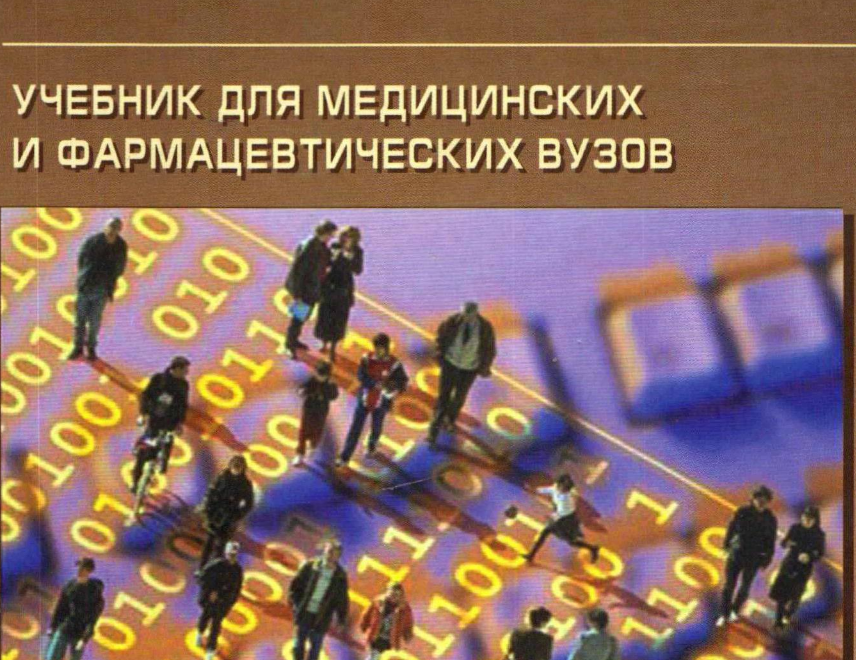
ОСНОВЫ ВЫСШЕЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

УЧЕБНИК ДЛЯ МЕДИЦИНСКИХ И ФАРМАЦЕВТИЧЕСКИХ ВУЗОВ



ИЗДАТЕЛЬСКАЯ ГРУППА
«ГЭОТАР-Медиа»

ОСНОВЫ ВЫСШЕЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

УЧЕБНИК ДЛЯ ВУЗОВ

2-е издание, исправленное

Рекомендовано УМО по медицинскому
и фармацевтическому образованию
вузов России в качестве учебника
для студентов медицинских
и фармацевтических вузов

Москва
ИЗДАТЕЛЬСКАЯ ГРУППА
«ГЭОТАР-Медиа»
2008

УДК 51(075.8)
ББК 22.1я73
O75

Рецензенты:

Д-р физ.-мат. наук, профессор математико-механического факультета Санкт-Петербургского государственного университета *В.Б. Невзоров*;

Кандидат физ.-мат. наук, доцент кафедры медицинской и биологической физики Московской медицинской академии им. И.М. Сеченова *Е.Ю. Смирнов*.

Авторы:

Кандидат физ.-мат. наук, доцент, зав. кафедрой высшей математики Санкт-Петербургской химико-фармацевтической академии И.В. Павлушков; д-р физ.-мат. наук, профессор Л.В. Розовский; кандидат техн. наук, доцент А.Е. Капульцевич; кандидат физ.-мат. наук, доцент Л.А. Кулонен; старшие преподаватели А.М. Камоцкая, И.Л. Степанова, Н.М. Тышко, Т.Ю. Ивановская, В.Д. Маслова.

O75 **Основы высшей математики и математической статистики:** Учебник / И.В. Павлушков и другие. — М.: ГЭОТАР-Медиа, 2008. — 424 с.

ISBN 978-5-9704-0880-3

В учебнике изложен курс высшей математики фармацевтического факультета, включающий основные элементарные фунуции, дифференциальное исчисление функции одной переменной, элементы дифференциального исчисления функций нескольких переменных, интегральное исчисление функции одной переменной, дифференциальные уравнения первого и второго порядка, основы теории вероятностей и математической статистики.

Учебник содержит подробные пояснения теоретического материала, а также большое количество примеров и задач.

Предназначен для студентов медицинских и фармацевтических вузов.

УДК 51(075.8)
ББК 22.1я73

ISBN 978-5-9704-0880-3

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Курс высшей математики на фармацевтическом факультете состоит из общего курса и специальных разделов. В общий курс входят: основные элементарные функции, дифференциальное исчисление функции одной переменной, элементы дифференциального исчисления функций нескольких переменных, интегральное исчисление функции одной переменной, дифференциальные уравнения первого и второго порядка, основы теории вероятностей и математической статистики.

Данный учебник содержит подробные пояснения теоретического материала, а также большое количество примеров и задач. В нем указаны методы решения типовых задач и приведены примеры. По каждому разделу учебник содержит большое количество задач для самостоятельного решения и может быть использован как задачник по общему курсу высшей математики для фармацевтических факультетов. Данным учебником с успехом могут пользоваться также и студенты заочной формы обучения фармацевтических вузов.

ГЛАВА 1
ВВЕДЕНИЕ В МАТЕМАТИЧЕСКИЙ АНАЛИЗ

§ 1.1. Функции

1.1.1. Определение функции, числовых промежутков и окрестностей точек

Одним из основных математических понятий является понятие функции, устанавливающее зависимость между элементами двух множеств.

Определение. Пусть X, Y — некоторые множества, элементами которых являются некоторые числа. Если каждому числу $x \in$ $\in X$ по некоторому закону или правилу f ставится в соответствие число $y \in Y$, то говорят, что на множестве X задана числовая функция f и записывают эту функциональную зависимость формулой $y = f(x)$ или, более наглядно, в виде диаграммы:

$$X \xrightarrow{f} Y. \tag{1.1}$$

Переменная x называется *независимой переменной* или *аргументом*, а переменная y называется *зависимой переменной* (от x) или *функцией*.

Множество X — область изменения аргумента — называется *областью определения функции*. Множество Y, содержащее все значения, которые принимает y, *называется областью изменения функции*.

При дальнейшем изложении множества X и Y часто оказываются конечными или бесконечными промежутками.

а. Конечные промежутки:

открытый интервал (или просто интервал) (a, b) — множество вещественных чисел, удовлетворяющих неравенствам $a < x < b$, или $(a, b) \Leftrightarrow (a < x < b)$, где $\Leftrightarrow$ знак эквивалентности;

замкнутый интервал (или отрезок) $[a, b]$: $[a, b] \Leftrightarrow (a \leqslant x \leqslant b)$;

полуоткрытые интервалы $(a, b]$ и $[a, b)$: $(a, b] \Leftrightarrow (a < x \leqslant b)$ и $[a, b) \Leftrightarrow (a \leqslant x < b)$ соответственно.

б. Бесконечные промежутки:

$(-\infty, +\infty) = \Re$ — множество всех вещественных чисел, т. е. $\Re \Leftrightarrow (-\infty < x < +\infty)$; аналогично, $(a, +\infty) \Leftrightarrow (a < x < +\infty)$ и т. д.

Числа a, b называются соответственно левым и правым *концами* этих промежутков.

Символы $-\infty$ и $+\infty$ не числа, это обозначение процесса неограниченного удаления точек числовой оси влево и вправо от начала 0.

Пусть x_0 — любое действительное число (точка на числовой прямой). *Окрестностью точки* x_0 называется любой интервал (a, b), содержащий точку x_0, интервал $(x_0 - \varepsilon, x_0 + \varepsilon)$, где $\varepsilon > 0$, симметричный относительно x_0, называется *ε-окрестностью точки* x_0.

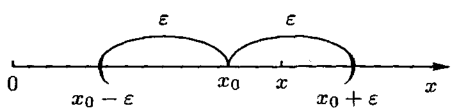


Рис. 1.1. Изображение ε-окрестности точки x_0.

Если $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$, то справедливы неравенства

$$x_0 - \varepsilon < x < x_0 + \varepsilon,$$

что равносильно

$$|x - x_0| < \varepsilon.$$

Частное значение функции $f(x)$ при $x = a$ можно найти, подставив a вместо аргумента: $f(a)$. При этом a может быть как буквенным выражением, числом, так и некоторой функцией, например $\varphi(t)$. В последнем случае $f(\varphi(t))$ будет сложной функцией, с ней мы ознакомимся в п. 1.1.4.

Примеры.

Найти область определения и область значений функций.

1. $y = \sqrt{1 - x^2}$.

Решение. Область определения этой функции состоит из всех x, для которых она имеет смысл. Таким образом, $X = \{|x| \leqslant \leqslant 1\} \Leftrightarrow [-1, 1]$, $Y = [0, 1]$, т. е. $[-1, 1] \xrightarrow{f} [0, 1]$.

2. $y_n = \left(\frac{1}{2}\right)^n$.

Решение. Здесь независимая переменная n принимает целые положительные значения $n \in N = \{1, 2, \ldots\}$, следовательно, y является функцией натурального аргумента и вычисляется по заданной формуле $Y = \left\{\frac{1}{2}, \frac{1}{4}, \ldots, \left(\frac{1}{2}\right)^n, \ldots\right\}$, $N \xrightarrow{f} Q$ (Q — множество рациональных чисел).

3. Построить ε-окрестность точки $x_0 = 2$ ($\varepsilon = 0{,}1$).

Решение. По определению, ε-окрестностью точки 2 будет интервал $|x - 2| < 0{,}1$ или $-0{,}1 < x - 2 < 0{,}1 \Rightarrow 1{,}9 < x < 2{,}1$.

Самостоятельная работа

Построить интервалы изменения переменной x, удовлетворяющей неравенствам:

1. $|x| < 4$. **2.** $x^2 \leqslant 9$. **3.** $|x - 4| < 1$.

4. $-1 < x - 3 \leqslant 2$. **5.** $x^2 > 9$. **6.** $(x - 2)^2 \leqslant 4$.

Найти области определения функций:

1. $y = \sqrt{x + 2}$. **2.** $y = \sqrt{9 - x^2}$. **3.** $y = \sqrt{4x - x^2}$.

4. $y = \sqrt{-x} + \sqrt{4 + x}$. **5.** $y = \arcsin\frac{x - 1}{2}$.

6. $y = -\sqrt{2\sin x}$. **7.** $y = -\frac{x\sqrt{16 - x^2}}{2}$.

8. $y = \sqrt{x + 1} - \sqrt{3 - x}$.

Вычислить значения функций в заданных точках:

1. $f(x) = x^2 - x + 1; \quad f(2), \quad f(a+1).$

2. $\varphi(x) = \dfrac{2x-3}{x^2+1}; \quad \varphi(3/2), \quad \varphi(1/x), \quad \dfrac{1}{\varphi(x)}.$

3. $F(x) = x^2; \quad \dfrac{F(b)-F(a)}{b-a}, \quad F\left(\dfrac{a+h}{2}\right) - F\left(\dfrac{a-h}{2}\right).$

1.1.2. Некоторые свойства функций и их графиков

Пусть задана функция $f\colon X \to Y$. Правило, по которому можно находить y, зная x, может быть задано графиком функции.

Определение. *Графиком функции* в декартовой прямоугольной системе координат называется множество всех точек, абсциссы которых являются значениями аргумента, а ординаты — соответствующими значениями функции.

Пример. Графиком функции $y = x^2$ является парабола, ось симметрии которой совпадает с положительной полуосью ординат, а вершина с началом координат.

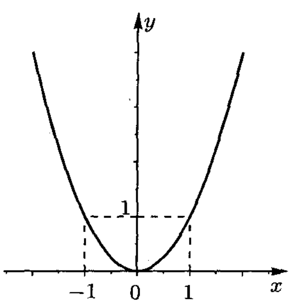


Рис. 1.2. График функции $y = x^2$.

Часто графики автоматически вычерчиваются самопишущими приборами или изображаются на экране дисплея. Преимуществом такого представления функции является наглядность, а недостатком — неточность.

Функцию можно задавать также с помощью таблицы или формулы (аналитически).

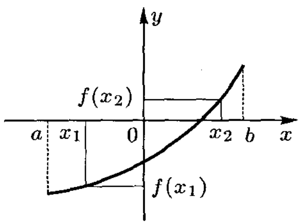
Табличный способ применяется на практике при обработке результатов наблюдений приближенных значений функции.

Аналитический способ задания функции является наиболее удобным для полного исследования функции при помощи методов математического анализа.

Отметим основные характеристики функции: монотонность, ограниченность, четность (нечетность), периодичность.

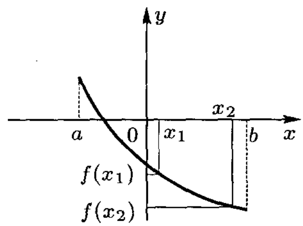
Определение. Функция называется *возрастающей* (*убывающей*) в интервале, если большему значению аргумента из этого интервала соответствует большее (меньшее) значения функции.

График возрастающей на интервале (a, b) функции, если его рассматривать слева направо, поднимается вверх (рис. 1.3а), а для убывающей функции — опускается вниз (рис. 1.3б).



$x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$

Рис. 1.3а. График возрастающей функции на (a, b).



$x_1 < x_2 \Rightarrow f(x_1) > f(x_2)$

Рис. 1.3б. График убывающей функции на (a, b).

Определение. Интервал независимой переменной, в котором функция возрастает (убывает), называется *интервалом возрастания* (*убывания*). Как интервал возрастания, так и интервал убывания называют *интервалами монотонности* функции, а функцию в этом интервале — *монотонной функцией*.

Определение. Значение аргумента, при котором функция обращается в нуль, называется *нулем функции*.

Если функция задана формулой $y = f(x)$, то для нахождения нуля (или нулей) функции следует решить уравнение $f(x) = 0$.

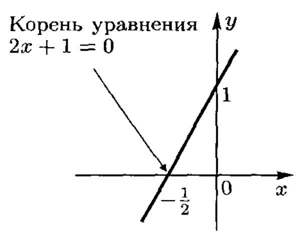


Рис. 1.4. График функции $y = 2x + 1$.

При графическом задании нулями функции являются точки пересечения ее графиком оси абсцисс.

Пример. Найти нули функции $y = 2x + 1$.

Решение. $2x + 1 = 0 \Rightarrow x = -1/2$.

Определение. Функция называется *четной*, если при изменении знака допустимого аргумента значение функции не изменяется. Функция называется *нечетной*, если при изменении знака допустимого аргумента значение функции меняет знак на противоположный.

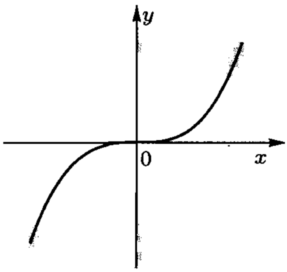


Рис. 1.5. График функции $y = x^3$.

Таким образом, если функция $f(x)$ — четная, то для всех x из ее области определения должно выполняться равенство $f(-x) = f(x)$, как это происходит, например, при $f(x) = x^2$, а если $f(x)$ — нечетная, то $f(-x) = -f(x)$ для любого x из области определения функции, как, например, в случае $f(x) = x^3$.

Обращаем внимание на то, что четные или нечетные функции должны быть обязательно определены в области, симметричной относительно начала координат.

При этом график четной функции симметричен относительно оси абсцисс (как на рис. 1.2), а график нечетной функции симметричен относительно начала координат (как на рис. 1.5).

Заметим, что не все функции являются четными либо нечетными; функции, не являющиеся ни четными, ни нечетными, будем называть функциями общего вида.

Определение. Функция $f(x)$, определенная на множестве X, называется *ограниченной* на этом множестве, если существует

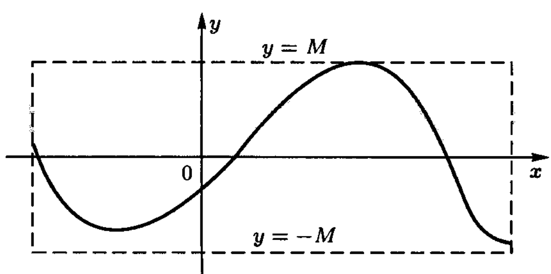


Рис. 1.6. График ограниченной функции.

такое число $M > 0$, что для всех $x \in X$ выполняется неравенство $|f(x)| \leqslant M$.

График ограниченной функции лежит между прямыми $y = M$ и $y = -M$.

Определение. Функция $f(x)$ называется *периодической*, если существует такое положительное число a, что $f(x+a) = f(x) = f(x-a)$ для любого x из ООФ (точки x, $x+a$, $x-a \in$ области определения функции). При этом наименьшее положительное a с таким свойством (если таковое существует) называется *периодом* функции.

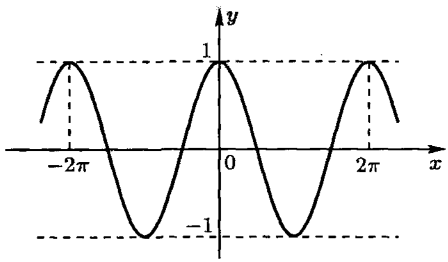


Рис. 1.7. График функции $y = \cos x$.

График периодической функции получается путем повторения части графика, соответствующего интервалу оси абсцисс, равному по длине периоду функции.

Примером периодической функции служит определенная на всей оси функция $y = \cos x$, период которой равен 2π.

Таким образом, сдвиг графика периодической функции вдоль оси абсцисс на интервал, длина которого кратна периоду, не приводит к изменению этого графика. В частности, область определения периодической функции неограничена.

Самостоятельная работа

Указать, какие из следующих функций четные и какие нечетные:

1. $\dfrac{\sin x}{x}$. **2.** $\dfrac{a^x - 1}{a^x + 1}$. **3.** $a^x + \dfrac{1}{a^x}$. **4.** $a^x - \dfrac{1}{a^x}$.

5. $x\sin^2 x - x^3$. **6.** $x + x^2$. **7.** $|x|$. **8.** $\dfrac{\operatorname{tg} x}{\sin 2x}$.

Найти нули функции:

1. $y = ax + b$. **2.** $y = x^2 + px + q$. **3.** $y = x^4 + px^2 + q$.

4. $y = 2\lg(x+1)$. **5.** $y = a^{2x} - a^2$ $(a > 0)$.

6. $y = 2\sin x - 1$. **7.** $y = \operatorname{tg} x + 1$.

Найти период функции:

1. $y = \operatorname{tg} 2x$. **2.** $y = \sin\dfrac{x}{2}$. **3.** $y = \dfrac{\operatorname{ctg} x}{\cos 2x}$.

Воспользовавшись свойствами графиков четных/нечетных функций и результатами п. 1.3, построить графики функций:

1. $y = |x|$. **2.** $y = -x + |x|$. **3.** $y = -|x-2|$.

4. $y = x - 4 + |x-2|$, $x \in [-2, 5]$. **5.** $y = \lg(x+2)$.

6. $y = 2^{-x}$. **7.** $y = x^2 + 2x + 2$. **8.** $y = -x^2 + 4x$.

1.1.3. Сложная функция. Обратная функция

Определение. Сложной функцией называется функция, аргумент которой также является функцией, т. е. $F(x) = f(\varphi(x))$, или, в виде диаграммы (см. (1.1)):

$$\begin{tikzcd} X \arrow[r, "\varphi"] \arrow[rr, bend left, "F"] & U \arrow[r, "f"] & Y \end{tikzcd}$$

Иначе говоря, чтобы сосчитать значение в точке x сложной функции $f(\varphi(x))$, составленной из функций f и φ, следует сначала найти частное значение $u = \varphi(x)$ внутренней функции φ, а затем подставить его в качестве аргумента во внешнюю функцию f.

При этом область определения функции $F(x)$ следует выбирать таким образом, чтобы промежуточное множество U, с одной стороны, было областью значений функции $\varphi(x)$, а с другой стороны, являлось областью определения функции $f(u)$.

Пример 1. Рассмотрим сложную функцию $y = \lg(1 - x^2)$. Здесь $y = f(u) = \lg u$, в то время как $u = \varphi(x) = 1 - x^2$.

Областью определения функции y является интервал $(-1, 1)$, в котором как функция $\varphi(x)$, так и функция $f(u)|_{u=\varphi(x)}$ имеют смысл.

Рассмотрим функцию с областью определения X и областью значений Y. Предположим, что каждому значению $y \in Y$ соответствует одна определенная точка $x \in X$, такое что $y = f(x)$. Тогда существует функция $\varphi\colon Y \to X$, переводящая любое $y \in Y$ в $x \in X$, удовлетворяющее вышеуказанному свойству $y = f(x)$.

Функции f и φ с вышеперечисленными свойствами называются *взаимно-обратными*, а функция φ называется *обратной* по отношению к f. С учетом того что символ x соответствует, как правило, независимой переменной, обычно вместо записи $x = \varphi(y)$ используют запись $y = \varphi(x)$.

Из определения обратной функции вытекает, что любая строго монотонная функция имеет обратную.

Между графиками функций $y = f(x)$ и $y = \varphi(x)$ имеется простая связь: график обратной функции $y = \varphi(x)$ симметричен графику данной функции $y = f(x)$ относительно биссектрисы первого и третьего координатных углов.

Отметим, что взаимно-обратные функции f и φ удовлетворяют соотношению

$$f(\varphi(x)) = \varphi(f(x)) = x \tag{1.2}$$

и могут с помощью него вычисляться.

Пример 2. Пусть $y = f(x) = x^3$. Тогда $f(\varphi(x)) = \varphi^3(x)$ и равенство (1.2) дает $\varphi^3(x) = x$, или $\varphi(x) = x^{1/3}$, что, впрочем, легко следует непосредственно из соотношения $y = x^3$.

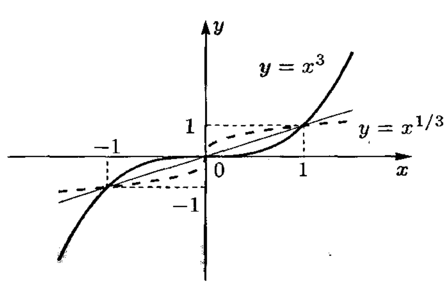


Рис. 1.8. График взаимно-обратных функций $y = x^3$ и $y = x^{1/3}$.

Важно иметь в виду, что функция $f(x)$, возрастающая или убывающая на X, заведомо имеет обратную функцию (определение возрастания и убывания функции дано в п. 1.3).

В противоположном случае однозначность соответствия между X и Y нарушается, и обратной функции не существует. Однако, как правило, область определения X можно разбить на участки возрастания и убывания функции $f(x)$, на каждом из которых обратная функция уже может быть определена.

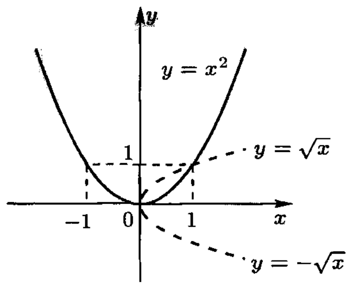


Рис. 1.9. Пример функции, не имеющей взаимно-обратной.

Пример 3. Пусть $y = x^2$. Тогда $X = (-\infty, \infty)$, а $Y = [0, +\infty)$. Таким образом, взаимно-однозначного соответствия между X и Y нет (каждому $y \neq 0$ соответствует два значения x, отличающиеся знаками), следовательно, нет и обратной функции.

Если же разбить X на $(-\infty, 0]$ и $[0, +\infty)$, то на каждой полупрямой зависимость $y = x^2$ является взаимно-однозначной, и, следовательно, на $(-\infty, 0]$ функция $y = x^2$ имеет обратную функцию $y = -\sqrt{x}$, а на $[0, +\infty)$ обратной к ней является функция $y = \sqrt{x}$.

Пример 4. Пусть функция y связана с независимой переменной x линейной зависимостью $3x + 2y - 6 = 0$. Найти обратную функцию и построить графики прямой и обратной функций.

Решение. Для нахождения обратной функции в общей с прямой функцией системе координат достаточно в соответствующем уравнении поменять обозначения x и y местами.

Таким образом, в нашем примере обратная зависимость выражается соотношением $3y + 2x - 6 = 0$ и также является линейной.

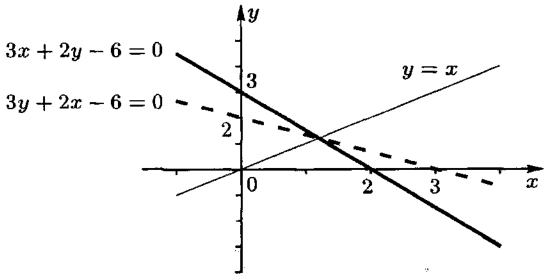


Рис. 1.10. Пример функции, не имеющей взаимно-обратной.

При построении графиков принималось во внимание то, что прямая линия однозначно определяется любой парой различных лежащих на ней точек. В частности, прямая $3x + 2y - 6 = 0$ определяется точками $(0, 3)$ и $(2, 0)$.

Заметим, что в соответствии со свойствами взаимно-обратных функций прямые на рис. 1.10 симметричны относительно биссектрисы 1-го и 3-го координатных углов.

1.1.4. Элементарные функции

Определение. Основными *элементарными функциями* называются:

1) степенная функция: $y = x^n$, где n — действительное число, $x > 0$ (в некоторых случаях, в частности при натуральном n, степенная функция определена на всей оси);

2) показательная функция: $y = a^x$, где $a > 0$, $a \neq 1$, и $X = \Re$;

3) логарифмическая функция: $y = \log_a x$, где основание логарифмов $a > 0$, $a \neq 1$, и $X = (0, +\infty)$.

4) тригонометрические функции: $y = \sin x$, $y = \cos x$, $y = \operatorname{tg} x$ и $y = \operatorname{ctg} x$.

5) обратные тригонометрические функции: $y = \arcsin x$, $y = \arccos x$, $y = \operatorname{arctg} x$ и $y = \operatorname{arcctg} x$.

К множеству элементарных функций относятся все основные элементарные функции и постоянные, а также все функции, получающиеся из них с помощью четырех арифметических действий и операции взятия функции от функции, примененных последовательно конечное число раз.

Так, элементарными являются функции $y = \lg(x + \sqrt{1+x^2})$, $y = \operatorname{arctg}\dfrac{2x^4}{1-x^2}$ и $y = \operatorname{tg} x - \sqrt{x}$. Функция $y = \begin{cases} x, & x \leqslant 1, \\ x^2, & x > 1 \end{cases}$ элементарной не является.

Областью определения элементарной функции являются все значения аргумента, при котором эта функция имеет смысл.

Например, областью определения функции $y = \sqrt{x^2-1}$ является множество $X = (-\infty < x \leqslant -1 \cup 1 \leqslant x < +\infty)$. Здесь символ $\cup$ обозначает объединение.

1. Степенная функция $y = x^n$ при целом n определена на всей оси; четная, если n четное, и нечетная, если n нечетное (см. рис. 1.2 и 1.5).

При произвольном n рассматривается в области $x > 0$. Если $n > 0$, то графики функции $y = x^n$ возрастают от нуля до бесконечности в интервале $(0, +\infty)$, проходят через точки $(0, 0)$ и $(1, 1)$ и разделяются прямой $y = x$ на кривые, обращенные выпуклостью вниз при $n > 1$ и вверх при $0 < n < 1$ (рис. 1.11а).

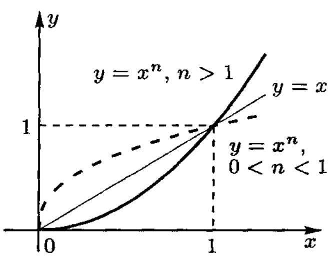


Рис. 1.11а. График функции $y = x^n$, $n > 0$.

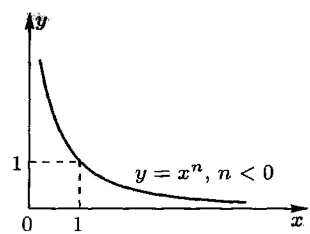


Рис. 1.11б. График функции $y = x^n$, $n < 0$.

Если $n < 0$, то график функции $y = x^n = (1/x)^{|n|}$ убывает от бесконечности до нуля (см. рис. 1.11б). Обратной к функции $y = x^n$, $x > 0$, является функция $y = x^{1/n}$.

Напоминаем, что

$$(x^a)^b = x^{ab}, \qquad x^{-a} = \frac{1}{x^a}; \tag{1.3}$$

$$x^{n/m} = \sqrt[m]{x^n} \tag{1.4}$$

(n и m целые положительные числа) и, в частности, $\sqrt[n]{x} = x^{1/n}$.

2. Показательная $y = a^x$, $-\infty < x < \infty$, и логарифмическая $y = \log_a x$, $x > 0$, функции при одном и том же параметре a являются взаимно обратными. Их графики симметричны относительно биссектрисы 1-го и 3-го координатных углов (рис. 1.12).

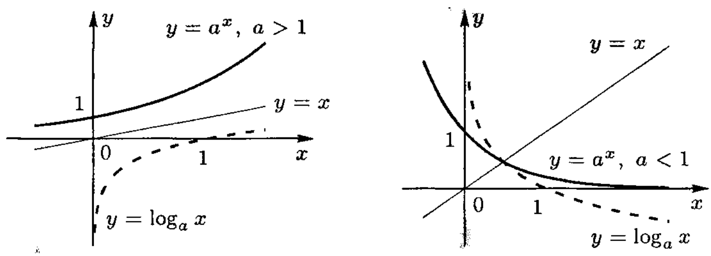


Рис. 1.12. Графики показательной и логарифмической функций.

Показательная функция всегда положительна, поэтому ее график расположен над осью OX. Кроме того, поскольку $a^0 = 1$, он проходит через точку $(0, 1)$. При $a > 1$ показательная функция возрастает от нуля до бесконечности, при $a < 1$ — убывает от бесконечности до нуля. Отметим, что график показательной функции с основанием a симметричен относительно оси OY графику показательной функции с основанием $1/a$, что следует из равенства $(1/a)^x = a^{-x}$.

Напоминаем, что функция $y = e^x$ ($e = 2{,}718\ldots$) называется *экспоненциальной*, а ее график — *экспонентой*; логарифмы с основанием e обозначают через $\ln x$ и называют *натуральными*. Логарифмы с основанием 10 обозначают через $\lg x$ и называют *десятичными*. Таким образом, $\log_e x = \ln x$, $\log_{10} x = \lg x$.

Принимая во внимание, что логарифмическая и показательная функции взаимно обратны, имеем (см. (1.2))

$$\log_a a^x = x, \qquad a^{\log_a x} = x, \tag{1.5}$$

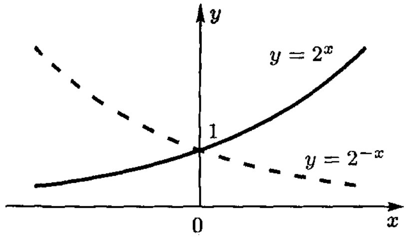


Рис. 1.13. Графики функций $y = 2^x$ и $y = 2^{-x}$.

где первое равенство справедливо при любом x, а второе — при $x > 0$.

В частности, $x = e^{\ln x}$ и, следовательно (см. (1.3)),

$$x^n = e^{n \ln x}, \quad x > 0, \tag{1.6}$$

— представление степенной функции с помощью показательной.

Справедлива формула

$$\log_a x = \frac{\log_b x}{\log_b a}, \tag{1.7}$$

т. е. логарифмы чисел при разных основаниях (a и b соответственно) пропорциональны друг другу с коэффициентом пропорциональности (*модулем перехода*) $\dfrac{1}{\log_b a} = \log_a b$.

Примеры.

1. Выразить $\log_2 x$ через $\lg x$ и через $\ln x$.

Решение. $\log_2 x = \dfrac{\lg x}{\lg 2} = \log_2 10 \cdot \lg x$, $\quad \log_2 x = \dfrac{\ln x}{\ln 2} = \log_2 e \cdot \ln x$.

2. Записать функцию $y = 2^x$ в виде показательной функции с основанием 10.

Решение. $2^x = 10^{\lg 2^x} = 10^{\lg 2 \cdot x}$.

1.1.5. Тригонометрические функции

Прежде всего напомним, что в качестве аргумента тригонометрических функций в математическом анализе всегда принимается радианная мера дуги или угла, т. е. число, равное отношению длины этой дуги к радиусу окружности. Таким образом,

$$\alpha_{рад.} = \alpha^\circ \frac{\pi}{180}, \quad \text{или} \quad \alpha^\circ = \frac{180}{\pi}\alpha_{рад.}, \tag{1.8}$$

где α° — градусная, а $\alpha_{рад.}$ — радианная мера угла. В частности,

$$\frac{\pi}{4} \Leftrightarrow 45^\circ, \quad \frac{\pi}{3} \Leftrightarrow 60^\circ, \quad \frac{\pi}{6} \Leftrightarrow 30^\circ, \quad \frac{\pi}{2} \Leftrightarrow 90^\circ \quad \text{и т. д.}$$

Тригонометрические функции являются периодическими: $\sin x$ и $\cos x$ имеют период 2π, а $\operatorname{tg} x$ и $\operatorname{ctg} x$ — период π (см. рис. 1.14 и 1.15).

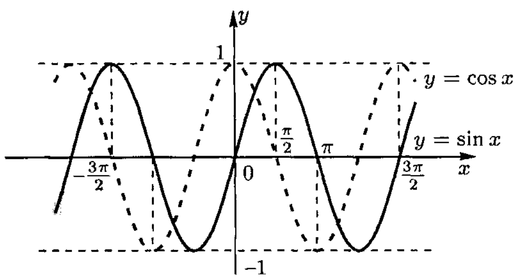


Рис. 1.14. Графики функций $y = \sin x$ и $y = \cos x$.

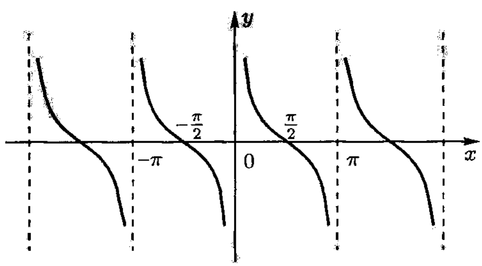


Рис. 1.15. График функции $y = \operatorname{ctg} x$.

График косинусоиды отличается от графика синусоиды сдвигом влево по оси OX на $\pi/2$, поскольку $\cos x = \sin(x + \pi/2)$ (см. п. 1.3).

Функции $\sin x$, $\operatorname{tg} x$ и $\operatorname{ctg} x$ являются нечетными, а функция $\cos x$ — четная. Наконец, $\sin x$ и $\cos x$ определены при любом x,

$\operatorname{tg} x$ — при всех x, кроме точек вида $(2k+1)\pi/2$, где k — любое целое число, а $\operatorname{ctg} x$ — при всех x, кроме точек вида $k\pi$.

Напоминаем, что

$$\cos x = \sin\left(\frac{\pi}{2} \pm x\right), \quad \sin x = \cos\left(\frac{\pi}{2} - x\right),$$

$$\sin^2 x + \cos^2 x = 1, \quad \sin 2x = 2\sin x \cos x,$$

$$\cos 2x = \cos^2 x - \sin^2 x = 1 - 2\sin^2 x = 2\cos^2 x - 1, \tag{1.9}$$

$$\operatorname{tg} x = \frac{\sin x}{\cos x}, \quad \operatorname{ctg} x = \frac{\cos x}{\sin x}, \quad \operatorname{ctg} x = \operatorname{tg}\left(\frac{\pi}{2} - x\right),$$

$$\operatorname{tg} 2x = \frac{2\operatorname{tg} x}{1 - \operatorname{tg}^2 x}.$$

1.1.6. Обратные тригонометрические функции

Поскольку тригонометрические функции являются периодическими, каждому значению функции соответствует бесчисленное множество значений аргумента. Таким образом, взаимнооднозначного соответствия между x и y нет и, следовательно, однозначная обратная функция не может быть определена.

Для решения проблемы нахождения x по y в случае зависимости $y = \sin x$, поступают следующим образом.

Функция $y = \sin x$ на интервале $-\pi/2 \leqslant x \leqslant \pi/2$ является возрастающей (см. рис. 1.16) и, следовательно, имеет обратную, обозначаемую через $y = \arcsin x$.

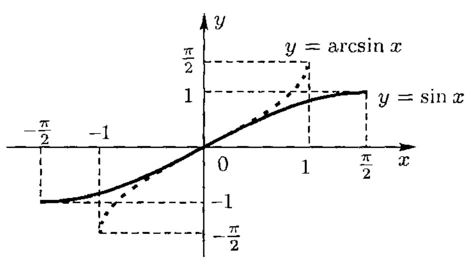


Рис. 1.16. График функции $y = \arcsin x$.

Учитывая, что график обратной функции симметричен с графиком прямой функции относительно биссектрисы 1-го и 3-го координатных углов, находим график функции $\arcsin x$.

Областью определения X функции $y = \arcsin x$ является отрезок $[-1, 1]$, а областью ее значений Y — отрезок $[-\pi/2, \pi/2]$, т. е.

$$[-1, 1] \xrightarrow{\arcsin x} [-\pi/2, \pi/2].$$

Кроме того, $y = \arcsin x$ является нечетной и возрастающей функцией.

Значение функции $\arcsin x$ есть радианная мера угла, синус которого равен данному значению независимой переменной x; при этом из всех углов, удовлетворяющих этому условию, выбирается угол из отрезка $[-\pi/2, \pi/2]$, т. е.

$$y = \arcsin x \Leftrightarrow \sin y = x, |y| \leqslant \pi/2. \tag{1.10}$$

Все значения y, удовлетворяющие уравнению $\sin y = x$, находят по формуле

$$y = \pi k + (-1)^k \arcsin x, \quad k = 0, \pm 1, \pm 2, \ldots \tag{1.11}$$

Например, решением уравнения $\sin y = 1/2$, $|y| \leqslant \pi/2$, является $\arcsin(1/2)$, т. е. число $\pi/6$. Общим же решением уравнения $\sin y = 1/2$ будут числа $y = \pi k + (-1)^k \pi/6$, или числа

$$\ldots, -1\frac{5}{6}\pi, \ -1\frac{1}{6}\pi, \ \frac{\pi}{6}, \ \frac{5}{6}\pi, \ 2\frac{1}{6}\pi, \ldots$$

Функция, обратная к $y = \cos x$, определяется аналогично. Имеем

$$[-1, 1] \xrightarrow{\arccos x} [0, \pi],$$

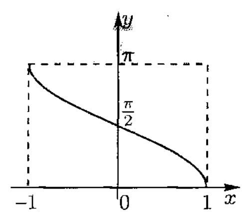


Рис. 1.17. График функции $y = \arccos x$.

или

$$y = \arccos x \Leftrightarrow \cos y = x, \quad 0 \leqslant y \leqslant \pi.$$

Функция $y = \arccos x$ убывает и удовлетворяет равенству

$$\arccos(-x) = \pi - \arccos x. \tag{1.12}$$

Общее решение уравнения $\cos y = x$, $|x| \leqslant 1$, имеет вид

$$y = 2\pi k \pm \arccos x. \tag{1.13}$$

Отметим формулы

$$\sin(\arccos x) = \cos(\arcsin x) = \sqrt{1-x^2}, \quad \arcsin x + \arccos = \frac{\pi}{2}. \tag{1.14}$$

Функция $y = \operatorname{arctg} x$ определена на всей оси; область ее значений — открытый интервал $(-\pi/2, \pi/2)$, т. е.

$$\Re \xrightarrow{\operatorname{arctg} x} [-\pi/2, \pi/2].$$

Эта функция является возрастающей и нечетной.

Уравнению $\operatorname{tg} y = x$ удовлетворяют числа вида

$$y = \pi k + \operatorname{arctg} x. \tag{1.15}$$

Аналогично, функция $y = \operatorname{arcctg} x$ определена на всей оси и принимает значения в интервале $(0, \pi)$. Она убывает и удовле-

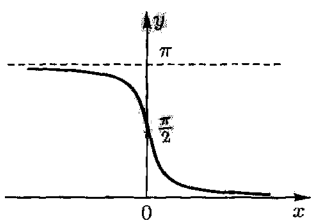


Рис. 1.18. График функции $y = \operatorname{arcctg} x$.

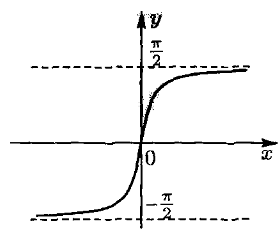


Рис. 1.19. График функции $y = \operatorname{arctg} x$.

творяет равенству

$$\operatorname{arcctg}(-x) = \pi - \operatorname{arcctg} x. \tag{1.16}$$

Общее решение уравнения $\operatorname{ctg} y = x$ имеет вид

$$y = \pi k + \operatorname{arcctg} x. \tag{1.17}$$

Отметим формулы:

$$\begin{aligned} &\operatorname{tg}(\operatorname{arcctg} x) = \operatorname{ctg}(\operatorname{arctg} x) = \frac{1}{x}, \\ &\operatorname{arctg} x + \operatorname{arcctg} x = \frac{\pi}{2}, \\ &\operatorname{arctg} x = \operatorname{arcctg}\frac{1}{x}, \quad \operatorname{arcctg} x = \operatorname{arctg}\frac{1}{x}. \end{aligned} \tag{1.18}$$

§ 1.2. Пределы

1.2.1. Предел функции

Теория пределов играет основополагающую роль в математическом анализе. С ее помощью устанавливаются такие свойства функции, как непрерывность, дифференцируемость, интегрируемость и т. д.

Рассмотрим пример. Пусть задана функция $f(x) = \dfrac{x^2 - 1}{x - 1}$, определенная при всех x, кроме $x = 1$. Исследуем поведение функции при значениях x, мало отличающихся от 1. Для этого составим таблицу значений функции в интересующем нас интервале:

x	0,97	0,98	0,99	1,01	1,02
$f(x)$	1,97	1,98	1,99	2,01	2,02

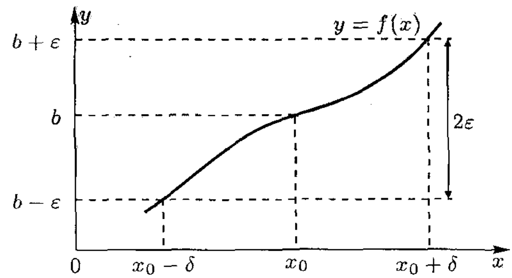
Мы видим, что чем ближе x приближается к 1, тем значения $f(x)$ ближе к 2. В подобных случаях говорят, что число 2 является пределом функции $f(x)$ при x, стремящемся к 1 (или более кратко: $f(x) \to 2$ при $x \to 1$).

Дадим теперь строгое определение предела функции.

Определение. Пусть функция $f(x)$ определена в некоторой окрестности точки x_0, кроме, может быть, самой точки x_0. Число b называется *пределом функции в точке* x_0 (или при $x \to \to x_0$), если для любого положительного ε, как бы мало оно не было, можно найти такое положительное δ, что для всех $x \neq \neq x_0$, удовлетворяющих неравенству $|x - x_0| < \delta$, выполняется неравенство $|f(x) - b| < \varepsilon$.

Записывают так: $\lim\limits_{x \to x_0} f(x) = b$.

Геометрический смысл этого определения: *если для любой ε-окрестности точки b найдется такая δ-окрестность точки x_0, что для всех $x \neq x_0$ из этой δ-окрестности соответствующие точки графика $f(x)$ лежат внутри полосы шириной 2ε, ограниченной прямыми $y = b+\varepsilon$, $y = b-\varepsilon$.* Заметим, что $b = \lim\limits_{x \to x_0} f(x)$.



В данном определении не оговаривается способ приближения x к x_0: слева, справа или колеблясь около x_0. Бывают случаи, что это существенно, поэтому вводится понятие одностороннего предела.

Определение. Число b_1 называется *пределом функции $y = f(x)$ слева в точке* x_0, если для любого наперед заданного сколь угодно малого $\varepsilon > 0$ существует такое число $\delta > 0$, что при всех $x \in (x_0 - \delta, x_0)$ выполняется неравенство $|f(x) - b_1| < \varepsilon$.

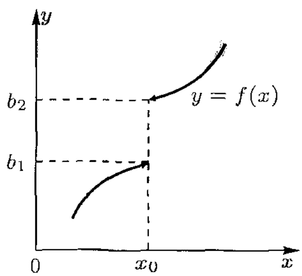


Рис. 1.20. Иллюстрация к определению односторонних пределов.

Записывается так: $\lim\limits_{x \to x_0 - 0} f(x) = b_1$.

Аналогично определяется *предел функции справа*.

$\lim\limits_{x \to x_0 + 0} f(x) = b_2$, если $f(x) \to b_2$, когда $x \to x_0$ оставаясь больше x_0.

Пределы функции слева и справа называются односторонними пределами. Очевидно, что если существует $\lim\limits_{x \to x_0} f(x) = b$, то существуют оба односторонних предела, тоже равных b.

Справедливо и обратное утверждение: если существуют односторонние пределы, оба равные b, то существует и $\lim\limits_{x \to x_0} f(x)$, также равный b.

Если же односторонние пределы не равны между собой $b_1 \neq \neq b_2$, то $\lim\limits_{x \to x_0} f(x)$ не существует.

Если функция $y = f(x)$ определена в промежутке $(a, +\infty)$, то можно определить предел функции при $x \to +\infty$.

Определение. Число b называется *пределом функции при* $x \to +\infty$, если для любого наперед заданного сколь угодно малого $\varepsilon > 0$ существует такое число $E > 0$, что при всех x, удовлетворяющих неравенству $x > E$, выполняется неравенство $|f(x) - b| < \varepsilon$, т. е. $\lim\limits_{x \to +\infty} f(x) = b$.

Иными словами, при достаточно больших x значения функции попадают в ε-окрестность точки b (рис. 1.21).

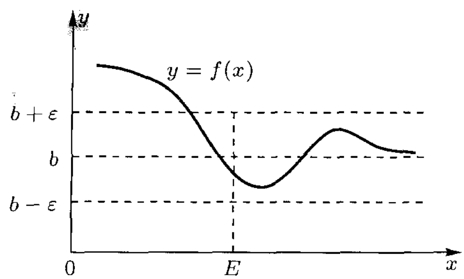


Рис. 1.21. Предел функции при $x \to +\infty$.

Аналогично определяется предел функции при $x \to -\infty$.

Если пределы функции $f(x)$ при $x \to +\infty$ и $x \to -\infty$ существуют и равны, то говорят, что $f(x)$ имеет предел при $x \to \infty$, совпадающий с ними, и обозначается $\lim\limits_{x \to \infty} f(x)$.

Во всех данных выше определениях функции имели конечные пределы, но может оказаться, что функция неограниченно возрастает или убывает, следовательно, не имеет конечного предела.

Среди функций, имеющих пределы (в точке или на ∞), выделяют класс функций, имеющих предел, равный 0. Такие функции называются *бесконечно малыми функциями* или бесконечно малыми величинами. Обозначаются буквами α, β, γ и т. д.

При вычислении пределов используется понятие эквивалентности бесконечно малых величин.

Определение. Пусть $\alpha(x)$ и $\beta(x)$ — бесконечно малые. Если $\lim\limits_{x\to x_0}\dfrac{\alpha(x)}{\beta(x)}=1$, то $\alpha(x)$ и $\beta(x)$ называются *эквивалентными.*

1.2.2. Основные теоремы о пределах

Этот пункт посвящен основным свойствам пределов функций, которые позволяют вычислять пределы функций, определяемых алгебраическими действиями над переменной. В приводимых ниже теоремах будем считать, что функции $f(x)$, $g(x)$ имеют общую область определения, содержащую точку x_0, и обладают пределами в этой точке.

Теорема 1 Предел суммы двух функций равен сумме их пределов:

$$\lim_{x\to x_0}(f(x)+g(x))=\lim_{x\to x_0}f(x)+\lim_{x\to x_0}g(x).$$

Теорема 2 Предел произведения двух функций равен произведению их пределов.

$$\lim_{x\to x_0}(f(x)\cdot g(x))=\lim_{x\to x_0}f(x)\cdot\lim_{x\to x_0}g(x).$$

Следствие. Постоянный множитель можно выносить за знак предела:

$$\lim_{x\to x_0}(cf(x))=c\lim_{x\to x_0}f(x).$$

Теорема 3 Предел отношения двух функций равен отношению пределов этих функций, если предел знаменателя отличен от 0:

$$\lim_{x\to x_0}\frac{f(x)}{g(x)}=\frac{\lim\limits_{x\to x_0}f(x)}{\lim\limits_{x\to x_0}g(x)}.$$

Теорема 4 Предел положительной функции неотрицателен.

Сформулируем некоторые признаки существования предела функции.

Теорема 5 Если функция $h(x)$ заключена между двумя функциями $f(x)$ и $g(x)$ и пределы $f(x)$ и $g(x)$ при x, стремящемся к x_0, равны между собой, то $\lim\limits_{x\to x_0} h(x)$ существует и совпадает с ними, т. е. если $f(x) \leqslant g(x) \leqslant h(x)$и $\lim\limits_{x\to x_0} f(x) = \lim\limits_{x\to x_0} g(x) = b$, то $\lim\limits_{x\to x_0} h(x) = b$.

Теорема 6 Если функция $h(x)$ монотонна и ограничена на интервале, содержащем точку x_0, то она имеет в этой точке односторонние пределы.

Эти утверждения справедливы и при x, стремящемся к ∞.

1.2.3. Специальные пределы

Если при применении основных теорем о пределах функции (т. 1, 2, 3 в п. 1.2.2) получаем выражения вида $[0/0]$, $[\infty/\infty]$, $[\infty - \infty]$, $[0 \cdot \infty]$, $[1^\infty]$, $[0^0]$, $[\infty^0]$, которые носят название неопределенностей, то требуются специальные методы для получения ответа (раскрытие неопределенности).

Для решения примеров используются следующие пределы:

$$\lim_{x\to 0} \frac{\sin x}{x} = 1, \tag{1.19}$$

$$\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x = \lim_{\alpha\to 0}(1+\alpha)^{1/\alpha} = e = 2{,}71828\ldots \tag{1.20}$$

(первый и второй классические пределы).

Напоминаем (см. п. 1.1.5), что логарифм числа x по основанию e называется *натуральным логарифмом* и обозначается $\ln x$.

При решении примеров полезно иметь в виду следующие равенства:

$$\begin{gathered} \lim_{x\to 0} \frac{\ln(1+x)}{x} = 1, \qquad \lim_{x\to 0} \frac{a^x - 1}{x} = \ln a, \\ \lim_{x\to 0} \frac{e^x - 1}{x} = 1. \end{gathered} \tag{1.21}$$

Самостоятельного названия эти пределы не имеют, поскольку могут быть выведены из второго классического предела, записанного во второй форме:

$$\lim_{x\to 0}(1+x)^{1/x} = e.$$

1.2.4. Примеры нахождения некоторых пределов

1. Найти $\lim\limits_{x\to 3}(x^2-7x+4)$.

Решение. Применяя теоремы о пределах (теорема 1) и заменяя в аналитическом выражении x его предельным значением, получаем

$$\lim_{x\to 3}(x^2-7x+4)=\lim_{x\to 3}x^2-7\lim_{x\to 3}x+4=9-21+4=-8.$$

2. Найти $\lim\limits_{x\to 3}\dfrac{x^2-9}{x^2-3x}$.

Решение. Числитель и знаменатель дроби стремится к нулю при x стремящемся к 3 (принято говорить, что получается неопределенность вида $[0/0]$). Имеем

$$\lim_{x\to 3}\frac{x^2-9}{x^2-3x}=\left[\frac{0}{0}\right]=\lim_{x\to 3}\frac{(x-3)(x+3)}{x(x-3)}.$$

Так как в определении предела функции отмечено, что при отыскании предела функции значения функции в предельной точке можно не рассматривать, то $x-3\neq 0$, вследствие чего можно разделить числитель на знаменатель и получить $\lim\limits_{x\to 3}\dfrac{x+3}{x}=$ $=\dfrac{3+3}{3}=2$.

3. Найти $\lim\limits_{x\to 0}\dfrac{\sqrt{x+4}-2}{x}$ (неопределенность $[0/0]$).

Решение. Умножим числитель и знаменатель дроби на сумму $\sqrt{x+4}+2$ (сопряженное выражение). Воспользуемся известной формулой алгебры $a^2-b^2=(a-b)(a+b)$. Получим

$$\lim_{x\to 0}\frac{(\sqrt{x+4}-2)(\sqrt{x+4}+2)}{x(\sqrt{x+4}+2)}=\lim_{x\to 0}\frac{x+4-4}{x(\sqrt{x+4}+2)}=$$
$$=\lim_{x\to 0}\frac{1}{\sqrt{x+4}+2}=\frac{1}{4}.$$

4. Найти $\lim\limits_{x\to 3}\dfrac{\sqrt{3x+7}-\sqrt{2x+10}}{\sqrt{4x+13}-\sqrt{x+22}}$ (неопределенность $[0/0]$).

Решение. Числитель и знаменатель дроби нужно одновременно умножить на их сопряженные выражения, т. е. на выражение

$$(\sqrt{3x+7}+\sqrt{2x+10})(\sqrt{4x+13}+\sqrt{x+22}).$$

Получаем

$$\lim_{x\to 3}\frac{\sqrt{3x+7}-\sqrt{2x+10}}{\sqrt{4x+13}-\sqrt{x+22}} = \lim_{x\to 3}\frac{\sqrt{3x+7}-\sqrt{2x+10}}{\sqrt{4x+13}-\sqrt{x+22}}\times$$
$$\times\frac{(\sqrt{3x+7}+\sqrt{2x+10})(\sqrt{4x+13}+\sqrt{x+22})}{(\sqrt{3x+7}+\sqrt{2x+10})(\sqrt{4x+13}+\sqrt{x+22})} =$$
$$= \lim_{x\to 3}\frac{(\sqrt{4x+13}+\sqrt{x+22})}{3(\sqrt{3x+7}+\sqrt{2x+10})} = \frac{5}{12}.$$

5. Найти $\lim\limits_{x\to 7}\dfrac{\sqrt[3]{x-6}-1}{x-7}$ (неопределенность $[0/0]$).

Решение. Воспользуемся известной формулой алгебры $(a-b)(a^2+ab+b^2)=a^3-b^3$. Положим $a=\sqrt[3]{x-6}$, $b=1$, значит, чтобы получить в числителе разность кубов, надо его умножить на $(\sqrt[3]{(x-6)^2}+\sqrt[3]{x-6}+1)$. Умножая числитель и знаменатель на эту величину, получаем

$$\lim_{x\to 7}\frac{\sqrt[3]{x-6}-1}{x-7} = \lim_{x\to 7}\frac{x-7}{(x-7)(\sqrt[3]{(x-6)^2}+\sqrt[3]{x-6}+1)} =$$
$$= \lim_{x\to 7}\frac{1}{\sqrt[3]{(x-6)^2}+\sqrt[3]{x-6}+1} = \frac{1}{1+1+1} = \frac{1}{3}.$$

6. Найти $\lim\limits_{x\to\infty}\dfrac{2x^3+x^2+5}{3x^3+x-1}$ (неопределенность $[\infty/\infty]$).

Решение. Разделим и числитель и знаменатель на высшую степень x, встречающуюся в членах дроби, а после этого перейдем к пределу:

$$\lim_{x\to\infty}\frac{2x^3+x^2+5}{3x^3+x-1} = \lim_{x\to\infty}\frac{2+1/x+5/x^3}{3+1/x^2-1/x^3} =$$
$$= \frac{\lim\limits_{x\to\infty}2+\lim\limits_{x\to\infty}\dfrac{1}{x}+\lim\limits_{x\to\infty}\dfrac{5}{x^3}}{\lim\limits_{x\to\infty}3+\lim\limits_{x\to\infty}\dfrac{1}{x^2}-\lim\limits_{x\to\infty}\dfrac{1}{x^3}} = \frac{2}{3},$$

так как при $x \to \infty$, $1/x$ — величина бесконечно малая, $1/x^2 = 1/x \cdot 1/x$ и $1/x^3 = 1/x \cdot 1/x \cdot 1/x$ тоже бесконечно малые (что следует из свойств бесконечно малых величин), и пределы этих величин равны нулю, когда $x \to \infty$. Теперь возможно применить теорему о пределе частного.

7. Найти $\lim\limits_{x\to\infty}\left(\dfrac{x^3}{5x^2+1} - \dfrac{3x^2}{15x+1}\right)$ (неопределенность $[\infty - \infty]$).

Решение. Приведем выражение к общему знаменателю, а затем разделим числитель и знаменатель на высшую степень x, встречающуюся в дроби.

$$\lim_{x\to\infty}\left(\frac{x^3}{5x^2+1} - \frac{3x^2}{15x+1}\right) = \lim_{x\to\infty}\frac{15x^4 + x^3 - 15x^4 - 3x^2}{(5x^2+1)(15x+1)} =$$
$$= \lim_{x\to\infty}\frac{x^3 - 3x^2}{(5x^2+1)(15x+1)} = \lim_{x\to\infty}\frac{1 - 3/x}{(5 + \frac{1}{x^2})(15 + 1/x)} = \frac{1}{75}.$$

8. Найти $\lim\limits_{x\to\infty}\dfrac{3x^4 - 2}{\sqrt{x^8 + 3x + 2}}$ (неопределенность$[\infty/\infty]$).

Решение. Разделим числитель и знаменатель на высшую степень x, встречающуюся в примере, т. е. на x^4:

$$\lim_{x\to\infty}\frac{3x^4 - 2}{\sqrt{x^8 + 3x + 2}} = \lim_{x\to\infty}\frac{3 - \frac{2}{x^4}}{\sqrt{1 + \frac{3}{x^7} + \frac{4}{x^8}}} = \frac{3}{1} = 3.$$

9. Найти $\lim(\sqrt{x^2 + 8x + 3} - \sqrt{x^2 + 4x + 3})$ (неопределенность $[\infty - \infty]$).

Решение. Умножим и разделим рассматриваемое выражение на сопряженное:

$$\lim_{x\to\infty}(\sqrt{x^2 + 8x + 3} - \sqrt{x^2 + 4x + 3}) =$$
$$= \lim_{x\to\infty}(\sqrt{x^2 + 8x + 3} - \sqrt{x^2 + 4x + 3}) \times$$
$$\times \frac{(\sqrt{x^2 + 8x + 3} + \sqrt{x^2 + 4x + 3})}{\sqrt{x^2 + 8x + 3} + \sqrt{x^2 + 4x + 3}} =$$

$$= \lim_{x\to\infty} \frac{x^2+8x+3-x^2-4x-3}{\sqrt{x^2+8x+3}+\sqrt{x^2+4x+3}} =$$

$$= \lim_{x\to\infty} \frac{4x}{\sqrt{x^2+8x+3}+\sqrt{x^2+4x+3}} =$$

$$= \lim_{x\to\infty} \frac{4}{\sqrt{1+8/x+3/x^2}+\sqrt{1+4/x+3/x^2}} = \frac{4}{2} = 2.$$

10. Найти $\lim\limits_{x\to 0} \dfrac{\sin 4x}{x}$ (неопределенность $[0/0]$).

Решение. Используем I замечательный предел.

$$\lim_{x\to 0} \frac{\sin 4x}{x} = \lim_{x\to 0} \frac{\sin 4x}{4x} \cdot 4 = 4.$$

11. Найти $\lim\limits_{h\to 0} \dfrac{\cos(x+h)-\cos(x-h)}{h}$ (неопределенность $[0/0]$).

Решение. Применяя формулу

$$\cos\alpha - \cos\beta = -2\sin\frac{\alpha+\beta}{2} \times \sin\frac{\alpha-\beta}{2},$$

получим

$$\lim_{h\to 0} \frac{\cos(x+h)-\cos(x-h)}{h} = \lim_{h\to 0} \frac{-2\sin x \sin h}{h} =$$

$$= -2\sin x \lim_{h\to 0} \frac{\sin h}{h} = -2\sin x.$$

12. Найти $\lim\limits_{x\to 0} \dfrac{\arcsin x}{x}$ (неопределенность $[0/0]$).

Решение. Заменим $\arcsin x$ на y; если $x \to 0$, то $\arcsin x \to 0$, $y \to 0$ и $\arcsin x = y$, отсюда $\sin y = x$.

$$\lim_{x\to 0} \frac{\arcsin x}{x} = \lim_{y\to 0} \frac{y}{\sin y} = 1.$$

13. Найти $\lim\limits_{x\to 0}(1+kx)^{1/x}$, где k — любое число, не равное нулю.

Решение. Произведем замену переменной $kx = y$, $x \to 0$, $y \to 0$, $x = y/k$.

$$\lim_{x\to 0}(1+kx)^{1/x} = \lim_{y\to 0}(1+y)^{k/y} = \left[\lim_{y\to 0}(1+y)^{1/y}\right]^k = e^k.$$

Аналогично можно доказать равенство: $\lim\limits_{x\to\infty}(1+k/x)^x = e^k$.

14. Найти $\lim\limits_{x\to\infty}\left(\dfrac{2x-1}{2x+1}\right)^{2x}$ (неопределенность $[1^\infty]$).

Решение.

$$\lim_{x\to\infty}\left(\frac{2x-1}{2x+1}\right)^{2x} = \lim_{x\to\infty}\left(\frac{1-1/2x}{1+1/2x}\right)^{2x} =$$

$$= \frac{\left[\lim\limits_{x\to\infty}(1-1/2x)^x\right]^2}{\left[\lim\limits_{x\to\infty}(1+1/2x)^x\right]^2} = \frac{[e^{-1/2}]^2}{[e^{1/2}]^2} = \frac{e^{-1}}{e^1} = e^{-2}.$$

15. Найти $\lim\limits_{n\to\infty} n\,[\ln(n+3)-\ln n]$ (неопределенность $[\infty-\infty]$).

Решение.

$$\lim_{n\to\infty} n\,[\ln(n+3)-\ln n] = \lim_{n\to\infty} n\ln\frac{n+3}{n} =$$

$$= \lim_{n\to\infty} n\ln(1+\frac{3}{n}) = \lim_{n\to\infty}\ln(1+\frac{3}{n})^n = \ln e^3 = 3.$$

16. Найти $\lim\limits_{x\to 0}\dfrac{\ln(1+mx)}{x}$ (неопределенность $[0/0]$).

Решение.

$$\lim_{x\to 0}\frac{\ln(1+mx)}{x} = \lim_{x\to 0}\frac{1}{x}\ln(1+mx) =$$

$$= \lim_{x\to 0}\ln(1+mx)^{1/x} = \ln\lim_{x\to 0}(1+mx)^{1/x} = \ln e^m = m.$$

17. Найти $\lim\limits_{x\to 0}\dfrac{\sin 3x-\sin x}{\ln(1+3x)}$ (неопределенность $[0/0]$).

Решение.

$$\lim_{x\to 0}\frac{\sin 3x-\sin x}{\ln(1+3x)} = \lim_{x\to 0}\frac{2\sin x\cos 2x}{\ln(1+3x)} =$$

$$= 2\lim_{x\to 0}\frac{\sin x}{x}\times\lim_{x\to 0}\frac{x}{\ln(1+3x)}\times\lim_{x\to 0}\cos 2x =$$

$$= 2\times 1\times\frac{1}{3}\times 1 = \frac{2}{3}.$$

18. Найти $\lim\limits_{x\to 0}\dfrac{1-5^x}{1-e^x}$ (неопределенность [0/0]).

Решение.

$$\lim_{x\to 0}\frac{1-5^x}{1-e^x} = \lim_{x\to 0}\frac{5^x-1}{x}\lim_{x\to 0}\frac{x}{e^x-1} = \ln 5.$$

19. Найти $\lim\limits_{x\to e}\dfrac{\ln x - 1}{x-e}$.

Решение. Произведем замену переменной $x - e = y$, $x\to e$, $y\to 0$, $x = y + e$. Получим

$$\lim_{x\to e}\frac{\ln x-1}{x-e} = \lim_{y\to 0}\frac{\ln(y+e)-\ln e}{y} = \lim_{y\to 0}\frac{\ln\dfrac{y+e}{e}}{y} =$$

$$= \lim_{y\to 0}\frac{\ln(1+y/e)}{y} = \frac{1}{e}.$$

Самостоятельная работа

1. $\lim\limits_{x\to -1}\dfrac{x^2-x+1}{2x^3-x^2+x+2}$. **2.** $\lim\limits_{x\to 3}\dfrac{x^2+x-12}{2x^2-9x+9}$.

3. $\lim\limits_{x\to 1}\dfrac{x^3-x^2+1}{x^3-3x+2}$. **4.** $\lim\limits_{x\to 0}\dfrac{5x^3-6x^2}{4x^5+2x^3+x^2}$.

5. $\lim\limits_{x\to 2}\dfrac{x^3-8}{x^2-6x+8}$. **6.** $\lim\limits_{x\to 3}\dfrac{\sqrt{x+6}-3}{x-3}$.

7. $\lim\limits_{x\to 1}\dfrac{\sqrt{5-x^2}-2}{1-x}$. **8.** $\lim\limits_{x\to 1}\dfrac{\sqrt{x+8}-\sqrt{8x+1}}{\sqrt{5-x}-\sqrt{7x-3}}$.

9. $\lim\limits_{x\to 4}\dfrac{\sqrt{x+5}-\sqrt{3x^2-39}}{\sqrt{x^2-3}-\sqrt{2x^2-19}}$. **10.** $\lim\limits_{x\to 0}\dfrac{\sqrt[3]{1+x}-\sqrt[3]{3x+1}}{6x}$.

11. $\lim\limits_{x\to 0}\dfrac{\sqrt[3]{1+x^3}-\sqrt[3]{1+x^2}}{\sqrt{1+2x}-\sqrt{1-2x}}$. **12.** $\lim\limits_{x\to 0}\dfrac{\sqrt{1+3x^4}-\sqrt{1-2x}}{x+x^2+2x^3}$.

13. $\lim\limits_{x\to\infty} \dfrac{5x^4+7x^2-5x+4}{x^4+x^2+x+1}$. **14.** $\lim\limits_{x\to\infty} \dfrac{x^3+x+2}{x^2-x+1}$.

15. $\lim\limits_{x\to\infty} \left(\dfrac{5x^2-x}{x^2-3}-\dfrac{3x^3-4}{x^3-x}\right)^4$.

16. $\lim\limits_{x\to\pm\infty} \left(\sqrt{x^2+4}-\sqrt{x^2-4}\right)$.

17. $\lim\limits_{x\to\pm\infty} \dfrac{\sqrt{x^2+4}}{x+2}$. **18.** $\lim\limits_{x\to\infty} \left((x+1)^{2/3}-(x-1)^{2/3}\right)$.

19. $\lim\limits_{x\to\pm\infty} \left(\sqrt{x^2+x+1}-\sqrt{x^2-x+1}\right)$.

20. $\lim\limits_{x\to\infty} x\left(\sqrt[3]{x^2+1}-x^{2/3}\right)$. **21.** $\lim\limits_{x\to 0} \dfrac{1-\cos 5x}{1-\cos 3x}$.

22. $\lim\limits_{x\to 0} \dfrac{\operatorname{tg} x-\sin x}{x^3}$. **23.** $\lim\limits_{x\to 0} \dfrac{\operatorname{arctg} x}{x}$.

24. $\lim\limits_{x\to 1/2} \dfrac{\arcsin(1-2x)}{4x^2-1}$. **25.** $\lim\limits_{h\to 0} \dfrac{\sin(x+h)-\sin(x-h)}{h}$.

26. $\lim\limits_{x\to 0} \dfrac{\sin 4x}{\sqrt{1+x}-1}$. **27.** $\lim\limits_{x\to 0} \dfrac{1-\cos 2x+\operatorname{tg}^2 x}{x\cdot\sin x}$.

28. $\lim\limits_{x\to 1} (1-x)\operatorname{tg}\dfrac{\pi}{2}x$. **29.** $\lim\limits_{x\to\pi/2} \left(\dfrac{\sin x}{\cos^2 x}-\operatorname{tg}^2 x\right)$.

30. $\lim\limits_{x\to\pi} \dfrac{\cos\dfrac{x}{2}}{x-\pi}$. **31.** $\lim\limits_{x\to 1} \dfrac{\sin(1-x)}{\sqrt{x}-1}$.

32. $\lim\limits_{x\to -1} \cos\dfrac{\pi(x+1)}{\sqrt[3]{x}+1}$. **33.** $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x(\sqrt{1+x}-1)}$.

34. $\lim\limits_{x\to\infty} x\cdot\sin\dfrac{1}{x}$. **35.** $\lim\limits_{x\to\infty} \left(\dfrac{x^2+1}{x^2}\right)^{x^2+1}$.

36. $\lim\limits_{x\to\infty} \left(1+\dfrac{3}{x}\right)^x$. **37.** $\lim\limits_{x\to\infty} \left(\dfrac{x+8}{x-2}\right)^x$.

38. $\lim\limits_{x\to\infty} \left(\dfrac{3x-2}{3x+1}\right)^{2x-1}$. **39.** $\lim\limits_{x\to 0} (1+\sin x)^{1/x}$.

40. $\lim\limits_{x\to 0}(1+\sin^2 x)^{1/x}$. **41.** $\lim\limits_{n\to+\infty}\dfrac{1-10^n}{1+10^{n+1}}$.

42. $\lim\limits_{x\to 0}\dfrac{1-10^x}{1-e^{-x}}$. **43.** $\lim\limits_{x\to 0}\dfrac{\cos 4x-\cos 2x}{\ln(1+x)}$. **44.** $\lim\limits_{x\to 0}\dfrac{5^{2x}-1}{\sin x}$.

45. $\lim\limits_{n\to+\infty} n[\ln n-\ln(n+2)]$. **46.** $\lim\limits_{x\to 0}\dfrac{\ln(1+2x)}{\sin 3x}$.

47. $\lim\limits_{t\to 0}\dfrac{t}{\ln(1+xt)}$. **48.** $\lim\limits_{t\to\infty} t\,(\sqrt[t]{a}-1)\quad (t>0)$.

49. $\lim\limits_{x\to 0}\dfrac{5^x-4^x}{x^2+x}$. **50.** $\lim\limits_{x\to\infty}\left(\dfrac{x+a}{x+b}\right)^{x+c}$.

§ 1.3. Непрерывность функций

1.3.1. Основные понятия и определения

Рассмотрим функцию $y=f(x)$, определенную в интервале $[a,b]$. Пусть x_0 и x — два произвольных значения из этого интервала. Обозначим $x-x_0=\Delta x$, откуда $x=x_0+\Delta x$. Говорят, что для перехода от значения аргумента x_0 к значению x первоначальному значению придано приращение Δx. *Приращением Δy функции $y=f(x)$*, соответствующим приращению Δx аргумента x в точке x_0, называется разность

$$\Delta y=f(x_0+\Delta x)-f(x_0). \tag{1.22}$$

Пример. Найти приращение функции $y=x^2$, соответствующее приращению Δx аргумента x в точке x_0.

Решение. По формуле (1.22)

$$(x_0+\Delta x)^2-x_0^2=(x_0+\Delta x-x_0)(x_0+\Delta x+x_0)=(2x_0+\Delta x)\cdot\Delta x.$$

С понятием предела функции тесно связано еще одно понятие математического анализа — понятие непрерывности функции.

Определение. Говорят, что функция $f(x)$ *непрерывна в точке* $x=x_0$, если

$$\lim_{x\to x_0} f(x)=f(x_0). \tag{1.23}$$

Определение непрерывности функции в точке можно сформулировать также следующим образом.

Определение. Функция называется непрерывной в данной точке, если в этой точке бесконечно малому приращению аргумента соответствует бесконечно малое приращение функции, т. е.

$$\lim_{\Delta x \to 0} \Delta y = 0. \tag{1.24}$$

При этом предполагается, что бесконечно малая величина Δx пробегает лишь те значения, для которых $f(x_0 + \Delta x)$ имеет смысл.

Оба определения непрерывности функции в точке эквивалентны.

Определение. Функция $f(x)$ называется *непрерывной в интервале*, если эта функция непрерывна в каждой точке этого интервала.

Для функции, непрерывной в интервале (a, b), для каждого значения x_0 из интервала (a, b) выполнено равенство (1.24). Так как $x_0 = \lim\limits_{x \to x_0} x$, то отсюда получим $\lim\limits_{x \to x_0} f(x) = f(\lim\limits_{x \to x_0} x)$, т. е. если функция непрерывна, то знаки предела и функции можно менять местами.

Отметим, что, вычисляя предел $f(x)$, мы можем приближать x к x_0 как справа, так и слева, т. е. вычислять

$$\lim_{x \to x_0 - 0} f(x) \quad \text{и} \quad \lim_{x \to x_0 + 0} f(x).$$

Непрерывность функции в точке x_0 равносильна ее непрерывности в этой точке одновременно слева и справа, то есть должны выполняться следующие четыре условия непрерывности.

1. $f(x)$ должна быть определена в некоторой окрестности точки x_0.

2. Должны существовать конечные пределы слева и справа: $\lim\limits_{x \to x_0 - 0} f(x)$, $\lim\limits_{x \to x_0 + 0} f(x)$.

3. Эти пределы слева и справа должны быть равны: $\lim\limits_{x \to x_0 - 0} f(x) = \lim\limits_{x \to x_0 + 0} f(x) = A$.

4. Эти пределы должны быть равны значению функции в точке x_0, т. е. $A = f(x_0)$.

Если в точке x_0 не выполняется хотя бы одно условие, то в этой точке функция терпит разрыв, а сама эта точка называется точкой разрыва.

Основные теоремы о непрерывности

Теорема 7 Если две функции $f(x)$ и $g(x)$ определены в одном и том же промежутке и обе непрерывны в точке x_0 из этого промежутка, то в той же точке будут непрерывны и функции

$$f(x) \pm g(x), \quad f(x) \cdot g(x), \quad \frac{f(x)}{g(x)} \quad (g(x_o) \neq 0).$$

Доказательство непосредственно вытекает из теорем 1–3 из п. 1.2.2 о пределе суммы, разности, произведения и частного двух функций.

Теорема 8 Все основные элементарные функции непрерывны там, где они определены.

Доказательство. Докажем для примера непрерывность функции $y = \sin x$ при любом значении $x = x_0$.

$$\Delta y = \sin(x_0 + \Delta x) - \sin x_0 = 2\cos\left(x_0 + \frac{\Delta x}{2}\right) \cdot \sin\frac{\Delta x}{2}.$$

Отсюда

$$\lim_{\Delta x \to 0} \Delta y = \lim_{\Delta x \to 0} 2\cos\left(x_0 + \frac{\Delta x}{2}\right) \cdot \sin\frac{\Delta x}{2} \cdot \frac{\Delta x}{\Delta x} =$$
$$= \lim_{\Delta x \to 0} \cos\left(x_0 + \frac{\Delta x}{2}\right) \cdot \Delta x = 0.$$

Функция $y = \cos x$ ограниченная, а произведение ограниченной величины на бесконечно малую есть бесконечно малая величина (см. свойство 3 бесконечно малых функций), и согласно первому замечательному пределу (см. (1.19)) $\lim\limits_{\Delta x \to 0} \dfrac{\sin\frac{\Delta x}{2}}{\Delta x} = \dfrac{1}{2}$.

Согласно определению 2 функция $y = \sin x$ непрерывна в точке x_0. Поскольку x_0 произвольная точка промежутка $(-\infty, +\infty)$, то по определению она непрерывна на всей оси.

Теорема 9 Пусть сложная функция $f(\varphi(x))$ определена на множестве X (см. п. 1.1.4). Если функция $U = \varphi(x)$ непрерывна в точке $x_0 \in X$, а $f(U)$ непрерывна в соответствующей точке

$U_0 = \varphi(x_0)$, то сложная функция $f(\varphi(x))$ будет непрерывна в точке x_0.

Свойства функций, непрерывных на отрезке

1) Функция $f(x)$, непрерывная в точке x_0 и не равная нулю в этой точке, сохраняет знак $f(x_0)$ в некоторой окрестности этой точки.

2) Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и на концах его принимает значения разных знаков, то между точками a и b найдется по крайней мере одна точка c такая, что $f(c) = 0$.

Это утверждение имеет очень простой геометрический смысл: если непрерывная кривая переходит с одной стороны оси OX на другую, то она пересекает эту ось. Им можно пользоваться для приближенного вычисления корней алгебраических уравнений.

Пример. Пусть $f(x) = x^4 - x - 1$. Так как $f(1) = -1$, $f(2) = 13$, то многочлен имеет корень между 1 и 2. Разделим этот промежуток пополам и найдем значение $f(1{,}5) = 2{,}56$. Следовательно, корень лежит в промежутке $(1,\ 1{,}5)$. Найдем значение функции $f(1{,}25) = 0{,}19$. Следовательно, корень в промежутке $(1,\ 1{,}78)$. Далее, $f(1{,}39) = 1{,}34$, следовательно, корень в промежутке $(1,\ 1{,}25)$; $f(1{,}125) = -0{,}52$, следовательно, корень в промежутке $(1{,}125,\ 1{,}25)$. $f(1{,}19) = -0{,}2$, т. е. корень в промежутке $(1{,}19,\ 1{,}125)$, $f(1{,}22) = -0{,}004$, т. е. корень в $(1{,}22,\ 1{,}25)$, т. е. $x \approx 1{,}235$.

3. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она ограничена на нем, т. е. существует такое положительное число M, что $|f(x)| \leqslant M$ при $a \leqslant x \leqslant b$.

4. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то на этом отрезке найдутся точки x_1 и x_2 такие, что значение функций $f(x_1)$, $f(x_2)$ будут соответственно наибольшим и наименьшим из всех значений функции $f(x)$ на отрезке $[a, b]$.

Определение. Если функция $f(x)$ не является непрерывной в точке x_0, то говорят, что в точке x_0 функция $f(x)$ *разрывна*, а точка x_0 называется *точкой разрыва* функции.

В качестве конкретного примера функции, имеющей точку разрыва, рассмотрим изменение биомассы микроорганизмов, чувствительных к температурным колебаниям.

При изменении температуры общая биомасса m, как правило, увеличивается — тепло способствует размножению. Однако когда температура слишком высока, практически все колонии гибнут; значение m, скачкообразно меняясь, становится равным нулю. Примерно то же самое происходит и при понижении температуры: как только она достигнет некоторого нижнего предела, микроорганизмы погибают. В реальных условиях температура меняется в зависимости от времени, то повышаясь, то понижаясь. Поэтому графическим изображением изменения биомассы в зависимости от времени может быть разрывная кривая. Точки разрыва t_1, t_2, t_3 соответствуют тем моментам времени, когда температура стала слишком высокой или слишком низкой.

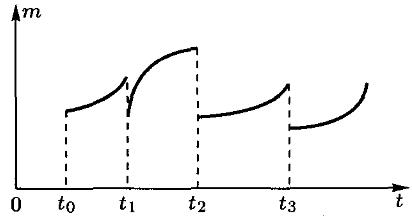


Рис. 1.22. Зависимость биомассы микроорганизмов от температуры.

Классификация точек разрыва

Пусть x_0 является внутренней точкой отрезка $[a, b]$. Если существуют конечные пределы $f(x)$ при стремлении x к x_0 слева и справа, но нарушены условия 3 или 4, то из определения непрерывности функции (см. п. 1.3.1) точку x_0 называют *точкой разрыва первого рода.*

Если хотя бы один из пределов слева или справа бесконечен или его вовсе нет, тогда говорят о *разрыве второго рода.*

Легко понять, при каких обстоятельствах для функции $f(x)$ в точке a справа появляется разрыв. Может случиться, что, хотя конечный предел $f(a+0)$ и существует, но он не равен значению $f(a)$. Аналогично, в точке b появляется разрыв, если существует конечный предел $f(b-0)$, но он не равен $f(b)$.

Если существуют конечные пределы функции $f(x)$ при стремлении x к x_0 слева и справа, они равны между собой, но не равны значению функции в этой точке, то восстановить непрерывность в точке x_0 можно, положив $f(x_0) = f(x_0 - 0) = f(x_0 + 0)$. В этом случае точка x_0 называется устранимой точкой разрыва.

Если функция $f(x)$ в точке x_0 имеет разрыв второго рода, в этом случае, как бы дополнительно не определять функцию при $x = x_0$, она неизбежно будет иметь разрыв.

Примеры.

1. Указать точку разрыва функции $f(x) = \dfrac{2}{x-3}$, найти $\lim\limits_{x\to 3\mp 0} f(x)$, определить какого рода разрыв терпит функция в этой точке.

Решение.

$$f(3+0) = \lim_{x\to 3+0} f(x) = \lim_{x\to 3+0} \frac{2}{x-3} = +\infty,$$

$$\lim(3-0) = \lim_{x\to 3-0} f(x) = \lim_{x\to 3-0} \frac{2}{x-3} = -\infty.$$

В точке $x = 3$ разрыв 2-го рода.

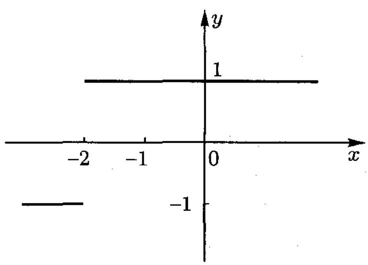


Рис. 1.23. График функции к примеру 2.

2. Построить график функции $y = \dfrac{x+2}{|x+2|}$. Какие из условий непрерывности в точках разрыва выполнены и какие не выполнены?

Решение.

$$y = \begin{cases} 1, & x > -2, \\ -1, & x < -2, \end{cases}$$

$$\lim_{x\to -2+0} f(x) = 1, \qquad \lim_{x\to -2-0} f(x) = -1.$$

Конечные пределы слева и справа существуют, но не равны между собой. В точке $x = -2$ разрыв 1-го рода.

3. Построить график функции

$$y = \begin{cases} x^2, & |x| < 1, \\ 2, & |x| = 1, \\ 3, & |x| > 1 \end{cases}$$

и указать точки ее разрыва.

Решение.

$$\lim_{x\to -1+0} f(x) = 1, \qquad \lim_{x\to -1-0} f(x) = 3, \qquad f(-1) = 2.$$

Пределы слева и справа конечны, но не равны между собой. И не равны значению функции в этой точке. В точке $x = -1$ разрыв 1-го рода.

Аналогично, в точке $x = 1$ разрыв 1-го рода.

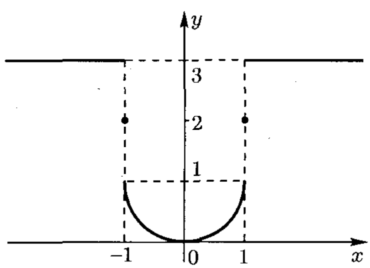


Рис. 1.24. График функции к примеру 3.

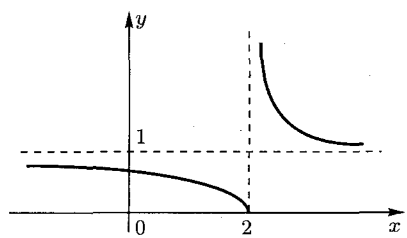


Рис. 1.25. График функции к примеру 4.

4. Найти точку разрыва и построить график функции $y = 2^{\frac{1}{x-2}}$.

Решение.

$$\lim_{x\to 2+0} 2^{\frac{1}{x-2}} = \infty \qquad \lim_{x\to 2-0} 2^{\frac{1}{x-2}} = 0.$$

В точке $x = 2$ разрыв 2-го рода.

5. Пусть $f(x) = \begin{cases} e^x, & x < 0, \\ a + x, & x \geqslant 0. \end{cases}$ При каком выборе числа a функция будет непрерывной?

Решение.

$$\lim_{x\to-0} f(x) = \lim_{x\to-0} e^x = 1,$$

$$\lim_{x\to+0} f(x) = \lim_{x\to+0} (a+x) = a.$$

Функция будет непрерывна при $a = 1$.

Самостоятельная работа

1. Пусть $f(x) = \begin{cases} x+2, & x \leqslant 2, \\ 4-ax^2, & x > 2. \end{cases}$ При каком выборе числа a функция будет непрерывной? (Построить ее график.)

2. Исследовать характер разрыва функции $f(x) = \dfrac{3^{1/x}-1}{3^{1/x}+1}$ в точке $x = 0$.

3. Исследовать непрерывность функции $y = \dfrac{1}{1+3^{\operatorname{tg} x}}$.

4. Исследовать характер разрыва функции $y = 2^{1/(1-x)}$.

5. Используя свойства непрерывных функций, убедиться в том, что уравнение $x^5 - 4x = 1$ имеет по меньшей мере один корень, заключенный между 1 и 2.

6. Исследовать непрерывность и построить график функции $y = 5^{|x|/x}$.

7. Найти точку разрыва функции $y = 4 \operatorname{arctg} \dfrac{4}{x-4}$.

8. Исследовать на непрерывность функцию и изобразить графически $f(x) = \begin{cases} \dfrac{1}{(1+x)^2}, & x \neq -1, \\ C, & x = -1, \end{cases}$ где C — произвольная постоянная.

9. Исследовать на непрерывность $f(x) = \begin{cases} \dfrac{\sin x}{|x|}, & x \neq 0, \\ 1, & x = 0. \end{cases}$

10. $f(x) = \begin{cases} \left|\dfrac{\sin x}{x}\right|, & x \neq 0, \\ 1, & x = 0. \end{cases}$

11. Определить точки разрыва функции $f(x) = \dfrac{x}{(1+x)^2}$ и исследовать характер этих точек.

12. Определить точки разрыва функции $f(x) = \dfrac{x^2 - 1}{x^3 - 3x + 2}$ и исследовать характер этих точек.

13. Исследовать на непрерывность и нарисовать эскиз графика функции $f(x) = e^{1/x}$.

14. Исследовать на непрерывность и построить график функции $f(x) = 1 - e^{-1/x^2}$.

15. Является ли непрерывной функция

$$f(x) = \begin{cases} 2x, & 0 \leqslant x \leqslant 1, \\ 2 - x, & 1 < x \leqslant 2? \end{cases}$$

ГЛАВА 2
ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ

§ 2.1. Производная функции

2.1.1. Задачи, приводящие к понятию производной

Задача 1 (о скорости движущейся точки).

Пусть материальная точка движется по прямолинейной траектории, которую примем за ось OX (рис. 2.1).

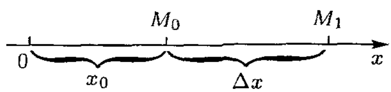


Рис. 2.1. Движение точки вдоль прямолинейной траектории.

Положение точки на траектории будет тогда определяться ее абсциссой x, которая является функцией времени t: $x = f(t)$ (последнее равенство называется уравнением движения точки).

Пусть в момент времени t_0 движущаяся точка занимала на траектории положение M_0 и имела абсциссу x_0 , а по прошествии времени Δt переместилась в положение M_1 и имеет абсциссу $x_0 + \Delta x$. Таким образом, если за время Δt точка не меняла направление движения, то $|\Delta x|$ — путь, пройденный точкой за время Δt. Очевидно, что $\Delta x = f(t_0 + \Delta t) - f(t_0)$.

Назовем *средней скоростью* движения точки за время Δt отношение $\Delta x / \Delta t$.

Определение. *Мгновенной скоростью* движения точки в момент времени t_0 назовем предел v, к которому стремится средняя скорость точки за промежуток времени Δt, когда $\Delta t \to 0$:

$$v = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = \lim_{\Delta t \to 0} \frac{f(t_0 + \Delta t) - f(t)}{\Delta t}. \qquad (2.1)$$

Пример. Точка свободно падает в пустоте. Вычислить скорость точки в произвольный момент t.

Решение. Уравнение свободного движения (падения) в этом случае имеет вид

$$x = \frac{1}{2}gt^2.$$

Тогда по формуле (2.1) находим

$$v = \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[\frac{1}{2}g(t+\Delta t)^2 - \frac{1}{2}gt^2\right] = \lim_{\Delta t \to 0}\left(gt + \frac{1}{2}g\Delta t\right) = gt.$$

Задача 2 (о проведении касательной к плоской кривой).

Пусть M_0 — некоторая точка данной кривой (рис. 2.2). Возьмем на этой кривой другую точку M и проведем секущую MM_0. Пусть теперь точка M приближается вдоль кривой к точке M_0 по любому закону так, что расстояние между этими точками стремится к нулю. Если при этом секущая MM_0, поворачиваясь вокруг точки M_0, будет приближаться к некоторой прямой M_0T так, что угол между прямыми MM_0 и M_0T будет стремиться к нулю, то в соответствии с определением прямая M_0T будет называться *касательной* к данной кривой в точке M_0.

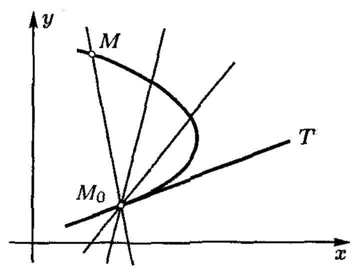


Рис. 2.2. Иллюстрация к определению касательной.

Заметим, что не всякая кривая обладает в каждой своей точке касательной. Прямая M_0T будет касательной к кривой в точке M_0, если секущая MM_0 всегда стремится в указанном смысле к этой единственной прямой M_0T, по какому бы закону точка M ни стремилась по кривой к точке M_0.

На рис. 2.3 представлена кривая, не имеющая касательной в точке M_0.

Здесь секущая MM_0 стремится к прямой M_0N, если $M \to$ $\to M_0$ слева. Если же $M \to M_0$ справа, то секущая MM_0 стремится к другой прямой M_0L.

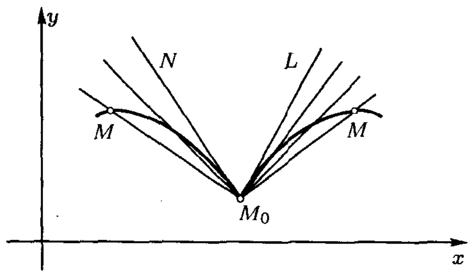


Рис. 2.3. Пример кривой, не имеющей касательной в точке M_0.

Пусть теперь в прямоугольной декартовой системе координат XOY задана кривая своим уравнением $y = f(x)$.

Требуется найти касательную к этой кривой в некоторой ее точке $M_0(x_0, y_0)$ (рис. 2.4).

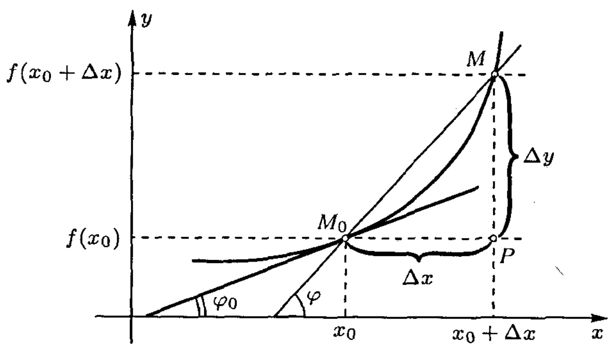


Рис. 2.4. Иллюстрация к задаче о проведении касательной к кривой $y = f(x)$ в точке $M(x_0, y_0)$.

Легко написать искомое уравнение касательной, зная угловой коэффициент κ_0 касательной к кривой в точке M_0.

Возьмем на кривой другую точку $M(x_0+\Delta x, y_0+\Delta y)$ и проведем секущую MM_0. Пусть φ — угол наклона этой секущей к оси OX. Наличие в точке касательной M_0T, образующей с осью OX угол φ_0, очевидно, эквивалентно равенству $\varphi_0 = \lim\limits_{\Delta x \to 0} \varphi$, которое, в свою очередь, эквивалентно равенству

$$\operatorname{tg} \varphi_0 = \lim_{\Delta x \to 0} \operatorname{tg} \varphi. \tag{2.2}$$

Но

$$\operatorname{tg} \varphi = \frac{PM}{M_0 P} = \frac{\Delta y}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x},$$

и, следовательно, формула (2.2) дает

$$\kappa_0 = \operatorname{tg} \varphi_0 = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}. \tag{2.3}$$

Значит, уравнения касательной примет вид

$$y - y_0 = k_0(x - x_0).$$

П р и м е р. Вычислить угловой коэффициент касательной к кривой $y = x^3$ в точке с абсциссой $x_0 = 2$. Написать уравнение касательной.

По формуле (2.3) находим

$$\kappa_0 = \lim_{\Delta x \to 0} \frac{(2 + \Delta x)^3 - 2^3}{\Delta x} = \lim_{\Delta x \to 0} (12 + 6\Delta x + (\Delta x)^2) = 12,$$

и, следовательно, уравнение касательной имеет вид

$$y - 8 = 12(x - 2) \quad (\text{здесь } y_0 = y(x_0) = 2^3 = 8) \text{ или } y = 12x - 16.$$

Заметим, что предел (2.3) имеет ту же структуру, что и предел (2.1). В обоих случаях отыскивается предел отношения приращения функции к приращению аргумента, вызвавшему это приращение, при условии, что приращение аргумента стремится к нулю.

2.1.2. Понятие производной

Пусть $y = f(x)$ определена на некотором промежутке и пусть x_0 — некоторая точка этого промежутка. Пусть Δx — приращение к значению аргумента такое, что $(x_0 + \Delta x)$ не выходит за пределы упомянутого промежутка, а $\Delta y = f(x_0 + \Delta x) - f(x_0)$ — соответствующее приращение функции.

Определение. Если существует $\lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x}$, то этот предел называется производной от функции $y = f(x)$ по переменной x в точке x_0 (обозначения: $\dfrac{dy}{dx}$ или y'_x). Итак:

$$y'_x = \frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}. \tag{2.4}$$

Если предел (2.4) конечен, то производная называется *конечной*, если же этот предел бесконечен, то y'_x — бесконечная производная.

Если конечная производная существует в каждой точке некоторого множества, то она оказывается функцией от x , заданной на этом множестве.

Пример. Найдем производную $y = x^2$ на основании определения производной.

Решение.

$$\Delta y = (x + \Delta x)^2 - x^2 = 2x \cdot \Delta x + (\Delta x)^2.$$

Тогда

$$y'_x = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} (2x + \Delta x) = 2x.$$

Геометрическая и механическая интерпретации производной

Используя результаты предыдущих пунктов, можно сформулировать следующие два предложения, содержащие механический и геометрический смысл производной.

1. Если $x = f(t)$ есть уравнение прямолинейного движения точки, то производная $f'(t)$ представляет собой скорость точки в момент времени t.

Быстрота протекания физических, химических, биологических и других процессов, например скорость охлаждения тела, скорость химической реакции и т. п., также выражается при помощи производной. Поясним это на примерах.

Пример. Предположим, что температура тела U есть убывающая функция времени: $U = f(t)$. Пусть t — фиксированный момент времени. Если t получает приращение Δt, температура U уменьшается на ΔU; тогда отношение $\Delta U / \Delta t$ представляет собой среднюю скорость охлаждения тела. Предел этого отношения при $\Delta t \to 0$, т. е. $\lim\limits_{\Delta t \to 0} \dfrac{\Delta U}{\Delta t} = f'(t)$ выражает скорость охлаждения тела в данный момент t.

Таким образом, скорость охлаждения тела равна производной температуры тела по времени.

Пример. Обозначим через m количество вещества, образовавшегося при химической реакции за промежуток времени t.

Очевидно, m есть функция времени: $m = f(t)$. Если t получает приращение Δt, то m получает приращение Δm. Тогда отношение $\Delta m/\Delta t$ представляет собой скорость химической реакции в данный момент t.

Таким образом, скорость химической реакции равна производной реагирующей массы по времени.

2. Производная $f'(x)$ функции $y = f(x)$ геометрически представляет собой угловой коэффициент касательной к графику этой функции в точке с абсциссой x.

При этом если существует касательная, то существует и производная, и наоборот. Случаю касательной, не параллельной оси

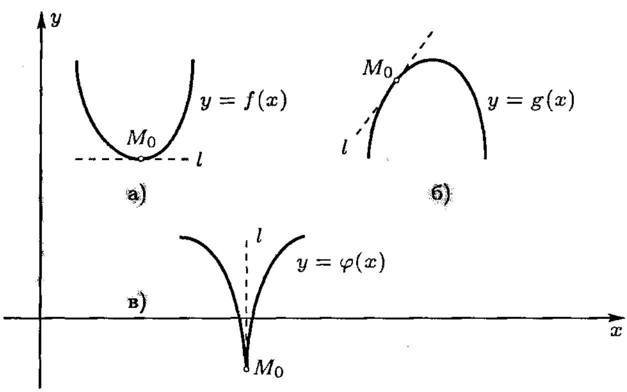


Рис. 2.5. Существование конечной и бесконечной производных: а), б) — конечные производные в точке M_0 (l — касательная, l не параллельна OY); в) — бесконечная производная в точке M_0 ($l \| OY$, $k_0 = \operatorname{tg} \varphi_0 = \operatorname{tg} \pi/2$; $k_0 = \infty$).

OY, отвечает конечная производная, параллельной оси OY — бесконечная производная (рис. 2.5).

Дифференцируемость функции в точке

Определение. Если приращение функции $y = f(x)$ в точке x можно представить в форме

$$\Delta y = A\Delta x + \alpha\Delta x, \tag{2.5}$$

где A не зависит от Δx и α — бесконечно малое в точке x (см. п. 1.2), то эта функция называется дифференцируемой в точке x.

Из последнего равенства следует, что

$$A = \frac{\Delta y}{\Delta x} - \alpha.$$

Перейдя здесь к пределу при $\Delta x \to 0$, получим

$$A = \lim_{\Delta x \to 0}\left(\frac{\Delta y}{\Delta x} - \alpha\right) = \lim_{\Delta x \to 0}\frac{\Delta y}{\Delta x} - \lim_{\Delta x \to 0}\alpha = \lim_{\Delta x \to 0}\frac{\Delta y}{\Delta x} = y'_x.$$

Итак, если $y = f(x)$ дифференцируема в точке x, то приращение этой функции можно представить в виде

$$\Delta y = f'(x)\Delta x + \alpha\Delta x, \tag{2.6}$$

где α — бесконечно малое в точке x.

Отсюда следует, что если функция $y = f(x)$ дифференцируема в точке x, то она обладает в этой точке конечной производной. Покажем, что справедливо и обратное утверждение.

Пусть $f'(x)$ конечна в точке x. По определению

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

Но функцию, стремящуюся к конечному пределу, можно представить в виде суммы этого предела и бесконечно малой функции. Тогда из предыдущего соотношения следует, что

$$\frac{\Delta y}{\Delta x} = f'(x) + \alpha, \qquad \alpha \to 0 \quad \text{при} \quad \Delta x \to 0,$$

из чего сразу же вытекает (2.6), что и означает дифференцируемость функции в точке x.

Условие дифференцируемости функции в точке эквивалентного условию конечности производной $f'(x)$ в этой точке.

Связь дифференцируемости и непрерывности

Теорема 1 (необходимое условие дифференцируемости).

Если функция $y = f(x)$ дифференцируема в точке x, то она непрерывна в ней.

Обратное утверждение, вообще говоря, неверно, т. е. из непрерывности не следует дифференцируемость. Для подтверждения сказанного рассмотрим две функции, графики которых представлены на рис. 2.6.

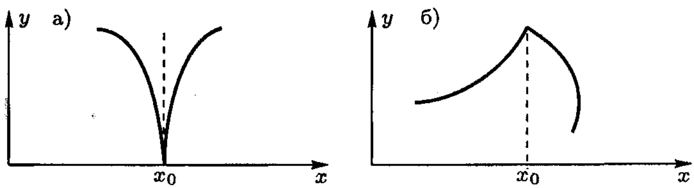


Рис. 2.6. Функции, не имеющие производной в точке x_0: а) — бесконечная производная; б) — производной нет.

Обе эти функции непрерывны в точке x_0, но не будут дифференцируемы в этой точке. Касательная к графику на рис. 2.6а в точке с абсциссой x_0 параллельна OY, т. е. функция обладает в точке x_0 бесконечной производной. Вторая функция (рис. 2.6б) не имеет производной в точке x_0.

2.1.3. Правила дифференцирования

1. Производная от постоянной величины равна нулю, т. е. если $y = C$, то $y' = 0$:

$$C' = 0. \tag{2.7}$$

Доказательство. По определению производной $y' = \lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x}$. Очевидно, что $\Delta y = 0$, следовательно $\Delta y / \Delta x = 0$; $y' = 0$.

2. Производная алгебраической суммы конечного числа функций равна сумме производных слагаемых:

$$(u + v + w + \ldots)' = u' + v' + w' + \ldots \tag{2.8}$$

Доказательство. Очевидно, $\Delta(u + v + w + \ldots) = \Delta u + \Delta v + \Delta w + \ldots$ Остается поделить обе части этого равенства на Δx, перейти к пределу при $\Delta x \to 0$ и воспользоваться тем, что предел алгебраической суммы функций равен алгебраической сумме пределов слагаемых (см. п. 1.2).

3. Производная произведения двух функций определяется формулой

$$(u \cdot v)' = u' \cdot v + u \cdot v'. \tag{2.9}$$

Доказательство. По (2.4) и теоремам из п. 1.2

$$\begin{aligned}(u \cdot v)' &= \lim_{\Delta x \to 0} \frac{\Delta (u \cdot v)}{\Delta x} = \lim_{\Delta x \to 0} \frac{(u + \Delta u)(v + \Delta v) - u \cdot v}{\Delta x} = \\ &= \lim_{\Delta x \to 0} \left(v \frac{\Delta u}{\Delta v} + u \cdot \frac{\Delta v}{\Delta x} + \frac{\Delta u \cdot \Delta v}{\Delta x}\right) = \\ &= v \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta v} + u \cdot \lim_{\Delta x \to 0} \frac{\Delta v}{\Delta x} + \lim_{\Delta x \to 0} \frac{\Delta u \cdot \Delta v}{\Delta x} = \\ &= u' \cdot v + u \cdot v' + \lim_{\Delta x \to 0} \Delta u \cdot v',\end{aligned}$$

но функция u дифференцируема, а значит непрерывна. Поэтому (см. п. 1.2) $\lim\limits_{\Delta x \to 0} \Delta u = 0$.

Совершенно аналогично может быть доказано правило дифференцирования произведения трех функций:

$$(u \cdot v \cdot w)' = u' \cdot v \cdot w + u \cdot v' \cdot w + u \cdot v \cdot w'.$$

4. Производная частного от деления двух функций определяется формулой

$$\left(\frac{u}{v}\right)' = \frac{u' \cdot v - u \cdot v'}{v^2}. \tag{2.10}$$

Доказательство.

$$\begin{aligned}(u/v)' &= \lim_{\Delta x \to 0} \frac{\Delta (u/v)}{\Delta x} = \lim_{\Delta x \to 0} \frac{(u + \Delta u)/(v + \Delta v) - u/v}{\Delta x} = \\ &= \lim_{\Delta x \to 0} \frac{(u + \Delta u)v - (v + \Delta v)u}{v(v + \Delta v)\Delta x} = \lim_{\Delta x \to 0} \frac{(v \cdot \Delta u - u \cdot \Delta v)}{v \cdot (v + \Delta v)\Delta x} = \\ &= \left(v \lim_{\Delta x \to 0} \Delta u/\Delta x - u \cdot \lim_{\Delta x \to 0} \Delta v/\Delta x\right) \Big/ \lim_{\Delta x \to 0} v(v + \Delta v) = \\ &= (u' \cdot v - v' \cdot u)/v^2.\end{aligned}$$

Что и требовалось доказать.

Дифференцирование сложной функции

Рассмотрим сложную функцию с одним промежуточным аргументом: $y = f(u)$, $u = \varphi(x)$, предполагая при этом, что функция y дифференцируема по аргументу u, а функция u дифференцируема по аргументу x. Требуется вывести правило дифференцирования этой сложной функции. На основании определения производной имеем

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

так как предел равен произведению пределов.

Если учесть, что $\lim\limits_{\Delta x \to 0} \dfrac{\Delta u}{\Delta x} = \dfrac{du}{dx}$ и $\lim\limits_{\Delta x \to 0} \Delta u = 0$, поскольку всякая дифференцируемая функция непрерывна, то получим, что

$$\frac{dy}{dx} = \lim_{\Delta u \to 0} \frac{\Delta y}{\Delta u} \cdot \frac{du}{dx} = \frac{dy}{du} \cdot \frac{du}{dx},$$

или, в других обозначениях,

$$y'_x = y'_u \cdot u'_x. \tag{2.11}$$

Совершенно аналогично может быть выведено правило дифференцирования сложной функции с двумя промежуточными аргументами:

$$y = y(u(\nu(x))): \quad y'_x = y'_u \cdot u'_\nu \cdot \nu'_x.$$

Пример. Из школьного курса математики известны производные функций:

$$(x^n)' = nx^{n-1}, \quad (\ln x)' = \frac{1}{x}, \quad (\sin x)' = \cos x, \quad (e^x)' = e^x$$

(в дальнейшем мы выведем эти формулы).

Найти y', если: $y = \sin^3 5x$.

Производную этой сложной функции будем находить по формуле (2.11). Предположим, что данное y нужно вычислить при конкретном значении x:

Вычисление Y		Нахождение производной
1. Ввести $\boxed{x}$	Направление вычислений ↓	↑ Вычисление производной $\boxed{1}$
2. Найти $\boxed{5x}$		$\boxed{5}$
3. Найти $\boxed{\sin}\ 5x$		$\boxed{\cos}\ 5x$
4. Найти $(\sin 5x)^{\boxed{3}}$		$\boxed{3}\ \sin^{\boxed{2}} 5x$

Таким образом, видно, что нахождение производной идет в порядке, обратном порядку вычисления функции.

Тогда по формуле (2.11) нам остается перемножить результаты, полученные на всех четырех ступенях, и

$$(\sin^3 5x)' = 3\sin^2 5x \cdot \cos 5x \cdot 5.$$

П р и м е р. $y = \ln\sin e^{3x}$. Найти y'.

Действуем в соответствии со схемой, приведенной выше:

Вычисление Y		Нахождение Y'
1. $\boxed{x}$	Направление вычислений ↓	↑ Вычисление производной 1
2. $\boxed{3x}$		$\boxed{3}$
3. $\boxed{e}^{3x}$		$\boxed{e}^{3x}$
4. $\boxed{\sin}\ e^{3x}$		$\boxed{\cos}\ e^{3x}$
5. $\boxed{\ln}\ \sin e^{3x}$		$\dfrac{1}{\boxed{\sin e^{3x}}}$

Значит, $y' = \dfrac{1}{\sin e^{3x}} \cdot \cos e^{3x} \cdot e^{3x} \cdot 3.$

Производная обратной функции

Предположим, что функция $y = f(x)$ определена, монотонна и дифференцируема в некоторой области, причем производная dy/dx нигде не равна нулю. Наша функция $y = f(x)$ имеет обратную функцию $x = \varphi(y)$, и нам надо получить правило дифференцирования этой функции. Для вывода можно воспользоваться правилом дифференцирования сложной функции с одним промежуточным аргументом. Согласно представлениям о сложной функции функцию y можно рассматривать как сложную функцию от самой себя с промежуточным аргументом x: $y = f(x)$, $x = \varphi(y)$.

На основании правила дифференцирования сложной функции получаем: $y'_y = y'_x \cdot x'_y$.

Поскольку $y'_y = 1$, получаем правило дифференцирования обратной функции:

$$x'_y = \frac{1}{y'_x}. \tag{2.12}$$

Производная от функции, заданной параметрически

Рассмотрим параметрически заданную функцию ($t \in T$)

$$\begin{cases} x = \varphi(t), \\ y = \psi(t). \end{cases}$$

Теорема 2 Если в некотором промежутке изменения t функция $x = \varphi(t)$ возрастает (убывает), функции $x = \varphi(t)$ и $y = \psi(t)$ дифференцируемы, причем $\varphi'(t) \neq 0$, то на соответствующем промежутке изменения x переменная y будет однозначной и дифференцируемой функцией от x; при этом

$$y'_x = \frac{\psi'(t)}{\varphi'(t)} = \frac{y'_t}{x'_t}. \tag{2.13}$$

Доказательство. Поскольку $x = \varphi(t)$ монотонна в рассматриваемом промежутке изменения t , то она допускает однозначную и дифференцируемую (в силу условия $\varphi'(t) \neq 0$) обратную функции $t = g(x)$, причем

$$g'(x) = \frac{1}{\varphi'(t)} \quad (\text{см. } (2.12)).$$

Учитывая это, на основании теоремы о дифференцировании сложной функции находим производную от $y = \psi[g(t)] = f(x)$:

$$y'_x = \psi'_t \cdot g'_x = \frac{\psi'_t}{\varphi'_t}, \tag{2.14}$$

что и требовалось доказать.

Пример 3.

$$\begin{cases} x = \ln t, \\ y = e^{2t}. \end{cases}$$

Найдем y'_x.

В соответствии с формулой (2.13) получим

$$y'_x = \frac{(e^{2t})'_t}{(\ln t)'} = \frac{e^{2t} \cdot 2}{1/t} = 2t \cdot e^{2t}.$$

2.1.4. Производные от основных элементарных функций

1. Рассмотрим сначала производную от функции $y = \ln x$:

$$(\ln x)' = \lim_{\Delta x \to 0} \frac{\ln(x + \Delta x) - \ln x}{\Delta x} =$$
$$= \lim_{\Delta x \to 0} \frac{\ln((x + \Delta x)/x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\ln(1 + \Delta x/x)}{\dfrac{\Delta x}{x} \cdot x} = \frac{1}{x},$$

поскольку (см. п. 1.1.2) $\lim\limits_{z \to 0} \dfrac{\ln(1+z)}{z} = 1$.

Итак, $(\ln x)' = 1/x$.

2. Рассмотрим производную от степенной функции $y = x^n$.

Сначала логарифмируем: $\ln y = n \ln x$. Дифференцируя и пользуясь правилом дифференцирования сложной функции, получаем

$$(\ln y)'_x = \frac{1}{y} \cdot y'_x = n(\ln x)' = n \cdot \frac{1}{x}.$$

Окончательно: $y'_x = y \cdot n \cdot 1/x = x^n \cdot n \cdot \dfrac{1}{x} = nx^{n-1}$.

Итак, $(x^n)' = nx^{n-1}$.

3. Рассмотрим производную от показательной функции $y = a^x$.

Совершенно аналогично предыдущему выводится формула $(a^x)' = a^x \ln a$.

4. Рассмотрим производные от тригонометрических функций:

$$(\sin x)' = \lim_{\Delta x \to 0} \frac{\sin(x + \Delta x) - \sin x}{\Delta x} =$$
$$= \lim_{\Delta x \to 0} \frac{2 \sin \Delta x/2 \cdot \cos(x + \Delta x/2)}{\Delta x} =$$
$$= \lim_{\Delta x \to 0} \frac{\sin \Delta x/2}{\Delta x/2} \cdot \lim_{\Delta x \to 0} \cos(x + \Delta x/2) = \cos x,$$

с учетом I замечательного предела (см. п. 1.2) и непрерывности функции $\cos x$;

$$(\cos x)' = \left[\sin(\frac{\pi}{2} - x)\right]' = \cos(\frac{\pi}{2} - x)(-1) = -\sin x,$$

$$(\operatorname{tg} x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{(\sin x)' \cdot \cos x - \sin x \cdot (\cos x)'}{\cos^2 x} = \frac{1}{\cos^2 x} =$$
$$= 1 + \operatorname{tg}^2 x.$$

Итак,

$$(\sin x)' = \cos x, \quad (\cos x)' = -\sin x, \quad (\operatorname{tg} x)' = \frac{1}{\cos^2 x}.$$

Совершенно аналогично выводится формула

$$(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x} = -(1 + \operatorname{ctg}^2 x).$$

Перейдем к изложению вопроса о производных от обратных тригонометрических функций: $y = \arcsin x$, $y = \arccos x$, $y = \operatorname{arctg} x$, $y = \operatorname{arcctg} x$.

Функции $y = \arcsin x$ соответствует обратная функция $x = \sin y$ при условии $-\pi/2 \leqslant y \leqslant \pi/2$.

Согласно правилу (2.12) дифференцирования обратной функции имеем

$$y'_x = \frac{1}{x'_y} = \frac{1}{\cos y} = \frac{1}{\sqrt{1 - x^2}},$$

поскольку $\cos y = \sqrt{1 - \sin^2 y}$. (Здесь перед корнем выбран знак «+» так как $\cos y > 0$ для $y \in [-\pi/2, \pi/2]$.)

Итак,

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}.$$

Совершенно аналогично получается формула

$$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}.$$

Функции $y = \operatorname{arctg} x$ соответствует обратная функция $x = \operatorname{tg} y$ при условии $-\pi/2 < y < \pi/2$.

Согласно правилу дифференцирования обратной функции имеем

$$(\operatorname{arctg} x)'_x = \frac{1}{(\operatorname{tg} y)'_y} = \cos^2 y = \frac{1}{1+\operatorname{tg}^2 y} = \frac{1}{1+x^2} \Rightarrow$$

$$\Rightarrow (\operatorname{arctg} x)' = \frac{1}{1+x^2}.$$

Совершенно аналогично выводится формула

$$(\operatorname{arcctg} x)' = -\frac{1}{1+x^2}.$$

Составим сводную таблицу производных от основных элементарных функций.

Таблица 2.1 Производные основных элементарных функций

№	y	y'	№	y	y'
1	x^n	nx^{n-1}	7	$\operatorname{tg} x$	$1/\cos^2 x$
2	$\ln x$	$1/x$	8	$\operatorname{ctg} x$	$-1/\sin^2 x$
3	a^x	$a^x \ln a$	9	$\arcsin x$	$1/\sqrt{1-x^2}$
4	e^x	e^x	10	$\arccos x$	$-1/\sqrt{1-x^2}$
5	$\sin x$	$\cos x$	11	$\operatorname{arctg} x$	$1/(1+x^2)$
6	$\cos x$	$-\sin x$	12	$\operatorname{arcctg} x$	$-1/(1+x^2)$

$\log_a x = \ln x / \ln a$ (формула перехода к другому основанию); $\ln a$ — постоянная. Следовательно, $(\log_a x)' = \left(\frac{\ln x}{\ln a}\right)' = \frac{1}{\ln a} \times$ $\times (\ln x)' = \frac{1}{x \ln a}$.

Полезно также помнить формулу для $y = \sqrt{x}$, $\sqrt{x} = x^{1/2}$. Тогда $(\sqrt{x})' = \frac{1}{2}x^{-1/2} = \frac{1}{2\sqrt{x}}$.

2.1.5. Логарифмическое дифференцирование

Дифференцирование многих функций значительно упрощается, если их предварительно прологарифмировать. Если функция $y = f(x)$ есть произведение нескольких сомножителей, то можно сначала логарифмировать обе части уравнения $y = f(x)$ (например, по основанию e): $\ln y = \ln f(x) = \varphi(x)$, затем дифференцировать обе части полученного равенства: $(\ln y)' = \varphi'(x)$, $y'/y = \varphi'(x)$.

Отсюда, умножив обе части последнего уравнения на $y = f(x)$, получим искомую производную y': $y' = y \cdot \varphi'(x) = f(x) \cdot \varphi'(x)$

Логарифмическое дифференцирование полезно применять, когда заданная функция содержит логарифмирующиеся операции (умножение, деление, возведение в степень, извлечение корня) и, в частности, для дифференцирования показательно-степенной функции $y = U^V$, где U и V — функции от x.

Например, надо найти производные от функций:

$$1.\ \ y = x^x \quad (x > 0).$$

$$2.\ \ y = \frac{x^2 \cdot \sqrt[3]{x+1} \cdot a^x}{\sqrt{x^2+1}} \quad (a > 0, a \neq 1).$$

Решение.

1) $y = x^x$. Сначала логарифмируем: $\ln y = x \ln x$. Теперь дифференцируем: $(\ln y)' = (x \ln x)'$. На основании формул и правил дифференцирования получаем: $y'/y = \ln x + x \cdot 1/x = \ln x + 1$. Отсюда: $y' = y(\ln x + 1) = x^x(\ln x + 1)$.

2) $y = \dfrac{x^2\sqrt[3]{x+1} \cdot a^x}{\sqrt{x^2+1}}$; $\ln y = 2\ln x + \dfrac{1}{3}\ln(x+1) + x\ln a - \dfrac{1}{2}\ln(x^2+1)$; $\dfrac{y'}{y} = \dfrac{2}{x} + \dfrac{1}{3(x+1)} + \ln a - \dfrac{x}{x^2+1}$.

$$y' = \frac{x^2\sqrt[3]{x+1} \cdot a^x}{\sqrt{x^2+1}} \cdot \left(\frac{2}{x} + \frac{1}{3(x+1)} + \ln a - \frac{x}{x^2+1}\right).$$

2.1.6. Производные неявных функций

Допустим, что надо дифференцировать функцию, заданную уравнением $f(x, y) = 0$, не разрешенным относительно y. Для этого сначала дифференцируют по x обе части равенства $f(x, y) = 0$, помня, что y есть функция от x, а затем разрешают полученное равенство относительно искомой производной. Как правило, она будет зависеть от x и y: $y' = \varphi(x, y)$.

Поясним на примере:

1. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1, \quad y' = ?$

Дифференцируем обе части равенства:

$$\frac{2x}{a^2} + \frac{2y \cdot y'}{b^2} = 0.$$

Отсюда $y' = -\dfrac{b^2}{a^2}\dfrac{x}{y}$.

Пример. Пусть, например, задано уравнение окружности с центром в точке (a, b) и радиусом R: $(x-a)^2 + (y-b)^2 = R^2$.

Дифференцируя обе стороны равенства по переменной x, найдем: $2(x-a) + 2(y-b)y' = 0$. Отсюда $y' = -\dfrac{x-a}{y-b}$.

Самостоятельная работа

1. $y = \dfrac{1+\sqrt{x}}{1+2\sqrt{x}}$. **2.** $y = \dfrac{\operatorname{tg} x}{x}$. **3.** $y = \dfrac{\sin a}{x} + \dfrac{x}{\sin x}$.

4. $y = x \arcsin x$. **5.** $y = \dfrac{\arccos x}{x}$.

6. $y = \dfrac{x}{1+x^2} - \operatorname{arctg} x$. **7.** $y = x^2 \log_3 x$. **8.** $y = \dfrac{x}{\ln x}$.

9. $y = 10^x \cdot e^x$. **10.** $y = \dfrac{1}{3^x}$. **11.** $y = \dfrac{e^x}{\sin x}$.

12. $y = \dfrac{\cos x}{x^2}$. **13.** $y = x^3 - 4 - 3^x$. **14.** $y = \dfrac{\ln 4}{\sin x}$.

15. $y = x \cdot \sin 2x$. **16.** $y = (x^3 + 5)\lg x$. **17.** $y = \dfrac{1 - \ln x}{1 + \ln x}$.

18. $y = x^7 \cdot \ln x$. **19.** $y = \operatorname{arcctg} x \cdot 3x^2$.

20. $y = \dfrac{x^5 + 3^x}{e^x}$. **21.** $y = \dfrac{x^2 - 2}{4^x}$. **22.** $y = \dfrac{5x}{\sin x + \ln x}$.

23. $y = \dfrac{x \sin x}{1 + \operatorname{tg} x}$. **24.** $y = \dfrac{\sin t}{5 + \cos t}$. **25.** $y = \dfrac{e^x}{1 - \cos x}$.

26. $y = 2 \cdot \sqrt[3]{x} \cdot \arccos x$. **27.** $y = \dfrac{\operatorname{ctg} x}{x - 5}$.

28. $y = \dfrac{\sin x + 1}{1 - \cos x}$. **29.** $y = \dfrac{\log_2 x}{\ln 2}$. **30.** $y = \sin x \cdot e^3$.

Дифференцирование сложной функции

1. $y = \arcsin^2 x$. **2.** $y = \sqrt{1 - \sin^2 x}$.

3. $y = \operatorname{ctg}^3 \sqrt{1 + x^2}$. **4.** $y = \sqrt{1 + \operatorname{tg}(x + \frac{1}{x})}$.

5. $y = \sin^2(\cos 5x)$. **6.** $y = \cos^2 \dfrac{1 - \sqrt{x}}{1 + \sqrt{x}} - \operatorname{arctg} x$.

7. $y = \cos^3(4x + 5)$. **8.** $y = \operatorname{tg}^5 \dfrac{x + 1}{2}$.

9. $y = e^{\sqrt{\ln x}} \cdot \sin x$. **10.** $y = \ln \sin \sqrt[3]{\operatorname{arctg} e^{3x}}$.

11. $y = \ln \operatorname{tg}^3 x$. **12.** $y = \ln^4 \sin x$.

13. $y = \ln \operatorname{arctg} \sqrt{1 + x^5}$. **14.** $y = \sqrt[3]{\ln \sin \dfrac{x + 3}{4}}$.

15. $y = \arcsin \sqrt{\dfrac{1 - x^2}{1 + x^2}}$. **16.** $y = \operatorname{arctg}^2 \dfrac{1}{x}$.

17. $y = \arcsin(\sin^2 5x)$. **18.** $y = \operatorname{arctg}(x - \sqrt{1 + x^2})$.

19. $y = \log_3^2(x^2 - 4x + 5)$. **20.** $y = 2^{\frac{x^3}{\ln x}}$.

21. $y = \sin(2^{x+3} \cdot e^x)$. **22.** $y = 2^{3^x}$. **23.** $y = e^{\sqrt[5]{\ln^2 3x}}$.

24. $y = 10^{1-\sin^4 3x}$. **25.** $y = \ln\sin\sqrt[3]{\operatorname{arctg} e^{3x}}$.

26. $y = 2^x \cdot x^2 \cdot 5$. **27.** $y = 5^{\sin^3(3x^2-8)}$.

28. $y = \dfrac{1}{2}\sqrt[4]{\arcsin\sqrt{x^2+2x}}$. **29.** $y = \operatorname{arctg}(x + 3\sqrt{1+x^5})$.

30. $y = \operatorname{ctg}\sqrt{\sin(\sin x)}$. **31.** $y = \log_2(\log_3(\log_4(x-2)))$.

Логарифмическое дифференцирование

1. $y = x^{x^2}$. **2.** $y = (\sin 5x)^{\cos x}$. **3.** $y = (x+2)^{3/x}$.

4. $y = (\ln 3x)^{x^2}$. **5.** $y = x^3 \cdot e^{x^4} \cdot \sin 2x$. **6.** $y = x^{\ln x}$.

7. $y = \dfrac{(x-1)^2\sqrt[3]{x+2}}{(x+6)^3}$. **8.** $y = \sqrt{x\sin x\sqrt{4-e^x}}$.

9. $y = \left(\dfrac{x}{1+x}\right)^x$. **10.** $y = \sqrt[3]{\dfrac{x(x^2+1)}{(x^2-1)^2}}$.

Функции, заданные неявно

1. $\dfrac{x^2}{9} + \dfrac{y^2}{4} = 1$. **2.** $x^4 + y^4 - 5axy = 0$.

3. $y^2\cos x = a^2\sin 3x$. **4.** $3^x + 3^y = 3^{x+y}$.

5. $y = \cos(x+y)$. **6.** $x\sin y - \cos y + \cos 2y = 0$.

7. $y\ln y = x$. **8.** $y = 1 + xe^y$. **9.** $\operatorname{tg}\dfrac{y}{2} = \sqrt{a}\operatorname{tg}\dfrac{x}{2}$.

10. $y = x + \operatorname{arctg} y$. **11.** $x^4 + y^3 + xy = 2$.

12. $\sin(xy) + \cos(xy) = \operatorname{tg}(x+y)$. **13.** $\sqrt[3]{x^2} + \sqrt[3]{y^2} = \sqrt[3]{a^2}$.

14. $x - y = \arcsin x - \arcsin y$. **15.** $y^2 - 2xy = 1$.

16. $x^y = y$. **17.** $\sin\dfrac{x}{y} = \cos\dfrac{y}{x}$. **18.** $y \cdot \operatorname{arctg} x = \dfrac{y^3-4}{5}$.

19. $\cos^2 x - \sin(xy) = y$. **20.** $e^{3xy} + \dfrac{y}{e^x} = 2xy^2$.

§ 2.2. Дифференциал функции

2.2.1. Определение

Согласно определению производной от функции $y = f(x)$ имеем

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

На основании определения предела (п. 1.2) это означает, что $\Delta y/\Delta x = y' + \alpha(\Delta x)$, где $\lim\limits_{\Delta x \to 0} \alpha(\Delta x) = 0$. Отсюда $\Delta y = y' \cdot \Delta x + \Delta x \cdot \alpha(\Delta x)$. Первое слагаемое в правой части этого равенства стремится к нулю, как Δx (если $y' \neq 0$), а второе слагаемое кроме Δx содержит в себе множитель $\alpha(\Delta x)$, который тоже стремится к нулю при $\Delta x \to 0$. Таким образом, первое слагаемое стремится к нулю медленнее второго, и поэтому его называют главной частью приращения функции Δy.

Определение. Главная часть приращения функции Δy, равная произведению $y' \cdot \Delta x$, называется дифференциалом первого порядка от функции $y = f(x)$, соответствующим выбранным значениям x и Δx.

Обозначается так:

$$dy = y' \cdot \Delta x. \tag{2.15}$$

Это так называемая первая форма записи дифференциала.

Выясним геометрический смысл дифференциала. На схематическом графике функции $y = f(x)$ (рис. 2.7) отмечены две точки: точка A с абсциссой x и точка B с абсциссой $x + \Delta x$. Производная от функции $y = f(x)$ в точке A равна угловому коэффициенту касательной к графику $y = f(x)$ в этой точке. Поскольку $CD = AC \cdot \operatorname{tg} \varphi = y' \cdot \Delta x$, заключаем, что дифференциал dy равен приращению ординаты касательной к графику $y = f(x)$, соответствующему значениям x и $x + \Delta x$.

Каков же механический смысл дифференциала?

Если $s = f(t)$ есть путь, пройденный материальной точкой за время t, то, как известно, производная ds/dt есть скорость

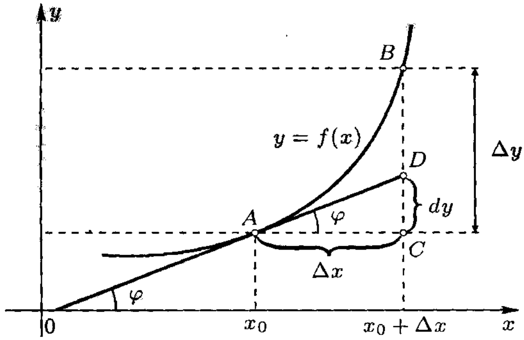


Рис. 2.7. Геометрический смысл дифференциала.

движения в момент времени t. Тогда *дифференциал* пути $ds = f'(t) \cdot \Delta t$ *приближенно равен пути, пройденному материальной точкой от момента времени* t *до момента времени* $t + \Delta t$, если пренебречь изменением скорости движения на этом промежутке времени.

Кроме первой формы записи дифференциала существует и вторая. Дифференциалом аргумента называется дифференциал функции $y = x$, т.е. $dy = dx$. Однако на основании первой формы дифференциала имеем $dy = \Delta x$. Следовательно, $\Delta x = dx$, и мы получаем, таким образом, вторую форму записи дифференциала:

$$dy = y' \cdot dx. \tag{2.16}$$

Вторая форма дифференциала обладает свойством инвариантности относительно аргумента, т.е. не зависит от того, является ли аргумент x окончательным или промежуточным. Поясним это на примере.

Пример. Пусть требуется вычислить дифференциал функции $y = (1 + \operatorname{tg} x)^8$. Это можно осуществить двумя способами:

1. Найдем $y'(x)$ по правилам дифференцирования сложной функции:

$$y'(x) = \frac{8(1 + \operatorname{tg} x)^7}{\cos^2 x}.$$

Тогда

$$dy = y'(x)dx = \frac{8(1 + \operatorname{tg} x)^7}{\cos^2 x} dx.$$

2. Введем новую функцию $U = 1 + \operatorname{tg} x$. Тогда $y = U^8$, $dy = 8U^7 dU$ (как дифференциал функции $y = y(u)$).

Вычислим dU:

$$dU = U'_x\, dx = (1 + \operatorname{tg} x)'\, dx = \frac{1}{\cos^2 x}\, dx.$$

Следовательно,

$$dy = \underbrace{8(1 + \operatorname{tg} x)^7}_{y'_u} \cdot \underbrace{\frac{1}{\cos^2 x}}_{U'_x}\, dx,$$

т. е. результат совпадает с результатом, вычисленным по первому способу.

Первая форма дифференциала таким свойством не обладает.

Поскольку формально дифференциал отличается от производной лишь множителем Δx или dx, его свойства являются отражением соответствующих свойств производной.

Например, $d(c) = 0$, $d(x)^n = nx^{n-1}\, dx$, $d(\sin x) = \cos x\, dx$ и т. д.

Пользуясь таблицей производных, мы можем написать таблицу дифференциалов от основных элементарных функций.

Аналогично могут быть получены формулы:

$$d(u \pm v) = du \pm dv, \quad d(uv) = v\, du + u\, dv, \quad d\left(\frac{u}{v}\right) = \frac{v\, du - u\, dv}{v^2}.$$

2.2.2. Приближенные вычисления с помощью дифференциала

Теорема 1 Если функция $y = f(x)$ дифференцируема в точке x, причем $f'(x) \neq 0$, то при $\Delta x \to 0$ приращение Δy и дифференциал dy функции являются эквивалентными бесконечно малыми.

Доказательство. По условию теоремы функция $y = f(x)$ дифференцируема в точке x, т. е.

$$\Delta y = f'(x)\Delta x + \alpha\Delta x, \qquad \text{где} \quad \alpha \to 0 \quad \text{при} \quad \Delta x \to 0.$$

Тогда

$$\lim_{\Delta x \to 0} \frac{\Delta y}{dy} = \lim_{\Delta x \to 0} \frac{f'(x)\Delta x + \alpha\Delta x}{f'(x)\Delta x} = \lim_{\Delta x \to 0} \left(1 + \frac{\alpha}{f'(x)}\right) = 1,$$

что и требовалось доказать.

На этой теореме и основано применение дифференциала к приближенным вычислениям. Известно, что любую из двух эквивалентных бесконечно малых можно приближенно заменить другой. Следовательно,

$$\Delta y \approx dy. \tag{2.17}$$

Абсолютная и относительная погрешности этого равенства могут быть сделаны сколь угодно малыми при достаточно малом $|\Delta x|$. Структура дифференциала обычно значительно проще структуры приращения функции, в силу чего формула (2.17) широко применяется в приближенных вычислениях.

Пример 1. Дан куб с ребром $x = 2$ м. Вычислить, на сколько возрастет его объем, если ребро увеличится на 1 см. Объем куба $V = x^3$; Δx по условию задачи равно 0,01 м.

Точное решение:

$$\Delta V = (x + \Delta x)^3 - x^3 = 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3 =$$
$$= 3 \cdot 2^2 \cdot 0{,}01 + 3 \cdot 2 \cdot (0{,}01)^2 + (0{,}01)^3 = 0{,}120601 \text{ м}^3.$$

Таким образом, объем куба увеличится на $0{,}120601$ м^3.

Приближенное решение:

$$\Delta V \approx dV = V'\,dx = 3x^2\Delta x = 3 \cdot 2^2 \cdot 0{,}01 = 0{,}12 \text{ м}^3.$$

Абсолютная погрешность этого результата

$$|dy - \Delta y| = 0{,}000601 \text{ м}^3,$$

а относительная погрешность

$$\frac{dV - \Delta V}{\Delta V} = \frac{0{,}000601}{0{,}120601} \approx 0{,}005 \quad (\text{или } 0{,}5\%).$$

Пример 2. Вычислить приближенное значение $\ln 0{,}9$.

Решение. $y = \ln x$; Возьмем $x = 1$, тогда $\Delta x = -0{,}1$. Из формулы (2.17): $f(x + \Delta x) - f(x) \approx f'(x)\Delta x$ или

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x, \tag{2.18}$$

$$y = \ln x\big|_{x=1} = 0, \quad y'(x) = \frac{1}{x}\bigg|_{x=1} = 1, \quad \Delta x = -0{,}1.$$

Эти результаты подставим в формулу (2.18). Тогда

$$\ln 0{,}9 \approx 1 \cdot (-0{,}1) = -0{,}1.$$

Пример 3. Приближенно вычислить $\sqrt{16{,}02}$.

Работая с формулой (2.17), будем иметь

$$y = \sqrt{x}\big|_{x=16} = 4, \quad \Delta x = 0,02, \quad y'(x) = \frac{1}{2\sqrt{x}}\bigg|_{x=16} = \frac{1}{8}.$$

Тогда

$$\sqrt{16{,}02} \approx 4 + \frac{1}{8} \cdot 0{,}02 \approx 4{,}0025.$$

2.2.3. Производные и дифференциалы высших порядков

Пусть функция $y = f(x)$ дифференцируема на некотором промежутке. Тогда производная $f'(x)$ этой функции будет новой функцией от x, заданной на этом промежутке, и может в свою очередь иметь производную. Эту производную называют производной второго порядка и обозначают одним из символов:

$$y'', \quad f''(x), \quad y^{(2)}, \quad f^{(2)}(x).$$

Производную от второй производной называют третьей производной и т. д.

Вообще,

$$y^{(n)} = \left(y^{(n-1)}\right)'. \tag{2.19}$$

Пример.

$$\begin{aligned}
&y = 4x^3 + 2x - 8,\\
&y' = 12x^2 + 2,\\
&y'' = 24x,\\
&y''' = 24,\\
&y^{IV} = y^{V} = \ldots = 0.
\end{aligned}$$

Отметим, что для многочлена степени n все производные начиная с $(n+1)$-го порядка равны нулю.

Теперь рассмотрим дифференциалы высшего порядка.

Дифференциал функции $y = f(x)$, где x — независимая переменная, определяется формулой $dy = f'(x)\,dx$, где $dx = \Delta x$ — произвольное приращение аргумента x. Зафиксируем dx; тогда dy будет функцией от x. Дифференциал от этой функции называют *дифференциалом второго порядка* и обозначают d^2y или $d^2f(x)$.

Поскольку dx зафиксирован (постоянен), то

$$d^2y = d(dy) = d[f'(x)\,dx] = [f'(x)\,dx]'dx = f''(x)\,dx^2.$$

Аналогично определяются дифференциалы для любого $n > 2$:

$$d^ny = f^{(n)}(x)\,dx^n. \tag{2.20}$$

Итак, дифференциал n-го порядка функции равен произведению n-й производной этой функции на n-ю степень дифференциала независимой переменной.

Если имеется сложная функция $y = f(u)$, $u = \varphi(x)$, то $dy = f'(u)\,du$, но теперь уже $f'(u)$ и $du = \varphi'(x)\,dx$ являются функциями от x. Поэтому при отыскании повторных дифференциалов здесь нельзя выносить du за знак производной, а надо дифференцировать произведение. Формулы для дифференциалов высших порядков сложной функции отличаются от последней формулы, и дифференциалы второго и высших порядков свойством инвариантности по отношению к аргументу уже не обладают.

Самостоятельная работа

Приложения дифференциала к приближенным вычислениям

1. Найти приращение функции $y = x^2$, соответствующее приращению Δx независимой переменной. Вычислить Δy, если $x = 1$ и $\Delta x = 0{,}1$.

2. Найти приращение Δv объема v шара при изменении радиуса $R = 2$ на $\Delta R = 0{,}1$.

3. $y = x^3+2x$. Найти значения приращения и его линейной главной части, соответствующие изменению x от $x = 2$ до $x = 2{,}1$.

4. Найти приращение и дифференциал функции $y = x^2 - x$ при $x = 10$ и $\Delta x = 0{,}1$. Вычислить абсолютную и относительную ошибки, которые получаются при замене приращения дифференциалом.

5. $\sqrt{4,041} \approx$?

6. Сторона квадрата равна 8 см. На сколько увеличивается его площадь, если каждую его сторону увеличить на 0,1 см?

7. Вычислить значение дифференциала функции:

а) $y = \dfrac{1}{(\operatorname{tg} x + 1)^2}$ при изменении x от $\dfrac{\pi}{6}$ до $\dfrac{61\pi}{360}$;

б) $y = \cos^2 \varphi$ при изменении φ от 60° до $60^\circ 30'$.

8. $y = 3^{1/x} + 1/2^{2x} + 6^{\sqrt{x}}$. Вычислить dy при $x = 1$ и $dx = 0{,}2$.

9. $\sin 60^\circ 18' \approx ?$ **10.** $\operatorname{arctg} 1{,}02 \approx ?$

11. $\sqrt{\dfrac{(2{,}037)^2 - 3}{(2{,}037)^2 + 5}} \approx ?$ **12.** $\arcsin 0{,}4983 \approx ?$

13. $f(x) = e^{0{,}1x(1-x)}$, $f(1{,}05) \approx ?$

Повторное дифференцирование

1. $y = x^2 - 3x + 2$, $y''' = ?$ **2.** $y = \operatorname{arctg} x$, $y''(1) = ?$

3. $y = (x^2 + 1)^3$, $y'' = ?$ **4.** $y = \cos^2 5x$, $y''' = ?$

5. $y = x^3 \cdot \ln x$, $y^{IV} = ?$ **6.** $y = \dfrac{1}{1-x}$, $y^V = ?$

7. $y = e^{2x-1}$, $y''(0) = ?$ **8.** $y = (1 + x^2) \operatorname{arctg} x$, $y''(x) = ?$

9. $y = \ln(x + \sqrt{1 + x^2})$, $y''(x) = ?$

10. $\sqrt{1 - x^2} \arcsin x$, $y''(x) = ?$

§ 2.3. Приложения производной

2.3.1. Касательная и нормаль к плоской кривой

Если плоская кривая, являющаяся графиком функции $y = f(x)$ в декартовой системе координат, имеет в точке $M_0\,(x_0, f(x_0))$ касательную и нормаль, ни одна из которых не параллельна оси OY, то их уравнения, соответственно, имеют вид

$$y - f(x_0) = f'(x_0)(x - x_0), \quad y - f(x_0) = -\frac{1}{f'(x_0)}(x - x_0).$$

Направление кривой в каждой точке определяется направлением касательной к ней в этой точке.

Рассмотрим примеры.

Составить уравнения касательной и нормали:

1) к параболе $y = x^2 - 3x$, где $x = 1$;

2) к окружности $x^2+y^2-2x+4y-3 = 0$ в точках пересечения ее с осью OX.

Решение.

1) $y = x^2 - 3x$. Если $x = 1$, то $y = 1 - 3 = -2$, $y' = 2x - 3x$. В точке $x = 1$ $y' = 2 - 3 = -1$.

Уравнение касательной: $y+2 = -1\cdot(x-1)$, или $x+y+1 = 0$; уравнение нормали $y + 2 = x - 1$, или $y - x + 3 = 0$.

2) $x^2 + y^2 - 2x + 4y - 3 = 0$. В точках пересечения с осью OX $y = 0$, значит $x^2 - 2x - 3 = 0$; $x_1 = -1$; $x_2 = 3$. Найдем угловые коэффициенты касательных:

$$2x + 2y \cdot y' - 2 + 4y' = 0, \quad x + y \cdot y' - 1 + 2y' = 0, \quad y' = \frac{1-x}{y+2}.$$

В точке $(-1, 0)$

$$y' = \frac{1+1}{2} = 1.$$

Уравнение касательной: $y = (x + 1)$, или $x - y + 1 = 0$; уравнение нормали $y = -(x + 1)$, или $x + y + 1 = 0$.

В точке $(3, 0)$

$$y' = \frac{1-3}{2} = -1;$$

Уравнение касательной: $y = -(x - 3)$, или $x + y - 3 = 0$; уравнение нормали $y = (x - 3)$, или $x - y - 3 = 0$.

2.3.2. Скорость изменения переменной величины. Скорость и ускорение прямолинейного движения

Если z есть функция времени t, то скорость ее изменения определяется производной dz/dt.

Если точка движется прямолинейно, то ее скорость V и ускорение W определяются первой и второй производными от пути S по времени t:

$$V = \frac{dS}{dt}, \quad W = \frac{dV}{dt} = \frac{d^2S}{dt^2}.$$

Поясним эти положения примерами.

1. Точка движется по кубической параболе $y = x^3$. Какая из ее координат изменяется быстрее?

Решение. Если $y = \varphi(t)$, $x = \psi(t)$ — параметрические уравнения кубической параболы $y = x^3$, соответствующие закону движения точки, то:

$$\frac{dy}{dt} = 3x^2\frac{dx}{dt}; \quad \text{отсюда} \quad \frac{dy}{dt} : \frac{dx}{dt} = 3x^2.$$

Если $|x| < \frac{1}{\sqrt{3}}$, то $\frac{dy}{dt} : \frac{dx}{dt} < 1$, т. е. $\frac{dy}{dt} < \frac{dx}{dt}$, и, значит, абсцисса x изменяется быстрее ординаты y.

Если $|x| > \frac{1}{\sqrt{3}}$, то $\frac{dy}{dt} : \frac{dx}{dt} > 1$, т. е. $\frac{dy}{dt} > \frac{dx}{dt}$, и, значит, ордината y изменяется быстрее абсциссы x.

2. Зависимость количества Q вещества, получаемого в химической реакции, от времени t определяется формулой $Q = a(1 + b \cdot e^{-mt})$. Определить скорость реакции.

Решение. Скорость реакции есть производная

$$\frac{dQ}{dt} = a \cdot b(-m)e^{-mt} = -abme^{-mt}.$$

3. Скорость прямолинейного движения тела пропорциональна корню квадратному из пройденного пути. Доказать, что движение происходит под действием постоянной силы.

Решение. По закону Ньютона сила F, вызывающая движение, пропорциональна ускорению:

$$F = k\frac{d^2S}{dt^2}.$$

Согласно условию $\frac{dS}{dt} = \lambda\sqrt{S}$. Отсюда

$$\frac{d^2S}{dt^2} = \frac{\lambda}{2\sqrt{S}}\frac{dS}{dt} = \frac{\lambda}{2\sqrt{S}} \cdot \lambda\sqrt{S} = \frac{\lambda^2}{2}.$$

С другой стороны, $\frac{d^2S}{dt^2} = \frac{F}{k}$, т. е. $\frac{F}{k} = \frac{\lambda^2}{2}$.

Это значит, что действующая сила F постоянная и равна $\frac{\lambda^2 k}{2}$.

Самостоятельная работа

1. Найти уравнения касательных и нормалей к гиперболе $y^2 - 2x^2 = 1$ в точках, где $x = 2$.

2. Точка движется по параболе $y = 5 - x^2$ так, что ее абсцисса x изменяется с течением времени t по закону $x = at^2$. С какой скоростью изменяется ордината точки?

Рассмотрим теперь еще два важных приложения производной:

1) правило Лопиталя (вычисление пределов с помощью производных);

2) исследование функций с помощью производных и построение графиков функций.

2.3.3. Правило Лопиталя

Пусть требуется вычислить предел $\lim\limits_{x \to a} f(x)$. Если $f(x)$ — элементарная функция и a принадлежит области ее определения, то для вычисления этого предела нужно перейти к пределу под знаком функции $f(x)$, что приводит к подстановке в $f(x)$ вместо x предельного значения a:

$$\lim_{x \to a} f(x) = f\left(\lim_{x \to a} x\right) = f(a).$$

Если же a не принадлежит области определения функции $f(x)$, то формальное применение правил предельного перехода приводит к одному их следующих символов:

$$\left[\frac{0}{0}\right], \quad \left[\frac{\infty}{\infty}\right], \quad [0 \cdot \infty], \quad [\infty - \infty], \quad [0^0], \quad [\infty^0], \quad [1^\infty].$$

В этом случае говорят, что в точке a имеет место неопределенность соответствующего типа, а вычисление предела в этом случае называют раскрытием неопределенности.

Теорема 1. Если функции $\varphi(x)$ и $g(x)$ дифференцируемы в некоторой окрестности точки a, при $x \to a$ одновременно являются бесконечно малыми и $g'(x) \neq 0$ в данной окрестности (исключая, может быть, саму точку a), то

$$\lim_{x\to a}\frac{\varphi(x)}{g(x)}=\lim_{x\to a}\frac{\varphi'(x)}{g'(x)} \tag{2.21}$$

при условии, что предел в правой части (2.21) существует (заметим, что a здесь может быть числом или одним из символов: $+\infty, -\infty, \infty$).

Из данной теоремы вытекает следующее практическое правило раскрытия неопределенностей первых двух типов, называемое правилом Лопиталя:

Для раскрытия неопределенностей вида $[0/0]$ или $[\infty/\infty]$ нужно от предела отношения двух функций перейти к пределу отношения производных. Если отношение производных стремится к некоторому пределу (конечному или бесконечному), то к этому же пределу стремится и отношение функций.

Примеры.

1. $\lim\limits_{x\to 2}\dfrac{x^2-5x+6}{x^2-3x+2}=\left[\dfrac{0}{0}\right]=\lim\limits_{x\to 2}\dfrac{2x-5}{2x-3}=-1.$

2. $\lim\limits_{x\to+\infty}\dfrac{\ln x}{x}=\left[\dfrac{\infty}{\infty}\right]=\lim\limits_{x\to\infty}\dfrac{1/x}{1}=0.$

Может случиться так, что отношение производных снова приведет к неопределенности вида $[0/0]$ или $[\infty/\infty]$. Тогда, рассматривая производные как исходные функции, перейдем к пределу отношения вторых производных и т. д.:

$$\lim_{x\to a}\frac{\varphi(x)}{g(x)}=\frac{[0/0]}{[\infty/\infty]}=\lim_{x\to a}\frac{\varphi'(x)}{g'(x)}=\frac{[0/0]}{[\infty/\infty]}=\lim_{x\to a}\frac{\varphi''(x)}{g''(x)}=\ldots$$

Если на некотором шаге мы получим предел, то его значение и будет искомым пределом отношения исходных функций.

3. $$\lim_{x\to\infty}\frac{3x^3+4x^2+2x-5}{5x^4+8x-1}=\left[\frac{\infty}{\infty}\right]=$$
$$=\lim_{x\to\infty}\frac{9x^2+8x+2}{20x^3+8}=\left[\frac{\infty}{\infty}\right]=\lim_{x\to\infty}\frac{18x+8}{60x^2}=\left[\frac{\infty}{\infty}\right]=$$
$$=\lim_{x\to\infty}\frac{18}{60x}=0.$$

4. $$\lim_{x\to 0}\frac{1-\cos 5x}{x^2}=\left[\frac{0}{0}\right]=\lim_{x\to 0}\frac{5\sin 5x}{2x}=\left[\frac{0}{0}\right]=$$
$$=\lim_{x\to 0}\frac{5\cos 5x\cdot 5}{2}=\frac{25}{2}.$$

При применении формулы (2.21) может оказаться, что предел в правой ее части не существует. Это говорит лишь о том, что правило Лопиталя неприменимо и нужно вычислять исходный предел каким-либо другим способом.

5. $$\lim_{x\to 0}\frac{x^2\sin(1/x)}{\sin x}=\left[\frac{0}{0}\right]=\lim_{x\to 0}\frac{2x\ \sin(1/x)-\cos(1/x)}{\cos x}.$$

Этот последний предел не существует, так как $\cos(1/x)$ при $x\to 0$ не стремится ни к какому пределу (конечному или бесконечному), поэтому правило Лопиталя неприменимо.

Вычислим предел иным способом:

$$\lim_{x\to 0}\frac{x^2\sin\dfrac{1}{x}}{\sin x}=\lim_{x\to 0}\left(\frac{x^2}{\sin x}\sin\frac{1}{x}\right)=\left[\text{так как}\quad\lim_{x\to 0}\frac{x}{\sin x}=1\right]=$$
$$=\lim_{x\to 0}\left(x\sin\frac{1}{x}\right)=0\left[\begin{array}{l}x\text{ — бесконечно малая}\\ \text{при } x\to 0,\\ \text{а } \sin\dfrac{1}{x}\text{ — ограничена}\end{array}\right].$$

К неопределенности вида $[0/0]$ или $[\infty/\infty]$ можно легко свести и другие виды неопределенностей:

$$\lim_{x\to a}h(x)\,m(x)=[0\cdot\infty]=\left[\begin{array}{l}\displaystyle\lim_{x\to a}\frac{h(x)}{\dfrac{1}{m(x)}}=\left[\frac{0}{0}\right]=\ldots,\\[4ex] \displaystyle\lim_{x\to a}\frac{m(x)}{\dfrac{1}{h(x)}}=\left[\frac{\infty}{\infty}\right]=\ldots,\end{array}\right.$$

т. е. неопределенность вида $[0\cdot\infty]$ можно привести к одной из стандартных неопределенностей для дальнейшего применения правила Лопиталя.

а) $\lim\limits_{x\to 1}(1-x)\ \operatorname{tg}\dfrac{\pi x}{2} = [0\cdot\infty] =$

$$= \lim_{x\to 1}\frac{1-x}{\operatorname{ctg}\dfrac{\pi x}{2}} = \left[\frac{0}{0}\right] = \lim_{x\to 1}\frac{-1}{-\dfrac{\pi}{2}\,\dfrac{1}{\sin^2\dfrac{\pi x}{2}}} = \lim_{x\to 1}\frac{2}{\pi}\sin^2\frac{\pi x}{2} = \frac{2}{\pi}.$$

б) $\lim\limits_{x\to a}\left[k(x)-h(x)\right] = [\infty-\infty] =$

$$= \lim_{x\to a}k(x)\cdot h(x)\cdot\left[\frac{1}{h(x)}-\frac{1}{k(x)}\right] = [0\cdot\infty].$$

в) Неопределенности видов $[1^\infty]$, $[0^\infty]$, $[\infty^0]$ раскрываются с помощью основного логарифмического тождества:

$$\lim_{x\to a}f(x)^{g(x)} = \lim_{x\to a}e^{\ln f(x)^{g(z)}} = \lim_{x\to a}e^{g(x)\ln f(x)} = e^{\lim\limits_{x\to a}g(x)\ln(x)}.$$

Для всех трех рассматриваемых случаев вычисление предела $\lim\limits_{x\to a}g(x)\,f(x)$ снова приведет к раскрытию неопределенности вида $[0\cdot\infty]$, о которой мы говорили в пункте (а) настоящего замечания.

7. $\lim\limits_{x\to 0}(\cos x)^{1/x^2} = [1^\infty] = ?$

$$\lim_{x\to 0}(\cos x)^{1/x^2} = \lim_{x\to 0}e^{\ln(\cos x)^{1/x^2}} = \lim_{x\to 0}e^{(1/x^2)\ln\cos x}.$$

Вычисление последнего предела свелось к вычислению предела от его показателя:

$$\lim_{x\to 0}e^{(1/x^2)\ln\cos x} = \exp\left(\lim_{x\to 0}\frac{1}{x^2}\ln\cos x\right) = [0\cdot\infty] =$$

$$= \exp\left(\lim_{x\to 0}\frac{\ln\cos x}{x^2}\right) = \left(\frac{0}{0}\right) = \exp\left(\lim_{x\to 0}\frac{\dfrac{1}{\cos x}(-\sin x)}{2x}\right) =$$

$$= \exp\left(-\lim_{x\to 0}\frac{\operatorname{tg}x}{2x}\right) = \left[\frac{0}{0}\right] = \exp\left(-\lim_{x\to 0}\frac{1}{\cos^2 x\cdot 2}\right) = e^{-1/2}.$$

Правило Лопиталя часто используется при отыскании асимптот плоских кривых.

Самостоятельная работа

Правило Лопиталя

1. $\lim\limits_{x\to 0}\dfrac{e^x+e^{-x}-2}{\sin^2 x}$. **2.** $\lim\limits_{x\to -1}\dfrac{x^3+1}{\sin(x+1)}$.

3. $\lim\limits_{x\to 0}\dfrac{\sqrt{1+x\sin x}-1}{e^{x^2}-1}$. **4.** $\lim\limits_{x\to 0}\dfrac{1+x\sin x-\cos 2x}{\sin^2 x}$.

5. $\lim\limits_{x\to 0}\dfrac{x^2(e^x-e^{-x})}{e^{x^3+1}-e}$. **6.** $\lim\limits_{x\to 0}\dfrac{\sin 2x-2\sin x}{x\ln\cos 3x}$.

7. $\lim\limits_{x\to 0}\dfrac{1-\cos 2x+\operatorname{tg}^2 x}{x\sin 3x}$. **8.** $\lim\limits_{x\to \pi/4}\dfrac{\sin x-\cos x}{\ln\operatorname{tg} x}$.

9. $\lim\limits_{x\to 0}\dfrac{e^{\sin 2x}-e^{\sin x}}{\operatorname{tg} x}$. **10.** $\lim\limits_{x\to 1}\dfrac{1-x}{\log_2 x}$. **11.** $\lim\limits_{x\to 1}\dfrac{2^x-2}{\ln x}$.

12. $\lim\limits_{x\to 0}\dfrac{3^{x+1}-3}{\ln(1+x\sqrt{1+xe^x})}$. **13.** $\lim\limits_{x\to \frac{\pi}{2}}\dfrac{1-\sin^3 x}{\cos^2 x}$.

14. $\lim\limits_{x\to 1}\dfrac{e^x-e}{\sin(x^2-1)}$. **15.** $\lim\limits_{x\to 3}\dfrac{\log_3 x-1}{\operatorname{tg}\pi x}$.

16. $\lim\limits_{x\to 0}\dfrac{1+1/x}{1-1/x}$. **17.** $\lim\limits_{x\to 0}\dfrac{x+1/\sqrt{x}}{\sqrt{x+1/x}}$.

18. $\lim\limits_{x\to +0}\dfrac{\ln\sin 3x}{\ln\operatorname{tg}(\pi/2-x)}$. **19.** $\lim\limits_{x\to +0}\dfrac{3^{1/x}}{\log_3 x}$.

20. $\lim\limits_{x\to +\pi/2}\dfrac{\operatorname{tg} x}{5^{\operatorname{tg} x}}$. **21.** $\lim\limits_{x\to -2/3}\dfrac{\ln(2+3x)}{\operatorname{tg}\dfrac{3\pi}{4}x}$. **22.** $\lim\limits_{x\to \infty}\dfrac{3^x}{x^3}$.

23. $\lim\limits_{x\to +0}\dfrac{\ln\sin x}{\ln x}$. **24.** $\lim\limits_{x\to +0}\dfrac{\ln\operatorname{tg} x}{\operatorname{ctg} x}$. **25.** $\lim\limits_{x\to \infty}\dfrac{x^{10}}{10^x}$.

26. $\lim\limits_{x\to \infty}\dfrac{x^2}{e^x}$. **27.** $\lim\limits_{x\to \infty}\dfrac{\ln(x^2+e^x)}{\ln(x^4+e^{2x})}$. **28.** $\lim\limits_{x\to \infty}\dfrac{x^5+x^3}{5^x+x}$.

29. $\lim\limits_{x\to +0}\dfrac{\ln\sin x}{1/x}$. **30.** $\lim\limits_{x\to \pi/2}\dfrac{\operatorname{ctg}(x-\pi/2)}{\operatorname{tg} 3x}$.

2.3.4. Отыскание асимптот плоских кривых

Пусть кривая является графиком функции $y = f(x)$. Пусть при $x \to +\infty$ ($x \to -\infty$) эта кривая неограниченно приближается к некоторой прямой $y = kx + b$. В этом случае данная прямая называется асимптотой кривой, причем при $k \neq 0$ — наклонной асимптотой, а при $k = 0$ — горизонтальной.

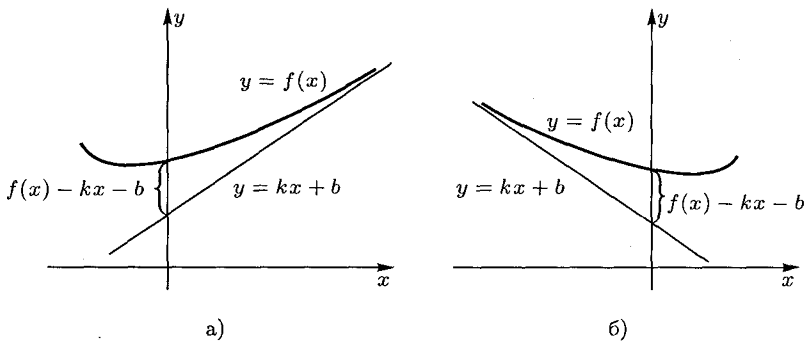


Рис. 2.8. Наклонные асимптоты кривой: а) при $x \to +\infty$; б) при $x \to -\infty$.

Определение. Прямая $y = kx + b$ называется *асимптотой кривой* $f(x)$ при $x \to +\infty$ ($x \to -\infty$), если

$$\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} [f(x) - kx - b] = 0. \tag{2.22}$$

Пусть кривая $y = f(x)$ имеет асимптоту $y = kx + b$ при $x \to +\infty$. Тогда, согласно определению, имеет место равенство (2.22). Разделим (2.22) на x:

$$\lim_{x \to +\infty} \left[\frac{f(x)}{x} - k - \frac{b}{x}\right] = 0,$$

или

$$\lim_{x \to +\infty} \frac{f(x)}{x} - k - \lim_{x \to +\infty} \frac{b}{x} = 0,$$

откуда

$$k = \lim_{x \to +\infty} \frac{f(x)}{x}. \tag{2.23}$$

Из (2.22) следует, что

$$b = \lim_{x \to +\infty} [f(x) - kx]. \tag{2.24}$$

Таким образом, если кривая имеет асимптоту $y = kx + b$ при $x \to +\infty$, то справедливы равенства (2.23), (2.24), и наоборот, если выполняются равенства (2.23), (2.24), то кривая имеет асимптоту $y = kx + b$.

Те же самые рассуждения справедливы при $x \to -\infty$.

Заметим, что если по крайней мере один из пределов (2.23), (2.24) не существует или бесконечен, то кривая не имеет асимптоты при $x \to +\infty$ ($x \to -\infty$).

Пример. $y = xe^{-x}$.

Находим асимптоту при $x \to +\infty$:

$$k = \lim_{x \to +\infty} \frac{f(x)}{x} = \lim_{x \to +\infty} \frac{x}{e^x \cdot x} = \lim_{x \to +\infty} \frac{1}{e^x} = 0,$$

$$b = \lim_{x \to +\infty} [f(x) - kx] = \lim_{x \to +\infty} \left[\frac{x}{e^x} - 0 \cdot x\right] =$$

$$= \lim_{x \to +\infty} \frac{x}{e^x} = \left[\frac{\infty}{\infty}\right] = \left| \begin{matrix} \text{воспользуемся} \\ \text{правилом} \\ \text{Лопиталя} \end{matrix} \right| = \lim_{x \to +\infty} \frac{1}{e^x} = 0.$$

Следовательно, данная функция имеет горизонтальную (так как $k = 0$) асимптоту $y = 0$.

Найдем асимптоту при $x \to -\infty$:

$$k = \lim_{x \to -\infty} \frac{f(x)}{x} = \lim_{x \to -\infty} \frac{1}{e^x} = +\infty,$$

т. е. при $x \to -\infty$ асимптоты график не имеет.

Пример. $y = \dfrac{x^2}{x-1}$.

Здесь

$$k = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{x^2}{(x-1) \cdot x} = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{x}{x-1} = \left[\frac{\infty}{\infty}\right] = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{1}{1} = 1,$$

$$b = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(\frac{x^2}{x-1} - 1 \cdot x \right) = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{x^2 - x^2 + x}{x-1} =$$

$$= \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{x}{x-1} = \left[\frac{\infty}{\infty}\right] = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{1}{1} = 1.$$

Итак, $k = 1$, $b = 1$. Следовательно, график имеет асимптоту $y = x + 1$, причем как в случае $x \to +\infty$, так и в случае $x \to -\infty$.

Пусть a — некоторое число. Если кривая $y = f(x)$ неограниченно приближается к вертикальной прямой $x = a$ при $x \to a$, то эта прямая, параллельная оси OY, называется *вертикальной асимптотой* (см. рис. 2.9)

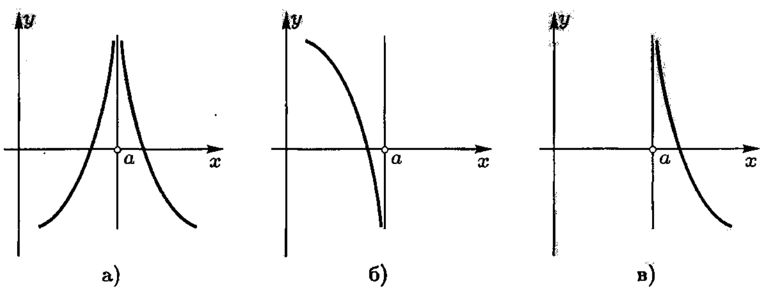


Рис. 2.9. Вертикальные асимптоты: а) при $x \to a$; б) при $x \to a - 0$; в) при $x \to a + 0$.

Итак, прямая $x = a$ называется вертикальной асимптотой кривой $y = f(x)$, если выполняется одно из трех условий:

$$\left[\begin{array}{l} \lim\limits_{x \to a} f(x) = \infty, \\ \lim\limits_{x \to a-0} f(x) = \infty, \\ \lim\limits_{x \to a+0} f(x) = \infty. \end{array}\right.$$

Следовательно, для отыскания вертикальных асимптот нужно найти точки $x_1, x_2, \ldots$, вблизи которых $f(x)$ неограниченно возрастает по модулю. Тогда прямые $x = x_1$, $x = x_2$, ... будут вертикальными асимптотами.

Пример. $y = \dfrac{x^2}{x-1}$.

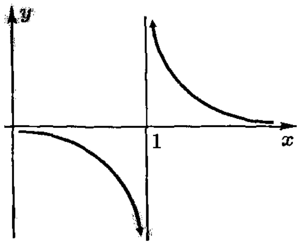


Рис. 2.10. Вертикальная асимптота функции $y = x^2/(x-1)$ (схематически).

Точка $x = 1$ не входит в область определения данной функции.

Вычислим пределы: $\lim\limits_{x\to 1-0} f(x) = \lim\limits_{x\to 1-0} \dfrac{x^2}{x-1} = -\infty$ и $\lim\limits_{x\to 1+0} f(x) = \lim\limits_{x\to 1+0} \dfrac{x^2}{x-1} = +\infty$. Следовательно, $x = 1$ — уравнение вертикальной асимптоты, причем вблизи точки $x = 1$ функция будет вести себя, как показано на рис. 2.10.

Самостоятельная работа

Раскрыть неопределенность $[\infty - \infty]$.

1. $\lim\limits_{x\to 1}\left(\dfrac{x}{x-1} - \dfrac{1}{\ln x}\right)$. **2.** $\lim\limits_{x\to 1}\left(\dfrac{1}{\ln x} - \dfrac{1}{x-1}\right)$.

3. $\lim\limits_{x\to 0}\left(\dfrac{1}{x} - \dfrac{1}{e^x-1}\right)$. **4.** $\lim\limits_{x\to 1}\left(\dfrac{1}{\ln x} - \dfrac{x}{\ln x}\right)$.

5. $\lim\limits_{x\to 0}\left(\dfrac{1}{x^2} - \operatorname{ctg}^2 x\right)$.

Раскрыть неопределенность $[0\cdot\infty]$.

1. $\lim\limits_{x\to 0} \operatorname{ctg} 2x \cdot \arcsin x$. **2.** $\lim\limits_{x\to \pi} (x-\pi)\cdot \operatorname{tg}\dfrac{x}{2}$.

3. $\lim\limits_{x\to \infty} \ln x (e^{1/x} - 1)$. **4.** $\lim\limits_{x\to \infty} e^{-x^2}\cdot x^2$.

5. $\lim\limits_{x\to \pi/2} (\pi/2 - x)\operatorname{tg} x$.

Раскрыть неопределенность $[1^\infty]$.

1. $\lim\limits_{x\to 0} (\cos x)^{1/x^2}$. **2.** $\lim\limits_{x\to 0} (1+x)^{\ln x}$.

3. $\lim\limits_{x\to 0} (\operatorname{tg}(\pi/4 - x))^{\operatorname{ctg} x}$.

4. $\lim\limits_{x\to 0}(\cos\pi x)^{\frac{1}{x\sin\pi x}}$. **5.** $\lim\limits_{x\to 0}(2-e^{\sin x})\operatorname{ctg}\pi x$.

Раскрыть неопределенность $[0^0; \infty^0]$.

1. $\lim\limits_{x\to 0}(\operatorname{tg}x)^{3x}$. **2.** $\lim\limits_{x\to 0}(\sin x)^{x}$. **3.** $\lim\limits_{x\to 0}(\operatorname{tg}2x)^{\arcsin x}$.

4. $\lim\limits_{x\to 0}(e^x-e^{-x})^{\sin 2x}$. **5.** $\lim\limits_{x\to\infty}(x+2)^{1/x}$.

6. $\lim\limits_{x\to\pi/2}(\operatorname{tg}x+1)^{\operatorname{ctg}x}$.

§ 2.4. Применение производной к исследованию функции

2.4.1. Признак постоянства функции. Признаки возрастания и убывания функций

Теорема 1. Для того чтобы функция $y=f(x)$ была постоянна на некотором промежутке, необходимо и достаточно, чтобы на этом промежутке $f'(x)=0$.

Предположим теперь, что $y=f(x)\neq\text{const}$.

Тогда можно выделить интервалы, в которых функция возрастает или убывает (см. п. 1.1).

Теорема 2. (Необходимый признак монотонности.)

а) Если $y=f(x)$ дифференцируема и возрастает на некотором промежутке, то $f'(x)\geqslant 0$.

б) Если $y=f(x)$ дифференцируема и убывает на некотором промежутке, то $f'(x)\leqslant 0$ на этом промежутке.

Доказательство. Пусть x — некоторая точка рассматриваемого промежутка, а Δx — величина настолько малая по модулю, что $x+\Delta x$ также принадлежит этому промежутку. Воспользуемся определением производной:

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

Если $f(x)$ возрастает на рассматриваемом промежутке, то $f(x+\Delta x)>f(x)$ при $\Delta x>0$ и $f(x+\Delta x)<f(x)$ при $\Delta x<0$ (см. рис. 2.11.).

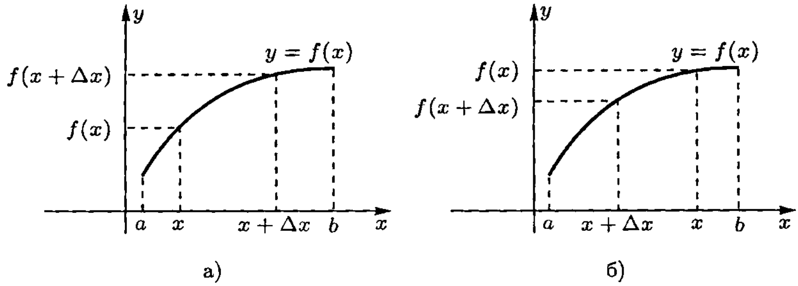


Рис. 2.11. Графики возрастающей на (a, b) функции: а) $\Delta x > 0$; б) $\Delta x < 0$.

Таким образом, дробь под знаком предела в определении производной положительна. Но пределом положительной величины может быть либо положительное число, либо ноль. Поэтому $f'(x) \geqslant 0$.

Аналогично можно доказать теорему для случая убывания функции $y = f(x)$.

Теорема 3. (Достаточный признак монотонности.)

а) Если на некотором интервале (a, b) $f'(x) > 0$, то $f(x)$ возрастает на (a, b).

б) Если $f'(x) < 0$ на некотором интервале (a, b), то $f(x)$ убывает на данном интервале.

Данную теорему примем без доказательства.

Ранее мы останавливались на геометрическом смысле производной. Исходя из этого утверждения теоремы 2 и 3 могут быть проиллюстрированы геометрически. В любой точке промежутка возрастания функции $y = f(x)$ касательная к ее графику об-

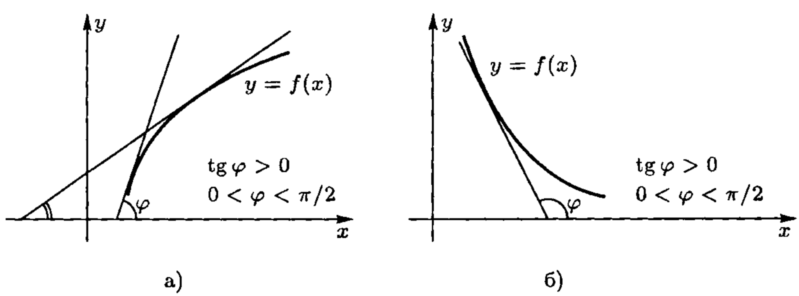


Рис. 2.12. Связь между монотонностью функции и углом наклона касательной: а) для возрастающей функции; б) для убывающей функции.

разует острый угол φ с осью OX (рис. 2.12а), на промежутке убывания функции угол φ тупой (рис. 2.12б).

Таким образом, промежутку возрастания функции $f(x)$ соответствует случай $\operatorname{tg}\varphi = f'(x) > 0$, а промежутку убывания — случай $\operatorname{tg}\varphi = f'(x) < 0$.

Пример. $y = x^2 - 2x + 1$.

Область определения данной функции: $x \in (-\infty, +\infty)$.

Найдем $y' = 2x - 2 = 2(x - 1)$. Следовательно, $y' < 0$ при $x < 1$ и $y' > 0$ при $x > 1$.

Таким образом, на интервале $(-\infty, 1)$ функция убывает, а при $x \in (1, +\infty)$ возрастает. В точке $x = 1$ убывание переходит в возрастание.

2.4.2. Экстремумы функции. Необходимое условие экстремума

Функция $y = f(x)$ может убывать или возрастать не во всей своей области определения. Эта область часто распадается на промежутки, в одних из которых функция убывает, в других — возрастает.

Точка x_0, отделяющая промежуток возрастания от промежутка убывания и принадлежащая области определения функции, называется точкой экстремума.

Точки экстремума бывают двух типов: точки максимума функции, где функция переходит от возрастания к убыванию (рис. 2.13; точки x_1 и x_3) и точки минимума (см. рис. 2.13; точки x_2 и x_4), где функция переходит от убывания к возрастанию.

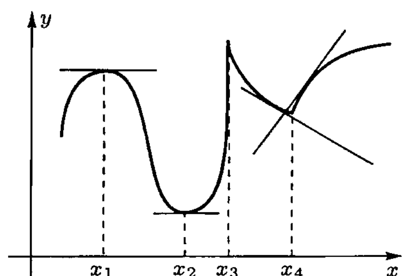


Рис. 2.13. Точки экстремума.

В точках максимума величина $f(x)$ больше, а в точках минимума — меньше, чем во всех соседних достаточно близких точках.

Таким образом, логично ввести следующие определения.

Определение. Точка x_0 называется *точкой максимума* функции $y = f(x)$, если для любой точки x, принадлежащей достаточно малой окрестности точки x_0,

$$f(x) < f(x_0) \tag{2.25}$$

и x_0 — *точка минимума*, если в малой окрестности этой точки

$$f(x) > f(x_0). \tag{2.26}$$

Определение. Значения функции в точках максимума и минимума называются, соответственно, максимальным и минимальным значением функции.

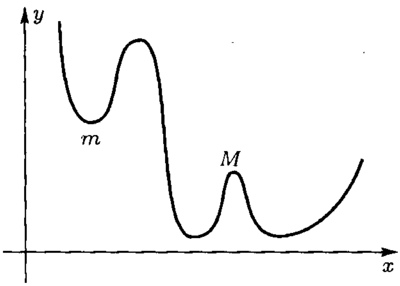


Рис. 2.14. Точки экстремума.

Отметим, что речь идет о локальных экстремумах, т. е. $\max M$ может быть меньше $\min m$ (рис. 2.14).

Таким образом, максимум и минимум функции — это не наибольшее и наименьшее значение функции во всей области определения, а только max или min значения функции по сравнению со значениями функции во всех соседних, близких точках.

Теорема 4. Если $y = f(x)$ дифференцируема в точке x_0 и имеет в этой точке экстремум, то $f'(x_0) = 0$.

Доказательство. Пусть в точке x_0 функция имеет максимум. Тогда по определению (2.25) для произвольного достаточно малого по модулю Δx будет

$$\left.\begin{aligned} &\frac{f(x_0 - \Delta x) - f(x_0)}{\Delta x} < 0 \quad \text{при} \quad \Delta x > 0 \\ \text{и}\quad & \\ &\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} > 0 \quad \text{при} \quad \Delta x < 0. \end{aligned}\right\} \tag{2.27}$$

При $\Delta x \to 0$ оба эти отношения стремятся к общему пределу $f'(x_0)$, откуда следует, что $f'(x_0) = 0$.

Если x_0 — точка минимума, то, воспользовавшись определением (2.26), можно провести аналогичное доказательство.

Эту теорему можно истолковать геометрически, используя вновь смысл производной: если в точке экстремума график функции $y = f(x)$ имеет касательную и эта касательная не параллельна OY (как для точки x_3 на рис. 2.13), то эта касательная обязательно параллельна оси OX (точки x_1 и x_2 на рис. 2.13).

Кроме точек экстремума, где функция дифференцируема и где, как показано выше, $f'(x_0) = 0$ (точки x_1 и x_2 на рис. 2.13), могут быть такие точки экстремума, где функция $f(x)$ недифференцируема.

В точке x_3 на рис. 2.13 — максимум (касательная параллельна оси ординат, $f'(x_0) = \infty$), в точке x_4 — минимум (касательная к графику вообще не существует).

Из всего сказанного выше следует практическое правило: точки, в которых функция имеет экстремумы, надлежит искать среди точек, в которых:

1) $f'(x) = 0$ либо

2) $f'(x) = \infty$ либо

3) $f'(x)$ не существует, причем предполагается, что точки эти принадлежат области определения функции. Точки указанных типов называются критическими точками I рода.

Отметим, что не в каждой критической точке функция имеет экстремум.

Пример. $y = x^3$.

$$y' = 3x^2.$$

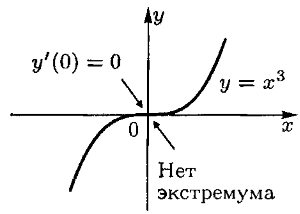


Рис. 2.15. Кубическая парабола.

Точка $x = 0$ будет для данной функции критической точкой I рода, так как $y'(0) = 0$. Однако в этой точке экстремума нет (рис. 2.15).

Как же решить вопрос о наличии экстремума в критической точке?

Ответ на этот вопрос мы найдем в следующем параграфе.

2.4.3. Достаточное условие экстремума (по I производной)

Теорема 5 Пусть x_0 будет критической точкой I рода для функции $y = f(x)$. Тогда, если в окрестности точки x_0 будет меняться знак производной, то точка x_0 — точка экстремума причем, если знак меняется с «+» на «−», то в точке x_0 —

максимум, если с «−» на «+», то минимум. Иначе, если

$$\left.\begin{array}{l} 1)\ x_0 \text{ — критическая точка I рода,} \\ 2)\ f'(x) > 0 \text{ при } x < x_0 \text{ и} \\ 3)\ f'(x) < 0 \text{ при } x > x_0 \end{array}\right\} \Rightarrow x_0 \text{ — точка максимума,}$$

если

$$\left.\begin{array}{l} 1)\ x_0 \text{ — критическая точка I рода,} \\ 2)\ f'(x) < 0 \text{ при } x < x_0 \text{ и} \\ 3)\ f'(x) > 0 \text{ при } x > x_0 \end{array}\right\} \Rightarrow x_0 \text{ — точка минимума.}$$

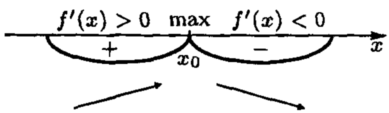


Рис. 2.16. Схема нахождения максимума.

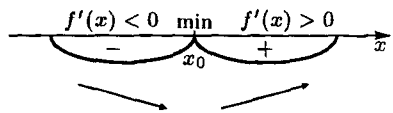


Рис. 2.17. Схема нахождения минимума.

Доказательство этой теоремы вытекает из теоремы 3 п. 2.4.1 (так, если при $x < x_0$ функция возрастает, а при $x > x_0$ убывает, то в точке x_0 функция переходит от возрастания к убыванию, т. е. x_0 — точка максимума).

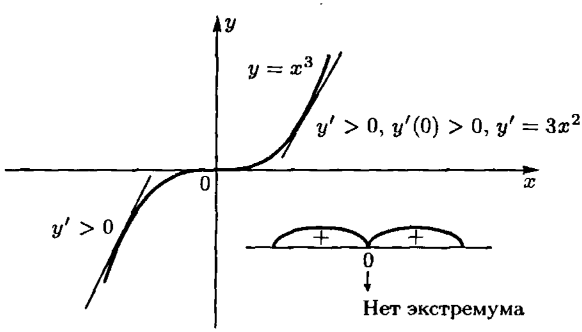


Рис. 2.18. Схема, иллюстрирующая отсутствие экстремума у кубической параболы.

Заметим, что если при переходе через критическую точку $f'(x)$ не меняет знака, то в критической точке экстремума нет (см. пример с кубической параболой).

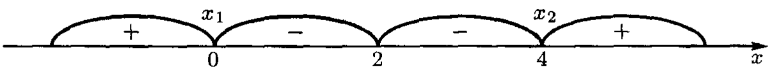
Пример. $y = \dfrac{x^2}{x-2}$.

Функция не определена при $x = 2$.

$$y' = \frac{x^2 - 4x}{(x-2)^2} = \frac{x(x-4)}{(x-2)^2},$$

$y' = 0$ при $x_1 = 0$ и $x_2 = 4$.

Нанесем на ось OX точки, не принадлежащие области определения функции, а также критические точки I рода:



Тогда ось разобьется на 4 интервала, в каждом из которых исследуем знак y'.

Расставим эти знаки над интервалами. Тогда, согласно выше изложенной теории, $f(x)$ возрастает при $x \in (-\infty, 0) \cup (4, +\infty)$ и убывает при $x \in (0, 2) \cup (2, 4)$.

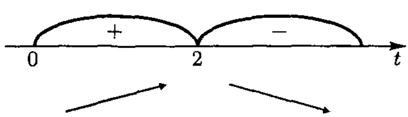
Так как в окрестности точки $x_1 = 0$ знак y' поменялся с «+» на «−», то $x_1 = 0$ — точка максимума, а в окрестности точки $x_2 = 4$ знак y' изменился с «−» на «+», и, следовательно, $x_2 = 4$ — точка минимума. Тогда $y_{\min} = f(4) = 8$, а $y_{\max} = f(0) = 0$.

Пример. Дано уравнение прямолинейного движения точки $x = 6t^2 - t^3$ (t — в секундах, x — в метрах). Найти наибольшую скорость точки.

Решение. Скорость точки $v = x'(t) = 12t - 3t^2$. Исследуем, когда $v = v_{\max}$. Итак, нужно найти наибольшее значение функции $v(t) = 12t - 3t^2$ при $t \in (0, +\infty)$.

На данном промежутке $v(t)$ дифференцируема и $v'(t) = 12 - 6t$. Найдем критические точки I рода: $v'(t) = 0$, т. е. $12 - 6t = 0$, откуда $t = 2$.

Так как в окрестности точки $t = 2$ изменился знак $v'(t)$, то в точке $t = 2$ — экстремум, причем изменение знака произошло с «+» на «−», и, следовательно, $t = 2$ — точка максимума.



Тогда $v(2) = 12 \cdot 2 - 3 \cdot 2^2 = 12$ (м/с) — максимальное значение скорости.

2.4.4. Выпуклость и вогнутость кривых

Кривая, являющаяся графиком функции $y = f(x)$, называется выпуклой на некотором промежутке, если она целиком лежит под касательной, проведенной к ней в любой точке этого промежутка (рис. 2.19а).

Кривая называется вогнутой на некотором промежутке, если она целиком лежит над касательной, проведенной к ней в любой точке этого промежутка (рис. 2.19б).

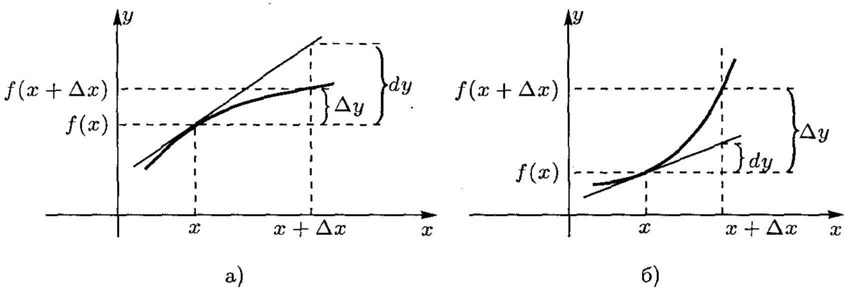


Рис. 2.19. Характер выпуклости кривых: а) — кривая выпукла; б) — кривая вогнута.

Если график выпуклый на промежутке, то $\Delta y < dy$, а если вогнутый, то $\Delta y > dy$ в любой точке x и для произвольного $\Delta x \neq 0$, настолько малого по модулю, чтобы точка $x + \Delta x$ не вышла за пределы рассматриваемого промежутка.

Теорема 6. Если функция $y = f(x)$дважды дифференцируема на некотором промежутке и на данном промежутке $f''(x) < 0$, то график функции выпуклый, если же $f''(x) > 0$, график функции вогнутый на данном промежутке.

Пример. $y = x^3$, $y' = 3x^2$, $y'' = 6x$.

$$y'' > 0 \quad \text{при} \quad x > 0, \qquad y'' < 0 \quad \text{при} \quad x < 0.$$

Следовательно, согласно последней теореме, $y = f(x)$ выпукла для $x < 0$ и вогнута для $x > 0$.

2.4.5. Точки перегиба графика функции

График функции $y = f(x)$ может быть выпуклым или вогнутым не во всей области определения этой функции. Область определения функции часто распадается на промежутки выпуклости и вогнутости (I и II на рис. 2.20). Точки, отделяющие друг от друга промежутки выпуклости и вогнутости, называются точками перегиба функции. Геометрически это значит, что кривая лежит над касательной с одной стороны от точки перегиба и под касательной с другой.

Аналогично определению критической точки первого рода можно ввести определение критической точки второго рода.

Точка x_0, принадлежащая области определения функции, называется критической точкой второго рода, если в этой точке выполняется одно из 3 условий:

1) $f''(x_0) = 0$ либо
2) $f''(x_0) = \infty$ либо
3) $f''(x_0)$ не существует.

Точки перегиба следует искать среди критических точек второго рода. Но не всякая критическая точка второго рода является

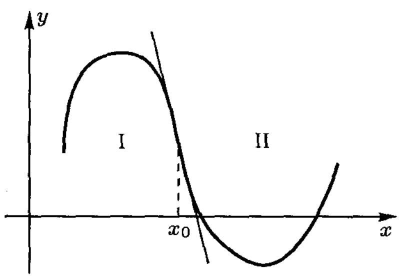


Рис. 2.20. Точки перегиба как границы промежутков выпуклости и вогнутости.

точкой перегиба. Однако если x_0 — критическая точка второго рода и в ее окрестности изменяется знак второй производной, то точка x_0 — точка перегиба.

Если x_0 — критическая точка второго рода и в ее окрестности изменился знак второй производной, то точка x_0 — точка перегиба.

Пример. $y = \dfrac{4x-1}{x^2}; \quad x \in (-\infty, 0) \cup (0, \infty)$

$$y' = \frac{(-2x+1)\cdot 2}{x^3} = \frac{-4x+2}{x^3},$$

$$y'' = \frac{8x - 6}{x^4}.$$

$y'' = 0$, если $x = 3/4$ — критическая точка второго рода.

На ось OX нанесем $x = 3/4$ и выколем точку $x = 0$ ($x = 0 \notin$ ООФ).

В каждом из получившихся трех интервалов исследуем знак второй производной и расставим над интервалами соответствующие знаки.

В окрестности точки $x = 3/4$ изменился знак y''. Следовательно, это точка перегиба.

Итак, в пунктах 2.4.1–2.4.5 мы рассмотрели задачи об исследовании функций по первой и второй производной. Теперь объединим эти знания, чтобы строить графики функций.

2.4.6. Полная схема исследования функций с помощью производных и построения графиков

Знание промежутков возрастания (убывания) функции, промежутков выпуклости (вогнутости), точек экстремума, перегиба, а также асимптот позволяет иметь достаточно полное представление о характере поведения функции, что дает возможность построить ее график.

I. Элементарное исследование

Для построения графика удобно провести следующее исследование:

1) Установить область определения функции.

2) Найти точки пересечения графика с осями координат.

3) Исследовать функцию на четность, отметив какую-либо симметрию графика (если она существует).

Напомним, что график четной функции симметричен относительно оси OY, а график нечетной функции — относительно начала координат.

II. Нахождение асимптот графика

См. рис. 2.21.

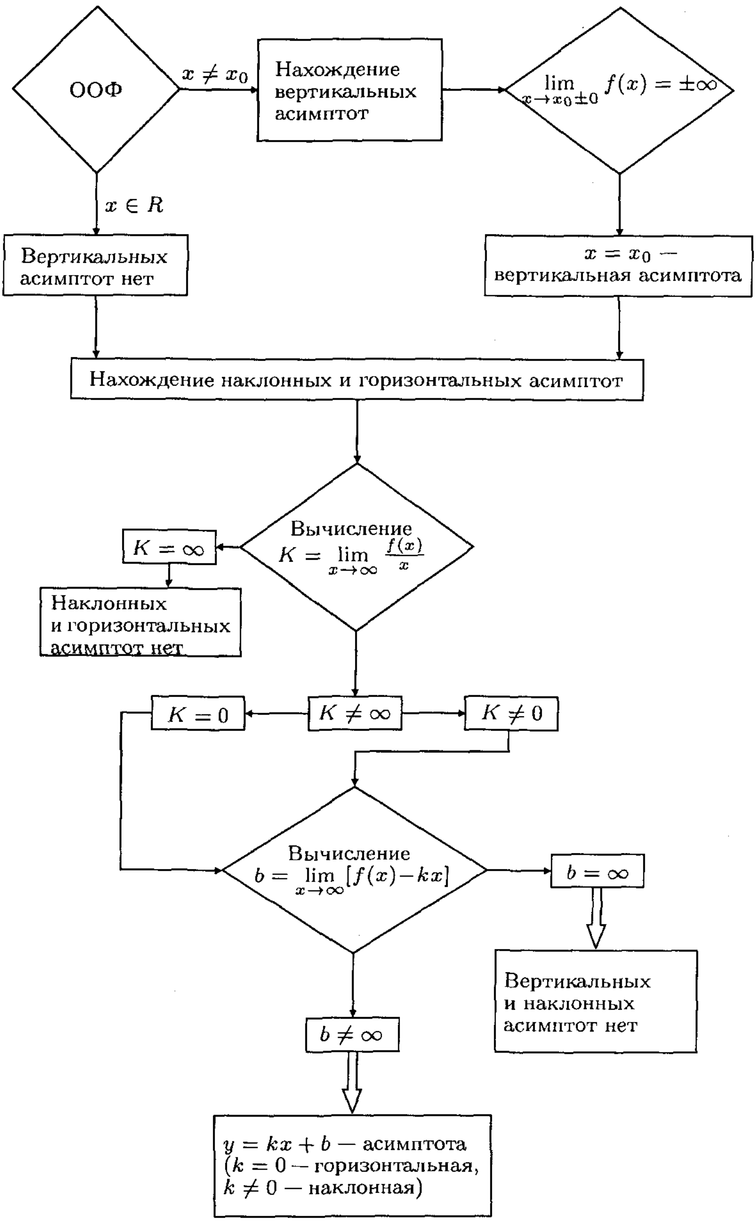


Рис. 2.21. Схема вычисления асимптот графиков функции.

III. Исследование функции по первой производной (промежутки монотонности, экстремумы)

1) Найти $y'(x)$.

2) Найти критические точки I рода.

3) На ось OX нанести точки, не принадлежащие области определения функции (выколов их), а также критические точки I рода.

4) Тогда ось OX разобьется на ряд интервалов. В каждом из интервалов определить знак $y'(x)$ (для этого достаточно найти знак y' в любой точке каждого интервала, поскольку внутри любого интервала знак $y'(x)$ не меняется) и отметить полученные результаты над соответствующими интервалами.

Пусть, например, $f(x)$ имеет пять критических точек первого рода $x_1, x_2, \ldots, x_5$, а в точке x_0 функция не определена (рис. 2.22).

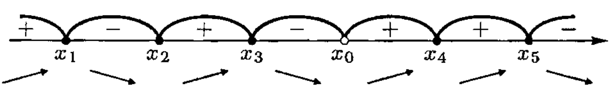


Рис. 2.22. Схема для исследования $y = f(x)$ по y'.

Здесь «+» означает, что $f'(x) > 0$, а «−» , что $f'(x) < 0$.

По знаку $y'(x)$ найти интервалы возрастания и убывания (на рис. 2.22 символ «$\searrow$» означает убывание функции в данном интервале, а «$\nearrow$» — возрастание $f(x)$).

Так, функция $y = f(x)$, рассмотренная на рис. 2.22, ведет себя следующим образом: она, возрастает для $x \in (-\infty, x_1) \cup$ $\cup (x_2, x_3) \cup (x_0, x_4) \cup (x_4, x_5)$ и убывает при $x \in (x_1, x_2) \cup (x_3, x_0) \cup$ $\cup\, (x_5, +\infty)$.

5) Найти точки экстремумов функции (при переходе через такие точки $y'(x)$ меняет знак). Однако заметим, что точки экстремума — это точки, которые обязательно принадлежат области определения функции.

Так, на рис. 2.22 точки x_1, x_2, x_3, x_5 — точки экстремумов, причем точки x_1, x_3, x_5 — точки максимума (так как в их окрестностях знак y' поменялся с «+» на «−»), а точка x_2 — точка минимума (знак y' в окрестности данной точки изменился с «−» на «+»). В окрестности точки x_0 также менялся знак производной, но эта точка не является точкой экстремума, так как она не входит в область определения функции. Точка x_4 не является точкой экстремума, так как и слева, и справа от нее знак y' один и тот же («+»).

6) Вычислим значения функции в точках экстремума: $M_1(x_1, f(x_1))$, $M_2(x_3, f(x_3))$, $M_3(x_5, f(x_5))$ — максимумы, и $m(x_0, f(x_0))$, $m(x_2, f(x_2))$ — минимумы.

IV. Исследование функции по второй производной (характер выпуклости и точки перегиба)

1) Найти $y''(x)$.

2) Найти критические точки II рода.

3) На ось OX нанести точки, не принадлежащие области определения, а также критические точки II рода.

4) В каждом из получившихся интервалов определить знак $y''(x)$. Пусть, например, $f(x)$ имеет 3 критические точки II рода x_6, x_7, x_8, а в точке x_0 функция не определена, и знаки y'' распределились по интервалам как показано на рис. 2.23. Знаками «∪» и «∩» здесь показано направление выпуклости:

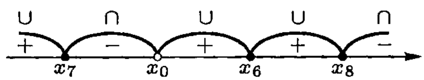


Рис. 2.23. Схема для исследования $y = f(x)$ по y''.

«∪» — функция вогнута $(f''(x) > 0)$,

«∩» — функция выпукла $(f''(x) < 0)$.

Как следует из Рис. 2.23 $y = f(x)$ выпукла для $x \in (x_7, x_0) \cup \cup (x_8, +\infty)$ и выгнута для $x \in (-\infty, x_7) \cup (x_0, x_6) \cup (x_6, x_8) \cup \cup (x_4, x_5)$.

5) Найти точки перегиба (точки, в окрестности которых $y''(x)$ меняет знак).

Как видно из рис. 2.23, точками перегиба будут: x_7, x_8. Точка x_0 не является точкой перегиба, так как она не принадлежит области определения, а точка x_6 — так как в ее окрестности знак второй производной постоянен.

6) Вычислить значения функции в точках перегиба.

$(\Pi_1(x_7, f(x_7), \Pi_2(x_8, f(x_8)$ — перегибы функции).

V. Объединение результатов пунктов 3 и 4

Очевидно, что если, например, функция $f(x)$ возрастала в некотором интервале (a, b) и была вогнутой, то для построения графика в данном интервале нужно взять аналог правой ветви параболы (рис. 2.24а).

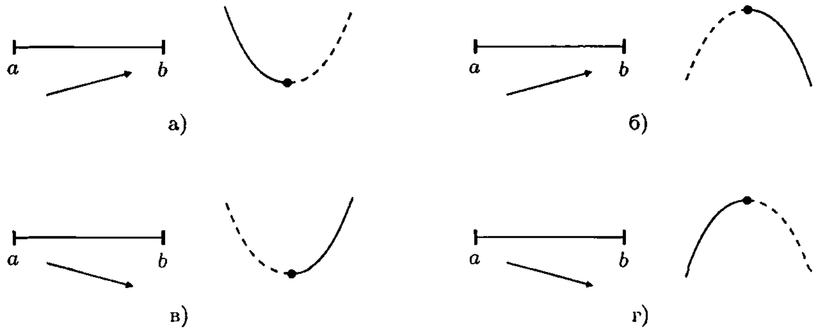


Рис. 2.24. Схема для исследования $y = f(x)$ по y''. Обозначения в тексте.

Если же функция возрастает на (a, b), но выпукла, берем аналог левой ветви параболы (см. рис. 2.24б) и т. д.

VI. Построение графика по результатам исследования 1–5

Отметим, что иногда для уточнения вида графика целесообразно вычислить дополнительно еще несколько значений функции в «обыкновенных» точках.

Изложенную схему проиллюстрируем следующим примером:

$$y = \frac{1 - 2x}{x^2 - x - 2}.$$

I. Элементарное исследование

1) Данная функция представляет собой дробь, а значит, она не определена, если знаменатель $x^2 - x - 2 = 0$.

Найдем корни знаменателя : $x_1 = -1$, $x_2 = 2$.

Следовательно, область определения функции имеет вид

$$x \in (-\infty, x - 1) \cup (-1, 2) \cup (2, +\infty).$$

2) Определим точки пересечения функции с осями координат.

Пусть $x = 0$. Тогда $y = -1/2$, т. е. точка $(0, -1/2)$ — точка пересечения графика с осью OY.

Если $y = 0$, то $1 - 2x = 0$, и, следовательно, $x = 1/2$. Точка $(1/2, 0)$ — пересечение графика с осью OX.

3) Рассмотрим функцию на аргументе $(-x)$:

$$f(-x) = \frac{1 - 2(-x)}{(-x)^2 - (-x) - 2} = \frac{1 + 2x}{x^2 + x - 2}.$$

Как видно, $f(-x) \neq f(x)$ и $f(-x) \neq -f(x)$, а значит, $y = f(x)$ — функция общего вида, и график ее не обладает симметрией ни относительно точки $O(0, 0)$, ни относительно оси OY.

II. Асимптоты

Так как точки $x = -1$ и $x = 2$ не принадлежат области определения функции, то следует рассматривать предел функции при стремлении аргумента к данным точкам слева и справа от них.

Итак,

$$\lim_{x \to -1+0} f(x) = \lim_{x \to -1+0} \frac{1 - 2x}{(x + 1)(x - 2)} = -\infty,$$
$$\lim_{x \to -1-0} f(x) = \lim_{x \to -1-0} \frac{1 - 2x}{(x + 1)(x - 2)} = +\infty.$$

Тогда по определению уравнение вертикальных асимптот имеет вид: $x = 2$ и $x = -1$

$$\lim_{x \to 2+0} f(x) = \lim_{x \to 2+0} \frac{1 - 2x}{(x + 1)(x - 2)} = -\infty,$$
$$\lim_{x \to 2-0} f(x) = \lim_{x \to 2-0} \frac{1 - 2x}{(x + 1)(x - 2)} = +\infty.$$

Теперь рассмотрим вопрос о существовании наклонных (горизонтальных) асимптот.

Вычислим

$$k = \lim_{x \to +\infty} \frac{f(x)}{x} = \lim_{x \to +\infty} \frac{1 - 2x}{x^3 - x^2 - 2x} = 0$$

(по правилу Лопиталя).

Тогда

$$b = \lim_{x \to +\infty} [f(x) - kx] = \lim_{x \to +\infty} \frac{1 - 2x}{x^2 - x - 2} = 0.$$

Итак, при $x \to +\infty$ мы получим: $k = 0$, $b = 0$, и, следовательно, существует горизонтальная асимптота $y = 0$.

Аналогичным образом, при $x \to -\infty$ $k = 0$, $b = 0$, а значит, $y = 0$ — уравнение горизонтальной асимптоты и при $x \to -\infty$. Наклонных ($k \neq 0$) асимптот нет.

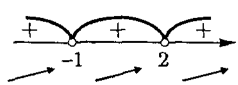
III. Исследование функции по $y'(x)$.

Вычислим $y'(x)$ по правилу дифференцирования дробей:

$$y'(x) = \frac{(1-2x)'(x^2-x-2)-(1-2x)(x^2-x-2)'}{(x^2-x-2)^2} =$$
$$= \frac{-2(x^2-x-2)-(1-2x)(2x-1)}{(x^2-x-2)^2} =$$
$$= \frac{-2(x^2-x-2)+(2x-1)^2}{(x^2-x-2)^2} =$$
$$= \frac{-2x^2+2x+4+4x^2-4x+1}{(x^2-x-2)^2} = \frac{2x^2-2x+5}{(x^2-x-2)^2}.$$

Приравняем уравнение $y'(x)$ к нулю. Тогда $2x^2-2x+5=0$. Но дискриминант этого уравнения отрицателен, и, следовательно, корней нет. Значит, критических точек I рода нет, и на ось OX нанесем только те точки, которые не принадлежат области определения функции:

В каждом из интервалов рассмотрим знак $y'(x)$:



Таким образом, в каждом из получившихся интервалов функция возрастает, а точек экстремума нет.

IV. Исследование функции по $y''(x)$.

1) Вычислим

$$y''(x) = \left(\frac{2x^2-2x+5}{(x^2-x-2)^2}\right)' =$$

$$= \frac{(4x-2)(x^2-x-2)^2 - (2x^2-2x+5)\cdot 2(x^2-x-2)\cdot(2x-1)}{(x^2-x-2)^4} =$$

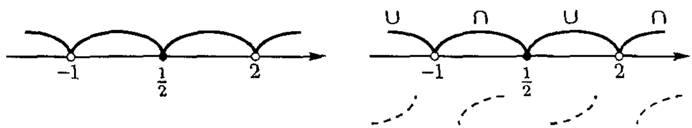
$$= \frac{(4x-2)\cdot(x^2-x-2-2x^2-2x-5)}{(x^2-x-2)^3} = \frac{(4x-2)\cdot(-x^2+x-7)}{(x^2-x-2)^3}.$$

2) Найдем критические точки II рода. Найдем корни уравнения $y''(x) = 0$:

$$(4x-2)(-x^2+x-7) = 0,$$

или

$$(2x-1)(x^2-x+7) = 0.$$

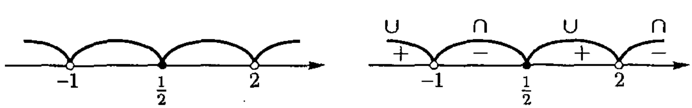


Первая скобка равна нулю, если $x = 1/2$, а вторая в 0 не обращается, так как дискриминант этого квадратного уравнения меньше нуля.

Итак $x = 1/2$ — критическая точка II рода.

3) Нанесем эту точку, а также точки $x = -1$ и $x = +2$ (не принадлежат области определения функции) на ось OX.

В каждом из полученных интервалов определим знак $y''(x)$, а также направление выпуклости:



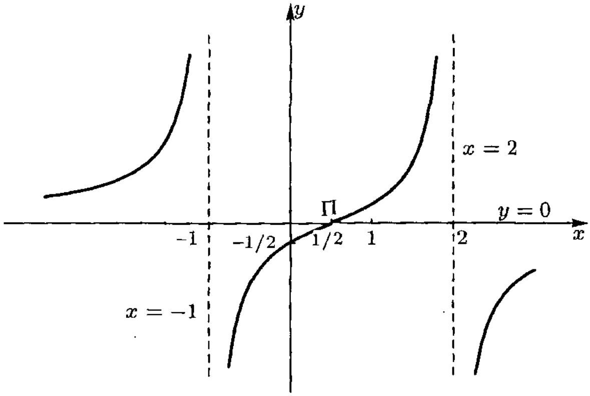
Итак, функция вогнута при $x \in (-\infty, -1) \cup (1/2, 2)$ и выпукла при $x \in (-1, 1/2) \cup (2, +\infty)$.

Точки $x = -1$ и $x = +2$ не являются точками перегиба, а $x = 1/2$ — точка перегиба.

4) Найдем $f(1/2)$; $\Pi(1/2, 0)$ — точка перегиба.

V. Объединим исследования пунктов III и IV.

VI. Построим график $y = f(x)$:



Самостоятельная работа

Построить график функции:

1. $y = \dfrac{x^2}{x^2 - 9}$. **2.** $y = \sqrt[3]{1 - x^2}$. **3.** $y = \dfrac{x^2}{x^2 - 4}$.

4. $y = \dfrac{3x^4 + 1}{x^3}$. **5.** $y = \dfrac{x^3}{x^2 - 1}$. **6.** $y = \dfrac{x^3 - 1}{4x^2}$.

7. $y = \dfrac{x^2 + 4}{x}$. **8.** $y = \dfrac{1 - 2x}{x^2}$. **9.** $y = \dfrac{4 - x^3}{x^2}$.

10. $y = \dfrac{2x^2}{4x^2 - 1}$. **11.** $y = \dfrac{x^2}{x^2 + 4}$.

12. $y = \dfrac{(x + 1)(x + 2)}{x^2}$. **13.** $y = \dfrac{x^2}{x^2 - 2}$.

14. $y = \dfrac{2}{x^2 + x + 1}$. **15.** $y = \dfrac{x^3}{x^2 + 1}$. **16.** $y = \dfrac{(x - 3)^2}{4(x - 1)}$.

17. $y = x^4 - 6x^2$. **18.** $y = \dfrac{x^3 + 2x^2 + 7x - 3}{2x^2}$.

19. $y = \dfrac{x^2 - x - 1}{x^2 - 2x}$. **20.** $y = \dfrac{x^2 - 6x + 13}{x - 3}$.

ГЛАВА 3
ФУНКЦИЯ ДВУХ ПЕРЕМЕННЫХ

§ 3.1. Понятие функции двух переменных

Пусть заданы два непустых множества $D \in R^2$ и $E \in R$. Если каждой паре значений (x, y) из данной области D соответствует вполне определенное единственное значение $z \in E$, то переменная z называется функцией 2 переменных, а x и y называются независимыми переменными или аргументами. Область D называется областью определения функции, множество $E = \{z = f(x, y), (x, y) \in D\}$ — множеством значений функции. Обозначение функции 2 переменных: $z = f(x, y)$, или $z = F(x, y)$, или $z = z(x, y)$.

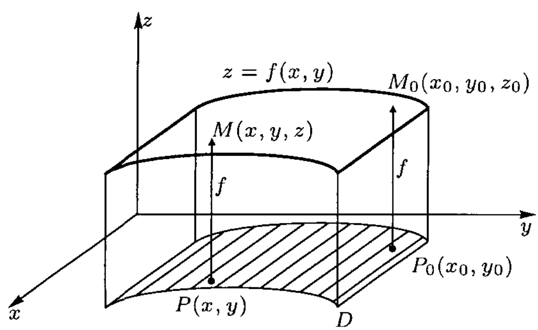


Рис. 3.1. Изображение функции двух переменных в трехмерном пространстве.

Частным значением функции $z = f(x, y)$ называют число, соответствующее какой-либо определенной паре значений аргументов.

Для обозначения частного значения функции $z = f(x, y)$ при $x = x_0$ и $y = y_0$ употребляется символ $f(x_0, y_0)$, или, что тоже самое, $z_0 = z|_{y=y_0}^{x=x_0}$. Каждая пара значений аргументов (x, y) геометрически определяет точку P на плоскости xOy, а значение функции в этой точке есть аппликата z пространственной точки $M(x, y, z)$. Геометрическое место всех точек M есть поверхность, взаимно однозначно проектирующаяся в область D на плоскости xOy (рис. 3.1). Эта поверхность служит геометрическим изображением (графиком) функции $f(x, y)$. Иногда используют запись $z = f(p)$.

Переменная величина U называется функцией независимых переменных $x_1, x_2, \ldots, x_n$, если каждой системе значений $(x_1, x_2, \ldots, x_n)$ этих переменных из данной области их изменения соответствует единственное значение величины U. Обозначение: $U = f(x_1, x_2, \ldots, x_n)$ или $U = U(x_1, x_2, \ldots, x_n)$ и т. д.

В случае функции 3 переменных пишут $U = f(x, y, z)$; каждая система (x, y, z) значений аргументов определяет некоторую точку M трехмерного пространства.

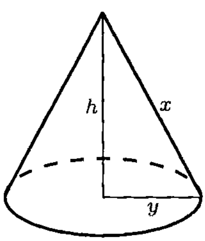


Рис. 3.2. Круговой конус.

Пример 1. Выразить объем кругового конуса V как функцию его образующей x и радиуса основания y.

Решение. Объем кругового конуса (рис. 3.2) равен одной трети произведения площади основания S на высоту h. Так как $S = \pi y^2$, $h = \sqrt{x^2 - y^2}$, то $V = \frac{1}{3}\pi y^2 \times \times \sqrt{x^2 - y^2}$. Это и есть искомая функция.

Пример 2. Найти частное значение функции $f(x, y) = \frac{2xy}{x^2 + y^2}$ в точке $P(2, -1)$.

Решение.

$$f(2, -1) = \frac{2 \cdot 2 \cdot (-1)}{2^2 + (-1)^2} = -\frac{4}{5}.$$

§ 3.2. Область определения функции

Если функция задана аналитическим выражением (формулой) без каких-либо дополнительных условий, то под ее областью определения понимают область существования аналитического выражения, т. е. совокупность всех точек, в которых данное аналитическое выражение определено и принимает только действительные значения. Область D называется замкнутой, если она включает в себя все свои границы.

Область определения функции 3 переменных представляет собой некоторую пространственную область, в частности некоторый объем.

Пример 1. Указать область определения функции, выражающей объем кругового конуса V через образующую x и радиус основания y.

Решение. Функция, найденная в примере 1 (п. 3.1), выглядит так: $V = \frac{1}{3}\pi y^2\sqrt{x^2 - y^2}$. По смыслу задачи переменные x и y могут принимать только положительные значения, и при этом всегда $x > y$, так как гипотенуза больше катета (рис. 3.2). Следовательно, область определения D задается неравенствами $x > 0$, $y > 0$, $x > y$, т. е. состоит из всех тех точек $P(x, y)$ первой четверти на плоскости xOy, которые лежат ниже биссектрисы $x = y$ (рис. 3.3). Границами области D служат прямые $y = 0$, $y = x$, которые сами в область D не входят, так что эта область незамкнутая.

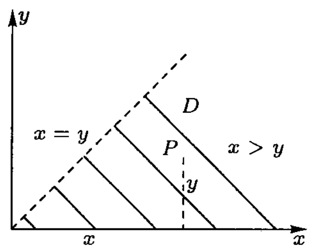


Рис. 3.3. К задаче о нахождении области определения.

Пример 2. Найти область определения функции $z = \frac{\pi}{3}y^2\sqrt{x^2 - y^2}$.

Решение. Поскольку никаких дополнительных ограничений на аргументы x и y не наложено, область определения D будет состоять из всех тех точек плоскости, для которых данное аналитическое выражение принимает действительные значения.

Для этого подкоренное выражение должно быть неотрицательным, т. е. $x^2 - y^2 \geqslant 0$, или $x^2 \geqslant y^2$.

Если оставить здесь только знак равенства, то получится уравнение границы области D: $x^2 = y^2$ или $x = \pm y$. Эта граница состоит из двух биссектрис $x = y$ и $x = -y$ координатных углов. Для внутренних точек области D должно соблюдаться неравенство $x > y$, или $|x| > |y|$. Следовательно, эти точки расположены между биссектрисами ближе к оси OX, так как $|y|$ — расстояние точки $P(x, y)$ до оси OX и оно меньше расстояния $|x|$ точки P до оси OY. Таким образом, область D состоит из всех точек 2 углов между биссектрисами $x = \pm y$, заключающими внутри себя ось OX (рис. 3.4). Область замкнутая, так как включает в себя обе свои границы.

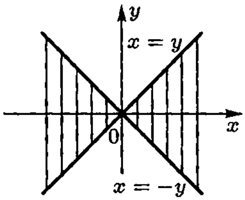


Рис. 3.4. Область определения функции $z = \frac{\pi}{3} y^2 \sqrt{x^2 - y^2}$.

Замечание. Хотя аналитические выражения функции в примерах 1 и 2 одинаковые, их области определения разные. На переменные x и y в примере 1 были наложены дополнительные условия $x > 0$, $y > 0$, $x > y$, вытекающие из их геометрического смысла.

Пример 3. Найти область определения функции $z = \dfrac{x + y}{2x - y}$.

Решение. Выражение, стоящее справа, теряет смысл при тех значениях x и y, при которых знаменатель обращается в нуль. Отсюда областью определения нашей функции является вся плоскость, из которой выброшена прямая $y = 2x$ (рис. 3.5).

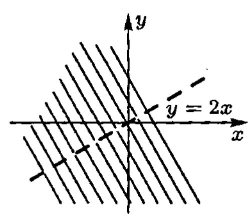


Рис. 3.5. Область определения функции $z = (x + y)/(2x - y)$.

Пример 4. Найти область определения функции $z = \sqrt{(x-1)(y+2)}$.

Решение. Для того чтобы квадратный корень имел вещественные значения, его подкоренное выражение должно быть неотрицательным. Решая неравенство $(x-1)(y+2) \geqslant 0$, находим, что либо $\begin{cases} x-1 \geqslant 0, \\ y+2 \geqslant 0, \end{cases}$ либо $\begin{cases} x-1 \leqslant 0, \\ y+2 \leqslant 0. \end{cases}$ Решением первой системы неравенств является $x \geqslant 1$, $y \geqslant -2$. Чтобы получить изображение искомой области на координатной плоскости, достаточно провести две прямые $x = 1$ и $y = -2$. Область состоит из 2 квадрантов с общей вершиной в точке $(1, -2)$ (рис. 3.6).

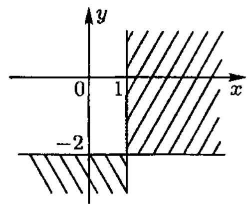


Рис. 3.6. Область определения функции $z = \sqrt{(x-1)/(y+2)}$.

Пример 5. Найти область определения функции $z = \ln(9 + 9x - y^2)$.

Решение. Логарифм определен только при положительном значении его аргумента, поэтому $9 + 9x - y^2 > 0$, или $9 + 9x > y^2$. Чтобы изобразить геометрически область D, найдем сначала ее границу $9 + 9x = y^2$, или $y^2 = 9(x+1)$. Полученное уравнение определяет параболу, вершина которой расположена в точке $O(-1, 0)$, а ось направлена в положительную сторону оси OX. Точки пересечения параболы с осью OY получаются из условия $x = 0$, откуда $y^2 = 9$, т.е. $y = \pm 3$ (рис. 3.7).

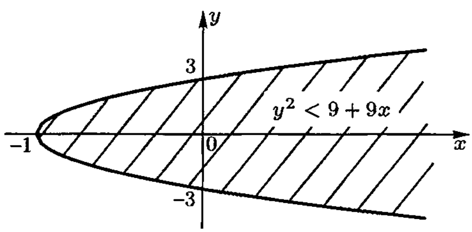


Рис. 3.7. Область определения функции $z = \ln(9 + 9x - y^2)$.

Парабола делит всю плоскость на две части — внутреннюю и внешнюю по отношению к параболе. Для точек одной из этих частей выполняется неравенство $y^2 < 9(x+1)$, а для другой $y^2 > 9(x+1)$ (на самой параболе $y^2 = 9(x+1)$). Чтобы установить, какая из этих 2 частей является областью определения данной функции, т.е. удовлетворяет условию $y^2 < 9(x+1)$, достаточно проверить это условие для какой-нибудь одной точки, не лежащей на параболе. Например, начало координат $(0,0)$ лежит внутри параболы и удовлетворяет нужному условию $0 < 9(0+1)$. Следовательно, рассматриваемая область D состоит из внутренних точек параболы. Сама парабола в область D входить не может, так как для точек параболы $9 + 9x - y^2 = 0$, и логарифм не определен.

Самостоятельная работа

1. Выразить площадь треугольника с данным периметром $2p$ как функцию длин двух его сторон x и y.

2. Выразить объем правильной шестиугольной пирамиды как функцию ее высоты x и бокового ребра y.

3. Найти $f(0{,}5;\ 1)$ и $f(2;\ -1)$, если $f(x,y) = \sqrt{xy^2 + x + 1}$.

Найти области определения функций:

4. $z = \dfrac{1}{x^2 + y^2}$. **5.** $z = \ln(x^2 + y^2 - 1)$.

6. $z = \sqrt{1 + x - y^2} + \sqrt{1 - x - y^2}$.

7. $z = \sqrt{x + y}\ \ln(y^2 - x^2)$.

8. $z = \sqrt{\sin \pi(x^2 + y^2)}$. **9.** $z = \arcsin(y/x)$.

10. $u = \sqrt{1 - \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2}}$. **11.** $z = \dfrac{1}{x + y}$.

12. $z = xy$. **13.** $z = \sqrt{3 - x^2 - y^2}$.

14. $z = \sqrt{-x} + \sqrt{y}$. **15.** $z = y/x$.

16. $z = \sqrt{x + y}$. **17.** $z = \sqrt{y \sin x}$.

18. $z = \sqrt{x - \sqrt{y}}$. **19.** $z = \ln(x^2/9 - y^2/4 - 1)$.

20. $u = \ln(xyz)$. **21.** $z = \sqrt{\dfrac{x^2 + 2x + y^2}{x^2 - 2x + y^2}}$.

22. $z = \arcsin(x/y^2) + \arcsin(1 - y)$.

23. Указать область определения функции, выражающей площадь треугольника с данным периметром через длины двух его сторон.

§ 3.3. Частные производные и полный дифференциал

3.3.1. Частные производные

Пусть (x, y) — произвольная фиксированная точка из области определения $z = f(x, y)$. Рассмотрим предел

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}.$$

Этот предел (если он существует) называется частной производной (1-го порядка) данной функции z по переменной x в точке (x, y). Производная обозначается одним из символов: $\dfrac{\partial z}{\partial x}$, $\dfrac{\partial f}{\partial x}$, z'_x, f'_x, $f'_x(x, y)$.

Аналогично, $\lim\limits_{\Delta y \to 0} \dfrac{f(x, y + \Delta y) - f(x, y)}{\Delta y} = \dfrac{\partial z}{\partial y} = \dfrac{\partial f}{\partial y} = z'_y = f'_y = f'_y(x, y)$.

Частные производные функции $z = f(x, y)$ сами представляют собой некоторые функции переменных x и y. Поэтому, если нас интересуют значения частных производных в какой-либо точке $P(x, y)$, надо сначала по общим правилам найти $\dfrac{\partial z}{\partial x}$, $\dfrac{\partial z}{\partial y}$, а затем подставить в полученные функции координаты точки P_0. Значения частных производных в точке P_0 обозначаются одним из символов: $\left.\dfrac{\partial z}{\partial x}\right|_{P_0} = \left.\dfrac{\partial f(x, y)}{\partial x}\right|_{P_0} = z'_x|_{P_0}$, или $\left.\dfrac{\partial f(x, y)}{\partial x}\right|_{(x_0, y_0)} = \left.\dfrac{\partial f(x, y)}{\partial x}\right|_{\substack{x = x_0 \\ y = y_0}} = \dfrac{\partial f(x_0, y_0)}{\partial x} = f'_x(x_0, y_0)$, и аналогично — для производных по y.

Таким образом, частная производная функции $z = f(x, y)$ по аргументу x есть производная этой функции по x при постоянном значении y. Аналогично, $\dfrac{\partial z}{\partial y}$ есть производная функции $z = f(x, y)$ по y в предположении, что x является константой.

Частные производные функции нескольких переменных определяются как производные этой функции по одному из них при условии, что остальные переменные считаются постоянными. Например, производная функции $u = f(x, y, z)$ по x определяется формулой

$$\frac{\partial u}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x}.$$

Пример 1. Найти частные производные функции $z = x^3 + y^3 - 3axy$. Вычислить их значения в точке $P_0(1, 1)$.

Решение. Считая y постоянным, находим $\frac{\partial z}{\partial x}$:

$$\frac{\partial z}{\partial x} = (x^3 + y^3 - 3axy)'_x = 3x^2 + 0 - 3ay \cdot 1 = 3x^2 - 3ay.$$

При нахождении $\frac{\partial z}{\partial y}$ фиксируется аргумент x, т. е.

$$\frac{\partial z}{\partial y} = (x^3 + y^3 - 3axy)'_y = 0 + 3y^2 + 0 - 3ax \cdot 1 = 3y^2 - 3ax.$$

Значения производных в точке $P_0(1, 1)$ следующие:

$$\left.\frac{\partial z}{\partial x}\right|_{(1,1)} = 3 - 3a, \quad \left.\frac{\partial z}{\partial y}\right|_{(1,1)} = 3 - 3a.$$

Пример 2. Найти частные производные функции $z = e^{-xy}$.

Решение. Находим частные производные, используя формулу дифференцирования сложной функции $(e^u)' = e^u u'$:

$$\frac{\partial z}{\partial x} = (e^{-xy})'_x = e^{-xy}(-xy)'_x = e^{-xy}(-y) = -ye^{-xy},$$

$$\frac{\partial z}{\partial y} = (e^{-xy})'_y = e^{-xy}(-xy)'_y = e^{-xy}(-x) = -xe^{-xy}.$$

Пример 3. Доказать, что функция $z = \ln(x^2 + xy + y^2)$ удовлетворяет уравнению

$$x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = 2.$$

Решение. Находим частные производные:

$$\frac{\partial z}{\partial x} = \frac{1}{x^2 + xy + y^2}(x^2 + xy + y^2)'_x = \frac{2x + y}{x^2 + xy + y^2};$$

$$\frac{\partial z}{\partial y} = \frac{1}{x^2+xy+y^2}(x^2+xy+y^2)'_y = \frac{x+2y}{x^2+xy+y^2}.$$

Подставляя $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ в данное уравнение

$$x\frac{2x+y}{x^2+xy+y^2} + y\frac{x+2y}{x^2+xy+y^2} = 2,$$

получим тождество.

3.3.2. Полный дифференциал

Пусть $P(x, y)$ — данная точка, а $P'(x+\Delta x, y+\Delta y)$ — близкая точка, отвечающая приращениям аргументов Δx и Δy. Полным приращением функции $z = f(x, y)$ в точке P называется разность $\Delta z = f(P') - f(P) = f(x+\Delta x, y+\Delta y) - f(x, y)$. Если приращение Δz можно представить в виде $\Delta z = A\Delta x + B\Delta y + \varepsilon$, где ε — бесконечно малая более высокого порядка по сравнению с расстоянием $\rho = \sqrt{\Delta x^2 + \Delta y^2}$ между точками P и P' (т. е. $\varepsilon/\rho \to 0$ при $\rho \to 0$), то функция $z = f(x, y)$ называется дифференцируемой в точке P, а главная линейная часть ее приращения $A\Delta x + B\Delta y = dz$ называется полным дифференциалом функции в точке P. Функция, имеющая дифференциал в каждой точке некоторой области D, называется дифференцируемой в этой области.

Если функция дифференцируема, то необходимо, чтобы выполнялись равенства:

$$A = \frac{\partial z}{\partial x}, \quad B = \frac{\partial z}{\partial y}.$$

Достаточным условием дифференцируемости является наличие непрерывных частных производных. Так как приращения независимых переменных совпадают с их дифференциалами, т. е. $\Delta x = dx$, $\Delta y = dy$, то дифференциал функции $z = f(x, y)$ вычисляется по формуле

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy. \tag{3.1}$$

Последняя формула остается справедливой также и в том случае, когда x и y в свою очередь являются функциями каких-либо других аргументов (свойство инвариантности полного дифференциала).

Пример. Найти полный дифференциал функции $z = x^2y - y^2x$.

Решение. Находим частные производные:

$$\frac{\partial z}{\partial x} = (x^2y - y^2x)'_x = 2xy - y^2, \quad \frac{\partial z}{\partial y} = (x^2y - y^2x)'_y = x^2 - 2xy.$$

По формуле (3.1) имеем

$$dz = (2xy - y^2)\,dx + (x^2 - 2xy)\,dy.$$

3.3.3. Применение дифференциала к приближенным вычислениям

Приращение функции Δz и ее полный дифференциал dz связаны равенством $\Delta z = dz + \varepsilon$, где ε — бесконечно малая более высокого порядка малости по сравнению с $\rho = \sqrt{\Delta x^2 + \Delta y^2}$; при достаточно малых приращениях аргументов можно величиной ε пренебречь и считать $\Delta z \approx dz$. Это приводит к приближенному равенству $\Delta z \approx dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy$, или (подробно):

$$f(x+\Delta x, y+\Delta y) - f(x,y) \approx df(x,y) = \frac{\partial f(x,y)}{\partial x}\Delta x + \frac{\partial f(x,y)}{\partial y}\Delta y. \tag{3.2}$$

Этой формулой можно пользоваться для приближенного вычисления значения $f(x + \Delta x, y + \Delta y)$ по известным значениям функции $z = f(x,y)$ и ее частных производных в данной точке $P(x,y)$.

Пример 1. Вычислить приближенно число $(1{,}04)^{2{,}03}$.

Решение. Рассмотрим функцию $f(x,y) = x^y$. Данное число есть приращенное значение этой функции в точке $P_0(1,2)$ при $\Delta x = 0{,}04$, $\Delta y = 0{,}03$. Дифференциал данной функции

$$df = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y = y \cdot x^{y-1} \cdot \Delta x + x^y \cdot \ln x \cdot \Delta y.$$

Его значение в точке $P_0(1,2)$ при данных приращениях

$$df|_{(1,2)} = 2 \cdot 1^{2-1} \cdot 0{,}04 + 1^2 \cdot \ln 1 \cdot 0{,}03 = 0{,}08,$$

поэтому по формуле (3.2) имеем

$$f(1{,}04,\ 2{,}03) \approx f(1,2) + df|_{(1,2)} = 1^2 + 0{,}08 = 1{,}08,$$

т. е. $(1{,}04)^{2{,}03} \approx 1{,}08$.

Пример 2. Вычислить приближенно $\sqrt{(4{,}05)^2 + (2{,}93)^2}$.

Решение. Рассмотрим функцию $f(x,y) = \sqrt{x^2+y^2}$. Данное число есть приращенное значение этой функции в точке $P_0(4{,}3)$ при $\Delta x = 0{,}05$, $\Delta y = -0{,}07$. Дифференциал данной функции

$$df = \frac{2x}{2\sqrt{x^2+y^2}}\,\Delta x + \frac{2y}{2\sqrt{x^2+y^2}}\,\Delta y.$$

Его значение в точке $P_0(4,3)$ при данных приращениях

$$df|_{(4,3)} = \frac{4}{\sqrt{4^2+3^2}} \cdot 0{,}05 + \frac{3}{\sqrt{4^2+3^2}} \cdot (-0{,}07) = -0{,}002,$$

поэтому по формуле (3.2) имеем

$$f(4{,}05,\ 2{,}93) \approx f(4,3) + (df)|_{(4,3)} = 5 + (-0{,}002) = 4{,}998,$$

так как $f(4,3) = \sqrt{4^2+3^2} = 5$.

Итак, $\sqrt{(4{,}05)^2 + (2{,}93)^2} \approx 4{,}998$.

Пример 3. Вычислить приближенно $\operatorname{arctg}(1{,}97/1{,}02 - 1)$.

Решение. Данное число есть приращенное значение этой функции в точке $P_0(2,1)$ при $\Delta x = -0{,}03$, $\Delta y = 0{,}02$. Найдем частные производные функции $f(x,y)$ и их значения в точке $P_0(2,1)$:

$$\left.\frac{\partial f}{\partial x}\right|_{P_0} = \left.\frac{1 \cdot 1/y}{1+(x/y-1)^2}\right|_{P_0} = \frac{1}{2},$$

$$\left.\frac{\partial f}{\partial y}\right|_{P_0} = \left.\frac{(-x/y^2)}{1+(x/y-1)^2}\right|_{P_0} = -1.$$

Значение функции в точке $P_0(2,1)$ равно

$$f(2,1) = \operatorname{arctg}(2/1-1) = \operatorname{arctg} 1 = \pi/4.$$

Подставляя найденные значения функции и частных производных в формулу (3.2), получим

$$\operatorname{arctg}(1{,}97/1{,}02 - 1) \approx \pi/4 + 1/2 \cdot (-0{,}03) + (-1) \cdot 0{,}02 \approx 0{,}75.$$

Самостоятельная работа

Найти частные производные следующих функций:

1. $z = 2x^2y + 3xy^2 + x^3$. **2.** $z = xy - \dfrac{x}{y}$.

3. $z = (x^2 + y^2)^3$. **4.** $z = x - 3\sin y$. **5.** $z = \dfrac{x}{\sqrt{x+y}}$.

6. $z = \operatorname{tg}\dfrac{y}{x}$. **7.** $z = \ln(x^2 + y)$. **8.** $z = \ln\sqrt{x^2 + y^2}$.

9. $z = \sin\dfrac{x}{y}\cos\dfrac{y}{x}$. **10.** $z = y\sqrt{x} + \dfrac{x}{\sqrt{y}}$. **11.** $z = \left(\dfrac{y}{x}\right)^x$.

12. $z = (\sin x)^{\cos y}$. **13.** $z = \ln\operatorname{tg}\dfrac{y}{x}$.

14. $u = \sqrt{1 - \left(\dfrac{x+y}{xy}\right)^2} + \arcsin\dfrac{x+y}{xy}$. **15.** $z = \dfrac{x}{y}e^{xy}$.

16. Показать, что функция $z = \sqrt{x} \cdot \sin\dfrac{y}{x}$ удовлетворяет уравнению $2x\dfrac{\partial z}{\partial x} + 2y\dfrac{\partial z}{\partial y} = z$.

Найти полные дифференциалы функций:

17. $z = \sqrt{x^2 + y^2}$. **18.** $z = \sin^2 y + \cos^2 x$.

19. $z = \dfrac{x+y}{xy}$. **20.** $z = e^{xy}$. **21.** $z = \dfrac{x^2 - y^2}{x^2 + y^2}$.

22. $z = \ln(x^2 + y^2)$. **23.** $z = xy\ln(xy)$. **24.** $z = \operatorname{arctg}\dfrac{xy}{x+y}$.

25. $z = x\sin y + y\sin x$. **26.** $z = x^y + y^x$.

Вычислить приближенно следующие числа:

27. $(1{,}03)^{2{,}02}$. **28.** $(0{,}97)^{2{,}02}$. **29.** $\sqrt{(6{,}03)^2 + (8{,}04)^2}$.

30. $\cos 31^\circ \cdot \sin 28^\circ$. **31.** $\sqrt{(1{,}02)^3 + (1{,}97)^3}$.

32. $\sin 29^\circ \cdot \operatorname{tg} 46^\circ$. **33.** $\sqrt{7{,}01 + (3{,}02)^2}$.

34. $z = \arcsin\left(\frac{5{,}02}{1{,}95} - 2\right)$.

§ 3.4. Дифференцирование сложных и неявных функций двух переменных

3.4.1. Сложная функция

Пусть $z = f(x, y)$, где $x = x(t)$, $y = y(t)$ (при этом считается, что $f(x, y)$, $x(t)$, $y(t)$ — дифференцируемые функции). Тогда сложная функция $z = z(t) = f(x(t), y(t))$ также дифференцируема, а ее производная находится по формуле

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}. \tag{3.3}$$

Если $z = f(x, y)$, а функция y зависит от аргумента x: $y = y(x)$, тогда в формуле (3) надо положить $t = x$ и $\frac{dx}{dt} = 1$, и полная производная $\frac{\partial z}{\partial x}$ вычисляется по формуле

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx}. \tag{3.4}$$

Если же x и y зависят от нескольких переменных, например $x = x(u, v)$, $y = y(u, v)$, то формулы частных производных сложной функции $z = z(u, v) = f(x(u, v), y(u, v))$ имеют вид

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}, \quad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}. \tag{3.5}$$

Пример 1. Найти $\frac{dz}{dt}$, если $z = x^2 + xy + y^2$, где $x = t^2$, $y = t^3$.

Решение. По формуле (3.3) имеем

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} = (2x+y)2t + (x+2y)3t^2 =$$
$$= (2t^2+t^3)2t + (t^2+2t^3)3t^2 = 4t^3+5t^4+6t^5.$$

Тот же результат можно получить и иным путем: сначала выразить явно через t, а затем найти производную полученной функции:

$$z = z(t) = (t^2)^2+t^2\cdot t^3+(t^3)^2 = t^4+t^3+t^6, \qquad \frac{dx}{dt} = 4t^3+5t^4+6t^5.$$

Пример 2. Найти $\frac{\partial z}{\partial x}$ и $\frac{dz}{dx}$, если $z = \arcsin\frac{x}{y}$, а $y = \sqrt{1+x^2}$.

Решение. Имеем

$$\frac{\partial z}{\partial x} = \left(\arcsin\frac{x}{y}\right)'_x = \frac{1}{\sqrt{1-\left(\frac{x}{y}\right)^2}}\cdot\left(\frac{x}{y}\right)'_x = \frac{1}{\sqrt{1-\frac{x^2}{y^2}}}\cdot\frac{1}{y} =$$
$$= \frac{1}{\sqrt{y^2-x^2}}.$$

По формуле (3.4)

$$\frac{dz}{dx} = \frac{\partial z}{\partial x}+\frac{\partial z}{\partial y}\cdot\frac{dy}{dx} = \frac{1}{\sqrt{y^2-x^2}}+\frac{1}{\sqrt{1-\frac{x^2}{y^2}}}\cdot\left(-\frac{x}{y^2}\right)\cdot\frac{1}{2\sqrt{1+x^2}}\cdot 2x =$$
$$= \frac{1}{\sqrt{y^2-x^2}} - \frac{x^2}{y\sqrt{y^2-x^2}\sqrt{1+x^2}}.$$

Поскольку $y = \sqrt{1+x^2}$, то $y^2 - x^2 = 1$ и

$$\frac{dz}{dx} = 1 - \frac{x^2}{\sqrt{1+x^2}\sqrt{1+x^2}} = 1 - \frac{x^2}{1+x^2} = \frac{1}{1+x^2}.$$

Пример 3. Найти частные производные сложной функции $z = x^2\ln y$, $x = \frac{u}{v}$, $y = uv$.

Решение. Частные производные данных функций имеют вид

$$\frac{\partial z}{\partial x} = 2x \cdot \ln y, \quad \frac{\partial z}{\partial y} = \frac{x^2}{y}, \quad \frac{\partial x}{\partial u} = \frac{1}{v},$$

$$\frac{\partial x}{\partial v} = -\frac{u}{v^2}, \quad \frac{\partial y}{\partial u} = v, \quad \frac{\partial y}{\partial v} = u.$$

По формуле (3.5) находим производные от сложной функции:

$$\frac{\partial z}{\partial u} = 2x \cdot \ln y \frac{1}{v} + \frac{x^2}{y} v = 2\frac{u}{v} \cdot \ln(uv) \cdot \frac{1}{v} + \frac{u^2}{v^2} \cdot \frac{1}{uv} \cdot v = \frac{u}{v^2}(1 + 2\ln(uv)),$$

$$\frac{\partial z}{\partial v} = 2x \cdot \ln y \cdot \left(-\frac{u}{v^2}\right) + \frac{x^2}{y} u = -2\frac{u}{v} \cdot \ln(uv) \cdot \frac{u}{v^2} + \frac{u^2}{v^2} \cdot \frac{1}{uv} \cdot u =$$

$$= \frac{u^2}{v^3}(1 - 2\ln(uv)).$$

3.4.2. Неявная функция одной переменной

Функция y называется неявной функцией от x, если она задана уравнением

$$F(x, y) = 0, \tag{3.6}$$

не разрешенным относительно y. Это означает, что при каждом значении $x = x_0$, при котором неявная функция определена, она принимает такое значение y_0, для которого $F(x_0, y_0) = 0$.

Если $F(x, y)$ — дифференцируемая функция переменных x и y, и $F'_y(x, y) \neq 0$, то определяемая уравнением (3.6) неявная функция имеет производную, которая вычисляется по формуле $\dfrac{dy}{dx} = -\dfrac{F'_x(x, y)}{F'_y(x, y)}$ (альтернативный способ нахождения производной неявной функции был приведен в п. 2.2).

Пример 1. Функция $y(x)$ задана уравнением $1 + xy - \ln(e^{xy} + e^{-xy}) = 0$. Найти $\dfrac{dy}{dx}$.

Решение. В данном случае $F(x, y) = 1 + xy - \ln(e^{xy} + e^{-xy})$, поэтому

$$F'_x = y - \frac{1}{e^{xy} + e^{-xy}} \cdot (e^{xy} - e^{-xy}) \cdot y = \frac{2ye^{-xy}}{e^{xy} + e^{-xy}},$$

$$F'_y = x - \frac{1}{e^{xy} + e^{-xy}} \cdot \left(e^{xy} - e^{-xy}\right) \cdot x = \frac{2xe^{-xy}}{e^{xy} + e^{-xy}},$$

и $\dfrac{dy}{dx} = -\dfrac{F'_x}{F'_y} = -\dfrac{y}{x}$.

3.4.3. Неявная функция двух переменных

Функция z называется неявной функцией от x и y, если она задана уравнением

$$F(x, y, z) = 0, \tag{3.7}$$

не разрешенным относительно z. Это значит, что при любых $x = x_0$, $y = y_0$ из области определения функции она принимает такое значение z_0 , для которого $F(x_0, y_0, z_0) = 0$. Если $F(x, y, z)$ — дифференцируемая функция трех переменных x, y, z и $F'_z(x, y, z) \neq 0$, то определяемая уравнением (3.7) функция $z = z(x, y)$ также дифференцируема и ее частные производные находятся по формулам

$$\frac{\partial z}{\partial x} = -\frac{F'_x(x, y, z)}{F'_z(x, y, z)}, \quad \frac{\partial z}{\partial y} = -\frac{F'_y(x, y, z)}{F'_z(x, y, z)}. \tag{3.8}$$

Пример 1. Найти $\dfrac{\partial z}{\partial x}$, $\dfrac{\partial z}{\partial y}$ и полный дифференциал функции $z = z(x, y)$, заданной уравнением $z^3 - 3xyz = 10$.

Решение. Здесь $F(x, y, z) = z^3 - 3xyz - 10$ и

$$F'_x = -2yz, \quad F'_y = -3xz, \quad F'_z = 3z^2 - 3xy.$$

По формуле (3.8) найдем частные производные:

$$\frac{\partial z}{\partial x} = \frac{yz}{z^2 - xy}, \quad \frac{\partial z}{\partial y} = \frac{xz}{z^2 - xy},$$

следовательно,

$$dz = \frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy = \frac{yz}{z^2 - xy}\, dx + \frac{xz}{z^2 - xy}\, dy.$$

Самостоятельная работа

Найти dz/dt:

1. $z = e^{2x^2-2y^2}$, $x = \cos t$, $y = \sin t$.

2. $z = \ln\sin\dfrac{x}{\sqrt{y}}$, $x = 3t^2$, $y = \sqrt{t^2+1}$.

3. $z = x^y$, $x = \sin t$, $y = 2t$.

4. $z = e^{xy}\cdot\ln(x+y)$, $x = t^3$, $y = 1 - t^3$.

5. $z = \cos(2x + 4x^2 - y)$, $x = \dfrac{1}{t}$, $y = \dfrac{\sqrt{t}}{\ln t}$.

Найти dz/dx:

6. $z = \operatorname{arctg}\dfrac{x+1}{y}$, $y = e^{(1+x)^2}$.

7. $z = \ln(e^x + e^y)$, $y = x^2$.

Найти $\partial z/\partial u$ и $\partial z/\partial v$:

8. $z = e^{x+y}$, $x = 2u - v$, $y = 2u + v$.

9. $z = x\sin y + y\cos x$, $x = \dfrac{u}{v}$, $y = uv$.

10. $z = xy\operatorname{arctg}(xy)$, $x = u^2 + 1$, $y = v^3$.

11. $z = x^y + y^x$, $x = u^2 + v^2$, $y = u^2 - v^2$.

12. $z = \arccos\dfrac{x}{y}$, $x = u\sin v$, $y = \cos v$.

13. $z = \sin^2(x+y)$, $x = \dfrac{3u}{v}$, $y = v + 1$.

Найти dy/dx:

14. $x^2 - 3xy^2 = 6$. **15.** $\dfrac{x^2}{9} + \dfrac{y^2}{4} = 1$.

16. $\ln\sqrt{x^2+y^2} = \operatorname{arctg}\dfrac{y}{x}$.

17. $e^x\sin y - e^y\cos x = 0$. **18.** $y = 1 + y^x$.

19. $x^2 - 1 + \cos xy = 0$. **20.** $x + y = e^{x-y}$.

21. $x^3 + 4y^3 - 3x^2y = 2.$ **22.** $x^2 + 2xy.$

Найти $\partial z/\partial x$ и $\partial z/\partial y$ неявных функций:

23. $x^2 + y^2 + z^2 - 2xz = 1.$

24. $5(x^2 + y^2 + z^2) - 2(xy + yz + xz) = 72.$

25. $3x^2z + z^3 = 2xy.$ **26.** $xyz = x + y = z.$

27. $\cos^2 x + \cos^2 y + \cos^2 z = 1.$

28. $z = ye^{x/y}.$ **29.** $z\ln(x+z) - \dfrac{xy}{z} = 0.$

30. $x + y + z = e^{-(x+y+z)}.$ **31.** $\dfrac{x}{z} = \ln\dfrac{x}{y}.$

§ 3.5. Производные и дифференциалы высших порядков

Частными производными второго порядка функции $z = f(x, y)$ называются частные производные от ее первых производных $\dfrac{\partial z}{\partial x}$, $\dfrac{\partial z}{\partial y}$, т. е.

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right), \quad \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right),$$
$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right), \quad \frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right).$$

Частные производные второго порядка обозначаются также символами: $f''_{xx}(x, y)$, $f''_{xy}(x, y)$, $f''_{yx}(x, y)$, $f''_{yy}(x, y)$. Частные производные второго порядка вида $\dfrac{\partial^2 z}{\partial x \partial y}$ и $\dfrac{\partial^2 z}{\partial y \partial x}$ называются смешанными производными. Они не зависят от порядка дифференцирования, т. е. равны между собой, если выполняется условие следующей теоремы (Шварца): Если функция $z = f(x, y)$ определена в некоторой области D и ее частные производные $\dfrac{\partial z}{\partial x}$, $\dfrac{\partial z}{\partial y}$, $\dfrac{\partial^2 z}{\partial x \partial y}$, $\dfrac{\partial^2 z}{\partial y \partial x}$ определены и непрерывны в этой области D, то значения смешанной производной не зависят от порядка дифференцирования, т. е. $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$.

Аналогично определяются и обозначаются частные производные более высоких порядков.

Дифференциалы второго, третьего и более высоких порядков от функции $z = f(x, y)$ определяются формулами $d^2z = d(dz)$, $d^3z = d(d^2z)$ и выражаются через частные производные следующим образом:

$$d^2z = \frac{\partial^2 z}{\partial x^2}(dx)^2 + 2\frac{\partial^2 z}{\partial x \partial y}dxdy + \frac{\partial^2 z}{\partial y^2}(dy)^2, \tag{3.9}$$

$$d^3z = \frac{\partial^3 z}{\partial x^3}(dx)^3 + 3\frac{\partial^3 z}{\partial x^2 \partial y}(dx)^2dy + 3\frac{\partial^3 z}{\partial x \partial y^2}dx(dy)^2 + \frac{\partial^2 z}{\partial y^3}(dy)^3 \tag{3.10}$$

и т. д.

При этом приращении dx и dy рассматриваются как постоянные и остаются одними и теми же при переходе от дифференциала первого порядка к дифференциалу более высокого порядка.

Пример 1. Найти частные производные второго порядка функции $z = \ln(x^2 + y^2)$. Убедиться, что $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$.

Решение. Находим сначала первые производные:

$$\frac{\partial z}{\partial x} = (\ln(x^2 + y^2))'_x = \frac{1}{x^2 + y^2}(x^2 + y^2)'_x = \frac{2x}{x^2 + y^2},$$
$$\frac{\partial z}{\partial y} = (\ln(x^2 + y^2))'_y = \frac{1}{x^2 + y^2}(x^2 + y^2)'_y = \frac{2y}{x^2 + y^2}.$$

Дифференцируя каждую из полученных функций по x и y, получим вторые частные производные:

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \left(\frac{2x}{x^2 + y^2}\right)'_x = 2\frac{x^2 + y^2 - x \cdot 2x}{(x^2 + y^2)^2} = \frac{2(y^2 - x^2)}{(x^2 + y^2)^2},$$
$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \left(\frac{2x}{x^2 + y^2}\right)'_y = -\frac{2x}{(x^2 + y^2)^2} \cdot 2y = -\frac{4xy}{(x^2 + y^2)^2},$$
$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \left(\frac{2y}{x^2 + y^2}\right)'_x = -\frac{2y}{(x^2 + y^2)^2} \cdot 2x = -\frac{4xy}{(x^2 + y^2)^2},$$

$$\frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) = \left(\frac{2y}{x^2+y^2}\right)'_y = 2\frac{x^2+y^2-y\cdot 2y}{(x^2+y^2)^2} = \frac{2(x^2-y^2)}{(x^2+y^2)^2}.$$

Отсюда видно, что смешанные частные производные $\frac{\partial^2 z}{\partial x \partial y}$ и $\frac{\partial^2 z}{\partial y \partial x}$ равны. Интересно отметить, что для данной функции производные $\frac{\partial^2 z}{\partial x^2}$ и $\frac{\partial^2 z}{\partial y^2}$ отличаются только знаком, т. е. $\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0$. Это уравнение называется уравнением Лапласа. Мы показали, что функция $z = \ln(x^2+y^2)$ есть одно из решений уравнения Лапласа.

Пример 2. Найти дифференциал второго порядка функции $z = 3x^2y - 2xy + y^2 - 1$.

Решение. Находим частные производные первого и второго порядков:

$$\frac{\partial z}{\partial x} = 6xy - 2y, \quad \frac{\partial z}{\partial y} = 3x^2 - 2x + 2y,$$

$$\frac{\partial^2 z}{\partial x^2} = 6y, \quad \frac{\partial^2 z}{\partial x \partial y} = 6x - 2, \quad \frac{\partial^2 z}{\partial y^2} = 2.$$

Следовательно, по формуле (3.9)

$$d^2 z = 6y\,(dx)^2 + 2(6x-2)\,dxdy + 2(dy)^2.$$

Тот же результат можно получить и другим способом. Найдем сначала первый дифференциал данной функции:

$$dz = 3\,d(x^2y) - 2\,d(xy) + (y^2-1) =$$
$$= 3x^2\,dy - 6xy\,dx - 2x\,dy - 2y\,dx + 2y\,dy =$$
$$= (6xy - 2y)\,dx + (3x^2 - 2x + 2y)\,dy$$

Беря дифференциал от dz, найдем d^2z, при этом dx и dy надо считать постоянным:

$$d^2z = d[(6xy-2y)\,dx + (3x^2-2x+2y)\,dy] =$$
$$= (6x\,dy + 6y\,dx - 2\,dy)\,dx + 6x\,dx - 2\,dx + 2\,dy)\,dy =$$
$$= 6y(dx)^2 + (12x-4)dxdy + 2(dy)^2.$$

Самостоятельная работа

Найти вторые частные производные следующих функций:

1. $z = x^2y - xy^2 + 3.$ **2.** $z = xy - \dfrac{y}{x}.$

3. $z = (x^2 + y^2)^3.$ **4.** $z = x - 3\sin y.$

5. $z = \ln(x^2 + y).$ **6.** $z = \dfrac{2x + 3y}{x - y}.$

7. $z = \sqrt{2xy + y^2}.$ **8.** $z = \operatorname{arctg}\dfrac{x + y}{1 - xy}.$

9. $z = (\sin x)^{\cos y}.$ **10.** $z = \ln\sqrt{x^2 + y^2}.$

11. $z=\left(\dfrac{y}{x}\right)^x.$ **12.** $z = \ln\operatorname{tg}\dfrac{y}{x}.$ **13.** $z = \operatorname{arctg}\sqrt{xy}.$

Найти d^2z следующих функций:

14. $z = x + xy.$ **15.** $z = e^{x^2+y^2}.$ **16.** $z = e^x\cos y.$

17. $z = x\sin^2 y.$ **18.** $z = x^3 + y^3 - 3xy(x - y).$

19. $z = x\ln\dfrac{y}{x}.$ **20.** $z = \operatorname{arctg}\sqrt{xy}.$

Найти d^3z следующих функций:

21. $z = x^3 + y.$ **22.** $z = e^y\cos x.$

ГЛАВА 4
НЕОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

§ 4.1. Простейшие методы интегрирования

4.1.1. Понятие неопределенного интеграла

При изучении дифференциального исчисления рассматривалась задача нахождения производной или дифференциала по заданной функции $y = F(x)$, т. е. необходимо было найти $f(x) = F'(x)$ или $dF(x) = F'(x)\,dx = f(x)\,dx$. Можно поставить обратную задачу: восстановить продифференцированную функцию, т. е., зная производную $f(x)$ (или дифференциал $f(x)\,dx$), найти такую функцию $F(x)$, чтобы $F'(x) = f(x)$. Например, пусть известна скорость перемещения точки $v = v(t)$, а надо найти закон ее перемещения $S = S(t)$, причем $S'(t) = v(t)$. Эта задача оказывается значительно более трудной, чем задача дифференцирования. Для решения подобных задач вводятся новые понятия и действия.

Определение 1. Дифференцируемая функция $F(x)$ называется первообразной для функции $f(x)$ на (a, b), если $F'(x) = f(x)$ на (a, b).

Например, для $f(x) = x^2$ первообразная $F(x) = x^3/3$, так как $F'(x) = (x^3/3)' = x^2$; для $f(x) = \cos x$ первообразной будет $F(x) = \sin x$, потому что $F'(x) = (\sin x)' = \cos x$, что совпадает с $f(x)$.

Всегда ли существует первообразная для заданной функции $f(x)$? Ответ положителен, если эта функция непрерывна на (a, b). Кроме того, первообразных бесчисленное множество и отличаются они друг от друга только постоянным слагаемым. Действительно, $\sin x + 2$, $\sin x - 2$, $\sin x + c$, — все эти функции будут первообразными для $\cos x$ (производная от постоянной величины равна 0).

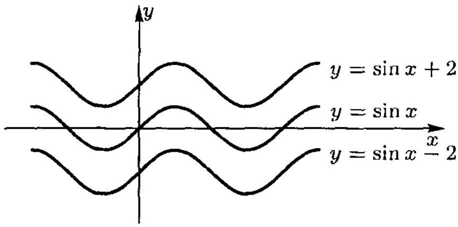


Рис. 4.1. Пример семейства интегральных кривых.

Определение 2. Выражение $F(x) + C$, где C — произвольная постоянная величина, определяющее множество первообразных для функции $f(x)$, называется неопределенным интегралом и обозначается символом $\int f(x)\,dx$, т. е. $\int f(x)\,dx = F(x) + C$, где знак $\int$ — знак неопределенного интеграла, $f(x)$ — называется подынтегральной функцией, $f(x)\,dx$ — подынтегральным выражением, x — переменной интегрирования.

Определение 3. Операция нахождения первообразной по заданной производной или дифференциалу называется интегрированием этой функции.

Интегрирование — действие, обратное дифференцированию, его можно проверить дифференцированием, причем дифференцирование однозначно, а интегрирование дает ответ с точностью до постоянной. Придавая постоянной величине C конкретные значения C_1, C_2, C_3, получим различные функции:

$$y_1(x) = F(x) + C_1, \quad y_2(x) = F(x) + C_2, \quad y_3(x) = F(x) + C_3,$$

каждая из которых задает на координатной плоскости кривую, называемую интегральной. Все графики интегральных кривых сдвинуты относительно друг друга вдоль оси OY. Следовательно, геометрически неопределенный интеграл представляет собой семейство интегральных кривых.

Итак, введены новые понятия (первообразной и неопределенного интеграла) и новое действие (интегрирование), но как все-таки находить первообразную? Чтобы легко было ответить на этот вопрос, надо в первую очередь составить и выучить наизусть таблицу неопределенных интегралов от основных элементарных функций. Она получается в результате обращения соответствующих формул дифференцирования. Например, если $(\sin x)' = \cos x$, то $\int \cos x\, dx = \sin x + C$.

Обычно в таблицу включаются и некоторые интегралы, полученные после применения простейших методов интегрирования. Эти формулы будут помечены в таблице символом «*» и доказаны при дальнейшем изложении материала.

Таблица 4.1. Таблица основных неопределенных интегралов

1. $\int x^n\, dx = \frac{x^{n+1}}{n+1} + C,$ $(n \neq -1)$	9.* $\int \operatorname{tg} x\, dx = -\ln\|\cos x\| + C$
2. $\int \frac{dx}{x} = \ln\|x\| + C$	10.* $\int \operatorname{ctg} x\, dx = \ln\|\sin x\| + C$
3. $\int e^x\, dx = e^x + C$	11. $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C$
4. $\int a^x\, dx = \frac{a^x}{\ln a} + C,$ $a \neq 1, a > 0$	12.* $\int \frac{dx}{\sqrt{a^2-x^2}} = \arcsin \frac{x}{a} + C$
5. $\int \sin x\, dx = -\cos x + C$	13. $\int \frac{dx}{1+x^2} = \operatorname{arctg} x + C$
6. $\int \cos x\, dx = \sin x + C$	14.* $\int \frac{dx}{a^2+x^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C$
7. $\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C$	15.* $\int \frac{dx}{a^2-x^2} = \frac{1}{2a} \ln \left\|\frac{a+x}{a-x}\right\| + C$
8. $\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C$	16.* $\int \frac{dx}{\sqrt{x^2+m}} =$ $= \ln\|x + \sqrt{x^2+m}\| + C$

Формула (11) может иметь вид $\int \frac{dx}{\sqrt{1-x^2}} = -\arccos x + C$, так как $\arcsin x + \arccos x = \pi/2$. Аналогичное замечание и по поводу формулы (13):

$$\int \frac{dx}{1+x^2} = -\operatorname{arcctg} x + C.$$

4.1.2. Свойства неопределенных интегралов

Рассмотрим простейшие свойства неопределенного интеграла, которые позволят интегрировать не только основные элементарные функции.

1. Производная от неопределенного интеграла равна подынтегральной функции:

$$\left(\int f(x)\,dx\right)' = (F(x)+C)' = F'(x) + C' = f(x).$$

2. Дифференциал от неопределенного интеграла равен подынтегральному выражению:

$$d\left(\int f(x)\,dx\right) = d\,(F(x)+C) = F'(x)\,dx = f(x)\,dx.$$

3. Неопределенный интеграл от дифференциала функции равен этой функции, сложенной с произвольной постоянной:

$$\int d(F(x)) = \int F'(x)\,dx = \int f(x)\,dx = F(x) + C$$

Пример 1. $\int dx = x + C$.

Пример 2. $\int d(e^x) = e^x + C$.

4. Постоянный множитель можно выносить за знак интеграла:

$$\int kf(x)\,dx = k\int f(x)\,dx = k\,(F(x)+C)$$

Пример 3. $\int \frac{5dx}{x} = 5\int \frac{dx}{x} = 5\ln|x| + C$.

5. Интеграл от алгебраической суммы двух функций равен алгебраической сумме интегралов от этих функций:

$$\int (f_1(x) \pm f_2(x))\,dx = \int f_1(x)\,dx \pm \int f_2(x)\,dx$$

Пример 4.

$$\int \left(3x^2 + \frac{\sin x}{2}\right) dx = \int 3x^2\,dx + \int \frac{\sin x}{2}\,dx = 3\frac{x^3}{3} - \frac{1}{2}\cos x + C.$$

Формула интегрирования остается справедливой, если переменная интегрирования является функцией: если $\int f(x)\,dx = F(x) + C$, то $\int f(u)\,du = F(u) + C$, где $u = \varphi(x)$ — произвольная функция, имеющая непрерывную производную. Это свойство называется инвариантностью.

Пример 5. $\int e^x\,dx = e^x + C$, поэтому

$$\int e^{x^2}\,d(x^2) = e^{x^2} + C, \qquad \int e^{\operatorname{tg} x}\,d(\operatorname{tg} x) = e^{\operatorname{tg} x} + C.$$

Сравнить с $\int e^x\,d(e^x) = \dfrac{e^{2x}}{2} + C$.

В интегральном исчислении нет универсального способа интегрирования. Применение различных методов приводит данный интеграл к табличному, который надо узнать с учетом свойства инвариантности. Полезно прочитать табличный интеграл, обращая внимание на то, где находится переменная интегрирования (в показателе степени, в знаменателе, под знаком синуса и т. д.).

4.1.3. Непосредственное интегрирование

Этот метод заключается в прямом использовании табличных интегралов и свойств.

Примеры.

1. $$\int \left(6x^5 + \frac{\sin x}{2} - 10\right) dx = 6\int x^5\,dx + \frac{1}{2}\int \sin x\,dx -$$

$$-10\int dx = 6\frac{x^6}{6} - \frac{1}{2}\cos x - 10x + C = x^6 - \frac{1}{2}\cos x - 10x + C.$$

2. $\displaystyle\int\left(\frac{3}{\cos^2 x}-\frac{1}{4x}+e^2\right)dx = 3\operatorname{tg} x-\frac{1}{4}\ln|x|+e^2x+C.$

3. $\displaystyle\int(\sin^2 x+1)\,d(\sin x) =$

$$=\int\sin^2 x d(\sin x)+\int d(\sin x)=\frac{\sin^3 x}{3}+\sin x+C.$$

4. $\displaystyle\int\frac{d(\operatorname{tg} x)}{\operatorname{tg} x}=\ln|\operatorname{tg} x|+C.$

5. $\displaystyle\int\frac{d(\ln x)}{\sin^2\ln x}=-\operatorname{ctg}\ln x+C.$

4.1.4. Метод разложения

Этот метод заключается в разложении подынтегральной функции в линейную комбинацию более простых функций с использованием известных формул.

Примеры.

1. $\displaystyle\int\left(\frac{x^4-x^2+x+1}{x^2}\right)dx=\int\left(x^2-1+\frac{1}{x}+\frac{1}{x^2}\right)dx=$

$$=\frac{x^3}{3}-x+\ln|x|-\frac{1}{x}+C.$$

2. $\displaystyle\int\frac{dx}{\sin^2 x\cdot\cos^2 x}=\int\frac{\sin^2 x+\cos^2 x}{\sin^2 x\cos^2 x}\,dx=$

$$=\int\left(\frac{1}{\cos^2 x}+\frac{1}{\sin^2 x}\right)dx=\operatorname{tg} x-\operatorname{ctg} x+C.$$

3. $\displaystyle\int\sin^2\frac{x}{2}\,dx=\int\frac{1-\cos x}{2}\,dx=\frac{1}{2}(x-\sin x)+C$

4. $\displaystyle\int\frac{2^{x-1}-5^{x+2}}{10^x}\,dx=\int\left(\frac{1}{2}\left(\frac{1}{5}\right)^x-25\left(\frac{1}{2}\right)^x\right)dx=$

$$=\frac{\frac{1}{2}\left(\frac{1}{5}\right)^x}{\ln\frac{1}{5}}-\frac{25\left(\frac{1}{2}\right)^x}{\ln\frac{1}{2}}+C=\frac{25}{2^x\ln 2}-\frac{1}{2\cdot 5^x\ln 5}+C.$$

4.1.5. Метод подведения под знак дифференциала

Для приведения данного интеграла к табличному бывает удобно сделать преобразования дифференциала.

а) Подведение под знак дифференциала линейной функции $d(ax+b) = a\,dx$, отсюда $dx = \frac{1}{a} d\,(ax+b)$, $a \neq 0$, в частности $dx = d(x+b)$, $dx = \frac{1}{a}\,d(ax)$, т. е. дифференциал не меняется, если к переменной прибавить или отнять постоянную величину, а если переменная увеличивается в несколько раз, то дифференциал умножается на обратную величину.

Примеры с решениями.

1. $\int (x+1)^{17}\,dx = \int (x+1)^{17}\,d(x+1) = \frac{(x+1)^{18}}{18} + C.$

2. $\int \sin(x-3)\,dx = \int \sin(x-3)\,d(x-3) = -\cos(x-3) + C.$

3. $\int \frac{dx}{x+7} = \int \frac{d(x+7)}{x+7} = \ln|x+7| + C.$

4. $\int e^{4x+1}\,dx = \frac{1}{4}\int e^{4x+1}\,d(4x+1) = \frac{1}{4}e^{4x+1} + C.$

5. $\int \frac{dx}{1+4x^2} = \frac{1}{2}\int \frac{d(2x)}{1+(2x)^2} = \frac{1}{2}\operatorname{arctg} 2x + C.$

6. $\int \frac{dx}{\sqrt{1-9x^2}} = \frac{1}{3}\int \frac{d(3x)}{\sqrt{1-(3x)^2}} = \frac{1}{3}\arcsin 3x + C.$

7. $\int \frac{dx}{1+\cos x} = \int \frac{dx}{2\cos^2(x/2)} = \int \frac{d\,(x/2)}{\cos^2(x/2)} = \operatorname{tg}\frac{x}{2} + C.$

8. $\int \frac{dx}{1-\cos x} = \int \frac{d\,(x/2)}{\sin^2(x/2)} = -\operatorname{ctg}\frac{x}{2} + C.$

9. $\int \frac{dx}{\sqrt{3+4x^2}} = \frac{1}{2}\int \frac{d(2x)}{\sqrt{3+(2x)^2}} = \frac{1}{2}\ln\left|2x+\sqrt{3+4x^2}\right| + C.$

10. $\int \cos(5x+2)\,dx = \frac{1}{5}\int \cos(5x+2)d(5x+2) =$

$$= \frac{1}{5}\sin(5x+2) + C.$$

Докажем формулы № 9*, 12* и 14* из таблицы интегралов, используя метод подведения под знак дифференциала:

$$\int \operatorname{tg} x\,dx = \int \frac{\sin x\,dx}{\cos x} = -\int \frac{d(\cos x)}{\cos x} = -\ln|\cos x| + C,$$

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \int \frac{d(x/a)}{\sqrt{1-(x/a)^2}} = \arcsin\frac{x}{a} + C,$$

$$\int \frac{dx}{a^2 + x^2} = \frac{1}{a}\int \frac{d(x/a)}{1+(x/a)^2} = \frac{1}{a}\operatorname{arctg}\frac{x}{a} + C,$$

что и требовалось доказать.

б) Подведение под знак дифференциала основных элементарных функций:

$$e^x\,dx = d(e^x) \qquad a^x\,dx = \frac{d(a^x)}{\ln a} \qquad \cos x\,dx = d(\sin x)$$

$$\sin x\,dx = -d(\cos x) \qquad \frac{dx}{x} = d(\ln|x|) \qquad \frac{dx}{\sin^2 x} = -d(\operatorname{ctg} x)$$

$$\frac{dx}{\cos^2 x} = d(\operatorname{tg} x) \qquad \frac{dx}{1+x^2} = d(\operatorname{arctg} x)$$

$$\frac{dx}{\sqrt{1-x^2}} = d(\arcsin x) \qquad x\,dx = \frac{d(x^2)}{2} \qquad x^2\,dx = \frac{d(x^3)}{3}$$

$$\frac{dx}{\sqrt{x}} = 2d(\sqrt{x}) \qquad \frac{dx}{x^2} = -d\left(\frac{1}{x}\right).$$

Примеры.

1. $$\int x\sqrt{x^2+1}\,dx = \frac{1}{2}\int (x^2+1)^{\frac{1}{2}}\,d(x^2+1) = \\ = \frac{1}{2}\cdot\frac{2}{3}(x^2+1)^{\frac{3}{2}} + C = \frac{1}{3}(x^2+1)\sqrt{x^2+1} + C.$$

2. $$\int x^2\sqrt{x^3+1}\,dx = \frac{1}{3}\int (x^3+1)^{\frac{1}{2}}\,d(x^3+1) = \\ = \frac{1}{3}\cdot\frac{2}{3}(x^3+1)^{\frac{2}{3}} + C = \frac{2}{9}(x^3+1)\sqrt{x^3+1} + C.$$

3. $$\int \frac{\sin\sqrt{x}}{\sqrt{x}}\,dx = 2\int \sin\sqrt{x}\cdot d(\sqrt{x}) = -2\cos\sqrt{x} + C.$$

4. $$\int x\cdot e^{-x^2}\,dx = -\frac{1}{2}\int e^{-x^2}(dx^2) = -\frac{1}{2}e^{-x^2} + C.$$

5. $\displaystyle\int x^2 \operatorname{tg}(x^3+1)\,dx = \frac{1}{3}\int \operatorname{tg}(x^3+1)\,d(x^3+1) = -\frac{1}{3}\ln|\cos(x^3+1)| + C.$

6. $\displaystyle\int \frac{dx}{\sqrt{(3+\arcsin x)(1-x^2)}} = \int \frac{dx}{\sqrt{1-x^2}\cdot\sqrt{3+\arcsin x}} = \int \frac{d(3+\arcsin x)}{\sqrt{3+\arcsin x}} = 2\sqrt{3+\arcsin x} + C.$

7. $\displaystyle\int \sin^3 x \cdot \cos x\,dx = \int \sin^3 x\,d(\sin x) = \frac{1}{4}\sin^4 x + C.$

8. $\displaystyle\int \cos^3 x \cdot \sin 2x\,dx = 2\int \cos^4 x \cdot \sin x\,dx = -2\int \cos^4 x\,d(\cos x) = -\frac{2}{5}\cos^5 x + C.$

9. $\displaystyle\int \frac{(\operatorname{arctg} x)^2\,dx}{1+x^2} = \int (\operatorname{arctg} x)^2\,d(\operatorname{arctg} x) = \frac{1}{3}(\operatorname{arctg} x)^3 + C.$

10. $\displaystyle\int \frac{dx}{(\arcsin x)^3\sqrt{1-x^2}} = \int (\arcsin x)^{-3}\,d(\arcsin x) = \frac{(\arcsin x)^{-2}}{-2} + C = -\frac{1}{2(\arcsin x)^2} + C.$

11. $\displaystyle\int \frac{dx}{\cos^2 x\sqrt{1+\operatorname{tg} x}} = \int \frac{d(1+\operatorname{tg} x)}{\sqrt{1+\operatorname{tg} x}} = 2\sqrt{1+\operatorname{tg} x} + C.$

12. $\displaystyle\int \frac{dx}{x\cos^2(1+\ln x)} = \int \frac{d(1+\ln x)}{\cos^2(1+\ln x)} = \operatorname{tg}(1+\ln x) + C.$

13. $\displaystyle\int e^x(\sin e^x)\,dx = \int \sin e^x \cdot d(e^x) = -\cos e^x + C.$

14. $\displaystyle\int \frac{dx}{\sqrt{(1-x^2)}\cdot\arcsin x} = \int \frac{d(\arcsin x)}{\arcsin x} = \ln|\arcsin x| + C.$

15. $\displaystyle\int \frac{e^x\,dx}{e^x+1} = \int \frac{d(e^x+1)}{e^x+1} = \ln(e^x+1) + C.$

16. $$\int \frac{\cos x\, dx}{\sqrt{1+\sin^2 x}} = \int \frac{d(\sin x)}{\sqrt{1+\sin^2 x}} = \\ = \ln\left|\sin x + \sqrt{1+\sin^2 x}\right| + C.$$

17. $$\int e^{\sin x} \cdot \cos x\, dx = \int e^{\sin x}\, d(\sin x) = e^{\sin x} + C.$$

18. $$\int \frac{x\, dx}{1+9x^2} = \frac{1}{18}\int \frac{d(1+9x^2)}{1+9x^2} = \frac{1}{18}\ln(1+9x^2) + C.$$

Замечания

1. Формула 16* в рамках учебной программы для фармацевтического факультета доказательству не подлежит, но проверяется дифференцированием:

$$\left(\ln\left|x+\sqrt{x^2+m}\right|\right)' = \frac{1}{x+\sqrt{x^2+m}}\cdot\left(1+\frac{2x}{2\sqrt{x^2+m}}\right) = \\ = \frac{x+\sqrt{x^2+m}}{(x+\sqrt{x^2+m})\sqrt{x^2+m}} = \frac{1}{\sqrt{x^2+m}},$$

а это и есть подынтегральная функция из формулы 16*.

2. Формула 15* будет доказана при изложении интегрирования дробно-рациональных функций.

4.1.6. Метод выделения полного квадрата из квадратичного трехчлена

При интегрировании выражений типа $\int \frac{dx}{ax^2+bx+c}$ или $\int \frac{dx}{\sqrt{ax^2+bx+c}}$ выделением полного квадрата из квадратного трехчлена ax^2+bx+c удается свести их к табличным 12*, 14*, 15* или 16*:

$$ax^2+bx+c = a\left(x^2+\frac{b}{a}x+\frac{c}{a}\right) = \\ = a\left(\left(x^2+2\cdot\frac{b}{2a}x+\left(\frac{b}{2a}\right)^2\right)-\left(\frac{b}{2a}\right)^2+\frac{c}{a}\right) = \\ = a\left(\left(x+\frac{b}{2a}\right)^2-\left(\frac{b}{2a}\right)^2+\frac{c}{a}\right).$$

В общем виде эта операция выглядит сложнее, чем на самом деле.

Например, надо выделить полный квадрат из квадратного трехчлена $x^2 + 6x + 5 = (x^2 + 6x + 9) - 9 + 5 = (x + 3)^2 - 4$.

Тогда, используя метод подведения под знак дифференциала, можно вычислить следующие интегралы

1. $$\int \frac{dx}{x^2 + 6x + 5} = \int \frac{dx}{(x+3)^2 - 4} = \int \frac{d(x+3)}{(x+3)^2 - 4} =$$
$$= \frac{1}{4} \ln \left| \frac{(x+3) - 2}{(x+3) + 2} \right| + C = \frac{1}{4} \ln \left| \frac{x+1}{x+5} \right| + C.$$

2. $$\int \frac{dx}{x^2 + 4x + 5} = \int \frac{dx}{(x^2 + 4x + 4) - 4 + 5} =$$
$$= \int \frac{d(x+2)}{(x+2)^2 + 1} = \operatorname{arctg}(x+2) + C.$$

3. $$\int \frac{dx}{\sqrt{3 - x^2 - 2x}} = \int \frac{dx}{\sqrt{-(x^2 + 2x - 3)}} =$$
$$= \int \frac{dx}{\sqrt{-((x^2 + 2x + 1) - 1 - 3)}} = \int \frac{d(x+1)}{\sqrt{4 - (x+1)^2}} =$$
$$= \arcsin \frac{x+1}{2} + C.$$

4. $$\int \frac{dx}{\sqrt{4x^2 + x - 6}} = \int \frac{dx}{\sqrt{4\left(x^2 + \frac{x}{4} - \frac{6}{4}\right)}} =$$
$$= \int \frac{dx}{\sqrt{4\left(\left(x^2 + \frac{x}{4} + \left(\frac{1}{8}\right)^2\right) - \frac{1}{64} - \frac{3}{2}\right)}} =$$
$$= \int \frac{d\left(x + \frac{1}{8}\right)}{\sqrt{4\left(\left(x + \frac{1}{8}\right)^2 - \frac{97}{64}\right)}} = \frac{1}{2} \ln \left| x + \frac{1}{8} + \sqrt{x^2 + \frac{x}{4} - \frac{3}{2}} \right| + C.$$

На заключительном этапе интегрирования была использована формула 16*.

4.1.7. Основные методы интегрирования

Таких методов два: метод замены переменной или подстановка и интегрирование по частям.

Метод замены переменной

Существуют две формулы замены переменной в неопределенном интеграле:

1. $\int f(x)\,dx = \int f\,[\varphi(t)]\,\varphi'(t)\,dt$, где $x = \varphi(t)$;

2. $\int f\,[\varphi(x)]\,\varphi'(x)\,dx = \int f(t)dt$, где $\varphi(x) = t$.

Здесь $x = \varphi(t)$ и $t = \varphi(x)$ суть монотонные дифференцируемые функции своих переменных.

Искусство применения метода состоит, в основном, в выборе функций $x = \varphi(t)$ или $t = \varphi(x)$ так, чтобы новые интегралы являлись табличными или сводились к ним. В окончательном ответе следует вернуться к старой переменной.

Замечание. Подведение под знак дифференциала является частным случаем замены переменной, так как выполняются те же действия, только не вводится новая переменная. Это производится в уме.

Примеры.

1. $\int \frac{dx}{1+\sqrt{x+1}} = \left| x+1 = t^2, \sqrt{x+1} = t, dx = 2tdt \right|$

Здесь следует ввести новую переменную t так, чтобы избавиться ~~от~~ квадратного корня. Положим $x + 1 = t^2$, тогда $x = t^2 - 1$, а $dx = 2t\,dt$:

$$2\int \frac{t\,dt}{1+t} = 2\int \frac{(1+t)-1}{1+t}\,dt = 2\int dt - 2\int \frac{dt}{1+t} =$$

$$= 2t - 2\int \frac{d(1+t)}{1+t} = 2t - 2\ln|1+t| + C = \left| t = \sqrt{x+1} \right| =$$

$$= 2\sqrt{x+1} - 2\ln\left|1+\sqrt{x+1}\right| + C.$$

2. $\int \frac{4x+3}{(x-2)^3}\,dx = \left| x-2 = t, x = t+2, dx = dt \right|.$

Заменив $x - 2$ на t , получим в знаменателе одночлен и после почленного деления интеграл сведется к табличному от степенной функции:

$$\int \frac{4(t+2)+3}{t^3}\,dt = \int \frac{4t+11}{t^3}\,dt = 4\int t^{-2}\,dt + 11\int t^{-3}\,dt =$$
$$= -\frac{4}{t} - \frac{11}{2t^2} + C = -\frac{4}{x-2} - \frac{11}{2(x-2)^2} + C.$$

3. $$\int \frac{dx}{1+\sqrt[3]{x+1}} = \left| x+1=t^3,\ \sqrt[3]{x+1}=t,\ dx=3t^2\,dt \right| =$$
$$= 3\int \frac{t^2\,dt}{t+1} = 3\int \frac{(t^2-1)+1}{t+1}\,dt = 3\int (t+1)\,dt + 3\int \frac{d(t+1)}{t+1} =$$
$$= \frac{3}{2}t^2 + 3t + 3\ln|t+1| + C = \frac{3}{2}\sqrt[3]{(x+1)^2} + 3\sqrt[3]{x+1} +$$
$$+ 3\ln|\sqrt[3]{x+1}+1| + C.$$

4. $$\int \frac{e^{2x}\,dx}{\sqrt[4]{e^x+1}} = \int \frac{e^x e^x\,dx}{\sqrt[4]{e^x+1}} =$$
$$= \left| e^x+1=t^4,\ e^x\,dx=4t^3\,dt,\ e^x=t^4-1 \right| = 4\int \frac{(t^4-1)t^3dt}{t} =$$
$$= 4\int (t^4-1)t^2\,dt = 4\int t^6\,dt - 4\int t^2\,dt = \frac{4}{7}t^7 - \frac{4}{3}t^3 + C =$$
$$= \frac{4}{7}(e^x+1)\sqrt[4]{(e^x+1)^3} - \frac{4}{3}\sqrt[4]{(e^x+1)^3} + C =$$
$$= \sqrt[4]{(e^x+1)^3}\cdot\left(\frac{4}{7}e^x - \frac{16}{21}\right) + C.$$

5. $$\int \frac{\sqrt{1+\ln x}}{x\ln x}\,dx = \left| 1+\ln x = t^2,\ \frac{dx}{x} = 2t\,dt,\ \ln x = t^2-1 \right| =$$
$$= 2\int \frac{t^2\,dt}{t^2-1} = 2\int \frac{(t^2-1)+1}{t^2-1}dt = 2\int dt + 2\int \frac{dt}{t^2-1} =$$
$$= 2t + \ln\left|\frac{(t-1)}{t+1}\right| + C = 2\sqrt{1+\ln x} + \ln\left|\frac{(\sqrt{1+\ln x}-1)}{(\sqrt{1+\ln x}+1)}\right| + C.$$

6. $$\int \frac{x^5\,dx}{\sqrt{a^3 - x^3}} = \left| \begin{array}{l} a^3 - x^3 = t^2,\ -3x^2\,dx = 2t\,dt, \\ x^5\,dx = x^3 \cdot x^2\,dx = \dfrac{2}{3}(t^2 - a^3)t\,dt \end{array} \right| =$$

$$= \frac{2}{3} \int \frac{(t^2 - a^3)t\,dt}{t} = \frac{2}{3} \int t^2\,dt - \frac{2}{3}a^3 \int dt =$$

$$= \frac{2}{9}t^3 - \frac{2a^3 t}{3} + C = \frac{2}{9}(a^3 - x^3)\sqrt{a^3 - x^3} - \frac{2a^3}{3}\sqrt{a^3 - x^3} + C.$$

7. $$\int \frac{dx}{x^2\sqrt{x^2 + 4}} = \left| x = \frac{1}{t},\ dx = -\frac{dt}{t^2} \right| = -\int \frac{t\,dt}{\sqrt{1 + 4t^2}} =$$

$$= -\frac{1}{8} \int (1 + 4t^2)^{-1/2} d(1+4t^2) = -\frac{\sqrt{1 + 4t^2}}{4} + C = -\frac{\sqrt{x^2 + 4}}{|x|} + C.$$

8. $$\int \frac{\sqrt{x^2 + 4}}{x^4}\,dx = \left| \begin{array}{l} x = 2\operatorname{tg} t,\ dx = \dfrac{2}{\cos^2 t} \cdot dt, \\ \sqrt{x^2 + 4} = \dfrac{2}{\cos t} \end{array} \right| = \frac{1}{4} \int \frac{\cos^4 t}{\cos t \cdot \cos^2 t \cdot \sin^4 t}\,dt,$$

$$\frac{1}{4} \int (\sin t)^{-4}\, d(\sin t) = -\frac{1}{12\sin^3 t} + C =$$

$$= \left| \sin t = \frac{\operatorname{tg} t}{\sqrt{1 + \operatorname{tg}^2 t}} = \frac{x\sqrt{x^2 + 4}}{4} \right| =$$

$$= -\frac{64}{12x^3(x^2 + 4)\sqrt{x^2 + 4}} + C = -\frac{16}{3x^3(x^2 + 4)\sqrt{x^2 + 4}} + C.$$

9. $$\int \frac{x^2 dx}{\sqrt{9 - x^2}} = \left| x = 3\sin t, 9 - x^2 = 3\cos t, dx = 3\cos t\,dt \right| =$$

$$= 9 \int \frac{\sin^2 t \cdot \cos t\,dt}{\cos t} = \frac{9}{2} \int (1 - \cos 2t)\,dt = \frac{9}{2}t - \frac{9}{4}\sin 2t + C =$$

$$= \frac{9}{2}\arcsin\frac{x}{3} - \frac{1}{2}x\sqrt{9 - x^2} + C.$$

При переходе к переменной x использованы формулы:

$$\sin 2\alpha = 2\sin\alpha \cdot \cos\alpha, \quad \sin^2\alpha + \cos^2\alpha = 1.$$

Метод интегрирования по частям

Дифференциал произведения двух функций определяется формулой $d(u \cdot v) = u\,dv + v\,du$.

Интегрируя это равенство, получим выражение:

$$u \cdot v = \int u\,dv + \int v\,du.$$

Отсюда $\int u\,dv = uv - \int v\,du$. Это и есть формула *интегрирования по частям.*

Применение этого метода предполагает субъективное представление подынтегрального выражения в виде $u \cdot dv$, и при этом интеграл $\int v\,du$ должен быть не труднее, чем $\int u\,dv$. В противном случае применение метода не имеет смысла.

Итак, искусство применения метода интегрирования по частям предполагает умение выделять из подынтегральной функции сомножители u и dv с учетом вышеизложенных требований. Конечно, не все интегралы могут быть найдены этим методом.

Примеры.

1) Интегралы вида

$$\int P(x)e^{kx}\,dx, \quad \int P(x)\sin kx\,dx, \quad \int P(x)\cos kx\,dx,$$

где $P(x)$ — многочлен, k — константа вычисляются по частям, обозначив $u = P(x)$, а dv — все остальные сомножители.

$$\mathbf{1.}\quad \int xe^{3x}\,dx = \left| u = x; dv = e^{3x}\,dx; du = dx; v = \frac{1}{3}e^{3x} \right| =$$
$$= \frac{1}{3}xe^{3x} - \frac{1}{3}\int e^{3x}dx = \frac{1}{3}\left(x - \frac{1}{3}\right)e^{3x} + C.$$

2) Интегралы типа

$$\int P(x)\arcsin x\,dx, \quad \int P(x)\,\mathrm{arctg}\,x\,dx, \quad \int P(x)\ln x\,dx.$$

Здесь положим $dv = P(x)\,dx$, а u — другие сомножители.

2. $\displaystyle\int x^2\ln x\,dx = \left| u=\ln x;\, dv = x^2\,dx;\, du=\frac{1}{x}\,dx; \right.$

$$\left. v=\frac{1}{3}x^3 \right| = \frac{1}{3}x^3\ln x-\frac{1}{3}\int x^2\,dx=\frac{1}{3}x^3\ln x-\frac{x^3}{9}+C=$$

$$=\frac{1}{3}x^3\left(\ln x-\frac{1}{3}\right)+C.$$

3. $\displaystyle\int x\operatorname{arctg}x\,dx = \left| u=\operatorname{arctg}x;\, dv=x\,dx;\, du=\frac{dx}{1+x^2}; \right.$

$$\left. v=\frac{x^2}{2}\right| = \frac{x^2}{2}\operatorname{arctg}x-\frac{1}{2}\int\frac{x^2}{1+x^2}\,dx=\frac{x^2}{2}\operatorname{arctg}x-$$

$$-\frac{1}{2}\int\frac{(1+x^2-1)}{(1+x^2)}\,dx=\frac{x^2}{2}\operatorname{arctg}x-\frac{1}{2}\int dx+\frac{1}{2}\int\frac{dx}{(1+x^2)}=$$

$$=\frac{(x^2+1)\operatorname{arctg}x}{2}-\frac{x}{2}+C.$$

4. $\displaystyle\int\arcsin x\,dx = \left| u=\arcsin x;\, dv=dx;\, du=\frac{dx}{\sqrt{1-x^2}}; \right.$

$$\left. v=x\right| = x\arcsin x-\int\frac{x\,dx}{\sqrt{1-x^2}}=x\arcsin x+\int\frac{d(1-x^2)}{2\sqrt{1-x^2}}=$$

$$=x\arcsin x+\sqrt{1-x^2}+C.$$

Любой результат можно проверить дифференцированием. Например, в данном случае

$$(x\arcsin x+\sqrt{1-x^2}+C)'=\arcsin x+\frac{x}{\sqrt{1-x^2}}-\frac{x}{\sqrt{1-x^2}}=$$

$$=\arcsin x.$$

Результат верен.

3) Интегралы вида

$$\int e^{ax}\sin bx\,dx,\qquad \int e^{ax}\cos bx\,dx,$$

где a, b — const. За u можно взять e^{ax}, $\sin bx$ или $\cos bx$.

5. $\displaystyle\int e^x \sin x\,dx = \left| \begin{matrix} u = \sin x;\, dv = e^x\,dx; \\ du = \cos x\,dx;\, v = e^x \end{matrix} \right| =$

$$= e^x \sin x - \int e^x \cos x\,dx = \left| \begin{matrix} u = \cos x;\, dv = e^x\,dx; \\ du = -\sin x\,dx;\, v = e^x \end{matrix} \right| =$$

$$= e^x \sin x - e^x \cos x - \int e^x \sin x\,dx.$$

Отсюда получаем, что

$$\int e^x \sin x\,dx = e^x(\sin x - \cos x)/2 + C.$$

6. $\displaystyle\int \sqrt{1+x^2}\,dx = \left| \begin{matrix} u = \sqrt{1+x^2};\, dv = dx; \\ du = \dfrac{x}{\sqrt{1+x^2}}dx;\, v = x \end{matrix} \right| =$

$$= x\sqrt{1+x^2} - \int \frac{x^2dx}{\sqrt{1+x^2}}.$$

Отсюда

$$2\int \sqrt{1+x^2}\,dx = x\sqrt{1+x^2} + \int \frac{dx}{\sqrt{1+x^2}},$$

$$\int \sqrt{1+x^2}\,dx = \frac{(x\sqrt{1+x^2})}{2} + \frac{1}{2}\ln(x + \sqrt{1+x^2}) + C.$$

7. $\displaystyle\int \frac{x\,dx}{\cos^2 x} = \left| u = x;\, dv = \frac{dx}{\cos^2 x};\, du = dx;\, v = \operatorname{tg} x \right| =$

$$= x \operatorname{tg} x - \int \operatorname{tg} x\,dx = x \operatorname{tg} x + \ln|\cos x| + C.$$

8. $\displaystyle\int e^{\sqrt{x}}\,dx$. Здесь надо сперва сделать замену переменной, а потом интегрировать по частям:

$$\int e^{\sqrt{x}}\,dx = \left| x = t^2;\, dx = 2t\,dt \right| = 2\int te^t\,dt = 2(t-1)e^t + C =$$

$$= 2(\sqrt{x} - 1)e^{\sqrt{x}} + C.$$

9. $\displaystyle\int \sin\sqrt{x}\,dx = \left| x = t^2;\, dx = 2t\,dt \right| = 2\int t\sin t\,dt =$

$$= \left| \begin{array}{l} u = t;\, dv = \sin t\,dt; \\ du = dt;\, v = -\cos t \end{array} \right| = -2t\cos t + 2\int \cos t\,dt =$$

$$= -2t\cos t + 2\sin t = -2\sqrt{x}\cos\sqrt{x} + 2\sin\sqrt{x} + C.$$

10. $\displaystyle\int \frac{x^3dx}{\sqrt{1+x^2}}$. Этот интеграл с равным успехом может быть найден как в результате замены переменной $1 + x^2 = t^2$, так и методом интегрирования по частям:

$$\int \frac{x^3\,dx}{\sqrt{1+x^2}} = \left| \begin{array}{l} u = x^2;\, dv = \dfrac{x\,dx}{\sqrt{1+x^2}}; \\ du = 2x\,dx;\, v = \sqrt{1+x^2} \end{array} \right| = x^2\sqrt{1+x^2} -$$

$$-2\int x\sqrt{1+x^2}\,dx = x^2\sqrt{1+x^2} - \int (1+x^2)^{1/2}\,d(1+x^2) =$$

$$= x^2\sqrt{1+x^2} - \frac{2}{3}(1+x^2)\sqrt{1+x^2} + C = \frac{(x^2-2)}{3\sqrt{1+x^2}} + C.$$

Самостоятельная работа

Непосредственное интегрирование

1. $\displaystyle\int \frac{d\sqrt{x-4}}{\sqrt{x-4}}$. **2.** $\displaystyle\int \sin^3 2x\,d(\sin 2x)$. **3.** $\displaystyle\int \frac{d(\cos 5x)}{\cos^2 5x - 4}$.

4. $\displaystyle\int \left(\frac{1}{\operatorname{tg}^2 x + 9} + \operatorname{tg}^3 x\right) d(\operatorname{tg} x)$. **5.** $\displaystyle\int \frac{d(\arcsin x)}{\arcsin x}$.

6. $\displaystyle\int 2^{3x-4}\,d(3x-4)$. **7.** $\displaystyle\int \frac{d(\ln x)}{\sqrt{4-\ln^2 x}}$.

8. $\displaystyle\int \left(\frac{3}{\sqrt[3]{5x-1}} - \frac{1}{5x-1}\right) d(5x-1)$.

9. $\displaystyle\int (\sin(e^x) + 2e^x)\,d(e^x)$. **10.** $\displaystyle\int \frac{d(2^x)}{\sqrt{4^x+9}}$.

Простейшие методы интегрирования

11. $\int \frac{dx}{x+3} = \ln|x+3| + C.$

12. $\int \frac{dx}{\sqrt{1+(x-1)^2}} = \ln\left|x-1+\sqrt{x^2-2x+2}\right| + C.$

13. $\int \frac{dx}{(x+2)^2-3} = \frac{1}{2\sqrt{3}} \ln\left|\frac{x+2-\sqrt{3}}{x+2+\sqrt{3}}\right| + C.$

14. $\int 5^{3x}\,dx = \frac{1}{3\ln 5} 5^{3x} + C.$ **15.** $\int \frac{dx}{4+9x^2} = \frac{1}{6} \operatorname{arctg} \frac{3x}{2} + C.$

16. $\int \frac{dx}{\sqrt{4-25x^2}} = \frac{1}{5} \arcsin \frac{5x}{2} + C.$

17. $\int \frac{dx}{1-\cos x} = -\operatorname{ctg} \frac{x}{2} + C.$

18. $\int \frac{dx}{\cos 4x + \sin^2 2x} = \frac{1}{2} \operatorname{tg} 2x + C.$

19. $\int \cos(1-2x)\,dx = -\frac{1}{2} \sin(1-2x) + C.$

20. $\int \sqrt[5]{(8-3x)^4}\,dx = -\frac{5}{27}(8-3x)\sqrt[5]{(8-3x)^4} + C.$

21. $\int x^2\sqrt{x^3+10}\,dx = \frac{2}{9}(x^3+10)\sqrt{x^3+10} + C.$

22. $\int \frac{\cos\sqrt{x+1}}{\sqrt{x+1}} = 2\sin\sqrt{x+1} + C.$

23. $\int \operatorname{tg}(2x+3)\,dx = -\frac{1}{2} \ln|\cos(2x+3)| + C.$

24. $\int \operatorname{ctg}(7x-1)\,dx = -\frac{1}{7} \ln|\sin(7x-1)| + C.$

25. $\int \frac{dx}{\sqrt{x(1-x)}} = 2\arcsin\sqrt{x} + C.$

26. $\int \sin 4x \cdot \cos^2 2x = -\frac{\cos^4 2x}{4} + C.$

27. $\displaystyle\int \frac{e^{\operatorname{arctg} x}}{1+x^2}\,dx = e^{\operatorname{arctg} x} + C.$

28. $\displaystyle\int \frac{dx}{\arcsin 2x \cdot \sqrt{1-4x^2}} = \frac{1}{2}\ln|\arcsin 2x| + C.$

29. $\displaystyle\int \frac{dx}{(\cos 2x + \sin^2 x)(3+\operatorname{tg} x)} = \ln|3+\operatorname{tg} x| + C.$

30. $\displaystyle\int \frac{\cos x\,dx}{4-\sin^2 x} = \frac{1}{4}\ln\left|\frac{2+\sin x}{2-\sin x}\right| + C.$

31. $\displaystyle\int \frac{\cos x\,dx}{\sqrt{1+\sin^2 x}} = \ln\left|\sin x + \sqrt{1+\sin^2 x}\right| + C.$

32. $\displaystyle\int \frac{dx}{\sin^2 x\sqrt{1+\operatorname{ctg}^2 x}} = -\ln\left|\operatorname{ctg} x + \sqrt{1+\operatorname{ctg}^2 x}\right| + C.$

33. $\displaystyle\int \frac{e^x\,dx}{e^{2x}+4} = \frac{1}{2}\operatorname{arctg}\frac{e^x}{2} + C.$

34. $\displaystyle\int \frac{\sin 2x}{1+\sin^2 x}\,dx = \ln|1+\sin^2 x| + C.$

35. $\displaystyle\int \frac{x\,dx}{1+x^4} = \frac{1}{2}\operatorname{arctg} x^2 + C.$

36. $\displaystyle\int \frac{(\ln x)^m}{x(\ln x)^{m+1}} = \ln|\ln x| + C$, если $m=-1$; если $m \neq -1$.

37. $\displaystyle\int \frac{dx}{\cos^4 x} = \operatorname{tg} x + \frac{1}{3}\operatorname{tg}^2 x + C.$

38. $\displaystyle\int \frac{1-\sin x}{\cos x}\,dx = \ln|1+\sin x| + C.$

39. $\displaystyle\int \frac{dx}{1+\sin x} = \frac{\sin x - 1}{\cos x} + C.$

40. $\displaystyle\int \frac{dx}{1-\cos x} = -\frac{1+\cos x}{\sin x} + C.$

41. $\displaystyle\int \frac{dx}{9x^2+6x+2} = \frac{1}{3}\operatorname{arctg}(3x+1) + C.$

42. $\displaystyle\int \frac{dx}{x^2+6x+8} = \frac{1}{2}\ln\left|\frac{x+2}{x+4}\right| + C.$

43. $\displaystyle\int \frac{dx}{24-x^2-2x} = \frac{1}{10}\ln\left|\frac{x+6}{x-4}\right| + C.$

44. $\displaystyle\int \frac{dx}{\sqrt{9x^2+6x+2}} = \frac{1}{3}\ln\left|3x+1+\sqrt{9x^2+6x+2}\right| + C.$

45. $\displaystyle\int \frac{dx}{\sqrt{8-x^2+2x}} = \arcsin\frac{x-1}{3} + C.$

46. $\displaystyle\int \frac{x\,dx}{x^4+2x^2+2} = \frac{1}{2}\operatorname{arctg}(x^2+1) + C.$

Методы замены переменной и интегрирования по частям

47. $\displaystyle\int \frac{dx}{1+\sin^2 x} = \frac{1}{\sqrt{2}}\operatorname{arctg}(\sqrt{2}\operatorname{tg} x) + C.$

48. $\displaystyle\int \sqrt[3]{1+3\sin x}\cos x\,dx = \frac{(1+3\sin x)\sqrt[3]{1+3\sin x}}{4} + C.$

49. $\displaystyle\int \frac{\sin 2x}{1+\sin^2 x}\,dx = \ln(1+\sin^2 x) + C.$

50. $\displaystyle\int \frac{1+\ln x}{3+x\ln x} = \ln|3+x\ln x| + C.$

51. $\displaystyle\int \frac{x\,dx}{3-x^4} = \frac{1}{2}\arcsin\frac{x^2}{\sqrt{3}} + C.$

52. $\displaystyle\int \frac{x^{n-1}}{x^{2n}+a^2}\,dx = \frac{1}{an}\operatorname{arctg}\frac{x^n}{a} + C.$

53. $\displaystyle\int \left(\ln x + \frac{1}{\ln x}\right)\frac{dx}{x} = \frac{1}{2}\ln^2 x + \ln|\ln x| + C.$

54. $\displaystyle\int \cos^5 x\sqrt{\sin x}\,dx = \frac{2}{3}\sin x\sqrt{\sin x} - \frac{4}{7}\sin^3 x\sqrt{\sin x} +$

$\displaystyle+\frac{2}{11}\sin^5 x\sqrt{\sin x} + C = \sqrt{\sin x}\left(\frac{2}{3}\sin x - \frac{4}{7}\sin^3 x + \frac{2}{11}\sin^5 x\right) + C.$

55. $\displaystyle\int \operatorname{arctg}\sqrt{x}\,dx = (x+1)\operatorname{arctg}\sqrt{x} - \sqrt{x} + C.$

56. $\displaystyle\int e^{\sqrt{x}}\,dx = 2e^{\sqrt{x}}(\sqrt{x}-1)+C.$

57. $\displaystyle\int \ln(x+\sqrt{1+x^2})\,dx = x\ln(x+\sqrt{1+x^2})-\sqrt{1+x^2}+C.$

58. $\displaystyle\int \frac{x\cos x\,dx}{\sin^3 x} = -\frac{1}{2}\left(\frac{x}{\sin^2 x}+\operatorname{ctg} x\right)+C.$

59. $\displaystyle\int x\operatorname{tg}^2 x\,dx = x\operatorname{tg} x-\frac{x^2}{2}+\ln|\cos x|+C.$

60. $\displaystyle\int \frac{x\operatorname{arctg} x\,dx}{\sqrt{1+x^2}} = \sqrt{1+x^2}\operatorname{arctg} x-\ln(x+\sqrt{1+x^2})+C.$

61. $\displaystyle\int \ln(x^2+1)\,dx = x\ln(x^2+1)-2x+2\operatorname{arctg} x+C.$

62. $\displaystyle\int \frac{\arcsin\sqrt{x}\,dx}{\sqrt{1-x}} = 2(\sqrt{x}-\sqrt{1-x}\arcsin\sqrt{x})+C.$

63. $\displaystyle\int \ln^2 x\,dx = x(\ln^2 x-2\ln x+2)+C.$

64. $\displaystyle\int \sin\ln x\,dx = \frac{x}{2}(\sin\ln x-\cos\ln x)+C.$

65. $\displaystyle\int \cos\ln x\,dx = \frac{x}{2}(\cos\ln x+\sin\ln x)+C.$

66. $\displaystyle\int \frac{\ln\operatorname{tg} x}{\sin x\cdot\cos x}\,dx = \frac{1}{2}\ln^2\operatorname{tg} x+C.$

67. $\displaystyle\int \frac{\ln x\,dx}{x(1-\ln^2 x)} = -\frac{1}{2}\ln|1-\ln^2 x|+C.$

68. $\displaystyle\int \frac{\operatorname{ctg} x}{\ln\sin x}\,dx = \ln|\ln\sin x|+C.$ **69.** $\displaystyle\int e^{e^x+x}\,dx = e^{e^x}+C.$

70. $\displaystyle\int \frac{x^7\,dx}{(1+x^4)^2} = \frac{1}{4}\ln(1+x^4)+\frac{1}{4(1+x^4)}+C.$

71. $\displaystyle\int \frac{\ln(x+1)-\ln x}{x(x+1)}\,dx = -\frac{\ln^2(1+t)}{2}+C.$

72. $$\int \frac{x^7\,dx}{(x^2-1)^3} = -\frac{1}{4(x^2-1)^2} - \frac{3}{2(x^2-1)} + \frac{3}{2}\ln|x^2-1| + \frac{1}{2}x^2 + C.$$

73. $$\int \sqrt{\frac{1+x}{1-x}}\,dx = \arcsin x - \sqrt{1-x^2} + C.$$

74. $$\int \ln\left|\frac{1}{x-\sqrt{1+x^2}}\right| dx = x\ln(x+\sqrt{1+x^2}) - \sqrt{1+x^2} + C.$$

75. $$\int \frac{\operatorname{arctg}\sqrt{x}}{\sqrt{x}}\,dx = 2\sqrt{x}\operatorname{arctg}\sqrt{x} - \ln(1+t^2) + C.$$

§ 4.2. Интегрирование дробно-рациональных функций

Дробно-рациональной функцией называется отношение двух алгебраических многочленов (полиномов): $P_n(x)/Q_m(x)$. Здесь $P_n(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$, а $Q_{m(x)} = b_0 + b_1x + \ldots + b_mx^m$ — алгебраические многочлены степени n и m соответственно.

Коэффициенты многочленов $a_0, a_1, a_2, \ldots, a_n, b_0, b_1, b_2, \ldots, b_m$ суть некоторые действительные числа.

В дальнейшем для кратности будем называть дробно-рациональную функцию просто рациональной дробью.

Рациональная дробь называется правильной, если степень числителя меньше степени знаменателя, т. е. $n < m$. В противном случае рациональная дробь называется неправильной. Любая неправильная рациональная дробь может быть представлена в виде суммы алгебраического полинома и правильной рациональной дроби. Например, неправильная рациональная дробь $\dfrac{x^4+x-1}{x^3+1}$ равна сумме x и правильной рациональной дроби $\dfrac{-1}{x^3+1}$. Алгебраический многочлен называется целой частью неправильной рациональной дроби. В общем случае выделение целой части производится следующим образом. Сначала в числителе и знаменателе степени x располагаются в порядке убывания. Например, $\frac{1+x^3+2x-3x^2}{x+x^2+1} = \frac{x^3-3x^2+2x+1}{x^2+x+1}$. В качестве следующего шага делим старшую степень числителя на старшую степень знаменателя: $\frac{x^3}{x^2} = x$. Затем умножаем знаменатель на полученный результат и вычитаем это произведение из числителя: $x^3-3x^2+2x+1-x(x^2+x+1) = -4x^2+x+1$. Продолжаем этот процесс: $-\frac{4x^2}{x^2} = -4$, $-4x^2+x+1+4(x^2+x+1) = 5x+5$. В дальнейшем $5x$ уже нельзя разделить на x^2. Следовательно, в результате мы получим равенство: $\frac{x^3-3x^2+2x+1}{x^2+x+1} = (x-4)+ + \frac{5x+5}{x^2+x+1}$. Рассмотрим еще один пример без подробных объяснений отдельных действий. Выделим целую часть из неправильной рациональной дроби $\frac{x^6-3x^5+2x+1}{x^3-x+1}$. Последовательность работы следующая:

$$
\begin{array}{r|l}
-\; x^6-3x^5+2x+1 & x^3-x+1 \\
x^6-x^4+x^3 \qquad\qquad & \overline{x^3-3x^2+x-4} \\
\hline
-\; -3x^5+x^4-x^3+2x+1 & \\
-3x^5+3x^3-3x^2 \qquad\quad & \\
\hline
-\; x^4-4x^3+3x^2+2x+1 & \\
x^4-x^2+x \qquad\qquad & \\
\hline
-\; -4x^3+4x^2+x+1 & \\
-4x^3+4x-4 & \\
\hline
4x^2-3x+5. &
\end{array}
$$

Итак, неправильная рациональная дробь представлена в виде суммы алгебраического полинома (x^3-3x^2+x-4) и правильной рациональной дроби $\frac{4x^2-3x+5}{x^3-x+1}$. Задача интегрирования неправильной рациональной дроби приводится к задаче интегрирования целой части, т. е. алгебраического многочлена, и правильной рациональной дроби.

4.2.1. Методика разложения правильной рациональной дроби на простейшие

Любая правильная рациональная дробь может быть представлена в виде суммы простейших рациональных дробей следующих четырех типов:

1) $\dfrac{A}{x-a}$.

2) $\dfrac{A}{(x-a)^k}$, где $k = 2, 3, \ldots$

3) $\dfrac{Ax+B}{x^2+px+q}$, где $\dfrac{p^2}{4} - q < 0$.

4) $\dfrac{Ax+B}{(x^2+px+q)^k}$, где $k = 2, 3, \ldots$

Простейшая дробь четвертого типа не входит в учебную программу и поэтому не будет рассматриваться. Ограничимся рассмотрением случая, когда знаменатель правильной рациональной дроби представлен в виде произведения линейных сомножителей в любой степени и квадратичных сомножителей в первой степени.

Рассмотрим порядок разложения правильной рациональной дроби на простейшие. Если знаменатель представляет собой произведение простых линейных сомножителей, то разложение содержит только простейшие дроби первого типа. Для примера разложим на простейшие правильную дробь $\dfrac{2x-3}{x(x^2-4)}$. Разложение имеет вид: $\dfrac{2x-3}{x(x-2)(x+2)} = \dfrac{A}{x} + \dfrac{B}{x-2} + \dfrac{C}{x+2}$. Для нахождения A сперва умножим последнее равенство на x, а затем положим $x = 0$. В результате получим аналогично: $\dfrac{2x-3}{x(x+2)} = \dfrac{A}{x}(x-2) + B + \dfrac{C}{x+2}(x-2)$; если $x = 2$, то $B = \dfrac{1}{8}$. Совершенно аналогично находим, что $C = \dfrac{7}{8}$. Окончательно: $\dfrac{2x-3}{x(x^2-4)} = \dfrac{3}{4x} + \dfrac{1}{8(x-2)} + \dfrac{7}{8(x+2)}$. Если знаменатель содержит множитель $(x-a)^k$, то этому множителю соответствует k простейших дробей: одна дробь первого и $k-1$ дробей второго типа. Для примера разложим на простейшие правильную рациональную дробь $\dfrac{x^2+3x+4}{x(x-1)^2} = \dfrac{A}{x} + \dfrac{B}{(x-1)^2} + \dfrac{C}{x-1}$. Коэффициенты A и B можно найти аналогично предыдущему:

$$A = \left.\frac{x^2+3x+4}{(x-1)^2}\right|_{x=0} = 4, \qquad B = \left.\frac{x^2+3x+4}{x}\right|_{x=1} = 7,$$

$$\frac{x^2+3x+4}{x(x-1)^2} = \frac{4}{x} + \frac{7}{(x-1)^2} + \frac{C}{x-1}.$$

Для нахождения коэффициента C можно принять в последнем равенстве, например, $x=2$. Это дает уравнение для коэффициента C:

$$\frac{4+6+4}{2} = \frac{4}{2} + 7 + C, \qquad C = -2,$$

$$\frac{x^2+3x+4}{x(x-1)^2} = \frac{4}{x} + \frac{7}{(x-1)^2} + \frac{2}{x-1}.$$

Если знаменатель содержит множитель x^2+px+q, где $\dfrac{p^2}{4} - q < 0$, то разложение правильной дроби на простейшие содержит простейшую дробь третьего типа. Изложим этот случай на примере разложения на простейшие дроби $\dfrac{4x-3}{x^2(x^2+2x+2)}$.

$$\frac{4x-3}{x^2(x^2+2x+2)} = \frac{A}{x^2} + \frac{B}{x} + \frac{Cx+D}{x^2+2x+2}.$$

Сразу видно, что $A = -1{,}5$.

Для нахождения остальных коэффициентов вновь сложили простейшие рациональные дроби в правой части равенства:

$$\frac{4x-3}{x^2(x^2+2x+2)} = \frac{A(x^2+2x+2) + Bx(x^2+2x+2) + x^2(Cx+D)}{x^2(x^2+2x+2)}.$$

Если две дроби равны и у них равны знаменатели, то у них равны и числители.

Два алгебраических многочлена равны, если у них совпадают коэффициенты при соответствующих степенях x. Это приводит к системе линейных алгебраических уравнений:

$$\left.\begin{aligned} \text{при}\quad x^3 &: \quad 0 = B + C, \\ x^2 &: \quad 0 = A + 2B + D, \\ x\ &: \quad 4 = 2A + 2B, \\ x^0 &: \quad -3 = 2A. \end{aligned}\right\}.$$

Четвертое уравнение дает $A = -1{,}5$, что подтверждает найденное раннее значение. Подставим в третье уравнение найденное значение A: $4 = -3 + 2B$. Отсюда $B = 3{,}5$. Подставляя найденные A и B во второе уравнение, получим: $0 = -1{,}5 + 7 + D$; $D = -5{,}5$. На основании первого уравнения $C = -3{,}5$. Окончательно получаем разложение правильной рациональной дроби на простейшие: $\dfrac{4x-3}{x^2(x^2+2x+2)} = -\dfrac{1{,}5}{x^2} + \dfrac{3{,}5}{x} - \dfrac{3{,}5x+5{,}5}{x^2+2x+2}$.

Вывод параграфа: интегрирование правильной рациональной дроби всегда можно свести к интегрированию суммы простейших рациональных дробей.

4.2.2. Методика интегрирования простейших рациональных дробей

Простейшие дроби первого и второго типов интегрируются путем применения простейшего метода интегрирования, изложенного в п. 4.1.5.:

$$\int \frac{A}{x-a}\,dx = A\int \frac{d(x-a)}{x-a} = A\ln|x-a| + C,$$

$$\int \frac{A}{(x-a)^k}dx = A\int (x-a)^{-k}d(x-a) = \frac{A(x-a)^{-k+1}}{(-k+1)} + C =$$
$$= \frac{A}{(1-k)(x-a)^{k-1}} + C \quad (k \neq 1).$$

Простейшая дробь третьего типа интегрируется путем применения методов интегрирования, изложенных в гл. 4.1.

Итак, надо найти неопределенный интеграл $\displaystyle\int \frac{Ax+B}{x^2+px+q}\,dx$ при условии $\dfrac{p^2}{4} - q < 0$. Выделим полный квадрат в знаменателе

$x^2+px+q = \left(x+\frac{p}{2}\right)^2+q-\frac{p^2}{4}$, затем сделаем замену переменной $x+\frac{p}{2}=t$, тогда $x=t-\frac{p}{2}$, $dx=dt$. Обозначим $q-\frac{p^2}{4}=a^2$ и получим

$$\int \frac{Ax+B}{x^2+px+q}\,dx = \int \frac{A\left(t-\frac{p}{2}\right)+B}{t^2+a^2}\,dt =$$

$$= A\int \frac{t\,dt}{t^2+a^2} + \left(B-\frac{Ap}{2}\right)\int \frac{dt}{t^2+a^2} =$$

$$= \frac{1}{2}A\ln(t^2+a^2) + \left(B-\frac{Ap}{2}\right)\frac{1}{a}\operatorname{arctg}\frac{t}{a}+C =$$

возвращаясь к старой переменной, получим

$$= \frac{1}{2}A\ln\left(x^2+px+q\right) + \frac{\left(B-\frac{Ap}{2}\right)}{\sqrt{q-\frac{p^2}{4}}}\operatorname{arctg}\frac{x+\frac{p}{2}}{\sqrt{q-\frac{p^2}{4}}}+C.$$

Окончательный результат интегрирования:

$$\int \frac{Ax+B}{x^2+px+q}\,dx =$$

$$= \frac{A}{2}\ln(x^2+px+q) + \frac{2B-Ap}{\sqrt{4q-p^2}}\operatorname{arctg}\frac{2x+p}{\sqrt{4q-p^2}}+C.$$

Разумеется, нет смысла заучивать эту формулу наизусть. Лучше всего запомнить последовательность действий или включить эту формулу в *личный справочник*.

В заключение этого параграфа интегрируем правильную рациональную дробь, рассмотренную в конце п. 4.5.1.:

$$\int \frac{4x-3}{x^2(x^2+2x+2)}\,dx = -1,5\int\frac{dx}{x^2}+3,5\int\frac{dx}{x}-\int\frac{3,5x+5,5}{x^2+2x+2}\,dx =$$

$$= \frac{1,5}{x}+3,5\ln|x| - \left|\begin{array}{r} x^2+2x+2=x^2+2x+1+1= \\ =(x+1)^2+1=t^2+1, \\ x+1=t,\quad x=t-1,\quad dx=dt \end{array}\right| -$$

$$-\int \frac{3{,}5(t-1)+5{,}5}{t^2+1}\,dt = \frac{1{,}5}{x} + 3{,}5\ln|x| - 3{,}5\int\frac{t\,dt}{t^2+1} - 2\int\frac{dt}{t^2+1} =$$

$$= \frac{1{,}5}{x} + 3{,}5\ln|x| - \frac{3{,}5}{2}\ln(t^2+1) - 2\operatorname{arctg} t =$$

$$= \frac{1{,}5}{x} + 3{,}5\ln|x| - 1{,}75\ln(x^2+2x+2) - 2\operatorname{arctg}(x+1) + C.$$

4.2.3. Примеры интегрирования дробно-рациональных функций (рациональных дробей)

1. Найти интеграл $\int \frac{3x+2}{x^2+3x-4}\,dx$. Поскольку подынтегральная функция — правильная несократимая дробь, разложим ее на простейшие дроби:

$$\frac{3x+2}{x^2+3x-4} = \frac{3x+2}{(x-1)(x+4)} = \frac{A}{x-1} + \frac{B}{x+4},$$

$$3x+2 = A(x+4) + B(x-1),$$

при $x=1$: $\quad 5 = 5A \Rightarrow A = 1$,
при $x=-4$: $\quad -10 = -5B \Rightarrow B = 2$,

$$\frac{3x+2}{x^2+3x-4} = \frac{1}{x-1} + \frac{2}{x+4}.$$

Теперь можно интегрировать:

$$\int \frac{3x+2}{x^2+3x-4}\,dx = \int\frac{dx}{x-1} + 2\int\frac{dx}{x+4} =$$

$$= \ln|x-1| + 2\ln|x+4| + \ln|C| = \ln\left|C(x-1)(x+4)^2\right|.$$

2. Найти интеграл $\int \frac{dx}{(x-1)(x-2)(x-3)}$.

Как и в первом примере, подынтегральная функция — правильная и несократимая дробь, где знаменатель представляет собой произведение линейных двучленов в первой степени. Поэтому ее разложение на простейшие дроби содержит лишь простейшие дроби первого типа:

$$\frac{1}{(x-1)(x-2)(x-3)} = \frac{A}{x-1} + \frac{B}{x-2} + \frac{C}{x-3},$$

$$A = \frac{1}{(1-2)(1-3)} = \frac{1}{2}, \quad B = \frac{1}{(2-1)(2-3)} = -1,$$

$$C = \frac{1}{(3-1)(3-2)} = \frac{1}{2},$$

$$\frac{1}{(x-1)(x-2)(x-3)} = \frac{1}{2} \cdot \frac{1}{x-1} - \frac{1}{x-2} + \frac{1}{2} \cdot \frac{1}{x-3},$$

$$\int \frac{dx}{(x-1)(x-2)(x-3)} = \frac{1}{2} \int \frac{dx}{x-1} - \int \frac{dx}{x-2} + \frac{1}{2} \int \frac{dx}{x-3} =$$

$$= \frac{1}{2} \ln|x-1| - \ln|x-2| + \frac{1}{2} \ln|x-3| + \ln|C| =$$

$$= \ln \left| \frac{C\sqrt{|(x-1)(x-3)|}}{x-2} \right|.$$

3. Найти интеграл $\displaystyle\int \frac{3x^3 + x^2 - 1}{x^2 - 1}\, dx$.

Под знаком интеграла — неправильная рациональная дробь. Поэтому сперва надо выделить из нее целую часть, а затем остаток разложить на сумму простейших дробей.

$$\begin{array}{r|l} -\,3x^3 + x^2 - 1 & x^2 - 1 \\ \cline{2-2} 3x^3 - 3x \qquad & 3x + 1 \\ \cline{1-1} -\,x^2 + 3x - 1 & \\ x^2 \qquad - 1 & \\ \cline{1-1} 3x, \qquad & \end{array}$$

$$\frac{3x^3 + x^2 - 1}{x^2 - 1} = 3x + 1 + \frac{3x}{x^2 - 1},$$

$$\frac{3x}{x^2 - 1} = \frac{3x}{(x-1)(x+1)} = \frac{A}{x-1} + \frac{B}{x+1},$$

$$A = \frac{3}{2}, \quad B = \frac{3}{2}.$$

Заметим однако, что в данном конкретном случае можно обойтись и без разложения остатка на простейшие дроби:

$$\int \frac{3x^3 + x^2 - 1}{x^2 - 1}\, dx = \int (3x + 1)\, dx + 3 \int \frac{x\, dx}{x^2 - 1} =$$

$$= \frac{3}{2}x^2 + x + \frac{3}{2}\int \frac{d(x^2-1)}{x^2-1} = \frac{3}{2}x^2 + x + \frac{3}{2}\ln|x^2-1| + C =$$

$$= \frac{3}{2}(x^2 + \ln|x^2-1|) + x + C.$$

4. Найти интеграл $\int \frac{dx}{x^3(x^2-1)}$.

Наличие множителя x^3 в знаменателе говорит о том, что в разложении подынтегральной функции на простейшие появляются простейшие дроби и первого и второго типов:

$$\frac{1}{x^3(x^2-1)} = \frac{A}{x^3} + \frac{B}{x^2} + \frac{C}{x} + \frac{D}{x-1} + \frac{E}{x+1},$$

$$1 = Ax^2(x-1)(x+1) + Bx(x-1)(x+1) + Cx^2(x-1)(x+1) +$$
$$+ Dx^3(x+1) + Ex^3(x-1).$$

Подставляя в это тождество $x = 0, 1, -1$, получим

$$A = -1, \quad D = \frac{1}{2}, \quad E = \frac{1}{2},$$

$$\frac{1}{x^3(x^2-1)} = -\frac{1}{x^3} + \frac{B}{x^2} + \frac{C}{x} + \frac{1}{2}\left(\frac{1}{x-1} + \frac{1}{x+1}\right).$$

Для нахождения B и C можно взять два любых значения x, но не 0 и не 1. Возьмем, например, $x = -2$ и $x = +2$,

$$x = -2 \Rightarrow -\frac{1}{24} = \frac{1}{8} + \frac{3}{4} - \frac{C}{2} + \frac{1}{2}\cdot\left(-\frac{1}{3} - 1\right),$$

$$x = 2 \Rightarrow \frac{1}{24} = -\frac{1}{8} + \frac{3}{4} + \frac{C}{2} + \frac{1}{2}\cdot\left(1 + \frac{1}{3}\right).$$

Следовательно, коэффициенты B и C будут найдены в результате решения системы двух линейных алгебраических уравнений с двумя неизвестными:

$$\begin{cases} B - 2C = 2, \\ B + 2C = -2. \end{cases}$$

Отсюда получаем $B = 0$, $C = -1$. Окончательное разложение подынтегральной функции на простейшие имеет вид

$$\frac{1}{x^3(x^2-1)} = -\frac{1}{x^3} - \frac{1}{x} + \frac{1}{2}\left(\frac{1}{x-1} + \frac{1}{x+2}\right).$$

Приступаем к интегрированию:

$$\int \frac{dx}{x^3(x^2-1)} = -\int \frac{dx}{x^3} - \int \frac{dx}{x} + \frac{1}{2}\int \frac{dx}{x-1} + \frac{1}{2}\int \frac{dx}{x+2} =$$

$$= \frac{1}{2x^2} - \ln|x| + \frac{1}{2}\ln|x-1| + \frac{1}{2}\ln|x+2| + \ln|C| =$$

$$= \frac{1}{2x^2} + \ln\left|\frac{C\sqrt{|(x-1)(x+2)|}}{x}\right|.$$

5. Найти интеграл $\int \frac{dx}{x^2(x^2-9)}$.

$$\frac{1}{x^2(x^2-9)} = |x^2 = t| = \frac{1}{t(t-9)} = \frac{A}{t} + \frac{B}{t-9}, \quad A = -\frac{1}{9}, \quad B = \frac{1}{9},$$

$$\frac{1}{x^2(x^2-9)} = \frac{1}{9}\left(-\frac{1}{x^2} + \frac{1}{x^2-9}\right),$$

$$\int \frac{dx}{x^2(x^2-9)} = -\frac{1}{9}\int \frac{dx}{x^2} + \frac{1}{9}\int \frac{dx}{x^2-9} =$$

$$= \frac{1}{9x} + \frac{1}{9} \cdot \frac{1}{6}\ln\left|\frac{3-x}{3+x}\right| + C = \frac{1}{9x} + \frac{1}{54}\ln\left|\frac{3-x}{3+x}\right| + C.$$

Использованы формулы №№ 1, 15* из таблицы основных неопределенных интегралов.

6. Найти интеграл $\int \frac{dx}{(x+1)(x^2+4)}$.

Поскольку в знаменателе имеется квадратный множитель, в разложении подынтегральной функции появится простейшая дробь третьего типа.

$$\frac{1}{(x+1)(x^2+4)} = \frac{A}{x+1} + \frac{Bx+C}{x^2+4}.$$

Коэффициенты A, B и C будем искать методом сравнения коэффициентов. Если сложить две дроби с неопределенными коэффициентами в правой части равенства, то получим дробь следующего вида:

$$\frac{(A+B)x^2+(C+B)x+(4A+C)}{(x+1)(x^2+4)}.$$

Если две дроби равны и у них равны знаменатели, то должны быть равны и числители. Это приводит нас к системе трех линейных алгебраических уравнений с тремя неизвестными A, B, C:

$$\begin{cases} A+B=0, \\ B+C=0, \\ 4A+C=1, \end{cases}$$

$$A=\frac{1}{5}, \quad B=-A=-\frac{1}{5}, \quad C=-B=A=\frac{1}{5}.$$

Итак,

$$\frac{1}{(x+1)(x^2+4)}=\frac{1}{5}\cdot\frac{1}{x+1}+\frac{1}{5}\cdot\frac{-x+1}{x^2+4}.$$

Переходим к интегрированию:

$$\int\frac{dx}{(x+1)(x^2+4)}=\frac{1}{5}\int\frac{dx}{x+1}-\frac{1}{5}\int\frac{x\,dx}{x^2+4}+\frac{1}{5}\int\frac{dx}{x^2+4}=$$
$$=\frac{1}{5}\ln|x+1|-\frac{1}{10}\ln(x^2+4)+\frac{1}{10}\operatorname{arctg}\frac{x}{2}+C.$$

7. Найти интеграл $\displaystyle\int\frac{x^4+3}{x^3+2x^2+5x}\,dx$.

Подынтегральная функция — неправильная несократимая дробь. Прежде всего нужно выделить из нее целую часть:

$$\begin{array}{r|l} x^4+3 & x^3+2x^2+5x \\ \cline{2-2} -\underline{x^4+2x^3+5x^2} & x-2 \\ -2x^3-5x^2+3 & \\ -\underline{-2x^3-4x^2-10x} & \\ -x^2+10x+3, & \end{array}$$

$$\frac{x^4+3}{x^3+2x^2+5x}=(x-2)+\frac{-x^2+10x+3}{x(x^2+2x+5)}.$$

Перейдем к интегрированию:

$$\int \frac{x^4+3}{x^3+2x^2+5x}\,dx = \int (x-2)\,dx + \int \frac{-x^2+10x+3}{x(x^2+2x+5)}\,dx.$$

Интеграл от целой части труда не представляет. Разложим на простейшие дроби подынтегральную функцию второго интеграла:

$$\frac{-x^2+10x+3}{x(x^2+2x+5)} = \frac{A}{x} + \frac{Bx+C}{x^2+2x+5}.$$

$$\begin{cases} -1 = A+B, \\ 10 = 2A+C, \\ 3 = 5A \end{cases} \quad \Rightarrow \quad A = \frac{3}{5}, \quad B = -\frac{8}{5}, \quad C = \frac{44}{5}.$$

$$\int \frac{-x^2+10x+3}{x(x^2+2x+5)}\,dx = \frac{3}{5}\int \frac{dx}{x} - \frac{1}{5}\int \frac{8x-44}{x^2+2x+5}\,dx = \frac{3}{5}\ln|x| -$$

$$-\left| \begin{array}{r} x^2+2x+5 = x^2+2x+1+4 = \\ = (x+1)^2+4 = t^2+4 \\ x+1=t \quad x=t-1 \quad dx=dt \end{array} \right| = -\frac{1}{5}\int \frac{8(t-1)-44}{t^2+4}\,dt =$$

$$= \frac{3}{5}\ln|x| - \frac{8}{5}\int \frac{t\,dt}{t^2+4} + \frac{52}{5}\int \frac{dt}{t^2+4} = \frac{3}{5}\ln|x| - \frac{4}{5}\ln(t^2+4) +$$

$$+\frac{26}{5}\operatorname{arctg}\frac{t}{2} + C = \frac{3}{5}\ln|x| - \frac{4}{5}\ln(x^2+2x+5) + \frac{26}{5}\operatorname{arctg}\frac{x+1}{2} + C.$$

Окончательный ответ:

$$\int \frac{x^4+3}{x^3+2x^2+5x}\,dx =$$

$$= \frac{x^2}{2} - 2x + \frac{3}{5}\ln|x| - \frac{4}{5}\ln(x^2+2x+5) + \frac{26}{5}\operatorname{arctg}\frac{x+1}{2} + C.$$

Самостоятельная работа

Найти интеграл:

1. $\displaystyle\int \frac{9x^2+20x-5}{(x-1)(x+2)(x+3)}\,dx.$ **2.** $\displaystyle\int \frac{2x^3-4x-5}{x^2-x-2}\,dx.$

3. $\displaystyle\int \frac{4x^2-3x-2}{x^3+x^2}\,dx.$ **4.** $\displaystyle\int \frac{x^3+2x^2+x+1}{x(x+1)^2}\,dx.$

5. $\displaystyle\int \frac{7x^2+13x+34}{(x-1)(x^2+4x+13)}\,dx.$ **6.** $\displaystyle\int \frac{x^4-3x^3+x^2+1}{x^2+2x+2}\,dx.$

7. $\displaystyle\int \frac{10x^3+16x^2+16x+4}{(x^2+1)(x^2+2x+2)}\,dx.$ **8.** $\displaystyle\int \frac{5x^3+6x^2-6x+3}{(x+1)^2(x^2-2x+2)}\,dx.$

ГЛАВА 5
ОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

§ 5.1. Понятие определенного интеграла

К понятию определенного интеграла приводят разнообразные задачи математики, физики, химии и других точных наук, в том числе вычисление площадей плоских фигур, длин дуг, объема произведенной работы, количества вещества, образовавшегося в результате химической реакции. Далее рассмотрим некоторые из этих задач более подробно.

5.1.1. Вычисление площади криволинейной трапеции

Криволинейной трапецией будем называть плоскую фигуру, ограниченную осью OX, графиком непрерывной функции $y = f(x)$ и двумя вертикальными прямыми $x = a$ и $x = b$. Ниже для удобства будем считать $f(x) \geqslant 0$.

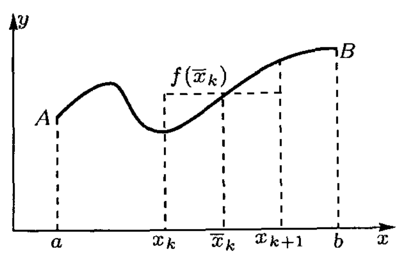


Рис. 5.1. Криволинейная трапеция.

Пример. Фигура $aABb$, изображенная на рис. 5.1, является криволинейной трапецией. Чтобы вычислить площадь криволинейной трапеции, проделаем следующие действия. Сначала разделим основание трапеции $[a, b]$ на n частичных интервалов $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$, считая что

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b. \tag{5.1}$$

Проведем через точки разбиения прямые, параллельные оси OY, тогда фигура $aABb$ разделится на n элементарных криволинейных трапеций. Обозначим $\Delta x_k = x_{k+1} - x_k$, $k = 0, 1, \ldots$ $\ldots, n-1$.

Вычислим площадь прямоугольника с основанием Δx_k и высотой $f(\overline{x}_k)$:

$$s_k = f(\overline{x}_k) \cdot \Delta x_k,$$

что приближенно равняется площади k-й элементарной криволинейной трапеции с тем же основанием (см. рис. 5.1). Учитывая, что площадь фигуры, составленной из нескольких непересекающихся фигур, равна сумме площадей этих фигур, получим

$$S_n = \sum_{k=0}^{n-1} f(\overline{x}_k)\Delta x_k. \tag{5.2}$$

Эта сумма является приближением для искомой площади, причем чем Δx_k меньше (а следовательно, n больше), тем это приближение точнее, т. е.

$$S = \text{площадь } aABb = \lim S_n, \tag{5.3}$$

где переход к пределу совершается при условии $\max\limits_{0 \leqslant k \leqslant n} \Delta x_k \to 0$.

5.1.2. Определение пути

Предположим, что материальная точка совершает поступательное движение по прямой линии, причем в любой момент времени $t \in [T_1, T_2]$ известна величина скорости $\nu = \nu(t)$. Так же как в (5.1), разобьем интервал $[T_1, T_2]$ точками $T_1 = t_0 < t_1 <$ $< \ldots < t_{n-1} < t_n = T_2$ на n непересекающихся промежутков.

Путь s, пройденный точкой за промежуток времени $[T_1, T_2]$, определим как

$$s = \lim_{\max\limits_{0 \leqslant k < n}(t_{k+1} - t_k) \to 0} \sum_{k=0}^{n-1} v(\overline{t}_k)(t_{k+1} - t_k), \tag{5.4}$$

где, так же как в (5.2), предполагается, что $t_k \leqslant \overline{t}_k \leqslant t_{k+1}$.

Такое определение пути выглядит вполне естественным, поскольку каждое слагаемое в (5.4) дает путь, который прошла бы точка за промежуток $[t_k, t_{k+1}]$, двигаясь с постоянной скоростью $v(\overline{t}_k)$.

5.1.3. Количество вещества, образовавшегося в результате химической реакции

Предположим, что в результате химической реакции образуется некоторое вещество. Скорость химической реакции $V = f(t)$, где t — время, предполагается известной.

Наша задача состоит в том, чтобы определить количество вещества, образовавшегося от момента времени $t = a$ до момента времени $t = b$, где $b > a$. Разобьем временной промежуток $[a, b]$ точками $t = t_0 = a$, $t = t_1 > t_0$, $t = t_2 > t_1, \ldots, t = t_n = b$. Таким образом, отрезок $[a, b]$ разбивается на n отрезков $[t_k, t_{k+1}]$, где $k = 0, 1, \ldots, n-1$.

Если отрезок $[t_k, t_{k+1}]$ достаточно мал, то можно пренебречь изменением скорости химической реакции на этом отрезке и приближенно считать, что скорость химической реакции на отрезке $[t_k, t_{k+1}]$ постоянна и равна $f(\overline{t_k})$, где $\overline{t_k}$ — любая точка отрезка $[t_k, t_{k+1}]$. Тогда масса вещества, образовавшегося в результате химической реакции за время от момента $t = t_k$ до момента $t = t_{k+1}$, приближенно находится по формуле

$$\Delta m_k \approx f(\overline{t_k}) \cdot \Delta t_k,$$

где $\Delta t_k = t_{k+1} - t_k$, а масса всего вещества, образовавшегося от момента времени $t = a$ до момента времени $t = b$, определяется приближенной формулой

$$m \approx \sum_{k=0}^{n-1} \Delta m_k = \sum_{k=0}^{n-1} f(\overline{t_k}) \cdot \Delta t_k.$$

Эта формула тем точнее, чем короче промежутки времени Δt_k. Точный результат получается из приближенного путем перехода к пределу при условии $\max \Delta t_k \to 0$, где $0 \leqslant k \leqslant n-1$, т. е.

$$m = \lim_{\max\limits_k \Delta t_k \to 0} \sum_{k=0}^{n-1} f(\overline{t_k}) \cdot \Delta t_k. \tag{5.5}$$

5.1.4. Определенный интеграл. Теорема существования

Рассмотренные примеры, если абстрагироваться от физического смысла переменных и их обозначений, приводят к одной и той же математической задаче: найти предел *интегральной суммы*

$$\mathrm{I}_n = \sum_{k=0}^{n-1} f(\overline{x}_k)\Delta x_k \tag{5.6}$$

при $\max\limits_{0\leqslant k<n} \Delta x_k \to 0$, где $f(x)$ — функция, непрерывная на промежутке $[a, b]$, а прочие обозначения такие же, как в (5.1) и (5.2).

Другими словами:

1) отрезок $[a, b]$ разбиваем точками $a = x_0 < x_1 < x_2 < \ldots$ $\ldots < x_{n-1} < x_n = b$ на интервалы $[x_k, x_{k+1}]$, $k = 0, 1, \ldots, n-1$;

2) берем произвольные точки $\overline{x}_k$ из каждого интервала $[x_k, x_{k+1}]$ и вычисляем $f(\overline{x}_k)$;

3) составляем произведения $f(\overline{x}_k)\Delta x_k$, где $\Delta x_k = x_{k+1} - x_k$;

4) вычисляем сумму этих произведений $\mathrm{I}_n = \sum\limits_{k=0}^{n-1} f(\overline{x}_k)\Delta x_k$;

5) находим предел I этой суммы при $\max\limits_{0\leqslant k<n} \Delta x_k \to 0$.

В общем случае такой предел называется *определенным интегралом* от функции $f(x)$ в пределах от a до b и обозначается

$$\mathrm{I} = \int\limits_a^b f(x)dx \tag{5.7}$$

(значок $\int$ напоминает букву S, первую букву слова summa).

В (5.7), как ранее в неопределенном интеграле, функция $f(x)$ называется подынтегральной функцией; $f(x)dx$ — подынтегральным выражением; x — переменной интегрирования; a — нижним, b — верхним пределами интеграла; $[a, b]$ — промежутком интегрирования.

Определение. Определенным интегралом называется предел, к которому стремится интегральная сумма (5.6) при стремлении к нулю длины наибольшего частичного интервала Δx_k.

Обращаем внимание, что в обозначениях (5.7) площадь криволинейной трапеции $S = \int\limits_a^b f(x)dx$, длина пути $s = \int\limits_{T_1}^{T_2} v(t)dt$, а масса вещества $m = \int\limits_a^b f(x)dx$ (см. (5.2), (5.4) и (5.7)). Отсюда, в частности, следует, что обозначение переменной интегрирования (т. е. x или t в данном случае) не играет роли и служит лишь интересам удобства.

Возникает естественный вопрос, для каких классов функций $f(x)$ предел (5.6), посредством которого определяется интеграл (5.7), существует независимо от способа разбиения интервала $[a, b]$ на частичные интервалы Δx_k и от выбора значений точек $\overline{x}_k$ внутри этих интервалов.

Ответом является следующая теорема, которую мы приводим без доказательства.

Теорема (существования определенного интеграла). Если функция $f(x)$ непрерывна в замкнутом интервале $[a, b]$, то ее n-я интегральная сумма стремится к пределу при стремлении к нулю наибольшего частичного интервала. Этот предел не зависит от способа разбиения интервала интегрирования на частичные интервалы и от выбора в них промежуточных точек.

Итак, в классе непрерывных функций интеграл $\int\limits_a^b f(x)dx$ корректно определен и всегда существует.

§ 5.2. Свойства определенного интеграла

Исходя из определения можно установить следующие свойства определенного интеграла.

1. Интеграл от суммы конечного числа функций равен сумме интегралов от слагаемых:

$$\int\limits_a^b (f_1(x) + \ldots + f_m(x))\,dx = \int\limits_a^b f_1(x)\,dx + \ldots + \int\limits_a^b f_m(x)\,dx.$$

2. Постоянный множитель подынтегральной функции можно выносить за символ интеграла:

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx.$$

В частности, $\int_a^b 0\,dx = 0$.

3. Если верхний и нижний пределы интеграла поменять местами, то интеграл изменит знак:

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

В частности, $\int_a^a f(x)\,dx = 0$.

4. Если интервал интегрирования $[a, b]$ разбить на две части $[a, c]$ и $[c, b]$, то

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

(свойство аддитивности определенного интеграла).

5. Если подынтегральная функция в интервале интегрирования не меняет знака, то интеграл представляет собой число того же знака, что и функция. В частности, если $f(x) \geqslant 0$ в интервале $[a, b]$, где $a < b$, то

$$\int_a^b f(x)\,dx \geqslant 0.$$

6. Геометрический смысл определенного интеграла.

В соответствии с разделом 5.1.1. То есть $\int_a^b f(x)\,dx$ равен площади под графиком $f(x) \geqslant 0$. В случае, если $f(x)$ на $[a, b]$ имеет разные знаки (см. рис. 5.2), общая площадь будет вычисляться

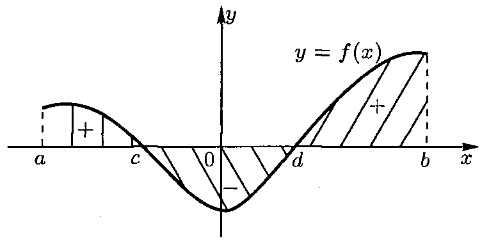


Рис. 5.2. Геометрический смысл определенного интеграла.

по формуле $S = S_1 - S_2 + S_3$, или $S = \int\limits_a^c f(x)dx - \int\limits_c^d f(x)dx + \int\limits_d^b f(x)dx = \int\limits_a^b |f(x)|\,dx$.

7. Если в интервале $[a, b]$, где $a < b$, $f_1(x) \leqslant f(x) \leqslant f_2(x)$, то

$$\int\limits_a^b f_1(x)\,dx \leqslant \int\limits_a^b f(x)\,dx \leqslant \int\limits_a^b f_2(x)\,dx, \tag{5.8}$$

т. е. неравенства можно интегрировать.

В частности, если $m \leqslant f(x) \leqslant M$ на $[a, b]$, то

$$m(b-a) \leqslant \int\limits_a^b f(x)\,dx \leqslant M(b-a). \tag{5.9}$$

Заметим, что в (5.9) можно положить $m = \min\limits_{[a,b]} f(x)$ и $M = \max\limits_{[a,b]} f(x)$.

Доказательство соотношения (5.9). Из того что $f(x) \geqslant m$, следует неравенство $f(x) - m \geqslant 0$, откуда по свойствам 5, 1 и 2

$$0 \leqslant \int\limits_a^b (f(x) - m)\,dx = \int\limits_a^b f(x)dx + \int\limits_a^b (-m)\,dx =$$

$$= \int\limits_a^b f(x)\,dx - m\int\limits_a^b dx,$$

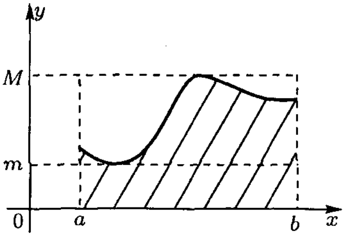


Рис. 5.3. Геометрический смысл оценки (5.9).

или

$$m\int_a^b dx = m(b-a) \leqslant \int_a^b f(x)\,dx.$$

Оценка сверху проверяется аналогично.

Соотношение (5.9) имеет следующую геометрическую интерпретацию: площадь криволинейной трапеции заключена между площадями прямоугольников, каждый из которых имеет основание, равное основанию трапеции, и высоты, равные соответственно наименьшей и наибольшей ординате трапеции (рис. 5.3).

Теорема (о среднем). Внутри интервала $[a, b]$ существует по крайней мере одно значение $x = c$, при котором

$$\int_a^b f(x)\,dx = f(c)(b-a). \tag{5.10}$$

Доказательство. Из соотношения (5.9) при $m = \min\limits_{[a,b)} f(x)$ и $M = \max\limits_{[a,b)} f(x)$ следует, что $\dfrac{1}{b-a}\displaystyle\int_a^b f(x)dx = d$, где d — некоторое число из интервала $[m, M]$. Известно, однако, что каждая непрерывная функция $f(x)$, $a \leqslant x \leqslant b$, обязательно, по крайней мере единожды, принимает любое значение между m и M, т. е. $f(c) = d$ при некотором c, $a \leqslant x \leqslant b$.

Определение. Величина $f(c)$ называется средним значением функции $f(x)$ на интервале $[a, b]$.

Геометрически теорема о среднем означает, что площадь криволинейной трапеции равна площади прямоугольника, основание которого совпадает с основанием трапеции, а высота равна значению ординаты кривой, ограничивающей эту трапецию, в некоторой промежуточной точке.

Термин «среднее значение» можно обосновать следующим образом. Пусть функция $f(x)$ задана и непрерывна на интервале $[a, b]$. Разобьем этот интервал на n равных частичных интервалов и возьмем значения $f(x)$ в серединах этих интервалов — точках $x_1, \ldots, x_n$. Затем составим среднее арифметическое указанных значений $\overline{f_n} = \dfrac{f(x_1) + \ldots + f(x_n)}{n}$. Имеем: $\overline{f_n} = = \dfrac{1}{b-a} \sum\limits_{k=1}^{n} f(x_k)\Delta x_k$, где $\Delta x_k = \dfrac{b-a}{n}$. В соответствии с (5.6), (5.7) $\lim\limits_{n\to\infty} \overline{f_n} = \dfrac{1}{b-a} \lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} f(x_n)\Delta x_n = \dfrac{1}{b-a} \int\limits_a^b f(x)\,dx$. Следовательно, вполне логично назвать (см. (5.10)) $f(c) = \dfrac{1}{b-a} \times \times \int\limits_a^b f(x)\,dx$ средним значением $f(x)$ на $[a, b]$, как это и было сделано.

Продолжим изучение свойств определенного интеграла.

Для $a \leqslant x \leqslant b$ положим $\mathrm{I}(x) = \int\limits_a^x f(x)\,dx$. Таким образом, нами введена функция верхнего предела интеграла, заданная на интервале $[a, b]$.

Более строго было бы употреблять запись $\mathrm{I}(x) = \int\limits_a^x f(u)\,du$, однако достаточно просто помнить о различном смысле переменных интегрирования и верхнего предела.

8. Производная интеграла по его верхнему пределу равна подынтегральной функции, вычисленной на верхнем пределе:

$$\mathrm{I}'(x) = \left(\int\limits_a^x f(x)\,dx\right)' = f(x), \quad a \leqslant x \leqslant b. \tag{5.11}$$

Другими словами, интеграл с переменным верхним пределом суть первообразная для подынтегральной функции.

Доказательство. Из свойств 4 и 8 (аддитивность и теорема о среднем) следует

$$\mathrm{I}'(x) = \lim_{\Delta x\to 0} \frac{\mathrm{I}(x+\Delta x) - \mathrm{I}(x)}{\Delta x} =$$
$$= \lim_{\Delta x\to 0} \frac{1}{\Delta x} \int\limits_x^{x+\Delta x} f(x)\,dx = \lim_{\Delta x\to 0} f(c),$$

но $\lim\limits_{\Delta x \to 0} f(c) = \lim\limits_{c \to x} f(c) = f(\lim\limits_{c \to x} c) = f(x)$, поскольку $f(x)$ непрерывна на $[a, b]$, а $c \in (x, x + \Delta x)$, и, следовательно, $c \to$ $\to x$ при $\Delta x \to 0$.

§ 5.3. Формула Ньютона–Лейбница

Непосредственное вычисление интеграла как предела соответствующих интегральных сумм (см. (5.6)) затруднительно, да и не требуется, поскольку для этой цели можно воспользоваться следующей теоремой.

Теорема. Значение определенного интеграла равно разности значений любой первообразной от подынтегральной функции, взятых при верхнем и нижнем пределах интеграла:

$$\int\limits_a^b f(x)\,dx = F(b) - F(a), \text{ где } F'(x) = f(x). \tag{5.12}$$

Равенство (5.12) называется *формулой Ньютона–Лейбница.*

Доказательство. В соответствии с равенством (5.11) функция $\mathrm{I}(x) = \int\limits_a^x f(x)\,dx$ является первообразной от функции $f(x)$ и, поэтому, имеет вид $F(x) + C$, где $F(x)$ — некоторая первообразная от $f(x)$, а C — определенная постоянная, т. е. $\mathrm{I}(x) = F(x) + C$. По свойству 3

$$\mathrm{I}(a) = \int\limits_a^a f(x)\,dx = 0,$$

откуда $\mathrm{I}(a) = 0 = F(a) + C$, или $C = -F(a)$. Другими словами, $\mathrm{I}(x) = F(x) - F(a)$, $a \leqslant x \leqslant b$.

При $x = b$ отсюда получаем (5.12). Теорема доказана.

Разность значений функции часто записывают так:

$$F(b) - F(a) = F(x)\Big|_a^b. \tag{5.13}$$

В случае использования обозначения (5.13) формуле Ньютона–Лейбница можно придать вид

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b, \text{ где } F'(x) = f(x). \tag{5.14}$$

Формула Ньютона–Лейбница дает нам альтернативный способ вычисления определенных интегралов. Она позволяет находить их по формуле

$$\int_a^b f(x)\,dx = \int f(x)\,dx\Big|_a^b. \tag{5.15}$$

В качестве иллюстрации приведем несколько примеров:

$$\int_1^2 x^3\,dx = \frac{1}{4}x^4\Big|_1^2 = \frac{16}{4} - \frac{1}{4} = \frac{15}{4},$$

$$\int_a^b \frac{1}{x}\,dx = \ln|x|\Big|_a^b = \ln|b| - \ln|a| = \ln\left|\frac{b}{a}\right|,$$

$$\int_0^{\pi/2} \cos x\,dx = \sin x\Big|_0^{\pi/2} = 1.$$

§ 5.4. Методы вычисления определенных интегралов

5.4.1. Метод замены переменной

Если в интервале $[\alpha, \beta]$ функции $\Psi(u)$, $\Psi'(u)$ и $f(\Psi(u))$ непрерывны и $\Psi(\alpha) = a$, $\Psi(\beta) = b$, то

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(\Psi(u))\Psi'(u)\,du. \tag{5.16}$$

Равенство (5.16) следует из соотношения (5.15) и формулы замены переменной в неопределенном интеграле (принимается $x = \Psi(u)$). Отметим, что новые пределы интегрирования α и β являются корнями уравнений $a = \Psi(u)$ и $b = \Psi(u)$ соответственно. При этом желательно (хотя и не обязательно) в случае, когда уравнения $a = \Psi(u)$, $b = \Psi(u)$ решаются неоднозначно, брать наименьший интервал изменения u. Например, если $\Psi(u) = \sin u$, $a = 0$, $b = 1/2$, то предпочтительнее выбор $\alpha = 0$, $\beta = \pi/6$, а не $\alpha = 0$, $\beta = 5/6\pi$, хотя в обоих случаях $\sin\beta = 1/2$ (см. пример 1 ниже).

Часто замена переменной в интеграле производится по формуле $u = g(x)$, выражающей новую переменную через заданную (а не наоборот, как в (5.16)). Тогда новые пределы α и β вычисляются непосредственно по формулам $\alpha = g(a)$, $\beta = g(b)$. Заметим, что при такой замене переменных требуется, чтобы функция $g(x)$ была обратимой в интервале $[a, b]$: для этого достаточно, чтобы $g'(x) \neq 0$ при всех $x \in (a, b)$.

Примеры.

1. $$\int\limits_0^{1/2} \frac{x^2\,dx}{\sqrt{1-x^2}} = \left| \begin{array}{l} x = \sin u,\ 0 = \sin 0, \\ \dfrac{1}{2} = \sin\dfrac{\pi}{6},\ dx = \cos u\,du \end{array} \right| = \int\limits_0^{\pi/6} \sin^2 u\,du =$$
$$= \frac{1}{2}\int\limits_0^{\pi/6} (1-\cos 2u)\,du = \frac{1}{2}\Big(u - \frac{\sin 2u}{2}\Big)\Big|_0^{\pi/6} = \frac{\pi}{12} - \frac{\sqrt{3}}{8}.$$

2. $$\int\limits_0^{\pi/2} \frac{\sin x}{1+\cos^2 x}\,dx = \Big|\cos x = u, -\sin x\,dx = du\Big| =$$
$$= -\int\limits_1^0 \frac{du}{1+u^2} = -\arcsin u\Big|_1^0 = \frac{\pi}{4}.$$

3. $$\mathrm{I} = \int\limits_{-a}^{a} f(x)\,dx = \int\limits_0^a (f(x)+f(-x))\,dx.$$

Действительно,
$$\mathrm{I} = \mathrm{I}_+ + \mathrm{I}_-,$$

где

$$\mathrm{I}_+ = \int\limits_0^a f(x)\,dx, \qquad \mathrm{I}_- = \int\limits_{-a}^0 f(x)\,dx,$$

но

$$\mathrm{I}_- = \left| \begin{array}{l} x = -u, \\ dx = -du \end{array} \right| = -\int\limits_a^0 f(-u)\,du = \int\limits_0^a f(-u)\,du = \int\limits_0^a f(-x)\,dx.$$

Следствием доказанного равенства является полезная формула

$$\int\limits_{-a}^a f(x)dx = \begin{cases} 0, & \text{если } f(x) \text{ нечетная функция,} \\ 2\int\limits_0^a f(x)\,dx, & \text{если } f(x) \text{ четная функция,} \end{cases} \tag{5.17}$$

позволяющая находить значения некоторых определенных интегралов без вычислений.

4. $\int\limits_{-1}^1 x^5 e^{x^2}\,dx = 0$, так как $x^5 e^{x^2}$ — нечетная функция.

5. $$\int\limits_4^9 \frac{dx}{\sqrt{x}-1} = \left| \begin{array}{l} x = u^2,\ dx = 2u\,du \\ 4 = (2)^2,\ 9 = (3)^2 \end{array} \right| = \int\limits_2^3 \frac{2u\,du}{u-1} =$$

$$= 2\int\limits_2^3 \left(1 + \frac{1}{u-1}\right) du = 2(u + \ln(u-1))\Big|_2^3 = 2(1 + \ln 2).$$

6. $$\int\limits_{\pi/4}^{\pi/3} \frac{1+\mathrm{tg}^2 x}{(1+\mathrm{tg}\,x)^2}\,dx = \left| \begin{array}{l} \mathrm{tg}\,x = t,\ (1+\mathrm{tg}^2 x)\,dx = dt \\ \mathrm{tg}\,\frac{\pi}{4} = 1, \mathrm{tg}\,\frac{\pi}{3} = \sqrt{3} \end{array} \right| =$$

$$= \int\limits_1^{\sqrt{3}} \frac{dt}{(1+t)^2} = -\frac{1}{1+t}\Big|_1^{\sqrt{3}} = -\frac{1}{1+\sqrt{3}} + \frac{1}{2} = 1 - \frac{\sqrt{3}}{2}.$$

7. $$\int\limits_1^{\sqrt{3}} \frac{dx}{\sqrt{(1+x^2)^3}} = \left| \begin{array}{l} x = \operatorname{tg} u, dx = (1+\operatorname{tg}^2 u)\,du, \\ 1 = \operatorname{tg}\frac{\pi}{4}, \sqrt{3} = \operatorname{tg}\frac{\pi}{3} \end{array} \right| =$$

$$= \int\limits_{\pi/4}^{\pi/3} \frac{du}{(1+\operatorname{tg}^2 u)^{1/2}} = \int\limits_{\pi/4}^{\pi/3} \cos u\,du = \sin u\Big|_{\pi/4}^{\pi/3} = \frac{\sqrt{3}-\sqrt{2}}{2}.$$

8. $$\int\limits_0^{\pi} \sin^6\frac{x}{2}\,dx = |x = 2u| = \int\limits_0^{\pi/2} \sin^6 u\,du = \mathrm{I}_6 = \frac{5}{16}\pi$$

(см. (5.19), (5.20)).

9. $$\int\limits_{\pi/2}^{\pi/6} \sqrt{1-\sin^2 t}\cos t\,dt = \int\limits_{\pi/2}^{\pi/6} \frac{1+\cos 2t}{2}\,dt = \frac{t}{2} + \frac{\sin 2t}{4}\Big|_{\pi/2}^{\pi/6} =$$

$$= \frac{\sqrt{3}}{8} - \frac{\pi}{6}.$$

10. $$\int\limits_0^{\pi/2} \sin^3 x\cos^2 x\,dx = \left| \begin{array}{l} \cos x = u \\ -\sin x\,dx = du \end{array} \right| =$$

$$= -\int\limits_1^0 (1-u^2)u^2\,du = \frac{u^5}{5} - \frac{u^3}{3}\Big|_1^0 = \frac{2}{15}.$$

Самостоятельная работа

1. $\displaystyle\int\limits_0^1 \frac{dx}{e^x+1} = \ln\frac{2e}{1+e}$. **2.** $\displaystyle\int\limits_0^{1/2} \sqrt{\frac{x}{1-x}}\,dx = \frac{\pi-2}{4}$.

3. $\displaystyle\int\limits_0^{\pi/2} \sin\cos^2 x\,dx = \frac{1}{3}$. **4.** $\displaystyle\int\limits_0^2 x^2\sqrt{4-x^2}\,dx = \pi$.

5. $\int\limits_0^1 \sqrt{1+x^2}\,dx = \frac{\sqrt{2}+\ln(1+\sqrt{2})}{2}$. **6.** $\int\limits_0^{\pi/4} \operatorname{tg}^3 x\,dx = \frac{1-\ln 2}{2}$.

7. $\int\limits_1^{\sqrt{2}} \sqrt{2-x^2}\,dx = \frac{\pi-2}{2}$. **8.** $\int\limits_0^1 \frac{e^x\,dx}{1+e^{2x}} = \operatorname{arctg} e - \frac{\pi}{4}$.

9. $\int\limits_0^{\sqrt{3}} \frac{x\,dx}{\sqrt{4-x^2}} = 1$. **10.** $\int\limits_1^4 \frac{dx}{(1+\sqrt{x})^2} = 2\ln\frac{3}{2} - \frac{1}{3}$.

11. $\int\limits_1^5 \frac{x\,dx}{\sqrt{4x+5}} = \frac{17}{6}$. **12.** $\int\limits_{\pi/2}^{2\pi/3} \frac{dx}{\sin x} = \frac{1}{2}\ln 6$.

13. $\int\limits_{\ln 2}^{\ln 3} \frac{e^x+1}{e^x-1}\,dx = 3\ln 2 - \ln 3$. **14.** $\int\limits_1^2 \frac{\sqrt{x^2-1}}{x}\,dx = \sqrt{3} - \frac{\pi}{3}$.

15. $\int\limits_0^1 \frac{(x-3)\,dx}{(x^2-6x+10)^2} = -\frac{3}{200}$. **16.** $\int\limits_0^1 \frac{dx}{\sqrt{x}+\sqrt[4]{x}} = 4\ln 2 - 2$.

17. $\int\limits_0^{\ln 3} \frac{dx}{\sqrt{e^x+1}} = \ln(1+\frac{2}{3}\sqrt{2})$. **18.** $\int\limits_0^{\ln 2} \frac{dx}{e^x(3+e^{-x})} = \ln\frac{8}{7}$.

19. $\int\limits_{\pi/4}^{\pi/3} \frac{\ln \operatorname{tg} x}{\sin x \cdot \cos x}\,dx = \frac{1}{8}(\ln 3)^2$. **20.** $\int\limits_1^4 \frac{e^{\sqrt{x}}}{\sqrt{x}}\,dx = 2e(e-1)$.

5.4.2. Метод интегрирования по частям

Справедливо равенство

$$\int\limits_a^b u\,d\nu = u\nu\Big|_a^b - \int\limits_a^b \nu\,du, \tag{5.18}$$

где u и ν — функции независимой переменной.

Доказательство. Имеем по (5.15)

$$\int_a^b u d\nu = \int u d\nu \Big|_\alpha^b = \left(u\nu - \int \nu du\right)\Big|_a^b,$$

откуда следует (5.18).

В качестве примера использования формулы (5.18) вычислим интеграл $\mathrm{I}_n = \int_0^{\pi/2} \sin^n x \, dx$, где n — натуральное число.

При $n = 0$ или 1 интеграл легко берется:

$$\mathrm{I}_0 = \int_0^{\pi/2} dx = \frac{\pi}{2}, \quad \mathrm{I}_1 = \int_0^{\pi/2} \sin x \, dx = -\cos x \Big|_0^{\pi/2} = 1. \qquad (5.19)$$

Пусть $n \geqslant 2$. Положим $d\nu = \sin x \, dx$, $u = \sin^{n-1} x$. Тогда $\nu = -\cos x$, $du = (n-1)\sin^{n-2} x \cdot \cos x \, dx$,

$$\mathrm{I}_n = -\cos x \cdot \sin^{n-1} x \Big|_0^{\pi/2} + (n-1) \int_0^{\pi/2} \sin^{n-2} x \cos^2 x \, dx.$$

Первое слагаемое здесь равно нулю. Заменяя во втором слагаемом множитель $\cos^2 x$ на $1 - \sin^2 x$, получим

$$\mathrm{I}_n = (n-1) \int_0^{\pi/2} \sin^{n-2} x \, dx - (n-1) \int_0^{\pi/2} \sin^n x \, dx =$$
$$= (n-1)\mathrm{I}_{n-2} - (n-1)\mathrm{I}_n,$$

откуда следует, что

$$\mathrm{I}_n = \frac{n-1}{n} \mathrm{I}_{n-2}. \qquad (5.20)$$

Пользуясь этой рекуррентной формулой, мы в состоянии понизить показатель степени n до 0 (если n — четное число) или до 1 (при нечетном n). С учетом (5.19) задача решается до конца. Так,

$$\mathrm{I}_3 = \frac{2}{3}\,\mathrm{I}_1 = \frac{2}{3}, \quad \mathrm{I}_4 = \frac{3}{4}\,\mathrm{I}_2 = \frac{3}{4}\cdot\frac{1}{2}\,\mathrm{I}_0 = \frac{3}{16}\,\pi.$$

Примеры.

1. $$\int_0^1 (x-1)\cos\pi x\,dx = \left|\begin{array}{l} u = x-1;\, du = dx \\ d\nu = \cos\pi x\,dx;\, v = \dfrac{1}{\pi}\sin\pi x \end{array}\right| =$$

$$= (x-1)\,\frac{\sin\pi x}{\pi}\bigg|_0^1 - \frac{1}{\pi}\int_0^1 \sin\pi x\,dx = 0 + \frac{1}{\pi^2}\cos\pi x\bigg|_0^1 =$$

$$= \frac{1}{\pi^2}(\cos\pi - \cos 0) = -\frac{2}{\pi^2}.$$

2. $$\int_1^{e^2} \frac{\ln x}{x^2}\,dx = \left|\begin{array}{l} u = \ln x;\, du = \dfrac{1}{x}\,dx \\ d\nu = \dfrac{dx}{x^2};\, \nu = -\dfrac{1}{x} \end{array}\right| =$$

$$= -\frac{1}{x}\ln x\bigg|_1^{e^2} + \int_1^{e^2} \frac{1}{x^2}\,dx = -\frac{1}{e^2}\ln e^2 - \frac{1}{x}\bigg|_1^{e^2} = 1 - \frac{3}{e^2}.$$

3. $$\int_0^{\pi/2} x\sin x\,dx = \left|\begin{array}{l} u = x;\, du = dx \\ d\nu = \sin x\,dx;\, \nu = -\cos x \end{array}\right| =$$

$$= -x\cos x\bigg|_0^{\pi/2} + \int_0^{\pi/2} \cos x\,dx = 0 + \sin x\bigg|_0^{\pi/2} = 1.$$

4. $$\int_0^1 x^3 e^{x^2}\,dx = \left|\begin{array}{l} u = \dfrac{1}{2}x^2;\, du = x\,dx \\ d\nu = 2xe^{x^2};\, \nu = e^{x^2} \end{array}\right| =$$

$$= \frac{x^2}{2}e^{x^2}\bigg|_0^1 - \int_0^1 xe^{x^2}\,dx = \frac{1}{2}e - \frac{1}{2}e^{x^2}\bigg|_0^1 = \frac{1}{2}.$$

5. $$\int_0^{\pi/2} x^2\cos x\,dx = \left|\begin{array}{l} u = x^2;\, du = 2x\,dx \\ dv = \cos x\,dx;\, v = \sin x \end{array}\right| =$$

$$= x^2 \sin x \Big|_0^{\pi/2} - 2 \int_0^{\pi/2} x \sin x \, dx = \frac{\pi^2}{4} - 2,$$

(см. пример 3).

6. $$\mathrm{I} = \int_0^{\pi/2} e^x \cos x \, dx = \left| \begin{array}{l} u = \cos x;\, du = -\sin x \, dx \\ d\nu = e^x \, dx;\, \nu = e^x \end{array} \right| =$$

$$= \cos x e^x \Big|_0^{\pi/2} + \int_0^{\pi/2} e^x \sin x \, dx = -1 + J,$$

где

$$J = \int_0^{\pi/2} e^x \sin x \, dx = \left| \begin{array}{l} u = \sin x;\, du = \cos x \, dx \\ d\nu = e^x dx;\, \nu = e^x \end{array} \right| =$$

$$= \sin x e^x \Big|_0^{\pi/2} - \int_0^{\pi/2} e^x \cos x \, dx = e^{\pi/2} - \mathrm{I}.$$

Отсюда $\mathrm{I} = J - 1 = e^{\pi/2} - 1 - \mathrm{I}$, т. е.

$$\mathrm{I} = \frac{1}{2}(e^{\pi/2} - 1).$$

7. $$\int_{\pi/4}^{\pi/3} \frac{x}{\sin^2 x} \, dx = \left| \begin{array}{l} u = x;\, du = dx \\ d\nu = \dfrac{dx}{\sin^2 x};\, \nu = -\operatorname{ctg} x \end{array} \right| = -x \operatorname{ctg} x \Big|_{\pi/4}^{\pi/3} +$$

$$+ \int_{\pi/4}^{\pi/3} \operatorname{ctg} x \, dx = -\frac{\pi}{3} \cdot \frac{\sqrt{3}}{3} + \frac{\pi}{4} \cdot 1 + \ln \sin x \Big|_{\pi/4}^{\pi/3} = \pi \left(\frac{1}{4} - \frac{\sqrt{3}}{9} \right) + \frac{1}{2} \ln \frac{3}{2}.$$

8. $$\int_0^1 x \operatorname{arctg} x \, dx = \left| \begin{array}{l} u = \operatorname{arctg} x;\, du = \dfrac{dx}{1+x^2} \\ d\nu = x \, dx;\, \nu = \dfrac{1}{2} x^2 \end{array} \right| =$$

$$= \frac{1}{2} x^2 \operatorname{arctg} x \Big|_0^1 - \frac{1}{2} \int_0^1 \frac{x^2}{1+x^2} \, dx = \frac{1}{2} \cdot \frac{\pi}{4} - \frac{1}{2} \int_0^1 \left(1 - \frac{1}{1+x^2} \right) dx =$$

$$= \frac{\pi}{8} - \frac{1}{2}(x - \operatorname{arctg} x)\Big|_0^1 = \frac{\pi}{4} - \frac{1}{2}.$$

9. $\mathrm{I} = \int\limits_1^e \sin \ln x \, dx = \left| \begin{array}{l} u = \sin \ln x;\, du = \frac{1}{x} \cos \ln x \, dx \\ d\nu = dx;\, \nu = x \end{array} \right| =$

$$= x \sin \ln x \Big|_1^e - \int\limits_1^e \cos \ln x \, dx = e \sin 1 - J,$$

где

$J = \left| \begin{array}{l} u = \cos \ln x;\, du = -\frac{1}{x} \sin \ln x \, dx \\ d\nu = dx;\, \nu = x \end{array} \right| =$

$$= x \cos \ln x \Big|_1^e + \int\limits_1^e \sin \ln x \, dx = e \cos 1 - 1 + \mathrm{I}.$$

Отсюда $\mathrm{I} = e \sin 1 - J = 1 + e \sin 1 - e \cos 1 - \mathrm{I}$, т. е.

$$\mathrm{I} = \frac{1}{2}(1 + e \sin 1 - e \cos 1).$$

10. $\int\limits_0^1 \ln(x^2 + 1) \, dx = \left| \begin{array}{l} u = \ln(x^2 + 1);\, du = \frac{2x}{x^2 + 1} \\ d\nu = dx;\, \nu = x \end{array} \right| =$

$$= x \ln(x^2 + 1)\Big|_0^1 - 2 \int\limits_0^1 \frac{x^2}{x^2 + 1} \, dx = \ln 2 - 2\left(1 - \frac{\pi}{4}\right)$$

(см. пример 8).

Самостоятельная работа

1. $\int\limits_0^1 x e^{-x} \, dx = 1 - 2e^{-1}$. **2.** $\int\limits_0^{e-1} \ln(x + 1) \, dx = 1$.

3. $\int\limits_1^2 x \log_2 x \, dx = 2 - \frac{3}{4 \ln 2}$. **4.** $\int\limits_1^e \ln^3 x \, dx = 6 - 2e$.

5. $\displaystyle\int\limits_0^{1/2} \arccos x\,dx = 1 - \frac{\sqrt{3}}{2} + \frac{\pi}{6}.$

6. $\displaystyle\int\limits_0^{\pi/4} x\cos^2 x\,dx = \frac{\pi}{16} + \frac{\pi^2}{64} - \frac{1}{8}.$

7. $\displaystyle\int\limits_2^{e+1} x\ln(x-1)\,dx = \frac{(e+1)^2}{4} + 1 - \frac{e}{2}.$

8. $\displaystyle\int\limits_0^1 \frac{\arcsin x\,dx}{\sqrt{1+x}} = \pi\sqrt{2} - 4.$ **9.** $\displaystyle\int\limits_0^1 (\arcsin x)^2\,dx = \frac{\pi^2}{4} - 2.$

10. $\displaystyle\int\limits_1^e \frac{\ln^3 x}{x^2}\,dx = 6 - \frac{16}{e}.$

11. $\displaystyle\int\limits_0^1 (\operatorname{arctg} x)^2\,dx = \frac{1}{2}\ln 2 - \frac{\pi}{4} + \frac{\pi^2}{16}.$

12. $\displaystyle\int\limits_0^1 x^2\ln(1+x)\,dx = \frac{1}{3}\left(2\ln 2 - \frac{5}{6}\right).$

13. $\displaystyle\int\limits_0^{\pi/4} x\operatorname{tg}^2 x\,dx = \frac{\pi}{4} - \frac{\pi^2}{16} - \frac{1}{2}\ln 2.$

14. $\displaystyle\int\limits_0^1 \operatorname{arctg}\sqrt{x}\,dx = \frac{\pi}{2} - 1.$

15. $\displaystyle\int\limits_0^{1/4} \frac{\arcsin\sqrt{x}}{\sqrt{1-x}}\,dx = 1 - \frac{\pi\sqrt{3}}{6}.$ **16.** $\displaystyle\int\limits_1^{\sqrt{2}} \frac{\lg x}{x^3}\,dx = \frac{1}{8}\lg\frac{e}{2}.$

17. $\displaystyle\int\limits_0^1 x^3 e^x\,dx = 6 - 2e.$ **18.** $\displaystyle\int\limits_0^2 \frac{x^2 e^x\,dx}{(x+2)^2} = 2.$

§ 5.5. Несобственные интегралы

5.5.1. Несобственные интегралы с бесконечными пределами интегрирования

Понятие определенного интеграла было установлено для функции $f(x)$, непрерывной на конечном интервале $[a, b]$. Пусть теперь функция $f(x)$ непрерывна на бесконечном интервале $[a, \infty)$.

Определение. *Несобственным интегралом* от функции $f(x)$ на интервале $[a, +\infty)$ называется предел интеграла $\int\limits_a^b f(x)\,dx$ при $b \to +\infty$, т. е. $\int\limits_a^{+\infty} f(x)\,dx = \lim\limits_{b \to +\infty} \int\limits_a^b f(x)\,dx$.

Если этот предел существует, то несобственный интеграл называется *сходящимся*, а если не существует (т. е. стремится к бесконечности или колеблется), то *расходящимся*.

Аналогично определяются несобственные интегралы и для прочих бесконечных интервалов $(-\infty, a]$ и $(-\infty, \infty)$.

Обозначив через $F(x)$ первообразную от функции $f(x)$, условно можно записать

$$\int\limits_a^{+\infty} f(x)\,dx = F(+\infty) - F(a), \qquad \int\limits_{-\infty}^{a} f(x)\,dx = F(a) - F(-\infty),$$

$$\int\limits_{-\infty}^{\infty} f(x)\,dx = F(+\infty) - F(-\infty),$$

понимая под символом $F(\pm\infty)$ предел, к которому стремится $F(x)$ при $x \to \pm\infty$.

Примеры.

1. $\int\limits_a^{\infty} \frac{1}{x^2}\,dx = -\left.\frac{1}{x}\right|_a^{\infty} = \frac{1}{a} - \frac{1}{\infty} = \frac{1}{a} - 0 = \frac{1}{a}.$

2. $\int\limits_{1}^{+\infty} \frac{1}{x}\,dx = \ln x\Big|_{1}^{+\infty} = \ln(+\infty) = \infty$ — интеграл расходится.

Заметим, что несобственный интеграл можно, как и ранее, интерпретировать как площадь соответствующей бесконечной криволинейной трапеции. Таким образом, в первом примере пло-

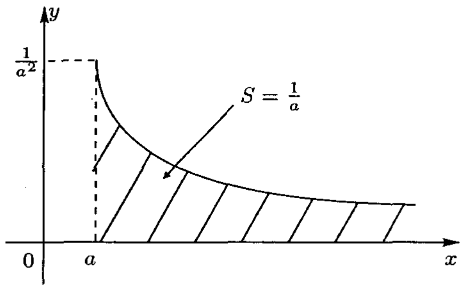


Рис. 5.4. Площадь бесконечной криволинейной трапеции.

щадь фигуры, расположенная между кривой $y = 1/x^2$ и осью OX на интервале $[a, \infty)$ (рис. 5.4), равна $1/a$, а во втором — равна бесконечности.

3. $\int\limits_{-\infty}^{\infty} \frac{dx}{1+x^2} = \operatorname{arctg} x\Big|_{-\infty}^{+\infty} = \operatorname{arctg}(+\infty) - \operatorname{arctg}(-\infty) =$

$$= \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi.$$

4. $\int\limits_{0}^{\infty} \sin x\,dx = -\cos x\Big|_{0}^{\infty}$ не имеет предела (т. е. интеграл расходится).

Самостоятельная работа

Установить сходимость или расходимость несобственных интегралов.

1. $\int\limits_{0}^{\infty} \frac{2x\,dx}{1+x^2}$. **2.** $\int\limits_{0}^{\infty} e^{-ax}dx \quad (a > 0)$. **3.** $\int\limits_{0}^{\infty} x\sin x\,dx$.

4. $\dfrac{1}{\sqrt{2\pi}}\int\limits_{a}^{\infty} e^{-x^2/2}\,dx.$ **5.** $\int\limits_{-\infty}^{0} \dfrac{dx}{1-x^3}.$ **6.** $\int\limits_{e^2}^{\infty} \dfrac{dx}{\ln\ln x}.$

5.5.2. Интегралы от разрывных функций

Если в интервале (конечном или бесконечном) интегрирования, функция $f(x)$ имеет некоторое число разрывов I рода, то интеграл определяется как сумма интегралов по частичным интервалам, на которые точки разрыва разбивают основной промежуток интегрирования.

Другими словами, если функция $f(x)$ определена в интервале $[a, b]$ и имеет в нем точки разрыва: $a < c_1 < \ldots < c_k < b$, то

$$\int\limits_{a}^{b} f(x)\,dx = \int\limits_{a}^{c_1} f(x)\,dx + \ldots + \int\limits_{c_k}^{b} f(x)\,dx. \tag{5.21}$$

Последняя формула имеет очевидную геометрическую иллюстрацию (рис. 5.5).

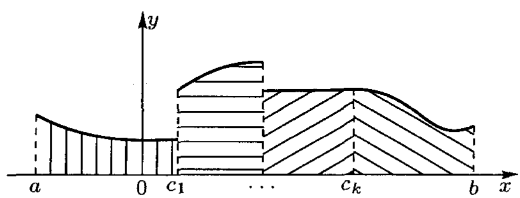


Рис. 5.5. Площадь криволинейной трапеции, соответствующей функции $y = f(x)$ с конечным числом точек разрыва I рода.

Если функция $f(x)$ в интервале интегрирования претерпевает один или несколько разрывов II рода, то интеграл определяется следующим образом.

Определение. Пусть функция $f(x)$ непрерывна в интервале $[a, b)$ и при $x \to b$, $x < b$, не имеет конечного предела. *Несобственным интегралом* от функции $f(x)$ в интервале $[a, b)$

называется предел интеграла $\int\limits_a^{b-\varepsilon} f(x)\,dx$ при $0<\varepsilon\to 0$, т. е.

$$\int\limits_a^b f(x)\,dx = \lim_{0<\varepsilon\to 0} \int\limits_a^{b-\varepsilon} f(x)\,dx. \tag{5.22}$$

Если конечный предел (5.22) существует, то несобственный интеграл называется сходящимся, иначе — расходящимся.

Аналогично определяется несобственный интеграл от функции $f(x)$, претерпевающей разрыв лишь в левом конце интервала интегрирования:

$$\int\limits_a^b f(x)\,dx = \lim_{0<\varepsilon\to 0} \int\limits_{a+\varepsilon}^{b} f(x)\,dx. \tag{5.23}$$

В случае, когда функция $f(x)$ непрерывна в интервале $[a, b)$ и имеет разрывы II рода в его обоих концах, полагаем

$$\int\limits_a^b f(x)\,dx = \int\limits_a^c f(x)\,dx + \int\limits_c^b f(x)\,dx, \tag{5.24}$$

где c — любая точка из интервала (a, b).

Если же функция $f(x)$ имеет разрывы II рода в некоторых промежуточных точках $a < c_1 < \ldots < c_k < b$, то пользуемся равенством

$$\int\limits_a^b f(x)\,dx = \int\limits_a^{c_1} f(x)\,dx + \ldots + \int\limits_{c_k}^b f(x)\,dx$$

и вычисляем слагаемые в правой части по формулам (5.22)–(5.24) соответственно.

Примеры. 1. $I_n = \int\limits_0^1 \frac{dx}{x^n} =?$

Если $n < 1$, то

$$I_n = \lim_{0<\varepsilon\to 0} \frac{x^{1-n}}{1-n}\bigg|_{\varepsilon}^{1} = \frac{x^{1-n}}{1-n}\bigg|_{0}^{1} = \frac{1}{n-1},$$

если $n = 1$, то

$$I_n = \lim_{0<\varepsilon\to 0} \ln x\bigg|_{\varepsilon}^{1} = \lim_{0<\varepsilon\to 0} \ln\frac{1}{\varepsilon} = \infty,$$

если $n > 1$, то

$$I_n = \lim_{0<\varepsilon\to 0} \frac{x^{1-n}}{1-n}\bigg|_{\varepsilon}^{1} = \frac{1}{n-1}\left(1 - \lim_{0<\varepsilon\to 0}\left(\frac{1}{x}\right)^{n-1}\right) = \infty.$$

Таким образом, несобственный интеграл I_n сходится при $n <$ < 1 и расходится, если $n \geqslant 1$.

2. $$\int\limits_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} = \int\limits_{-1}^{0} \frac{dx}{\sqrt{1-x^2}} + \int\limits_{0}^{1} \frac{dx}{\sqrt{1-x^2}} =$$
$$= \arcsin x\bigg|_{-1}^{0} + \arcsin x\bigg|_{0}^{1} = \frac{\pi}{2} + \frac{\pi}{2} = \pi.$$

Интеграл сходится.

3. $$\int\limits_{0}^{2} \frac{dx}{\sqrt[3]{(x-1)^2}} = \int\limits_{0}^{1} \frac{dx}{\sqrt[3]{(x-1)^2}} + \int\limits_{1}^{2} \frac{dx}{\sqrt[3]{(x-1)^2}} =$$
$$= 3(x-1)^{1/3}\Big|_{0}^{1} + 3(x-1)^{1/3}\Big|_{1}^{2} = -3\sqrt[3]{-1} + 3 = 6.$$

Интеграл сходится.

Самостоятельная работа

1. $\displaystyle\int\limits_{2}^{6} \frac{dx}{\sqrt[3]{(4-x)^2}}$. **2.** $\displaystyle\int\limits_{0}^{2} \frac{dx}{(x-1)^2}$.

3. $\displaystyle\int\limits_{a}^{b} \frac{dx}{(b-x)^n}$ при $b > a$.

§ 5.6. Геометрические приложения определенных интегралов

Рассмотрим приложения определенных интегралов для вычисления площади, длины дуги и объема тела вращения в зависимости от формы задания кривых.

5.6.1. Вычисление площади плоской фигуры

Если плоская фигура ограничена прямыми $x = a$, $x = b$ $(a < < b)$ и кривыми $y = f_1(x)$, $y = f_2(x)$, причем $f_1(x) < f_2(x)$ $(a < x < b)$, то ее площадь вычисляется по формуле (см. также п. 5.2, рис. 5.2):

$$S = \int_a^b [f_2(x) - f_1(x)]\, dx. \tag{5.25}$$

В частном случае, когда плоская фигура ограничена снизу осью OX, формула (5.25) несколько упрощается:

$$S = \int_a^b f_2(x)\, dx. \tag{5.26}$$

Примеры.

1. Вычислить площадь фигуры, ограниченной кривыми (рис. 5.6) $y = \sqrt{x}$ и $y = x^2$.

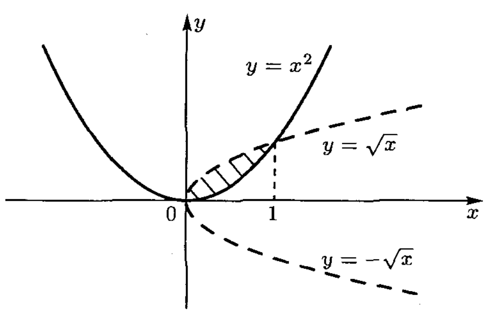


Рис. 5.6. Площадь, ограниченная двумя параболами.

Решение. Находим точки пересечения кривых: $\sqrt{x} = x^2$, следовательно $x = x^4$, откуда $x_1 = 0$, $x_2 = 1$. По формуле (5.25) имеем

$$S = \int\limits_0^1 (\sqrt{x} - x^2)\, dx = \frac{2}{3} x^{3/2}\Big|_0^1 - \frac{x^3}{3}\Big|_0^1 = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.$$

2. Найти площадь фигуры, заключенной между параболой $x^2 = 4y$ и локоном Аньези $y = 8/(x^2+4)$ (рис. 5.7).

Решение. Найдем абсциссы точек A и C пересечения кривых.

Для этого исключим y из системы уравнений:

$$\begin{cases} y = 8/(x^2+4), \\ y = x^2/4, \end{cases}$$

откуда $8/(x^2+4) = x^2/4$, или $x^4 + 4x^2 - 32 = 0$.

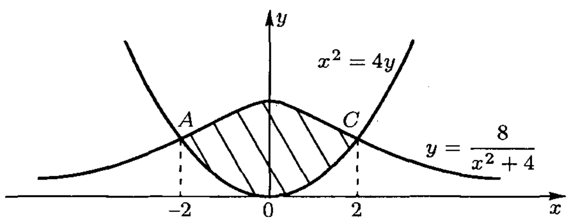


Рис. 5.7. Площадь между параболой и локоном Аньези.

Действительными корнями этого уравнения являются точки $x_1 = -2$ и $x_2 = 2$. Функция $8/(x^2+4) \geqslant x^2/4$ на отрезке $[-2, 2]$ (в чем можно убедиться прямым подсчетом значений этих функций в любой точке внутри отрезка, например, в точке $x = 0$). Следовательно, по формуле (5.25) получим

$$S = \int\limits_{-2}^{2} \left(\frac{8}{x^2+4} - \frac{x^2}{4} \right) dx = \left(4 \operatorname{arctg} \frac{x}{2} - \frac{x^3}{12} \right)\Big|_{-2}^{2} = 2\pi - \frac{4}{3}.$$

3. Вычислить площадь фигуры, ограниченной прямыми $x = 0$, $x = 2$ и кривыми $y = 2^x$, $y = 2x - x^2$ (рис. 5.8).

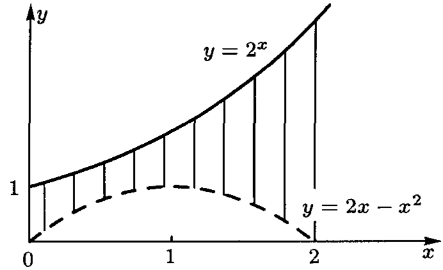


Рис. 5.8. Площадь между параболой и показательной кривой.

Решение. Так как максимум функции $y = 2x - x^2$ достигается в точке $x = 1$ и равен 1, а функция $y = 2^x \geqslant 1$ на отрезке $[0, 2]$, то

$$S = \int_0^2 [2^x - (2x - x^2)]\,dx = \left.\frac{2^x}{\ln 2}\right|_0^2 - \left.\left(x^2 - \frac{x^3}{3}\right)\right|_0^2 = \frac{3}{\ln 2} - \frac{4}{3}.$$

5.6.2. Вычисление длины дуги плоской кривой

Пусть некоторая кривая задана уравнением $y = f(x)$. Длина дуги AB этой кривой, заключенной между вертикальными прямыми $x = a$ и $x = b$, находится по формуле

$$L = \int_a^b \sqrt{1 + [f'(x)]^2}\,dx = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\,dx. \tag{5.27}$$

Примеры.

1. Определить длину окружности $x^2 + y^2 = r^2$.

Решение. Вычислим сначала длину четвертой части окружности, лежащей в первом квадранте. Тогда уравнение дуги AB будет $y = \sqrt{r^2 - x^2}$, откуда $\dfrac{dy}{dx} = -\dfrac{x}{\sqrt{r^2 - x^2}}$, следовательно,

$$\frac{1}{4}L = \int_0^r \sqrt{1 + \frac{x^2}{r^2 - x^2}}\,dx = \int_0^r \frac{r}{\sqrt{r^2 - x^2}}\,dx = r \cdot \arcsin \left.\frac{x}{r}\right|_0^r = r\frac{\pi}{2}.$$

Длина всей окружности $L = 2\pi r$.

2. Найти длину дуги кривой $y^2 = x^3$ от $x = 0$ до $x = 1$ $(y \geqslant 0)$.

Решение.

Дифференцируя уравнение кривой, найдем $y' = (3/2)x^{1/2}$, откуда

$$L = \int_0^1 \sqrt{1 + \frac{9}{4}x}\, dx = \frac{4}{9} \cdot \frac{2}{3} \left(1 + \frac{9}{4}x\right)^{3/2} \Bigg|_0^1 =$$

$$= \frac{8}{27}\left(\frac{13}{4}\right)^{3/2} - \frac{8}{27} = \frac{8}{27}\left(\frac{13}{8}\sqrt{13} - 1\right).$$

3. Вычислить длину дуги кривой $y = \ln\cos x$, заключенной между точками с абсциссами $x = 0$, $x = \pi/4$.

Решение.

Так как $y' = -\operatorname{tg} x$, то $\sqrt{1 + (y')^2} = \sqrt{1 + \operatorname{tg}^2 x} = \dfrac{1}{\cos x}$.

Следовательно,

$$L = \int_0^{\pi/4} \frac{1}{\cos x}\, dx = \ln\operatorname{tg}\left(\frac{\pi}{4} + \frac{x}{2}\right)\Bigg|_0^{\pi/4} = \ln\operatorname{tg}\frac{3\pi}{8}.$$

5.6.3. Вычисление объема тела вращения

Пусть дана криволинейная трапеция, ограниченная кривой $y = f(x)$, осью OX и прямыми $x = a$, $x = b$. Вращая эту фигуру вокруг оси OX, получим тело вращения. Для вычисления объема тела вращения применяется формула:

$$V = \pi \int_a^b y^2\, dx = \pi \int_a^b [f(x)]^2\, dx. \qquad (5.28)$$

Примеры

1. Вычислить объем тела, образуемого вращением вокруг оси OX фигуры, ограниченной линиями $y = x/2 + 4$, $y = 0$, $x = 0$, $x = 6$ (рис. 5.9).

Решение. В первую очередь следует построить эту фигуру и изобразить на том же чертеже тело вращения. В данной задаче

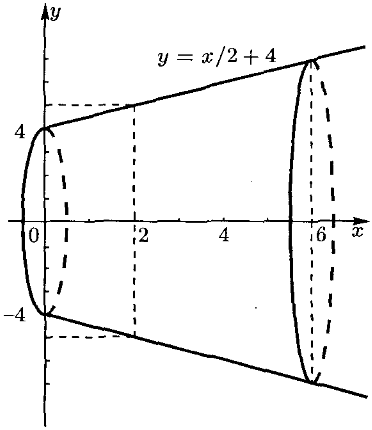


Рис. 5.9. Объем тела, образованного вращением линии $y = x/2+4$ вокруг оси OX.

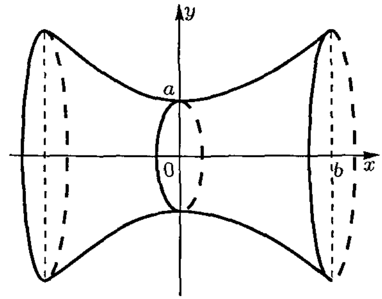


Рис. 5.10. Объем тела, образованного вращением вокруг оси OX цепной линии

$a = 0$, $b = 6$, $y = f(x) = x/2 + 4$. По формуле (5.28) находим

$$V = \pi \int_a^b y^2\,dx = \pi \int_0^6 \left(\frac{1}{2}x + 4\right)^2 dx =$$

$$= \pi \left[\int_0^6 \frac{1}{4}x^2 dx + \int_0^6 4x\,dx + \int_0^6 16\,dx\right] =$$

$$= \pi \left[\frac{1}{12}\left(6^3 - 0\right) + 2\left(6^2 - 0\right) + 16\,(6 - 0)\right] = 186\pi \text{ куб. ед.}$$

Результат можно проверить по формуле объема усеченного конуса:

$$V = \frac{\pi\left(R^2 + Rr + r^2\right)H}{3}.$$

Проверку рекомендуется выполнить самостоятельно.

2. Найти объем тела, образуемого вращением цепной линии (рис. 5.10) $y = \dfrac{a}{2}\left(e^{x/a} + e^{-x/a}\right)$ вокруг оси OX на участке $x = 0$ до $x = b$.

Решение.

$$V = \pi\frac{a^2}{4}\int\limits_0^b \left(e^{x/a}+e^{-x/a}\right)^2\,dx = \frac{\pi a^2}{4}\int\limits_0^b \left(e^{2x/a}+2+e^{-2x/a}\right)\,dx =$$

$$= \frac{\pi a^2}{4}\left[\frac{a}{2}e^{2x/a}+2x-\frac{a}{2}e^{-2x/a}\right]_0^b = \frac{\pi a^3}{8}\left(e^{2b/a}-e^{-2b/a}\right)+\frac{\pi a^2 b}{2}.$$

3. Фигура, ограниченная дугой синусоиды $y = \sin x$, осью ординат и прямой $y = 1$, вращается вокруг оси OY (рис. 5.11). Определить объем V получающегося тела вращения.

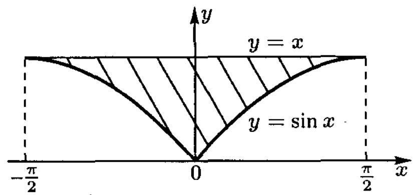


Рис. 5.11. Объем тела, образованного вращением вокруг оси OY дуги синусоиды.

Решение. Обратная функция $x = \arcsin y$ рассматривается на отрезке $[0, 1]$. Поэтому $V = \pi\int\limits_{y_1}^{y_2} x^2\,dy = \pi\int\limits_0^1 (\arcsin y)^2\,dy$.

Применим подстановку $\arcsin y = t$.

y	t
0	0
1	$\pi/2$

Отсюда

$$y = \sin t, \quad dy = \cos t\,dt,$$

значит, $V = \pi\int\limits_0^{\pi/2} t^2\cos t\,dt$. Интегрируя по частям, получим

$$V = \pi\left(\pi^2-8\right)/4.$$

Самостоятельная работа

1. Найти площадь, заключенную между параболой $y = x^2 - 2x + 2$, касательной к ней в точке $M(3, 5)$ и осью ординат.

2. Вычислить площадь фигуры, ограниченной линиями $y = (x-4)^2$, $y = 16 - x^2$ и осью OX.

3. Вычислить площадь фигуры, ограниченной параболами $y = x^2$, $x = (3/4)y^2 + 1$.

4. Вычислить площадь частей эллипса $x^2 + 4y^2 = 8$, отсеченных гиперболой $x^2 - 3y^2 = 1$.

5. Вычислить площадь фигуры, ограниченной замкнутой кривой $y^2 = (1 - x^2)^3$.

6. Вычислить площадь фигуры, ограниченной кривой $\sqrt{x} + \sqrt{y} = 1$ и прямой $x + y = 1$.

7. Вычислить площадь фигуры, ограниченной кривой $y = x^2$, осью OX и прямыми $x = 1$ и $x = 3$.

8. Вычислить площадь фигуры, ограниченной кривой $y = 4x^2 - x^3$ и осью OX.

9. Вычислить площадь фигуры, ограниченной линиями $y = 4 - x^2$, $y = x + 2$.

10. Вычислить длину дуги кривой $y = x^2/2 - 1$, отсеченной осью OX.

11. Вычислить длину дуги кривой $y = \ln(2\cos x)$ между соседними точками пересечения с осью OX.

12. Вычислить длину дуги кривой $3y^2 = x(x-1)^2$ между точками пересечения с осью OX (половину длины петли).

§ 5.7. Численное интегрирование

Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и известна ее первообразная $F(x)$, то определенный интеграл от этой функции в пределах от a до b может быть вычислен по формуле Ньютона–Лейбница (5.12).

Однако во многих случаях первообразная функция не может быть найдена с помощью элементарных средств или является слишком сложной. Вследствие этого вычисление определенного интеграла по формуле может быть затруднительным или практически невозможным.

Для того чтобы все же найти значение $\int\limits_a^b f(x)\,dx$, широко применяются *численные методы*, суть которых основывается, по существу, на определении (см. (5.6)) интеграла. Рассмотрим метод, основанный на формуле площади трапеции.

Пусть для функции $y = f(x)$ в $n+1$ точках $x_0, x_1, \ldots, x_n$ отрезка $[a, b]$ вычислены соответствующие значения

$$f(x_i) = y_i, \quad i = 0, 1, 2 \ldots, n,$$

причем $x_i - x_{i-1} = h = \mathrm{const}$ — шаг разбиения отрезка $[a, b]$. Точки y_0, y_1, $y_2, \ldots, y_n$ соединим отрезками прямых, т. е. заменим непрерывную функцию $y = f(x)$ ломанной линией y_0,

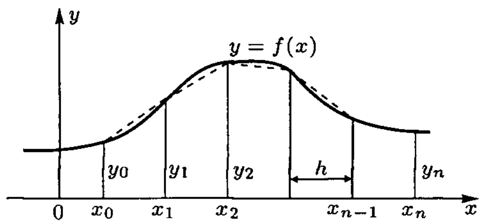


Рис. 5.12. К выводу формулы трапеции.

$y_1, y_2, \ldots, y_n$. Площадь криволинейной трапеции на $[a, b]$ будет приближенно выражаться суммой площадей трапеций с основаниями y_0y_1, $y_1y_2, \ldots, y_{n-1}y_n$ и высотами h.

Следовательно,

$$\int\limits_{x_0}^{x_1} y\,dx \approx \frac{h}{2}(y_0 + y_1),$$

а

$$\int\limits_a^b y\,dx \approx h\left(\frac{y_0 + y_n}{2} + y_1 + y_2 + \ldots + y_{n-1}\right). \tag{5.29}$$

Это выражение позволяет выполнить численное интегрирование методом трапеций, для чего достаточно лишь найти $n+1$ значение подынтегральной функции на отрезке $[a, b]$ и знать шаг разбиения этого отрезка:

$$h = \frac{b-a}{n},$$

где n — количество точек разбиения.

Нетрудно заметить, что сложность подынтегральной функции $y = f(x)$ не имеет существенного значения, а сами величины $y_0, y_1, y_2, \ldots, y_n$ можно найти как с помощью калькулятора, так и на компьютере.

Пример. Найти площадь фигуры, заключенной между параболой $x^2 = 4y$ и локоном Аньези $y = \dfrac{8}{(x^2+4)}$ (см. рис. 5.7).

Точками пересечения этих кривых будут абсциссы: $x_1 = -2$, $x_2 = 2$. Следовательно, искомую площадь запишем как

$$S = \int_{-2}^{2} \left(\frac{8}{x^2+4} - \frac{x^2}{4} \right) dx.$$

Примем $n = 10$, тогда $h = \dfrac{2-(-2)}{10} = 0{,}4$. Составим таблицу значений подынтегральной функции y_i.

x_i	-2	$-1{,}6$	$-1{,}2$	$-0{,}8$	$-0{,}4$
y_i	0	0,5794	1,1105	1,5641	1,8830

x_i	0	0,4	0,8	1,2	1,6	2,0
y_i	2	1,8830	1,5641	1,1105	0,5795	0

Подставляя значения y_i из таблицы в формулу (5.33), получим

$$\int_{-2}^{2} \left(\frac{8}{x^2+4} - \frac{x^2}{4} \right) dx = 0{,}4\left[\frac{0+0}{2} + 2 + 2(0{,}5795 + 1{,}1105 + {} \right.$$
$$\left. {} + 1{,}5641 + 1{,}8830)\right] = 4{,}90968.$$

Точное значение, вычисленное ранее аналитическим способом, равно $2\pi - 4/3 = 4{,}94985$.

Образовавшаяся погрешность, равная 0,04, объясняется малым числом точек деления отрезка $[-2, 2]$ $(n = 10)$.

ГЛАВА 6

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

§ 6.1. Основные понятия и определения

При решении различных задач математики и физики, биологии и медицины довольно часто не удается сразу установить функциональную зависимость в виде формулы, связывающей переменные величины, которые описывают исследуемый процесс. Обычно приходится использовать уравнения, содержащие, кроме независимой переменной и неизвестной функции, еще и ее производные.

Определение. Уравнение, связывающее независимую переменную, неизвестную функцию и ее производные различных порядков, называется *дифференциальным.*

Неизвестную функцию обычно обозначают $y(x)$ или просто y, а ее производные — y', y'' и т. д.

Возможны и другие обозначения, например: если $y = x(t)$, то $x'(t)$, $x''(t)$ — ее производные, а t — независимая переменная.

Определение. Если функция зависит от одной переменной, то дифференциальное уравнение называется обыкновенным. Общий вид *обыкновенного дифференциального уравнения*:

$$F(x, y, y', y'', \dots y^{(n)}) = 0 \tag{6.1}$$

или

$$y^{(n)} = f(x, y, y^{(n-1)}). \tag{6.2}$$

Функции F и f могут не содержать некоторых аргументов, но для того чтобы уравнения были дифференциальными, существенно наличие производной.

Определение. *Порядком дифференциального уравнения* называется порядок старшей производной, входящей в него.

Например: $x^2y' - y = 0$, $y' + \sin x = 0$ — уравнения первого порядка, а $y'' + 2y' + 5y = x$ — уравнение второго порядка.

Определение. Решением дифференциального уравнения называется такая функция, которая обращает уравнение в тождество после подстановки этой функции и ее производных в уравнение.

При решении дифференциальных уравнений используется операция интегрирования, что связано с появлением произвольной постоянной. Если операция интегрирования применяется n раз, то очевидно, что в решении будет содержаться n произвольных постоянных.

§ 6.2. Дифференциальные уравнения первого порядка

Общий вид дифференциального уравнения 1-го порядка определяется выражением

$$F(x, y, y') = 0. \tag{6.3}$$

Уравнение может не содержать в явном виде x и y, но обязательно содержит y'.

Если уравнение можно записать в виде

$$y' = f(x, y), \tag{6.4}$$

то получим дифференциальное уравнение 1-го порядка, разрешенное относительно производной.

Определение. *Общим решением* дифференциального уравнения 1-го порядка (6.3) (или (6.4)) является множество решений $y = f(x, C)$, где C — произвольная постоянная.

График решения дифференциального уравнения называется *интегральной кривой*.

Придавая произвольной постоянной C различные значения, можно получить частные решения. На плоскости XOY общее решение представляет собой семейство интегральных кривых, соответствующих каждому частному решению.

Если задать точку $A(x_0, y_0)$, через которую должна проходить интегральная кривая, то, как правило, из множества функций $y = \varphi(x, C)$ можно выделить одну — частное решение.

Определение. *Частным решением* дифференциального уравнения называется его решение, не содержащее произвольных постоянных.

Если функция $y = \varphi(x, C)$ является общим решением, то из условия $y_0 = \varphi(x_0, C)$ можно найти постоянную C.

Условие $y = y_0$ при $x = x_0$ называют *начальным условием.*

Задача нахождения частного решения дифференциального уравнения (6.3) или (6.4), удовлетворяющего начальному условию $y = y_0$ при $x = x_0$, называется *задачей Коши.*

Всегда ли эта задача имеет решение? Ответ содержится в теореме Коши.

Теорема Коши (теорема существования и единственности решения). Пусть в дифференциальном уравнении $y' = f(x, y)$ функция $f(x, y)$ и ее частная производная $f'_y(x, y)$ определены и непрерывны в некоторой области D, содержащей точку $A(x_0, y_0)$. Тогда в области D существует единственное решение этого уравнения, удовлетворяющее начальному условию $y = y_0$ при $x = x_0$.

Теорема Коши утверждает, что при определенных условиях существует единственная интегральная кривая $y = f(x)$, проходящая через точку $A(x_0, y_0)$. Точки, в которых не выполняются условия теоремы Коши, называются *особыми.* В этих точках терпит разрыв $f(x, y)$ или $\dfrac{df}{dy}$. Через особую точку проходит либо несколько интегральных кривых, либо ни одной.

Определение. Если решение (6.3), (6.4) найдено в виде, не разрешенном относительно y, то оно называется *общим интегралом* дифференциального уравнения $f(x, y, C) = 0$.

Теорема Коши гарантирует только существование решения, а как найти это решение? Оказывается, что единого метода нахождения решения нет. Мы будем рассматривать только некоторые типы дифференциальных уравнений 1-го порядка, интегрируемые в *квадратурах*.

Определение. Дифференциальное уравнение называется интегрируемым в квадратурах, если отыскание его решения сводится к интегрированию функций.

6.2.1. Дифференциальные уравнения первого порядка с разделяющимися переменными

Определение. Дифференциальное уравнение первого порядка называется уравнением с *разделяющимися переменными*, если его можно представить в виде

$$y' = f_1(x) \cdot f_2(y). \tag{6.5}$$

Правая часть уравнения (6.5) представляет собой произведение двух функций, каждая из которых зависит только от одной переменной.

Например, уравнение $y' = \dfrac{y}{x+1}$ является уравнением с разделяющимися переменными ($f_1(x) = \dfrac{1}{x+1}$, $f_2(y) = y$), а уравнение $x^2y' = 2y - x^2$ нельзя представить в виде (6.5).

Учитывая, что $y' = \dfrac{dy}{dx}$, перепишем (6.5) в виде $\dfrac{dy}{dx} = f_1(x) \times$ $\times f_2(y)$.

Из этого уравнения получим дифференциальное уравнение с разделяющимися переменными, в котором при дифференциалах стоят функции, зависящие от соответствующей переменной:

$$\frac{dy}{f_2(y)} = f_1(x)\,dx.$$

Интегрируя почленно, имеем

$$\begin{gathered} \int \frac{dy}{f_2(y)} + C_1 = \int f_1(x)\,dx + C_2, \\ \int \frac{dy}{f_2(y)} = \int f_1(x)\,dx + C, \end{gathered} \tag{6.6}$$

где $C = C_2 - C_1$ — произвольная постоянная. Выражение (6.6) представляет собой общий интеграл уравнения (6.5).

Разделив обе части уравнения (6.5) на $f_2(y)$, мы можем потерять те решения, при которых $f_2(y) = 0$. Действительно, если $f_2(y) = 0$ при $y = y_0$, то $y = y_0$, очевидно, является решением уравнения (6.5).

Пример 1. Найти решение $y' = \frac{y}{x+1}$, удовлетворяющее условию: $y = 6$ при $x = 2$ $(y(2) = 6)$.

Решение. Заменим y' на $\frac{dy}{dx}$, тогда $\frac{dy}{dx} = \frac{y}{x+1}$. Умножим обе части на dx, так как при дальнейшем интегрировании нельзя оставлять dx в знаменателе:

$$dy = \frac{y}{x+1}\,dx,$$

а затем, разделив обе части на y $(y \neq 0)$, получим уравнение $\frac{dy}{y} = \frac{dx}{x+1}$, которое можно проинтегрировать. Интегрируем:

$$\int \frac{dy}{y} = \int \frac{dx}{x+1}.$$

Тогда $\ln|y| = \ln|x+1| + \ln C$; потенцируя, получим $y = C(x+1)$ — общее решение.

По начальным данным определяем произвольную постоянную, подставив их в общее решение $6 = C(2+1) \Rightarrow C = 2$.

Окончательно получаем $y = 2(x+1)$ — частное решение.

Рассмотрим еще несколько примеров решения уравнений с разделяющимися переменными.

Пример 2.

$$x^2y' + y = 0.$$

Учитывая, что $y' = \frac{dy}{dx}$, получим $\frac{dy}{y} = -\frac{dx}{x^2}$.

Проинтегрировав обе части уравнения, будем иметь

$$\ln|y| = \ln C + \frac{1}{x},$$

откуда

$$y = Ce^{1/x}.$$

Пример 3.

$$\sin x \cdot \sin y\, dx + \cos x \cdot y\, dy = 0.$$

Делим обе части уравнения на те сомножители, которые зависят от переменной, не совпадающей с переменной под знаком дифференциала, т. е. на $\sin x \cdot \cos y \neq 0$ и интегрируем.

Тогда получим

$$\int \frac{\sin x}{\cos x}\, dx + \int \frac{\cos y}{\sin y}\, dy = C, \qquad \int \frac{d(\cos x)}{\cos x} + \int \frac{d(\sin y)}{\sin y} = C,$$

$$-\ln|\cos x| + \ln|\sin y| = \ln C$$

и, наконец,

$$\frac{\sin y}{\cos x} = C.$$

Пример 4.

$$x + xy + yy'(1 + x) = 0.$$

Зная, что $y' = \dfrac{dy}{dx}$, получим $x(1 + y) + y(1 + x)\dfrac{dy}{dx} = 0$.
Разделим переменные.
Тогда

$$\frac{x}{x+1} dx + \frac{y}{y+1} dy = 0, \qquad \frac{x}{x+1}\, dx = -\frac{y}{y+1}\, dy$$

Интегрируя, будем иметь

$$x + y = \ln\left[C(x+1)(y+1)\right].$$

Замечание. В примерах 1 и 2 искомая функция y выражена явно (общее решение).

В примерах 3 и 4 — неявно (общий интеграл). В дальнейшем форма решения оговариваться не будет.

Пример 5.

$$x\sqrt{1+y^2} + y\sqrt{1+x^2} \cdot y' = 0, \qquad \int \frac{x\, dx}{\sqrt{1+x^2}} + \int \frac{y\, dy}{\sqrt{1+y^2}} = C,$$

$$\frac{1}{2}\int \frac{d(x^2+1)}{\sqrt{1+x^2}} + \frac{1}{2}\int \frac{d(y^2+1)}{\sqrt{1+y^2}} = C, \qquad \sqrt{1+x^2} + \sqrt{1+y^2} = C.$$

Пример 6. Найти решение уравнения $y' = 2\sqrt{y}\ln x$, удовлетворяющее условию $y(e) = 1$.

Запишем уравнение в виде $\dfrac{dx}{dy} = 2\sqrt{y}\ln x$.

Умножая обе части уравнения на dx и на $2\sqrt{y}$, получим

$$\frac{dy}{2\sqrt{y}} = \ln x\, dx.$$

Интегрируя обе части уравнения (интеграл в правой части берется по частям), получим

$$\sqrt{y} = x\ln x - x + C.$$

Но по условию $y = 1$ при $x = e$.

Тогда $1 = e\ln e - e + C$, $C = 1$.

Подставим найденные значения C в общее решение:

$$\sqrt{y} = x\ln x - x + 1.$$

Полученное выражение называется частным решением дифференциального уравнения.

6.2.2. Однородные дифференциальные уравнения первого порядка

Определение. Дифференциальные уравнения 1-го порядка называется *однородным*, если его можно представить в виде $y' = f\,(y/x)$.

В частности, уравнение, записанное в виде $y' = f(x, y)$, является однородным, если $f(x, y)$ есть отношение двух однородных многочленов одного измерения.

(Однородными многочленами измерения n называются многочлены, у которых сумма показателей степеней переменных в каждом члене равна n. Например, $x^4+x^3y-3x^2y^2+4xy^3-y^4$ — однородный многочлен четвертого измерения.) В однородном уравнении переменные, вообще говоря, не разделяются.

Однако оно легко может быть преобразовано в уравнение с разделяющимися переменными.

Введем новую функцию $u = \dfrac{y}{x}$ или $y = u\cdot x$.

Отсюда $y' = u'x + u$ или $dy = u\,dx + x\,du$.

Эта замена сводит однородные уравнения к уравнениям с разделяющимися переменными.

Пример 1. Решить уравнение $2xyy' = x^2 + y^2$.

Запишем уравнение в виде $y' = \dfrac{x^2+y^2}{2xy}$ или $y' = \dfrac{1+(y/x)^2}{2\cdot(y/x)}$.

Производим подстановку:

$$\frac{y}{x} = u, \quad y' = u'x + u.$$

Тогда

$$u'x + u = \frac{1+u^2}{2u}, \quad u'x = \frac{1+u^2}{2u} - u, \quad u'x = \frac{1-u^2}{2u}.$$

Заменим

$$u' = \frac{du}{dx} \Rightarrow x\frac{du}{dx} = \frac{1-u^2}{2u}.$$

Умножим на dx:

$$x\,du = \frac{1-u^2}{2u}\,dx.$$

Разделим на x и на $\dfrac{1-u^2}{2u}$, тогда

$$\frac{2udu}{1-u^2} = \frac{dx}{x}.$$

Проинтегрировав обе части уравнения по соответствующим переменным, будем иметь

$$\ln\left|\frac{C}{u^2-1}\right| = \ln x, \quad \frac{C}{u^2-1} = x$$

или, возвращаясь к старым переменным, окончательно получим

$$x^2 + cx - y^2 = 0.$$

Пример 2.

$$2x^2y' = \left(x^2+y^2\right).$$

Пусть $u = \dfrac{y}{x}$, тогда $y = u\cdot x$, $y' = u'x + u$,

$$2x^2\left(u'x+u\right) = x^2 + u^2x^2.$$

Поделим обе части уравнения на x^2:

$$2\left(u'x + u\right) = 1 + u^2.$$

Раскроем скобки и перегруппируем слагаемые:

$$2xu' = (u-1)^2, \quad 2x\frac{du}{dx} = (u-1)^2.$$

Далее, интегрируя, получим

$$\int \frac{2du}{(u-1)^2} = \int \frac{dx}{x},$$

или

$$\frac{-2}{u-1} = \ln Cx.$$

Переходя к старым переменным, придем к окончательному результату:

$$Cx = e^{-\frac{2x}{y-x}}.$$

Пример 3. Найти решение уравнения $xy' = y\left(1 + \ln\frac{y}{x}\right)$ при условии $y(1) = e^{-1/2}$.

Выполняя стандартную замену $u = \dfrac{y}{x}$, получаем

$$u + x\frac{du}{dx} = u\ (1 + \ln u)$$

или

$$x\frac{du}{dx} = u\ln u.$$

$$\int \frac{du}{u\ln u} = \int \frac{dx}{x}, \quad \ln|\ln u| = \ln|Cx|, \quad \ln u = Cx, \quad u = e^{Cx}$$

или

$$y = x \cdot e^{Cx}.$$

Но $y = e^{-1/2}$ при $x = 1$. Откуда $e^{-1/2} = 1 \cdot e^C$, $C = -1/2$. Значит, частное решение имеет вид $y = x \cdot e^{-x/2}$.

Пример 4.

$$x\,dy - y\,dx = \sqrt{x^2 + y^2}\,dx,$$

$$x\left(u'x+u-u\right)=x\sqrt{1+u^2},\qquad x\frac{du}{dx}=\sqrt{1+u^2},$$

$$\frac{du}{\sqrt{1+u^2}}=\frac{dx}{x},\qquad \ln\left|u+\sqrt{1+u^2}\right|=\ln|Cx|,$$

$$u+\sqrt{u^2+1}=Cx\quad\text{или}\quad y+\sqrt{x^2+y^2}=Cx^2.$$

Пример 5.

$$y'=\frac{x^2+xy+y^2}{x^2},\qquad u'x+u=\frac{x^2+x^2u+u^2x^2}{x^2},$$

$$u'x+u=1+u+u^2,\qquad x\frac{du}{dx}=1+u^2,\qquad \int\frac{du}{1+u^2}=\int\frac{dx}{x},$$

$$\operatorname{arctg}u=\ln|xc|,\qquad \operatorname{arctg}\frac{y}{x}=\ln|xc|.$$

6.2.3. Линейные дифференциальные уравнения первого порядка

Определение. Дифференциальное уравнение первого порядка называется *линейным*, если его можно представить в виде

$$y'+p(x)y=q(x). \tag{6.7}$$

Здесь y, y' — входят в 1-й степени, $p(x)$, $q(x)$ — неизвестные функции, в частности, они могут быть постоянными величинами.

Если $q(x)=0$, то уравнение $y'+p(x)y=0$ называется *однородным*.

Уравнение (6.7) решается методом Бернулли с помощью специального приема. Представим y в виде произведения двух функций $y=u(x)v(x)$, где u, v — неизвестные функции, причем одну из них можно выбрать произвольно.

Итак, если $y=uv$, то $y'=u'v+uv'$ и, подставляя в (6.7), получим

$$u'v+uv'+p(x)uv=q(x),$$

$$u'v+u(v'+p(x)v)=q(x).$$

Функции u, v нам неизвестны, определим одну из них, например v, из условия

$$v' + p(x)v = 0. \tag{6.8}$$

Учитывая это, придем к уравнению

$$u'v = q(x). \tag{6.9}$$

Найдем $v(x)$ из (6.8):

$$v' = -p(x)v, \quad \frac{dv}{dx} = -p(x)v, \quad \frac{dv}{v} = -p(x)\,dx,$$

$$\ln|v| = -\int p(x)\,dx, \quad v = e^{\int p(x)\,dx}.$$

Подставим найденное решение в (6.9):

$$\frac{du}{dx}e^{-\int p(x)\,dx} = q(x), \quad du = q(x)e^{\int p(x)\,dx}dx,$$

$$u = \int q(x)e^{\int p(x)dx}dx + C.$$

Общее решение (6.7):

$$y = e^{-\int p(x)\,dx}\left(C + \int q(x)e^{\int p(x)\,dx}\right).$$

Пример 1.

$$y' - \frac{xy}{x^2+1} = x.$$

После подстановки в уравнение получим $u'v + uv' - \dfrac{xuv}{x^2+1} = x$.

Сгруппируем второе и третье слагаемое в левой части уравнения:

$$u'v + u\left(v' - \frac{xv}{x^2+1}\right) = x.$$

Пользуясь тем, что одна из неизвестных функций (например, v) может быть выбрана произвольно, потребуем, чтобы $v' - \dfrac{xv}{x^2+1} = 0$.

Тогда из уравнения получим $u'v = x$.

Таким образом, решение линейного уравнения свелось к решению системы дифференциальных уравнений с разделяющимися переменными:

$$\begin{cases} v' - \dfrac{xv}{x^2+1} = 0, \\ u'v = x. \end{cases}$$

Решая первое уравнение системы, получим

$$\ln|v| = \frac{1}{2}\ln|x^2+1|, \quad v = \sqrt{x^2+1}.$$

Подставим найденную функцию v во второе уравнение системы. Тогда это уравнение примет вид: $u'\sqrt{x+1} = x$ или $du = \dfrac{x}{\sqrt{x^2+1}}dx$, откуда $u = \sqrt{1+x^2} + C$.

Искомое решение будет иметь вид

$$y = \left(\sqrt{1+x^2} + C\right)\sqrt{x^2+1}.$$

Пример 2. Найти общее решение $xy' + y = e^{-x}$ $(x \neq 0)$.

Уравнение линейное, так как y, y' входят в него в первой степени и нет члена с произведением yy'.

Полагаем $y = uv$, тогда $y' = u'v + uv'$,

$$u'v + uv' + \frac{1}{x}uv = \frac{e^{-x}}{x}, \quad u'v + u\left(v' + \frac{v}{x}\right) = \frac{e^{-x}}{x},$$

$$\begin{cases} v' + \dfrac{v}{x} = 0, \\ u'v = \dfrac{e^{-x}}{x}, \end{cases}$$

$$\frac{dv}{dx} = -\frac{v}{x} \Rightarrow \frac{dv}{v} = -\frac{dx}{x}, \quad v = \frac{1}{x}.$$

Подставим во второе уравнение:

$$u'\frac{1}{x} = \frac{e^{-x}}{x} \quad \text{или} \quad du = e^{-x}\,dx,$$

$$\int du = \int e^{-x}\,dx \Rightarrow u = -e^{-x} + C.$$

Итак, $v = 1/x$, $u = -e^{-x} + C$, тогда $y = uv = 1/x\,(C - e^{-x})$.

Пример 3.

$$y' - \frac{2x}{1+x^2}y = 1 + x^2.$$

После введения функций u и v получим

$$u'v + u\left(v' - \frac{2xy}{1+x^2}\right) = 1 + x^2$$

или

$$\begin{cases} v' - \dfrac{2xv}{1+x^2} = 0, \\ u'v = 1 + x^2. \end{cases}$$

Интегрируя первое уравнение системы, получим: $v = 1 + x^2$, тогда $u' = 1$, откуда $u = x + C$, тогда $y = (1 + x^2)\,(x + C)$.

Пример 4.

$$y' - y\operatorname{tg} x = \frac{2x}{\cos x}.$$

Применяя указанный выше способ решения, получим систему

$$\begin{cases} v' - v\operatorname{tg} x = 0, \\ u'v = \dfrac{2x}{\cos x}. \end{cases}$$

Из первого уравнения находим $v = 1/\cos x$.

Решая второе уравнение системы, получим $u = x^2 + C$.

Тогда искомое решение примет вид: $y = \dfrac{1}{\cos x}\,(x^2 + C)$.

Пример 5.

$$x^2y' + 2xy = \ln x.$$

$$x^2u'v + u\,(x^2v' + 2xv) = \ln x,$$

$$\begin{cases} x^2v' + 2xv = 0, \\ x^2u'v = \ln x. \end{cases}$$

После деления на $x \neq 0$ первое уравнение примет вид $xv' + 2v = 0$.

Разделяя переменные в этом уравнении, будем иметь $v = 1/x^2$.

Второе уравнение системы после подстановки функции v примет вид $u' = \ln x$, откуда $u = x\ln x - x + C$.

Следовательно,

$$y = \frac{x\ln x - x + C}{x^2}.$$

Самостоятельная работа

Найти решение следующих дифференциальных уравнений с разделяющимся переменными:

1. $y' = \dfrac{\sqrt{y-1}}{x}$. **2.** $xy(1+x^2)y' - 1 - y^2 = 0$.

3. $y'\cos x = \dfrac{y}{\ln y}$, $y(0) = 1$.

4. $(1+x^2)y' + y = 0$, $y(1) = 1$.

5. $y' + \sqrt{\dfrac{1-y^2}{1-x^2}} = 0$. **6.** $yy' = \dfrac{1-2x}{y}$. **7.** $y'y^2 + 2x = 1$.

8. $y' = \dfrac{2^{3x}}{5^{2y}}$. **9.** $(x+1)^3 \cdot y' - (y-2)^2 = 0$.

10. $y' + \sqrt{\dfrac{1-y^2}{1-x^2}} = 0$.

Найти решение однородных дифференциальных уравнений:

1. $xy'\cos\dfrac{y}{x} = y\cos\dfrac{y}{x} - \dfrac{x^2}{y}$. **2.** $x^3y' = y\left(y^2 + x^2\right)$.

3. $yy' = 2y - x$. **4.** $y' = \dfrac{x+3y}{2x}$.

5. $y\,dx + (2\sqrt{xy} - x)\,dy = 0$.

6. $xye^{x/y}\,dy + y^2\,dx = x^2e^{x/y}\,dy$.

7. $xy' - y = \sqrt{x^2 + y^2}$. **8.** $y' = \dfrac{y^2}{x^2} - \dfrac{y}{x}$. **9.** $y' - \dfrac{y}{x} = 1$.

Найти решение линейных дифференциальных уравнений:

1. $y' + y \operatorname{tg} x = \sin 2x$. **2.** $xy' - 2y = x^3 \ln x$.

3. $y' - 2y = x^2 - x$. **4.** $y' + xy = e^{-x^2/2}$.

5. $x^3 y' + 4y = 12$. **6.** $y' - \dfrac{y}{x} = -\dfrac{\ln x}{x}$.

7. $(2 + x^2)\, y' + xy = 1$. **8.** $y' - \dfrac{2}{x} y = -x^4$.

9. $xy' + y = \ln x + 1$. **10.** $y' - y \operatorname{tg} x = \operatorname{ctg} x$.

6.2.4. Некоторые приложения дифференциальных уравнений первого порядка

Задача о радиоактивном распаде

Скорость распада Ra (радия) в каждый момент времени пропорциональна его наличной массе. Найти закон радиоактивного распада Ra, если известно, что в начальный момент имелось m_0 Ra и период полураспада Ra равен 1590 лет.

Р е ш е н и е. Пусть в момент t масса Ra составляет x г. Тогда скорость распада Ra равна

$$\frac{d(m_0 - x)}{dt} = -\frac{dx}{dt}.$$

По условию задачи

$$-\frac{dx}{dt} = kx,$$

где k — коэффициент пропорциональности.

Разделяя в последнем уравнении переменные и интегрируя, получим

$$\ln x = -kt + \ln C,$$

откуда $x = Ce^{-kt}$.

Для определения C используем начальное условие: при $t = 0$ $x = m_0$.

Тогда $C = m_0$ и, значит, $x(t) = m_0 e^{-kt}$.

Коэффициент пропорциональности k определяем из дополнительного условия: при $t = 1590$ $x = \frac{m_0}{2}$.

Имеем $\frac{m_0}{2} = m_0 e^{-1590k}$ или $e^{1590k} = 2$.

Отсюда $e^k = 2^{1/1590}$ и искомая формула

$$x(t) = m_0 2^{-t/1590}.$$

Задача о скорости размножения бактерий

Скорость размножения бактерий пропорциональна их количеству. В начальный момент имелось 100 бактерий. В течение 3 часов их число удвоилось. Найти зависимость количества бактерий от времени. Во сколько раз увеличится количество бактерий в течение 9 часов?

Решение. Пусть x — количество бактерий в момент t. Тогда согласно условию

$$\frac{dx}{dt} = kx$$

где k — коэффициент пропорциональности.

Отсюда $x = ce^{kt}$. Из условия известно, что $x|_{t=0} = 100$. Значит,

$$C = 100, \quad x = 100e^{kt}.$$

Из дополнительного условия $x|_{t=3} = 200$. Тогда $200 = 100e^{3k}$, $2 = e^{3k}$, $e^k = 2^{1/3}$.

Искомая функция:

$$x = 100 \cdot 2^{t/3}.$$

Значит, при $t = 9$ $x = 800$, т. е. в течение 9 часов количество бактерий увеличилось в 8 раз.

Задача об увеличении количества фермента

В культуре пивных дрожжей быстрота прироста действующего фермента пропорциональна его начальному количеству x. Первоначальное количество фермента a в течение часа удвоилось. Найти зависимость $x(t)$.

Решение. По условию дифференциальное уравнение процесса имеет вид

$$\frac{dx}{dt} = kx$$

отсюда $x = ce^{kt}$.

Но $x\,|_{t=0} = a$. Значит, $C = a$, и тогда $x = ae^{kt}$.

Известно также, что $x\,|_{t=1} = 2a$.

Следовательно, $2a = ae^{k}$, $e^{k} = 2 \Rightarrow x(t) = a2^{t}$.

Задача. Динамика численности популяции

Рассмотрим колонию организмов, обитающих в условиях неограниченных ресурсов питания. Предположим, что колония не подавляется никаким другим видом. В силу размножения и смертности число живых организмов в колонии будет меняться с течением времени. Найти закон этого изменения.

Решение. Пусть $x = x\,(t)$ — число живых организмов в момент t, $x\,(t + \Delta t)$ — в момент $t + \Delta t$. Тогда $\Delta x = x\,(t + \Delta t) - x\,(t)$ — приращение функции $x\,(t)$ за промежуток Δt.

Из чего складывается приращение?

За время Δt взрослые особи (или их часть) произведут потомство, а часть особей может погибнуть.

Таким образом,

$$\Delta x = G - H,$$

где G — число развившихся особей за период Δt, H — число погибших особей за это время.

G зависит от длины промежутка Δt (чем больше Δt, тем больше G) и от количества «родителей» (чем больше взрослых особей, тем больше их потомство):

$$G = \Phi\,(x, \Delta t)\,,$$

где $\Phi\,(x, \Delta t)$ растет с ростом x или Δt и равна нулю, если равна нулю одна из переменных.

Из экспериментов известно, что Δt должна входить линейно: если промежуточные наблюдения увеличить, например, в 2 раза, то и прирост потомства микроорганизмов увеличивается в 2 раза, т. е. $\Phi\,(x, \Delta t) = f\,(x)\,\Delta t$.

Характер $f(x)$ определить сложнее. Но известно, что $f(x)$ монотонно возрастает с ростом xи $f(0) = 0$.

Но каков рост $f(x)$? Он существенно зависит от биологических особенностей исследуемого вида, и для его описания могут понадобиться те или иные степени x, рациональная функция и т. п.

Мы рассмотрим простейший случай, когда численность потомства пропорциональна количеству «родителей»: $f(x) = \alpha x$, $\alpha = \text{const}$ (например, такой случай реализуется при делении клеток).

Итак, $G = \alpha x \Delta t$.

Аналогично, $H = \beta x \Delta t$.

Следовательно,

$$\Delta x = \gamma x \Delta t, \tag{6.10}$$

где $\gamma = \alpha - \beta$.

Разделим обе части равенства (6.10) на Δt и перейдем к пределу при $\Delta t \to 0$:

$$\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = \frac{dx}{dt},$$

и, значит,

$$\frac{dx}{dt} = \gamma x$$

Тогда, после интегрирования и разделения переменных, будем иметь $x(t) = Ce^{\gamma t}$.

Используем начальное условие: $x(t_0) = x_0$ (t_0 — время начала наблюдения за колонией; x_0 — количество организмов).

Тогда искомый закон будет иметь такой вид:

$$C = x_0 e^{-\gamma t_0}$$

$$x(t) = x_0 e^{\gamma(t-t_0)}$$

Нужно отметить, что найденный закон носит только предположительный пока характер. Для удвоения количества живых организмов требуется всегда одно и то же время, независимо от первоначального количества (кстати, население Земли удваивается примерно через каждые 40 лет).

§ 6.3. Дифференциальные уравнения второго порядка

6.3.1. Основные понятия

Определение. Дифференциальным уравнением второго порядка называется соотношение, связывающее независимую переменную, искомую функцию и ее первую и вторую производные.

В частных случаях в уравнении могут отсутствовать x, y или y'. Однако уравнение 2-го порядка обязательно должно содержать y''. В общем случае дифференциальное уравнение 2-го порядка записывается в виде

$$F(x, y, y', y'') = 0, \tag{6.11}$$

или, если это возможно, в виде, разрешенном относительно второй производной:

$$y'' = f(x, y, y'). \tag{6.12}$$

Как и в случае уравнения 1-го порядка, для уравнения 2-го порядка могут существовать общее и частное решения. Общее решение имеет вид

$$y = \varphi(x, C_1, C_2).$$

Частное решение

$$y = \varphi(x, C_{10}, C_{20}) \tag{6.13}$$

находится при начальных условиях $y = y_0$, $y' = y_1$ при $x = x_0$ (x_0, y_0, y_1 — заданные числа).

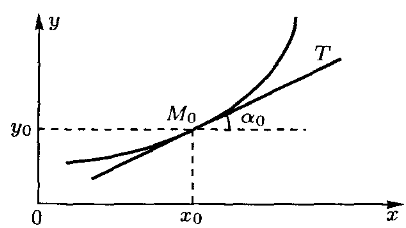


Рис. 6.1. Интегральная кривая.

Нахождение частного решения при указанных начальных данных называется *задачей Коши*. Геометрически это означает, что требуется найти интегральную кривую $y = y(x)$, проходящую через заданную точку M_0 (x_0, y_0) и имеющую в этой точке касательную M_0T, которая образует с положительным направлением оси OX заданный угол α_0, т. е. $\operatorname{tg} \alpha_0 = y_1$.

Задача Коши имеет единственное решение, если правая часть уравнения (6.12), $f(x, y, y')$, непрерывна и имеет непрерывные частные производные по y, y' в некоторой окрестности начальной точки (x_0, y_0, y_1).

Для нахождения постоянных C_1, C_2, входящих в частное решение, надо разрешить систему

$$\begin{cases} y_0 = \varphi(x_0, C_1, C_2), \\ y_1 = \varphi'(x_0, C_1, C_2) \end{cases}$$

относительно произвольных постоянных C_1, C_2: $C_1 = C_{10}$, $C_2 = C_{20}$. Подставим найденные значения произвольных постоянных в общее решение, получим $y = \varphi(x, C_{10}, C_{20})$ — частное решение. Других решений с начальными данными x_0, y_0, y_1 — нет.

Общее решение, записанное в виде, не разрешенном относительно искомой функции y, $\Phi(x, y, C_1, C_2) = 0$, называется общим интегралом.

6.3.2. Дифференциальные уравнения второго порядка, допускающие понижение порядка

Простейшее дифференциальное уравнение 2-го порядка

Это уравнение имеет вид

$$y'' = f(x). \tag{6.14}$$

Правая часть уравнения — функция одной переменной x. Общее решение уравнения найдем двукратным интегрированием, последовательно понижая порядок уравнения на единицу.

Перепишем уравнение (6.14) в виде

$$\frac{d(y')}{dx} = f(x).$$

Тогда

$$dy' = f(x)\,dx \quad \text{и} \quad y' = \int f(x)\,dx + C_1.$$

Затем

$$dy = \left(\int f(x)dx + C_1\right) dx \quad \text{и} \quad y = \int \left(\int f(x)\, dx + C_1\right) dx + C_2$$

— будет общим решением (6.14).

6.3.3. Дифференциальное уравнение второго порядка, не содержащее в явном виде искомую функцию y

Допустим, что $F(x, y', y'') = 0$. Порядок уравнения можно понизить заменой $y' = z(x)$, $y'' = z'(x)$:

$$F(x, z, z') = 0.$$

Решив это уравнение, найдем $p = \varphi(x, C_1)$.

Выполним теперь обратную замену: так как $z = y'$, то $y' = \varphi(x, C_1)$. Это дифференциальное уравнение с разделяющимися переменными, из которого получим

$$dy = \varphi(x, C_1)dx,$$

$$y = \int \varphi(x, C_1)dx + C_2 = \varphi_1(x, C_1, C_2)$$

— общее решение.

6.3.4. Дифференциальное уравнение второго порядка, не содержащее в явном виде независимой переменной x

Допустим, что $F(y, y', y'') = 0$. Порядок этого уравнения можно понизить на единицу, выполнив подстановку

$$y' = p(y).$$

Так как исходное уравнение не содержит в явном виде x, будем считать, что y', y'', p — являются функциями от y, а y — функция от x. Следовательно, p — сложная функция, зависящая от x через y.

Найдем вторую производную:

$$y'' = \frac{dy'}{dx} = \frac{d(p(y))}{dx} = p'_y y'_x = \frac{dp}{dy}p.$$

Выполним замены $F\left(y,\ p(y),\ p\dfrac{dp}{dy}\right) = 0$. Решив полученное дифференциальное уравнение 1-го порядка, находим $p = \varphi(y, C_1)$. Затем заменим p на $y' = \dfrac{dy}{dx}$:

$$\frac{dy}{dx} = \varphi(y, C_1), \quad dy = \varphi(y, C_1)dx, \quad \int \frac{dy}{\varphi(y, C_1)} = \int dx.$$

Решая это уравнение, получим $y = \varphi(x,\ C_1,\ C_2)$ — общее решение.

Примеры.

1. Решить уравнение: $y'' = 4\cos 2x$.

Интегрируя один раз, будем иметь

$$y' = 4\int \cos 2x dx + C_1 = 2\sin 2x + C_1.$$

После повторного интегрирования получим

$$y = \int (2\sin 2x + C_1)dx = -\cos 2x + C_1 x + C_2.$$

2. Решить уравнение: $2xy'' = y'$.

Это уравнение не содержит y, следовательно, пусть $y' = z$, $y'' = z'$. Тогда относительно функции z мы получим уравнение 1-го порядка с разделяющимися переменными: $2xz' = z$. Разделяя переменные и интегрируя, получим: $\ln z = 1/2 \ln x + \ln C_1$, откуда $z = \sqrt{x}C_1 = y'$. По найденному y' восстановим функцию y:

$$y = \int \sqrt{x}C_1\, dx_1 = \frac{2C_1 x^{3/2}}{3} + C_2.$$

3. $y'' + y' \operatorname{tg} x = \sin 2x$.

Переходя к функции z, как и в предыдущем примере, получим следующее уравнение 1-го порядка: $z' + \operatorname{tg} xz = \sin 2x$. Это уравнение является линейным. Для его решения воспользуемся методом Бернулли (будем искать неизвестную функцию z как произведение двух функций): $z = u \cdot v$. Тогда $z' = u'v + uv'$ и уравнение примет вид

$$u'v + uv' + \operatorname{tg} xuv = \sin 2x.$$

Объединяем второе и третье слагаемые: $u'v + u(v' + \operatorname{tg} xv) = \sin 2x$. Потребуем, чтобы $v' + \operatorname{tg} xv = 0$, а значит, $u'v = \sin 2x$. Итак:

$$\begin{cases} v' + \operatorname{tg} xv = 0, \\ u'v = \sin 2x. \end{cases}$$

Решим первое уравнение системы, являющееся уравнением с разделяющимися переменными:

$$\frac{dv}{dx} = -\operatorname{tg} xv, \qquad \frac{dv}{v} = -\operatorname{tg} x\, dx,$$

и после интегрирования получим $\ln|v| = \ln|\cos x|$, $v = \cos x$ Подставляя найденное решение во второе уравнение системы, получим: $u' \cos x = 2\sin x \cos x$ или $u' = 2\sin x$. Откуда $\int du = \int 2\sin x\, dx$, и, следовательно, $u = -2\cos x + C_1$.

Значит,

$$z = (-2\cos x + C_1)\cos x = -2\cos^2 x + C_1 \cos x = y,$$

$$y = \int y'\, dx = \int (-2\cos^2 x + C_1 \cos x)\, dx =$$
$$= -\int (1 + \cos 2x)\, dx + C_1 \sin x = -x - \frac{\sin 2x}{2} + C_1 \sin x + C_2.$$

4. Решить уравнение: $2(y')^2 = (y-1)y''$.

В отличие от уравнений из примеров 3 и 4, это уравнение в явном виде не содержит x. Следовательно, полагая $y' = p(y)$ (тогда $y'' = p\dfrac{dp}{dy}$), будем иметь

$$2p^2 = (y-1)p\frac{dp}{dy}.$$

1) Пусть $p = 0$. Тогда $dy/dx = 0$ и $y = C$.

2) Пусть $p \neq 0$. Разделим обе части уравнения на p:

$$2p = (y-1)\cdot\frac{dp}{dy}.$$

Это уравнение с разделяющимися переменными:

$$\int \frac{dy}{y-1} = \frac{1}{2}\int \frac{dp}{p},$$

откуда

$$\ln|y-1| = \ln\sqrt{p}C_1, \quad y-1 = \sqrt{p}C_1.$$

Следовательно, $\sqrt{p} = \widetilde{C}_1(y-1)$, где $\widetilde{C}_1 = \dfrac{1}{C_1}$. Тогда $p = \widetilde{C}_1^2(y-1)^2$, $\dfrac{dy}{dx} = \widetilde{C}_1^2(y-1)^2$. Таким образом, мы вновь пришли к решению дифференциального уравнения с разделяющимися переменными:

$$\int \frac{dy}{\widetilde{C}_1(y-1)^2} = \int dx,$$

$$(1-y)(x+C_2) = C, \quad C = \frac{1}{\widetilde{C}_1^2}.$$

Самостоятельная работа

1. $2yy'' = 1 + (y')^2$. 2. $y''\operatorname{tg} y = 2(y')^2$. 3. $y''\operatorname{tg} x = y' + 1$.
4. $(y''x - y')y' = x^3$. 5. $y'' = e^{5x}$. 6. $y'' = \ln x$.

6.3.5. Линейные дифференциальные уравнения второго порядка с постоянными коэффициентами

Рассмотрим уравнение вида

$$y'' + py' + qy = 0,$$

где p, q — натуральные числа.

Для нахождения общего решения линейного однородного уравнения 2-го порядка достаточно знать фундаментальную систему частных решений $y_1(x)$, $y_2(x)$. Тогда общее решение получается как линейная комбинация $y(x)$: $y(x) = C_1y_1 + C_2y_2$.

Чтобы найти фундаментальную систему решений, поступим следующим образом. Ищем частное решение в виде $y = e^{kx}$, где k — подлежащее определению число. Подставим e^{kx} в уравнение, учитывая, что $y' = ke^{kx}$, $y'' = k^2e^{kx}$.

Тогда

$$k^2e^{kx}+pk\ {}^{kx}+qe^{kx}=0,\quad e^{kx}(k^2+pk+q)=0.$$

Так как $e^{kx}\neq 0$, то сокращая, получим уравнение

$$k^2+pk+q=0.$$

Это уравнение называется *характеристическим*, а его корни — характеристическими числами, которые определяют те значения k, которые дают решение уравнения в виде e^{kx}.

Характеристическое уравнение может быть получено из исходного заменой y', y'', y соответственно на k, k^2 и 1, т. е. степень k совпадает с порядком производной y, если считать, что производная порядка 0 равна самой функции.

Характеристическое уравнение является уравнением второй степени и имеет, следовательно, два корня. Эти корни могут быть либо действительными различными, либо действительными равными, либо комплексными сопряженными.

Рассмотрим, какой вид имеет фундаментальная система частных решений в каждом из этих случаев,

$$k_{1,2}=-\frac{p}{2}\pm\sqrt{\frac{p^2}{4}-q}.$$

1. Корни характеристического уравнения действительные и различные,

$$D=\frac{p^2}{4}-q>0,\quad k_1\neq k_2.$$

В этом случае $y_1=e^{k_1x}$, $y_2=e^{k_2x}$ — линейно независимые частные решения, поэтому они образуют фундаментальную систему, а общее решение:

$$y=C_1e^{k_1x}+C_2e^{k_2x}.$$

2. Корни характеристического уравнения действительные равные,

$$D=0,\quad k_1=k_2=-\frac{p}{2}.$$

Тогда $y_1=e^{k_1x}$ и $y_2=xe^{k_1x}$ будут линейно независимыми решениями.

Проверим, что y_2 является решением. Действительно:

$$\begin{aligned} y_2'' + py_2' + qy_2 &= (xe^{k_1x})'' + p(xe^{k_1x})' + qxe^{k_1x} = \\ &= 2k_1e^{k_1x} + k_1^2xe^{k_1x} + p(e^{k_1x} + xk_1e^{k_1x}) + qxe^{k_1x} = \\ &= e^{k_1x}(2k_1 + k_1^2x + p + pxk_1 + qx) = \\ &= e^{k_1x}(x(k_1^2 + pk_1 + q) + p + 2k_1). \end{aligned}$$

$k_1^2 + pk_1 + q = 0$, так как k_1 — корень. По теореме Виета $p = -(k_1 + k_2) = -2k_1$, поэтому $p + 2k_1 = 0$, значит, xe^{k_1x} — является решением. Тогда общее решение:

$$y = e^{k_1x}(C_1 + C_2x).$$

3. Корни характеристического уравнения комплексные сопряженные,

$$D < 0 \quad k_1 = \alpha + i\beta, k_2 = \alpha - i\beta.$$

В этом случае можно записать частные решения в виде

$$\bar{y}_1 = e^{(\alpha+i\beta)x} = e^{\alpha x}e^{i\beta x}, \quad \bar{y}_2 = e^{(\alpha-i\beta)x} = e^{\alpha x}e^{-i\beta x}.$$

Пользуясь формулой Эйлера, получим

$$\bar{y}_1 = e^{\alpha x}(\cos\beta x + i\sin\beta x), \quad \bar{y}_2 = e^{\alpha x}(\cos\beta x - i\sin\beta x).$$

Эти решения являются комплексными. Чтобы получить действительные решения, составим новые функции:

$$y_1 = \frac{1}{2}(\bar{y}_1 + \bar{y}_2) = e^{\alpha x}\cos\beta x, \quad y_2 = \frac{1}{2i}(\bar{y}_1 - \bar{y}_2) = e^{\alpha x}\sin\beta x.$$

Эти функции тоже будут частными решениями как линейные комбинации $\bar{y}_1$, $\bar{y}_2$, и, кроме того, они линейно независимы. Поэтому

$$y = C_1e^{\alpha x}\cos\beta x + C_2e^{\alpha x}\sin\beta x = e^{\alpha x}(C_1\cos\beta x + C_2\sin\beta x).$$

Составим таблицу формул общего решения в зависимости от вида корней характеристического уравнения.

Таблица формул общего решения

<table>
<tr><td colspan="3">Дифференциальное уравнение</td></tr>
<tr><td colspan="3">$y'' + py' + qy = 0$</td></tr>
<tr><td colspan="3">Характеристическое уравнение</td></tr>
<tr><td colspan="3">$k^2 + pk + q = 0$</td></tr>
<tr><td colspan="3">Корни характеристического уравнения</td></tr>
<tr><td>$k_1 \neq k_2$</td><td>$k_1 = k_2$</td><td>$k_1 = \alpha + i\beta$
$k_2 = \alpha - i\beta$</td></tr>
<tr><td colspan="3">Фундаментальная система решений</td></tr>
<tr><td>$e^{k_1 x}$
$e^{k_2 x}$</td><td>$e^{k_1 x}$
$xe^{k_1 x}$</td><td>$e^{\alpha x}\cos\beta x$
$e^{\alpha x}\sin\beta x$</td></tr>
<tr><td colspan="3">Вид общего решения</td></tr>
<tr><td>$C_1 e^{k_1 x} + C_2 e^{k_2 x}$</td><td>$e^{k_1 x}(C_1 + C_2 x)$</td><td>$e^{\alpha x}(C_1\cos\beta x + C_2\sin\beta x)$</td></tr>
</table>

Примеры.

1. $2y'' + 5y' + 2y = 0$.

Составим характеристическое уравнение: $2k^2 + 5k + 2 = 0$, откуда $k_1 = -2$, $k_2 = -1/2$, и, следовательно, $y = C_1 e^{-2x} + C_2 e^{-x/2}$.

2. $y'' - 2y' + y = 0$.

Этому дифференциальному уравнению соответствует характеристическое уравнение вида: $k^2 - 2k + 1 = 0$, $k_1 = k_2 = 1$, и общее решение уравнения имеет вид: $y = e^x(C_1 + C_2 x)$.

3. $y'' + 6y' + 13y = 0$.

Характеристическое уравнение $k^2 + 6k + 13 = 0$ будет иметь в данном случае комплексно-сопряженные корни: $k_{1,2} = -3 \pm 2i$ ($\alpha = -3$, $\beta = 2$), а значит, $y = e^{-3x}\ (C_1\cos 2x + C_2\sin 2x)$.

4. $y'' + y = 0$.

Характеристическое уравнение имеет вид: $k^2 + 1 = 0$, и его корни: $k_{1,2} = \pm i$ ($\alpha = 0$, $\beta = 1$). Общее решение $y = C_1\cos x + C_2\sin x$.

Самостоятельная работа

Решить уравнения:

1. $2y'' + y' + 2\sin^2 15^\circ \cos^2 15^\circ \cdot y = 0$. **2.** $4y'' - 8y' + 5y = 0$.

3. $3y'' - 2y' - 8y = 0$. **4.** $y'' + a^2 y = 0$.

5. $y'' + 5y' + 4y = 0$. **6.** $y'' - y' + 2y = 0$.

7. $y'' - 4y' + 13y = 0$. **8.** $y'' - 4y' + 3y = 0$.

9. $y'' - 10y' + 25y = 0$. **10.** $y'' - 2y' + 10y = 0$.

11. $y'' + 3y' - 10y = 0;\quad y = 2,\quad y' = -3$ при $x = 0$.

12. $y'' + 4y' - 32y = 0;\quad y = 8,\quad y' = -4$ при $x = 0$.

13. $y'' + 8y' + 16y = 0;\quad y = 3,\quad y' = 1$ при $x = 0$.

14. $y'' - 8y' + 25y = 0;\quad y = 2,\quad y' = 11$ при $x = 0$.

15. $y'' + 4y' + 29y = 0;\quad y = 3,\quad y' = 4$ при $x = 0$.

ГЛАВА 7
ОСНОВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Теория вероятностей как научная дисциплина занимается изучением закономерностей в случайных явлениях. Она изучает модели экспериментов, результат которых нельзя предсказать определенно. Погода на завтрашний день, доля отбракованных деталей при их массовом производстве, прогнозирование курса акций при устойчивом финансовом положении рынка и т. п., — все этого предмет приложений теории вероятностей.

§ 7.1. Случайные события и их вероятности

7.1.1. Элементы комбинаторики

Комбинаторика — это раздел математики, изучающий вопросы о том, сколько комбинаций определенного типа можно составить из данных элементов.

Теорема 1 (о числе комбинаций).

Пусть имеется m множеств по $n_1, n_2, \ldots, n_m$ элементов в каждом. Выбрать по одному объекту из каждого множества можно $n_1 \cdot n_2 \cdot \ldots \cdot n_m$ способами.

Пример 1. Из трех классов спортшколы надо составить команду по одному ученику от класса. Сколько различных команд можно составить, если в классах 18, 20 и 22 ученика соответственно.

Ответ на вопрос примера 1 вытекает из теоремы 1 при $m = 3$, $n_1 = 18$, $n_2 = 20$, $n_3 = 22$. Всего можно составить $18 \cdot 20 \cdot 22 = 7920$ команд.

Определение. Множество из n элементов называется *упорядоченным*, если каждому его элементу поставлен в соответствие свой номер от 1 до n.

Если $n = \infty$, то множество называется *счетным*.

Определение. Упорядоченные наборы, составленные из всех элементов данного конечного множества, называются *перестановками* этого множества.

Пример 2. Множество из трех элементов $\{1, 2, 3\}$ имеет следующие перестановки: $(1, 2, 3)$, $(1, 3, 2)$, $(2, 1, 3)$, $(2, 3, 1)$, $(3, 1, 2)$, $(3, 2, 1)$. Всего их шесть.

Теорема 2 (о числе перестановок).

Число всех перестановок множества из n элементов определяется формулой $P_n = n!$, где $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ (считаем, что по определению $P_0 = 0! = 1$).

Пример 3. Цифры 0, 1, 2, 3 записаны на 4 карточках. Сколько различных 4-значных цифр можно составить из этих карточек?

Решение. Число различных комбинаций из 4 цифр равно $4!$; $3!$ комбинаций, начинающихся с 0, следует исключить. В результате количество различных 4-значных чисел равно $4! - 3! = 18$.

Определение. Упорядоченные наборы, состоящие из k элементов, взятых из данных n элементов, называются *размещениями* из n элементов по k.

Размещения отличаются друг от друга либо элементами, либо порядком.

Теорема 3 (о числе размещений).

Число всех размещений из n элементов по k определяется формулой $A_n^k = n(n-1)(n-2) \cdot \ldots \cdot (n-k+1) = \dfrac{n!}{(n-k)!}$.

Например, из множества $\{1,\ 2,\ 3\}$ по два элемента можно (с учетом порядка) выбрать $A_3^2 = 6$ способами: $(1, 2)$, $(2, 1)$, $(1, 3)$, $(3, 1)$, $(2, 3)$, $(3, 2)$.

При $k = n$ из теоремы 3 следует теорема 2, т. е. $P_n = A_n^n$.

Пример 4. Сколько имеется вариантов занятия трех призовых мест 8 спортсменами одного уровня?

Ответ дается формулой $A_8^3 = 336$.

Определение. Неупорядоченные наборы, состоящие из k элементов, взятых из данных n элементов, называются *сочетаниями* из n элементов по k.

Сочетания отличаются друг от друга лишь элементами.

Так, для множества $\{1, 2, 3\}$ сочетаниями по два элемента являются комбинации $(1,2)$, $(1,3)$, $(2,3)$.

Теорема 4 (о числе сочетаний).

Число всех сочетаний из n элементов по k определяется формулой $C_n^k = A_n^k / P_k = \dfrac{n!}{k!(n-k)!}$.

Коэффициенты C_n^k называют *биномиальными коэффициентами*, так как они входят в формулу бинома Ньютона:

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}.$$

Свойства коэффициентов C_n^k:

$$1)\ \sum_{k=0}^{n} C_n^k = 2^n, \quad 2)\ C_n^k = C_n^{n-k}, \quad 3)\ C_n^k = C_{n-1}^{k-1} + C_{n-1}^k.$$

Пример 5. В турнире участвует 6 команд. Каждая команда должна сыграть с каждой одну игру. Сколько игр сыграно в турнире?

Решение. Различные пары команд образуют сочетания из 6 по 2, поскольку порядок среди двух команд, играющих в одной игре, нам безразличен. Следовательно, число игр будет равно $C_6^2 = \dfrac{6!}{2! \cdot 4!} = 15$.

Пример 6. Сколькими способами можно выбрать 2 ампулы из упаковки, содержащей 10 ампул.

Решение. По аналогии с предыдущим примером,

$$C_{10}^2 = \frac{A_{10}^2}{P_2} = \frac{10 \cdot 9}{1 \cdot 2} = 45 \text{ способами.}$$

Пример 7. У 6 мальчиков и 11 девочек имеются признаки инфекционного заболевания. Чтобы проверить наличие заболевания, требуется взять выборочный анализ крови у 2 мальчиков и 2 девочек. Сколькими способами можно это сделать?

Решение. Количество способов выбора двух мальчиков:

$$n_1 = C_6^2 = \frac{6!}{2! \cdot 4!} = 15 \text{ способов.}$$

Количество способов выбора двух девочек:

$$n_2 = C_{11}^2 = \frac{11!}{2! \cdot 9!} = 55 \text{ способов.}$$

Поскольку каждая пара мальчиков может быть взята с каждой парой девочек, то по теореме 1 имеем

$$N = n_1 \cdot n_2 = 15 \cdot 55 = 825 \text{ способов.}$$

7.1.2. Случайные события и их свойства

В основе теории вероятностей лежит понятие случайного эксперимента, т. е. эксперимента, который может закончиться любым из некоторого множества известных результатов, но заранее (до осуществления эксперимента) не известно, каким именно.

Так, бросая игральную кость, нельзя быть заранее уверенным, что выпадает грань с номером 6.

В дальнейшем, *исходами* будем называть различные результаты эксперимента.

Определение. *Пространством элементарных исходов* называется множество всех взаимно исключающих (т. е. не могущих наступить одновременно) исходов эксперимента.

Пространство элементарных исходов мы будем обозначать буквой Ω.

Событие может состоять из одного или нескольких исходов, а может включать счетное или даже несчетное число исходов. События будут обозначаться большими буквами A, B, ... или же словами.

Будем говорить, что *событие A наступило*, если эксперимент заканчивается одним из исходов, входящих в событие A.

Пример 1. Бросание игральной кости.

Здесь $\Omega = \{1, 2, \dots, 6\}$ состоит из шести исходов, заключающихся в выпадении грани с соответствующим номером. Событие $A = \{\text{выпало четное число}\}$ состоит из трех исходов, т. е. $A = \{2, 4, 6\}$. Считаем, что A наступило, если при бросании игральной кости выпала грань с номером 2, 4 или 6.

Пример 2. Стрельба по мишени до первого попадания.

Здесь $\Omega = \{п, нп, ннп, нннп, \ldots\}$, где исход $п = \{$попадание при первом выстреле$\}$, а исход $н \ldots нп$, где $н$ встречается k раз означает, что стрелок первые k раз промахнулся, а в $(k+1)$-й раз мишень была поражена. В данном случае, пространство элементарных исходов Ω является счетным множеством.

Заметим, возвращаясь несколько назад, что выбор Ω субъективен и диктуется условиями задачи, а следовательно, может быть разным в одном и том же эксперименте.

Так, если в условиях примера 1 нас интересует лишь, четное или нечетное число очков выпало, то логично в качестве Ω взять $\{A, B\}$, где $A = \{2, 4, 6\}$, $B = \{1, 3, 5\}$.

Определение. *Суммой* $A \cup B$ ($A + B$; A или B) называют событие, состоящее из всех исходов, принадлежащих либо A, либо B.

Определение. *Произведением* AB двух событий A и B называют событие, состоящее только из тех исходов, которые входят в A и B одновременно.

Определение. *Разностью* $A - B$ двух событий A и B называется событие, состоящее из исходов, входящих в A, но не входящих в B.

Если события изобразить множествами на плоскости, то результат определенных выше операций над событиями выглядит следующим образом:

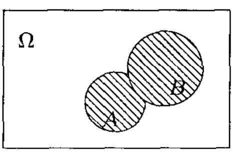


Рис. 7.1. Сумма событий $A + B$.

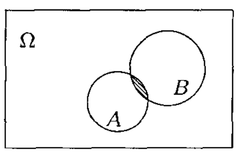


Рис. 7.2. Произведение событий $A \cdot B$.

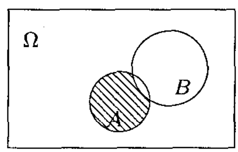


Рис. 7.3. Разность событий $A - B$.

Пример 3. Пусть при бросании игральной кости (см. пример 1) событие $A = \{1, 3, 5\} = \{$выпало нечетное число очков$\}$, а событие $B = \{3, 6\} = \{$выпало число очков, кратное трем$\}$. Тогда $A + B = \{1, 3, 5, 6\}$, $AB = \{3\}$, $A - B = \{1, 5\}$, $B - A = \{6\}$.

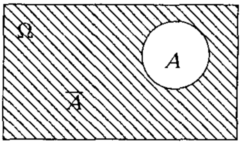


Рис. 7.4. События A и $\overline{A}$.

Определение. *Противоположным* (дополнительным) для события A называется событие $\overline{A}$, состоящее из всех исходов, не входящих в A. Отметим, что $\overline{A} = \Omega - A$.

Пример 4. Противоположными событиям A и B из примера 2 являются события $\overline{A} = \{1, 3, 5\}$ и $\overline{B} = \{2, 4, 6\}$.

Определение. Событие Ω называется *достоверным* (оно обязательно происходит). Событие $\varnothing = \overline{\Omega}$ называется *невозможным*.

Примером невозможного события при бросании игральной кости на ровную поверхность является событие, заключающееся в том, что игральная кость встала на вершину.

Определение. События A и B называются *несовместными*, если они не могут наступать одновременно.

Примером несовместных событий служат события $A - B$ и B.

Отметим, что если $AB = \varnothing$, то события A и B обязательно несовместны.

Определение. Говорят, что событие A содержится в событии B ($A \subset B$), если все исходы события A входят в событие B.

Так, $A \subset A + B$, $AB \subset A$.

Приведем некоторые свойства операций над событиями :

1) $A + B = B + A$, 2) $AB = BA$, 3) $A + \overline{A} = \Omega$,

4) $A\Omega = A$, 5) $AB \subset A$, 6) $A\overline{A} = \varnothing$, 7) $\overline{\overline{A}} = A$,

8) $A - B = A\overline{B}$, 9) $(A + B)C = AC + BC$, 10) $\overline{A + B} = \overline{A}\,\overline{B}$,

11) $\overline{AB} = \overline{A} + \overline{B}$, 12) $(A - B)C = AC - BC$.

Равенства 9–11 можно распространить на более общий случай:

13) $B \sum\limits_{k=1}^{n} A_k = \sum\limits_{k=1}^{n} BA_k$, 14) $\overline{\sum\limits_{k=1}^{n} A_k} = \prod\limits_{k=1}^{n} \overline{A_k}$,

15) $\overline{\prod\limits_{k=1}^{n} A_k} = \sum\limits_{k=1}^{n} \overline{A_k}$.

Эти равенства верны и при $n = \infty$.

Следует иметь в виду, что событие $\sum\limits_{k=1}^{n} A_k$ наступает, если происходит «хотя бы одно» событие из $A_1, A_2, \ldots, A_n$, а наступление события $\prod\limits_{k=1}^{n} \overline{A_k}$ означает, что ни одного события из A_1, $A_2, \ldots, A_n$ не произошло.

7.1.3. Классическое определение вероятностей

Основы теории вероятностей будем изучать, рассматривая случайные эксперименты с конечным или счетным числом исходов. Пусть пространство элементарных исходов Ω имеет вид $\Omega = \{\omega_1, \ldots, \omega_n\}$, где $n \leqslant \infty$. Поставим в соответствие каждому исходу ω_i неотрицательное число $P(\omega_i)$, так чтобы было выполнено условие нормировки

$$\sum_{i=1}^{n} P(\omega_i) = 1. \tag{7.1}$$

Тем самым мы задаем на Ω *распределение вероятностей.*

Определение. Вероятностью события $A = \{\omega_{l_1}, \omega_{l_2}, \ldots, \omega_{l_k}\}$ называется число $P(A)$, равное сумме вероятностей исходов, составляющих событие A, т. е.

$$P(A) = \sum_{1 \leqslant l \leqslant n} p(\omega_{i_l}) = \sum_{\omega_l \in A} p(\omega_l). \tag{7.2}$$

Далее рассмотрим специальную модель теории вероятностей, носящую название «классическая схема». Эта схема возникает тогда, когда в силу некоторых свойств, обычно связанных с симметрией, каждый исход эксперимента равновозможен. В таком случае вероятность каждого исхода одинакова и в силу условия (7.1) равна $1/n$. Действительно, если $p(\omega_1) = \ldots = p(\omega_n) = p$, то $p + \ldots + p = np = 1 \Rightarrow p = 1/n$.

Если для любого множества A символом $|A|$ обозначить число его элементов, то в «классической схеме» из равенства (7.2) следует

$$P(A) = p \cdot |A| = \frac{|A|}{|\Omega|}.$$

Другими словами, в случае, когда все исходы равновозможны, вероятность события A равна отношению числа m благоприятных исходов для события A к общему числу n исходов эксперимента:

$$P(A) = \frac{m}{n}. \tag{7.3}$$

Вероятность (7.3) часто называют «классической».

Пример 1. При бросании правильной (т. е. симметричной) монеты $|\Omega| = 2$, а вероятности выпадения герба или решетки одинаковы и равны $1/2$.

При бросании игральной кости $|\Omega| = 6$, а вероятности выпадения любой грани одинаковы и равны $1/6$;

$$P(\textit{выпало четное число очков}) = P(\{2, 4, 6\}) = 3/6 = 1/2.$$

Пример 2. Какова вероятность угадать в лотерее 6 чисел из 49?

Решение. Здесь Ω =*совокупность всех сочетаний из* 49 *чисел по* 6, $|\Omega| = C_{49}^6$. Благоприятствует выигрышу лишь одно событие. Поэтому $P(\textit{угадать } 6 \textit{ из } 49) = 1/C_{49}^6 = (13983816)^{-1}$.

Пример 3. В клетке имеется 10 мышей: 3 белых и 7 серых. Наудачу вытаскивается одна мышь. Какова вероятность события A, что эта мышь белая.

Решение. В этом примере $n = |\Omega| = 10$, а $m = |A| = 3$. В соответствии с формулой (7.3) вероятность вытащить белую мышь равна $3/10$.

Замечание. Слово «наудачу», употребленное в примере 3, является синонимом выражения «случайным образом» и означает выбор, при котором все комбинации элементов, которые могут быть выбраны, равновозможны.

Пример 4. Технический контроль из партии в 10 изделий проверяет взятые наудачу 3 изделия. Партия не принимается, если среди трех проверяемых изделий окажется хотя бы одно бракованное. Определить вероятность приемки партии, если в ней окажется 5 бракованных изделий.

Решение. Обозначим через A событие, состоящее в том, что партия изделий будет принята. Общее число исходов, состоящее в проверке трех наудачу взятых изделий из 10, равно $n = C_{10}^3$, т. е. числу сочетаний из 10 по 3. Событию A благоприятствуют

$m = C_5^3$ исходов. Таким образом, получаем $P(A) = C_5^3/C_{10}^3 = 1/12$.

Пример 5. При наборе телефонного номера абонент забыл две последние цифры и наудачу набрал их, помня только, что эти цифры нечетные и различные. Найти вероятность того, что номер набран правильно.

Решение. В этой задаче порядок набора цифр важен. Поэтому для определения числа всех возможных исходов следует использовать теорему 3. Количество благоприятных исходов здесь равно 1, так как правильным является только один номер. Следовательно, искомая вероятность равна $P(A) = 1/A_5^2 = 1/20$.

Пример 6. Пусть в группе из 10 человек — четверо мужчин. Если случайным образом выбирают двух человек, то какова вероятность, что это:

а) оба мужчины, б) обе женщины, в) один мужчина и одна женщина.

Решение. Обозначим события:

A — выбирают двоих мужчин, B — выбирают двух женщин, C — выбирают одного мужчину и одну женщину.

Число всех исходов

$$|\Omega| = n = C_{10}^2 = \frac{10!}{2! \cdot 8!} = 45.$$

Число выбора двоих мужчин из четырех и двух женщин из шести:

$$m_1 = C_4^2 = 6, \quad m_2 = C_6^2 = 15.$$

По формуле (7.3) имеем

$$P(A) = \frac{6}{45} = \frac{2}{15}, \quad P(B) = \frac{15}{45} = \frac{1}{3}.$$

Наконец, существует 4 способа выбора одного мужчины и 6 способов выбора одной женщины. По теореме 1 о числе комбинаций количество способов выбора двух человек обоего пола равно $m_3 = 6 \cdot 4 = 24$. Таким образом, $P(C) = 24/45 = 8/15$.

Пример 7. Из 20 человек, одновременно заболевших гриппом, 15 выздоровели полностью за 3 дня. Предположим, что из этих 20 человек случайным образом выбирают 5. Какова вероятность того, что за 3 дня из этих человек выздоравливают:

а) 5 человек, б) 4 человека, в) никто не выздоравливает.

Решение. Определим события:

A — за 3 дня выздоравливают 5 человек.

B — за 3 дня выздоравливают 4 человека.

C — за 3 дня не выздоравливает никто.

Число всех исходов равно $n = C_{20}^5 = \dfrac{20 \cdot 19 \cdot 18 \cdot 17 \cdot 16}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 15504$.

Число способов выбора из 15 выздоровевших пяти человек — $m_1 = C_{15}^5 = 3003$. Таким образом, $P(A) = 3003/15504 = 0{,}194$.

Число способов выбора из 15 выздоровевших четырех человек — $C_{15}^4 = 1365$, одного невыздоровевшего из пяти — $C_5^1 = 5$. По теореме 1, число способов выбора 5 человек, из которых выздоравливают за 3 дня только четверо, равно $m_2 = C_{15}^4 \cdot C_5^1 = 1365 \cdot 5 = 6825$. Следовательно $P(B) = 6825/15504 = 0{,}440$.

Число способов выбора пяти невыздоровевших из пяти человек равно 1, откуда $P(C) = 1/15504 = 0{,}00006$.

Пример 8. На полке 6 книг по математике и 4 по физике. Найти вероятность того, что среди выбранных наудачу пяти книг будут три по математике и две по физике.

Решение. Три книги по математике можно выбрать $m_1 = C_6^3 = 20$ способами, а две книги по физике — $m_2 = C_4^2 = 6$ способами.

Пять книг, среди которых три по математике и две по физике, по теореме о числе комбинаций можно выбрать $m = m_1 \cdot m_2 = 20 \cdot 6 = 120$ способами.

Число всех исходов равно $n = C_{10}^5 = 252$. Искомая вероятность по формуле (7.3) равна $P = 120/252 = 10/21$.

Пример 9. В клетке 6 белых и 4 серых мыши. Случайным образом извлекают 3 мыши. Вычислить вероятность для 4 возможных комбинаций цвета мышей.

Решение. Обозначим: *б* – белая мышь, *с* – серая мышь, A_1, A_2, A_3, A_4 — события, состоящие в появлении одной из возможных комбинаций цвета мышей: *ббб*, *ссс*, *бсс*, *ббс*.

Число всех исходов равно $n = C_{10}^3 = \dfrac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3} = 120$ (всего мышей 10, выбираем 3).

Число способов выбора из шести белых мышей трех равно

$$m_1 = C_6^3 = \frac{6 \cdot 5 \cdot 4}{1 \cdot 2 \cdot 3} = 20,$$

и

$$P(A_1) = \frac{20}{120} = \frac{1}{6}.$$

Число способов выбора из четырех серых мышей трех равно

$$m_2 = C_4^3 = \frac{4 \cdot 3 \cdot 2}{1 \cdot 2 \cdot 3} = 4,$$

и

$$P(A_2) = \frac{4}{120} = \frac{1}{30}.$$

Событие A_3 состоит в том, что из шести белых мышей нужно выбрать одну — это $C_6^1 = 6$ способов, и из четырех серых мышей выбрать две — это $C_4^2 = 6$ способов. Общее количество вариантов будет равно $m_3 = C_6^1 \cdot C_4^2 = 36$, а $P(A_3) = 36/120 = 3/10$.

Рассуждая аналогично, найдем общее число способов выбора двух белых и одной серой мыши: $m_4 = C_6^2 \cdot C_4^1 = 15 \cdot 4 = 60$, и $P(A_4) = 60/120 = 1/2$.

Пример 10. В коробке 5 одинаковых изделий, причем 3 из них окрашены. Наудачу извлечены 2 изделия. Найти вероятность того, что среди отобранных изделий:

а) одно окрашено, б) два окрашены, в) хотя бы одно окрашено.

Решение. Обозначим события:

A — окрашено одно изделие из двух извлеченных,

B — окрашены два,

C — оба не окрашены.

Число всех исходов определяется как $n = C_5^2 = \dfrac{5 \cdot 4}{1 \cdot 2} = 10$.

а) Так как окрашено одно изделие, то второе не окрашено, и, следовательно, $m_1 = C_3^1 \cdot C_2^1 = 3 \cdot 2 = 6$, а $P(A) = 6/10 = 0{,}6$.

б) Поскольку из двух извлеченных изделий окрашены два, то неокрашенных нет вовсе и $m_2 = C_3^2 = 3$, а $P(B) = 3/10 = 0{,}3$.

в) Здесь $m = C_2^2 = 1$ и $P(C) = m/n = C_2^2/C_5^2 = 0{,}1$.

Пример 11. Из пяти карточек с буквами *А*, *Б*, *В*, *Г*, *Д* наугад выбираются три и располагаются в ряд в порядке появления. Какова вероятность того, что получится слово «два».

Решение. Так как положение буквы в комбинации имеет значение, то для определения общего числа возможных исходов следует пользоваться размещениями

$$n = A_5^3 = 5 \cdot 4 \cdot 3 = 60.$$

Здесь $m = 1$, так как требуется получить только одно слово «два». Следовательно, $p = 1/60$.

7.1.4. Свойства вероятностей

1. $\mathrm{P}(\Omega) = 1$ в силу условия нормировки (7.1).
2. $\mathrm{P}(\varnothing) = 0$, так как в (7.2) сумма не содержит слагаемых.
3. Если $A \subset B$, то $P(A) \leqslant P(B)$ (в сумме (7.2), определяющей $P(B)$, присутствуют дополнительные по сравнению с $P(A)$ неотрицательные слагаемые).
4. $0 \leqslant P(A) \leqslant 1$, что следует из свойств 3 и 1.
5. $P(\overline{A}) = 1 - P(A)$, так как по свойству 1

$$\sum_{\omega_e \in \overline{A}} P(\omega_e) + \sum_{\omega_e \in A} P(\omega_e) = \sum_{\omega_e \in \Omega} P(\omega_e) = 1.$$

Свойство 5 важно в практическом отношении, так как позволяет вычислить вероятность события A, если известна вероятность противоположного события $\overline{A}$.

Пример 1. Из полного набора костей домино наугад берутся пять костей. Найти вероятность того, что среди них будет хотя бы одна с шестеркой.

Решение. При решении этой задачи удобно использовать свойство 5 вероятностей, перейдя к противоположному событию «среди пяти отобранных костей домино нет ни одной шестерки».

Число всех исходов

$$n = C_{28}^5 = \frac{28 \cdot 27 \cdot 26 \cdot 25 \cdot 24}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 98280.$$

Поскольку количество костей домино, не содержащих шестерку, равно 21, то число исходов, благоприятствующих противоположному событию, будет

$$m = C_{21}^5 = \frac{21 \cdot 20 \cdot 19 \cdot 18 \cdot 17}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 20349.$$

Отсюда (см.(7.3)) $q = \frac{m}{n} = \frac{20349}{98280} = 0{,}207$ и $p = 1 - q = 0{,}793$.

Из свойства 5 вероятностей и свойства 13 операций над событиями вытекает следующее соотношение:

для любого набора событий A_i, $i = 1, 2, \ldots, n$,

$$P\left(\sum_{i=1}^{n} A_i\right) = 1 - P\left(\prod_{i=1}^{n} \overline{A_i}\right). \tag{7.4}$$

Равенство (7.4) означает, что вероятность того, что произошло «хотя бы одно» событие из $A_1, A_2, \ldots, A_n$, можно найти, вычитая из 1 вероятность того, что не произошло ни одного события из $A_1, A_2, \ldots, A_n$. В некоторых случаях таким приемом удобно пользоваться.

Для любых событий A и B имеет место *формула сложения вероятностей*:

$$P(A + B) = P(A) + P(B) - P(AB). \tag{7.5}$$

При проверке равенства (7.5) полезно воспользоваться геометрической иллюстрацией событий $A + B$ и AB (см. рис. 7.1, 7.2), интерпретируя вероятности событий как площади соответствующих им изображений.

Пример 2. В марте 7 дней шел снег, 10 — дождь, из них 4 дня — снег с дождем. Найти вероятность того, что в наугад выбранный день шел дождь или снег.

Решение. Пусть $A = \{$*шел снег*$\}$ и $B = \{$*шел дождь*$\}$. По формуле (7.5) $P($*шел дождь или снег*$) = P(A) + P(B) - P(AB) = 7/31 + 10/31 - 4/31 = 13/31$.

Пример 3. Из 100 студентов, находящихся в аудитории, 50 человек знают английский язык, 40 — французский, 35 — немецкий. Английский и французский знают 20 человек, английский и немецкий — 8, французский и немецкий — 10. Все три языка знают 5 человек. Один из студентов вышел. Определить вероятности следующих событий:

1. Вышедший студент знает английский и французский языки.

Решение. Введем события:

A — знает английский;

B — знает французский.

$$P(D) = P(A+B) = P(A)+P(B)-P(AB) = 0{,}5+0{,}4-0{,}2 = 0{,}7.$$

2. Знает английский или немецкий языки.

Решение. Событие:

C — знает немецкий.

$$P(D) = P(A+C) = P(A) + P(C) - P(AC) = \\ = 0{,}35 + 0{,}5 - 0{,}08 = 0{,}77.$$

3. Знает немецкий или французский языки.

Решение.

$$P(D) = P(C+B) = P(C) + P(B) - P(CB) = \\ = 0{,}35 + 0{,}4 - 0{,}1 = 0{,}65.$$

4. Кроме английского языка знает или французский, или немецкий.

Решение.

$$P(D) = P(AB + AC) = P(AB) + P(AC) - P(ABC) = \\ = 0{,}2 + 0{,}08 - 0{,}05 = 0{,}23.$$

Из формулы сложения вероятностей можно вывести несколько следствий.

Следствие 1. Если события A и B несовместны, то

$$P(A+B) = P(A) + P(B). \tag{7.6}$$

Пример 4. Определить вероятность того, что партия из 15 изделий, среди которых 5 бракованных, будет принята после выборочного контроля 4 вместе взятых изделий. Условием приемки является обнаружение не более одного бракованного среди 4 проверяемых изделий.

Решение. Введем обозначения:

A — событие, состоящее в том, что партия изделий принята;

B — все четыре проверяемые изделия качественные;

C — среди четырех проверяемых одно изделие бракованное.

Очевидно, что $A = B + C$ (причем события B и C несовместны),

$$P(B) = \frac{C_{10}^4}{C_{15}^4} = \frac{2}{13}, \quad P(C) = 5 \cdot \frac{C_{10}^3}{C_{15}^4} = \frac{40}{91},$$

$$P(A) = P(B) + P(C) = \frac{2}{13} + \frac{40}{91} = \frac{54}{91}.$$

Следствие 2. Если события A_i, $i = 1, 2, \ldots, n$, попарно несовместны, т. е. $A_i A_j = \varnothing$ для любых $i \neq j$, то

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i). \tag{7.7}$$

Пример 5. В клетке находится 10 мышей, причем 4 из них серые. Наугад берут трех мышей. Найти вероятность того, что хотя бы одна мышь серая.

Решение. Определим события:

A — одна мышь серая, две — нет;

B — две серые, одна — нет;

C — все серые;

D — ни одной серой;

E — хотя бы одна мышь серая.

События A, B и C являются несовместными, поэтому

$$P(E) = P(A + B + C) = P(A) + P(B) + P(C) =$$
$$= \frac{C_4^1 \cdot C_6^2}{C_{10}^3} + \frac{C_4^2 \cdot C_6^1}{C_{10}^3} + \frac{C_4^3}{C_{10}^3} = \frac{5}{6},$$

или (с учетом того, что $A + B + C + D = \Omega$)

$$P(A + B + C) = 1 - P(D) = 1 - \frac{C_6^3}{C_{10}^3} = \frac{5}{6}.$$

7.1.5. Статистическое определение вероятностей

Классическое определение вероятностей, по существу, сводит понятие «вероятности» к понятию «равновозможности». Однако «равновозможность», представляя собой объективное свойство испытаний, определяемое условиями эксперимента, может быть установлена лишь с известной степенью точности.

Наше представление о симметричных игральных костях, монетах и т. п. было бы только иллюзией, не подтверждай правоту сделанных предположений данные опыта.

Многочисленные эксперименты открыли путь для *статистического* подхода к численному определению вероятностей, который особенно важен тогда, когда из теоретических соображений, подобных соображениям симметрии, значение вероятности заранее установить нельзя.

В соответствии с этим подходом вероятностью события A называется объективно существующая величина, около которой группируются относительные частоты наступления этого события при неограниченном увеличении числа испытаний.

Другими словами, $P(A) = \lim\limits_{n\to\infty} \dfrac{m_A}{n}$, где n — общее число испытаний, а m_A — число испытаний, в которых событие A наступило. Разумеется, условия, при которых проводится эксперимент, не должны изменяться от опыта к опыту.

Близость частоты и вероятности при многократном повторении одних и тех же экспериментов объясняется так называемым законом больших чисел.

Отметим, что статистическое и классическое определение вероятностей, как, впрочем, и все остальные (геометрическое, аксиоматическое) согласованы, т. е. при решении одних и тех же задач приводят к одинаковым результатам.

Пример 1. Еще в XVIII веке было замечено, что среди обычной корреспонденции письма без адреса обладают определенной устойчивостью, приблизительно 25–27 на каждый миллион.

Пример 2. Относительная частота рождений мальчиков не зависит от местности или этнического состава населения и приблизительно равняется 0,515.

Пример 3. При бросании симметричной монеты вероятность выпадения герба на основании классического определения будет, очевидно, равна 0,5. Английский статистик К. Пирсон (1857–1936) бросил симметричную монету 24 000 раз, при этом герб выпал 12 012 раз, следовательно, относительная частота выпадения герба оказалась равной 0,5005, т. е. как и ожидалось, оказалась примерно равной 0,5.

7.1.6. Условные вероятности. Независимость событий

Рассмотрим вопрос о том, как определить вероятность какого-либо события A при условии, что уже произошло другое событие B.

Пусть, для примера, брошена игральная кость и нам неизвестен результат, однако известно, что выпало четное число очков. Мы же хотим, располагая этой информацией, найти вероятность того, что выпало число больше трех. Речь, следовательно, идет об условной вероятности события $A = \{$*выпало число больше трех*$\}$ при условии, что произошло событие $B = \{$*выпало четное число*$\}$ (обозначается $P(A/B)$).

Поскольку нам известно, что выпало либо 2, либо 4, либо 6 очков и эти исходы равновозможны, а в такой ситуации событию A благоприятствуют лишь исходы 4 и 6, естественно считать, что $P(A/B) = 2/3$. Формально это можно объяснить следующим образом: так как событие B произошло, мы в качестве нового пространства элементарных исходов берем все исходы, которые принадлежат B, а благоприятствующими наступлению событию A считаем исходы, входящие в A и B одновременно.

Таким образом, по формуле классической вероятности

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

Принимая во внимание то, что последнее равенство можно переписать в виде

$$P(A/B) = \frac{|AB| : |\Omega|}{|B| : |\Omega|} = \frac{P(AB)}{P(B)}, \qquad P(B) > 0,$$

приходим к следующему, согласующемуся с интуицией, определению.

Определение. Условная вероятность события A при условии, что произошло событие B с $P(B) > 0$, задается формулой

$$P(A/B) = \frac{P(AB)}{P(B)}. \tag{7.8}$$

В уже рассмотренном примере $AB = \{4, 6\} \Rightarrow P(A/B) = = \dfrac{2/6}{3/6} = \dfrac{2}{3}$, как и было сказано.

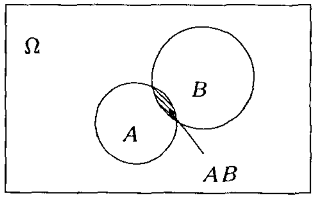


Рис. 7.5. Иллюстрация к процедуре вычисления условной вероятности $P(A/B)$.

Из формулы (7.8) следует, что условная вероятность — это обычная вероятность, заданная на более узком пространстве элементарных исходов, поэтому для нее справедливы все свойства обычной вероятности.

В частности,

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

$$P(A + C/B) = P(A/B) + P(C/B) - P(AC/B)$$

(формула сложения условных вероятностей).

Из равенства (7.8) вытекает *формула умножения вероятностей*:

$$P(AB) = P(B)\,P(A/B) = P(A)\,P(B/A). \tag{7.9}$$

Пример 1. В ящике m белых и n черных шаров. Какова вероятность вытащить подряд два белых шара?

Решение. Событие, вероятность которого требуется вычислить, можно представить в виде произведения событий AB, где $A = \{\textit{первый шар белый}\}$, $B = \{\textit{второй шар белый}\}$. Вероятность события A вычисляется по формуле (7.3) и равна $P(A) = \frac{m}{n+m}$, условная вероятность $P(B/A) = \frac{m-1}{n+m-1}$ (если A произошло, то в ящике осталось $(m-1)$ белых шаров, а всего — $(m+n-1)$ шаров). Остается воспользоваться формулой умножения вероятностей (7.9):

$$P(\textit{оба шара белые}) = P(AB) = \frac{m(m-1)}{(m+n)(m+n-1)}.$$

Тот же результат может быть получен с помощью формулы (7.3):

$$P(\textit{оба шара белые}) = \frac{C_m^2}{C_{m+n}^2}.$$

Формулу умножения вероятностей можно обобщить на случай n произвольных событий $A_1, A_2, \ldots, A_n$:

$$P(A_1A_2\ldots A_n) = P(A_1)\cdot P(A_2/A_1)\times \\ \times P(A_3/A_1A_2)\cdot\ldots\cdot P(A_n/A_1A_2\ldots A_{n-1}). \quad (7.10)$$

Определение. События A и B называются *независимыми*, если

$$P(AB) = P(A)P(B). \quad (7.11)$$

Отсюда и из формулы умножения вероятностей следует, что $P(A/B) = P(A)$ и $P(B/A) = P(B)$, т. е. вероятность наступления события A не зависит от того, произошло B или нет (и наоборот). Так, если в ситуации примера 1, вытащив белый шар, вернуть его в ящик, то $P(B/A) = \dfrac{n}{n+m} = P(B)$.

Заметим, что из формулы сложения вероятностей (7.5) и из (7.11) следует, что если события A и B являются независимыми, то

$$P(A+B) = P(A) + P(B) - P(A)\,P(B). \quad (7.12)$$

Определение. События $A_1, A_2, \ldots, A_n$ называются *взаимно-независимыми*, если для любого набора индексов $1 \leqslant l_1 < < l_2 < \ldots < l_k \leqslant n$ выполняется равенство

$$P(A_{l_1}A_{l_2}\ldots A_{l_k}) = P(A_{l_1})\cdot P(A_{l_2})\cdot\ldots\cdot P(A_{l_k}). \quad (7.13)$$

Следствие. Если события $A_1, A_2, \ldots, A_n$ взаимно-независимы, то

$$P\left(\sum_{i=1}^{n} A_i\right) = 1 - \prod_{i=1}^{n}(1-P(A_i)). \quad (7.14)$$

С помощью формулы (7.14), которая вытекает из (7.4) и (7.13), очень просто сосчитать вероятность наступления хотя бы одного из n взаимно-независимых событий.

Отметим, что понятие независимости отвечает всем интуитивным представлениям о независимости. В частности, если независимы события A и B, то независимы также события A и $\overline{B}$, $\overline{A}$ и $\overline{B}$ и т. д.

Пример 2. Некоторая вакцина эффективна на 75% в формировании иммунитета. Вакцинировали 2 человек. Пусть A и B — независимые события, состоящие в том, что соответственно первый и второй человек приобретают иммунитет. Найти вероятность того, что:

а) оба человека приобрели иммунитет;

б) первый приобрел иммунитет, а второй нет;

в) оба не приобрели иммунитета.

Решение.

а) $P(AB) = P(A) \cdot P(B) = 0{,}75 \cdot 0{,}75 = (0{,}75)^2$;

б) $P(A\overline{B}) = P(A) \cdot (1 - P(B)) = 0{,}75 \cdot 0{,}25$;

в) $P(\overline{A}\overline{B}) = (1 - P(A)) \cdot (1 - P(B)) = 0{,}25 \cdot 0{,}25 = (0{,}25)^2$.

7.1.7. Формула полной вероятности и формула Байеса

Определение. Набор попарно несовместных событий H_1, $H_2, \ldots, H_n$, сумма которых составляет достоверное событие, называется *полной группой событий.*

Теорема 1 (формула полной вероятности).

Пусть события H_1, $H_2, \ldots, H_n$ являются полной группой событий. Тогда для любого события A

$$P(A) = \sum_{j=1}^{n} P(H_j)P(A/H_j). \tag{7.15}$$

Доказательство. В силу свойств операций над событиями (см. п. 7.2)

$$A = A\Omega = A(H_1 + \ldots + H_n) = AH_1 + \ldots + AH_n.$$

Поскольку события H_j попарно несовместны, то и события AH_j обладают тем же свойством. Поэтому в силу следствия 2 (см. (7.7))

$$P(A) = P(AH_1 + \ldots + AH_n) = \sum_{j=1}^{n} P(AH_j).$$

Остается применить к слагаемым $P(AH_j)$ формулу умножения вероятностей.

Пример 1. Команда на хорошем поле выигрывает с вероятностью $2/3$, а на плохом — с вероятностью $1/2$. Известно, что $3/4$ игр проводится на хорошем поле. Какова вероятность выиграть в наудачу выбранном матче?

Решение. Введем события $A = \{\textit{выигрыш}\}$, $H_1 = \{\textit{плохое поле}\}$, $H_2 = \{\textit{хорошее поле}\}$. По условиям примера

$$P(A/H_1) = 2/3, \quad P(A/H_2) = 1/2,$$

$$P(H_1) = 3/4, \quad P(H_2) = 1/4 \Rightarrow$$

$$\Rightarrow P(A) = 2/3 \cdot 3/4 + 1/2 \cdot 1/4 = 5/8.$$

Пример 2. Вероятность брака при изготовлении детали равна 0,04. Приемка деталей производится по следующей системе контроля: годная деталь принимается с вероятностью 0,98, а бракованная — с вероятностью 0,1. Найти вероятность приемки детали.

Решение. Пусть событие A — деталь принята, H_1 — изготовлена хорошая деталь, H_2 — изготовлена бракованная деталь.

Из условий найдем соответствующие вероятности:

$$p(H_1) = 0{,}96, \quad p(H_2) = 0{,}04,$$

$p(A/H_1) = 0{,}98$ — вероятность того, что принятая деталь хорошая,

$p(A/H_2) = 0{,}1$ — вероятность того, что принятая деталь бракованная.

Вероятность того, что деталь будет принята, на основании формулы (7.15) равна

$$P(A) = p(H_1) \cdot p(A/H_1) + p(H_2) \cdot p(A/H_2) =$$
$$= 0{,}96 \cdot 0{,}98 + 0{,}04 \cdot 0{,}1 = 0{,}9448.$$

Пример 3. Имеются три одинаковых ящика с шарами. В первом ящике 30 белых и 20 черных шаров, во втором — 15 белых и 15 черных шаров, в третьем — 5 белых и 15 черных шаров. Какова вероятность вытащить из случайно выбранного ящика черный шар.

Решение. Обозначим события:

H_1 — выбран первый ящик, H_2 — выбран второй ящик, H_3 — выбран третий ящик.

Поскольку выбор того или иного ящика — события равновероятные, то $p(H_1) = p(H_2) = p(H_3) = 1/3$;

$p(A/H_1) = 20/50 = 2/5$ — вероятность того, что черный шар вытащили из первого ящика;

$p(A/H_2) = 15/30 = 1/2$ — вероятность того, что черный шар вытащили из второго ящика;

$p(A/H_3) = 15/20 = 3/4$ — вероятность того, что черный шар вытащили из третьего ящика.

По формуле (7.15) найдем искомую вероятность:

$$P(A) = p(H_1) \cdot p(A/H_1) + p(H_2) \cdot p(A/H_2) + p(H_3) \cdot p(A/H_3) = \\ = \frac{1}{3} \cdot \frac{2}{5} + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{3}{4} = 0{,}55.$$

Пусть имеется полная группа событий $H_1, H_2, \ldots, H_n$, и пусть в результате произведенного испытания появилось некоторое событие A.

В практических приложениях часто требуется высказать суждение о том, какое из событий H_i $(i = 1, \overline{n})$ имеет место. Решить этот вопрос с полной определенностью невозможно. Поэтому ставится такая задача: какую вероятность нужно приписать каждому из событию H_i $(i = 1, \overline{n})$, если известно, что произошло событие A. При этом предполагается, что известны априорные (доопытные) вероятности $P(H_i)$ и условные вероятности $P(A/H_i)$ события A $(i = 1, \overline{n})$.

Таким образом, поставленная задача сводится к переоценке вероятностей событий в связи с наличием нового фактора — наступления события A. Решение задачи дается следующей теоремой.

Теорема 2 (формула Байеса).

Пусть дана полная группа событий H_i $(i = 1, \overline{n})$ и некоторое событие A, имеющее положительную вероятность. Тогда условная вероятность события H_i $(i = 1, \overline{n})$ при условии, что событие

A произошло, определяется по формуле

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{\sum\limits_{i=1}^{n} P(H_i) \cdot P(A/H_i)}. \tag{7.16}$$

Пример 4. В группе 10 студентов, пришедших на экзамен. Трое подготовлены отлично, 4 — хорошо, 2 — посредственно, 1 — плохо. В экзаменационных билетах имеется 20 вопросов. Отлично подготовленный студент может ответить на все 20 вопросов, хорошо подготовленный — на 16, посредственно — на 10, плохо подготовленный — на 5. Вызванный наугад студент ответил на 3 заданных вопроса. Найти вероятность того, что этот студент подготовлен:

1) отлично, 2) плохо.

Решение. Рассмотрим полную группу событий:

H_1 — подготовлен отлично,

H_2 — подготовлен хорошо;

H_3 — подготовлен посредственно;

H_4 — подготовлен плохо.

Вероятности этих событий равны:

$$P(H_1) = 0{,}3, \quad P(H_2) = 0{,}4, \quad P(H_3) = 0{,}2, \quad P(H_4) = 0{,}1.$$

Найдем соответствующие вероятности:

$$P(A/H_1) = 1, \quad P(A/H_2) = \frac{16}{20} \cdot \frac{15}{19} \cdot \frac{14}{18} = 0{,}491,$$

$$P(A/H_3) = \frac{10}{20} \cdot \frac{9}{10} \cdot \frac{8}{18} = 0{,}105, \quad P(A/H_4) = \frac{5}{20} \cdot \frac{4}{19} \cdot \frac{3}{18} = 0{,}009,$$

$$P(A) = 0{,}3 \cdot 0{,}1 + 0{,}491 \cdot 0{,}4 + 0{,}105 \cdot 0{,}2 + 0{,}009 \cdot 0{,}01 = 0{,}518.$$

По формуле (7.16) найдем:

1) $P(H_1/A) = 0{,}3 \cdot \dfrac{0{,}1}{0{,}518} = 0{,}58;$

2) $P(H_4/A) = 0{,}1 \cdot \dfrac{0{,}009}{0{,}518} = 0{,}002.$

Пример 5. В первой коробке 5 белых шаров и 4 черных, во второй — 3 белых и 1 черный, в третьей — 2 белых и 1 черный. Наугад выбирается коробка и из нее два шара. Оказалось, что оба шара белые. Найти условную вероятность того, что шары извлекались из первой коробки.

Решение. Имеют место следующая полная группа событий:

H_1 — выбрана первая коробка;

H_2 — выбрана вторая коробка;

H_3 — выбрана третья коробка.

Выбор каждой из коробок равновероятен:

$$p(H_1) = p(H_2) = p(H_3) = \frac{1}{3}.$$

Найдем условные вероятности.

1 коробка:

$$n = C_9^2 = \frac{9 \cdot 8}{1 \cdot 2} = 36, \quad m = C_5^2 = \frac{5 \cdot 4}{1 \cdot 2} = 10,$$

$$p(A/H_1) = \frac{10}{36} = \frac{5}{18}.$$

2 коробка:

$$n = C_4^2 = 6, \quad m = C_3^2 = 3, \quad p(A/H_2) = \frac{1}{2}.$$

3 коробка:

$$n = C_3^2 = 3, \quad m = C_2^2 = 1, \quad p(A/H_3) = \frac{1}{3}.$$

Подставляя эти данные в формулу (7.16), найдем

$$p(A/H_1) = \frac{\frac{1}{3} \cdot \frac{5}{18}}{\frac{1}{3}\left(\frac{5}{18} + \frac{1}{2} + \frac{1}{3}\right)} = \frac{1}{4}.$$

7.1.8. Формулы Бернулли и Пуассона

Схемой Бернулли будем называть схему одинаковых независимых испытаний, каждое из которых заканчивается одним из двух несовместных исходов: или наступает некоторое событие A, или оно не наступает, т. е. наступает противоположное ему событие $\bar{A}$. Независимость испытаний означает (см. (7.13)), что вероятность события A в любом испытании не зависит от того, какими сведениями мы располагаем относительно результатов других испытаний. Кроме того, в схеме Бернулли предполагается, что вероятность события A в одном испытании остается постоянной во всей серии испытаний.

Многие практически важные задачи приводят к отысканию вероятности $P_n(m)$ того, что при n испытаниях, произведенных по схеме Бернулли, событие A произойдет m раз $(m = 0, 1, 2, \ldots$ $\ldots, n)$, если известна вероятность p события A в каждом испытании.

Справедлива следующая *формула Бернулли*:

$$P_n(m) = C_n^m p^m \cdot q^{n-m}, \tag{7.17}$$

где $q = 1 - p$, $m = 0, 1, 2, \ldots, n$.

Пример 1. Что вероятнее выиграть у равносильного противника:

1) три партии из четырех или пять из восьми?

2) не менее трех партий из четырех или не менее пяти партий из восьми?

Решение. Так как противники равносильны, то вероятность выигрыша и проигрыша в каждой партии одинаковы.

1) Вероятность выиграть три партии из четырех:

$$P_4(3) = C_4^3 \cdot \left(\frac{1}{2}\right)^3 \cdot \frac{1}{2} = \frac{1}{4}.$$

Вероятность выиграть пять партий из восьми:

$$P_8(5) = C_8^5 \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^3 = \frac{7}{32}.$$

Так как $1/4 > 7/32$, то вероятнее выиграть три партии их четырех.

2) Вероятность выиграть не менее трех партий из четырех:

$$p = P_4(3) + P_4(4) = \frac{1}{4} + \frac{1}{16} = \frac{5}{16}.$$

Вероятность выиграть не менее пяти партий из восьми:

$$p = P_8(5) + P_8(6) + P_8(7) + P_8(8) =$$
$$= \frac{7}{32} + \left(\frac{8 \cdot 7}{2} + 8 + 1\right) \cdot \left(\frac{1}{2}\right)^8 = \frac{93}{256}.$$

Поскольку $93/256 > 5/16$, то вероятнее выиграть не менее пяти партий из восьми.

Пример 2. Студент отвечает на 4 дополнительных вопроса при сдаче экзамена. Вероятность правильного ответа на каждый вопрос будем считать равной $p = 1/4$. Предполагая, что все ответы — события независимые, найти вероятность того, что будет дано хотя бы два правильных ответа.

Решение. $A = \{$*хотя бы два правильных ответа*$\}$ — это 2, 3 или 4.

Так как $p = 1/4$, то $q = 1 - 1/4 = 3/4$ — вероятность неправильного ответа на вопрос. Эту задачу удобно решать, используя противоположные события, т. е. пользуясь равенством: $P(A) = 1 - [P_4(0) + P_4(1)]$:

$$P_4(0) = C_4^0 p^0 q^4 = 1 \cdot 1 \cdot \left(\frac{3}{4}\right)^4 = \frac{81}{256};$$

$$P_4(1) = C_4^1 p^1 q^3 = 4 \cdot \frac{1}{4} \cdot \left(\frac{3}{4}\right)^3 = \frac{27}{64};$$

$$P(A) = 1 - \left[\frac{81}{256} + \frac{27}{64}\right] = 1 - \frac{189}{256} = \frac{67}{256}.$$

В случаях, когда n велико, а p мало, использование формулы Бернулли вызывает затруднения. Часто ограничиваются формулой Пуассона

$$P_n(m) \approx \frac{\mu^m}{m!} e^{-\mu}, \quad \mu = np. \tag{7.18}$$

Пример 3. При перевозке 1000 стеклянных колб вероятность разбить 1 колбу равна 0,002. Какова вероятность, того что будут разбиты 4 колбы?

Решение. $P_{1000}(4) \approx \dfrac{(1000 \cdot 0{,}002)^4}{4! e^{1000 \cdot 0{,}002}} = \dfrac{2^4}{1 \cdot 2 \cdot 3 \cdot 4 e^2} \approx 0{,}08.$

Самостоятельная работа

Комбинаторика

1. Сколькими способами можно выбрать три различных краски из имеющихся пяти?

2. Сколькими способами можно составить трехцветный полосатый флаг, если имеется материал 5 различных цветов?

3. Сколько словарей надо издать, чтобы можно было непосредственно выполнять переводы с любого из пяти языков — русского, английского, французского, немецкого и итальянского — на любой другой из этих пяти языков?

4. Составляются знаки, состоящие из геометрической фигуры (окружность, квадрат, треугольник или шестиугольник), буквы и цифры. Сколько таких знаков можно составить?

5. Для эксперимента по определению скорости роста требуется выбрать 4 штамма бактерий из имеющихся восьми. Сколькими способами это можно сделать?

Классическая вероятность

6. Студент знает 40 вопросов из 60, включенных в программу. Найти вероятность того, что студент знает 2 вопроса, содержащихся в его экзаменационном билете?

7. В одном аквариуме находятся 5 белых, 4 красных и 3 голубых рыбки. Двух случайно выбранных рыбок переносят в другой аквариум. Какова вероятность того, что обе рыбки голубые.

8. Из десяти билетов выигрышными являются два. Определить вероятность того, что среди взятых наудачу пяти билетов:

1) один выигрышный, 2) два выигрышных,

3) хотя бы один выигрышный.

9. Набирается четырехзначный номер телефона. Какова вероятность набрать его правильно с первого раза, если все цифры номера нечетные и ровно три из них пятерки.

10. Среди кандидатов в студенческий совет факультета три первокурсника, пять второкурсников и семь третьекурсников. Из этого состава наугад выбирают пять человек на конференцию. Найти вероятности следующих событий:

A — будут выбраны одни третьекурсники,

B — все первокурсники попадут на конференцию,

C — не будет выбрано ни одного второкурсника,

D — выберут одного первокурсника, двоих второкурсников и двоих третьекурсников.

Сложение и умножение вероятностей

11. На клумбе растут 20 красных, 30 синих и 40 белых астр. Какова вероятность сорвать в темноте цветную астру, если срывают одну астру?

12. Студент пришел на зачет, зная из 30 вопросов только 24. Какова вероятность сдать зачет, если после первого отказа отвечать на вопрос, преподаватель задает еще один вопрос?

13. Шесть человек больны заболеванием, для которого коэффициент выздоровления составляет 98%. Какова вероятность того, что:

а) выздоровеют все шестеро;

б) выздоровеют только пятеро?

14. В читальном зале 6 учебников по теории вероятностей, из которых 3 в переплете. Библиотекарь взял наугад 2 учебника. Найти вероятность того, что оба учебника окажутся в переплете.

15. Два стрелка стреляют по мишени. Вероятность попадания в мишень при одном выстреле для 1-го стрелка равна 0,7, а для 2-го — 0,8. Найти вероятность того, что при одном залпе в мишень попадет только один из стрелков.

16. Студент знает 20 из 25 вопросов программы. Найти вероятность того, что студент знает предложенные ему экзаменатором 3 вопроса.

17. В цехе работает 7 мужчин и 3 женщины. По табельным номерам наугад отобраны 3 человека. Найти вероятность того, что все отобранные лица окажутся мужчинами.

18. Некто написал 3 письма, запечатал их в конверты, а затем наугад на каждом из них написал различные адреса. Определить вероятность того, что на всех конвертах написаны правильные адреса.

Формула полной вероятности

19. В первой коробке 20 ампул, из них 18 стандартных, во второй — 10 ампул, из них 9 стандартных. Из второй коробки взята наугад ампула и переложена в первую коробку. Найти вероятность того, что случайным образом извлеченная из первой коробки ампула окажется стандартной.

20. В вычислительной лаборатории имеются 6 клавишных автоматов и 4 полуавтомата. Вероятность того, что за время выполнения расчета автомат не выйдет из строя, равна 0,95, а для полуавтомата эта вероятность равна 0,8. Студент выполняет расчет на наугад выбранной машине. Найти вероятность того, что до окончания расчета машина не выйдет из строя.

21. В первой коробке 3 белых шарика и два черных, во второй — 2 белых и 3 черных, в третьей — 1 белый и 2 черных. Наудачу выбирается коробка, и из нее два шарика. Какова вероятность того, что они разных цветов?

Формула Байеса

22. В первой коробке 3 белых шарика и 2 черных, во второй — 6 белых и 5 черных, в третьей — 7 белых и 5 черных. Наугад выбирается коробка и из нее два шарика. Оба шарика черные. Найти условную вероятность того, что шары извлекались из первой коробки.

23. В одной лечебнице, согласно оценкам, 50% мужчин и 30% женщин имеют серьезные нарушения сердечной деятельности. В этой лечебнице женщин вдвое больше, чем мужчин. У случайно выбранного пациента оказалось серьезное нарушение сердечной деятельности. Какова вероятность того, что этот пациент — мужчина?

Формула Бернулли

24. Монету бросают шесть раз. Найти вероятность того, что герб выпадает:

а) менее двух раз,

б) не менее двух раз.

25. Производится серия из 4 выстрелов по мишеням с вероятностью попадания в каждом выстреле $p = 1/4$. Предполагая, что результаты выстрелов — события независимые, найти вероятность того, что будет хотя бы два попадания.

26. Монету подбрасывают восемь раз подряд. Какова вероятность того, что герб выпадет 5 раз?

27. В библиотеке имеются учебники по химии и по биологии. Вероятность того, что любой читатель возьмет книгу по химии, равна 0,7, а по биологии — 0,3. Определить вероятность того, что 5 читателей подряд возьмут книги или только по химии, или только по биологии, если каждый берет только одну книгу по этим предметам.

Формула Пуассона

28. Завод отправил на базу 5000 доброкачественных изделий. Вероятность того, что в пути изделие повредится, равна 0,0002. Найти вероятность того, что на базу прибудут три негодных изделия (остальные — годные!).

29. Среди семян ржи 0,4% семян сорняков. Какова вероятность при случайном отборе 5000 семян обнаружить 5 семян сорняков?

§7.2. Случайные величины и их законы распределения

До сих пор мы изучали способы вычисления вероятностей различных случайных событий. При этом под случайными событиями понималась любая качественная характеристика результата испытания. Однако существует много задач, решение которых требует изучения явлений, характер которых определяется не фактом появления или не появления того или иного случайного события, а зависит от некоторых величин, могущих принять иные случайные значения. Такие величины в теории вероятностей называются случайными. Таким образом, если случайное событие есть качественная характеристика испытания, то случайная величина является количественной характеристикой испытания. Следует заметить, что понятие случайной величины является в известном смысле обобщением понятия случайного события, так как каждому случайному событию A можно всегда поставить в соответствие случайную величину, принимающую два значения, например 1 или 0, в зависимости от того, произошло событие A в результате данного испытания или не произошло.

Приведем примеры случайных величин:

1) продолжительность жизни любого живого существа,

2) число дефектных изделий в любой выборке данного объема,

3) прибыль предприятия за любой фиксированный отрезок времени,

4) погрешность любого измерения,

5) время безотказной работы любого технического устройства,

6) отклонения параметров данного устройства от номинальных значений.

Приведенные примеры с достаточной определенностью показывают, что со случайными величинами приходится иметь дело в самых разнообразных областях науки и техники.

Для полной характеристики случайной величины необходимо знать:

1) множество ее возможных значений;

2) как часто, т. е. с какой вероятностью, случайная величина принимает те или иные значения.

Такая характеристика случайной величины называется ее *законом распределения*. Закон распределения случайной величины может быть задан в различных формах в зависимости от типа случайной величины. Тип случайной величины определяется видом множества значений, которые может принимать рассматриваемая случайная величина. Особую важность для технических приложений имеют два типа случайных величин: дискретного и непрерывного типа.

Определение. Случайная величина называется величиной *дискретного типа* или просто *дискретной*, если она может принимать конечное или счетное множество значений.

Определение. Случайная величина называется величиной *непрерывного типа*, если она может принимать любые значения в одном или нескольких заданных интервалах.

7.2.1. Закон распределения вероятностей дискретной случайной величины

Рассмотрим дискретную случайную величину X с конечным или счетным множеством значений $x_1, x_2, \dots, x_n$. Закон распределения этой случайной величины задан, если известны все вероятности $P(X = x_i) = p_i$ $(i = \overline{1, n})$.

Совокупность всех значений x_i случайной величины X и вероятностей p_i называется рядом распределения величины X, который может быть задан с помощью таблицы следующего вида:

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

Очевидно, что всегда должно быть выполнено условие $\sum p_i = 1$, так как события $\{X = x_i\}$ $(i = 1, \dots, n)$ образуют полную группу событий.

Важный пример дискретной случайной величины дает число m появлений некоторого события A в серии последовательных испытаний, проведенных по схеме Бернулли.

Обозначим через X число появлений события A в n испытаниях, и пусть вероятность наступления события A в каждом испытании равна p.

Говорят, что случайная величина X имеет *биномиальное распределение* с параметрами n и p, если

$$P(X = m) = C_n^m p^m (1 - p)^{n-m}, \quad m = 0, 1, 2, \dots, n.$$

Еще одним примером дискретной случайной величины является случайная величина X, имеющая *распределение Пуассона* с параметром $\lambda > 0$:

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

Напоминаем, что распределение Пуассона является предельным для биномиального распределения (см. п. 7.1.3). Отметим также, что случайная величина X, имеющая распределение Пуассона, принимает бесконечное число значений.

Закон распределения дискретной случайной величины можно задавать по-другому, а именно с помощью так называемого интегрального закона распределения, или функции распределения вероятностей. Функцией распределения $F(x)$ случайной величины X называется вероятность того, что величина X примет значение, меньшее чем x: $F(x) = P(X < x)$, где x — произвольное вещественное число.

Заметим сразу же, что функция распределения вероятностей $F(x)$ случайной величины X является универсальной характеристикой X, она существует для случайной величины любого типа.

Очевидно, что если X — дискретная случайная величина, то ее функция распределения определяется посредством равенства

$$F(x) = \sum_{x_i < x} p_i. \tag{7.19}$$

Пример 1. Построить функцию распределения дискретной случайной величины X, если известен ее ряд распределения:

X	0	1	2
P	0,3	0,2	0,5

Решение. Если $x \leqslant 0$, то $F(x) \equiv 0$, поскольку в данном случае событие $\{X < x\}$ является невозможным.

Если $0 < x \leqslant 1$, то событие $\{X < x\}$ равносильно событию $\{X = 0\}$. Поэтому для данного промежутка $F(x) = P(X = 0) = 0{,}3$.

Аналогично, если $1 < x \leqslant 2$, то $F(x) = P(X = 0) + P(X = 1) = 0{,}5$.

Наконец, если $x > 2$, то $F(x) \equiv 1$, поскольку теперь событие $\{X < x\}$ становится достоверным. Итак, функция распределения такова:

$$F(x) = \begin{cases} 0, & x \leqslant 0, \\ 0{,}3, & 0 < x \leqslant 1, \\ 0{,}5, & 1 < x \leqslant 2, \\ 1, & x > 2. \end{cases}$$

График данной функции распределения имеет вид

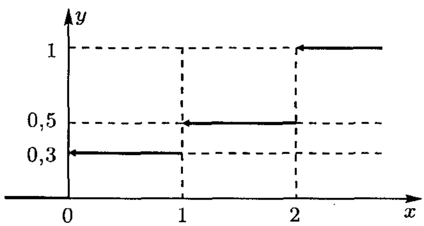


Рис. 7.8. Функция распределения ДСВ.

График $F(x)$ дискретной случайной величины и в общем случае оказывается ступенчатым со скачками в точках $x_1, x_2, \ldots$ $\ldots, x_n$, равными соответственно $p_1, p_2, \ldots, p_n$. Само распределение дискретной случайной величины также называется дискретным.

Пример 2. Построить ряд распределения случайной величины X, представляющей собой число выпадения герба при двукратном бросании монеты.

Решение.

X	0	1	2
p	1/4	1/2	1/4

$$p(X = 0) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \quad p(X = 1) = 2\left(\frac{1}{2} \cdot \frac{1}{2}\right) = \frac{1}{2},$$

$$p(X = 2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Пример 3. Из партии в 25 изделий, среди которых имеются 6 бракованных, выбраны случайным образом 3 изделия для проверки их качества. Построить ряд распределения случайного числа бракованных изделий, содержащихся в выборке.

Решение.

$$p(X=0) = \frac{C_{19}^3}{C_{25}^3} \approx 0{,}42, \quad p(X=1) = \frac{6C_{19}^2}{C_{25}^3} \approx 0{,}45,$$

$$p(X=2) = \frac{19C_6^2}{C_{25}^3} \approx 0{,}12, \quad p(X=3) = \frac{C_6^3}{C_{25}^3} \approx 0{,}01.$$

X	0	1	2	3
p	0,42	0,45	0,12	0,01

7.2.2. Функция распределения и плотность вероятности непрерывной случайной величины

Пусть X является случайной величиной непрерывного типа. Множество значений такой случайной величины несчетно, поэтому мы не можем выписать в последовательность и занумеровать все ее значения и, следовательно, не можем задать закон ее распределения рядом распределения. В этом случае на помощь приходит функция распределения, с помощью которой можно характеризовать закон распределения случайной величины любого типа. Случайная величина X (любого типа) считается заданной, если известна ее функция распределения, т. е. функция $F(x) = P(X < x)$.

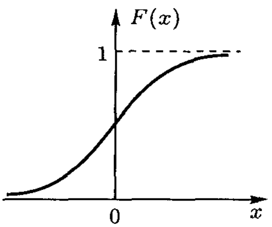


Рис. 7.9. Функция распределения НСВ.

Рассмотрим некоторые основные свойства функции распределения $F(x)$ произвольной случайной величины X. Прежде всего, как всякая вероятность, $F(x)$ при любом x удовлетворяет неравенству $0 \leqslant F(x) \leqslant 1$, так что график функции $y = F(x)$ всегда лежит в полосе между прямыми $y = 0$, $y = 1$.

С помощью функции распределения $F(x)$ можно всегда определить вероятность попадания случайной величины X в промежуток $[\alpha, \beta)$ при любых значениях α и β.

В самом деле, если через A обозначить событие $\{X < \beta\}$, через B — событие $\{X < \alpha\}$, и , наконец, через C — событие $\{\alpha \leqslant x < \beta\}$, то очевидно, что $A = B + C$. Следовательно, по правилу сложения вероятностей для несовместных событий $P(A) = P(B) + P(C)$, откуда $P(C) = P(A) - P(B)$. Но $P(A) = P(X < \beta) = F(\beta)$, $P(B) = F(\alpha)$.

Поэтому окончательно

$$P(\alpha \leqslant X < \beta) = F(\beta) - F(\alpha), \tag{7.20}$$

т. е. вероятность попадания случайной величины в промежуток $[\alpha, \beta)$ равна приращению функции распределения на этом промежутке.

Так как вероятность $P(\alpha \leqslant X < \beta)$ неотрицательна, как всякая вероятность, то из последнего равенства следует, что $F(\beta) \geqslant F(\alpha)$, если $\beta > \alpha$, т.е. функция распределения любой случайной величины является неубывающей функцией.

Если случайная величина X является ограниченной, т. е. все ее возможные значения лежат в каком-либо конечном отрезке $[a, b]$, мы будем иметь равенства

$$F(x) = \begin{cases} 0, & x \leqslant a, \\ 1, & x > b. \end{cases}$$

В общем случае, когда величина X может быть и неограниченной, будут иметь место следующие соотношения:

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

Примером может служить функция распределения случайной величины, распределенной по *показательному закону* ($\mu > 0$):

$$F(x) = \begin{cases} 0, & x \leqslant 0, \\ 1 - e^{-\mu x}, & x > 0. \end{cases} \tag{7.21}$$

Если случайная величина X принадлежит к непрерывному типу, т. е. имеет непрерывную функцию распределения, то для любого x имеем равенство

$$P(X = x) = 0.$$

Напомним, что для дискретных случайных величин вероятность $P(X = x)$ отлична от нуля, если x совпадает с каким-либо частным значением этой случайной величины.

Таким образом, любая непрерывная случайная величина X не может быть характеризована вероятностями своих значений, поскольку все эти вероятности равны нулю.

Чтобы преодолеть эту трудность, вводится понятие *плотности вероятности*.

Плотностью вероятности, или дифференциальным законом распределения вероятностей случайной величины X называется функция $f(x)$, определяемая равенством

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x)}{\Delta x}.$$

Рассмотрим свойства плотности вероятности.

Во-первых, очевидно, что $f(x) \geqslant 0$, поскольку вероятность неотрицательна.

Во-вторых, числитель правой части последней формулы может быть заменен разностью $F(x + \Delta x) - F(x)$.

На основании определения производной получим выражение плотности вероятности через функцию распределения:

$$f(x) = F'(x). \tag{7.22}$$

Третье свойство плотности вероятности может быть получено, если сопоставить равенство $P(\alpha < X \leqslant \beta) = F(\beta) - F(\alpha)$ для непрерывной случайной величины с формулой Ньютона–Лейбница

$$\int_{\alpha}^{\beta} f(x)\,dx = F(\beta) - F(\alpha). \tag{7.23}$$

Это сопоставление позволяет получить выражение для вероятности попадания случайной величины X непрерывного типа в интервал (α,β) через ее плотность вероятности $f(x)$:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x)\,dx. \tag{7.24}$$

Заменяя в последнем равенстве β на x, устремляя $\alpha \to -\infty$ и пользуясь определением функции распределения, получим четвертое свойство плотности вероятности, позволяющее найти функцию распределения, если известна плотность вероятности:

$$F(x) = \int\limits_{-\infty}^{x} f(x)\,dx. \tag{7.25}$$

Последнее, пятое, свойство получим, если верхний предел x интеграла последней формулы устремим к $+\infty$:

$$\int\limits_{-\infty}^{+\infty} f(x)\,dx = 1. \tag{7.26}$$

Это означает, что площадь, ограниченная осью абсцисс и графиком плотности вероятности, равна 1.

Отметим, что не всякая непрерывная случайная величина X имеет плотность вероятности, но лишь та, функция распределения которой дифференцируема.

П р и м е р 1. Дана функция распределения непрерывной случайной величины X:

$$F(x) = \begin{cases} 0, & x \leqslant 0, \\ \sin x, & 0 < x \leqslant \pi/2, \\ 1, & x > \pi/2. \end{cases}$$

Найти плотность вероятности $f(x)$ и построить ее график.

Р е ш е н и е.

$$f(x) = F'(x) = \begin{cases} 0, & x \leqslant 0, \\ \cos x, & 0 < x \leqslant \pi/2, \\ 1, & x > \pi/2. \end{cases}$$

П р и м е р 2. Дана плотность распределения случайной величины X:

$$f(x) = \begin{cases} 0, & x \leqslant 0, \\ a\sin x, & 0 < x \leqslant \pi, \\ 0, & x > \pi. \end{cases}$$

Определить коэффициент a и функцию распределения $F(x)$, построить график функции $F(x)$.

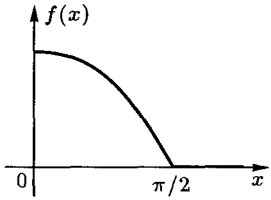


Рис. 7.10. Плотность распределения к примеру 2.

Решение.

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} dx + \int_{0}^{\pi} a\sin x\,dx + \int_{\pi}^{\infty} 0\cdot dx = 1 \Rightarrow a = \frac{1}{2}.$$

$$x \leqslant 0, \quad F(x) = \int_{-\infty}^{x} 0\cdot dx = 0.$$

$$0 < x \leqslant \pi, \quad F(x) = \int_{-\infty}^{0} 0\cdot dx + \int_{0}^{x} \frac{1}{2}\cdot \sin x\,dx = -\frac{1}{2}\cos x\Big|_{0}^{x} =$$

$$= -\frac{1}{2}\cos x + \frac{1}{2} = \frac{1}{2}(1-\cos x).$$

$$x > \pi, \quad F(x) = \int_{-\infty}^{0} 0\cdot dx + \int_{0}^{\pi} \frac{1}{2}\sin x\,dx + \int_{\pi}^{x} 0\cdot dx = 1,$$

т. е.

$$F(x) = \int_{-\infty}^{x} f(x)\,dx = \begin{cases} 0, & x \leqslant 0, \\ \frac{1}{2}(1-\cos x), & 0 < x \leqslant \pi, \\ 1, & x > \pi. \end{cases}$$

7.2.3. Числовые характеристики случайных величин

Для полной характеристики распределения случайной величины X необходимо знать, как мы видели, ее функцию распределения или, в случае непрерывной X, ее плотность вероятности. Однако в ряде практических задач нет необходимости в таком полном исследовании случайной величины X и бывает достаточно иметь хотя бы грубое представление о распределении X, описать его посредством немногих простых параметров. В ряде случаев достаточно знать:

1) примерное расположение того интервала значений, в котором находится основная масса вероятности случайной величины X, а так же положение «центра группирования» на числовой оси;

2) насколько широко разбросаны значения случайной величины X по каждую сторону от «центра группирования».

Параметры первого рода, характеризующие положение центра группирования, называются характеристиками расположения (или положения); параметры второго рода, характеризующие разброс значений случайной величины около центра группирования, называются характеристиками рассеивания.

Наиболее часто употребляемой характеристикой рассеивания является так называемое математическое ожидание случайной величины X, которое будем обозначать символом $M(X)$. Рассмотрим сначала математическое ожидание дискретной случайной величины X, заданной таблицей распределения (см. п. 7.2.1). Математическое ожидание такой величины есть по определению сумма произведений всех ее возможных частных значений на соответствующие им вероятности:

$$M(X) = \sum_{i=1}^{n} x_i p_i. \tag{7.27}$$

Если возможные значения дискретной случайной величины X составляют счетное множество и она задана рядом распределения $P(X = x_i) = p_i$ $(i = 1, 2, 3, \ldots)$, то по определению математическое ожидание X равно сумме ряда:

$$M(X) = \sum_{i=1}^{\infty} x_i p_i, \tag{7.28}$$

если этот ряд сходится, и полагается равным бесконечности, если ряд расходится.

Математическое ожидание непрерывной случайной величины, имеющей плотность вероятности $f(x)$, определяется равенством

$$M(X) = \int_{-\infty}^{+\infty} x f(x)\, dx. \tag{7.29}$$

Пример 1. Найти математическое ожидание случайной величины, принимающей только два значения: 1 и 0, с вероятностью p и $q = 1 - p$ соответственно.

Решение. По определению математического ожидания имеем

$$M(X) = 1 \cdot p + 0 \cdot q = p.$$

Перейдем к перечислению основных свойств математического ожидания.

1. Математическое ожидание постоянной равно этой постоянной, т. е. если C — константа, то $M(C) = C$.

2. Математическое ожидание суммы двух произвольных случайных величин X и Y равно сумме их математических ожиданий: $M(X + Y) = M(X) + M(Y)$.

Следствие. Математическое ожидание суммы конечного числа произвольных случайных величин равно сумме их математических ожиданий, т. е. $M(\sum\limits_i X_i) = \sum\limits_i M(X_i)$.

Пример 2. Найти математическое ожидание числа X наступлений события A при n независимых испытаниях, в каждом из которых событие A появляется с вероятностью p или не появляется с вероятностью $q = 1 - p$. Другими словами, X имеет биномиальное распределение с параметрами n и p (см. п. 7.2.1).

Решение. Очевидно, что X может быть представлено в виде суммы: $X = \sum\limits_{k=1}^{n} X_k$, где X_k есть случайная величина, равная единице, если событие A произошло в k-м испытании, и нулю, если событие A не произошло. Таким образом, $P(X_k = 1) = p$ и $P(X_k = 0) = q$.

На основании последнего свойства математического ожидания имеем: $M(X) = \sum\limits_{k=1}^{n} M(X_k)$.

Как было доказано в примере 1, $M(X_k) = p$. Следовательно, $M(X) = n \cdot p$.

Итак, математическое ожидание случайной величины, распределенной по закону Бернулли (с параметрами n и p), равно $n \cdot p$.

3. Математическое ожидание произведения *независимых случайных величин* равно произведению их математических ожиданий:

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

(выражаясь не совсем строго, случайные величины X и Y считаются независимыми, если значения, принимаемые величиной X, не зависят от того, какое значение приняла величина Y, и наоборот; формальное определение независимости случайных величин X и Y можно найти в п. 8.4.2).

Следствием 1-го и 3-го свойств является утверждение: постоянный множитель можно выносить за знак математического ожидания, т.е. $M(C \cdot X) = C \cdot M(X)$.

Наиболее употребительной характеристикой рассеивания случайной величины X является так называемая дисперсия, которую будем обозначать символом $D(X)$.

Определение. Дисперсией случайной величины называется математическое ожидание квадрата отклонения X от своего математического ожидания:

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

Таким образом, если X — дискретная случайная величина, принимающая конечное или счетное число значений $x_1, x_2, \ldots$ $\ldots, x_n$ с вероятностями, соответственно равными $p_1, p_2, \ldots, p_n$, где $n \leqslant \infty$, то по определению

$$D(X) = \sum_{k=1}^{n} (x_k - M(x))^2 p_k. \tag{7.30}$$

По аналогии с формулой (7.29) дисперсия $D(X)$ случайной величины с плотностью вероятности $f(x)$ определяется с помощью равенства

$$D(X) = \int_{-\infty}^{+\infty} (x - M(X))^2 f(x)\, dx. \tag{7.31}$$

Перечислим теперь основные свойства дисперсии:

1. Дисперсия постоянной равна нулю, т. е. если $X = C$, то $D(C) = 0$.

2. Дисперсия любой случайной величины есть число неотрицательное, т. е. $D(X) \geqslant 0$.

3. Дисперсия случайной величины равна разности математического ожидания квадрата случайной величины и квадрата математического ожидания, т. е. $D(X) = M(X^2) - (M(X))^2$.

4. Если C — константа, то $D(CX) = C^2 D(X)$.

5. Дисперсия суммы или разности независимых случайных величин равна сумме их дисперсий: $D(X \pm Y) = D(X) + D(Y)$.

Последнее свойство остается справедливым для любого конечного числа попарно независимых случайных величин X_1, $X_2, \ldots, X_n$, т. е. $D(X_1 \pm X_2 \pm X_3 \pm \ldots \pm X_n) = D(X_1) + \ldots$ $\ldots + D(X_n)$.

Пример 3. Найти дисперсию случайной величины X, которая имеет биномиальное распределение с параметрами n и p (см. пример 2).

Решение. Как указано в примере 2, справедливо равенство $X = \sum\limits_{k=1}^{n} X_k$, где все X_k $(k = 1, 2, \ldots, n)$ попарно независимы и имеют закон распределения

X_k	0	1
p_k	q	P

На основании 5-го свойства имеем равенство $D(X) = \sum\limits_{k=1}^{n} D(X_k)$.

На основании 3-го свойства имеем равенство $D(X_k) = M(X_k^2) - M^2(X_k)$.

Выше было найдено, что $M(X_k) = p$. Очевидно, что $M(X_k^2) = M(X_k) = p$, следовательно, $D(X_k) = p - p^2 = p(1 - p) = p \cdot q$. Окончательно получаем: $D(X) = n \cdot p \cdot q$, т. е. дисперсия случайной величины, имеющей биномиальное распределение с параметрами n и p, равна $n \cdot p \cdot q$.

В вычислительном отношении более удобна не дисперсия, а другая мера рассеивания случайной величины X, которая чаще всего и используется, — корень квадратный из дисперсии, взятый с положительным знаком, — которая называется *средним квадратичным отклонением* или *стандартным отклонением* случайной величины.

Стандартное отклонение обычно обозначается символом σ_x или $\sigma(X)$:

$$\sigma(X) = \sqrt{D(X)}. \tag{7.32}$$

Удобство стандартного отклонения состоит в том, что оно имеет размерность самой случайной величины X, в то время как дисперсия имеет размерность, представляющую квадрат размерности X.

В заключение отметим, что дисперсия среднего арифметического $\overline{X}$ из n независимых одинаково распределенных случайных величин X_1, X_2, ..., X_n равна дисперсии какой-либо из этих величин, деленной на общее число величин:

$$D(\overline{X}) = D((X_1 + X_2 + \ldots + X_n)/n) = D(X)/n.$$

Отсюда следует, что $\sigma(\overline{X}) = \sigma(X)/\sqrt{n}$.

Именно так обстоит дело, когда мы производим n измерений какой-либо физической величины. Важно то, что результат каждого измерения не зависит от результатов отдельных измерений; тогда последовательные измерения можно рассматривать как значения независимых и одинаково распределенных случайных величин. Средняя квадратичная ошибка одного измерения (стандартное отклонение X) будет при этом $\sigma(X)$, а средняя квадратичная ошибка среднего арифметического X из n измерений будет в $\sqrt{n}$ раз меньше средней квадратичной ошибки одного измерения. Следует заметить, что если в опыте наш прибор давал систематическую ошибку, то никакими повторениями опыта от нее не избавиться.

Замечание. Совместный закон распределения и числовые характеристики двух случайных величин обсуждаются в п. 8.4.2.

7.2.4. Нормальный закон распределения

Среди непрерывных распределений наиболее важную роль играет *нормальное распределение.* К нормальному закону распределения, называемому также *законом Гаусса*, при весьма часто встречающихся условиях приближаются другие законы. В частности, если мы имеем сумму большого числа независимых величин, подчиненных каким угодно законам распределения, то при некоторых общих условиях она будет приближенно подчиняться нормальному закону.

Случайная величина, распределенная по нормальному закону, имеет плотность вероятности вида

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-a)^2}{2\sigma^2}}, \tag{7.33}$$

где постоянные a и σ — параметры распределения; график этой плотности схематически изображен на рис. 7.11.

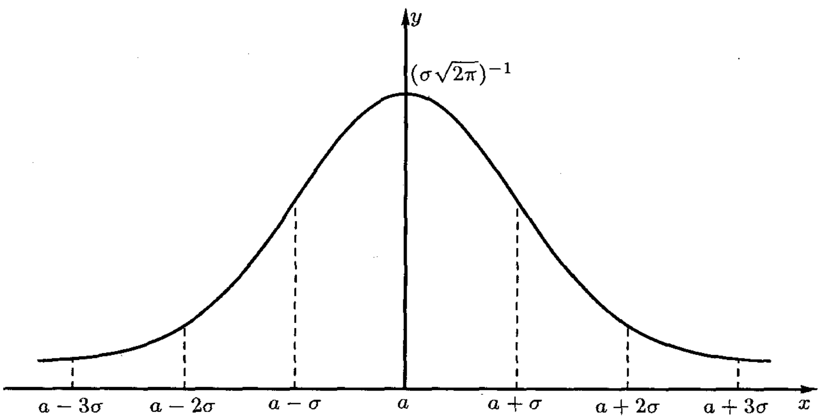


Рис. 7.11. Плотность вероятности нормального распределения.

Легко видеть, что $f(a + x) = f(a - x)$, так что кривая плотности симметрична относительно прямой $x = a$. Методами дифференциального исчисления можно установить, что кривая плотности имеет единственный максимум при $x = a$ и две точки перегиба при $x = a \pm \sigma$.

Если случайная величина подчинена нормальному закону распределения вероятности с параметрами a и σ, то этот факт для краткости может быть записан в виде $X \sim N(a, \sigma)$.

Вероятностный смысл параметров a и σ выясняется после нахождения математического ожидания и дисперсии. Оказывается, что $M(X) = a$, $D(X) = \sigma^2$.

Определенный интеграл с переменным верхним пределом вида

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2}\, dt. \qquad (7.34)$$

носит название нормированной *функции Лапласа* или просто функции Лапласа (см. приложение 1).

Функция Лапласа обладает следующими свойствами:

$$\Phi(0) = 0, \quad \Phi(\infty) = 0{,}5, \quad \Phi(-x) = -\Phi(x).$$

Функция распределения случайной величины $X \sim N(a, \sigma)$ может быть выражена через функцию Лапласа следующим образом:

$$F(x) = 1/2 + \Phi((x - a)/\sigma). \qquad (7.35)$$

Принимая во внимание формулу (7.35), найдем вероятность попадания случайной величины $X \sim N(a, \sigma)$ в промежуток (α,β):

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = \Phi((\beta - a)/\sigma) - \Phi((\alpha - a)/\sigma). \qquad (7.36)$$

Если промежуток симметричен относительно точки a $(= M(X))$, то для случайной величины $X \sim N(a, \sigma)$ получаем формулу следующего вида:

$$P\{|X - a| < \varepsilon\} = 2\Phi(\varepsilon/\sigma). \qquad (7.37)$$

При $\varepsilon = t\sigma$ получим

$$P\{|X - a| < t \cdot \sigma\} = 2\Phi(t). \qquad (7.38)$$

В частности, пользуясь приложением 1, при $t = 1, 2, 3$ из формулы (7.38) получим

$$P\{|X - a| < \sigma\} = 2\Phi(1) = 0{,}6827, \quad P\{|X - a| < 2\sigma\} = 0{,}9545,$$

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

Последнее равенство дает основание к следующему практическому правилу, которое часто называют *правилом «3σ»*: все практически возможные значения случайной величины, подчиненной нормальному закону $N(a, \sigma)$, заключены в интервале $(a - 3\sigma,\ a + 3\sigma)$.

Самостоятельная работа

Дискретные случайные величины

1. Построить закон распределения и функцию распределения числа попаданий мячом в корзину при двух бросках, если вероятность попадания равна 0,4.

2. В партии из 10 деталей имеется 8 стандартных. Наугад отобраны 2 детали. Составить закон распределения стандартных деталей среди отобранных.

3. В партии из 6 деталей имеется 4 стандартных. Наугад отобраны 3 детали. Составить закон распределения дискретной случайной величины X — числа стандартных деталей среди отобранных.

4. Две игральные кости одновременно бросают два раза. Написать биноминальный закон распределения дискретной случайной величины X — числа выпадений четного числа очков на обеих игральных костях.

5. В партии деталей 10% нестандартных. Наугад отобраны 4 детали. Написать биноминальный закон распределения дискретной случайной величины X — числа нестандартных деталей среди четырех отобранных.

Непрерывные случайные величины

6. Дана плотность распределения вероятности:

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

Найти $F(x)$.

7. Дана плотность распределения вероятности:

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

Найти $F(x)$.

8. Дана плотность распределения вероятности:

$$f(x) = \begin{cases} 0, & x \leqslant \pi/6, \\ 3\sin 3x, & \pi/6 < x \leqslant \pi/3, \\ 0, & x > \pi/3. \end{cases}$$

Найти $F(x)$.

9. Дана плотность распределения случайной величины X:

$$f(x) = \begin{cases} 0, & x \leqslant 0, \\ \cos x, & 0 < x \leqslant \pi/2, \\ 0, & x > \pi/2. \end{cases}$$

Найти функцию распределения $F(x)$.

10. Непрерывная случайная величина X задана плотностью вероятности $f(x) = 2/3\sin 3x$ в интервале $(0, \pi/3)$, вне этого интервала $f(x) = 0$. Найти вероятность того, что X примет значение, принадлежащее интервалу $(\pi/6, \pi/4)$.

11. Случайная величина X задана функцией распределения:

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

Найти вероятность того, что в результате четырех независимых испытаний величина X ровно три раза примет значение, принадлежащее интервалу (0,25, 0,75).

Числовые характеристики

12. В партии деталей 10% нестандартных. Наугад отобраны 4 детали. Написать биноминальный закон распределения дискретной случайной величины X — числа нестандартных деталей среди четырех отобранных.

13. Случайная величина X задана законом распределения:

X	0	1	2
p	0,25	0,25	0,5

Найти математическое ожидание, дисперсию, среднее квадратичное отклонение, построить функцию распределения.

14. Найти дисперсию дискретной случайной величины X — числа появлений события A в пяти независимых испытаниях, если вероятность появления события A в каждом испытании равна 0,2.

15. Дискретная случайная величина X принимает только два возможных значения x_1 и x_2, $x_1 > x_2$, $P(X = x_1) = 0{,}6$. Найти закон распределения величины X, если $M(X) = 1{,}4$, $D(X) = 0{,}24$.

16. Случайная величина X задана функцией плотности вероятности:

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

Найти математическое ожидание величины X.

17. Случайная величина X задана функцией плотности вероятности:

$$f(x) = \begin{cases} 0, & x \leqslant 3, \\ \dfrac{3}{4}x^2 + 6x - \dfrac{45}{4}, & 3 < x \leqslant 5, \\ 0, & x > 5. \end{cases}$$

Найти математическое ожидание величины X.

18. Случайная величина X задана функцией плотности вероятности:

$$f(x) = \begin{cases} 0, & x \leqslant -3, \\ \dfrac{1}{\pi\sqrt{9 - x^2}}, & -3 < x \leqslant 3, \\ 0, & x > 3. \end{cases}$$

Найти дисперсию X.

19. Плотность вероятности случайной величины X, равномерно распределенной на $[a, b]$:

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

Найти:

1) функцию распределения $F(x)$ и начертить ее график;

2) математическое ожидание, дисперсию и среднее квадратичное отклонение случайной величины X.

Нормальный закон

20. Масса взрослого животного некоторого вида является нормально распределенной случайной величиной со средним значением 100 кг и стандартным отклонением 8 кг. Наудачу выбирают взрослое животное. Найти вероятности следующих событий:

1) масса животного меньше 90 кг,

2) больше 110 кг,

3) находится в интервале от 95 до 105 кг,

4) находится в интервале от 97 до 112 кг.

Замечание. Любое нормальное распределение может принимать как положительные, так и отрицательные значения. Поэтому можно лишь приближенно считать, что случайная величина в этой и других задачах имеет нормальный закон.

21. Диастолическое давление крови выпускников некоторого училища является нормально распределенной случайной величиной со средним значением 80 мм и стандартным отклонением 5 мм. Измеряют давление крови у случайно выбранного выпускника. Определить вероятности того, что:

1) давление ниже 70 мм,

2) выше 85 мм,

3) выше 90 мм, но при дополнительном условии, что пациент выбран из числа тех, у кого на день проверки диастолическое давление оказалось выше 85 мм.

22. Предприятие выпускает стеклянные ампулы, размеры которых будем считать распределенными по нормальному закону. Средняя длина 100 мм, а стандартное отклонение 1 мм. Ампула считается бракованной, если она короче 98 мм или длиннее 101 мм. Найти среднее число бракованных ампул среди 3 наудачу взятых ампул.

23. В условиях задачи 22 найти интервал, симметричный относительно среднего значения бракованных ампул, в который попадает реальное число бракованных ампул с вероятностью не менее 0,96.

24. Определить среднеквадратическую ошибку весов, если систематических ошибок весы не имеют, а случайные распределены по нормальному закону и с вероятностью 0,9 не выходят за пределы ± 20 мг.

25. Амперметр имеет систематическую ошибку 0,5 А и среднеквадратическую ошибку 0,1 А. Найти вероятность того, что ошибка измерения не превзойдет 0,5 А.

26. Весы имеют среднеквадратическую ошибку 60 мг, систематические ошибки отсутствуют. Сколько необходимо произвести взвешиваний, чтобы с вероятностью не менее 0,9 средняя ошибка не превосходила по абсолютной величине 7,5 мг.

27. Производится два независимых взвешивания на весах, имеющих среднеквадратическую ошибку 30 мг и систематическую ошибку $+10$ мг. Какова вероятность того, что обе ошибки измерения, имея разные знаки, по абсолютной величине превзойдут 10 мг?

28. Изделие считается высшего качества, если отклонение его длины от номинала не превосходит по абсолютной величине 3,45 мм. Случайное отклонение длины изделия от номинала подчиняется закону со среднеквадратическим отклонением в 3 мм. Систематические отклонения отсутствуют. Определить среднее число изделий высшего сорта если изготовляются 4 изделия.

29. При большом числе измерений установлено, что 75% ошибок:

1) не превосходит 1,25 мм,

2) не превосходит по абсолютной величине 1,25 мм.

Заменяя частоты появления ошибок их вероятностями, определить в обоих случаях среднеквадратическую ошибку измерения, если систематические ошибки отсутствуют.

30. Весы не имеют систематических ошибок. Случайная ошибка распределена по нормальному закону, причем $\sigma = 10$ мг. Сколько нужно произвести взвешиваний одного и того же предмета, чтобы с вероятностью не менее 0,98 средний арифметический результат имел ошибку в пределах ± 3 мг.

31. В условиях задачи 28 сколько нужно взять изделий, чтобы среди них с вероятностью не менее 0,97 было хотя бы одно изделие не высшего качества?

32. Средняя жирность данной партии молока 2,5%. Отклонение жирности в отдельных пакетах от номинала составляет $\pm 0{,}1\%$ с вероятностью 0,9973. Сколько пакетов молока нужно проверить, чтобы среди них был хотя бы один с молоком жирности менее 2,4% с вероятностью не менее 0,9.

33. Рост взрослых мужчин является случайной величиной, распределенной по нормальному закону. Пусть математическое ожидание ее равно 175 см, а среднее квадратическое отклонение — 6 см. Определить вероятность того, что хотя бы один из 3 наудачу выбранных мужчин будет иметь рост от 170 до 180 см.

34. Случайная величина X распределена нормально с математическим ожиданием $a = 25$. Вероятность попадания в интервал $(10, 15)$ равна 0,2. Чему равна вероятность попадания X в интервал $(35, 40)$?

ГЛАВА 8

ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Математической статистикой называют раздел математики, посвященный математическим методам систематизации, обработки и использования *статистических данных* для научных и практических выводов. Статистические данные здесь понимаются как сведения о числе объектов в какой-либо более или менее обширной совокупности, обладающих теми или иными признаками.

В первом приближении можно сказать, что главная цель математической статистики — получение осмысленных, научно обоснованных выводов из подверженных случайному разбросу данных. При этом само изучаемое явление, генерирующее эти данные, чаще всего слишком сложно, чтобы можно было составить его полное описание, отражающее все детали. Поэтому статистические выводы делаются на основе некоторой математической вероятностной модели реального случайного явления, которая должна воспроизводить его существенные черты и исключать те, которые предполагаются несущественными. Методы математической статистики позволяют по наблюдениям над изучаемым явлением определить вероятностные характеристики случайных величин, участвующих в математической модели, описывающей это явление.

Применительно к предстоящему изложению, мы в первую очередь будем интересоваться задачами точечного и интервального оценивания параметров распределений исследуемых случайных величин, а также задачами статистической проверки гипотез относительно этих параметров.

§ 8.1. Выборочный метод

8.1.1. Генеральная совокупность. Выборка

Всякое каким-то образом выделенное множество объектов, которые могут отличаться друг от друга значением некоторой определенной характеристики, называется *генеральной совокупностью.*

Число элементов генеральной совокупности называется ее *объемом.*

Часть генеральной совокупности, случайным образом отобранная для наблюдений, называется *случайной выборкой* или, для краткости, *выборкой.*

Число элементов выборки называется ее *объемом.*

Так, если из ста тысяч упаковок некоторого лекарства (генеральная совокупность) для контроля качества отобрано сто упаковок (выборка), то объем генеральной совокупности составляет 100 000, а объем выборки — 100.

Поскольку мы рассчитываем с помощью статистических методов высказать определенное суждение о свойствах генеральной совокупности по свойствам выборки, последняя должна быть *репрезентативной* (представительной), т. е. должна быть организована таким образом, чтобы, по возможности, отражать все интересующие нас свойства генеральной совокупности.

Например, при обследовании на предмет успеваемости по физиологии студентов медицинских институтов A, B и C, в которых обучаются 500, 200 и 300 студентов соответственно, выборку объемом 100 следует строить так, чтобы в нее входило 50 случайным образом выбранных студентов института A, 20 студентов института B и 30 студентов института C. Говоря короче, пропорции в выборке должны соответствовать пропорциям генеральной совокупности.

Для обеспечения репрезентативности выборка должна быть достаточно объемной с тем, чтобы охватывать всю генеральную совокупность, и производиться беспристрастно по отношению к отдельным ее частям. Последнее в определенной мере обеспечивается случайностью отбора элементов выборки.

В дальнейшем мы будем придерживаться простой и надежной схемы, в соответствии с которой интересующая нас характеристика объектов генеральной совокупности является случайной величиной X, с частично или полностью неизвестной функцией распределения, а выборка объема n, по которой эту функцию распределения требуется определить, представляет собой n независимых случайных величин $X_1, X_2, \ldots, X_n$, распределенных так же, как X. Такая схема возникает тогда, когда исследователь проводит n независимых друг от друга опытов, в каждом из которых наблюдается некоторое значение одного и того же интересующего его признака, причем предполагается, что во всех n опытах условия эксперимента оставались неизменными. Заметим, что последнее требование достаточно условно и необходимая степень его выполнения устанавливается экспериментатором в соответствии с целями опыта.

Расположив члены выборки $X_1, X_2, \ldots, X_n$ (т. е. n чисел в определенных физических единицах) в порядке возрастания, получим *вариационный ряд*

$$X_{(1)} \leqslant X_{(2)} \leqslant \ldots \leqslant X_{(n)}. \tag{8.1}$$

Пример 1.1. Вариационный ряд, полученный по выборке 2, −1, 0, 2, 0, 3, 3, 0, 4, 2, имеет вид −1, 0, 0, 0, 2, 2, 2, 3, 3, 4.

Необходимо отметить, что на деле исследователь располагает не собственно случайной выборкой, а лишь ее реализацией, т. е. набором чисел $x_1, x_2, \ldots, x_n$ (их еще называют *вариантами*), полученных в результате наблюдений величины X при n независимых повторениях случайного эксперимента в одинаковых условиях. Таким образом, при использовании соответствующих формул математической статистики в них вместо случайных величин X_k следует подставлять их конкретные реализации x_k. Однако в дальнейшем, чтобы не загромождать изложение материала, мы иногда будем обозначать посредством X_k как случайную величину, так и ее значение.

8.1.2. Статистическое распределение выборки

Пусть в выборке объема n из некоторой генеральной совокупности варианты принимают $m \leqslant n$ различных упорядоченных по возрастанию значений $x_1 < x_2 < \ldots < x_m$, причем значения x_k встречаются n_k раз. Числа $n_1, n_2, \ldots, n_m$ называются *частотами* вариант $x_1, x_2, \ldots, x_m$ соответственно, а величины $w_k = n_k/n$, $1 \leqslant k \leqslant m$, их *относительными частотами*.

Очевидно, что $n_1+n_2+\ldots+n_m = n$, а $w_1+w_2+\ldots+w_m = 1$.

Определение. *Статистическим (эмпирическим) законом распределения выборки*, или просто *статистическим распределением выборки* называется набор вариант x_k и соответствующих им частот n_k или относительных частот w_k.

Статистическое распределение выборки удобно представлять в форме таблицы распределения частот (частотной таблицы), иногда называемой *статистическим рядом распределения*,

Таблица 8.1

x_1	x_2	$\ldots$	x_m
n_1	n_2	$\ldots$	n_m

или в виде таблицы распределения относительных частот:

Таблица 8.2

x_1	x_2	$\ldots$	x_m
w_1	w_2	$\ldots$	w_m

Отметим, что с формальной точки зрения табл. 8.2 описывает закон распределения некоторой гипотетической дискретной случайной величины.

Графическими аналогами табл. 8.1 и 8.2 являются столбцовые диаграммы, т. е. последовательности вертикальных отрезков (ординат) длины n_k или w_k с абсциссами x_k, $k = 1, 2, \ldots, m$, которые отличаются друг от друга лишь масштабом по оси ординат (рис. 8.1, 8.2).

Другая эквивалентная форма графического изображения табл. 8.1 и 8.2, называемая *полигоном частот* (или *полигоном относительных частот*), представлена на рис. 8.1 и 8.2 пунктирной линией, соединяющей вершины вертикальных отрезков, т. е. точки (x_i, n_i) или (x_i, w_i) соответственно.

Пример 1.2. Найти статистическое распределение выборки из примера 1.1 и представить его в форме таблицы частот и отно-

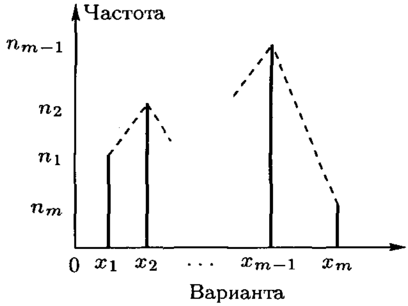


Рис. 8.1. Объяснения в тексте.

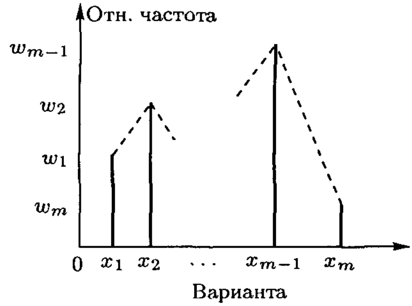


Рис. 8.2. Объяснения в тексте.

сительных частот. Построить график и полигон относительных частот.

Решение. Статистическое распределение выборки представлено в табл. 8.3,

Таблица 8.3

x_k	−1	0	2	3	4
n_k	1	3	3	2	1
w_k	0,1	0,3	0,3	0,2	0,1

а график и полигон относительных частот изображены на рис. 8.3. и 8.4. соответственно.

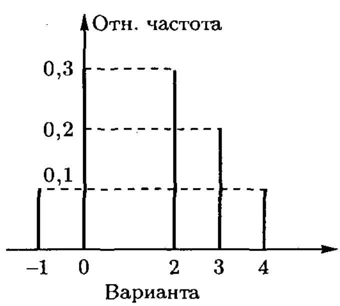


Рис. 8.3. Объяснения в тексте.

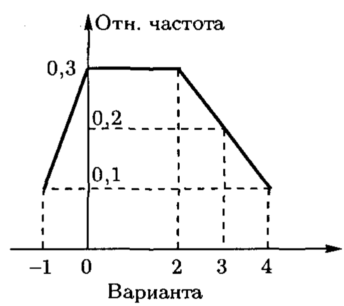


Рис. 8.4. Объяснения в тексте.

8.1.3. Эмпирическая функция распределения

Определение. Эмпирической функцией распределения, построенной по случайной выборке X_1, $X_2, \ldots, X_n$, называется функция $F_n(x)$, равная доле тех элементов выборки, которые меньше x, где x — любое вещественное число.

Из приведенного определения следует, что если статистическое распределение выборки задано частотной таблицей (см. табл. 8.1, 8.2), то

$$F_n(x) = \frac{1}{n}\sum_{x_k<x} n_k = \sum_{x_k<x} w_k. \tag{8.2}$$

Эмпирическая функция распределения $F_n(x)$ обладает всеми свойствами обычной функции распределения, в частности не убывает по x.

Замечание 1.1. Если статистическое распределение выборки задано табл. 8.2, то эмпирическая функция распределения равна нулю при $x \leqslant x_1$, равна единице при $x > x_m$ и $F_n(x) = \sum\limits_{1\leqslant k\leqslant i} w_k$, если $x_i < x \leqslant x_{i+1}$, $i = 1, 2, \ldots, m-1$.

Отметим, что статистическое распределение выборки и ее эмпирическая функция распределения взаимно-однозначно определяют друг друга.

Пример 1.3. Найти эмпирическую функцию распределения, соответствующую выборке из примера 1.1, и построить ее график.

Решение. Объем выборки равен $n = 10$, а соответствующая ей частотная таблица есть табл. 8.3. Таким образом, по замечанию 1.1

$$F_n(x) = \begin{cases} 0 & \text{при} \quad x \leqslant -1, \\ 0{,}1 & \text{при} \quad -1 < x \leqslant 0, \\ 0{,}4 & \text{при} \quad 0 < x \leqslant 2, \\ 0{,}7 & \text{при} \quad 2 < x \leqslant 3, \\ 0{,}9 & \text{при} \quad 3 < x \leqslant 4, \\ 1 & \text{при} \quad x > 4. \end{cases}$$

График функции $F_n(x)$ представлен на рис. 8.5.

Репрезентативная выборка подобна генеральной совокупности, из которой она извлечена, в силу чего эмпирическая функция распределения $F_n(x)$ должна быть в некотором смысле близка к функции распределения $F(x)$ признака X генеральной

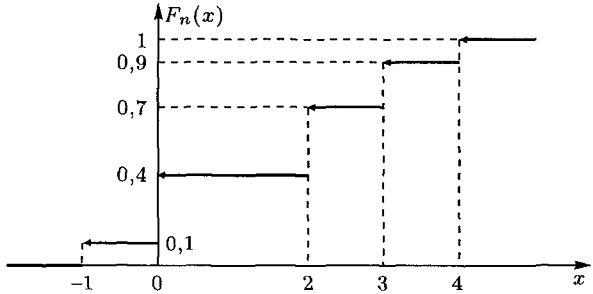


Рис. 8.5. График функции распределения.

совокупности. Это действительно так. Методами теории вероятностей доказывается, что эмпирическая функция распределения $F_n(x)$, которая, как и прочие оценки, построенные по случайным выборкам, является случайной величиной, с увеличением объема выборки сходится (в частности, по вероятности) к теоретической функции распределения $F(x)$ и является хорошим ее приближением.

8.1.4. Статистический интервальный ряд распределения. Гистограмма

Статистическим дискретным рядом (или эмпирической функцией распределения) обычно пользуются в том случае, когда отличных друг от друга вариант в выборке не слишком много, или тогда, когда дискретность по тем или иным причинам существенна для исследователя. Если же интересующий нас признак генеральной совокупности X распределен непрерывно или его дискретность нецелесообразно (или невозможно) учитывать, то варианты группируются в интервалы.

Типичная процедура выглядит следующим образом. Чтобы сгруппировать данные, содержащиеся в выборке объема n из наблюдений над случайной величиной X, следует разделить отрезок $[a, b]$, содержащий все варианты, на k интервалов (h_{i-1}, h_i), $i = 1, \ldots, k$:

$$a = h_0 < h_1 < \ldots < h_k = b \tag{8.3}$$

(обычно, $a \approx x_1 = X_{(1)}$, $b \approx x_m = X_{(n)}$ (см. табл. 8.1 и 8.2)).

Затем надо определить *частоту* каждого *интервала*, т. е. количество наблюдений n_i, попавших в i-й интервал. Для определенности будем полагать n_i равным числу вариант, принадлежащих полуинтервалу $[h_{i-1}, h_i)$, включая все варианты, попавшие на правую границу, т. е. равные h_i, в следующий промежуток во всех случаях, кроме последнего. Количество интервалов k, а также значения их границ h_i, до некоторой степени произвольны и являются компромиссом между требованиями экономии и точности.

Замечание 1.2. Часто $h_i - h_{i-1} = h$ при всех i, т. е. группировку осуществляют с равным шагом h. В этой ситуации можно руководствоваться следующими эмпирическими рекомендациями по выбору a, k и h_i:

1) k — наименьшее целое такое, что $k \geqslant 1 + 3{,}32 \lg n$ (так, если $n = 50$, то $k = 7$),

2) $a = X_{(1)}$ (или чуть меньше),

3) $h = (X_{(n)} - X_{(1)})/k$ (или чуть больше),

4) $h_i = a + ih$, $i = 0, 1, \ldots, k$ (в частности, $b = a + kh$).

Так же, как ранее в п. 8.1.2, полученную группировку удобно представить в форме частотной таблицы, которая носит название *статистический интервальный ряд распределения*:

Таблица 8.4

Интервалы группировки	$[h_0, h_1)$	$[h_1, h_2)$	$\ldots$	$[h_{k-2}, h_{k-1})$	$[h_{k-1}, h_k]$
Частоты	n_1	n_2	$\ldots$	n_{k-1}	n_k

(граничная точка h_k включается в последний интервал, чтобы избежать проблем в случае, когда $b = x_m$ (см. (8.3)). Аналогичную таблицу можно образовать, заменяя в табл. 8.4 частоты n_i относительными частотами $w_i = n_i/n$:

Таблица 8.5

Интервалы группировки	$[h_0, h_1)$	$[h_1, h_2)$	$\ldots$	$[h_{k-2}, h_{k-1})$	$[h_{k-1}, h_k]$
Отн. частоты	w_1	w_2	$\ldots$	w_{k-1}	w_k

Напоминаем, что $n_1 + n_2 + \ldots + n_k = n$ — объему выборки, а $w_1 + w_2 + \ldots + w_k = 1$.

Наиболее информативной графической формой частотной табл. 8.4 является специальный график, называемый *гистограммой частот.*

Гистограмма частот состоит из прямоугольников с основаниями (h_{i-1}, h_i), высота которых равна $n_i/(h_i - h_{i-1})$. Таким образом, площадь каждого прямоугольника равна n_i, а их сумма равна n.

Аналогично определяется графическая форма табл. 8.5 — *гистограмма относительных частот*, состоящая из прямоугольников с основаниями (h_{i-1}, h_i) и высотами $w_i/(h_i - h_{i-1})$. Заметим, что площадь той части гистограммы относительных частот, что лежит между h_i и h_m $(i < m)$, равна относительному числу вариант, попавших в интервал $[h_i, h_m)$, и в соответствии с статистическим определением вероятностей может быть интерпретирована как грубая оценка вероятности $P(h_i \leqslant X < h_m)$. Следовательно, с определенными оговорками можно утверждать, что гистограмма относительных частот является выборочным аналогом графика плотности вероятностей распределения признака генеральной совокупности.

Пример 1.4. Из очень большой партии огурцов извлечена случайная выборка объема 50. Интересующий нас признак X — длины огурцов, измеренные с точностью до 1 см, представлен следующим вариационным рядом: 22, 24, 26, 26, 27, 28, 28, 31, 31, 31, 32, 32, 33, 33, 33, 33, 34, 34, 34, 34, 34, 35, 35, 36, 36, 36, 36, 36, 37, 37, 37, 37, 37, 37, 38, 38, 40, 40, 40, 40, 40, 41, 41, 43, 44, 44, 45, 45, 47, 50. Найти статистический интервальный ряд распределения и построить гистограмму относительных частот с равным шагом.

Решение. Определим характеристики группировки с помощью замечания 1.2. Имеем, $a = 22$, $k = 7$, $h = (50 - 22)/7 = 4$, $h_i = 22 + 4i$, $i = 0, 1, \ldots, 7$. Таким образом, частотная таблица представлена табл. 8.6 (напоминаем, что при попадании на границу интервала варианта записывается в правый интервал во всех случаях, кроме последнего).

Гистограмма относительных частот строится по данным первой и третьей строк табл. 8.6:

Таблица 8.6

Интервалы группировки	22–26	26–30	30–34	34–38	38–42	42–46	46–50
Частоты n_i	2	5	9	18	9	5	2
Отн. частоты w_i	0,04	0,1	0,18	0,36	0,18	0,1	0,04

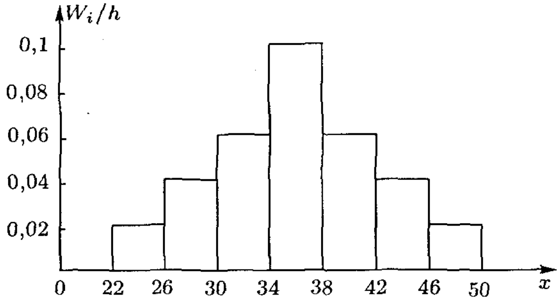


Рис. 8.6. Гистограмма относительных частот

Самостоятельная работа

1. Построить полигоны частот и эмпирические функции распределения по нижеследующим частотным таблицам:

1)

x_k	25	28	30	32	34	36	40
n_k	5	15	17	24	25	16	10

2)

x_k	1	2	3	4	5	6
n_k	5	10	10	25	40	10

2. Изучалось содержание кальция (мг%) в сыворотке крови больных обезьян. По случайной выборке объема 50: 9, 10, 8, 7, 9, 10, 10, 10, 8, 7, 8, 7, 10, 7, 8, 6, 8, 10, 7, 6, 8, 5, 7, 8, 8, 8, 8, 9, 5, 9, 7, 7, 7, 8, 7, 7, 9, 6, 11, 9, 8, 8, 7, 7, 5, 11, 8, 6, 8, найти статистический интервальный ряд распределения и построить гистограмму относительных частот.

3. Изучалось содержание кальция (мг%) в сыворотке крови здоровых обезьян. По случайной выборке объема 44: 8, 10, 8, 8, 9, 8, 5, 6, 7, 8, 9, 7, 8, 9, 8, 8, 7, 7, 8, 9, 11, 9, 8, 8, 9, 10, 7, 9, 7, 8, 8, 8, 7, 7, 7, 8, 7, 7, 9, 8, 6, 10, 8, 10, определить то же, что в предыдущей задаче.

4. Изучалось среднее артериальное давление (мм рт.ст.) в начальной стадии шока. По случайной выборке объема 50: 112, 110, 107, 103, 108, 109, 111, 110, 103, 103, 109, 102, 113, 106, 108, 105, 108, 104, 99, 112, 112, 103, 101, 98, 100, 97, 98, 100, 98, 107, 108, 99, 98, 92, 98, 110, 106, 105, 102, 100, 101, 100, 95, 100, 105, 100, 102, 102, 99, 97, определить то же, что в предыдущей задаче.

5. Изучался рост (см) мужчин возраста 25 лет для сельской местности. По случайной выборке объема 35: 175, 167, 168, 169, 168, 170, 174, 173, 177, 172, 174, 167, 173, 172, 171, 171, 170, 167, 174, 177, 171, 172, 173, 169, 171, 173, 173, 168, 173, 172, 166, 164, 168, 172, 174, определить то же, что в предыдущей задаче.

§ 8.2. Оценки характеристик распределения по данным выборки

В ряде практических ситуаций форма закона распределения изучаемой характеристики объектов генеральной совокупности, или случайной величины X, предполагается известной с точностью до одного или нескольких числовых параметров. В таких случаях перед исследователем стоит задача получения оценок этих параметров на основании извлеченной из генеральной совокупности случайной выборки.

Так, центральная предельная теорема позволяет погрешности измерений рассматривать как случайную величину X, распределенную по нормальному закону; в теории надежности систем радиоэлектронной аппаратуры срок службы элемента системы часто рассматривают как случайную величину X, распределенную показательно. В первом случае, для нахождения закона распределения X достаточно определить (оценить) два параметра a и σ, во втором — один параметр λ.

Ясно, что разумная процедура оценивания не должна ограничиваться лишь выбором приближенного численного значения для неизвестного параметра, но должна также что-то говорить и о надежности этого приближения. Хотя эти два аспекта единой проблемы тесно связаны, удобно обсуждать их отдельно. Соответственно мы будем говорить о точечном оценивании (п. 8.2.1) и об интервальном оценивании (п. 8.2.2).

8.2.1. Точечные оценки

Пусть X_1, $X_2, \ldots, X_n$ — выборка, отвечающая случайной величине X с функцией распределения $F(x)$, относительно которой предполагается, что она зависит от некоторого неизвестного параметра θ. Таким параметром может быть *генеральное среднее*, т. е. математическое ожидание $M(X)$, или *генеральная дисперсия*, т. е. дисперсия $D(X)$. Требуется оценить параметр θ по наблюдениям выборки.

Назовем оценкой неизвестного параметра θ величину θ_n, зависящую от наблюдений выборки и приближенно равную θ:

$$\theta_n \approx \theta \tag{8.4}$$

(такого рода оценки параметров распределений называют *точечными*). Заметим, что оценка θ_n, как функция выборки, является случайной величиной и будет меняться от выборки к выборке.

Для того чтобы хорошо аппроксимировать θ, оценка θ_n в соответствии с требованиями математической статистики должна удовлетворять ряду критериев, основными из которых являются требования *состоятельности* и *несмещенности*.

Определение. Оценка θ_n называется *состоятельной оценкой* параметра θ, если при $n \to \infty$

$$\theta_n \to \theta \tag{8.5}$$

по вероятности, т. е. $P(|\theta_n - \theta| > \varepsilon) \to 0$ при любом $\varepsilon > 0$.

Другими словами, вероятность отклонения оценки от истинного значения параметра можно сделать сколь угодно малой, увеличивая объем выборки.

Определение. Оценка θ_n называется *несмещенной оценкой* параметра θ, если при любом n

$$M(\theta_n) = \theta. \tag{8.6}$$

Последнее означает, что отклонение θ_n от θ не содержит систематической ошибки.

Если равенство (8.6) не выполняется, оценка θ_n называется *смещенной*.

Поскольку наиболее важными числовыми характеристиками большинства случайных величин являются их математические ожидания и дисперсии, далее приведем точечные оценки этих параметров.

В качестве оценки математического ожидания $M(X)$ используется оценка

$$\overline{X} = \frac{1}{n} \sum_{1 \leqslant k \leqslant n} X_k, \tag{8.7}$$

называемая *выборочным средним.*

Оценкой дисперсии $D(X)$ является так называемая *исправленная выборочная дисперсия*

$$S^2 = \frac{1}{n-1} \sum_{1 \leqslant k \leqslant n} (X_k - \overline{X})^2. \tag{8.8}$$

Можно показать, что выборочное среднее $\overline{X}$ и исправленная выборочная дисперсия S^2 являются несмещенными состоятельными оценками генерального среднего $M(X)$ и генеральной дисперсии $D(X)$ соответственно.

Наряду с оценкой дисперсии (8.8) в математической статистике используется также *выборочная дисперсия*

$$\bar{S}^2 = \frac{1}{n} \sum_{1 \leqslant k \leqslant n} (X_k - \overline{X})^2. \tag{8.9}$$

Таким образом,

$$S^2 = \frac{n}{n-1} \bar{S}^2. \tag{8.10}$$

Заметим, что оценка (8.9), в отличие от оценки (8.8), является смещенной, хотя и состоятельной, оценкой генеральной дисперсии, причем занижает (и довольно существенно в случае малых n) ее величину (наличие в определении исправленной выборочной дисперсии множителя $1/(n-1)$ вызвано именно требованием несмещенности).

Замечание 1. Если известен статистический закон распределения выборки (см. табл. 8.1, 8.2), то оценки $\bar{X}$ и $\bar{S}^2$ можно вычислить по формулам:

$$\bar{X} = \frac{1}{n} \sum_{1 \leqslant k \leqslant m} n_k x_k = \sum_{1 \leqslant k \leqslant m} w_k x_k, \tag{8.11}$$

$$\bar{S}^2 = \frac{1}{n} \sum_{1 \leqslant k \leqslant m} n_k (x_k - \bar{X})^2 = \sum_{1 \leqslant k \leqslant m} w_k (x_k - \bar{X})^2, \qquad (8.12)$$

где (см. п. 8.1.2) x_k — набор всех отличающихся друг от друга вариант и соответствующих им частот n_k или относительных частот $w_k = n_k/n$.

Равенство (8.12) может быть записано в следующем более удобном для расчетов виде:

$$\bar{S}^2 = \frac{1}{n} \sum_{1 \leqslant k \leqslant m} n_k x_k^2 - (\bar{X})^2 = \sum_{1 \leqslant k \leqslant m} w_k x_k^2 - (\bar{X})^2. \qquad (8.13)$$

Следует также отметить, что при «ручных» вычислениях по формулам (8.7) и (8.8) полезно применять тождества

$$\begin{gathered} \bar{X} = c + \frac{1}{n} \sum_{1 \leqslant k \leqslant n} (X_k - c), \\ S^2 = \frac{1}{n-1} \sum_{1 \leqslant k \leqslant n} (X_k - c)^2 - \frac{n}{n-1} (\bar{X} - c)^2, \end{gathered} \qquad (8.14)$$

выбирая число с наиболее подходящим способом. Аналогичным образом можно преобразовать формулы (8.11) и (8.13):

$$\begin{gathered} \bar{X} = c + \sum_{1 \leqslant k \leqslant m} w_k (x_k - c), \\ \bar{S}^2 = \sum_{1 \leqslant k \leqslant m} w_k (x_k - c)^2 - (\bar{X} - c)^2. \end{gathered}$$

Пример 2.1. По статистическому распределению выборки из примера 1.2 найти несмещенные оценки генерального среднего и генеральной дисперсии.

Решение. Для нахождения оценки генерального среднего воспользуемся формулой (8.11) и данными табл. 8.3. Имеем при $n = 10$

$$\bar{X} = 0{,}1 \cdot (-1) + 0{,}3 \cdot 0 + 0{,}3 \cdot 2 + 0{,}2 \cdot 3 + 0{,}1 \cdot 4 = 1{,}5.$$

Оценку генеральной дисперсии вычисляем по формулам (8.10) и (8.13):

$$S^2 = \frac{10}{9}(0{,}1 \cdot (-1)^2 + 0{,}3 \cdot 0^2 + 0{,}3 \cdot 2^2 + 0{,}2 \cdot 3^2 + 0{,}1 \cdot 4^2 - 1{,}5^2) =$$
$$= \frac{10}{9}\, 2{,}45 \approx 2{,}722.$$

8.2.2. Интервальные оценки

Пусть, как и в предыдущем параграфе, X_1, $X_2, \ldots, X_n$ — выборка, отвечающая случайной величине X с функцией распределения $F(x)$, зависящей от некоторого параметра θ, а θ_n — несмещенная и состоятельная оценка этого параметра. Поскольку θ_n является случайной величиной, желательно уметь оценивать «качество» равенства (8.4) с тем, чтобы иметь представление к каким ошибкам может привести замена параметра θ его точечной оценкой и с какой степенью уверенности можно ожидать, что эти ошибки не выйдут за известные пределы. Для ответа на эти вопросы в математической статистике вводится понятие *доверительного интервала.*

Пусть α — некоторое число из интервала $(0, 1)$, а $\underline{\theta}$ и $\bar{\theta}$ — две функции, зависящие от выборки, такие что $\underline{\theta} < \bar{\theta}$.

Определение. Интервал $(\underline{\theta}, \bar{\theta})$ называется доверительным интервалом для оценки параметра θ, отвечающим *доверительной вероятности* α, если $P(\underline{\theta} < \theta < \bar{\theta}) = \alpha$.

Иными словами, доверительный интервал — это интервал со случайными границами, который накрывает оцениваемый параметр с наперед заданной вероятностью. При этом, границы доверительного интервала будут зависеть от вариант, объема выборки и доверительной вероятности α (заметим, что число α обычно выбирается достаточно близким к единице; в фармации, как правило, полагают $\alpha = 0{,}95$ или $0{,}99$).

С практической точки зрения можно утверждать, что если мы извлечем из генеральной совокупности, скажем, сто случайных выборок одинакового объема и построим по ним соответствующие доверительные интервалы (с одной и той же доверительной вероятностью α), то следует ожидать, что приблизительно в 100 α случаях (если $\alpha = 0{,}95$, то в 95), эти интервалы будут содержать параметр θ.

Рассмотрим методы построения доверительного интервала для генерального среднего $M(X)$, предполагая, что интересующая нас характеристика объектов X генеральной совокупности, как это часто бывает на практике, имеет нормальное распределение, т. е. $X \sim N(m, \sigma)$ (напоминаем, что если $X \sim N(m, \sigma)$, то математическое ожидание $M(X) = m$, а дисперсия $D(X) = \sigma^2$).

Сначала рассмотрим случай известной дисперсии $\sigma^2 = \mathrm{D(X)}$.

Теорема 2.1 Доверительным интервалом с доверительной вероятностью α для неизвестного математического ожидания $m = M(X)$ при известной дисперсии σ^2 является случайный интервал

$$\left(\bar{X} - \frac{z_\alpha \sigma}{\sqrt{n}}, \bar{X} + \frac{z_\alpha \sigma}{\sqrt{n}}\right). \tag{8.15}$$

Здесь z_α является решением уравнения $2\Phi(z_\alpha) = \alpha$, $\Phi(x)$ — функция Лапласа (см. (7.34)), численные значения которой приведены в приложении 1. В частности,

Таблица 8.7

α	0,95	0,99
z_α	1,96	2,576

Доказательство теоремы 2.1 опирается на следующий факт из теории вероятностей: случайная величина $\dfrac{\sqrt{n}}{\sigma}(\bar{X} - m) \sim N(0, 1)$.

Пример 2.2. Отобранные случайным образом из некоторой популяции (генеральной совокупности) 10 крыс откармливались по специальному рациону. Прибавление веса регистрировалось и, по прошествии определенного периода, оказалось следующим: 113,4, 112,0, 88,6, 49,1, 133,7, 57,8, 44,2, 124,3, 89,5, 93,4. Предполагая, что прибавление веса имеет нормальное распределение со среднеквадратичной ошибкой 40, определить доверительный интервал с доверительной вероятностью 0,95 для среднего прибавления веса в популяции крыс при этом рационе.

Решение. Применим теорему 2.1 при $n = 10$, $\alpha = 0{,}95$ и $\sigma = 40$. Учитывая, что в настоящем примере выборочное среднее

$\bar{X} = 90{,}6$, а $z_\alpha = 1{,}96$, найдем границы искомого случайного интервала (см. соотношение (8.15)) $90{,}6 \pm 1{,}96 \cdot 40/\sqrt{10} = 90{,}6 \pm$ $\pm\, 24{,}48$ и сам доверительный интервал (66,12, 115,4). Таким образом, при изучении большого числа выборок по 10 крыс в каждой в предполагаемых условиях вычисленный в результате интервал должен в 95% случаях включать истинное среднее прибавление веса в популяции крыс.

Обратимся теперь к более естественной ситуации, когда оба параметра нормального распределения m и σ предполагаются неизвестными.

Теорема 2.2 Доверительным интервалом с доверительной вероятностью α для неизвестного математического ожидания $m =$ $= M(X)$ при неизвестной дисперсии σ^2 является случайный интервал

$$\left(\bar{X} - \frac{z_{\alpha,n-1}S}{\sqrt{n}}, \bar{X} + \frac{z_{\alpha,n-1}S}{\sqrt{n}}\right). \tag{8.16}$$

Здесь $S = \sqrt{S^2}$ (см. (8.8)), а $z_{\alpha,n-1} = t_{\text{кр}}(1-\alpha, n-1)$, где $t_{\text{кр}}(1-\alpha, k)$ можно определить по приведенной в приложении 2 таблице правых критических точек *распределения Стьюдента* с $k = n-1$ степенями свободы, отвечающих уровню значимости $1-\alpha$ (двусторонняя критическая область).

Замечание 2. Последовательность значений $z_{\alpha,k}$ убывает с ростом k, и $z_{\alpha,k} > z_{\alpha,\infty} = z_\alpha$. Это приводит к тому, что в типичных случаях длина доверительного интервала (8.16) больше длины доверительного интервала (8.15).

При доказательстве теоремы 2.2 используется то, что случайная величина $\frac{\sqrt{n}}{S}(\bar{X} - m)$, где S^2 — исправленная выборочная дисперсия, определенная в (8.8), имеет распределение, не зависящее от параметров m и σ нормального распределения, из которого извлекается выборка. В курсе математической статистики доказывается, что указанное распределение, называемое распределением Стьюдента, имеет симметричную относительно нуля плотность, которая с ростом n сходится к плотности нормального распределения $N(0,1)$.

Пример 2.3. Найти доверительный интервал с доверительной вероятностью 0,95 для среднего прибавления веса в популяции крыс по выборке из примера 2.2, считая, что генеральная дисперсия неизвестна.

Решение. Исправленная выборочная дисперсия S^2 по данным примера 2.2, рассчитанная по формуле (8.8), равна 992,8. Применяя теорему 2.2 и учитывая, что $z_{0,95,9} = 2{,}262$, найдем границы случайного интервала (см. (8.16)):

$$90{,}6 \pm 2{,}262\sqrt{\frac{992{,}8}{10}} = 90{,}6 \pm 22{,}5,$$

и сам доверительный интервал (68,06, 113,14).

Замечание 3. При построении доверительных интервалов (8.15) и (8.16) предполагается, что варианты X_i имеют нормальное распределение $N(m, \sigma)$, что предопределяет одно из основных достоинств доверительных интервалов — их неасимптотический характер (теорема 2.1 имеет место при любом $n \geqslant 1$, а теорема 2.1 — при любом $n \geqslant 2$). При большом n, однако, можно строить приближенные доверительные интервалы для неизвестного математического ожидания $m = M(X)$ без предположения нормальности. Так, случайный интервал

$$\left(\bar{X} - z_\alpha \frac{\widetilde{S}}{\sqrt{n}}, \bar{X} + z_\alpha \frac{\widetilde{S}}{\sqrt{n}}\right) \tag{8.17}$$

является доверительным интервалом для математического ожидания m с доверительной вероятностью, приблизительно равной α. Здесь z_α — то же, что в (8.15), а в качестве $\widetilde{S}^2$ может выступать какая-либо состоятельная оценка дисперсии σ^2, например S^2 или $\overline{S}^2$.

8.2.3. Статистические оценки случайных погрешностей измерений

Измерения неизвестных величин приходится осуществлять в самых различных областях человеческой деятельности. При этом результаты измерений X_1, $X_2, \ldots, X_n$ одной и той же величины, как правило, несколько отличаются друг от друга, так как невозможно полностью воспроизвести условия проведения различных измерений. Общая ошибка измерения часто складывается из большого числа незначительных ошибок. Если измерительный прибор не дает систематической ошибки, то в такой ситуации (с учетом центральной предельной теоремы) становится правдоподобным предположение, что каждое отдельное измерение X_i можно представить в виде суммы $X_i = m + \varepsilon_i$, где m представляет неизвестное истинное значение измеряемой величины, а $\{\varepsilon_i\}$ являются независимыми нормально распределенными случайными ошибками с нулевыми средними (следствие отсутствия систематической ошибки) и некоторыми одинаковыми, вообще говоря неизвестными, дисперсиями σ^2 (таким образом, $X_i \sim N(m, \sigma)$).

Другими словами, наблюдаемые величины $X_1, X_2, \ldots, X_n$ можно считать случайной выборкой объема n из генеральной совокупности всевозможных, проведенных в одних и тех же условиях измерений интересующей нас случайной величины $X \sim N(m, \sigma)$, и применять для их обработки такие статистические методы, как точечное и интервальное оценивание, и, в частности, использовать теоремы 2.1 и 2.2.

Принимая во внимание вышесказанное, будем считать, что справедливы следующие утверждения: если в одних и тех же условиях осуществлено n независимых измерений некоторой интересующей нас величины, точное значение m которой неизвестно, и результаты этих измерений обозначены через $x_1, x_2, \ldots, x_n$, то

$$m \approx \bar{x}, \tag{8.18}$$

где величина $\bar{x}$ определена формулой

$$\bar{x} = \frac{1}{n} \sum_{1 \leqslant k \leqslant n} x_k; \tag{8.19}$$

кроме того, с доверительной вероятностью α истинное значение m содержится в интервале

$$(\bar{x} - \Delta x, \bar{x} + \Delta x), \tag{8.20}$$

где

$$\Delta x = z_{\alpha, n-1} \sqrt{\frac{1}{n(n-1)} \sum_{1 \leqslant k \leqslant n} (x_k - \overline{x})^2}, \tag{8.21}$$

а множитель $z_{\alpha, n-1}$ такой же, как в (8.16).

Отметим, что величину Δx называют *абсолютной погрешностью* измерений, а соотношение (8.20) обычно записывают как $m = \bar{x} \pm \Delta x$.

Наряду с абсолютной погрешностью Δx используют также *относительную погрешность*

$$\delta x\,\% = \frac{\Delta x}{|\bar{x}|}\,100\%. \qquad (8.22)$$

Пример 2.4. Произведено пять независимых измерений веса таблетки лекарственного вещества. Получены следующие значения (в мг): 159,4, 159,7, 159,6, 159,3, 159,0.

Предполагая, что ошибки измерения распределены по нормальному закону, оценить истинный вес таблетки и найти погрешности его измерения при доверительной вероятности 0,99.

Решение. Для решения задачи воспользуемся формулами (8.18)–(8.22). Имеем, $\bar{x} = 159{,}4$, $z_{0,99,4} \approx 4{,}6$, $\Delta x \approx 4{,}6 \times 0{,}1225 \approx 0{,}5634 \approx 0{,}56$, $\delta x\,\% \approx 0{,}35\,\%$. Таким образом, истинный вес таблетки m оценивается числом 159,4. Кроме того, $m = 159{,}4 \pm 0{,}56$, т. е. с доверительной вероятностью 0,99 истинный вес таблетки может колебаться в пределах от 158,84 до 159,96.

Самостоятельная работа

1. По данным примеров 1.1 и 1.2 из упражнений к п. 8.1 найти выборочные средние и выборочные дисперсии, а также приближенные доверительные интервалы для математических ожиданий с доверительной вероятностью 0,95.

2. При измерении содержание кальция (мг%) в сыворотке крови подопытных животных получили следующие результаты: 14,5, 14,7, 14,8, 14,9, 15,1, 15,3, 15,5, 15,8, 15,9. Определить доверительный интервал для среднего значения содержание кальция с доверительной вероятностью 0,9 (предполагается, что показатель распределен по нормальному закону).

3. При подсчете количества листьев у шести лекарственных растений одного вида были получены следующие результаты: 5, 8, 10, 11, 12, 14. Определить доверительный интервал для среднего числа листьев с доверительной вероятностью 0,95 (предполагается, что показатель распределен по нормальному закону).

4. По данным девяти равноточных измерений физической величины 34, 40, 45, 43, 39, 41, 37, 48, 36 найти погрешность измерения истинного значения этой величины при доверительной вероятности 0,95 (предполагается, что погрешности измерения распределены по нормальному закону).

§ 8.3. Метод наименьших квадратов и сглаживание экспериментальных зависимостей

Предположим, что результаты некоторого эксперимента систематизированы в виде таблицы,

Таблица 8.8

x	x_1	x_2	...	x_n
y	y_1	y_2	...	y_n

в которой $y_i = y(x_i)$ является функцией, а x — аргументом, и требуется сгладить эту заданную таблично зависимость многочленом или некоторой другой функцией, известной нам с точностью до нескольких подлежащих определению параметров.

Задача сглаживания экспериментальных зависимостей достаточно типична для практики. Решая эту задачу, обычно рассчитывают освободить экспериментальные данные от случайных ошибок, допущенных в каждом отдельном опыте, и свести большое количество этих данных к нескольким параметрам (в частности, к коэффициентам многочлена), одновременно получив возможность обрабатывать полученную функциональную зависимость аналитически (например, дифференцировать).

Отметим, что практически любая табличная зависимость может быть сглажена различными функциями с приблизительно одинаковой точностью. В этой связи сглаживающую функцию следует выбирать исходя из физических соображений, а при отсутствии таковых использовать наиболее простые варианты формул, с минимальным числом подлежащих определению коэффициентов. Определенную помощь в выборе сглаживающей функции может также оказать графическое представление табл. 8.8.

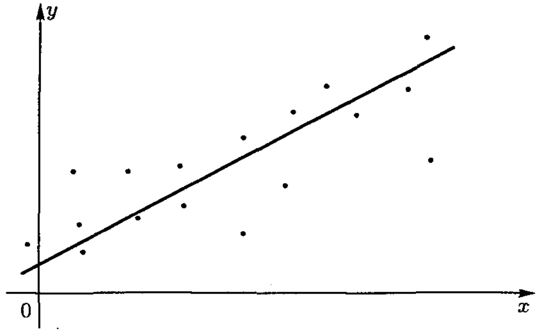


Рис. 8.7. Графическое изображение табл. 8.8.

Пусть, для определенности, функция, с помощью которой будет осуществляться сглаживание, является многочленом

$$Q_r(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_r x^r \tag{8.23}$$

известной степени $r \geqslant 1$ $(r < n)$ с числовыми коэффициентами $\alpha_0, \alpha_1, \ldots, \alpha_r$, подлежащими определению.

Составим разности

$$\varepsilon_i = y_i - Q_r(x_i), \quad i = 1, \ldots, n, \tag{8.24}$$

характеризующие близость табличных данных и значений, полученных при помощи сглаживающей функции. Нам следует подобрать коэффициенты $\alpha_0, \alpha_1, \ldots, \alpha_r$ таким образом, чтобы величины ε_i в совокупности были минимальными.

Эффективной процедурой для решения подобного сорта задач является так называемый *метод наименьших квадратов*, который гласит, что наилучшими во многих отношениях оценками для $\alpha_0, \alpha_1, \ldots, \alpha_r$ являются оценки, минимизирующие сумму квадратов разностей $\sum\limits_{1 \leqslant i \leqslant n} \varepsilon_i^2$. Иными словами, в качестве оценки для неизвестных параметров α_i следует взять такие значения a_i, при которых функция

$$S(\alpha_0, \alpha_1, \ldots, \alpha_r) = \sum_{1 \leqslant i \leqslant n} (y_i - Q_r(x_i))^2$$

достигает минимума. Поскольку $S = S(\alpha_0, \alpha_1, \ldots, \alpha_r)$ представляет дифференцируемую функцию r переменных, необходимым условием ее минимизации является равенство нулю частных производных $\dfrac{\partial S}{\partial \alpha_i}$, $0 \leqslant i \leqslant r$, вычисленных при $\alpha_0 = a_0, \ldots, \alpha_r = a_r$. Решение системы из $r+1$ *нормального* уравнения

$$\frac{\partial S}{\partial \alpha_i} = 0, \quad 0 \leqslant i \leqslant r, \tag{8.25}$$

в типичных случаях единственное, дает нам искомые коэффициенты $a_0, a_1, \ldots, a_r$.

Рассмотрим более детально линейную аппроксимацию экспериментальных зависимостей между величинами, т. е. сглаживание с помощью функции $Q_1(x) = \alpha_0 + \alpha_1 x$. В этом случае

$$S = \sum_{1 \leqslant i \leqslant n} (y_i - \alpha_o - \alpha_1 x_i)^2,$$

$$\frac{\partial S}{\partial \alpha_o} = \sum_{i=1}^{n} 2\,(y_i - \alpha_0 - \alpha_1 x_i)(-1),$$

$$\frac{\partial S}{\partial \alpha_1} = \sum_{i=1}^{n} 2\,(y_i - \alpha_0 - \alpha_1 x_i)(-x_i),$$

вследствие чего система нормальных уравнений принимает вид

$$\begin{cases} \displaystyle\sum_{i=1}^{n} (y_i - a_0 - a_1 x_i) = 0, \\ \displaystyle\sum_{i=1}^{n} (y_i - a_0 - a_1 x_i) x_i = 0. \end{cases} \tag{8.26}$$

Положим

$$\begin{aligned} \overline{y} &= \frac{1}{n}\sum_{i=1}^{n} y_i, \qquad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \\ \overline{xy} &= \frac{1}{n}\sum_{i=1}^{n} x_i y_i, \qquad \overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2. \end{aligned} \tag{8.27}$$

Тогда уравнения (8.26) можно переписать в виде системы двух линейных уравнений с двумя неизвестными:

$$\begin{cases} a_0 + \overline{x}\,a_1 = \overline{y}, \\ \overline{x}\,a_0 + \overline{x^2}a_1 = \overline{xy} \end{cases} \tag{8.28}$$

Решив эту систему, найдем

$$a_0 = \bar{y} - a_1\bar{x}, \quad a_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}. \tag{8.29}$$

Таким образом, наилучшая в смысле метода наименьших квадратов линейная сглаживающая функция выражается уравнением $y = a_0 + a_1 x$.

Замечание 1. Сумма $\sum\limits_{1\leqslant i\leqslant n} \varepsilon_i$ оптимальных разностей равняется нулю, что может быть использовано для контроля правильности вычислений.

Пример 3.1. Сгладить линейной зависимостью от x следующие табличные данные:

Таблица 8.9

x	-6	-2	-1	1	3	5
y	0,2	1	1,5	2	3,4	3,9

Вычислить разности и отобразить на графике табличные данные и сглаживающую прямую.

Решение. Используя формулы (8.27), найдем $\overline{x} = 0$, $\overline{y} = 2$, $\overline{xy} = 4{,}5$, $\overline{x^2} = 12{,}5$ и, следовательно, $a_0 = 2$, $a_1 = 0{,}36$. Таким

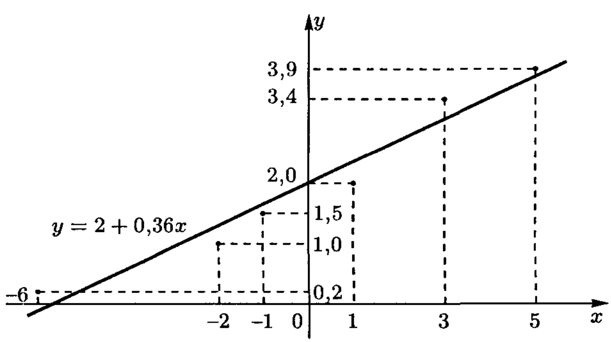


Рис. 8.8. График сглаживающей прямой.

образом, уравнение сглаживающей прямой имеет вид $y = 2 + 0{,}36x$ (рис. 8.8). Теперь дополним табл. 8.9:

Таблица 8.10

x	-6	-2	-1	1	3	5
y	0,2	1	1,5	2	3,4	3,9
$2+0{,}36x$	−0,16	1,28	1,64	2,36	3,08	3,8
Разности	0,36	−0,28	−0,14	−0,36	0,32	0,1

Заметим, что сумма разностей равна нулю.

Замечание 2. Метод наименьших квадратов можно с некоторой потерей точности использовать для сглаживания функциональных зависимостей, приводящихся к линейной с помощью замены переменных. Так, зависимость $y = \alpha_0 e^{\alpha_1 x}$ может быть переписана в виде $z = \beta + \alpha_1 x$, где $z = \ln y$ и $\beta = \ln \alpha_0$. Применяя метод наименьших квадратов для нахождения коэффициентов линейного уравнения $z = \beta + \alpha_1 x$ и возвращаясь к первоначальной зависимости, получим в качестве оценки коэффициента α_0 величину e^{β} (оценка для α_1 в данном случае остается без изменений).

Пример 3.2. Пусть данные некоторых измерений представлены таблицей:

Таблица 8.11

x	-1	0	2	3	5
z	1,8	2	2,4	2,7	3,3

Требуется сгладить их при помощи формулы $z = \alpha_0 e^{\alpha_1 x}$ и вычислить разности между табличными и сглаженными значениями с точностью до тысячных.

Решение. Воспользуемся замечанием 2.

1) Заменим числа z из табл. 8.11 числами $y = \ln z$:

Таблица 8.12

x	-1	0	2	3	5
y	0,5877	0,6931	0,8754	0,9932	1,193

2) По формулам (8.27) и (8.29) найдем коэффициенты $a_0 = 0{,}6876$ и $a_1 = 0{,}1006$ линейного уравнения $a_0 + a_1 x$, наилучшим образом приближающего y.

3) Учитывая, что коэффициенты α_0 и α_1 связаны с a_0 и a_1 соотношениями $\alpha_0 = e^{a_0}$, $\alpha_1 = a_1$ окончательно получим $\alpha_0 = 1{,}988$, $\alpha_1 = 0{,}1006$, и, следовательно, $z = 1{,}988\, e^{0{,}1006x}$.

Разности между табличными и сглаженными значениями z, сведены в таблицу:

Таблица 8.13

x	-1	0	2	3	5
ε	0,002	0,012	$-0{,}031$	0,012	0,012

Заметим, что сумма разностей здесь отлична от нуля.

Самостоятельная работа

Используя метод наименьших квадратов, сгладить с помощью функций заданного вида следующие табличные зависимости:

1. $y = a\, e^{bx}$

x	-1	0	2	3	5
y	1	2	5	9	24

x	1,8	2	2,4	2,7	3,3
y	5	4	7	8	10

2. $y = ax^k$

x	1	2	4	7	8	10
y	2	8	25	55	68	80

x	0,5	1	2	4	6
y	1	3	9	22	45

3. $y = \dfrac{1}{ax + b}$

x	-2	-1	1	2	3	5
y	1	0,8	0,4	0,1	0	0

x	-2	0	1	1,2	1,5
y	1	0,3	0,2	0,2	0,2

Примечание. Перед использованием метода наименьших квадратов сглаживающие зависимости следует линеаризовать (по параметрам a, b или a, k, используя в первых двух случаях операцию логарифмирования, а в третьем случае взяв $1/y$ вместо y).

§ 8.4. Элементы корреляционно-регрессионного анализа

Объекты ряда генеральных совокупностей обладают несколькими подлежащими изучению признаками X, $Y, \ldots$, которые можно интерпретировать как систему взаимосвязанных величин. Примерами являются:

– масса животных (X) и количество гемоглобина (Y) в их крови,

– рост (X) мужчины и объем (Y) его грудной клетки

– количество (X) вводимого объекту препарата и его концентрация в крови (Y).

Отметим, что в двух первых примерах как X, так и Y представляют собой случайные величины, а в третьем — параметр X не является случайным, поскольку экспериментатор может его контролировать. Очевидно, что в приведенных примерах величины Y и X не связаны функционально, поскольку на значения, принимаемые случайной величиной Y, влияют помимо X многие другие факторы, и в то же время они не являются независимыми. В таких случаях говорят, что Y и X связаны *стохастической* (т. е. случайной) *зависимостью.* Понятие стохастической зависимости обобщает понятие функциональной (или детерминированной) зависимости, подобно тому как понятие случайной величины обобщает понятие (детерминированной) величины.

В дальнейшем, в основном, будем исследовать зависимость между двумя признаками X и Y объектов генеральной совокупности, где случайная величина Y играет роль функции, а X — случайного или детерминированного аргумента. Кроме того, мы ограничимся частным случаем стохастической зависимости величин X и Y, а именно *регрессионной зависимостью*, при которой изменение величины X влечет изменение математического (п. 8.4.1) или *условного математического* (п. 8.4.2 и 8.4.3) ожидания случайной величины Y.

8.4.1. Линейная и полиномиальная регрессии

Во многих экспериментальных исследованиях вид зависимости между различными величинами можно, в той или иной степени, выявить лишь посредством ряда наблюдений. Рассмотрим достаточно типичный случай. Предположим, что изучается зависимость между величинами x и Y (скажем, затратами на функционирование и величиной годовой прибыли некоторой аптеки), и из условий эксперимента следует, что при значениях $x_1, x_2, \ldots, x_n$, известных с устраивающей исследователя точностью, соответствующие значения $Y_1, Y_2, \ldots, Y_n$ определяются с некоторыми несистематическими (т. е. имеющими нулевое среднее) случайными ошибками (мы специально используем обозначение x, а не X, чтобы подчеркнуть тот факт, что величина x не является случайной).

Будем считать, что наблюдаемые величины Y_i представимы в виде

$$Y_i = Q_r(x_i) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{8.30}$$

где $Q_r(x)$ — некоторая функция, определенная с точностью до $r \geqslant 1$ $(r < n)$ неизвестных нам числовых параметров $\alpha_0, \alpha_1, \ldots, \alpha_r$, а ошибки ε_i — независимые случайные величины с нулевыми средними $M(\varepsilon_i) = 0$ и конечными одинаковыми дисперсиями $D(\varepsilon_i) = \sigma^2$.

Из условия (8.30) следует, что $Y_1, Y_2, \ldots, Y_n$ являются независимыми случайными величинами, дисперсии которых одинаковы, а средние значения $M(Y_i)$ задаются *функцией регрессии* $M(Y_i) = Q_r(x_i)$.

Заметим, что вид функции регрессии должен быть определен заранее, из физических или иных соображений. Определенной подсказкой здесь может служить *корреляционое поле* точек (см. рис. 8.7 при $y = Y$), т. е. графическое представление наблюдений (x_i, Y_i), дающее возможность визуально оценить, с каким типом зависимости (линейной, экспоненциальной и т. п.) между величинами x и Y мы имеем дело.

В дальнейшем будем считать, что

$$Q_r(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_r x^r \tag{8.31}$$

— многочлен известной степени $r \geqslant 1$ $(r < n)$ с неизвестными нам числовыми коэффициентами $\alpha_0, \alpha_1, \ldots, \alpha_r$, т. е. уравнение регрессии имеет вид $M(Y_i) = \alpha_0 + \alpha_1 x + \ldots + \alpha_r x_i^r$. Такая модель носит название *полиномиальной регрессионной*. Если $M(Y_i) = \alpha_0 + \alpha_1 x_i$, говорят о *линейной регрессии*.

Нам требуется по наблюдениям (x_i, Y_i) определить *коэффициенты регрессии* $\alpha_0, \alpha_1, \ldots, \alpha_r$.

Поскольку ошибки ε_i в равенстве (8.30) предполагаются случайными, а число уравнений n больше числа параметров r, подлежащих определению, поставленная задача не имеет единственного решения. Мы можем лишь оценить неизвестные параметры $\alpha_0, \alpha_1, \ldots, \alpha_r$, используя содержащуюся в уравнениях (8.30) информацию. С этой целью воспользуемся примененным для решения аналогичной задачи в п. 8.3 методом наименьших квадратов, в соответствии с которым оценками для $\alpha_0, \alpha_1, \ldots$ $\ldots, \alpha_r$ являются оценки, минимизирующие сумму квадратов ошибок $\sum\limits_{1\leqslant i\leqslant n} \varepsilon_i^2$.

Такая оценка параметров полиномиальной регрессионной модели (подробнее см. п. 8.3, формулы (8.25) и далее) называется *оценкой по методу наименьших квадратов* (МНК-оценкой). Важно отметить, что в случае полиномиальной регрессионной модели (и вообще, при линейном оценивании) МНК-оценки являются несмещенными (см. (8.6)) и, в некотором смысле, самыми точными.

В случае линейной регрессии, т. е. модели, которая в наших обозначениях имеет вид

$$\begin{gathered} Y_i = a_0 + a_1 x_i + \varepsilon_i, \\ M(\varepsilon_i) = 0, \quad D(\varepsilon_i) = \sigma^2, \\ i = 1, \ldots, n, \end{gathered} \tag{8.32}$$

рассуждения, позволяющие вычислить МНК-оценки a_0 и a_1 для параметров α_0 и α_1 из (8.32), дословно совпадают с соответствующими выкладками из п. 8.3. Единственное отличие заключается в том, что в формулах (8.26)–(8.29) следует положить Y_i вместо y_i. Таким образом,

$$a_0 = \overline{Y} - a_1 \overline{x}, \quad a_1 = \frac{\overline{xY} - \overline{x}\,\overline{Y}}{\overline{x^2} - \overline{x}^2}, \tag{8.33}$$

где

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$
$$\overline{xY} = \frac{1}{n}\sum_{i=1}^{n} x_i Y_i, \quad \overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2. \tag{8.34}$$

Пример 4.1. Предположим, что x — затраты на функционирование и Y — величина годовой прибыли некоторой аптеки в течение 5 лет представлены в условных единицах следующей таблицей:

Таблица 8.14

x	6	3	7	5	10
Y	33	27	32	28	42

Считая, что между затратами и прибылью имеет место линейная зависимость вида $Y = \alpha_0 + \alpha_1 x + \varepsilon$ с постоянными коэффициентами α_0, α_1 и величиной независимого от года к году случайного влияния ε с нулевым средним и конечной дисперсией, найти МНК-оценки параметров α_0 и α_1.

Решение. Зависимость y от x, представленная по годам, имеет вид

$$Y_i = \alpha_0 + \alpha_1 x_i + \varepsilon, \quad i = 1, \ldots, 5, \tag{8.35}$$

т. е. описывается линейной регрессионной моделью (8.32). По формулам (8.33) и (8.34) для данных из табл. 8.14 найдем $\overline{x} = 6{,}2$, $\overline{Y} = 32{,}4$, $\overline{xY} = 212{,}6$, $\overline{x^2} = 43{,}8$, откуда $\alpha_0 = 18{,}84$, $\alpha_1 = 2{,}186$.

8.4.2. Совместный закон распределения и числовые характеристики двух случайных величин

Для того чтобы в дальнейшем уметь правильно определять выборочные характеристики взаимосвязанных случайных величин, нам понадобятся некоторые сведения из теории вероятностей.

Рассмотрим случайные величины X и Y, получающие свои значения в результате одного и того же случайного эксперимента.

Сначала будем предполагать, что X и Y являются дискретными случайными величинами и принимают значения $x_1, x_2, \ldots$ $\ldots, x_m$ и $y_1, y_2, \ldots, y_n$ соответственно, где m и n — некоторые целые положительные числа. Для $1 \leqslant i \leqslant m$ и $1 \leqslant j \leqslant n$ определим вероятности P_{ij}, положив

$$P_{ij} = P(X = x_i, Y = y_j), \tag{8.36}$$

т. е. P_{ij} равно вероятности того, что события $\{X = x_i\}$ и $\{Y = y_j\}$ произойдут одновременно.

Определение. *Совместным законом распределения* пары случайных величин X и Y, или законом распределения двумерного случайного вектора (X, Y), называется соответствие, которое каждой паре значений (x_i, y_j) дискретных случайных величин X и Y ставит в соответствие ее вероятность P_{ij}.

Указанные соответствия можно записать в виде таблицы

Таблица 8.15

X \ Y	y_1	y_2	$\ldots$	y_j	$\ldots$	y_n	P_x
x_1	P_{11}	P_{12}	$\ldots$	P_{1j}	$\ldots$	P_{1n}	$P_{1\cdot}$
x_2	P_{21}	P_{22}	$\ldots$	P_{2j}	$\ldots$	P_{2n}	$P_{2\cdot}$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
x_i	P_{i1}	P_{i2}	$\ldots$	P_{ij}	$\ldots$	P_{in}	$P_{i\cdot}$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
x_m	P_{m1}	P_{m2}	$\ldots$	P_{mj}	$\ldots$	P_{mn}	$P_{m\cdot}$
P_y	$P_{\cdot 1}$	$P_{\cdot 2}$		$P_{\cdot j}$		$P_{\cdot n}$	1

в которой

$$\begin{aligned} P_{i\cdot} &= \sum_{j=1}^{n} P_{ij}, \quad 1 \leqslant i \leqslant m, \\ P_{\cdot j} &= \sum_{i=1}^{m} P_{ij}, \quad 1 \leqslant j \leqslant n, \end{aligned} \tag{8.37}$$

т. е. $\mathrm{P}_{i\cdot}$ равняется сумме вероятностей, расположенных в i-й строке, а $P_{\cdot j}$ равняется сумме вероятностей из j-го столбца.

Очевидно, что

$$\sum_{i=1}^{m}\sum_{j=1}^{n} P_{ij} = \sum_{i=1}^{m} P_{i\cdot} = \sum_{j=1}^{n} P_{\cdot j} = 1;$$

кроме того,

$$P_{i\cdot} = P(X = x_i), \quad P_{\cdot j} = P(Y = y_j). \tag{8.38}$$

Таким образом, индивидуальные законы распределения собственно случайных величин X и Y можно записать в следующем виде:

Таблица 8.16

X	x_1	x_2	...	x_i	...	x_m
P	$P_{1\cdot}$	$P_{2\cdot}$	...	$P_{i\cdot}$	...	$P_{m\cdot}$

Таблица 8.17

Y	y_1	y_2	...	y_j	...	y_n
P	$P_{\cdot 1}$	$P_{\cdot 2}$	...	$P_{\cdot j}$	...	$P_{\cdot n}$

Определение. Дискретные случайные величины X и Y называются независимыми, если при любых i $(1 \leqslant i \leqslant m)$ и j $(1 \leqslant j \leqslant n)$ выполняются соотношения

$$P(X = x_i, Y = y_j) = P(X = x_i)\, P(Y = y_j), \tag{8.39}$$

или (см. (8.38)), равносильно, $P_{ij} = P_{i\cdot}P_{\cdot j}$.

Таким образом, если случайные величины X и Y независимы, то их совместный закон распределения взаимно-однозначно определяется по индивидуальным законам распределения (см. табл. 8.15–8.17). Более интересен случай, когда условие (8.39) не выполняется и случайные величины X и Y связаны стохастической зависимостью.

Далее приведем некоторые числовые характеристики рассматриваемых дискретных случайных величин X и Y (см. (8.36) и табл. 8.15). Предварительно напомним, что для любой функции двух переменных $h(x, y)$ математическое ожидание $M(h(X, Y))$

случайной величины $h(X, Y)$ вычисляется по формуле

$$M(h(X,Y)) = \sum_{i=1}^{m}\sum_{j=1}^{n} h(x_i, y_j)P_{ij}. \tag{8.40}$$

Математические ожидания $M(X)$, $M(Y)$ случайных величин X и Y, а также их дисперсии $D(X)$, $D(Y)$ и стандартные отклонения σ_x, σ_y, с учетом условия (8.38), рассчитываются по стандартным (см. (7.27), (7.30)) формулам:

$$m_x = M(X) = \sum_{i=1}^{m} x_i P_{i\cdot}, \quad m_y = M(Y) = \sum_{j=1}^{n} y_j P_{\cdot j}, \tag{8.41}$$

$$\begin{aligned} D(X) &= \sum_{i=1}^{m}(x_i - m_x)^2 = \sum_{i=1}^{m} x_i^2 - m_x^2, \quad \sigma_x = \sqrt{D(X)}, \\ D(Y) &= \sum_{j=1}^{n}(y_j - m_y)^2 = \sum_{j=1}^{n} y_j^2 - m_y^2, \quad \sigma_y = \sqrt{D(Y)}, \end{aligned} \tag{8.42}$$

в которых используются лишь данные из табл. 8.16 и 8.17.

Две следующие характеристики случайных величин X и Y, в отличие от предыдущих, в известной степени учитывают связь между этими величинами.

Определение. *Ковариацией*, или *корреляционным моментом*, случайных величин X и Y называется число

$$\mathrm{cov}(X,Y) = M((X - M(X))(Y - M(Y))).$$

Отметим, что $\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$ и

$$\mathrm{cov}(X,Y) = M(XY) - M(X)M(Y). \tag{8.43}$$

Важным свойством ковариации является то, что для независимых случайных величин X и Y их ковариация обращается в нуль. Отсюда, в частности, следует, что если ковариация двух случайных величин не равна нулю, то они стохастически зависимы. В то же время, условие $\mathrm{cov}(X, Y) = 0$ не означает независимость величин X и Y.

Принимая во внимание формулу (8.40), видим, что

$$\operatorname{cov}(X,Y)=\sum_{i=1}^{m}\sum_{j=1}^{n}(x_i-m_x)(y_j-m_y)P_{ij}=$$

$$=\sum_{i=1}^{m}x_i\sum_{j=1}^{n}y_jP_{ij}-m_xm_y. \qquad (8.44)$$

Пример 4.2. Пусть X и ε — независимые случайные величины, $D(X)=\sigma_x^2$, $M(\varepsilon)=0$, $D(\varepsilon)=\sigma_\varepsilon^2$ и

$$Y=a+bX+\varepsilon. \qquad (8.45)$$

Требуется найти ковариацию случайных величин X и Y.

Решение.

$$M(Y)=M(a+bX+\varepsilon)=M(a)+M(bX)+M(\varepsilon)=a+bM(X),$$

$$M(XY)=M(aX+bX^2+X\varepsilon)=aM(X)+bM(X^2)+M(X\varepsilon).$$

Учитывая, что $M(X\varepsilon)=M(X)M(\varepsilon)=0$ (по независимости случайных величин X и ε), и используя формулу (8.43), получим

$$\operatorname{cov}(X,Y)=b(M(X^2)-M^2(X))=bD(X)=b\sigma_x^2.$$

Итак, $\operatorname{cov}(X,Y)=b\sigma_x^2$.

Определение. *Коэффициентом корреляции* случайных величин X и Y называется число

$$r=r(X,Y)=\frac{\operatorname{cov}(X,Y)}{\sigma_x\sigma_y}. \qquad (8.46)$$

Пример 4.3. По данным примера 4.2 найти коэффициент корреляции r случайных величин X и Y.

Решение. Используя свойства дисперсии и, в частности, то, что дисперсия суммы независимых случайных величин равняется сумме их дисперсий, получим

$$D(Y)=D(a+bX)+D(\varepsilon)=D(bX)+D(\varepsilon)=b^2\sigma_x^2+\sigma_\varepsilon^2.$$

Отсюда по формуле (8.46) найдем

$$r=\frac{b\sigma_x^2}{\sqrt{\sigma_x^2(b^2\sigma_x^2+\sigma_\varepsilon^2)}}=\frac{b\sigma_x}{\sqrt{b^2\sigma_x^2+\sigma_\varepsilon^2}}. \qquad (8.47)$$

Коэффициент корреляции является безразмерной характеристикой, которая используется в качестве меры *линейной зависимости* случайных величин. Чем его модуль ближе к единице, тем, вообще говоря, *теснее* линейная зависимость между величинами.

Вышесказанное подтверждают результаты примеров 4.2 и 4.3. В самом деле, из формулы (8.47) вытекает, что если модуль коэффициента корреляции стремится к единице, то σ_ε^2 стремится к нулю. В то же время из уравнения (8.45) следует, что случайная величина Y тем «ближе» к $a + bX$, чем меньше $D(\varepsilon) = \sigma_\varepsilon^2$. Графически это означает, что случайным образом выбранные точки (X, Y) с ростом модуля коэффициента корреляции будут в среднем теснее группироваться относительно прямой $a + bX$.

Корреляционное поле точек (X, Y), изображенное на рис. 8.10, указывает на более тесную линейную зависимость между X и Y (и соответствует большему значению коэффициента корреляции), чем аналогичное корреляционное поле на рис. 8.9.

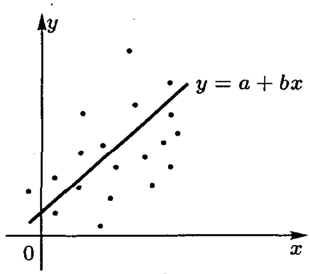


Рис. 8.9. Объяснения в тексте.

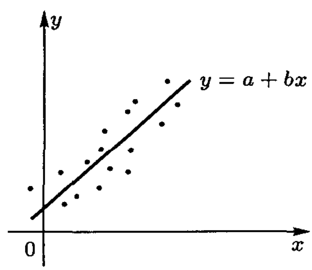


Рис. 8.10. Объяснения в тексте.

Понятно, что коэффициент корреляции и ковариация обращаются в нуль одновременно; следовательно, коэффициент корреляции независимых случайных величин равен нулю.

Определение. Если коэффициент корреляции случайных величин X и Y равен нулю, то такие величины называются *некоррелированными*.

Перечислим основные свойства коэффициента корреляции.

1) $|r| \leqslant 1$.

2) $|r| = 1$ тогда и только тогда, когда $Y = a + bX$ с вероятностью единица, где a и b — некоторые постоянные.

3) Если величины X и Y независимы, то они некоррелированы.

На практике некоррелированность часто отождествляют с независимостью, что, как уже отмечалось, неправомерно. Тем не менее в важном частном случае *совместного нормального распределения* величин X и Y некоррелированность и независимость являются равносильными понятиями.

Определение. Прямая, имеющая уравнение

$$y = m_y + r\frac{\sigma_y}{\sigma_x}(x - m_x), \tag{8.48}$$

называется *прямой средней квадратической регрессии* Y на X.

Эта прямая наилучшим (в смысле среднего квадратического) образом приближает случайную величину Y линейной функцией $a + bX$ (другими словами, дисперсия $D(Y - a - bX)$ становится наименьшей, когда уравнение $y = a + bx$ имеет вид (8.48)).

Меняя в соотношении (8.48) X и Y ролями, мы получаем прямую средней квадратической регрессии X на Y:

$$x = m_x + r\frac{\sigma_x}{\sigma_y}(y - m_y). \tag{8.49}$$

Легко понять, что линии регрессии (8.48) и (8.49) будут совпадать тогда и только тогда, когда $r = \pm 1$.

Замечание. Совместный закон распределения может быть определен для пары случайных величин X и Y общего вида и, в частности, для непрерывных случайных величин. В последнем случае равенство (8.36) заменяется его дифференциальным (при $0 < dx,\, dy \to 0$) аналогом

$$P(x \leqslant X < x + dx, y \leqslant Y < y + dy) \sim f(x, y)\,dxdy \tag{8.50}$$

Функция $f(x, y)$ из соотношения (8.50) называется *совместной плотностью* распределения случайных величин X и Y. Она должна удовлетворять определенным условиям, в частности быть неотрицательной. Из ее свойств вытекает, что функции

$$f_x(x) = \int\limits_{-\infty}^{\infty} f(x,y)\,dy \quad \text{и} \quad f_y(y) = \int\limits_{-\infty}^{\infty} f(x,y)\,dx \tag{8.51}$$

являются плотностями распределений случайных величин X и Y соответственно.

По аналогии с (8.39) непрерывные случайные величины X и Y называются независимыми, если $f(x,y) = f_x(x)f_y(y)$ при любых x и y.

Естественно, что, также как ранее в случае (8.39), совместный закон распределения независимых величин X и Y взаимно-однозначно определяется по их индивидуальным плотностям. Наконец, отметим, что определения ковариации, коэффициента корреляции и прямых средней квадратической регрессии не меняются, а их свойства сохраняются.

Введем еще одно понятие, характеризующее степень зависимости между случайными величинами X и Y.

Определение. *Функцией регрессии* (или просто *регрессией*) Y на X называется *условное математическое ожидание* $M(Y|X = x)$ случайной величины Y при условии, что случайная величина X приняла значение x.

График $M(Y|X = x)$, как функции x, называют *линией регрессии.*

Регрессия Y на X обладает оптимальным свойством. Она, как функция X, наилучшим образом приближает случайную величину Y в смысле среднего квадратического, т. е. дисперсия $D(Y - g(X))$ принимает минимальное значение, когда $g(x) = M(Y|X = x)$.

В дискретном случае (8.36) регрессия Y на X определена лишь в точках x_i, $1 \leqslant i \leqslant m$, и принимает в них значения

$$M\,(Y|X = x_i) = \sum_{j=1}^{n} y_j \frac{P_{ij}}{P_{i\cdot}} = \frac{1}{P_{i\cdot}} \sum_{j=1}^{n} y_j P_{ij}, \quad 1 \leqslant i \leqslant m, \tag{8.52}$$

являясь, таким образом, средневзвешенной зависимостью Y от значений случайной величины X.

В непрерывном случае (8.50) при некоторых дополнительных предположениях функция регрессии Y на X задается формулой

$$M(Y|X=x) = \frac{1}{f_x(x)} \int_{-\infty}^{\infty} y\, f(x,y)\, dy$$

Регрессия X на Y определяется аналогично. Так, в дискретном случае

$$M(X|Y=y_j) = \frac{1}{P_{\cdot j}} \sum_{i=1}^{m} x_i P_{ij}, \quad 1 \leqslant j \leqslant n.$$

Если регрессия Y на X является линейной по x, то она называется *линейной регрессией* Y на X. Линия регрессии в этом случае является прямой и задается уравнением (8.48), т. е. совпадает с прямой средней квадратической регрессии Y на X.

Можно показать, что для зависимости $Y = a + bX + \varepsilon$ из примера 4.2 условное математическое ожидание $M(Y|X=x) = a + bx$, т. е. регрессия является линейной. Другим примером пары случайных величин Y и X, регрессии Y на X (и X на Y) которых линейны, являются непрерывные случайные величины Y, X с совместным нормальным распределением.

Пример 4.4. По закону распределения двумерного случайного вектора (X, Y)

Таблица 8.18

X	Y		
	3	5	P_x
2	0,3	0,1	0,4
3	0,4	0,2	0,6
P_y	0,7	0,3	1

найти:

1) математические ожидания и стандартные отклонения X и Y;

2) корреляционный момент и коэффициент корреляции X и Y;

3) уравнение прямой средней квадратической регрессии Y на X;

4) значения регрессии Y на X;

5) сравнить между собой при каждом значении X приближения средних значений Y, полученные по функции регрессии и по уравнению прямой средней квадратической регрессии.

Решение.

1) Формулы (8.41) и (8.42):

$$M(X) = 2 \cdot 0{,}4 + 3 \cdot 0{,}6 = 2{,}6, \quad M(Y) = 3 \cdot 0{,}7 + 5 \cdot 0{,}3 = 3{,}6,$$

$$D(X) = 2^2 \cdot 0{,}4 + 3^2 \cdot 0{,}6 - 2{,}6^2 = 0{,}24, \quad \sigma_x = \sqrt{0{,}24} = 0{,}4898,$$

$$D(Y) = 3^2 \cdot 0{,}7 + 5^2 \cdot 0{,}3 - 3{,}6^2 = 0{,}84, \quad \sigma_y = \sqrt{0{,}84} = 0{,}9165.$$

2) Формулы (8.44) и (8.46):

$$\operatorname{cov}(X, Y) = 2 \cdot (3 \cdot 0{,}3 + 5 \cdot 0{,}1) + 3 \cdot (3 \cdot 0{,}4 + 5 \cdot 0{,}2)) - 2{,}6 \cdot 3{,}6 = 0{,}04,$$

$$r = \frac{0{,}04}{0{,}4898 \cdot 0{,}9165} = 0{,}0891.$$

3) Формула (8.48) приводит к уравнению линейной регрессии

$$y_{\text{лин}} = 3{,}6 + 0{,}0891 \frac{0{,}9165}{}0{,}4898(x - 2{,}6) = 3{,}157 + 0{,}1667x;$$

$$y_{\text{лин}}|_{x=2} = 3{,}5, \quad y_{\text{лин}}|_{x=3} = 3{,}667$$

4) Формула (8.52) дает

$$M\,(Y|X = 2) = \frac{1}{0{,}4}(3 \cdot 0{,}3 + 5 \cdot 0{,}1) = 3{,}5,$$

$$M\,(Y|X = 3) = \frac{1}{0{,}6}(3 \cdot 0{,}4 + 5 \cdot 0{,}2) = 3{,}667.$$

Таким образом,

Таблица 8.19

X	$y_{\text{лин}}$	$M(Y\|X)$
2	3,5	3,5
3	3,667	3,667

Заключение. Величины, вычисленные путем подстановки возможных значений X в уравнение прямой средней квадратической регрессии и в функцию регрессии, практически совпадают.

8.4.3. Построение выборочной линии регрессии

В настоящем параграфе исследуется случай, когда интересующие нас признаки объектов генеральной совокупности можно рассматривать как двумерный случайный вектор (X, Y) с частично или полностью неизвестным совместным законом распределения. Нашей целью является вычисление статистических оценок основных характеристик этого распределения по наблюдениям выборки из генеральной совокупности, и, в частности, *выборочной ковариации, выборочного коэффициента корреляции* и *выборочных условных средних* (статистических аналогов функции регрессии), а также выявление стохастической зависимости между случайными величинами X и Y.

Определение. Случайной выборкой объема n, отвечающей паре случайных величин (X, Y), назовем набор n независимых пар случайных величин $(X_1, Y_1), \ldots, (X_n, Y_n)$, каждая из которых имеет такой же закон распределения, как и пара величин (X, Y).

Другими словами, случайной выборкой объема n можно считать величины (X_i, Y_i), $1 \leqslant i \leqslant n$, полученные в результате n независимых «одинаковых» случайных экспериментов.

Оценками для математических ожиданий $M(Y)$ и $M(X)$, построенными по выборке $(X_1, Y_1), \ldots, (X_n, Y_n)$, являются хорошо знакомые выборочные средние $\overline{Y}$, $\overline{X}$ случайных величин Y и X соответственно:

$$\overline{Y} = \frac{Y_1 + \ldots + Y_n}{n}, \qquad \overline{X} = \frac{X_1 + \ldots + X_n}{n}, \tag{8.53}$$

а оценками для дисперсий $D(Y)$, $D(X)$ — исправленные выборочные дисперсии Y и X (см. п. 8.2.1):

$$S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2, \quad S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2; \tag{8.54}$$

Оценкой для ковариации $\operatorname{cov}(X,Y)$ случайных величин X и Y является *исправленная* выборочная ковариация

$$l_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}). \tag{8.55}$$

Отметим, что справедливо равенство

$$l_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}X_iY_i - \frac{n}{n-1}\overline{X}\,\overline{Y}.$$

Можно доказать, что l_{xy} является несмещенной состоятельной оценкой ковариации $\operatorname{cov}(X,Y)$.

В качестве оценки для коэффициента корреляции $r(X,Y)$ используется выборочный коэффициент корреляции

$$\overline{r} = \frac{l_{XY}}{\sqrt{S_X^2 S_Y^2}}, \tag{8.56}$$

являющийся состоятельной оценкой коэффициента корреляции $r(X,Y)$.

Существенность выборочного коэффициента корреляции приводит к предположению о близкой к линейной зависимости между величинами X и Y. Последнее возможно визуально оценить по виду корреляционного поля точек. Например, корреляционное поле точек, изображенных на рис. 8.10, позволяет высказать гипотезу о линейной регрессии Y на X, т.е. дает основание рассчитывать на то, что теоретическое *уравнение регрессии* имеет вид $M(Y \mid X = x) = \alpha_0 + \alpha_1 x$.

Следующей задачей является нахождение оценок коэффициентов регрессии α_0, α_1. Вновь, как и в п. 8.4.1, обратимся к методу наименьших квадратов и будем искать такие значения a_0, a_1 величин α_0, α_1, которые минимизируют сумму

$$\sum_{i=1}^{n}(Y_i - \alpha_0 - \alpha_1 X_i)^2. \tag{8.57}$$

Имеем (см. (8.53) и формулы (8.33), (8.34) с заменой в их обозначениях x на X)

$$a_0 = \overline{Y} - a_1\overline{X}, \quad a_1 = \frac{\overline{XY} - \overline{X}\,\overline{Y}}{\overline{X^2} - \overline{X}^2}, \tag{8.58}$$

где использованы обозначения

$$\overline{X^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2, \quad \overline{Y^2} = \frac{1}{n}\sum_{i=1}^{n} Y_i^2, \quad \overline{XY} = \frac{1}{n}\sum_{i=1}^{n} X_i Y_i. \tag{8.59}$$

Принимая во внимание формулы

$$\frac{S_Y}{S_X} = \sqrt{\frac{\overline{Y^2} - \overline{Y}^2}{\overline{X^2} - \overline{X}^2}}, \quad \overline{r} = \frac{\overline{XY} - \overline{X}\,\overline{Y}}{\sqrt{(\overline{X^2} - \overline{X}^2)(\overline{Y^2} - \overline{Y}^2)}}, \tag{8.60}$$

видим, что выборочное уравнение прямой средней квадратической регрессии Y на X имеет вид

$$y = \overline{Y} + \overline{r}\frac{S_Y}{S_X}(x - \overline{X}) \tag{8.61}$$

и отличается от своего теоретического аналога (8.48) лишь заменой параметров m_y, m_x, r, σ_y, σ_x их оценками. То же можно сказать о выборочном уравнении прямой среднеквадратической регрессии X на Y

$$x = \overline{X} + \overline{r}\frac{S_X}{S_Y}(y - \overline{Y}) \tag{8.62}$$

и уравнении (8.49).

Отметим, что прямые, имеющие уравнения (8.60) и (8.61), проходят через точку $(\overline{X}, \overline{Y})$, а их угловые коэффициенты совпадают по знаку с $\overline{r}$ и что в практическом плане коэффициенты этих уравнений удобно вычислять при помощи формул (8.60).

Пример 4.5. Изучалась зависимость между систолическим давлением Y мужчин в начальной стадии шока и возрастом X. Результаты наблюдений приведены в таблице в виде двумерной выборки объема 11:

Таблица 8.20

X	68	37	50	53	75	66	52	65	74	65	54
Y	114	149	146	141	114	112	124	105	141	120	124

Требуется вычислить выборочный коэффициент корреляции и найти выборочное уравнение линейной регрессии Y на X.

Решение. Применяя формулы (8.59) и (8.34) с X вместо x, найдем, округляя до сотых:

$$\overline{X^2} = 3711{,}73, \quad \overline{Y^2} = 1681{,}33, \quad \overline{XY} = 7471{,}2,$$

$$\overline{X} = 59{,}91, \quad \overline{Y} = 126{,}36.$$

Отсюда, по (8.60) и (8.58)

$$\overline{r} = \frac{7471{,}2 - 59{,}91 \cdot 126{,}36}{\sqrt{(3711{,}73 - 59{,}91^2)(1681{,}33 - 126{,}36^2)}} = \\ = \frac{-99{,}05}{\sqrt{122{,}53 \cdot 214{,}33}} = \frac{-99{,}05}{162{,}06} = -0{,}61,$$

$$a_1 = \frac{-99{,}05}{122{,}53} = -0{,}81,$$

$$a_0 = 126{,}36 - (-0{,}81) \cdot 59{,}91 = 174{,}89.$$

Таким образом, выборочное уравнение прямой средней квадратической регрессии Y на X имеет вид $y = 174{,}98 - 0{,}81x$.

Если объем n выборки достаточно велик, то перед нахождением тех или иных статистических оценок по наблюдениям выборки ее обычно сводят в *корреляционную таблицу*:

Таблица 8.21

X	Y					
	x_1	...	x_i	...	x_m	n_y
y_1	n_{11}	...	n_{i1}	...	n_{m1}	$n_{\cdot 1}$
...	...	...	...	...	...	...
y_j	n_{1i}	...	n_{ij}	...	n_{mj}	$n_{\cdot j}$
...	...	...	...	...	...	...
y_k	n_{1k}	...	n_{ik}	...	n_{mk}	$n_{\cdot k}$
n_x	$n_{1\cdot}$	...	$n_{i\cdot}$	...	$n_{m\cdot}$	n

Здесь

$$n_{i\cdot} = \sum_{j=1}^{k} n_{ij}, \quad 1 \leqslant i \leqslant m, \text{ — сумма элементов } i\text{-го столбца},$$

$$n_{\cdot j} = \sum_{i=1}^{m} n_{ij}, \quad 1 \leqslant j \leqslant k, \text{ — сумма элементов } j\text{-й строки},$$

$$\sum_{i=1} n_{i\cdot} = \sum_{j=1}^{k} n_{\cdot j} = n, \tag{8.63}$$

$x_1, x_2, \ldots, x_m$ — середины интервалов группировки (см. п. 8.1.4) по X; $y_1, y_2, \ldots, y_k$ — середины интервалов группировки по Y, n_{ij} — число точек выборки, попавших в прямоугольник с центром (x_i, y_j).

Как правило, группировка осуществляется с равным шагом h_x по x и равным шагом h_y по y, т. е.

$$x_{i+1} - x_i = h_x, \quad 1 \leqslant i \leqslant m; \quad y_{j+1} - y_j = h_y, \quad 1 \leqslant j \leqslant k. \tag{8.64}$$

Аналоги формул (8.53) и (8.59), полученные по данным корреляционной таблицы, выглядят так:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{m} x_i n_{i\cdot}, \quad \overline{Y} = \frac{1}{n}\sum_{j=1}^{k} y_j n_{\cdot j}, \quad \overline{XY} = \frac{1}{n}\sum_{i=1}^{m} x_i \sum_{j=1}^{k} y_j n_{ij},$$
$$\overline{Y^2} = \frac{1}{n}\sum_{j=1}^{k} y_j^2 n_{\cdot j}, \quad \overline{X^2} = \frac{1}{n}\sum_{i=1}^{m} x_i^2 n_{i\cdot}. \tag{8.65}$$

Оценки дисперсий и ковариации рассчитываются по формулам:

$$S_x^2 = \overline{X^2} - \overline{X}^2, \quad S_y^2 = \overline{Y^2} - \overline{Y}^2,$$
$$l_{xy} = \overline{XY} - \overline{X}\,\overline{Y}. \tag{8.66}$$

Выборочный коэффициент корреляции $\overline{r}$, выборочные коэффициенты регрессии a_0, a_1 и выборочные уравнения прямой средней квадратической регрессии находятся по тем же формулам (8.56), (8.58), (8.61) и (8.62), что и ранее, с учетом обозначений (8.65) и (8.66). Напоминаем, что при вычислениях удобно пользоваться равенствами (8.60).

Представление наблюдений выборки в виде корреляционной таблицы позволяет определить выборочный аналог функции регрессии (см. (8.52)):

$$M\,(Y|X = x_i) = \overline{Y}(x_i) = \frac{1}{n_{i\cdot}}\sum_{j=1}^{k} y_j n_{ij}, \quad 1 \leqslant j \leqslant m. \qquad (8.67)$$

Графическое представление точек $(x_i, \overline{Y}(x_i))$ в случае, если стохастическая зависимость Y от X является регрессионной, позволяет проверить справедливость предположений о виде теоретической функции регрессии Y на X.

В самом деле, если выборочная линия регрессии имеет тенденцию к расположению точек вдоль прямой, то весьма вероятно, что теоретическая линия регрессии Y на X также является прямой (см. рис. 8.11), а если эти точки расположены вдоль, скажем, параболы, как на рис. 8.12, то теоретическая линия регрессии предположительно является параболой.

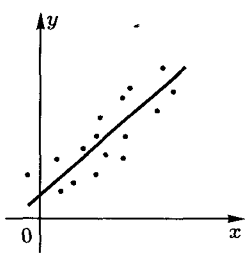


Рис. 8.11. Объяснения в тексте.

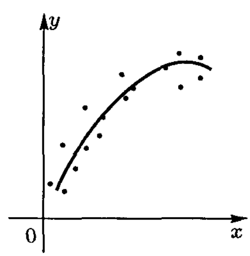


Рис. 8.12. Объяснения в тексте.

Замечание 1. Пусть имеются основания считать, что теоретическая регрессия задается функцией $Q_r(x)$ (см. (8.30)), зависящей от числовых параметров $\alpha_0, \alpha_1, \ldots, \alpha_r$, подлежащих определению. Поскольку случай линейной регрессии уже был рассмотрен, нас в первую очередь интересует *нелинейная регрессия*.

Можно показать, что наилучшими в смысле метода наименьших квадратов оценками неизвестных параметров регрессии α_i являются величины a_i, при которых функция (см. (8.63) и (8.67))

$$S(\alpha_0, \alpha_1, \ldots, \alpha_r) = \sum_{1 \leqslant i \leqslant m} n_{i\cdot}(\overline{Y}(x_i) - Q_r(x_i))^2$$

достигает минимума. Процедура, позволяющая найти такие оценки, описана в п. 8.1.3 и в случае полиномиальной регрессии (8.31) приводит к линейной системе из $r+1$ нормального уравнения, подробно исследованной при $r = 1$.

Замечание 2. Для упрощения вычислений в табл. 8.21 удобно от (x_i, y_j) перейти к новым переменным (u_i, v_j), положив

$$u_i = \frac{x_i - x_0}{h_1}, \quad v_j = \frac{y_j - y_0}{h_2}, \tag{8.68}$$

и выбирая числа x_0, y_0, h_1, h_2 таким образом, чтобы по формулам (8.65), (8.66), в которых x_i заменено на u_i, а y_j на v_j, было проще считать (при условии (8.64), например, удобно положить $h_1 = h_x$, $h_2 = h_y$). Обратный пересчет осуществляется по формулам

$$\begin{gathered}\overline{X} = h_1\overline{U} + x_0, \quad \overline{Y} = h_2\overline{V} + y_0, \quad S_x^2 = h_1^2 S_u^2,\\ S_y^2 = h_2^2 S_v^2, \quad l_{xy} = h_1 h_2 l_{uv}.\end{gathered} \tag{8.69}$$

Заметим, что при линейном преобразовании (8.68) значение выборочного коэффициента корреляции $\overline{r}$ не изменяется.

Аналогично, $M(Y|X = x_i)$ (см. (8.67)), перейдя к новым координатам, можно вычислить по формуле

$$M(Y|X = x_i) = y_0 + h_2\frac{1}{n_{i\cdot}}\sum_{j=1}^{k} v_j n_{ij}, \quad 1 \leqslant i \leqslant m. \tag{8.70}$$

Пример 4.6. Изучалась зависимость между количеством гемоглобина в крови Y (%) и массой животных X (кг). Результаты наблюдений приведены в виде корреляционной таблицы (пропуски означают нули):

Таблица 8.22

Y	X			
	18	22	26	30
70	5			
75	7	46	1	
80		29	72	
85			29	8
90				3

Требуется определить выборочные аналоги функции регрессии и уравнения прямой средней квадратической регрессии Y на X.

Решение. Для упрощения вычислений перейдем к новым переменным U и V, воспользовавшись формулами (8.68) при $h_1 = 4$, $h_2 = 5$, $x_0 = 26$, $y_0 = 80$. Для удобства перепишем табл. 8.22 в новых обозначениях, добавив справа столбец с суммой частот по строкам, а снизу строку с суммой частот по столбцам:

Таблица 8.23

U	V				
	-2	-1	0	1	n_v
-2	5				5
-1	7	46	1		54
0		29	72		101
1			29	8	37
2				3	3
n_u	12	75	102	11	$n = 200$

Имеем (см. (8.65), (8.66) при $x_i = u_i$ и $y_j = v_j$)

$$\overline{U} = \frac{1}{200}(-2 \cdot 12 + (-1) \cdot 75 + 0 \cdot 102 + 1 \cdot 11) = -0{,}44,$$

$$\overline{V} = \frac{1}{200}(-2 \cdot 5 + (-1) \cdot 54 + 0 \cdot 101 + 1 \cdot 37 + 2 \cdot 3) = -0{,}105,$$

$$\overline{U^2} = \frac{1}{200}((-2)^2 \cdot 12 + (-1)^2 \cdot 75 + 0^2 \cdot 102 + 1^2 \cdot 11) = 0{,}67,$$

$$\overline{UV} = \frac{1}{200}((-2)((-2)\cdot 5 + (-1)\cdot 7) + (-1)((-1)\cdot 46 + \\ + 0\cdot 29) + 0\cdot((-1)\cdot 1 + 1\cdot 29) + 1\cdot(1\cdot 8 + 2\cdot 3)) = 0{,}47,$$

$$S_U^2 = 0{,}67 - (-0{,}44)^2 = 0{,}4764, \quad S_U = 0{,}6902,$$

$$S_V^2 = 0{,}615 - (-0{,}105)^2 = 0{,}604, \quad S_V = 0{,}7772,$$

$$l_{UV} = 0{,}47 - (-0{,}44)(-0{,}105) = 0{,}426,$$

$$\overline{r}(U,V) = \frac{0{,}426}{0{,}6902\cdot 0{,}7772} = 0{,}7941.$$

Таким образом (см. (8.69)),

$$\overline{X} = 4\cdot(-0{,}44) + 26 = 24{,}24, \quad \bar{Y} = 5\cdot(-0{,}105) + 80 = 79{,}475,$$

$$S_X = 4\cdot 0{,}6902, \quad S_Y = 5\cdot 0{,}7772 = 3{,}886, \quad \overline{r} = \overline{r}(U,V) = 0{,}7941$$

выборочное уравнение прямой средней квадратической регрессии Y на X выражается формулой (8.61):

$$y_{\text{лин}} = 79{,}475 + 0{,}7941\frac{3{,}886}{2{,}7608}(x - 24{,}24) = 52{,}381 + 1{,}118x, \quad (8.71)$$

откуда

Таблица 8.24

X	18	22	26	30
$y_{\text{лин}}$	72,5	76,98	81,45	85,92

Формула (8.70) дает

$$M(Y|X=18) = 80 + 5\cdot\frac{1}{12}((-2)\cdot 5 + (-1)\cdot 7) = 72{,}92,$$

$$M(Y|X=22) = 80 + 5\cdot\frac{1}{75}((-1)\cdot 46 + 0\cdot 29) = 77{,}91,$$

$$M(Y|X=26) = 80 + 5\cdot\frac{1}{102}((-1)\cdot 1 + 0\cdot 72 + 1\cdot 29) = 81{,}37,$$

$$M(Y|X=30) = 80 + 5\cdot\frac{1}{11}(1\cdot 8 + 2\cdot 3) = 86{,}36.$$

Сопоставляя полученные результаты, приходим к выводу, что значения, вычисленные по уравнению выборочной регрессии и по линейной зависимости (8.71), хорошо согласуются.

В заключение коснемся случая, когда наблюдаемых признаков генеральной совокупности больше двух. Можно, например, рассматривать три взаимосвязанные случайные переменные (X_1, X_2, X_3) и исследовать стохастическую зависимость X_3 от переменных X_1 и X_2. Так, минимизация по β_0, β_1 и β_2 дисперсии $D(X_3-\beta_0-\beta_1X_1-\beta_2X_2)$, позволяет найти уравнение *плоскости средней квадратической регрессии* X_3 на X_1 и X_2:

$$x_3 = m_3 + b_1(x_1 - m_1) + b_2(x_2 - m_2), \tag{8.72}$$

наилучшим образом приближающее случайную величину X_3 линейной комбинацией случайных величин X_1 и X_2 (см. (8.48) и (8.49)). Здесь $m_i = M(X_i)$, b_i определяются формулами

$$b_1 = \frac{r_{13} - r_{12}r_{23}}{1 - r_{12}^2}\frac{\sigma_3}{\sigma_1}, \qquad b_2 = \frac{r_{23} - r_{12}r_{13}}{1 - r_{12}^2}\frac{\sigma_3}{\sigma_2}, \tag{8.73}$$

в которых $r_{ij} = r\,(X_i, X_j)$, $i \neq j$, — коэффициенты корреляции X_i и X_j (см. (8.46)), $\sigma_i = \sqrt{D(X_i)}$ — стандартные отклонения случайных величин X_i.

Для оценки тесноты линейной зависимости между случайной величиной X_3 и парой случайных величин X_1, X_2 вводится коэффициент множественной корреляции, являющийся обобщением понятия коэффициента корреляции (8.46).

Величину R можно вычислить по формуле

$$R = \sqrt{\frac{r_{13}^2 + r_{23}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{12}^2}}.$$

Отметим, что всегда $R \geqslant |r_{i3}|$, $i = 1, 2$, и если $R = 1$, то между X_3 и X_1, X_2 имеется линейная функциональная связь, а именно (см. (8.72))

$$X_3 = m_3 + b_1(X_1 - m_1) + b_2(X_2 - m_2)$$

с вероятностью единица.

Подобным же образом обобщается понятие регрессии. Регрессия X_3 на X_1, X_2, например, представляет собой функцию

$g(x_1, x_2) = M\left(X_3 | X_1 = x_1,\ X_2 = x_2\right)$, т. е. является условным математическим ожиданием случайной величины X_3 при условии, что случайные величины X_1, X_2 приняли значения x_1 и x_2 соответственно. Также как в случае двух случайных переменных, функция $g(X_1, X_2)$ наилучшим образом приближает случайную величину X_3 в смысле среднего квадратического, и если функция регрессии $g(x_1, x_2)$ является линейной по x_1 и x_2, то уравнения $x_3 = g(x_1, x_2)$ и (8.72) совпадают.

Выборочные аналоги уравнения (8.72) и коэффициента множественной корреляции R вычисляются путем замены соответствующих характеристик в формулах (8.73) и для R их оценками, полученными по выборке (X_{1i}, X_{2i}, X_{3i}), $1 \leqslant i \leqslant n$, состоящей из n независимых наборов случайных величин (X_1, X_2, X_3). Таким образом (см. формулы (8.43)–(8.46)), m_k заменяется на $\overline{X}_k = \dfrac{X_{k1} + \ldots + X_{kn}}{n}$, $k = 1, 2, 3$; вместо r_{ik} подставляется $\overline{r}_{ik}$ — выборочный коэффициент корреляции между случайными величинами X_i и X_k, и т. п.

Самостоятельная работа

1. В 100 частях воды растворяется следующее число условных частей азотнокислого натрия $NaNO_3$ (признак Y) при соответствующих температурах (X):

X	0	4	10	15	21	29	36	51	68
Y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

На количество растворившегося $NaNO_3$ влияют случайные факторы. Предполагается наличие стохастической линейной зависимости между температурой и количеством растворившегося $NaNO_3$ вида (8.32). Найти МНК-оценку коэффициентов линейной модели.

2. В нижеследующих задачах требуется вычислить выборочный коэффициент корреляции между переменными Y и X и найти выборочное уравнение прямой средней квадратической регрессии Y на X.

№ 1. Изучалась зависимость между содержанием коллагена Y и эластина X в магистральных артериях головы (г/100 г сухого вещества) (возраст 51–75 лет). Результаты наблюдений приведены в таблице в виде двумерной выборки объема 5:

X	13,50	13,09	6,45	7,26	8,80
Y	33,97	38,07	53,98	46,00	48,61

№ 2. Изучалась зависимость между массой новорожденных павианов-гамадрилов X (кг) и массой их матерей Y (кг). Результаты наблюдений приведены в таблице в виде двумерной выборки объема 9:

X	0,7	0,73	0,75	0,7	0,65	0,7	0,61	0,70	0,63
Y	10	10,8	11,3	10	11,1	11,3	10,2	13,5	12

№ 3. Изучалась зависимость между объемом грудной клетки мужчин Y (см) и ростом X (см). Результаты наблюдений приведены в таблице в виде двумерной выборки объема 7:

X	162	164	179	172	182	188	168
Y	88	94	98	100	102	108	112

№ 4. Изучалась зависимость между минутным объемом сердца Y (л/мин) и средним давлением в левом предсердии X (мм рт. ст.). Результаты наблюдений приведены в таблице в виде двумерной выборки объема 5:

X	4,8	6,4	9,3	11,2	17,7
Y	0,4	0,69	1,29	1,64	2,4

3. По закону распределения двумерного случайного вектора (X, Y)

Y	X		
	−2	3	4
0	0,1	0,2	0,2
2	0,2	0,2	0,1

найти математические ожидания и стандартные отклонения X и Y, корреляционный момент и коэффициент корреляции X и Y, уравнение прямой средней квадратической регрессии Y на X и вычислить значения регрессии Y на X.

4. По таблицам сгруппированных данных, предполагая, что координаты X и Y имеют совместное нормальное распределение:

1) вычислить выборочный коэффициент корреляции X и Y;

2) написать уравнение линейной регрессии Y на X;

3) вычислить значения в точках $x_1, x_2, \ldots, x_m$ выборочной функции регрессии Y на X;

4) сравнить между собой приближения Y, полученные по функции регрессии и уравнению линейной регрессии.

№ 1

Y	X						
	2	10	18	26	34	42	50
45	1	1					
50		1	1	5	4		
55			5	3	4		
60			10	3	3	2	
65			2		1		1
70	1	1		1			

№ 2

Y	X						
	22	32	42	52	62	72	82
15	1	1	2				
19	2		3				
23		4	2	10			
27		2		3	7	2	
31					5	4	
35					1		1

№ 3

Y	X						
	4	8	12	16	20	24	28
115			2		2		
120	1		2	2			1
125				13		4	
130	2	2	4	2	1		1
135		2	1		5	4	
140					1	2	

№ 4

Y	X						
	2	5	8	11	14	17	20
115	2		2				
130		2		2		2	
145			3	12	3		1
160		2		4	5		
175		1	4		3		
190	1						1

§ 8.5. Проверка статистических гипотез

8.5.1. Выбор из двух гипотез. Введение

В прикладных задачах часто требуется по наблюдениям выборки высказать некоторое суждение (*гипотезу*) относительно интересующих экспериментатора характеристик генеральной совокупности, из которой эта выборка извлечена.

В таких случаях говорят, что речь идет о *проверке статистических гипотез.*

Правила, согласно которым выясняется, соответствует или нет интересующая нас гипотеза опытным данным, называются *статистическими критериями* или просто *критериями.*

Следует сказать, что статистические критерии — это, пожалуй, наиболее широко применяемые статистические средства. Они дают возможность экспериментатору найти разумный ответ на вопрос, подобный следующему.

В двух однородных группах больных гриппом, одной из которых проведена вакцинация средством A, а другой — средством B, среднее время выздоровления неодинаково. Указывает ли это обстоятельство на то, что одно противогриппозное средство по эффективности превосходит другое или же выявленное различие случайно?

Итак, пусть случайная величина X является интересующей нас характеристикой генеральной совокупности и мы хотим по наблюдениям выборки подтвердить или опровергнуть некоторую гипотезу, скажем о виде функции распределения этой случайной величины. Такая гипотеза носит название *нулевой гипотезы* и обозначается H_0.

Наряду с гипотезой H_0 необходимо также рассмотреть *конкурирующую гипотезу* H_1, которая является альтернативной по отношению к H_0, т. е. принимается в том случае, если H_0 не верна.

Например, если $X \sim N(\mu, 1)$, где параметр μ предполагается неизвестным, то в качестве нулевой гипотезы можно выбрать H_0: $\mu = 0$, а в качестве конкурирующей гипотезы рассмотреть H_1: $\mu > 0$ или H_1: $\mu \neq 0$.

Пусть нам необходимо проверить справедливость гипотезы H_0 при альтернативе H_1 (предполагая при этом, что либо H_0, либо H_1 имеет место). Поскольку выборочные наблюдения, на основе которых вопрос решается, являются случайными величинами, то абсолютно достоверно этого сделать нельзя. Всегда остается риск отвергнуть истинную гипотезу H_0, тем самым совершив так называемую *ошибку первого рода*, или же принять ложную гипотезу H_0, сделав *ошибку второго рода*. Таким образом, любой критерий будет в той или иной мере субъективным, и нам следует по возможности минимизировать меру этой субъективности в принятии решения. Поступим следующим образом.

Зафиксируем некоторое малое положительное число α (например, $\alpha = 0{,}01$), называемое *уровнем значимости*.

Выберем некоторую функцию Λ, зависящую от выборки, которую будем называть *статистикой критерия*.

Среди всех возможных значений, принимаемых Λ, определим множество K_α (зависящее от вида статистики критерия Λ и от уровня значимости α), для которого вероятность события $\{\Lambda \in \in K_\alpha\}$, если верна гипотеза H_0, равна α, или

$$P(\Lambda \in K_\alpha | H_0) = \alpha. \qquad (8.74)$$

Статистический критерий состоит в следующем: если статистика критерия Λ, подсчитанная по выборке, попадает в множество $K_\alpha (\Lambda \in K_\alpha)$, то гипотеза H_0 отвергается с вероятностью ошибки первого рода, равной α (говорят, что H_0 отклоняется на уровне α), в противном случае (т. е. когда $\Lambda \notin K_\alpha$), она принимается. В силу этого, K_α называется *критической областью* размера α.

Сконструированный таким способом критерий отвергает истинную гипотезу H_0 лишь в $100\,\alpha$ случаях из 100 (при уровне значимости $\alpha = 0{,}01$, например, лишь в одном случае из ста). Другими словами, попадание статистики критерия в критическую область дает нам веские основания считать, что нулевая гипотеза H_0 не является истинной и справедлива конкурирующая гипотеза H_1.

Соотношение, формально определяющее вероятность β ошибки второго рода (т. е. вероятность принять H_0, если верна гипотеза H_1), можно переписать в виде условия

$$P(\Lambda \notin K_\alpha | H_1) = \beta. \qquad (8.75)$$

Заметим, что (8.75) равносильно равенству

$$P(\Lambda \in K_\alpha | H_1) = 1 - \beta = W. \qquad (8.76)$$

Величина W называется *мощностью критерия*.

Таким образом, при заданной вероятности ошибки первого рода среди нескольких критериев лучшим является тот, у которого вероятность ошибки второго рода меньше (или мощность больше).

Замечание. Согласно требованиям фармакопеи в биологических исследованиях принимается $\alpha = 0{,}05$, а при разработке биологических стандартов — $\alpha = 0{,}01$.

8.5.2. Проверка значимости коэффициента корреляции

Предположим, что интересующие нас характеристики X и Y объектов генеральной совокупности имеют двумерное нормальное распределение с неизвестным коэффициентом корреляции ρ. Требуется по наблюдениям выборки объема n, извлеченной из этой совокупности, проверить нулевую гипотезу H_0: $\rho = 0$ при конкурирующей гипотезе H_1: $\rho \neq 0$, т. е. выяснить будут ли случайные величины X и Y некоррелированными (и, следовательно, независимыми) или нет.

Обозначив через r выборочный коэффициент корреляции (см. 8.56), возьмем в качестве статистики критерия функцию $\Lambda = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ и положим $K_\alpha = \{x : |x| > t_{\text{кр}}(\alpha, n-2)\}$, где $t_{\text{кр}}(\alpha, k)$ — критическая точка распределения Стьюдента, отвечающая уровню значимости α, k степеням свободы и двусторонней критической области (см. приложение 2).

Известно, что при нулевой гипотезе случайная величина Λ имеет распределение Стьюдента с $n-2$ степенями свободы, и условие (8.74) выполняется. Можно также показать, что вероятность ошибки второго рода $\beta = P(|\Lambda| \leqslant t_{\text{кр}}(\alpha, n-2) | \rho \neq 0)$ стремится к нулю при $n \to \infty$.

Статистический критерий формулируется следующим образом.

По извлеченным из генеральной совокупности наблюдениям выборки объема n следует вычислить выборочный коэффициент корреляции $r_{\text{набл}}$ и соответствующее ему наблюдаемое значение статистики критерия $\Lambda_{\text{набл}} = \dfrac{r_{\text{набл}}\sqrt{n-2}}{\sqrt{1-r^2_{\text{набл}}}}$. Если $|\Lambda_{\text{в}}| > t_{\text{кр}}(\alpha, n-$ $-2)$, то нулевая гипотеза отклоняется на уровне α, в противном случае, при $|\Lambda_{\text{в}}| \leqslant t_{\text{кр}}(\alpha, n-2)$, она принимается.

Пример 5.1. По данным примера 4.5, предполагая, что пара случайных величин X и Y имеет двумерное нормальное распределение, при уровне значимости 0,01, проверить нулевую гипотезу о равенстве нулю коэффициента корреляции ρ между X и Y при конкурирующей гипотезе H_1: $\rho \neq 0$.

Решение. Найдем наблюдаемое значение статистики критерия $\Lambda_{\text{набл}} = -0{,}61\dfrac{\sqrt{11-2}}{\sqrt{1-0{,}61^2}} = -2{,}31$. В таблице критических точек распределения Стьюдента (приложение 2) по уровню значимости 0,01 и числу степеней свободы $11-2=9$ находим критическую точку двусторонней критической области $t_{\text{кр}}(0{,}01,\ 9) = 3{,}25$. Поскольку $|\Lambda_{\text{набл}}| = 2{,}31 < 3{,}25$, оснований отвергать нулевую гипотезу нет. Таким образом, выборочные данные не противоречат предположению о некоррелированности случайных величин X и Y.

8.5.3. Сравнение средних двух нормально распределенных генеральных совокупностей с неизвестными одинаковыми дисперсиями

Пусть $X_1, X_2, \ldots, X_m$ и $Y_1, Y_2, \ldots, Y_n$ — независимые выборки объема m и n, извлеченные из генеральных совокупностей, нормально распределенных с неизвестными параметрами (μ_x, σ) и (μ_y, σ) соответственно. Обращаем внимание на то, что неизвестные генеральные дисперсии по предположению равны. Проверяется нулевая гипотеза о равенстве генеральных средних μ_x и μ_y, т. е. H_0: $\mu_x = \mu_y$.

Сделаем следующее. Вычислим средние выборок $\overline{X}$ и $\overline{Y}$, а также их исправленные выборочные дисперсии S_x^2 и S_y^2 (см. формулы (8.7), (8.8)). Затем образуем результирующую оценку общей дисперсии

$$S^2 = \frac{(m-1)S_x^2 + (n-1)S_y^2}{m+n-2} \tag{8.77}$$

и статистику критерия

$$\Lambda = \frac{\bar{X} - \bar{Y}}{S\sqrt{\dfrac{1}{m} + \dfrac{1}{n}}}. \tag{8.78}$$

При нулевой гипотезе статистика Λ имеет распределение Стьюдента с $m+n-2$ степенями свободы, что предопределяет дальнейший выбор критических областей K_α при различных конкурирующих гипотезах H_1.

Сначала приведем статистический критерий проверки нулевой гипотезы H_0: $\mu_x = \mu_y$ при конкурирующей гипотезе H_1: $\mu_x \neq \mu_y$.

По формулам (8.77), (8.78) следует вычислить наблюдаемое значение статистики критерия $\Lambda_{\text{набл}}$.

Если $|\Lambda_{\text{набл}}| > t_{\text{кр}}(\alpha, m+n-2)$, то нулевая гипотеза отклоняется на уровне α; если же $|\Lambda_{\text{набл}}| \leqslant t_{\text{кр}}(\alpha, m+n-2)$ — нулевая гипотеза принимается.

Здесь величина $t_{\text{кр}}(\alpha, k)$ так же, как в п. 8.4.2, определяется по таблице критических точек распределения Стьюдента (уровень значимости α, число степеней свободы k, двусторонняя область) из приложения 2.

Пример 5.2. Две группы детей, одинаковых по оценке умственных способностей, независимо обучались по двум различным методикам преподавания. Затем их подвергли выборочному тестированию, давшему следующие результаты:

1) объем выборки из первой группы равен 20, $\overline{X} = 29{,}233$, $S_x^2 = 5{,}62$;

2) объем выборки из второй группы равен 10, $\overline{Y} = 27{,}562$, $S_y^2 = 2{,}19$.

В предположении, что изучаемые показатели в каждой группе имеют нормальное распределение с неизвестными средними и неизвестными, но одинаковыми дисперсиями, проверить при уровне значимости 0,05, существенно ли отличаются средние показания групп?

Решение. Проверяем H_0: $\mu_x = \mu_y$ при двусторонней альтернативе H_1: $\mu_x \neq \mu_y$. Наблюдаемые значения статистики критерия

$$\Lambda_{\text{набл}} = \frac{29{,}233 - 27{,}562}{\sqrt{\dfrac{(19 \cdot 5{,}62 + 9 \cdot 2{,}19)}{28} \cdot \left(\dfrac{1}{20} + \dfrac{1}{10}\right)}} = 2{,}03,$$

а критическая точка $t_{\text{кр}}(0{,}05,\ 28)$ распределения Стьюдента, соответствующая двусторонней области (см. приложение 3, верхняя строка) равна 2,05.

Поскольку $|\Lambda_{\text{набл}}| = 2{,}03 < t_{\text{кр}}(0{,}05,\ 28) = 2{,}05$, у нас нет оснований отвергать нулевую гипотезу о равенстве средних в группах.

Далее сформулируем критерий проверки нулевой гипотезы H_0: $\mu_x = \mu_y$ при альтернативах H_1: $\mu_x > \mu_y$ или H_1: $\mu_x < \mu_y$.

В случае проверки нулевой гипотезы H_0: $\mu_x = \mu_y$:

1) при конкурирующей гипотезе H_1: $\mu_x > \mu_y$ нулевая гипотеза на уровне α отвергается, если

$$\Lambda_{\text{набл}} > t_{\text{одн.кр}}(\alpha, m+n-2); \tag{8.79}$$

2) при конкурирующей гипотезе H_1: $\mu_x < \mu_y$ нулевая гипотеза отвергается на уровне α, если

$$\Lambda_{\text{набл}} < -t_{\text{одн.кр}}(\alpha, m+n-2).$$

Здесь $t_{\text{одн.кр}}(\alpha, k)$ — критическая точка распределения Стьюдента с уровнем значимости α, числом степеней свободы k, отвечающая правосторонней критической области (см. приложение 2, нижняя строка). Заметим, что критические точки $t_{\text{одн.кр}}(\alpha, k)$ и $t_{\text{кр}}(\alpha, k)$ связаны соотношением

$$t_{\text{одн.кр}}(\alpha, k) = t_{\text{кр}}(2\alpha, k), \quad 0 < \alpha < 1/2.$$

Например, $t_{\text{одн.кр}}(0{,}05,\ 28) = 1{,}7 = t_{\text{кр}}(0{,}1,\ 28)$.

Пример 5.3. По данным примера 5.2 проверить при уровне значимости 0,05 нулевую гипотезу о равенстве средних H_0: $\mu_x = \mu_y$ при конкурирующей гипотезе H_1: $\mu_x > \mu_y$.

Решение. Так как наблюдаемое значение статистики $\Lambda_{\text{набл}} = 2{,}03$, а соответствующее значение «односторонней» критической точки из приложения 3 $t_{\text{одн.кр}}(0{,}05,\ 28) = 1{,}7$, находим $\Lambda_{\text{набл}} > t_{\text{одн.кр}}(0{,}05,\ 28)$. Таким образом (см. (8.79)), статистика критерия попадает в критическую область и, следовательно, нулевая гипотеза отвергается в пользу предположения $\mu_x > \mu_y$. Другими словами, выборочные данные не подтверждаются нулевой гипотезой, и мы, в отличие от вывода примера 5.2, вынуждены заключить, что средние показания первой группы существенно превышают средние показания второй группы.

Отметим, что сравнение примеров 5.2 и 5.3 наглядно демонстрирует важность выбора конкурирующей гипотезы при проверке статистических гипотез.

8.5.4. Критерий равенства двух дисперсий

Пусть имеются две независимых выборки, первая — объема m из наблюдений случайной величины $X \sim N(\mu_1, \sigma_1)$, и вторая — объема n из наблюдений случайной величины $Y \sim$ $\sim N(\mu_2, \sigma_2)$. Требуется проверить, согласуются ли выборочные данные с нулевой гипотезой H_0: $\sigma_1^2 = \sigma_2^2$.

Заметим, что $\sigma_1^2 = D(X)$, а $\sigma_2^2 = D(Y)$, и, следовательно, нулевая гипотеза предполагает равенство генеральных дисперсий $D(X)$ и $D(Y)$, т.е. H_0: $D(X) = D(Y)$.

По выборкам, отвечающим случайным величинам X и Y, вычислим исправленные выборочные дисперсии S_x^2 и S_y^2 и составим статистику критерия

$$\Lambda = \frac{S_x^2}{S_y^2}, \tag{8.80}$$

считая без потери общности, что

$$S_x^2 \geqslant S_y^2 \tag{8.81}$$

(в противном случае следует поменять обозначения X и Y местами).

Можно доказать, что при нулевой гипотезе статистика критерия Λ является случайной величиной, имеющей *распределение Фишера-Снедекора* (F — распределение) с $m-1$ и $n-1$ степенями свободы (напоминаем, что m — объем выборки, соответствующей числителю S_x^2 дроби Λ, в то время как n — объем выборки, определяющей знаменатель S_y^2 этой дроби).

Для проверки на уровне значимости α нулевой гипотезы H_0: $D(X) = D(Y)$ при конкурирующей гипотезе H_1: $D(X) > D(Y)$ следует по выборочным данным с учетом договоренности (8.81) вычислить наблюдаемое значение статистики критерия $\Lambda_{\text{набл}}$ и сравнить его с величиной $f_{\text{кр}}(\alpha; m-1, n-1)$, которая является критической точкой F-распределения, отвечающей уровню значимости α и степеням свободы $k_1 = m-1$ и $k_2 = n-1$. Значения

$f_{\text{кр}}(\alpha; k_1, k_2)$ при $\alpha = 0{,}01$, $0{,}05$ и некоторых k_1, k_2 содержатся в таблицах из приложения 3.

Если

$$\Lambda_{\text{набл}} \leqslant f_{\text{кр}}(\alpha; m-1, n-1), \tag{8.82}$$

то нулевая гипотеза принимается;
если же

$$\Lambda_{\text{набл}} > f_{\text{кр}}(\alpha; m-1, n-1),$$

то нулевая гипотеза отклоняется на уровне α (и, следовательно, принимается конкурирующая гипотеза H_1: $D(X) > D(Y)$).

В случае конкурирующей гипотезы H_1: $D(X) \neq D(Y)$ нулевая гипотеза принимается, если

$$\Lambda_{\text{набл}} \leqslant f_{\text{кр}}(\alpha/2; m-1, n-1), \tag{8.83}$$

и отвергается на уровне α в противоположном случае.

Заметим, что вышеизложенная процедура проверки равенства дисперсий двух генеральных совокупностей чувствительна к нарушениям предположения о нормальности.

Пример 5.4. По данным примера 5.3 проверить предположение о равенстве дисперсий в тестируемых группах при H_1: $\sigma_1 \neq \sigma_2$ и уровне значимости $\alpha = 0{,}05$.

Решение. Имеем: $\Lambda_{\text{набл}} = \dfrac{5{,}62}{2{,}19} = 2{,}57$. Значения $f_{\text{кр}}(0{,}025; 19{,}9)$ в таблицах из приложения 3 нет, однако в силу того что величина $f_{\text{кр}}(\alpha; k_1, k_2)$ убывает с ростом α, $f_{\text{кр}}(0{,}025; 19, 9) > f_{\text{кр}}(0{,}05; 19, 9) > f_{\text{кр}}(0{,}05; \infty, 9) = 2{,}71$. Поскольку $\Lambda_{\text{набл}} = 2{,}57 < 2{,}71$, условие (8.83) заведомо выполняется, откуда следует, что оснований отвергать предположение о равенстве дисперсий в тестируемых группах у нас нет.

8.5.5. Проверка непараметрических статистических гипотез. Критерий знаков

Выводы, которые мы делали в предыдущих параграфах, существенно используют предположение о нормальности наблюдаемых величин, или, в широком смысле, предположение о принадлежности распределения наблюдений некоторому семейству распределений, определяемых с точностью до одного или нескольких параметров. Следовательно, ценность любых наших заключений зависит от адекватности такого предположения. Например, статистические критерии, такие как критерий равенства дисперсий, строго говоря, пригодны только в том случае, когда предположение о нормальности верно, хотя многие из них довольно устойчивы к отклонениям от нормальности.

Было бы полезно уметь строить статистические модели и критерии, не зависящие от априорных вероятностных предположений о, скажем, нормальности наблюдаемых характеристик генеральной совокупности. Такие модели и критерии называются *непараметрическими* (напоминаем, что предположение типа $X \sim N(m, \sigma)$ относится к параметрическим, поскольку семейство нормальных распределений полностью описывается двумя параметрами).

Ниже будет рассмотрен один из наиболее простых непараметрических критериев, так называемый *критерий знаков.*

Определение. *Медианой* случайной величины X с непрерывной функцией распределения $F(x)$ называется такое значение m этой случайной величины, для которого $F(m) = 0{,}5$.

Таким образом, $P(X < m) = P(X > m) = 0{,}5$.

Заметим, что медиана равномерно или нормально распределенной случайной величины совпадает с ее математическим ожиданием.

В целом, в непараметрической статистике математическое ожидание во многом теряет свою привлекательность (оно может просто не существовать, как у распределения Коши например), и на первый план выходит именно медиана.

Вернемся к критерию знаков. Пусть случайная выборка X_1, $X_2, \ldots, X_n$ состоит из наблюдаемых значений случайной величины X с неизвестной непрерывной функцией распределения $F(x)$ и мы хотим на основе этой выборки проверить нулевую гипотезу H_0, что медиана m равна определенному значению m_0 (H_0: $m = m_0$), против конкурирующей (двусторонней) гипотезы H_1: $m \neq m_0$.

В критерии знаков роль статистики критерия Λ выполняет число вариант, больших m_0. При этом, если какие-нибудь варианты в точности равны m_0, они отбрасываются с соответствующим уменьшением объема выборки n. Критическая область K_α

(размера α) при двусторонней альтернативе имеет вид

$$K_\alpha = (0, 1, \ldots, s_\alpha, n - s_\alpha, n - s_\alpha + 1, \ldots, n), \tag{8.84}$$

где s_α — наибольшее целое s, такое что

$$P_s = \left(\frac{1}{2}\right)^n \sum_{k=0}^{s} C_n^k \leqslant \frac{1}{2}\alpha \tag{8.85}$$

(таким образом, $P_{s_\alpha} \leqslant \alpha/2 < P_{s_\alpha+1}$).

Можно показать, что (см. (8.74)) $P(\Lambda \in K_\alpha | H_0) \leqslant \alpha$, т. е. уровень значимости критерия не превосходит α.

При небольших значениях n для нахождения s_α пользуются непосредственно формулой (8.85), а при $n \geqslant 10$ применяют приближенное равенство, использующее нормальную аппроксимацию

$$s_\alpha \cong \left[\frac{n - 1 - z_\alpha\sqrt{n}}{2}\right], \tag{8.86}$$

в котором $[x]$ — целая часть x, а $z_\alpha = t_{\text{кр}}(\alpha, \infty)$ и может быть найдена по таблице критических точек распределения Стьюдента (см. приложение 2) или же как решение уравнения $\Phi(z_\alpha) = \frac{1-\alpha}{2}$ (см. приложение 1).

Для проверки гипотезы H_0 при альтернативе H_1 следует по выборочным данным вычислить число наблюдений $\Lambda_{\text{набл}} > s_0$. В том случае, когда $\Lambda_{\text{набл}} \in K_\alpha$ (см. (8.84)), т. е. $\Lambda_{\text{набл}} \leqslant s_\alpha$ или $\Lambda_{\text{набл}} \geqslant n - s_\alpha$, нулевая гипотеза отвергается на уровне $\leqslant \alpha$. В противном случае, т. е. когда

$$s_\alpha < \Lambda_{\text{набл}} < n - s_\alpha, \tag{8.87}$$

она принимается.

В тех случаях, когда H_1 утверждает, что $m > m_0$ (или $m < m_0$), критическая область размера α имеет вид $(n - s_{2\alpha}, n - s_{2\alpha} + 1, \ldots, n)$ (или $(0,\ 1, \ldots, s_{2\alpha})$), где (см. (8.85)) $s_{2\alpha}$ — наибольшее целое s, такое что $P_s < \alpha$ (см. (8.85)).

На практике критерий знаков применяется в тех случаях, когда по результатам эксперимента мы в качестве наблюдений получаем независимые пары взаимосвязанных измерений $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$. Для таких данных обычно переходят к рассмотрению разностей $X_i - Y_j$, $1 \leqslant i \leqslant n$, к которым и применяют критерий знаков. Надо только учесть, что в этой ситуации при гипотезе H_0 величина $m_0 = 0$.

Пример 5.5. Пусть выборка 1,4, 0,9, −0,1, −0,2, 1, −0,3, 2, 0,5, 2,1, 0,6 из непрерывного распределения содержит 10 наблюдений, выражающих различие в длительности действия лекарств, полученных двумя разными методами (если, например, длительности действия лекарств были равны 89 и 88, то соответствующее различие $89 - 88 = 1$).

Требуется при уровне значимости $\alpha = 0{,}05$ проверить нулевую гипотезу H_0: $m = 0$ при конкурирующей гипотезе H_1: $m \neq 0$.

Решение. Подсчет положительных вариант дает $\Lambda_{\text{набл}} = 7$. По формуле (8.85) найдем $P_1 = (1/2)^{10}(1 + 10) = 0{,}011$, $P_2 = P_1 + (1/2)^{10} 45 = 0{,}055$, откуда следует, что $P_1 < 1/2 \cdot 0{,}5 = 0{,}25 < P_2$ и поэтому $s_{0,05} = 1$. Таким образом, условие (8.87) выполняется ($1 < 7 < 9$). Следовательно, на уровне значимости 0,05 оснований отвергнуть нулевую гипотезу нет.

Отметим, что вычисление числа $s_{0,05}$ по формуле (8.86) также приводит к равенству $s_{0,05} = 1$. В самом деле, $z_{0,95} = 1{,}98$ (см. приложение 2 или 1) и $s_{0,05} = [4{,}5 - 1/2 \cdot 1{,}96\sqrt{10}] = [1{,}401] = 1$.

Пример 5.6. По данным примера 5.5 проверить нулевую гипотезу H_0: $m = 0$ против альтернативы H_1: $m > 0$, уровень значимости 0,05.

Решение. Вычисления, проделанные в примере 5.5, показывают, что $P_1 < 0{,}5 < P_2$, т. е. $s_{0,1} = 1$. Поскольку односторонняя критическая область $[9, 10]$ не включает точку $\Lambda_{\text{набл}} = 7$, делаем вывод, что нулевая гипотеза принимается.

Заметим, что формула (8.86), $\alpha = 0{,}1$, $z_\alpha = 1{,}645$, также дает верный результат: $s_{0,1} = \left[4{,}5 - \frac{1}{2} \cdot 1{,}645\sqrt{10}\right] = [1{,}899] = 1$.

8.5.6. Критерий согласия Пирсона (хи-квадрат критерий)

Пусть интересующий нас признак генеральной совокупности является случайной величиной X и мы хотим по наблюдениям выборки из генеральной совокупности проверить некоторую гипотезу о функции распределения признака X, ответив на вопрос, совместима ли эта гипотеза с выборочными данными. В такой ситуации говорят о проверке *статистической гипотезы согласия.*

Далее будет изложена одна довольно универсальная процедура, позволяющая по статистическому интервальному ряду распределения выборки (см. табл. 8.4) проверить гипотезу о том, что функция распределения признака X генеральной совокупности совпадает с некоторой наперед данной функцией распределения $F(x)$. Эта процедура называется *критерием согласия Пирсона* или *критерием хи-квадрат* (критерий χ^2).

Исходные данные для применения критерия χ^2 представлены в табл. 8.6, содержащей информацию о том, сколько элементов n_i выборки объема n из наблюдений над случайной величиной X попадает в интервалы группировки $[h_{i-1}, h_i)$, где $i = 1, 2, \ldots$ $\ldots, k$, причем $n_1 + n_2 + \ldots + n_k = n$.

Далее для удобства будем считать, что

$$h_0 = -\infty, \quad h_k = +\infty. \tag{8.88}$$

Каждой наблюдаемой частоте n_i $(1 \leqslant i \leqslant k)$ ставится в соответствие ожидаемая теоретическая частота np_i, где

$$p_i = F(h_i) - F(h_{i-1}) \tag{8.89}$$

(с учетом соглашения (8.88)), а $F(x)$ — предполагаемая (теоретическая) функция распределения, гипотеза о совпадении которой с функцией распределения случайной величины X проверяется.

Таблица 8.25

Интервалы группировки	Наблюдаемые частоты n_i	Ожидаемые частоты np_i	Расхождение $\delta_i = n_i - np_i$
$< h_1$	n_1	np_1	δ_1
$[h_1, h_2)$	n_2	np_2	δ_2
...	...	...	...
$[h_{k-2}, h_{k-1})$	n_{k-1}	np_{k-1}	δ_{k-1}
$> h_{k-1}$	n_k	np_k	δ_k

Понятно, что нужен способ объединения индивидуальных расхождений δ_i в некоторую единую статистику, которую можно было бы использовать как статистику критерия согласия при проверке нулевой гипотезы H_0: $P(X \in [h_{i-1}, h_i)) = P_i$, $1 \leqslant i \leqslant k$, при конкурирующей гипотезе H_1, противоположной H_0.

Возьмем в качестве такой статистики величину

$$\chi^2 = \sum_{i=1}^{k} \frac{\delta_i^2}{np_i} = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}, \tag{8.90}$$

которая называется статистикой Пирсона или статистикой (хи-квадрат).

Можно показать, что при гипотезе H_0 статистика χ^2, определяемая формулой (8.90), приближенно имеет распределение χ^2 с числом степеней свободы $r = k - 1 - l$, где k — число интервалов группировки, и l — число неизвестных параметров теоретического распределения $F(x)$, определяемых по выборке.

Так, если предполагается, что распределение $F(x)$ является нормальным с параметрами (μ, σ) и эти параметры известны, то $l = 0$ и $r = k - 1$; если же μ и σ заменяются их выборочными аналогами $\overline{X}$ и S (см. п. 8.2.1), то $l = 2$ и $r = k - 3$.

Критерий согласия Пирсона выглядит следующим образом.

В таблице критических точек распределения χ^2 (см. приложение 4) по заданному уровню значимости α и числу степеней свободы $r = k-1-l$, где k — число интервалов группировки, l — число параметров теоретического распределения, определяемых по выборке, находят критическую точку $\chi^2_{\text{кр}}(\alpha, r)$. Затем по формуле (8.90) (см. также (8.89) и табл. 8.25) находят наблюдаемое значение $\chi^2_{\text{набл}}$ статистики Пирсона.

Теперь, если $\chi^2_{\text{набл}} \leqslant \chi^2_{\text{кр}}(\alpha, r)$, то оснований отвергнуть гипотезу H_0 не имеется, и мы принимаем допущение, что $P(h_{i-1} \leqslant X < h_i) = p_i$ при всех i, $1 \leqslant i \leqslant k$. Если же $\chi^2_{\text{набл}} > \chi^2_{\text{кр}}(\alpha, r)$, то гипотезу H_0 отклоняют на уровне, приближенно равном α, как несовместимую с выборочными данными.

Можно показать, что ошибка второго рода в критерии согласия Пирсона стремится к нулю при неограниченном возрастании объема выборки n.

Замечание 1. Справедлива формула $\chi^2 = \dfrac{1}{n}\sum\limits_{i=1}^{k} \dfrac{n_i^2}{p_i} - n$.

Замечание 2. Перед непосредственным применением критерия χ^2 следует объединять интервалы (и складывать соответствующие частоты) группировки на «хвостах» таким образом, чтобы наименьшая ожидаемая частота была не меньше 1, т. е. если, например, в табл. 8.25 величина $np_1 < 1$, то интервалы $(-\infty, h_1)$ и $[h_1, h_2)$ объединяются в один общий интервал $(-\infty, h_2)$, содержащий n_1+n_2 наблюдений с ожидаемой частотой $n(p_1 + p_2)$.

Отметим, что в случае объединения интервалов число степеней свободы r статистики χ^2, подсчитанной по преобразованной группировке, соответствующим образом уменьшится.

Приведем два примера.

Пример 5.7. В препарате через равные промежутки времени регистрируется число бактерий, попавших в поле зрения микроскопа. Получены следующие данные:

Таблица 8.26

i	0	1	2	3	4	5	6	7
n_i	112	168	130	68	32	5	1	1

Здесь i — число бактерий и n_i — число моментов времени, соответствующих наблюдению ровно i бактерий.

Используя критерий χ^2, проверить при уровне значимости 0,05, что число бактерий, попадающих в поле зрения микроскопа в любой момент регистрации, распределено по закону Пуассона с параметром $\lambda = 1{,}5$.

Решение. Будем считать, что признаком генеральной совокупности является неотрицательная случайная величина X, равная числу бактерий, попадающих в поле зрения микроскопа в произвольный момент регистрации.

Тогда данные из табл. 8.26 представляют статистический ряд распределения выборки (см. табл. 8.1) объема $n = n_0 + n_1 + \ldots$ $\ldots + n_7 = 517$, отвечающей случайной величине X.

Разобьем ось на 8 интервалов Δ_i, $0 \leqslant i \leqslant 7$: $(-\infty, 1)$, $[1, 2)$, $[2, 3), \ldots, [6, 7)$, $[7, +\infty)$. Тогда n_i означает число вариант, попавших в интервал Δ_i. Учитывая, что случайная величина X_λ, имеющая распределение Пуассона с параметром λ, принимает лишь значения $l = 0, 1, 2, \ldots$ с вероятностями $\Pi_l = P\,(X_\lambda = l) =$ $= \dfrac{\lambda^l e^{-\lambda}}{l!}$, заключаем, что теоретические вероятности P_l (см.

(8.89)) попадания соответствующего числа бактерий в поле зрения микроскопа с точностью до 10^{-4} равны

$P_0 = \Pi_0 = 0{,}2231$	$P_1 = \Pi_1 = 0{,}3347$
$P_2 = \Pi_2 = 0{,}2510$	$P_3 = \Pi_3 = 0{,}1255$
$P_4 = \Pi_4 = 0{,}0471$	$P_5 = \Pi_5 = 0{,}0141$
$P_6 = \Pi_6 = 0{,}0035$	$P_7 = \Pi_7 + \Pi_8 + \ldots = 0{,}0045$

Таким образом, ожидаемые частоты равны $517\,P_i$, и аналог табл. 8.25 имеет следующий вид:

Таблица 8.27

Число бактерий	Наблюдаемые частоты	Ожидаемые частоты	Расхождение
0	112	115,34	−3,34
1	168	173,04	−5,04
2	130	129,77	0,23
3	68	64,88	3,12
4	32	24,35	7,65
5	5	7,29	−2,29
6 } 7 }	1 } 1 }	2,33 }	−0,33 }

В соответствии с рекомендациями замечания 2 элементы двух последних строк табл. 8.27 должны быть объединены.

В результате число интервалов группировки с 8 уменьшается до 7, что приводит к уменьшению на единицу числа степеней свободы r (с $7 = 8 - 1$ до $6 = 7 - 1$).

Критическая точка $\chi^2_{\text{кр}}(0{,}05, 6) = 12{,}59$ (см. приложение 4), а наблюдаемое значение статистики Пирсона, подсчитанное по формуле (8.90) (при $k = 6$ и с модифицированными частотами n_6 и np_6),

$$\chi^2_{\text{набл}} = \frac{(-3{,}34)^2}{115{,}34} + \frac{(-5{,}04)^2}{173{,}04} + \frac{0{,}23^2}{129{,}77} + \frac{3{,}12^2}{64{,}88} + \\ + \frac{7{,}65^2}{24{,}35} + \frac{(-2{,}29)^2}{7{,}29} + \frac{(-0{,}33)^2}{2{,}33} \approx 4{,}3.$$

Поскольку $4{,}3 < 12{,}59$, гипотеза о том, что число бактерий, попадающих в поле зрения микроскопа, распределено по закону Пуассона с параметром $\lambda = 1{,}5$, принимается.

Пример 5.8. Проверить по критерию χ^2 при уровне значимости 0,05 гипотезу о том, что распределение признака X из примера 1.4 п. 8.4 (см. табл. 8.6) является нормальным с параметрами (μ, σ), взяв в качестве μ и σ их оценки, полученные по сгруппированным данным.

Решение. Оценки для μ и σ^2 найдем по формулам (8.11) и (8.12), взяв в них в качестве x_i середины интервалов группировки. Имеем,

$$\overline{X} = \frac{1}{50} \cdot (2 \cdot 24 + 5 \cdot 28 + 9 \cdot 32 + 18 \cdot 36 + 9 \cdot 40 + 5 \cdot 44 + \\ + 2 \cdot 48) = 36,$$

$$\tilde{S}^2 = \frac{1}{50} \cdot (2 \cdot (-12)^2 + 5 \cdot (-8)^2 + 9 \cdot (-4)^2 + 18 \cdot 0^2 + \\ + 9 \cdot 4^2 + 5 \cdot 8^2 + 2 \cdot 12^2) = 30{,}08.$$

Принимая во внимание *поправку Шеппарда* на группировку, в соответствии с которой рекомендуется значение $\tilde{S}^2$ уменьшать на $(1/12)h^2$, где h — шаг группировки, окончательно получим

$$\sigma = \sqrt{\tilde{S}^2 - \frac{1}{12}h^2} = \sqrt{30{,}08 - \frac{1}{12} \cdot 4^2} = \sqrt{28{,}747} \approx 5{,}36,$$
$$\mu = 36.$$

Составим таблицу:

Таблица 8.28

i	$s_{0,05} = 1$	$z_i = \dfrac{h_i - \mu}{\sigma}$	$\Phi(z_i)$	$p_i = \Phi(z_i) - \Phi(z_{i-1})$	np_i	n_i	$np_i - n_i$
1	$(-\infty, 26)$	$-1{,}87$	$-0{,}4693$	0,0307	1,535	2	0,465
2	$(26, 30)$	$-1{,}12$	$-0{,}3686$	0,1007	5,035	5	$-0{,}035$
3	$[30, 34)$	$-0{,}37$	$-0{,}1443$	0,2243	11,215	9	$-2{,}215$
4	$[34, 38)$	0,37	0,1443	0,2886	14,43	18	3,57
5	$[38, 42)$	1,12	0,3686	0,2243	11,215	9	$-2{,}215$
6	$[42, 44)$	1,87	0,4693	0,1007	5.035	5	$-0{,}035$
7	$[44, +\infty)$	∞	0,5	0,0307	1,535	2	0,465

Здесь i $(1 \leqslant i \leqslant 7)$, $[h_{i-1}, h_i)$ и n_i — номер интервала группировки (с учетом соглашения (8.88)), i-й интервал группировки и его частота соответственно; $\Phi(z_i)$ вычисляется по приложению 1 (предполагается, что $z_0 = -\infty$), а вероятности p_i — по формуле (8.89).

Вычислим наблюдаемое значение $\chi^2_{\text{набл}}$. Имеем

$$\chi_{\text{набл}} = 2 \cdot \left(\frac{0{,}465^2}{1{,}535} + \frac{(-0{,}035)^2}{5{,}035} + \frac{(-2{,}215)^2}{11{,}215} \right) + \frac{3{,}57^2}{14{,}43} \approx 2{,}04.$$

Учитывая, что критическая точка распределения χ^2 с $7 - 1 - 2 = 4$ степенями свободы и уровнем значимости 0,05 $\chi^2_{\text{кр}}(4,\ 0{,}05) = 9{,}5$ превышает $\chi^2_{\text{набл}}$, делаем вывод о том, что гипотеза $X \sim N(36,\ 5{,}36)$ не противоречит статистическим данным.

Самостоятельная работа

1. По данным самостоятельных работ 2.1–2.4 из п. 8.4, предполагая, что пары наблюдаемых случайных величин X и Y имеют двумерное нормальное распределение, при уровнях значимости 0,05 и 0,01, проверить нулевую гипотезу о равенстве нулю коэффициента корреляции ρ между X и Y при конкурирующей гипотезе H_1: $\rho \neq 0$.

2. Средняя длина скорлупы 8 мелководных крабов оказалась равной 8,40 см при исправленной выборочной дисперсии этих длин 0,0016 см^2, а средняя длина скорлупы 12 глубоководных крабов оказалось равной 8,61 см при исправленной выборочной дисперсии этих длин 0,0038 см^2. Предполагается (с известной степенью точности), что обе выборки независимы и извлечены из нормально распределенных генеральных совокупностей. Определить, значимо или нет различаются между собой дисперсии длины скорлупы мелководных и глубоководных крабов при уровнях значимости $\varepsilon = 0{,}01$, $\varepsilon = 0{,}05$. Если различия между дисперсиями незначимы, проверить также гипотезу о равенстве генеральных средних.

3. Обследование 10 яиц, положенных кукушками в гнезда одного определенного вида птиц, показало, что средняя длина яйца $\overline{x} = 21{,}9$ мм, а $S = 0{,}79$ мм (здесь через S^2 обозначена исправленная выборочная дисперсия). Обследование 15 яиц, положенных кукушками в гнезда другого определенного вида птиц, показало, что $\overline{x} = 22{,}6$ мм, а $S = 0{,}86$ мм. Считается, что длина яйца распределена приближенно нормально. Определить значимость различий между дисперсиями длин яиц, положенных кукушками в гнезда разного вида птиц и, если различия между дисперсиями незначимы, проверить гипотезу о равенстве средних длин яиц. Уровни значимости принимаются $\varepsilon = 0{,}01$, $\varepsilon = 0{,}05$.

4. Даны результаты измерений частоты сердечных сокращений 11 студентов, проведенных сразу после окончания занятий по физкультуре, и 10 студентов — через полчаса после окончания этих занятий: исправленные выборочные дисперсии равны 139,9 и 74,2, соответственно. При уровнях значимости $\varepsilon = 0{,}01$, $\varepsilon = 0{,}05$ в предположении приближенной нормальности проверить гипотезу о равенстве генеральных дисперсий.

5. Пусть выборка 0,3, 6,3, 3,7, 2,8, 5,8, −1,4, 1,7, 2,3, −1,7, 1,6, −1,8, 0,6, 4,5, 1,9, 2,4, 6,8 объема 16 из непрерывного распределения содержит различия в измерениях артериального давления пациентов, сделанных до и после врачебного обхода. Требуется при уровне значимости $\alpha = 0{,}1$ проверить гипотезу о равенстве нулю медианы m теоретического распределения наблюдаемых разностей при конкурирующей гипотезе $m > 0$.

6. По группировкам, полученным в процессе решения самостоятельных работ 2–5 из п. 8.1, используя критерий χ^2, проверить при уровнях значимости 0,05 и 0,01 гипотезу о нормальном распределении соответствующих признаков, взяв в качестве значений параметров нормального распределения их оценки, полученные по сгруппированным данным.

§ 8.6. Основы дисперсионного анализа

В п. 8.5.3 была приведена статистическая процедура оценки значимости различия между средними двух нормальных генеральных совокупностей, основанная на критерии Стьюдента. Для сравнения средних в случае трех и более генеральных совокупностей используется другой подход, известный как дисперсионный анализ.

Типичной задачей, решаемой средствами дисперсионного анализа, является проверка существенности различий средних значений той или иной характеристики жизнедеятельности нескольких групп организмов, каждая из которых подвергалась влиянию одного или нескольких *факторов*, например некоторому медикаментозному воздействию при различных дозах и (или) модификациях применявшихся препаратов. При этом сами факторы принимают несколько значений (*уровней*), которые могут быть как фиксированными, так и случайными.

Ниже будет рассмотрен случай одного (п. 8.6.1) и двух (п. 8.6.3) фиксированных факторов.

8.6.1. Однофакторный дисперсионный анализ

Пусть $x_1^{(q)}, \ldots, x_n^{(q)}$, где $q = 1, 2, \ldots, k$, — независимые выборки объема n каждая, отвечающие случайным величинам $X^{(q)}$, нормально распределенным с некоторыми неизвестными параметрами $(\mu_q,\ \sigma)$. Здесь мы имеем дело с единственным фактором, принимающим k уровней $\mu_1, \mu_2, \ldots, \mu_q$.

Нулевая гипотеза имеет вид H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$, т. е. предполагает, что влияние всех уровней одинаково, и, следовательно, случайные величины $X^{(q)}$ имеют не только общую дисперсию σ^2, но и общее математическое ожидание.

Замечание. Изначальное предположение о равенстве дисперсий $D(\overline{X^q})$, $q = 1, 2, \ldots, k$, вообще говоря, требует проверки, которая может быть проведена при помощи критерия Кочрена (см. п. 8.6.2).

По каждой из k выборок составим (см. (8.7), (8.8)) выборочные средние $\overline{X^{(q)}} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i^{(q)}$, исправленные выборочные дисперсии

$$S_q^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i^{(q)} - \overline{X^{(q)}})^2 \tag{8.91}$$

и общее выборочное среднее $\overline{X} = \frac{1}{k}\sum\limits_{q=1}^{k} \overline{X^q}$, а также *факторную дисперсию*

$$S_{\text{факт}}^2 = \frac{n\sum\limits_{q=1}^{k} (\overline{X^{(q)}} - \overline{X})^2}{k-1}$$

и остаточную дисперсию

$$S_{\text{ост}}^2 = \frac{1}{k}\sum_{q=1}^{k} S_q^2. \tag{8.92}$$

Можно показать, что при гипотезе H_0 факторная и остаточная дисперсии являются независимыми и несмещенными оценками дисперсии σ^2. При этом $S^2_{\text{факт}}$ характеризует влияние фактора на случайные величины $X^{(q)}$ и имеет тенденцию при нарушении нулевой гипотезы в среднем расти с ростом n, в то время как математическое ожидание $S^2_{\text{ост}}$ не изменяется.

Применим в качестве статистики критерия случайную величину

$$\Lambda = S^2_{\text{факт}} / S^2_{\text{ост}}, \tag{8.93}$$

при нулевой гипотезе распределенную по закону Фишера–Снедекора с $k-1$ и $k(n-1)$ степенями свободы.

Проверка нулевой гипотезы сводится к проверке равенства двух дисперсий и осуществляется аналогично тому, как это делалось в п. 8.4.4: следует вычислить по формуле (8.94) $\Lambda_{\text{набл}}$ и сравнить его значение с критической точкой F-распределения $f_{\text{кр}}\,(\alpha; k-1, kn-k)$, отвечающей уровню значимости α и степеням свободы $k_1 = k-1$ и $k_2 = kn-k$ (см. приложение 3). В результате, если

$$\Lambda_{\text{набл}} \leqslant f_{\text{кр}}(\alpha; k-1, kn-k), \tag{8.94}$$

то нулевая гипотеза о том, что влияние всех уровней одинаково, принимается (заметим, что если $\Lambda_{\text{набл}} < 1$, то гипотеза H_0 заведомо справедлива); если же неравенство (8.94) не выполняется, то нулевая гипотеза отклоняется на уровне α и принимается конкурирующая гипотеза, заключающаяся в том, что по крайней мере один уровень фактора значимо отличается от других.

Следующие формулы упрощают вычисления:

$$S^2_{\text{факт}} = \frac{n}{k-1}\left(\sum_{q=1}^{k}(\overline{X^{(q)}} - \beta)^2 - k(\overline{X} - \beta)^2\right),$$

где β — любое число, взятое для облегчения расчетов;

$$S^2_{\text{ост}} = \frac{\sum\limits_{i=1}^{n}\sum\limits_{q=1}^{k}(x_i^{(q)} - \beta_q)^2 - n\sum\limits_{q=1}^{k}(\overline{X}^{(q)} - \beta_q)^2}{k(n-1)},$$

β_q, $q = 1, 2, \ldots, k$ — любые числа.

Пример 6.1. Две группы белых мышей по 10 животных в каждой группе в течении 86 дней подвергались воздействию биопрепаратов. По отношению к животным первой группы применялся лизат сердца $N/10$, а по отношению к животным второй группы — экстракт мышц $10N$. По прошествии 86 дней животные были забиты и определено отношение веса сердца к общему весу каждого в %. Данные опытов сведены в таблицу.

Биопрепарат	Номер опыта									
	1	2	3	4	5	6	7	8	9	10
Лизат сердца $N/10$	0,46	0,48	0,45	0,49	0,47	0,50	0,44	0,48	0,46	0,43
Экстракт мышц $10N$	0,47	0,46	0,48	0,50	0,51	0,48	0,52	0,45	0,50	0,49
Контроль	0,55	0,58	0,60	0,62	0,61	0,57	0,60	0,59	0,58	0,56

Провести дисперсионный анализ влияния указанных биопрепаратов на относительный вес сердца белых мышей при уровне значимости $\alpha = 0{,}01$.

Решение. Примем $\beta = 0{,}5$. Все необходимые расчеты приведены в табл. 8.29.

$$S_{\text{ост}}^2 = \frac{\sum\limits_{q=1}^{k} B_q - n\sum\limits_{q=1}^{k} D_q}{k(n-1)} =$$

$$= \frac{\sum\limits_{q=1}^{3} B_q - 10\sum\limits_{q=1}^{3} D_q}{27} = \frac{1{,}332\cdot 10^{-2}}{27} = 4{,}93\cdot 10^{-4},$$

$$\Lambda_{\text{набл}} = \frac{S_{\text{факт}}^2}{S_{\text{ост}}^2} = 44{,}75.$$

Принимая $\alpha = 0{,}01$, число степеней свободы (2, 27) по таблице критических точек F-распределения находим, что $f_{\text{кр}}(0{,}01; 2, 27) = 5{,}53$ значительно меньше, чем $\Lambda_{\text{набл}}$, т. е. налицо факт существенного влияния указанных биопрепаратов на относительный вес сердца белых мышей.

Таблица 8.29

№ опыта	Уровни фактора F (биопрепа					
	F_1 (лизат сердца) $N/10$			F_2 (экстракт мышц) 10		
	$x_i^{(1)}$	$x_i^{(1)} - \beta$	$(x_i^{(1)} - \beta)^2$	$x_i^{(2)}$	$x_i^{(2)} - \beta$	$(x_i^{(2)} -$
1	0,46	−0,04	$16 \cdot 10^{-4}$	0,47	−0,03	$9 \cdot 10^{-}$
2	0,48	−0,02	$4 \cdot 10^{-4}$	0,46	−0,04	$16 \cdot 10^{-}$
3	0,45	−0,05	$25 \cdot 10^{-4}$	0,48	−0,02	$4 \cdot 10^{-}$
4	0,49	−0,01	$1 \cdot 10^{-4}$	0,50	0	0
5	0,47	−0,03	$9 \cdot 10^{-4}$	0,51	0,01	$1 \cdot 10^{-}$
6	0,50	0	0	0,48	−0,02	$4 \cdot 10^{-}$
7	0,44	−0,06	$36 \cdot 10^{-4}$	0,52	0,02	$4 \cdot 10^{-}$
8	0,48	−0,02	$4 \cdot 10^{-4}$	0,45	−0,05	$25 \cdot 10^{-}$
9	0,46	−0,04	$16 \cdot 10^{-4}$	0,50	0	0
10	0,43	−0,07	$49 \cdot 10^{-4}$	0,49	−0,01	$1 \cdot 10^{-}$
	$\beta = 0{,}5$ $A_1 = \sum_{i=1}^{10} (x_i^{(1)} - \beta) = -0{,}34$ $B_1 = \sum_{i=1}^{10} (x_i^{(1)} - \beta)^2 = 0{,}016$ $C_1 = A_1/10 = -0{,}034$ $\overline{x^{(1)}} = C_1 + \beta = 0{,}466$ $D_1 = C_1^2 = 1{,}156 \cdot 10^{-3}$			$A_2 = -0{,}14$ $B_2 = 0{,}0064$ $C_2 = -0{,}014$ $\overline{x^{(2)}} = 0{,}486;$ $D_2 = 1{,}96 \cdot 10^{-4}$		

8.6.2. Проверка равенства нескольких дисперсий по выборкам равного объема. Критерий Кочрена

Пусть, так же как в п. 8.6.1, $x_1^{(q)}, \ldots, x_n^{(q)}$, $q = 1, 2, \ldots, k$, представляют собой независимые выборки объема n каждая, отвечающие нормально распределенным случайным величинам $X^{(q)}$. Требуется на уровне значимости α проверить гипотезу H_0: $D(X^{(1)}) = D(X^{(2)}) = \ldots = D(X^{(q)})$ при конкурирующей гипотезе H_1, противоположной H_0. Пусть S_q^2 — исправленные выборочные дисперсии, определенные формулой (8.91). В качестве статистики критерия возьмем отношение максимальной исправленной выборочной дисперсии к сумме всех исправленных выборочных дисперсий:

$$\Lambda = \frac{\max S_q^2}{\sum\limits_{q=1}^{k} S_q^2}. \tag{8.95}$$

Распределение случайной величины Λ при нулевой гипотезе зависит лишь от уровня значимости α, числа степеней свободы $n - 1$ и количества k выборок и называется распределением Кочрена.

Для проверки на уровне значимости α гипотезы H_0 о равенстве дисперсий k нормально распределенных совокупностей следует по формуле (8.95) вычислить наблюдаемое значение статистики критерия $\Lambda_{\text{набл}}$ и сравнить его с критической точкой $C_{\text{кр}}(\alpha; n-1, k)$ распределения Кочрена (приложение 5). Если $\Lambda_{\text{набл}} \leqslant C_{\text{кр}}(\alpha; n-1, k)$, то статистические данные подтверждают нулевую гипотезу, в противоположном случае нулевая гипотеза отвергается.

Пример 6.2. По данным примера 6.1 на уровне значимости 0,05 проверить гипотезу о равенстве дисперсий нормально распределенных случайных величин $X^{(q)}$, $q = 1, 2, 3$, отражающих значение изучаемого показателя в каждой из трех групп белых мышей.

Решение. Исправленные выборочные дисперсии в каждой группе с учетом обозначений и выкладок примера 6.1 вычисляются по формуле $S_q^2 = \dfrac{B_q - nD_q}{n-1}$, $n = 10$, $1 \leqslant q \leqslant k = 3$, и равны $4{,}932 \cdot 10^{-4}$. Следовательно, $\Lambda_{\text{набл}} = 0{,}333$. Поскольку $C_{\text{кр}}(0{,}05; 9, 3) = 0{,}6167 > 0{,}333$, заключаем, что нулевая гипотеза о равенстве дисперсий не опровергается.

8.6.3. Двухфакторный дисперсионный анализ

В исследованиях, включающих много факторов, неэффективно изучать воздействие каждого фактора на результат поочередно, тем более что такой способ не дает достаточной информации о возможных факторных взаимодействиях. Таким образом, следует рассмотреть эксперименты, в которых учитывается сразу два и более факторов.

Ниже будет проанализирован наиболее простой случай с двумя факторами, воздействующими на результат эксперимента. При этом мы ограничимся факторами F и G с фиксированными уровнями $F_1, F_2, \ldots, F_p$ и $G_1, G_2, \ldots, G_q$ соответственно и, кроме того, будем предполагать, что взаимодействие между факторами отсутствует и что при каждой комбинации уровней F_i, G_j, производится лишь по одному наблюдению x_{ij}. Модель такого эксперимента можно записать в виде

$$x_{ij} = F_i + G_j + \varepsilon_{ij}, \quad 1 \leqslant i \leqslant p, \quad 1 \leqslant j \leqslant q, \tag{8.96}$$

где ε_{ij} — независимые нормально распределенные случайные ошибки с нулевым средним и общей дисперсией σ^2.

Результаты наблюдений удобно представить в виде таблицы:

Таблица 8.30

Фактор G	Фактор F				
	F_1	F_2	...	F_p	Суммы по строкам
G_1	x_{11}	x_{21}	...	x_{p1}	$x_{\cdot 1}$
G_2	x_{12}	x_{22}	...	x_{p2}	$x_{\cdot 2}$
...	...	...	...	...	...
G_q	x_{1q}	x_{2q}	...	x_{pq}	$x_{\cdot q}$
Суммы по столбцам	$x_{1\cdot}$	$x_{2\cdot}$	...	$x_{p\cdot}$	$x_{\cdot\cdot}$

Здесь приняты обозначения $x_{\cdot j}$ $(j = 1, 2, \ldots, q)$ для сумм наблюдений по строкам и $x_{i\cdot}$ $(i = 1, 2, \ldots, p)$ для сумм наблюдений по столбцам, $x_{\cdot\cdot}$ — общая сумма всех наблюдений.

Факторные S_F^2, S_G^2 и остаточную $S_{\text{ост}}^2$ дисперсии можно вычислить по формулам

$$S_F^2 = \frac{1}{p-1}\left(\frac{1}{q}\sum_{i=1}^{p}(x_{i\cdot} - \beta)^2 - \frac{1}{pq}(x_{\cdot\cdot} - p\beta)^2\right),$$

$$S_G^2 = \frac{1}{q-1}\left(\frac{1}{p}\sum_{i=1}^{q}(x_{\cdot j} - \beta)^2 - \frac{1}{pq}(x_{\cdot\cdot} - q\beta)^2\right), \tag{8.97}$$

$$S_{\text{ост}}^2 = \frac{1}{(p-1)(q-1)}\left(S^2 - (p-1)S_F^2 - (q-1)S_G^2\right),$$

где

$$S^2 = \sum_{i,j}(x_{ij} - \beta)^2 - \frac{1}{pq}(x_{\cdot\cdot} - pq\beta)^2, \tag{8.98}$$

β — произвольное число, которое может быть разным в различных формулах.

При этом (см. (8.96)) $M(S_{\text{ост}}^2) = \sigma^2$, т. е. $S_{\text{ост}}^2$ является несмещенной оценкой дисперсии σ^2. Аналогичным свойством, в предположениях H_F: $F_1 = F_2 = \ldots = F_p$ и H_G: $G_1 = G_2 = \ldots = G_q$ соответственно, обладают факторные дисперсии S_F^2 и S_G^2.

Таким образом, так же как в п. 8.6.1, мы можем составить отношения $\Lambda_F = \dfrac{S_F^2}{S_{\text{ост}}^2}$ или $\Lambda_G = \dfrac{S_G^2}{S_{\text{ост}}^2}$ и с их помощью проверить нулевые гипотезы H_F или H_G, сравнивая Λ_F или Λ_G с критической точкой $f_{\text{кр}}(\alpha; k_1, k_2)$ F- распределения (приложение 3) со степенями свободы $k_1 = q - 1$ или $p - 1$, и $k_2 = (p-1)(q-1)$.

Например, чтобы проверить существенность влияния фактора F на результаты эксперимента, следует, задавшись уровнем значимости α, сравнить наблюдаемое значение $\Lambda_{F,\text{набл}}$ статистики критерия Λ_F с величиной $f_{\text{кр}} = f_{\text{кр}}(\alpha; p-1, (p-1)(q-1))$ из приложения 3.

Если $\Lambda_{F,\text{набл}} < f_{\text{кр}}$, то нулевая гипотеза H_F об отсутствии влияния фактора F на результаты эксперимента подтверждается, если же $\Lambda_{\text{набл}} \geqslant f_{\text{кр}}$, то гипотеза H_F отвергается на уровне α и делается вывод, что фактор F воздействует на результаты.

Заметим, что если $\Lambda_{\text{ост}} \leqslant 1$, то нулевая гипотеза принимается сразу.

Пример 6.3. В условиях модели (8.96) определить по данным табл. 8.31 достоверность влияния препаратов F и G на массу подопытных животных при уровне значимости $\alpha = 0{,}05$.

Таблица 8.31

Фактор G	Фактор F			
	F_1	F_2	F_3	Суммы по строкам
G_1	30	35	40	105
G_2	32	39	38	109
G_3	34	38	44	116
G_4	28	36	42	106
Суммы по столбцам	124	148	164	436

Решение. Сначала найдем S_F^2. Имеем (см. (8.97) при $p = 3$, $q = 4$ и $\beta = 148$)

$$S_F^2 = \frac{1}{2}\left(\frac{1}{4}(-24)^2 + 0^2 + 16^2) - \frac{1}{12}(436 - 3 \cdot 148)^2\right) = 101{,}33.$$

Аналогично при $\beta = 109$

$$S_G^2 = \frac{1}{3}\left(\frac{1}{3}(-4)^2 + 0^2 + 7^2 + (-3)^2) - \frac{1}{12}(436 - 4 \cdot 109)^2\right) = 8{,}22$$

и (при $\beta = 38$)

$$S^2 = (-8)^2 + (-3)^2 + 2^2 + (-6)^2 + 1^2 + 0^2 + (-4)^2 + 0^2 + 6^2 + \\ + (-10)^2 + (-2)^2 + 4^2 - \frac{1}{12}(436 - 12 \cdot 38)^2 = 252{,}67,$$

$$S_{\text{ост}}^2 = \frac{1}{2 \cdot 3}(252{,}67 - 2 \cdot 101{,}33 - 3 \cdot 8{,}22) = \frac{25{,}35}{6} = 4{,}22.$$

Отсюда $\Lambda_{F,\text{набл}} = 24{,}01$ и $\Lambda_{G,\text{набл}} = 1{,}95$. Учитывая (см. приложение 3), что $f_{\text{кр}}(005; 3, 6) = 4{,}76 > 1{,}95$, а $f_{\text{кр}}(005; 2, 6) = 5{,}14 < 24{,}01$, делаем вывод, что препарат F оказывает влияние на массу подопытных животных, а препарат G такого влияния не оказывает.

Самостоятельная работа

1. Проводились опыты по изменению веса тела и отдельных органов у млекопитающих путем обработки их биопрепаратами. В каждой серии опыта на белых мышах было по 10 животных. Использовались молодые мыши весом 16–18 г. Каждый опыт продолжался 56 или 86 дней. В результате опытов после забоя животных определялось изменение общего веса и веса отдельных органов, вызванное воздействием биопрепаратов. Результаты опытов сведены в таблицы, приведенные ниже.

№ 1. Изменение общего веса после обработки животных биопрепаратами (прирост общего веса в граммах).

Биопрепарат	Номер опыта									
	1	2	3	4	5	6	7	8	9	10
Лизат печени N	3,83	3,55	3,61	3,75	3,52	3,43	3,24	3,66	3,37	3,94
Лизат селезенки $N/10$	3,02	3,30	3,09	3,12	3,25	2,92	3,20	2,82	3,36	3,52
Экстракт мозга $10N$	3,32	3,33	3,16	3,40	3,62	3,73	3,03	3,22	3,47	3,52

Продолжительность опыта 56 дней.

№ 2. Изменение веса отдельных органов после обработки животных биопрепаратами. Измерялось отношение веса органа к общему весу тела животного в %.

Печень

Биопрепарат	Номер опыта									
	1	2	3	4	5	6	7	8	9	10
Экстракт печени $N/10$	6,06	6,27	6,15	6,32	6,11	5,92	6,18	6,41	6,22	6,01
Лизат мозга $10N$	5,85	5,94	6,03	5,71	6,20	5,91	6,13	6,98	6,07	5,81
Лизат мозга $N/100$	6,15	6,27	6,09	6,35	6,20	6,45	6,31	5,95	6,21	6,05

Продолжительность опыта 56 дней.

Мозг

Биопрепарат	Номер опыта									
	1	2	3	4	5	6	7	8	9	10
Экстракт печени $N/10$	1,66	1,68	1,72	1,69	1,71	1,64	1,70	1,68	1,69	1,67
Экстракт печени N	1,68	1,69	1,66	1,70	1,65	1,70	1,73	1,67	1,68	1,69
Экстракт мышц $10N$	1,55	1,57	1,56	1,58	1,53	1,61	1,61	1,58	1,59	1,57
Лизат сердца $N/10$	1,59	1,65	1,62	1,57	1,60	1,64	1,60	1,63	1,62	1,61

Продолжительность опыта 86 дней.

Сердце

Биопрепарат	Номер опыта									
	1	2	3	4	5	6	7	8	9	10
Лизат сердца $N/100$	0,48	0,49	0,48	0,49	0,50	0,47	0,49	0,50	0,51	0,48
Лизат сердца $N/10$	0,49	0,50	0,49	0,50	0,49	0,50	0,51	0,53	0,51	0,48
Лизат селезенки $N/100$	0,46	0,48	0,48	0,49	0,50	0,52	0,50	0,48	0,50	0,49
Лизат селезенки $10N$	0,47	0,49	0,49	0,50	0,50	0,51	–0,52	–0,49	0,54	0,50

Продолжительность опыта 56 дней.

Для каждой из таблиц провести сравнительный дисперсионный анализ влияния биопрепаратов на изменение веса тела или отдельных органов белых мышей при уровнях значимости $\alpha = 0{,}01$ и $\alpha = 0{,}05$. Предварительно проверить предположение о равенстве дисперсий изучаемых показателей каждой из групп по критерию Кочрена на уровне значимости 0,05.

3. Методом двухфакторного дисперсионного анализа проверить значимость влияния каждого фактора на результирующий признак.

№ 1. Выявить влияние реагентов F и G на синтез лекарственного препарата (выход в условных единицах) при уровне значимости 0,05.

Фактор G	Фактор F			
	F_1	F_2	F_3	F_4
1	4,5	3	4	3,5
2	3,5	2,5	3,5	2
3	6,5	5,5	4,5	6
4	7,5	7	8,5	7

№ 2. Проверить значимость влияния двух факторов: pH (фактор F) и концентрации (%) лиганда (фактор G) на экстракцию комплекса металла с лигандом из водной в органическую фазу при уровне значимости 0,1.

Фактор G	Фактор F		
	F_1	F_2	F_3
3	63	42	71
6	69	73	82
9	91	68	60

№ 3. Проверить существенность влияния температуры (фактор F) и фермента (фактор G) на продукт биохимического анализа при уровне значимости 0,01.

Фактор G	Фактор F	
	F_1	F_2
27	100	90
3	86	87

§ 8.7. Временные ряды. Основные понятия

Временной ряд — это совокупность измерений некоторой переменной x_t, производимых по мере возрастания времени. Значение x_t в каждый момент времени t является случайной величиной. С временными рядами связаны многие задачи науки и техники. Примерами служат:

1) запись электрокардиограммы,

2) изменение уровня жидкости в капельнице,

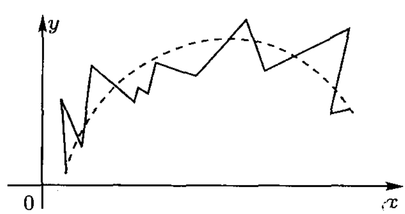


Рис. 8.13. Графическое изображение фрагмента временного ряда с параболическим трендом.

3) количество поступивших в медучреждение больных по дням,

4) измерение артериального давления через каждый час.

Еще одним примером временного ряда может служить переменная x_t, полученная при наложении случайных флуктуаций на детерминированную (неслучайную) зависящую от времени составляющую, которую называют *трендом* временного ряда (тренд изображен на рис. 8.13 пунктирной линией).

В примерах 1 и 2 измерения могут регистрироваться (по крайней мере, теоретически) непрерывно, в силу чего соответствующие временные ряды называют *непрерывными*, примеры 3 и 4 относятся к группе *дискретных временных рядов*.

Далее мы будем рассматривать временные ряды лишь при целых значениях t, поскольку измерения обычно осуществляются через некоторые равные промежутки времени и нумеруются аналогично элементам выборки $x_1, x_2, \ldots, x_n$ объема n.

При изучении временных рядов основной интерес представляет моделирование их структуры с дальнейшим применением модели для экстраполяции или прогнозирования. При исследовании временных рядов необходим статистический подход из-за ошибок измерений и случайных флуктуаций, свойственных практически любой наблюдаемой системе, относится ли она к медицине, биологии, окружающей среде или технике. Важной составляющей статистических методов анализа временных рядов является оценка тренда, который, при наличии информации об его виде, можно моделировать при помощи компонент, являющихся детерминированными функциями времени. Отметим, что эксперименты, в которых осуществляются наблюдения, как правило, не являются независимыми, и последовательные ошибки модели должны, вообще говоря, рассматриваться как статистически связанные. Обычной практикой является предположение, что последовательные ошибки модели представляют *стационарный временной ряд*.

Определение. Временной ряд x_t называют стационарным, если:

1) математическое ожидание и дисперсия ряда постоянны во времени:

$$M(x_t) = \mu, \quad D(x_t) = \sigma^2 \tag{8.99}$$

при всех t;

2) ковариация между любыми членами ряда (см. п. 8.4.2) зависит только от расстояния во времени между этими наблюдениями, или (в дискретном случае) от разности между их номерами, т. е. $\operatorname{cov}(x_t, x_{t+k})$ для всех t зависит лишь от k.

Заметим, что если совместное распределение членов стационарного временного ряда $x_1, x_2, \ldots, x_n$ при любом n является нормальным, то совместное распределение $x_{k+1}, x_{k+2}, \ldots, x_{k+n}$ не зависит от k, или, другими словами, вероятностное поведение любой группы членов временного ряда x_t не зависит от сдвига по t. В частности, любые индивидуальные вероятностные характеристики x_t не изменяются во времени.

8.7.1. Оценки вероятностных характеристик временных рядов

Пусть x_t является стационарным временным рядом и мы хотим оценить параметры μ и σ^2 (см.(8.99)) по наблюдениям (выборке) $x_1, x_2, \ldots, x_n$. Можно показать, что при некоторых весьма слабых предположениях несмещенной и состоятельной оценкой (см. п. 8.2.1) для μ является выборочное среднее

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + x_2 + \ldots + x_n), \tag{8.100}$$

а состоятельной оценкой для σ^2 будет выборочная дисперсия S^2, которую можно определить как среднее арифметическое квадратов отклонений наблюдавшихся значений временного ряда от их среднего значения:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2. \tag{8.101}$$

Заметим, что формулы (8.100) и (8.101) совпадают с формулами (8.7), (8.9), что неудивительно, поскольку и те, и другие оценивают одни и те же вероятностные характеристики.

Моделью некоторых *нестационарных* временных рядов служат процессы вида

$$y_t = \mu_t + x_t, \tag{8.102}$$

где детерминированная функция μ_t зависит лишь от t, а x_t — стационарный временной ряд с нулевым средним $M(x_t) = 0$.

Очевидно, что $M(y_t) = \mu_t$, а $\operatorname{cov}(y_t, y_{t+k}) = \operatorname{cov}(x_t, x_{t+k})$. Поэтому детерминированная компонента μ_t характеризует тенденцию изменения временного ряда y_t со временем в среднем, т. е. его тренд, а слагаемое x_t определяет случайные, не зависящие от t ошибки модели и структуру их зависимости. Типичный вид графического изображения подобного нестационарного временного ряда представлен на рис. 8.13.

8.7.2. Сглаживание временных рядов

Как уже говорилось выше, одной из задач теории временных рядов является оценивание тренда временного ряда x_t по результатам наблюдений $x_1, x_2, \ldots, x_n$ его отрезка длины n. Эту процедуру можно осуществить при помощи операции *сглаживания*, целью которой является уменьшение дисперсии временного ряда x_t, а, по существу, амплитуды случайных флуктуаций вокруг его детерминированной составляющей. Сглаженный временной ряд обычно получают как линейную комбинацию элементов x_t, $x_{t\pm1}, \ldots, x_{t\pm m}, \ldots$ с некоторыми неотрицательными весами a_0, $a_{\pm1}, \ldots, a_{\pm m}, \ldots$, в сумме равными единице, которые желательно выбирать, используя информацию о ковариационной структуре ряда. На практике в основном ограничиваются *методом экспоненциального сглаживания* или его частным случаем — *методом скользящего среднего*. При любом использованном методе сглаживания исследователь рассчитывает, что полученный в результате сглаживания новый временной ряд y_t будет иметь более четко выраженный тренд, мало отличающийся от тренда первоначального ряда x_t и, следовательно, в первом приближении могущий его заменить.

Приведем формулу для метода экспоненциального сглаживания по трем точкам:

$$y_t = \frac{1}{1+2q}\left(qx_{t-1}+x_t+qx_{t+1}\right), \quad 2 \leqslant t \leqslant n-1, \quad q \geqslant 0. \tag{8.103}$$

Если в (8.103) положить $q = 1$, получим формулу для метода скользящего среднего по трем точкам:

$$y_t = \frac{1}{3}\left(x_{t-1} + x_t + x_{t+1}\right), \quad 2 \leqslant t \leqslant n-1; \tag{8.104}$$

устремляя q к бесконечности, придем к сглаживающей формуле

$$y_t = \frac{1}{2}\left(x_{t-1} + x_{t+1}\right), \quad 2 \leqslant t \leqslant n-1. \tag{8.105}$$

Соотношения (8.103) и (8.104) легко обобщаются с трех на произвольное нечетное число точек $2l + 1$. Соответствующий вариант формулы (8.104), например, имеет вид

$$\begin{gathered} y_t = \frac{1}{1+2l}\left(x_{t-l} + \ldots + x_t + \ldots + x_{t+l}\right), \\ l+1 \leqslant t \leqslant n-l. \end{gathered} \tag{8.106}$$

Приведенные методы сглаживания оставляют *линейный тренд* временного ряда x_t без изменения и, вообще говоря, уменьшают его дисперсию. В частности, дисперсия сглаженного по формуле (8.106) временного ряда, определенного равенством (8.102), в том случае, когда случайные величины x_t при разных t являются независимыми, уменьшается ровно в $2l + 1$ раз.

Отметим также, что при сглаживании, как правило, происходит некоторая потеря информации. Так, при использовании метода скользящего среднего по $2l + 1$ точкам отрезок сглаженного ряда будет содержать $n - 2l$ элементов вместо n.

Пример 7.1. Данные о динамике роста объема производства x_t некоторого препарата (в тоннах) на фармацевтической фабрике за 10 последовательных лет представлены в табл. 8.32:

Таблица 8.32

t	1	2	3	4	5	6	7	8	9	10
x_t	12	10	17	13	20	18	25	27	24	30

Провести сглаживание временного ряда x_t, используя формулы (8.104), (8.105) и (8.103) при $q = 1/2$. На одном графике построить изображение ряда x_t и трех сглаженных его вариантов $y_t(1)$, $y_t(2)$, $y_t(3)$.

Решение. Сглаженные значения $y_t(1)$, $y_t(2)$, $y_t(3)$, подсчитанные при $t = 2, \ldots, 9$ по указанным формулам, запишем в табл. 8.33:

Таблица 8.33

T	2	3	4	5	6	7	8	9
$y_t(1)$	13	13	17	17	21	23	25	27
$y_t(2)$	14,5	11,5	18,5	15,5	22,5	22,5	24,5	28,5
$y_t(3)$	12,7	13,7	16,3	17,3	20,7	23,5	25,5	26,7

Например, столбцы таблицы, соответствующие $t = 2$ и $t = 9$, заполняются следующим образом:

$$y_2(1) = (12 + 10 + 17)/3 = 13, \quad y_2(2) = (12 + 17)/2 = 14{,}5,$$

$$y_2(3) = (12 \cdot 3 + 10 \cdot 4 + 17 \cdot 3)/10 = 12{,}7,$$

$$y_9(1) = (27 + 24 + 30)/3 = 27, \quad y_9(2) = (27 + 30)/2 = 28{,}5,$$

$$y_9(3) = (27 \cdot 3 + 24 \cdot 4 + 30 \cdot 3)/10 = 26{,}7.$$

Графики, построенные по таблицам 8.32 и 8.33, изображены на рис. 8.14 Их визуальный анализ дает основания полагать, что наблюдаемый временной ряд имеет линейный тренд.

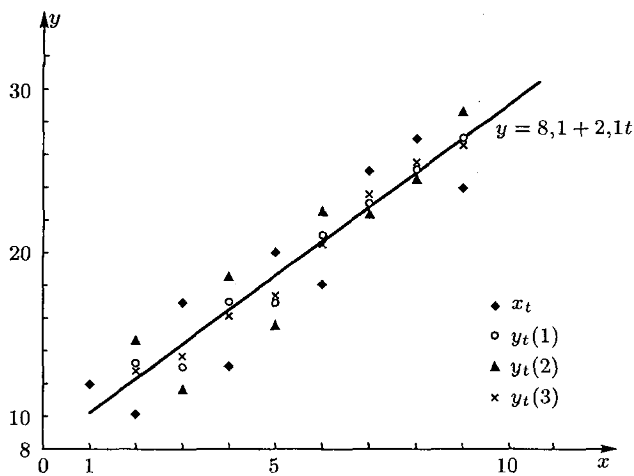


Рис. 8.14. Графики к примерам 7.1 и 7.2.

Если с помощью, скажем, метода скользящего среднего или из каких-либо других, например теоретических, соображений сделан вывод, что тренд μ_t временного ряда (см. (8.102)) является функцией, известной с точностью до нескольких параметров, то оценки последних можно определить, используя метод наименьших квадратов.

В случае, когда тренд μ_t является линейной функцией времени, т. е. $\mu_t = a_0 + a_1 t$, МНК-оценки параметров a_1 и a_0 выражаются формулами

$$a_1 = \frac{\overline{t \cdot x} - \overline{t} \cdot \overline{x}}{\overline{t^2} - \overline{t}^2}, \qquad a_0 = \overline{x} - a_1 \overline{t}, \tag{8.107}$$

в которых

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \qquad \overline{t} = \frac{1}{n} \sum_{i=1}^{n} t_i,$$

$$\overline{tx} = \frac{1}{n} \sum_{i=1}^{n} t_i x_i, \qquad \overline{t^2} = \frac{1}{n} \sum_{i=1}^{n} t_i^2,$$

а x_i является значением временного ряда в момент времени t_i, $i = 1, 2, \ldots, n$ (n — объем выборки). Обращаем внимание на то, что формулы (8.107) с точностью до обозначений совпадают с формулами (8.14) и (8.33).

Пример 7.2. По данным табл. 8.32 методом наименьших квадратов оценить коэффициенты уравнения линейного тренда для зависимости объема производства от номера года.

Решение. Для оценки параметров уравнения тренда $\mu_t = a_0 + a_1 t$ применим формулы (8.107):

$$a_1 = \frac{10(12 \cdot 1 + \ldots + 30 \cdot 10) - (12 + \ldots + 30)(1 + \ldots + 10)}{10\,(1^2 + \ldots + 10^2) - (1 + \ldots + 10)^2} \approx 2{,}1,$$

$$a_0 = \frac{(12 + \ldots + 30)(1^2 + \ldots + 10^2) - (12 \cdot 1 + \ldots + 30 \cdot 10)(1 + \ldots + 10)}{10(1^2 + \ldots + 10^2) - (1 + \ldots + 10)^2} \approx$$
$$\approx 8{,}1.$$

В заключение заметим, что знание тренда μ_t временного ряда позволяет с известной степенью надежности и точности *прогнозировать* его значения. Предположительные значения ряда для моментов времени t, выходящих за границу проведенных наблюдений, считаются равными μ_t. При этом важно понимать, что оценки подобного сорта предполагают, что основная тенденция изменения временного ряда в течение интервала времени между моментом наблюдения и моментом времени, для которого оценивается значение временного ряда, сохраняется. Кроме того, надо иметь в виду, что точность прогноза, как правило, снижается с ростом этого интервала.

Пример 7.3. Используя результат решения примера 7.2, оценить объем производства указанного препарата (в тоннах) на фармацевтической фабрике в 11-м году от начала наблюдений.

Решение. Уравнение тренда для зависимости объема производства от номера года наблюдений из примера 7.2 имеет вид $\mu_t = 8{,}1 + 2{,}1 \cdot t$. Подставляя в это уравнение значение $t = 11$, получим прогнозируемое на 11-й год от начала наблюдений значение объема производства $\mu_{11} = 8{,}1 + 2{,}1 \cdot 11 = 31{,}2$.

Самостоятельная работа

1. В табл. 8.34 приведены данные об объемах реализации некоторого препарата (в тыс. упаковок) в аптеке за 10 последовательных месяцев:

Таблица 8.34

N	5,1	4,9	5,0	4,8	5,0	5,2	5,0	4,9	5,1

В предположении стационарности временного ряда, образуемого этими значениями, оценить математическое ожидание, дисперсию и среднее квадратичное отклонение для этого ряда.

2. В табл. 8.35 приведены результаты измерения массы лабораторного животного (в граммах), полученные последовательно в течение 12 дней.

Таблица 8.35

M	50	52	48	50	49	50	51	53	47	50	52	48

В предположении стационарности временного ряда, образуемого этими значениями, оценить математическое ожидание, дисперсию и среднее квадратичное отклонение для этого ряда.

3. Данные по дисциплине роста товарооборота аптеки (в десятках миллионов рублей) за 10 последовательных лет представлены в табл. 8.36:

Таблица 8.36

t	1	2	3	4	5	6	7	8	9	10
V_t	10	9	16	13	23	21	28	27	34	33

Составить уравнение тренда для зависимости товарооборота аптеки от года в предположении линейности этой зависимости и оценить товарооборот аптеки в 11-м году от начала наблюдений.

4. Методом скользящего среднего по трем точкам произвести сглаживание временного ряда из упражнения 3.

ГЛАВА 9

ВВЕДЕНИЕ В МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОПТИМИЗАЦИИ

Математические методы оптимизации используются при решении экономических задач, в которых требуется выбрать оптимальное из нескольких решений. Потребности практики, связанные с планированием целенаправленной человеческой деятельности, вызвали к жизни специальные научные методы, одним из которых является *математическое программирование* (английское слово programming означает составление плана или программы действий).

Чтобы представить специфику этой науки, приведем несколько типичных для нее задач.

1) Продажа сезонных товаров

Для реализации определенной массы сезонных товаров создается сеть временных торговых точек. Требуется выбрать разумным образом: число точек, их размещение, товарные запасы и количество персонала на каждой из них так, чтобы обеспечить максимальную экономическую эффективность распродажи.

2) Медицинское обследование

Известно, что в каком-то районе обнаружены случаи опасного заболевания. С целью выявления заболевших (или носителей инфекции) организуется медицинское обследование жителей района. На это выделены материальные средства, оборудование, медицинский персонал. Требуется разработать такой план обследования (число медпунктов, их размещение, последовательность осмотров специалистами, виды анализов и т. д.), который позволит выявить, по возможности, максимальный процент заболевших и носителей инфекции.

3) Составление раствора

Пусть, например, требуется приготовить раствор, содержащий некоторые вещества A и B, используя два стандартных готовых раствора, в первом из которых концентрации этих веществ составляют a_1 и b_1, а во втором — a_2 и b_2, причем стоимость единицы объема первого раствора составляет c_1, а второго раствора — c_2. При этом стоимость приготовленного раствора должна быть минимальной, тогда как вещества A в нем должно быть не меньше α, а вещества B — не меньше β.

§ 9.1. Линейное программирование

9.1.1. Основные определения и понятия

Общая задача математического программирования состоит в нахождении вектора $x = (x_1, x_2, \ldots, x_n)$, удовлетворяющего системе ограничений

$$\begin{aligned} g_i(x_1, \ldots, x_n) = b_i, \quad & i = 1, \ldots, k \\ g_i(x_1, \ldots, x_n) \leqslant b_i, \quad & i = k+1, \ldots, m, \end{aligned} \tag{9.1}$$

и доставляющего экстремум функции

$$z = f(x_1, \ldots, x_n). \tag{9.2}$$

Функции f, g_i считаются известными. Обычно на некоторые переменные x_i накладывается условие неотрицательности (из прикладного смысла задачи). Кроме того, ограничением может служить условие целочисленности решения для ряда переменных.

Определение. Функция f, для которой отыскивается экстремум, называется *целевой*, функции g_i — функциями ограничений.

Определение. Вектор $x = (x_1, \ldots, x_n)$, удовлетворяющий ограничениям (9.1), называется *допустимым вектором*.

Определение. Множество всех допустимых векторов называется допустимым множеством.

Определение. Допустимый вектор, доставляющий экстремум целевой функции — *оптимальное решение* (оптимальный план) задачи.

Определение. Если функции (9.1), (9.2) являются линейными, то соответствующая задача называется *задачей линейного программирования* (ЛП), в противном случае — *нелинейного программирования* (НП).

Итак, задача линейного программирования в математической формулировке имеет следующий вид.

Найти значения вектора $x = (x_1, \ldots, x_n)$, удовлетворяющие системе ограничений

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

и доставляющие экстремум целевой функции

$$z = c_1x_1 + c_2x_2 + \ldots + c_nx_n.$$

Отметим, что

1) существует общий метод решения задачи линейного программирования, который называется симплекс-методом, разработаны стандартные программы для решения задачи на ЭВМ;

2) не каждая задача линейного программирования имеет допустимое решение, так как не всякая система уравнений или неравенств имеет решение;

3) не всегда задача линейного программирования имеет оптимальное решение, даже если существует допустимое решение;

4) если оптимальное решение существует, то оно достигается на границе допустимой области.

Несмотря на относительно простую математическую постановку, задачи ЛП довольно часто встречаются на практике, например при решении проблем, связанных с распределением ресурсов, планированием производства, организацией работы транспорта. Это достаточно естественно, поскольку во многих практических задачах «расходы» и «доходы» линейно зависят от количества закупленных или использованных средств (так, оплата перевозок может быть пропорциональна их длине и (или) весам перевозимых грузов и т. д.). Кроме того, линейные модели зачастую являются неплохим приближением при изучении задач, требующих использования методов НП — в отличие от НП в ЛП нет принципиальных затруднений при нахождении решения.

9.1.2. Некоторые прикладные задачи

Основной областью приложения линейного программирования являются экономические задачи.

1. Задача планирования производства

Пусть некоторое предприятие производит n типов товаров, затрачивая m типов ресурсов. При этом: a_{ij} — количество i-го ресурса, необходимое для производства единичного количества j-го товара, $a_{ij} \geqslant 0$ $(i = 1, \ldots, m;\ j = 1, \ldots, n)$, b_i — запас i-го ресурса на предприятии, c_j — цена единичного количества j-го товара.

Предполагается, что технология производства линейна, т. е. затраты ресурсов растут прямо пропорционально объему производства. Пусть число x_j показывает планируемый объем производства j-го товара. Тогда допустимым является только такой набор производимых товаров $(x_1, \ldots, x_n)$, при котором суммарные затраты каждого i-го ресурса не превосходят его запаса:

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i, \quad i = 1, \ldots, m.$$

При этом имеем естественные ограничения

$$x_j \geqslant 0, \quad j = 1, \ldots, n.$$

Стоимость набора товаров выражается величиной

$$z = \sum_{j=1}^{n} c_j x_j.$$

Требуется максимизировать доход при указанных ограничениях.

2. Задача о рационе

Пусть имеется n продуктов питания (хлеб, мясо, молоко, картофель, ...), в которых учитывается m полезных веществ (жиры, белки, углеводы, витамины, ...). Известны следующие параметры: a_{ij} — содержание i-го вещества в единичном количестве j-го продукта, $a_{ij} \geqslant 0$, $(i = 1, \ldots, m;\ j = 1, \ldots, n)$, b_i — минимальное количество i-го вещества, которое должно потребляться индивидуумом в расчете, скажем, на месяц, $b_i > 0$, c_j — цена единичного количества продукта, $c_j > 0$, x_j — количество j-го продукта, потребляемого индивидуумом в течение месяца.

Среди всех рационов питания $x = (x_1, \ldots, x_n)$, покрывающих минимальные потребности индивидуума в полезных веществах, необходимо выбрать наиболее дешевый. Задача о рационе формулируется следующим образом:

$$\sum_{j=1}^{n} c_j x_j \to \min, \quad \sum_{j=1}^{n} a_{ij} x_j \geqslant b_i, \quad i = 1, \ldots, m,$$

$$x_j \geqslant 0, \quad j = 1, \ldots, n.$$

Аналогичен пример о составе шихты. Известно, что для получения легированной стали нужно использовать шихту определенного химического состава. Многие ингредиенты шихты весьма дорогостоящие. В то же время в состав шихты входят малоценные материалы — чугун, лом, отходы с определенным известным содержанием присадок. Возникает задача о выборе такой шихты, в состав которой входили бы в заданных количествах необходимые химические вещества, а стоимость ее была бы минимальной.

Еще один пример задачи о смеси. Бензины разных сортов получают путем смешивания нефтепродуктов, имеющих различные технические характеристики. Заданные показатели качества бензина (октановое число, степень очистки и др.) должны выдерживаться очень точно, так как они играют важную роль для потребителя. От того, какие нефтепродукты при этом смешиваются, зависит рентабельность производства. Требуется построить такой план смешивания нефтепродуктов, который обеспечивал бы максимальную рентабельность производства и позволял в то же время получать бензины заданных сортов в нужных пропорциях.

3. Задача о смешивании двух растворов

Приведем математическую постановку задачи, рассмотренной в начале главы.

Пусть количество первого стандартного раствора, необходимого для приготовления требуемого раствора, имеет объем V_1, а второго — объем V_2. Поскольку в объеме V_1 первого раствора содержится a_1V_1 вещества A и b_1V_1 вещества B, а в объеме V_2 второго раствора содержится a_2V_2 вещества A и b_2V_2 вещества B, на основании заданных ограничений на количества веществ A и B в окончательном растворе можно составить следующие неравенства:

$$a_1V_1 + a_2V_2 \geqslant \alpha, \quad b_1V_1 + b_2V_2 \geqslant \beta,$$

где объемы V_1 и V_2, естественно, не могут быть отрицательными, т. е.

$$V_1 \geqslant 0, \quad V_2 \geqslant 0.$$

Общая стоимость C приготовленного раствора, равная суммарной стоимости используемых стандартных растворов, должна быть минимальна, откуда получаем условие

$$C = c_1V_1 + c_2V_2 = \min .$$

9.1.3. Геометрическая интерпретация задачи линейного программирования

Задачу линейного программирования с двумя переменными x_1 и x_2

$$z = c_1x_1 + c_2x_2 \to \max$$

$$a_{11}x_1 + a_{12}x_2 \leqslant b_1$$

$$\dots\dots\dots\dots\dots$$

$$a_{m1}x_1 + a_{m2}x_2 \leqslant b_m$$

можно решить геометрически.

Введем на плоскости декартову систему координат с осями x_1, x_2. Изобразим допустимую область, задаваемую системой линейных неравенств. Каждое ограничение определяет на плоскости одну из двух частей (полуплоскостей), на которые прямая $a_1x_1+a_2x_2 = b$ разбивает плоскость. При этом соответствующая полуплоскость включает и граничную прямую. Для уточнения полуплоскости достаточно подставить в неравенство координаты какой-нибудь точки, не лежащей на этой прямой. Если неравенство удовлетворяется, то искомая полуплоскость та, в которой лежит взятая точка, а если не удовлетворяется — то противоположная ей.

Пример. Найти полуплоскость, определяемую неравенством: $4x_1 - 6x_2 \leqslant 12$. Построим прямую $4x_1 - 6x_2 = 12$. Для этого составим таблицу с координатами точек пересечения прямой с осями.

x_1	0	3
x_2	−2	0

Так как прямая не проходит через начало координат (свободный член не равен 0), то для определения полуплоскости достаточно взять точку $0(0,0)$. Подставив ее координаты в неравенство, получаем верное соотношение $0 \leqslant 3$. Значит, определяемая полуплоскость та, которая содержит точку 0. Допустимое множество — это те точки плоскости, координаты которых удовлетворяют всем неравенствам — ограничениям, т. е. принадлежат пересечению полуплоскостей, определяемых отдельными нера-

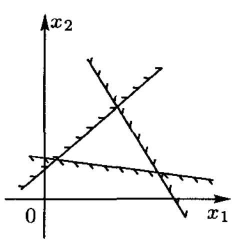


Рис. 9.1. Пустая допустимая область.

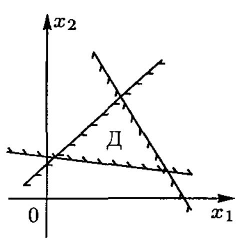


Рис. 9.2. Ограниченная допустимая область.

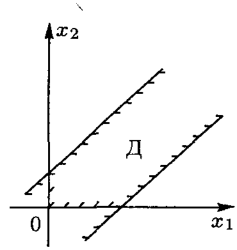


Рис. 9.3. Неограниченная допустимая область.

венствами. Допустимая область может быть пустой (рис. 9.1), непустой и ограниченной (рис. 9.2), непустой и неограниченной (рис. 9.3). Допустимая область, если она непуста, представляет собой выпуклый многоугольник.

Предположим, что допустимая область существует и мы ее построили. Чтобы найти оптимальное решение, т. е. выбрать из допустимой области точки, доставляющие максимальное значение целевой функции, изобразим геометрически условие: $z \to$ $\to \max$.

1. Положим сначала $z = 0$ и построим прямую по уравнению

$$c_1x_1 + c_2x_2 = 0.$$

Она проходит через начало координат. Назовем ее *опорной прямой.*

2. Построим вектор нормали $\overline{c} = (c_1, c_2)$. Направление этого вектора задает направление возрастания целевой функции. Если z будет принимать значения $z_1, z_2, \ldots$, то опорная прямая будет перемещаться параллельно самой себе и образует семейство линий уровня.

3. Будем перемещать опорную прямую в направлении $\overline{c}$ параллельно самой себе до тех пор, пока она не выйдет из допустимой области. Последняя точка касания этой прямой с допустимой областью будет являться оптимальным планом. Обозначим его $x^0 = (x_1^0, x_2^0)$.

4. Вычислим значение z в точке x^0 — это значение и будет максимальным значением целевой функции.

На рис. 9.4 точка $A(x_1^0, x_2^0)$ доставляет целевой функции максимум. Заметим, что max достигается в одной из вершин многоугольника.

В случае, когда целевая функция не ограничена в допустимой области сверху, оптимального решения не существует (рис. 9.5). Действительно, как долго мы ни перемещали бы опорную прямую в указанном направлении, она всегда будет иметь непустое пересечение с допустимой областью.

Рассмотрим случай, когда оптимальное решение существует, но не единственно, когда $\max z$ достигается не в одной точке A, а на целом отрезке AB, параллельном опорной прямой. Но

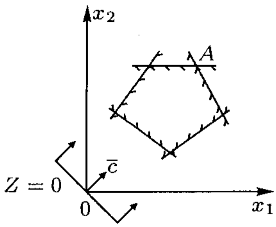


Рис. 9.4. $A(x_1^0, x_2^0)$ — точка max целевой функции.

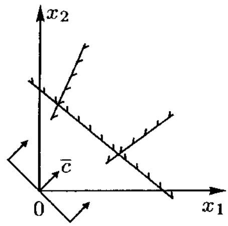


Рис. 9.5. Отсутствие оптимального плана.

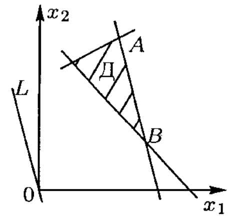


Рис. 9.6. Бесконечное множество оптимальных решений ($L||AB$).

и здесь можно считать, что max достигается в какой-то из вершин допустимой области (A или B — безразлично) (см. рис. 9.6).

Пример решения практической задачи

Фирма производит 2 продукта A и B, рынок сбыта которых неограничен. Каждый продукт должен быть обработан каждой из машин 1, 2, 3 типа. Время обработки в часах для каждого из изделий A и B дано в таблице:

	1	2	3
A	0,5	0,4	0,2
B	0,25	0,3	0,4

Время работы машин 40, 36, 36 часов в неделю соответственно. Прибыль от изделий A и B составляет соответственно 5 и 3 руб. Фирме надо определить недельные нормы выпуска изделий A и B, максимизирующие прибыль.

Решение. Пусть x_1 — количество выпускаемого продукта A, x_2 — количество продукта B. Требуется максимизировать целевую функцию $z = 5x_1 + 3x_2 \to \max$ при ограничениях:

$$0{,}5x_1 + 0{,}25x_2 \leqslant 40, \quad 0{,}4x_1 + 0{,}3x_2 \leqslant 36,$$

$$0{,}2x_1 + 0{,}4x_2 \leqslant 36, \quad x_1, x_2 \geqslant 0.$$

Построим допустимую область, ограниченную прямыми l_1, l_2, l_3 и осями координат x_1, x_2.

$$l_1 : 2x_1 + x_2 = 160, \quad l_2 : 4x_1 + 3x_2 = 360, \quad l_3 : x_1 + 2x_2 = 180.$$

Строим вектор $\overline{c} = (5, 3)$ и опорную прямую $z \perp \overline{c}$. Перемещая прямую z по направлению $\overline{c}$, заметим, что последний раз опорная прямая пересечет допустимую область, когда пройдет через точку A, координаты которой можно найти, решив систему уравнений:

$$\begin{cases} 2x_1 + x_2 = 160, \\ 4x_1 + 3x_2 = 360, \end{cases}$$

$$x_2 = 160 - 2x_1, \quad 4x_1 + 480 - 6x_1 = 360, \quad 2x_1 = 120,$$

$$x_1^0 = 60, \quad x_2^0 = 40.$$

Значение целевой функции $z(60, 40) = 5 \cdot 60 + 3 \cdot 40 = 420$.

Иногда требуется заменить задачу отыскания max эквивалентной ей задачей на min. Известно, что $\max z$ и $\min(-z)$ до-

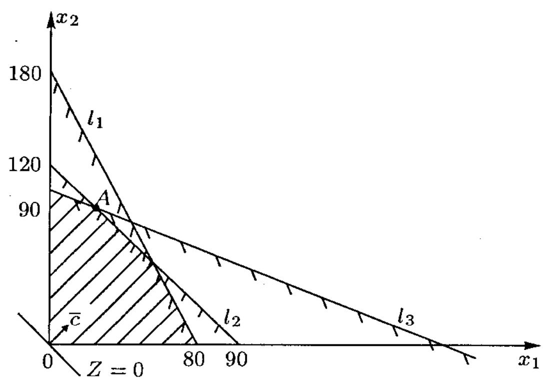


Рис. 9.7. Графическое решение задачи.

стигаются при одних и тех же значениях переменных. Поэтому можно рассматривать и задачу: $z = f(x_1, x_2) \to \min$. Тогда при решении графическим методом алгоритм, описанный в п. 3. сохраняется, но опорную прямую нужно перемещать в направлении, противоположном вектору нормали, т. е. в направлении вектора $\overline{c} = (-c_1, -c_2)$.

Самостоятельная работа

1. По данным справочника дети должны получать за завтраком по крайней мере 1 мг тиамина, 5 мг ниацина и 400 калорий. Для удовлетворения этой потребности хозяйка располагает двумя видами крупяных изделий K и C. Можно выбрать либо первый вид, либо второй, либо смесь обоих видов. Этикетки на каждой пачке сообщают, что 1 унция K содержит 0,1 мг тиамина, 1 мг ниацина и 110 кал., в то время как 1 унция C содержит 0,25 мг тиамина, 0,25 мг нацина и 120 кал. Составить так меню завтрака, чтобы дети получили необходимое количество питательных веществ при минимальных затратах, если известно, что унция K стоит 3,8 цента, а унция C — 4,2 цента.

2. Чаеразвесочная фабрика выпускает чай сорта A и B, смешивая 3 ингредиента: индийский, грузинский и краснодарский чай. В таблице приведены нормы расхода ингредиентов, объем запасов каждого ингредиента и прибыль от реализации 1 т чая сорта A и B.

Требуется составить план производства чая сорта A и B с целью максимизации суммарной прибыли.

Ингредиенты	Нормы расхода (т)		Объем запасов
	A	B	
индийский чай	0,5	0,2	600
грузинский чай	0,2	0,6	870
краснодарский чай	0,3	0,2	430
прибыль от 1 т. продукции (дол.)	320	290	

3. Для изготовления двух видов продукции P_1, P_2 используют 3 вида сырья S_1, S_2, S_3. Запасы сырья, количество единиц сырья, затрачиваемых на изготовление единицы продукции, а также величина прибыли, получаемой от реализации продукции, приведены в таблице:

Вид сырья	Количество ед. сырья для изготовления ед. продукции		Запас сырья
	P_1	P_2	
S_1	2	5	20
S_2	8	5	40
S_3	5	6	30
Прибыль от ед. продукции	50	40	

Спланировать выпуск продукции таким образом, чтобы получить максимальную прибыль от реализации продукции.

4. Цех предприятия выпускает продукцию двух видов: A и B. Трудоемкость обработки каждого изделия, а также полезный фонд времени даются в таблице:

Машины	Продукт A	Продукт B	Полезный фонд времени
1	2	3	6000
2	1	1	4000
Доход от реализации 1 ед. пр. (дол.)	50	40	

Необходимо разработать план работы предприятия на неделю, позволяющий максимизировать доход.

5. Средства очистки пола оценивают по следующим трем показателям: а) очищающие свойства, б) дезинфицирующие свойства, в) раздражающее воздействие на кожу. Каждый из этих показателей измеряется по линейной шкале от 0 до 100 единиц.

Продукт на рынке должен иметь по крайней мере 60 единиц очищающих свойств и по крайней мере 60 единиц дезинфицирующих свойств по соответствующей шкале. При этом раздражающее воздействие на кожу должно быть минимальным. Конечный продукт должен быть смесью трех основных очистителей, характеристики которых приводятся в таблице.

Очиститель	Очищающие свойства	Дезинфицирующие свойства	Раздражающее воздействие на кожу
A	90	30	70
B	65	85	50
C	45	70	10

Требуется создать продукт с минимальным раздражающим воздействием на кожу.

§ 9.2. Нелинейное программирование

Класс задач нелинейного программирования значительно шире класса задач линейного программирования. Это связано с тем, что большинство реальных задач не может быть адекватно описано с помощью моделей линейного программирования из-за нелинейности целевой функции или ограничений.

Отметим, что:

1) общих методов решения задачи нелинейного программирования нет;

2) основные результаты в нелинейном программировании получены при рассмотрении задач, в которых система ограничений линейна, а целевая функция нелинейная, но даже в таких задачах оптимальное решение может быть найдено только для узкого класса целевых функций;

3) если в задачах линейного программирования точки экстремума являются вершинами допустимой области, то в задачах с нелинейной целевой функцией они могут лежать и внутри области;

4) допустимое множество не обязательно выпукло;

5) если и ограничения, и целевая функция нелинейны, то найти оптимальное решение возможно только при особых условиях.

9.2.1. Некоторые прикладные задачи

1. Задача планирования производства

Рассматривая задачу о планировании производства, следует отметить, что величина дохода от единицы произведенной продукции на самом деле не является постоянной, а зависит от объема производства: по мере освоения технологии производства продукции затраты уменьшаются, с увеличением числа произведенных продуктов спрос на них падает. Поэтому целевая функция, описывающая доход, является нелинейной функцией.

2. Задача о капиталовложениях

Имеется n различных отраслей, производящих продукцию. Эффективность капиталовложений в j-ю $(j = 1, \dots, n)$ отрасль определяется известной функцией дохода $q_j(x_j)$, зависящей от величины x_j. Необходимо распределить ресурсы величины c между отраслями таким образом, чтобы максимизировать суммарный доход. Если $q_j(x_j)$ — нелинейная функция, то мы имеем задачу отыскания экстремума нелинейной функции:

$$\max f(x) = \sum_{j=1}^{n} q_j(x_j),$$

$$\sum_{j=1}^{n} x_j \leqslant c, \quad x_j \geqslant 0, \quad j = 1, \ldots, n.$$

3. Задача планирования строительства

Строительная организация выполняет заказы по возведению n различных промышленных сооружений. Сооружение j-го ($j = 1, \ldots, n$) типа может быть построено по одной из технологических схем. Возведение j-го здания по s-й ($s = 1, \ldots, k$) технологической схеме приводит к расходу l-го ресурса a^l_{js}, $l = 1, \ldots, m$, и позволяет получить строительной организации доход величины $c(x_{js})$. Требуется выбрать такие схемы возведения промышленных зданий, которые позволяют максимизировать прибыль строительной организации, располагающей ресурсами l-го вида в количестве b_l единиц, $l = 1, \ldots, m$.

Обозначим $x_{js} = 1$ — факт строительства j-го здания по s-й схеме (если $x_{js}=0$, то s-я схема для строительства j-го здания не выбирается). Требуется найти:

$$\max f(x) = \sum_{j=1}^{n} \sum_{s=1}^{k} c(x_{js}),$$

$$\sum_{j=1}^{n} \sum_{s=1}^{k} a^l_{js} x_{js} \leqslant b_l, \quad l = 1, \ldots, m,$$

$$x_{js} = \begin{cases} 1, \\ 0, \quad j = 1, \ldots, n, \quad s = 1, \ldots, k. \end{cases}$$

4. Задача складирования

Пусть в каждом из n периодов известен спрос $d_1, \ldots, d_n$ на некоторый продукт. Для удовлетворения спроса в j-м периоде может быть использован как произведенный в этом периоде, так и хранимый на складе продукт. В j-м периоде выпускается x_j единиц продукта. При этом производственные затраты задаются функцией $f(x_j)$, $j = 1, \ldots, n$. Выпущенный в некотором периоде в избытке продукт можно хранить на складе. Емкость склада равна K, а стоимость хранения единицы продукции в течение одного периода равна c рублям. Считая заданным начальный запас I_0, хранимый на складе, задачу планирования производства можно сформулировать как задачу нелинейного программирования.

Задачи оптимизации, связанные с оптимальным управлением, выбором «портфеля ценных бумаг», проектированием строительных, механических конструкций, управлением водными ресурсами, размещением оборудования, экологического контроля за окружающей средой и т. д., могут быть формально представлены в виде моделей нелинейного программирования.

9.2.2. Решение задач нелинейного программирования геометрическим методом

Обычно задачу нелинейного программирования пытаются свести к классической задаче на условный экстремум о минимизации функции $z(x_1, \ldots, x_n)$ при ограничениях в форме равенств, которую решают, используя критерии выпуклости (вогнутости) функций и теоремы об экстремумах или используя метод множителей Лагранжа.

В простейшем случае, когда целевая функция z и функции ограничений зависят лишь от двух переменных x_1, x_2 , задача нелинейного программирования может быть решена геометрическим методом.

Пример 1. Случай, когда целевая функция нелинейна, а функции ограничений — линейны.

Найти $\min$ и $\max$ целевой функции $z = (x_1 + 1)^2 + (x_2 - 2)^2$ при ограничениях:

$$\begin{cases} x_1 + x_2 \geqslant 1, \\ 2x_1 + 3x_2 \leqslant 12, \\ x_1, x_2 \geqslant 0. \end{cases}$$

Решение.

1) Построим в системе координат $(x_1 0 x_2)$ допустимую область. Уравнения прямых, образующих границу области, имеют

вид

а) $x_1 + x_2 = 1$; б) $2x_1 + 3x_2 = 12$; в) $x_1 = 0$; г) $x_2 = 0$.

Каждый раз, выбирая нужную полуплоскость, задаваемую соответствующим неравенством, и далее взяв пересечение этих областей, получим

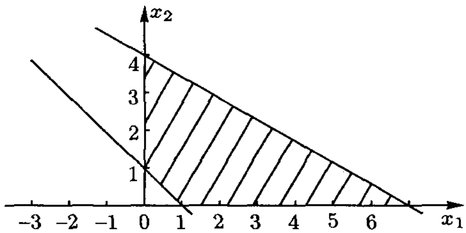


Рис. 9.8. Допустимая область к примеру 1.

2) Сравним целевую функцию из условия задачи $z = (x_1 + 1)^2 + (x_2 - 2)^2$ и уравнение окружности: $(x_1+1)^2+(x_2-2)^2 = R^2$ (окружность с центром $(-1, 2)$ и радиусом R). Заметим, что целевую функцию можно рассматривать как квадрат радиуса окружности. Поэтому $z \to \min$, если $R \to \min$, а радиус окружности минимален, если окружность стягивается в точку — центр окружности. Итак, $z \to \min$, если $x_1 = -1$, $x_2 = 2$. Однако мы видим, что точка $(-1, 2)$ не принадлежит допустимой области. Поэтому будем постепенно «раздувать» окружность, увеличивая ее радиус, пока не коснемся допустимой области. Нанесем на чертеж эти окружности с различными радиусами, представляющие

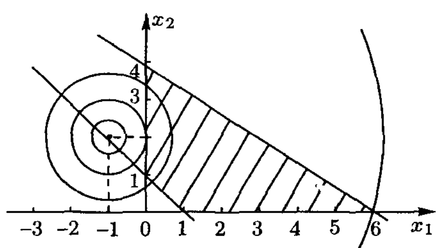


Рис. 9.9. Графическое решение примера 1.

собой линии уровня целевой функции. Как видно из рис. 9.9, это касание осуществится в точке $(0, 2)$: $z_{\min} = (0+1)^2+(2-2)^2 = 1$.

Далее, ищем $z \to \max$, продолжая увеличивать радиус R. Следует помнить о том, что радиус можно увеличивать не до бесконечности, а лишь до тех пор, пока мы будем находиться в допустимой области. Последняя точка, в которой произойдет пересечение линии уровня целевой функции с допустимой областью, будет точка $(6, 0)$. Поэтому $z_{\max} = (6+1)^2 + (0-2)^2 = 53$.

Пример 2. Случай, когда целевая функция z линейна, а функции ограничений — нелинейны.

Найти $z = 3x_1 + x_2 \to \max$ при ограничениях:

$$\begin{cases} x_1 \cdot x_2 \geqslant 2, \\ x_1^2 + x_2^2 \leqslant 16. \end{cases}$$

1) Построим допустимое множество решений. Уравнение $x_1 \cdot x_2 = 2$ задает гиперболу, строим ее по точкам. Уравнение $x_1^2 + x_2^2 = 16$ — уравнение окружности с центром в начале

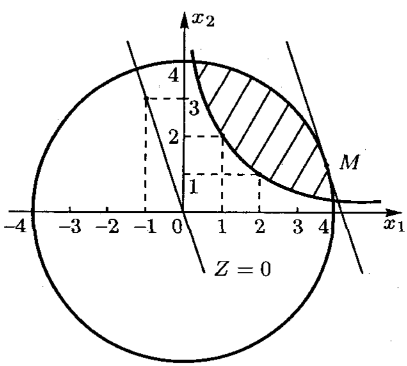


Рис. 9.10. Графическое решение примера 2.

координат и радиусом, равным 4. Искомая допустимая область на рис. 9.10 заштрихована.

2) Рассмотрим целевую функцию $z = 3x_1 + x_2$, которая представляет собой уравнение прямой, задающее линии уровня целевой функции. Построим опорную прямую при $Z = 0$: $3x_1 + x_2 = 0$.

С увеличением Z имеем $z \to \max$, поэтому будем перемещать опорную прямую параллельно самой себе до тех пор, пока не произойдет последнее касание линии уровня с допустимой областью. Это будет точка M. Для определения координат этой точки рассмотрим систему уравнений:

$$\begin{cases} z = 3x_1 + x_2, \\ x_1^2 + x_2^2 = 16. \end{cases}$$

Из второго уравнения выразим x_2 и, подставив в первое, получим:

$$x_2 = \sqrt{16 - x_1^2}, \quad z = 3x_1 + \sqrt{16 - x_1^2}.$$

Функция z стала функцией лишь одной переменной, x_1, и можно говорить об ее экстремуме, найдя $\frac{dz}{dx_1}$ и приравнивая эту производную нулю (по необходимому условию существования экстремума).

$$\begin{cases} \dfrac{dz}{dx_1} = 3 + \dfrac{(-2x_1)}{2\sqrt{16 - x_1^2}} = 0, \\ 3\sqrt{16 - x_1^2} = x_1. \end{cases}$$

Возведем обе части уравнения в квадрат, выполним преобразования и получим: $10x_1^2 = 9 \cdot 16$, откуда находим $x_1 = \pm\frac{3 \cdot 4}{\sqrt{10}}$. Однако, из рис. 9.10 видно, что точка M находится в первом квадранте, следовательно, $x_1 \geqslant 0$, и поэтому

$$x_1 = \frac{12}{\sqrt{10}}, \qquad x_2 = \sqrt{16 - \frac{9 \cdot 16}{10}} = \frac{4}{\sqrt{10}}.$$

Итак, max достигается в точке $\left(\frac{12}{\sqrt{10}}, \frac{4}{\sqrt{10}}\right)$. Следовательно,

$$z_{\max} = 3x_1 + x_2|_M = \frac{36}{\sqrt{10}} + \frac{4}{\sqrt{10}} = 4\sqrt{10}.$$

Пример 3. Случай, когда и целевая функция, и функции ограничений нелинейны.

Найти $z = x_1^2 + x_2^2$ при ограничениях:

$$\begin{cases} x_1 \cdot x_2 \leqslant 4, \\ x_1 + x_2 \geqslant 5, \\ x_1 \leqslant 7, \\ x_2 \leqslant 6, \\ x_1, x_2 \geqslant 0. \end{cases}$$

Решение.

1) Построим допустимую область (заштрихована на рис. 9.11) $ABCD$ и $A_1B_1C_1D_1$, а также линии уровня целевой функции;

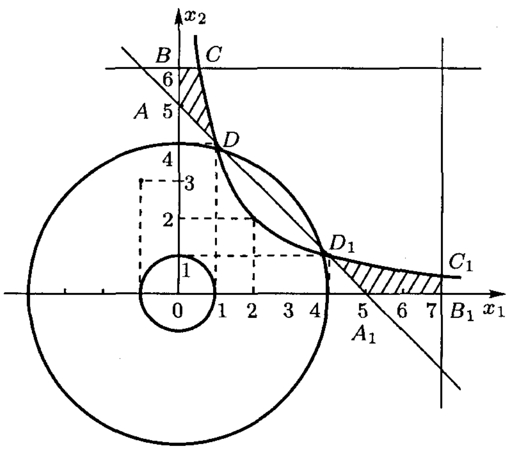


Рис. 9.11. Графическое решение примера 3.

$z \to \max$, если $R \to \max$ (из рис. 9.11 это точка C_1),

$z \to \min$, если $R \to \min$ (из рис. 9.11 это точки D_1, D_2).

Найдем координаты этих точек. Точка C_1 является точкой пересечения прямой $x_1 = 7$ и гиперболы. Решим систему уравнений:

$$\begin{cases} x_1 = 7, \\ x_1 \cdot x_2 = 4, \end{cases}$$

откуда $C_1(7,\ 4/7)$.

Точки D_1, D_2 — точки пересечения гиперболы и прямой $x_1 + x_2 = 5$.

Решаем систему уравнений:

$$\begin{cases} x_1 + x_2 = 5, \\ x_1 \cdot x_2 = 4, \end{cases}$$

откуда $x_1 = 5 - x_2$, а значит, из второго уравнения системы получим $x_2^2 - 5x_2 + 4 = 0$ и тогда $x_2^{(1)} = 1$; $x_2^{(2)} = 4$. Следовательно, $x_1^{(1)} = 4$; $x_1^{(2)} = 1$ и $D(1,4)$, $D_1(4,1)$, $z_{\min} = z|_D = z|_{D1} = 17$.

Если возникает сомнение, в какой именно точке достигается, например, max — в точке B_1 или C_1, то вычисляем координаты этих точек и значения целевой функции, а затем выбираем наибольшее из них.

Самостоятельная работа

Найти экстремальное значение z при указанных ограничениях:

1. $z = x_1 + x_2 \to \max, \min$ $\quad \begin{cases} x_1^2 + x_2^2 \leqslant 9, \\ x_1 + 2x_2 \geqslant 2. \end{cases}$

2. $z = x_1 + x_2 \to \max, \min$ $\quad \begin{cases} (x_1 - 2)^2 + (x_2 - 3)^2 \leqslant 4, \\ x_1 + 4x_2 \leqslant 16. \end{cases}$

3. $z = 3x_1 + x_2 \to \max, \min$ $\quad \begin{cases} x_1 \cdot x_2 \geqslant 2, \\ x_1^2 + x_2^2 \leqslant 16. \end{cases}$

4. $z = 2x_1 + x_2 \to \max, \min$ $\quad \begin{cases} x_1 \cdot x_2 \leqslant 4, \\ 3x_1 + x_2 \leqslant 3, \\ x_1, x_2 \geqslant 0. \end{cases}$

5. $z = 2x_1 - x_2 \to \max, \min$ $\quad \begin{cases} x_1^2 \leqslant x_2 - 2, \\ 4x_1 + x_2 \leqslant 12. \end{cases}$

6. $z = (x_1 - 5)^2 + (x_2 - 4)^2 \to \max, \min$ $\quad \begin{cases} x_1 + x_2 \geqslant 1, \\ x_1 + 3x_2 \leqslant 9, \\ x_1, x_2 \geqslant 0. \end{cases}$

7. $z = x_1^2 + x_2^2 \to \max, \min$ $\quad \begin{cases} x_1 \cdot x_2 \leqslant 4, \\ x_1 + x_2 \geqslant 5, \\ 0 \leqslant x_2 \leqslant 6, \\ 0 \leqslant x_1 \leqslant 7. \end{cases}$

8. $z = x_1^2 + x_2^2 \to \max, \min$ $\quad \begin{cases} x_2 - x_1^2 \leqslant 2, \\ x_1 \geqslant 1, \\ x_2 \leqslant 5. \end{cases}$

9. $z = (x_1 - 1)^2 + (x_2 - 1)^2 \to \max, \min$ $\quad \begin{cases} x_1 + x_2 \geqslant 2, \\ x_1^2 + x_2^2 \leqslant 9, \\ x_1 \leqslant 2. \end{cases}$

10. $z = x_1^2 + x_1^2 \to \max, \min$ $\quad \begin{cases} (x_1 - 12)^2 + (x_2 - 12)^2 \leqslant 144, \\ x_1 + x_2 \leqslant 15, \\ x_2 \geqslant 5. \end{cases}$

§ 9.3. Транспортная задача линейного программирования

9.3.1. Математическая постановка транспортной задачи

Одним из важных частных типов задач линейного программирования является так называемая *транспортная задача* (ТЗ). Она может быть описана следующим образом: имеется m пунктов отправления $A_1, A_2, \ldots, A_m$, в которых сосредоточены запасы какого-то однородного груза в количестве соответственно a_1,

$a_2, \ldots, a_m$ единиц. Весь груз следует доставить в n пунктов назначения $B_1, B_2, \ldots, B_n$ в количествах $b_1, b_2, \ldots, b_n$ единиц соответственно. При этом сумма всех запасов в пунктах отправления равна суммарной потребности во всех пунктах назначения, т. е.

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j. \tag{9.3}$$

Предполагается, что каждый пункт отправления связан с каждым пунктом назначения некоторым маршрутом. Заданы стоимости перевозки c_{ij} единицы груза от пункта отправления A_i до пункта назначения B_j $(i = 1, \ldots, m;\ j = 1, \ldots, n)$.

Стоимость перевозки по любому маршруту пропорциональна количеству перевозимых единиц груза.

Требуется составить такой план перевозок (откуда, куда и сколько единиц груза везти), чтобы потребности пунктов назначения были полностью удовлетворены, а общая стоимость всех перевозок была минимальной.

Формализуем вышесказанное. Обозначим x_{ij} — количество единиц груза, отправляемого из пункта отправления A_i в пункт назначения B_j. Эти неотрицательные переменные, которые называются «планом перевозок», можно записать в виде таблицы (x_{ij}):

$$\begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix}. \tag{9.4}$$

Каким же условиям перевозки x_{ij} должны удовлетворять?

1. Из каждого пункта отправления вывозится весь имеющийся там груз. Это дает m равенств:

$$\begin{gathered} x_{11} + x_{12} + \ldots + x_{1n} = a_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{m1} + x_{m2} + \ldots + x_{mn} = a_m. \end{gathered} \tag{9.5}$$

2. Потребности каждого пункта назначения полностью удовлетворяются. Это дает n равенств:

$$\begin{gathered}x_{11}+x_{21}+\ldots+x_{m1}=b_1,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ x_{1n}+x_{2n}+\ldots+x_{mn}=b_n.\end{gathered} \tag{9.6}$$

3. Перевозки x_{ij} должны быть неотрицательными:

$$\begin{gathered}x_{ij}\geqslant 0\\ (i=1,\ldots,m;\ \ j=1,\ldots,n),\end{gathered} \tag{9.7}$$

4. Суммарная стоимость всех перевозок, т. е. сумма величин x_{ij}, умноженных на соответствующие стоимости c_{ij}, должна быть минимальной:

$$L=\sum_{i=1}^{m}\sum_{j=1}^{n}c_{ij}x_{ij}\Rightarrow \min\,. \tag{9.8}$$

Итак, перед нами — задача ЛП с условиями-равенствами (9.5), (9.6) и минимизируемой линейной функцией (9.8), постановка которой выглядит следующим образом: минимизировать по x_{ij} $(i=1,\ldots,m;\ j=1,\ldots,n)$ сумму $\sum\limits_{i,j}c_{ij}x_{ij}$ при ограничениях (9.5)–(9.7).

Специфика этой задачи — все коэффициенты в условиях (9.5), (9.6) равны единице, каждое неизвестное входит лишь в два уравнения — позволяет существенно упростить процесс ее решения, не прибегая к общим методам решения задач ЛП.

Следует сказать, что круг практических ситуаций, приводящих к ТЗ, не обязан иметь отношения к перевозкам. Так, например, пункты назначения можно интерпретировать как различные виды каналов обслуживания, а пункты отправления — как различные классы заявок. Каждое число b_j при этом показывает, сколько каналов содержит данный вид, а числа a_i — сколько заявок класса i имеется. Числа c_{ij} могут характеризовать время или любые другие издержки обслуживания заявки i-го класса каналом обслуживания j-го вида. Целью же решения задачи является распределение заявок между каналами обслуживания, минимизирующее суммарные издержки.

9.3.2. Открытая модель транспортной задачи

Так называется ТЗ с нарушенным балансом запасов $\sum\limits_{i=1}^{m} a_i$ и потребностей $\sum\limits_{j=1}^{n} b_j$, когда условие (9.3) не выполняется. Подобные задачи возникают при определении наиболее рационального размещения подлежащих строительству производственных мощностей, что является частным случаем так называемой задачи размещения.

С формальной точки зрения открытая модель ТЗ несложно сводится к обычной замкнутой модели. Пусть для определенности

$$\delta = \sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j > 0.$$

Введем фиктивный $(n+1)$-й пункт назначения B_{n+1}, для которого положим $b_{n+1} = \delta$ и $c_{i,n+1} = 0$ $(i = 1, \ldots, m)$. В результате придем к ТЗ с m пунктами отправления и $(n+1)$-м пунктом назначения, в которой запасы груза и потребности в нем сбалансированы. Можно показать, что оптимальный план перевозок из пунктов отправления в пункты назначения с номерами $j = 1, \ldots, n$ в этой обычной ТЗ совпадает с оптимальным планом перевозок исходной задачи. Груз же, формально вывозимый в фиктивный $(n+1)$-й пункт назначения должен оставаться в соответствующих пунктах отправления.

Итак, чтобы решить открытую модель транспортной задачи в случае (9.8), достаточно найти решение задачи, полученной введением фиктивного пункта назначения, а груз, запланированный к перевозке в этот пункт, не перевозить.

9.3.3. Решение транспортной задачи

Для решения ТЗ было разработано несколько методов, максимально учитывающих ее специфику. Мы остановимся на одном из них — *методе потенциалов*.

В силу особой структуры ТЗ все операции по ее решению сводятся к манипуляциям с таблицей вида 9.1, где в определенном порядке записаны условия ТЗ: номера пунктов отправления (i) и пунктов назначения (j), запасы a_i, заявки b_j, и стоимости перевозок c_{ij}.

Таблица 9.1

№ пункта отправ. (i)	№ пункта назначения (j)					Запасы a_i
	1	2	3	4	5	
1	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	a_1
2	c_{21}	c_{22}	c_{23}	c_{24}	c_{25}	a_2
3	c_{31}	c_{32}	c_{33}	c_{34}	c_{35}	a_3
4	c_{41}	c_{42}	c_{43}	c_{44}	c_{45}	a_4
Заявки b_j	b_1	b_2	b_3	b_4	b_5	

В эту же таблицу по мере ее заполнения вписываются величины перевозок x_{ij}. Таблица состоит из m строк и n столбцов. В правом верхнем углу каждой клетки (всего их $m \cdot n$) проставляется соответствующая ей стоимость перевозки c_{ij}. В левом нижнем углу в дальнейшем помещается сама перевозка x_{ij}.

Назовем *циклом* совокупность клеток таблицы, устроенную таким образом, что каждая пара соседних клеток расположена либо в одной строке, либо в одном столбце, причем никакие три клетки не лежат в одной строке (столбце), а последняя клетка расположена в одном столбце (строке) с первой.

Так, в табл. 9.1 изображен цикл $(1,2)$, $(1,4)$, $(3,4)$, $(3,5)$, $(2,5)$, $(2,2)$; совокупность клеток $(1,4)$, $(1,2)$, $(3,2)$, $(3,5)$, $(2,5)$, $(2,4)$ также является циклом.

Будем называть план перевозок (9.4) *опорным*, если он удовлетворяет условиям (9.5)–(9.7) и если совокупность клеток с элементами x_{ij} этого плана не содержит циклов.

Можно показать, что число N положительных элементов любого опорного плана удовлетворяет неравенствам

$$\max(n, m) \leqslant N \leqslant n + m - 1$$

и что оптимальное решение (*оптимальный план*) ТЗ, т. е. план перевозок x_{ij}, удовлетворяющий условию (9.8), обязательно будет опорным. Другими словами, оптимальный план перевозок следует искать среди всех опорных планов.

Все дальнейшие рассуждения будем осуществлять на примере конкретной таблицы.

Таблица 9.2

№ пункта отправ. (i)	№ пункта назначения (j)				Запасы a_i
	1	2	3	4	
1	3	5	7	11	100
2	1	4	6	3	130
3	5	8	12	7	170
Заявки b_j	150	120	80	50	400

В соответствии с табл. 9.2 и условиями (9.3), (9.5), (9.6), имеем

$$a_1 = 100, \quad a_2 = 130, \quad a_3 = 170 \quad (m = 3);$$

$$b_1 = 150, \quad b_1 = 120, \quad b_3 = 80, \quad b_4 = 50 \quad (n = 4);$$

$$\sum_{i=1}^{3} a_i = \sum_{j=1}^{4} b_j (= 400),$$

$$c_{23} = 6, \quad c_{32} = 8$$

и т. д.

Первоначальный опорный план

Займемся составлением *первоначального опорного плана*. Для этой цели мы применим *метод минимального элемента*, в котором существенно учитываются стоимости перевозок c_{ij} и который, как правило, приводит к опорным планам, более близким к оптимальным, чем другие методы («северо-западного угла», например).

Суть метода минимального элемента заключается в том, что из всей таблицы выбирают клетку с минимальной стоимостью перевозки c_{ij} и помещают в нее перевозку $x_{ij} = \min(a_i, b_j)$.

Далее из рассмотрения исключают строку, если запас в i-м пункте отправления полностью израсходован, или столбец, если заявка в j-м пункте назначения полностью удовлетворена, или же и строку и столбец, если израсходованы запасы пункта отправления и потребности пункта назначения.

С оставшейся частью таблицы поступают аналогично, выбирая наименьшую стоимость перевозки и продолжая распределять оставшиеся запасы. Процесс заканчивается, когда все запасы полностью распределены, а потребности удовлетворены.

Применительно к данным табл. 9.2 сказанное выглядит следующим образом.

Находим в таблице элемент c_{ij} с минимальной стоимостью. Им будет элемент $c_{12} = 1$ из клетки $(2, 1)$, соответствующий второму пункту отправления с запасом груза $a_2 = 130$ и первому пункту назначения с заявкой $b_1 = 150$. Таким образом, в клетку $(2, 1)$ следует поместить число $x_{21} = \min(a_2, b_1) = \min(130, 150) = 130$ и исключить вторую строку из дальнейшего рассмотрения, имея при этом в виду, что в первый пункт назначения надо завести еще $\widetilde{b}_1 = b_1 - a_2 = 150 - 130 = 20$ единиц груза.

Таблица 9.2а

№ пункта отправ. (i)	№ пункта назначения (j)				Запасы a_i
	1	2	3	4	
1	3	5	7	11	100
2	1 130	–	–	–	–
3	5	8	12	7	170
Заявка b_j	20	120	80	50	

В табл. 9.2а наименьшая стоимость соответствует клетке $(1, 1)$: $c_{11} = 3$. Поскольку $\widetilde{b}_1 = b_1 = 20 < a_1 = 100$, в клетку $(1, 1)$ вписываем величину $x_{11} = 20$ и исключаем из рассмотрения столбец с номером 1, отмечая попутно, что $\widetilde{a}_1 = a_1 - \widetilde{b}_1 = 100 - 20 = 80$.

В табл. 9.2б минимальная стоимость $c_{12} = 5$ содержится в клетке $(1, 2)$ и, следовательно, в эту клетку необходимо поместить перевозку $x_{12} = \min(a_1, b_2) = \min(80, 120) = 80$, в дальнейшем не учитывая строку с номером 1.

Таблица 9.2б

№ пункта отправ. (i)	№ пункта назначения (j)				Запасы a_i
	1	2	3	4	
1	3 20	5	7	11	80
2	1 130	–	–	–	–
3	–	8	12	7	170
Заявка b_j	–	120	80	50	

Отметим, что во второй пункт назначения предстоит завести еще $\widetilde{b}_2 = b_2 - \widetilde{a}_1 = 120 - 80 = 40$ единиц груза. Следующей на очереди является клетка $(3, 4)$ (в нее завозим $x_{34} = \min(a_3, b_4) = \min(170, 50) = 50$ единиц), затем клетка $(3, 2)$ ($x_{32} = \min(a_3 - b_4, \widetilde{b}_2) = \min(170 - 50, 40) = 40$) и, наконец, клетка $(3, 3)$ с $x_{33} = 80$.

Таблица 9.2в

№ пункта отправ. (i)	№ пункта назначения (j)				Запасы a_i
	1	2	3	4	
1	3 20	5 80	–	–	–
2	1 130	–	–	–	–
3	–	8	12	7	170
Заявка b_j	–	40	80	50	

Занесем все найденные нами методом минимального элемента перевозки в табл. 9.3:

Таблица 9.3

№ пункта отправ. (i)	№ пункта назначения (j)				Запасы a_i
	1	2	3	4	
1	3 20	5 80	7	11	100
2	1 130	4	6	3	130
3	5	8 40	12 80	7 50	170
Заявка b_j	150	120	80	50	400

Является ли этот план опорным? Да, потому что в нем сумма перевозок по каждой строке равна запасу соответствующего пункта отправления, а сумма перевозок по столбцу-заявке соответствующего пункта назначения (значит, все заявки удовлетворены, все запасы израсходованы) и число перевозок x_{ij} в нашем плане равно $6 = m + n - 1$ $(m = 3, n = 4)$; кроме того, совокупность клеток, занятых планом, не содержит циклов.

Замечание. Если число N элементов плана, полученного методом минимального элемента, меньше $m + n - 1$ (тогда задача называется *вырожденной*), то в этот план следует ввести дополнительно $m + n - 1 - N$ новых клеток с перевозками $x_{ij} = 0$, но так, чтобы он не содержал циклов (см. определение опорного плана).

Здесь и в дальнейшем в таблице будут проставляться только величины перевозок, отвечающие опорному плану (в том числе, нулевые), а остальные клетки оставляться свободными.

Теперь пора подумать о том, является ли план из табл. 9.3 оптимальным, т. е. минимальна ли для него общая стоимость перевозок L (см. (9.8))?

Условия оптимальности плана перевозок

Имеет место следующий критерий оптимальности плана перевозок:

План перевозок $X = (x_{ij})$ является оптимальным тогда и только тогда, когда ему соответствует система из $m + n$ чисел $u_1, u_2, \ldots, u_m$; $v_1, v_2, \ldots, v_n$, удовлетворяющая условиям

$$v_j - u_i \leqslant c_{ij} \quad (i = 1, \ldots, m; \;\; j = 1, \ldots, n), \tag{9.9}$$

причем

$$v_j - u_i = c_{ij} \tag{9.10}$$

для всех $x_{ij} > 0$ $(x_{ij} \in X)$.

Числа u_i и v_j называются *потенциалами* соответственно пункта отправления и пункта назначения, а условия (9.9), (9.10) — условиями потенциальности системы $\{u_i, v_j\}$, что собственно, и привело к названию метода.

Итак, нам следует проверить план перевозок из табл. 9.3 на потенциальность. С этой целью определим систему потенциалов, соответствующих нашей задаче. Воспользовавшись условием (9.10) (напоминаем, что индекс i соответствует поставщикам, а индекс j — потребителям) для занятых клеток табл. 9.3, получим систему уравнений:

$$\begin{cases} v_1 - u_1 = c_{11} = 3, \\ v_1 - u_2 = c_{21} = 1, \\ v_2 - u_1 = c_{12} = 5, \\ v_2 - u_3 = c_{32} = 8, \\ v_3 - u_3 = c_{33} = 12, \\ v_4 - u_3 = c_{34} = 7. \end{cases} \tag{9.11}$$

Поскольку число неизвестных $m + n = 7$ на единицу превышает число уравнений, (опорный план состоит из $m + n - 1 = 6$ элементов), одну из переменных можно положить равной нулю. Выберем для этой цели переменную u_3, которая чаще всего входит в систему (9.11). Положив $u_3 = 0$, легко находим значения остальных переменных:

$$v_2 = 8, \quad v_3 = 12, \quad v_4 = 7,$$

$$u_1 = 3, \quad v_1 = 6, \quad u_2 = 5.$$

Введем их в табл. 9.3, попутно несколько ее модифицировав:

Таблица 9.4

<table>
<tr><td rowspan="4">i</td><td colspan="5">j</td></tr>
<tr><td></td><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td rowspan="2">u_i</td><td colspan="4">v_j</td></tr>
<tr><td>6</td><td>8</td><td>12</td><td>7</td></tr>
<tr><td>1</td><td>3</td><td>3
20</td><td>5
80</td><td>7</td><td>11</td></tr>
<tr><td>2</td><td>5</td><td>1
130</td><td>4</td><td>6</td><td>3</td></tr>
<tr><td>3</td><td>0</td><td>5</td><td>8
40</td><td>12
80</td><td>7
50</td></tr>
</table>

Теперь исследуем построенную систему $\{u_i, v_j\}$ на потенциальность, т. е. проверим, будут ли потенциалы свободных клеток табл. 9.4 удовлетворять условию (9.9).

Имеем:

$$\begin{cases} v_3 - u_1 = 9 > 7 = c_{13}^*, \\ v_4 - u_1 = 4 < 11 = c_{14}, \\ v_2 - u_2 = 3 < 4 = c_{22}, \\ v_3 - u_2 = 7 > 6 = c_{23}^*, \\ v_4 - u_2 = 2 < 3 = c_{24}, \\ v_1 - u_3 = 6 > 5 = c_{31}^*. \end{cases} \tag{9.12}$$

В строках, отмеченных *, условие (9.9) нарушается. Таким образом, наш первоначальный опорный план оптимальным не является.

Улучшение плана

В теории ТЗ доказывается, что при опорном плане *для каждой свободной клетки «транспортной» таблицы существует единственный цикл, вершина которого лежит в данной свободной клетке, а остальные — в клетках опорного плана.*

Такие клетки еще называют *базисными*.

Для улучшения первоначального опорного плана нам следует определить среди свободных клеток наиболее «выгодную» и ввести ее в план. Чтобы план оставался опорным, мы должны при этом сделать одну из базисных клеток свободной, изменив перевозки таким образом, чтобы баланс запасов и заявок не нарушался.

Номер (i, j) «выгодной» клетки, подлежащей введению в план, определяется как номер максимальной положительной разности $\alpha_{ij} = v_j - u_i - c_{ij}$ (см.(9.9), (9.12)). Если таковых наберется несколько, выбирается номер любой из них.

Дальнейшее изложение проведем на примере табл. 9.4.

1) Согласно равенствам (9.12) наибольшая разность равна $\max(\alpha_{13}, \alpha_{23}, \alpha_{31}) = \alpha_{13} = 2$. Следовательно, в план вводится клетка $(1, 3)$ (или клетка $(1, 3)$ табл. 9.3).

2) В табл. 9.4 находим (единственный) цикл с первой вершиной в клетке $(1, 3)$, а остальными — базисных клетках: это последовательность клеток $(1,3) \to (1,2) \to (3,2) \to (3,3)$ $(\to (1,3))$. Обходим этот цикл против часовой стрелки, выходя из клетки $(1, 3)$, и отмечаем его клетки попеременно знаки «+» и «−» начиная со знака «+».

Цикл показан в табл. 9.5:

Таблица 9.5

i		j			
		1	2	3	4
	u_i	v_j			
		6	8	12	7
1	3	3 20	− 5 80	+7	11
2	5	1 130	4	6	3
3	0	5	8 40+	12 80 −	7 50

Обозначим через θ наименьшую из перевозок x_{ij}, расположенных в «отрицательных» клетках и сделаем клетку, соответствующую этой перевозке, свободной, исключив ее из плана.

В данном случае $\theta = \min(x_{12}, x_{32}) = 80$, $x_{12} = x_{33}$, и из базисных клеток *исключается именно клетка с номером* $(3, 3)$ *поскольку* стоимость перевозки $c_{33} = 12$ в этой клетке больше, чем $c_{12} = 5$. Для сохранения баланса запаса и заявок вычитаем $\theta = 80$ из перевозок в «отрицательных» и прибавляем к перевозкам в «положительных» клетках. Найденный нами новый план заносим в табл. 9.6.

Таблица 9.6

i		j			
		1	2	3	4
	u_i	v_j			
		6	8	12	7
1	3	3 20	5 0	7 80	11
2	5	1 130	4	6	3
3	0	5	8 120	12	7 50

Новый план лучше прежнего: можно подсчитать, что общая стоимость перевозки по нему на 160 ед. меньше, чем стоимость плана из табл. 9.5.

Действительно, алгебраическая, с учетом знака клетки, сумма стоимостей, стоящих в вершинах цикла из табл. 9.5 (так называемая *цена цикла*) равна $7 - 5 + 8 - 12 = -2$. Это значит, что при переносе одной единицы груза по этому циклу, стоимость перевозок уменьшается на 2. А мы перенесли целых 80 ед.; значит, стоимость должна была уменьшиться на $2 \cdot 80 = 160$, что и произошло.

Далее с планом из табл. 9.6 следует поступить точно так же, как мы ранее поступили с первоначальным опорным планом из табл. 9.3, т. е. для начала проверить его оптимальность. Для этого надо построить новую систему потенциалов. Однако обычно проще не решать уравнения (9.10) заново, а подправить старые потенциалы. Покажем на нашем примере, как это делается.

Для введенной в план новой базисной клетки $(1, 3)$ необходимое условие потенциальности $v_3 - u_1 = c_{13}$ не выполняется

($\alpha_{13} = v_3 - u_1 - c_{13} = 2$). Чтобы исправить ситуацию, можно либо уменьшить v_3 на α_{13}, либо увеличить u_1 на α_{13}. Сделаем первое, поскольку в столбце 3 меньше базисных клеток, чем в строке 1.

Имеем: $v_3' = v_3 - \alpha_{13} = 12 - 2 = 10$.

В третьем столбце, кроме $(1, 3)$, нет базисных клеток. Поэтому замена v_3 на v_3' не нарушает условий (9.10) для остальных базисных клеток. Получим новую транспортную таблицу (табл. 9.7), отличающуюся от табл. 9.6 лишь значением потенциала v_3 и изображением «цикла», который будет получен в дальнейшем.

Таблица 9.7

i		j			
		1	2	3	4
	u_i	v_j			
		6	8	10	7
1	3	− 3 20	+ 5 0	7 80	11
2	5	1 130 +	4 −	6	3
3	0	5	8 120	12	7 50

Последующий анализ табл. 9.7 проводится аналогично тому, как это делалось с табл. 9.5.

Для свободных клеток табл. 9.7 вычисляются разности $v_j - u_i$:

$$\begin{cases} v_1 - u_3 = 6 - 0 = 6 > c_{31} = 5^*, \\ v_2 - u_2 = 8 - 5 = 3 < c_{22} = 4, \\ v_3 - u_2 = 10 - 5 = 5 < c_{23} = 6, \\ v_3 - u_3 = 10 - 0 = 10 < c_{33} = 12, \\ v_4 - u_1 = 7 - 3 = 4 < c_{14} = 11, \\ v_4 - u_2 = 7 - 5 = 2 < c_{24} = 3. \end{cases} \tag{9.13}$$

Поскольку в первой (отмеченной *) строке неравенство (9.9) нарушается, план из табл. 9.7 все еще не является оптимальным, и мы вновь возвращаемся к подпунктам 1 и 2.

Для клетки $(3, 1)$ (с единственной положительной невязкой), которая вводится в план, строим цикл $(3, 1) \to (3, 2) \to (1, 2) \to$ $\to (1, 1)$ $(\to (3, 1))$ и отмечаем его вершины знаками «+» и «−» попеременно (см. табл. 9.7). Наименьшая перевозка, стоящая в вершине цикла, равна 20 и расположена в клетке $(1, 1)$, которая из плана исключается. Добавляя 20 к перевозкам, стоящим в положительных вершинах цикла (т. е. в клетках $(1, 2)$, $(3, 1)$) и вычитая это число из перевозок, находящихся в отрицательных вершинах (клетка $(1, 1)$), получаем новый план перевозок:

Таблица 9.8

i		j			
		1	2	3	4
	u_i	v_j			
		6	8	10	7
1	3	3	5 20	7 80	11
2	5	1 130	4	6	3
3	0	5 20	8 100	12	7 50

«Цена» этого цикла равна $5 - 8 + 5 - 3 = -1$. Следовательно, общая стоимость перевозок уменьшилась еще на $1 \cdot 20 = 20$.

Теперь исправим систему потенциалов в табл. 9.8.

Для введенной в план клетки $(3, 1)$ равенство $v_1 - u_3 = c_{31}$ не выполняется, поэтому следует положить $v_1' = v_1 - 1 = 6 - 1 = 5$ (в столбце 1 меньше базисных клеток, чем в строке 3). Однако замена v_1 на v_1' нарушает условие (9.9) для входящей в первый столбец базисной клетки $(2, 1)$, поэтому исправим потенциал u_2: $u_2' = u_2 - 1 = 5 - 1 = 4$ (из u_2 вычитается то же число, что ранее из v_1). Поскольку во второй строке базисных клеток больше нет, потенциалы v_2, v_3, v_4 в исправлении не нуждаются, как впрочем и потенциал u_1.

Табл. 9.8 с учетом измененной системы потенциалов выглядит теперь следующим образом:

Таблица 9.9

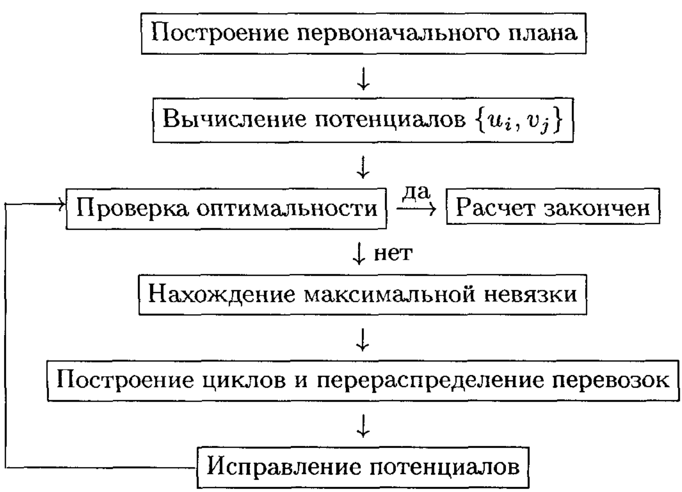
i	j				
		1	2	3	4
	u_i	v_j			
		5	8	10	7
1	3	3	5 20	7 80	11
2	4	1 130	4	6	3
3	0	5 20	8 100	12	7 50

Проверка условий (9.8), (9.9) показывает, что они выполняются. Таким образом, план в табл. 9.9 является оптимальным. Соответствующая ему минимальная общая стоимость перевозок составляет

$$L = 5 \cdot 20 + 7 \cdot 80 + 1 \cdot 130 + 5 \cdot 20 + 8 \cdot 100 + 7 \cdot 50 = 2040.$$

Замечание. Если среди всевозможных опорных планов ТЗ хотя бы один содержит менее $m+n-1$ положительных элементов

Блок-схема вычислений по методу потенциалов



x_{ij}, при использовании метода потенциалов возможно зацикливание, т. е. возвращение к ранее встретившемуся опорному плану. Такая ситуация возникает в случае наличия группы из $p < m$ пунктов отправления, суммарный запас груза в котором в точности равен суммарным потребностям некоторой группы из $q < n$ пунктов назначения. Для предотвращения возможности зацикливания в этом случае следует слегка подкорректировать числа $\{a_i,\ b_i\}$, не нарушая, разумеется, равенства (9.3).

Самостоятельная работа

1. Используя метод минимального элемента, составить опорный план для перевозки фармацевтического сырья некоторого вида с четырех оптовых складов на пять перерабатывающих фабрик. Запасы $\{a_i\}$ оптовых складов по поставкам сырья (в тоннах), а также потребности $\{b_i\}$ фабрик и стоимости транспортировок тонны сырья с i-го оптового склада j-ю на фармацевтическую фабрику (где $i = 1, 2, 3, 4$; $j = 1, 2, 3, 4, 5$), приведены в табл. 9.10.

Используя метод потенциалов, составить оптимальный план таких перевозок и найти его минимальную стоимость.

Таблица 9.10

i	j					a_i
	1	2	3	4	5	
1	16	30	17	10	16	5
2	30	27	26	9	23	5
3	13	4	22	3	1	10
4	3	1	5	4	24	10
b_j	7	7	7	7	2	30

Таблица 9.11

i	j					a_i
	1	2	3	4	5	
A_1	17	20	29	26	25	15
A_2	3	4	5	15	24	15
A_3	19	2	22	4	13	15
A_4	20	27	1	17	19	15
b_j	11	11	11	11	16	60

Таблица 9.12

i	j					a_i
	1	2	3	4	5	
1	4	21	12	8	1	19
2	20	8	25	15	23	23
3	17	1	11	5	3	23
4	23	10	24	6	5	23
b_j	22	22	22	11	11	88

Таблица 9.13

i	j					a_i
	1	2	3	4	5	
1	3	25	11	22	12	23
2	9	15	4	26	12	25
3	13	22	15	12	27	12
4	6	19	8	11	8	30
b_j	18	18	18	18	18	90

2. Решить предыдущую задачу для исходных данных, приведенных в табл. 9.11–9.13.

§ 9.4. Элементы сетевого планирования и управления

Если выполняется сложный комплекс работ (строительство дома, проведение избирательной компании), требуется согласовать отдельные работы, с тем чтобы завершить их в наиболее короткий срок, или распределить финансы, оборудование, материалы, прочие ресурсы, так чтобы вся работа обошлась дешевле. Ответы на поставленные вопросы можно получить методами сетевого планирования.

Назовем намеченный комплекс работ, необходимый для достижения некоторой цели, *проектом*. Примерами может служить проект строительства здания, сервировки стола, написания контрольной и т. д.

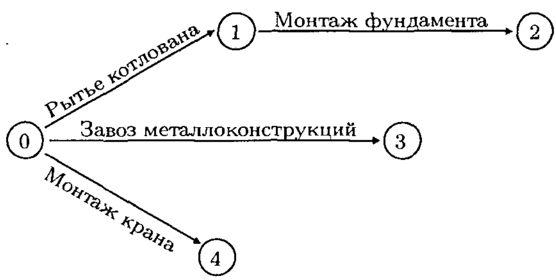


Рис. 9.12. Схема выполнения работ при начале строительства здания.

Каждая отдельная работа, входящая в проект, требует затраты определенного времени. Некоторые работы можно выполнить только в определенном порядке (сначала купить гуся, потом зажарить), другие — одновременно (приготовление салата и жарка гуся).

Если каждому отдельному событию, входящему в проект, поставить в соответствие некоторую точку (вершину), а каждой работе — отрезок с учетом его направления (ориентированное ребро), то получится некая фигура, отражающая последовательность выполнения отдельных работ в общем проекте:

0 — начало,

1 — котлован готов,

2 — фундамент подведен,

3 — металлоконструкции завезены,

4 — смонтирован кран.

Если над ребрами поставить время, необходимое для завершения соответствующей работы, то получится *сеть*. Ее изображение называется *сетевым графиком*. Если для начала какой-то работы требуется завершение нескольких работ, для отражения очередности их выполнения используются штриховые стрелки.

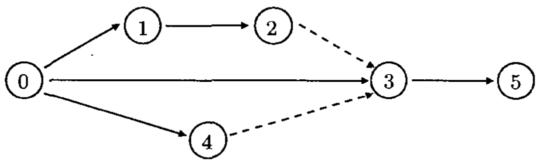


Рис. 9.13а. Пример сетевого графика.

Сетевой график является графической моделью всего проекта. Он отражает взаимосвязь всех работ.

Любая стрелка на сетевом графике соединяет только две вершины и отражает переход от одного события к другому. Сами стрелки, как уже говорилось, используются для отображения:

а) *действительной* работы (любой трудовой процесс, требующий затрат труда, времени и материальных ресурсов),

б) *ожидания* (пассивный процесс, требующий временных затрат — «что-то сохнет ...»),

в) *фиктивной* работы (чисто условная зависимость между событиями, не связанная с какими либо затратами, и вводимая для удобства изображения сети).

При этом для отображения действительной работы и ожидания используют сплошные стрелки, а фиктивная работа отображается штриховыми стрелками.

До события 3 надо произвести работы 1, 2, 4–6.

Событие 0 — исходное;

Событие 7 — завершающее;

Остальные — промежуточные.

При построении сетевых графиков следят, чтобы:

1) из всех вершин, кроме «завершающей», выходили стрелки;

2) во все вершины, кроме «исходной», входили стрелки;

3) если одно событие служит началом для двух и более работ, после завершения которых начинается выполнение следующей работы, то вводится штриховая стрелка и дополнительное событие со своим номером.

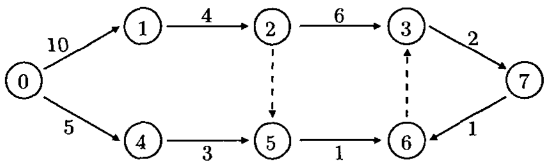
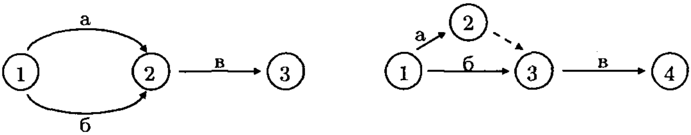


Рис. 9.13б. Пример сетевого графика.

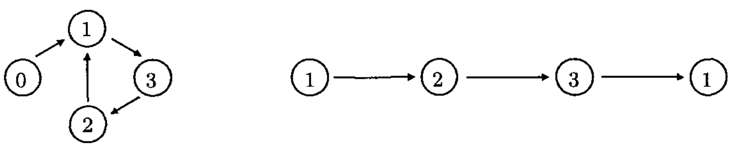


Неправильно Правильно

Рис. 9.13в. Пример сетевых графиков.

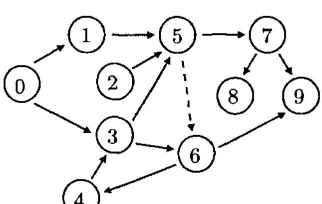
Вообще, на сетевом графике надо четко отражать последовательность выполнения отдельных работ и их взаимосвязи.

4) Не допускаются циклы, т.е. пути, у которых совпадают начальные и конечные вершины.



Нельзя Нельзя

Рис. 9.13г. Примеры циклов.



Ошибки:
1. У вершины 2 нет входа;
2. У вершины 8 нет выхода;
3. Имеется цикл 3 → 6 → 4 → 3.

Рис. 9.14. Пример сетевого графика, имеющего ошибки.

Пусть сетевой график некоторого проекта построен. За какое время его можно осуществить?

Назовем *путем* в сети последовательность ребер, такую что конец каждого ребра совпадает с началом следующего и ни одно ребро не встречается больше одного раза.

Назовем *продолжительностью* пути время, необходимое для выполнения всех работ, лежащих на нем.

Рассмотрим все пути, ведущие от исходного события к завершающему. Как правило, их несколько.

Путь с наибольшей продолжительностью называется *критическим* путем, а работы на нем лежащие — *критическими*.

Оказывается, минимальное время осуществления всего проекта равно продолжительности критического пути.

Действительно, все работы, в том числе критические, необходимо выполнить в порядке, который определяет сетевой график, т. е. движение по критическому пути неизбежно. Кстати, в сети может быть несколько критических путей.

Понятно, что сокращение или увеличение сроков выполнения критических работ соответственно сокращает или увеличивает общее время выполнения проекта.

На сетевом графике критический путь выделяют более жирными стрелками. Заметим, что направление штриховой стрелки влияет на продолжительность критического пути, что видно из следующего примера:

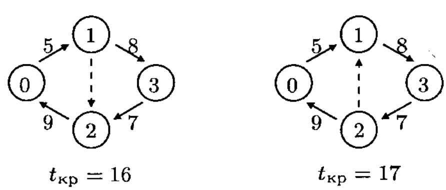


Рис. 9.15. Пример сетевого графика, иллюстрирующий влияние направления штриховой стрелки на продолжительность критического пути.

Что касается работ, не лежащих на критическом пути, то они, вообще говоря, допускают некоторое запаздывание в их выполнении, которое не задержит сроков реализации всего проекта.

Для отыскания критического пути на сети и его продолжительности применяется алгоритм пошаговой оптимизации. Этот алгоритм основан на принципе оптимальности Р. Беллмана из теории динамического программирования: критический путь обладает тем свойством, что, каков бы ни был его отрезок от исходной до некоторой вершины сети, оставшаяся часть критического пути между этой и завершающей вершинами должна быть наиболее продолжительной.

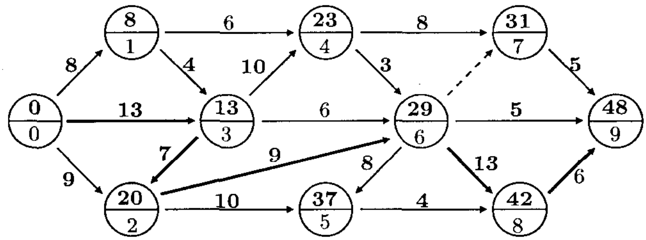


Рис. 9.16. Пример расчета критического пути.

Рассмотрим использование этого алгоритма на примере конкретного сетевого графика (цифры $0, 1, \ldots, 9$ в нижних половинках кружков, являются номерами этих вершин).

Обозначим через t_i, $i = 0, \ldots, 9$, продолжительность наиболее «долгого» пути, ведущего от вершины 0 к вершине i. Тогда $t_0 = 0$, $t_9 = t_{\text{кр}}$. В соответствии с принципом оптимальности Беллмана получаем

$$t_i = \max(t_k + l_k(i)), \quad i = 1, \ldots, 9, \quad k \in S_i, \tag{9.14}$$

где S_i — множество вершин сети, непосредственно предшествующих i-й вершине (при $i = 5$, например, $S_i = \{2, 6\}$), $l_k(i)$, $k \in S_i$, — продолжительность работы между k-й и i-й вершинами ($l_2(5) = 10$, $l_6(5) = 8$).

В нашем случае

$t_0 = 0$, $t_1 = 8$,

$t_3 = \max(t_0 + 13, t_1 + 4) = \max(0 + 13, 8 + 4) = \max(13, 12) = 13$ ($S_3 = \{0, 1\}$),

$t_2 = \max(9, t_3 + 7) = \max(9, 13 + 7) = 20$ ($S_2 = \{0, 3\}$),

$t_4 = \max(t_1 + 6, t_3 + 10) = \max(8 + 6, 13 + 10) = 23$ ($S_4 = \{1, 3\}$),

$t_6 = \max(t_2 + 9, t_3 + 8, t_4 + 3) = \max(30 + 9, 13 + 8, 23 + 3) = 29$ $(S_6 = \{2, 3, 4\})$,
$t_5 = \max(t_2 + 10, t_6 + 8) = \max(20 + 10, 29 + 8) = 37$ $(S_5 = \{2, 6\})$,
$t_7 = \max(t_4 + 8, t_6 + 0) = \max(23 + 8, 29 + 0) = 31$ $(S_7 = \{4, 6\})$,
$t_8 = \max(t_5 + 4, t_6 + 13) = \max(37 + 4, 29 + 13) = 42$ $(S_8 = \{5, 6\})$,
$t_9 = \max(t_7 + 5, t_6 + 5, t_8 + 6) = \max(31 + 5, 29 + 5, 42 + 6) = 48 = t_{\text{кр}}$ $(S_9 = \{6, 7, 8\})$.

Заметим, что в рассмотренном примере вершина 3 при вычислениях предшествовала вершине 2, а вершина 6 — вершине 5.

Итак, время (наименьшее) выполнения всего проекта равно 48 единицам. Соответствующий этому времени путь несложно восстановить, двигаясь назад от завершающей вершины 9 к исходной 0.

В вершину 9 можно попасть лишь из вершины 8 $(48 = 42 + 6)$, в 8 — из 6 $(42 = 29 + 13)$, в 6 — из 2 $(29 = 20 + 9)$, в 2 — из 3 $(20 = 13 + 7)$ и, наконец, в 0 — из 3 $(13 = 13 + 0)$.

Таким образом, критический путь: $0 \to 3 \to 2 \to 6 \to 8 \to 9$.

Аналогичным образом можно определить *кратчайший путь* на сети. Для этого в формуле (9.14) достаточно max заменить на min.

В нашем случае этот путь, точнее, пути $0 \to 1 \to 4 \to 6 \to 7 \to 9$ и $0 \to 1 \to 4 \to 6 \to 9$ имеют продолжительность 22 единицы. Действительно,
$t_0 = 0$, $t_1 = 8$, $t_3 = \min(0 + 13, 8 + 4) = 12$,
$t_2 = \min(9, 12 + 7) = 9$,
$t_4 = \min(8 + 6, 12 + 10) = 14$,
$t_6 = \min(9 + 9, 12 + 8, 14 + 3) = 17$,
$t_5 = \min(9 + 10, 17 + 8) = 19$,
$t_7 = \min(14 + 8, 17 + 0) = 17$,
$t_8 = \min(19 + 4, 17 + 13) = 23$,
$t_9 = \min(17 + 5, 17 + 5, 17 + 6) = 22 = t_{\text{кр}}$.

В заключение отметим, что все вычисления можно выполнять непосредственно на сети. При этом полезно вычеркивать ребра, прошедшие «обработку», и последовательно приписывать каждой рассмотренной вершине наиболее продолжительное время ее достижения из 0, как это сделано на последнем рисунке.

Самостоятельная работа

Построить сетевой график, найти сроки промежуточных итогов в процессе реализации проекта, установить критический путь и вычислить время, необходимое для реализации всего проекта.

1. $t_{12} = 9$; $t_{13} = 5$; $t_{14} = 1$; $t_{29} = 4$; $t_{26} = 4$; $t_{32} = 8$; $t_{36} = 10$; $t_{35} = 8$; $t_{43} = 3$; $t_{45} = 0$; $t_{48} = 15$; $t_{37} = 3$; $t_{58} = 7$; $t_{65} = 9$; $t_{67} = 3$; $t_{69} = 6$; $t_{79} = 2$; $t_{7,10} = 6$; $t_{87} = 11$; $t_{8,10} = 12$; $t_{9,10} = 7$.

2. $t_{12} = 7$; $t_{13} = 4$; $t_{25} = 11$; $t_{26} = 10$; $t_{32} = 5$; $t_{34} = 6$; $t_{35} = 8$; $t_{48} = 4$; $t_{54} = 2$; $t_{58} = 7$; $t_{57} = 9$; $t_{5,10} = 15$; $t_{65} = 8$; $t_{6,10} = 20$; $t_{79} = 14$; $t_{7,10} = 10$; $t_{87} = 3$; $t_{89} = 6$; $t_{9,10} = 7$.

3. $t_{12} = 13$; $t_{13} = 10$; $t_{14} = 7$; $t_{15} = 9$; $t_{27} = 9$; $t_{32} = 5$; $t_{36} = 0$; $t_{37} = 7$; $t_{43} = 8$; $t_{46} = 4$; $t_{49} = 7$; $t_{52} = 10$; $t_{57} = 13$; $t_{5,10} = 20$; $t_{69} = 5$; $t_{76} = 4$; $t_{79} = 10$; $t_{78} = 15$; $t_{8,10} = 12$; $t_{98} = 8$.

4. $t_{12} = 5$; $t_{13} = 4$; $t_{14} = 10$; $t_{25} = 9$; $t_{26} = 15$; $t_{32} = 7$; $t_{36} = 9$; $t_{37} = 3$; $t_{43} = 8$; $t_{47} = 12$; $t_{56} = 5$; $t_{59} = 7$; $t_{68} = 3$; $t_{69} = 4$; $t_{76} = 0$; $t_{78} = 2$; $t_{89} = 1$; $t_{8,10} = 1$; $t_{9,10} = 3$.

5. $t_{12} = 15$; $t_{13} = 10$; $t_{14} = 17$; $t_{15} = 20$; $t_{25} = 12$; $t_{26} = 13$; $t_{27} = 9$; $t_{32} = 7$; $t_{34} = 5$; $t_{37} = 13$; $t_{47} = 4$; $t_{49} = 10$; $t_{4,10} = 16$; $t_{58} = 5$; $t_{65} = 4$; $t_{76} = 8$; $t_{7,8} = 11$; $t_{79} = 17$; $t_{89} = 12$; $t_{8,10} = 19$; $t_{9,10} = 14$.

§ 9.5. Введение в теорию массового обслуживания. Формулы Эрланга

С работой своеобразных систем, называемых системами массового обслуживания (сокращенно СМО), приходится сталкиваться повседневно. Примерами таких СМО могут служить телефонные станции, ремонтные службы, билетные кассы, справочные бюро, магазины, аптеки, парикмахерские, т. е. любые *системы, предназначенные для обслуживания (в том или ином смысле) некоторого потока заявок (или «требований»), поступающих в какие-то, вообще говоря случайные, моменты времени.*

Каждая СМО состоит из некоторого числа обслуживающих единиц (или «приборов»), называемых *каналами обслуживания.* Каналами могут быть: линии связи, лифты, продавцы, кассиры и т. д.

Время обслуживания потока заявки длится какое-то, как правило случайное, время, после чего канал освобождается и готов к приему следующей заявки. Случайный характер потока заявок и времен обслуживания приводит к тому, что в какие-то периоды времени на входе СМО создается очередь, в другие же периоды СМО будет работать с недогрузкой.

Таким образом, процесс работы СМО представляет собой случайный процесс с дискретными состояниями и непрерывным временем: состояние СМО меняется скачком в моменты появления прихода новой заявки или окончания обслуживания (клиент пришел — клиент ушел).

Предметом теории массового обслуживания (ТМО) является построение математических моделей, связывающих данные условия работы СМО (характер потока заявок, число каналов и их производительность, дисциплина обслуживания) с показателями эффективности СМО.

В качестве таких показателей могут использоваться разные характеристики: среднее число заявок, обслуживаемых в единицу времени; среднее число занятых каналов; вероятность отказа в обслуживании.

Рассмотрим следующий пример.

Пусть речь идет об аптеке, в которой работает несколько служащих (каналов обслуживания). Клиенты, обратившиеся за медикаментами, образуют поток требований. Представьте, что в аптеку забежал покупатель, готовый приобрести дорогое лекарство, но не располагающий временем или желанием стоять в очереди. Надо уметь вычислять вероятность того, что он не будет обслужен — ведь если большинство клиентов уйдет без покупки, вряд ли стоит держать аптеку вообще. Полезно также знать степень загрузки каждого работника — это характеризует рентабельность аптеки. Поскольку число потенциальных клиентов и время обслуживания — величины случайные, задача решается далеко не просто.

В примере условие ухода клиента, если его обслуживание не началось немедленно, выглядит несколько искусственным — большинство покупателей могут подождать. Однако если вместо аптеки рассматривать АТС (автоматическую телефонную станцию), а обслуживанием считать продолжительность телефонного разговора, то вышеупомянутое условие выполняется.

Если абстрагироваться от реального наполнения моделей СМО (мастерская, аптека, АТС, лифты в доме и т. д.), СМО можно описать, задавая следующие ее составляющие:

1) входящий поток требований,
2) дисциплину очереди,
3) механизм обслуживания,
4) выходящий поток требований.

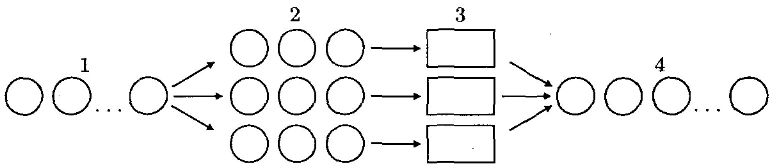


Рис. 9.17. Модель теории массового обслуживания.

В некоторых системах «очередь» отсутствует.

СМО делится на классы по ряду признаков.

1. *СМО с отказами* (как в телефонии) и *СМО с очередью.*

На практике чаще встречаются и имеют большее значение СМО с очередью: недаром ТМО иногда называют «теория очередей».

СМО с очередью: длина очереди и (или) время ожидания могут быть ограничены или нет; обслуживание может быть с *приоритетом* или без него, в порядке поступления или случайным.

Приоритет может быть *абсолютным* или *относительным.*

СМО могут быть *открытыми* и *закрытыми.* В первой — поток заявок не зависит от состояния самой СМО (сколько каналов занято), во второй — зависит. Пример — наладка группы станков одним рабочим. Здесь интенсивность «требований» со стороны станков зависит от того, сколько их уже неисправно.

Классификация СМО не ограничивается приведенными разновидностями.

Возвращаясь к компонентам СМО, рассмотрим более подробно *входящий поток* требований как одно из наиболее важных понятий ТМО.

Потоком требований называется совокупность заявок на обслуживание, поступающих в обслуживающую систему.

Он может быть *регулярным* или *стохастическим* (т. е. случайным). В 1-м случае требования следуют друг за другом через равные промежутки времени Δt, во 2-м моменты появления требований — случайные величины.

Важной характеристикой потока требований является его *интенсивность* λ — среднее число требований, поступающих в систему в единицу времени.

В случае регулярного потока $\lambda = 1/\Delta t$; в общем случае интенсивность может быть как постоянной, так и зависящей от t. Например, поток машин ночью не так интенсивен, как днем.

Входящий поток называется *стационарным*, если вероятность поступления определенного количества требований в течение определенного промежутка времени зависит *лишь* от длины этого промежутка.

В частности, интенсивность λ стационарного потока должна быть постоянной, т.е. *в среднем* на интервалах равной длины должно быть одинаковое количество требований.

Свойством стационарности обладают многие реальные потоки требований, по крайней мере на ограниченном участке времени (нагрузка на АТС меняется в течение суток, но не между, скажем, часом и двумя).

Поток требований называется потоком *без последействия*, если для любых двух непересекающихся участков времени τ_1 и τ_2 число требований, поступивших в систему за τ_2, не зависит от того, сколько требований поступило за промежуток τ_1.

Другими словами, прошлое не влияет на настоящее!

По сути, это означает, что требования, образующие поток, появляются в те или иные моменты времени независимо друг от друга (как, например, поток пассажиров, входящих в метро).

Пусть случайная величина $X(t)$ обозначает число требований на интервале $[0, t]$.

Поток называется *ординарным*, если

$$(1/t)P(X(t) > 1) \to 0 \tag{9.15}$$

при $t \to 0$.

Заметим, что $P(X(t) > 1) = \sum\limits_{k=2}^{\infty} p_k(t) = 1 - p_0(t) - p_1(t)$, где $p_k(t) = P(X(t) = k)$.

В ординарном потоке появление двух и более требований за малый промежуток времени практически невозможно.

Поток клиентов в аптеку обычно ординарен.

Поток требований называется *простейшим*, если он стационарен, ординарен и не имеет последействия.

Потоки такого типа часто встречаются на практике. Термин «простейший» связан с простым математическим описанием этих потоков.

Можно показать, что для простейшего потока число требований $X(t)$ в промежутке времени длиной t распределено по закону Пуассона с параметром λt (см. п. 7.2.1), т. е.

$$P_k(t) = \frac{(\lambda \cdot t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \ldots \tag{9.16}$$

Стационарность и отсутствие последействия налицо, ординарность (т. е. условие (9.15)) вытекает из равенства

$$\frac{1}{t} P_{k>1}(t) = \frac{1 - e^{-\lambda t} - \lambda \cdot t e^{-\lambda t}}{t} \to 0, \quad t \to 0,$$

которое можно проверить по правилу Лопиталя.

Параметр λ здесь характеризует интенсивность потока. Действительно,

$$\frac{1}{t} MX(t) = \frac{1}{t} \sum_{k=0}^{\infty} k P_k(t) = \lambda \sum_{k=1}^{\infty} k \frac{(\lambda \cdot t)^{k-1}}{k!} e^{-\lambda t} =$$

$$= \lambda \cdot e^{-\lambda t} \sum_{l=0}^{\infty} \frac{(\lambda \cdot t)^l}{l!} = \lambda \sum_{l=0}^{\infty} p_l(t) = \lambda.$$

Простейший поток еще называют *стационарным пуассоновским.*

Пример 1. Наладка станков одним рабочим. Предполагается, что все станки приблизительно в одинаковом состоянии (последнее обеспечивает стационарность потока поломок); вероятность поломки одного станка невелика (двух, ..., тем более) — отсюда следует ординарность; кроме того, если станков много, а среднее время ремонта мало, можно считать, что поток поломок не имеет последействия. Другими словами, он является простейшим.

Пусть интенсивность $\lambda = 3$ поломки/час. По формуле (9.16) при $\lambda = 3$ и $t = 1$ найдем вероятность к поломок в течение часа $p_k(1) = 3^k e^{-3}/k!$, $k = 0, 1, \ldots$ ($e^{-3} \approx 0{,}0498$).

K	0	1	2	3	4	5	6	...
$P_k(1)$	0,05	0,15	0,22	0,22	0,17	0,1	0,05	...

Следующее важное понятие ТМО — это *время обслуживания.*

Оно является характеристикой функционирования каждого отдельного канала обслуживающей системы и отражает его пропускную способность. Время обслуживания — случайная величина.

Для простоты будем рассматривать систему, состоящую из однотипных обслуживающих аппаратов, имеющих общий закон распределения. При этом будем предполагать, что этот закон распределения — показательный, с функцией распределения времени обслуживания (см. (7.21))

$$F(t) = (1 - e^{-\mu t}). \tag{9.17}$$

Параметр μ (аналогично параметру λ входящего потока) определяет интенсивность обслуживания; величина $1/\mu$ является средним временем обслуживания t одной заявки:

$$Mt = \int\limits_0^\infty tF'(t)\,dt = \frac{1}{\mu}. \tag{9.18}$$

Показательный закон имеет большое значение как в теоретических исследованиях, так и во многих приложениях. Важнейшим его свойством является то, что при таком законе распределения времени обслуживания оставшееся время обслуживания не зависит от того, сколько времени обслуживание уже длилось.

Далее коротко опишем n-канальную систему массового обслуживания с отказами.

Это «классическая» задача ТМО, возникшая из практических нужд телефонии и решенная в начале XX века датским математиком Эрлангом. Задача ставится следующим образом.

Имеется n каналов, на которые поступает простейший поток заявок с интенсивностью λ. Если в момент поступления очередного требования имеется хотя бы один свободный аппарат, то любой из них немедленно приступает к обслуживанию. В противном случае заявка получает отказ и покидает систему.

Все каналы работают независимо друг от друга и от входящего потока.

Время обслуживания каждого требования распределено по показательному закону (см. (9.17)) с параметром μ (т. е. среднее время обслуживания $\tau_{\text{об}} = 1/\mu$). Требуется найти характеристики эффективности работы СМО в стационарном (установившемся) режиме, т. е. при неограниченно возрастающем времени ее работы. Конкретнее, нас интересуют:

A — абсолютная пропускная способность, т. е. среднее число заявок, обслуживаемых в единицу времени;

Q — относительная пропускная способность, или средняя доля пришедших заявок, обслуживаемых системой;

$P_{\text{отк}}$ — вероятность отказа, или того, что заявка покинет СМО необслуженной;

K — среднее число занятых каналов;

P_k, $0 \leqslant k \leqslant n$, — вероятность того, что занято ровно k каналов, и, в частности, P_0 — вероятность простоя системы;

$K_3 = \overline{K}/n \cdot 100\%$ — коэффициент простоя каналов в %;

$K_n = (1 - \overline{K}/n) \cdot 100\% = (100 - K_3)\%$ — коэффициент простоя каналов в %.

Обозначим

$$\lambda/\mu = \alpha. \tag{9.19}$$

Величина α обычно называется «приведенной интенсивностью потока заявок» и ее смысл — среднее число заявок, приходящее за среднее время обслуживания одной заявки. Пользуясь этим обозначением, можно показать, что вероятность P_0 того, что все n каналов СМО свободны, выражается формулой

$$P_0 = \left(1 + \alpha + \frac{\alpha^2}{2!} + \ldots + \frac{\alpha^n}{n!}\right)^{-1} \tag{9.20}$$

а вероятности P_k при $1 \leqslant k \leqslant n$ имеют вид

$$P_k = \frac{\alpha k}{k!} P_0. \tag{9.21}$$

Формулы (9.20), (9.21) для вероятностей P_k называются *формулами Эрланга* — в честь основателя ТМО. С их помощью можно вычислить остальные интересующие нас характеристики СМО. Так, вероятность $P_{\text{отк}} = P_n$. Действительно, для того чтобы пришедшая заявка получила отказ, необходимо, чтобы все n каналов были заняты. Итак,

$$P_{\text{отк}} = P_n = \frac{\alpha^n}{n!} P_0. \tag{9.22}$$

Отсюда находим относительную пропускную способность, т. е. вероятность, что заявка будет обслужена:

$$Q = 1 - P_{\text{отк}} = 1 - \frac{\alpha^n}{n!} P_0. \tag{9.23}$$

Абсолютную пропускную способность получим, умножая интенсивность потока заявок на Q:

$$A = \lambda \cdot Q = \lambda \cdot \left(1 - \frac{\alpha^n}{n!} P_0\right). \tag{9.24}$$

Среднее число занятых каналов $\overline{K}$ по определению математического ожидания с учетом формул (9.20) и (9.21) равно

$$\begin{aligned} 0 \cdot P_0 + 1 \cdot P_1 + \ldots + nP_n &= \sum_{K=1}^{n} \frac{k\alpha^k}{k!} P_0 = \\ &= \sum_{l=0}^{n-1} \frac{\alpha^{l+1}}{l!} P_0 = \alpha \sum_{l=0}^{n-1} P_l = \alpha(1 - P_n); \end{aligned}$$

$$\overline{K} = \alpha \cdot \left(1 - \frac{\alpha^n}{n!} P_0\right). \tag{9.25}$$

Отметим, что, зная вероятность отказа $P_{\text{отк}}(n) = P_n$ в обслуживании системы с n каналами обслуживания (см. (9.22)), аналогичную вероятность для системы с $(n \pm 1)$ каналом можно вычислить, пользуясь несложно проверяемыми равенствами

$$P_{\text{отк}}(n+1) = \frac{P_n}{P_n + \dfrac{n+1}{\alpha}} \quad \text{и} \quad P_{\text{отк}}(n-1) = \frac{n}{\alpha} \cdot \frac{P_n}{1-P_n}. \qquad (9.26)$$

Приведем два примера, использующих рассмотренную теорию.

Пример 2. Пусть имеется АТС с пятью линиями связи. Поток вызовов, поступающий на АТС, предполагается простейшим с интенсивностью $\lambda = 1$ вызова в минуту, а время разговора — распределенным по показательному закону со средним временем разговора $t_{\text{ср}} = 2$ мин. Предполагается также, что требование получает отказ, если в момент его поступления все 5 линий заняты. Требуется вычислить основные характеристики эффективности СМО в установившемся режиме.

Решение. Используя формулы (9.19)–(9.25) найдем, что $\alpha = \lambda/\mu = \lambda t_{\text{ср}} = 1 \cdot 2 = 2$, $P_0 = 0 \cdot 138$, $P_1 = P_2 = 0{,}275$, $P_3 = 0{,}092$, $P_4 = 0{,}092$, $P_5 = 0{,}138 \cdot 2^5/5! = 0{,}037$.

Таким образом, $P_{\text{отк}} = 0{,}037$, $\overline{K} = 2(1 - 0{,}037) \approx 1{,}93$, $Q = 1 - 0{,}037 \approx 0{,}96$, $A = 1 \cdot Q \approx 0{,}96$, $K_3 = 1.93/5 \cdot 100\% \approx 39\%$, $K_n = (100 - 39)\% = 61\%$.

Отсюда заключаем, что на АТС в среднем занято 2 линии из 5, каждая линия загружена всего на 39%, теряется приблизительно 4 вызова из 100. Таким образом, АТС работает не слишком эффективно и вполне можно сократить общее число линий и (или) увеличить интенсивность потока заявок.

Пример 3. Следующий пример возвращает нас к задаче об эффективности работы аптеки.

Пусть имеется аптека с обслуживающим персоналом из 3 человек. Статистическое обследование показало, что среднее число клиентов, обращающихся в аптеку в течение часа равно 24, а среднее время обслуживания каждого клиента занимает 5 минут. Выясним, какова вероятность, что вас не обслужат (предполагается, что если все окошки заняты, то клиент уходит) и насколько продавцы загружены работой.

Решение. Будем предполагать, что клиенты образуют простейший поток (если аптека расположена на бойком месте, это можно эвристически обосновать) и воспользуемся формулами Эрланга для решения.

Имеем: $\lambda = 24$, $n = 3$, $1/\mu = t_{\text{ср}} = 5$ мин. $= 1/12$ часа, и, следовательно, $\alpha = \lambda/\mu = 2$, $P_0 = (1 + 2 + 2^2/2! + 2^3/3!)^{-1} = 0{,}16$, $P_{\text{отк}} = P_3 = 2^3/3! \cdot 0{,}16 \approx 0{,}21$. Таким образом, $Q = 0{,}79$ и, следовательно, будет обслужено в среднем 79 клиентов из 100.

Далее $\overline{K} = \alpha(1 - P_3) = 2(1 - 0{,}21) \approx 1{,}58$, т. е. каждый продавец в среднем загружен $\overline{K}/n = 1{,}58/3 \approx 0{,}53$ рабочего дня.

Казалось бы, одного продавца можно и даже нужно сократить.

Проведенные расчеты, однако, этого не подтверждают. Действительно, пользуясь формулой (9.26), найдем $P_{\text{отк}}(2) = (3/2) \times 0{,}21/(1 - 0{,}21) \approx 0{,}4$, $\overline{K}_{n=2} = 2(1 - 0{,}4) = 1{,}2$. Таким образом, загрузка каждого из двух оставшихся продавцов немного вырастет (с 0,53 до $1/2 \cdot 1{,}2 = 0{,}6$ рабочего дня), зато «коэффициент полезного действия» аптеки упадет с 0,79 до 0,6, поскольку в сложившейся ситуации будет обслужено лишь 60% $(= (1 - 0{,}4) \cdot 100\%)$ потенциальных клиентов, а не 79%, как ранее при трех продавцах.

ПРИЛОЖЕНИЯ

Приложение 1

Функция Лапласа $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_0^x e^{-t^2/2}\,dt$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0040	0080	0120	0159	0199	0239	02799	0319	0359
0,1	0398	0438	0477	0517	0557	0596	0636	0675	0714	0753
0,2	0793	0832	0870	0909	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2257	2291	2324	2356	2389	2421	2454	2486	2517	2549
0,7	2580	2611	2642	2673	2703	2731	2764	2793	2823	2852
0,8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3240	3365	3389
1,0	3413	3437	3461	3485	3508	3531	3554	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3906	3925	3943	3962	3980	3997	4014
1,3	4032	4049	4066	4082	4099	414 5	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4277	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1,6	4452	4463	4474	4484	4595	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4648	4656	4664	4671	4678	4686	4693	4699	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2,0	4772	4778	4783	4788	4793	4798	4803	4807	4812	4817
2,1	4861	4826	4830	4834	4838	4842	4846	4850	4854	4857
2,2	4861	4864	4868	4871	4874	4873	4881	4884	4887	4890
2,3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2,4	4918	4920	4922	4924	4927	4929	4930	4932	4934	4936
2,5	4938	4940	4941	4943	4945	4946	4948	4949	4957	4952
2,6	4953	4955	4956	4957	4958	4960	4961	4962	4963	4964
2,7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2,8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981
2,9	4281	4982	4982	4983	4984	4984	4985	4985	4986	4986
3,0	0,4986		3,1	4990	3,2	4993	3,3	4995	3,4	4997
3,5	4998		3,6	4998	3,7	4999	3,8	4999	3,9	4999
4,0	49997									
4,5	499997									
5,0	499999									

Пример: $\Phi(1{,}32) = 0{,}4066$, $z_{0,9} = \{x : \Phi(x) = 0{,}9/2\} \approx 1{,}645$.

Приложение 2

Критические точки распределения Стьюдента при уровне значимости α в случае двусторонней критической области или $\alpha/2$ в случае односторонней критической области (число степеней свободы k)

k	α								
	0,10	0,05	0,025	0,02	0,01	0,005	0,003	0,002	0,001
1	6,314	12,706	25,452	31,821	63,657	127,3	212,2	318,3	636,6
2	2,920	4,303	6,205	6,965	9,925	14,089	18,216	22,327	31,600
3	2,353	3,182	4,177	4,541	5,841	7,453	8,891	10,214	12,922
4	2,132	2,776	3,495	3,747	4,604	5,597	6,435	7,173	8,610
5	2,015	2,571	3,163	3,365	4,032	4,773	5,376	5,893	6,869
6	1,943	2,447	2,969	3,143	3,707	4,317	4,800	5,208	5,959
7	1,895	2,365	2,841	2,998	3,499	4,029	4,442	4,785	5,408
8	1,860	2,306	2,752	2,896	3,355	3,833	4,199	4,501	5,041
9	1,833	2,262	2,685	2,821	3,250	3,690	4,024	4,297	4,781
10	1,812	2,228	2,634	2,764	3,169	3,581	3,892	4,144	4,587
11	1,796	2,20	2,6	2,72	3,11	3,5	3,8	4,03	4,44
12	1,782	2,179	2,560	2,681	3,055	3,428	3,706	3.930	4,318
13	1,771	2,16		2,65	3,01			3,85	4,22
14	1,761	2,145	2,510	2,624	2,977	3,326	3,583	3,787	4,140
15	1,753	2,13		2,60	2,95			3,73	4,07
16	1,746	2,120	2,473	2,583	2,921	3,252	3,494	3,686	4,015
17	1,740	2,11		2,57	2,90			3,65	3,96
18	1,734	2,101	2,445	2,552	2,878	3,193	3,428	3,610	3,922
19	1,729	2,09		2,54	2,86			3,58	3,88
20	1,725	2,086	2,423	2,528	2,845	3,153	3,376	3,552	3,849
21	1,721	2,08		2,52	2,83			3,53	3,82
22	1,717	2,074	2,405	2,508	2,819	3,119	3,335	3,505	3,792
23	1,714	2,07		2,50	2,81			3,49	3,77
24	1,711	2,064	2,391	2,492	2,797	3,092	3,302	3,467	3,745
25	1,708	2,06		2,49	2,79			3,45	3,72
26	1,706	2,056	2,379	2,479	2,779	3,067	3,274	3,435	3,704
27	1,703	2,05		2,47	2,77			3,42	3,69
28	1,701	2,048	2,369	2,467	2,763	3,047	3,250	3,408	3,674
29	1,699	2,05		2,46	2,76			3,40	3,66
30	1,697	2,042	2,360	2,457	2,750	3,030	3,230	3,386	3,646
40	1,684	2,02		2,42	2,70			3,31	3,55
60	1,671	2,00		2,39	2,66			3,23	3,46
120	1,658	1,98		2,36	2,62			3,17	3,37
$+\infty$	1,645	1,960	2,241	2,326	2,576	2,807	2,968	3,090	3,291

Критические точки распределения Фишера–Снедекора при уровне значи критической области или $\alpha/2$ в случае двусторонней критической облас большей дисперсии, k_2 — число степеней свободы мень

$\alpha = 0{,}01$

k_2	k_1												
	1	2	3	4	5	6	7	8	9	10	11	12	1
1	4052	4999	5403	5625	5764	5889	5928	5981	6022	6056	6082	6106	61
2	98,49	99,01	90,17	99,25	99,33	99,30	99,34	99,36	99,36	99,40	99,41	99,42	99
3	34,12	30,81	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05	26
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,45	14,37	14
5	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,27	10,15	10,05	9,96	9,89	9,
6	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72	7,
7	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,54	6,47	6,
8	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,74	5,67	5,
9	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,18	5,11	5,
10	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,78	4,71	4,
11	9,86	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,46	4,40	4,
12	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,22	4,16	4,
13	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96	3,
14	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80	3,
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67	3,
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,61	3,55	3,
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,45	3,
18	8,28	6,01	5,09	4,58	4,25	4,01	3,85	3,71	3,60	3,51	3,44	3,37	3,
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,36	3,30	3,
20	8,10	5,85	4,94	4,43	4,10	3,87	3,71	3,56	3,45	3,37	3,30	3,23	3,

Критические точки распределения Фишера–С

$\alpha = 0{,}05$

k_2	k_1												
	1	2	3	4	5	6	7	8	9	10	11	12	1
1	161	200	216	225	230	234	237	239	241	242	243	244	2
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40	19,41	19,
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74	8,
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91	5,
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68	4,
6	5,99	5,14	4,76	4,53	4,39	4,28	4.21	4,15	4,10	4,06	4,03	4,00	3,
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57	3,
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28	3,
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07	3,
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91	2,
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79	2,
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69	2,
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60	2,
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53	2,
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48	2,
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42	2,
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38	2,
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,34	2,31	2,
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31	2,28	2,

Правые критические точки распределения χ^2 (уровень значимости

k	α									
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10
1	0,00016	0,0006	0,0039	0,016	0,064	0,148	0,455	1,07	1,64	2,7
2	0,020	0,040	0,103	0,211	0,446	0,713	1,386	2,41	3,22	4,6
3	0,115	0,185	0,352	0,584	1,005	1,424	2,366	3,66	4,64	6,3
4	0,30	0,43	0,71	1,06	1,65	2,19	3,36	4,9	6,0	7,8
5	0,55	0,75	1,14	1,61	2,34	3,00	4,35	6,1	7,3	9,2
6	0,87	1,13	1,63	2,20	3,07	3,83	5,35	7,2	8,6	10,6
7	1,24	1,56	2,17	2,83	3,82	4,67	6,35	8,4	9,8	12,0
8	1,65	2,03	2,73	3,49	4,59	5,53	7,34	9,5	11,0	13,4
9	2,09	2,53	3,32	4,17	5,38	6,39	8,34	10,7	12,2	14,7
10	2,56	3,06	3,94	4,86	6,18	7,27	9,34	11,8	13,4	16,0
11	3,1	3,6	4,6	5,6	7,0	8,1	10,3	12,9	14,6	17,3
12	3,6	4,2	5,2	6,3	7,8	9,0	11,3	14,0	15,8	18,5
13	4,1	4,8	5,9	7,0	8,6	9,9	12,3	15,1	17,0	19,8
14	4,7	5,4	6,6	7,8	9,5	10,8	13,3	16,2	18,2	21,1
15	5,2	6,0	7,3	8,5	10,3	11,7	14,3	17,3	19,3	22,3
16	5,8	6,6	8,0	9,3	11,2	12,6	15,3	18,4	20,5	23,5
17	6.4	7,3	8,7	10,1	12,0	13,5	16,3	19,5	21,6	24,8
18	7,0	7,9	9,4	10,9	12,9	14,4	17,3	20,6	22,8	26,0
19	7,6	8,6	10,1	11,7	13,7	15,4	18,3	21,7	23,9	27,2
20	8,3	9,2	10,9	12,4	14,6	16,3	19,3	22,8	25,0	28,4
21	8,9	9,9	11,6	13,2	15,4	17,2	20,3	23,9	26,2	29,6
22	9,5	10,6	12,3	14,0	16,3	18,1	21,3	24,9	27,3	30,8

Правые критические точки распределения χ^2 (уровень значимости

k	α										
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05
23	10,2	11,3	13,1	14,8	17,2	19,0	22,3	26,0	28,4	32,0	35,2
24	10,9	12,0	13,8	15,7	18,1	19,9	23,3	27,1	29,6	33,2	36,4
25	11,5	12,7	14,6	16,5	18,9	20,9	24,3	28,1	30,7	34,4	37,7
26	12,2	13,4	15,4	17,3	19,8	21,8	25,3	29,3	31,8	35,6	38,9
27	12,9	14,1	16,2	18,1	20,7	22,7	26,3	30,3	32,9	36,7	40,1
28	13,6	14,8	16,9	18,9	21,6	23,6	27,3	31,4	34,0	37,9	41,3
29	14,3	15,6	17,7	19,8	22,5	24,6	28,3	32,5	35,1	39,1	42,6
30	15,0	16,3	18,5	20,6	23,4	25,5	29,3	33,5	36,3	40,3	43,8

Критические значения распределения Кочрена (k — число степеней своб

$\alpha = 0{,}01$

k	m									
	1	2	3	4	5	6	7	8	9	
2	0,9999	0,9950	0,9794	0,9586	0,9373	0,9172	0,8988	0,8823	0,8674	0
3	9933	9423	8831	8335	7933	7606	7335	7107	6912	
4	9676	8643	7814	7212	6761	6410	6129	5897	5702	
5	0,9279	0,7885	0,6957	0,6329	0,5875	0,5531	0,5259	0,5037	0,4854	0
6	8828	7218	6258	5635	5195	4866	4608	4401	4229	
7	8376	6644	5685	5080	4659	4347	4105	3911	3751	
8	0,7945	0,6152	0,5209	0,4627	0,4226	0,3932	0,3704	0,3522	0,3373	0
9	7544	5727	4810	4251	3870	3592	3378	3207	3067	
10	7175	5358	4469	3934	3572	3308	3106	2945	2813	
12	0,6528	0,4751	0,3919	0,3428	0,3099	0,2861	0,2680	0,2535	0,2419	0
15	5747	4069	3317	2882	2593	2386	2228	2104	2002	
20	4799	3297	2654	2288	2048	1877	1748	1646	1567	
24	0,4247	0,2871	0,2295	0,1970	0,1759	0,1608	0,1495	0,1406	0,1338	0
30	3632	2412	1913	1635	1454	1327	1232	1157	1100	
40	2940	1915	1508	1281	1135	1033	0957	0898	0853	
60	0,2151	0,1371	0,1069	0,0902	0,0796	0,0722	0,0668	0,0625	0,0594	0
120	1225	0759	0585	0489	0429	0387	0357	0334	0316	
∞	0000	0000	0000	0000	0000	0000	0000	0000	0000	

Критические значения распределения Кочрена (k — число степеней сво

$\alpha = 0{,}05$

k	m								
	1	2	3	4	5	6	7	8	9
2	0,9985	0,9750	0,9392	0,9057	0,8772	0,8534	0,8332	0,8159	0,8010
3	9669	8709	7977	7457	7071	6771	6530	6333	6167
4	9065	7679	6841	6287	5895	5598	5365	5175	5017
5	0,8412	0,6338	0,5981	0,5440	0,5063	0,4783	0,4564	0,4387	0,4241
6	7808	6161	5321	4803	4447	4184	3980	3817	3682
7	7271	5612	4800	4307	3974	3726	3535	3384	3259
8	0,6798	0,5157	0,4377	0,3910	0,3595	0,3362	0,3185	0,3043	0,2926
9	6385	4775	4027	3584	3286	3067	2901	2868	2659
10	6020	4450	3733	3311	3029	2823	2666	2541	2439
12	0,5410	0,3924	0,3624	0,2880	0,2624	0,2439	0,2299	0,2187	0,2098
15	4709	3346	2758	2419	2195	2034	1911	1815	1736
20	3894	2705	2205	1921	1735	1602	1501	1422	1357
24	0,3434	0,2854	0,1907	0,1656	0,1493	0,1374	0,1286	0,1216	0,1160
30	2929	1980	1593	1377	1237	1137	1061	1002	0958
40	2370	1576	1259	1082	0968	0887	0827	0780	0745
60	0,1737	0,1131	0,0895	0,0765	0,0682	0623	0,0583	0,0552	0,0520
120	0998	0632	0495	0419	0371	0,0337	0312	0292	0279
∞	0000	0000	0000	0000	0000	0000	0000	0000	0000

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

К

Л

М

Н

О

П

СПИСОК ЛИТЕРАТУРЫ

1. *Бермант А.Ф., Араманович И.Г.* Краткий курс математического анализа для втузов. — М.: Наука, 1967.
2. *Бородин А.Н.* Элементарный курс теории вероятностей и математической статистики. — СПб.: Лань, 1998.
3. *Браунли К.А.* Статистическая теория и методология в науке и технике. — М.: Наука, 1977.
4. *Вентцель Е.С.* Теория вероятностей. — М., 1962.
5. *Гмурман В.Е.* Руководство к решению задач по теории вероятностей и математической статистике. — М.: Высш. школа, 1997.
6. *Гмурман В.Е.* Теория вероятностей и математическая статистика. — М.: Высш. школа, 1998.
7. *Зуховицкий С.И., Авдеева Л.И.* Линейное и выпуклое программирование. — М.: Наука, 1967.
8. *Кофман А., Фор Р.* Займемся исследованием операций. — М.: Мир, 1966.
9. *Морозов Ю.В.* Основы высшей математики и статистики. — М.: Медицина, 1998. — 232 с.
10. *Омельченко В.П., Курбатова Э.В., Антоненко Г.В.* Тестовые задания по высшей математике и информатике. — Р-н/Д, 2002. — 136 с.
11. Справочник по прикладной статистике, т. 1. — М.: Финансы и статистика, 1989.
12. Справочник по прикладной статистике, т. 2. — М.: Финансы и статистика, 1990.
13. *Тутубалин В.Н.* Теория вероятностей. — М.: Изд-во Моск. ун-та, 1972.
14. *Уилкс С.* Математическая статистика. — М., 1967.
15. *Хьютсон А.* Дисперсионный анализ. — М.: Статистика, 1971.

Учебное издание

Иван Васильевич Павлушков и другие.

ОСНОВЫ ВЫСШЕЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Подписано в печать 03.06.08. Формат $60x90^1/_{16}$. Бумага офсетная.
Печать офсетная. Объем 26,5 п.л. Тираж 2000 экз. Заказ № 762
Издательская группа «ГЭОТАР-Медиа».
119828, Москва, ул. Малая Пироговская, 1а,
тел./факс: (495) 921-39-07, 246-39-47
e-mail: info@geotar.ru, http://www.geotar.ru
Отпечатано в ООО «Чебоксарская типография №1».
428019, г. Чебоксары, пр. И. Яковлева, 15.

ISBN 978-5-9704-0880-3

9 785970 408803

ОСНОВЫ ВЫСШЕЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

В учебнике изложен курс высшей математики фармацевтического факультета, включающий основные элементарные функции, дифференциальное исчисление функции одной переменной, элементы дифференциального исчисления функций нескольких переменных, интегральное исчисление функции одной переменной, дифференциальные уравнения первого и второго порядка, основы теории вероятностей и математической статистики. Учебник содержит подробные пояснения теоретического материала, а также большое количество примеров и задач.

Предназначен студентам медицинских и фармацевтических вузов.

ISBN 978-5-9704-0880-3
9 785970 408803