Задание

Уважаемые студенты, на предыдущих занятиях вы изучили понятие дискретной случайной величины, ее способы задания и числовые характеристики. В этой лекции Вы познакомитесь с непрерывной случайной величиной, ее способами задания и числовыми характеристиками. В лекции представлена информация и о дискретной случайной величине и о непрерывной случайной величине.

Вам необходимо прочесть лекцию, записать определение непрерывной случайной величины, определение функции распределения, ее свойства, определение плотности распределения, составить сравнительную таблицу числовых характеристик случайных величин.

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

План

1. **Случайные величины и способы их задания.**
2. **Характеристики случайных величин.**
3. **Основные законы распределения случайных величин.**

<u>Цели занятия:</u> закрепить знание дискретной случайной величины, дать определение непрерывной случайной величины, выполнить сравнительный анализ случайных величин; вырабатывать интерес к учебной дисциплине; воспитывать внимательность, усидчивость.

Достаточно часто на практике рассматриваются такие испытания, в результате реализации которых случайным образом получается некоторое число.

Например, число родившихся мальчиков (или девочек) среди ста новорожденных, время безотказной работы некоторого медицинского прибора, температура воздуха на следующий день, число появлений герба при четырех бросаниях монеты, температура воздуха на следующий день и т.д.

Числовая величина, принимающая то или иное значение в результате реализации испытания случайным образом, называется *случайной величиной.*

Понятие случайной величины играет весьма важную роль в теории вероятностей. Если «классическая» теория вероятностей изучала главным образом случайные события, то современная теория вероятностей преимущественно имеет дело со случайными величинами.

Сами случайные величины обозначаются прописными латинскими буквами X, Y, Z и т.д., а их возможные значения – соответствующими строчными x, y, z. Например, если случайная величина имеет три возможных значения, то будем обозначать их так: x_1, x_2, x_3.

Если случайная величина принимает конечное или счетное множество значений, то она называется *дискретной (дискретно распределенной)*.

Непрерывной случайной величиной называется такая случайная величина, которая может принимать все значения из некоторого конечного или бесконечного промежутка.

Для задания случайной величины необходимо знать не только, какие значения может она принимать, но и как часто, то есть с какой вероятностью она принимает эти значения. Для того, чтобы задавать вероятности значений случайных величин, столь разнообразных по своей природе, и притом задавать их одним и тем же способом, вводят понятие функции распределения вероятностей случайной величины или просто функции распределения случайной величины.

Функцией распределения называют функцию $F(x)$, определяющую вероятность того, что случайная величина X в результате испытания примет значение, меньшее x, то есть $F(x) = P\{X < x\}$.

Иногда вместо термина “функция распределения” используют термин “интегральная функция ”.

Свойства интегральной функции распределения:

1. $0 \leq F(x) \leq 1$.
2. $F(x_1) \leq F(x_2)$при$x_1 < x_2$, т.е. $F(x)$ – неубывающая функция.
3. $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$; $F(+\infty) = \lim\limits_{x \to +\infty} F(x) = 1$.
4. $P\{x_1 \leq X < x_2\} = F(x_2) - F(x_1)$.
5. $F(x) = F(x - 0)$, где $F(x - 0) = \lim\limits_{y \to x - 0} F(y)$, т.е.$F(x)$- непрерывная слева функция.

Таким образом, каждая функция распределения удовлетворяет свойствам $1 - 5$. Верно и обратное: каждая функция, удовлетворяющая свойствам $1 - 5$, может рассматриваться как функция распределения некоторой случайной величины.

Непрерывную случайную величину можно задать с помощью функции распределения $F(x)$. Этот способ задания не является единственным. Непрерывную случайную величину можно также задать, используя другую функцию, которую называют *плотностью распределения* или *плотностью вероятности* (иногда её называют дифференциальной функцией).

Плотность распределения случайной величины – производная от функции распределения$F(x)$:

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

Из этого определения следует, что функция распределения является первообразной для плотности распределения. Заметим, что для описания распределения вероятностей дискретной случайной величины плотность распределения неприменима.

Свойства плотности вероятности

Дифференциальная функция распределения – неотрицательная, т.е. $f(x) \geq 0$.

Площадь фигуры, ограниченной кривой распределения и осью абсцисс, равна единице.

Вероятность попадания непрерывной случайной величины в интервал $[a; b]$равна определенному интегралу от ее плотности вероятности в пределах от a до b:

$$P(a \leq x \leq b) = \int_a^b f(x) \cdot dx.$$

Несобственный интеграл от плотности распределения в пределах от$-\infty$ до ∞ равен единице:

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 1.$$

Последнее равенство называется также *условием нормировки непрерывной случайной величины*.

Пример 1. Задана плотность вероятности случайной величины X

$$f(x) = \begin{cases} 0 & \text{при} \quad x \leq 0, \\ 2x & \text{при} \quad 0 < x \leq 1, \\ 0 & \text{при} \quad x > 1. \end{cases}$$

Найти вероятность того, что в результате испытания X примет значение, принадлежащие интервалу$(0{,}5; 1)$.

Решение: Искомая вероятность

$$P(0{,}5 < X < 1) = 2\int_{0,5}^{1} x \cdot dx = x^2|_{0,5}^{1} = 1 - 0{,}25 = 0{,}75.$$

Зная плотность распределения $f(x)$, можно найти функцию распределения $F(x)$по формуле

$$F(x) = \int_{-\infty}^{x} f(x) \cdot dx.$$

Действительно,$F(x) = P(X < x) = P(-\infty < X < x)$.

Следовательно,

$$P(-\infty < X < x) = \int_{-\infty}^{x} f(x) \cdot dx \text{или } F(x) = \int_{-\infty}^{x} f(x) \cdot dx.$$

Таким образом, *зная плотность распределения, можно найти функцию распределения.*

Пример 2. Найти функцию распределения по заданной плотности распределения:

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

Решение: Воспользуемся формулой$F(x) = \int_{-\infty}^{x} f(x) \cdot dx$

Если$x \leq a$, то $f(x) = 0$, следовательно,$F(x) = 0$. Если$a < x \leq b$, то $f(x) = \frac{1}{b-a}$, следовательно,

$$F(x) = \int_{-\infty}^{x} f(x) \cdot dx = \int_{-\infty}^{a} 0 \cdot dx + \int_a^x \frac{1}{b-a} dx = \frac{x-a}{b-a}.$$

Если $x > b$, то

$$F(x) = \int_{-\infty}^{a} 0 \cdot dx + \int_a^b \frac{dx}{b-a} + \int_b^x 0 \cdot dx = \frac{b-a}{b-a} = 1.$$

Итак, искомая функция распределения

$$F(x) = \begin{cases} 0 \quad \text{при} \quad x \le a, \\ (x-a)/(b-a) \text{ при} \quad a < x \le b, \\ 1 \quad \text{при} \quad x > b. \end{cases}$$

Пример 3. Плотность распределения случайной величины имеет следующий вид:

$$f(x) = \begin{cases} asinx, 0 \le x \le \pi \\ 0, x < 0, x > \pi \end{cases}.$$

Найти постоянный параметр a.

Решение:Плотность распределения должна удовлетворять условию нормировки:

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 1.$$

Поэтому потребуем, чтобы выполнялось равенство

$$\int_0^{\pi} asinx \cdot dx = 1.$$

Отсюда$a = \frac{1}{\int_0^{\pi} sinx \cdot dx}$. Найдём определённый интеграл:

$$\int_0^{\pi} sinx \cdot dx = -cosx_0^{\pi} = -(cos\pi - cos0) = -(-1-1) = 2.$$

Таким образом, искомый параметр

$$\alpha = \frac{1}{2}.$$

Следовательно, функция плотности вероятности примет вид:

$$f(x) = \begin{cases} 2sinx, 0 \le x \le \pi \\ 0, x < 0, \quad x > \pi \end{cases}$$

Способы задания дискретной случайной величины

Для задания дискретной случайной величины достаточно задать семейство вероятностей $p_i = P(X = x_i)$, где$i = \overline{1,n}$или$i = \overline{1,\infty}$.

Законом распределения дискретной случайной величины называют соответствие между возможными значениями и их вероятностями.

Закон распределения можно задать в виде таблицы, формулы или графически. При табличном задании закона распределения в первой строке таблицы перечислены все значения случайной величины в порядке возрастания, а в нижней – соответствующие им вероятности.

X	x_1	x_2	x_3	...	x_n
P	p_1	p_2	p_3	...	p_n

Такую таблицу будем называть *рядом распределения* случайной величины X. Поскольку все возможные значения можно рассматривать как полную группу несовместных событий, то сумма соответствующих вероятностей обязательно должна быть равна единице (условие нормировки):

$$p_1 + p_2 + p_3 + \cdots p_n = \sum_{i=1}^{n} p_i = 1$$

Для наглядности ряд распределения случайной величины можно изобразить графически. Для этого в прямоугольной системе координат по оси OX откладывают значения случайной величины, а по оси ординат OY – соответствующие им вероятности. Полученные точки соединяются отрезками прямых. Построенная таким образом фигура называется *многоугольником или полигоном распределения* вероятностей.

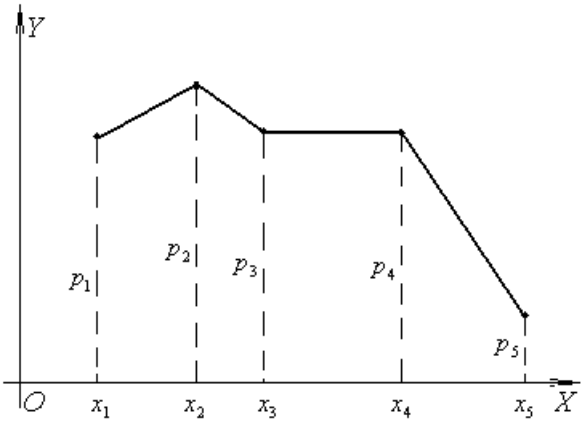

Рис. 1. Многоугольник или полигон распределения вероятностей.

Многоугольник распределения, также как и ряд распределения, полностью характеризует случайную величину. Он является одним из форм закона распределения.

Задать закон распределения дискретной случайной величины можно в виде функции распределения вероятностей (интегральной функции распределения) $F(x)$, где$F(x) = P(X < x) = \sum_{i:x_i<x} P(X = x_i)$.

График функции распределения дискретной случайной величины имеет ступенчатый вид:

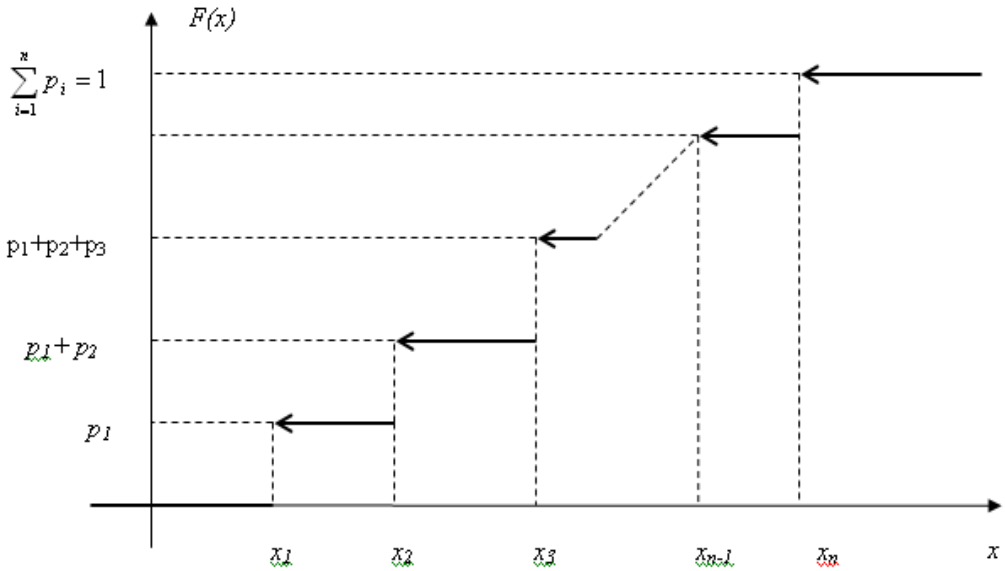

Рис.2 График функции распределения дискретной случайной величины

Пример 4. В денежной лотерее выпущено 100 билетов. Разыгрывается один выигрыш в 10000 рублей и десять выигрышей по 1000 рублей. Найти ряд распределения, функцию распределения случайной величины X – стоимости возможного выигрыша для владельца одного лотерейного билета. Построить многоугольник распределения.

Решение: Случайная величина X принимает значения 0,1000,10000 с вероятностями: $P(X=0)=\frac{89}{100}=0{,}89$, $P(X=1000)=\frac{10}{100}=0{,}1$, $P(X=10000)=\frac{1}{100}=0{,}01$.

Таким образом, ряд распределения имеет следующий вид:

X	0	1000	10000
P	0,89	0,1	0,01

Условие нормировки выполняется:

$$\sum_{i=1}^{3} p_i = 0{,}89 + 0{,}1 + 0{,}01 = 1.$$

Найдем функцию распределения данной случайной величины X:

Если $x \le 0$, то $F(x)=0$. Если $0 < x \le 100$, то $F(x)=0{,}89$. Действительно, X может принять значение 0 с вероятностью 0,89 . Если $1000 < x \le 10000$, то$F(x)=0{,}99$.

Действительно, если x_1 удовлетворяет неравенству $1000 < x \le 10000$, то $F(x_1)$ равно вероятности события $X < x_1$, которое может быть осуществлено, когда X примет значение 0 с вероятностью 0,89 или 1000 с вероятностью 0,1. Поскольку эти два события несовместны, то по теореме сложения вероятность события $X < x_1$ равна сумме вероятностей $0{,}89 + 0{,}1 = 0{,}99$. Если $x > 10000$, то $F(x)=1$. Итак, функция распределения аналитически может быть записана следующим образом:

$$F(x)=\begin{cases} 0, & \text{если } x \le 0, \\ 0{,}89 & \text{если } 0 < x \le 1000, \\ 0{,}99 & \text{если } 1000 < x \le 10000, \\ 1 & \text{если } x > 10000. \end{cases}$$

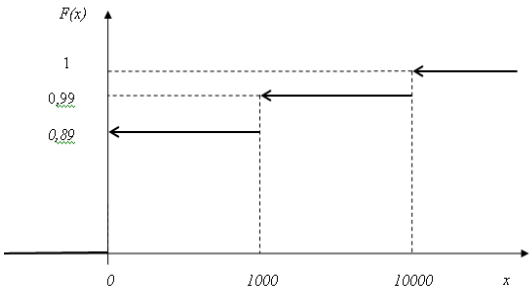

Рис.3 График функции распределения

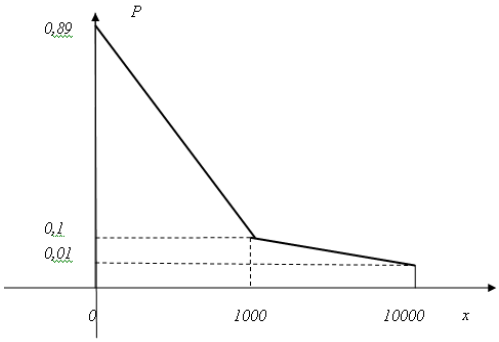

Рис. 4 Многоугольник распределения

Характеристики случайных величин.

Числовые характеристики случайных величин количественно определяют различные свойства случайных величин. Они позволяет проводить сравнительный анализ случайных величин, давать оценку ожидаемым результатам опыта, находить связь и определять зависимость между различными случайными величинами и многое другое. К числовым характеристикам случайной величины относятся характеристики положения и характеристики разброса.

Из характеристик положения в теории вероятностей важнейшую роль играет *математическое ожидание* случайной величины, которое иногда называют просто *средним значением* случайной величины.

Рассмотрим дискретную случайную величинуX, имеющую возможные значения$x_1, x_2, \dots, x_n$ с вероятностями $p_1, p_2, \dots, p_n$.

Математическое ожидание для дискретной случайной величины – сумма произведений всех возможных значений случайной величины на вероятности этих значений или среднее взвешенное значение случайной величины:

$$M(x) = x_1p_1 + x_2p_2 + \dots + x_np_n = \sum_{i=1}^{n} x_ip_i.$$

Математическое ожидание удовлетворяет следующим свойствам:

1) $M(x_1 + x_2) = M(x_1) + M(x_2)$;

2) $M(kx) = kM(x)$для всех k;

3) если $P(x_1 + x_2) = 1$, то $M(x_1) = M(x_2)$;

4) если$x \geq 0$, то$M(x) \geq 0$;

5) если$P\{x = c\} = 1$, то$M(x) = c$.

Для вычисления математического ожидания непрерывной случайной величины используются формулы:

$$M(x) = \int_{-\infty}^{\infty} xdF(x) \text{или} \qquad M(x) = \int_{-\infty}^{\infty} xf(x)dx$$

Пример 5. Случайная величина представлена законом распределения:

x_i	3	5	2
p_i	0.1	0.6	0.3

Определитьматематическоеожидание.

Решение:

По определению $M(x) = \sum_{i=1}^{n} x_ip_i$, следовательно, математическое ожидание будет равно:

$$M(x) = 3 \cdot 0{,}1 + 5 \cdot 0{,}6 + 2 \cdot 0{,}3 = 3{,}9.$$

Модой Mo называют значение случайной величины, имеющее у дискретной величины наибольшую вероятность, а у непрерывной – наибольшую плотность вероятности.

Если кривая распределения имеет один максимум, то мода равна значению случайной величины, соответствующей этому максимуму. Такая кривая называется унимодальной (одномодальной). Если кривая распределения имеет два или несколько случайной величины одинаковых

максимумов, то она соответственно называется двухмодульной, или многомодальной.

Пример 6.Распределение пациентов по палатам характеризуется следующим образом:

Номер палаты	1	2	3	4	5	6	7	8
Количество пациентов	1	4	2	3	7	8	5	9

Определить моду данной случайной величины.

Решение:

В этом ряду распределения модой является 8 палата, т.е. $Mo = 9$.

Пример 7. Найти моду непрерывной случайной величины, представленной функцией распределения:

$$f(x) = ae^{2x-x^2}, (a > 0)$$

Решение: По определению медианы непрерывной случайной величины, необходимо определить максимальное значение функции плотности вероятности, т.е. найти экстремум функции $f(x)$.

Находим производную $f'(x)$ и прораниваем ее к нулю:

$$f'(x) = 2a(1-x)e^{2x-x^2} = 0.$$

Из последнего выражения получаем $(1-x) = 0$, следовательно$x = 1$.

Убедимся, что в точке $x = 1$ действительно максимум, для этого найдем вторую производную:

$$f''(x) = 2ae^{2x-x^2} + 4a(1-x)^2e^{2x-x^2} = 0.$$

Определим значение второй производной в точке $x = 1$:

$$f''(x) = 2ae.$$

Вторая производная $f''(x)$ в точке $x = 1$ всегда меньше нуля, т.к. $a > 0$ и $e > 0$, следовательно, в точке $x = 1$ функция имеет максимум. Таким образом $Mo = 1$.

Медианой случайной величины называют такое ее значение Me, для которого функция распределения равна $0{,}5$. Это означает, что вероятность случайной величины принять значение меньше медианы в точности равна вероятности этой величины принять значение, большее медианы.

Для непрерывной случайной величины медиана определяется из соотношения

$$\int_{-\infty}^{Me} f(x)dx = \int_{Me}^{\infty} f(x)dx$$

Геометрически медиана представляет собой абсциссу точки, которая делит площадь, ограниченную кривой распределения пополам.

Для дискретной случайной величины необходимо расположить ее значения в порядке возрастания и в качестве медианы принять такое срединное значение, для которого выполняется условие:

$$\sum_{i=1}^{m-1} p(x_i) = \sum_{i=m}^{n} p(x_i)$$

Пример 8. Определить медиану: 1) Стаж пяти рабочих составил: 2, 4, 7, 9 и 10 лет. 2) Стаж работы шести рабочих составил 1, 3, 4, 5, 10 и 11лет.

Решение:

В первом ряду медиана равна 7, т.е. $Me = 7$

Дискретный упорядоченный ряд состоит из чётного числа членов, следовательно, медианой будет средняя арифметическая из двух смежных вариант, стоящих в центре ряда, это варианты 4 и 5. Средняя арифметическая из этих значений и будет медианой ряда:

$$Me = \frac{4+5}{2} = 4{,}5$$

Наряду с характеристиками положения используются числовые характеристики, по которым судят о рассеивании случайной величины. К ним, в частности, относят дисперсию и среднеквадратическое отклонение.

Дисперсией случайной величины X называют математическое ожидание квадрата отклонения случайной величины от ее математического ожидания:

$$D(x) = M\big(x - M(x)\big)^2 \quad \text{или } D(x) = M(x^2) - [M(x)]^2$$

Дисперсия непрерывной случайной величины определяется выражением:

$$D(x) = \int_{-\infty}^{\infty} x^2 f(x)dx - [M(x)]^2.$$

Дисперсия характеризует рассеивание значений случайной величины, в окрестности ее математического ожидания. Само слово «дисперсия» означает рассеивание.

Свойства дисперсии:

1. Дисперсия постоянной величины равна нулю:

$$D(C) = 0,\ C\ - const.$$

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

$$D(kx) = k^2 D(x)$$

3. Дисперсия алгебраической суммы независимых случайных величин равна сумме дисперсий слагаемых:

$$M(x_1 \pm x_2 \pm \cdots \pm x_n) = D(x_1) + D(x_2) + \cdots + D(x_n)$$

Наряду с дисперсией в качестве меры рассеивания широко используется среднее квадратичное отклонение случайной величиныx.

Средним квадратичным отклонением случайной величиныx называют квадратный корень из дисперсии:

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

Дисперсия имеет размерность, равную квадрату размерности случайной величины. Поэтому в тех случаях, когда желательно, чтобы оценка рассеяния имела размерность случайной величины, вычисляют среднее квадратическое

отклонение, а не дисперсию. Например, еслиX выражается в метрах, то $\sigma(x)$ будет выражаться тоже в метрах, а $D(x)$ – в квадратных метрах.

Пример 9. Дискретная случайная величина представлена законом распределения:

x_i	3	5	2
p_i	0,1	0,6	0,3

Определить дисперсию и среднее квадратическое отклонение.

Решение:

Найдем сначала математическое ожидание:

$$M(x) = \sum_{i=1}^{n} x_i p_i = 3 \cdot 0{,}1 + 5 \cdot 0{,}6 + 2 \cdot 0{,}3 = 3{,}9$$

Тогда дисперсия будет равна:

$$D(x) = M\big(x - M(x)\big)^2 p_i = (3 - 3{,}9)^2 \cdot 0{,}1 + (5 - 3{,}9)^2 \cdot 0{,}6 + (2 - 3{,}9)^2 \cdot 0{,}3 = 1{,}89.$$

Находим среднее квадратическое отклонение:

$$\sigma(x) = \sqrt{D(x)} = \sqrt{1{,}89} = 1{,}37$$

Пример 10. Найти математическое ожидание, дисперсию и среднее квадратическое отклонение непрерывной случайной величины, заданной функцией распределения:

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

Решение:

Найдем дифференциальную функцию распределения:

$$f(x) = F'(x) = \begin{cases} 0, & x \le 0 \\ 2x, & 0 < x \le 1 \\ 0, & x > 1 \end{cases}$$

Определяем математическое ожидание:

$$M(x) = \int_0^1 x \cdot 2x \cdot dx = 2\int_0^1 x^2 dx = 2\left.\frac{x^3}{3}\right|_0^1 = \frac{2}{3}(1^3 - 0^3) = \frac{2}{3}.$$

Дисперсия непрерывной величины определяется выражением:

$$D(x) = \int_{-\infty}^{\infty} x^2 f(x) dx - [M(x)]^2.$$

Подставляя численные значения, получаем:

$$D(x) = \int_0^1 x^2 \cdot 2x \cdot dx - \left(\frac{2}{3}\right)^2$$

$$= 2\int_0^1 x^3 \cdot dx - \frac{4}{9} = 2 \cdot \left.\frac{x^4}{4}\right|_0^1 = 2\left(\frac{1}{4} - 0\right) - \frac{4}{9} = \frac{1}{18}$$

Среднее квадратическое отклонение равно: $\sigma(x) = \sqrt{D(x)} = \sqrt{\frac{1}{18}} = 0{,}236.$

Обобщением основных числовых характеристик случайных величин является понятие моментов случайной величины.

*Моменты случайных величин*служат для описания свойств плотности распределения случайной величины. Моменты содержат меньше информации о случайной величине по сравнению с плотностью распределения, но часто более удобны при решении прикладных задач.

Начальным моментом порядка k случайной величины X называют математическое ожидание величины X^k, т.е.

$$\nu_k = M(X^k).$$

В частности, $\nu_1 = M(X)$, $\nu_2 = M(X^2)$. Следовательно,$D(X) = \nu_2 - \nu_1^2$.

Центральным моментом порядка k случайной величины X называют математическое ожидание величины $\left(X - M(X)\right)^k$:

$M_k = M[X - M(X)]^k$или $M_k = \sum_{i=1}^{n}\left(x_i - M(x)\right)^k p_i$

В частности, $M_1 = 0$,$M_2 = D(X)$. Следовательно,$M_2 = \nu_2 - \nu_1^2$.

Для непрерывных случайных величин центральный момент -го порядка определяется выражением:

$$M_k = \int_{-\infty}^{\infty}\left(x - M(x)\right)^k f(x)dx.$$

Обычно для практических целей ограничиваются вычислением моментов не выше четвертого порядка.

Асимметрией называют отношение центрального момента третьего порядка к кубу среднего квадратического отклонения:

$$A_S(X) = \frac{M_3(X)}{\sigma(X)^3}.$$

Асимметрия положительна, если “длинная часть” кривой распределения расположена справа от математического ожидания; асимметрия отрицательна, если “длинная часть” кривой расположена слева от математического ожидания. Практически определяют знак асимметрии по расположению кривой распределения относительно моды.

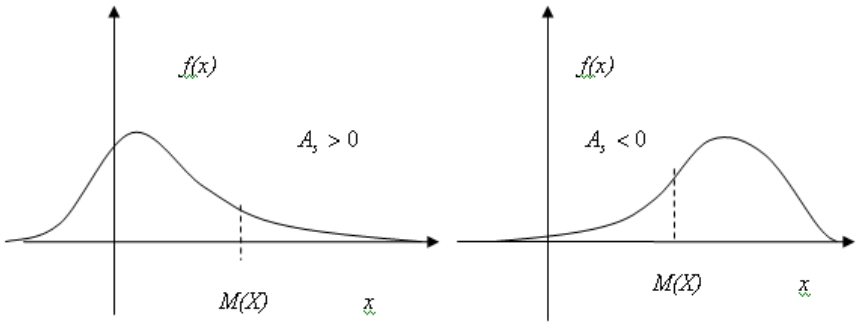

Рис. 5

Эксцессом называют характеристику, которая определяется равенством:

$$E_k(X) = \frac{M_4(X)}{\sigma(X)^4} - 3.$$

Для нормального распределения эксцесс равен нулю. Следовательно, если эксцесс некоторого распределения отличен от нуля, то кривая этого распределения отличается от нормальной кривой: если эксцесс положительный, то кривая имеет более высокую и “острую” вершину, чем нормальная кривая; если эксцесс отрицательный, то сравниваемая кривая имеет более низкую и “плоскую” вершину, чем нормальная кривая:

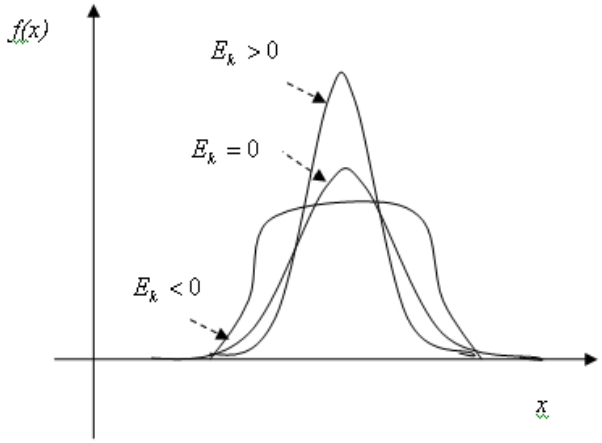

Рис. 6

Пример 11. Определить характеристики положения, характеристики разброса, асимметрию и эксцесс дискретной случайной величины, представленной законом распределения:

x_i	0	1	2	3
p_i	0,1	0,3	0,4	0,2

Решение:

Найдем характеристики положения.

Математическое ожидание дискретной случайной величины:

$$M(x) = \sum_{i=1}^{n} x_i p_i = 0 \cdot 0{,}1 + 1 \cdot 0{,}3 + 2 \cdot 0{,}4 + 3 \cdot 0{,}2 = 1{,}7$$

Мода случайной величины $Mo = X_i(p_{max}) = 2$.

Медиана дискретной случайной величины: $Me = \frac{1+2}{2} = 1{,}5$.

Найдем характеристики разброса.

Дисперсия дискретной случайной величины:

$$D(x) = \left(x - M(x)\right)^2 p_i = (0 - 1{,}7)^2 \cdot 0{,}1 + (1 - 1{,}7)^2 \cdot 0{,}3 + (2 - 1{,}7)^2 \cdot 0{,}4 + (3 - 1{,}7)^2 \cdot 0{,}2 = 0{,}81.$$

Среднее квадратическое отклонение:

$$\sigma(x) = \sqrt{D(x)} = \sqrt{0{,}81} = 0{,}9$$

Определим коэффициент асимметрии:

$$A_S(X) = \frac{M_3(X)}{\sigma(X)^3}$$

Найдем сначала третий центральный момент:

$$M_3 = \sum_{i=1}^{n} \left(x - M(x)\right)^3 p_i =$$
$$= (0 - 1{,}7)^3 \cdot 0{,}1 + (1 - 1{,}7)^3 \cdot 0{,}3 + (2 - 1{,}7)^3 \cdot 0{,}4 + (3 - 1{,}7)^3 \cdot 0{,}2 = -0{,}4913 - 0{,}1029 + 0{,}0108 + 0{,}4394 = -0{,}144$$

Тогда коэффициент асимметрии будет равен:

$$A_S(X) = \frac{M_3(X)}{\sigma(X)^3} = \frac{-0{,}144}{0{,}9^3} = -0{,}198$$

Находим эксцесс:

$$E_k(X) = \frac{M_4(X)}{\sigma(X)^4} - 3.$$

Найдем сначала четвертый центральный момент:

$$M_4 = \sum_{i=1}^{n} \left(x - M(x)\right)^4 p_i = (0 - 1{,}7)^4 \cdot 0{,}1 + (1 - 1{,}7)^4 \cdot 0{,}3 + (2 - 1{,}7)^4 \cdot 0{,}4 + (3 - 1{,}7)^4 \cdot 0{,}2 = 0{,}8352 + 0{,}0720 + 0{,}0324 + 0{,}5712 = 1{,}4817$$

Окончательно величина эксцесса будет равна:

$$E_k(X) = \frac{M_4(X)}{\sigma(X)^4} - 3 = \frac{1{,}4817}{0{,}9^4} - 3 = 2{,}583 - 3 = -0{,}7417$$

Основные законы распределения случайных величин

Наиболее полно случайная величина может быть охарактеризована законом распределения случайной величины в виде функции распределения или плотности распределения.

В зависимости от характера самих объектов имеют место следующие наиболее распространенные законы распределения случайных величин:

1) биномиальный закон распределения;
2) распределение Пуассона;
3) геометрическое распределение;
4) нормальный закон распределения (закон Гаусса);
5) показательное распределение;
6) распределение «хи-квадрат»;
7) распределение Стьюдента;
8) распределение Фишера-Снедекора.

Биномиальное распределение

В некоторых случаях вероятности значений дискретных случайных величин подчиняются определенным законами и могут быть заданы *аналитически* (в виде формул).

Среди законов распределения для дискретных случайных величин наиболее распространенным является *биномиальное распределение*.

Рассмотрим последовательность n идентичных повторных испытаний, удовлетворяющих следующим условиям, которые носят название схемы Бернулли:

1). Все n испытания – независимы. Это значит, что вероятность наступления события в любом из n повторных испытаний не зависит от результатов других испытаний.

2). В каждом испытании может наступить или не наступить некоторое событие A, вероятность наступления которого p остается неизменной.

3). Противоположное событие $\bar{A}$ имеет вероятность $q = 1 - p$, которая тоже остается неизменной от испытания к испытанию.

4). Эти события A и $\overline{A}$ - взаимно несовместные и противоположные, называемые успех и неуспех.

Тогда для вычисления вероятности того, что в n независимых повторных испытаниях, удовлетворяющих условиям $1 - 4$, событие A наступит ровно m раз, при $X = m = 0,1,2, \dots, n$ (в любой последовательности), вычисляется по формуле Бернулли:

$$P_{n,m} = C_n^m p^m q^{n-m},$$

где p– вероятность успеха в каждом испытании; $q = 1 - p$- вероятность неуспеха в каждом испытании; C_n^m - число сочетаний из nпо m.

Так как правая часть формулы Бернулли представляет общий член биномиального разложения $(q + p)^n$, то этот закон распределения называют *биномиальным*.

Используя формулу Бернулли, можно построить ряд распределения дискретной случайной величины, подчиняющейся биноминальному закону распределению, т.е., удовлетворяющей условиям 1-4. Вероятность того, что случайная величина X принимает значение, равное k, определяется по формуле

$$P(X = k) = C_n^k p^k q^{n-k}.$$

Биномиальный закон задается рядом распределения, который имеет вид, аналогичный приведенному в таблице:

Число успехов $X = m$	0	1	2	...	m	...	n
Вероятность $P_{n,m}$	$C_n^0 p^0 q^n = P_{n,0}$	$C_n^1 p^1 q^{n-1} = P_{n,1}$	$C_n^2 p^2 q^{n-2} = P_{n,2}$	...	$C_n^m p^m q^{n-m} = P_{n,m}$	$\dots$	$C_n^m p^n q^0 = P_{n,m}$

Пример 12. По мишени производится три выстрела, причем вероятность попадания при каждом выстреле равна $0,8$. Рассматривается случайная величина X– число попаданий в мишень. Найти ее ряд распределения.

Решение:Случайная величина X принимает значения $0,1,2,3$ с вероятностями, вычисленными по формуле Бернулли, где $n = 3$, $p = 0,8$ (вероятность попадания),$q = 1 - 0,8 = 0,2$ (вероятность непопадания).

Тогда

$$P_3(0) = P(X = 0) = C_3^0 \cdot 0{,}8^0 \cdot 0{,}2^{3-0} = 0{,}2^3 = 0{,}008,$$
$$P_3(1) = P(X = 1) = C_3^1 \cdot 0{,}8^1 \cdot 0{,}2^{3-1} = 3 \cdot 0{,}2^3 \cdot 0{,}8 = 0{,}096,$$
$$P_3(2) = P(X = 2) = C_3^2 \cdot 0{,}8^2 \cdot 0{,}2^{3-2} = 3 \cdot 02 \cdot 0{,}8^2 = 0{,}384,$$
$$P_3(3) = P(X = 3) = C_3^3 \cdot 0{,}8^3 \cdot 0{,}2^{3-3} = 0{,}8^3 = 0{,}512.$$

Таким образом, ряд распределения имеет следующий вид:

0	1	2	3
0,008	0,096	0,384	0,512

Распределение Пуассона

Дискретную случайную величину называют *Пуассоновской,* если ее закон распределения имеет следующий вид:

$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$,где$k = 0, \overline{\infty}$ и $\lambda = const$(постоянное значение).

Придавая k целые неотрицательные значения $k = 0,1,2, \dots, n$, можно записать ряд распределения вероятностей, который называется *законом распределения Пуассона*.

m	0	1	2	…	m	…	n
$P_{n,m}$	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2 e^{-\lambda}}{2!}$	…	$\frac{\lambda^m e^{-\lambda}}{m!}$	…	$\frac{\lambda^n e^{-\lambda}}{n!}$

Примеры Пуассоновских случайных величин:

Число вызовов на автоматическую станцию за промежуток времени T.

Число частиц распада некоторого радиоактивного вещества за промежуток времени T.

Число телевизоров, которые поступают в мастерскую за промежуток времени T в большом городе.

Число автомобилей, которые поступят к стоп-линии перекрестка в большом городе.

Замечание: В сериях независимых испытаний (когда n велико, p мало) для вычисления вероятности наступления события ровно k раз используют формулу Пуассона: $P_n(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$, где $\lambda = np$, то есть среднее число появлений событий остается постоянным.

Пример 13. Завод отправил на базу 5000 доброкачественных изделий. Вероятность того, что в пути изделие повредится, равна 0,0002. Найти вероятность, что на базу прибудут ровно три негодных изделия.

Решение: По условию$n = 5000, p = 0{,}0002, k = 3$. Найдем$\lambda$:

$$\lambda = n \cdot p = 5000 \cdot 0{,}0002 = 1.$$

По формуле Пуассона искомая вероятность равна:

$$P(X = 3) = \frac{1^3}{3!} e^{-1} = \frac{1}{6e} \cong 0{,}06,$$

где случайная величина X – число негодных изделий.

Геометрическое распределение

Пусть производятся независимые испытания, в каждом из которых вероятность появления события А равна $p(0 < p < 1)$ и, следовательно, вероятность его «непоявления» $q = 1 - p$. Испытания заканчиваются, как только появится событие А. Таким образом, если событиеА появилось в -м испытании, то в предшествующих $k - 1$ испытаниях оно не появлялось.

Обозначим через X дискретную случайную величину – число испытаний, которые нужно провести до первого появления события А. Очевидно, возможными значениями X являются натуральные числа $x_1 = 1, x_2 = 2, \ldots$ Пусть в первых $k - 1$ испытаниях событие А не наступило, а в -м испытании появилось. Вероятность этого “сложного события”, по теореме умножения вероятностей независимых событий,$P(X = k) = q^{k-1}p$.

Дискретная случайная величина имеет *геометрическое распределение*, если ее закон распределения имеет следующий вид:

$$P(X = k) = q^{k-1}p, \text{где} k = \overline{1, \infty}.$$

Замечание 1: Полагая $k = 1,2, \ldots$, получим геометрическую прогрессию с первым членом p и знаменателем q $(0 < q < 1)$. По этой причине распределение называют геометрическим.

Замечание 2: Ряд $\sum_{i=1}^{\infty} q^{k-1}p$ сходится и сумма его равна единице. Действительно сумма ряда равна $\frac{p}{(1-q)} = \frac{p}{p} = 1$.

Пример 14. Из орудия производится стрельба по цели до первого попадания. Вероятность попадания в цель $p = 0{,}6$. Найти вероятность того, что попадание произойдет при третьем выстреле.

Решение: По условию $p = 0{,}6; q = 1 - 0{,}6 = 0{,}4; k = 3$. Искомая вероятность равна:

$$P(X = 3) = 0{,}4^2 \cdot 0{,}6 = 0{,}096.$$

Нормальный закон распределения

Нормальный закон распределения является самым распространенным из законов распределения, наиболее часто встречающимся в случайных явлениях природы. Объясняется это тем, что, как правило, случайная величина принимает те или иные значения под воздействием большого числа взаимно независимых малых причин, другими словами, *является суммой очень большого числа взаимно независимых случайных величин*, малых по сравнению со всей суммой.

Нормальный закон распределения задается дифференциальной функцией:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\alpha)^2}{2\sigma^2}}.$$

Параметры $\alpha = M(X)$ и $\sigma^2 = D(X)$ являются соответственно математическим ожиданием и средним квадратическим отклонением случайной величины X, распределенной по нормальному закону.

В этом случае говорят, что X имеет нормальное распределение с параметрами α и σ^2.

Таким образом, если известны математическое ожидание и дисперсия случайной величины и известно, что она подчиняется нормальному закону, то известен полностью закон распределения этой величины – дифференциальная функция распределения.

Нормальный закон распределения иногда называют законом Гаусса. График плотности нормального распределения называют *нормальной кривой* или *кривой Гаусса*.

Максимум функции плотности вероятностей нормального распределения равен $\frac{1}{\sigma\sqrt{2\pi}}$. Отсюда следует, что с возрастанием σ максимальная ордината нормальной кривой убывает, а сама кривая становится более пологой, то есть сжимается к оси Ox; при убывании σ нормальная кривая становится более "островершинной" и растягивается в положительном направлении оси Oy:

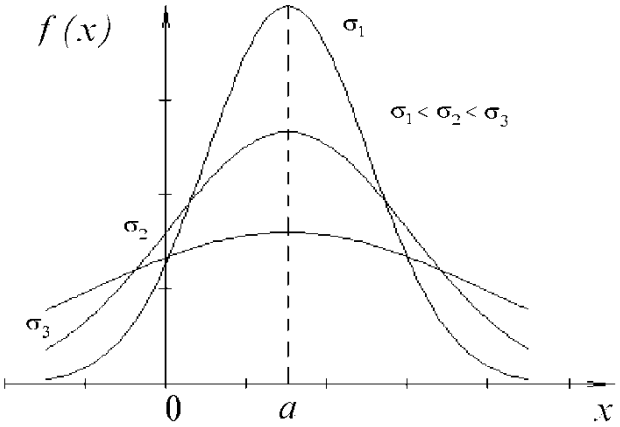

Рис. 7 График нормального закона распределения

Практически достоверно, что величина, распределения по нормальному закону, отклоняется от математического ожидания не более чем на утроенное среднее квадратическое отклонение, т.е.

$$P\{|x-\alpha|<3\sigma\}\approx 1.$$

Таким образом, имеет место, так называемое, *правило «трех сигм»*: если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения.

Серединным, или вероятным, отклонением случайной величины называют такое число α_0, что вероятности неравенств

$$|x-M(x)|<\alpha_0 \text{ и } |x-M(x)|\geq\alpha_0$$

одинаковы (т. е. равны по $\frac{1}{2}$ каждая).

Вероятность попадания в заданный интервал нормальной случайной величины

Пусть случайная величина X распределена по нормальному закону. Тогда вероятность того, что X примет значение, принадлежащее интервалу(α, β), равна:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x) \cdot dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-(x-\alpha)^2/2\sigma^2} \cdot dx.$$

Введем новую переменную $z = (x - \alpha)/\sigma$. Отсюда, $x = \sigma z + \alpha$, $dx = \sigma dz$. Найдем новые пределы интегрирования. Если $x = \alpha$ то$z = (\alpha - \alpha)/\sigma$; если $x = \beta$ то $z = (\beta - \alpha)/\sigma$.

Таким образом, имеем

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{(\alpha-\alpha)/\sigma}^{(\beta-\alpha)/\sigma} e^{-z^2/2}(\sigma dz) =$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{(\alpha-\alpha)/\sigma}^{0} e^{-z^2/2}\, dz + \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{(\beta-\alpha)/\sigma} e^{-z^2/2}\, dz =$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{(\beta-\alpha)/\sigma} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{(\alpha-\alpha)/\sigma} e^{-z^2/2}\, dz.$$

Пользуясь функцией Лапласа$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} \cdot dz$, получим

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta-\alpha}{\sigma}\right) - \Phi\left(\frac{\alpha-\alpha}{\sigma}\right).$$

Пример 15. Случайная величина X распределена по нормальному закону с $M(X) = 30$ и $\sigma(X) = 10$. Найти вероятность того, что случайная величина X примет значение, принадлежащее интервалу $(10{,}50)$.

Решение:

$$P(10 < X < 50) = \Phi\left(\frac{50-30}{10}\right) - \Phi\left(\frac{10-30}{10}\right) = 2\Phi(2).$$

По таблице приложения 2 находим $\Phi(2) = 0{,}4772$. Отсюда искомая вероятность:

$$P(10 < X < 50) = 2 \cdot 0{,}4772 = 0{,}9544.$$

Показательное распределение

Показательным (экспоненциальным) называют распределение вероятностей непрерывной случайной величины X, которое описывается плотностью:

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

Гдеλ – постоянная положительная величина.

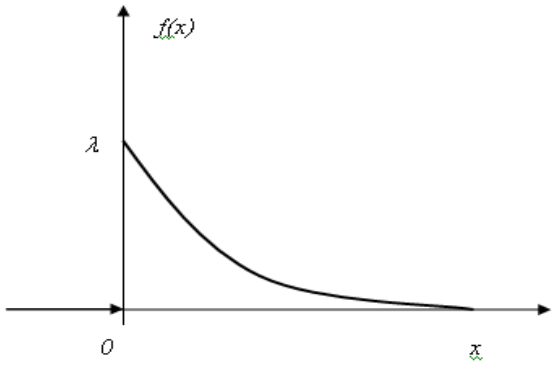

Рис. 8 График функции показательного распределения

Например, время безотказной работы компьютерной системы есть случайная величина, имеющая показательное распределение с параметром λ, физический смысл которого – среднее число отказов в единицу времени. Интервал между последовательными поступлениями вызовов на автоматическую телефонную станцию, интервал между посещениями врача-педиатра – это примеры показательных случайных величин.

Интегральная функция распределения показательного закона имеет вид:

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

Пример 16. Написать плотность и функцию распределения показательного закона, если параметр $\lambda = 8$.

Решение. Очевидно, искомая плотность распределения

$$f(x) = 8e^{-8x} \text{ при} x \geq 0;\ f(x) = 0 \text{ при } x < 0.$$

Искомая функция распределения

$$F(x) = 1 - e^{-8x} \text{ при } x \geq 0; F(x) = 0 \text{ при } x < 0.$$

Распределение «Хи-квадрат»

Пусть $x_1, x_2, x_3, \dots, x_n$ – нормально распределенные независимые случайные величины, причем математическое ожидание каждой из них равно нулю, а дисперсия равна единице, т.е. закон распределения, имеет вид:

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

Возьмем сумму квадратов этих случайных величин:

$$\chi^2 = \sum_{i=1}^{n} x_i^2.$$

Принято считать, что сума квадратов этих величин распределена по закону «хи-квадрат» с $k = n$ степенями свободы.

Плотность распределения χ^2 с $k = n$ степенями свободы имеет вид:

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} \cdot e^{-\frac{x}{2}} \cdot x^{\frac{k}{2}-1}, & x > 0, \end{cases}$$

где $\Gamma(x) = \int t^{x-1}e^{-t}dt$ – гамма функция.

Распределение χ^2 определяется одним параметром – числом степеней свободыk. С ростом числа степеней свободы, распределение χ^2 достаточно медленно приближается к нормальному распределению.

Распределение Стьюдента

Распределение Гаусса хорошо описывает отклонение измеряемой величины в эксперименте при большом количестве повторных измерений. На практике нередко обходятся двумя-тремя измерениями. Результаты образуют очень малую выборку. В этом случае распределение Стьюдента гораздо лучше описывает получаемые результаты.

Распределение Стьюдента – это непрерывное одномерное распределение с одним параметром – количеством степеней свободы. Форма распределения Стьюдента похожа на форму нормального распределения (чем больше число степеней свободы, тем ближе распределение к нормальному). Отличием является то, что хвосты распределения Стьюдента медленнее стремятся к нулю, чем хвосты нормального распределения.

Пусть Z – нормально распределенная случайная величина с математическим ожиданием равным нулю и дисперсией равной единице. Пусть V– независимая случайная величина, которая распределена по закону χ^2 с k степенями свободы. Тогда величина $t = \frac{Z}{\sqrt{\frac{V}{k}}}$, имеет распределение, называемое t– распределением или распределением Стьюдента с k степенями свободы. Плотность распределения Стьюдента имеет вид:

$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k}\cdot\Gamma\left(\frac{k}{2}\right)} \cdot \left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}}.$$

С ростом числа степеней свободы распределение Стьюдента достаточно быстро приближается к нормальному распределению.

Таблицы распределения Стьюдента позволяют при данном числе степеней свободыk по вероятности q определить значение t_q, для которого выполняется соотношение: $P\left(|t| > t_q\right) = q$.

Пример 17. Найти симметричный интервал, в который случайная величина, распределенная по закону Стьюдента с 12-ю степенями свободы, попадает вероятностью 0,9.

Решение: Очевидны соотношения:

$$P(-x < t < x) = P(|t| < x) = 1 - P(|t| \geq x) = 0{,}9$$

Из последнего равенства следует:

$$P(|t| \geq x) = 0{,}1 \quad (n = 12).$$

Определяем из таблицы IV в приложении: $x = 1{,}782$.

Распределение Фишера-Снедекора

Пусть U и V – независимые случайные величины, распределенные по закону χ^2 со степенями свободы k_1 и k_2. Тогда величина $F = \frac{U/k_1}{U/k_2}$ имеет

распределение, которое называется распределением Фишера-Снедекора со степенями свободы k_1и k_2.

Плотность распределения Фишера-Снедекора со степенями свободы k_1и k_2 имеет вид:

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \dfrac{\Gamma\left(\dfrac{k_1+k_2}{2}\right) \cdot k_1^{\frac{k_1}{2}} \cdot k_2^{\frac{k_2}{2}}}{\Gamma\left(\dfrac{k_1}{2}\right) \cdot \Gamma\left(\dfrac{k_2}{2}\right)} \cdot \dfrac{x^{\frac{k_1-2}{2}}}{(k_2+k_1x)^{\frac{k_1+k_2}{2}}}, & x > 0. \end{cases}$$