}\right)} \cdot \frac{8x^3(1-x^8) - (-8x^7)2x^4}{(1-x^8)^2} = \frac{(1-x^8)^2(8x^3 - 8x^{11} + 16x^{11})}{(1+x^8)^2(1-x^8)^2} = \frac{8x^3 + 8x^{11}}{(1+x^8)^2} =$$

$$= \frac{8x^3(1+x^8)}{(1+x^8)^2} = \frac{8x^3}{1+x^8}$$

Пример. Найти производную функции $y = x^2 e^{x^2} \ln x$

$$y' = \left(x^2 e^{x^2}\right)' \ln x + x^2 e^{x^2} \frac{1}{x} = \left(2xe^{x^2} + x^2 e^{x^2} 2x\right)\ln x + xe^{x^2} = 2xe^{x^2}(1+x^2)\ln x + xe^{x^2} =$$

$$= xe^{x^2}(1 + 2\ln x + 2x^2 \ln x)$$

ГЛАВА II. ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

§ 8. Элементы теории множеств

Два математика в ресторане поспорили, насколько хорошо знают математику большинство людей. Один (пессимист) утверждал, что большинство ее вообще не знает, а другой (оптимист) – что хоть и немного, но знают. Когда пессимист отошел в туалет, оптимист подозвал симпатичную официантку – блондинку и говорит:

«Когда мой коллега вернется, я задам вам вопрос. Суть не важна. Все, что вы должны сделать – это сказать: "Треть икс куб".

- Как-как? Третий скуп? – переспрашивает официантка?

- Да нет, Треть Икс Куб, понятно?

- А-а! Третик скуп? – повторяет официантка.

- Да, да. Это все о чем я вас прошу».

Официантка уходит, твердя про себя как заклинание фразу «Третик скуп».

Тут возвращается пессимист. Оптимист говорит – давай спросим у нашей официантки чему равен какой-нибудь простенький интеграл. Пессимист, со смехом соглашается. Оптимист вызывает официантку и спрашивает:

- Извините, вы не помните, чему равен интеграл x^2 по dx.

- Треть икс куб.. – отвечает официантка.

Пессимист сильно удивлен, оптимист весело смеется. Официантка отходит на несколько шагов, и обернувшись через плечо добавляет:

- Плюс константа.

- Товарищи студенты! Начинаем изучать дисциплину «Теория вероятностей и математическая статистика».

Методы теории вероятностей широко применяются в различных отраслях естествознания и техники: в теории надежности, теории массового обслуживания, в теоретической физике, геодезии, астрономии, теории стрельбы, теории ошибок наблюдений, теории автоматического управления, общей теории связи и во многих других теоретических и прикладных науках. Теория вероятностей служит также для обоснования математической и прикладной статистики, которая в свою очередь используется при планировании и организации производства, при анализе технологических процессов, предупредительном и приемочном контроле качества продукции и для многих других целей.

В последние годы методы теории вероятностей все шире и шире проникают в различные области науки и механики, способствуя их прогрессу.

8.1. Основные понятия и предмет теории вероятностей

Теорией вероятностей называется наука, изучающая закономерности в случайных явлениях.

Наблюдаемые нами события (явления) можно подразделить на следующие три вида: достоверные, невозможные и случайные. Достоверным называют событие, которое обязательно произойдет, если будет осуществлена определенная совокупность условий *S*. Невозможным, называют событие, которое заведомо не произойдет, если будет осуществлена совокупность условий *S*. Например, событие «вода в сосуде находится в твердом состоянии» заведомо не произойдет, если будет осуществлена совокупность условий предыдущего примера. Случайным называют событие, которое при осуществлении совокупности условий S может либо произойти, либо не произойти. Например, если брошена монета, то она может упасть так, что сверху будет либо герб, либо надпись. Поэтому событие «при бросании монеты выпал «герб» – случайное. Большое число однородных случайных событий независимо от их конкретной природы подчиняется определенным закономерностям, а именно вероятностным закономерностям. Установлением этих закономерностей и занимается теория вероятностей.

Основным интуитивным понятием классической теории вероятностей является случайное событие. События, которые могут произойти в результате опыта, можно подразделить на три вида:

а) достоверное событие – событие, которое всегда происходит при проведении опыта;

б) невозможное событие – событие, которое в результате опыта произойти не может;

в) случайное событие – событие, которое может либо произойти, либо не произойти. Например, при броске игральной кости достоверным событием является выпадение числа очков, не превышающего 6, невозможным – выпадение 10 очков, а случайным – выпадение 3 очков.

Предметом теории вероятностей является изучение вероятностных закономерностей массовых однородных случайных событий.

Знание закономерностей, которым подчиняются массовые случайные события, позволяет предвидеть, как эти события будут протекать. Например, хотя, как было уже сказано, нельзя наперед определить результат одного бросания монеты, но можно предсказать, причем с небольшой погрешностью, число появлений «герба», если монета будет брошена достаточно большое число раз. При этом предполагается, конечно, что монету бросают в одних и тех же условиях.

1. Возникновение теории вероятностей относят обычно к XVII веку и связывают с комбинаторными задачами азартных игр. Азартные игры трудно считать серьезным занятием. Но именно они привели к задачам, которые не укладывались в рамки существовавших тогда математических моделей и стимулировали тем самым введение новых понятий, подходов и идей. Эти новые элементы можно встретить уже у Ферма, Паскаля, Гюйгенса и, в более развитой форме, чуть позже, у Якоба Бернулли, Лапласа, Гаусса и др. Имена перечисленных ученых, безусловно, украшают родословную теории вероятностей, в какой-то мере связанную, как мы видели, с пороками общества. Впрочем, как выяснилось, иногда именно это обстоятельство может сообщать ей в глазах читателей некоторую дополнительную привлекательность.

Первым руководством по теории вероятностей был трактат Гюйгенса «О расчетах в азартной игре», вышедший в 1657 году. Предмет этого трактата тот же, что и в работах Ферма и Паскаля: игральные кости и карточные игры. Словно предвидя дальнейшее развитие событий, Гюйгенс писал: «Я полагаю, что при внимательном изучении предмета читатель заметит, что имеет дело не только с игрой, но что здесь закладываются основы очень интересной и глубокой теории». Трактат Гюйгенса, известный также тем, что в нем впервые вводится понятие математического ожидания, был включен позднее Якобом Бернулли в знаменитую книгу «Искусства предположений», опубликованную уже после его смерти в 1713 году. С этой книгой связаны понятие схемы Бернулли и доказательство первой предельной теоремы теории вероятностей — закона больших чисел. Доказательство Бернулли хотя и громоздко, но математически безупречно.

В конце прошлого и начале этого века стали возникать более серьезные проблемы, порожденные нуждами естествознания и приведшие к развитию большого, в значительной мере самостоятельного раздела математики, именуемого сегодня теорией вероятностей. Эта область знаний вплоть до настоящего времени находится в состоянии интенсивного развития. Своей стройностью, современной формой и многими своими достижениями теория вероятностей во многом обязана трудам наших замечательных соотечественников П. Л. Чебышева, А. А. Маркова, А. Н. Колмогорова и др.

То обстоятельство, что увеличение наших знаний о природе предъявляет все новые требования к теории вероятностей, на первый взгляд представляется парадоксом. Ведь основной объект теории вероятностей, как читатель, по-видимому, уже знает, есть случайность или неопределенность, как правило, связанная с незнанием. Именно так обстоит дело в классическом примере с подбрасыванием монеты, где трудно учесть все факторы, влияющие на положение монеты после падения.

Однако отмеченный парадокс является лишь видимым. На самом деле точных, детерминированных количественных законов в природе почти не существует. Скажем, классический пример таких законов – закон о зависимости давления газа от его температуры – есть на самом деле результат вероятностного характера о числе соударений частиц о стенки сосуда и их скоростях. Просто в области обычных температур и давлений случайные отклонения, которые тут имеют место, с большой вероятностью очень малы и не регистрируются нашими приборами. Иначе обстоит дело при изучении более редких потоков частиц, скажем, космического излучения, хотя качественной разницы между этими двумя примерами нет.

Можно пойти и в несколько ином направлении и указать на принцип неопределенности, в силу которого для любой пары физических характеристик, связанных этим принципом, фиксация одной из них делает невозможным точное определение другой. Тут уже случайность появляется не как следствие недостаточности наших знаний, а как принципиальное явление и отражение природы вещей. Например, время жизни радиоактивного ядра случайно по существу, и эта случайность не может быть устранена увеличением наших познаний.

Таким образом, неопределенность стояла в начале процесса познания, она будет стоять и на всем пути его. Эти замечания носят, конечно, общий характер. Но вопрос о том, когда следует применять методы теории вероятностей и когда нет, по-видимому, всегда будет определяться соотношением между степенью точности, с которой мы хотим изучать данное явление, и сведениями о его природе, которыми мы располагаем.

2. Почти во всех областях человеческой деятельности существуют ситуации, когда те или иные эксперименты или наблюдения могут быть повторены большое число раз в одинаковых условиях. Теорию вероятностей интересуют те эксперименты, результат которых, выраженный каким-либо образом, может меняться от опыта к опыту. События, относящиеся к результату эксперимента, которые при этом могут происходить или не происходить, называют обычно случайными событиями.

Мы можем, например, подбрасывать монету. Эксперимент имеет всего два исхода: монета падает либо гербом вверх, либо решеткой. При этом до получения результата эксперимента нельзя сказать, какой именно исход осуществится. Как уже отмечалось, это происходит оттого, что мы практически не в состоянии учесть все факторы, влияющие на положение монеты в момент падения. Примерно то же самое будет происходить, если вы при каждой лотерее будете покупать один билет и будете пытаться предугадать, выиграет он или нет. Или, наблюдая за работой достаточно сложного механизма, будете пытаться заранее определить, выйдет он из строя до назначенного срока или после. В таких ситуациях при рассмотрении результатов отдельных экспериментов бывает очень трудно обнаружить какие-либо закономерности. И, стало быть, здесь мало оснований строить какую-либо теорию.

Однако, если обратить внимание на последовательность большого числа такого рода одинаковых экспериментов, то обнаружится интересное явление. Если индивидуальные результаты опытов ведут себя «неправильно», то средние результаты обнаруживают устойчивость. Будем, например, повторять наш эксперимент с подбрасыванием монеты и обозначим через $n_{\text{г}}$ число выпадений герба после первых n испытаний. Построим следующий график: на оси абсцисс отложим число проведенных экспериментов, а на оси ординат – отношение ${}^{n_{\text{г}}}/_{n}$ (рис. 1; график построен для последовательности исходов грргрргггрггр, где г — герб, р — решетка).

Мы заметим, что ломаная, соединяющая точки $\left(n, {}^{n_{\text{г}}}/_{n}\right)$, с ростом n очень быстро прижимается к прямой ${}^{n_{\text{г}}}/_{n} = {}^{1}/_{2}$. Чтобы проверить это обстоятельство, Бюффон в XVIII веке провел 4040 подбрасываний монеты.

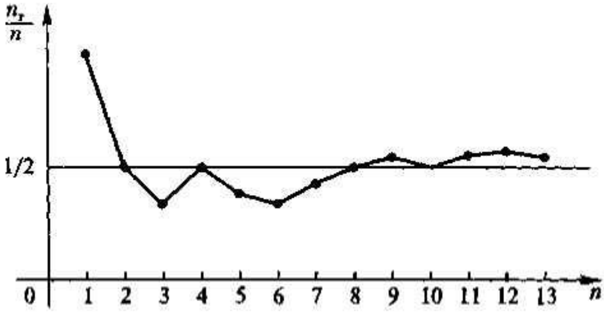


Рис. 1

Из них герб выпал 2048 раз, так что частота выпадения герба $^{n_г}/_{n}$ оказалась равной 0,508. Пирсон провел 24000 бросаний монеты; герб выпал 12012 раз, $^{n_г}/_{n} = 0{,}5005$.

Оказывается, это явление имеет общий характер: частота осуществления какого-либо исхода в последовательности повторяемых в одинаковых условиях экспериментов приближается к некоторому числу $p \in [0,1]$ при росте числа экспериментов. Это объективный закон природы, который лежит в основе теории вероятностей.

Естественно было бы принять за вероятность некоторого исхода в эксперименте как раз то число p, к которому приближается частота появления этого исхода. Но такое определение вероятности (его связывают обычно с именем Мизеса) оказывается неудобным. Прежде всего, последовательность частот $\{^{n_г}/_{n}\}$ появления некоторого исхода при проведении одной серии экспериментов будет, как правило, отличаться от последовательности частот появления того же исхода при другой серии экспериментов. Кроме того, на самом деле мы будем иметь не бесконечную последовательность частот, а только конечное число ее элементов. Получить всю последовательность невозможно.

По-видимому, вероятность нужно определять иначе, но так, чтобы имело место отмеченное желательное свойство частоты: в каком-то смысле отношение $n_г/_n$ должно приближаться при увеличении и к вероятности рассматриваемого события. Отсюда можно извлечь, в частности, что вероятностями могут быть только числа $p \in [0,1]$, при этом вероятность достоверного события следует считать равной 1.

8.2. Сигма-алгебра событий

Современная теория вероятностей основывается на понятии *вероятностного пространства*, одним из трех компонентов которого является *сигма-алгебра событий*.

Назовем событием любое подмножество *пространства элементарных исходов* Ω. Такое определение допустимо, если Ω является конечным или счетным множеством. Оказывается, однако, что в случае несчетного множества элементарных исходов уже нельзя построить логически непротиворечивую теорию, называя событием произвольное подмножество множества Ω. Поэтому событиями в этом случае называют не любые подмножества элементарных исходов, а только подмножества из Ω, принадлежащие некоторому классу Ђ. Этот класс в теории множеств принято называть сигма-алгеброй событий (пишут σ-алгебра).

С точки зрения здравого смысла событие – это то, что мы наблюдаем после проведения опыта. В частности, если можно после опыта установить, произошли или нет события A и B, то можно также сказать, произошли или нет события $\bar{A}$ и $\bar{B}$, *объединение, пересечение и разность событий* A и B. Таким образом, σ-алгебра событий обязана быть классом подмножеств, замкнутым относительно приведенных операций над подмножествами, т. е. указанные операции над элементами (подмножествами) данного класса приводят к элементам (подмножествам) того же класса.

Дадим теперь строгое определение σ-алгебры событий

Определение 1. Сигма-алгеброй (σ-алгеброй) Ђ называют непустую систему подмножеств некоторого множества J, удовлетворяющую следующим двум условиям.

1. Если подмножество A принадлежит $Ђ$, то дополнение $\bar{A}$ принадлежит $Ђ$.

2. Если подмножества $A_1, A_2, \ldots, A_n, \ldots$ принадлежат $Ђ$, то их объединение $A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots$ и их пересечение $A_1, A_2, \ldots, A_n, \ldots$ принадлежат $Ђ$.

Поскольку $J = A \cup \overline{A}$ и $\varnothing = \bar{J}$, то множество J и пустое множество $\varnothing$ принадлежат $Ђ$.

Рассмотрим пространство элементарных исходов Ω. Элементы некоторой σ-алгебры $Ђ$, заданной на Ω, будем называть ***событиями***. В этом случае σ-алгебру $Ђ$ принято называть ***сигма-алгеброй (σ-алгеброй) событий***.

Любая σ-алгебра событий содержит *достоверное событие* Ω и *невозможное событие* $\varnothing$.

В случае конечного или счетного пространства элементарных исходов Ω в качестве σ-алгебры событий обычно рассматривают множество всех подмножеств Ω.

Замечание 1. Если в условии 2 счетное множество событий заменить на конечное, то получим определение ***алгебры событий***. Любая σ-алгебра событий обязательно является алгеброй событий. Обратное утверждение, вообще говоря, не верно.

Пример 1. Пусть опыт состоит в подбрасывании один раз тетраэдра, каждая грань которого помечена одним из чисел 1, 2, 3 и 4.

Очевидно, что пространство элементарных исходов Ω в этом опыте имеет вид

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\},$$

где ω_i – падение тетраэдра на грань с числом i, $i = \overline{1}, \overline{4}$.

Поскольку в рассматриваемом опыте может происходить одно из следующих событий $\emptyset$,

$$\{\omega_1\},\quad \{\omega_2\},\quad \{\omega_3\},\quad \{\omega_4\},$$

$$\{\omega_1,\omega_2\},\quad \{\omega_1,\omega_3\},\quad \{\omega_1,\omega_4\},\quad \{\omega_2,\omega_3\},\quad \{\omega_2,\omega_4\},\quad \{\omega_3,\omega_4\}$$

$$\{\omega_1,\omega_2,\omega_3\},\quad \{\omega_1,\omega_2,\omega_4\},\quad \{\omega_1,\omega_3,\omega_4\},\quad \{\omega_2,\omega_3,\omega_4\},$$

$$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\}$$

то алгебра событий будет содержать все подмножества *Ω*, включая *Ω* (достоверное событие) и Ø (невозможное событие).

Пример 2. Пусть опыт состоит в случайном бросании точки на числовую прямую $\boldsymbol{R}^1 = (-\infty, +\infty)$, которая в данном случае будет представлять собой пространство элементарных исходов *Ω*. Ясно, что, зная результат опыта, всегда можно установить, попала или нет точка в любой из промежутков $[a,\ b]$, $[a,\ b)$, $(a,\ b]$, $(a,\ b)$. Поэтому относительно σ-алгебры событий *Ъ* предполагают, что она содержит все эти промежутки.

В принципе могут существовать различные σ-алгебры, удовлетворяющие этому требованию. Но среди них есть одна σ-алгебра, элементы которой принадлежат всем остальным. Ее называют ***минимальной***, или ***борелевской***, ***σ-алгеброй*** на числовой прямой.

Аналогично определяют борелевскую σ-алгебру и в $\boldsymbol{R}^n$, $n > 1$.

Заметим, что с точки зрения повседневной практики подмножества пространства *Ω*, элементарных исходов, не являющиеся событиями, представляют собой чистую математическую абстракцию и в практических задачах никогда не встречаются. Даже само доказательство их существования представляет весьма сложную задачу. Поэтому мы предлагаем при первоначальном знакомстве с теорией вероятностей под событием понимать произвольное подмножество пространства Ω элементарных исходов, а под σ-алгеброй событий – совокупность всех подмножеств множества Ω.

8.3. Множества и операции над ними

Определение 1. Суммой $A+B$ двух событий A и B называют событие, состоящее в том, что произошло хотя бы одно из событий A и B. Суммой нескольких событий, соответственно, называется событие, заключающееся в том, что произошло хотя бы одно из этих событий.

Пример 1. Два стрелка делают по одному выстрелу по мишени. Если событие A – попадание первого стрелка, а событие B – второго, то сумма $A+B$ – это хотя бы одно попадание при двух выстрелах.

Пример 2. Если при броске игральной кости событием A_i назвать выпадение i очков, то выпадение нечетного числа очков является суммой событий $A_1+A_2+A_3$.

Назовем все возможные результаты данного опыта его исходами и предположим, что множество этих исходов, при которых происходит событие A (исходов, благоприятных событию A), можно представить в виде некоторой области на плоскости. Тогда множество исходов, при которых произойдет событие $A+B$, является объединением множеств исходов, благоприятных событиям A или B (рис. 1).

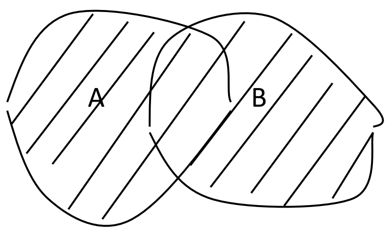


Рис. 1 - $A+B$

Определение 2. Произведением AB событий A и B называется событие, состоящее в том, что произошло и событие A, и событие B. Аналогично произведением нескольких событий называется событие, заключающееся в том, что произошли все эти события.

Пример 3. В примере 1 (два выстрела по мишени) событием *AB* будет попадание обоих стрелков.

Пример 4. Если событие *A* состоит в том, что из колоды карт извлечена карта пиковой масти, а событие *B* – в том, что из колоды вынута дама, то событием *AB* будет извлечение из колоды дамы пик.

Геометрической иллюстрацией множества исходов опыта, благоприятных появлению произведения событий *A* и *B*, является пересечение областей, соответствующих исходам, благоприятным *A* и *B*.

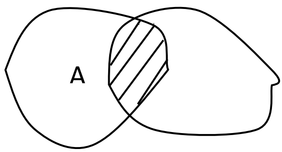


Рис. 2 - AB

Определение 3. Разностью *A*-*B* событий *A* и *B* называется событие, состоящее в том, что *A* произошло, а *B* – нет.

Пример 5. Вернемся к примеру 1, где *A*-*B* – попадание первого стрелка при промахе второго.

Пример 6. В примере 4 *A*-*B* – извлечение из колоды любой карты пиковой масти, кроме дамы. Наоборот, *B*-*A* – извлечение дамы любой масти, кроме пик.

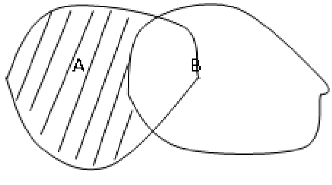


Рис. 3 - A-B.

Введем еще несколько категорий событий.

Определение 4. События A и B называются совместными, если они могут произойти оба в результате одного опыта. В противном случае (то есть если они не могут произойти одновременно) события называются несовместными.

Примеры: совместными событиями являются попадания двух стрелков в примере 1 и появление карты пиковой масти и дамы в примере 4; несовместными – события $A_1 - A_6$ в примере 2.

Пример 7. Из ящика с деталями наудачу извлечена деталь. Появление стандартной детали исключает появление нестандартной детали. События «появилась стандартная деталь» и «появилась нестандартная деталь» – несовместные.

Пример 8. Брошена монета. Появление «герба» исключает появление надписи. События «появился герб» и «появилась надпись» – несовместные.

Замечание 1. Если изобразить графически области исходов опыта, благоприятных несовместным событиям, то они не будут иметь общих точек.

Замечание 2. Из определения несовместных событий следует, что их произведение является невозможным событием.

Определение 5. Говорят, что события $A_1, A_2, \ldots, A_n$ образуют полную группу, если в результате опыта обязательно произойдет хотя бы одно из событий этой группы.

Замечание 3. В частности, если события, образующие полную группу, попарно несовместны, то в результате опыта произойдет одно и только одно из них. Такие события называют элементарными событиями.

Пример 9. В примере 4 события $A_1 - A_6$ (выпадение одного, двух,…, шести очков при одном броске игральной кости) образуют полную группу несовместных событий.

<u>Пример 10.</u> Приобретены два билета денежно-вещевой лотереи. Обязательно произойдет одно и только одно из следующих событий: «выигрыш выпал на первый билет и не выпал на второй», «выигрыш не выпал на первый билет и выпал на второй», «выигрыш выпал на оба билета», «на оба билета выигрыш не выпал». Эти события образуют полную группу попарно несовместных событий.

<u>Пример 11.</u> Стрелок произвел выстрел по цели. Обязательно произойдет одно из следующих двух событий: попадание, промах. Эти два несовместных события образуют полную группу.

Определение 6. События называются равновозможными, если нет оснований считать, что одно из них является более возможным, чем другое.

<u>Примеры:</u> выпадение любого числа очков при броске игральной кости, появление любой карты при случайном извлечении из колоды, выпадение герба или цифры при броске монеты и т. п.

<u>Пример 12.</u> Появление «герба» и появление надписи при бросании монеты – равновозможные события. Действительно, предполагается, что монета изготовлена из однородного материала, имеет правильную цилиндрическую форму и наличие чеканки не оказывает влияния на выпадение той или иной стороны монеты.

<u>Пример 13.</u> Появление того или иного числа очков на брошенной игральной кости – равновозможные события. Действительно, предполагается, что игральная кость изготовлена из однородного материала, имеет форму правильного многогранника и наличие очков не оказывает влияния на выпадение любой грани.

Аксиоматика теории вероятностей

Вероятность – одно из основных понятий теории вероятностей. Существует несколько определений этого понятия. Приведем определение, которое называют классическим. Далее укажем слабые стороны этого определения и приведем другие определения, позволяющие преодолеть недостатки классического определения.

Рассмотрим пример. Пусть в урне содержится 6 одинаковых, тщательно перемешанных шаров, причем 2 из них – красные, 3 – синие и 1 – белый. Очевидно, возможность вынуть наудачу из урны цветной (т. е. красный или синий) шар больше, чем возможность извлечь белый шар. Можно ли охарактеризовать эту возможность числом? Оказывается, можно. Это число и называют вероятностью события (появления цветного шара). Таким образом, вероятность есть число, характеризующее степень возможности появления события.

§9. Классическое определение вероятности

9.1. Классическое определение вероятности

Поставим перед собой задачу дать количественную оценку возможности того, что взятый наудачу шар цветной. Появление цветного шара будем рассматривать в качестве события A. Каждый из возможных результатов испытания (испытание состоит в извлечении шара из урны) назовем элементарным исходом (элементарным событием). Элементарные исходы обозначим через $\omega_1, \omega_2, \omega_3$ и т. д. В нашем примере возможны следующие 6 элементарных исходов: ω_1– появился белый шар; ω_2, ω_3 – появился красный шар; $\omega_4, \omega_5, \omega_6$ – появился синий шар. Легко видеть, что эти исходы образуют полную группу попарно несовместных событий (обязательно появится только один шар) и они равновозможны (шар вынимают наудачу, шары одинаковы и тщательно перемешаны).

Те элементарные исходы, в которых интересующее событие наступает, назовем благоприятствующими этому событию. В нашем примере благоприятствуют событию A (появлению цветного шара) следующие 5 исходов: $\omega_2, \omega_3, \omega_4, \omega_5, \omega_6$.

Таким образом, событие A наблюдается, если в испытании наступает один, безразлично какой, из элементарных исходов, благоприятствующих A; в нашем примере A наблюдается, если наступит ω_2, или ω_3, или ω_4, или ω_5, или ω_6. В этом смысле событие A подразделяется на несколько элементарных событий $(\omega_2, \omega_3, \omega_4, \omega_5, \omega_6)$; элементарное же событие не подразделяется на другие события. В этом состоит различие между событием A и элементарным событием (элементарным исходом).

При изучении случайных событий возникает необходимость количественно сравнивать возможность их появления в результате опыта. Например, при последовательном извлечении из колоды пяти карт более возможна ситуация, когда появились карты разных мастей, чем появление пяти карт одной масти; при десяти бросках монеты более возможно чередование гербов и цифр, нежели выпадение подряд десяти гербов, и т. д. Поэтому с каждым таким событием связывают по определенному правилу некоторое число, которое тем больше, чем более возможно событие. Это число называется вероятностью события и является вторым основным понятием теории вероятностей.

Отметим, что само понятие вероятности, как и понятие случайного события, является аксиоматическим и поэтому не поддается строгому определению. То, что в дальнейшем будет называться различными определениями вероятности, представляет собой способы вычисления этой величины.

Определение 1. Если все события, которые могут произойти в результате данного опыта,

а) попарно несовместны;

б) равновозможны;

в) образуют полную группу,

то говорят, что имеет место схема случаев.

Можно считать, что случаи представляют собой все множество исходов опыта. Пусть их число равно n (число возможных исходов), а при m из них происходит некоторое событие А (число благоприятных исходов).

Определение 2. Вероятностью события А называется отношение числа исходов опыта, благоприятных этому событию, к числу возможных исходов:

$$p(A) = \frac{m}{n} \qquad (1.1)$$

– классическое определение вероятности.

Из определения 2 вытекают следующие свойства вероятности:

Свойство 1. Вероятность достоверного события равна единице.

Доказательство. Так как достоверное событие всегда происходит в результате опыта, то все исходы этого опыта являются для него благоприятными, то есть $m = n$, следовательно,

$$P(A) = 1.$$

Свойство 2. Вероятность невозможного события равна нулю.

Доказательство. Для невозможного события ни один исход опыта не является благоприятным, поэтому $m = 0$ и $p(A) = 0$.

Свойство 3. Вероятность случайного события есть положительное число, заключенное между нулем и единицей.

Доказательство. Случайное событие происходит при некоторых исходах опыта, но не при всех, следовательно, 0 < m < n, и из (1.1) следует, что 0 < $p(A)$ < 1.

Пример 1. Из урны, содержащей 6 белых и 4 черных шара, наудачу вынут шар. Найти вероятность того, что он белый.

Решение. Будем считать элементарными событиями, или исходами опыта, извлечение из урны каждого из имеющихся в ней шаров. Очевидно, что эти события удовлетворяют всем условиям, позволяющим считать их схемой случаев. Следовательно, число возможных исходов равно 10, а число исходов, благоприятных событию A (появлению белого шара) – 6 (таково количество белых шаров в урне). Значит,

$$p(A) = \frac{m}{n} = \frac{6}{10} = 0{,}6$$

Замечание. Современные строгие курсы теории вероятностей построены на теоретико-множественной основе. Ограничимся изложением на языке теории множеств тех понятий, которые рассмотрены выше.

Пусть в результате испытания наступает одно и только одно из событий ω_i (i = 1, 2, ..., n). События ω_i – называют элементарными событиями (элементарными исходами). Уже отсюда следует, что элементарные события попарно несовместны. Множество всех элементарных событий, которые могут появиться в испытании, называют пространством элементарных событий Ω, а сами элементарные события – точками пространства Ω.

Событие A отождествляют с подмножеством (пространства Ω), элементы которого есть элементарные исходы, благоприятствующие A; событие B есть подмножество Ω, элементы которого есть исходы, благоприятствующие B, и т. д. Таким образом, множество всех событий, которые могут наступить в испытании, есть множество всех подмножеств Ω. Само Ω наступает при любом исходе испытания, поэтому Ω – достоверное событие; пустое подмножество пространства Ω возможное событие (оно не наступает ни при каком исходе испытания).

Заметим, что элементарные события выделяются из числа всех событий тем, что каждое из них содержит только один элемент Ω.

Каждому элементарному исходу ω_i ставят в соответствие положительное число p_i – вероятность этого исхода, причем $\sum_i p_i = 1$. По определению, вероятность $P(A)$ события A равна сумме вероятностей A элементарных исходов, благоприятствующих А. Отсюда легко получить, что вероятность события достоверного равна единице, невозможного – нулю, произвольного – заключена между нулем и единицей.

Рассмотрим важный частный случай, когда все исходы равновозможные. Число исходов равно n, сумма вероятностей всех исходов равна единице; следовательно, вероятность каждого исхода равна $1/n$. Пусть событию A благоприятствует m исходов. Вероятность события A равна сумме вероятностей исходов, благоприятствующих A:

$$P(A) = \frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}$$

Учитывая, что число слагаемых равно m, имеем

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

Получено классическое определение вероятности.

Построение логически полноценной теории вероятностей основано на аксиоматическом определении случайного события и его вероятности. В системе аксиом, предложенной А.Н. Колмогоровым[1], неопределяемыми

---

[1] К о л м о г о р о в А . Н . Основные понятия теории вероятностей. М., «Наука», 1974.

понятиями являются элементарное событие и вероятность. Приведем аксиомы, определяющие вероятность:

Каждому событию A поставлено в соответствие неотрицательное действительное число *P(A)*. Это число называется вероятностью события A.

Вероятность достоверного события равна единице:

$$P(\Omega) = 1$$

Вероятность наступления хотя бы одного из попарно несовместных событий равна сумме вероятностей этих событий.

Исходя из этих аксиом, свойства вероятностей и зависимости между ними выводят в качестве теорем.

9.2. Статистическое определение вероятности

Классическое определение вероятности предполагает, что число элементарных исходов испытания конечно. На практике же весьма часто встречаются испытания, число возможных исходов которых бесконечно. В таких случаях классическое определение неприменимо. Уже это обстоятельство указывает на ограниченность классического определения. Отмеченный недостаток может быть преодолен, в частности, введением геометрических вероятностей и, конечно, использованием аксиоматической вероятности.

Наиболее слабая сторона классического определения состоит в том, что очень часто невозможно представить результат испытания в виде совокупности элементарных событий. Еще труднее указать основания, позволяющие считать элементарные события равновозможными. Обычно о равновозможности элементарных исходов испытания говорят из соображений симметрии. Так, например, предполагают, что игральная кость имеет форму правильного многогранника (куба) и изготовлена из однородного материала. Однако задачи, в которых можно исходить из соображений симметрии, на практике встречаются весьма редко. По этой причине наряду с классическим определением вероятности используют и другие определения, в частности статистическое определение.

Введем вначале понятие относительной частоты *W(A)* события A как отношения числа опытов, в которых наблюдалось событие A, к общему количеству проведенных испытаний:

$$W(A) = \frac{M}{N}, \tag{1.2}$$

где N – общее число опытов, M – число появлений события A.

Сопоставляя определения вероятности и относительной частоты, заключаем: определение вероятности не требует, чтобы испытания производились в действительности; определение же относительной частоты предполагает, что испытания были произведены фактически. Другими словами, вероятность вычисляют до опыта, а относительную частоту – после опыта.

Пример 1. Отдел технического контроля обнаружил 3 нестандартных детали в партии из 80 случайно отобранных деталей. Относительная частота появления нестандартных деталей

$$W(A) = \frac{3}{80}$$

Пример 2. По цели произвели 24 выстрела, причем было зарегистрировано 19 попаданий. Относительная частота поражения цели

$$W(A) = \frac{19}{24}$$

Пример 3. По данным шведской статистики, относительная частота рождения девочек за 1935 г. по месяцам характеризуется следующими числами (числа расположены в порядке следования месяцев, начиная с января): 0,486; 0,489; 0,490; 0,471; 0,478; 0,482; 0.462; 0,484; 0,485; 0,491; 0,482; 0,473.

Относительная частота колеблется около числа 0,482, которое можно принять за приближенное значение вероятности рождении девочек.

Заметим, что статистические данные различных стран дают примерно то же значение относительной частоты.

Пример 4. Многократно проводились опыты бросания монеты, которые подсчитывали число появления «герба». Результаты нескольких опытов приведены в табл. 1.1.

Здесь относительные частоты незначительно отклоняются от числа 0,5, причем тем меньше, чем больше число испытаний. Например, при 4040 испытаниях отклонение равно 0, 0069, а при 24 000 испытаний – лишь 0, 0005. Приняв во внимание, что вероятность появления «герба» при бросании монеты равна 0,5, мы вновь убеждаемся, что относительная частота колеблется около вероятности.

Таблица 1.1

Число бросаний	Число появления «герба»	Относительная частота
4 040	2 048	0, 5069
12 000	6 019	0, 5016
24 000	12 012	0, 5005

Большое количество экспериментов показало, что если опыты проводятся в одинаковых условиях, то для большого количества испытаний относительная частота изменяется мало, колеблясь около некоторого постоянного числа. Это число можно считать вероятностью рассматриваемого события.

Определение 1. Статистической вероятностью события считают его относительную частоту или число, близкое к ней.

Замечание 1. Из формулы (1.2) следует, что свойства вероятности, доказанные для ее классического определения, справедливы и для статистического определения вероятности.

Замечание 2. Для существования статистической вероятности события A требуется:

1) возможность производить неограниченное число испытаний;

2) устойчивость относительных частот появления A в различных сериях достаточно большого числа опытов.

Замечание 3. Недостатком статистического определения является неоднозначность статистической вероятности.

Пример 5. Если в задаче задается вероятность попадания в мишень для данного стрелка (скажем, $p = 0{,}7$), то эта величина получена в результате изучения статистики большого количества серий выстрелов, в которых этот стрелок попадал в мишень около семидесяти раз из каждой сотни выстрелов.

9.3. Геометрические вероятности

Чтобы преодолеть недостаток классического определения вероятности, состоящий в том, что оно неприменимо к испытаниям с бесконечным числом исходов, вводят геометрические вероятности – вероятности попадания точки в область (отрезок, часть плоскости и т. д.). Пусть отрезок l составляет часть отрезка L. На отрезок L наудачу поставлена точка. Это означает выполнение следующих предположений: поставленная точка может оказаться в любой точке отрезка L, вероятность попадания точки на отрезок l пропорциональна длине этого отрезка и не зависит от его расположения относительно отрезка L. В этих предположениях вероятность попадания точки на отрезок l определяется равенством

$$P = \frac{\text{Длина } l}{\text{Длина } L}$$

<u>Пример 1.</u> На отрезок ОА длины L числовой оси Оx наудачу поставлена точка *B(x)*. Найти вероятность того, что меньший из отрезков ОВ и ВА имеет длину, большую L/3. Предполагается, что вероятность попадания точки на отрезок пропорциональна длине отрезка и не зависит от его расположения на числовой оси.

Решение. Разобьем отрезок ОА точками С и D на 3 равные части. Требование задачи будет выполнено, если точка В(x) попадет на отрезок CD длины L/3. Искомая вероятность

$$P = \frac{L/_{3}}{L} = \frac{1}{3}$$

Пусть плоская фигура g составляет часть плоской фигуры G. На фигуру G наудачу брошена точка. Это означает выполнение следующих предположений: брошенная точка может оказаться в любой точке фигуры G, вероятность попадания брошенной точки на фигуру g пропорциональна площади этой фигуры и не зависит ни от ее расположения относительно G, ни от формы g. В этих предположениях вероятность попадания точки в фигуру g определяется равенством

$$P = \frac{\text{Площадь } g}{\text{Площадь } G}$$

<u>Пример 2.</u> На плоскости начерчены две концентрические окружности, радиусы которых 5 и 10 см соответственно. Найти вероятность того, что точка, брошенная наудачу в большой круг, попадет в кольцо, образованное построенными окружностями. Предполагается, что вероятность попадания точки в плоскую фигуру пропорциональна площади этой фигуры и не зависит от ее расположения относительно большого круга.

Решение. Площадь кольца (фигуры g)

$$S_g = \pi(10^2 - 5^2) = 75\pi$$

Площадь большого круга (фигуры G)

$$S_0 = \pi \cdot 10^2 = 100\pi$$

Искомая вероятность

$$P = \frac{75\pi}{100\pi} = 0{,}75$$

Пример 3. В сигнализатор поступают сигналы от двух устройств, причем поступление каждого из сигналов равновозможно в любой момент промежутка времени длительностью T. Моменты поступления сигналов независимы один от другого. Сигнализатор срабатывает, если разность между моментами поступления сигналов меньше t(t < T). Найти вероятность того, что сигнализатор сработает за время T, если каждое из устройств пошлет по одному сигналу.

Решение. Обозначим моменты поступления сигналов первого и второго устройств соответственно через x и y. В силу условия задачи должны выполняться двойные неравенства: $0 \leq x \leq T$, $0 \leq y \leq T$. Введём в рассмотрение прямоугольную систему координат xOy. В этой системе двойным неравенствам удовлетворяют координаты любой точки квадрата ОTАT (рис. 1). Таким образом, этот квадрат можно рассматривать как фигуру G, координаты точек которой представляют возможные значения моментов поступления сигналов.

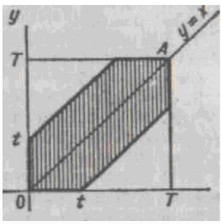


Рис. 1

Сигнализатор срабатывает, если разность между моментами поступления сигналов меньше t, т. е. если $y - x < t$ при $y > x$ и $x - y < t$ при $x > y$, или что то же,

$$y < x + t \quad при \quad y > x, \qquad (*)$$

$$y > x - t \quad при \quad y < x. \qquad (**)$$

Неравенство (*) выполняется для тех точек фигуры G, которые выше прямой $y = x$ и ниже прямой $y = x + t$ неравенство (**) имеет место для точек, расположенных ниже прямой $y = x$ и выше прямой $y = x - t$.

Как видно из рис. 1, все точки, координаты которых удовлетворяют неравенствам (*) и (**), принадлежат заштрихованному прямоугольнику. Таким образом, этот шестиугольник можно рассматривать как фигуру g, координаты точек которой являются благоприятствующими моментами времени x и y.

Искомая вероятность

$$P = \frac{\text{Пл}. g}{\text{Пл}. G} = \frac{t(2T - t)}{T^2}$$

Замечание 1. Приведенные определения являются частными случаями общего определения геометрической вероятности. Если обозначить меру (длину, площадь, объем) области через mes, то вероятность попадания точки, брошенной наудачу (в указанном выше смысле) в область g – часть области G, равна

$$P = \frac{mes\ g}{mes\ G}$$

Замечание 2. В случае классического определения вероятность достоверного (невозможного) события равна единице (нулю); справедливы и обратные утверждения (например, если вероятность события равна нулю, то событие невозможно). В случае геометрического определения вероятности обратные утверждения не имеют места. Например, вероятность попадания брошенной точки в одну определённую точку области *G* равна нулю, однако это событие может произойти, и, следовательно, не является невозможным.

<u>Пример 4.</u> Найти вероятность того, что точка, наудачу брошенная в круг, не попадет в правильный шестиугольник, вписанный в него.

Решение. Пусть радиус круга равен *R*, тогда сторона шестиугольника тоже равна *R*. При этом площадь круга $S = \pi R^2$, а площадь шестиугольника $s = \frac{3\sqrt{3}}{2}R^2$. Следовательно,

$$p = \frac{S - s}{S} = \frac{\pi R^2 - \frac{3\sqrt{3}}{2}R^2}{\pi R^2} = \frac{\pi - 3\sqrt{3}}{2\pi} \approx 0{,}174$$

<u>Пример 5.</u> На отрезок *AB* случайным образом брошены три точки: *C, D* и *M*. Найти вероятность того, что из отрезков *AC, AD* и *AM* можно построить треугольник.

Решение. Обозначим длины отрезков *AC, AD* и *AM* через *x, y* и *z* и рассмотрим в качестве возможных исходов множество точек трехмерного пространства с координатами (*x, y, z*). Если принять длину отрезка равной 1, то это множество возможных исходов представляет собой куб с ребром, равным 1. Тогда множество благоприятных исходов состоит из точек, для координат которых выполнены неравенства треугольника: $x + y > z$, $x + z > y$, $y + z > x$. Это часть куба, отрезанная от него плоскостями $x + y = z$, $x + z = y$, $y + z = x$.

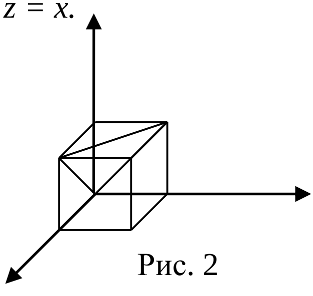


Рис. 2

(одна из них, плоскость $x + y = z$, проведена на рис. 2). Каждая такая плоскость отделяет от куба пирамиду, объем которой равен $\frac{1}{3}\cdot\frac{1}{2}\cdot 1 = \frac{1}{6}$. Следовательно, объем оставшейся части $v = 1 - 3\cdot\frac{1}{6} = \frac{1}{2}$. Тогда $p = \frac{v}{V} = \frac{1}{2} : 1 = \frac{1}{2}$.

9.4. Основные формулы комбинаторики

При вычислении вероятностей часто приходится использовать некоторые формулы комбинаторики – науки, изучающей комбинации, которые можно составить по определенным правилам из элементов некоторого конечного множества. Определим основные такие комбинации.

Определение 1. Перестановки – это комбинации, составленные из всех n элементов данного множества и отличающиеся только порядком их расположения. Число всех возможных перестановок

$$P_n = n! \tag{1.3}$$

Пример 1. Сколько различных списков (отличающихся порядком фамилий) можно составить из 7 различных фамилий?

Решение. $P_7 = 7! = 2\cdot3\cdot4\cdot5\cdot6\cdot7 = 5040$.

Определение 2. Размещения – комбинации из m элементов множества, содержащего n различных элементов, отличающиеся либо составом элементов, либо их порядком. Число всех возможных размещений

$$A_n^m = n(n-1)(n-2)\ldots(n-m+1 \tag{1.4}$$

Пример 2. Сколько возможно различных вариантов пьедестала почета (первое, второе, третье места), если в соревнованиях принимают участие 10 человек?

Решение. $A_n^m = 10\cdot 9\cdot 8 = 720.$

Определение 3. Сочетания – неупорядоченные наборы из m элементов множества, содержащего n

различных элементов (то есть наборы, отличающиеся только составом элементов). Число сочетаний

$$C_n^m = \frac{n!}{m!(n-m)!} \qquad (1.5)$$

Пример 3. В отборочных соревнованиях принимают участие 10 человек, из которых в финал выходят трое. Сколько может быть различных троек финалистов?

Решение. В отличие от предыдущего примера, здесь не важен порядок финалистов, следовательно, ищем число сочетаний из 10 по 3:

$$C_{10}^3 = \frac{10!}{3! \cdot 7!} = \frac{8 \cdot 9 \cdot 10}{6} = 120$$

При решении задач комбинаторики используют следующие правила:

Правило суммы. Если некоторый объект A может быть выбран из совокупности объектов m способами, а объект B может быть выбран n способами, то выбрать либо A, либо B можно $m + n$ способами.

Правило произведения. Если объект A можно выбрать из совокупности объектов m способами и после каждого такого выбора объект B можно выбрать n способами, то пара объектов (A, B) в указанном порядке может быть выбрана mn способами.

Теоремы сложения и умножения вероятностей и их следствия

Понятие условной вероятности играет важнейшую роль в современной теории вероятностей. Условная вероятность позволяет учитывать дополнительную информацию при определении вероятности события. В ряде случаев при помощи условной вероятности можно существенно упростить вычисление вероятности.

§10. Теоремы сложения и умножения вероятностей

10.1. Теоремы сложения вероятностей

Теорема 1 (теорема сложения). Вероятность p ($A + B$) суммы событий A и B равна

$$P(A+B) = p(A) + p(B) - p(AB) \qquad (1.1)$$

Доказательство. Докажем теорему сложения для схемы случаев. Пусть n – число возможных исходов опыта, m_A – число исходов, благоприятных событию A, m_B – число исходов, благоприятных событию B, а m_{AB} – число исходов опыта, при которых происходят оба события (то есть исходов, благоприятных произведению AB). Тогда число исходов, при которых имеет место событие $A + B$, равно $m_A + m_B - m_{AB}$ (так как в сумме ($m_A + m_B$) m_{AB} учтено дважды: как исходы, благоприятные A, и исходы, благоприятные B). Следовательно, вероятность суммы можно определить по формуле классического определения вероятности:

$$P(A+B) = \frac{m_A + m_B - m_{AB}}{n} = \frac{m_A}{n} + \frac{m_B}{n} - \frac{m_{AB}}{n}$$
$$= p(A) + p(B) - p(AB)$$

что и требовалось доказать.

Следствие 1. Теорему 1 можно распространить на случай суммы любого числа событий. Например, для суммы трех событий A, B и C:

$$P(A + B + C) = p(A) + p(B) + p(C) - p(AB) - p(AC) - p(BC) + p(ABC) \qquad (1.2)$$

и т. д.

Следствие 2. Если события A и B несовместны, то $m_{AB} = 0$, и, следовательно, вероятность суммы несовместных событий равна сумме их вероятностей:

$$P(A + B) = p(A) + p(B). \qquad (1.3)$$

Определение 1. **Противоположными событиями** называют два несовместных события, образующих полную группу. Если одно из них назвать A, то второе принято обозначать $\overline{A}$.

Замечание 1. Таким образом, $\overline{A}$ заключается в том, что событие A не произошло.

Теорема 2. Сумма вероятностей противоположных событий равна 1:

$$p(A) + p(\overline{A}) = 1 \quad (1.4)$$

Доказательство. Так как A и $\overline{A}$ образуют полную группу, то одно из них обязательно произойдет в результате опыта, то есть событие $A + \overline{A}$ является достоверным. Следовательно, $P(A + \overline{A}) = 1$. Но, так как A и $\overline{A}$ несовместны, из (1.3) следует, что $P(A + \overline{A}) = p(A) + p(\overline{A})$. Значит, $p(A) + p(\overline{A}) = 1$, что и требовалось доказать.

Замечание 2. В ряде задач проще искать не вероятность заданного события, а вероятность события, противоположного ему, а затем найти требуемую вероятность по формуле (1.4).

<u>Пример 1.</u> Из урны, содержащей 2 белых и 6 черных шаров, случайным образом извлекаются 5 шаров. Найти вероятность того, что вынуты шары разных цветов.

Решение. Событие $\overline{A}$, противоположное заданному, заключается в том, что из урны вынуто 5 шаров одного цвета, а так как белых шаров в ней всего два, то этот цвет может быть только черным. Множество возможных исходов опыта найдем по определению сочетания:

$$n = C_8^5 = \frac{8!}{5! \cdot 3!} = \frac{6 \cdot 7 \cdot 8}{6} = 56$$

а множество исходов, благоприятных событию $\overline{A}$ – это число возможных наборов по 5 шаров только из шести черных:

$$m_{\overline{A}} = C_6^5 = 6$$

Тогда $p(\overline{A}) = \frac{6}{56} = \frac{3}{28}$, а $p(A) = 1 - \frac{3}{28} = \frac{25}{28}$.

Определение 2. Назовем **условной вероятностью *p*(*B**A*) события *B*** вероятность события B при условии, что событие A произошло.

Замечание 3. Понятие условной вероятности используется в основном в случаях, когда осуществление события A изменяет вероятность события B.

Примеры:

1) пусть событие A – извлечение из колоды в 32 карты туза, а событие B – то, что и вторая вынутая из колоды карта окажется тузом. Тогда, если после первого раза карта была возвращена в колоду, то вероятность вынуть вторично туз не меняется: $p(B) = p(A) = \frac{4}{32} = \frac{1}{8} = 0{,}125$. Если же первая карта в колоду не возвращается, то осуществление события A приводит к тому, что в колоде осталась 31 карта, из которых только 3 туза. Поэтому $p(B\backslash A) = \frac{3}{31} \approx 0{,}097$.

2) если событие A – попадание в самолет противника при первом выстреле из орудия, а B – при втором, то первое попадание уменьшает маневренность самолета, поэтому $p(B\backslash A)$ увеличится по сравнению с $p(A)$.

10.2. Теоремы умножения вероятностей

Теорема 3 (теорема умножения). Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого при условии, что первое событие произошло:

$$p\,(AB) = p\,(A) \cdot p\,(B\backslash A). \qquad (1.5)$$

Доказательство. Воспользуемся обозначениями теоремы 1. Тогда для вычисления $p(B\backslash A)$ множеством возможных исходов нужно считать m_A (так как A произошло), а множеством благоприятных исходов – те, при которых произошли и A, и B (m_{AB}). Следовательно,

$$p(B\backslash A) = \frac{m_{AB}}{m_A} = \frac{m_{AB}}{n} \cdot \frac{n}{m_A} = p(AB) : p(A),$$ откуда следует утверждение теоремы.

Пример 2. Для поражения цели необходимо попасть в нее дважды. Вероятность первого попадания равна 0,2, затем она не меняется при промахах, но после первого попадания увеличивается вдвое. Найти вероятность того, что цель будет поражена первыми двумя выстрелами.

Решение. Пусть событие A – попадание при первом выстреле, а событие B – попадание при втором. Тогда $p\ (A) = 0{,}2$, $p\ (B \backslash A) = 0{,}4$, $p\ (AB) = 0{,}2 \cdot 0{,}4 = 0{,}08$.

Следствие. Если подобным образом вычислить вероятность события BA, совпадающего с событием AB, то получим, что $p\ (BA) = p\ (B) \cdot p\ (A \backslash B)$. Следовательно,

$$p\ (A) \cdot p\ (B \backslash A) = p\ (B) \cdot p\ (A \backslash B). \qquad (1.6)$$

Определение 3. Событие B называется **независимым** от события A, если появление события A не изменяет вероятности B, то есть $p\ (B \backslash A) = p\ (B)$.

Замечание 4. Если событие B не зависит от A, то и A не зависит от B. Действительно, из (1.6) следует при этом, что $p\ (A) \cdot p\ (B) = p\ (B) \cdot p\ (A/B)$, откуда $p\ (A \backslash B) = p\ (A)$. Значит, **свойство независимости событий взаимно**.

Теорема умножения для независимых событий имеет вид:

$$p\ (AB) = p\ (A) \cdot p\ (B)\ , \qquad (1.7)$$

то есть вероятность произведения независимых событий равна произведению их вероятностей.

При решении задач теоремы сложения и умножения обычно применяются вместе.

Пример 3. Два стрелка делают по одному выстрелу по мишени. Вероятности их попадания при одном выстреле равны соответственно 0,6 и 0,7. Найти вероятности следующих событий:

A – хотя бы одно попадание при двух выстрелах;

B – ровно одно попадание при двух выстрелах;

C – два попадания;

D – ни одного попадания.

Решение. Пусть событие H_1 – попадание первого стрелка, H_2 – попадание второго. Тогда $A = H_1 + H_2$, $B = H_1 \cdot \overline{H}_2 + \overline{H}_1 \cdot H_2$, $C = H_1 \cdot H_2$, $D = \overline{H}_1 \cdot \overline{H}_2$. События H_1 и H_2 совместны и независимы, поэтому теорема сложения применяется в общем виде, а теорема умножения – в виде (1.7). Следовательно, $p(C) = 0{,}6{\cdot}0{,}7 = 0{,}42$, $p(A) = 0{,}6 + 0{,}7 - 0{,}42 = 0{,}88$, $p(B) = 0{,}6{\cdot}0{,}3 + 0{,}7{\cdot}0{,}4 = 0{,}46$ (так как события $H_1 \cdot \overline{H}_2$ и $\overline{H}_1 \cdot H_2$ несовместны), $p(D) = 0{,}4{\cdot}0{,}3 = 0{,}12$. Заметим, что события A и D являются противоположными, поэтому $p(A) = 1 - p(D)$.

Теорема 4. Вероятность появления хотя бы одного из попарно независимых событий $A_1, A_2, \ldots, A_n$ равна

$$p\,(A) = 1 - q_1 q_2 \ldots q_n\,, \tag{1.8}$$

где q_i – вероятность события $\overline{A_i}$, противоположного событию A_i .

Доказательство. Если событие A заключается в появлении хотя бы одного события из $A_1, A_2, \ldots, A_n$, то события A и $\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_n$ противоположны, поэтому по теореме 2.2 сумма их вероятностей равна 1. Кроме того, поскольку $A_1, A_2, \ldots, A_n$ независимы, то независимы и $\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_n$, следовательно, $p(\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_n) = p(\overline{A}_1)p(\overline{A}_2) \ldots p(\overline{A}_n) = q_1 q_2 \ldots q_n$. Отсюда следует справедливость формулы (1.8).

Пример 4. Сколько нужно произвести бросков монеты, чтобы с вероятностью не менее 0,9 выпал хотя бы один герб?

Решение. Вероятность выпадения герба при одном броске равна вероятности противоположного события (выпадения цифры) и равна 0,5. Тогда вероятность выпадения хотя бы одного герба при n выстрелах равна 1- $(0{,}5)^n$. Тогда из решения неравенства 1- $(0{,}5)^n > 0{,}9$ следует, что $n > \log_2 10 \geq 4$.

10.3. Следствия теорем сложения и умножения

Теорема 1. Вероятность появления хотя бы одного из двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного появления:

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

Доказательство. Поскольку события A и B по условию совместны, то событие $A + B$ наступит, если насупит одно из следующих трех несовместных событий: $A\bar{B}, \bar{A}B$ или AB. По теореме сложения вероятностей несовместных событий

$$P\,(A + B) = P\,(A\bar{B}) + P\,(\bar{A}B) + P\,(AB). \qquad (*)$$

Событие A произойдет, если наступит одно из двух несовместных событий: $A\bar{B}$ или $\bar{A}B$. По теореме сложения вероятностей несовместных событий имеем

$$P\,(A) = P\,(A\bar{B}) + P\,(AB).$$

Отсюда

$$P\,(A\bar{B}) = P\,(A) - P\,(AB). \qquad (**)$$

Аналогично имеем

$$P\,(B) = P\,(\bar{A}B) + P\,(AB).$$

Отсюда

$$P\,(\bar{A}B) = P\,(B) - P\,(AB). \qquad (***)$$

Подставив (**) и (***) в (*), окончательно получим

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

Определение 1. Пусть событие A может произойти только совместно с одним из событий $H_1, H_2, \ldots, H_n$, образующих полную группу несовместных событий. Тогда события $H_1, H_2, \ldots, H_n$ называются **гипотезами.**

Теорема 2. Вероятность события A, наступающего совместно с гипотезами $H_1, H_2, \ldots, H_n$, равна:

$$p(A)=\sum_{i=1}^{n} p(H_i)\,p(A \setminus H_i), \qquad (1.9)$$

где $p(H_i)$ – вероятность i-й гипотезы, а $p(A \backslash H_i)$ – вероятность события A при условии реализации этой гипотезы. Формула (1.9) носит название **формулы полной вероятности**.

Доказательство. Можно считать событие A суммой попарно несовместных событий $AH_1, AH_2, \ldots, AH_n$. Тогда из теорем сложения и умножения следует, что

$$p(A) = p(AH_1 + AH_2 + \ldots + AH_n) = p(AH_1) + p(AH_2) + \ldots + p(AH_n) = \sum_{i=1}^{n} p(H_i)p(A \backslash H_i),$$

что и требовалось доказать.

<u>Пример 1.</u> Имеются три одинаковые урны с шарами. В первой из них 3 белых и 4 черных шара, во второй – 2 белых и 5 черных, в третьей – 10 черных шаров. Из случайно выбранной урны наудачу вынут шар. Найти вероятность того, что он белый.

Решение. Будем считать гипотезами H_1, H_2 и H_3 выбор урны с соответствующим номером. Так как по условию задачи все гипотезы равновозможны, то $p(H_1) = p(H_2) = p(H_3) = \frac{1}{3}$. Найдем условную вероятность A при реализации каждой гипотезы: $p(A \backslash H_1) = \frac{3}{7}$, $p(A \backslash H_2) = \frac{2}{7}$, $p(A \backslash H_3) = 0$. Тогда $p(A) = \frac{1}{3} \cdot \frac{3}{7} + \frac{1}{3} \cdot \frac{2}{7} + \frac{1}{3} \cdot 0 = \frac{5}{21} \approx 0{,}238$.

Пусть известен результат опыта, а именно то, что произошло событие A. Этот факт может изменить априорные (то есть известные до опыта) вероятности гипотез. Например, в предыдущем примере извлечение из урны белого шара говорит о том, что этой урной не могла быть третья, в которой нет белых шаров, то есть $p\ (H_3/A) = 0$. Для переоценки вероятностей гипотез при известном результате опыта используется **формула Байеса**:

$$p(H_i \backslash A) = \frac{p(H_i)p(A \backslash H_i)}{p(A)} \qquad (1.10)$$

Действительно, из следствия теоремы умножения получим, что $p(A)p(H_i\backslash A) = p(H_i)p(A\backslash H_i)$, откуда следует справедливость формулы (1.10).

Пример 2. После двух выстрелов двух стрелков, вероятности попаданий которых равны 0,6 и 0,7, в мишени оказалась одна пробоина. Найти вероятность того, что попал первый стрелок.

Решение. Пусть событие A – одно попадание при двух выстрелах, а гипотезы: H_1 – первый попал, а второй промахнулся, H_2 – первый промахнулся, а второй попал, H_3 – оба попали, H_4 – оба промахнулись. Вероятности гипотез: $p(H_1) = 0{,}6{\cdot}0{,}3 = 0{,}18$, $p(H_2) = 0{,}4{\cdot}0{,}7 = 0{,}28$, $p(H_3) = 0{,}6{\cdot}0{,}7 = 0{,}42$, $p(H_4) = 0{,}4{\cdot}0{,}3 = 0{,}12$. Тогда $p(A\backslash H_1) = p(A\backslash H_2) = 1$, $p(A\backslash H_3) = p(A\backslash H_4) = 0$. Следовательно, полная вероятность $p(A) = 0{,}18{\cdot}1 + 0{,}28{\cdot}1 + 0{,}42{\cdot}0 + 0{,}12{\cdot}0 = 0{,}46$. Применяя формулу Байеса, получим:

$$p(H_i\backslash A) = \frac{0{,}18\cdot 1}{0{,}46} = \frac{9}{23} \approx 0{,}391.$$

10.4. Повторение испытаний

Рассмотрим серию из n испытаний, в каждом из которых событие A появляется с одной и той же вероятностью p, причем результат каждого испытания не зависит от результатов остальных. Подобная постановка задачи называется **схемой повторения испытаний**. Найдем вероятность того, что в такой серии событие A произойдет ровно $к$ раз (неважно, в какой последовательности). Интересующее нас событие представляет собой сумму равновероятных несовместных событий, заключающихся в том, что A произошло в некоторых $к$ испытаниях и не произошло в остальных $n - к$ испытаниях. Число таких событий равно числу сочетаний из n по $к$, то есть C_n^k, а вероятность каждого из них: $p^k q^{n-k}$, где $q = 1 - p$ – вероятность того, что в данном опыте A не произошло. Применяя теорему сложения для несовместных событий, получим **формулу Бернулли**:

$$p_n(k) = C_n^k \cdot p^k \cdot q^{n-k} \qquad (1.11)$$

<u>Пример 1.</u> Для получения приза нужно собрать 5 изделий с особым знаком на этикетке. Найти вероятность того, что придется купить 10 изделий, если этикетки с этим знаком имеют 5% изделий.

Решение. Из постановки задачи следует, что последнее купленное изделие имеет особый знак. Следовательно, из предыдущих девяти эти знаки имели 4 изделия. Найдем вероятность этого по формуле Бернулли: $p_9(4) = C_9^4 \cdot (0{,}05)^4 \cdot (0{,}95)^5 = 0{,}0006092.$ Тогда p = 0,0006092·0,05 = 0,0000304.

Формула Бернулли требует громоздких расчетов при большом количестве испытаний. Можно получить более удобную для расчетов приближенную формулу, если при большом числе испытаний вероятность появления A в одном опыте мала, а произведение $np = \lambda$ сохраняет постоянное значение для разных серий опытов (то есть среднее число появлений события A в разных сериях испытаний остается неизменным). Применим формулу Бернулли:

$$p_n(k) = \frac{n(n-1)(n-2)...(n-k+1)}{k!} p^k (1-p)^{n-k} = \frac{n(n-1)...(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}.$$

Найдем предел полученного выражения при $n \to \infty$:

$$p_n(k) \approx \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1 \cdot \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-k}\right) = \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1.$$

Таким образом, **формула Пуассона**

$$p_n(k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad (1.12)$$

позволяет найти вероятность k появлений события A для массовых (n велико) и редких (p мало) событий.

При выводе формулы Бернулли мы предполагали, что вероятность появления события в каждом испытании постоянна. Легко видеть, что пользоваться формулой Бернулли при больших значениях n достаточно трудно, так как формула требует выполнения действий громадными числами. Например, если $n = 50$, $k = 30$, $p = 0{,}1$, то для отыскания вероятности $P_{50}(30)$ надо вычислить выражение $P_{50}(30) = \frac{50!}{30! \cdot 20!} \cdot (0{,}1)^{30} \cdot (0{,}9)^{20}$, где $50! = 30414093 \cdot 10^{57}$, $30! = 26525286 \cdot 10^{25}$, $20! = 24329020 \cdot 10^{11}$. Правда, можно несколько упростить вычисления, пользуясь специальными таблицами логарифмов факториалов. Однако и этот путь остается громоздким и к тому же имеет существенный недостаток: таблицы содержат приближенные значения логарифмов, поэтому в процессе вычислений накапливаются погрешности; в итоге окончательный результат может значительно отличаться от истинного.

Естественно возникает вопрос: нельзя ли вычислить интересующую нас вероятность, не прибегая к формуле Бернулли? Оказывается, можно. Локальная теорема Лапласа и дает асимптотическую[2] формулу, которая позволяет приближенно найти вероятность появления события ровно k раз в n испытаниях, если число испытаний достаточно велико.

Заметим, что для частного случая, а именно для p = 1/2, асимптотическая формула была найдена в 1730 г. Муавром; в 1783 г. Лаплас обобщил формулу Муавра произвольного p, отличного от 0 и 1. Поэтому теорему, о которой здесь идет речь, иногда называют теоремой Муавра–Лапласа.

[2] Функцию $\varphi(x)$ называют асимптотическим приближением функции $f(x)$, если $\lim\limits_{x \to \infty} \frac{f(x)}{\varphi(x)} = 1$.

Доказательство локальной теоремы Лапласа довольно сложно, поэтому мы приведем лишь формулировку теоремы и примеры, иллюстрирующие ее использование.

Теорема 1 (Локальная теорема Лапласа). Если вероятность p появления события A в каждом испытании постоянна и отлична от нуля и единицы, то вероятность $P_n(k)$ того, что событие A появится в n испытаниях ровно k раз, приближенно равна (тем точнее, чем больше n) значению функции

$$y = \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \frac{1}{\sqrt{npq}} \cdot \varphi(x)$$

при $x = \frac{k-np}{\sqrt{npq}}$.

Имеются таблицы, в которых помещены значения функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, соответствующие положительным значениям аргумента x. Для отрицательных значений аргумента пользуются теми же таблицами, так как функция $\varphi(x)$ четна, т. е. $\varphi(-x) = \varphi(x)$.

Итак, вероятность того, что событие A появится в n независимых испытаниях ровно k раз, приближенно равна

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

где $x = \frac{k-np}{\sqrt{npq}}$..

<u>Пример 2.</u> Найти вероятность того, что событие A наступит ровно 80 раз в 400 испытаниях, если вероятность появления этого события в каждом испытании равна 0,2.

Решение. По условию, $n = 400$; $k = 80$; $p = 0,2$; $q = 0,8$. Воспользуемся асимптотической формулой Лапласа:

$$P_{400}(80) \approx \frac{1}{\sqrt{400 \cdot 0,2 \cdot 0,8}} \cdot \varphi(x) = \frac{1}{8} \cdot \varphi(x).$$

Вычислим определяемое данными задачи значение x:

$$x = (k - np)/\sqrt{npq} = (80 - 400 \cdot 0{,}2)/8 = 0.$$

По таблице находим $\varphi(x) = 0.3989$.

Искомая вероятность

$$P_{400}(80) = (1/8) \cdot 0{,}3989 = 0{,}046986.$$

Формула Бернулли приводит примерно к такому же результату (выкладки ввиду их громоздкости опущены):

$$P_{400}(80) = 0{,}0498.$$

Пример 3. Вероятность поражения мишени стрелком при одном выстреле $p = 0{,}75$. Найти вероятность того, что при 10 выстрелах стрелок поразит мишень 8 раз.

Решение. По условию, $n = 10$; $k = 8$; $p = 0{,}75$; $q = 0{,}25$. Воспользуемся асимптотической формулой Лапласа:

$$P_{10}(8) \approx \frac{1}{\sqrt{10 \cdot 0{,}75 \cdot 0{,}25}} \cdot \varphi(x) = 0{,}7301 \cdot \varphi(x).$$

Вычислим определяемое данными задачи значение x:

$$x = \frac{k - np}{\sqrt{npq}} = \frac{8 - 10 \cdot 0{,}75}{\sqrt{10 \cdot 0{,}75 \cdot 0{,}25}} \approx 0{,}36.$$

По таблице находим $\varphi(0{,}36) = 0.3739$.

Искомая вероятность

$$P_{10}(8) = 0{,}7301 \cdot 0{,}3739 = 0{,}273.$$

Формула Бернулли приводит к иному результату, а именно $P_{10}(8) = 0{,}282$. Столь значительное расхождение ответов объясняется тем, что в настоящем примере n имеет малое значение (формула Лапласа дает достаточно хорошие приближения лишь при достаточно больших значениях n).

Вновь предположим, что производится n испытаний, в каждом из которых вероятность появления события A постоянна и равна p ($0 < p < 1$). Как вычислить вероятность $P_n(k_1, k_2)$ того, что событие A появится в n испытаниях не менее k_1 и не более k_2 раз (для краткости будем говорить «от k_1 до k_2 раз»)? На этот вопрос отвечает интегральная теорема Лапласа, которую мы приводим ниже, опустив доказательство.

Теорема 2. Если вероятность p наступления события A в каждом испытании постоянна и отлична от нуля и единицы, то вероятность $P_n(k_1, k_2)$ того, что событие A появится в n испытаниях от k_1 до k_2 раз, приближенно равна определенному интегралу

$$P_n(k_1,k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-x^2/2} dz \quad , \tag{1.13}$$

где $x' = (k_1 - np)/\sqrt{npq}$ и $x'' = (k_2 - np)/\sqrt{npq}$.

При решении задач, требующих применения интегральной теоремы Лапласа, пользуются специальными таблицами, так как неопределенный интеграл $\int e^{-x^2/2}$ не выражается через элементарные функции. Таблица для интеграла $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int_0^x e^{-x^2/2} dz$ обычно приводится в приложении к учебному пособию. В таблице даны значения функции $Ф(x)$ для положительных значений x и для $x = 0$; для $x < 0$ пользуются той же таблицей [функция $Ф(x)$ нечетна, т. е. $Ф(-x) = -Ф(x)$]. В таблице приведены значения интеграла лишь до $x = 5$, так как для $x > 5$ можно принять $Ф(x) = 0{,}5$. Функцию $Ф(x)$ часто называют функцией Лапласа.

Для того чтобы можно было пользоваться таблицей функции Лапласа, преобразуем соотношение (1.1) так:

$$P_n(k_1,k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x'}^{0} e^{-x^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_{0}^{x''} e^{-x^2/2} dz = \frac{1}{\sqrt{2\pi}} \int_{0}^{x''} e^{-x^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_{0}^{x'} e^{-x^2/2} dz = \Phi(x'') - \Phi(x') .$$

Итак, вероятность того, что событие A появится в n независимых испытаниях от k_1 до k_2 раз,

$$P_n(k_1,k_2) \approx \Phi(x'') - \Phi(x'),$$

где $x' = (k_1 - np)/\sqrt{npq}$ и $x'' = (k_2 - np)/\sqrt{npq}$.

Приведем примеры, иллюстрирующие применение интегральной теоремы Лапласа.

<u>Пример 4.</u> Вероятность того, что деталь не прошла проверку ОТК, равна p = 0,2. Найти вероятность того, что среди 400 случайно отобранных деталей окажется непроверенных от 70 до 100 деталей.

Решение. По условию, p = 0,2; q = 0,8; n = 400; k_1 = 70; k_2 = 100. Воспользуемся интегральной теоремой Лапласа:

$$P_{400}(70{,}100) \approx \Phi(x'') - \Phi(x').$$

Вычислим нижний и верхний пределы интегрирования:

$$x' = \frac{k_1 - np}{\sqrt{npq}} = \frac{70 - 400 \cdot 0{,}2}{\sqrt{400 \cdot 0{,}2 \cdot 0{,}8}} = -1{,}25;$$

$$x'' = \frac{k_2 - np}{\sqrt{npq}} = \frac{100 - 400 \cdot 0{,}2}{\sqrt{400 \cdot 0{,}2 \cdot 0{,}8}} = 2{,}5.$$

Таким образом, имеем

$$P_{400}(70{,}100) = \Phi(2{,}5) - \Phi(-1{,}5) = \Phi(2{,}5) + \Phi(1{,}5).$$

По таблице приложения 2 находим:

$$\Phi(2{,}5) = 0{,}4938;\ \Phi(1{,}25) = 0{,}3944.$$

Искомая вероятность

$$P_{400}(70{,}100) = 0{,}4938 + 0{,}3944 = 0{,}8882.$$

Замечание. Обозначим через m число появлений события A при n независимых испытаниях, в каждом из которых вероятность наступления события A постоянна и равна p. Если число m изменяется от k_1 до k_2, то дробь $(m - np)/\sqrt{npq}$ изменяется от $x' = (k_1 - np)/\sqrt{npq}$ до

$x''=(k_2-np)/\sqrt{npq}$. Следовательно, интегральную теорему Лапласа можно записать и так:

$$P(x'\leq \frac{m-np}{\sqrt{npq}}\leq x'')\approx \frac{1}{\sqrt{2\pi}}\int_{x'}^{x''} e^{-x^2/2}dz.$$

Эта форма записи используется ниже.

Вновь будем считать, что производится *n* независимых испытании, в каждом из которых вероятность появления события *A* постоянна и равна *p(0<p<1)*. Поставим перед собой задачу найти вероятность того, что отклонение относительной частоты *m/n* от постоянной вероятности *p* по абсолютной величине не превышает заданного числа $\varepsilon>0$. Другими словами, найдем вероятность осуществления неравенства

$$|m/n-p|\leq\varepsilon. \qquad (1.14)$$

Эту вероятность будем обозначать так: $P(|m/n-p|\leq\varepsilon)$. Заменим неравенство (1.14) ему равносильными:

$$-\varepsilon\leq m/n-p\leq\varepsilon \text{ или } -\varepsilon\leq(m/n-p)\leq\varepsilon$$

Умножая эти неравенства на положительный множитель $\sqrt{n/(pq)}$, получим неравенства, равносильные исходному:

$$-\varepsilon\sqrt{n/(pq)}\leq(m/n-p)/\sqrt{npq}\leq\varepsilon\sqrt{n/(pq)}.$$

Воспользуемся интегральной теоремой Лапласа в форме, указанной в замечании к ней. Положив $x'=-\varepsilon\sqrt{n/(pq)}$ и $x''=\varepsilon\sqrt{n/(pq)}$, имеем

$$P(-\varepsilon\sqrt{n/(pq)} \le (m/n-p)/\sqrt{npq} \le \varepsilon\sqrt{n/(pq)}) \approx \frac{1}{\sqrt{2\pi}}\int_{-\varepsilon\sqrt{n/(pq)}}^{\varepsilon\sqrt{n/(pq)}} e^{-x^2/2}dz =$$

$$= \frac{2}{\sqrt{2\pi}}\int_{0}^{\varepsilon\sqrt{n/(pq)}} e^{-x^2/2}dz = 2\Phi(\varepsilon\sqrt{n/(pq)})$$

.

Наконец, заменив неравенства, заключенные в скобках, равносильным им исходным неравенством, окончательно получим

$$P(|\,m/n-p\,| \le \varepsilon) \approx 2\Phi(\varepsilon\sqrt{n/(pq)}).$$

Итак, вероятность осуществления неравенства $|\,m/n-p\,| \le \varepsilon$ приближенно равна значению удвоенной функции Лапласа 2*Ф(x)* при $x = \varepsilon\sqrt{n/(pq)}$

Пример 5. Вероятность того, что деталь нестандартна, *p* = 0,1 Найти вероятность того, что среди случайно отобранных 400 деталей относительная частота появления нестандартных деталей отклонится от вероятности *p* = 0,1 по абсолютной величине не более чем на 0,03.

Решение. По условию, *n* = 400; *p* = 0,1; *q* = 0,9; ε =0,03. Требуется найти вероятность $P(|\,m/400-0{,}1\,| \le 0{,}03)$. Пользуясь формулой $P(|\,m/n-p\,| \le \varepsilon) \approx 2\Phi(\varepsilon\sqrt{n/(pq)})$, имеем

$$P(|\,m/400-0{,}1\,| \le 0{,}03) \approx 2\Phi(0{,}03\sqrt{400/(0{,}1\cdot 0{,}9)}) = 2\Phi(2).$$

По таблице находим Ф(2) = 0,4772. Следовательно, 2Ф(2) = 0,9544.

Итак, искомая вероятность приближенно равна 0,9644.

Смысл полученного результата таков: если взять достаточно большое число проб по 400 деталей в каждой, то примерно в 95,44% этих проб отклонение относительной частоты от постоянной вероятности p = 0,1 по абсолютной величине не превысит 0,03.

Пример 6. Вероятность того, что деталь не стандартна, p = 0,1 Найти, сколько деталей надо отобрать, чтобы с вероятностью, равной 0,9544, можно было утверждать, что относительная частота появления нестандартных деталей (среди отобранных) отклонится от постоянной вероятности p по абсолютной величине не более чем на 0,03.

Решение. По условию, p = 0,1; q = 0,9; ε = 0,03; $P(|m/400-0{,}1|\le 0{,}03)=0{,}9544$. Требуется найти n.

Воспользуемся формулой

$$P(|m/n-p|\le\varepsilon)\approx 2\Phi(\varepsilon\sqrt{n/(pq)}).$$

В силу условия

$$2\Phi(0{,}03\sqrt{n/(0{,}1\cdot 0{,}9)}=2\Phi(0{,}1\sqrt{n})=0{,}9544.$$

Следовательно, $\Phi(0{,}1\sqrt{n})=0{,}4772$.

По таблице находим $\Phi(2)$ = 0,4772.

Для отыскания числа n получаем уравнение $0{,}1\sqrt{n}=2$. Отсюда искомое число деталей n = 400.

Смысл полученного результата таков: если взять достаточно большое число проб по 400 деталей, то в 96,44% этих проб относительная частота появления нестандартных деталей будет отличаться от постоянной вероятности p = 0,1 по абсолютной величине не более чем на 0,03, т. е. относительная частота заключена в границах от 0,07 (0,1 - 0,03 = 0,07) до 0,13 (0,1 + 0,03 = 0,13). Другими словами, число нестандартных деталей в 96,44% проб будет заключено между 28 (7% от 400) и 62 (13% от 400).

Если взять лишь одну пробу на 400 деталей, то с большой уверенностью можно ожидать, что в этой пробе будет нестандартных деталей не менее 28 и не более 63.

Возможно, хотя и маловероятно, что нестандартных деталей окажется меньше 26 либо больше 52.

§11. Случайные величины и случайные векторы

Ранее приводились события, состоящие в появлении того или иного числа. Например, при бросании игральной кости могли появиться числа 1, 2, 3, 4, 5 и 6. Наперед определить число выпавших очков невозможно, поскольку оно зависит от многих случайных причин, которые полностью не могут быть учтены. В этом смысле число очков есть величина случайная, а числа 1, 2, 3, 4, 5 и 6 есть возможные значения этой величины.

Данная величина может принимать отдельные, изолированные возможные значения или принимать любое из значений некоторого промежутка (a, b). Здесь нельзя отделить одно возможное значение от другого промежутком, не содержащим возможных значений случайной величины. Таким образом, можно заключить о целесообразности различать случайные величины, принимающие лишь отдельные, изолированные значения, и случайные величины, возможные значения которых сплошь заполняют некоторый промежуток.

11.1. Случайные величины

Наряду с понятием случайного события в теории вероятности используется и более удобное понятие случайной величины.

Определение 1. Случайной величиной называется величина, принимающая в результате опыта одно из своих возможных значений, причем заранее неизвестно, какое именно.

Будем обозначать случайные величины заглавными буквами латинского алфавита $(X, Y, Z, \ldots)$, а их возможные значения – соответствующими малыми буквами $(x_i, y_i, \ldots)$.

Например, число очков, выпавших при броске игральной кости; число появлений герба при 10 бросках монеты; число выстрелов до первого попадания в цель; расстояние от центра мишени до пробоины при попадании.

Можно заметить, что множество возможных значений для перечисленных случайных величин имеет разный вид: для первых двух величин оно конечно (соответственно 6 и 11 значений), для третьей величины множество значений бесконечно и представляет собой множество натуральных чисел, а для четвертой – все точки отрезка, длина которого равна радиусу мишени. Таким образом, для первых трех величин множество значений из отдельных (дискретных), изолированных друг от друга значений, а для четвертой оно представляет собой непрерывную область. По этому показателю случайные величины подразделяются на две группы: дискретные и непрерывные.

Определение 2. Случайная величина называется дискретной, если она принимает отдельные, изолированные возможные значения с определенными вероятностями.

Определение 3. Случайная величина называется непрерывной, если множество ее возможных значений целиком заполняет некоторый конечный или бесконечный промежуток.

Для задания дискретной случайной величины нужно знать ее возможные значения и вероятности, с которыми принимаются эти значения. Соответствие между ними называется законом распределения случайной величины. Он может иметь вид таблицы, формулы или графика.

Таблица, в которой перечислены возможные значения дискретной случайной величины и соответствующие им вероятности, называется рядом распределения:

x_i	x_1	x_2	…	x_n	…
p_i	p_1	p_2	…	p_n	…

Заметим, что событие, заключающееся в том, что случайная величина примет одно из своих возможных значений, является достоверным, поэтому $\sum_{i=1}^{n(\infty)} p_i = 1.$

<u>Пример 1.</u> Два стрелка делают по одному выстрелу по мишени. Вероятности их попадания при одном выстреле равны соответственно 0,6 и 0,7. Составить ряд распределения случайной величины X – числа попаданий после двух выстрелов.

Решение. Очевидно, что X может принимать три значения: 0, 1 и 2. Их вероятности найдены в примере, рассмотренном в **§9**. Следовательно, ряд распределения имеет вид:

x_i	*0*	*1*	*2*
p_i	*0,12*	*0,46*	*0,42*

Графически закон распределения дискретной случайной величины можно представить в виде многоугольника распределения – ломаной, соединяющей точки плоскости с координатами (x_i , p_i) (рис. 1).

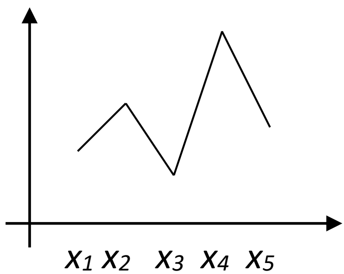


Рис. 1

11.2. Функция и плотность распределения

Определение 1. Функцией распределения $F(x)$ случайной величины X называется вероятность того, что случайная величина примет значение, меньшее x:

$F(x) = p(X < x)$. (1.1)

Свойства функции распределения:

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

Действительно, так как функция распределения представляет собой вероятность, она может принимать только те значения, которые принимает вероятность.

2) Функция распределения является неубывающей функцией, то есть $F(x_2) \ge F(x_1)$ при $x_2 > x_1$. Это следует из того, что $F(x_2) = p(X < x_2) = p(X < x_1) + p(x_1 \le X < x_2) \ge F(x_1)$.

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

В частности, если все возможные значения X лежат на интервале $[a, b]$, то $F(x) = 0$ при $x \le a$ и $F(x) = 1$ при $x \ge b$. Действительно, $X < a$ – событие невозможное, а $X < b$ – достоверное.

4) Вероятность того, что случайная величина примет значение из интервала $[a, b]$, равна разности значений функции распределения на концах интервала: $p(a < X < b) = F(b) - F(a)$.

Справедливость этого утверждения следует из определения функции распределения (см. свойство 2).

Для дискретной случайной величины значение $F(x)$ в каждой точке представляет собой сумму вероятностей тех ее возможных значений, которые меньше аргумента функции.

<u>Пример 1.</u> Найдем $F(x)$ для предыдущего примера:

$$F(x) = \begin{cases} 0, \ x \le 0 \\ 0{,}12, \ 0 < x \le 1 \\ 0{,}12 + 0{,}46 = 0{,}58, \ 1 < x \le 2 \\ 0{,}58 + 0{,}42 = 1, \ x > 2 \end{cases}$$

Соответственно график функции распределения имеет ступенчатый вид (рис. 2):

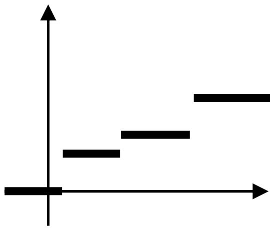


Рис. 2

Определение и свойства функции распределения сохраняются и для непрерывной случайной величины, для которой функцию распределения можно считать одним из видов задания закона распределения. Но для непрерывной случайной величины вероятность каждого отдельного ее значения равна 0. Это следует из свойства 4 функции распределения: $p(X = a) = F(a) - F(a) = 0$. Поэтому для такой случайной величины имеет смысл говорить только о вероятности ее попадания в некоторый интервал.

Вторым способом задания закона распределения непрерывной случайной величины является так называемая плотность распределения (плотность вероятности, дифференциальная функция).

Определение 2. Функция $f(x)$, называемая плотностью распределения непрерывной случайной величины, определяется по формуле:

$$f(x) = F'(x), \tag{1.2}$$

то есть является производной функции распределения.

Свойства плотности распределения:

1) $f(x) \geq 0$, так как функция распределения является неубывающей.

2) $F(x) = \int\limits_{-\infty}^{x} f(t)dt$, что следует из определения плотности распределения.

3) Вероятность попадания случайной величины в интервал (*a,* *b*) определяется формулой $p(a \le X < b) = \int\limits_a^b f(x)dx.$

Действительно,

$$p(a \le X < b) = F(b) - F(a) = \int\limits_{-\infty}^{b} f(x)dx - \int\limits_{-\infty}^{a} f(x)dx = \int\limits_{a}^{b} f(x)dx.$$

4) $\int\limits_{-\infty}^{+\infty} f(x)dx = 1$ (условие нормировки).

Его справедливость следует из того, что $\int\limits_{-\infty}^{+\infty} f(x)dx = F(+\infty),$ а $\lim\limits_{x\to+\infty} F(x) = 1.$

5) $\lim\limits_{x\to\pm\infty} f(x) = 0,$ так как $F(x) \to const$ при $x \to \pm\infty.$

Таким образом, график плотности распределения представляет собой кривую, расположенную выше оси *Ox*, причем эта ось является ее горизонтальной асимптотой при $x \to \pm\infty$ (последнее справедливо только для случайных величин, множеством возможных значений которых является все множество действительных чисел). Площадь криволинейной трапеции, ограниченной графиком этой функции, равна единице.

Замечание. Если все возможные значения непрерывной случайной величины сосредоточены на интервале [*a*, *b*], то все интегралы вычисляются в этих пределах, а вне интервала [*a, b*] $f(x) \equiv 0$.

<u>Пример 2.</u> Плотность распределения непрерывной случайной величины задана формулой $f(x) = \dfrac{C}{1+x^2},\ -\infty < x < +\infty.$

Найти: а) значение константы С; б) вид функции распределения; в) $p(-1 < x < 1)$.

Решение. а) значение константы С найдем из свойства 4:

$$\int_{-\infty}^{+\infty} \frac{C}{1+x^2} dx = Carctgx \Big|_{-\infty}^{+\infty} = C\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = C\pi = 1,$$

откуда

$C = \dfrac{1}{\pi}$.

б)

$$F(x) = \frac{1}{\pi} \int_{-\infty}^{x} \frac{1}{1+t^2} dt = \frac{1}{\pi} arctg\ t \Big|_{-\infty}^{x} = \frac{1}{\pi}\left(arctgx + \frac{\pi}{2}\right) = \frac{1}{\pi} arctgx + \frac{1}{2}.$$

в)

$$p(-1 < x < 1) = \frac{1}{\pi} \int_{-1}^{1} \frac{1}{1+x^2} dx = \frac{1}{\pi} arctgx \Big|_{-1}^{1} = \frac{1}{\pi}\left(\frac{\pi}{4} + \frac{\pi}{4}\right) = 0{,}5.$$

<u>Пример 1.4.</u> Функция распределения непрерывной случайной величины имеет вид:

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

Найти плотность распределения.

Решение.

$$f(x)=\begin{cases} 0', & x\le 2 \\ \left(\dfrac{x-2}{2}\right)', & 2<x\le 4 \\ 1', & x>4 \end{cases}=\begin{cases} 0, & x\le 2 \\ 0{,}5, & 2<x\le 4 \\ 0, & x>4. \end{cases}$$

11.3. Основные дискретные и абсолютно непрерывные распределения

Вернемся к схеме независимых испытаний и найдем закон распределения случайной величины X – числа появлений события A в серии из n испытаний. Возможные значения A: 0, 1, ..., n. Соответствующие им вероятности можно вычислить по формуле Бернулли:

$$p(X=k)=C_n^k p^k q^{n-k} \qquad (1.3)$$

(p – вероятность появления A в каждом испытании).

Такой закон распределения называют биномиальным, поскольку правую часть равенства (1.2) можно рассматривать как общий член разложения бинома Ньютона:

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

Пример 1. Составим ряд распределения случайной величины X – числа попаданий при 5 выстрелах, если вероятность попадания при одном выстреле равна 0,8.

Решение. p(X=0) = $1\cdot(0{,}2)^5$ = 0,00032; p(X=1) = $5\cdot 0{,}8\cdot(0{,}2)^4$ = 0,0064; p(X=2) = $10\cdot(0{,}8)^2\cdot(0{,}2)^3$ = 0,0512; p(X=3) = $10\cdot(0{,}8)^3\cdot(0{,}2)^2$ = 0,2048; p(X=4) = $5\cdot(0{,}8)^4\cdot 0{,}2$ = 0,4096; p(X=5) = $1\cdot(0{,}8)^5$ = 0,32768. Таким образом, ряд распределения имеет вид:

x	*0*	*1*	*2*	*3*	*4*	*5*
p	*0.00032*	*0.0064*	*0.0512*	*0.2048*	*0.4096*	*0.32728*

Пусть производятся независимые испытания, в каждом из которых вероятность появления события A равна p ($0<p<1$) и, следовательно, вероятность его непоявления

q=1-p. Испытания заканчиваются, как только появится событие A. Таким образом, если событие A появилось в k-м испытании, то в предшествующих k-1 испытаниях оно не появилось.

Обозначим через X дискретную случайную величину – число испытаний, которые нужно провести до первого появления события A. Очевидно, возможными значениями X являются натуральные числа: $x_1 = 1, x_2 = 2, \ldots$

Пусть в первых k-1 испытаниях событие A не наступило, а в k-м испытании появилось. Вероятность этого «сложного события», по теореме умножения вероятностей независимых событий,

$$P(X = k) = q^{k-1} p .$$

Полагая k = 1, 2, ... в формуле Бернулли, получим геометрическую прогрессию с первым членом p и знаменателем q (0<q< 1):

$$p, \quad qp, \quad q^2 p, \quad \ldots, q^{k-1} p, \quad \ldots \tag{1.4}$$

По этой причине данное распределение называют геометрическим.

Легко убедиться, что ряд (1.4) сходится и сумма его равна единице. Действительно, сумма ряда (1.4) $p/(1-q) = p/p = 1$.

<u>Пример 2.</u> Из орудия производится стрельба по цели до первого попадания. Вероятность попадания в цель p = 0,6. Найти вероятность того, что попадание произойдет при третьем выстреле.

Решение. По условию, p = 0,6, q = 0,4, k = 3. Искомая вероятность по формуле Бернулли

$$P = q^{k-1} \cdot p = 0{,}4^2 \cdot 0{,}6 = 0{,}096 .$$

Рассмотрим дискретную случайную величину X, принимающую только целые неотрицательные значения (0, 1, 2,..., m,...), последовательность которых не ограничена.

Такая случайная величина называется распределенной по закону Пуассона, если вероятность того, что она примет значение m, выражается формулой:

$$p(X = m) = \frac{a^m}{m!} e^{-a} \quad , \tag{1.5}$$

где a – некоторая положительная величина, называемая параметром закона Пуассона.

Покажем, что сумма всех вероятностей равна 1:

$$\sum_{m=0}^{\infty} p(X = m) = e^{-a} \sum_{m=0}^{\infty} \frac{a^m}{m!} = e^{-a} \cdot e^{a} = 1$$

(использовано разложение в ряд Тейлора функции e^x).

Рассмотрим типичную задачу, приводящую к распределению Пуассона. Пусть на оси абсцисс случайным образом распределяются точки, причем их распределение удовлетворяет следующим условиям:

1) вероятность попадания некоторого количества точек на отрезок длины l зависит только от длины отрезка и не зависит от его расположения на оси (то есть точки распределены с одинаковой средней плотностью);

2) точки распределяются независимо друг от друга (вероятность попадания какого-либо числа точек на данный отрезок не зависит от количества точек, попавших на любой другой отрезок);

3) практическая невозможность совпадения двух или более точек.

Тогда случайная величина X – число точек, попадающих на отрезок длины l – распределена по закону Пуассона, где a – среднее число точек, приходящееся на отрезок длины l.

Замечание. Ранее говорилось о том, что формула Пуассона выражает биномиальное распределение при большом числе опытов и малой вероятности события.

Поэтому закон Пуассона часто называют законом редких явлений.

Часто на практике мы имеем дело со случайными величинами, распределенными определенным типовым образом, то есть такими, закон распределения которых имеет некоторую стандартную форму. В прошлых параграфах были рассмотрены примеры таких законов распределения для дискретных случайных величин (биномиальный и Пуассона). Для непрерывных случайных величин тоже существуют часто встречающиеся виды закона распределения, и в качестве первого из них рассмотрим равномерный закон.

Определение 1. Закон распределения непрерывной случайной величины называется равномерным, если на интервале, которому принадлежат все возможные значения случайной величины, плотность распределения сохраняет постоянное значение $f(x) = const$ при $a \le x \le b$, $f(x) = 0$ при $x < a$, $x > b$.

Найдем значение, которое принимает $f(x)$ при $x \in [a,b]$. Из условия нормировки следует, что $\int_a^b f(x)dx = \int_a^b cdx = c(b-a) = 1,$ откуда $f(x) = c = \frac{1}{b-a}$.

Вероятность попадания равномерно распределенной случайной величины на интервал $[\alpha, \beta]$ $(a \le \alpha < \beta \le b)$ равна при этом $\int_\alpha^\beta \frac{1}{b-a}dx = \frac{\beta-\alpha}{b-a}.$

Вид функции распределения для нормального закона:

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

Пример 3. Автобусы некоторого маршрута идут с интервалом 5 минут. Найти вероятность того, что пришедшему на остановку пассажиру придется ожидать автобуса не более 2 минут.

Решение. Время ожидания является случайной величиной, равномерно распределенной в интервале [0, 5]. Тогда $f(x)=\dfrac{1}{5},\ \ p(0\le x\le 2)=\dfrac{2}{5}=0{,}4.$

Числовые характеристики случайных величин

Закон распределения случайной величины (функция распределения и ряд распределения или плотность вероятности) полностью описывают поведение случайной величины. Но в ряде задач достаточно знать некоторые числовые характеристики исследуемой случайной величины (например, ее среднее значение и возможное отклонение от него), чтобы ответить на поставленный вопрос. В следующих параграфах рассмотрим основные числовые характеристики дискретных и непрерывных случайных величин.

Следует отметить, что основную роль на практике играют математическое ожидание, задающее «центральное» значение случайной величины, и дисперсия, характеризующая «разброс» значений случайной величины вокруг ее математического ожидания. В математической статистике для построения доверительных интервалов и проверки статистических гипотез широко используются квантили.

§12. Основные числовые характеристики дискретных случайных величин

12.1. Основные числовые характеристики дискретных случайных величин

Определение 1. Математическим ожиданием дискретной случайной величины называется сумма произведений ее возможных значений на соответствующие им вероятности:

$$M(X) = x_1p_1 + x_2p_2 + \dots + x_np_n . \qquad (1.1)$$

Если число возможных значений случайной величины бесконечно, то $M(X)=\sum_{i=1}^{\infty} x_i p_i$, причем математическое ожидание существует, если полученный ряд сходится абсолютно.

Замечание 1. Математическое ожидание называют иногда взвешенным средним, так как оно приближенно равно среднему арифметическому наблюдаемых значений случайной величины при большом числе опытов.

Замечание 2. Из определения математического ожидания следует, что его значение не меньше наименьшего возможного значения случайной величины и не больше наибольшего.

Замечание 3. Математическое ожидание дискретной случайной величины есть неслучайная (постоянная) величина. В дальнейшем увидим, что это же справедливо и для непрерывных случайных величин.

Пример 1. Найти математическое ожидание случайной величины X – числа стандартных деталей среди трех, отобранных из партии в 10 деталей, среди которых 2 бракованных.

Решение. Составим ряд распределения для X. Из условия задачи следует, что X может принимать значения 1, 2, 3.

$$p(1)=\frac{C_8^1\cdot C_2^2}{C_{10}^3}=\frac{1}{15},\quad p(2)=\frac{C_8^2\cdot C_2^1}{C_{10}^3}=\frac{7}{15},\quad p(3)=\frac{C_8^3}{C_{10}^3}=\frac{7}{15}.$$

Тогда $$M(X) = 1 \cdot \frac{1}{15} + 2 \cdot \frac{7}{15} + 3 \cdot \frac{7}{15} = 2{,}4.$$

Пример 2. Определить математическое ожидание случайной величины X – числа бросков монеты до первого появления герба.

Решение. Эта величина может принимать бесконечное число значений (множество возможных значений есть множество натуральных чисел). Ряд ее распределения имеет вид:

X	*1*	*2*	...	*n*	...
p	*0,5*	*(0,5)*^{2}	...	*(0,5)*^{n}	...

Тогда

$$M(X) = \sum_{n=1}^{\infty} \frac{n}{2^n} = \frac{1}{2} + 2 \cdot \left(\frac{1}{2}\right)^2 + 3 \cdot \left(\frac{1}{2}\right)^3 + \ldots + n \cdot \left(\frac{1}{2}\right)^n + \ldots = \sum_{n=1}^{\infty} \frac{1}{2^n} + \frac{1}{2} \sum_{n=1}^{\infty} \frac{1}{2^n} + .$$

..+

$$+ \frac{1}{2^n} \sum_{n=1}^{\infty} \frac{1}{2^n} + \ldots = 1 \cdot \left(1 + \frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^n} + \ldots\right) = 1 \cdot 2 = 2$$

(при вычислении дважды использовалась формула суммы бесконечно убывающей геометрической прогрессии: $S = \dfrac{b_1}{1-q}$, откуда

$$\frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^n} + \ldots = 1, \; 1 + \frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^n} + \ldots = 2$$

Свойства математического ожидания:

1) Математическое ожидание постоянной равно самой постоянной:

$$M(C) = C \tag{1.2}$$

Доказательство. Если рассматривать C как дискретную случайную величину, принимающую только одно значение C с вероятностью $p = 1$, то M(C) = C·1 = C.

2) Постоянный множитель можно выносить за знак математического ожидания:

$$M(CX) = C\,M(X) \qquad (1.3)$$

Доказательство. Если случайная величина X задана рядом распределения

x_i	x_1	x_2	…	x_n
p_i	p_1	p_2	…	p_n

то ряд распределения для СХ имеет вид:

Cx_i	Cx_1	Cx_2	…	Cx_n
p_i	p_1	p_2	…	p_n

Тогда $M(CX) = Cx_1p_1 + Cx_2p_2 + \ldots + Cx_np_n = C(x_1p_1 + x_2p_2 + \ldots + x_np_n) = CM(X)$.

Определение 2. Две случайные величины называются независимыми, если закон распределения одной из них не зависит от того, какие значения приняла другая. В противном случае случайные величины зависимы.

Определение 3. Назовем произведением независимых случайных величин X и Y случайную величину XY, возможные значения которой равны произведениям всех возможных значений X на все возможные значения Y, а соответствующие им вероятности равны произведениям вероятностей сомножителей.

3) Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий:

$$M(XY) = M(X)M(Y) \qquad (1.4)$$

Доказательство. Для упрощения вычислений ограничимся случаем, когда X и Y принимают только по два возможных значения:

x_i	x_1	x_2
p_i	p_1	p_2

y_i	y_1	y_2
g_i	g_1	g_2

Тогда ряд распределения для XY выглядит так:

XY	x_1y_1	x_2y_1	x_1y_2	x_2y_2

p	p_1g_1	$p_2\,g_1$	p_1g_2	p_2g_2

Следовательно, $M(XY) = x_1y_1{\cdot}p_1g_1 + x_2y_1{\cdot}p_2g_1 + x_1y_2{\cdot}p_1g_2 + x_2y_2{\cdot}p_2g_2 = y_1g_1(x_1p_1 + x_2p_2) + y_2g_2(x_1p_1 + x_2p_2) = (y_1g_1 + y_2g_2)\,(x_1p_1 + x_2p_2) = M(X){\cdot}M(Y)$.

Замечание 4. Аналогично можно доказать это свойство для большего количества возможных значений сомножителей.

Замечание 5. Свойство 3 справедливо для произведения любого числа независимых случайных величин, что доказывается методом математической индукции.

Определение 4. Определим сумму случайных величин X и Y как случайную величину $X + Y$, возможные значения которой равны суммам каждого возможного значения X с каждым возможным значением Y; вероятности таких сумм равны произведениям вероятностей слагаемых (для зависимых случайных величин – произведениям вероятности одного слагаемого на условную вероятность второго).

4) Математическое ожидание суммы двух случайных величин (зависимых или независимых) равно сумме математических ожиданий слагаемых:

$$M\,(X + Y) = M\,(X) + M\,(Y) \qquad (1.5)$$

Доказательство. Вновь рассмотрим случайные величины, заданные рядами распределения, приведенными при доказательстве свойства 3. Тогда возможными значениями $X + Y$ являются $x_1 + y_1,\ x_1 + y_2,\ x_2 + y_1,\ x_2 + y_2$. Обозначим их вероятности соответственно как p_{11}, p_{12}, p_{21} и p_{22}. Найдем $M(X +Y) = (x_1 + y_1)p_{11} + (x_1 + y_2)p_{12} + (x_2 + y_1)p_{21} + (x_2 + y_2)p_{22} =$

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

Докажем, что $p_{11} + p_{22} = p_1$. Действительно, событие, состоящее в том, что $X + Y$ примет значения $x_1 + y_1$ или $x_1 +$

y_2 и вероятность которого равна $p_{11} + p_{22}$, совпадает с событием, заключающемся в том, что $X = x_1$ (его вероятность – p_1). Аналогично доказывается, что $p_{21} + p_{22} = p_2$, $p_{11} + p_{21} = g_1$, $p_{12} + p_{22} = g_2$. Значит,

$$M(X + Y) = x_1p_1 + x_2p_2 + y_1g_1 + y_2g_2 = M(X) + M(Y).$$

Замечание 6. Из свойства 4 следует, что сумма любого числа случайных величин равна сумме математических ожиданий слагаемых.

Пример 3. Найти математическое ожидание суммы числа очков, выпавших при броске пяти игральных костей.

Решение. Найдем математическое ожидание числа очков, выпавших при броске одной кости: $\mathrm{M(X_1)} = (1 + 2 + 3 + 4 + 5 + 6) \cdot \frac{1}{6} = \frac{7}{2}$. Тому же числу равно математическое ожидание числа очков, выпавших на любой кости. Следовательно, по свойству 4 $M(X) = 5 \cdot \frac{1}{6} = \frac{5}{6}$.

Для того, чтобы иметь представление о поведении случайной величины, недостаточно знать только ее математическое ожидание. Рассмотрим две случайные величины: *X* и *Y*, заданные рядами распределения вида

X	49	50	51
p	0,1	0,8	0,1

Y	0	100
p	0,5	0,5

Найдем $M(X) = 49 \cdot 0{,}1 + 50 \cdot 0{,}8 + 51 \cdot 0{,}1 = 50$, $M(Y) = 0 \cdot 0{,}5 + 100 \cdot 0{,}5 = 50$. Как видно, математические ожидания обеих величин равны, но если для *X* *M(X)* хорошо описывает поведение случайной величины, являясь ее наиболее вероятным возможным значением (причем остальные значения ненамного отличаются от 50), то значения *Y* существенно отстоят от *M(Y)*. Следовательно, наряду с математическим ожиданием желательно знать, насколько значения случайной величины отклоняются от него. Для характеристики этого показателя служит дисперсия.

Определение 5. Дисперсией (рассеянием) случайной величины называется математическое ожидание квадрата ее отклонения от ее математического ожидания:

$$D(X) = M\,(X - M(X))^2. \qquad (1.6)$$

Пример 4. Найти дисперсию случайной величины *X* (числа стандартных деталей среди отобранных) в примере 1.

Решение. Вычислим значения квадрата отклонения каждого возможного значения от математического ожидания:

$(1 - 2{,}4)^2 = 1{,}96$; $(2 - 2{,}4)^2 = 0{,}16$; $(3 - 2{,}4)^2 = 0{,}36$.

Следовательно,

$$D(X) = 1{,}96 \cdot \frac{1}{15} + 0{,}16 \cdot \frac{7}{15} + 0{,}36 \cdot \frac{7}{15} = \frac{28}{75} \approx 0{,}373.$$

Замечание 7. В определении дисперсии оценивается не само отклонение от среднего, а его квадрат. Это сделано для того, чтобы отклонения разных знаков не компенсировали друг друга.

Замечание 8. Из определения дисперсии следует, что эта величина принимает только неотрицательные значения.

Замечание 9. Существует более удобная для расчетов формула для вычисления дисперсии, справедливость которой доказывается в следующей теореме.

Теорема 1.

$$D(X) = M(X^2) - M^2(X). \qquad (1.7)$$

Доказательство. Используя то, что *M(X)* – постоянная величина, и свойства математического ожидания, преобразуем формулу (1.6) к виду:

$D(X) = M(X - M(X))^2 = M(X^2 - 2X \cdot M(X) + M^2(X)) = M(X^2) - 2M(X) \cdot M(X) + M^2(X) = M(X^2) - 2M^2(X) + M^2(X) = M(X^2) - M^2(X)$, что и требовалось доказать.

<u>Пример 5.</u> Вычислить дисперсии случайных величин *X* и *Y*, рассмотренных в начале этого раздела.

Решение. *M(X)* = $(49^2 \cdot 0{,}1 + 50^2 \cdot 0{,}8 + 51^2 \cdot 0{,}1) - 50^2 = 2500{,}2 - 2500 = 0{,}2$.

M(Y) = $(0^2 \cdot 0{,}5 + 100^2 \cdot 0{,}5) - 50^2 = 5000 - 2500 = 2500$. Итак, дисперсия второй случайной величины в несколько тысяч раз больше дисперсии первой. Таким образом, даже не зная законов распределения этих величин, по известным значениям дисперсии мы можем утверждать, что *X* мало отклоняется от своего математического ожидания, в то время как для *Y* это отклонение весьма существенно.

Свойства дисперсии:

1) Дисперсия постоянной величины С равна нулю:

$$D(C) = 0. \qquad (1.8)$$

Доказательство. $D(C) = M((C - M(C))^2) = M((C - C)^2) = M(0) = 0$.

2) Постоянный множитель можно выносить за знак дисперсии, возведя его в квадрат:

$$D(CX) = C^2D(X). \qquad (1.9)$$

Доказательство. $D(CX) = M((CX - M(CX))^2) = M((CX - CM(X))^2) = M(C^2(X - M(X))^2) = C^2D(X)$.

3) Дисперсия суммы двух независимых случайных величин равна сумме их дисперсий:

$$D(X + Y) = D(X) + D(Y). \qquad (1.10)$$

Доказательство. $D(X + Y) = M(X^2 + 2XY + Y^2) - (M(X) + M(Y))^2 = M(X^2) + 2M(X)M(Y) + M(Y^2) - M^2(X) - 2M(X)M(Y) - M^2(Y) = (M(X^2) - M^2(X)) + (M(Y^2) - M^2(Y)) = D(X) + D(Y)$.

Следствие 1. Дисперсия суммы нескольких взаимно независимых случайных величин равна сумме их дисперсий.

Следствие 2. Дисперсия суммы постоянной и случайной величин равна дисперсии случайной величины.

4) Дисперсия разности двух независимых случайных величин равна сумме их дисперсий:

$$D(X - Y) = D(X) + D(Y). \qquad (1.11)$$

Доказательство. D(X – Y) = D(X) + D(-Y) = D(X) + (-1)²D(Y) = D(X) + D(X).

Дисперсия дает среднее значение квадрата отклонения случайной величины от среднего; для оценки самого отклонения служит величина, называемая средним квадратичным отклонением.

Определение 6. Средним квадратичным отклонением σ случайной величины X называется квадратный корень из дисперсии:

$$\sigma = \sqrt{D(X)}. \qquad (1.12)$$

<u>**Пример 6.**</u> В предыдущем примере средние квадратичные отклонения X и Y равны соответственно $\sigma_x = \sqrt{0{,}2} \approx 0{,}447; \quad \sigma_y = \sqrt{2500} = 50.$

12.2. Основные числовые характеристики непрерывных случайных величин

Распространим определения числовых характеристик случайных величин на непрерывные случайные величины, для которых плотность распределения служит в некотором роде аналогом понятия вероятности.

Определение 1. Математическим ожиданием непрерывной случайной величины называется

$$M(X) = \int_{-\infty}^{+\infty} x f(x) dx. \qquad (1.13)$$

Замечание 1. Общее определение дисперсии сохраняется для непрерывной случайной величины таким же, как и для дискретной, а формула для ее вычисления имеет вид:

$$D(X)=\int_{-\infty}^{+\infty} x^2 f(x)dx - M^2(X). \qquad (1.14)$$

Среднее квадратичное отклонение вычисляется по формуле (1.12).

Замечание 2. Если все возможные значения непрерывной случайной величины не выходят за пределы интервала $[a, b]$, то интегралы в формулах (1.13) и (1.14) вычисляются в этих пределах.

Пример 1. Плотность распределения случайной величины X имеет вид:

$$f(x)=\begin{cases} 0, & x<2 \\ -\dfrac{3}{4}(x^2-6x+8), & 2\le x\le 4 \\ 0, & x>4. \end{cases}$$

Найти $M(X)$, $D(X)$, σ.

Решение.

$$M(X)=-\frac{3}{4}\int_2^4 x(x^2-6x+8)dx=-\frac{3}{4}\left(\frac{x^4}{4}-2x^3+4x^2\right)\Bigg|_2^4=3;$$

$$D(X)=-\frac{3}{4}\int_2^4 x^2(x^2-6x+8)dx-9=-\frac{3}{4}\left(\frac{x^5}{5}-\frac{3x^4}{2}+\frac{8x^3}{3}\right)\Bigg|_2^4-9=9{,}2-9=0{,}2;$$

$$\sigma=\sqrt{0{,}2}\approx 0{,}447.$$

12.3. Числовые характеристики случайных величин, имеющих некоторые стандартные законы распределения

1. Биномиальное распределение.

Для дискретной случайной величины X, представляющей собой число появлений события A в серии из n независимых испытаний, $M(X)$ можно найти, используя свойство 4 математического ожидания. Пусть X_1 – число появлений A в первом испытании, X_2 – во втором и т.д. При этом каждая из случайных величин X_i задается рядом распределения вида

X_i	0	1
p_i	q	p

Следовательно, $M(X_i) = p$. Тогда $M(X) = \sum_{i=1}^{n} M(X_i) = \sum_{i=1}^{n} p = np.$

Аналогичным образом вычислим дисперсию: $D(X_i) = 0^2 \cdot q + 1^2 \cdot p - p^2 = p - p^2 = p(1 - p)$, откуда по свойству 4

$$D(X) = \sum_{i=1}^{n} D(X_i) = np(1-p) = npq.$$

дисперсии

2. Закон Пуассона.

Если $p(X = m) = \dfrac{a^m}{m!} e^{-a}$, то $M(X) = \sum_{m=1}^{\infty} m \dfrac{a^m}{m!} e^{-a} = ae^{-a} \sum_{m=1}^{\infty} \dfrac{a^{m-1}}{(m-1)!} = ae^{-a} e^{a} = a$

(использовалось разложение в ряд Тейлора функции e^x).

Для определения дисперсии найдем вначале $M(X^2) = \sum_{m=1}^{\infty} m^2 \dfrac{a^m}{m!} e^{-a} = a \sum_{m=1}^{\infty} m \dfrac{a^{m-1}}{(m-1)!} e^{-a} =$

$$= a\sum_{m=1}^{\infty} \left((m-1)+1\right) \frac{a^{m-1}}{(m-1)!} e^{-a} = a\left(\sum_{m=1}^{\infty} (m-1) \frac{a^{m-1}}{(m-1)!} e^{-a} + \sum_{m=1}^{\infty} \frac{a^{m-1}}{(m-1)!} e^{-a}\right) = a(a+1).$$

Поэтому $D(X) = a^2 + a - a^2 = a$.

Замечание. Таким образом, обнаружено интересное свойство распределения Пуассона: математическое ожидание равно дисперсии (и равно единственному параметру a, определяющему распределение).

3. Равномерное распределение.

Для равномерно распределенной на отрезке $[a,\ b]$ непрерывной случайной величины

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

то есть математическое ожидание равномерно распределенной случайной величины равно абсциссе середины отрезка [a, b] .

Дисперсия

$$D(X)=\int_a^b x^2\frac{1}{b-a}dx-\frac{(a+b)^2}{4}=\frac{b^3-a^3}{3(b-a)}-\frac{(a+b)^2}{4}=\frac{a^2+ab+b^2}{3}-\frac{a^2+2ab+b^2}{4}==\frac{(b-a)^2}{12}.$$

4. Нормальное распределение.

Для вычисления математического ожидания нормально распределенной случайной величины воспользуемся тем, что интеграл Пуассона $\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz=\sqrt{2\pi}$.

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

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sigma ze^{-\frac{z^2}{2}}dz+\frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz=0+\frac{a}{\sqrt{2\pi}}\sqrt{2\pi}=a$$

(первое слагаемое равно 0, так как подынтегральная функция нечетна, а пределы интегрирования симметричны относительно нуля).

$$D(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-a)^2e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}z\cdot ze^{-\frac{z^2}{2}}dz=(u=z,dv=ze^{-\frac{z^2}{2}})=$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\left(\left.-z\cdot e^{-\frac{z^2}{2}}\right|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz\right)=\frac{\sigma^2}{\sqrt{2\pi}}\left(-0+\sqrt{2\pi}\right)=\sigma^2$$

.

Следовательно, параметры нормального распределения (a и σ) равны соответственно математическому ожиданию и среднему квадратичному отклонению исследуемой случайной величины.

§13. Основные понятия математической статистики

Раздел «Математическая статистика» общего курса математики обычно изучается последним, так как основывается на теории вероятностей и всех других ранее изученных математических разделов. Сами статистические данные и выводы, полученные на их основе, используются в естественных и гуманитарных науках, в инженерной практике, экономике. Особенно велика роль статистики в решении задач управления производством, социальными группами людей, ибо без знания состояния управляемого объекта разумное управление этим объектом невозможно. Эти знания об объекте несут обработанные и осмысленные статистические данные.

Слово «статистика» происходит от слова «status» – состояние, государство. Статистика – одна из древнейших наук. Еще в глубокой древности люди накапливали и анализировали сведения о природных и общественных явлениях с целью их познания и прогноза. Существуют производственная, экономическая, социальная, медицинская, демографическая и другие отраслевые статистики. Математическая статистика изучает математическую сторону работы с числовыми данными независимо от конкретной отраслевой специфики.

13.1. Предмет и основные задачи математической статистики

Определение 1. *Математической статистикой* называется наука, разрабатывающая методы регистрации, описания и анализа данных наблюдений и экспериментов с целью получения вероятностно-статистических моделей случайных явлений.

Математическая статистика – абстрактная наука. Ее методы применимы для обработки данных наблюдений и экспериментов любой природы, поэтому используются во всех конкретных естественных и гуманитарных науках, экономике, технике, медицине и т. д., т. е. во всех отраслевых статистиках.

Основными задачами математической статистики являются следующие:

- Приближенное определение вероятности события по относительной частоте.
- Нахождение приближенного закона распределения случайной величины по данным экспериментов.
- Оценивание числовых характеристик или параметров распределения случайной величины по данным экспериментов.
- Проверка статистических гипотез о свойствах изучаемого случайного явления.
- Определение эмпирической (регрессионной) зависимости между переменными, описывающими случайное явление, на основе экспериментальных данных.

Рассмотрим типичную схему исследований при решения указанных задач. Эти исследования естественно делятся на две части.

Сначала путем наблюдений и экспериментов собираются, регистрируются статистические данные, составляющие *выборку,* – это числа, называемые также *выборочными элементами.* Затем они упорядочиваются, представляются в компактной, наглядной или функциональной форме. Вычисляются различного рода средние величины, характеризующие выборку. Часть математической статистики, обеспечивающая эту работу, называется *описательной статисткой* (descriptive statistics).

Вторая часть работы исследователя состоит в получении на основе найденных сведений о выборке достаточно обоснованных выводов о свойствах исследуемого случайного явления. Эта часть работы обеспечивается статистическими методами, составляющими *статистику выводов* (inferential statistics).

Краткие исторические сведения

Несомненно, что создание теории вероятностей в XVII веке и ее развитие стимулировались требованиями улучшения статистической обработки данных экономики, демографии, страхового дела. Сам термин «статистика» возник в XVIII в. К этому же времени относится начало преподавания статистики в университетах Германии.

Методы теории вероятностей стали оказывать влияние и на статистику. Она перестает быть чисто описательной и наполняется математическими моделями изучаемых случайных явлений.

Рождение математической статистики, как полноправной науки, ее выделение из ряда отраслевых и хозяйственных статистик произошло после того, как она стала строиться на основе теории вероятностей. Это случилось в двадцатом столетии, когда и сама теория вероятностей стала строиться на основе аксиоматики, теории множеств, функционального анализа.

Основоположниками математической статистики были Я. Бернулли (1654–1705), К. Гаусс (1777–1855), П. Лаплас (1749–1827). В XIX веке большой вклад в нее сделали российские математики П. Л. Чебышёв (1821–

1894), А. А. Марков (1856–1922), А. М. Ляпунов (1857–1918). В XX веке важные результаты были получены К. Пирсоном (1857–1936), Р. Фишером (1890–1962), Г. Крамером (1893–1985), Р. Мизесом (1883–1953), а также отечественными математиками А. Н. Колмогоровым (1903–1987), Н. В. Смирновым (1900–1966) и др.

13.2. Описательная статистика

Генеральной совокупностью называется множество возможных значений изучаемой случайной величины *X* с приписанным этому множеству законом распределения *X: L(X).*

Числа, составляющие генеральную совокупность, называются ее *элементами.* Закон *L(X)* распределения случайной величины *X* называется *генеральным законом распределения,* а числовые характеристики *X* – *генеральными числовыми характеристиками.*

Выборкой называется множество измеренных значений $x_1, x_2, ..., x_n$ случайной величины *X*.

Выборка записывается в виде *n*-мерной точки ($x_1, x_2, ..., x_n$). Числа, составляющие выборку, называются ее *элементами;* их количество *n –объемом выборки.*

Повторной называют выборку, при которой отобранный объект (перед отбором следующего) возвращается в генеральную совокупность.

Бесповторной называют выборку, при которой отобранный объект в генеральную совокупность не возвращается. На практике обычно пользуются бесповторным случайным отбором.

Для того чтобы по данным выборки можно было достаточно уверенно судить об интересующем признаке генеральной совокупности, необходимо, чтобы объекты выборки правильно его представляли. Другими словами, выборка должна правильно представлять пропорции генеральной совокупности. Это требование коротко формулируют так: выборка должна быть *репрезентативной (представительной).*

Пусть из генеральной совокупности извлечена выборка, причем x_1 наблюдалось n_1 раз, $x_2 - n_2$ раз, $x_k - n_k$ раз и $\sum n_i = n$ – объем выборки. Наблюдаемые значения x_i – называют *вариантами,* а последовательность вариант, записанных в возрастающем порядке, – *вариационным рядом.* Числа наблюдений называют *частотами,* а их отношения к объему выборки $n_i / n = W_i$ – *относительными частотами.*

Элементы вариационного ряда называются *порядковыми статистиками.* Минимальный и максимальный элементы называются *крайними,* иначе – *экстремальными элементами* вариационного ряда:

$$x_{\min} = x_{(1)}, x_{\max} = x_{(n)}$$

Разность между максимальным и минимальным элементами называется *размахом,* или *широтой* выборки:

$$R = x_{\max} - x_{\min}$$

Средний элемент вариационного ряда, если n – нечетное, или полусумма двух средних элементов, если n – четное, называется *медианой выборки* и обозначается med.

Элементы вариационного ряда, на четверть отстоящие от краев, называются соответственно *нижней* и *верхней квартилями* и обозначаются $Z_{1/4}$ и $Z_{3/4}$.

Математически точно квартили определяются по формулам:

$$z_{1/4} = x_i; i = \begin{cases} \left[n/4\right] + 1 \text{ при } n/4 \text{ дробном} \\ n/4 \text{ при } n/4 \text{ целом} \end{cases} \qquad z_{3/4} = x_{(n-i-1)}$$

Здесь [*a*] – целая часть числа *a,* т. е. наибольшее целое число, не превосходящее *a.*

Если в выборке много повторяющихся элементов, то вариационный ряд можно преобразовать в статистический ряд.

Статистическим рядом называется последовательность различных элементов z_i, вариационного ряда с указанием частот n_i, повторения элементов.

В общем случае статистический ряд можно записать в виде таблицы $(n_1+n_2+ \ldots + = n)$.

z_i	z_1	z_2	...	...	z_k
n_i	n_1	n_2	...	...	n_k

Статистический ряд можно изобразить графически в виде полигона (многоугольника), откладывая по оси абсцисс элементы статистического ряда, а по оси ординат – частоты (или относительные частоты). Полученные точки плоскости соединяются отрезками.

Элемент, отвечающий наибольшей частоте по сравнению с соседними элементами статистического ряда, называется *выборочной модой* (*mod*).

Эмпирической функцией распределения (функцией распределения выборки) называют функцию $F^*(x)$, определяющую для каждого значения x относительную частоту события $X< x$.

Итак, по определению,

$$F^*(x) = n_x / n,$$

где n_x – число вариант, меньших x; n – объем выборки. Таким образом, для того чтобы найти, например, $F^*(x_2)$, надо число вариант, меньших x_2, разделить на объем выборки:

$$F^*(x_2) = n_{x_2} / n .$$

В отличие от эмпирической функции распределения выборки функцию распределения $F(x)$ генеральной совокупности называют *теоретической функцией распределения.* Различие между эмпирической и теоретической функциями состоит в том, что теоретическая функция $F(x)$ определяет вероятность события $X < x$, а эмпирическая функция $F^*(x)$ определяет относительную частоту этого же события. Из теоремы Бернулли следует, что относительная частота события $X < x$, т. е. $F^*(x)$ стремится по вероятности к вероятности $F(x)$ этого события. Другими словами, при больших n числа $F^*(x)$ и $F(x)$ мало отличаются одно от другого в том смысле, что $\lim\limits_{n\to\infty} P[|F(x) - F^*(x)| < \varepsilon] = 1 (\varepsilon > 0)$. Уже отсюда следует целесообразность использования эмпирической функции распределения выборки для приближенного представления теоретической (интегральной) функции распределения генеральной совокупности.

Это функция распределения дискретной случайной величины X^* заданной таблицей распределения

X^*	z_1	z_2	…	…	z_k
P	n_1/n	n_2/n	…	...	n_k/n

Ее графиком является восходящая ступенчатая линия, называемая кумулятой (линия накопленных относительных частот).

С помощью выборки образуются ее числовые характеристики. Это числовые характеристики случайной величины X^* с равномерным законом распределения, который означает, что каждый элемент выборки $x_k (k = 1,...,n)$ принимается с вероятностью $1/n$, ибо предполагается, что выборка образована с помощью простого случайного выбора.

Числовые характеристики случайной величины X^* называются *выборочными числовыми характеристиками.* Случайная величина X^* аппроксимирует изучаемую случайную величину X в силу того, что $F^*_n\ (x)$ по вероятности стремится к $F_X(x)$ при $n \to \infty$. При этом следует ожидать, что и выборочные числовые характеристики будут аппроксимировать соответствующие генеральные характеристики, т. е. являться их *оценками.* Такой метод образования оценок генеральных числовых характеристик называется методом *аналогии* (или *подстановки).* Вместо числовых характеристик X рассматриваются аналогичные числовые характеристики X^*. Это означает также, что во все формулы для генеральных числовых характеристик вместо X подставляется случайная величинах X^*, ее аппроксимирующая.

Выборочной оценкой генеральной числовой характеристики называется ее приближенное значение, найденное по выборке. Основные оценки:

1. Генеральная средняя $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ является оценкой генерального математического ожидания $\bar{x} = MX$.

2. Выборочная средняя $\bar{x}_в = \frac{1}{n}\sum_{i=1}^{n} n_i x_i$ является оценкой математического ожидания $\bar{x} = M[X_B]$.

3. Выборочный начальный момент порядка l: $a_l = \frac{1}{n}\sum_{i=1}^{n} n_i x_i^l$ является оценкой генерального начального момента порядка l: $\alpha_l = \mathbf{M}[X^l]$.

4. Выборочный центральный момент порядка l: $m_l = \frac{1}{n}\sum_{i=1}^{n} n_i (x_i - \bar{x})^l$ является оценкой генерального центрального момента порядка l: $\mu_l = \mathbf{M}[(X = m)^l]$.

5. Выборочная дисперсия $s^2 = m_2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2$ является оценкой генеральной дисперсии $\sigma^2 = \mu_2 = \mathbf{M}[(X - m)^2]$.

6. Выборочное среднее квадратическое отклонение $s = \sqrt{s^2}$ является оценкой генерального среднего квадратического отклонения $\sigma = \sqrt{\sigma^2}$.

13.3. Группированный статистический ряд

Если выборка получена из непрерывной генеральной совокупности и объем ее большой, то вариационный и статистический ряды, как и сама выборка, будут трудно обозримыми множествами. Действительно, в этом случае при достаточно точном измерении практически не будет равных элементов выборки, ибо вероятность равенства значений непрерывной случайной величины равна нулю. Тогда прибегают к другому способу группирования элементов выборки.

Промежуток $[x_{\min}, x_{\max}]$ делится на некоторое число k равных по длине промежутков. Обозначим эти промежутки слева направо через $\Delta_1, \Delta_2, ..., \Delta_k$. Если точки, разделяющие промежутки, обозначить $a_0, a_1, ..., a_k$, то $\Delta_1 = [x_{\min}, a_1), \Delta_2 = [a_1, a_2), ..., \Delta_k = [a_{k-1}, x_{\max}]$. Пусть n_i – число элементов выборки, попавших в промежуток Δ_i. Числа $n_1, n_2, ..., n_k$ называются частотами попадания элементов выборки в рассматриваемые промежутки.

Совокупность промежутков $\Delta_1, \Delta_2, ..., \Delta_k$ и соответствующих им частот называется *группированным статистическим рядом.*

Для определения *к* можно использовать полуэмпирическую формулу

$k = 1.72n^{1/3}$. Здесь n – объем выборки; $30 \le n \le 1000$. При $n = 40 \Rightarrow k = 6$; при $n = 100 \Rightarrow k = 8$; при $n = 200 \Rightarrow k = 10$; при $n = 400 \Rightarrow k = 12$; при $n = 1000 \Rightarrow k = 17$.

Применяется также формула Старджесса: $k = 1 + 3.3\lg n$.

Длина промежутков $\Delta_1, \Delta_2, ..., \Delta_k$ определится по формуле $h = (x_{\max} - x_{\min})/k$

Вместо группы элементов, попавших в интервал Δ_i, рассматривается один их представитель. В качестве такого представителя обычно берут среднюю точку x_i^* промежутка Δ_i. Группированный статистический ряд можно оформить в виде таблицы.

С помощью группированного статистического ряда можно приближенно вычислить выборочные моменты. Так как группа элементов выборки, входящих в промежуток Δ_i заменяется средней точкой x_i^* промежутка, то следует считать, что элемент x_i^* встречается в выборке n_i, раз, т. е. имеет частоту n_i. Получаем следующие формулы:

$$\bar{x} = a_1 \approx \frac{1}{n}\sum_{i=1}^{k} n_i x_i^* \qquad a_1 \approx \frac{1}{n}\sum_{i=1}^{k} n_i (x_i^*)^l$$

Такое усреднение по промежуткам несколько искажает выборочные числовые характеристики, но при большом объеме выборки это искажение несущественно.

Полигоном частот называют ломаную, отрезки которой соединяют точки *(x_1, n_1), (x_2; n_2), ..., (x_k; n_k)*. Для построения полигона частот на оси абсцисс откладывают варианты x_i, а на оси ординат – соответствующие им частоты n_j. Точки *(x_i; n_i)* соединяют отрезками прямых и получают полигон частот.

Полигоном относительных частот называют ломаную, отрезки которой соединяют точки *(x_1; W_1), (x_2; W_2), ..., (x_k; W_k)*. Для построения полигона относительных частот на оси абсцисс откладывают варианты x_i, а на оси ординат – соответствующие им относительные частоты W_i. Точки *(x_i; W_i)* соединяют отрезками прямых и получают полигон относительных частот.

На рис. 1 изображен пример полигона относительных частот.

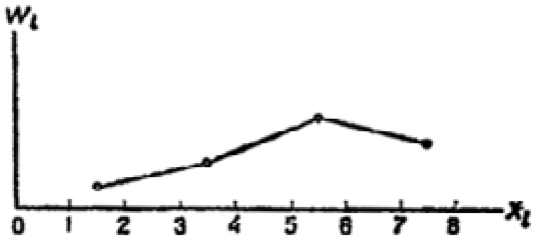


Рис. 1 – Полигон относительных частот

В случае непрерывного признака целесообразно строить гистограмму, для чего интервал, в котором заключены все наблюдаемые значения признака, разбивают на несколько частичных интервалов длиной h и находят для каждого частичного интервала n_i – сумму частот вариант, попавших в i-й интервал.

Гистограммой частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиною h, а высоты равны отношению n_i/h (плотность частоты).

Для построения гистограммы частот на оси абсцисс откладывают частичные интервалы, а над ними проводят отрезки, параллельные оси абсцисс на расстоянии n_i/h.

Площадь i-го частичного прямоугольника равна $hn_i/h = n_i$ – сумме частот вариант i-го интервала; следовательно, *площадь гистограммы частот равна сумме всех частот, т. е. объему выборки.*

На рис. 2 изображена гистограмма частот распределения объема n = 100.

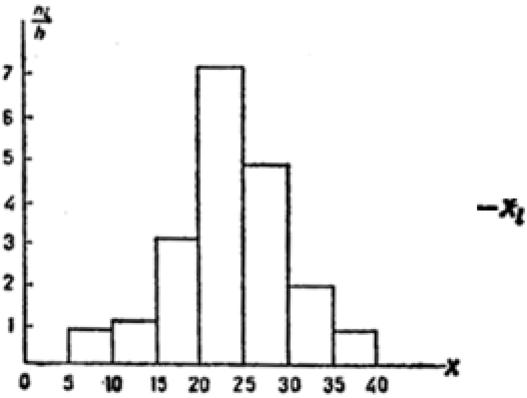


Рис. 2 – Гистограмма частот распределения

Гистограммой относительных частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиною h, а высоты равны отношению W_i/h (плотность относительной частоты).

Для построения гистограммы относительных частот на оси абсцисс откладывают частичные интервалы, а над ними проводят отрезки, параллельные оси абсцисс на расстоянии W_i/h. Площадь 1-го частичного прямоугольника равна $hW_i/h = W_i$ – относительной частоте вариант, попавших в 1-й интервал. Следовательно, *площадь гистограммы относительных частот равна сумме всех относительных частот, т. е. единице.*

Точечное оценивание числовых характеристик и параметров распределения генеральной совокупности

Одной из важнейших задач математической статистики является задача приближенного вычисления числовых характеристик и параметров закона распределения изучаемой случайной величины. Эта задача называется задачей оценивания неизвестных величин. Сформировались два направления в теории оценивания – точечное и интервальное. В параграфах рассматривается теория точечного оценивания.

§14. Точечное оценивание числовых характеристик и параметров распределения генеральной совокупности

14.1. Понятие точечной статистической оценки. Требования к оценкам

Определение 1. *Точечной статистической оценкой* неизвестной числовой характеристики или параметра θ распределения называется функция в $\bar{\theta}_n(x_1,...,x_n)$, зависящая от элементов выборки, приближенно равная:

$$\bar{\theta}_n(x_1,...,x_n) \approx \theta \qquad (1.1)$$

Для каждой конкретной выборки – это число, то есть точка на числовой оси.

Определение 2. *Статистикой* называется любая функция выборочных элементов (наблюдений).

Таким образом, статистическая точечная оценка – это статистика, по значениям которой можно судить о величине θ.

Для одной и той же неизвестной величины θ можно составить бесконечно много различных оценок. Например, в качестве оценки математического ожидания m нормального распределения могут служить выборочное среднее $\bar{x}$, выборочная медиана *med*, полусумма квартилей t_q, полусумма крайних элементов t_R.

В силу многообразия оценок, применяемых для оценивания одной и той же неизвестной величины, возникает задача выбора из них лучшей в определенном смысле. К оценкам предъявляются требования.

Заметим предварительно, что все статистические оценки являются случайными, так как случайными являются элементы выборки.

Определение 3. Оценка $\bar{\theta} \approx \bar{\theta}_n(x_1,...,x_n)$ называется *состоятельной оценкой* θ, если она стремится по вероятности к θ с ростом n:

$$\bar{\theta}_n(x) \xrightarrow[n \to \infty]{P} \theta \tag{1.2}$$

Это означает, что для любого $\varepsilon > 0$ выполняется соотношение

$$\lim_{n\to\infty} P(|\bar{\theta}_n - \theta| \geq \varepsilon) = 0 \tag{1.3}$$

Это требование означает сближение $\bar{\theta}_n$ и θ с ростом n в вероятностном смысле. В математической статистике, как правило, применяются только состоятельные оценки.

<u>Пример 1.</u> Из предельной теоремы Бернулли теории вероятностей следует, что относительная частота *P* (A)* события *A* является состоятельной оценкой вероятности *P(A)* этого события: $P^*(A) \xrightarrow[n \to \infty]{P} P(A)$

Определение 4. Оценка $\overline{\theta_n}$ называется *несмещённой оценкой* θ, если математическое ожидание оценки равно θ :

$$\mathbf{M}(\overline{\theta_n}) = \theta \quad . \qquad (1.4)$$

В противном случае оценка называется *смещённой*. Разность между ними называется *смещением оценки*.

Требование несмещённости означает, что выборочные значения $\widehat{\theta}_{n,i}$ оценок, полученные в результате повторения выборок, группируются не только около их математического ожидания, но и около оцениваемой величины θ (рис. 1).

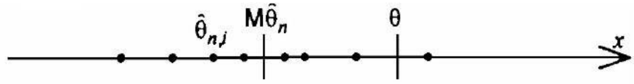


Рис. 1 – Группировка выборочных значений $\widehat{\theta}_{n,i}$ смещённой оценки $\overline{\theta_n}$ около своего математического ожидания $\mathbf{M}(\widehat{\theta}_n)$, а не около оцениваемой величины θ.

Определение 5. Оценка $\overline{\theta_n}$ величины θ называется *робастной*, если она устойчива по отношению к выбросам в статистических данных.

Выбросы в выборке могут появиться вследствие сбоев регистрирующего прибора, грубых ошибок оператора.

Выбросы группируются на концах вариационного ряда наблюдений. Поэтому оценки, не имеющие в своем составе элементов, близких к концам вариационного ряда, будут робастными. Это, например, выборочная медиана *med* и полусумма квартилей t_q.

Определение 6. Оценка $\widehat{\theta}^*{}_n$ числовой характеристики или параметра θ распределения называется *эффективной* в рассматриваемом классе Т состоятельных и несмещенных оценок, если она имеет в этом классе минимальную дисперсию:

$$\mathbf{D}\widehat{\theta}_n^* = \min_T \mathbf{D}\widehat{\theta}_n \qquad (1.5)$$

Замечание. Для рассматриваемого распределения и рассматриваемого класса оценок Г эффективная оценка может не существовать, а дается лишь определить нижнюю грань дисперсий оценок $\inf\limits_T \mathbf{D}\bar{\theta}_n$. Тогда возникнет задача построения оценок, дисперсии которых будут возможно ближе к этой грани.

Определение 7. Из двух оценок $\widehat{\theta}_{1n}$ и $\widehat{\theta}_{2n}$ одной и той же числовой

характеристики или параметра θ распределения в классе Т состоятельных и несмещенных оценок *более эффективной* считается та, дисперсия которой меньше.

Если имеет место неравенство

$$\mathbf{D}\bar{\theta}_{1n} < \mathbf{D}\bar{\theta}_{2n} \qquad (1.6)$$

то $\widehat{\theta}_{1n}$ – более эффективная оценка θ, чем $\widehat{\theta}_{2n}$. Отношение

$$\mathbf{D}\bar{\theta}_{1n} / \mathbf{D}\bar{\theta}_{1n} \qquad (1.7)$$

называется *относительной эффективностью* оценки $\widehat{\theta}_{2n}$ относительно

оценки $\widehat{\theta}_{1n}$, а отношение

$$\inf_T \mathbf{D}\bar{\theta}_n / \mathbf{D}\bar{\theta}_{2n} = eff\,\bar{\theta}_{2n} \qquad (1.8)$$

называется *эффективностью* оценки $\widehat{\theta}_{2n}$ в рассматриваемом классе оценок *Т*.

<u>Пример 2.</u> Для нормального распределения $N(m,\sigma)$ оценкой математического ожидания m могут служить выборочное среднее x и выборочная медиана med в силу симметричности нормального распределения.

Доказано, что $\mathbf{D}\bar{x} = \sigma^2 / n$ (для любого *n)* $\mathbf{D}med \approx \frac{\pi}{2} \cdot \frac{\sigma^2}{n}$ (при больших *n).* Следовательно, при больших n относительной эффективностью выборочной медианы относительно $\bar{x}$ будет $\mathbf{D}\bar{x} / \mathbf{D}med = 2/\pi \approx 0.64$.

Определение 8. Оценка $\hat{\theta}_n$ параметра θ распределения называется *асимптотически эффективной* в классе Т состоятельных и несмещённых оценок, если существует предел

$$\lim_{n\to\infty} eff\,\bar{\theta}_n = 1 \tag{1.9}$$

Асимптотически эффективные оценки дает метод максимального правдоподобия получения оценок.

В более общем случае, если отказаться от требования несмещенности оценки $\hat{\theta}_n$ относительно θ вместо дисперсии $\mathbf{D}\hat{\theta}_n$ обычно выбирается величина *среднего квадрата ошибки,* то есть второй момент вида $\mathbf{M}[(\hat{\theta}_n - \theta)^2]$. Тогда оценка $\hat{\theta}^*_n$ называется эффективной в классе T состоятельных оценок, если выполняется равенство

$$\mathbf{M}[(\bar{\theta}_n - \theta)^2] = \inf_T \mathbf{M}[(\bar{\theta}_n - \theta)^2] \tag{1.10}$$

Отношение

$$eff = \inf_T \mathbf{M}[(\bar{\theta}_n - \theta)^2] / \mathbf{M}[(\bar{\theta}_{1n} - \theta)^2] \tag{1.11}$$

называется *эффективностью оценки* $\hat{\theta}_{1n}$ в классе Т состоятельных оценок.

14.2. Свойства оценок. Методы получения оценок параметров генерального распределения. Свойства выборочного среднего и выборочной дисперсии

1°. Свойства $\bar{x}$. Выборочное среднее $\bar{x}$ определяется как $\left(\sum n_i x_i\right)/n$. Если элементы x_i большие, то для упрощения расчетов целесообразно вычесть из каждого элемента одно и то же число C (близкое к выборочному среднему), т. е. перейти к условным элементам $u_i = x_i - C$.

Тогда $\bar{x} = X + \left(\sum n_i x_i\right)/n$.

Свойство 1. Выборочное среднее $\bar{x}$ является состоятельной оценкой генерального математического ожидания $m = \mathbf{M}X$, что следует из предельной теоремы Чебышева:

$$\bar{x} \xrightarrow[n \to \infty]{P} m. \qquad (2.1)$$

Свойство 2. $\bar{x}$ является несмещенной оценкой *m:*

$$\mathbf{M}\bar{x} = m. \qquad (2.2)$$

Свойство 3. $\bar{x}$ не является робастной оценкой *m,* так как в своем составе имеет крайние элементы вариационного ряда.

Свойство 4.

$$\mathbf{D}\bar{x} = \frac{\sigma^2}{n}. \qquad (2.3)$$

$$\mathbf{D}\bar{x} = D\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{D}x_i = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

Этот результат означает, что с ростом *n* рассеяние $\bar{x}$ уменьшается обратно пропорционально *n.*

2°. Аналогично доказывается, что выборочный начальный момент α_l порядка l также является состоятельной и несмещенной оценкой генерального начального момента $\alpha_l \mathbf{M}X^l$ порядка l:

$$a_l \xrightarrow[n \to \infty]{\mathbf{P}} \alpha_l \quad . \tag{2.4}$$

$$\mathbf{M}a_l = \alpha_l \, . \tag{2.5}$$

3°. Свойства выборочной дисперсии s^2.

Выборочная дисперсия определяется как $s^2 = \dfrac{1}{n}\sum n_i (x_i - \bar{x})^2$. Если элементы x_i – большие, то для упрощения расчетов целесообразно вычесть из каждого элемента одно и то же число C (равное выборочному среднему или близкое к нему), т. е. перейти к условным элементам $u_i = x_i - C$. Дисперсия при этом не изменится. Если элементы x_i – являются десятичными дробями с k десятичными знаками после запятой, то, чтобы избежать действий с дробями, умножают первоначальные элементы на постоянное число С = 10^k, т. е. переходят к условным элементам $u_i = Cx_i$. При этом дисперсия увеличится в C^2 раз. Поэтому, найдя дисперсию условных элементов, надо разделить ее на C^2.

Свойство 1. Выборочная дисперсия является состоятельной оценкой генеральной дисперсии:

$$s^2 \xrightarrow[n \to \infty]{\mathbf{P}} \sigma^2 = \mathbf{D}X \, . \tag{2.6}$$

Свойство 2. Вспомогательная формула выборочной дисперсии

$$s^2 = \frac{\sum n_i x_i^2}{n} - \left[\frac{\sum n_i x_i}{n} \right]^2 . \tag{2.7}$$

Свойство 3. Выборочная дисперсия s^2 – смещённая оценка генеральной дисперсии σ^2 с отрицательным смещением:

$$\mathbf{M}s^2 = \sigma^2 - \frac{\sigma^2}{n}. \qquad (2.8)$$

Вследствие смещенности выборочной дисперсии возникает задача создания несмещенной оценки дисперсии.

Так как $\mathbf{M}s^2 = \sigma^2$, то смещение можно устранить, умножив s^2 на множитель $\frac{n-1}{n}$.

Образуем исправленную выборочную дисперсию:

$$s^{*2} = \frac{n}{n-1}s^2 = \frac{1}{n-1}\sum n_i(x_i - \bar{x})^2 \qquad (2.9)$$

s^{*2} является несмещенной оценкой σ^2. Действительно,

$$\mathbf{M}s^{*2} = \frac{n}{n-1}\mathbf{M}s^2 = \frac{n}{n-1}\cdot\frac{n-1}{n}\sigma^2 = \sigma^2.$$

Отметим, что s^2 не является робастной оценкой σ^2.

Более удобной является формула:

$$s^{*2} = \frac{\sum n_i x_i^2 - \frac{\left[\sum n_i x_i\right]^2}{n}}{n-1}.$$

14.3. Метод моментов получения оценок параметров генерального распределения

Пусть известен вид генерального закона распределения, а параметры $\theta_1, \theta_2, ..., \theta_n$, в него входящие, неизвестны. Возникает задача их статистического оценивания. *Метод моментов* Пирсона (К. Пирсон – англ. математик, 1857–1936) – один из первых методов получения таких оценок, основанный на сравнении выборочных и генеральных моментов распределения. Идейно он очень прост.

Предполагается, что имеется выборка (x_1 ..., x_n) из исследуемой генеральной совокупности. На ее основе вычисляются m начальных моментов $a_1, \ldots, a_m$,. Так как вид генерального закона известен, то, следовательно, можно найти m первых начальных генеральных моментов $\alpha_1(\theta_1,...,\theta_m),...,\alpha_m(\theta_1,...,\theta_m)$, которые выражаются через неизвестные параметры. Выборочные и генеральные моменты одинакового порядка приравниваются:

$$\begin{cases} \alpha_1(\theta_1,...,\theta_m) = a_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_m(\theta_1,...,\theta_m) = a_m. \end{cases} \tag{2.10}$$

Получили систему m уравнений с неизвестными величинами $\theta_1,...,\theta_m$. Решение $(\widehat{\theta}_1,...,\widehat{\theta}_m)$ этой системы дает оценки $\widehat{\theta}_i = \widehat{\theta}_i(x_i,...,x_n)$ неизвестных параметров $\theta_i = (i,...,m)$. При выполнении достаточно общих условий полученные оценки состоятельные:

$$\bar{\theta}_i \xrightarrow{P} \theta_i \qquad (i = 1, \ldots, m) \tag{2.11}$$

Среднее значение такой оценки отличается от истинного значения параметра на величину порядка $1/n$. В общем случае они – смещенные и не являются эффективными и асимптотически эффективными.

Замечание 1. Вместо начальных моментов можно использовать центральные.

<u>Пример 1.</u> ► Для показательного закона с плотностью

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

известно, что $\alpha_1 = MX = 1/\lambda$. Так как $a_1 = a_1 = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, то система (2.10) в этом случае сводится к одному уравнению $1/\lambda = \bar{x}$, из которого находим $\bar{\lambda} = 1/\bar{x}$ (2.12)◄

Пример 2. Для нормального закона $N(m,\sigma)$ известно, что $\alpha_1 = MX = m; \mu_2 = M[(X - m_X)^2] = \sigma^2$. Для этого случая удобно взять первый начальный и второй центральный моменты. Получаем систему из двух уравнений: $\begin{cases}\alpha_1 = a_1, \\ \mu_2 = m_2\end{cases} \Leftrightarrow \begin{cases} m = \bar{x} \\ \sigma^2 = s^2.\end{cases}$

Находим оценки двух параметров по методу моментов:

$$\overset{\frown}{m} = \bar{x}; \overset{\frown}{\sigma} = s \tag{2.13}$$

Пример 3. Для равномерного закона, определяемого плотностью

$$f(x) = \begin{cases} 1/(b-a), x \in [a,b]; \\ 0, x \notin [a,b], \end{cases}$$

неизвестными параметрами являются *a, b*. Известно, что $\alpha_1 = m_X (a+b)/2; \mu_2 = D_X = (b-a)^2/12$. Для нахождения оценок $\hat{a}, \hat{b}$ составляем систему уравнений

$$\begin{cases}\alpha_1 = a_1, \\ \mu_2 = m_2\end{cases} \Leftrightarrow \begin{cases}(a+b)/2 = \bar{x}, \\ (a+b)^2/12 = s^2\end{cases} \Leftrightarrow \begin{cases} b + a = 2\bar{x}, \\ b - a = 2\sqrt{3}\cdot s,\end{cases}$$ из

которой находим

$$\overset{\frown}{a} = \bar{x} - s\sqrt{3}; \overset{\frown}{b} = \bar{x} + s\sqrt{3} \tag{2.14}$$

14.4. Метод максимального правдоподобия получения оценок параметров генерального распределения

Метод максимального правдоподобия, созданный Фишером (Р. Фишер – англ. математик, 1890–1962), является достаточно универсальным и плодотворным методом оценивания.

Пусть имеется выборка $(x_1,...,x_n)$ из генеральной совокупности с плотностью вероятности *f(x, θ),* содержащей один неизвестный параметр θ.

Выборка является *n*-мерной случайной величиной, компоненты x_t которой взаимно независимы, одинаково распределены с плотностью *f(x, θ).* Тогда плотность распределения *n*-мерной случайной величины $(x_1, x_2,..., x_n)$ будет равна

$$L(x_1, x_2,..., x_n;\theta) = f(x_1,\theta) f(x_2,\theta)...f(x_n,\theta). \quad (2.15)$$

Эта функция называется *функцией правдоподобия* для рассматриваемой выборки.

Будем считать θ переменной неслучайной величиной, а элементы $x_1, x_2,..., x_n$ выборки фиксированными, так как выборка фактически осуществлена. Если придавать θ различные значения, то естественно ожидать, что плотность $L(x_1, x_2,..., x_n;\theta)$ примет максимальное значение в случае, когда θ окажется равным истинному его значению, так как при других значениях θ менее вероятно за один раз получить именно данную выборку.

Эти интуитивные соображения приводят к тому, что за оценку θ берут такое значение $\widehat{\theta}$, при котором функция правдоподобия достигает максимума.

Технически (так как *L* состоит из произведений) удобнее искать *max*In*L* (точка *в*, дающая максимум In *L*, даёт и максимум *L)*. Итак, для отыскания $\widehat{\theta}$ имеем уравнение

$$\frac{\partial \ln L}{\partial \theta} = 0, \qquad (2.16)$$

которое называется *уравнением правдоподобия,* а его решение $\widehat{\theta} = \widehat{\theta}(x_1, x_2, ..., x_n)$, зависящее от элементов выборки, *оценкой максимального правдоподобия.*

При выполнении достаточно общих условий оценки максимального правдоподобия являются состоятельными и асимптотически эффективными. В общем случае они являются смещенными.

В случае, когда генеральная плотность вероятности $f(x, \theta_1, ..., \theta_k)$ содержит k параметров, вместо одного уравнения правдоподобия решается система уравнений

$$\frac{\partial \ln L}{\partial \theta_1} = 0, ..., \frac{\partial \ln L}{\partial \theta_k} = 0 \qquad (2.17)$$

<u>Пример 1.</u> Рассмотрим показательный закон с плотностью

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

Функция правдоподобия при x > 0 имеет вид:

$$L(x_1, ..., x_n, \lambda) = \lambda^n e^{-\lambda x_1} ... e^{-\lambda x_n} = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right);$$

$$\ln L = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i ; \frac{\partial \ln L}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

Отсюда: $\dfrac{1}{\lambda} = \dfrac{1}{n} \sum_{i=1}^{n} x_i = \bar{x};$

$$\overset{\frown}{\lambda} = 1 / \bar{x} \,. \qquad (2.18)$$

Пример 2. ►Для нормального закона *N(m, σ)* плотность вероятности имеет вид
$$f(x,m,\sigma^2)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{(x-m)^2}{2\sigma^2}}.$$

Удобно считать, что здесь два параметра тис. Следовательно, функция правдоподобия равна
$$L=(\frac{1}{\sigma 2\pi})^n e^{\frac{(x_1-m)^2}{2\sigma^2}-\ldots-\frac{(x_n-m)^2}{2\sigma^2}}.$$

Тогда $\ln L=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}(\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2$.

Далее, дифференцируя ln *L* по *m* и σ^2, получаем систему уравнений правдоподобия
$$\left.\begin{aligned}\frac{\partial\ln L}{\partial m}&=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-m)=0,\\ \frac{\partial\ln L}{\partial\sigma^2}&=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-m)^2=0.\end{aligned}\right\}$$

Из первого уравнения находим $\sum_{i=1}^{n}x_i-nm=0$. Отсюда
$$\overset{\frown}{m}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}. \tag{2.19}$$

Из второго уравнения:
$$\begin{gathered}\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2=n;\\ (\overset{\frown}{\sigma^2})=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=s^2\end{gathered} \tag{2.20}$$

Эти оценки мы получили ранее методом моментов◄

Пример 3. Для равномерного закона плотность вероятности равна:

$$f(x,a,b)=\begin{cases}1/(b-a), x\in[a,b];\\ \\ 0, \qquad\qquad ,x\notin[a,b].\end{cases}$$

Функция правдоподобия для $x\in[x,b]$ имеет вид:

$$L=\frac{1}{(b-a)^n}\le\frac{1}{(x_{\max}-x_{\min})^n} \quad . \qquad (2.21)$$

так как $b-a\ge x_{\max}-x_{\min}$. Из неравенства (2.13) следует, что функция L принимает наибольшее значения при $b = x_{max}$ и $a = x_{min}$. Таким образом, оценками максимального правдоподобия в случае равномерного закона являются

$$\overset{\frown}{a}=x_{\min},\overset{\frown}{b}=x_{\max}. \qquad (2.22)$$

Они отличаются от оценок (2.5) метода моментов.

Замечание 1. В случае дискретного закона распределения $P(X=x_i)=p(x_i,\theta)$ функция правдоподобия определяется формулой:

$$L=\prod_{i=1}^{n}p(x_i,\theta) \qquad (2.23)$$

.

<u>Пример 4.</u> Найти оценку максимального правдоподобия параметра a закона Пуассона $P(X=k)-a^k e^{-a}/k!$

$$L=\frac{a^{x_1}}{x_1!}e^{-a}\dots\frac{a^{x_n}}{x_n!}e^{-a}=e^{-na}a^{\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}(x_i!)^{-1}$$

.

$$\ln L=-na+\ln a\cdot\left(\sum_{i=1}^{n}x_i\right)-\sum_{i=1}^{n}\ln(x_i!);\frac{\partial\ln L}{\partial a}=-n+\frac{1}{a}\sum_{i=1}^{n}x_i=0$$

.

$$\frac{1}{a}\sum_{i=1}^{n}x_i=n;a=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x};\overset{\frown}{a}=\bar{x}$$

.

Замечание 2. Ранее были рассмотрены два наиболее употребительных на практике метода получения оценок параметров закона распределения – методы моментов и максимума правдоподобия. Существуют и другие методы, освещенные в литературе. Назовем еще методы квантилей, минимума хи-квадрат, наименьших квадратов, наименьших абсолютных отклонений, минимакса.

§15. Интервальное оценивание числовых характеристик и параметров распределения генеральной совокупности

Точечные оценки, рассмотренные в предыдущих параграфах, хотя и являются численными, не дают желательной информации об оцениваемых генеральных характеристиках. Если, например, $\bar{x} = 10$, то совершенно неясно, насколько точно число 10 оценивает неизвестное математическое ожидание m. Мы лишь знаем некоторые качественные свойства, что $\bar{x}$ – хорошая оценка по сравнению с другими возможными. А следовало бы связать точечную оценку с объемом выборки, выработать показатели ее точности и надежности. Эти вопросы решаются в теории интервального оценивания.

15.1. Доверительный интервал. Точность и надежность оценки

Пусть θ – неизвестная числовая характеристика или параметр генерального распределения.

Определение 1. Если выполняется соотношение

$$P(\theta_1 < \theta < \theta_2) = \gamma, \tag{1.1}$$

то интервал (θ, θ_2) называется *доверительным интервалом*, который накрывает неизвестную генеральную характеристику θ с доверительной вероятностью γ.

Здесь $\theta_1 = \theta_1(x_1,...,x_n), \theta_2 = \theta_2(x_1,...,x_n)$ – известные функции выборочных элементов $x_1, \ldots, x_n$, то есть статистики. Они вычисляются по выборке.

Число γ называется также *надежностью,* с которой доверительный интервал накрывает θ. Число α = 1-γ называется *уровнем значимости.*

Статистики θ_1 и θ_1 в соотношении (1.1) являются точечными оценками θ. Одна дает левую, а другая – правую границы, между которыми содержится θ с надежностью γ.

Половину длины доверительного интервала

$$\varepsilon = (\theta_2 - \theta_1)/2 \tag{1.2}$$

назовем *точностью интервального оценивания.*

Пусть теперь известна одна точечная оценка $\widehat{\theta}$ генеральной числовой характеристики или параметра распределения θ.

Определение 2. Если выполняется соотношение

$$P(|\theta - \bar{\theta}|) < \varepsilon = \gamma \tag{1.3}$$

то число ε называется точностью, а число γ – надежностью оценки $\widehat{\theta}$ генеральной числовой характеристики θ.

Здесь $\widehat{\theta} = \widehat{\theta}\,(x_1, \ldots, x_n)$ – статистика, то есть функция выборочных элементов.

Если известны ε и γ, то легко построить доверительный интервал для θ с помощью ее точечной оценки $\widehat{\theta}$.

Действительно,

$$|\,\theta-\bar{\theta}\,|<\varepsilon \Leftrightarrow -\varepsilon<\theta-\bar{\theta}<\varepsilon \Leftrightarrow \bar{\theta}-\varepsilon<\theta<\bar{\theta}+\varepsilon\,.$$

Тогда $\theta_1 = \hat{\theta} - \varepsilon$; $\theta_2 = \hat{\theta} + \varepsilon$, и мы от соотношения (1.3) приходим к соотношению (1.1).

Как находить ε, γ, строить доверительный интервал (θ_1, θ_2) в конкретных случаях будет рассмотрено далее. Мы эти вопросы рассмотрим для практически наиболее важных случаев оценивания: вероятности события *p,* математического ожидания *m* и среднего квадратичного отклонения σ.

15.2. Точность и надежность оценивания вероятности события с помощью его относительной частоты

Пусть *p* – вероятность события *A,* $p^* = \dfrac{\mu}{n}$ – его относительная частота. По теореме Муавра-Лапласа теории вероятностей при больших *n* справедливо приближенное равенство

$$P\left(\frac{\mu-np}{\sqrt{npq}}<x\right)\approx \Phi(x), \tag{2.1}$$

где

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt \tag{2.2}$$

– функция Лапласа.

Из формулы (2.1) находим

$$P\left(\frac{|\,\mu-np\,|}{\sqrt{npq}}<x\right)=P\left(-x<\frac{\mu-np}{\sqrt{npq}}<x\right)\approx \Phi(x)-\Phi(-x)=\Phi(x)-[1-\Phi(x)$$

Отсюда

$$P(|\, p - \frac{\mu}{n}\,| < x\sqrt{\frac{pq}{n}}) \approx 2\Phi(x) - 1 \qquad (2.3)$$

Из формулы (2.3) видим, что в наших построениях p отличается от $p^* = \eta / n$ на величину порядка $1/\sqrt{n}$. Так как p неизвестно, то его заменяем на p^*, а q соответственно на $q^* = 1 - p^*$. Это означает, что под корнем в формуле (2.3) мы пренебрегаем малыми слагаемыми порядка $1/(n\sqrt{n})$. Получаем формулу

$$P(|\, p - p^*\,| < x\sqrt{\frac{p^*(1-p^*)}{n}}) \approx \Phi(x) - 1 \qquad (2.4)$$

Полагаем

$$\sqrt[x]{\frac{p^*(1-p^*)}{n}} = \varepsilon; 2\Phi(x) - 1 = \gamma$$

Отсюда

$$\Phi(x) = (1+\gamma)/2 \qquad (2.5)$$

Решая уравнение (2.5), находим его корень

$$u_{(1+\gamma)/2} = \Phi^{-1}((1+\gamma)/2) \qquad (2.6)$$

– квантиль нормального распределения N (0,1) порядка $(1+\gamma)/2$.

Тогда

$$\varepsilon = u_{(1+\gamma)/2}\sqrt{\frac{p^*(1-p^*)}{n}}\,. \qquad (2.7)$$

Формулы (2.5), (2.7) связывают три величины ε, γ, n. Задавая две из них, можно найти третью. Тем самым будет построен доверительный интервал для неизвестной вероятности p:

$$\mathbf{P}(p^* - \varepsilon < p < p^* + \varepsilon) = \gamma\,. \qquad (2.8)$$

<u>Пример 1.</u> Заданы n = 1600 и γ = 0.95. Требуется найти ε и построить доверительный интервал для вероятности p с помощью найденной по выборке относительной частоты $p^* = 0.2$.

Решая уравнение (2.5) $\Phi(x) = (1+\gamma)/2 = (1+0.95)/2 = 0.975$ с помощью таблицы квантилей нормального распределения, получим $u_{(1+\gamma)/2} = u_{0.975} = 1.96$. Далее по формуле (2.7) находим

$$\varepsilon = u_{0.975}\sqrt{\frac{0.2\cdot 0.8}{1600}} = u_{0.975}\frac{0.4}{40} = 0.01u_{0.975} = 0.01\cdot 1.96 \approx 0.02$$

Получаем доверительный интервал для неизвестной вероятности p:

$$\mathbf{P}(0.18 < p < 0.22) = 0.95 \quad . \tag{2.9}$$

Замечание 1. Решенной задаче 1 может быть придано, например, следующее реальное содержание. В результате проведенного социологического опроса n = 1600 человек рейтинг кандидата N в президенты составляет 20%. Тогда доверительный интервал (2.9) позволяет утверждать, что с надежностью γ = 0.95 действительный рейтинг кандидата N заключен в пределах 18% – 20%. Этот результат можно выразить и иначе: рейтинг N равен 20% ± 2% с 5-процентной ошибкой.

Замечание 2. Вероятность p, оцениваемая с помощью доверительного интервала (2.8) и точечной оценки p^*, является параметром биноминального закона распределения случайной величины X: $\mathbf{P}(X = k) = C_m^k p^k (1-p)^{m-k}, k = 0,1,\ldots,m.$

15.3. Доверительный интервал для математического ожидания *m* нормальной генеральной совокупности

Известно, что для выборки объема *n* из нормальной генеральной совокупности случайная величина $\sqrt{n-1}\dfrac{\overline{x-m}}{s}$ распределена по закону Стьюдента с *n*-1 степенями свободы. Здесь *s* – выборочное среднее квадратическое отклонение. Так как плотность $f_{n-1}(x)$ этого распределения –функция четная, то получаем:

$$P(-x<\sqrt{n-1}\frac{\bar{x}-m}{s}<x)=P(-x<\sqrt{n-1}\frac{m-\bar{x}}{s}<x)=\int_{-x}^{x}f_{n-1}(t)dt=2\int_{0}^{x}f_{n-1}(t)dt=2\int_{-\infty}^{x}f_{n-1}(t)dt-1=2F_{n-1}(x)-1$$

Здесь $F_{n-1}(x)$ – функция распределения закона Стьюдента с *n*-1 степенями свободы. Отсюда находим

$$P\left(\bar{x}-\frac{sx}{\sqrt{n-1}}<m<\bar{x}+\frac{sx}{\sqrt{n-1}}\right)=2F_{n-1}(x)-1 \qquad (3.1)$$

Полагаем 2 $F_{n-1}(x)$ -1 = γ. Тогда

$$F_{n-1}(x)=(1+\gamma)/2 \qquad (3.2)$$

По таблице квантилей распределения Стьюдента с *n*-1 степенями свободы находим квантиль порядка $(1+\gamma)/2$

$$t_{(1+\gamma)/2}(n-1)=F_{n-1}^{-1}((1+\gamma)/2) \qquad (3.3)$$

и получаем искомый доверительный интервал для *m:*

$$P\left(\bar{x}-\frac{st_{(1-\gamma)/2}(n-1)}{\sqrt{n-1}}<m<\bar{x}+\frac{st_{(1+\lambda)/2}(n-1)}{\sqrt{n-1}}\right)=\lambda \qquad (3.4)$$

<u>Пример 1.</u> По выборке объема *n = 20* из нормальной генеральной совокупности найдены $\bar{x}$ = 5.00 и *s* = 0.25. Требуется построить доверительный интервал для *m* при γ = 0.95. Число *m* можно интерпретировать как среднее значение контролируемого параметра производимого продукта. Контроль производится по результатам 20 измерений.

С помощью таблицы квантилей распределения Стьюдента решаем уравнение (3.2): $F_{19}(x) = (1+0.95)/2 = 0.975$. Получаем $t_{0.975}(19) = 2.09$. Формула 3.4 дает:

$$5.00 - \frac{0.25 \cdot 2.09}{\sqrt{19}} < m < 5.00 + \frac{0.25 \cdot 2.09}{\sqrt{19}} \Leftrightarrow 4.88 < m < 5.12.$$

Итак, P(4.88<m<5.12)=0.95.

§16. Статистические гипотезы

При проведении статистических исследований возникают различные вопросы о свойствах генерального распределения и выборки. Для ответов на эти вопросы выдвигаются гипотезы, требующие статистической проверки на основе полученной выборки.

Эти гипотезы могут быть выдвинуты непосредственно практикой, а могут возникнуть как дальнейший этап статистических исследований после разведочного анализа, обеспеченного описательной статистикой.

16.1. Виды статистических гипотез

Приведём примеры наиболее важных в практическом отношении гипотез.

1. Гипотеза о равенстве математических ожиданий двух генеральных совокупностей.
2. Гипотеза о равенстве дисперсий нескольких генеральных совокупностей.
3. Гипотеза о законе распределения генеральной совокупности. Эта гипотеза может возникнуть на основе теоретических соображений, имеющегося опыта исследований, на основе изучения гистограммы выборки.
4. Гипотеза об однородности выборки, об

отсутствии в ней выбросов.

Определение 1. *Статистической гипотезой* называется предположение о виде или свойствах генерального или выборочного распределений, которое можно проверить статистическими методами на основе имеющейся выборки.

Определение 2. *Статистическая гипотеза* о генеральном распределении называется простой, если она его полностью определяет. В противном случае гипотеза называется сложной.

Как правило, гипотезы о генеральном распределении – сложные.

16.2. Критерий значимости

Для проверки любой статистической гипотезы выбирается какой-либо критерий, называемый критерием значимости.

Определение 1. *Критерием значимости* называется правило проверки статистической гипотезы.

Выдвинутую гипотезу проверяют на основе имеющейся выборки. Для этого конструируется функция выборочных элементов, называемая статистикой, по величине которой судят о справедливости гипотезы.

Определение 2. *Статистикой критерия значимости* называется статистика, по значениям которой судят о справедливости статистической гипотезы. Часто ее для простоты тоже называют критерием.

Например, для проверки гипотезы о том, что вероятность интересующего нас события А равна *p,* можно взять статистику

$$Z = \frac{\mu}{n} - p\,, \qquad (2.1)$$

являющуюся отклонением относительной частоты μ / n от вероятности события А.

Если гипотеза верна, то при увеличении n относительная частота будет приближаться к p по вероятности, а следовательно, Z будет стремиться к нулю. При большом n маловероятно, что Z будет сильно отличаться от нуля. В этом примере и в общем случае следует знать закон распределения статистики критерия, чтобы судить, какие ее значения маловероятны, а какие – нет.

В основе большинства критериев значимости лежит следующий простой принцип: если сделана гипотеза о том, что событие имеет очень малую вероятность, но в результате одного лишь испытания это событие произошло, то следует подвергнуть сомнению справедливость выдвинутой гипотезы.

События с малой вероятностью *а,* которой в данной ситуации можно пренебречь, будем называть *практически невозможными,* а с вероятностью 1-*а*, близкой к единице, – *практически достоверными.*

Вероятности *а* и 1-α абстрактно выбрать нельзя. Их значения диктуются реальной ситуацией. Например, если α – вероятность нераскрытая парашюта или разрушения дорогостоящей плотины паводком, то α должно быть десятичной дробью с большим числом нулей после запятой. Это число обычно стандартизируется мировой практикой.

Определение 3. *Уровнем значимости* α называется столь малая вероятность, что событие с такой вероятностью является практически невозможным.

Обычно проверяемая гипотеза обозначается H_0, а ей альтернативная – H_1. Например, если вероятность брака (событие *А)* равна *p,* а после усовершенствования технологического процесса ожидается, что она будет меньше, то в качестве H_0 можно взять гипотезу: $\mathbf{P}(A) = p$, а в качестве $\mathbf{P}(A) < p$.

Если сформулированы гипотезы H_0 и H_1, выбрана статистика критерия Z, то следует указать еще область V_k маловероятных значений Z, попадание в которую статистики Z заставляет нас отвергнуть H_0 и принять H_1.

Определение 4. *Критической областью критерия значимости* называется подобласть V_k области V значение статистики Z, вероятность попадания в которую для этой статистики при условии истинности проверяемой гипотезы H_0 равна уровню значимости α:

$$\mathbf{P}(Z \in V_k / H_0) = \alpha. \quad (2.2)$$

Дополнительная область $V \backslash V_k$ называется *областью допустимых значений* статистики критерия. Если $Z \in (V \backslash V_k)$, то гипотеза H_0 при заданном уровне значимости a принимается. Обычно говорят более осторожно: H_0 не противоречит имеющейся выборке, то есть гипотеза H_0 правдоподобна.

Область V_k можно выбрать неоднозначно. Однако, зная закон распределения случайной величины Z, хотя бы асимптоматический, то есть при большом объеме выборки n, и, налагая на V_k дополнительные условия, можно критическую область найти однозначно, задав величину α.

16.3. Общая схема проверки статистических гипотез

1. Выдвигаются проверяемая и альтернативная гипотезы H_0, H_1.

2. Выбирается уровень значимости α (обычно 0.001; 0.01; 0.05; 0.1).

3. Выбирается статистика Z критерия значимости и соответствующая ей, уровню значимости и проверяемым

гипотезам H_0 и H_1 критическая область V_k, являющаяся частью области V значений статистики Z. При этом $V \setminus V_k$ будет областью допустимых значений Z.

4. Вычисляется выборочное значение Z_B статистики Z.

5. Формулируется критерий проверки. Если $Z_B \in V_k$, то гипотеза H_0 отвергается, так как в результате одного лишь испытания, получения выборки произошло практически невозможное событие $Z_B \in V_k$ с вероятностью α. Если $Z \in (V \setminus V_k)$, то гипотеза H_0 принимается.

Определение 1. *Критерием согласия* называется критерий значимости, применяемый для проверки гипотезы о генеральном законе распределения.

Заметим, что существуют и другие схемы проверки статистических гипотез.

16.4. Ошибки первого и второго рода

Суждения о принятии или отвержении выдвинутой статистической гипотезы не являются абсолютными, а носят лишь вероятностный характер, то есть являются правдоподобными. Принимая или отвергая гипотезу, мы можем совершить ошибку.

Определение 1. *Ошибкой первого рода* называется ошибка отвержения правильной гипотезы. *Ошибкой второго рода* называется ошибка принятия неверной гипотезы.

Вероятность ошибки первого рода равна уровню значимости, то есть

$$\alpha = \mathbf{P}(Z \in V_k / H_0) \qquad (3.1)$$

Эта формула означает, что гипотеза H_0 отвергается с вероятностью α, хотя эта гипотеза верна.

Вероятность ошибки второго рода обозначается β:

$$\beta = \mathbf{P}(Z \in V \setminus V_k / H_1) \qquad (3.2)$$

Формула (3.2) означает, что принимается гипотеза H_0, с вероятностью β, хотя верна альтернативная гипотеза H_1.

При той схеме проверки гипотез, которая сформулирована в вопросе 2, вероятность α задается. Вероятность же β приходится находить. Это удается в редких случаях, так как для этого нужно знать распределение статистики *Z* для случая альтернативной гипотезы H_1.

Принципы назначения уровня значимости α при проверке статистической гипотезы согласуются с опасностью совершения ошибок первого и второго рода. Эти принципы вообще находятся вне статистики. Они выдвигаются практикой.

Для того, чтобы проверяемая гипотеза была достаточно обоснованно отвергнута, уровень значимости выбирают достаточно малым; в практике: 0.01; 0.001. Напротив, если делается вывод о принятии гипотезы, то уровень значимости не должен быть очень малым, ибо в этом случае расширяется область допустимых значений $V \setminus V_k$ и даже при неверной гипотезе статистика *Z* критерия может попасть в эту область за счет случайных колебаний. Будет совершена ошибка второго рода. Уровень значимости в этом случае можно взять равным 0.05; 0.10. Чем меньше уровень значимости, тем меньше вероятность забраковать верную гипотезу, т. е. совершить ошибку первого рода, но при этом увеличивается вероятность принятия неверной гипотезы, т. е. совершения ошибки второго рода.

§31. Односторонний и двусторонний критерии

Пусть известен закон распределения статистики критерия *Z* (хотя бы асимптотический). Будем предполагать, что известна плотность вероятности $f(z/H_0)$ при условии, что справедлива проверяемая гипотеза H_0. График плотности изображен на рис. 3.

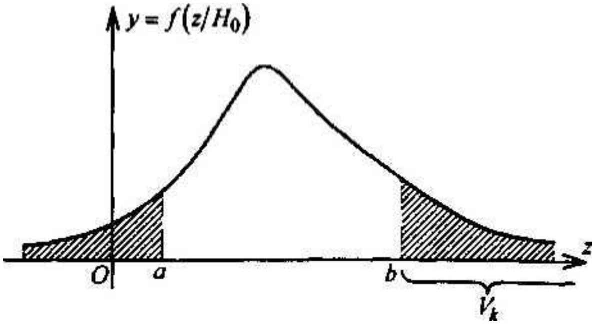


Рис. 3 – Правосторонняя критическая область

Пусть для простоты область значений Z – вся вещественная ось. Если критическая область V_k представляет собой промежуток $(-\infty, \alpha)$ или $(\beta, +\infty)$, то соответствующий критерий называется *односторонним (левосторонним* или *правосторонним)*.

Если же критическая область является объединением этих полубесконечных промежутков: $V_k = (-\infty, \alpha) \bigcup (\beta, +\infty)$, то соответствующий критерий называется *двусторонним*.

Пример 1. Проверяется гипотеза H_0 о том, что вероятность события $\mathbf{P}(A) = p$. Для статистики критерия (2.1): $Z = \frac{\mu}{n} - p$ целесообразен двусторонний критерий, так как большие по модулю отклонения относительной частоты μ / n от вероятности p заставляют отвергнуть гипотезу H_0.

КОНТРОЛЬНЫЕ ВОПРОСЫ ПО КУРСУ

1. Матрицы.
2. Операции над матрицами.
3. Определитель квадратной матрицы.
4. Свойства определителей.
5. Миноры и алгебраические дополнения квадратной матрицы.
6. Ранг матрицы.
7. Обратная матрица.
8. Свойства обратной матрицы.
9. Нахождение обратной матрицы с помощью элементарных преобразований (метод Гаусса).
10. Векторы. Линейные операции над векторами.
11. Скалярное произведение.
12. Основные задачи на прямую линию на плоскости.
13. Определение основных элементов в треугольнике.
14. Системы линейных уравнений.
15. Решение системы линейных уравнений методом обратной матрицы.
16. Формулы Крамера.
17. Исследование системы уравнений, с использованием определителей.
18. Производная функции, ее геометрический и физический смысл.
19. Основные правила дифференцирования.
20. Производная сложной функции.
21. Логарифмическое дифференцирование.
22. Производная показательно- степенной функции.
23. Производная обратных функций.
24. Дифференциал функции.
25. Элементы теории множеств.
26. Классическое определение вероятности.

27. Классическое определение вероятности.
28. Статистическое определение вероятности.
29. Геометрические вероятности.
30. Основные формулы комбинаторики.
31. Теоремы сложения и умножения вероятностей.
32. Случайные величины и случайные векторы.
33. Основные числовые характеристики дискретных случайных величин.
34. Основные числовые характеристики непрерывных случайных величин.
35. Основные понятия математической статистики.
36. Точечное оценивание числовых характеристик и параметров распределения генеральной совокупности.
37. Интервальное оценивание числовых характеристик и параметров распределения генеральной совокупности.
38. Статистические гипотезы.

СПИСОК ЛИТЕРАТУРЫ

1. Александров П.С. Курс аналитической геометрии и линейной алгебры. – М.: Наука, 1979. – 512 с.

2. Вентцель Е.С. Теория вероятностей и ее инженерные приложения: Учебное пособие для ВУЗов. - М.: Высшая школа, 2000. - 480 с.

3. Высшая математика для экономистов: Практикум для студентов вузов, обучающихся по экономическим специальностям / под ред. Н.Ш. Кремера. – 2-е изд., перераб. И доп. – М.: ЮНИТИ-ДАНА, 2007. – 479 с.

4. Высшая математика для экономистов: Учебник для вузов / Под редакцией Н.Ш. Кремера – 2-е изд., перераб. и доп. – М.: Банки и биржи, ЮНИТИ, 1999. – 471 с.

5. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. Юрайт, 2014. - 400 с.

6. Гмурман В.Е. Теория вероятностей и математическая статистика. Юрайт, 2014. – 400 с.

7. Данко П.Е., Попов А.Г., Кожевникова Т.Я. Высшая математика в упражнениях и задачах: Учеб. пособие для студентов втузов. В 2 ч-х. Ч. I. – 6-е изд., испр. и доп. − М.: Высшая школа, 2006. − 304 с.

8. Данко П.Е., Попов А.Г., Кожевникова Т.Я. Высшая математика в упражнениях и задачах: Учеб. пособие для студентов втузов. В 2 ч-х. Ч. II. – 6-е изд., испр. и доп. − М.: Высшая школа, 2007. − 416 с.

9. Ефимов Н.В. Краткий курс аналитической геометрии: учеб. для студентов вузов. – М.: Физматлит, 2004. – 238 с.

10. Зайцев И.Л. Высшая математика. – М.: Высшая школа, 1991. -– 427 с.

11. Зайцева И.В., Малафеев О.А. Линейная алгебра с приложениями к моделированию коррупционных систем и

процессов. - Ставрополь, Издательский дом «ТЭСЭРА», 2016. – 366 с.

12. Зайцева И.В., Малафеев О.А., Новожилова Л.М. Конечные множества: от концептов - до проектов. Рабочая тетрадь № 1 бакалавра-гуманитария: Учебное пособие / Ставрополь, 2020. – 53 с.

13. Зайцева И.В., Малафеев О.А., Новожилова Л.М. Элементы конечной комбинаторики: от концептов - до проектов. Рабочая тетрадь № 2 бакалавра-гуманитария: Учебное пособие / Ставрополь, 2020. – 40 с.

14. Клетеник Д. В. Сборник задач по аналитической геометрии. Учебное пособие. СПб: Специальная литература 1998. - 200 с.

15. Кудрявцев Л.Д. Краткий курс математического анализа. Т.1., Т.2. Дифференциальное и интегральное исчисления функций одной переменной. Ряды: Учебник. – 3-е изд., перераб. – М.: Физматлит, 2002. – 400 с.

16. Курош А. Г. Курс высшей алгебры: учебник. - М: Наука, 1971. - 431 с.

17. Маркович Э.С. Курс высшей математики с элементами теории вероятности и математической статистики. – М.: Наука, 1976. – 486 с.

18. Минорский В.Д. Сборник задач по высшей математике. – М.: Наука, 1987. – 352 с.

19. Натансон И. П. Краткий курс высшей математики. – СПб: Лань, 1997. – 727 с.

20. Письменный Д. Т. Конспект лекций по теории вероятностей, математической статистике и случайным процессам. М: АЙРИС Пресс 2013. - 287 с.

21. Письменный Д.Т. Конспект лекций по высшей математике. I часть. – 2-е изд., испр. – М.: Айрис-пресс, 2003. – 288 с.

22. Письменный Д.Т. Конспект лекций по высшей математике. 2 часть. – 2-е изд., испр. – М.: Айрис-пресс, 2003. – 256 с.

23. Сборник задач по высшей математике для экономистов. Учебное пособие под ред. В. И. Ермакова. М: ИНФРА-М, 2005. - 574 с.

24. Тарасенко Е.О., Зайцева И.В., Корнеев П.К., Гладков А.В. Теория вероятностей и математическая статистика: учебное пособие / Северо-кавказский федеральный университет. Ставрополь, 2018. - 229 с.

25. Таха, Хемди А. Введение в исследование операций. – М.: Издательский дом «Вильямс», 2005. – 912 с.