

Г.М. КУЛИКОВ, А.Д. НАХМАН, С.В. ПЛОТНИКОВА

**ЭЛЕМЕНТЫ ПРИКЛАДНОЙ
МАТЕМАТИКИ**

ИЗДАТЕЛЬСТВО ТГТУ

УДК 51 (075)
ББК В19я73
К903

Рецензенты:

Доктор физико-математических наук, профессор, заведующий кафедрой «Алгебра и геометрия» ТГУ им. Г.Р. Державина
А.И. Булгаков

Кандидат технических наук, доцент кафедры
«Высшая математика» ТГТУ
В.А. Попов

Куликов, Г.М.

К903

Элементы прикладной математики : учебное пособие / Г.М. Куликов, А.Д. Нахман, С.В. Плотникова. – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2008. – 160 с. – 300 экз. – ISBN 978-5-8265-0781-0.

Изложены элементы комплексного анализа, теории числовых и функциональных рядов, теории вероятностей и математической статистики. Особенностью материала является его прикладная направленность.

Предназначено для студентов 2 курса инженерно-технических специальностей дневного отделения.

УДК 51 (075)

ББК В19я73

ISBN 978-5-8265-0781-0 © ГОУ ВПО «Тамбовский государственный
технический университет» (ТГТУ), 2008

Министерство образования и науки Российской Федерации
ГОУ ВПО «Тамбовский государственный технический университет»

Г.М. КУЛИКОВ, А.Д. НАХМАН, С.В. ПЛОТНИКОВА

ЭЛЕМЕНТЫ ПРИКЛАДНОЙ МАТЕМАТИКИ

*Утверждено Учёным советом университета в качестве
учебного пособия для студентов 2 курса
инженерно-технических специальностей
дневного отделения*



Тамбов
Издательство ТГТУ
2008

Учебное издание

КУЛИКОВ Геннадий Михайлович,
НАХМАН Александр Давидович,
ПЛОТНИКОВА Светлана Валерьевна

ЭЛЕМЕНТЫ ПРИКЛАДНОЙ МАТЕМАТИКИ

Учебное пособие

Редактор Ю.В. Ш и м а н о в а
Инженер по компьютерному макетированию М.А. Филатова

Подписано в печать 24.12.2008
Формат 60 × 84/16. 9,3 усл. печ. л. Тираж 300 экз. Заказ № 592.

Издательско-полиграфический центр
Тамбовского государственного технического университета
392000, Тамбов, Советская, 106, к. 14

ВВЕДЕНИЕ

Основу настоящего пособия составляет содержание курса лекций для студентов инженерно-технических специальностей, читаемых кафедрой «Прикладная математика и механика» Тамбовского государственного технического университета. Целью пособия является помощь студентам в овладении элементами комплексного анализа и вероятностно-статистического материала, необходимых в изучении специальных дисциплин и в их будущей профессиональной деятельности.

Отличие настоящего пособия от имеющихся учебных изданий состоит в том, что «под одной обложкой» собраны сведения из теории числовых и функциональных рядов действительного и комплексного переменного, теории вероятностей и математической статистики. Такой «синтез» не является искусственным: известно, что ряды – это средство исследований счётных вероятностных пространств [1], некоторых специальных вероятностных распределений (например, геометрического и пуассоновского); глубокие связи комплексного анализа, функциональных рядов и теории вероятностей наблюдаются в эргодической теории [2], математической физике [3] и др. Настоящий подход, как представляется авторам, является современным и обеспечивает системность знаний, формирование у студентов представления о целостности математической науки и универсальности её методов.

Другая особенность состоит в том, что ряды с действительными членами рассматриваются как частный случай комплексных рядов: это позволяет сделать изложение более компактным, особенно в части степенных рядов.

Статистический материал «увязан» с задачами практического характера, что обеспечивает его прикладную направленность. Имеется также ряд других отличительных особенностей настоящего пособия.

Контрольные задания к излагаемому материалу представлены в изданиях [4], [5].

1. ФУНКЦИИ КОМПЛЕКСНОГО ПЕРЕМЕННОГО

1.1. КОМПЛЕКСНЫЕ ЧИСЛА

1.1.1. В выбранной прямоугольной системе координат точка $(1, 0)$ соответствует числу 1 на числовой оси абсцисс (оси действительных чисел), а точка $(0, 1)$ – числу 1 на оси ординат. Чтобы отличать по написанию эти две единицы, последнюю обозначим буквой i и назовём мнимой единицей. Итак, точка $(0, 1)$ отождествляется с мнимой единицей; всякое же другое число оси ординат, отвечающее точке $(0, y)$, теперь естественно записать в виде yi и назвать чисто мнимым; сама ось OY будет далее называться мнимой осью (тогда как OX – действительная ось).

1.1.2. Произвольную упорядоченную пару x, y действительных чисел («комплекс» из двух действительных чисел), соответствующую точке (x, y) координатной плоскости, назовём комплексным числом.

Перейдём к так называемой алгебраической записи (форме) комплексного числа, употребив по аналогии с разложением радиус-вектора $\vec{z} = x\vec{e}_1 + y\vec{e}_2$ по базису $\{\vec{e}_1, \vec{e}_2\}$ (где \vec{e}_1 и \vec{e}_2 – единичные направляющие вектора координатных осей) для точки z с координатами (x, y) запись

$$z = x + yi. \quad (1.1.1)$$

Итак, между точками (x, y) и комплексными числами вида (1.1.1) установлено взаимно однозначное соответствие. Сама же плоскость (со введённой в ней прямоугольной системой координат) называется *комплексной* плоскостью. В частности, для действительного числа x естественна запись $x = x + 0i$, что соответствует точке $(x, 0)$; и теперь мы не делаем различия между действительными числами x и комплексными числами вида $x + 0i$. Для чисто мнимого yi , соответствующего точке $(0, y)$, употребима запись $yi = 0 + yi$. Таким образом, множество \mathbf{C} всех комплексных чисел содержит своими подмножествами \mathbf{R} (множество всех действительных чисел) и множество всех чисто мнимых чисел.

1.1.3. Числа вида $z = x + yi$ и $\bar{z} = x - yi$ называются *комплексно-сопряжёнными*, они изображаются точками, симметричными относительно оси OX . Модулем комплексного числа называется длина (модуль) радиус-вектора точки (x, y) , т.е.

$$|z| = \sqrt{x^2 + y^2}. \quad (1.1.2)$$

В частности, модуль действительного числа $x = x + 0i$ есть $\sqrt{x^2}$, т.е. он равен абсолютной величине числа x ; аналогично, $|yi| = \sqrt{y^2} = |y|$.

1.1.4. Действительной частью числа $z = x + yi$ называется x , а мнимой частью – число y ; применяем обозначения: $x = \operatorname{Re} z, y = \operatorname{Im} z$.

Комплексные числа $z_1 = x_1 + y_1i$ и $z_2 = x_2 + y_2i$ называются равными, если совпадают их действительные и мнимые части. Геометрически, соотношение $z_1 = z_2$ означает совпадение соответствующих точек комплексной плоскости.

Комплексные числа не сравниваются, т.е. во множестве \mathbf{C} не вводятся отношения «больше» и «меньше».

1.1.5. В п. 1.1.2 мы отождествили любое комплексное число $z = x + yi$ с радиус-вектором точки (x, y) . В связи с этим операция сложения и вычитания комплексных чисел $z_1 = x_1 + y_1i$ и $z_2 = x_2 + y_2i$ вводится по аналогии с такой же операцией над векторами, которая, в свою очередь, выполняется *покоординатно*. Так, по определению

$$z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)i.$$

Другими словами, сложение и вычитание комплексных чисел производят по тем же правилам, по которым эти действия производят над многочленами.

Произведение двух комплексных чисел $z_1 = x_1 + y_1i$ и $z_2 = x_2 + y_2i$ определим в виде

$$z_1 z_2 = (x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)i.$$

Имеем, в частности, квадрат комплексного числа z^2 в виде $z^2 = (x^2 - y^2) + 2xyi$; следовательно, $i^2 = 0 - 1 + 0i$, $i^2 = -1$.

В связи с таким свойством числа i его удобно обозначать в виде $i = \sqrt{-1}$; ясно, что $i \notin \mathbf{R}$; теперь становится понятно, почему число i названо *мнимой* единицей. Заметим, что умножение комплексных чисел выполняется по правилу умножения многочленов с заменой i^2 на -1 .

1.1.6. Деление комплексных чисел определяется как действие, обратное умножению. Именно,

$$z = \frac{z_1}{z_2}, \text{ если } z_1 = zz_2,$$

где $z_2 \neq 0$.

Решая конкретные примеры, можно пользоваться способом одновременного умножения числителя и знаменателя дроби на число, сопряжённое знаменателю.

1.1.7. Совместим стандартным образом прямоугольную и полярную системы координат: полярную ось направим по оси OX , полюс системы совмещаем с точкой $O(0, 0)$; выбираем в обеих системах одинаковые единицы масштаба; ось OY направляем по лучу $\varphi = \frac{\pi}{2}$. В этом случае прямоугольные координаты (x, y) и полярные координаты (ρ, φ) одной и той же точки z связаны соотношениями (рис. 1.1.1)

$$\begin{cases} x = \rho \cos \varphi; \\ y = \rho \sin \varphi. \end{cases}$$

Теперь комплексное число $z = x + yi$ принимает вид $z = \rho \cos \varphi + (\rho \sin \varphi)i$ или

$$z = \rho(\cos \varphi + i \sin \varphi). \quad (1.1.3)$$

Форма записи (1.1.3) комплексного числа называется *тригонометрической*.

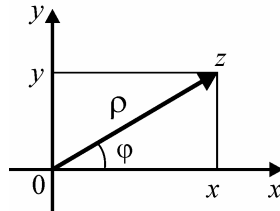


Рис. 1.1.1

Связь полярных и прямоугольных координат точки M может быть также представлена в виде

$$\operatorname{tg} \varphi = \frac{y}{x}, \quad \rho = \sqrt{x^2 + y^2}.$$

Следовательно, $\rho = |z|$ есть модуль числа z ; число φ назовём аргументом z . Обозначим через $\operatorname{arg} z$ одно из возможных значений φ : $-\pi < \operatorname{arg} z \leq \pi$; это значение назовём главным значением аргумента; иногда главное значение рассматриваем в интервале $[0, 2\pi)$; совокупность всех значений φ имеет вид

$$\operatorname{Arg} z = \operatorname{arg} z + 2\pi k, \quad k \in \mathbf{Z},$$

где \mathbf{Z} – множество всех целых чисел. Для точки $z=0$ значение аргумента не определено; очевидно, что $|0| = 0$.

1.1.8. Умножение, возведение в натуральную степень (т.е. умножение числа z на себя n раз) и деление комплексных чисел удобно выполнять, записав эти числа в тригонометрической форме. Легко проверить, что при умножении комплексных чисел в тригонометрической форме их модули перемножаются, а аргументы складываются; при делении – модули делятся, а аргументы вычитаются; при возведении в степень $n \in \mathbf{N}$ – модуль возводится в эту степень, а аргумент умножается на n . Так,

$$z^n = \rho^n (\cos n\varphi + i \sin n\varphi), \quad n = 1, 2, \dots \quad (1.1.4)$$

1.2. ФУНКЦИИ КОМПЛЕКСНОГО ПЕРЕМЕННОГО. ПРЕДЕЛ. НЕПРЕРЫВНОСТЬ

1.2.1. Определение. Пусть G – некоторое множество комплексных чисел. Говорят, что на множестве G (области определения G) задана *функция* вида $w = f(z)$, если каждому $z \in G$ поставлено в соответствие одно или несколько комплексных чисел w . В последнем случае мы говорим, что функция f *многозначна*.

Если, в частности, все значения w – действительные числа, то говорим о *действительнозначной функции комплексного переменного*. Если G – множество на «действительной оси» (оси абсцисс), т.е. $z = x \in \mathbf{R}$, то $w = f(x)$ – *комплекснозначная функция действительного переменного*.

Поскольку $w = f(z) = f(x + yi)$ определяется парами значений (x, y) , то можно говорить об f как функции двух действительных переменных, заданной на некотором множестве G . В то же время $w = u + vi$, тогда $u = \operatorname{Re} f(x + yi) = u(x, y)$, $v = \operatorname{Im} f(x + yi) = v(x, y)$ – две действительнозначные функции действительных переменных x и y . Таким образом,

$$w = f(z) = u(x, y) + iv(x, y), \quad (1.2.1)$$

т.е. задание $f(z)$ есть задание пары функций $u = u(x, y)$, $v = v(x, y)$, и этим облегчаются многие формулировки и доказательства в теории функций комплексного переменного.

Комплекснозначная функция вида $w = f(n)$, $n \in \mathbf{N}$ называется последовательностью комплексных чисел. Множество её значений имеет вид $\{w_1, w_2, \dots, w_n, \dots\}$. Часто термин «последовательность» употребляется и для обозначения множества $\{w_n\}$ всех получаемых значений функции. Согласно (1.2.1)

$$w = f(n) = w_n = u_n + iv_n, \quad n = 1, 2, \dots,$$

т.е. одновременно с $f(n)$ задаются две последовательности действительных чисел $\{u_n\}$ и $\{v_n\}$.

1.2.2. Функция вида $w = z^n$, $n \in \mathbf{Z}$ является примером однозначной, а $w = \text{Arg}z$, $w = \sqrt[n]{z}$ – примерами многозначных (бесконечнозначной и n -значной, соответственно) функций.

В основу определения комплекснозначной показательной функции положим известное свойство соответствующей (действительнозначной) функции $\varphi(x)$ (например, $\varphi(x) = e^x$) в случае действительного переменного: $\varphi(x+y) = \varphi(x)\varphi(y)$. Комплекснозначная функция $\varphi(y) = \cos y + i \sin y$, как легко проверить, обладает этим же свойством; в силу указанной причины положим по определению

$$e^{iy} = \cos y + i \sin y, \quad y \in \mathbf{R};$$

последнее соотношение называется формулой Эйлера. Примем теперь для всякого $z = x + iy$, по определению,

$$e^z = e^x (\cos y + i \sin y), \quad x, y \in \mathbf{R}.$$

В частности, случай действительного переменного $z = x + i0$, приводит к равенству $e^z = e^x$, так что имеем обобщение понятия показательной функции на комплексный случай. С выходом в комплексную плоскость экспонента приобретает некоторые новые непривычные свойства: она периодична с периодом $T = 2\pi i$, её значения могут быть отрицательными действительными числами, однако сохраняется свойство $e^z \neq 0$ при любых z .

Из формулы Эйлера легко вытекают соотношения

$$\cos y = \frac{e^{iy} + e^{-iy}}{2}, \quad \sin y = \frac{e^{iy} - e^{-iy}}{2i},$$

которые положим в основу определения тригонометрических функций

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

При переходе к комплексному аргументу сохраняются основные тригонометрические формулы, однако может нарушаться привычная ограниченность единицей модулей значений $\sin z$ и $\cos z$. Например,

$$\cos i = \frac{e^{i^2} + e^{-i^2}}{2} = \frac{1}{2} \left(e + \frac{1}{e} \right)$$

– действительное число, большее единицы.

По определению полагаем также

$$\text{tg } z = \frac{\sin z}{\cos z}, \quad \text{ctg } z = \frac{\cos z}{\sin z}$$

во всех точках z , где знаменатель соответствующей дроби не обращается в ноль. Гиперболические синус, косинус, тангенс и котангенс определяется, соответственно, в виде

$$\begin{aligned} \text{ch } z &= \frac{e^z + e^{-z}}{2}, & \text{sh } z &= \frac{e^z - e^{-z}}{2}, \\ \text{th } z &= \frac{\text{sh } z}{\text{ch } z}, & \text{cth } z &= \frac{\text{ch } z}{\text{sh } z}; \end{aligned}$$

в последних двух случаях исключаются из рассмотрения те значения z , для которых знаменатели обращаются в ноль.

Натуральным логарифмом числа z называется число w , обладающее свойством $e^w = z$, где $z \neq 0$. Формула для вычисления логарифма получается из следующих рассуждений. Если $w = u + iv$ и $z = |z|(\cos \varphi + i \sin \varphi)$, то, согласно определению,

$$e^w (\cos v + i \sin v) = |z|(\cos \varphi + i \sin \varphi).$$

Модули равных комплексных чисел равны, а аргументы могут отличаться на $2\pi k$:

$$e^u = |z|, \text{ откуда } u = \ln |z| \text{ и } v = \varphi + 2\pi k, \quad k = 0, \pm 1, \dots$$

Следовательно,

$$w = \ln |z| + i(\varphi + 2\pi k) \text{ или } \text{Ln } z = \ln |z| + i \text{Arg} z;$$

употребление заглавной буквы (в символе логарифма) означает многозначность результата.

Логарифмическая функция $w = \text{Ln } z$ определена при всех $z \neq 0$ и многозначна; при $k=0$ получаем так называемое

главное значение логарифма $\ln z = \ln|z| + i \arg z$.

В основу определения показательной функции положим известное (для действительного переменного) свойство

$$a^x = (e^{\ln a})^x = e^{x \ln a}, \quad a > 0.$$

Положим по определению для *любых комплексных* $a \neq 0$ и z

$$a^z = e^{z \ln a}.$$

Эта функция также оказывается многозначной в силу многозначности логарифма.

Обратные тригонометрические функции определяются как функции, обратные по отношению к синусу, косинусу, тангенсу, котангенсу. Они также оказываются многозначными; формулы для вычисления их значений читатель может найти в более подробных курсах.

1.2.3. Определение предела последовательности комплексных чисел вводится так же, как в случае последовательности действительных чисел: число (комплексное) A называется пределом последовательности $\{w_n\}$, если для любого $\varepsilon > 0$ найдётся такой номер N , что $|w_n - A| < \varepsilon$ для всех номеров $n > N$; применяют обозначение

$$\lim_{n \rightarrow \infty} w_n = A.$$

Описанное «предельное поведение» последовательности $\{w_n\}$ равносильно, очевидно, выполнению соотношения

$$\lim_{n \rightarrow \infty} |w_n - A| = 0. \quad (1.2.2)$$

В свою очередь, если $w_n = u_n + i v_n$ (т.е. u_n и v_n — соответственно, действительная и мнимая части w_n) и $A = a + ib$, то (1.2.2) равносильно одновременному выполнению соотношений

$$\lim_{n \rightarrow \infty} u_n = a \text{ и } \lim_{n \rightarrow \infty} v_n = b.$$

Это утверждение очевидным образом вытекает из оценки (см. (1.1.2))

$$\max\{|u_n - a|, |v_n - b|\} \leq \sqrt{(u_n - a)^2 + (v_n - b)^2} = |w_n - A|.$$

Говорят, также, что последовательность w_n имеет бесконечный предел, и записывают

$$\lim_{n \rightarrow \infty} w_n = \infty,$$

если

$$\lim_{n \rightarrow \infty} |w_n| = +\infty.$$

1.2.4. Пусть функция $w = f(z)$ определена в окрестности точки z_0 (т.е. в некотором круге с центром в z_0). Говорят, что число w_0 есть предел функции в этой точке, если

$$\lim_{|z - z_0| \rightarrow 0} |f(z) - w_0| = 0. \quad (1.2.3)$$

Предельный переход вида (1.2.3) равносильно тому, что одновременно

$$\lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} u(x, y) = u_0, \quad \lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} v(x, y) = v_0.$$

Другими словами, предельный переход совершается по отдельности в действительной и мнимой части функции $w = f(z)$. Отсюда вытекает, что простейшие свойства пределов (вынесение постоянного множителя за знак предела, предельный переход в сумме, произведении и т.п.) переносятся и на случай функций комплексного переменного.

В теории функций комплексного переменного вводятся также определения предела функции на бесконечности и определение бесконечного предела, на которых мы здесь не останавливаемся.

1.2.5. Функция $f(z)$ называется непрерывной в точке z_0 , внутренней для области определения G , если

$$\lim_{z \rightarrow z_0} f(z) = f(z_0). \quad (1.2.4)$$

Другими словами, непрерывность в точке z_0 есть возможность предельного перехода под знаком функции при $z \rightarrow z_0$.

Согласно п. 1.2.3 для $f(z) = u(x, y) + i v(x, y)$ непрерывность в точке $z_0 = x_0 + i y_0$ означает непрерывность действительной части u и мнимой части v как функций от x и y .

Как в случае функций действительного переменного, определению (1.2.4) можно придать иную форму. Если обозначить $\Delta z = z - z_0$, $\Delta w = f(z) - f(z_0)$, то непрерывность функции f в точке z_0 означает, что

$$\lim_{\Delta z \rightarrow 0} \Delta w = 0,$$

т.е. бесконечно — малому приращению аргумента (в точке z_0) соответствует бесконечно малое приращение функции f .

Функция $f(z)$ называется непрерывной в области G , если она непрерывна в каждой точке z этой области. Понятия области, односвязной, многосвязной областей читателю известны из курса математического анализа; см. также [6].

Вместе с $f(z)$ и $g(z)$ непрерывными будут (в их общей области непрерывности G) их сумма, произведение, частное. Справедливо утверждение о непрерывности сложной функции и др.

1.3. ПРОИЗВОДНАЯ

1.3.1. Пусть однозначная функция $w = f(z)$ определена в точке $z = x + iy$ и в некоторой её окрестности, а переменные x и y получают, соответственно, приращения Δx и Δy . Тогда $\Delta z = \Delta x + i\Delta y$ – соответствующее приращение переменной z . При переходе от точки z к точке $z + \Delta z$ (значения Δx , Δy предполагаем столь малыми, что точка $z + \Delta z$ расположена в той же окрестности) значение $w = f(z)$ получает некоторое приращение $\Delta w = f(z + \Delta z) - f(z)$.

Определение. Пусть существует предел вида

$$\lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z}. \quad (1.3.1)$$

Он называется производной функции $f(z)$ в точке z и обозначается $f'(z)$ либо w' , $\frac{dw}{dz}$, $\frac{df}{dz}$. Функция же $f(z)$ называется дифференцируемой в точке z .

Заметим, что в случае функции действительного переменного $\varphi(x)$ существование производной есть существование предела $\frac{\Delta \varphi}{\Delta x}$, когда Δx приближается к нулю вдоль оси абсцисс. В случае же (1.3.1) Δz приближается к нулю в комплексной плоскости по *любому* пути. Это и является причиной появления некоторых новых дополнительных свойств дифференцируемых функций в сравнении со случаем функций действительного переменного.

1.3.2. Из свойств пределов и определения (1.3.1) вытекает, что дифференцируемость $f(z)$ в точке z эквивалентна равенству

$$\frac{\Delta w}{\Delta z} - f'(z) = \alpha(z, \Delta z),$$

где $\alpha(z, \Delta z) \rightarrow 0$ при $\Delta z \rightarrow 0$. Следовательно, существование производной равносильно соотношению

$$\Delta w = f'(z)\Delta z + \alpha(z, \Delta z)\Delta z. \quad (1.3.2)$$

Выражение $dw = f'(z)\Delta z$ называется *дифференциалом* функции $f(z)$ в точке z .

Из (1.3.2) вытекает соотношение $\Delta w \rightarrow 0$ при $\Delta z \rightarrow 0$, а это означает, что дифференцируемость в точке z влечёт за собой *непрерывность* $f(z)$ в той же точке.

1.3.3. Имеют место те же правила дифференцирования, что и в случае функций действительного переменного; например, если $f(z) = C$, где $C = \text{const}$ (постоянное комплексное число), то $f'(z) = 0$; $(Cf(z))' = Cf'(z)$ и т.д. Сохраняется таблица производных в том же виде, что для функций действительного переменного.

Если $w = f(z)$ осуществляет взаимно-однозначное соответствие области G (в комплексной плоскости точек z) на \tilde{G} (в комплексной плоскости точек w), то определено (однозначное) обратное соответствие $z = \varphi(w)$, называемое обратной функцией. Справедлива формула

$$\varphi'(w) = \frac{1}{f'(z)}.$$

1.3.4. Пусть $z = x + iy$, и $w = f(z)$ определена в точке z и в некоторой её окрестности. Запишем $f(z)$ в виде

$$f(z) = u(x, y) + iv(x, y). \quad (1.3.3)$$

Необходимое условие дифференцируемости f в точке z содержится в следующем утверждении.

Теорема 1. Пусть $f(z)$ дифференцируема в точке z . Тогда существуют частные производные функций u и v по обоим переменным в точке (x, y) , причём в этой точке

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \quad (1.3.4)$$

Соотношения (1.3.4) называются условиями Коши-Римана.

▷ Пусть существует $f'(z)$, определяемая как предел вида (1.3.1). Выше отмечалось (см. п. 1.3.1), что характер стремления к нулю величины $\Delta z = \Delta x + i\Delta y$ может быть произвольным. Выберем, в частности, случаи:

1) $\Delta y = 0$, т.е. $\Delta z = \Delta x$, тогда $\Delta z \rightarrow 0$ означает, что $\Delta x \rightarrow 0$;

2) $\Delta x = 0$, т.е. $\Delta z = i\Delta y$, тогда $\Delta z \rightarrow 0$ одновременно с $\Delta y \rightarrow 0$.

В обоих случаях переход от точки z к точке Δz вызывает приращение $\Delta w = \Delta u(x, y) + i\Delta v(x, y)$. В первом случае каждое из приращений Δu и Δv есть *приращение по переменной x* , следовательно (см. (1.3.3))

$$\begin{aligned}
f'(z) &= \lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta x \rightarrow 0} \frac{\Delta_x u + i \Delta_x v}{\Delta x} = \\
&= \lim_{\Delta x \rightarrow 0} \left(\frac{\Delta_x u}{\Delta x} + i \frac{\Delta_x v}{\Delta x} \right) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}.
\end{aligned}
\tag{1.3.5}$$

При этом само существование $\frac{\partial u}{\partial x}$ и $\frac{\partial v}{\partial y}$ вытекает из существования пределов действительной и мнимой части (при $\Delta z \rightarrow 0$) функции («разностного отношения») $\frac{\Delta w}{\Delta z}$, тогда как сама эта функция имеет предел по условию теоремы.

Аналогично, во втором случае, $\Delta u = \Delta_y u$, $\Delta v = \Delta_y v$, т.е.

$$\begin{aligned}
f'(z) &= \lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta y \rightarrow 0} \frac{\Delta_y u + i \Delta_y v}{i \Delta y} = \\
&= \lim_{\Delta y \rightarrow 0} \left(\frac{\Delta_y v}{\Delta y} + \frac{i \Delta_y u}{i^2 \Delta y} \right) = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y}.
\end{aligned}
\tag{1.3.6}$$

Правые части соотношений (1.3.5) и (1.3.6) совпадают, так как выражают собою одну и ту же $f'(z)$:

$$\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \left(-\frac{\partial u}{\partial y} \right).$$

По определению равенства комплексных чисел имеем отсюда соотношения (1.3.4), что и требовалось доказать. \triangleleft

Замечание 1. В результате более детального рассмотрения можно было бы доказать, что для дифференцируемой в точке z функции f не только существуют указанные в (1.3.4) частные производные, но функции u и v дифференцируемы в точке (x, y) .

Замечание 2. Как установлено выше, $f'(z)$ можно вычислить по любой из указанных в (1.3.5) и (1.3.6) формул.

1.3.5. Достаточное условие дифференцируемости $f(z)$ в точке z содержится в следующем утверждении, которое мы приведём без доказательства.

Теорема 2. Если $u(x, y)$ и $v(x, y)$ дифференцируемы в точке (x, y) и выполнены условия Коши-Римана (1.3.4), то $f'(z)$ существует в точке $z = x + iy$.

Согласно теоремам 1 и 2 и замечанию 1 дифференцируемость функций u, v в точке (x, y) и выполнение условий Коши-Римана необходимы и достаточны для существования производной $f'(z)$.

1.3.6 *Определение.* Функция $w = f(z)$, дифференцируемая в точке z_0 и некоторой её окрестности, называется *аналитической в точке z_0* .

Функция, аналитическая во всех точках некоторой области G , называется *аналитической в этой области*.

Точки z комплексной плоскости, в которых однозначная $f(z)$ является аналитической, называются *правильными* точками этой функции, а все остальные точки (в частности, те, где $f(z)$ не определена) – *особыми* для $f(z)$.

Согласно п. 1.3.5 критерием аналитичности $f(z)$ в данной точке z (в данной области G) является дифференцируемость u, v и выполнение условий Коши-Римана в этой точке и некоторой её окрестности (в области G).

Так, например, функции $w = e^z$, $w = \sin z$, $w = \cos z$, $w = \operatorname{sh} z$, $w = \operatorname{ch} z$ оказываются аналитическими во всей комплексной плоскости, поскольку для каждой из них выполняются условия Коши-Римана в любой точке (в чём можно убедиться непосредственной проверкой, выделив соответствующие действительные и мнимые части).

1.4. ИНТЕГРАЛ ОТ ФУНКЦИИ КОМПЛЕКСНОГО ПЕРЕМЕННОГО

1.4.1. Понятие интеграла функции $w = f(z)$ по линии L вводится аналогично понятию криволинейного интеграла функции действительного переменного.

Пусть дуга $\cup AB$ линии L задаётся параметрически:

$$\begin{cases} x = x(t); \\ y = y(t), \end{cases} \quad \alpha \leq t \leq \beta;$$

при этом точка $M(x, y)$ совершает движение из положения A в положение B при изменении t от α до β . Будем считать, что $x'(t)$ и $y'(t)$ существуют, непрерывны на отрезке $[\alpha, \beta]$ и одновременно не обращаются в ноль. Иными словами, дуга $\cup AB$ задаётся с помощью уравнения

$$z = z(t), \text{ где } z(t) = x(t) + iy(t), \quad \alpha \leq t \leq \beta;$$

при этом $z'(t) = x'(t) + iy'(t)$ непрерывна на $[\alpha, \beta]$; дугу, заданную таким образом, называем *гладкой*.

Пусть далее $f(z)$ непрерывна в открытой области G и $\cup AB \subset G$. Разобьём произвольным образом эту дугу на части точками z_0, z_1, \dots, z_n в направлении от A к B , при этом z_0 совпадает с точкой A , z_n с точкой B . На каждой из частичных дуг $\cup z_{k-1}z_k$ произвольным образом выберем по точке η_k ($k = 1, 2, \dots, n$) и составим сумму (рис. 1.4.1)

$$\sum_{k=1}^n f(\eta_k) \Delta z_k, \quad (1.4.1)$$

где $\Delta z_k = z_k - z_{k-1}$ – вектор, идущий из точки z_{k-1} в точку z_k (рис. 1.4.1). Обозначим через s наибольшую из длин этих векторов (хорд): $s = \max_k |\Delta z_k|$.

Сумма (1.4.1) называется интегральной, а предел вида

$$\lim_{s \rightarrow 0} \sum_{k=1}^n f(\eta_k) \Delta z_k \quad (1.4.2)$$

– интегралом от функции $f(z)$ по дуге $\cup AB$; он обозначается символом $\int_{\cup AB} f(z) dz$.

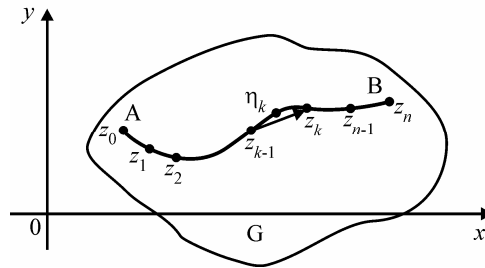


Рис. 1.4.1

Можно доказать, что при сформулированных выше условиях на функцию $f(z)$ и дугу линии L предел (1.4.2) существует и не зависит от способа разбиения $\cup AB$ на части точками z_k и от выбора «промежуточных» точек η_k .

1.4.2. Из определения п. 1.4.1 вытекают свойства интеграла, аналогичные свойствам криволинейных интегралов по координатам; например:

$$\int_{\cup AB} f(z) dz = - \int_{\cup BA} f(z) dz,$$

где $\cup BA$ – та же самая дуга, но с противоположным направлением обхода (от точки B к точке A);

$$\left| \int_{\cup AB} f(z) dz \right| \leq M \ell,$$

где M – любая постоянная, определяемая условием $|f(z)| \leq M$, $z \in \cup AB$, ℓ – длина дуги AB и др.

1.4.3. Вычисление интеграла (1.4.2) сводится к вычислению определённого интеграла комплексной функции действительной переменной t :

$$\int_{\cup AB} f(z) dz = \int_{\alpha}^{\beta} f(z(t)) z'(t) dt. \quad (1.4.3)$$

Формулой (1.4.3) следует пользоваться, применяя известные нам свойства интеграла (те же, что в случае действительной функции) и оперируя с числом i , как с обычной константой.

Формальная подстановка

$$f(z) = u(x, y) + iv(x, y) \quad \text{и} \quad dz = dx + idy$$

и вычисление произведения

$$f(z) dz = u(x, y) dx - v(x, y) dy + i(v(x, y) dx + u(x, y) dy)$$

приводят нас также к формуле

$$\int_{\cup AB} f(z) dz = \int_{\cup AB} u(x, y) dx - v(x, y) dy + i \int_{\cup AB} v(x, y) dx + u(x, y) dy. \quad (1.4.4)$$

Доказательство формул (1.4.3) и (1.4.4) производится сравнением интегральных сумм для интегралов в их правых частях с суммой (1.4.1).

1.4.4. Вычислим (в качестве примера) интеграл

$$J = \oint_{|z-z_0|=R} (z-z_0)^n dz,$$

где n – любое целое число, а обход контура интегрирования – окружности $|z - z_0| = R$ (z_0 и $R > 0$ – данные числа) – ведётся в направлении против часовой стрелки.

Решение. Уравнение окружности с центром в точке $z_0 = x_0 + iy_0$ радиуса R может быть записано в виде

$$\begin{cases} x - x_0 = R \cos t; \\ y - y_0 = R \sin t, \end{cases} \quad 0 \leq t \leq 2\pi.$$

Иначе говоря, $z - z_0 = R(\cos t + i \sin t)$, а тогда (см. (1.1.4)) при $n \neq -1$.

$$\begin{aligned} (z - z_0)^n &= R^n (\cos nt + i \sin nt), \quad dz = R(i \cos t - \sin t) dt, \text{ и} \\ J &= \int_0^{2\pi} R^{n+1} (\cos nt + i \sin nt) (i \cos t - \sin t) dt = \\ &= R^{n+1} \int_0^{2\pi} (-\sin t \cos nt - \sin nt \cos t + i(\cos t \cos nt - \sin t \sin nt)) dt = \\ &= -R^{n+1} \int_0^{2\pi} \sin(n+1)t dt + i R^{n+1} \int_0^{2\pi} \cos(n+1)t dt = 0. \end{aligned}$$

Если $n = -1$, то

$$\begin{aligned} J &= \int_0^{2\pi} \frac{R(i \cos t - \sin t)}{R(\cos t + i \sin t)} dt = \int_0^{2\pi} \frac{\cos(t + \pi/2) + i \sin(t + \pi/2)}{\cos t + i \sin t} dt = \\ &= \int_0^{2\pi} (\cos \pi/2 + i \sin \pi/2) dt = i \int_0^{2\pi} dt = 2\pi i. \end{aligned}$$

Итак,

$$\int_{|z-z_0|=R} (z - z_0)^n dz = \begin{cases} 0, & n \in \mathbf{Z}, \quad n \neq -1; \\ 2\pi i, & n = -1. \end{cases}$$

1.5. ИНТЕГРАЛЬНАЯ ТЕОРЕМА КОШИ

1.5.1. Пусть L – замкнутый контур, целиком расположенный в области G . Будем считать, что L задан уравнением $z = z(t)$ с непрерывной $z'(t)$, т.е. контур гладкий или хотя бы кусочно-гладкий.

Теорема Коши. Пусть $f(z)$ аналитична в G , и контур L ограничивает односвязную область $D \subset G$. Тогда

$$\oint_L f(z) dz = 0.$$

▷ Эту теорему легко доказать при дополнительном предположении, что $f'(z)$ – непрерывна. В силу формул (1.3.5) и (1.3.6) тогда будут непрерывными в G все частные производные первого порядка функций $u(x, y) = \operatorname{Re} f(z)$ и $v(x, y) = \operatorname{Im} f(z)$. При этих предположениях к каждому из криволинейных интегралов в представлении (см. (1.4.4))

$$\oint_L f(z) dz = \oint_L u(x, y) dx + (-v(x, y)) dy + i \oint_L v(x, y) dx + u(x, y) dy \quad (1.5.1)$$

можно применить формулу Грина:

$$\oint_L u dx + (-v) dy = \iint_D \left(\frac{\partial(-v)}{\partial x} - \frac{\partial u}{\partial y} \right) dx dy, \quad (1.5.2)$$

$$\oint_L v dx + u dy = \iint_D \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) dx dy, \quad (1.5.3)$$

обход контура L в криволинейных интегралах происходит против часовой стрелки.

Согласно условиям Коши-Римана (1.3.4) интегралы в правых частях (1.5.2) и (1.5.3) оба равны нулю. Следовательно, равны нулю и оба криволинейных интеграла; они остаются нулевыми, если изменить направление обхода L на противоположное.

Теперь, в силу равенства (1.5.1), получаем утверждение теоремы. ◁

1.5.2. Имеет место следующая *теорема Коши* для двусвязной области.

Пусть сохраняются предположения п. 1.5.1 области причём выбрано направление обхода функция $f(z)$ аналитична в области G и непрерыв-

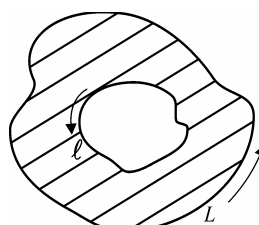


Рис. 1.5.1

для границ L (внешней) и l (внутренней) двусвязной каждой границы против часовой стрелки. Если на на L и l , то имеет место равенство (рис. 1.5.1)

$$\int_L f(z) dz = \int_l f(z) dz.$$

1.6. ФОРМУЛА КОШИ

1.6.1. Пусть функция $f(z)$ однозначна и аналитична в области G , L – контур (удовлетворяющий условиям п. 1.5.1), ограничивающий односвязную область D , целиком лежащий в G и обходимый против часовой стрелки. Тогда для любой точки z_0 , лежащей в D (т.е. расположенной внутри L) имеет место следующая интегральная формула:

$$f(z_0) = \frac{1}{2\pi i} \oint_L \frac{f(z) dz}{z - z_0}. \quad (1.6.1)$$

▷ Для любого $\varepsilon > 0$ выберем в D окружность $z = z_0 + \rho e^{i\varphi}$ (рис. 1.6.1) столь малого радиуса ρ , чтобы

$$|f(z) - f(z_0)| < \varepsilon; \quad (1.6.2)$$

это возможно, так как $f(z)$, будучи аналитической, является и непрерывной в D . Обозначим через γ указанную окружность, выберем на ней направление обхода против часовой стрелки и заметим, что

$$\frac{1}{2\pi i} \oint_\gamma \frac{dz}{z - z_0} = 1 \quad (1.6.3)$$

согласно примеру п. 1.4.4. Теперь по теореме Коши для двухсвязной области имеем

$$\oint_L \frac{f(z) dz}{z - z_0} = \oint_\gamma \frac{f(z) dz}{z - z_0}. \quad (1.6.4)$$

Указанная теорема применима, поскольку $\varphi(z) = \frac{f(z)}{z - z_0}$ аналитична вместе с $f(z)$ в области D , из которой исключена точка z_0 , так что, в частности, $\varphi(z)$ аналитична в области между контурами γ и L и на самих этих контурах.

Оценим разность между интегралом (1.6.1) и $f(z_0)$, воспользовавшись (1.6.3) и (1.6.4):

$$\begin{aligned} \left| \frac{1}{2\pi i} \oint_L \frac{f(z) dz}{z - z_0} - f(z_0) \right| &= \left| \frac{1}{2\pi i} \oint_\gamma \frac{f(z) dz}{z - z_0} - f(z_0) \frac{1}{2\pi i} \oint_\gamma \frac{dz}{z - z_0} \right| = \\ &= \left| \frac{1}{2\pi i} \oint_\gamma \frac{f(z) - f(z_0)}{z - z_0} dz \right|. \end{aligned} \quad (1.6.5)$$

Оценивая модуль интеграла (см. п. 1.4.2), имеем правую часть (1.6.5) не превосходящей

$$\frac{1}{2\pi i} \frac{\varepsilon}{\rho} 2\pi\rho = \varepsilon,$$

в силу неравенства (1.6.2) и соотношения $|z - z_0| = \rho$ на окружности γ . Теперь, в силу произвольности ε , имеем левую часть полученного неравенства

$$\left| \frac{1}{2\pi i} \oint_L \frac{f(z) dz}{z - z_0} - f(z_0) \right| < \varepsilon$$

равной нулю. ◁

1.6.2. В предположениях п. 1.6.1 в любой точке z_0 производная $f'(z)$ также оказывается аналитической функцией и имеет место формула

$$f'(z_0) = \frac{1}{2\pi i} \oint_L \frac{f(z) dz}{(z - z_0)^2}.$$

Вообще, для любого n в каждой точке $z_0 \in D$ существует производная $f^{(n)}(z_0)$ и для неё справедливо соотношение

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_L \frac{f(z)}{(z - z_0)^{n+1}} dz. \quad (1.6.6)$$

Формулу (1.6.6) можно получить формально дифференцированием по z_0 соотношения (1.6.1) n раз под знаком интеграла.

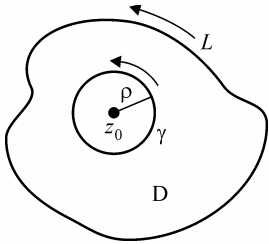


Рис. 1.6.1

2. ЧИСЛОВЫЕ И ФУНКЦИОНАЛЬНЫЕ РЯДЫ

2.1. ЧИСЛОВЫЕ РЯДЫ. ОСНОВНЫЕ ПОНЯТИЯ. ПРОСТЕЙШИЕ СВОЙСТВА

2.1.1. Пусть дана бесконечная последовательность комплексных чисел $w_1, w_2, w_3, \dots, w_n, \dots$. Формально составленная бесконечная сумма вида

$$w_1 + w_2 + w_3 + \dots + w_n + \dots$$

или, коротко,

$$\sum_{n=1}^{\infty} w_n \quad (2.1.1)$$

называется числовым рядом; общий член последовательности $\{w_n\}$ называется общим членом ряда (2.1.1).

Обозначим через

$$S_n = w_1 + w_2 + w_3 + \dots + w_n$$

n -ю частичную сумму ряда (2.1.1); при этом, по определению, $S_1 = w_1$.

Если существует предел вида

$$S = \lim_{n \rightarrow \infty} S_n, \quad (2.1.2)$$

то числовой ряд (2.1.1) называется сходящимся, а в противном случае – расходящимся.

Число S назовём суммой сходящегося ряда; говорят также, что ряд (2.1.1) сходится к сумме S и применяют запись

$$S = \sum_{n=1}^{\infty} w_n.$$

Главная задача, которая решается в теории числовых рядов – сходится или расходится данный ряд; вопрос о его сумме можно ставить лишь тогда, когда доказана сходимости. Сумму же сходящегося ряда всегда можно вычислить приближённо, взяв достаточно большое количество n членов в составе его частичной суммы S_n ; при этом точность вычисления увеличивается с ростом n .

2.1.2. Пример. Исследовать сходимость ряда

$$\sum_{n=1}^{\infty} \frac{1}{n}. \quad (2.1.3)$$

Решение. Данный ряд состоит из действительных чисел; исследование разобьём на несколько этапов.

1. Поведение частичных сумм ряда (2.1.3) определится следующей оценкой его общего члена:

$$\frac{1}{n} > \ln(n+1) - \ln n.$$

2. Доказательство этой оценки основано на неравенстве

$$\ln(1+x) < x, \quad x > 0, \quad (2.1.4)$$

которое мы сейчас установим (читателю рекомендуется изобразить графики левой и правой части неравенства). Разность левой и правой частей (2.1.4)

$$y(x) = \ln(1+x) - x$$

– убывающая функция, поскольку $y'(x) = \frac{1}{x+1} - 1 < 0$ при $x > 0$.

Кроме того очевидно, что $y(0) = 0$; значит разность $y(x)$ остаётся отрицательной: $\ln(1+x) - x < 0$ при всех $x > 0$. Это и утверждалось в соотношении (2.1.4).

3. Выбирая $x = \frac{1}{n}$ в (2.1.4), имеем неравенство

$$\frac{1}{n} > \ln\left(1 + \frac{1}{n}\right) = \ln \frac{n+1}{n} = \ln(n+1) - \ln n,$$

для общего члена ряда, которое мы и хотели установить.

4. Теперь частичная сумма (2.1.3) n -го порядка

$$S_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n-1} + \frac{1}{n}$$

имеет оценку снизу:

$$S_n > (\ln 2 - \ln 1) + (\ln 3 - \ln 2) + \dots + (\ln n - \ln(n-1)) + \\ + (\ln(n+1) - \ln n) = \ln(n+1),$$

откуда вытекает, что $S_n \rightarrow \infty$ вместе с $\ln(n+1)$ при $n \rightarrow \infty$.

Итак, исследуемый ряд расходится.

Замечание. Указанный ряд называется *гармоническим*. Ниже будет рассмотрен более общий случай: обобщенный гармонический ряд (ряд Дирихле).

2.1.3. Установим некоторые свойства сходящихся рядов. Пусть даны произвольные комплексные числа τ, ρ и сходящиеся числовые ряды

$$\sum_{n=1}^{\infty} u_n, \quad \sum_{n=1}^{\infty} v_n, \quad (2.1.5)$$

суммы которых равны U и V соответственно.

Тогда ряд

$$\sum_{n=1}^{\infty} (\tau u_n + \rho v_n) \quad (2.1.6)$$

сходится и его сумма равна $\tau U + \rho V$.

▷ Доказательство легко следует из определений сходимости и суммы ряда. Исключим из рассмотрения случай $\tau = \rho = 0$, в котором утверждение становится очевидным (сумма ряда, состоящего из нулей, равна нулю) и запишем n -ю частичную сумму исследуемого ряда (2.1.6):

$$\begin{aligned} S_n &= (\tau u_1 + \rho v_1) + (\tau u_2 + \rho v_2) + \dots + (\tau u_n + \rho v_n) = \\ &= \tau(u_1 + u_2 + \dots + u_n) + \rho(v_1 + v_2 + \dots + v_n) = \tau U_n + \rho V_n, \end{aligned}$$

где U_n, V_n – частичные суммы соответствующих рядов (2.1.5). В силу их сходимости имеем

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \tau U_n + \lim_{n \rightarrow \infty} \rho V_n = \tau U + \rho V.$$

Итак, ряд (2.1.6) оказался (на основании определения) сходящимся к сумме $\tau U + \rho V$. ◁

В частности, при $\rho = 0$ получаем, что ряд

$$\sum_{n=1}^{\infty} \tau u_n$$

имеет то же поведение, что и

$$\sum_{n=1}^{\infty} u_n.$$

Если исходный ряд был сходящимся, то его сумма умножится на τ .

2.1.4. Пусть $w_n = u_n + i v_n$, $n = 1, 2, \dots$, так что u_n – действительная часть, а v_n – мнимая часть w_n . Ряд (2.1.1) тогда можно записать в виде

$$\sum_{n=1}^{\infty} (u_n + i v_n).$$

Применяя доказанное в п. 2.1.3 свойство, получаем следующее утверждение.

Если сходятся ряды

$$\sum_{n=1}^{\infty} u_n, \quad \sum_{n=1}^{\infty} v_n, \quad (2.1.7)$$

составленные из действительных и мнимых частей последовательности w_n , то сходится и ряд (2.1.1).

Верно и обратное: если сходится ряд (2.1.1), то имеет место сходимость обоих рядов (2.1.7); утверждение вытекает из свойств п. 1.2.3, поскольку последовательности частичных сумм (n -го порядка) рядов (2.1.7) представляют собою, соответственно, действительную и мнимую часть сходящейся последовательности S_n .

2.1.5. По заданной бесконечной последовательности $\{w_n\}$ построим теперь ряд вида

$$w_{N+1} + w_{N+2} + \dots, \quad (2.1.8)$$

где $N = 1, 2, \dots$ и назовем его N -м остатком ряда (2.1.1); иными словами, N -й остаток (2.1.1) есть ряд, полученный отбрасыванием первых N членов.

Обозначим при $n > N$ через $S_{N,n}$ n -ю частичную сумму ряда-остатка (2.1.8):

$$S_{N,n} = w_{N+1} + \dots + w_n$$

и, в случае его сходимости, через $R = R_N$ – сумму этого ряда, т.е.

$$R_N = \lim_{n \rightarrow \infty} S_{N,n}. \quad (2.1.9)$$

Докажем, что ряд (2.1.1) сходится тогда и только тогда, когда сходится каждый его ряд-остаток (2.1.8). Другими словами, установим, что отбрасывание или добавление конечного числа (первых) членов не влияет на сходимость данного ряда.

Действительно, при $n > N$ частичная сумма ряда (2.1.1) есть

$$S_n = w_1 + \dots + w_N + w_{N+1} + \dots + w_n = S_N + S_{N, n}, \quad (2.1.10)$$

откуда следует, что S_n и $S_{N, n}$ отличаются на фиксированную величину S_N , а значит одновременно имеют или не имеют предел при $n \rightarrow \infty$. Итак, ряды (2.1.1) и (2.1.8) сходятся или расходятся одновременно.

Заметим также, что путём предельного перехода при $n \rightarrow \infty$ в соотношении (2.1.10) получается равенство

$$S = S_N + R_N, \quad N = 1, 2, \dots \quad (2.1.11)$$

Установим также следующее свойство остатка: если (2.1.1) является сходящимся, то

$$\lim_{N \rightarrow \infty} R_N = 0.$$

Действительно, если R_N сумма N -го остатка (см. (2.1.9)), то, устремляя $N \rightarrow \infty$ в (2.1.11), имеем стремление к нулю последовательности R_N , что и утверждалось.

2.2. НЕОБХОДИМЫЙ ПРИЗНАК СХОДИМОСТИ РЯДА. СУММА ГЕОМЕТРИЧЕСКОЙ ПРОГРЕССИИ

2.2.1. Теорема (необходимый признак сходимости ряда). Если ряд (2.1.1) сходится, то существует предел (при $n \rightarrow \infty$) его общего члена w_n , причём

$$\lim_{n \rightarrow \infty} w_n = 0. \quad (2.2.1)$$

Обратное утверждение неверно.

▷ Так как, очевидно, соотношение $S = \lim_{n \rightarrow \infty} S_n$ может быть записано и в виде $S = \lim_{n \rightarrow \infty} S_{n-1}$, то, вычисляя для $w_n = S_n - S_{n-1}$ предел разности, имеем

$$\lim_{n \rightarrow \infty} w_n = \lim_{n \rightarrow \infty} (S_n - S_{n-1}) = \lim_{n \rightarrow \infty} S_n - \lim_{n \rightarrow \infty} S_{n-1} = S - S = 0,$$

что и требовалось установить. ◁

Ряд с общим членом $w_n = \frac{1}{n} + i0$ является известным нам расходящимся гармоническим рядом (см. п. 2.1.2), и для таких w_n выполнено соотношение (2.2.1). Значит утверждение, обратное сформулированному в теореме, неверно.

Следствие (достаточный признак расходимости). Если

$$\lim_{n \rightarrow \infty} |w_n| \neq 0 \quad (2.2.2)$$

(или если предел не существует), то ряд расходится.

Действительно, в противном случае мы имели бы существование и равенство нулю предела вида (2.2.1), откуда бы следовало, что $|w_n| \rightarrow 0$, что противоречит условию (2.2.2).

2.2.2. Пусть a и q – ненулевые комплексные числа. Рассмотрим бесконечную геометрическую прогрессию

$$a, aq, aq^2, \dots, aq^n, \dots,$$

ряд, составленный из её членов

$$a + aq + aq^2 + \dots + aq^n + \dots = \sum_{n=0}^{\infty} aq^n \quad (2.2.3)$$

и исследуем его сходимость.

Случай 1: $|q| \geq 1$; в этом случае $|aq^n| = |a| \cdot |q|^n \geq |a|$. Могут представиться две возможности: либо

$$\lim_{n \rightarrow \infty} |aq^n|$$

не существует, либо он существует и согласно неравенству $|q| \geq 1$ его значение не меньше числа $|a| > 0$. В обоих случаях, по достаточному признаку расходимости ряда, получаем, что (2.2.3) расходится.

Случай 2: $|q| < 1$; в этом случае n -я частичная сумма ряда (2.2.3) имеет вид

$$S_n = \frac{a(1 - q^{n+1})}{1 - q}$$

(формула суммы первых членов геометрической прогрессии известна из школьного курса, причём её доказательство сохра-

няется и для прогрессий с комплексными членами). Вычислим теперь предел последовательности частичных сумм:

$$\lim_{n \rightarrow \infty} S_n = \frac{a}{1-q} \left(1 - \lim_{n \rightarrow \infty} q^n \right)$$

Последний предел существует, так как очевидно, что $\lim_{n \rightarrow \infty} q^n = 0$ при $|q| < 1$. Теперь

$$\lim_{n \rightarrow \infty} S_n = \frac{a}{1-q},$$

т.е. ряд оказался сходящимся к сумме:

$$S = \frac{a}{1-q}.$$

Итак, мы установили, что ряд (2.2.3) с $a \neq 0$ является сходящимся тогда и только тогда, когда $|q| < 1$, и нашли в этом случае его сумму.

2.3. СХОДИМОСТЬ РЯДОВ С ПОЛОЖИТЕЛЬНЫМИ ЧЛЕНАМИ: ПРИЗНАКИ СРАВНЕНИЯ

2.3.1. Рассмотрим тот важнейший случай, когда ряд

$$\sum_{n=1}^{\infty} a_n \tag{2.3.1}$$

составлен из действительных положительных чисел, т.е. порождён последовательностью $\{a_n\}$, $a_n \in \mathbf{R}$, $a_n > 0$, $n = 1, 2, \dots$; такой ряд называют знакоположительным. Обозначим через S_n частичную сумму ряда n -го порядка.

В вопросах исследования знакоположительных рядов потребуется следующее вспомогательное утверждение.

Лемма. Если последовательность $\{S_n\}$ ограничена сверху, то ряд (2.3.1) сходится.

▷ С ростом n последовательность $\{S_n\}$ возрастает, так как в частичной сумме будут добавляться положительные члены. Кроме того, по условию, эта последовательность ограничена. Но, как известно из анализа, всякая возрастающая ограниченная последовательность имеет предел; в нашем случае существует (конечный) предел вида $\lim_{n \rightarrow \infty} S_n$. Это и означает сходимость ряда (2.3.1). ◁

2.3.2. Одним из способов исследования сходимости знакоположительного ряда является сравнение его общего члена с общим членом некоторого ряда с известным поведением («эталонного ряда»). Примером эталонного является ряд, составленный из членов бесконечной геометрической прогрессии (см. п. 2.2.2). Другие примеры см. в параграфе 2.5.

Теорема 1 (сравнения). Пусть даны два знакоположительных ряда:

$$\sum_{n=1}^{\infty} a_n \tag{2.3.2}$$

и

$$\sum_{n=1}^{\infty} b_n \tag{2.3.3}$$

и при всех $n = 1, 2, \dots$ имеет место неравенство

$$a_n \leq b_n. \tag{2.3.4}$$

Тогда: 1) если сходится ряд (2.3.3) (к некоторой сумме B), то сходится и ряд (2.3.2) (к некоторой сумме A); при этом для их сумм имеет место соотношение $A \leq B$; 2) если ряд (2.3.2) расходится, то расходится и ряд (2.3.3).

Замечание. Согласно свойству п. 2.1.5 (отбрасывание или добавление конечного числа членов не влияет на сходимость ряда) утверждение теоремы имеет место, если соотношение (2.3.4) выполняется не при всех n , а лишь начиная с некоторого номера N .

▷ 1. Если ряд (2.3.3) сходится, то последовательность $\{B_n\}$ его частичных сумм (как сходящаяся последовательность) ограничена сверху некоторой постоянной C : $B_n \leq C$. Если также A_n – последовательность частичных сумм ряда (2.3.2), то из неравенства (2.3.4) вытекает, что

$$A_n \leq B_n \leq C \tag{2.3.5}$$

при всех n . Следовательно, последовательность A_n ограничена сверху, а тогда по лемме п. 2.3.1 ряд (2.3.2) сходится. Переходя к пределу в неравенстве (2.3.5), получаем также $A \leq B$. Утверждение 1 установлено.

2) Если ряд (2.3.2) расходится, то (2.3.3) не может быть, согласно (2.3.4), сходящимся по доказанному утверждению 1: тогда, обязан сходиться и ряд (2.3.2). Утверждение 2 доказано. ◁

2.3.3. *Теорема 2* (сравнения в предельной форме). Пусть даны два знакоположительных ряда (2.3.2) и (2.3.3), причём существует предел вида

$$L = \lim_{n \rightarrow \infty} \frac{a_n}{b_n}, \quad L > 0, \quad L < \infty. \tag{2.3.6}$$

Тогда ряды (2.3.2) и (2.3.3) сходятся или расходятся одновременно.

▷ Согласно (2.3.6) и определению предела, для каждого ε , такого что $0 < \varepsilon < L$ существует номер N , такой что неравенство

$$\left| \frac{a_n}{b_n} - L \right| < \varepsilon$$

имеет место при всех $n > N$. Из последнего соотношения (при указанных n) вытекает, что

$$-\varepsilon < \frac{a_n}{b_n} - L < \varepsilon$$

или, одновременно,

$$a_n < (L + \varepsilon)b_n, \quad b_n < \frac{1}{L - \varepsilon} a_n. \quad (2.3.7)$$

Согласно выбору ε имеют место оценки $L + \varepsilon > 0$ и $L - \varepsilon > 0$. Тогда, согласно свойствам п. 2.1.3, ряд с общим членом $a_n / (L - \varepsilon)$ ведёт себя так же, как (2.3.2), а ряд с членами $(L + \varepsilon)b_n$ – как (2.3.3). Теперь, в силу теоремы 1 и замечанию к ней, первое неравенство в (2.3.7) будет означать, что из сходимости (2.3.3) вытекает сходимость (2.3.2), а из расходимости (2.3.2) – расходимость (2.3.3). Аналогичные утверждения следует из второго неравенства в (2.3.7), если «поменять ролями» (2.3.2) и (2.3.3). Таким образом, поведение рядов (3.3.2) и (3.3.3) – одинаково, что и утверждалось. ◁

2.3.4. Использование признаков сравнения знакоположительных рядов предполагает наличие некоторого эталона для сравнения. Было бы полезно дополнить список признаков такими, которые позволяли бы исследовать поведение ряда, исходя лишь из вида его общего члена. Такие признаки предлагаются в следующих параграфах.

2.4. СХОДИМОСТЬ РЯДОВ С ПОЛОЖИТЕЛЬНЫМИ ЧЛЕНАМИ: ПРИЗНАКИ КОШИ И ДАЛАМБЕРА

2.4.1. Пусть дан знакоположительный ряд (2.3.1).

Теорема (радикальный признак Коши). Пусть существует предел вида

$$K = \lim_{n \rightarrow \infty} \sqrt[n]{a_n}. \quad (2.4.1)$$

Если $K < 1$, то ряд (2.3.1) сходится; если же $K > 1$, то ряд расходится.

Замечание. В случае $K = 1$ признак Коши не даёт ответа на вопрос о сходимости ряда: существуют примеры как сходящихся, так и расходящихся рядов, для которых $K = 1$.

▷ Согласно (2.4.1) и определению предела, для каждого $\varepsilon > 0$ существует номер N , такой что неравенство

$$\left| \sqrt[n]{a_n} - K \right| < \varepsilon$$

имеет место при всех $n > N$. Из последнего соотношения (при указанных n) вытекает, что

$$K - \varepsilon < \sqrt[n]{a_n} < K + \varepsilon. \quad (2.4.2)$$

Случай 1: $K < 1$. Ввиду произвольности выбора ε положим $0 < \varepsilon < 1 - K$ и обозначим $q = K + \varepsilon$, так что $0 < q < 1$. Из (2.4.2) вытекает тогда, что $a_n < (K + \varepsilon)^n$ или $a_n < q^n$ при всех $n > N$. Поскольку сумма членов геометрической прогрессии

$$\sum_{n=N+1}^{\infty} q^n$$

является сходящимся рядом, то по первой теореме сравнения (см. также замечание к ней) сходится и ряд (2.3.1).

Случай 2: $K > 1$. В этом случае выберем ε так, что $0 < \varepsilon < K - 1$ и обозначим $Q = K - \varepsilon$, так что $Q > 1$. Из (2.4.2) вытекает тогда, что $a_n > (K - \varepsilon)^n$ или $a_n > Q^n$ при всех n , начиная с некоторого номера N . Но в этом случае члены ряда (2.3.1) не могут стремиться к нулю и ряд расходится по достаточному признаку расходимости.

Теорема полностью доказана. ◁

2.4.2. *Теорема* (признак Даламбера). Пусть существует предел вида

$$D = \lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}. \quad (2.4.3)$$

Если $D < 1$, то ряд (2.3.1) сходится; если же $D > 1$, то ряд расходится.

Замечание. В случае $D = 1$ (подобно признаку Коши) теорема 2 не даёт ответа на вопрос о сходимости ряда.

Доказательство мы не приводим, но его идея та же, что и в случае теоремы 1. Отметим только (это потребуется в дальнейшем), что при $D > 1$ расходимость ряда имеет место на основании достаточного признака расходимости (см. доказательство признака Коши).

Пример. Исследовать сходимость ряда.

$$\sum_{n=1}^{\infty} \left(\frac{n+1}{n} \right)^{n^2}.$$

Решение. Имеем знакоположительный ряд с общим членом

$$a_n = \left(\frac{n+1}{n} \right)^{n^2},$$

вид которого позволяет использовать признак Коши:

$$K = \lim_{n \rightarrow \infty} \sqrt[n]{a_n} = \lim_{n \rightarrow \infty} \left(\frac{n+1}{n} \right)^n = e.$$

Поскольку $K = e > 1$, то данный ряд расходится.

2.5. СХОДИМОСТЬ РЯДОВ С ПОЛОЖИТЕЛЬНЫМИ ЧЛЕНАМИ: ИНТЕГРАЛЬНЫЙ ПРИЗНАК КОШИ

2.5.1. Следующий признак позволяет свести вопрос об исследовании сходимости знакоположительного ряда к более знакомой задаче об исследовании сходимости несобственного интеграла.

Рассмотрим аналитическое выражение общего члена a_n (формулу, которой он задан) ряда (2.3.1) и заменим в ней n на x . В результате получим некоторую функцию $a(x)$. Пусть эта функция непрерывна и убывает на промежутке $[1, \infty)$.

Теорема (интегральный признак Коши). Если несобственный интеграл

$$\int_1^{\infty} a(x) dx \tag{2.5.1}$$

сходится, то сходится и ряд (2.3.1); если же интеграл (2.5.1) расходится, то расходится и ряд.

▷ Доказательство основано на двойном неравенстве

$$S_n - a_1 < \int_1^n a(x) dx < S_{n-1}.$$

Чтобы его доказать, используем следующие рассуждения геометрического характера (рис. 2.5.1).

Значение $\int_1^n a(x) dx$ равно площади криволинейной трапеции, ограниченной сверху графиком $y = a(x)$, основанием которой является отрезок $[1, n]$. Точки с координатами (n, a_n) расположены на графике $y = a(x)$, а

$$S_n - a_1 = a_2 + a_3 + \dots + a_n.$$

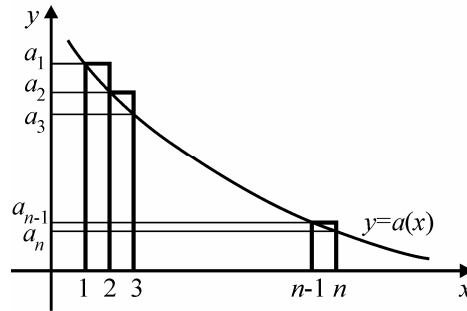


Рис. 2.5.1

Первое слагаемое $a_2 = 1 \cdot a_2$ численно равно площади прямоугольника, основание которого есть отрезок $[1, 2]$ оси абсцисс, а высота h равна a_2 ; второй член – площади прямоугольника с основанием $[2, 3]$ и высотой $h = a_3$; ...; последний член суммы a_n численно совпадает с площадью прямоугольника, имеющего основанием отрезок $[n-1, n]$ и высоту, равную a_n . Полученная ступенчатая фигура, состоящая из указанных прямоугольников, является «вписанной» по отношению к криволинейной трапеции (см. рис. 2.5.1), а значит имеет площадь, меньшую, чем криволинейная трапеция, так что неравенство

$$S_n - a_1 < \int_1^n a(x) dx \tag{2.5.2}$$

доказано.

Аналогичны рассуждения в случае второго неравенства: сумма $S_{n-1} = 1 \cdot a_1 + 1 \cdot a_2 + \dots + 1 \cdot a_{n-1}$ численно равна площади «описанной» ступенчатой фигуры, состоящей из прямоугольников, основания которых – отрезки $[1, 2], [2, 3], \dots, [n-1, n]$, а высоты равны, соответственно, a_1, a_2, \dots, a_{n-1} . Следовательно, площадь этой фигуры больше, чем площадь криволинейной трапеции, т.е.

$$S_{n-1} > \int_1^n a(x) dx \tag{2.5.3}$$

В случае сходимости несобственного интеграла (2.5.1) из неравенства (см. (2.5.2))

$$S_n < a_1 + \int_1^n a(x) dx < a_1 + \int_1^\infty a(x) dx$$

получаем ограниченность последовательности частичных сумм, а значит (на основании леммы п. 2.3.1) и её сходимости. В случае же расходимости несобственного интеграла и оценки (2.5.3) имеем неограниченность S_{n-1} , и, следовательно, расходимость ряда (2.3.1).

Итак, ряд ведёт себя так же, как несобственный интеграл (2.5.1), что и требовалось доказать. \triangleleft

2.5.2. Рассмотрим обобщенный гармонический ряд (называемый также рядом Дирихле)

$$\sum_{n=0}^{\infty} \frac{1}{n^p}, \quad p \in \mathbf{R}. \quad (2.5.4)$$

Докажем, что ряд сходится при $p > 1$ и расходится при остальных действительных значениях p .

Начнём со случая $p > 0$, $p \neq 1$ и применим интегральный признак Коши. Заменяя в записи общего члена ряда n на x , получим функцию $a(x) = \frac{1}{x^p}$, $x \in [1, \infty)$. Ясно, что на указанном промежутке функция $a(x)$ непрерывна и убывает. Исследуем теперь несобственный интеграл (2.5.1).

Если $p > 1$, то

$$\begin{aligned} \int_1^\infty a(x) dx &= \int_1^\infty \frac{1}{x^p} dx = \\ &= \lim_{T \rightarrow \infty} \int_1^T x^{-p} dx = \lim_{T \rightarrow \infty} \left. \frac{x^{1-p}}{1-p} \right|_1^T = \frac{1}{1-p} \lim_{T \rightarrow \infty} \left(\frac{1}{T^{p-1}} - 1 \right) = \frac{1}{p-1}. \end{aligned}$$

Итак, исследуемый интеграл оказался сходящимся, откуда и следует сходимость ряда (2.5.4).

Если $0 < p < 1$, то тот же интеграл вычисляется в виде

$$\int_1^\infty a(x) dx = \frac{1}{1-p} \lim_{T \rightarrow \infty} (T^{1-p} - 1) = \infty,$$

откуда следует расходимость ряда (3.5.4).

В случае $p = 1$ снова применяем интегральный признак Коши с $a(x) = \frac{1}{x}$:

$$\begin{aligned} \int_1^\infty a(x) dx &= \int_1^\infty \frac{dx}{x} = \\ &= \lim_{T \rightarrow \infty} \ln x \Big|_1^T = \lim_{T \rightarrow \infty} \ln T = \infty, \end{aligned}$$

так что ряд (2.5.4) расходится; тем самым ещё раз установлена расходимость гармонического ряда.

Наконец, при $p \leq 0$ имеем $a_n = n^{-p} \geq 1$, так что общий член ряда не стремится к нулю, а тогда (2.5.4) расходится по достаточному признаку расходимости ряда.

Замечание. Если к гармоническому ряду применить признаки Коши и Даламбера, то, как нетрудно проверить, получится соответственно $K=1$ и $D=1$. В то же время условия $K=1$ и $D=1$ выполняются и для членов сходящегося ($p > 1$) ряда Дирихле. Приведённые примеры подтверждают ранее сделанный вывод, что по одной только информации вида $K=1$ и $D=1$ о поведении ряда судить нельзя; следует провести дополнительное исследование: например, применить другие признаки.

2.6. ЗНАКОЧЕРЕДУЮЩИЕСЯ РЯДЫ

2.6.1. Рассмотрим знакоположительную последовательность

$$\{a_n\}, \quad a_n \in \mathbf{R}, \quad a_n > 0, \quad n = 1, 2, \dots \quad (2.6.1)$$

Ряд вида

$$a_1 - a_2 + a_3 - \dots + (-1)^{n-1} a_n + \dots = \sum_{n=1}^{\infty} (-1)^{n-1} a_n \quad (2.6.2)$$

называется *знакопередающим*. Достаточным признаком его сходимости является следующий признак Лейбница.

Теорема. Если последовательность (2.6.1) является убывающей, т.е.

$$a_1 > a_2 > \dots > a_n > \dots \quad (2.6.3)$$

и

$$\lim_{n \rightarrow \infty} a_n = 0, \quad (2.6.4)$$

то знакопередающий ряд (2.6.2) сходится и его сумма S удовлетворяет условиям

$$0 \leq S \leq a_1. \quad (2.6.5)$$

\triangleright Пусть S_n , $n = 1, 2, 3, \dots$ – последовательность частичных сумм ряда (2.6.2). Сходимость ряда означает существование такого числа S , что

$$\lim_{n \rightarrow \infty} S_n = S. \quad (2.6.6)$$

При этом параметр n пробегает последовательно чётные и нечётные натуральные значения: $n = 2m-1$ и $n = 2m$, $m = 1, 2, 3, \dots$. Значит достаточно установить, что обе последовательности частичных сумм чётного и нечётного порядка имеют один и тот же предел, равный S :

$$\lim_{m \rightarrow \infty} S_{2m} = \lim_{m \rightarrow \infty} S_{2m-1} = S. \quad (2.6.7)$$

Начнём с рассмотрения сумм чётного порядка; докажем что последовательность таких сумм возрастает и ограничена сверху.

В силу чередования знаков (см. (2.6.2)) имеем

$$\begin{aligned} S_{2m} &= u_1 - u_2 + u_3 - u_4 + \dots + u_{2m-1} - u_{2m} = \\ &= (u_1 - u_2) + (u_3 - u_4) + \dots + (u_{2m-1} - u_{2m}). \end{aligned} \quad (2.6.8)$$

Согласно (2.6.3) разности, записанные в каждой скобке, положительны; следовательно,

$$S_{2m} > 0. \quad (2.6.9)$$

Кроме того, с ростом в сумме (2.6.8) будут возникать новые положительные скобки-разности, так что она оказывается возрастающей. Чтобы доказать ограниченность S_{2m} , перепишем (2.6.8) в виде

$$S_{2m} = u_1 - (u_2 - u_3) - (u_4 - u_5) - \dots - (u_{2m-2} - u_{2m-1}) - u_{2m}.$$

Теперь из u_1 вычитаются положительные разности и положительное u_{2m} , а тогда справедлива оценка $S_{2m} < u_1$. В силу (2.6.9) имеем теперь

$$0 < S_{2m} < u_1. \quad (2.6.10)$$

Итак, установлены возрастание и ограниченность последовательности $\{S_{2m}\}$, откуда следует её сходимость к некоторому действительному числу S .

Далее, частичные суммы S_{2m-1} нечётного и S_{2m} чётного порядка отличаются членом u_{2m} :

$$S_{2m-1} = S_{2m} - u_{2m}.$$

Согласно условию (2.6.4), последовательность $\{S_{2m-1}\}$ имеет теперь тот же предел S , что и $\{S_{2m}\}$. Соотношение (2.6.7) установлено, чем и доказана сходимость ряда (2.6.2).

Наконец, переходя к пределу в двойном неравенстве (2.6.10), получаем оценку (2.6.5) суммы ряда. \triangleleft

2.6.2. Пусть выполнены условия признака Лейбница. Рассмотрим N -й остаток ряда (2.6.2)

$$(-1)^N (a_{N+1} - a_{N+2} + a_{N+3} - \dots), \quad (2.6.11)$$

который также является знакопеременным рядом. Следовательно, (2.6.11) сходится, и его сумма R_N положительна при чётном N и отрицательна, если N нечётно; при этом сумма ряда, записанного в скобках в (2.6.11) не превосходит a_{N+1} . Таким образом установлено утверждение.

Следствие. Если выполнены условия (2.6.3), (2.6.4), то сумма R_N остатка (2.6.11) знакопеременного ряда имеет оценку

$$|R_N| \leq a_{N+1}.$$

В частности, сумму знакопеременного ряда (2.6.2) можно приближённо заменить значением частичной суммы S_N первых N членов; погрешность при этом не превосходит модуля первого из отбрасываемых членов ряда.

2.6.3. Пример. Ряд вида

$$1 - \frac{1}{2} + \frac{1}{3} - \dots + (-1)^{n-1} \frac{1}{n} + \dots \quad (2.6.12)$$

является сходящимся, поскольку выполнены оба условия признака Лейбница: последовательность $\left\{\frac{1}{n}\right\}$ является, очевидно, убывающей и имеет место соотношение

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

Сумма этого ряда приближённо, с точностью, например, до 0,1, равна сумме первых его девяти членов (точность определяется значением первого отбрасываемого члена), т.е. $S \approx 0,7$.

2.7. АБСОЛЮТНАЯ И УСЛОВНАЯ СХОДИМОСТЬ ЗНАКОПЕРЕМЕННЫХ РЯДОВ С ДЕЙСТВИТЕЛЬНЫМИ ЧЛЕНАМИ. АБСОЛЮТНАЯ И УСЛОВНАЯ СХОДИМОСТЬ РЯДОВ С КОМПЛЕКСНЫМИ ЧЛЕНАМИ

2.7.1. Рассмотрим ряд из действительных чисел

$$\sum_{n=1}^{\infty} u_n, \quad (2.7.1)$$

среди членов которого имеются как положительные, так и отрицательные числа; такой ряд называется *знакопеременным*.

Рассмотрим также ряд, составленный из абсолютных величин членов (2.7.1):

$$\sum_{n=1}^{\infty} |u_n|. \quad (2.7.2)$$

Заметим, что если количество только положительных или только отрицательных членов ряда (2.7.1) оказывается конечным, то вопрос о сходимости сводится к случаю знакоположительных рядов. В самом деле, если конечным будет, например, количество положительных членов в (2.7.1), то, начиная с некоторого номера, все члены ряда будут отрицательными. Тогда поведение ряда определяется поведением этого остатка (свойство п. 2.1.5), состоящего только из отрицательных членов. Если изменить знаки всех членов ряда-остатка на противоположные (т.е. умножить все члены на (-1)), то его поведение не изменится в силу свойства п. 2.1.3. Таким образом, для ряда (2.7.1) вопрос о его сходимости сведён к исследованию полученного знакоположительного ряда.

Будем считать поэтому, что количество как положительных, так и отрицательных членов в (2.7.1) является бесконечным.

Теорема 1. Если сходится ряд (2.7.2), то сходится и ряд (2.7.1). Сходимость ряда (2.7.1) в этом случае называется *абсолютной*.

Обратное утверждение неверно: знакопеременный ряд может быть сходящимся, тогда как (2.7.2) – расходящийся. Примером служит (2.6.12), для которого ряд из абсолютных величин – это расходящийся (п. 2.1.2) гармонический ряд.

▷ Пусть

$$S_n = u_1 + u_2 + \dots + u_n$$

n -я частичная сумма ряда (2.7.1), а

$$\sigma_n = |u_1| + |u_2| + \dots + |u_n| \quad (2.7.3)$$

n -я частичная сумма ряда из абсолютных величин (2.7.2).

Выделим в частичной сумме S_n (2.7.1) сумму всех положительных членов, и обозначим её через S_n^+ , а сумму абсолютных величин всех отрицательных членов (в составе той же суммы) обозначим через S_n^- . Суммы S_n^+ и S_n^- , составленные из положительных чисел, возрастают с ростом n . Очевидно, что

$$S_n = S_n^+ - S_n^-, \quad \sigma_n = S_n^+ + S_n^-.$$

Последовательность (2.7.3) имеет предел (ввиду сходимости ряда (2.7.2)), а значит является ограниченной, т.е. существует постоянная $C > 0$, такая что $\sigma_n \leq C$ при всех n . Ясно, что тогда $S_n^+ \leq S_n^+ + S_n^- = \sigma_n \leq C$, и, точно так же, $S_n^- \leq \sigma_n \leq C$. Значит, последовательности S_n^+ и S_n^- , будучи возрастающими и ограниченными, имеют конечные пределы. Тогда имеет предел их разность S_n , что и означает сходимость ряда (2.7.1). ◁

2.7.2. Вернёмся к рассмотрению ряда с комплексными членами:

$$\sum_{n=1}^{\infty} w_n, \quad (2.7.4)$$

одновременно рассматривая соответствующий ряд из модулей

$$\sum_{n=1}^{\infty} |w_n|. \quad (2.7.5)$$

Рассмотрим также два ряда, составленные из действительных частей и мнимых частей последовательности $\{w_n\}$:

$$\sum_{n=1}^{\infty} u_n, \quad \sum_{n=1}^{\infty} v_n, \quad (2.7.6)$$

где $u_n = \operatorname{Re} w_n$, $v_n = \operatorname{Im} w_n$, $n = 1, 2, \dots$.

Теорема 2. Если сходится ряд (2.7.5), то сходится и ряд (2.7.4).

Как и в случае рядов с действительными членами, ряд (2.7.4) называется в этом случае абсолютно сходящимся.

Заметим, что утверждение, обратное теореме 2, неверно.

▷ Поскольку

$$|u_n| \leq \sqrt{u_n^2 + v_n^2} = |w_n|,$$

то сходится и ряд $\sum_{n=1}^{\infty} |u_n|$ (по теореме сравнения рядов с положительными членами). Аналогично, сходится и ряд $\sum_{n=1}^{\infty} |v_n|$.

Следовательно, оба ряда (2.7.6) сходятся абсолютно. Отсюда и вытекает сходимость ряда (2.7.4); см. п. 2.1.4. ◁

Выше показано (на примере ряда из действительных чисел), что числовой ряд может сходиться, тогда как ряд из модулей расходится. В этом случае сходимость ряда (2.7.4) называют условной.

Теорема 3. Ряд (2.7.4) абсолютно сходится тогда и только тогда, когда абсолютно сходятся оба ряда (2.7.6).

▷ При доказательстве теоремы 2 уже установлено, что из сходимости (2.7.5) вытекает абсолютная сходимость обоих рядов (2.7.6). Остаётся установить обратное утверждение. Заметим, что при каждом n наибольшее из двух чисел $|u_n|$ и $|v_n|$ не превосходит их суммы $|u_n| + |v_n|$, а тогда

$$|w_n| = \sqrt{u_n^2 + v_n^2} \leq |u_n| + |v_n|.$$

Если теперь абсолютно сходятся оба ряда (2.7.6), то будет сходящимся и ряд

$$\sum_{n=1}^{\infty} (|u_n| + |v_n|)$$

(см. простейшие свойства сходящихся рядов, п. 2.1.3). Согласно теореме сравнения (п. 2.3.1) будем иметь сходимость (2.7.5). ◁

2.7.3. Как следует из результата теоремы 2, достаточные условия сходимости ряда из модулей (2.7.5) являются одновременно и достаточными условиями сходимости ряда комплексных чисел (2.7.4). Поэтому признаки сходимости знакоположительных рядов, которым мы выше уделили столь значительное внимание, выступают здесь признаками сходимости (абсолютной) рядов с комплексными членами. Уточним последнюю мысль.

Пусть существует предел вида

$$K = \lim_{n \rightarrow \infty} \sqrt[n]{|w_n|}$$

(будем называть его числом Коши). Если $K < 1$, то ряд (2.7.4) сходится абсолютно. Если же $K > 1$, то ряд (2.7.4) расходится.

Стоит отметить, что при $K > 1$ ряд из модулей (2.7.5) расходится ввиду того, что не выполнен необходимый признак сходимости (см. доказательство теоремы 1 п. 2.4.1), но тогда не могут стремиться к нулю и члены w_n ; таким образом и ряд (2.7.4) оказывается расходящимся.

Аналогично обстоит дело и с «числом Даламбера»:

$$D = \lim_{n \rightarrow +\infty} \frac{|w_{n+1}|}{|w_n|};$$

если $D < 1$, то ряд (2.7.4) сходится абсолютно; если же $D > 1$, то (2.7.4) расходится.

2.8. РАВНОМЕРНАЯ СХОДИМОСТЬ ФУНКЦИОНАЛЬНОГО РЯДА

2.8.1. Пусть в области G задана бесконечная последовательность однозначных функций $\{u_n(z)\}$, $n = 1, 2, \dots$. Выражение вида

$$\sum_{n=1}^{\infty} u_n(z) \tag{2.8.1}$$

называется функциональным рядом. При каждом $z = z_0 \in G$ имеем *числовой* ряд из комплексных чисел $u_n(z_0)$. Если получаемый числовой ряд сходится, то z_0 называется его точкой сходимости, а если расходится – то точкой расходимости. На множестве $G_0 \subset G$ всех точек сходимости ряда (2.8.1) задана тогда функция $S = S(z)$, называемая суммой ряда.

2.8.2. Привычные свойства конечных сумм функций могут не сохраняться при переходе к рядам. Положение может быть «исправлено» требованием равномерной сходимости ряда.

Пусть $S(z)$ есть сумма ряда (2.8.1) на замкнутой ограниченной области G и при каждом n существует наибольшее значение модуля отклонения $S_n(z)$ от $S(z)$:

$$\rho_n = \max_{z \in G} |S_n(z) - S(z)|, \quad n = 1, 2, \dots$$

Если $\lim_{n \rightarrow \infty} \rho_n = 0$, то ряд (2.8.1) называется равномерно сходящимся на G к сумме $S(z)$.

Теорема 1 (признак Вейерштрасса). Если существует последовательность $\{\alpha_n\}$ действительных положительных чисел, такая что для всех $z \in G$, $n = 1, 2, \dots$ имеют место оценки

$$|u_n(z)| \leq \alpha_n \tag{2.8.2}$$

и ряд

$$\sum_{n=1}^{\infty} \alpha_n \tag{2.8.3}$$

сходящийся, то ряд (2.8.1) равномерно сходится на G .

При выполнении условий теоремы 1 говорят, что ряд (2.8.1) *мажорируем* на G , а знакоположительный ряд (2.8.3) называется *мажорантным*. В этих терминах теорема может быть сформулирована так: *мажорируемый на G функциональный ряд сходится равномерно на G .*

Отметим также (не приводя здесь соответствующих примеров), что условие мажорируемости является лишь достаточным для равномерной сходимости, но не является необходимым.

▷ Ввиду соотношения (2.8.2), выполненного на G , имеем абсолютную сходимость (на G) ряда (2.8.1) к некоторой сумме $S(z)$; при этом

$$S(z) - S_n(z) = r_n(z),$$

где $r_n(z)$ – сумма ряда-остатка; см. (2.1.11). По определению суммы ряда и ввиду сохранения для функций комплексного переменного привычных свойств пределов (предельный переход под знаком модуля и предельный переход в неравенстве) имеем

$$\begin{aligned} |r_n(z)| &= \lim_{m \rightarrow \infty} (u_{n+1}(z) + \dots + u_m(z)) = \lim_{m \rightarrow \infty} |u_{n+1}(z) + \dots + u_m(z)| \leq \\ &\leq \lim_{m \rightarrow \infty} (|u_{n+1}(z)| + \dots + |u_m(z)|) \leq \lim_{m \rightarrow \infty} (\alpha_{n+1} + \dots + \alpha_m). \end{aligned} \quad (2.8.4)$$

Сумма под знаком последнего написанного предела представляет собою m -ю частичную сумму n -го остатка числового ряда (2.8.3), а значение предела – сумма его n -го остатка, которую мы обозначим через r_n^* :

$$|r_n(z)| \leq r_n^*.$$

Ввиду сходимости ряда (2.8.3) имеем (п. 2.1.5), что $r_n^* \rightarrow 0$ при $n \rightarrow \infty$. Согласно (2.8.4) и условию теоремы тогда имеем

$$\rho_n = \max_{z \in G} |S_n(z) - S(z)| = \max_{z \in G} |r_n(z)| \leq r_n^*$$

и, следовательно,

$$\lim_{n \rightarrow \infty} \rho_n = 0,$$

что и означает равномерную (на G) сходимость ряда (2.8.1). \triangleleft

2.8.3. Теорема 2. Если ряд (2.8.1), составленный из функций $u_n(z)$, непрерывных на замкнутой ограниченной области G , равномерно сходится в этой области, то его сумма $S(z)$ непрерывна в каждой точке $z_0 \in G$, т.е.

$$\lim_{z \rightarrow z_0} S(z) = S(z_0). \quad (2.8.5)$$

\triangleright Оценим $|S(z) - S(z_0)|$. Имеем, в силу (2.8.5),

$$\begin{aligned} |S(z) - S(z_0)| &= |(S_n(z) + r_n(z)) - (S_n(z_0) + r_n(z_0))| \leq \\ &\leq |S_n(z) - S_n(z_0)| + |r_n(z)| + |r_n(z_0)| \leq |S_n(z) - S_n(z_0)| + 2 \max_{z \in G} |r_n(z)| = \\ &= |S_n(z) - S_n(z_0)| + 2\rho_n, \end{aligned} \quad (2.8.6)$$

где, по определению равномерной сходимости ряда, $\rho_n \rightarrow 0$ при $n \rightarrow \infty$. Поскольку конечная сумма $S_n(z)$ непрерывных (на G) функций является непрерывной, имеем

$$\lim_{z \rightarrow z_0} S_n(z) = S_n(z_0) \text{ или } \lim_{z \rightarrow z_0} |S_n(z) - S_n(z_0)| = 0,$$

а тогда, в силу (2.8.6),

$$\lim_{z \rightarrow z_0} |S(z) - S(z_0)| \leq \lim_{z \rightarrow z_0} |S_n(z) - S_n(z_0)| + 2\rho_n = 2\rho_n. \quad (2.8.7)$$

Левая часть (2.8.7) не зависит от n , и, следовательно, сохраняет свой вид при предельном переходе (по n), тогда как правая стремится к нулю. Переходя к пределу при $n \rightarrow \infty$ в обеих частях (2.8.7), получаем

$$\lim_{z \rightarrow z_0} |S(z) - S(z_0)| = 0,$$

откуда и следует (2.8.5). \triangleleft

2.8.4. Установим возможность почленного интегрирования равномерно сходящегося ряда.

Теорема 3. Пусть члены равномерно сходящегося в замкнутой ограниченной области G ряда (2.8.1) непрерывны и L – некоторая гладкая дуга, расположенная в этой области. Тогда, если $f(z)$ – сумма ряда, то возможно почленное интегрирование по дуге L :

$$\int_L f(z) dz = \int_L u_1(z) dz + \int_L u_2(z) dz + \dots + \int_L u_n(z) dz + \dots \quad (2.8.8)$$

\triangleright Во-первых (согласно теореме 2), $f(z)$ – непрерывна в указанной выше области, т.е. интеграл от неё (вдоль L) существует. Во-вторых, достаточно доказать (по определению сходимости ряда), что разность

$$\int_L f(z) dz - \left(\int_L u_1(z) dz + \int_L u_2(z) dz + \dots + \int_L u_n(z) dz \right)$$

является при $n \rightarrow \infty$ бесконечно малой последовательностью. Действительно, модуль этой разности обладает оценкой

$$\begin{aligned} &\left| \int_L f(z) dz - \int_L (u_1(z) + u_2(z) + \dots + u_n(z)) dz \right| = \\ &= \left| \int_L (f(z) - S_n(z)) dz \right| \leq \ell \max_{z \in L} |f(z) - S_n(z)|, \end{aligned} \quad (2.8.9)$$

где $S_n(z)$ – частная сумма ряда (2.8.1), ℓ – длина дуги L (см. п. 1.4.2). Ввиду равномерной сходимости правая часть (2.8.9) стремится к нулю (при $n \rightarrow \infty$), чем и доказано соотношение (2.8.8). \triangleleft

2.8.5. Теорема 4. Если члены ряда (2.8.1) аналитичны в замкнутой ограниченной области G и ряд равномерно сходится к сумме $f(z)$ в этой области, то и $f(z)$ аналитична в G .

▷ Пусть γ – некоторая окружность с центром в точке z , целиком лежащая внутри G ; тогда

$$\frac{f(s)}{s-z} = \sum_{n=0}^{\infty} \frac{f_n(s)}{s-z}, \quad s \in \gamma,$$

причём члены ряда аналитичны (как функции от s) вместе с $f_n(s)$. Согласно теореме 3 возможно почленное интегрирование ряда по окружности γ (обход – в направлении против часовой стрелки):

$$\frac{1}{2\pi i} \int_{\gamma} \frac{f(s)ds}{s-z} = \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_{\gamma} \frac{f_n(s)ds}{s-z} \right),$$

или, в силу интегральной формулы Коши,

$$\frac{1}{2\pi i} \int_{\gamma} \frac{f(s)ds}{s-z} = \sum_{n=0}^{\infty} f_n(z).$$

Тогда сумма $f(z)$ ряда (2.8.1) оказывается совпадающей с интегралом вида

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(s)ds}{s-z}.$$

Нетрудно доказать, что этот интеграл представляет собой аналитическую функцию в G , т.е. $f(z)$, ввиду произвольности z , аналитична в указанной области.

2.9. СТЕПЕННЫЕ РЯДЫ

2.9.1. Пусть $\{z^n\}$, $n=1, 2, \dots$ – последовательность степенных функций; $\{a_n\}$, $n=0, 1, \dots$ – последовательность комплексных чисел.

Ряд вида

$$a_0 + a_1 z + a_2 z^2 + \dots + a_n z^n + \dots \quad (2.9.1)$$

называется *степенным*, для (2.9.1) употребляем также обозначение

$$\sum_{n=0}^{\infty} a_n z^n.$$

Очевидно, что любой степенной ряд сходится в точке $z_0 = 0$, так как все его частичные суммы $S_n(z_0) = a_0$, и, следовательно, предел последовательности $\{S_n(z_0)\}$ существует и равен a_0 . Нахождение других точек сходимости будет опираться на следующую теорему.

Теорема Абеля. Если степенной ряд (2.9.1) сходится в некоторой точке $z_0 \neq 0$, то он абсолютно сходится в круге $U(0; |z_0|) = \{z \mid |z| < |z_0|\}$. Если же z_0' – точка расходимости, то ряд (2.9.1) расходится при всех z , таких что $|z| > |z_0'|$.

▷ 1. Ряд из модулей для (2.9.1) имеет вид

$$\sum_{n=0}^{\infty} |a_n| \cdot |z|^n. \quad (2.9.2)$$

Общий член ряда (2.9.2) можно представить следующим образом:

$$u_n(z) = |a_n| \cdot |z|^n = |a_n z_0^n| \cdot \left| \frac{z}{z_0} \right|^n, \quad (2.9.3)$$

где $z_0 \neq 0$ – точка сходимости ряда (2.9.1). Поскольку в этой точке выполнен необходимый признак сходимости, т.е.

$$\lim_{n \rightarrow \infty} |a_n z_0^n| = 0,$$

то для всех n существует постоянная $M > 0$, такая что $|a_n z_0^n| \leq M$.

При условии $|z| < |z_0|$ имеем для $q = \left| \frac{z}{z_0} \right|$, что $0 \leq q < 1$. Следовательно, в силу (2.9.3), имеет место оценка

$0 \leq u_n(z) \leq Mq^n$ ($0 \leq q < 1$), и ряд

$$M + Mq + Mq^2 + \dots + Mq^n + \dots$$

является сходящимся (сумма бесконечно убывающей геометрической прогрессии). По теореме сравнения знакоположительных рядов тогда сходится и ряд (2.9.2). Значит, в круге $U(0; |z_0|)$ ряд (2.9.1) сходится абсолютно, что и утверждалось.

2. В случае $|z| > |z'_0|$ ряд (2.9.1) не может сходиться в точке z . Действительно, имеем $z'_0 \in U(0; |z|)$, и, если (2.9.1) сходится в точке z , то по первой части теоремы Абеля, ряд сходится и в точке z'_0 . Но это противоречит условию. Итак, во всех точках z , таких что $|z| > |z'_0|$, ряд (2.9.1) расходится. \triangleleft

2.9.2 Из теоремы Абеля вытекает, что всякая точка сходимости z_0 степенного ряда ближе к началу координат, чем любая точка расходимости (если такая имеется). Следовательно, должно существовать некоторое "пограничное" число R , такое что при $|z| < R$ (т.е. в каждом таком круге) имеет место абсолютная сходимость, а при $|z| > R$ (вне круга) – расходимость ряда (2.9.1).

Число R называется *радиусом сходимости* степенного ряда, область $U(0, R)$ – его кругом сходимости, см. рис. 2.9.1.

При всяком $0 < \rho < R$ ряд (2.9.1) будет сходиться и равномерно в круге $|z| \leq \rho$. Действительно, взяв точку \tilde{z} , такую что $|\tilde{z}| = \rho$, имеем абсолютную сходимость в точке \tilde{z} , т.е. сходится ряд

$$\sum_{n=0}^{\infty} |a_n| \rho^n.$$

В то же время, для членов (2.9.1) имеем оценку

$$|a_n z^n| \leq |a_n| \cdot |z|^n < |a_n| \rho^n.$$

Согласно признаку Вейерштрасса, получаем равномерную сходимость при $|z| \leq \rho$.

В частности (вследствие непрерывности степенных функций $u_n(z) = z^n$), непрерывной в круге сходимости будет сумма ряда (2.9.1).

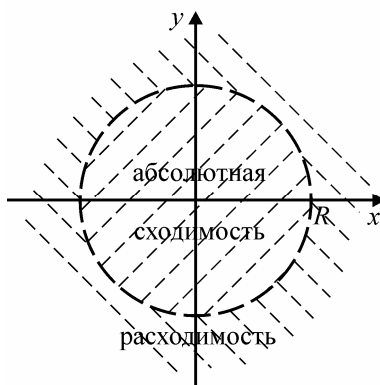


Рис. 2.9.1

2.9.3. Заметим, что если функция $\varphi(z)$ равномерно ограничена по модулю в круге $|z| \leq \rho$, т.е. если существует $M = \text{const}$, такая что $|\varphi(z)| \leq M$ в этом круге, то ряд

$$a_0 \varphi(z) + a_1 \varphi(z)z + \dots + a_n \varphi(z)z^n + \dots \quad (2.9.4)$$

остаётся мажорируемым в том же круге. Действительно, мажорантным для (2.9.4) является ряд вида

$$\sum_{n=0}^{\infty} M a_n,$$

если ряд с общим членом a_n будет мажорантным для (2.9.1).

Теперь возможность *почленного интегрирования степенных рядов* и рядов вида (2.9.4) в круге сходимости будет вытекать, как частный случай, из утверждения теоремы 3 п. 2.8.4.

Согласно же теореме 4 п. 2.8.5 сумма $f(z)$ степенного ряда (2.9.4) является аналитической в круге сходимости. Как оказывается, в каждой точке этого круга производная $f'(z)$ может быть получена путём почленного дифференцирования ряда:

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}.$$

Более того, сумма степенного ряда бесконечно дифференцируема в круге сходимости и производная этой суммы любого p -го порядка получается путём p -кратного почленного дифференцирования ряда.

2.9.4. Радиус сходимости R степенного ряда (2.9.1) можно найти по одной из формул:

$$R = \frac{1}{D} \quad \text{или} \quad R = \frac{1}{K}, \quad (2.9.5)$$

если существует соответствующее «число Даламбера»

$$D = \lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|}$$

или «число Коши»

$$K = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}.$$

Формулы (2.9.5) остаются справедливыми, если $D=0$ или $K=0$ – тогда $R=\infty$, т.е. областью сходимости ряда является вся комплексная плоскость. Если же $D=\infty$ или $K=\infty$, то $R=0$, т.е. «областью» сходимости является единственная точка $z_0=0$. Докажем, например, вторую из формул (2.9.5).

Согласно признаку Коши, ряд (2.9.1) сходится, если

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} \cdot |z|^n < 1, \text{ т.е. } |z|K < 1, \quad (2.9.6)$$

откуда получаем при $|z| < \frac{1}{K}$ сходимость ряда (2.9.1), а значит и абсолютную сходимость ряда (2.9.1). В то же время при

$|z| > \frac{1}{K}$ расходится не только ряд из модулей (2.9.2), но и сам ряд (2.9.1); см. замечание в п. 2.7.3. Итак, именно число $R = \frac{1}{K}$ оказалось радиусом сходимости согласно определению радиуса. Отметим также, что при $K=0$ условие (2.9.6) выполнено при всех z ($R=\infty$), а при $K=\infty$ условие (2.9.6) не выполнено при любом $z \neq 0$; точка же $z_0=0$, как упоминалось, служит точкой сходимости любого степенного ряда ($R=0$). Утверждение п. 2.9.3 полностью доказано.

2.9.5. Рассмотрим теперь ряд по степеням разности

$$(z - z_0),$$

где z_0 – данное комплексное число:

$$\sum_{n=0}^{\infty} C_n (z - z_0)^n = C_0 + C_1 (z - z_0) + \dots + C_n (z - z_0)^n + \dots \quad (2.9.7)$$

Если произвести в (2.9.7) замену переменных $s = z - z_0$, то из результатов п. 2.9.2 будет вытекать, что *областью сходимости* (2.9.7) *является некоторый круг* $|z - z_0| < R$; в этом круге по отношению к ряду (2.9.7) сохраняются свойства и результаты, приведённые в п. 2.9.3 и 2.9.4; в частности, имеют место формулы для радиуса сходимости (2.9.5).

2.10. СТЕПЕННЫЕ РЯДЫ: СЛУЧАЙ ДЕЙСТВИТЕЛЬНОГО ПЕРЕМЕННОГО

2.10.1. В настоящем параграфе мы выделим те специфические свойства, которые присущи степенным рядам в случае именно действительного аргумента. Рассмотрим начнём с общих свойств равномерно сходящихся рядов функций $f_n(x)$ ($n = 1, 2, \dots$) действительного переменного x . Понятие равномерной сходимости формулируется применительно к случаю задания последовательности этих функций на некотором отрезке $[a, b]$.

Если ряд

$$\sum_{n=1}^{\infty} f_n(x) \quad (2.10.1)$$

мажорируем на отрезке, то он обладает *равномерной сходимостью* на этом отрезке.

Ряд (2.10.1), составленный из функций, *непрерывных* на отрезке $[a, b]$, и *равномерно сходящийся* на $[a, b]$, *обладает непрерывной на этом отрезке суммой и допускает возможность почленного интегрирования* по всякому $[\alpha, \beta] \subset [a, b]$.

Приведённые утверждения являются частными случаями утверждений, установленных в п. 2.9.2–2.9.3.

2.10.2. Если ряды (2.10.1) и

$$\sum_{n=1}^{\infty} f'_n(x) \quad (2.10.2)$$

обладают *равномерной сходимостью* на отрезке $[a, b]$, и суммы их равны, соответственно, функциям $S(x)$ и $\varphi(x)$, то сумма $S(x)$ ряда (2.10.1) *дифференцируема* при всех $x \in (a, b)$; при этом $S'(x) = \varphi(x)$ и

$$\sum_{n=1}^{\infty} f'_n(x) = S'(x).$$

Чтобы доказать это утверждение, произведём почленное интегрирование (2.10.2) по некоторому отрезку $[\alpha, x] \subset [a, b]$, что возможно в силу его равномерной сходимости. Имеем

$$\sum_{n=1}^{\infty} \int_{\alpha}^x f'_n(x) dx = \int_{\alpha}^x \varphi(x) dx.$$

Поскольку

$$\int_{\alpha}^x f'_n(x) dx = f_n(x) - f_n(\alpha), \quad n = 1, 2, \dots,$$

то левая часть последнего соотношения есть разность значений суммы $S(x)$, вычисленных в точках x и α :

$$S(x) - S(\alpha) = \int_{\alpha}^x \varphi(x) dx.$$

Продифференцируем теперь обе части последнего соотношения и воспользуемся существованием производной интеграла с переменным верхним пределом; в нашем случае

$$\left(\int_{\alpha}^x \varphi(x) dx \right)' = \varphi(x), \quad x \in (a, b).$$

В результате получаем дифференцируемость $S(x)$ и соотношение

$$S'(x) - (S(\alpha))' = \varphi(x), \quad x \in (a, b);$$

при этом $(S(\alpha))' = 0$, так как α – постоянная величина. Утверждение п. 2.10.2 доказано.

2.10.3. Рассмотрим степенной ряд

$$a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots, \quad x \in \mathbf{R} \quad (2.10.3)$$

с действительными коэффициентами a_n , $n = 1, 2, \dots$; будем употреблять также запись

$$\sum_{n=0}^{\infty} a_n x^n.$$

Пересечением круга сходимости $\{z: |z| < R\}$ степенного ряда комплексной переменной с осью абсцисс является интервал $(-R, R)$, поэтому для ряда (2.10.3) следует вести речь об интервале сходимости, внутри которого ряд сходится абсолютно, а вне которого расходится. Радиус интервала (радиус сходимости) может быть, очевидно, и здесь найден по одной из формул: $R = \frac{1}{K}$ или $R = \frac{1}{D}$, где, соответственно случаю (2.10.3),

$$K = \sqrt[n]{|a_n|}, \quad D = \lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|}. \quad (2.10.4)$$

Стоит отметить, что общая теория не позволяет судить о поведении такого ряда в концевых точках интервала, и для каждого конкретного степенного ряда исследование в обеих концевых точках проводят дополнительно.

2.10.4. Как и в п. 2.9.2–2.9.3 устанавливаем, что степенной ряд (2.10.3) мажорируем на всяком отрезке вида $[-\rho, \rho] \subset (-R, R)$, а значит, равномерно сходится на этом отрезке. Отсюда вытекает возможность почленного интегрирования ряда по всякому отрезку, расположенному внутри интервала сходимости (см. п. 2.10.2).

Возможность же почленного дифференцирования будет обеспечена равномерной сходимостью ряда, составленного из производных; см. утверждение п. 2.10.2. Достаточно поэтому установить мажорируемость ряда

$$(a_0)' + (a_1 x)' + (a_2 x^2)' + \dots + (a_n x^n)' + \dots,$$

т.е.

$$\sum_{n=1}^{\infty} n a_n x^{n-1} \quad (2.10.5)$$

на любом отрезке $[-\rho, \rho] \subset (-R, R)$. Доказательство мажорируемости проведем в предположении, что существует предел, обозначенный через D в (2.10.4); следовательно, $R = \frac{1}{D}$.

Определим радиус сходимости \tilde{R} ряда (2.10.5). Соответствующее число Даламбера имеет вид

$$\tilde{D} = \lim_{n \rightarrow \infty} \frac{(n+1) |a_{n+1}|}{n |a_n|} = \lim_{n \rightarrow \infty} \frac{n+1}{n} \lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} = 1 \cdot D,$$

следовательно, $\tilde{R} = \frac{1}{D}$. Таким образом, $\tilde{R} = R$, и интервалы сходимости рядов (2.10.3), (2.10.5) совпадают. Окончательно имеем мажорируемость ряда (2.10.5) на всяком отрезке $[-\rho, \rho] \subset (-R, R)$, а значит и возможность почленного дифференцирования исходного степенного ряда (2.10.3).

Если только что приведённые рассуждения применить к ряду из производных (2.10.5), то получаем возможность и его почленного дифференцирования в интервале $(-R, R)$. Повторяя и далее указанные рассуждения, приходим к следующему важному выводу: *степенной ряд, обладающий суммой $S(x)$ в некотором интервале сходимости, можно почленно дифференцировать сколь угодно много раз в этом интервале, при этом сумма ряда из p -ых производных совпадает с $S^{(p)}(x)$.*

3. РАЗЛОЖЕНИЕ ФУНКЦИЙ В СТЕПЕННЫЕ РЯДЫ

3.1. РЯД МАКЛОРЕНА

3.1.1. В п. 2.9.3 установлено, что сумма степенного ряда в круге его сходимости является аналитической в этом круге функцией. Поставим теперь обратную задачу: функцию $w = f(z)$, однозначную и аналитическую в некотором круге $U(0; R)$, разложить в ряд по степеням z .

3.1.2. В условиях п. 3.1.1 для любой $z \in U(0; R)$ имеет место разложение

$$f(z) = f(0) + \frac{f'(0)}{1!}z + \frac{f''(0)}{2!}z^2 + \dots + \frac{f^{(n)}(0)}{n!}z^n + \dots \quad (3.1.1)$$

Это утверждение будет доказано ниже, в п. 3.1.4. Представление функции $f(z)$ в виде суммы ряда (3.1.1) называется её разложением в ряд Маклорена.

3.1.3. Другая форма (3.1.1) оказывается следующей:

$$f(z) = \sum_{n=0}^{\infty} C_n z^n, \quad (3.1.2)$$

где

$$C_n = \frac{1}{2\pi i} \int_{\gamma} \frac{f(s) ds}{s^{n+1}}; \quad (3.1.3)$$

γ – любая окружность с центром в точке $z_0 = 0$, обходимая против часовой стрелки и целиком лежащая в круге $U(0; R)$.

Разложения (3.1.1) и (3.1.2) эквивалентны, так как коэффициенты при z^n , записываемые в виде $\frac{f^{(n)}(0)}{n!}$, совпадают с правой частью (3.1.3):

$$\frac{f^{(n)}(0)}{n!} = \frac{1}{2\pi i} \int_{\gamma} \frac{f(s) ds}{(s-0)^{n+1}}$$

ввиду формулы (1.6.6) для $f^{(n)}(0)$.

3.1.4. Итак, достаточно доказать справедливость (3.1.2), (3.1.3). Так как $f(z)$ аналитична в $U(0; R)$, то по формуле Коши

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(s)}{s} \frac{1}{1 - \frac{z}{s}} ds, \quad (3.1.4)$$

где z – произвольная точка в $U(0; R)$, а окружность γ с центром в точке $z_0 = 0$ проведена так, что z расположена внутри её.

При таком выборе контура интегрирования (рис. 3.1.1) $|z| < |s|$, т.е. $q = \frac{z}{s}$ обладает свойством $|q| < 1$.

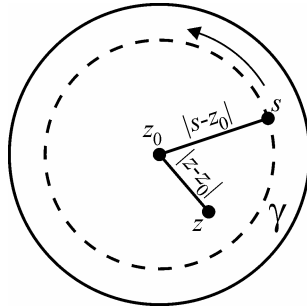


Рис. 3.1.1

Теперь

$$\frac{1}{1-q} = 1 + q + q^2 + \dots + q^n + \dots \quad (3.1.5)$$

является сходящимся рядом как сумма геометрической прогрессии. Под знаком интеграла (3.1.4) получается ряд (по степеням $\frac{z}{s}$), сходящийся при каждом z .

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \oint_{\gamma} \frac{f(s)}{s} \left(1 + \frac{z}{s} + \frac{z^2}{s^2} + \dots + \frac{z^n}{s^n} + \dots \right) ds = \\ &= \frac{1}{2\pi i} \oint_{\gamma} \frac{f(s)}{s} ds + \frac{1}{2\pi i} \oint_{\gamma} \frac{f(s)}{s^2} ds + \frac{1}{2\pi i} \oint_{\gamma} \frac{f(s)}{s^3} ds + \dots + \frac{1}{2\pi i} z^n \oint_{\gamma} \frac{f(s)}{s^{n+1}} ds + \dots; \end{aligned} \quad (3.1.6)$$

почленное интегрирование возможно в силу утверждения п. 2.9.3, условия которого выполнены, так как:

- $f(z)$ аналитична в $U(0; R)$, и, следовательно, существует постоянная M , такая что $|f(s)| \leq M$ на γ ;
- $|s| = r = \text{const}$, где r – радиус выбранной окружности γ ;
- ряд вида

$$\sum_{n=0}^{\infty} \frac{f(s)}{s} \left(\frac{z}{s} \right)^n$$

имеет тогда в качестве мажорантного сходящийся числовой ряд

$$\sum_{n=0}^{\infty} \frac{M}{r^n} |q|^n, \text{ где } |q| < 1.$$

Разложение (3.1.6) теперь совпадает с (3.1.2), если учесть определение (3.1.3) коэффициентов C_n . Соотношения (3.1.1) и (3.1.2) доказаны.

Заметим, что контур интегрирования в доказательстве мы выбрали так, чтобы точка z была заключена внутри его. Однако, согласно теореме Коши для двухсвязной области (см. п. 1.5.2), в качестве γ можно выбрать *любую* окружность с центром в 0, лежащую в $U(0; R)$.

3.1.5 Функции $w = e^z$, $w = \sin z$, $w = \cos z$ являются (см. п. 1.3.6) аналитическими во всей комплексной плоскости и, следовательно, могут быть разложены в ряд Маклорена во всяком круге с центром в начале координат, а значит, в каждой точке z комплексной плоскости. Соответствующие разложения имеют вид

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \dots + \frac{z^n}{n!} + \dots, \quad (3.1.7)$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots + (-1)^n \frac{z^{2n+1}}{(2n+1)!} + \dots, \quad (3.1.8)$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots + (-1)^n \frac{z^{2n}}{(2n)!} + \dots \quad (3.1.9)$$

Докажем, например, соотношение (3.1.7). Для этого вычислим коэффициенты $c_0 = f(0)$, $c_n = \frac{f^{(n)}(0)}{n!}$, $n = 1, 2, \dots$ ряда Маклорена. Имеем

$$f(z) = f'(z) = f''(z) = \dots = f^{(n)}(z) = \dots = e^z, \\ f(0) = f'(0) = f''(0) = \dots = f^{(n)}(0) = \dots = e^0 = 1,$$

так что $c_n = \frac{1}{n!}$. Теперь разложение (3.1.1) функции $f(z) = e^z$ принимает вид (3.1.7), что и утверждалось.

Рассуждения в случаях (3.1.8), (3.1.9) аналогичны.

3.1.6. Список разложений основных элементарных функций в степенные ряды можно дополнить. Например,

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \dots + (-1)^n \frac{z^{n+1}}{n+1} + \dots, \quad |z| < 1. \quad (3.1.10)$$

Действительно, положив в соотношении (3.1.5) $q = -z$, $|z| < 1$ и проинтегрировав результат почленно от точки 0 до точки z (вдоль любого гладкого пути, лежащего в указанном круге), приходим к разложению (3.1.10).

3.1.7. Разложение (3.1.1) функции, аналитической в круге $U(0; R)$, справедливо, в частности, для всех значений действительной переменной $x \in (-R, R)$:

$$f(x) = f(0) + \frac{f'(0)}{1!} x + \frac{f''(0)}{2!} x^2 + \dots + \frac{f^{(n)}(0)}{n!} x^n + \dots \quad (3.1.11)$$

В общем случае функции $f(x)$ действительного переменного x , дифференцируемой сколь угодно много раз в некоторой окрестности нуля, может быть доказано следующее утверждение: если существует постоянная M , такая что $|f(x)| + |f^{(n)}(x)| \leq M$ ($n = 1, 2, \dots$) для всех $x \in (-R, R)$, то в каждой такой точке x имеет место разложение Маклорена (3.1.11).

3.2. РЯД ТЕЙЛОРА

3.2.1. Если $w = f(z)$ однозначна и аналитична в некотором круге G с центром в точке z_0 , то рассуждения, аналогичные приведённым в п. 3.1.4, позволяют разложить $f(z)$ в этом круге в ряд по степеням разности $(z - z_0)$:

$$f(z) = f(z_0) + \frac{f'(z_0)}{1!} (z - z_0) + \frac{f''(z_0)}{2!} (z - z_0)^2 + \dots + \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n + \dots; \quad (3.2.1)$$

разложение (3.2.1) называется *рядом Тейлора*. Другая его форма оказывается следующей:

$$f(z) = \sum_{n=0}^{\infty} C_n (z - z_0)^n, \quad (3.2.2)$$

где

$$C_n = \frac{1}{2\pi i} \int_{\gamma} \frac{f(s) ds}{(s - z_0)^{n+1}};$$

γ – любая окружность с центром в точке z_0 , обходимая против часовой стрелки и целиком лежащая в G (см. рис. 3.1.1).

3.2.2. Пусть теперь $f(z)$ аналитична в произвольной точке z_0 . Её можно выбрать в качестве центра круга (достаточно малого радиуса), в котором $f(z)$ остаётся аналитичной, а затем разложить $f(z)$ в ряд Тейлора (3.2.1). Говорят, что (3.2.1) есть разложение $f(z)$ в окрестности z_0 .

Если $f(z_0) = f'(z_0) = \dots = f^{(n-1)}(z_0) = 0$, но $f^{(n)}(z_0) \neq 0$, то разложение $f(z)$ в окрестности z_0 , записанное в форме (3.2.2) имеет вид

$$f(z) = C_n(z-z_0)^n + C_{n+1}(z-z_0)^{n+1} + \dots \quad (3.2.3)$$

В этом случае точку z_0 называют нулем n -го порядка функции $f(z)$, а при $n=1$ – простым нулём.

3.2.3. Теорема. Пусть $f(z)$ аналитична в точке z_0 . Эта точка является нулём n -го порядка для $f(z)$ тогда и только тогда, когда существует аналитическая в точке z_0 функция $\varphi(z)$, такая что $\varphi(z_0) \neq 0$ и

$$f(z) = (z-z_0)^n \varphi(z). \quad (3.2.4)$$

▷ Если z_0 – нуль n -го порядка, то согласно (3.2.3)

$$f(z) = (z-z_0)^n (C_n + C_{n+1}(z-z_0) + \dots), \quad (3.2.5)$$

причём сумма $\varphi(z)$ степенного ряда, записанного в скобках в (3.2.5), аналитична в точке z_0 и $\varphi(z_0) = C_n \neq 0$. Теперь равенство (3.2.4) установлено.

Обратно, пусть $f(z)$ представима в виде (3.2.4), где $\varphi(z)$ аналитична в точке z_0 и $\varphi(z_0) \neq 0$. Разложив $\varphi(z)$ в степенной ряд в окрестности точки z_0 , получим $\varphi(z) = \tilde{C}_0 + \tilde{C}_1(z-z_0) + \dots$, где $\tilde{C}_0 \neq 0$ (так как $\varphi(z_0) \neq 0$) и, следовательно, в указанной окрестности,

$$f(z) = \tilde{C}_0(z-z_0)^n + \tilde{C}_1(z-z_0)^{n+1} + \dots; \quad \tilde{C}_0 \neq 0. \quad (3.2.6)$$

Значит, $f(z)$ имеет разложение вида (3.2.3), если соответствующим образом переобозначить коэффициенты в (3.2.6). Таким образом, z_0 оказалась нулём n -го порядка для $f(z)$, чем и завершается доказательство теоремы. ◁

3.3. РЯД ЛОРАНА

3.3.1. Рассмотрим ряд вида

$$\begin{aligned} \dots + \frac{C_{-n}}{(z-z_0)^n} + \dots + \frac{C_{-2}}{(z-z_0)^2} + \frac{C_{-1}}{z-z_0} + C_0 + C_1(z-z_0) + C_2(z-z_0)^2 + \dots \\ + C_n(z-z_0)^n + \dots = \sum_{n=-\infty}^{\infty} C_n(z-z_0)^n, \end{aligned} \quad (3.3.1)$$

содержащий как неотрицательные, так и отрицательные степени разности $z-z_0$. Если (3.3.1) записать в виде

$$\sum_{n=1}^{\infty} \frac{C_{-n}}{(z-z_0)^n} + \sum_{n=0}^{\infty} C_n(z-z_0)^n, \quad (3.3.2)$$

то первый из рядов можно рассматривать как степенной:

$$\sum_{n=1}^{\infty} C_{-n} w^n, \quad \text{где } w = \frac{1}{z-z_0}.$$

В своём круге сходимости $|w| < R$ его сумма есть некоторая однозначная аналитическая функция $\varphi(w) = \varphi\left(\frac{1}{z-z_0}\right)$; тривиальный случай $R=0$ исключим из рассмотрения. Будучи суперпозицией аналитических функций,

$\varphi\left(\frac{1}{z-z_0}\right)$ как функция от z (где $z \neq z_0$) также аналитична при $|w| = \frac{1}{|z-z_0|} < R$. Итак, для $|z-z_0| > \frac{1}{R}$, имеем:

φ аналитична во "внешности" круга $\bar{U}(z_0; R_1)$; не исключён и случай $R_1=0$ (если $R=\infty$). Точно также, второй ряд в (3.3.2) сходится при $|z-z_0| < R_2$ с некоторым $R_2 > 0$. Сумма этого ряда есть некоторая аналитическая в этом круге функция $\psi(z)$.

Если $R_1 < R_2$, то существует общая область, в которой сходятся оба ряда в (3.3.2), т.е. оказывается, что (3.3.1) имеет область сходимости некоторое кольцо с центром в точке z_0 : $R_1 < |z-z_0| < R_2$ (рис. 3.3.1). При этом сумма ряда

$$f(z) = \varphi\left(\frac{1}{z-z_0}\right) + \psi(z).$$

3.3.2. Теперь возникает обратная задача: пусть $f(z)$ однозначна и аналитична в некотором кольце $R_1 < |z-z_0| < R_2$. Можно ли её разложить в степенной ряд вида (3.3.1) и, если можно, то каковы коэффициенты этого ряда?

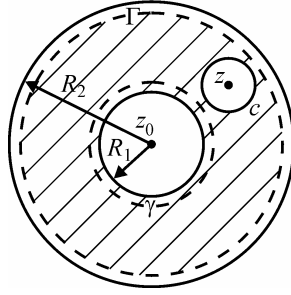


Рис. 3.3.1

Рассуждения, приводящие к разложению $f(z)$ в ряд (3.4.1), будут носить тот же характер, что в параграфе 3.1. Проведём внутри кольца окружность γ и Γ с центром в точке z_0 так, чтобы z оказалась внутренней точкой в новом кольце (т.е. между γ и Γ); пусть c – окружность (с центром в точке z) столь малого радиуса, что также расположена между γ и Γ . В этом случае справедлива следующая формула (аналог теоремы Коши для двухсвязной области):

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{f(s)}{s-z} ds = \frac{1}{2\pi i} \int_c \frac{f(s)}{s-z} ds + \frac{1}{2\pi i} \int_{\gamma} \frac{f(s)}{s-z} ds.$$

По интегральной формуле Коши

$$\frac{1}{2\pi i} \int_c \frac{f(s)}{s-z} ds = f(z),$$

следовательно,

$$f(z) = \frac{1}{2\pi i} \left(\int_{\Gamma} \frac{f(s)}{s-z} ds - \int_{\gamma} \frac{f(s)}{s-z} ds \right). \quad (3.3.3)$$

Представим дробь $\frac{1}{s-z}$ на окружности Γ в виде

$$\frac{1}{s-z} = \frac{1}{(s-z_0)-(z-z_0)} = \frac{1}{s-z_0} \frac{1}{1 - \frac{z-z_0}{s-z_0}} = \frac{1}{s-z_0} \sum_{n=0}^{\infty} \left(\frac{z-z_0}{s-z_0} \right)^n,$$

так как при $q = \frac{z-z_0}{s-z_0}$ на Γ имеем $|q| < 1$, а значит

$$\frac{1}{1-q} = \sum_{n=0}^{\infty} q^n.$$

Подобным образом на окружности γ :

$$\frac{1}{s-z} = \frac{1}{z-z_0} \left(- \frac{1}{1 - \frac{s-z_0}{z-z_0}} \right) = - \frac{1}{z-z_0} \sum_{m=0}^{\infty} \left(\frac{s-z_0}{z-z_0} \right)^m,$$

так как $q = \frac{s-z_0}{z-z_0}$ обладает свойством $|q| < 1$ на γ .

Почленно интегрируя в (3.3.3) полученные ряды (обоснование почленного интегрирования было приведено в параграфе 3.1), находим

$$f(z) = \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_{\Gamma} \frac{f(s)}{(s-z_0)^{n+1}} ds \right) (z-z_0)^n + \sum_{m=0}^{\infty} \left(\frac{1}{2\pi i} \int_{\gamma} \frac{f(s)}{(s-z_0)^{-m}} ds \right) \frac{1}{(z-z_0)^{m+1}}. \quad (3.3.4)$$

Если во второй сумме изменить нумерацию по формуле $n = m+1$ и обозначить

$$C_n = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(s)}{(s-z_0)^{n+1}} ds, \quad C_{-n} = \frac{1}{2\pi i} \int_{\gamma} \frac{f(s)}{(s-z_0)^{-n+1}} ds,$$

то приходим к доказываемому разложению вида (3.3.2). При вычислениях C_n контуром интегрирования (согласно теореме Коши для двусвязной области) может быть выбрана любая окружность ℓ (с центром в точке z_0), расположенная целиком в данном кольце (вместо ранее рассмотренных контуров γ и Γ).

Итак, согласно (3.3.4)

$$f(z) = \sum_{n=-\infty}^{\infty} C_n (z - z_0)^n, \quad (3.3.5)$$

где

$$C_n = \frac{1}{2\pi i} \int_{\ell} \frac{f(s)}{(s - z_0)^{n+1}} ds, \quad n = 0, \pm 1, \pm 2, \dots \quad (3.3.6)$$

3.3.3. Поставленная задача решена: $f(z)$, аналитичная в указанном кольце, может быть представлена в виде суммы ряда (3.3.5), называемой *рядом Лорана*, коэффициенты ряда вычисляются по формулам (3.3.6).

Впрочем, иногда можно применить другие, более простые приёмы. Продемонстрируем их на примерах.

Пример 1. Функцию $f(z) = \frac{1}{z+2}$ разложить в ряд по степеням $(z-1)$.

Решение. Центром кольца (или колец), в которых будет происходить разложение, должна быть, согласно условию, точка $z_0 = 1$. Особой точкой является $z_1 = -2$, лежащая на окружности с центром z_0 радиуса $R = 3$ (рис. 3.3.2), т.е. на окружности $|z-1| = 3$.

Следовательно, предстоит вести рассмотрение для $|z-1| < 3$ и $|z-1| > 3$ отдельно. В каждом случае воспользуемся суммой бесконечной геометрической прогрессии

$$\frac{1}{1+q} = 1 - q + q^2 - \dots + (-1)^n q^n + \dots; \quad |q| < 1.$$

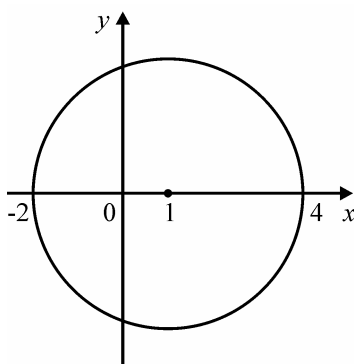


Рис. 3.3.2

В первом случае представим

$$f(z) = \frac{1}{3+(z-1)} = \frac{1}{3} \frac{1}{1 + \frac{z-1}{3}}.$$

Здесь $|q| = \left| \frac{z-1}{3} \right| < 1$ для $|z-1| < 3$; следовательно,

$$f(z) = \frac{1}{3} \left(1 - \frac{z-1}{3} + \frac{(z-1)^2}{3^2} - \dots + (-1)^n \frac{(z-1)^n}{3^n} + \dots \right).$$

Во втором случае

$$f(z) = \frac{1}{(z-1)+3} = \frac{1}{z-1} \frac{1}{1 + \frac{3}{z-1}}$$

для $|q| = \left| \frac{3}{z-1} \right| < 1$ (а это верно при $|z-1| > 3$) имеем

$$f(z) = \frac{1}{z-1} \left(1 - \frac{3}{z-1} + \frac{3^2}{(z-1)^2} - \dots + (-1)^n \frac{3^n}{(z-1)^n} + \dots \right).$$

Итак,

$$f(z) = \frac{1}{3} - \frac{z-1}{3^2} + \frac{(z-1)^2}{3^3} - \dots + (-1)^n \frac{(z-1)^n}{3^{n+1}} + \dots \quad \text{в круге } |z-1| < 3$$

и

$$f(z) = \frac{1}{z-1} - \frac{3}{(z-1)^2} + \frac{3^2}{(z-1)^3} - \dots + (-1)^n \frac{3^n}{(z-1)^{n+1}} + \dots \quad \text{для } |z-1| > 3.$$

Пример 2. $f(z) = z^3 \cos \frac{1}{z}$ разложить по степеням z .

Решение. Функция $f(z)$ аналитична при всех $z \neq 0$, т.е. для $|z| > 0$. В этой области воспользуемся разложением Маклорена функции $\cos w$ (см. (3.1.9)) при $w = \frac{1}{z}$:

$$\cos \frac{1}{z} = 1 - \frac{1}{2!z^2} + \frac{1}{4!z^4} - \dots + (-1)^n \frac{1}{(2n)!z^{2n}} + \dots$$

Почленно умножая на z^3 , получаем разложение

$$f(z) = z^3 - \frac{z}{2!} + \frac{1}{4!z} - \dots + (-1)^n \frac{1}{(2n)!z^{2n-3}} + \dots; |z| > 0.$$

3.4. ИЗОЛИРОВАННЫЕ ОСОБЫЕ ТОЧКИ

3.4.1. Особая точка z_0 функции $f(z)$, т.е. точка, в которой $f(z)$ не аналитична, называется *изолированной*, если в некоторой окрестности z_0 не существует других особых точек для $f(z)$. Другими словами, $f(z)$ аналитична в некоторой окрестности точки z_0 , но не в самой этой точке.

Окрестностью точки z_0 служит "кольцо" вида $R_1 < |z - z_0| < R_2$, для которого $R_1 = 0$. Разложение Лорана $f(z)$ в этом "кольце" называем *рядом Лорана в окрестности данной изолированной особой точки*.

Будем называть ряд

$$C_0 + C_1(z - z_0) + \dots + C_n(z - z_0)^n + \dots \quad (3.4.1)$$

правильной частью, а ряд

$$\frac{C_{-1}}{z - z_0} + \frac{C_{-2}}{(z - z_0)^2} + \dots + \frac{C_{-n}}{(z - z_0)^n} + \dots \quad (3.4.2)$$

– *главной* частью ряда Лорана

$$f(z) = \sum_{n=-\infty}^{\infty} C_n (z - z_0)^n. \quad (3.4.3)$$

Возможны случаи:

а) ряд Лорана (3.4.3) состоит из своей *правильной* части, т.е.

$$C_{-1} = C_{-2} = \dots = C_{-n} = \dots = 0,$$

тогда точку z_0 называем *устранимой* особой точкой (из дальнейших рассуждений станет ясно, почему выбран такой термин);

б) *главная* часть ряда Лорана содержит только конечное число ненулевых членов; например, она имеет вид

$$\frac{C_{-1}}{z - z_0} + \frac{C_{-2}}{(z - z_0)^2} + \dots + \frac{C_{-n}}{(z - z_0)^n}, \text{ т.е. } C_{-n-1} = C_{-n-2} = \dots = 0.$$

В этом случае точку z_0 называем *полюсом n -го порядка*; в частности, при $n = 1$ – *простым полюсом*;

в) если *главная* часть (3.4.2) имеет бесконечное количество ненулевых членов, то точка z_0 называется *существенно особой*.

3.4.2. Рассмотрим случай а). Имеет место

Теорема 1. Если в некоторой окрестности особой точки z_0 функция $f(z)$ ограничена, т.е. существует постоянная M , такая что $|f(z)| < M$ во всех точках этой окрестности, то z_0 – *устраняемая* особая точка.

▷ Достаточно доказать, что $C_{-n} = 0$ для всех $n = 1, 2, \dots$. Имеем согласно (3.3.6) для произвольной окружности ℓ с центром в точке z_0 радиуса ρ (столь малого, что окружность ℓ целиком расположена в указанной окрестности), что

$$|C_{-n}| = \left| \frac{1}{2\pi i} \int_{\ell} \frac{f(z)}{(z - z_0)^{-n+1}} dz \right| \leq \frac{1}{2\pi} M \frac{1}{\rho^{-n+1}} 2\pi\rho = M\rho^n,$$

(см. свойства интеграла в п. 1.4.2). В силу произвольной малости ρ все $C_{-n} = 0$, что и утверждалось. ◁

Итак, в окрестности *устраняемой* особой точки z_0

$$f(z) = C_0 + C_1(z - z_0) + \dots + C_n(z - z_0)^n + \dots \quad (3.4.4)$$

Если в точке z_0 доопределить функцию $f(z)$ суммой этого ряда при $z = z_0$, т.е. положить $f(z_0) = C_0$, то (3.4.4) как сумма степенного ряда в круге $|z - z_0| < R$ будет аналитична (в том числе, аналитична и в точке z_0). Тем самым мы как бы устранили "особенность" в точке z_0 , сделав эту точку *правильной*.

Например, при $z \neq 0$ из соотношения (3.1.8) получаем

$$\frac{\sin z}{z} = 1 - \frac{z^2}{3!} + \frac{z^4}{5!} - \dots + (-1)^n \frac{z^{2n}}{(2n+1)!} + \dots,$$

т.е. точка $z=0$ является особой и устранимой. Достаточно положить $\frac{\sin z}{z} = 1$ при $z=0$, чтобы эта точка стала правильной.

3.4.3. Рассмотрим случай б).

Теорема 2. Точка z_0 тогда и только тогда является полюсом n -го порядка, когда существует аналитическая в точке z_0 функция $\varphi(z)$, такая что $\varphi(z_0) \neq 0$ и

$$f(z) = \frac{\varphi(z)}{(z-z_0)^n}. \quad (3.4.5)$$

▷ Пусть z_0 – полюс n -го порядка, тогда (по определению)

$$f(z) = \frac{C_{-n}}{(z-z_0)^n} + \dots + \frac{C_{-2}}{(z-z_0)^2} + \frac{C_{-1}}{z-z_0} + \sum_{m=0}^{\infty} C_m (z-z_0)^m, \quad C_{-n} \neq 0;$$

иначе

$$f(z) = \frac{1}{(z-z_0)^n} \left\{ C_{-n} + (z-z_0)C_{n-1} + \dots + C_{-1}(z-z_0)^{n-1} + \right. \\ \left. + (z-z_0)^n \sum_{m=0}^{\infty} C_m (z-z_0)^m \right\}. \quad (3.4.6)$$

По условию (так как $f(z)$ в окрестности точки z_0 – аналитична), правильная часть ряда Лорана – сходящийся (в окрестности точки z_0) степенной ряд, сумму которого можно считать аналитичной и в точке z_0 , доопределив её значением C_0 . Следовательно, в фигурных скобках в (3.4.6) записана некоторая аналитическая функция $\varphi(z)$, т.е. мы пришли к (3.4.5); при этом $\varphi(z_0) = C_{-n} \neq 0$.

Обратно, пусть $f(z)$ имеет вид (3.4.5) и $\varphi(z_0) \neq 0$. Разложив $\varphi(z)$ в ряд вида (3.2.2), имеем

$$f(z) = \frac{1}{(z-z_0)^n} \left\{ C_0 + C_1(z-z_0) + \dots + C_n(z-z_0)^n + C_{n+1}(z-z_0)^{n+1} + \dots \right\} = \\ = \frac{C_0}{(z-z_0)^n} + \frac{C_1}{(z-z_0)^{n-1}} + \dots + C_n + C_{n+1}(z-z_0) + \dots, \quad (3.4.7)$$

где $C_0 = \varphi(z_0) \neq 0$. Таким образом, мы имеем в разложении Лорана (3.4.7) коэффициент C_0 перед $\frac{1}{(z-z_0)^n}$ не равным нулю,

т.е. z_0 оказывается полюсом n -го порядка. Теорема полностью доказана. ◁

Согласно (3.4.5), $z = z_0$ – полюс n -го порядка для $f(z)$ тогда и только тогда, когда эта точка является нулём n -го порядка для функции $\frac{1}{f(z)}$.

3.4.4 Теорема 3. Точка $z = z_0$ является полюсом n -го порядка функции $f(z)$ тогда и только тогда, когда

$$\lim_{z \rightarrow z_0} |f(z)| = \infty.$$

▷ Согласно формуле (3.4.5) в точке z_0 , являющейся полюсом n -го порядка,

$$\lim_{z \rightarrow z_0} |f(z)| = \lim_{z \rightarrow z_0} \frac{|\varphi(z)|}{|z-z_0|^n} = \infty,$$

так как $\varphi(z_0)$ существует в силу аналитичности $\varphi(z)$ в точке z_0 и $\varphi(z_0) \neq 0$.

Обратно, если

$$\lim_{z \rightarrow z_0} |f(z)| = \infty, \quad \text{то} \quad \lim_{z \rightarrow z_0} \frac{1}{|f(z)|} = 0,$$

а тогда точка z_0 являются нулём n -го порядка для функции

$$g(z) = \frac{1}{f(z)}, \quad \text{т.е.} \quad (z-z_0)^n \varphi(z) = \frac{1}{f(z)},$$

где $\varphi(z)$ аналитична в точке z_0 и $\varphi(z_0) \neq 0$ (см. п. 3.2.3).

Теперь

$$f(z) = \frac{1}{\frac{\varphi(z)}{(z-z_0)^n}},$$

где $\frac{1}{\varphi(z)}$ аналитична в точке z_0 , так как $\varphi(z_0) \neq 0$. Следовательно, $f(z)$ имеет вид (3.4.5), а тогда, по теореме 2, $z = z_0$ – полюс n -го порядка. \triangleleft

3.4.5. Рассмотрим, наконец, случай в) существенно особой точки $z = z_0$. Поведение $f(z)$ в её окрестности описывается следующей теоремой Сохоцкого-Вейерштрасса.

Теорема 4. Для любого $\varepsilon > 0$ и любого комплексного числа B в каждой окрестности существенно особой точки z_0 найдётся хотя бы одна точка z , такая что

$$|f(z) - B| < \varepsilon.$$

Смысл теоремы состоит в том, что в достаточно малых окрестностях существенно особой точки функция $f(z)$ может принять значение, сколь угодно близкое к наперёд заданному произвольному числу B .

\triangleright Предположим противное: для заданных B и $\varepsilon > 0$ найдётся такая окрестность U точки z_0 , что

$$|f(z) - B| \geq \varepsilon \text{ при всех } z \in U.$$

Но тогда функция

$$g(z) = \frac{1}{f(z) - B} \quad (3.4.8)$$

определена и ограничена в указанной окрестности точки z_0 , так как

$$|g(z)| = \frac{1}{|f(z) - B|} \leq \frac{1}{\varepsilon}.$$

Следовательно, согласно теореме 1, точка z_0 является устранимой особой точкой для $g(z)$. Тогда при $z \neq z_0$ функция $g(z)$ есть сумма степенного ряда (3.4.4), т.е. равна некоторой аналитической функции. При этом не исключено, что первые несколько коэффициентов C_0, C_1, \dots в (3.4.4) равны нулю, т.е.

$$g(z) = (z - z_0)^m \varphi(z) \text{ при некотором } m = 0, 1, \dots; \varphi(z_0) \neq 0.$$

Согласно (3.4.8)

$$f(z) = B + \frac{1}{g(z)} = B + \frac{1}{(z - z_0)^m \varphi(z)}. \quad (3.4.9)$$

Функция $\frac{1}{\varphi(z)}$ является аналитичной в окрестности точки z_0 . Действительно, вместе с $\varphi(z_0) \neq 0$ все значения $\varphi(z)$ из достаточно малой окрестности z_0 остаются ненулевыми, поскольку $\varphi(z)$ аналитична, а значит и непрерывна в этой окрестности. Согласно (3.4.9) для $m = 0$ $f(z)$ ограничена в окрестности точки z_0 , т.е. z_0 – устранимая особая точка; для $m \neq 0$ (3.4.9) означает, что z_0 – полюс m -го порядка. В обоих случаях получаем противоречие с условием теоремы. \triangleleft

В иной форме теорема Сохоцкого-Вейерштрасса звучит так: если z_0 – существенно особая точка функции $f(z)$, то для любого комплексного числа B найдётся последовательность точек z_k , такая что $z_k \rightarrow z_0$ и

$$\lim_{k \rightarrow \infty} f(z_k) = B.$$

Найдётся также и последовательность $z_k \rightarrow z_0$, такая что

$$\lim_{k \rightarrow \infty} |f(z_k)| = \infty.$$

3.4.6. Пример. Определить характер каждой особой точки функции

$$\text{а) } f(z) = \frac{z}{(z^2 + 4)^3}; \quad \text{б) } f(z) = \ln \frac{z+1}{z}.$$

Решение.

а) Запишем $f(z)$ в виде

$$f(z) = \frac{z}{(z^2 - (2i)^2)^3} = \frac{z}{(z-2i)^3(z+2i)^3}.$$

Изолированными особыми точками являются $z_1 = 2i$ и $z_2 = -2i$. Рассмотрим случай z_1 :

$$f(z) = \frac{1}{(z-2i)^3} \frac{z}{(z+2i)^3} = \frac{\varphi(z)}{(z-2i)^3}, \quad \text{где } \varphi(z) = \frac{z}{(z+2i)^3}.$$

В точке $z_1 = 2i$ функция $\varphi(z)$ аналитична, при этом $\varphi(z_1) = \frac{2i}{(4i)^3} = -\frac{1}{32} \neq 0$. Теперь, согласно теореме 2 (см. (3.4.5)) видим, что $z_1 = 2i$ – полюс третьего порядка. Аналогично, и точка $z_2 = -2i$ – полюс третьего порядка.

б) Запишем $f(z)$ в виде $f(z) = \ln\left(1 + \frac{1}{z}\right)$. Воспользовавшись разложением (3.1.10) Маклорена логарифмической функции для аргумента $\frac{1}{z}$ (вместо z) имеем

$$\ln\left(1 + \frac{1}{z}\right) = \frac{1}{z} - \frac{1}{2z^2} + \frac{1}{3z^3} - \dots + (-1)^n \frac{1}{(n+1)z^{n+1}} + \dots$$

Теперь мы видим, что $z = 0$ – существенно особая точка для $f(z)$.

4. ТЕОРИЯ ВЕРОЯТНОСТЕЙ. СЛУЧАЙНЫЕ СОБЫТИЯ

4.1. ЭЛЕМЕНТЫ КОМБИНАТОРИКИ

Рассмотрим конечное множество $X = \{x_1, x_2, \dots, x_n\}$, состоящее из n элементов. Из множества X можно образовать различные подмножества $V = \{x_{i_1}, x_{i_2}, \dots, x_{i_k}\}$, состоящие из k ($k \leq n$) элементов. Эти подмножества будем называть *выборками* объёма k из n различных элементов множества X .

Выборки, различающиеся лишь по виду входящих в них элементов называются *сочетаниями*. Число сочетаний может быть найдено по формуле

$$C_n^k = \frac{n!}{k!(n-k)!}, \quad (4.1.1)$$

где $n! = 1 \cdot 2 \cdot \dots \cdot n$. Например, из множества $X = \{A, B, B\}$, выбирая по 2 элемента ($n = 3, k = 2$), можно образовать 3 сочетания $\{A, B\}$, $\{A, B\}$ и $\{B, B\}$.

Выборки, различающиеся как видом элементов, так и порядком их вхождения в выборку, называются *размещениями*. Число размещений можно подсчитать по формуле

$$A_n^k = \frac{n!}{(n-k)!}. \quad (4.1.2)$$

Например, из множества $X = \{A, B, B\}$, выбирая по 2 элемента, можно образовать 6 размещений $\{A, B\}$, $\{B, A\}$, $\{A, B\}$, $\{B, A\}$, $\{B, B\}$ и $\{B, B\}$.

Размещения из n элементов по $k = n$ называются перестановками. Различные перестановки содержат одни и те же элементы, расположенные в разном порядке. Их число согласно (4.1.2) есть:

$$P_n = A_n^n = n!, \quad (4.1.3)$$

так как по определению $0! = 1$.

Размещения из n элементов по k , в которых элементы могут повторяться, называются *размещениями с повторениями*. Число таких размещений может быть подсчитано по формуле

$$Q_n^k = n^k. \quad (4.1.4)$$

Возвращаясь к нашему примеру, к 6 перечисленным выше размещениям без повторений необходимо добавить ещё 3 размещения $\{A, A\}$, $\{B, B\}$ и $\{B, B\}$.

4.2. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Из повседневной практики известно, что для широкого круга явлений наблюдается неоднозначность исхода при повторении опыта (*испытания*) с сохранением условий его проведения. События, связанные с такими явлениями, называются *случайными*. Случайными событиями являются, например, выпадение «герба» при подбрасывании монеты, результаты измерения температуры окружающей среды и т.д. Известно также, что одни случайные события наступают часто, другие – реже. Ясно, что такие характеристики событий довольно расплывчаты. Более объективной экспериментальной характеристикой случайного события, которое обозначим A , является *относительная частота*

$$h_m(A) = \frac{m_A}{m}, \quad (4.2.1)$$

где m_A – число испытаний, в которых событие A наступило, а m – общее число испытаний.

Опытным путём установлено, что при достаточно большом числе испытаний относительные частоты появления события A в различных сериях испытаний мало отличаются друг от друга. Говорят, что относительная частота при увеличении m перестаёт носить случайный характер, становясь почти постоянной. Это свойство называют статистической устойчивостью относительных частот. Таким образом, относительные частоты в различных сериях испытаний группи-

руются около некоторого числа, которое называется вероятностью случайного события A и обозначается символом $P(A)$.

Такое определение вероятности весьма неопределённо. Более точное описание вероятности возможно в случае построения математической модели случайного явления. Раздел математики, предметом изучения которого являются математические модели случайных явлений, называется теорией вероятностей.

4.3. ПРОСТРАНСТВО ЭЛЕМЕНТАРНЫХ СОБЫТИЙ

Предположим, что в результате некоторого испытания могут наступать как *элементарные (неразложимые) события* $\omega_1, \omega_2, \dots, \omega_n$, так и *составные события*, состоящие из нескольких элементарных. Например, событие $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}\}$ состоит из k ($k \leq n$) элементарных. В дальнейшем под составными событиями будем подразумевать просто события, обозначая их большими буквами латинского алфавита. Совокупность всех элементарных событий $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ называется *конечным пространством элементарных событий*.

Пример 1. Испытание состоит в подбрасывании игральной кости, на гранях которой отмечены числа от 1 до 6. Если через ω_i обозначить событие, состоящее в выпадении i очков, то элементарными событиями являются события $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6$, которые и образуют пространство элементарных событий Ω . В таком случае составное событие $A = \{\omega_1, \omega_3, \omega_5\}$ означает, что выпало нечётное число очков.

Пример 2. Испытание – бросание двух игральных костей. Пространство элементарных событий состоит из 6^2 элементарных событий ω_{ij} и может быть записано так: $\Omega = \{\omega_{ij} : i, j = 1, 2, \dots, 6\}$. Тогда составные события $A = \{\omega_{15}, \omega_{24}, \omega_{33}, \omega_{42}, \omega_{51}\}$ и $B = \{\omega_{16}, \omega_{23}, \omega_{32}, \omega_{61}\}$ означают, что сумма и произведение выпавших чисел равны 6.

Рассмотрим случай *счётного пространства элементарных событий* $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$, где n – натуральное число. Здесь уже составное событие $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}, \dots\}$ может и не состоять из конечного числа элементарных событий.

Пример 3. Артиллерийское орудие стреляет по цели до тех пор пока она не будет поражена. Под элементарным событием ω_n понимаем событие, когда на поражение цели потребовалось n выстрелов, т.е. ни в одном из первых $n - 1$ выстрелов цель не была поражена. Составное событие $A = \{\omega_2, \omega_4, \dots, \omega_{2k}, \dots\}$ означает, что на поражение цели потребуется чётное число выстрелов.

Более общий случай пространства элементарных событий изложим на следующем примере.

Пример 4. Испытание состоит в стрельбе по плоской квадратной мишени. Введём декартову систему координат xOy с центром в середине квадрата, направив оси координат параллельно его сторонам, длины которых равны $2a$. Элементарными событиями в этом испытании являются точки мишени $\omega = (x, y)$. Тогда пространство элементарных событий можно представить в виде

$$\Omega = \{(x, y) : -a \leq x \leq a, -a \leq y \leq a\}, \quad (4.3.1)$$

а событие, состоящее в попадании в круг радиуса $r < a$ с центром в точке O , можно записать так:

$$A = \{(x, y) : x^2 + y^2 \leq r^2\}. \quad (4.3.2)$$

4.4. ОПЕРАЦИИ НАД СОБЫТИЯМИ

Рассмотрим пространство элементарных событий Ω и некоторые два события A и B , состоящие из элементарных событий $\omega \in \Omega$.

Суммой событий A и B называется событие $A + B$ (используется также обозначение $A \cup B$), состоящее из элементарных событий, принадлежащих хотя бы одному из событий A или B , т.е., если $\omega \in A + B$, то $\omega \in A$ или $\omega \in B$. Событие $A + B$ означает, что произошло событие A или B , или оба события одновременно.

Введённое понятие суммы событий можно наглядно иллюстрировать с помощью примера 4. Пусть событиями A и B являются попадания соответственно в большой и малый круги (рис. 4.4.1). В таком случае событием $A + B$ является заштрихованная область на рис. 4.4.1, а.

Произведением событий A и B называется событие AB (или $A \cap B$), состоящее из элементарных событий, принадлежащих и A и B , т.е., если $\omega \in AB$, то $\omega \in A$ и $\omega \in B$. Событие AB означает, что события A и B произошли одновременно, что наглядно показано на рис. 4.4.1, б.

Разностью событий A и B называется событие $A - B$ (или $A \setminus B$), состоящее из элементарных событий, принадлежащих A и не принадлежащих B , т.е., если $\omega \in A - B$, то $\omega \in A$ и $\omega \notin B$. Событие $A - B$ означает, что событие A произошло, а B не произошло. Сказанное иллюстрируется рис. 4.4.1, в.

Событие \bar{A} называется *противоположным* событию A (рис. 4.4.1, г), если $\omega \in \bar{A}$ и $\omega \notin A$, т.е. $A + \bar{A} = \Omega$. Событие \bar{A} означает, что событие A не произошло.

Событие называется *достоверным*, если оно всегда происходит. Ясно, что Ω является достоверным событием. Событие \emptyset называется *невозможным*, если оно никогда не происходит. Отметим полезные соотношения, связывающие эти понятия:

$$\bar{\Omega} = \emptyset, \quad \bar{\emptyset} = \Omega, \quad A\bar{A} = \emptyset. \quad (4.4.1)$$

События A и B называются *несовместными*, если $AB = \emptyset$. Например, события A и B , описанные в примере 2, несовместны. Более общее определение формулируется следующим образом. События $A_1, A_2, \dots, A_n, \dots$ называются *попарно несовместными*, если $A_i A_j = \emptyset$ при любых $i \neq j$.

Под обозначением $A \subset B$ (или $B \subset A$) понимаем, что событие A влечёт за собой B , т.е. из наступления события A вытекает B (или наоборот). Если $A \subset B$ и $B \subset A$, то говорят, что события A и B *равносильны* по качеству упражнения, закрепляющего эти равенства:

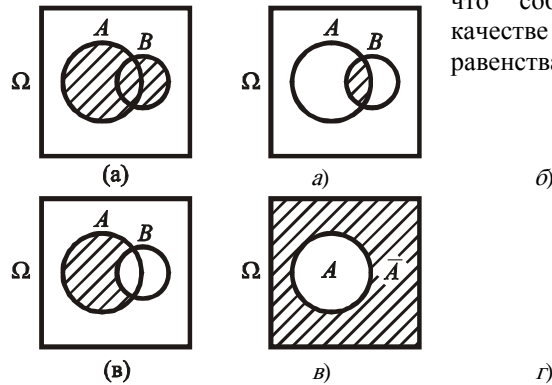


Рис. 4.4.1

$$\overline{A+B} = \overline{A}\overline{B}, \quad (A+B)C = AC + BC. \quad (4.4.2)$$

4.5. АКСИОМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Рассмотрим пространство элементарных событий Ω и его некоторый класс подмножеств E . Этот класс подмножеств называется *алгеброй событий*, если $\Omega \in E$, $A+B \in E$, $AB \in E$ и $A-B \in E$ для любых $A \in E$, $B \in E$.

Теперь мы можем дать определение вероятности. *Вероятностью называется числовая функция P , определённая на алгебре событий E* , при условии выполнения следующих аксиом:

A1. *Аксиома неотрицательности.*

$$P(A) \geq 0 \text{ для любого } A \in E. \quad (4.5.1)$$

A2. *Аксиома нормированности.*

$$P(\Omega) = 1. \quad (4.5.2)$$

A3. *Аксиома аддитивности.*

$$P(A+B) = P(A) + P(B), \quad (4.5.3)$$

если A и B несовместные события.

Замечание 1. Относительная частота $h_m(A)$ события A удовлетворяет аксиомам A1 – A3. Действительно, согласно (4.2.1) имеем

$$h_m(A) = \frac{m_A}{m} \geq 0, \quad h_m(\Omega) = \frac{m_\Omega}{m} = \frac{m}{m} = 1, \\ h_m(A+B) = \frac{m_{A+B}}{m} = \frac{m_A}{m} + \frac{m_B}{m} = h_m(A) + h_m(B),$$

где m_A , m_B и m_{A+B} обозначают число испытаний, в которых происходят события A , B и $A+B$. При этом $m_{A+B} = m_A + m_B$, так как события A и B несовместны.

В случае бесконечных последовательностей событий аксиомы A1 – A3 необходимо дополнить следующей аксиомой:

A4. *Расширенная аксиома аддитивности.*

$$P(A) = \sum_{n=1}^{\infty} P(A_n), \quad A = \sum_{n=1}^{\infty} A_n \in E, \quad (4.5.4)$$

если $A_1, A_2, \dots, A_n, \dots$ попарно несовместные события.

Тройку (Ω, E, P) называют *вероятностным пространством*, если вероятность P удовлетворяет аксиомам A1 – A4, а алгебра событий E дополнительно содержит счётные суммы и произведения событий.

Сформулируем следствия из аксиом теории вероятностей.

Следствие 1. Вероятность противоположного события равна

$$P(\overline{A}) = 1 - P(A). \quad (4.5.5)$$

▷ Поскольку $A + \overline{A} = \Omega$, то из аксиом A2 и A3 имеем

$$1 = P(\Omega) = P(A + \overline{A}) = P(A) + P(\overline{A}),$$

откуда вытекает формула (4.5.5). ◁

Следствие 2. Вероятность невозможного события равна нулю, т.е. $P(\emptyset) = 0$.

▷ Доказательство следует из равенств (4.4.1), (4.5.5) и аксиомы A2:

$$P(\emptyset) = P(\overline{\Omega}) = 1 - P(\Omega) = 1 - 1 = 0,$$

Следствие 3. Для произвольных событий A и B имеет место равенство $P(A+B) = P(A) + P(B) - P(AB)$. что и требовалось доказать. ◁ (4.5.6)

▷ Представим события $A+B$ и B в виде суммы несовместных событий: $A+B = A + \overline{A}B$ и $B = AB + \overline{A}B$, что можно проиллюстрировать с помощью примера 4, понимая под событиями A и B попадание соответственно в большой и малый круги (рис. 4.5.2). Используя далее аксиому A3, получим

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

Вычитая из первого равенства второе, находим

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

откуда следует требуемая формула. ◁

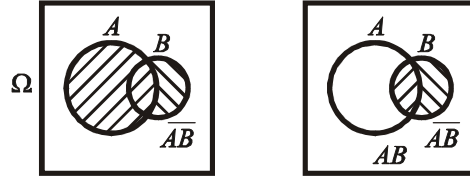


Рис. 4.5.1

Формулу (4.5.6) можно распространить и на большее число событий. Например, для трёх событий A, B, C имеем

$$P(A+B+C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC). \quad (4.5.7)$$

Следствие 4. Для произвольных событий A и B имеет место неравенство

$$P(A+B) \leq P(A) + P(B). \quad (4.5.8)$$

▷ Доказательство вытекает из (4.5.6) и аксиомы A1. ◁

4.6. КОНЕЧНОЕ ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО

В предыдущем параграфе было введено понятие вероятностного пространства (Ω, E, P) . Теперь мы рассмотрим несколько частных случаев вероятностных пространств.

Пусть $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ – конечное пространство элементарных событий, а E – алгебра событий, т.е. множество всех подмножеств Ω . Каждому элементарному событию ω_i поставим в соответствие некоторое число $p(\omega_i) \geq 0$, которое будем называть *элементарной вероятностью*. При этом элементарные вероятности должны удовлетворять условию

$$\sum_{i=1}^n p(\omega_i) = 1. \quad (4.6.1)$$

Рассмотрим составное событие $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}\}$. Вероятностью события $A \in E$ называется число $P(A)$, определяемое формулой

$$P(A) = \sum_{j=1}^k p(\omega_{i_j}) = \sum_{\omega \in A} p(\omega), \quad (4.6.2)$$

т.е. вероятность события A равна сумме вероятностей элементарных событий $\omega \in A$.

Нетрудно проверить, что в силу соотношений (4.6.1), (4.6.2) аксиомы A1 – A3 выполняются. Действительно,

$$P(A) \geq 0 \text{ для каждого } A \in E;$$

$$P(\Omega) = \sum_{i=1}^n p(\omega_i) = 1;$$

$$P(A+B) = \sum_{\omega \in A+B} p(\omega) = \sum_{\omega \in A} p(\omega) + \sum_{\omega \in B} p(\omega) = P(A) + P(B),$$

если $AB = \emptyset$.

Таким образом, конечное вероятностное пространство, которое будем также называть *конечной схемой*, построено.

Изучим подробнее важный случай конечной схемы с применением к явлениям, в которых элементарные вероятности одинаковы, т.е. $p(\omega) = 1/n$ для каждого $\omega \in \Omega$. Тогда из формулы (4.6.2) вытекает

$$P(A) = \sum_{j=1}^k \frac{1}{n} = \frac{k}{n}. \quad (4.6.3)$$

Формулу (4.6.3) нередко называют *классическим определением вероятности* и она выражает тот факт, что вероятность составного события A равна отношению числа элементарных событий $\omega \in A$ к общему числу элементарных событий.

Пример 5. В урне находятся L шаров, из них M белых и $L - M$ чёрных шаров. Какова вероятность того, что среди l выбранных шаров окажется m белых и $l - m$ чёрных шаров? В качестве элементарных событий рассматриваем все выборки из L шаров по l . Число элементарных событий l пространства Ω можно найти, воспользовавшись формулой для числа сочетаний из L элементов по l (п. 4.1), т.е. $n = C_L^l$. Число элементарных событий, входящих в интересующее нас событие A_m определим по следующей двухступенчатой схеме. Вначале составляем выборки из одних белых шаров, число которых есть C_M^m . Затем отдельно подсчитываем все выборки из чёрных шаров, которых будет C_{L-M}^{l-m} . Таким образом, находим $k = C_M^m C_{L-M}^{l-m}$. Поскольку элементарные события равновероятны, то в силу классического определения вероятности (4.6.3) имеем

$$P(A_m) = P(m, l, M, L) = \frac{C_M^m C_{L-M}^{l-m}}{C_L^l}. \quad (4.6.4)$$

4.7. СЧЁТНОЕ ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО

Рассмотрим более общий случай пространства элементарных событий. Пусть $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ – счётное пространство элементарных событий и E – множество всех его подмножеств. Элементарным событиям ω_n поставим в соответствие, как и прежде, элементарные вероятности $p(\omega_n) \geq 0$, удовлетворяющие следующему условию:

$$\sum_{n=1}^{\infty} p(\omega_n) = 1. \quad (4.7.1)$$

Вероятностью события $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}, \dots\} \in E$ называется число $P(A)$, определяемое формулой

$$P(A) = \sum_{k=1}^{\infty} p(\omega_{i_k}) = \sum_{\omega \in A} p(\omega), \quad (4.7.2)$$

т.е. в общем случае вероятность события A равна сумме ряда, составленного из элементарных вероятностей $p(\omega)$, $\omega \in A$.

Повторяя рассуждения конечной схемы (п. 4.6) можно убедиться, что все аксиомы, включая расширенную аксиому аддитивности, выполняются. Таким образом, счётное вероятностное пространство, которое нередко называют *счётной схемой*, построено.

Пример 6. Обратимся к рассмотренной в примере 3 задаче об артиллерийском орудии. Напомним, что элементарное событие ω_n означает, что в первых $n - 1$ выстрелах цель не поражена, в последнем же n -м выстреле – попадание в цель. Вероятность попадания в цель при каждом выстреле одинакова и равна p , тогда вероятность промаха будет $q = 1 - p$. В качестве элементарных вероятностей принимаем набор чисел $p(\omega_n) = pq^{n-1}$, справедливость которого легко обосновать, исходя из понятия последовательности независимых испытаний (п. 4.11). Заметим, что выбранные нами элементарные вероятности удовлетворяют условию (4.7.1). Действительно, с учётом известной формулы для суммы ряда, составленного из членов убывающей геометрической прогрессии, имеем

$$\sum_{n=1}^{\infty} p(\omega_n) = p \sum_{n=1}^{\infty} q^{n-1} = \frac{p}{1-q} = 1.$$

Далее, с использованием определения (4.7.2) находим вероятность описанного в примере 3 события A (на поражение цели потребуются чётное число выстрелов):

$$P(A) = \sum_{k=1}^{\infty} p(\omega_{2k}) = p \sum_{k=1}^{\infty} q^{2k-1} = \frac{pq}{1-q^2}.$$

4.8. НЕПРЕРЫВНОЕ ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО

Определение непрерывного вероятностного пространства в общем случае использует понятия, которые выходят за рамки стандартного курса математики для инженерно-технических специальностей вузов. Однако один частный случай может быть рассмотрен с помощью примера 4, где в качестве пространства элементарных событий Ω выбран квадрат со сторонами $2a$. При этом элементарными событиями являются точки квадрата $\omega = (x, y)$ (рис. 4.8.1).

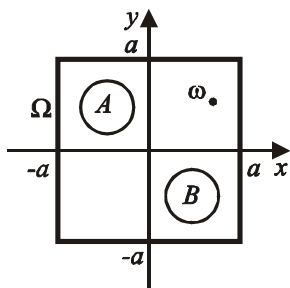


Рис. 4.8.1

Предположим, что существует некоторая непрерывная на Ω функция $p(x, y) \geq 0$, такая что двойной интеграл от нее по области Ω равен 1, т.е.

$$\int_{-a}^a \int_{-a}^a p(x, y) dx dy = 1. \quad (4.8.1)$$

Вероятностью события A назовём число $P(A)$, определяемое формулой

$$P(A) = \iint_A p(x, y) dx dy. \quad (4.8.2)$$

Используя свойства двойного интеграла, нетрудно проверить, что аксиомы $A1 - A3$ выполняются, в частности,

$$P(A + B) = \iint_{A+B} p(x, y) dx dy = \iint_A p(x, y) dx dy + \iint_B p(x, y) dx dy = P(A) + P(B),$$

если события A и B несовместны (рис. 4.8.1).

Рассмотрим простейший случай, когда

$$p(x, y) = \frac{1}{4a^2} \text{ на } \Omega, \quad (4.8.3)$$

что позволяет удовлетворить условию (4.8.1). Подставив функцию (4.8.3) в интеграл (4.8.2) и интегрируя, получим

$$P(A) = \frac{1}{4a^2} \iint_A dx dy = \frac{S_A}{S_\Omega}, \quad (4.8.4)$$

где S_A – площадь области A ; $S_\Omega = 4a^2$ – площадь квадрата. Такое определение вероятности называют *геометрическим*. Геометрическое определение вероятности (4.8.4) можно трактовать как обобщение классического определения вероятности (4.6.3), подразумевая при этом, что попадание точки в любую часть множества Ω равновероятно.

4.9. УСЛОВНАЯ ВЕРОЯТНОСТЬ. НЕЗАВИСИМОСТЬ СОБЫТИЙ

Здесь нас будет интересовать зависимость между вероятностями событий A , B и вероятностью их произведения AB . Чтобы осмыслить эту зависимость, обратимся снова к относительным частотам (4.2.1). Пусть проводится m испытаний, в которых m_A раз произошло событие A . Предположим, что среди m_A испытаний наряду с событием A m_{AB} раз произошло событие B . Рассмотрим отношение $h_m(B|A) = m_{AB}/m_A$, которое является относительной частотой события B по отношению к испытаниям с исходом A . Элементарные выкладки приводят к формуле

$$h_m(B|A) = \frac{m_{AB}/m}{m_A/m} = \frac{h_m(AB)}{h_m(A)},$$

Определение 1. Условной вероятностью события B при условии A называется число $h_m(B|A)$, которая может быть положена в основу определения условной вероятности.

$$P(B|A) = \frac{P(AB)}{P(A)}, \text{ если } P(A) > 0. \quad (4.9.1)$$

Условную вероятность иногда обозначают символом $P_A(B)$.

Из равенства (4.9.1) следует полезная формула

$$P(AB) = P(A)P(B|A), \quad (4.9.2)$$

которую можно распространить и на большее число событий. Так, для трёх событий A, B, C имеем

$$P(ABC) = P(A)P(B|A)P(C|AB). \quad (4.9.3)$$

Заметим, что формулы (4.9.2), (4.9.3) в общем случае не пригодны для вычисления вероятностей произведений событий, поскольку для этой цели необходимо знание условных вероятностей. Однако в некоторых случаях они могут оказаться полезными.

Пример 7. Пусть имеются две урны, в которых общее число шаров равно соответственно L_1 и L_2 , а число белых шаров – M_1 и M_2 . Испытание заключается в том, что из первой урны наудачу перекалывается один шар во вторую урну, затем из второй урны также наугад вынимается один шар. Найдём вероятность того, что оба выбранных шара являются белыми. В качестве событий A и B рассматриваем появление белого шара соответственно из первой и второй урн, тогда интересующее нас событие будет AB . Вероятность этого события находим, используя формулу (4.9.2) и определение классической вероятности (4.6.3):

$$P(AB) = P(A)P(B|A) = \frac{M_1}{L_1} \frac{M_2 + 1}{L_2 + 1}.$$

Введём теперь понятие независимости событий. Здесь возможны два подхода.

Определение 2. События A и B называются *независимыми*, если вероятность их произведения равна произведению вероятностей, т.е. справедливо равенство

$$P(AB) = P(A)P(B). \quad (4.9.4)$$

Другой подход связан с понятием условной вероятности. Сравнивая формулы (4.9.1) и (4.9.4), замечаем, что формула (4.9.4) равносильна следующей формуле:

$$P(B|A) = \frac{P(AB)}{P(A)} = P(B), \quad P(A) > 0.$$

По аналогии имеем

$$P(A|B) = \frac{P(AB)}{P(B)} = P(A), \quad P(B) > 0.$$

Определение 3. События A и B *независимы* тогда и только тогда, когда условные вероятности этих событий совпадают с их *безусловными* вероятностями.

Это определение можно обобщить на случай нескольких событий.

Определение 4. События A_1, A_2, \dots, A_n называются *независимыми в совокупности*, если для всех $1 \leq i_1 < i_2 < \dots < i_k \leq n$, где $k = 2, 3, \dots, n$, имеем

$$P(A_{i_1} A_{i_2} \dots A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k}). \quad (4.9.5)$$

Если равенство (4.9.5) справедливо только для $k = 2$, то события A_1, A_2, \dots, A_n называются *парно независимыми*.

Для практики полезно отметить, что независимость событий A и B влечёт за собой независимость противоположных событий \bar{A} и \bar{B} . Это соображение нередко позволяет упростить вычисление вероятностей.

Пример 8. Два стрелка независимо друг от друга стреляют по одной и той же цели. Вероятность попадания для первого стрелка равна p , для второго q . Определим систему событий: A – попадание первого стрелка в цель, B – попадание второго стрелка в цель. Так как в условии говорится о независимых событиях A и B , а значит о независимых противоположных событиях \bar{A} и \bar{B} , то вероятности событий AB (цель поражена обоими стрелками) и $\bar{A}\bar{B}$ (цель не поражена ни одним из стрелков) равны

$$\begin{aligned} P(AB) &= P(A)P(B) = pq, \\ P(\bar{A}\bar{B}) &= P(\bar{A})P(\bar{B}) = (1-p)(1-q). \end{aligned}$$

Для полноты картины найдём вероятности событий $A+B$ (хотя бы один стрелок попадет в цель) и $\bar{A}\bar{B} + \bar{A}B$ (только один стрелок попадет в цель). Применяя формулу (4.5.6) и аксиому аддитивности А3, поскольку события $\bar{A}\bar{B}$ и $\bar{A}B$ являются несовместными, получим

$$\begin{aligned} P(A+B) &= P(A) + P(B) - P(AB) = p + q - pq, \\ P(\bar{A}\bar{B} + \bar{A}B) &= P(\bar{A}\bar{B}) + P(\bar{A}B) = P(\bar{A})P(\bar{B}) + P(\bar{A})P(B) = p(1-q) + q(1-p). \end{aligned}$$

4.10. ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ. ФОРМУЛА БАЙЕСА

Рассмотрим пространство элементарных событий Ω и некоторую группу событий $A_i \in \Omega, i = 1, 2, \dots, m$.

Определение 5. События A_1, A_2, \dots, A_m образуют *полную группу*, если $\Omega = A_1 + A_2 + \dots + A_m$ и $A_i A_j = \emptyset$ при любых $i \neq j$.

Утверждение 1. Если события A_1, A_2, \dots, A_m образуют полную группу, то сумма их вероятностей равна 1, т.е.

$$\sum_{i=1}^m P(A_i) = 1. \quad (4.10.1)$$

▷ Формула (4.10.1) непосредственно вытекает из аксиом вероятностей А2, А3 и определения 5. Действительно,

$$1 = P(\Omega) = P\left(\sum_{i=1}^m A_i\right) = \sum_{i=1}^m P(A_i),$$

что и требовалось доказать. ◁

Полученная формула часто применяется для контроля расчёта вероятностей, позволяя обнаружить неверно введённую «полную» группу событий.

Утверждение 2. Если события A_1, A_2, \dots, A_m образуют полную группу и событие $B \in \Omega$, тогда имеет место *формула полной вероятности*

$$P(B) = \sum_{i=1}^m P(A_i)P(B|A_i). \quad (4.10.2)$$

▷ Для доказательства формулы (4.10.2) представим событие B в виде суммы попарно несовместных событий

$$B = A_1 B + A_2 B + \dots + A_m B,$$

отсюда, принимая во внимание аксиому аддитивности А3 и правило умножения вероятностей (4.9.2), получим требуемую формулу

$$P(B) = P\left(\sum_{i=1}^m A_i B\right) = \sum_{i=1}^m P(A_i B) = \sum_{i=1}^m P(A_i)P(B|A_i). \quad \triangleleft$$

Замечание 2. События A_1, A_2, \dots, A_m из формулы полной вероятности обычно называют *гипотезами*. Смысл этого названия проясняется в следующем примере.

Пример 9. Обратимся к сформулированной в примере 7 задаче об урнах. Теперь только нам неизвестен цвет шара, перекладываемого из первой урны во вторую. Найдём вероятность события B , состоящего в том, что выбранный из второй урны шар является белым. В качестве гипотез A_1 и A_2 введём выбор соответственно белого и чёрного шара из первой урны. Вероятности этих гипотез равны

$$P(A_1) = \frac{M_1}{L_1}, \quad P(A_2) = \frac{L_1 - M_1}{L_1},$$

и удовлетворяют условию (4.10.1). По формуле полной вероятности находим

$$\begin{aligned} P(B) &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) = \\ &= \frac{M_1}{L_1} \frac{M_2 + 1}{L_2 + 1} + \frac{L_1 - M_1}{L_1} \frac{M_2}{L_2 + 1}. \end{aligned} \quad (4.10.3)$$

Если предположить, что в обеих урнах одинаковый набор белых и чёрных шаров ($M_1 = M_2 = M$ и $L_1 = L_2 = L$), тогда приходим к любопытному результату $P(B) = M/L$. Дело в том, что вероятность извлечения белого шара из второй урны до перекладывания в неё шара из первой урны тоже равна M/L . Этот результат нетрудно обосновать, однако, мы этого делать не будем.

Рассмотрим одно интересное применение формулы полной вероятности. Если в выражении для условной вероятности

$$P(A_k | B) = \frac{P(A_k B)}{P(B)} = \frac{P(A_k)P(B|A_k)}{P(B)} \quad (k = 1, 2, \dots, m)$$

заменить вероятность $P(B)$ по формуле полной вероятности (4.10.2), то получим *формулу Байеса*

$$P(A_k | B) = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^m P(A_i)P(B|A_i)}. \quad (4.10.4)$$

Эта формула применяется для вычисления условных вероятностей $P(A_k | B)$ гипотез A_k после испытания, в котором произошло событие B . Иными словами, формула Байеса позволяет дать переоценку по результатам проведённого испытания *априорных вероятностей* $P(A_k)$, которые были предварительно известны до испытания.

Пример 10. Вернёмся к примеру 9 и пусть нам известно, что из второй урны извлечён белый шар. Вычислим условную вероятность гипотезы A_1 при этом условии, т.е. вероятность того, что из первой урны переложили во вторую белый шар, зная наперёд, что вынутый из второй урны шар также белый. По формуле Байеса, учитывая равенство (4.10.3), находим

$$P(A_1 | B) = \frac{P(A_1)P(B|A_1)}{P(B)} = \frac{M_1(M_2 + 1)}{M_1(M_2 + 1) + M_2(L_1 - M_1)}.$$

Если предположить, как в примере 9, что в обеих урнах содержатся одинаковые наборы белых и чёрных шаров, то снова приходим к интересному результату $P(A_1 | B) = (M + 1)/(L + 1)$.

4.11. СХЕМА БЕРНУЛЛИ

Пусть происходят n *независимых* испытаний, в каждом из которых может появиться событие A с вероятностью p или противоположное событие \bar{A} с вероятностью $q = 1 - p$. Если в определённом испытании появляется событие A , то принято говорить об «успехе», в противном случае – о «неудаче». Описанная схема независимых испытаний называется *схемой Бернулли*.

В качестве пространства элементарных событий $\Omega = \{\omega\}$ удобно принять множество всех возможных цепочек из нулей и единиц, т.е. $\omega = (\alpha_1, \alpha_2, \dots, \alpha_n)$, где $\alpha_i = 0$ или 1 , $i = 1, 2, \dots, n$. Ясно, что число успехов (число единиц) в цепочке $(\alpha_1, \alpha_2, \dots, \alpha_n)$ равно сумме $\alpha_1 + \alpha_2 + \dots + \alpha_n$. Это позволяет в силу независимости испытаний представить элементарное событие ω в виде произведения n независимых событий, каждое из которых есть результат соответствующего испытания, и, применяя правило умножения вероятностей (4.9.5), получить формулу для элементарных вероятностей:

$$p(\omega) = p^{\alpha_1 + \alpha_2 + \dots + \alpha_n} q^{n - (\alpha_1 + \alpha_2 + \dots + \alpha_n)}. \quad (4.11.1)$$

Пусть в n испытаниях наступили ровно m успехов. Обозначим это событие через B_m и с учётом наших обозначений представим его в виде

$$B_m = \{(\alpha_1, \alpha_2, \dots, \alpha_n) : \alpha_1 + \alpha_2 + \dots + \alpha_n = m\}.$$

Теорема 1. Вероятность того, что в n испытаниях схемы Бернулли наступили m успехов определяется формулой

$$P(B_m) = P_n(m) = C_n^m p^m q^{n-m}, \quad (4.11.2)$$

где p и $q = 1 - p$ – вероятности успеха и неудачи в отдельном испытании.

▷ Согласно формуле (4.11.1) для любого элементарного события $\omega \in B_m$ имеем $p(\omega) = p^m q^{n-m}$. Для того, чтобы опре-

делить число элементарных событий, входящих в B_m , заметим, что оно совпадает с числом способов выбора на m мест в цепочке $\omega = (\alpha_1, \alpha_2, \dots, \alpha_n)$ единиц, в то время как оставшиеся $n - m$ мест однозначно заполняются нулями. Ясно, что речь идёт о сочетаниях, поэтому число элементарных событий равно C_n^m . Таким образом, в силу определения вероятности в конечной схеме (4.6.2) имеем

$$P(B_m) = \sum_{\omega \in B_m} p(\omega) = \sum_{\omega \in B_m} p^m q^{n-m} = C_n^m p^m q^{n-m},$$

что и требовалось доказать. \triangleleft

Формулу (4.11.2) называют формулой Бернулли в честь знаменитого швейцарского математика Якоба Бернулли (1654 – 1705).

4.12. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ В СХЕМЕ БЕРНУЛЛИ

При больших значениях n , а также малых p или q использование формулы Бернулли (4.11.2) становится затруднительным. В таких случаях для вычисления вероятностей $P_n(m)$ применяют приближённые (асимптотические) формулы, которые формулируются в следующих предельных теоремах.

Теорема 2 (теорема Пуассона). Пусть n велико, p мало и $\lambda = np$, тогда имеет место формула

$$P_n(m) \approx \frac{\lambda^m}{m!} e^{-\lambda}. \quad (4.12.1)$$

\triangleright Формулу (4.11.2) после элементарных преобразований представим в виде

$$\begin{aligned} P_n(m) &= C_n^m p^m (1-p)^{n-m} = \frac{n(n-1)\dots(n-m+1)}{m!} \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} = \\ &= \frac{\lambda^m}{m!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{m-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{-m} \left(1 + \frac{(-\lambda)}{n}\right)^n. \end{aligned}$$

Используя второй замечательный предел, получим

$$P_n(m) \rightarrow \frac{\lambda^m}{m!} e^{-\lambda} \quad \text{при } n \rightarrow \infty,$$

откуда следует утверждение теоремы. \triangleleft

Замечание 3. Вероятность того, что при достаточно большом числе независимых испытаний произойдёт не менее m_1 и не более m_2 успехов можно найти по формуле

$$P_n(m_1 \leq m \leq m_2) \approx \sum_{m=m_1}^{m_2} \frac{\lambda^m}{m!} e^{-\lambda}, \quad (4.12.2)$$

непосредственно вытекающей из аксиомы аддитивности А3 и формулы (4.12.1).

Замечание 4. При малых значениях q асимптотическую формулу Пуассона можно использовать для числа неудач.

Теорема 3 (локальная теорема Муавра-Лапласа). Пусть n велико и $x_m = (m - np) / \sqrt{npq}$, тогда имеет место формула

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \varphi(x_m), \quad \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (4.12.3)$$

Доказательство теоремы 3 выходит за рамки нашего курса теории вероятностей.

Следующую предельную теорему также приведём без доказательства.

Теорема 4 (интегральная теорема Муавра-Лапласа). Вероятность того, что при достаточно большом числе испытаний схемы Бернулли произойдёт не менее m_1 и не более m_2 успехов выражается формулой

$$P_n(m_1 \leq m \leq m_2) \approx \Phi(x_{m_2}) - \Phi(x_{m_1}), \quad (4.12.4)$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt, \quad x_{m_i} = \frac{m_i - np}{\sqrt{npq}}, \quad i = 1, 2.$$

Замечание 5. Асимптотические формулы Муавра-Лапласа (4.12.3), (4.12.4) рекомендуется применять в случае, когда $npq > 10$. В противном случае к более точным результатам приводят формулы Пуассона (4.12.1), (4.12.2).

В заключение отметим, что вычисления по асимптотическим формулам (4.12.1) – (4.12.4) выполняются с использованием специальных таблиц, которые приведены практически во всех учебниках и сборниках задач по теории вероятностей.

5. ТЕОРИЯ ВЕРОЯТНОСТЕЙ. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

5.1. ПОНЯТИЕ СЛУЧАЙНОЙ ВЕЛИЧИНЫ.

ДИСКРЕТНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ.

ЗАКОН РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ.

ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Некоторые переменные величины могут принимать в результате испытания определённые значения из возможного множества своих значений в зависимости от случая. Например, число выпавших на игральной кости очков, число бракованных изделий в партии, продолжительность работы прибора до первой поломки и т.д. Приведённая в главе 4 трактовка основных понятий теории вероятностей позволяет дать следующее определение.

Случайной величиной X называется функция, заданная в пространстве элементарных событий, т.е. $X = f(\omega)$, $\omega \in \Omega$.

Например, если случайная величина X – количество выпавших на игральной кости очков, то пространство элементарных событий $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, где ω_i – событие, состоящее в выпадении i очков. Тогда $X(\omega_i) = i$ ($i = 1, \dots, 6$).

Случайная величина называется *дискретной*, если множество её значений конечно или счётно.

Рассмотренная выше случайная величина является дискретной, так как множество её значений $\{1, 2, 3, 4, 5, 6\}$ является конечным. Приведём пример дискретной случайной величины с бесконечным множеством значений. Пусть X – число выстрелов из артиллерийского орудия до первого попадания в цель. Пространство элементарных событий $\Omega = \{w_1, w_2, \dots, w_i, \dots\}$, где w_i состоит в том, что на поражение цели потребуется i выстрелов. Тогда $X(\omega_i) = i$ ($i \in \mathbb{N}$), т.е. множество возможных значений данной случайной величины является натуральным рядом чисел, и, следовательно, счётно.

Пусть случайная величина X – интервал времени между приходом на остановку автобусов. Множество возможных значений этой случайной величины – полупрямая $[0; +\infty)$, следовательно, данная случайная величина не является дискретной.

Следует понимать, что для задания случайной величины недостаточно знать только множество её возможных значений. Случайные величины могут иметь одинаковое множество значений, но принимать эти значения с различными вероятностями. Поэтому для полного описания случайной величины требуется задать её закон распределения.

Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями.

Простейшей формой задания закона распределения дискретной случайной величины является таблица, в которой перечислены в порядке возрастания все возможные значения случайной величины X и соответствующие им вероятности.

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

Такая таблица (она может содержать и бесконечное число столбцов) называется *рядом распределения* дискретной случайной величины. Так как в таблице перечислены все возможные значения случайной величины, и события $X = x_1$, $X = x_2$,

..., $X = x_n$ являются несовместными, то сумма их вероятностей равна 1: $\sum_{i=1}^n p_i = 1$. Если множество возможных значений

дискретной случайной величины X счётно, то ряд $\sum_{i=1}^{\infty} p_i$ сходится и его сумма равна 1.

Использование для описания случайной величины ряда распределения не является единственно возможным, а главное, не является универсальным. Если множество значений случайной величины представляет собой отрезок или прямую, то нельзя перечислить всё бесконечное несчётное множество её значений. Поэтому для описания закона распределения можно применить и другой подход: рассматривать не вероятности события $X = x$, как это имеет место в ряде распределения, а вероятности события $X < x$, где x – некоторое действительное число. Очевидно, что эта вероятность $P(X < x)$ зависит от x , т.е. является функцией от x . Обычно её обозначают $F(x)$.

Функцией распределения называют функцию $F(x)$, определяющую для каждого x вероятность того, что в результате испытания случайная величина X примет значение меньшее, чем x :

$$F(x) = P(X < x). \quad (5.1.1)$$

Функцию $F(x)$ также называют интегральной функцией распределения. Геометрически она интерпретируется как вероятность того, что случайная величина примет значение, лежащее на числовой прямой левее заданной точки x .

Пример 1. Дан ряд распределения дискретной случайной величины X :

X	1	3	5	7
P	0,2	0,5	0,2	0,1

Построим её функцию распределения. Для этого будем рассматривать различные значения x и находить для них $F(x) = P(X < x)$.

1. Если $x \leq 1$, то $F(x) = 0$.
2. Если $1 < x \leq 3$, то $F(x) = P(X < x) = P(X = 1) = 0,2$.
3. Если $3 < x \leq 5$, то $F(x) = P(X < x) = P(X = 1) + P(X = 3) = 0,2 + 0,5 = 0,7$.
4. Если $5 < x \leq 7$, то $F(x) = P(X < x) = P(X = 1) + P(X = 3) + P(X = 5) = 0,2 + 0,5 + 0,2 = 0,9$.
5. Если $x > 7$, то $F(x) = P(X < x) = P(X = 1) + P(X = 3) + P(X = 5) + P(X = 7) = 0,2 + 0,5 + 0,2 + 0,1 = 1$.

Итак, аналитически функция распределения может быть записана в виде

$$F(x) = \begin{cases} 0 & \text{при } x \leq 1; \\ 0,2 & \text{при } 1 < x \leq 3; \\ 0,7 & \text{при } 3 < x \leq 5; \\ 0,9 & \text{при } 5 < x \leq 7; \\ 1 & \text{при } x > 7. \end{cases}$$

Изобразим функцию $F(x)$ (рис. 5.1.1) графически. Заметим, что эта функция является непрерывной слева (сохраняет свои значения, если x стремится к точке разрыва слева).

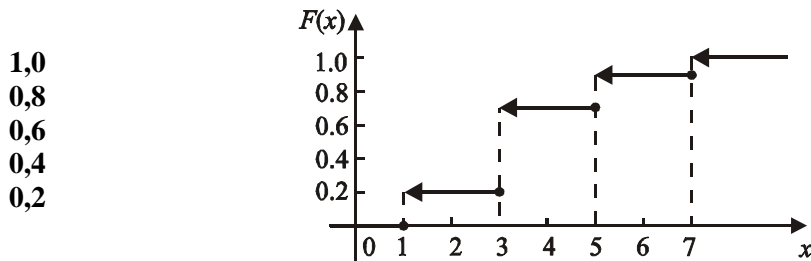


Рис. 5.1.1

Рассмотренный пример позволяет сформулировать свойство: функция распределения дискретной случайной величины является разрывной ступенчатой функцией, имеющей скачки в точках, соответствующих возможным значениям случайной величины. Значения скачков равны вероятностям этих значений, а сумма всех скачков равна 1. Рассмотрим общие свойства функции распределения.

1. Значения функции распределения принадлежат отрезку $[0, 1]$:

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

▷ Утверждение следует из определения функции распределения как вероятности. ◁

2. Функция распределения является неубывающей, т.е.

$$F(x_2) \geq F(x_1) \text{ при } x_2 > x_1.$$

▷ Пусть $x_2 > x_1$. Событие, состоящее в том, что X примет значение, меньшее чем x_2 , можно представить в виде суммы двух несовместных событий: $(X < x_1)$ и $(x_1 \leq X < x_2)$. По аксиоме аддитивности имеем

$$P(X < x_2) = P(X < x_1) + P(x_1 \leq X < x_2)$$

или

$$F(x_2) = F(x_1) + P(x_1 \leq X < x_2). \quad (5.1.2)$$

Так как $P(x_1 \leq X < x_2) \geq 0$, то $F(x_2) \geq F(x_1)$. ◁

3. Вероятность того, что случайная величина примет значение, заключённое в интервале $[x_1, x_2)$, равна приращению функции распределения на этом интервале:

$$P(x_1 \leq X < x_2) = F(x_2) - F(x_1).$$

▷ Утверждение следует из формулы (5.1.2). ◁

4. Для рассматриваемых в данном пособии дискретных и непрерывных случайных величин имеет место соотношение

$$F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0, \quad F(\infty) = \lim_{x \rightarrow \infty} F(x) = 1.$$

▷ $F(-\infty) = P(X < -\infty) = 0$ как вероятность невозможного события $X < -\infty$, а $F(\infty) = P(X < \infty) = 1$ как вероятность достоверного события $X < \infty$. ◁

5. Если все возможные значения случайной величины X принадлежат отрезку $[a, b]$, то: 1) $F(x) = 0$ при $x \leq a$; 2) $F(x) = 1$ при $x > b$.

▷ 1) Пусть $x \leq a$. Тогда $F(x) = P(X < x) = 0$ как вероятность невозможного события.

2) Пусть $x > b$. Тогда $F(x) = P(X < x) = 1$ как вероятность достоверного события. ◁

5.2. НЕПРЕРЫВНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ. ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ НЕПРЕРЫВНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Рассмотрим более подробно случайные величины, множеством значений которых является либо прямая $(-\infty, \infty)$, либо луч $(-\infty, a)$, либо луч (a, ∞) , либо интервал (a, b) (интервалы и лучи могут также быть замкнутыми). В качестве примеров

можно привести случайную величину X – продолжительность работы прибора до первой поломки, случайную величину X – расстояние, которое пролетает артиллерийский снаряд и т.д.

Случайная величина, имеющая бесконечное несчётное множество значений называется *непрерывной*, если её функция распределения непрерывна и кусочно-дифференцируема.

Учитывая, что для непрерывных случайных величин имеют место все свойства 1 – 5 из параграфа 5.1, можно представить, как выглядит график функции распределения непрерывных случайных величин (рис. 5.2.1). График расположен в полосе, ограниченной прямыми $y = 0$ и $y = 1$, и является гладким, за исключением некоторых точек излома.

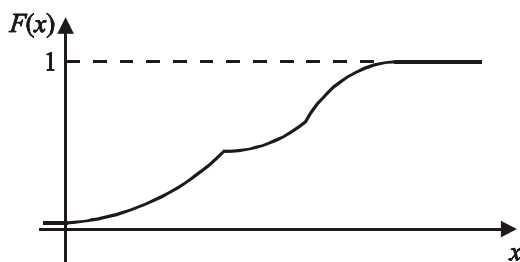


Рис. 5.2.1

Теорема 1. Вероятность того, что непрерывная случайная величина примет одно определенное значение, равна нулю.

▷ Воспользуемся формулой (5.1.2), положив в ней $x_2 = x_1 + \Delta x$, получаем

$$P(x_1 \leq X < x_1 + \Delta x) = F(x_1 + \Delta x) - F(x_1).$$

Устремляя Δx к нулю и учитывая непрерывность функции $F(x)$, получим

$$\begin{aligned} P(X = x_1) &= \lim_{\Delta x \rightarrow 0} P(x_1 \leq X < x_1 + \Delta x) = \\ &= \lim_{\Delta x \rightarrow 0} (F(x_1 + \Delta x) - F(x_1)) = F(x_1) - F(x_1) = 0. \triangleleft \end{aligned}$$

Таким образом, для непрерывных случайных величин не имеет смысла говорить о вероятности того, что случайная величина примет одно отдельно взятое значение – вероятность этого события равна нулю. Для непрерывных случайных величин следует рассматривать вероятность попадания случайной величины в какой-либо интервал, даже сколь угодно малый. Например, ищут вероятность того, что прочность материала не выходит за заданные границы, но не ставят вопрос о вероятности того, что прочность равна некоторому определённом значению.

Следствие 1. Если X – непрерывная случайная величина, то вероятность её попадания в интервал $(x_1; x_2)$ не зависит от того, является ли этот интервал открытым или закрытым, т.е.

$$P(x_1 \leq X \leq x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2).$$

$$\begin{aligned} \triangleright P(x_1 \leq X \leq x_2) &= P(x_1 < X \leq x_2) + P(X = x_1) = \\ &= P(x_1 < X \leq x_2) + 0 = P(x_1 < X \leq x_2). \end{aligned}$$

Аналогично доказываются и другие равенства. ◁

Задание непрерывной случайной величины с помощью функции распределения не является единственным. Её можно также задать, используя другую функцию, которую называют плотностью распределения, или дифференциальной функцией распределения.

Плотностью распределения вероятности (или плотностью вероятности) непрерывной случайной величины X называется производная от её функции распределения:

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

График плотности распределения называется *кривой распределения*. Плотность распределения вероятности $f(x)$ является одной из форм закона распределения, но для описания дискретных случайных величин она неприменима. Из определения следует:

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{P(x < X < x + \Delta x)}{\Delta x},$$

т.е. плотность вероятности можно определить как предел вероятности попадания случайной величины в интервал $(x; x + \Delta x)$, делённой на длину этого интервала (т.е. вероятность, приходящуюся на единицу длины), при стремлении длины интервала к нулю. Приведём основные *свойства плотности распределения вероятности*.

1. Плотность вероятности – неотрицательная функция, т.е.

$$f(x) \geq 0.$$

▷ Функция распределения – неубывающая функция, следовательно, её производная неотрицательна. ◁

2. Вероятность попадания непрерывной случайной величины в интервал $(x_1; x_2)$ равна определённому интегралу от её плотности в пределах от x_1 до x_2 :

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x) dx. \quad (5.2.1)$$

▷ По свойству 3 функции распределения $P(x_1 \leq X < x_2) = F(x_2) - F(x_1)$. Так как $F(x)$ является первообразной для плотности вероятности $f(x)$, то по формуле Ньютона-Лейбница $F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx$. Остаётся отметить, что по следствию из теоремы 1 вероятность попадания непрерывной случайной величины в интервал не зависит от того, замкнутый это интервал или открытый. Следовательно, формула (5.2.1) доказана. ◁

3. Функция распределения непрерывной случайной величины может быть выражена через плотность распределения по формуле

$$F(x) = \int_{-\infty}^x f(t) dt. \quad (5.2.2)$$

▷ Полагая в формуле (5.2.1) $a = -\infty$ и $b = x$, получим

$$P(-\infty < X < x) = \int_{-\infty}^x f(t) dt.$$

Учитывая, что $P(-\infty < X < x) = P(X < x) = F(x)$, получим формулу (5.2.2). ◁

4. Несобственный интеграл в бесконечных пределах от плотности распределения вероятности равен единице:

$$\int_{-\infty}^{+\infty} f(t) dt = 1. \quad (5.2.3)$$

▷ Из формулы (5.2.2) при $x \rightarrow +\infty$ получим

$$F(\infty) = \int_{-\infty}^{+\infty} f(t) dt.$$

Учитывая, что $F(\infty) = 1$, получим формулу (5.2.3). ◁

Геометрически последнее свойство означает, что площадь криволинейной трапеции, ограниченной осью Ox и кривой распределения, равна единице.

5.3. МАТЕМАТИЧЕСКИЕ ОПЕРАЦИИ НАД СЛУЧАЙНЫМИ ВЕЛИЧИНАМИ

Используя введённое в параграфе 4.9 понятие независимости случайных событий, введём понятие независимости случайных величин.

Две случайные величины называются *независимыми*, если закон распределения одной из них не зависит от того, какие возможные значения приняла другая величина. Несколько случайных величин называются *взаимно независимыми*, если законы распределения любого числа из них не зависят от того, какие возможные значения приняли остальные величины.

Так, если дискретная случайная величина X может принимать значения $x_i (i = 1, \dots, n)$, а дискретная случайная величина Y – значения $y_j (j = 1, \dots, m)$, то независимость случайных величин X и Y означает, что независимыми являются события $(X = x_i)$ и $(Y = y_j)$ для любых $i = 1, \dots, n$ и $j = 1, \dots, m$. В общем случае вводят *совместную функцию распределения* $F(x, y)$, определяемую равенством

$$F(x, y) = P((X < x) \cdot (Y < y)). \quad (5.3.1)$$

Геометрически функция распределения $F(x, y)$ означает вероятность попадания случайной точки (X, Y) в бесконечный квадрант, лежащий левее и ниже точки $M(x, y)$ (рис. 5.3.1). Если случайные величины X и Y являются непрерывными, то вводят *совместную плотность вероятности* случайных величин:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = F''_{xy}(x, y). \quad (5.3.2)$$

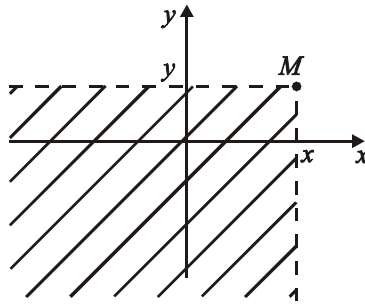


Рис. 5.3.1

Если случайные величины X и Y независимы, то согласно определению, события $(X < x)$ и $(Y < y)$ являются независимыми, следовательно, их совместная функция распределения равна произведению функций распределения этих случайных величин: $F(x, y) = P((X < x) \cdot (Y < y)) = P(X < x) \cdot P(Y < y) = F_1(x)F_2(y)$. Если равенство $F(x, y) = F_1(x)F_2(y)$ не выполняется, то случайные величины являются зависимыми.

Определим математические операции над случайными величинами. Ограничимся здесь случаем дискретных случайных величин. Пусть дискретная случайная величина X задана законом распределения:

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

Функцией $f(X)$ от случайной величины X называется случайная величина, которая принимает значения $f(x_i)$ с вероятностями p_i ($i = 1, \dots, n$). В частности, квадратом случайной величины X^2 является случайная величина, которая принимает значения $(x_i)^2$ с вероятностями p_i .

Пусть также имеется случайная величина Y , заданная законом распределения,

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

Суммой (разностью или произведением) случайных величин X и Y называется случайная величина, которая принимает все возможные значения вида $x_i + y_j$ ($x_i - y_j$ или $x_i \cdot y_j$) с вероятностями $p_{ij} = P((X = x_i) \cdot (Y = y_j))$, где $i = 1, \dots, n$; $j = 1, \dots, m$. Если случайные величины X и Y независимы, то $p_{ij} = P(X = x_i) \cdot P(Y = y_j) = p_i q_j$.

Приведённые выше определения нуждаются в уточнении, так как в ряде случаев одни и те же значения $f(x_i)$ могут быть получены при разных x_i , а одинаковые значения $x_i \pm y_j$ и $x_i \cdot y_j$ – при разных x_i и y_j .

Пример 2. Даны ряды распределения двух независимых случайных величин:

X	0	2	4
p	0,3	0,5	0,2
Y	-1	0	1
P	0,4	0,3	0,3

Найдём закон распределения случайной величины $Z = XY$. Для этого составим вспомогательную таблицу, в каждой клетке которой поместим значения случайной величины Z , а в скобках – вероятности этих значений.

y_j	-1 (0,4)	0 (0,3)	1 (0,3)
x_i			
0 (0,3)	0 (0,12)	0 (0,09)	0 (0,09)
2 (0,5)	-2 (0,20)	0 (0,15)	2 (0,15)
4 (0,2)	-4 (0,08)	0 (0,06)	4 (0,06)

Так как среди значений Z имеются одинаковые (нули встречаются в 6 клетках таблицы), то соответствующие вероятности складываем по теореме сложения вероятностей; поэтому $P(Z = 0) = 0,12 + 0,09 + 0,09 + 0,15 + 0,06 = 0,51$. В результате получим распределение:

Z	-4	-2	0	2	4
P	0,08	0,20	0,51	0,15	0,06

5.4. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ И ДИСПЕРСИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Закон распределения даёт исчерпывающее представление о случайной величине, так как позволяет вычислять вероятности любых событий, связанных со случайной величиной. Однако закон распределения не всегда бывает удобным для анализа, часто удобнее пользоваться числами, которые описывают случайную величину суммарно. Такие числа называются *числовыми характеристиками* случайной величины. Они помогают в сжатой форме выразить наиболее существенные черты распределения. Основными числовыми характеристиками случайной величины являются *математическое ожидание*, характеризующее среднее значение случайной величины, и *дисперсия*, характеризующая рассеяние случайной величины.

Математическим ожиданием $M(X)$ дискретной случайной величины, множество значений которой конечно, называется сумма произведений всех её значений на соответствующие им вероятности:

$$M(X) = \sum_{i=1}^n x_i p_i . \quad (5.4.1)$$

Математическим ожиданием $M(X)$ дискретной случайной величины, множество значений которой счётно, называется сумма ряда (если он абсолютно сходится):

$$M(X) = \sum_{i=1}^{\infty} x_i p_i . \quad (5.4.2)$$

Если ряд (5.4.2) расходится, то соответствующая случайная величина не будет иметь математического ожидания. Из определения следует, что математическое ожидание является неслучайной величиной (константой). С наибольшей полнотой смысл математического ожидания раскрывается в законе больших чисел. Можно заметить, что если случайная величина принимает значения из некоторого отрезка $[a, b]$, то и её математическое ожидание не может выйти за пределы этого отрезка.

Математическим ожиданием $M(X)$ непрерывной случайной величины с плотностью распределения $f(x)$ называется несобственный интеграл (если он абсолютно сходится):

$$M(X) = \int_{-\infty}^{+\infty} x f(x) dx . \quad (5.4.3)$$

Если непрерывная случайная величина может принимать значения только из отрезка $[a, b]$, то вне его плотность случайной величины равна нулю, поэтому вместо формулы (5.4.3) можно использовать

$$M(X) = \int_a^b x f(x) dx . \quad (5.4.4)$$

Приведём основные *свойства* математического ожидания.

1. Математическое ожидание постоянной величины равно самой этой величине:

$$M(C) = C . \quad (5.4.5)$$

▷ Постоянную величину C можно рассматривать как дискретную случайную величину, принимающую значение C с вероятностью 1. Поэтому её математическое ожидание $M(C) = C \cdot 1 = C$. ◁

2. Постоянный множитель можно выносить за знак математического ожидания:

$$M(CX) = CM(X) . \quad (5.4.6)$$

▷ Проведём доказательство для случая, когда случайная величина X является дискретной с конечным множеством значений. Так как случайная величина CX принимает значения Cx_i с вероятностями p_i , то

$$M(CX) = \sum_{i=1}^n (Cx_i) p_i = C \sum_{i=1}^n x_i p_i = CM(X) . \quad \triangleleft$$

3. Математическое ожидание суммы случайных величин равно сумме их математических ожиданий:

$$M(X_1 + \dots + X_m) = M(X_1) + \dots + M(X_m) . \quad (5.4.7)$$

4. Математическое ожидание произведения взаимно независимых случайных величин равно произведению их математических ожиданий:

$$M(X_1 \cdot \dots \cdot X_m) = M(X_1) \cdot \dots \cdot M(X_m) . \quad (5.4.8)$$

Свойства 3 и 4 приводим без доказательства.

Только математическое ожидание не может полностью характеризовать случайную величину. В большинстве практически важных случаях требуется знать не только средние показатели, но и отклонение (разброс, рассеяние) показателя от среднего значения.

Отклонением называется разность между случайной величиной и её математическим ожиданием $X - M(X)$. Приведём важное свойство отклонения.

Теорема 2. Математическое ожидание отклонения равно нулю:

$$M[X - M(X)] = 0.$$

▷ Воспользуемся свойствами 1 и 3 математического ожидания (заметим, что $-M(X)$ – это постоянная величина): $M[X - M(X)] = M(X) + M[-M(X)] = M(X) - M(X) = 0$. ◁

Доказанная выше теорема поясняет, что в качестве характеристики рассеяния использовать отклонение нельзя – его среднее значение равно нулю, так как одни возможные значения отклонения положительны, а другие отрицательны. Поэтому в качестве меры рассеяния используют среднее значение квадрата отклонения.

Дисперсией $D(X)$ случайной величины называется математическое ожидание квадрата отклонения:

$$D(X) = M[X - M(X)]^2. \quad (5.4.9)$$

Обозначим $a = M(X)$. Тогда формулу (5.4.9) можно переписать в виде $D(X) = M(X - a)^2$ и, в зависимости от вида случайной величины, можно предложить следующие формулы для вычисления дисперсии:

- если X – дискретная случайная величина с конечным множеством значений, то $D(X) = \sum_{i=1}^n (x_i - a)^2 p_i$;
- если X – дискретная случайная величина со счётным множеством значений, то $D(X) = \sum_{i=1}^{\infty} (x_i - a)^2 p_i$, если ряд в правой части равенства сходится;
- если X – непрерывная случайная величина, то $D(X) = \int_{-\infty}^{+\infty} (x - a)^2 f(x) dx$, если интеграл в правой части равенства сходится.

Дисперсия $D(X)$ имеет размерность квадрата случайной величины, что не всегда удобно. Поэтому в тех случаях, когда желательно, чтобы оценка рассеяния имела размерность случайной величины, используют величину $\sqrt{D(X)}$.

Среднеквадратическим отклонением $\sigma(X)$ (стандартным отклонением) случайной величины называется квадратный корень из дисперсии:

$$\sigma(X) = \sqrt{D(X)}. \quad (5.4.10)$$

Приведём основные *свойства* дисперсии случайной величины:

1. Дисперсия постоянной величины равна нулю:

$$D(C) = 0. \quad (5.4.11)$$

▷ По свойству (5.4.5) математического ожидания $M(C) = C$. Следовательно, из определения дисперсии имеем: $D(C) = M[C - M(C)]^2 = M[C - C]^2 = M(0) = 0$. ◁

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

$$D(CX) = C^2 D(X). \quad (5.4.12)$$

▷ По свойству (5.4.6) математического ожидания множитель C можно выносить за знак математического ожидания. Поэтому, используя определение дисперсии, получим:

$$\begin{aligned} D(CX) &= M[CX - M(CX)]^2 = \\ &= M[CX - CM(X)]^2 = M[C^2(X - M(X))]^2 = \\ &= C^2 M[X - M(X)]^2 = C^2 D(X). \quad \triangleleft \end{aligned}$$

3. Дисперсия случайной величины равна разности между математическим ожиданием квадрата случайной величины и квадратом её математического ожидания:

$$D(X) = M(X^2) - [M(X)]^2. \quad (5.4.13)$$

▷ Обозначим $a = M(X)$. Тогда $D(X) = M[X - a]^2 = M[X^2 - 2aX + a^2]$. Учитывая, что a – постоянная величина, а также используя свойство (5.4.7) математического ожидания, получим:

$$D(X) = M(X^2) - 2aM(X) + M(a^2) = M(X^2) - 2a \cdot a + a^2 = M(X^2) - a^2. \quad \triangleleft$$

Дисперсия суммы или разности взаимно независимых случайных величин равна сумме их дисперсий:

$$D(X_1 \pm \dots \pm X_m) = D(X_1) + \dots + D(X_m). \quad (5.4.14)$$

▷ Доказательство проведём для случая двух случайных величин, т.е. докажем равенство $D(X_1 \pm X_2) = D(X_1) + D(X_2)$. Обозначим $a_1 = M(X_1)$ и $a_2 = M(X_2)$. Тогда из (5.4.13) получим

$$\begin{aligned} D(X_1 \pm X_2) &= M(X_1 \pm X_2)^2 - [M(X_1 \pm X_2)]^2 = \\ &= M(X_1^2 \pm 2X_1X_2 + X_2^2) - (a_1 \pm a_2)^2. \end{aligned}$$

Далее используем свойства 3 и 4 математического ожидания:

$$\begin{aligned} D(X_1 \pm X_2) &= M(X_1^2) \pm 2M(X_1)M(X_2) + M(X_2^2) - a_1^2 \mp 2a_1a_2 - a_2^2 = \\ &= (M(X_1^2) - a_1^2) + (M(X_2^2) - a_2^2) = D(X_1) + D(X_2). \quad \triangleleft \end{aligned}$$

Кроме математического ожидания и дисперсии, для анализа и описания случайных величин используются и другие числовые характеристики (мода, медиана, асимметрия, эксцесс и др.), определения которых мы здесь не будем приводить. Дадим только определения моментов – начальных и центральных.

Начальным моментом k -го порядка случайной величины X называется математическое ожидание k -й степени этой величины:

$$\nu_k = M(X^k). \quad (5.4.15)$$

Центральным моментом k -го порядка случайной величины X называется математическое ожидание k -й степени отклонения этой величины:

$$\mu_k = M[X - M(X)]^k. \quad (5.4.16)$$

Нетрудно заметить, что при $k=1$ начальный момент (5.4.15) равен математическому ожиданию: $\nu_1 = M(X)$, а центральный момент равен нулю: $\mu_1 = 0$. При $k=2$ центральный момент (5.4.16) равен дисперсии: $\mu_2 = D(X)$. Моменты высших порядков служат для более детального описания случайной величины.

В следующих двух параграфах приведены основные законы распределения дискретных (параграф 5.5) и непрерывных (параграф 5.6) случайных величин, которые используются для построения вероятностных моделей реальных технических процессов, а также природных и социальных явлений.

5.5. ОСНОВНЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ ДИСКРЕТНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

5.5.1. Биномиальный закон распределения

Пусть производится n независимых испытаний, в каждом из которых событие A может произойти с вероятностью p или не произойти с вероятностью $q = 1 - p$. Рассмотрим случайную величину X – количество появлений события A в этих испытаниях. Очевидно, что множество значений случайной величины X – множество $\{0, 1, \dots, n\}$, причём вероятности событий $P(X = m)$, где $m = \overline{0, n}$, могут быть найдены по формуле Бернулли (4.11.2):

$$p_m = P(X = m) = C_n^m p^m q^{n-m} \quad (0 < p < 1; q = 1 - p; m = \overline{0, n}). \quad (5.5.1)$$

Определение 1. Закон распределения дискретной случайной величины X , заданный формулой (5.5.1), называется биномиальным законом распределения. Запишем биномиальный закон в виде таблицы:

X	0	1	...	m	...	n
p	$C_n^0 q^n$	$C_n^1 p^1 q^{n-1}$...	$C_n^m p^m q^{n-m}$...	$C_n^n p^n$

Закон назван биномиальным, так как $\sum_{i=0}^n p_i$ представляет собой сумму всех членов разложения бинома Ньютона:

$$C_n^0 q^n + C_n^1 p^1 q^{n-1} + C_n^2 p^2 q^{n-2} + \dots + C_n^m p^m q^{n-m} + \dots + C_n^n p^n = (p + q)^n = 1^n = 1.$$

Таким образом, мы доказали, что выполнено основное свойство ряда распределения $\sum_{i=0}^n p_i = 1$.

Теорема 3. Математическое ожидание случайной величины X , распределённой по биномиальному закону, равно произведению числа испытаний на вероятность появления события в одном испытании:

$$M(X) = np, \quad (5.5.2)$$

а её дисперсия равна произведению числа испытаний на вероятности появления и не появления события в одном испытании:

$$D(X) = npq. \quad (5.5.3)$$

▷ Случайную величину X – количество появлений события A в n испытаниях – можно представить в виде суммы n независимых случайных величин:

$$X = \sum_{i=1}^n X_i,$$

где случайная величина X_i – число появлений события A в i -м испытании ($i = \overline{1, n}$). Все случайные величины X_i имеют один и тот же закон распределения:

X_i	0	1
p_i	q	p

Следовательно, для всех $i = \overline{1, n}$ выполнено:

$$M(X_i) = 0 \cdot q + 1 \cdot p = p, \quad M(X_i^2) = 0^2 \cdot q + 1^2 \cdot p = p, \quad D(X_i) = M(X_i^2) - [M(X_i)]^2 = p^2 - p = p(1-p) = pq.$$

Тогда, в силу свойства 3 математического ожидания имеем: $M(X) = M\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n M(X_i) = \sum_{i=1}^n p = np$; а в силу свойства 4 дисперсии имеем: $D(X) = D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n D(X_i) = \sum_{i=1}^n pq = npq$. ◁

Введём случайную величину $Z_A = \frac{X}{n}$, имеющую смысл относительной частоты $h_n(A) = \frac{m}{n}$ события A .

Следствие. Математическое ожидание относительной частоты Z_A события A в n независимых испытаниях, в каждом из которых оно может наступить с вероятностью p , равно p , т.е.

$$M(Z_A) = p, \tag{5.5.4}$$

а её дисперсия

$$D(Z_A) = \frac{pq}{n}. \tag{5.5.5}$$

▷ Относительная частота $Z_A = \frac{m}{n} = \frac{X}{n}$, где случайная величина X распределена по биномиальному закону. Тогда

$$M(Z_A) = \frac{M(X)}{n} = \frac{np}{n} = p;$$

$$D(Z_A) = \frac{D(X)}{n^2} = \frac{npq}{n^2} = \frac{pq}{n}. \triangleleft$$

Биномиальный закон распределения широко используется при статистическом контроле качества продукции, при описании функционирования систем массового обслуживания, в теории стрельбы и т.д.

5.5.2. Закон распределения Пуассона

В параграфе 4.12 была доказана теорема Пуассона, в которой утверждается, что при достаточно больших n , при малых значениях p и при условии, что $\lambda = np$ – постоянная величина, хорошим приближением формулы Бернулли является формула Пуассона (4.12.1).

Определение 2. Дискретная случайная величина X имеет закон распределения Пуассона с параметром λ , если

$$p_m = P(X = m) = \frac{\lambda^m e^{-\lambda}}{m!} \quad (m = 0, 1, 2, \dots). \tag{5.5.6}$$

Как видим, закон распределения Пуассона является предельным случаем (при $n \rightarrow \infty$, $p \rightarrow 0$, $np \rightarrow \lambda$) биномиального закона. Так как вероятность появления события A в каждом испытании мала, то закон Пуассона называют законом редких явлений. По закону Пуассона распределены, например, число сбоев на автоматической линии, число бракованных изделий в большой партии товара и т.д. Однако распределение Пуассона применяют и в ряде других ситуаций. Так, например, число заявок на обслуживание, поступивших в единицу времени в системах массового обслуживания, имеет распределение Пуассона (где λ – интенсивность потока требований, т.е. среднее число заявок в единицу времени).

Приведём ряд распределения закона Пуассона:

X	0	1	2	...	m	...
P	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2 e^{-\lambda}}{2}$...	$\frac{\lambda^m e^{-\lambda}}{m!}$...

Докажем корректность определения закона Пуассона. Для этого вычислим сумму ряда:

$$\sum_{i=0}^{\infty} p_i = \sum_{i=0}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1$$

(использовали разложение функции e^x в ряд Маклорена).

Теорема 4. Математическое ожидание и дисперсия случайной величины, распределённой по закону Пуассона, совпадают и равны параметру λ этого распределения:

$$M(X) = \lambda, \quad (5.5.7)$$

$$D(X) = \lambda. \quad (5.5.8)$$

▷ Найдём математическое ожидание случайной величины X :

$$\begin{aligned} M(X) &= \sum_{i=0}^{\infty} x_i p_i = \sum_{i=0}^{\infty} i \frac{\lambda^i e^{-\lambda}}{i!} = \sum_{i=1}^{\infty} \frac{\lambda^i e^{-\lambda}}{(i-1)!} = \\ &= \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = \lambda e^{-\lambda} e^{\lambda} = \lambda. \end{aligned}$$

Для нахождения дисперсии сначала найдём математическое ожидание квадрата случайной величины X :

$$\begin{aligned} M(X^2) &= \sum_{i=0}^{\infty} x_i^2 p_i = \sum_{i=0}^{\infty} i^2 \frac{\lambda^i e^{-\lambda}}{i!} = \sum_{i=1}^{\infty} i \frac{\lambda^i e^{-\lambda}}{(i-1)!} = \\ &= e^{-\lambda} \sum_{i=1}^{\infty} \frac{((i-1)+1)\lambda^i}{(i-1)!} = e^{-\lambda} \lambda^2 \sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!} + \\ &+ e^{-\lambda} \lambda \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = e^{-\lambda} \lambda^2 e^{\lambda} + e^{-\lambda} \lambda e^{\lambda} = \lambda^2 + \lambda. \end{aligned}$$

Далее находим дисперсию

$$D(X) = M(X^2) - [M(X)]^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda. \triangleleft$$

5.5.3. Геометрический закон распределения

Пусть случайная величина X представляет собой число испытаний, проведённых по схеме Бернулли, до первого появления события A . В параграфе 4.7 была получена формула для нахождения вероятности p_m события ω_m , означающего, что при первых $m-1$ испытаний событие A не произошло, а в m -м произошло:

$$p_m = p(X = m) = q^{m-1} p, \quad (0 < p < 1; q = 1 - p; m = 1, 2, \dots) \quad (5.5.9)$$

Определение 3. Дискретная случайная величина X имеет геометрическое распределение, если её закон распределения задаётся формулой (5.5.9).

Ряд геометрического распределения имеет вид:

X	1	2	3	...	m	...
p	p	pq	pq^2	...	pq^{m-1}	...

Название распределения объясняется тем, что вероятности p_m образуют бесконечную геометрическую прогрессию с первым членом p и знаменателем q . Определение геометрического распределения корректно, так как сумма ряда

$$\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} pq^{i-1} = p \sum_{j=0}^{\infty} q^j = p \frac{1}{1-q} = \frac{p}{p} = 1 \quad (\text{использовали формулу для суммы бесконечной геометрической прогрессии}).$$

Теорема 5. Математическое ожидание случайной величины X , имеющей геометрическое распределение, равно величине, обратной появлению события в одном испытании:

$$M(X) = \frac{1}{p}, \quad (5.5.10)$$

а её дисперсия

$$D(X) = \frac{q}{p^2}. \quad (5.5.11)$$

▷ Найдём математическое ожидание случайной величины X :

$$M(X) = \sum_{i=1}^{\infty} x_i p_i = \sum_{i=1}^{\infty} i p q^{i-1} = p \sum_{i=1}^{\infty} \frac{dq^i}{dq} = p \frac{d}{dq} \left(\sum_{i=1}^{\infty} q^i \right) =$$

$$= p \frac{d}{dq} \left(\frac{q}{1-q} \right) = p \frac{(1-q) + q}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

(использовали свойство, что равномерно сходящиеся ряды можно почленно дифференцировать).

Для нахождения дисперсии сначала найдём математическое ожидание квадрата случайной величины X :

$$\begin{aligned} M(X^2) &= \sum_{i=1}^{\infty} x_i^2 p_i = \sum_{i=1}^{\infty} i^2 p q^{i-1} = \sum_{i=1}^{\infty} ((i^2 - i) + i) p q^{i-1} = \\ &= \sum_{i=1}^{\infty} i(i-1) p q^{i-1} + \sum_{i=1}^{\infty} i p q^{i-1} \end{aligned}$$

Вторую сумму мы вычисляли при нахождении математического ожидания и получили, что она равна $\frac{1}{p}$. Вычислим первую сумму:

$$\begin{aligned} \sum_{i=1}^{\infty} i(i-1) p q^{i-1} &= p q \sum_{i=2}^{\infty} i(i-1) q^{i-2} = p q \sum_{i=2}^{\infty} \frac{d^2 q^i}{dq^2} = \\ p q \frac{d^2}{dq^2} \sum_{i=2}^{\infty} (q^i) &= p q \frac{d^2}{dq^2} \left(\frac{q^2}{1-q} \right) = p q \frac{d}{dq} \left(\frac{2q(1-q) + q^2}{(1-q)^2} \right) = p q \frac{2}{(1-q)^3} = \frac{2q}{p^2}. \end{aligned}$$

$$\text{Тогда } M(X^2) = \frac{1}{p} + \frac{2q}{p^2} = \frac{p+2q}{p^2}.$$

Далее находим дисперсию:

$$D(X) = M(X^2) - [M(X)]^2 = \frac{p+2q}{p^2} - \frac{1}{p^2} = \frac{p+2q-1}{p^2} = \frac{2q-q}{p^2} = \frac{q}{p^2}. \triangleleft$$

5.5.4. Гипергеометрическое распределение

Рассмотрим ситуацию: имеется N объектов, из которых M обладают заданным свойством, а оставшиеся $N - M$ не обладают. Случайным образом без возврата извлекают n объектов ($n \leq N$). Пусть случайная величина X представляет собой число извлечённых объектов, которые обладают указанным свойством. Найдём вероятность события $X = m$. Общее число возможных элементарных исходов равно числу способов, которыми можно извлечь n объектов из N : C_N^n . Число элементарных исходов, благоприятствующих событию $X = m$, равно числу способов, сколькими мы извлекаем m «нужных» объектов из M имеющихся «нужных», умноженное на число способов, сколькими можно извлечь $n - m$ «ненужных» объектов из имеющихся в совокупности $N - M$ «ненужных», т.е. равно $C_M^m C_{N-M}^{n-m}$. Искомая вероятность равна отношению числа исходов, благоприятствующих событию $X = m$ к общему числу возможных исходов, т.е.

$$p_m = P(X = m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}. \quad (5.5.12)$$

Определение 4. Дискретная случайная величина X имеет гипергеометрическое распределение с параметрами N , M и n , если она принимает значения $0, 1, \dots, \min(n, M)$ с вероятностями, определяемыми формулой (5.5.12), где $M \leq N, n \leq N$; N, M и n – натуральные числа.

Приведём без доказательства теорему.

Теорема 6. Математическое ожидание случайной величины X , имеющей гипергеометрическое распределение с параметрами N , M и n , есть

$$M(X) = \frac{nM}{N}, \quad (5.5.13)$$

а её дисперсия

$$D(X) = \frac{nM}{N-1} \left(1 - \frac{M}{N} \right) \left(1 - \frac{n}{N} \right). \quad (5.5.14)$$

Заметим, что величина $p = \frac{M}{N}$ является вероятностью извлечения одного объекта, обладающего заданным свойством, из совокупности. Если N велико по сравнению с числом извлекаемых объектов n , то эта вероятность при последующем извлечении объектов будет меняться незначительно. Тогда вероятность, задаваемая формулой (5.5.12) будет близка к вероятности, вычисляемой по (5.5.1) для биномиального распределения. Поэтому биномиальное распределение можно рассматривать как модификацию гипергеометрического для случая бесконечной совокупности объектов.

5.6.1. Равномерный закон распределения

Определение 1. Непрерывная случайная величина X имеет равномерный закон распределения на отрезке $[a, b]$, если её плотность вероятности имеет вид:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{при } x \in [a, b], \\ 0 & \text{при } x \notin [a, b]. \end{cases} \quad (5.6.1)$$

Докажем, что приведённое определение является корректным, т.е. проверим выполнение свойства 4 плотности распределения (свойство 1, т.е. неотрицательность функции $f(x)$, очевидно). Имеем:

$$\int_{-\infty}^{+\infty} f(x) dx = \int_a^b \frac{1}{b-a} dx = \frac{1}{b-a} x \Big|_a^b = \frac{b-a}{b-a} = 1.$$

Найдём функцию распределения равномерно распределённой случайной величины. Для этого воспользуемся формулой (5.2.2). При $x \leq a$ функция распределения $F(x) = 0$. При $a < x \leq b$ получим:

$$F(x) = \int_{-\infty}^a 0 dt + \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}.$$

При $x > b$:

$$F(x) = \int_{-\infty}^a 0 dt + \int_a^b \frac{1}{b-a} dt + \int_b^x 0 dt = \frac{b-a}{b-a} = 1.$$

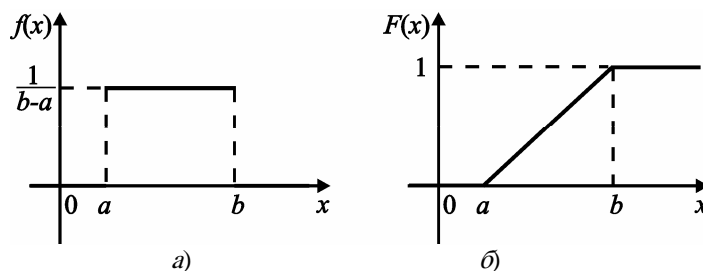


Рис. 5.6.1

На рис. 5.6.1 (а, б) приведены графики плотности распределения $f(x)$ и функции распределения $F(x)$. Очевидно, что если случайная величина распределена на отрезке $[a, b]$ равномерно, то её математическое ожидание должно быть равно абсциссе середины этого отрезка, дисперсия же должна зависеть от длины отрезка и быть пропорциональна квадрату этой длины. Выведем эти формулы, исходя из определения математического ожидания и дисперсии.

Теорема 7. Если случайная величина X распределена равномерно на отрезке $[a, b]$, то её математическое ожидание

$$M(X) = \frac{a+b}{2}, \quad (5.6.2)$$

а дисперсия

$$D(X) = \frac{(b-a)^2}{12}. \quad (5.6.3)$$

▷ Найдём математическое ожидание случайной величины X :

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

Найдём математическое ожидание квадрата случайной величины X :

$$M(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_a^b \frac{x^2}{b-a} dx = \frac{1}{b-a} \frac{x^3}{3} \Big|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}.$$

Далее находим дисперсию:

$$D(X) = M(X^2) - [M(X)]^2 = \frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4} =$$

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

Равномерный закон распределения используется при анализе ошибок округления при проведении числовых расчётов. Например, ошибка округления числа до целого распределена равномерно на отрезке $[-0,5; 0,5]$. Равномерное распределение на отрезке $[0; t_0]$ имеет время ожидания транспорта, если предположить, что пассажир приходит на остановку в случайный момент времени, а транспорт ходит регулярно с интервалом t_0 мин. При компьютерном моделировании случайных явлений используется так называемый «генератор случайных чисел», который генерирует значения случайной величины, распределённой равномерно на отрезке $[0; 1]$. Эти значения могут служить исходным материалом для получения значений случайных величин с любым законом распределения.

5.6.2. Показательный закон распределения

Определение 2. Непрерывная случайная величина X имеет показательный закон распределения с параметром $\lambda > 0$, если её плотность вероятности имеет вид:

$$f(x) = \begin{cases} 0 & \text{при } x < 0, \\ \lambda e^{-\lambda x} & \text{при } x \geq 0. \end{cases} \quad (5.6.4)$$

Докажем корректность данного определения:

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) dx &= \int_0^{+\infty} \lambda e^{-\lambda x} dx = \lambda \lim_{A \rightarrow \infty} \int_0^A e^{-\lambda x} dx = \lambda \lim_{A \rightarrow \infty} \left(-\frac{1}{\lambda} e^{-\lambda x} \Big|_0^A \right) = \\ &= \left(\lim_{A \rightarrow \infty} e^{-\lambda A} \right) - (-1) = 0 + 1 = 1. \end{aligned}$$

Найдём функцию распределения случайной величины X , распределённой по показательному закону. При $x \leq 0$ функция распределения $F(x) = 0$. При $x > 0$ получим:

$$F(x) = \int_{-\infty}^0 0 dt + \int_0^x \lambda e^{-\lambda t} dt = 0 - \frac{\lambda}{\lambda} e^{-\lambda t} \Big|_0^x = -(e^{-\lambda x} - e^0) = 1 - e^{-\lambda x}.$$

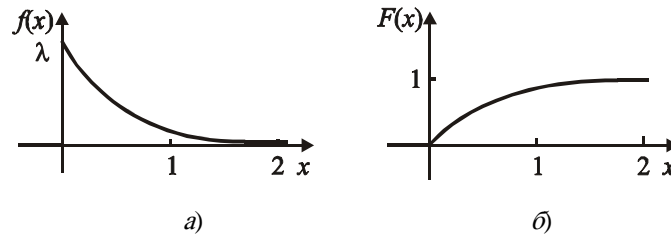


Рис. 5.6.2

На рис. 5.6.2 (а, б) приведены графики плотности распределения $f(x)$ и функции распределения $F(x)$. Найдём математическое ожидание и дисперсию показательного распределённой случайной величины.

Теорема 8. Если случайная величина X имеет показательный закон распределения с параметром λ , то её математическое ожидание

$$M(X) = \frac{1}{\lambda}, \quad (5.6.5)$$

а дисперсия

$$D(X) = \frac{1}{\lambda^2}. \quad (5.6.6)$$

▷ Найдём математическое ожидание случайной величины X , применяя метод интегрирования по частям:

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} x f(x) dx = \int_0^{+\infty} x \lambda e^{-\lambda x} dx = - \lim_{A \rightarrow \infty} \int_0^A x d e^{-\lambda x} = \\ &= - \lim_{A \rightarrow \infty} \left(x e^{-\lambda x} \Big|_0^A - \int_0^A e^{-\lambda x} dx \right) = - \lim_{A \rightarrow \infty} \left(\frac{A}{e^{\lambda A}} - 0 + \right. \\ &\left. + \frac{1}{\lambda} e^{-\lambda x} \Big|_0^A \right) = - \frac{1}{\lambda} \lim_{A \rightarrow \infty} (e^{-\lambda A} - 1) = \frac{1}{\lambda} \quad (\text{здесь } \lim_{A \rightarrow \infty} \frac{A}{e^{\lambda A}} = \lim_{A \rightarrow \infty} \frac{1}{\lambda e^{\lambda A}} = 0 \text{ по правилу Лопиталья}). \end{aligned}$$

Найдём математическое ожидание квадрата случайной величины X :

$$\begin{aligned}
M(X^2) &= \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx = - \lim_{A \rightarrow \infty} \int_0^A x^2 de^{-\lambda x} = \\
&= - \lim_{A \rightarrow \infty} \left(x^2 e^{-\lambda x} \Big|_0^A - \int_0^A 2xe^{-\lambda x} dx \right) = \\
&= - \lim_{A \rightarrow \infty} \left(\frac{A^2}{e^{\lambda A}} - 0 - \frac{2}{\lambda} \int_0^A \lambda x e^{-\lambda x} dx \right) = \frac{2}{\lambda} \int_0^{+\infty} \lambda x e^{-\lambda x} dx = \frac{2}{\lambda} \cdot \frac{1}{\lambda} = \frac{2}{\lambda^2} \quad (\text{учли то, что несобственный интеграл } \int_0^{+\infty} \lambda x e^{-\lambda x} dx \text{ есть} \\
M(X) &= \frac{1}{\lambda}, \quad \lim_{A \rightarrow \infty} \frac{A^2}{e^{\lambda A}} = 0).
\end{aligned}$$

Далее находим дисперсию:

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

Из доказанной теоремы следует важное свойство показательного распределения: для случайной величины, распределённой по показательному закону, математическое ожидание равно среднему квадратическому отклонению, т.е.

$$M(X) = \sigma(X) = \frac{1}{\lambda} .$$

Показательный закон распределения широко применяется в теории массового обслуживания и в теории надёжности. Показательный закон распределения, и только он, обладает важным свойством, называемым отсутствием последствия. Объясним это свойство на примере.

Пример 3. Установлено, что время работы прибора до первой поломки является случайной величиной T , распределённой по показательному закону с параметром λ .

Обозначим через A случайное событие, заключающееся в том, что прибор будет работать безотказно на интервале $[0, t]$. Вероятность этого события $P(A) = P(T > t) = 1 - P(T \leq t) = 1 - F(t) = 1 - (1 - e^{-\lambda t}) = e^{-\lambda t}$. Далее обозначим через B случайное событие, заключающееся в том, что прибор будет безотказно работать на интервале времени $[t, t + \tau]$, а через C – случайное событие, заключающееся в безотказной работе прибора на интервале времени $[0, t + \tau]$. Такими же рассуждениями, как для события A , можно получить, что $P(C) = e^{-\lambda(t+\tau)}$.

Из определения случайных событий A , B и C следует, что $C = AB$. Тогда $P(C) = P(A)P_A(B)$. Найдём $P_A(B)$, т.е. условную вероятность того, что прибор будет безотказно работать на интервале $[t, t + \tau]$ при условии, что он уже проработал безотказно на интервале $[0, t]$: $P_A(B) = \frac{P(C)}{P(A)} = \frac{e^{-\lambda(t+\tau)}}{e^{-\lambda t}} = e^{-\lambda \tau}$.

Полученная формула не содержит t , следовательно, событие B не зависит от A , или, другими словами, вероятность безотказной работы прибора на промежутке времени $[t, t + \tau]$ зависит только от длины этого промежутка τ , и не зависит от того, сколько времени прибор проработал до этого.

5.6.3. Нормальный закон распределения

Определение 3. Непрерывная случайная величина X имеет нормальный закон распределения с параметрами a и $\sigma > 0$ (обозначается $N_{a,\sigma}$), если её плотность вероятности имеет вид:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} . \quad (5.6.7)$$

Для доказательства корректности определения вспомним значение несобственного интеграла, называемого интегралом Пуассона:

$$\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi} . \quad (5.6.8)$$

Вычислим $\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx$. Введём новую переменную $z = \frac{x-a}{\sigma}$. Тогда $x = a + \sigma z$, $dx = \sigma dz$. Пределы интегрирования не меняются, следовательно, учитывая 5.6.8, получим

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sigma dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} = 1 .$$

Нормальный закон распределения наиболее часто встречается на практике. Объясняется это тем, что он является предельным законом, к которому приближаются другие законы распределения при типичных условиях. Именно, если случайная величина X представляет собой сумму большого числа взаимно независимых случайных величин $X = X_1 + \dots + X_n$, то при весьма общих условиях закон распределения случайной величины X близок к нормальному. Условиям, при которых возникает нормальный закон распределения, посвящён ряд теорем, называемых *центральной предельной теоремой*. Смысл этих условий состоит в том, что удельный вес каждого отдельного слагаемого должен стремиться к нулю при увеличении числа слагаемых. Например, если X – производственная погрешность, то на неё влияют множество факторов: погрешность материала, погрешность станка, инструмента и т.д., причём, если ни одна из этих погрешностей не является определяющей, то случайная величина X имеет закон распределения, близкий к нормальному.

Теорема 9. Если случайная величина X имеет нормальный закон распределения $N_{a,\sigma}$, то её математическое ожидание

$$M(X) = a, \quad (5.6.9)$$

а дисперсия

$$D(X) = \sigma^2. \quad (5.6.10)$$

▷ Найдём математическое ожидание случайной величины X :

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left| \begin{array}{l} z = (x-a)/\sigma; \\ x = a + \sigma z; dx = \sigma dz \end{array} \right| = \int_{-\infty}^{+\infty} (a + \sigma z) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sigma dz \\ &= \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ze^{-\frac{z^2}{2}} \sigma dz. \end{aligned}$$

Первый интеграл есть интеграл Пуассона, а второй интеграл равен нулю как интеграл от нечётной функции по симметричному отрезку. Следовательно, $M(X) = a \frac{\sqrt{2\pi}}{\sqrt{2\pi}} = a$.

Дисперсия случайной величины X :

$$\begin{aligned} D(X) &= \int_{-\infty}^{+\infty} (x-a)^2 f(x)dx = \int_{-\infty}^{+\infty} (x-a)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \\ &= \left| \begin{array}{l} z = (x-a)/\sigma; \\ x = a + \sigma z; dx = \sigma dz \end{array} \right| = \int_{-\infty}^{+\infty} (\sigma z)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2} \sigma dz = \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z^2 e^{-z^2/2} dz = -\frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z de^{-z^2/2}. \end{aligned}$$

Применяя метод интегрирования по частям, получим:

$$D(X) = -\frac{\sigma^2}{\sqrt{2\pi}} ze^{-\frac{z^2}{2}} \Big|_{-\infty}^{+\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz. \text{ Учитывая, что интеграл во втором слагаемом есть интеграл Пуассона, а}$$

$$\lim_{z \rightarrow \pm\infty} \frac{z}{e^{z^2/2}} = 0, \text{ получим } D(X) = \sigma^2 \frac{\sqrt{2\pi}}{\sqrt{2\pi}} = \sigma^2. \triangleleft$$

Теперь стало ясно, что обозначения параметров a и σ в определении нормального распределения не были случайными. Действительно, a является математическим ожиданием случайной величины, а σ – средним квадратическим отклонением. Нормальный закон $N_{0,1}$ распределения случайной величины X с параметрами $a = 0$ и $\sigma = 1$ называется *стандартным* или *нормированным*.

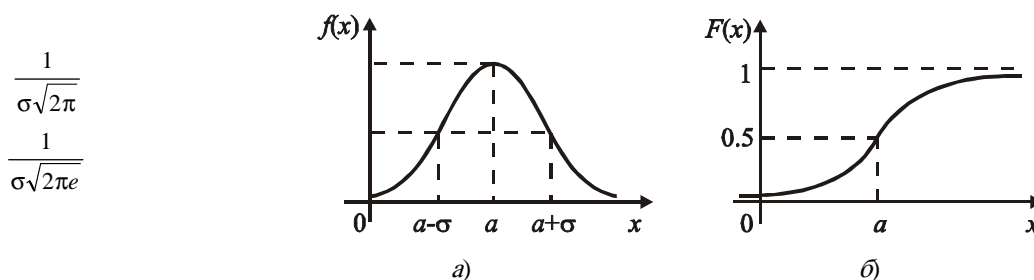


Рис. 5.6.3

Кривую нормального закона распределения называют *нормальной* или *гауссовой* кривой. На рис. 5.6.3 (а, б) приведены графики нормальной кривой с параметрами a и σ и функции распределения нормальной случайной величины. Отметим, что нормальная кривая симметрична относительно прямой $x = a$ и имеет максимум в точке $x = a$, равный $1/(\sigma\sqrt{2\pi})$, а также имеет точки перегиба $x = a \pm \sigma$ с ординатой $1/(\sigma\sqrt{2\pi e})$.

Выясним, как будет меняться нормальная кривая при изменении параметров a и σ . Если параметр σ остаётся постоянным, и меняется параметр a ($a_1 < a_2 < a_3$), то нормальная кривая смещается вдоль оси абсцисс, не меняя формы (рис. 5.6.4, а). Если параметр a остаётся постоянным, и меняется σ , то меняется ордината максимума кривой $1/(\sigma\sqrt{2\pi})$. При уменьшении σ (т.е. при уменьшении рассеяния случайной величины), ордината точки максимума увеличивается. Но так как площадь под любой кривой распределения должна оставаться равной единице, то кривая вытягивается вверх, одновременно сжимаясь с боков. При увеличении параметра σ наблюдается обратная картина (рис. 5.6.4, б). Таким образом, параметр a определяет положение, а параметр σ – форму нормальной кривой.

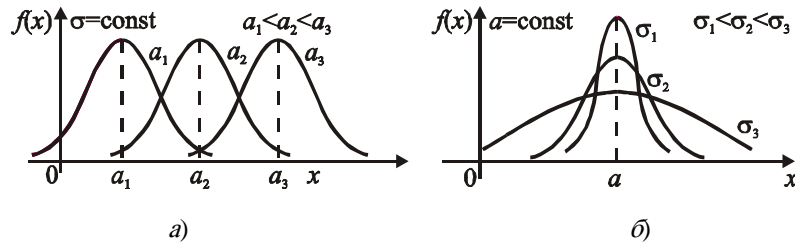


Рис. 5.6.4

Сложность нахождения функции распределения случайной величины, распределённой по нормальному закону, связана с тем, что интеграл от функции (5.6.7) является неберущимся в элементарных функциях. Поэтому функцию распределения $F(x)$ выражают через функцию Лапласа:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt, \quad (5.6.11)$$

для которой составлены таблицы.

Теорема 10. Если случайная величина X имеет нормальный закон распределения $N_{a,\sigma}$, то её функция распределения выражается через функцию Лапласа по формуле

$$F(x) = 0,5 + \Phi\left(\frac{x-a}{\sigma}\right). \quad (5.6.12)$$

▷ По формуле (5.2.2) функция распределения

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-a)^2}{2\sigma^2}} dt = \left| z = \frac{(t-a)}{\sigma}; \right. \\ &= \int_{-\infty}^{(x-a)/\sigma} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sigma dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_0^{(x-a)/\sigma} e^{-\frac{z^2}{2}} dz. \end{aligned}$$

Первый интеграл есть $\frac{1}{2}$ интеграла Пуассона (5.6.8) в силу чётности подынтегральной функции. Следовательно, первое слагаемое равно $\frac{1}{\sqrt{2\pi}} \frac{\sqrt{2\pi}}{2} = 0,5$. Второе слагаемое с учётом (5.6.11) равно $\Phi\left(\frac{x-a}{\sigma}\right)$. Следовательно, получаем (5.6.12). ◁

Рассмотрим свойства нормально распределённой случайной величины.

1. Вероятность попадания случайной величины X , распределённой по закону $N_{a,\sigma}$, в интервал $[\alpha; \beta]$, равна

$$P(\alpha \leq X \leq \beta) = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right).$$

▷ Учитывая, что вероятность попадания случайной величины в интервал равна приращению функции распределения на этом интервале, получим:

$$\begin{aligned} P(\alpha \leq X \leq \beta) &= F(\beta) - F(\alpha) = 0,5 + \Phi\left(\frac{\beta-a}{\sigma}\right) - \left(0,5 + \Phi\left(\frac{\alpha-a}{\sigma}\right)\right) = \\ &= \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right). \quad \triangleleft \end{aligned}$$

2. Вероятность того, что отклонение случайной величины X , распределённой по закону $N_{a,\sigma}$, от математического ожидания a не превысит по абсолютной величине заданное число $\delta > 0$, равна

$$P(|X - a| \leq \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right). \quad (5.6.13)$$

▷ Заменяем неравенство $|X - a| < \delta$ равносильным ему двойным неравенством $-\delta < X - a < \delta$ или $a - \delta < X < a + \delta$. Тогда, учитывая нечётность функции Лапласа, получим:

$$P(a - \delta \leq X \leq a + \delta) = \Phi\left(\frac{a + \delta - a}{\sigma}\right) - \Phi\left(\frac{a - \delta - a}{\sigma}\right) =$$

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

Применим формулу (5.6.13) при $\delta = 3\sigma$. Получим:

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

Отсюда вытекает «правило трёх сигм»: *если случайная величина распределена по нормальному закону, то практически достоверно, что абсолютная величина её отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения.*

3. Если случайная величина X_1 распределена по нормальному закону с параметрами a_1 и σ_1 , а случайная величина X_2 распределена по нормальному закону с параметрами a_2 и σ_2 , и они независимы, то их сумма $X = X_1 + X_2$ также распределена нормально с параметрами a и σ , причём $a = a_1 + a_2$, $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

5.6.4. Распределения некоторых случайных величин, связанных с нормальным

Коротко опишем несколько распределений, связанных с нормальным, которые будут использованы при изложении математической статистики.

Определение 4. Распределением χ^2 (*хи-квадрат*) с k степенями свободы называется распределение суммы квадратов k независимых случайных величин, каждая из которых имеет стандартное нормальное распределение, т.е.

$$\chi^2 = \sum_{i=1}^k X_i^2, \quad (5.6.1)$$

где X_i ($i = 1, \dots, k$) имеет нормальное распределение $N_{0,1}$.

Плотность распределения χ^2 выражается через так называемую гамма-функцию, и с увеличением числа степеней свободы плотность распределения χ^2 приближается к нормальному распределению.

Определение 5. Распределением Стьюдента (t -распределением) с k степенями свободы называется распределение случайной величины

$$t = \frac{X}{\sqrt{\chi^2/k}}, \quad (5.6.2)$$

где X имеет стандартное нормальное распределение $N_{0,1}$, а χ^2 – независимая от X случайная величина, имеющая χ^2 распределение с k степенями свободы.

Кривая распределения Стьюдента симметрична относительно оси ординат. С ростом числа степеней свободы t -распределение приближается к стандартному нормальному. Практически при $k > 30$ считают t -распределение приближенно нормальным.

Определение 6. Распределением Фишера-Снедекора (F -распределением) со степенями свободы k_1 и k_2 называется распределение случайной величины

$$F = \frac{\chi^2(k_1)/k_1}{\chi^2(k_2)/k_2}, \quad (5.6.3)$$

где $\chi^2(k_1)$, $\chi^2(k_2)$ – независимые случайные величины, имеющие χ^2 распределение соответственно с k_1 и k_2 степенями свободы.

5.7. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Как уже отмечалось, значения, которые принимает в результате испытания случайная величина, нельзя предсказать заранее. Однако при некоторых достаточно общих условиях, совокупное действие большого числа случайных величин приводит к результату, почти не зависящему от случая. Под законом больших чисел понимается ряд теорем, в каждой из которых для тех или иных условий устанавливается факт приближения средних значений большого числа случайных величин к некоторым постоянным. Прежде, чем перейти к этим теоремам, рассмотрим неравенство Чебышёва.

Теорема 1. Для любой случайной величины, имеющей математическое ожидание $a = M(X)$ и дисперсию $D(X)$, справедливо неравенство Чебышёва:

$$P(|X - a| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}, \quad (5.7.1)$$

где $\varepsilon > 0$.

▷ Проведём доказательство для случая, когда X – дискретная случайная величина с конечным числом значений. Запишем выражение дисперсии случайной величины X :

$$D(X) = \sum_{i=1}^n (x_i - a)^2 p_i = \sum_{i \in I_1} (x_i - a)^2 p_i + \sum_{i \in I_2} (x_i - a)^2 p_i. \quad (5.7.2)$$

Здесь множество I_1 – это такое подмножество индексов $\{1, 2, \dots, n\}$, что для всех $i \in I_1$ выполнено: $|x_i - a| < \varepsilon$, а множество I_2 – это подмножество всех таких индексов из $\{1, 2, \dots, n\}$, для которых выполнено: $|x_i - a| \geq \varepsilon$. Очевидно, что обе суммы в разложении (5.7.2) неотрицательны. Поэтому $D(X) \geq \sum_{i \in I_2} (x_i - a)^2 p_i \geq \sum_{i \in I_2} \varepsilon^2 p_i = \varepsilon^2 \sum_{i \in I_2} p_i$ или $\sum_{i \in I_2} p_i \leq \frac{D(X)}{\varepsilon^2}$. Учитывая оп-ределение множества I_2 , получим: $\sum_{i \in I_2} p_i = P(|X - a| \geq \varepsilon)$, откуда следует утверждение теоремы. ◁

Замечание. Учитывая, что события $|X - a| \geq \varepsilon$ и $|X - a| < \varepsilon$ являются противоположными, неравенство Чебышёва мож-но записать в другой форме:

$$P(|X - a| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}. \quad (5.7.3)$$

Пример 4. Оценим вероятность того, что для *любой* случайной величины вероятность её отклонения от математическо-го ожидания не превысит трёх среднеквадратических отклонений. Положим в формуле (5.7.3) $\varepsilon = 3\sigma$, где σ – среднеквадра-тическое отклонение случайной величины. Тогда

$$P(|X - a| < 3\sigma) \geq 1 - \frac{\sigma^2}{(3\sigma)^2} = \frac{8}{9} \approx 0,889.$$

Напомним, что для нормального закона распределения правило трёх сигм выполняется с вероятностью $P = 0,9973$. Таким образом, правило трёх сигм с достаточно большой вероятностью его выполнения можно применять на практике для произ-вольных случайных величин.

Пример 5. Найдём вероятность отклонения относительной частоты события $\frac{m}{n}$ от вероятности p менее чем на ε , т.е.

$P\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right)$, если испытания проводятся по схеме Бернулли. Для случайной величины $Z_A = \frac{m}{n}$ было получено (формулы (5.5.4) и (5.5.4)), что $M(Z_A) = p$, $D(Z_A) = \frac{pq}{n}$. Подставляя эти значения в (5.7.3) получим:

$$P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{pq}{n\varepsilon^2}.$$

Далее мы воспользуемся неравенством Чебышева для доказательства теоре-мы Чебышёва (закон больших чисел в форме Чебышева).

Теорема 2 (Чебышёва). Пусть X_1, \dots, X_n – взаимно независимые случайные величины, причём их дисперсии ограничены одной и той же постоянной C : $D(X_i) \leq C$ ($i = \overline{1, n}$). Тогда для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \frac{a_1 + a_2 + \dots + a_n}{n}\right| \leq \varepsilon\right) = 1, \quad (5.7.4)$$

где $a_i = M(X_i)$ ($i = \overline{1, n}$).

▷ Рассмотрим случайную величину X , равную среднему арифметическому n случайных величин:

$$X = \frac{X_1 + X_2 + \dots + X_n}{n}. \quad (5.7.5)$$

Тогда её математическое ожидание

$$\begin{aligned} a = M(X) &= \frac{1}{n}(M(X_1) + M(X_2) + \dots + M(X_n)) = \\ &= \frac{a_1 + a_2 + \dots + a_n}{n}. \end{aligned} \quad (5.7.6)$$

Так как случайные величины независимы, то

$$D(X) = \frac{1}{n^2}(D(X_1) + D(X_2) + \dots + D(X_n)) \leq \frac{C + C + \dots + C}{n^2} = \frac{nC}{n^2} = \frac{C}{n},$$

т.е. $1 - D(X) \geq 1 - \frac{C}{n}$. Применим к случайной величине X неравенство (5.7.3):

$$P(|X - a| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2} \geq 1 - \frac{C}{n\varepsilon^2}. \quad (5.7.7)$$

Перейдя в неравенстве (5.7.7) к пределу при $n \rightarrow \infty$, получим: $\lim_{n \rightarrow \infty} P(|X - a| < \varepsilon) \geq 1$. Наконец, учитывая, что вероятность не может превышать 1, окончательно можем написать: $\lim_{n \rightarrow \infty} P(|X - a| < \varepsilon) = 1$. Заменяя в последней формуле X и a по (5.7.5) и (5.7.6), получим утверждение теоремы. <

Подчеркнём смысл теоремы Чебышёва. При большом числе n случайных величин X_1, \dots, X_n практически достоверно, что их средняя арифметическая X – величина случайная сколь угодно мало отличается от неслучайной величины a , т.е. практически утрачивает случайный характер.

Следствие. Пусть X_1, \dots, X_n взаимно независимые одинаково распределённые случайные величины. Тогда для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - a\right| \leq \varepsilon\right) = 1,$$

где $a = M(X_i)$ ($i = 1, n$).

▷ В данном случае среднее арифметическое математических ожиданий $\frac{a + a + \dots + a}{n} = a$. Следовательно, из (5.7.4) получаем требуемое равенство. <

Из этого следствия ясно виден вероятностный смысл математического ожидания: при большом числе испытаний математическое ожидание приближённо равно среднему арифметическому наблюдаемых значений случайной величины.

6. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

6.1. ВВЕДЕНИЕ. ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Математическая статистика – это раздел математики, занимающийся систематизацией и обработкой результатов опытов и наблюдений, а также построением и проверкой подходящих математических моделей, опирающихся на теорию вероятностей. В отличие от теории вероятностей, которая изучает случайные явления на теоретическом уровне, не прибегая к эксперименту, математическая статистика позволяет находить закономерности в конкретных данных, и идёт от эксперимента (наблюдения) к построению вероятностной модели и её проверке.

Математическая статистика имеет дело со статистическими данными – это сведения о числе объектов в какой-либо достаточно обширной совокупности, обладающих теми или иными признаками. При большом числе наблюдений случайные воздействия на признак в значительной мере нейтрализуются, и получаемый результат оказывается предсказуемым (неслучайным). Этот принцип и является основой для практического использования статистических методов исследования. По сравнению с индивидуальным описанием объектов, статистические данные всегда обезличены, однако они позволяют изучать закономерности массовых явлений, прогнозировать их характеристики и воздействовать на них.

Широкому внедрению математико-статистических методов обработки данных способствовало появление электронных вычислительных машин, и, в особенности, персональных компьютеров. Статистические программные пакеты (среди которых можно выделить отечественные Статистик-консультант, Эвриста, СтатЭксперт и американские StatSoft Statistica, SysStat, StatGraphics и др.) сделали статистические методы более доступными и наглядными. Они освободили исследователя от рутинной работы по расчёту характеристик, построению таблиц и графиков, на его долю осталась творческая работа – постановка задачи и интерпретация результатов.

В практике статистических наблюдений различают два вида обследований: *сплошное*, когда обследуют каждый из объектов совокупности (например, перепись населения страны), и *выборочное*, когда исследуется часть объектов (например, для контроля качества проверяется часть изделий из выпущенной партии). На практике сплошное наблюдение применяют крайне редко, так как оно требует существенно больших затрат ресурсов по сравнению с выборочным. Кроме того, выборочное наблюдение является единственно возможным в случаях, когда наблюдаемая совокупность является бесконечной, или в случае, когда исследование объекта приводит к его уничтожению (например, исследование долговечности работы прибора).

Генеральной совокупностью называется вся подлежащая изучению совокупность объектов (наблюдений). Как правило, изучению подлежит некоторый количественный признак генеральной совокупности. В классической модели он рассматривается как одномерная случайная величина X с частично или полностью неизвестной нам функцией распределения вероятностей $F(x)$. Проведя n независимых испытаний (или измерений), получим набор случайных величин (X_1, X_2, \dots, X_n) , называемый *выборкой*. Каждая из этих случайных величин X_i имеет тот же закон распределения $F(x)$, что и случайная величина X . В серии уже проведённых экспериментов выборка – это набор чисел (x_1, x_1, \dots, x_n) . Однако, если серию экспериментов провести ещё раз, то получим уже другой набор чисел. Различные наблюдаемые значения признака называют *вариантами* (обозначаем их через x_i). *Объёмом* совокупности (выборочной или генеральной) называется число объектов в этой совокупности. Если объём генеральной совокупности достаточно велик, и его дальнейшее увеличение не сказывается на результатах обработки выборки, то допускают, что генеральная совокупность состоит из бесчисленного множества объектов. Ме-

тоды математической статистики позволяют по изучению некоторой части генеральной совокупности (выборки) выносить суждение об её свойствах в целом.

При составлении выборки возможны два способа: после того, как объект отобран и над ним произведено наблюдение, он может быть возвращён либо не возвращён в генеральную совокупность. *Повторной* называют выборку, при которой отобранный объект (перед отбором следующего) возвращается в генеральную совокупность. *Бесповторной* называют выборку, при которой отобранный объект в генеральную совокупность не возвращается. На практике обычно пользуются бесповторным случайным отбором. Если генеральная совокупность содержит бесконечное число объектов, а выборка конечна, то различие между повторной и бесповторной выборкой исчезает.

Для того, чтобы выборка позволяла правильно судить о генеральной совокупности, она должна достаточно хорошо воспроизводить её пропорции. Выборка, обладающая таким свойством, называется *репрезентативной*. В силу закона больших чисел, если каждый объект выборки будет отобран случайным образом, то выборка будет являться репрезентативной. На практике случайность отбора достигается тем, что извлечение объектов в выборку проводится путём жеребьёвки или с помощью датчика случайных чисел. Существуют различные методики отбора, которые применяются в зависимости от конкретной ситуации.

6.2. ВАРИАЦИОННЫЕ РЯДЫ И ИХ ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ

Рассмотрим произведённую выборку объёма n (x_1, x_2, \dots, x_n). Если среди чисел x_i есть одинаковые, то, как правило, их записывают один раз, и приписывают каждому из них число вхождений в выборку n_i , называемое *частотой*, тогда объём выборки можно определить как $n = \sum n_i$. Отношение частот к объёму выборки $w_i = n_i/n$ называют *относительными частотами*. Последовательность вариантов, записанных в порядке возрастания или убывания с соответствующими им частотами (или относительными частотами) называется *вариационным рядом*.

Пример 1. В результате 20 испытаний были получены данные (3, 2, 5, 4, 5, 5, 5, 5, 4, 4, 3, 3, 3, 4, 5, 2, 2, 3, 5, 4). Записать их в виде вариационного ряда.

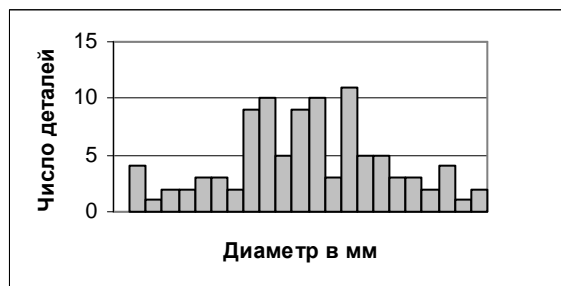
Имеем 4 варианты, расположив их в порядке возрастания, получим: $x_1 = 2, x_2 = 3, x_3 = 4, x_4 = 5$. Для каждой варианты определяем частоту n_i – количество вхождений в выборку. Получим следующий вариационный ряд:

x_i	2	3	4	5
n_i	3	5	5	7

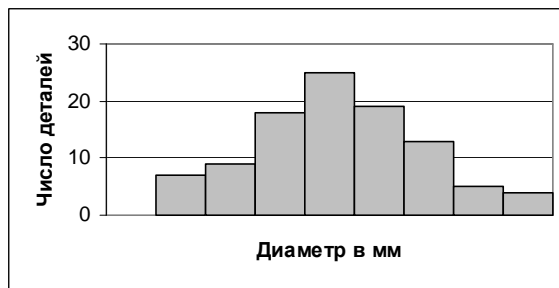
Вариационный ряд называется *дискретным*, если любые его варианты отличаются на постоянную величину, и *непрерывным*, если его значения могут отличаться одно от другого на сколь угодно малую величину. Пример дискретного вариационного ряда – отметки по математике у случайно отобранных студентов курса (пример 1). В качестве примера непрерывного вариационного ряда можно привести измерение диаметра некоторой детали при статистическом исследовании массовой продукции (пример 2).

При большом числе наблюдений n из соображений простоты и наглядности варианты разбивают на отдельные интервалы, т.е. производят их *группировку*. Обычно группировка по интервалам, в каждый из которых попадает не более 15...20 % значений выборки, оказывается достаточной для полного выявления существенных свойств распределения и надёжного вычисления по сгруппированным данным основных характеристик распределения. Согласно формуле Стерджесса (эта формула не является единственно возможной), рекомендуемое число интервалов $m = 1 + 3,322 \lg n$ (округляя сверху до целого, получим при $n = 50$ $m = 7$, при $n = 100$ $m = 8$, при $n = 200$ $m = 9$). Для интервального вариационного ряда *частотами* n_i называются числа, показывающие, сколько раз встречаются варианты из данного интервала.

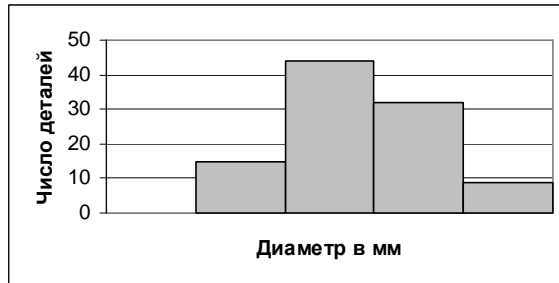
Для наглядного изображения интервальных вариационных рядов используют *гистограмму*, представляющую собой ступенчатую фигуру из прямоугольников с основаниями, равными интервалам значений признака (x_{i-1}, x_i), и высотами, равными частотам n_i (или относительным частотам w_i) интервалов. Гистограмма, составленная на основе группировки с маленькими интервалами, обычно многовершинная и не отражает наглядно характер распределения, а при слишком крупных интервалах теряется точность при вычислении характеристик распределения. В качестве примера на рис. 6.2.1, *а, б, в* приведены гистограммы, распределения 100 диаметров некоторой детали (в мм) при длине интервала группировки 0,5 мм (рис. 6.2.1, *а*), те же данные при длине интервала группировки 1,5 мм (рис. 6.2.1, *б*), и при длине интервала группировки 3 мм (рис. 6.2.1, *в*). Как видим, именно на рис. 6.2.1, *б*), где имеется 8 интервалов группировки, получена наиболее ясная картина характера распределения наблюдаемой величины.



а)



а)



б)

Рис. 6.2.1

Полигон, как правило, служит для изображения дискретного вариационного ряда, и представляет собой ломаную, в которой концы отрезков имеют координаты (x_i, n_i) (или (x_i, w_i)).

Весьма важным является понятие эмпирической функции распределения. Для её определения введём сначала понятие *накопленной частоты* (обозначаем $n_x^{\text{нак}}$) – это количество вариантов со значениями признака, меньшим x . Отношение накопленной частоты $n_x^{\text{нак}}$ к общему числу наблюдений n называют *накопленной относительной частотой* $w_x^{\text{нак}}$. *Эмпирической функцией распределения* называют функцию $F_n^*(x)$, определяющую для каждого значения x относительную частоту события $X < x$, или, по определению, $F_n^*(x) = w_x^{\text{нак}}$. Эмпирическая функция распределения определяется выборочными данными, и этим она отличается от функции распределения $F(x)$ генеральной совокупности. В то же время она сохраняет основные свойства функции распределения: её значения заключены между 0 и 1, и она является невозрастающей. Из закона больших чисел следует, что относительная частота события $X < x$, т.е. эмпирическая функция распределения $F_n^*(x)$, стремится по вероятности к вероятности этого же события, т.е. к функции вероятности $F(x)$. Следовательно, при больших n эмпирическая функция распределения может быть использована для приближённого представления интегральной функции распределения генеральной совокупности. Заметим, что для дискретного вариационного ряда эмпирическая функция распределения является разрывной, а для интервального вариационного ряда – непрерывной. В заключение приведем пример, иллюстрирующий эти понятия.

Пример 2. Построить эмпирическую функцию распределения для вариационного ряда.

Размеры деталей x	19,7...21,2	21,2...22,7	22,7...24,2	24,2...25,7	25,7...27,2	27,2...28,7	28,7...30,2	30,2...31,7
Кол-во n_i	7	9	18	25	21	9	8	3

Объём выборки $n = \sum_{i=1}^m n_i = 100$. Вычислим относительную частоту w_i , накопленную частоту $n_i^{\text{нак}}$ и накопленную относительную частоту $w_i^{\text{нак}}$.

w_i	0,07	0,09	0,18	0,25	0,21	0,09	0,08	0,03
$n_i^{\text{нак}}$	7	16	34	59	80	89	97	100
$w_i^{\text{нак}}$	0,07	0,16	0,34	0,59	0,80	0,89	0,97	1,00

По последней строке таблицы строим график эмпирической функции распределения (рис. 6.2.2).

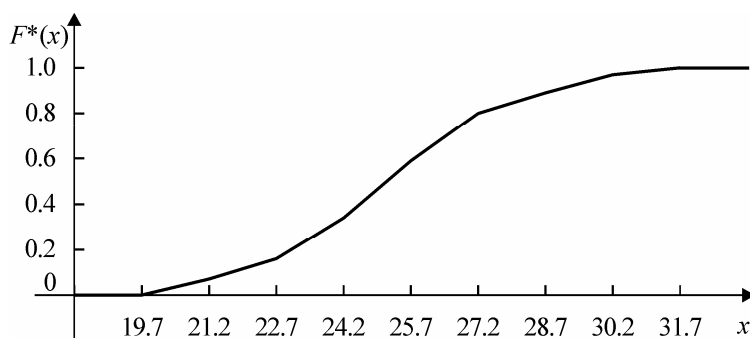


Рис.6.2.2

6.3. ОСНОВНЫЕ ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ВАРИАЦИОННОГО РЯДА

Средние величины характеризуют значения признака, вокруг которого концентрируются наблюдения, или, как говорят, центральную тенденцию распределения. Наиболее распространённой из средних величин является выборочная средняя, или средняя арифметическая вариационного ряда.

Определение 1. Выборочной средней \bar{X}_B называют среднее арифметическое значений признака выборочной совокупности:

$$\bar{X}_B = \frac{\sum_{i=1}^n X_i}{n}. \quad (6.3.1)$$

Следует понимать, что до проведения наблюдений, когда заранее неизвестно, какими они будут, \bar{X}_B рассматривается как случайная величина. После проведения наблюдений, когда получены конкретные значения, выборочная средняя становится уже неслучайной величиной (числом). В этом случае её обозначают \bar{x}_B . Если все значения x_1, x_2, \dots, x_n признака выборки объёма n не повторяются, то

$$\bar{x}_B = (x_1 + x_2 + \dots + x_n)/n. \quad (6.3.2)$$

Если же значения признака x_1, x_2, \dots, x_k имеют соответственно частоты n_1, n_2, \dots, n_k , причём $n = \sum n_i$, то

$$\bar{x}_B = (n_1 x_1 + n_2 x_2 + \dots + n_k x_k)/n. \quad (6.3.3)$$

Укажем, что для интервального вариационного ряда в качестве значений x_i берут середины соответствующих интервалов. Отметим основные свойства выборочной средней, аналогичные свойствам математического ожидания случайной величины.

1. Выборочная средняя постоянной равна самой этой постоянной.
2. Если все варианты умножить на какую-либо постоянную, то выборочная средняя также должна быть умножена на эту постоянную:

$$\overline{Cx}_B = C\bar{x}_B \quad \text{или} \quad \frac{\sum_{i=1}^m (Cx_i)n_i}{n} = C \frac{\sum_{i=1}^m x_i n_i}{n}.$$

3. Если ко всем вариантам прибавить какую-либо постоянную, то к выборочной средней также должна быть прибавлена эта постоянная:

$$\overline{C + x_B} = C + \bar{x}_B$$

или

$$\frac{\sum_{i=1}^m (C + x_i)n_i}{n} = \frac{\sum_{i=1}^m n_i}{n} + \frac{\sum_{i=1}^m x_i n_i}{n} = C + \frac{\sum_{i=1}^m x_i n_i}{n}.$$

В статистическом анализе применяются также и другие средние характеристики, наиболее распространенные из них – мода и медиана. *Медианой* Me_B вариационного ряда называется значение признака, приходящееся на середину ранжированного ряда наблюдений. Достоинство медианы как меры центральной тенденции состоит в том, что на неё не влияет изменение крайних членов вариационного ряда. Медиана предпочтительнее выборочной средней для ряда, у которого крайние варианты по сравнению с остальными оказались чрезмерно большими или чрезмерно малыми. *Модой* Mo_B вариационного ряда называется варианта, которому соответствует наибольшая частота. Отметим, что аналогично выборочной средней, до проведения наблюдений мода и медиана рассматриваются как случайные величины, а после проведения испытаний они уже являются неслучайными величинами (числами).

Пример 3. Вычислить выборочную среднюю для вариационного ряда из примера 1.

Здесь объём выборки $n = 20$. Получаем

$$\bar{x}_B = \frac{2 \cdot 3 + 3 \cdot 5 + 4 \cdot 5 + 5 \cdot 7}{20} = \frac{76}{20} = 3,8.$$

Пример 4. Вычислить выборочную среднюю для вариационного ряда из примера 2.

Здесь объём выборки $n = \sum_1^m n_i = 100$. Вычисляем среднюю выборочную по формуле (6.3.3), взяв в качестве вариант x_i середины интервалов.

$$\bar{x}_B = (20,45 \cdot 7 + 21,95 \cdot 9 + 23,45 \cdot 18 + 24,95 \cdot 25 + 26,45 \cdot 21 + 27,95 \cdot 9 + 29,45 \cdot 8 + 30,95 \cdot 3) / 100 = 25,22.$$

Рассмотренные нами средние величины не отражают изменчивости (вариации) значений признака. Простейшим, но весьма приближённым показателем вариации является *вариационный размах* R , равный разности между наибольшим и наименьшим вариантами ряда $R = x_{\max} - x_{\min}$. Наибольший интерес представляют меры вариации (рассеяния) наблюдаемых значений количественного признака выборки вокруг средней выборочной.

Определение 2. *Выборочной дисперсией* D_B называется среднее арифметической квадратов отклонения компонент выборки от выборочной средней:

$$D_B = \frac{\sum_{i=1}^n (X_i - \bar{X}_B)^2}{n}. \quad (6.3.4)$$

Определённая формулой (6.3.4) выборочная дисперсия является случайной величиной. После проведения испытаний, когда имеется конкретная реализация выборки, выборочная дисперсия, называемая также дисперсией вариационного ряда, становится неслучайной величиной, и может быть вычислена для несгруппированного вариационного ряда по формуле

$$D_B = \frac{\sum_{i=1}^n (x_i - \bar{x}_B)^2}{n}. \quad (6.3.5)$$

Если же значения признака x_1, x_2, \dots, x_k имеют соответственно частоты n_1, n_2, \dots, n_k , причём $n = \sum n_i$, то

$$D_B = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_B)^2}{n}. \quad (6.3.6)$$

Отметим, что формула (6.3.5) является частным случаем формулы (6.3.6), если положить все $n_i = 1$. Приведём основные свойства выборочной дисперсии, аналогичные свойствам дисперсии случайной величины. Указанные свойства будут также иметь место, если вместо вариант x_i рассматривать случайные величины X_i .

1. Дисперсия постоянной равна нулю.
2. Если все варианты умножить на одно и то же число C , то выборочная дисперсия должна быть умножена на C^2 :

$$D_B(Cx) = C^2 D_B(x)$$

или

$$\frac{\sum_{i=1}^m (Cx_i - \overline{Cx}_B)^2 n_i}{n} = C^2 \frac{\sum_{i=1}^m (x_i - \bar{x}_B)^2 n_i}{n}.$$

3. Если ко всем вариантам прибавить одно и то же число C , то выборочная дисперсия не изменится:

$$D_B(C + x) = D_B(x)$$

или

$$\frac{\sum_{i=1}^m (C + x_i - \overline{C + x}_B)^2 n_i}{n} = \frac{\sum_{i=1}^m (x_i - \bar{x}_B)^2 n_i}{n}.$$

4. Выборочная дисперсия равна разности между средней квадратов вариантов и квадратом средней выборочной:

$$D_B = \overline{x_B^2} - \bar{x}_B^2 = \frac{\sum_{i=1}^m x_i^2 n_i}{n} - (\bar{x}_B)^2. \quad (6.3.7)$$

$$\begin{aligned} \triangleright D_B &= \frac{\sum_{i=1}^m (x_i - \bar{x}_B)^2 n_i}{n} = \frac{\sum_{i=1}^m x_i^2 n_i}{n} - 2\bar{x}_B \frac{\sum_{i=1}^m x_i n_i}{n} + \bar{x}_B^2 \frac{\sum_{i=1}^m n_i}{n} = \\ &= \overline{x_B^2} - 2\bar{x}_B \cdot \bar{x}_B + \bar{x}_B^2 \frac{n}{n} = \overline{x_B^2} - (\bar{x}_B)^2. \triangleleft \end{aligned}$$

Кроме дисперсии для характеристики рассеяния используется *выборочное среднеквадратическое (стандартное) отклонение* $s = \sqrt{D_B}$. Его преимущество перед дисперсией состоит в том, что оно имеет те же единицы измерения, что и значения признака:

$$s = \sqrt{\frac{\sum_{i=1}^k n_i (x_i - \bar{x}_B)^2}{n}}. \quad (6.3.8)$$

Пример 5. Вычислить выборочную дисперсию, выборочное среднеквадратическое отклонение для вариационного ряда из примера 2.

Для нахождения дисперсии воспользуемся формулой (6.3.7). Учтём полученное в примере 4 значение выборочной средней $\bar{x}_B = 25,22$. Имеем

$$D_B = (20,45^2 \cdot 7 + 21,95^2 \cdot 9 + 23,45^2 \cdot 18 + 24,95^2 \cdot 25 + 26,45^2 \cdot 21 + 27,95^2 \cdot 9 + 29,45^2 \cdot 8 + 30,95^2 \cdot 3) / 100 - 25,22^2 = 6,5421.$$

Найдём выборочное среднеквадратическое отклонение:

$$s = \sqrt{D_B} = \sqrt{6,5421} \approx 2,56.$$

6.4. ОЦЕНКИ ПАРАМЕТРОВ

Пусть требуется изучить количественный признак генеральной совокупности. Допустим, что из теоретических предположений удалось установить, какое именно распределение имеет признак. Например, размер деталей, изготовленных по определённой технологии, имеет нормальное распределение $N_{a,\sigma}$ с известным параметром a и неизвестным σ . Количество покупателей в магазине в течение часа имеет распределение Пуассона с неизвестной интенсивностью λ и т.п.

Предположим, что имеется выборка объёма n , элементы которой X_1, X_2, \dots, X_n независимы (это условие выполнено для повторной выборки, а также для бесповторной выборки при бесконечном объёме генеральной совокупности) и имеют одинаковый закон распределения $F_\theta(x)$, известным образом зависящий от неизвестного параметра θ . Например, для всех $i = 1, \dots, n$

- X_i имеет распределение Пуассона Π_λ , где $\theta = \lambda > 0$ – неизвестный параметр;
- X_i имеет распределение Бернулли B_p , где $\theta = p \in (0; 1)$ – неизвестный параметр;
- X_i имеет равномерное распределение $U_{0,b}$, где $\theta = b > 0$ – неизвестный параметр;
- X_i имеет нормальное распределение $N_{a,\sigma}$, где $\theta = (a, \sigma) \in \mathfrak{R} \times (0; +\infty)$ – неизвестные параметры;
- X_i имеет нормальное распределение $N_{a,1}$, где $\theta = a \in \mathfrak{R}$ – неизвестный параметр.

Для вычисления параметра θ исследовать все элементы генеральной совокупности не представляется возможным. Поэтому о параметре θ пытаются судить по выборке. *Статистической оценкой (статистикой)* θ_n^* называют функцию от элементов выборки: $\theta_n^* = \theta_n^*(X_1, X_2, \dots, X_n)$. Поскольку X_1, X_2, \dots, X_n – случайные величины, то и оценка θ_n^* также является случайной величиной, в отличие от оцениваемого параметра θ , который является величиной неслучайной, детерминированной.

Всегда существует множество функций от результатов наблюдений, которые можно предложить в качестве оценки параметра θ . Например, если параметр θ является математическим ожиданием случайной величины X : $\theta = M(X)$, то в качестве его оценки θ_n^* по выборке можно предложить выборочную среднюю \bar{x}_B , моду Mo_B , медиану Me_B , полусумму наименьшего и наибольшего значений по выборке $\frac{x_{\max} + x_{\min}}{2}$ и т.д. Чтобы определить, какая из этих оценок является лучше других, нужно сформулировать свойства, какими должна обладать «доброкачественная» оценка. Так как θ_n^* случайная величина, невозможно предсказать конкретное значение оценки в каждом частном случае. Так что о качестве оценки следует судить не по индивидуальным её значениям, а по распределению её значений при большом числе испытаний.

Оценка θ_n^* называется *несмещённой*, если её математическое ожидание равно оцениваемому параметру, т.е. $M(\theta_n^*) = \theta$. В противном случае оценка называется смещённой. Если это равенство не выполняется, то оценка θ_n^* , полученная по разным выборкам, будет либо в среднем завышать значение θ (если $M(\theta_n^*) > \theta$) либо в среднем занижать её (если $M(\theta_n^*) < \theta$). Таким образом, требование несмещённости гарантирует от получения систематических ошибок.

Было бы ошибочным считать, что несмещённая оценка всегда даёт хорошее приближение оцениваемого параметра. Возможные значения θ_n^* могут быть сильно рассеяны вокруг своего среднего значения, т.е. иметь большую дисперсию $D(\theta_n^*)$. В этом случае найденная по одной выборке оценка может весьма отличаться от своего среднего значения $M(\theta_n^*)$, а значит, и от оцениваемого параметра θ . Если же потребовать, чтобы дисперсия $D(\theta_n^*)$ была мала, то возможность допустить грубую

ошибку уменьшится. Несмещённая оценка θ_n^* называется *эффективной*, если она имеет наименьшую дисперсию среди всех возможных несмещённых оценок параметра θ , вычисленных по выборкам одного и того же объёма n . Иногда в целях упрощения расчётов используют оценки, не обладающие высокой эффективностью. Так, например, в случае нормального распределения признака генеральную среднюю $\bar{x}_0 = M(X)$ часто оценивают медианой выборки Me_B , хотя её эффективной оценкой является выборочная средняя \bar{x}_B .

При рассмотрении выборок большого объёма к статистическим оценкам предъявляется требование состоятельности. Оценка θ_n^* называется *состоятельной*, если она удовлетворяет закону больших чисел, т.е. сходится по вероятности к оцениваемому параметру: $\lim_{n \rightarrow \infty} P(|\theta_n^* - \theta| < \varepsilon) = 1$ для любого $\varepsilon > 0$. В случае использования состоятельных оценок оправдывается увеличение объёма выборки, так как при достаточно большом n значительные ошибки при оценивании становятся маловероятными.

В качестве статистических оценок параметров генеральной совокупности желательно использовать оценки, удовлетворяющие одновременно требованиям несмещённости, состоятельности и эффективности. Однако иногда для простоты расчётов целесообразно применять оценки, обладающие большей дисперсией по сравнению с эффективными оценками, или незначительно смещённые оценки и т.п.

Теорема 1. Выборочная средняя \bar{X}_B есть несмещённая и состоятельная оценка генеральной средней $\bar{x}_0 = M(X)$.

▷ Докажем сначала несмещённость оценки. Найдём математическое ожидание \bar{X}_B :

$$M(\bar{X}_B) = M\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{\sum_{i=1}^n M(X_i)}{n} = \frac{nM(X)}{n} = M(X) = \bar{x}_0.$$

Найдём дисперсию выборочной средней, учитывая, что X_1, X_2, \dots, X_n – независимые случайные величины:

$$D(\bar{X}_B) = D\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{\sum_{i=1}^n D(X_i)}{n^2} = \frac{nD(X)}{n^2} = \frac{D(X)}{n}. \quad (6.4.1)$$

Из неравенства Чебышёва следует, что $P(|\bar{X}_B - \bar{x}_0| < \varepsilon) \geq 1 - \frac{D(\bar{X}_B)}{\varepsilon^2}$; в силу доказанного выше равенства запишем неравенство Чебышёва в виде $P(|\bar{X}_B - \bar{x}_0| < \varepsilon) \geq 1 - \frac{D(X)}{n\varepsilon^2}$. Следовательно, получаем, что для любого $\varepsilon > 0$ выполнено

$$\lim_{n \rightarrow \infty} P(|\bar{X}_B - \bar{x}_0| < \varepsilon) = 1. \triangleleft$$

Теорема 2. Выборочная дисперсия D_B является смещённой и состоятельной оценкой генеральной дисперсии $D(X) = \sigma^2$.

▷ Докажем смещённость оценки. Согласно (6.3.7) $D_B = \overline{X_B^2} - \bar{X}_B^2$. На основании свойства 3 средней арифметической и дисперсии, если из всех значений признака вычесть одно и то же число C , то средняя уменьшится на это число, а дисперсия не изменится, т.е.

$$D_B = D_B(X) = D_B(X - C) = \overline{(X - C)^2} - (\overline{X - C})^2 = \overline{(X - C)^2} - (\bar{X}_B - C)^2.$$

Полагая константу C равной генеральной средней $C = \bar{x}_0$, получим

$$D_B = \overline{(X - \bar{x}_0)^2} - (\bar{X}_B - \bar{x}_0)^2.$$

Найдём математическое ожидание выборочной дисперсии:

$$M(D_B) = M\left(\frac{\sum_{i=1}^n (X_i - \bar{x}_0)^2}{n}\right) - M(\bar{X}_B - \bar{x}_0)^2.$$

Рассмотрим первое слагаемое в правой части:

$$M\left(\frac{\sum_{i=1}^n (X_i - \bar{x}_0)^2}{n}\right) = \frac{\sum_{i=1}^n M(X_i - \bar{x}_0)^2}{n} = \frac{\sum_{i=1}^n D(X_i)}{n} = \frac{nD(X)}{n} = D(X).$$

Второе слагаемое в правой части, учитывая (6.3.10),

$$M(\bar{X}_B - \bar{x}_0)^2 = D(\bar{X}_B) = \frac{D(X)}{n}.$$

Поэтому $M(D_B) = D(X) - \frac{D(X)}{n} = \frac{n-1}{n}D(X)$. Состоятельность оценки примем без доказательства. <

Так как $\frac{n-1}{n} < 1$, то $M(D_B) < D(X)$ и выборочная дисперсия в среднем, при большом числе выборок, занижает генеральную дисперсию. Поэтому, заменяя $D(X)$ на D_B , мы допускаем систематическую погрешность. Чтобы её ликвидировать, достаточно ввести поправку. Сделав это, получим *исправленную выборочную дисперсию*

$$D_B^* = s^{*2} = \frac{n}{n-1}D_B \quad (6.4.2)$$

и *исправленное выборочное среднеквадратическое отклонение*

$$s^* = \sqrt{\frac{n}{n-1}D_B}. \quad (6.4.2)$$

Очевидно, что

$M(D_B^*) = M\left(\frac{n}{n-1}D_B\right) = \frac{n}{n-1}M(D_B) = \frac{n}{n-1} \frac{n-1}{n}D(X) = D(X)$, т.е. исправленная выборочная дисперсия является несмещённой и состоятельной оценкой генеральной дисперсии. Разница между D_B и D_B^* заметна при небольшом числе наблюдений n . Если же n велико (больше 30...40), то в качестве оценки $D(X)$ вполне можно использовать выборочную дисперсию D_B .

Что касается вопроса эффективности оценок параметров, то для широкого класса генеральных совокупностей доказательство эффективности оценок можно получить с помощью неравенства Рао-Крамера-Фреше, которое мы здесь приводить не будем. Приведём лишь некоторые полученные с использованием этого неравенства факты:

- если признак X в генеральной совокупности имеет нормальное распределение, то выборочная средняя \bar{X}_B является эффективной оценкой генеральной средней $\bar{x}_0 = M(X)$;

- если признак X в генеральной совокупности имеет нормальное распределение, то эффективной оценкой генеральной

дисперсии $D(X)$ является статистика $\hat{D}_B = \frac{\sum_{i=1}^n (x_i - \bar{x}_0)^2}{n}$, но для её нахождения нужно знать генеральную среднюю \bar{x}_0 , которая, как правило, неизвестна. Однако выборочная дисперсия D_B и исправленная выборочная дисперсия D_B^* являются асимптотически эффективными оценками генеральной дисперсии, т.е. при $n \rightarrow \infty$ они будут стремиться к эффективной оценке.

Вспомним, что математическое ожидание случайной величины мы называли начальным моментом первого порядка, а дисперсию – центральным моментом второго порядка. Аналогично вводят понятия эмпирических моментов. В отличие от теоретических моментов, эмпирические моменты вычисляют по данным наблюдений. *Начальным эмпирическим моментом k -го порядка* называют среднее значение k -х степеней вариантов:

$$\tilde{\nu}_k = \frac{\sum_{i=1}^m x_i^k n_i}{n}.$$

Центральным эмпирическим моментом k -го порядка называют среднее значение k -х степеней разностей $x_i - \bar{x}_B$:

$$\tilde{\mu}_k = \frac{\sum_{i=1}^m (x_i - \bar{x}_B)^k n_i}{n}.$$

Очевидно, что $\tilde{\nu}_1 = \bar{x}_B$, $\tilde{\mu}_1 = 0$, а $\tilde{\mu}_2 = D_B$. Можно доказать, что начальные и центральные эмпирические моменты являются состоятельными оценками соответственно начальных и центральных теоретических моментов того же порядка. На этом основан метод моментов точечной оценки неизвестных параметров заданного распределения. Он состоит в том, что определённое количество выборочных моментов (начальных или центральных) приравнивается к соответствующим теоретическим моментам случайной величины X .

Пусть задан вид функции распределения $F(x, \theta)$, определяемой одним неизвестным параметром θ . Для его оценки достаточно иметь одно уравнение относительно этого параметра. Согласно методу моментов, приравняем, например, начальный теоретический момент первого порядка начальному эмпирическому моменту первого порядка: $\nu_1 = \tilde{\nu}_1$. Учитывая,

что $v_1 = M(X)$, а $\tilde{v}_1 = \bar{x}_B$, получим уравнение $M(X) = \bar{x}_B$. При заданной функции распределения математическое ожидание может быть представлено как функция от неизвестного параметра $M(X) = \varphi(\theta)$, поэтому приходим к уравнению для нахождения параметра θ :

$$\varphi(\theta) = \bar{x}_B.$$

Пример 6. Найти методом моментов оценку параметра λ показательного распределения.

Приравняем начальный теоретический момент первого порядка начальному эмпирическому моменту первого порядка:

$$M(X) = \bar{x}_B.$$

Учтём, что для показательного распределения математическое ожидание $M(X) = \frac{1}{\lambda}$. Следовательно, искомая оценка параметра λ равна величине, обратной выборочной средней: $\lambda^* = \frac{1}{\bar{x}_B}$.

Если вид функции распределения определяется двумя неизвестными параметрами, т.е. имеет вид $F(x, \theta_1, \theta_2)$, то для их оценки необходимы два уравнения. Согласно методу моментов, приравняем, например, начальный теоретический момент первого порядка начальному эмпирическому моменту первого порядка и центральный теоретический момент второго порядка центральному эмпирическому моменту второго порядка: $v_1 = \tilde{v}_1$ и $\mu_2 = \tilde{\mu}_2$. Учитывая, что $v_1 = M(X)$, $\mu_2 = D(X)$, а $\tilde{v}_1 = \bar{x}_B$, $\tilde{\mu}_2 = D_B$, получим систему уравнений

$$\begin{cases} M(X) = \bar{x}_B; \\ D(X) = D_B. \end{cases}$$

Математическое ожидание и дисперсия являются функциями от неизвестных параметров θ_1 и θ_2 , поэтому указанную систему уравнений можно рассматривать как систему двух уравнений с двумя неизвестными θ_1 и θ_2 . Решив эту систему, найдём оценки неизвестных параметров θ_1^* и θ_2^* .

Пример 7. Найти методом моментов оценку параметров a и b равномерного распределения $U_{a,b}$.

Согласно методу моментов имеем систему уравнений

$$\begin{cases} M(X) = \bar{x}_B; \\ D(X) = D_B. \end{cases}$$

Учтём, что для равномерного распределения $M(X) = \frac{a+b}{2}$ и $D(X) = \frac{(b-a)^2}{12}$. Следовательно, для нахождения оценок неизвестных параметров нужно решить систему уравнений

$$\begin{cases} (a+b)/2 = \bar{x}_B; \\ (b-a)^2/12 = D_B. \end{cases}$$

Решив эту систему, получим искомые оценки

$$a^* = \bar{x}_B - s\sqrt{3}, \quad b^* = \bar{x}_B + s\sqrt{3},$$

где $s = \sqrt{D_B}$.

6.5. ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ. ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

Точечной называют оценку параметра θ генеральной совокупности, которая определяется одним числом. Оценки, рассмотренные выше – точечные. Однако точечная оценка θ_n^* является лишь приближённым значением неизвестного параметра θ , и при выборке малого объёма точечная оценка может значительно отличаться от оцениваемого параметра.

Чтобы получить представление о точности и надёжности оценки θ_n^* параметра θ пользуются интервальными оценками. *Интервальной оценкой* параметра θ называют числовой интервал $(\theta_n^{*(1)}, \theta_n^{*(2)})$ (рис. 6.5.1), для которого выполнено

$$P(\theta_n^{*(1)} < \theta < \theta_n^{*(2)}) = \gamma. \quad (6.5.1)$$

Границы доверительного интервала и его длина находятся по выборочным данным, и являются случайными величинами, поэтому формулу (6.5.1) читают как «интервал $(\theta_n^{*(1)}, \theta_n^{*(2)})$, который с заданной вероятностью γ накрывает неизвестное значение параметра θ », а не «параметр θ лежит в интервале...». Такой интервал называется *доверительным*, а вероятность γ называется *доверительной вероятностью*, или надёжностью оценки.

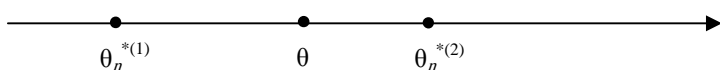


Рис. 6.5.1

Следует иметь в виду, что доверительный интервал при заданной надёжности γ можно выбрать различным способом. Пусть по данным выборки найдена точечная оценка θ_n^* параметра θ . Очень часто (но не всегда) доверительный интервал выбирают симметричным относительно параметра θ_n^* , т.е. $(\theta_n^* - \delta; \theta_n^* + \delta)$. Число δ , т.е. максимальное отклонение оценки от оцениваемого параметра, которое возможно с заданной доверительной вероятностью γ , называется *точностью оценки*. Величина доверительного интервала увеличивается с ростом доверительной вероятности γ , т.е. чем больше уверенности в том, что параметр лежит в указанном интервале, тем шире интервал. Обычно надёжность оценки задают заранее, причём в качестве γ берут число, близкое к единице, например 0,95 или 0,99 и т.д. Вероятностный смысл этого параметра – $P(|\theta - \theta_n^*| < \delta) = \gamma$, т.е. вероятность того, что оценка отклонится от истинного значения параметра не больше, чем на δ , равна γ .

Пусть количественный признак генеральной совокупности X имеет нормальное распределение $N_{a,\sigma}$, а оцениваемым параметром θ является генеральная средняя \bar{x}_0 . Доверительный интервал будем искать в виде $(\bar{x}_B - \delta; \bar{x}_B + \delta)$. Следовательно, нужно указать способ, как, зная объём выборки n и надёжность оценки γ , найти точность оценки δ .

Рассмотрим случай, когда генеральная дисперсия $D = \sigma^2$ является известной величиной (например, это заданная заранее ошибка измерительного прибора). По выборке объёма n находим значение случайной величины $\bar{X}_B = \left(\sum_{i=1}^n X_i \right) / n$. Так как все случайные величины X_i имеют нормальное распределение $N_{a,\sigma}$, то \bar{X}_B также имеет нормальное распределение, причём при доказательстве теоремы 1 было получено, что $M(\bar{X}_B) = \bar{x}_0 = a$, а $D(\bar{X}_B) = \frac{D(X)}{n} = \frac{\sigma^2}{n}$, т.е. \bar{X}_B имеет распределение $N_{a, \frac{\sigma}{\sqrt{n}}}$. Тогда, согласно (5.6.13), $P(|\bar{X}_B - a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma/\sqrt{n}}\right)$. Учитывая, что указанная вероятность задана и равна надёжности оценки γ , получим условие для определения числа δ :

$$\gamma = 2\Phi\left(\frac{\delta}{\sigma/\sqrt{n}}\right).$$

Приведём рабочие формулы, по которым может быть найден доверительный интервал для параметра a :

- определяем число t из равенства $\Phi(t) = \gamma/2$, т.е. по таблице функции Лапласа находят значение аргумента t , которому соответствует значение функции Лапласа $\gamma/2$;
- находим точность оценки δ по формуле

$$\delta = \frac{t\sigma}{\sqrt{n}}; \quad (6.5.2)$$

- строим доверительный интервал для параметра a :

$$(\bar{x}_B - \delta; \bar{x}_B + \delta).$$

Из формулы (6.5.2) можно сделать следующие выводы:

- при возрастании объёма выборки n число δ , а также и длина доверительного интервала, убывает, следовательно, точность оценки увеличивается;
- при увеличении надёжности γ увеличивается число t (напомним, что функция Лапласа является возрастающей), следовательно, число δ также увеличивается; т.е. увеличение надёжности оценки приводит к уменьшению её точности.

На практике обычно встречается случай, когда генеральная дисперсия неизвестна, а известна лишь её исправленная выборочная оценка $D_B^* = S^{*2}$. Оказывается, что по данным выборки можно построить случайную величину (статистику)

$$T = \frac{\bar{X}_B - a}{S^*/\sqrt{n}}. \quad (6.5.3)$$

Представим эту статистику в виде

$$T = \frac{(\bar{X}_B - a) / \frac{\sigma}{\sqrt{n}}}{\sqrt{\frac{(n-1)S^{*2}}{\sigma^2} \cdot \frac{1}{n-1}}}.$$

Можно доказать, что случайная величина $Z = (\bar{X}_B - a) / \frac{\sigma}{\sqrt{n}}$ имеет стандартное нормальное распределение $N_{0,1}$, а случайная

величина $H = \frac{(n-1)S^{*2}}{\sigma^2} - \chi^2$ распределение с $k = n - 1$ степенями свободы, причём эти случайные величины являются независимыми. Тогда, согласно (5.6.2), случайная величина T , определяемая (6.5.3), имеет распределение Стьюдента с $k = n - 1$ степенями свободы. Задав число γ – вероятность выполнения неравенства

$$P(|T| < t_\gamma) = \gamma, \quad (6.5.4)$$

можно по таблице критических точек распределения Стьюдента при уровне вероятности $\alpha = 1 - \gamma$ и числе степеней свободы $n - 1$ получить значение числа $t_\gamma = T(1 - \gamma; n - 1)$. Подставив в неравенство (6.5.4) значение статистики T из (6.5.3), и заменив случайные величины \bar{X}_B и S неслучайными величинами \bar{x}_B и s , найденными по выборке, получим

$$P(\bar{x}_B - t_\gamma s^* / \sqrt{n} < a < \bar{x}_B + t_\gamma s^* / \sqrt{n}) = \gamma.$$

Следовательно, при неизвестном σ точность оценки δ находится по формуле

$$\delta = \frac{t_\gamma s^*}{\sqrt{n}}. \quad (6.5.5)$$

Заметим, что при большом числе наблюдений ($n > 30$) точности оценок, найденные по формулам (6.5.2) и (6.5.5) различаются несущественно, поэтому для нахождения δ можно пользоваться формулой (6.5.2), заменив в ней точное значение среднеквадратического отклонения σ на исправленное выборочное значение среднеквадратического отклонения s^* . В случае же, когда объём выборки незначителен, предлагается для нахождения доверительного интервала для параметра a пользоваться следующим алгоритмом:

- определяем число t_γ по таблице критических точек распределения Стьюдента при уровне вероятности $\alpha = 1 - \gamma$ и числе степеней свободы $n - 1$;

- находим точность оценки δ по формуле (6.5.5);

- строим доверительный интервал для параметра a :

$$(\bar{x}_B - \delta; \bar{x}_B + \delta).$$

Без доказательства приведём один из способов построения доверительного интервала для среднеквадратического отклонения σ нормального распределения. Отметим, что в данном случае доверительный интервал не является симметричным относительно точечной оценки s^* указанного параметра.

- Находим числа χ_1^2 и χ_2^2 по таблице критических точек распределения χ^2 при числе степеней свободы $n - 1$ и уровнях вероятности $(1 + \gamma)/2$ и $(1 - \gamma)/2$ соответственно.

- Строим доверительный интервал

$$\sigma \in \left(\frac{s^* \sqrt{n-1}}{\chi_2}; \frac{s^* \sqrt{n-1}}{\chi_1} \right). \quad (6.5.6)$$

Пример 8. Найти доверительные интервалы для математического ожидания и среднеквадратического отклонения с надёжностью 95 % по данным выборки из примера 2. Предполагается, что признак X (размер деталей) в генеральной совокупности имеет нормальное распределение.

Построим доверительный интервал для генеральной средней с уровнем доверительной вероятности $\gamma = 0,95$. Значение генеральной дисперсии неизвестно, но объём выборки ($n = 100$) велик, поэтому для нахождения точности оценки можно воспользоваться формулой (6.5.2). Сначала из таблицы функции Лапласа найдём значение t из условия $\Phi(t) = \gamma/2 = 0,475$. Получим $t = 1,96$. Точечная оценка среднеквадратического отклонения s была найдена при решении примера 5: $s = 2,56$. Напомним, что несмещённой оценкой параметра σ является исправленное выборочное среднеквадратическое отклонение

$$s^* = \sqrt{\frac{n}{n-1}} s = \sqrt{\frac{100}{99}} \cdot 2,56 = 2,57.$$

Далее находим точность оценки δ , заменив в формуле (6.5.2) генеральное среднеквадратическое отклонение его точечной оценкой s^* : $\delta = \frac{ts^*}{\sqrt{n}} = \frac{1,96 \cdot 2,57}{\sqrt{100}} = 0,50$. Доверительный интервал для генеральной средней $\bar{x}_0 = a$ имеет вид $(\bar{x}_B - \delta; \bar{x}_B + \delta)$. Подставляя значения и учитывая, что выборочная средняя, найденная при решении примера 4, $\bar{x}_B = 25,22$, получим: $a \in (25,22 - 0,50; 25,22 + 0,50)$ или $24,72 < a < 25,72$. Напомним, что вероятность выполнения этого неравенства равна 0,95.

Построим доверительный интервал для генерального среднеквадратического отклонения с заданным уровнем доверительной вероятности $\gamma = 0,95$. Найдём значение χ_1^2 по таблице критических точек распределения χ^2 при уровне вероятности $(1 + \gamma)/2 = 0,975$ и числе степеней свободы $k = n - 1 = 99$. Получаем $\chi_1^2 = 73,36$, следовательно, $\chi_1 = 8,57$. Найдём значение χ_2^2 по таблице критических точек распределения χ^2 при уровне вероятности $(1 - \gamma)/2 = 0,025$ и числе степеней свободы $k = n - 1 = 99$. Получаем $\chi_2^2 = 128,42$, следовательно, $\chi_2 = 11,33$. Согласно (6.5.6), доверительный интервал для генерального среднеквадратического отклонения σ имеет вид: $\sigma \in \left(\frac{s\sqrt{n-1}}{\chi_2}; \frac{s\sqrt{n-1}}{\chi_1} \right)$. Подставляя значения, получаем, что с вероятностью 0,95 выполнено $\sigma \in \left(\frac{2,57\sqrt{99}}{11,33}; \frac{2,57\sqrt{99}}{8,57} \right)$ или $\sigma \in (2,26; 2,98)$.

6.6.1. Основные понятия

С теорией статистической оценки параметров тесно связана проверка статистических гипотез. Она используется в случаях, когда нужно сделать обоснованный вывод о преимуществах того или иного способа производства, той или иной технологии, об адекватности математической модели и т.п. Следует учитывать, что выводы, сделанные с помощью теории вероятностей и математической статистики, не являются окончательными. Математическая статистика позволяет ответить на заданный вопрос только с определённой долей уверенности. Например, на основании обработки статистических данных можно сделать вывод: «с вероятностью 95 % можно утверждать, что срок службы производимых приборов не меньше нормативного».

Статистической гипотезой называется любое предположение о виде неизвестного распределения или о параметрах закона распределения. Например, статистическими являются гипотезы:

- *генеральная совокупность распределена по нормальному закону;*
- дисперсии двух нормальных совокупностей равны между собой;
- математическое ожидание признака X нормально распределённой генеральной совокупности равно 25,5.

Различают простую и сложную статистические гипотезы. *Простая* гипотеза полностью определяет теоретическую функцию распределения случайной величины, а сложная гипотеза говорит о принадлежности функции распределения к некоторому классу. Проверяемую гипотезу обычно называют основной (нулевой) и обозначают H_0 . Наряду с выдвинутой гипотезой рассматривают и альтернативную гипотезу H_1 , противоречащую основной. Как правило, но не всегда, она является логическим отрицанием основной.

Для проверки нулевой гипотезы используют специально подобранную случайную величину (статистику) $\theta(X_1, \dots, X_n)$. В зависимости от вида распределения её обозначают через U (если она распределена по нормальному закону), T (если она распределена по закону Стьюдента), F (если она распределена по закону Фишера-Снедекора) и т.д. После проведения эксперимента может быть получено наблюдаемое значение статистики, т.е. $\theta_{\text{набл}} = \theta(x_1, \dots, x_n)$. Затем, используя знания о распределении случайной величины θ , множество всех её возможных значений разбивают на два подмножества: одно из них содержит значения критерия, при которых нулевая гипотеза отвергается (*критическая область* W), а другое (\bar{W}) содержит те значения критерия, при которых гипотеза принимается (*область принятия гипотезы*). *Правило, по которому гипотеза H_0 принимается или отвергается, называется статистическим критерием.* Иногда критерием называют также статистику $\theta(X_1, \dots, X_n)$.

Основной принцип проверки статистических гипотез можно сформулировать так: если наблюдаемое значение статистики критерия $\theta_{\text{набл}}$ принадлежит критической области, то нулевую гипотезу отвергают в пользу конкурирующей гипотезы; если наблюдаемое значение критерия принадлежит области принятия гипотезы, то нулевую гипотезу принимают. При этом возможны 4 случая:

Гипотеза H_0	Принимается	Отвергается
Верна	Правильное решение	Ошибка 1-го рода
Не верна	Ошибка 2-го рода	Правильное решение

Вероятность α допустить ошибку 1-го рода, т.е. отвергнуть гипотезу H_0 , когда она верна, называется *уровнем значимости критерия*.

$$\alpha = P(\theta(X_1, \dots, X_n) \in W | H_0).$$

Вероятность допустить ошибку 2-го рода, т.е. принять гипотезу H_0 , когда она неверна, обозначают β

$$\beta = P(\theta(X_1, \dots, X_n) \in \bar{W} | H_1).$$

Вероятность $1 - \beta$ не допустить ошибку 2-го рода называется *мощностью критерия*. Используя терминологию статистического контроля качества продукции, можно сказать, что вероятность α является «риском поставщика», когда по результатам выборочного контроля изделий партии бракуется удовлетворяющая стандарту партия. В этом примере вероятность β – «риск потребителя», когда в результате анализа выборки признаётся годной партия изделий, не удовлетворяющая стандарту.

Вероятности ошибок 1-го и 2-го рода определяются выбором критической области. При фиксированном объёме выборки уменьшение ошибки 1-го рода α неизбежно приводит к увеличению ошибки 2-го рода β , и наоборот. Лишь при увеличении объёма выборки возможно одновременное уменьшение вероятностей α и β .

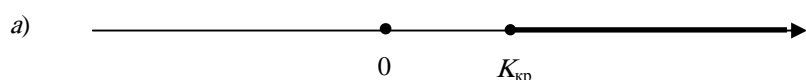
Поскольку статистика критерия θ – одномерная случайная величина, все её возможные значения принадлежат некоторому интервалу, поэтому критическую область и область принятия гипотезы можно разделить точками. *Критическими точками* $K_{\text{кр}}$ называют точки, отделяющие критическую область от области принятия гипотезы.

Различают одностороннюю (правостороннюю и левостороннюю) и двустороннюю критические области.

Правосторонней называют критическую область, определяемую неравенством $\theta > K_{\text{кр}}$ (рис. 6.6.1, а). При этом $P(\theta > K_{\text{кр}}) = \alpha$.

Левосторонней называют критическую область, определяемую неравенством $\theta < K_{\text{кр}}$ (рис. 6.6.1, б). При этом $P(\theta < K_{\text{кр}}) = \alpha$.

Двусторонней называют критическую область, определяемую неравенствами $\theta < K_{\text{кр}1}$ и $\theta > K_{\text{кр}2}$ (рис. 6.6.1, в). При этом $P(\theta > K_{\text{кр}2}) = P(\theta < K_{\text{кр}1}) = \alpha/2$.



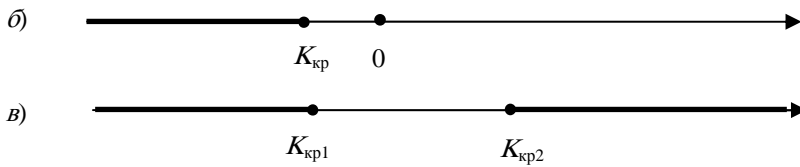


Рис. 6.6.1

Следует иметь в виду, что принцип проверки статистических гипотез не даёт математического доказательства её истинности. Принятие гипотезы H_0 лишь означает, что она не противоречит имеющимся опытным данным. При увеличении объёма выборки или при изменении уровня значимости критерия та же самая гипотеза H_0 может быть отвергнута. Рассмотрим способы проверки некоторых наиболее часто встречающихся гипотез.

6.6.2. Гипотеза о равенстве генеральной средней нормальной совокупности заданному числовому значению

Пусть количественный признак генеральной совокупности X распределён нормально, причём имеются основания предполагать, что генеральная средняя этой совокупности \bar{x}_0 равна некоторому значению a (например, если X – размер деталей, изготавливаемых станком-автоматом, то a – заданный проектный размер). Предполагаем, что дисперсия генеральной совокупности $D = \sigma^2$ известна (например, может быть найдена теоретически, или вычислена по выборке большого объёма). Кроме того, по произведённой выборке объёма n найдена выборочная средняя \bar{x}_B . Требуется по выборочной средней при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{X}_0 = a$.

Предположим, что гипотеза H_0 верна, т.е. X имеет распределение $N_{a,\sigma}$. Тогда все компоненты выборки (X_1, \dots, X_n) также имеют распределение $N_{a,\sigma}$. Выше мы показали, что в этом случае выборочная средняя \bar{X}_B также имеет нормальное распределение с математическим ожиданием a и дисперсией $\frac{\sigma^2}{n}$, т.е. распределена по закону $N_{a,\sigma/\sqrt{n}}$. В качестве статистики (критерия проверки нулевой гипотезы) принимают случайную величину

$$U = \frac{(\bar{X}_B - a)\sqrt{n}}{\sigma}, \quad (6.6.1)$$

которая имеет стандартное нормальное распределение $N_{0,1}$, так как

$$M(U) = \frac{(M(\bar{X}_B) - a)\sqrt{n}}{\sigma} = 0,$$

$$D(U) = D\left(\frac{(\bar{X}_B - a)\sqrt{n}}{\sigma}\right) = \frac{n}{\sigma^2} D(\bar{X}_B - a) = \frac{n}{\sigma^2} D(\bar{X}_B) = \frac{n}{\sigma^2} \frac{\sigma^2}{n} = 1.$$

Таким образом, наблюдаемое значение критерия находим по правилу

$$U_{\text{набл}} = \frac{(\bar{x}_B - a)\sqrt{n}}{\sigma}. \quad (6.6.2)$$

Рассмотрим три различных случая конкурирующих гипотез.

1. При конкурирующей гипотезе $H_1: \bar{x}_0 \neq a$ критическая область является двусторонней. Так как плотность распределения стандартной нормальной случайной величины U является чётной функцией, то правая и левая критические точки расположены симметрично относительно нуля. Обозначим их $U_{\text{кр}}$ и $-U_{\text{кр}}$. Тогда вероятности попадания в левую и правую части критической области одинаковы и равны $\alpha/2$, а вероятность выполнения неравенства $|U| < U_{\text{кр}}$ равна $1 - \alpha$ (рис. 6.6.1). С другой стороны имеем: $P(|U| < U_{\text{кр}}) = 2\Phi(U_{\text{кр}})$. Следовательно, выполнено $2\Phi(U_{\text{кр}}) = 1 - \alpha$, откуда получаем, что критическая точка $U_{\text{кр}}$ может быть найдена по таблице функции Лапласа из условия

$$\Phi(U_{\text{кр}}) = (1 - \alpha)/2. \quad (6.6.3)$$

Если для найденного по формуле (6.7.2) наблюдаемого значения критерия выполнено $|U_{\text{набл}}| < U_{\text{кр}}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2. В некоторых практических ситуациях имеет смысл принять в качестве конкурирующей гипотезу $H_1: \bar{x}_0 > a$. Например, если проверяют гипотезу о равенстве вредных примесей в продукте нормативу. Тогда критическую точку $U_{\text{кр}}$ правосторонней критической области находим по таблице функции Лапласа из условия $\Phi(U_{\text{кр}}) = (1 - 2\alpha)/2$.

Если $U_{\text{набл}} < U_{\text{кр}}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

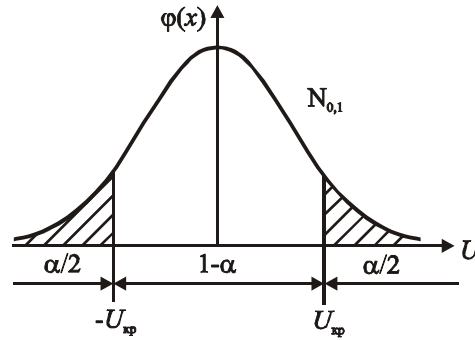


Рис. 6.6.1

3. При конкурирующей гипотезе $H_1: \bar{x}_0 < a$ критическую точку $U_{кр}$ левосторонней критической области находим по таблице функции Лапласа из условия $\Phi(U_{кр}) = (1 - 2\alpha)/2$.

Если $U_{набл} > U_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

Предположим теперь, что дисперсия генеральной совокупности $D = \sigma^2$ неизвестна, а известна только её исправленная выборочная оценка $D_B^* = s^{*2}$. При выборке небольшого объёма (менее 30 наблюдений) нахождение критической точки по таблице функции Лапласа может привести к существенной погрешности. В качестве критерия проверки нулевой гипотезы тогда принимают случайную величину

$$T = \frac{(\bar{X}_B - a)\sqrt{n}}{s}, \quad (6.6.4)$$

имеющую распределение Стьюдента с $n - 1$ степенями свободы.

Для того, чтобы при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x}_0 = a$, нужно вычислить наблюдаемое значение критерия

$$T_{набл} = \frac{(\bar{x}_B - a)\sqrt{n}}{s}. \quad (6.6.5)$$

1. При конкурирующей гипотезе $H_1: \bar{x}_0 < a$ критическую точку $T_{кр}(\alpha, n - 1)$ находим по таблице критических точек распределения Стьюдента при $n - 1$ степенях свободы и вероятности α . Если $T_{набл} < T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2. При конкурирующей гипотезе $H_1: \bar{x}_0 > a$ критическую точку $T_{кр}(2\alpha, n - 1)$ правосторонней критической области находим по таблице критических точек распределения Стьюдента при $n - 1$ степенях свободы и вероятности 2α . Если $T_{набл} < T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3. При конкурирующей гипотезе $H_1: \bar{x}_0 < a$ критическую точку $T_{кр}(2\alpha, n - 1)$ левосторонней критической области находим по таблице критических точек распределения Стьюдента при $n - 1$ степенях свободы и вероятности 2α . Если $T_{набл} > -T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

6.6.3. Гипотеза о равенстве генеральной дисперсии нормальной совокупности заданному числовому значению

Пусть генеральная совокупность распределена нормально, причём имеются основания предполагать, что генеральная дисперсия равна некоторому заданному значению σ_0^2 . Пусть по выборке объёма n получена исправленная выборочная дисперсия $D_B^* = s^{*2}$. Требуется по исправленной выборочной дисперсии при заданном уровне значимости α проверить нулевую гипотезу $H_0: \sigma^2 = \sigma_0^2$.

На практике рассматриваемая гипотеза проверяется, если нужно проверить точность настройки оборудования, устойчивость технологических процессов, точность приборов и т.п.

Для проверки гипотезы H_0 используют случайную величину $(n-1)s^{*2}/\sigma^2$, которая имеет распределение χ^2 с $n - 1$ степенями свободы.

Для того, чтобы при заданном уровне значимости α проверить нулевую гипотезу $H_0: \sigma^2 = \sigma_0^2$, нужно вычислить наблюдаемое значение критерия

$$\chi_{набл}^2 = \frac{(n-1)s^{*2}}{\sigma^2}. \quad (6.6.5)$$

1. При конкурирующей гипотезе $H_1: \sigma^2 \neq \sigma_0^2$ критическая область является двусторонней. Критические точки ищем по таблице критических точек распределения χ^2 с $n - 1$ степенями свободы: левую критическую точку при вероятности $1 - \alpha/2$, а правую критическую точку при вероятности $\alpha/2$. Таким образом, получим точки $\chi_{лев.кр.}^2(1 - \alpha/2, n - 1)$ и $\chi_{пр.кр.}^2(\alpha/2, n - 1)$.

Если $\chi_{\text{лев.кр.}}^2 < \chi_{\text{набл}}^2 < \chi_{\text{пр.кр.}}^2$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2. При конкурирующей гипотезе $H_1: \sigma^2 > \sigma_0^2$ критическую точку $\chi_{\text{кр.}}^2(\alpha, n-1)$ правосторонней критической области находим по таблице критических точек распределения χ^2 с $n-1$ степенями свободы при вероятности α . Если $\chi_{\text{набл}}^2 < \chi_{\text{кр.}}^2$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3. При конкурирующей гипотезе $H_1: \sigma^2 < \sigma_0^2$ критическую точку $\chi_{\text{кр.}}^2(1-\alpha, n-1)$ левосторонней критической области находим по таблице критических точек распределения χ^2 с $n-1$ степенями свободы при вероятности $1-\alpha$. Если $\chi_{\text{набл}}^2 > \chi_{\text{кр.}}^2$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

6.6.4. Гипотеза о равенстве двух средних нормальных генеральных совокупностей

Пусть генеральные совокупности X_1 и X_2 распределены нормально, причём генеральные средние этих совокупностей \bar{x}_1 и \bar{x}_2 нам неизвестны. По произведённым выборкам объёмов n_1 и n_2 найдены выборочные средние \bar{x}_{B1} и \bar{x}_{B2} .

Предполагаем, что дисперсии обеих генеральных совокупностей известны, и равны σ_1^2 и σ_2^2 (этот же способ можно применять, если известны только выборочные дисперсии, но объёмы выборок достаточно большие). Требуется при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x}_1 = \bar{x}_2$.

Сравнение средних двух совокупностей имеет большое практическое значение. Часто встречается случай, когда средний результат одной серии экспериментов отличается от среднего результата другой серии. Возникает вопрос, можно ли объяснить обнаруженное расхождение средних неизбежными случайными ошибками эксперимента, или оно вызвано некоторыми закономерностями. Задача сравнения средних возникает также при выборочном контроле качества изделий, изготовленных на разных установках или при разных технологических режимах.

В качестве критерия проверки нулевой гипотезы принимают случайную величину

$$U = \frac{\bar{X}_{B1} - \bar{X}_{B2}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}, \quad (6.6.6)$$

которая имеет нормальное распределение.

Для того, чтобы при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x}_1 = \bar{x}_2$, нужно вычислить наблюдаемое значение критерия

$$U_{\text{набл}} = \frac{\bar{x}_{B1} - \bar{x}_{B2}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}. \quad (6.6.7)$$

1. При конкурирующей гипотезе $H_1: \bar{x}_1 \neq \bar{x}_2$ критическую точку $U_{\text{кр}}$ находим по таблице функции Лапласа из условия $\Phi(U_{\text{кр}}) = (1-\alpha)/2$. Критическая область в этом случае является двусторонней. Если $|U_{\text{набл}}| < U_{\text{кр}}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2. При конкурирующей гипотезе $H_1: \bar{x}_1 > \bar{x}_2$ критическую точку $U_{\text{кр}}$ правосторонней критической области находим по таблице функции Лапласа из условия $\Phi(U_{\text{кр}}) = (1-2\alpha)/2$. Если $U_{\text{набл}} < U_{\text{кр}}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3. При конкурирующей гипотезе $H_1: \bar{x}_1 < \bar{x}_2$ критическую точку $U_{\text{кр}}$ левосторонней критической области находим по таблице функции Лапласа из условия $\Phi(U_{\text{кр}}) = (1-2\alpha)/2$. Если $U_{\text{набл}} > -U_{\text{кр}}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

Приведённое выше правило можно также применять, если генеральные совокупности не имеют нормального распределения, и их дисперсии неизвестны, а выборки имеют большой объём (не менее 30 каждая) и независимы. В этом случае наблюдаемое значение критерия

$$U_{\text{набл}} = \frac{\bar{x}_{B1} - \bar{x}_{B2}}{\sqrt{s_1^{*2}/n_1 + s_2^{*2}/n_2}}. \quad (6.6.8)$$

Предположим теперь, что дисперсии обеих генеральных совокупностей неизвестны, а известны только их исправленные выборочные оценки $D_{B1}^* = s_1^{*2}$ и $D_{B2}^* = s_2^{*2}$, а выборки имеют небольшой объём (меньше 30). Предполагается, что дисперсии двух генеральных совокупностей одинаковы (например, дисперсии определяются ошибкой измерительного прибора). Если же нет оснований считать дисперсии одинаковыми, то, прежде чем сравнивать средние, следует проверить гипотезу о равенстве генеральных дисперсий, пользуясь критерием Фишера-Снедекора. В этом случае в качестве критерия проверки нулевой гипотезы принимают случайную величину

$$T = \frac{\bar{X}_{B1} - \bar{X}_{B2}}{\sqrt{\frac{(n_1 - 1)D_{B1}^* + (n_2 - 1)D_{B2}^*}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}, \quad (6.6.9)$$

имеющую распределение Стьюдента с $n_1 + n_2 - 2$ степенями свободы.

Для того, чтобы при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x}_1 = \bar{x}_2$, нужно вычислить наблюдаемое значение критерия, заменив в формуле (6.6.9) случайные величины $\bar{X}_1, \bar{X}_2, D_{B1}^*, D_{B2}^*$ их выборочными значениями.

1. При конкурирующей гипотезе $H_1: \bar{x}_1 \neq \bar{x}_2$ критическую точку $T_{кр}(\alpha, n_1 + n_2 - 2)$ находим по таблице критических точек распределения Стьюдента при $n_1 + n_2 - 2$ степенях свободы и вероятности α . Если $|T_{набл}| < T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2. При конкурирующей гипотезе $H_1: \bar{x}_1 > \bar{x}_2$ критическую точку $T_{кр}(2\alpha, n_1 + n_2 - 2)$ правосторонней критической области находим по таблице критических точек распределения Стьюдента при $n_1 + n_2 - 2$ степенях свободы и вероятности 2α . Если $T_{набл} < T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. Иначе нулевую гипотезу отвергают.

3. При конкурирующей гипотезе $H_1: \bar{x}_1 < \bar{x}_2$ критическую точку $T_{кр}(2\alpha, n_1 + n_2 - 2)$ левосторонней критической области находим по таблице критических точек распределения Стьюдента при $n_1 + n_2 - 2$ степенях свободы и вероятности 2α . Если $T_{набл} > -T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. Иначе нулевую гипотезу отвергают.

6.7. ЛИНЕЙНАЯ И НЕЛИНЕЙНАЯ РЕГРЕССИЯ. КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

В естественных науках часто идёт речь о функциональной зависимости, когда каждому значению одной переменной соответствует вполне определённое значение другой переменной. Однако на практике между переменными величинами существуют зависимости, когда каждому значению одной переменной может соответствовать множество значений другой переменной, имеющее определённое распределение. Такая зависимость называется статистической.

Статистические связи между переменными можно изучать методами корреляционного и регрессионного анализа. Основной задачей регрессионного анализа является установление формы и изучение зависимости между переменными. Основной задачей корреляционного анализа – выявление связи между случайными переменными и оценка её тесноты.

В регрессионном анализе рассматривается зависимость случайного результативного признака Y от неслучайных факторных признаков X_1, X_2, \dots, X_n . В случае единственного факторного признака X уравнение взаимосвязи двух переменных может быть представлено в виде

$$Y = \varphi(X) + \varepsilon,$$

где ε – случайная величина, причём предполагается, что её математическое ожидание $M(\varepsilon) = 0$, а дисперсия постоянна и не зависит от X . Переменную ε называют *возмущающей переменной* или просто *возмущением*.

В зависимости от вида функции $\varphi(X)$ различают следующие виды регрессий: линейную, гиперболическую, показательную, степенную, логарифмическую, параболическую и т.д. Например, уравнение линейной регрессии имеет вид

$$Y = \alpha_0 + \alpha_1 X + \varepsilon. \quad (6.7.1)$$

Предположим, что для оценки параметров регрессии взята выборка, содержащая n пар значений (x_i, y_i) , где $i = 1, 2, \dots, n$. Оценкой уравнений регрессии являются выборочные уравнения регрессии, имеющие тот же тип, что и функция $\varphi(x)$:

- *линейное:* $\hat{y} = a_0 + a_1 x;$
- *гиперболическое:* $\hat{y} = a_0 + a_1 / x;$
- *показательное:* $\hat{y} = a_0 a_1^x;$
- *степенное:* $\hat{y} = a_0 x^{a_1};$
- *логарифмическое:* $\hat{y} = a_0 + a_1 \ln x;$
- *параболическое:* $\hat{y} = a_0 + a_1 x + a_2 x^2,$

где параметры a_0, a_1, a_2 являются точечными оценками соответствующих параметров $\alpha_0, \alpha_1, \alpha_2$ исходного уравнения и могут быть найдены на основе метода наименьших квадратов.

Сущность метода наименьших квадратов заключается в нахождении параметров модели a_0, a_1 , при которых минимизируется сумма квадратов отклонений эмпирических (фактических) значений результативного признака от теоретических, полученных по выборочному уравнению регрессии:

$$S = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min.$$

Для линейной модели

$$S = \sum_{i=1}^n (y_i - a_0 - a_1 x_i)^2 \rightarrow \min. \quad (6.7.2)$$

Функция двух переменных $S(a_0, a_1)$ может достигнуть экстремума в том случае, когда её частные производные равны нулю. Вычислим эти частные производные:

$$\begin{aligned}\frac{\partial S}{\partial a_0} &= -2 \sum_{i=1}^n (y_i - a_0 - a_1 x_i) = -2 \left[\sum_{i=1}^n y_i - a_0 \cdot n - a_1 \sum_{i=1}^n x_i \right]; \\ \frac{\partial S}{\partial a_1} &= -2 \sum_{i=1}^n (y_i - a_0 - a_1 x_i) x_i = -2 \left[\sum_{i=1}^n x_i y_i - a_0 \sum_{i=1}^n x_i - a_1 \sum_{i=1}^n x_i^2 \right].\end{aligned}$$

Приравняв частные производные к нулю, получим систему уравнений для нахождения параметров a_0, a_1 линейного уравнения регрессии $\hat{y} = a_0 + a_1 x$:

$$\begin{cases} na_0 + a_1 \sum x_i = \sum y_i; \\ a_0 \sum x_i + a_1 \sum x_i^2 = \sum x_i y_i. \end{cases} \quad (6.7.3)$$

В случае, когда возмущающая переменная ε имеет нормальное распределение, коэффициенты a_0, a_1 , полученные методом наименьших квадратов для линейной регрессии, являются несмещёнными эффективными оценками параметров α_0, α_1 уравнения регрессии $y = \alpha_0 + \alpha_1 x + \varepsilon$.

Для параболического уравнения регрессии $\hat{y} = a_0 + a_1 x + a_2 x^2$ система уравнений для нахождения параметров a_0, a_1, a_2 имеет вид

$$\begin{cases} na_0 + a_1 \sum x_i + a_2 \sum x_i^2 = \sum y_i; \\ a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 = \sum x_i y_i; \\ a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 = \sum x_i^2 y_i. \end{cases} \quad (6.7.4)$$

Заметим, что все предложенные выше виды нелинейных регрессий (кроме параболической) могут быть сведены к линейной путём какой-либо замены переменной. Для гиперболической регрессии вводится переменная $x' = 1/x$, для логарифмической регрессии $x' = \ln x$, уравнения показательной и степенной регрессии предварительно логарифмируют.

Для показательного уравнения получаем $\ln \hat{y} = \ln a_0 + x \ln a_1$, далее делаем замены $\hat{y}' = \ln \hat{y}$, $a'_0 = \ln a_0$, $a'_1 = \ln a_1$. Таким образом, получаем линейное уравнение регрессии $\hat{y}' = a'_0 + a'_1 x$. Найдя коэффициенты этого уравнения методом наименьших квадратов, получим коэффициенты исходного уравнения $a_0 = e^{a'_0}$, $a_1 = e^{a'_1}$.

Для степенного уравнения получаем $\ln \hat{y} = \ln a_0 + a_1 \ln x$, далее делаем замены $x' = \ln x$, $\hat{y}' = \ln \hat{y}$, $a'_0 = \ln a_0$. Таким образом, получаем линейное уравнение регрессии $\hat{y}' = a'_0 + a_1 x'$. Найдя коэффициенты этого уравнения методом наименьших квадратов, вычислим и коэффициент исходного уравнения $a_0 = e^{a'_0}$.

Регрессионную модель удобно представлять графически. Рассмотрим выборку объёма n . Отложим на координатной плоскости точки $P_i(x_i, y_i)$, ($i = 1, 2, \dots, n$) (рис. 6.7.1). Полученный график называется *диаграммой рассеивания*. Видим, что точки группируются около некоторой прямой $y = a_0 + a_1 x$. Коэффициенты a_0, a_1 этой прямой могут быть найдены методом наименьших квадратов с помощью системы (6.8.3). На этом же графике проведём линию, соответствующую уравнению регрессии.

Отклонение точки P_i от оцениваемой линии, измеренное по вертикали, будет равно $e_i = (y_i - a_0 - a_1 x_i)$. Значения e_i ($i = \overline{1, n}$), называемые *ошибками уравнения регрессии*, являются реализациями случайной величины ε . Функцию $S(a_0, a_1)$,

определённую по (6.7.2), можно представить как $S = \sum_{i=1}^n e_i^2$, т.е. если фактические данные точно лежали бы на некоторой прямой, то $S = 0$, а чем дальше точки отклоняются от линии регрессии, тем больше значение S .

Построив диаграмму рассеивания, можно подобрать вид уравнения регрессии. На рис. 6.7.2 для одних и тех же экспериментальных точек построены линейная и показательная регрессии. Видим, что экспериментальные точки «ближе» подходят к линии $y = a_0 a_1^x$, чем к прямой. Следовательно, можно сделать вывод, что показательная регрессия более адекватно описывает фактические данные, чем линейная.

Однако по графику можно только приближённо сделать вывод о качестве той или иной модели. Существуют способы более точной оценки адекватности (значимости) уравнения регрессии. Проверить значимость уравнения регрессии – значит установить, соответствует ли математическая модель, выражающая зависимость между переменными, экспериментальным данным.

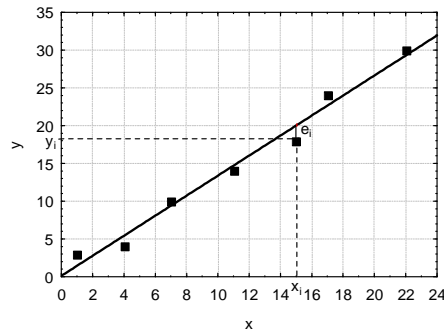


Рис. 6.7.1

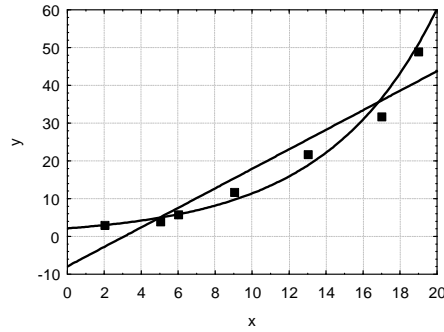


Рис. 6.7.2

Далее предполагаем, что возмущающая переменная ε имеет нормальный закон распределения. Задача проверки значимости уравнения регрессии сводится к проверке нулевой гипотезы $H_0: \alpha_1 = 0$ (напомним, что α_1 – коэффициент в уравнении регрессии (6.7.1)). Для проверки значимости уравнения регрессии на уровне значимости α вычисляют наблюдаемое значение случайной величины:

$$F_{\text{набл}} = \frac{\sigma_y^2 (n-2)}{\sigma_{\text{ост}}^2}, \quad (6.7.5)$$

где *остаточная дисперсия* $\sigma_{\text{ост}}^2$ и *дисперсия уравнения регрессии* σ_y^2 находятся по формулам:

$$\begin{aligned} \sigma_{\text{ост}}^2 &= \frac{\sum (y_i - \hat{y}_i)^2}{n-1}; \\ \sigma_y^2 &= \frac{\sum (\bar{y}_B - \hat{y}_i)^2}{n-1}, \end{aligned} \quad (6.7.6)$$

где $\bar{y}_B = \frac{\sum y_i}{n}$, $\hat{y}_i = a_0 + a_1 x$. Учитывая смысл величин $\sigma_{\text{ост}}^2$ и σ_y^2 , можно сказать, что значение $F_{\text{набл}}$ показывает, насколько лучше уравнение регрессии оценивает значение резульативного признака по сравнению с его средней. Далее находят критическое значение критерия $F(\alpha, 1, n-2)$ по таблице критических точек распределения Фишера при $k_1 = 1$, $k_2 = n-2$ степенях свободы и уровне значимости α . Если $F_{\text{набл}} > F(\alpha, 1, n-2)$, то нулевая гипотеза отвергается и уравнение регрессии признаётся значимым. В противном случае уравнение регрессии признаётся незначимым, т.е. статистически подтверждается отсутствие линейной связи между факторным и резульативным признаком.

По расположению точек на диаграмме рассеяния не всегда можно принять окончательное решение о виде уравнения регрессии. Если теоретические соображения не могут подсказать точного решения, то следует сделать расчёты по двум или более уравнениям. Предпочтение следует отдать той модели, для которой меньше величина остаточной дисперсии, или, что то же самое, меньше величина остаточной суммы квадратов $S = \sum e_i^2 = \sum (y_i - \hat{y}_i)^2$. Однако при незначительных расхождениях в остаточных дисперсиях следует остановиться на более простом уравнении.

Рассмотрим более подробно линейное уравнение регрессии. Коэффициент a_1 показывает величину изменения резульативного признака при изменении факторного признака на единицу. Следовательно, коэффициент a_1 зависит от единиц измерения переменных. Поэтому в качестве универсального показателя тесноты связи между величинами x и y используется *выборочный линейный коэффициент корреляции*

$$r = a_1 \frac{s_x}{s_y}. \quad (6.7.7)$$

Здесь s_x и s_y – исправленные средние квадратические отклонения соответствующих признаков (факторного и резульативного). Они вычисляются по формулам:

$$s_x^2 = \frac{\sum (x_i - \bar{x}_B)^2}{n-1}; \quad s_y^2 = \frac{\sum (y_i - \bar{y}_B)^2}{n-1}, \quad (6.7.8)$$

где \bar{x}_B, \bar{y}_B – выборочные средние значения.

Линейный коэффициент корреляции изменяется в пределах от -1 до 1 , т.е. $-1 \leq r \leq 1$. Если $r > 0$, то связь между переменными x и y прямая, если $r < 0$, то связь между переменными x и y обратная. В зависимости от того, насколько $|r|$ приближается к 1 , различают связь слабую, умеренную и сильную. При $|r| < 0,2$ связь между переменными практически отсутствует. При $|r| = 1$ связь между x и y функциональная, т.е. наблюдаемые значения располагаются точно на прямой.

Выборочный коэффициент корреляции r является оценкой генерального коэффициента корреляции ρ , который определяется по формуле

$$\rho = \frac{M(XY) - M(X)M(Y)}{\sigma_x \sigma_y}.$$

Величина ρ характеризует тесноту связи между величинами X и Y в генеральной совокупности. Однако в практических исследованиях о тесноте корреляционной зависимости между переменными судят фактически не по величине генерального коэффициента корреляции ρ , а по величине его выборочного аналога r .

Пусть вычисленное значение $r \neq 0$. Возникает вопрос, объясняется ли это действительно существующей линейной связью между переменными X и Y , или является следствием случайности отбора переменных в выборку. Обычно в этих случаях проверяется гипотеза H_0 об отсутствии линейной корреляционной связи между переменными. Иначе говоря, H_0 предполагает, что $\rho = 0$ при альтернативной гипотезе H_1 , состоящей в том, что $\rho \neq 0$. Для проверки этой гипотезы на уровне значимости α вычисляют наблюдаемое значение критерия

$$T_{\text{набл}} = \frac{|r|}{\sqrt{1-r^2}} \sqrt{n-2}. \quad (6.7.9)$$

Критическое значение критерия $T(1-\alpha, n-2)$ находят по таблице критических точек распределения Стьюдента для числа степеней свободы $n-2$ и уровня значимости α . Если $T_{\text{набл}} < T(\alpha, n-2)$, то гипотеза H_0 принимается, в противном случае гипотеза H_0 отвергается, т.е. коэффициент корреляции признается существенно отличающимся от нуля.

ЗАКЛЮЧЕНИЕ

Прикладная направленность математических знаний, умений и навыков являются существенной особенностью естественно-математической подготовки современного инженера. При этом числовые и функциональные ряды, вероятностно-статистические понятия, факты и методы наиболее востребованы как в процессе овладения общеинженерными и специальными дисциплинами, так и в практической деятельности специалиста, неотъемлемыми составляющими которой являются прогнозирование, проектирование, математическое моделирование.

В результате изучения основных положений курса прикладной математики, представленных в настоящем пособии, студент:

- приобретает знания в области комплексного анализа;
- овладевает способами представления функций степенными рядами;
- получает возможность использования различных вероятностных моделей в количественных оценках шансов наступления тех или иных событий;
- знакомится с основными видами распределений случайных величин и их практическим использованием;
- осваивает методы анализа выборочных совокупностей, статистической проверки гипотез и др.

Материал пособия, как надеются авторы, внесёт существенный вклад в развитие математической компетентности специалиста, приобщение его к методам научного познания, овладение общими логическими приёмами мышления, необходимыми как в профессиональной, так и повседневной деятельности.

СПИСОК ЛИТЕРАТУРЫ

1. Захаров, В.К. Теория вероятностей / В.К. Захаров, Б.А. Севостьянов – М. : Наука, 1983. – 166 с.
2. Венцель, Е.С. Теория вероятностей / Е.С. Венцель. – М. : Высшая школа, 1999. – 576 с.
3. Куликов, Г.М. Метод Фурье в уравнениях математической физики / Г.М. Куликов, А.Д. Нахман. – М. : Машиностроение, 2000. – 156 с.
4. Нахман, А.Д. Функции комплексного переменного : методические рекомендации и контрольные задания / А.Д. Нахман, Е.А. Петрова. – Тамбов : Изд-во тамб. гос. техн. ун-та, 2007. – 40 с.
5. Плотникова, С.В. Математическая статистика : методические разработки и контрольные задания / С.В. Плотникова. – Тамбов : Изд-во тамб. гос. техн. ун-та, 2005. – 52 с.
6. Сидоров, Ю.В. Лекции по теории функций комплексного переменного / Ю.В. Сидоров, М.В. Федорюк, М.И. Шабунин. – М. : Наука, 1989. – 478 с.
7. Гмурман, В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. – М. : Высшая школа, 2001. – 479 с.
8. Кремер, Н.Ш. Теория вероятностей и математическая статистика / Н.Ш. Кремер. – М. : Юнити-Дана, 2006. – 573 с.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	3
1. ФУНКЦИИ КОМПЛЕКСНОГО ПЕРЕМЕННОГО	4
1.1. Комплексные числа	4
1.2. Функции комплексного переменного. Предел. Непрерывность	7
1.3. Производная	12
1.4. Интеграл от функции комплексного переменного	15
1.5. Интегральная теорема Коши	18
1.6. Формула Коши	20
2. ЧИСЛОВЫЕ И ФУНКЦИОНАЛЬНЫЕ РЯДЫ	22
2.1. Числовые ряды. Основные понятия. Простейшие свойства	22
2.2. Необходимый признак сходимости ряда. Сумма геометрической прогрессии	26
2.3. Сходимость рядов с положительными членами: признаки сравнения	28
2.4. Сходимость рядов с положительными членами: признаки Коши и Даламбера	30
2.5. Сходимость рядов с положительными членами: интегральный признак Коши	32
2.6. Знакопеременные ряды	35
2.7. Абсолютная и условная сходимость знакопеременных рядов с действительными членами. Абсолютная и условная сходимость рядов с комплексными членами	37
2.8. Равномерная сходимость функционального ряда	40
2.9. Степенные ряды	44
2.10. Степенные ряды: случай действительного переменного	48
3. РАЗЛОЖЕНИЕ ФУНКЦИЙ В СТЕПЕННЫЕ РЯДЫ	51
3.1. Ряд Маклорена	51
3.2. Ряд Тейлора	54
3.3. Ряд Лорана	56
3.4. Изолированные особые точки	61
4. ТЕОРИЯ ВЕРОЯТНОСТЕЙ. СЛУЧАЙНЫЕ СОБЫТИЯ	66
4.1. Элементы комбинаторики	66
4.2. Основные понятия теории вероятностей	67
4.3. Пространство элементарных событий	68
4.4. Операции над событиями	69
4.5. Аксиомы теории вероятностей	71
4.6. Конечное вероятностное пространство	73
4.7. Счётное вероятностное пространство	74
4.8. Непрерывное вероятностное пространство	75
4.9. Условная вероятность. Независимость событий	77
4.10. Формула полной вероятности. Формула Байеса	79
4.11. Схема Бернулли	81
4.12. Предельные теоремы в схеме Бернулли	82
5. ТЕОРИЯ ВЕРОЯТНОСТЕЙ. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ	84
5.1. Понятие случайной величины. Дискретные случайные величины. Закон распределения случайной величины. Функция распределения случайной величины	84
5.2. Непрерывные случайные величины. Плотность распределения непрерывной случайной величины	88
5.3. Математические операции над случайными величинами ...	91
5.4. Математическое ожидание и дисперсия случайной величины	94
5.5. Основные законы распределения дискретных случайных величин	99
5.6. Основные законы распределения непрерывных случайных величин	105
5.7. Закон больших чисел	116
6. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА	119
6.1. Введение. Основные понятия математической статистики	119
6.2. Вариационные ряды и их графическое изображение	121
6.3. Основные числовые характеристики вариационного ряда	125
6.4. Оценки параметров	129
6.5. Интервальное оценивание. Доверительные интервалы	136
6.6. Проверка статистических гипотез	141
6.7. Линейная и нелинейная регрессия. Корреляционный ана-	150

ЛИЗ	
ЗАКЛЮЧЕНИЕ	157
СПИСОК ЛИТЕРАТУРЫ	158