

Конспект лекций по теории вероятностей и математической статистике.

Лектор – Лохвицкий Михаил Сергеевич

Часть 3+ предельные теоремы и коэффициент корреляции

2. Случайные величины

В жизни мы постоянно сталкиваемся с величинами, значения которых мы априорно не знаем (время ожидания транспорта, время безотказной работы аппаратуры, рост и вес человека в данный момент, курс акций на бирже, число студентов на лекции, продолжительность жизни любого субъекта). Всё это приводит нас к необходимости ввести в рассмотрение случайные величины.

К случайным величинам относятся результаты измерений (из-за погрешностей измерительной аппаратуры), вычислений (из-за ошибок, возникающих при округлении и применении приближённых алгоритмов), параметры аппаратуры (из-за разного рода помех и геомагнитных воздействий) и т.п.

Если значения, которые принимает случайная величина (СВ) дискретные или счетные $(x_1, x_2, \dots, x_k, \dots, x_n, \dots)$ то СВ называется дискретной (ДСВ). В противном случае СВ называется непрерывной.

2.1. Дискретные случайные величины

В результате опыта для ДСВ происходит событие $\{X = x_k\}$, а этому событию соответствует вероятность p_k . Соответствие между возможными значениями СВ и вероятностями их принятия называется **законом**

распределения. Закон распределения можно задать в виде таблицы, которая называется **рядом распределения**:

| | | | | | |
|---|-------|-------|-----|-------|-----|
| X | x_1 | x_2 | ... | x_k | ... |
| P | p_1 | p_2 | ... | p_k | ... |

Совокупность событий $\{X = x_k\}$ образует полную группу, поэтому выполняется следующее свойство:

Свойство ряда распределения (условие нормировки)

$$\sum_k p_k = 1 \quad (2.1)$$

Закон распределения можно задать и графически. Такой график называется **многоугольником распределения**.

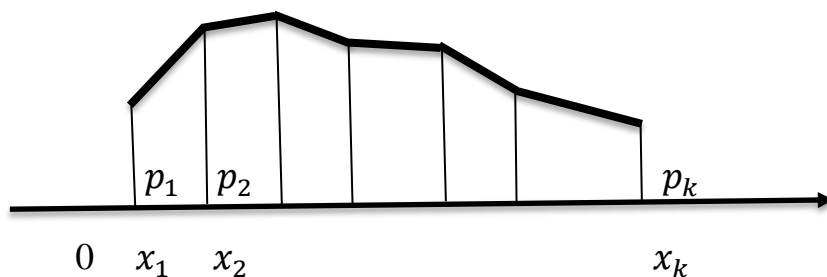


Рисунок. 2.2. Многоугольник распределения.

Пример 2.1. На экзамене по Теории вероятностей студент может получить оценки 2, 3, 4, 5 с вероятностями 0,05; 0,15; 0,5; 0,3. Построить ряд и многоугольник распределения.

Ряд распределения:

| | | | | |
|-----------|------|------|-----|-----|
| $X = x_k$ | 2 | 3 | 4 | 5 |
| p_k | 0,05 | 0,15 | 0,5 | 0,3 |

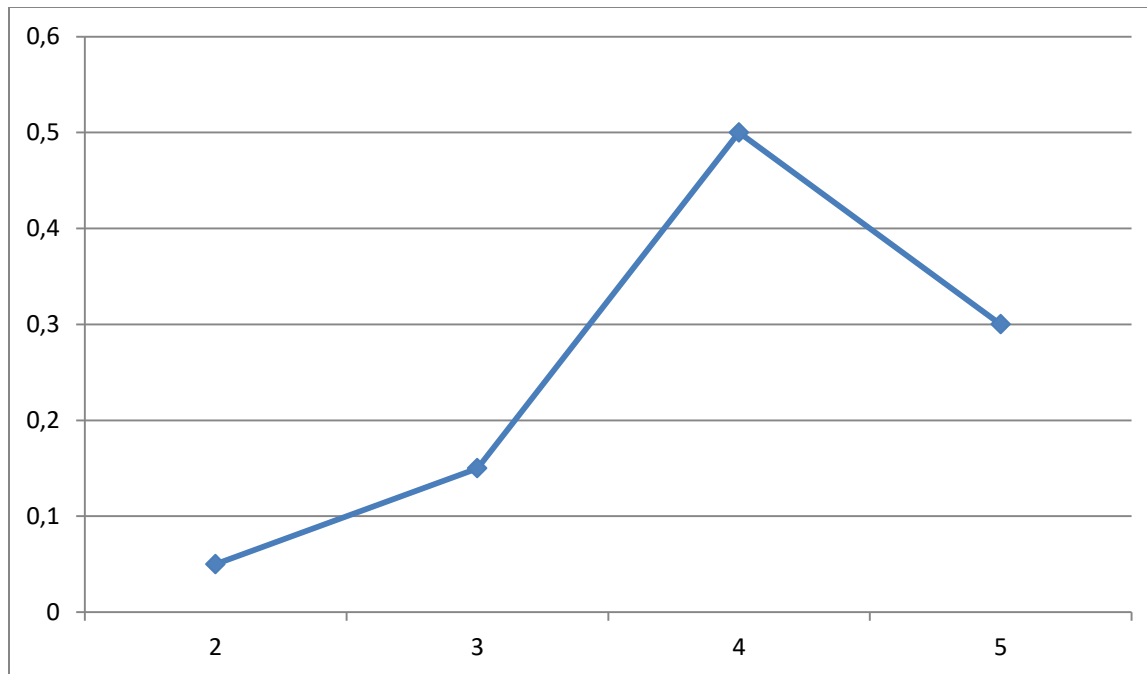


Рисунок 2.2. Многоугольник распределения СВ оценки на экзамене.

2.2. Функция распределения вероятностей

Определение 2.2. Функция

$$F(x) = P(X < x), \quad (2.2)$$

определенная при всех

$x \in R^1$ (т. е. на множестве действительных чисел), называется *функцией распределения вероятностей* случайной величины X .

Свойства функции $F(x)$.

1. $F(x)$ -неубывающая функция, т.е. если $x_1 < x_2$, то $F(x_1) \leq F(x_2)$.

Действительно, интервал $(-\infty, x_1)$ включен в интервал $(-\infty, x_2)$, поэтому вероятность попадания в меньший интервал (рисунок 2.3.) не может быть больше вероятности попадания в больший интервал: $F(x_1) = P(X < x_1) \leq P(X < x_2) = F(x_2)$

2. $F(\infty) = \lim_{x \rightarrow +\infty} F(x) = \lim_{x \rightarrow +\infty} P(X < x) = P(X < \infty) = 1.$

$$3. F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = \lim_{x \rightarrow -\infty} P(X < x) = P(X < -\infty) = 0$$

$$4. P(X \geq x) = 1 - P(X < x) = 1 - F(x)$$

$$5. P(x_1 \leq X < x_2) = F(x_2) - F(x_1). \quad (2.3)$$

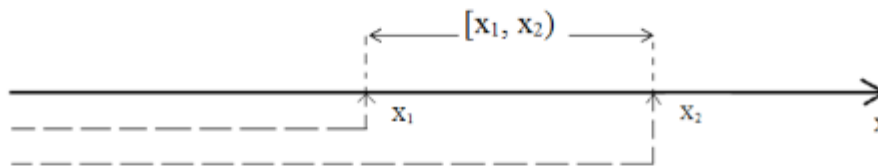


Рисунок 2.3. Вероятность попадания в интервал..

Событие попадания в интервал $(-\infty, x_2)$ можно представить в виде суммы двух событий

$$\{X < x_2\} = \{X < x_1\} + \{x_1 \leq X < x_2\}$$

Т.к слагаемые в правой части - несовместные события, то

$$P\{X < x_2\} = P\{X < x_1\} + P\{x_1 \leq X < x_2\}$$

Заменяя первые две из этих вероятностей на выражение через функцию распределения, получаем требуемый результат.

$$6. F(x) \text{ непрерывна слева, т.е. } \lim_{x \rightarrow x_0 - 0} F(x) = F(x_0).$$

Пример 2.2. Производятся наблюдения над состоянием определённого элемента технического устройства.

Если этот элемент включён в текущий процесс (состояние «on»),

| | | |
|-------|-------------|-----|
| x_k | 0 | 1 |
| p_k | $q = 1 - p$ | p |

случайная величина X равна 1, если не включён (состояние «off»), то $X = 0$.

Пусть вероятность состояния «on» равна $p \in (0;1)$, а состояния «off» $q = 1-p$.

Найдем $F(x)$. Очевидно, что

при всех $x \leq 0$ имеем $P(X < x) = F(x) = 0$, поскольку наименьшее возможное значение ξ равно 0.

При $0 < x \leq 1$ условие $\{X < x\}$ эквивалентно тому, что $\{X=0\}$, поэтому $F(x) = P(X = 0) = q$. При $x > 1$ имеем:

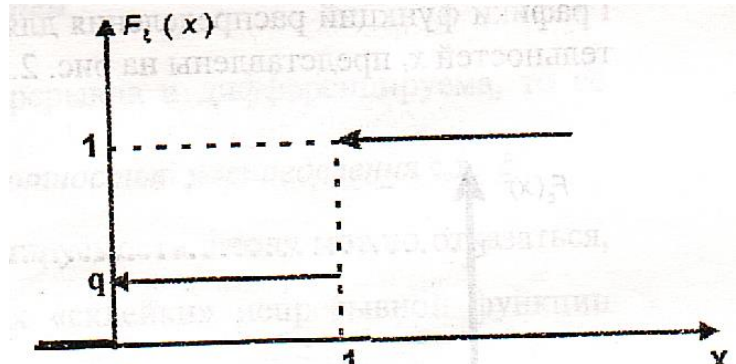


Рис. 2.4. График $F(x)$ примера 2.2.

$\{X < x\} = \{X=0\} + \{X=1\}$, причём события в правой части несовместны, поэтому $F(x) = P(X = 0) + P(X = 1) = q + p = 1$. График $F(x)$ представлен на рисунке 2.4. Проведённый анализ позволяет обобщить результат на случай многих дискретных значений С.В.

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

Пример 2.3. Случайная величина X имеет распределение вида:

$P(X = x_k) = p_k$, при некоторых значениях x_k , $p_k > 0$, $\sum_k p_k = 1$ Это позволяет описать её в виде *ряда распределения*.

| | | | | |
|-----|-------|-------|-----|-------|
| X | x_1 | x_2 | ... | x_n |
| p | p_1 | p_2 | ... | p_n |

В этом случае получаем для $x \in [x_i, x_{i+1})$:

$$\{X < x\} = \{X = x_1\} + \{X = x_2\} + \dots + \{X = x_i\}, \text{ поэтому } F(x) = p_1 + p_2 + \dots + p_i = \sum_{j=1}^i p_j.$$

Окончательно получаем:

$$F(x) = \begin{cases} 0, & x \leq x_1 \\ \sum_{j=1}^i p_j, & x_i < x \leq x_{i+1} \end{cases}$$

Если среди возможных значений случайной величины X имеется наибольшее, скажем, x_n , то

$$F_X(x) = \begin{cases} 0, & x \leq x_1 \\ \sum_{j=1}^i p_j, & x_i < x \leq x_{i+1}, i \leq n-1 \\ 1, & x > x_n \end{cases} \quad (2.4)$$

Графики функций распределения для бесконечной (а) и конечной (б) последовательностей x_i представлены на рис. 2.5.

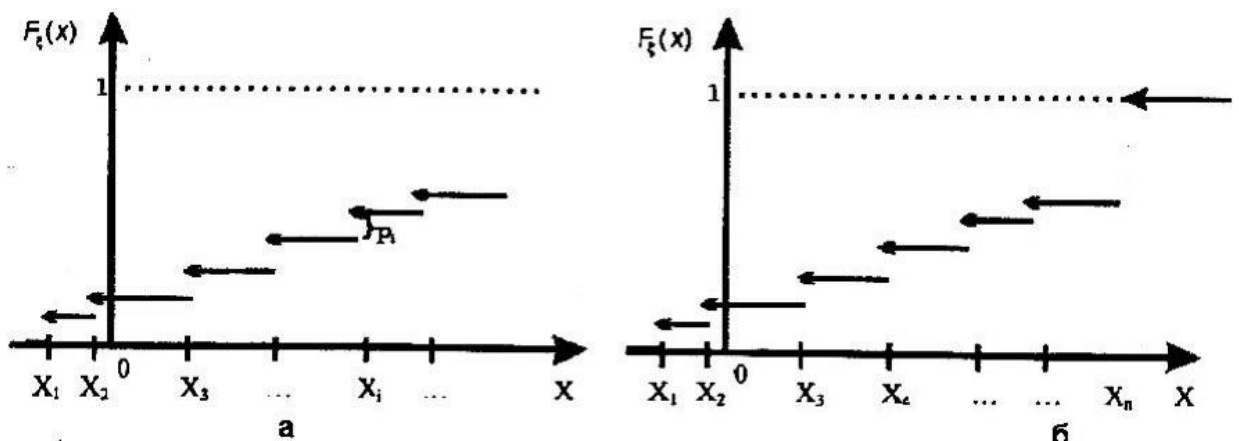


Рис. 2.5. Графики $F_\xi(x)$

примера 2.3.

В общем случае по виду функции распределения случайные величины делятся на дискретные ($F(x)$ кусочнопостоянна), непрерывные ($F(x)$ непрерывна) и смешанные. Пример графика $F(x)$ для с.в. смешанного типа приведён на рис. 2.6.

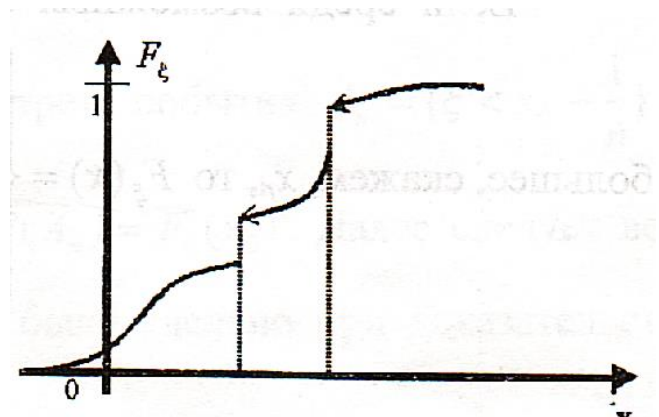


Рис. 2.6. График $F(x)$ смешанной с.в.

Пример. Построить функцию распределения для дискретной с. в., заданной рядом распределения:

| | | | | | |
|-------|-----|-----|-----|-----|-----|
| x_i | 0 | 2 | 4 | 6 | 8 |
| p_i | 0,1 | 0,2 | 0,4 | 0,2 | 0,1 |

Используем формулу (2.4):

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 0.1, & 0 < x \leq 2 \\ 0.1 + 0.2 = 0.3, & 2 < x \leq 4 \\ 0.3 + 0.4 = 0.7, & 4 < x \leq 6 \\ 0.7 + 0.2 = 0.9, & 6 < x \leq 8 \\ 1, & x > 8 \end{cases}$$

График функции распределения будет выглядеть следующим образом:

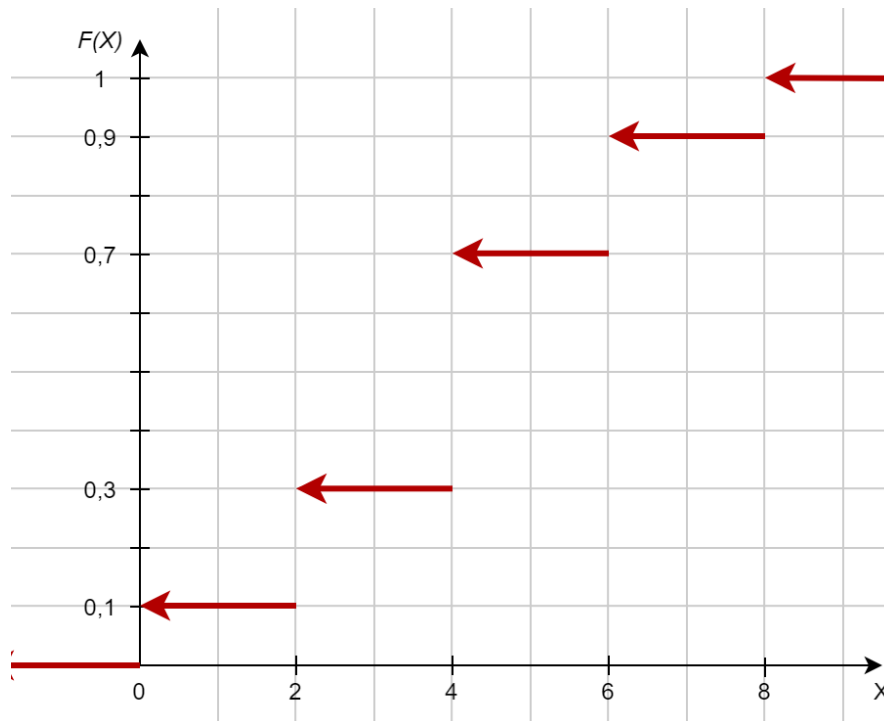


Рисунок 2.7 – График функции распределения

2.3. Плотность распределения вероятностей

Определение 2.3. Если $F(x)$ непрерывна и дифференцируема, то её производная

$$p(x) = \frac{d}{dx} F(x) \quad (2.5)$$

называется *плотностью распределения* с.в. X .

Свойства плотности распределения:

$$1. p(x) \geq 0$$

$$2. P(x < X < x + \Delta x) = p(x)\Delta x + o(\Delta x)$$

Доказательство: $\frac{d}{dx} F(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = p(x)$ по теореме из теории пределов получим равенство $\frac{F(x+\Delta x) - F(x)}{\Delta x} = p(x) + \alpha(x)$, где $\alpha(x)$ - бесконечно малая. Умножим обе части последнего равенства

на Δx , а приращение функции распределения (по свойству 5) заменим на вероятность попадания в интервал $P(x < X < x + \Delta x)$. В итоге получается равенство, которое мы хотели доказать (свойство 2 плотности).

$$3. P(\alpha < X < \beta) = \int_{\alpha}^{\beta} p(x) dx \quad (2.6)$$

Доказательство: интервал (α, β) разобьем на части Δx_k , вероятность попадания в каждый интервал определим по формуле из предыдущего свойства. Затем $\max \Delta x_k$ устремим к нулю. Предел интегральных сумм и будет интеграл $\int_{\alpha}^{\beta} p(x) dx$:

$$P(\alpha < X < \beta) = \sum_k P(x_k < X < x_k + \Delta x_k) = \lim_{\max \Delta x_k \rightarrow 0} \sum_k p(x_k) \Delta x_k + o(\Delta x) = \int_{\alpha}^{\beta} p(x) dx$$

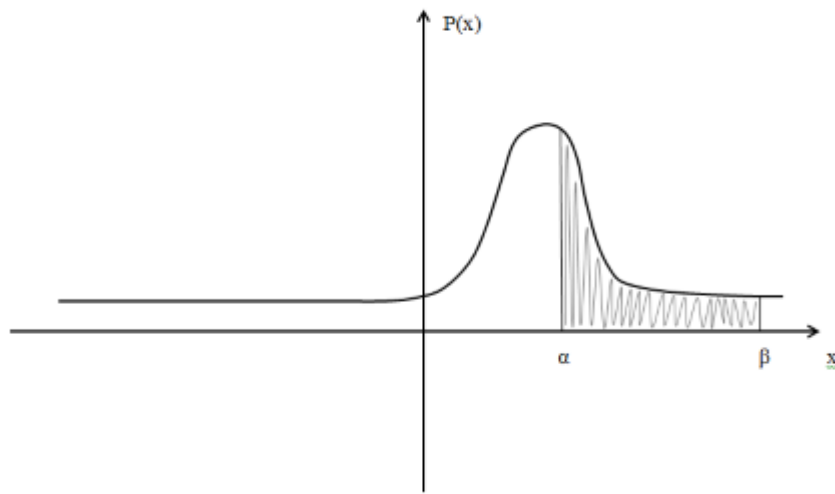


Рисунок 2.8. Вероятность попадания в интервал.

Т.е. вероятность попадания в интервал (α, β) численно равна площади под графиком плотности на этом интервале (рисунок 2.8).

$$4. F(x) = \int_{-\infty}^x p(y) dy \quad (2.7)$$

Доказательство: используем предыдущее свойство при $\beta = x$ и $\alpha \rightarrow -\infty$.

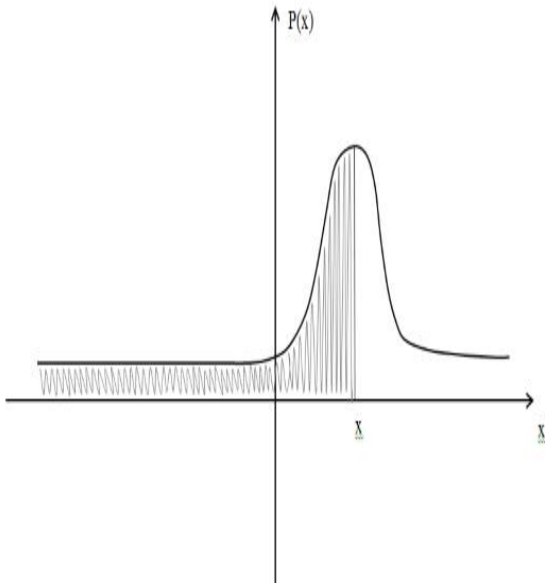


Рисунок 2.9. Связь функции распределения с плотностью.

Геометрически $F(x)$ численно равна площади под графиком функции $p(x)$, лежащей левее абсциссы x (заштрихованная область на рис. 2.9).

$$5. \lim_{x \rightarrow \pm\infty} p(x) = 0$$

$$6. \int_{-\infty}^{\infty} p(x) dx = 1 \text{ (условие нормировки)} \quad (2.8)$$

Условие нормировки: график $p(x)$ всегда ограничивает площадь, равную 1 (рисунок 2.10.)

Доказательство: используем свойство 3 при $\beta \rightarrow +\infty$ и $\alpha \rightarrow -\infty$. Вероятность, стоящая в левой части равенства из свойства 3, в этом случае стремится к 1.

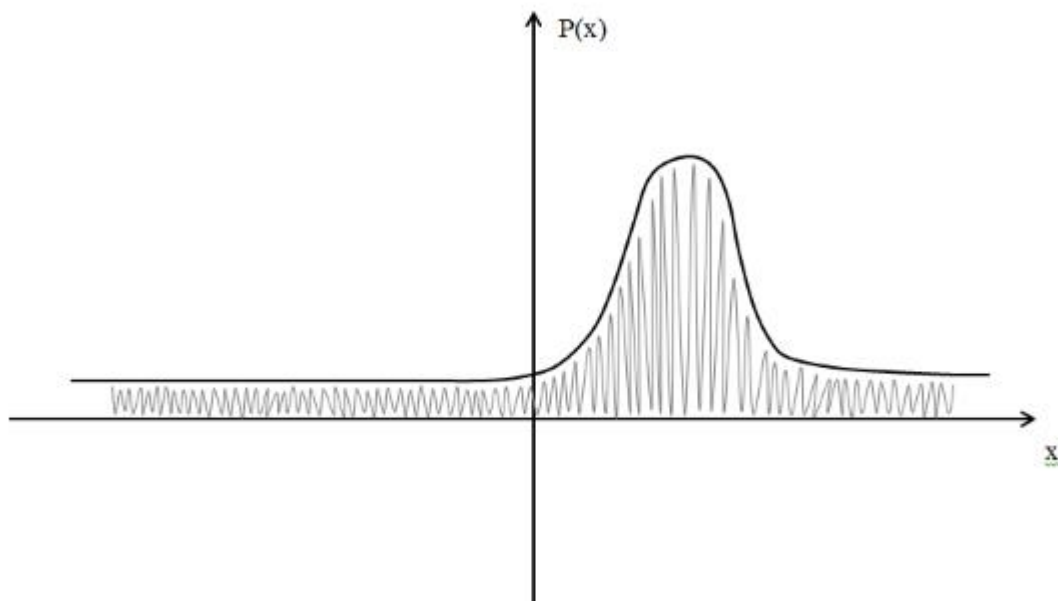


Рисунок 2.10. Условие нормировки плотности.

2.3. Числовые характеристики случайных величин

Закон распределения (ряд распределения или многоугольник) для ДСВ, функция распределения как для дискретной, так и для непрерывной СВ, плотность распределения для непрерывной СВ являются самыми полными характеристиками СВ. Однако, во многих задачах можно ограничиться только несколькими числовыми характеристиками такими как, например, среднее ожидаемое значение, среднее отклонение от этого среднего значения и т.д.

2.3.1. Математическое ожидание

Пример 2.4. В результате n испытаний СВ приняла значение x_1 - n_1 раз, значение x_2 - n_2 раза, ..., значение x_k - n_k раз. В математической статистике значения, которые принимает СВ называются *выборкой*, а n – *объёмом выборки*: $n = n_1 + n_2 + \dots + n_k$. Найдём среднее арифметическое всех

полученных значений. Оно называется *средним выборочным значением* и обозначается $\overline{x_{\text{выб.}}}$:

$$\overline{x_{\text{выб.}}} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{n} = \sum_1^k x_i \frac{n_i}{n} \quad (2.9)$$

Но величина n_i/n это относительная частота появления события $\{X = x_i\}$,

которая является оценкой вероятности этого события. Поэтому **средним ожидаемым значениям СВ по определению является следующая величина:**

Определение. 2.4. Математическим ожиданием ДСВ называется величина: $MX = \sum_1^k x_i p_i$. (2.10)

Определение .2.5. Математическим ожиданием непрерывной СВ называется величина: $MX = \int_{-\infty}^{\infty} xp(x)dx$ (2.11)

Замечание. В дальнейшем будем считать, если есть выражение $M\{\dots\}$, то это значит вычисляется среднее значение для $\{\dots\}$.

Свойства математического ожидания

$$1. M C = C; \quad (2.12)$$

Т.е. математическое ожидание постоянной равно самой этой постоянной.

Доказательство: можно считать, что в этом случае СВ принимает только одно значение C с вероятностью единица. $MC = C \cdot 1 = C$.

$$2. M(CX) = C MX; \quad (2.13)$$

Таким образом, постоянную можно выносить за знак математического ожидания (за знак усреднения).

Доказательство: Если случайная величина X принимает значение x_i , то CX принимает значение cx_i с вероятностью p_i . Поэтому

$$M(CX) = \sum_1^k (C \cdot x_i) p_i = C \cdot \sum_1^k x_i p_i = C \cdot MX$$

$$3. M(X+Y) = MX + MY; \quad (2.14)$$

4. Если случайные величины независимы, то

$$M(XY) = MX \cdot MY. \quad (2.15)$$

Замечание. Определение независимости СВ будет дано через несколько лекций, там же будет и доказано это свойство.

Пример 2.4 Рассмотрим две случайные величины X и Y со следующими законами распределения

| | | |
|-----|-----|-----|
| X | -1 | 1 |
| p | 0,5 | 0,5 |

| | | |
|-----|------|-----|
| Y | -100 | 100 |
| p | 0,5 | 0,5 |

Найти математические ожидания случайных величин X и Y .

Решение: $MX = -1 \cdot 0,5 + 1 \cdot 0,5 = 0$; $MY = -100 \cdot 0,5 + 100 \cdot 0,5 = 0$.

Вывод: Математические ожидания у этих величин равны, но из законов распределения видно насколько эти величины различны.

В теории вероятностей вводятся характеристики, которые характеризуют средние отклонения (разброс) СВ от их средних значений. Это дисперсия и среднее квадратическое отклонение.

2.3.2. Дисперсия и среднее квадратическое отклонение

Определение. 2.6. Дисперсией (рассеянием) случайной величины называется математическое ожидание квадрата отклонения случайной величины от математического ожидания:

$$DX = M(X - MX)^2 \quad (2.16)$$

Вероятностный смысл дисперсии: Дисперсия равна среднему значению квадрата разброса случайной величины от её математического ожидания.

Поэтому вводят определение **среднего квадратического отклонения:**

$$\sigma X = \sqrt{DX}. \quad (2.17)$$

Эта величина характеризует среднее отклонение СВ от математического ожидания.

Формула для вычисления дисперсии дискретной случайной величины:

$$DX = \sum_i (x_i - MX)^2 p_i \quad (2.18)$$

Формула для вычисления дисперсии непрерывной случайной величины:

$$DX = \int_{-\infty}^{+\infty} (x - MX)^2 p(x) dx. \quad (2.19)$$

Замечание. С целью сокращения числа вычислений дисперсию чаще вычисляют по вычислительной формуле:

Вычислительная формула для дисперсии:

$$DX = M(X^2) - (MX)^2. \quad (2.20)$$

Вывод формулы:

$$DX = M(X - MX)^2 = M(X^2 - 2X \cdot MX + (MX)^2) = M(X^2) - M(2X \cdot MX) + M(MX)^2 = M(X^2) - 2MXMX + (MX)^2 = M(X^2) - (MX)^2,$$

Величина $M(X^2)$ вычисляется по следующим формулам:

$$M(X^2) = \sum_{i=1}^n x_i^2 p_i, \quad M(X^2) = \int_{-\infty}^{\infty} x^2 p(x) dx, \quad (2.21)$$

для дискретной и непрерывной СВ соответственно.

Свойства дисперсии

1. Дисперсия постоянной величины равна нулю:

$$D(C) = 0 \quad (2.22)$$

Доказательство. $DC = M(C - MC)^2 = M(C - C)^2 = M(0)^2 = 0$

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

$$D(CX) = C^2 DX \quad (2.23)$$

Доказательство. $D(CX) = M(CX - MCX)^2 = M(CX - CMX)^2 = MC^2(X - MX)^2 = C^2 M(X - MX)^2 = C^2 DX$

3. Если случайные величины X и Y независимы, то:

$$D(X + Y) = DX + DY \quad (2.24)$$

Замечание. Определение независимости СВ будет дано через несколько лекций, там же будет и доказано это свойство.

Следствие 1. $D(X + C) = DX + DC = DX$

Следствие 2. $D(X - Y) = D(X + (-1)Y) = DX + D((-1)Y) =$

$DX + (-1)^2 DY = DX + DY$. (для независимых X, Y)

Свойства среднеквадратического отклонения:

$$1. \sigma c = 0$$

$$2. \sigma(cX) = |c| \sigma X$$

$$3. \sigma(c + X) = \sigma X$$

Достоинством σX является тот факт, что оно имеет ту же размерность, что и X .

Нормальное (гауссовское) распределение $N(m, \sigma^2)$

Плотность распределения нормального закона имеет вид

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (3.1)$$

Наиболее часто встречается во всех областях науки и техники, экономики и финансов, медицине и демографии. Например, величина помехи в радиоканале, ошибки в любых измерениях, коэффициент

интеллектуальности (IQ), рост человека, изменение напряжения в сети, колебания курса акций, температура тела, возраст въезжающих в страну эмигрантов имеют нормальное распределение (разумеется, в каждом случае значения параметров m и σ меняются). Важность этого закона следует из **Центральной предельной теоремы**. На качественном уровне ЦПТ утверждает, что при определенных условиях (достаточно часто выполненных на практике) сумма случайных величин имеет асимптотически нормальное распределение независимо от того, каково было первоначальное распределение отдельных слагаемых.

Параметры: $-\infty < m < \infty, \sigma > 0$.

Неприятной неожиданностью является тот факт, что функция распределения такой с. в. $F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(y-m)^2}{2\sigma^2}} dy$ аналитически явно не выражается. Поэтому для практического использования нормального распределения приходится прибегать к таблицам или специальным вычислительным процедурам для нахождения значений функции $F(x)$. При этом нет необходимости составлять таблицы для каждой пары параметров (m, σ) , а достаточно обойтись таблицами для $m = 0$ и $\sigma = 1$ (соответствующее распределение ещё называют *стандартным нормальным распределением*). Плотность распределения этого закона $N(0, 1)$ обозначается, как $\varphi(x)$.

Функция распределения этого закона имеет вид $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy$.

(в последнем интеграле мы поменяли местами пределы интегрирования, изменив знак перед интегралом).

ВАЖНАЯ ФОРМУЛА. Вероятность попадания в интервал для нормально распределённой С.В.:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta-m}{\sigma}\right) - \Phi\left(\frac{\alpha-m}{\sigma}\right) \quad (3.2)$$

$$\text{Где } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{y^2}{2}} dy - \text{функция Лапласа.} \quad (3.3)$$

(Напомним, что функция Лапласа – *нечетная* и при $x \geq 5$; $\Phi(x) \approx 0,5$)

Следствия из формулы (32):

1. Функция распределения равна

$$F(x) = P(-\infty < X < x) = \Phi(x) - \Phi(-\infty) = 0,5 + \Phi(x) \text{ (было использовано, что } \Phi(-\infty) = -\Phi(\infty) = -0,5).$$

2. **Формула вероятности попадания в интервал, симметричный относительно математического ожидания.** $P(|X - m| < \delta) =$
 $P(m - \delta < X < m + \delta) = \Phi\left(\frac{m+\delta-m}{\sigma}\right) - \Phi\left(\frac{m-\delta-m}{\sigma}\right) = 2\Phi\left(\frac{\delta}{\sigma}\right) \quad (3.4)$

(было использовано свойство нечетности функции Лапласа).

3. «Правило» трёх сигм. При $\delta = 3\sigma$ из предыдущей формулы получим $P(|X - m| < 3\sigma) = 2\Phi\left(\frac{3\sigma}{\sigma}\right) = 2\Phi(3) = 2 \cdot 0,49865 = 0,9973$

Т.е. с вероятностью близкой к 1 почти все значения нормального закона $N(m, \sigma^2)$ принимают значения в интервале с радиусом $\delta = 3\sigma$ и центром в точке m .

В обозначении $N(m, \sigma^2)$ первый параметр равен математическому ожиданию, а второй параметр равен дисперсии с.в.

Определение 6.2. Коэффициентом корреляции двух с. в. X и Y называется число

$$r(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \cdot \sigma_Y} \quad (3.5)$$

где $\text{cov}(X, Y) = M[(X - MX) \cdot (Y - MY)]$. (3.6)

Замечание. По аналогии с моментами одной с. в. ковариацию можно назвать смешанным центральным моментом второго порядка.

Вычислительная формула для ковариации.

$$\text{cov}(X, Y) = M(X \cdot Y) - MX \cdot MY$$

Свойства коэффициента корреляции:

$$1. \quad r(X, Y) = r(Y, X)$$

$$2. \quad r(cX, Y) = r(X, Y) \cdot \text{sign}(c)$$

$$\text{Здесь } \text{sign}(c) = \begin{cases} -1, & \text{если } c < 0; \\ 1, & \text{если } c > 0 \end{cases}$$

$$3. \quad r(X + c, Y) = r(X, Y)$$

$$4. \quad |r(X, Y)| \leq 1$$

5. Если X и Y независимы, то $r(X, Y) = 0$.

$$6. \quad \text{Если } Y = kX + b, \text{ то } |r(X, Y)| = 1; \quad r(X, Y) = \begin{cases} 1, & \text{если } k > 0; \\ -1, & \text{если } k < 0 \end{cases}$$

Из свойства 6 ковариации, в частности вытекает, что если две гауссовские случайные величины независимы, то $r = 0$. Верно и обратное: если $r = 0$, то две гауссовские с.в. независимы (докажите это). Но в общем случае из некоррелированности случайных величин (т. е. из того, что $\text{cov}(X, Y) = 0$) не следует их независимость.

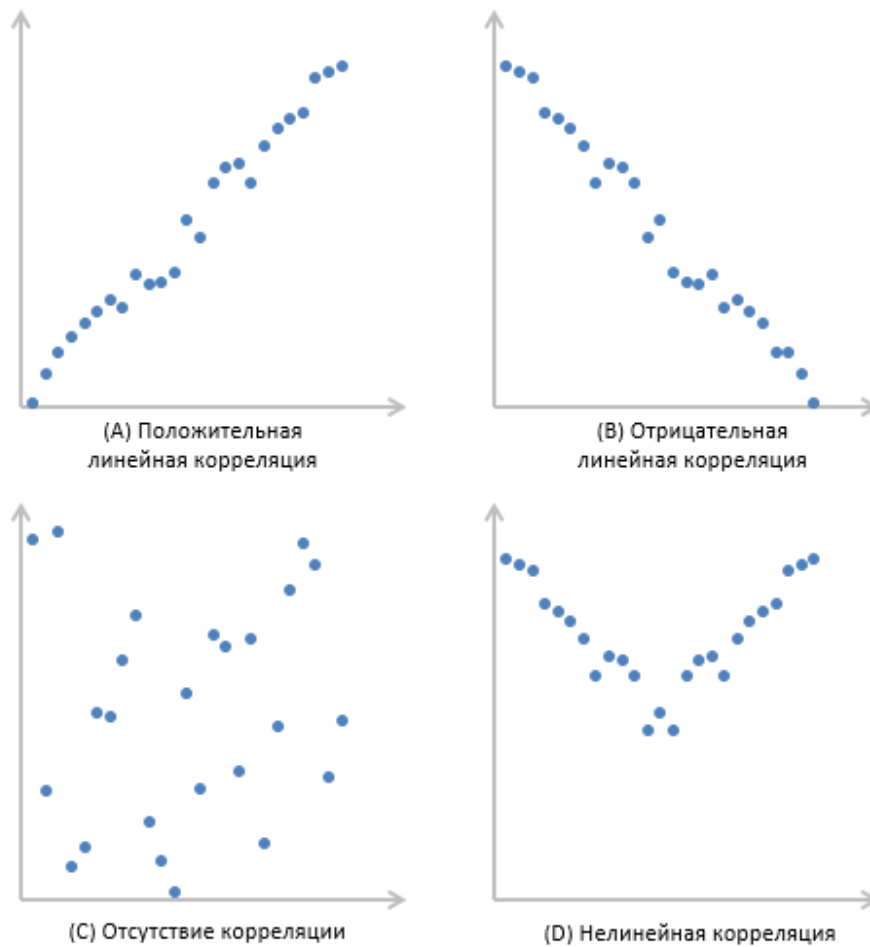
Чем лучше реальная зависимость Y от X аппроксимируется линейной, тем ближе по модулю к 1 будет их коэффициент корреляции.

Если $r(X, Y) < 0$ это означает, что с ростом одной величины другая имеет тенденцию убывать.

Если $r(X, Y) > 0$ это означает, что с ростом одной величины другая имеет тенденцию возрастать.

Шкала Чеддока для оценки линейной связи двух случайных величин

| Значение модуля коэффициента корреляции r | Теснота связи |
|---|----------------|
| 0,1 – 0,3 | Слабая |
| 0,3 – 0,5 | Умеренная |
| 0,5 – 0,7 | Заметная |
| 0,7 – 0,9 | Высокая |
| 0,9 – 0,99 | Весьма высокая |



ЗАКОН БОЛЬШИХ ЧИСЕЛ, **Утверждение теоремы Чебышева означает, что при выполнении определённых условий среднее арифметическое случайных величин в первом приближении сходится к среднему арифметическому их математических ожиданий, т.е. случайные величины $Y_n = \frac{1}{n} \sum_{k=1}^n X_k$ при $n \rightarrow \infty$ сходятся по вероятности к $a_n = \frac{1}{n} \sum_{k=1}^n M X_k$.**