Статистические ряды распределения

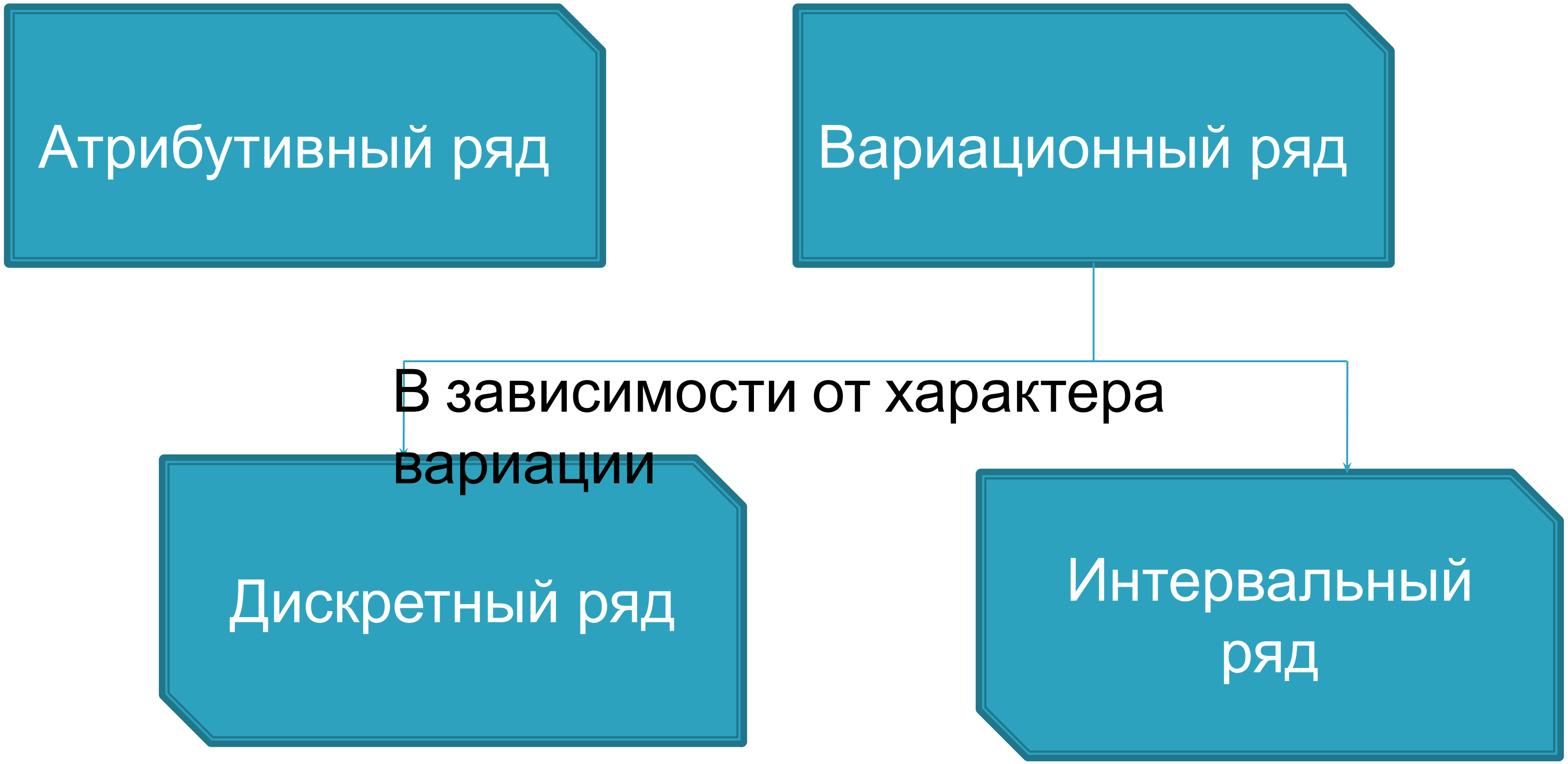
Большинство встречающихся на практике величин принимают неодинаковые значения у различных членов совокупности

- **Статистический ряд распределения** – это упорядоченное распределение единиц совокупности на группы по определенному варьирующемуся признаку (стаж работы, возраст, пол и т.д.)

С помощью статистического ряда распределения:

- Характеризуют состав (структуру), изучаемого явления
- Рассматривают вопрос об однородности совокупности
- Рассматривают вопрос о границах варьирования единиц совокупности и закономерностях ее распределения

Виды статистических рядов распределения и их элементы



Атрибутивный ряд

- Ряд построенный по атрибутивному признаку (пол, занятость, национальность, профессия и пр.)

Распределение студентов I курса экономического факультета по полу

Группа студентов, пол	**Число студентов**	**Удельный вес в общей численности, %**
Женщины	90	60,0
Мужчины	60	40,0
Всего	150	100,0

Вариационный ряд

Вариационный ряд – это ранжированный в порядке возрастания или убывания ряд вариантов с соответствующими им весами.

Применение дискретного ряда распределения

Число детей в семье	Количество семей	Удельный вес в общей численности, %
1	700	70,0
2	250	25,0
Более 2	50	5,0
Всего	1000	100,0

Характеристики вариационных рядов:

1. Варианты – это числовые значения количественного признака в вариационном ряду распределения (положительные, отрицательные, относительные, абсолютные)

2. Частоты – это численности отдельных вариантов или каждой группы вариационного ряда, т.е. числа, показывающие насколько часто встречаются те или иные варианты в ряду распределения

Сумма всех частот называется **объемом совокупности** и равна числу элементов всей совокупности

Характеристики вариационных рядов:

3. Частости – это частоты, выраженные в виде относительных величин (долях или процентах)

Сумма частостей равна 1 или 100%

Замена частот частостями позволяет сравнивать ряды с разным число наблюдений

Дискретный вариационный ряд

- В основе этого ряда лежит дискретный (прерывный) признак, т.е. значения признака отличаются друг от друга не менее чем на некоторую постоянную величину

Число детей в семье	Количество семей	Удельный вес в общей численности, %
1	700	70,0
2	250	25,0
Более 2	50	5,0
Всего	1000	100,0

Интервальный вариационный ряд

- В основе этого ряда лежит непрерывный признак, который может принимать любые значения (температура воздуха, объем выручки)

Численность работающих, чел.	Число торговых предприятий	Удельный вес, % к итогу
50-100	24	15,00
100-150	36	22,50
150-200	50	31,25
200-250	28	17,50
250 и выше	22	13,75
Всего	160	100,00

Первый шаг построения вариационного ряда распределения

- Ранжирование – расположение всех вариантов в возрастающем или убывающем порядке

Например стаж работы рабочих бригады:

2, 4, 5,3,15,6,5,9,7,14,8,5,9,10,11,4,2,3,4,6,5,13,10,1

Ранжированный ряд:

1,2,2,3,3,4,4,4,5,5,5,5,6,6,6,7,8,9,9,10,10,11,13,14,15

Строим дискретный ряд

Варианты (x)	Частоты (f)	Частости, в %	Частости, в долях
1	1	4,0	0,04
2	2	8,0	0,08
3	2	8,0	0,08
4	3	12,0	0,12
5	4	16,0	0,16
6	3	12,0	0,12
7	1	4,0	0,04
8	1	4,0	0,04
9	2	8,0	0,08
10	2	8,0	0,08
11	1	4,0	0,04
12	0	0,0	0
13	1	4,0	0,04
14	1	4,0	0,04
15	1	4,0	0,04
Итого:	25	100,0	1,00

Строим интервальный ряд (как группировку)

- Вычисляем количество интервалов по формуле Стерджесса $n = 1 + 3.322\lg(N)$
- Вычисляем величину интервала
- Строим таблицу:

n=1+3,322lg25=5,6 примерно 5

h=(15-1)/5=2,8 примерно 3

x	До 3 лет	3-6 года	6-9 лет	9-12 лет	12-15 лет
f	3	9	5	5	3

Графическое изображение рядов распределения

Полигон – графическое изображение вариационных дискретных рядов:

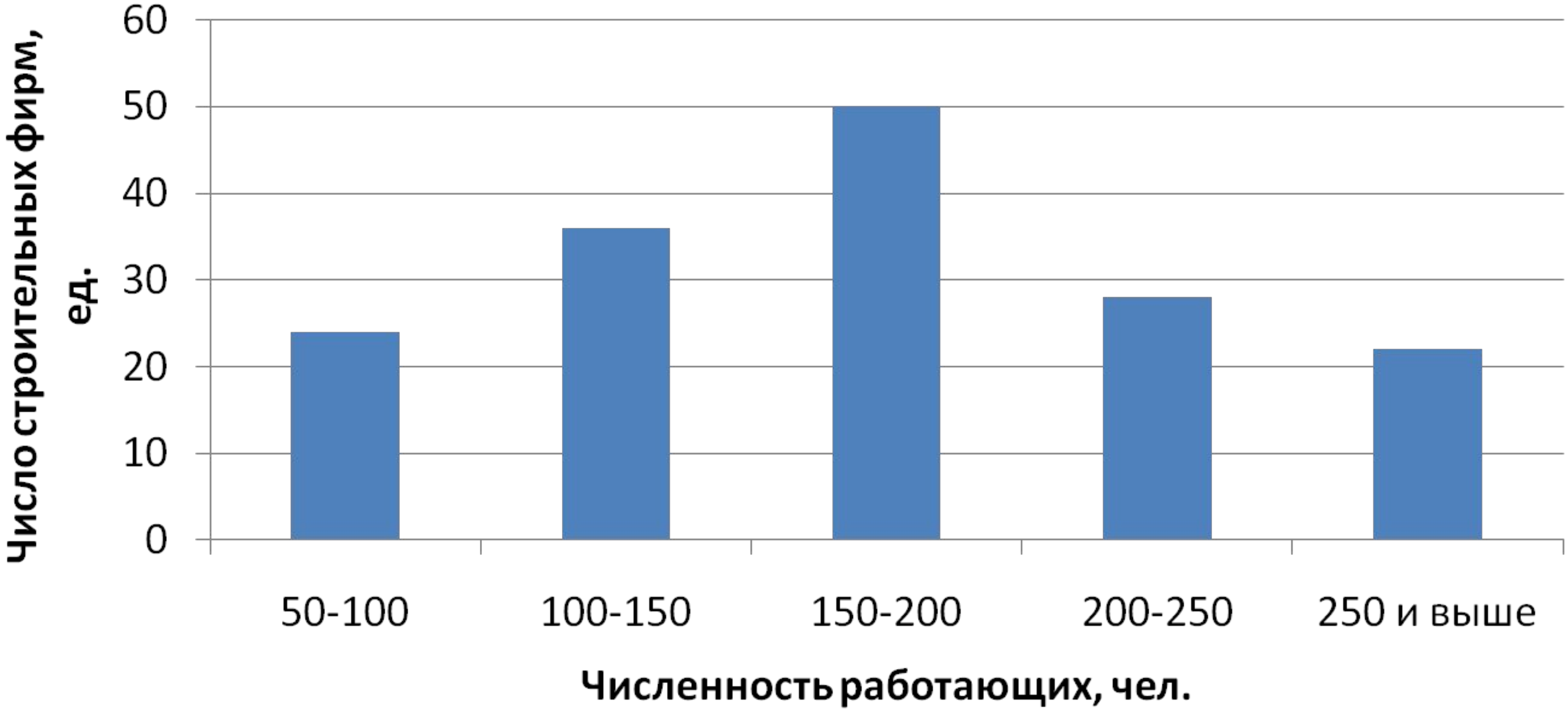
Ось абсцисс – ранжированные значения вариационного признака

Ось ординат – выражение численности каждого варианта (величины частот)

Полигон распределения работников по стажу работы

- Гистограмма - графическое изображение вариационных интервальных рядов
- Ось абсцисс – отображение величин интервалов
- Частоты описываются прямоугольниками, построенными на соответствующих интервалах, высота которых пропорциональна частотам

Гистограмма распределения торговых предприятий города по среднесписочной численности работающих



Формы статистических распределений

- Распределение называется симметричным если веса любых вариантов, равноотстоящих от среднего, равны между собой.
- Умеренно ассиметричные – это распределения у которых частоты, находящиеся по одну сторону от наибольшей, больше (или меньше) частот, находящихся по другую сторону

- Крайне ассиметричными называются распределения, у которых частоты или все время возрастают, или все время убывают
- При U- образном распределении частоты сначала убывают, а затем возрастают.

Эмпирическая функция распределения

Эмпирической функцией распределения *(функция распределения выборки) называется* $F^*(x)$*, определяющую для каждого значения* ***x*** *относительную частоту события* $X<x$.

$F^*(x)>n_x/n$; n_x – число вариант, меньше ***x***, *n* – объем выборки.

Свойства функции распределения

- значения $F^{*}(x)$ [0;1]
- $F^{*}(x)$ – функция неубывающая: $F^{*}(x_2) > F^{*}(x_1)$, если $x_2 > x_1$
- если x_1 – наименьшая варианта, $F^{*}(x_1)=0$

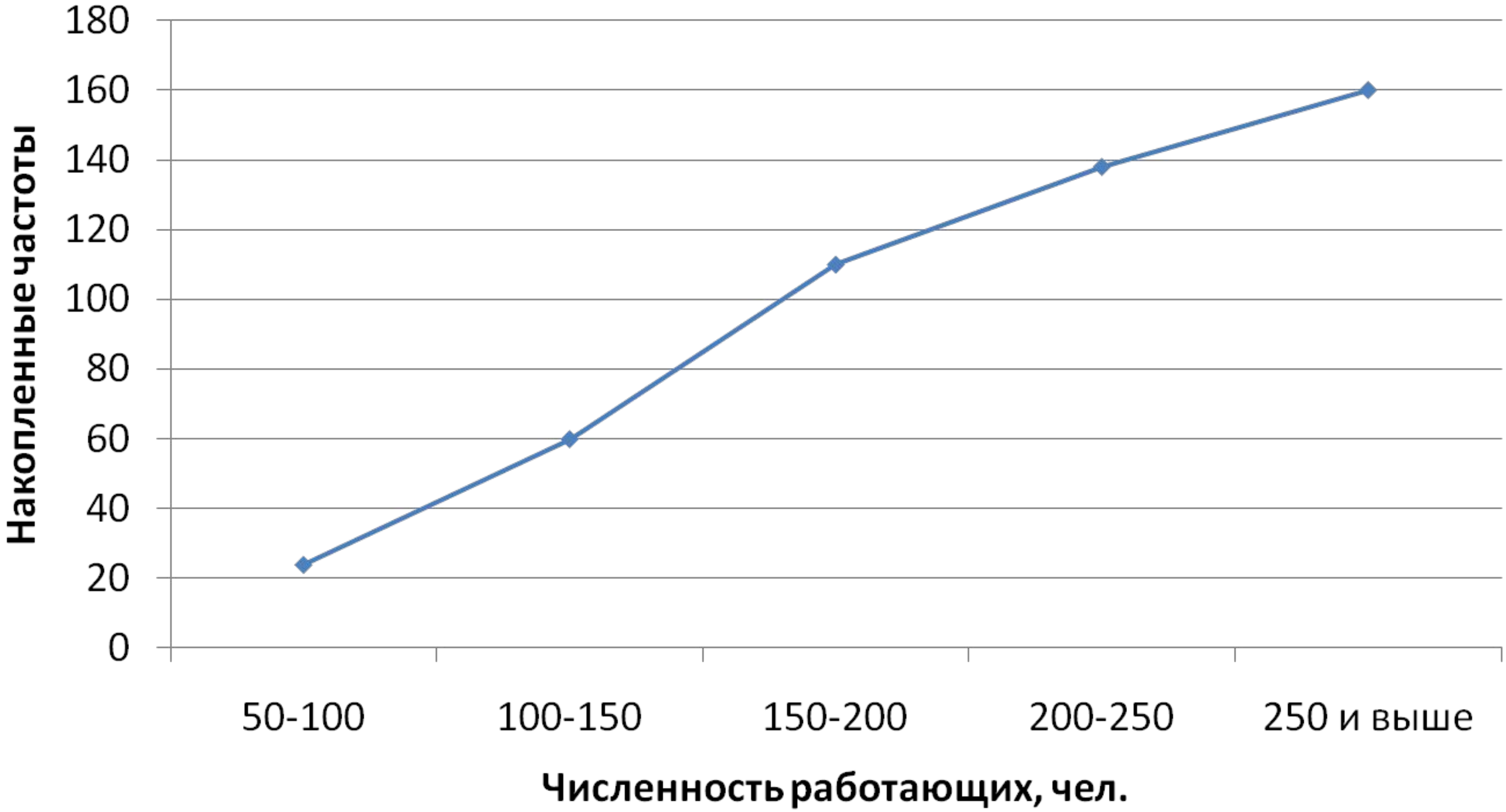
  если x_k – наибольшая, то $F^{*}(x_1)=1$.

Графическое представление

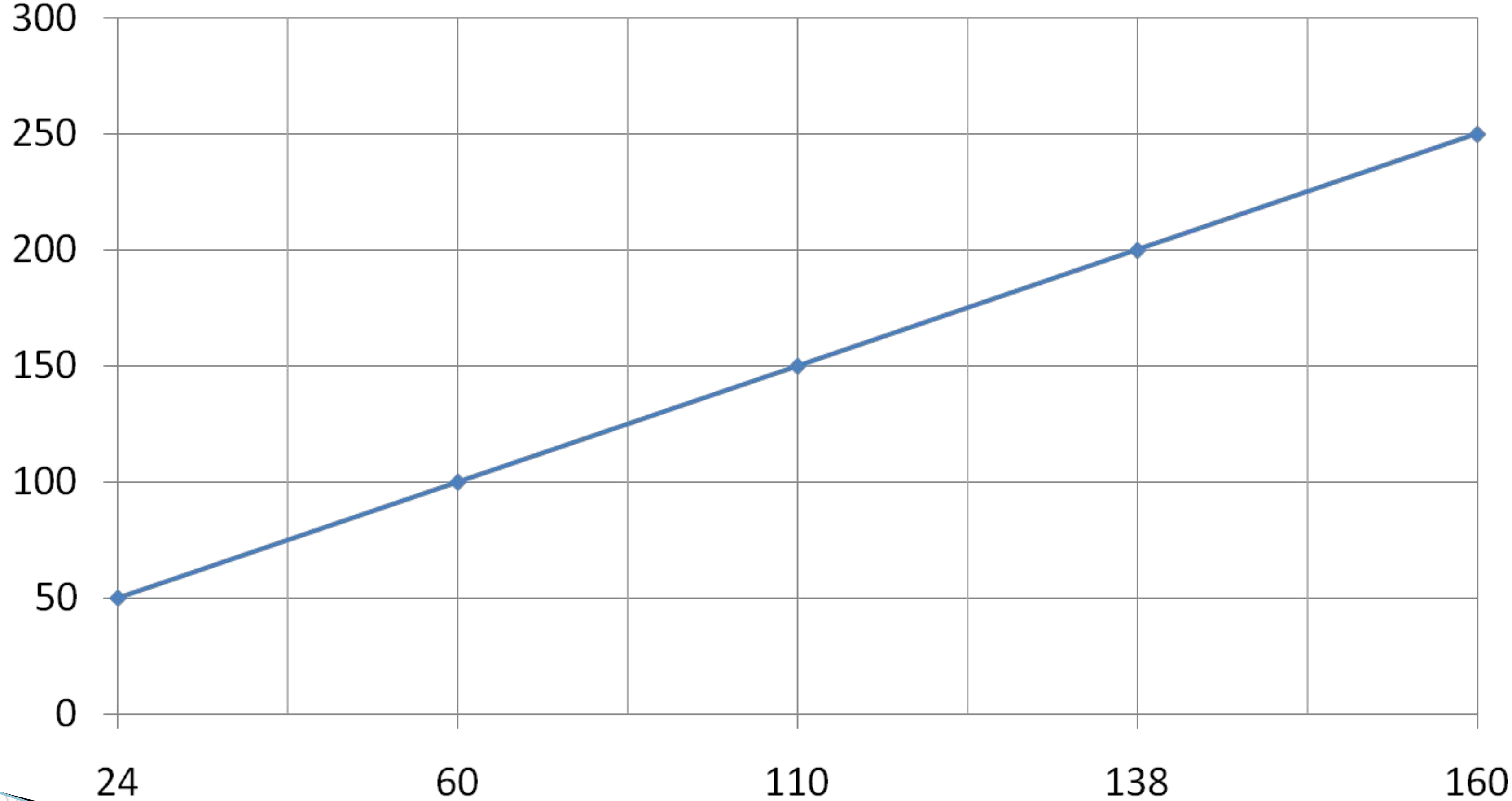
- Кумулята – для изображения ряда накопленных частот

- Огива – это кумулята, в которой оси поменяны местами

Пример кумуляты



Пример огивы



Меры уровня, или средние

- Наиболее употребительными в статистических исследованиях являются три вида средних: средняя арифметическая, мода и медиана.
- **средняя арифметическая**:

. Если вместо частоты заданы частости q_i, то формула имеет вид

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i q_i}{\sum_{i=1}^{k} q_i}, \qquad \text{где } \sum_{i=1}^{k} q_i = 1, \text{ или } 100\%.$$

Меры уровня

- **Медианой** (обозначим M_e) называется такое значение варьирующего признака, которое приходится на середину вариационного ряда.
- При нахождении медианы дискретного вариационного ряда могут возникнуть два случая: 1) число вариант нечетно ($k=2m+1$), 2) число вариант четно ($k=2m$). В первом случае $M_e=x_{m+1}$, т. е. медиана равна центральной (срединной) варианте ряда, во втором случае M_e,$=(x_m+x_{m+1})/2$, т.е. медиана принимается равной полу сумме находящихся в середине ряда вариант.

Меры уровня

- **Модой** (обозначим M_o) называется варианта, наиболее часто встречающаяся в данном вариационном ряду

Показатели вариации

- **Размах вариации** показывает разность между наибольшим и наименьшим значениями признака ($R=x_{max}-x_{min}$). Достоинством этого показателя является простота расчета. Однако возможности его применения ограничены, так как эта характеристика является наиболее грубой из всех мер рассеяния.

Показатели вариации

- **Дисперсия**, или **средний квадрат отклонения** (обозначим σ^2) есть средняя арифметическая из квадратов отклонений вариант от их средней арифметической, т. е. в математической записи

$$\sigma^2 = \frac{\sum_{i=1}^{k} (x_i - \overline{x})^2 q_i}{\sum_{i=1}^{k} q_i}, \qquad (4.6)$$

Показатели вариации

- Часто для исследования удобно представлять меру рассеяния в тех же единицах измерения, что и варианты. Тогда вместо дисперсии используют среднее квадратичное отклонение, которое является квадратным корнем из дисперсии, т. е. среднее квадратичное отклонение вычисляется по формуле

$$\sigma = \sqrt{\frac{\sum_{i=1}^{k} (x_i - \overline{x})^2 q_i}{\sum_{i=1}^{k} q_i}} \qquad (4.7)$$

Генеральная совокупность и выборка

- Вся подлежащая изучению совокупность объектов называется генеральной совокупностью
- Та часть объектов которая попала на проверку или исследование называется выборочной совокупностью или выборкой.
- Число элементов в генеральной совокупности и в выборке называется объемом.

Типы выборок

- **Собственно-случайная**
- **Механическая выборка** (члены из генеральной совокупности отбираются через определенный интервал)
- **Типическая** (генеральная совокупность разбита на непересекающиеся группы, а затем образуются собственно-случайные выборки из каждой группы)

Характеристики генеральной и выборочной совокупности

- Средняя арифметическая распределения признака генеральной совокупности называется генеральной средней, а дисперсия этого распределения – генеральной дисперсией

$$\overline{x_0} = \frac{\sum_{i=1}^{m} x_i N_i}{N} \qquad \sigma_0^2 = \frac{\sum_{i=1}^{m} (x_i - \overline{x_0})^2 N_i}{N}$$

Характеристики генеральной и выборочной совокупности

- Средняя арифметическая распределения признака в выборочной совокупности называется выборочной средней, а дисперсия этого распределения – выборочной дисперсией

$$\overline{x_0} = \frac{\sum_{i=1}^{m} x_i n_i}{n} \qquad \sigma_0^2 = \frac{\sum_{i=1}^{m} (x_i - \overline{x_0})^2 n_i}{n}$$

Характеристики генеральной и выборочной совокупности

- Генеральной долей *p* признака *A* называется отношение числа *M* членов генеральной совокупности с признаком *A* к ее объему

$$p = \frac{M}{N}$$

Выборочной долей признака *A* называется отношение числа *m* членов выборочной совокупности с признаком *A* к ее объему

$$\omega = \frac{m}{n}$$

Случайные величины

Случайной называют величину, которая в результате испытания примет одно и только одно возможное значение, наперед неизвестное и зависящее от случайных событий, которые заранее не могут быть учтены.

Обозначения случайных величин: *X, Y, Z;* значения — *x, y, z.*

Дискретной (прерывной) называется случайная величина, которая принимает отдельные, изолированные возможные значения с определенными вероятностями.

Число возможных значений дискретной случайной величины может быть как конечным, так и бесконечным (счетным).

Для задания дискретной случайной величины недостаточно перечислить все ее возможные значения,

нужно еще указать их вероятности.

Законом распределения дискретной случайной величины называют соответствие между возможными значениями и их вероятностями. Его можно задать в виде таблицы, аналитически и графически.
При табличном задании закона распределения дискретной случайной величины первая строка таблицы содержит возможные значения (как правило, в порядке возрастания), а вторая строка — их вероятности.

X	x_1	x_2	…	x_{k-1}	x_k
P	p_1	p_2	...	p_{k-1}	p_k

Поскольку в одном испытании случайная величина принимает одно и только одно возможное
значение, заключаем, что события $X = x_1$, $X = x_2$, …, $X = x_k$, образуют полную группу,
следовательно, сумма вероятностей этих событий равна единице:

$$\sum_{i=1}^{k} P(X = x_i) = \sum_{i=1}^{k} p_i = 1$$

Числовые характеристики дискретных случайных величин

Математическое ожидание дискретной случайной величины

Математическим ожиданием дискретной случайной величины *X* называется сумма произведений всех ее возможных значений на соответствующие вероятности

$$M(X) = \sum_{i=1}^{k} x_i p_i = x_i p_i + x_2 p_2 + \dots + x_k p_k$$

Замечание. Из определения следует, что математическое ожидание дискретной случайной величины есть **неслучайная** (постоянная) величина.

Вероятностный смысл математического ожидания

Пусть проведено n испытаний, в которых случайная величина X приняла m_1 раз значение x_1, m_2 раз значение x_2, …, m_k раз значение x_k, причем $m_1 + m_2 + \ldots + m_k = n$. Тогда сумма всех значений, принятых X равна $m_1x_1 + m_2 x_2 + \ldots + m_kx_k$. Среднее арифметическое всех значений, принятых этой случайной величиной

$$\overline{X} = \frac{x_1m_1 + x_2m_2 + \ldots + x_km_k}{n} = x_1\frac{m_1}{n} + x_2\frac{m_2}{n} + \ldots + x_k\frac{m_k}{n}$$

Заметим, что $m_i/n = w_i$ — относительной частоте значения x_i. допустим, что число испытаний велико. Тогда $w_i \approx p_i$. Заменяя в последнем выражении относительные частоты вероятностями, получим

$$\overline{X} \approx x_1p_1 + x_2p_2 + \ldots + x_kp_k$$

$$\overline{X} \approx M(X)$$

Замечание.

Математическое ожидание приближенно равно (тем точнее, чем больше число испытаний) среднему арифметическому наблюдаемых значений случайной величины.

$$x_C = \frac{\sum x_ip_i}{\sum p_i} \qquad \sum p_i = 1 \quad \Rightarrow x_C = M(X)$$

Свойства математического ожидания

1. Математическое ожидание постоянной величины *C* равно самой постоянной:
M(*C*) = *C.*

Доказательство. Рассмотрим постоянную величину *C* как дискретную случайную величину, которая имеет одно возможное значение *C* и принимает его с вероятностью *p* = 1. Следовательно, *M*(*C*) = *C* · 1 = *C*.

2. Постоянный множитель можно выносить за знак математического ожидания:
M(*CX*) = *C* · *M*(*X*).

Доказательство.

CX	Cx_1	Cx_2	…	Cx_k
P	p_1	p_2	...	p_k

$$M(CX) = Cx_1p_1 + Cx_2p_2 + \dots + Cx_kp_k = C(x_1p_1 + x_2p_2 + \dots + x_kp_k) = CM(X)$$

3. Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий:
M(*XY*) = *M*(*X*) · *M*(*Y*).

Доказательство.

XY	x_1y_1	x_1y_2	x_2y_1	x_2y_2
P	p_1g_1	p_1g_2	p_2g_1	p_2g_2

$$M(XY) = x_1y_1 \cdot p_1g_1 + x_1y_2 \cdot p_1g_2 + x_2y_1 \cdot p_2g_1 + x_2y_2 \cdot p_2g_2 =$$

$$= y_1g_1(x_1p_1 + x_2p_2) + y_2g_2(x_1p_1 + x_2p_2) = (x_1p_1 + x_2p_2)(y_1g_1 + y_2g_2) = M(X)M(Y)$$

4. Математическое ожидание суммы двух случайных величин равно сумме математических ожиданий слагаемых: *M*(*X* + *Y*) = *M*(*X*) + *M*(*Y*).

Доказательство.

$X+Y$	x_1+y_1	x_1+y_2	x_2+y_1	x_2+y_2
P	p_{11}	p_{12}	p_{21}	p_{22}

$$M(X+Y)=(x_1+y_1)p_{11}+(x_1+y_2)p_{12}+(x_2+y_1)\cdot p_{21}+(x_2+y_2)p_{22}=$$

$$=x_1(p_{11}+p_{12})+x_2(p_{21}+p_{22})+y_1(p_{11}+p_{21})+y_2(p_{12}+p_{22})$$

Докажем, что $p_{11}+p_{12}=p_1$.

Событие {$X = x_1$}влечет за собой событие {$X + Y = (x_1 + y_1$ или $x_1 + y_1)$} и обратно

$$P(X=x_1)=p_1$$

$$P\big(X+Y=(x_1+y_1\,\text{или}\,x_1+y_2)\big)=p_{11}+p_{12}$$

$$p_1=p_{11}+p_{12}$$

Аналогично,

$$p_2=p_{21}+p_{22},\ g_1=p_{11}+p_{21},\ \ g_2=p_{12}+p_{22}.$$

$$M(X+Y)=(x_1p_1+x_2p_2)+(y_1g_1+y_2g_2)=M(X)+M(Y)$$

Дисперсия дискретной случайной величины

Пусть X — случайная величина и $M(X)$ — ее математическое ожидание. *Отклонением* называют случайную величину X - $M(X)$, возможные значения которой равны разностям между возможными значениями случайной величины и ее математическим ожиданием, а вероятности величины X - $M(X)$ равны вероятностям величины X.

X	x_1	x_2	…	x_k
P	p_1	p_2	...	p_k

$X-M(X)$	$x_1-M(X)$	$x_2-M(X)$	…	$x_k-M(X)$
P	p_1	p_2	...	p_k

Теорема. Математическое ожидание отклонения равно 0: $M[X - M(X)] = 0$.

Дисперсией (рассеянием) дискретной случайной величины называют математическое ожидание
квадрата отклонения случайной величины от ее математического ожидания:

$$D(X) = M\left([X - M(X)]^2\right)$$

X	x_1	x_2	…	x_k
P	p_1	p_2	...	p_k

$[X-M(X)]^2$	$[x_1-M(X)]^2$	$[x_2-M(X)]^2$	…	$[x_k-M(X)]^2$
P	p_1	p_2	...	p_k

$$D(X) = M\left([X - M(X)]^2\right) = \sum\left([x_i - M(X)]^2 p_i\right)$$

Теорема. Дисперсия равна разности между математическим ожиданием квадрата случайной величины X и квадратом ее математического ожидания: $D(X) = M(X^2) - [M(X)]^2$.

Доказательство. Поскольку математическое ожидание $M(X)$ — есть величина постоянная, то $2\,M(X)$ и $[M(X)]^2$ — также постоянные величины. Поэтому

$$D(X) = M\left([X - M(X)]^2\right) = M[X^2 - 2X \cdot M(X) + M^2(X)] = M(X^2) - 2M(X) \cdot M(X) + M^2(X)$$

$$= M(X^2) - [M(X)]^2.$$

Пример. Найти дисперсию случайной величины X, заданной законом распределения

X	1	2	5
P	0,3	0,5	0,2

Решение. Математическое ожидание: $M(X) = 1 \cdot 0{,}3 + 2 \cdot 0{,}5 + 5 \cdot 0{,}2 = 2{,}3$.
Закон распределения квадрата случайной величины

X^2	1	4	25
P	0,3	0,5	0,2

Математическое ожидание квадрата случайной величины: $M(X) = 1 \cdot 0{,}3 + 4 \cdot 0{,}5 + 25 \cdot 0{,}2 = 7{,}3$.
Дисперсия: $D(X) = 7{,}3 - 2{,}3^2 = 2{,}01$.

1. Дисперсия постоянной величины *C* равна нулю: $D(C) = 0$.

2. Постоянный множитель можно выносить за знак дисперсии, возводя его в квадрат: $D(CX) = C^2D(X)$.

3. Дисперсия суммы двух независимых случайных величин равна сумме дисперсий этих величин: $D(X + Y) = D(X) + D(Y)$.

4. Дисперсия разности двух независимых случайных величин равна

сумме дисперсий этих величин:

$$D(X - Y) = D(X) + D(Y).$$

$$D(X-Y) = D(X+(-Y)) \overset{\substack{\text{св}-\text{во}\\ 3}}{=} D(X)+D(-Y) \overset{\substack{\text{св}-\text{во}\\ 2}}{=} D(X)+(-1)^2 D(Y) = D(X)+D(Y)$$

Среднее квадратическое отклонение

Дисперсия имеет размерность квадрата случайной величины. Для того чтобы иметь показатель рассеяния случайной величины той же размерности, что и размерность случайной величины, извлекают корень квадратный из дисперсии.

Средним квадратическим отклонением случайной величины *X* называют квадратный корень из дисперсии

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

Теорема. Среднее квадратическое отклонение суммы конечного числа взаимно независимых случайных величин равно
квадратному корню из суммы квадратов средних квадратических отклонений этих величин

$$\sigma(X_1 + X_2 + \dots + X_n) = \sqrt{\sigma^2(X_1) + \sigma^2(X_2) + \dots + \sigma^2(X_n)}.$$

Начальные и центральные теоретические моменты

X	1	2	5	100
P	0,6	0,2	0,19	0,01

$$M(X) = 1\cdot 0{,}6 + 2\cdot 0{,}2 + 5\cdot 0{,}19 + 100\cdot 0{,}01 = 2{,}95$$

X^2	1	4	25	10000
P	0,6	0,2	0,19	0,01

$$M(X) = 1\cdot 0{,}6 + 4\cdot 0{,}2 + 25\cdot 0{,}19 + 10000\cdot 0{,}01 = 106{,}15$$

Начальным моментом порядка k случайной величины *X* называют математическое ожидание величины X^k: $\nu_k = M(X^k)$

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

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

$$\mu_k = M\left[(X - M(X))^k\right]$$

В частности,

$$\mu_1 = M[(X - M(X))] = 0, \qquad \mu_2 = M\left[(X - M(X))^2\right] = D(X)$$

$$\mu_2 = \nu_2 - \nu_1^2,\ \mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3,\ \mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4.$$

Функция распределения

Функцией распределения называют функцию $F(x)$, определяющую вероятность того, что случайная величина X в результате испытания примет значение, меньшее x, т.е.

$$F(x) = P(X < x).$$

Свойства функции распределения

1. Значения функции распределения принадлежат отрезку [0; 1]:

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

2. Функция распределения непрерывна слева.

3. $F(x)$ — неубывающая функция, т.е. $F(x_1) \le F(x_2)$, если $x_1 < x_2$.

Доказательство. Пусть $x_1 < x_2$.

$$\{X < x_2\} = \{X < x_1 \text{ и } x_1 \le X < x_2\} \qquad (**)$$

$$\Rightarrow P(X < x_2) = P(X < x_1) + P(x_1 \le X < x_2)$$

$$\Rightarrow P(X < x_2) - P(X < x_1) = P(x_1 \le X < x_2)$$

$$\Rightarrow F(x_2) - F(x_1) = P(x_1 \le X < x_2)$$

$$\lim_{x\to-\infty} F(x) = 0; \quad \lim_{x\to\infty} F(x) = 1.$$

Поскольку $P(x_1 \le X < x_2) \ge 0$, то

$$\Rightarrow F(x_1) \le F(x_2)$$

Итак, *каждая функция распределения является неубывающей, непрерывной слева и удовлетворяющей условиям* $F(-\infty) = 0$, $F(+\infty) = 1$.
Верно и обратное: *каждая функция, удовлетворяющая указанным условиям, может рассматриваться как функция распределения некоторой случайной величины.*

Для дискретной случайной величины, заданной законом распределения

X	x_1	x_2	…	x_k
P	p_1	p_2	...	p_k

функция распределения $F(x)$ задается равенством

$$F(x) = \sum_{x_i < x} p_i .$$

Таким образом, функция распределения дискретной случайной величины является
ступенчатой функцией со скачками высотой p_i в точках x_i.

Непрерывная случайная величина

Случайная величина называется *непрерывной*, если существует неотрицательная функция $p(x)$, удовлетворяющая при любых x равенству

$$F(x) = \int_{-\infty}^{x} p(z)dz.$$

Функция $p(x)$ называется *плотностью распределения вероятностей*.
Если $F(x)$ абсолютно непрерывна, а тем более, дифференцируема при всех x, то ее производная и является плотностью распределения:

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

Функция распределения иногда называется *интегральной*, а плотность — *дифференциальной функцией распределения.*

Если возможные значения случайной величины принадлежат интервалу [*a, b*], то

$$\int_{a}^{b} p(x)dx = 1.$$

Если возможные значения случайной величины принадлежат всей числовой оси, то

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

Свойства функции распределения непрерывной случайной величины

1. Вероятность того, что **непрерывная** случайная величина X примет одно определенное значение равно 0.

Доказательство. Положим в (**) $x_2 = x_1 + \Delta x$. Тогда

$$P(x_1 \le X < x_1 + \Delta x) = F(x_1 + \Delta x) - F(x_1).$$

Пусть $\Delta x \to 0$. Тогда, в силу непрерывности $F(x)$

$$F(x_1 + \Delta x) - F(x_1) \to 0$$

$$\Rightarrow P(X = x_1) = 0.$$

$$P(a \le X < b) = P(a < X < b) = P(a < X \le b) = P(a \le X \le b).$$

2. Если возможные значения непрерывной случайной величины принадлежат интервалу (a, b), то

а) $F(x) = 0$ при $x \le a$;

б) $F(x) = 1$ при $b \le x$.

Вероятность попадания непрерывной случайной величины в заданный интервал

Теорема. Вероятность того, что непрерывная случайная величина *X* примет значение, принадлежащее интервалу (*a, b*), равна определенному интегралу от плотности распределения, взятому в пределах от *a* до *b*:

$$P(a < x < b) = \int_a^b p(x)dx.$$

Доказательство. Воспользуемся соотношением (**):

$P(a \le X < b) = F(b) - F(a)$

По формуле Ньютона-Лейбница

$$F(b) - F(a) = \int_a^b F'(x)dx = \int_a^b p(x)dx$$

Таким образом,

$$P(a \le x < b) = \int_a^b p(x)dx.$$

Поскольку $P(a \le X < b) = P(a < X < b)$, то

$$P(a < x < b) = \int_a^b p(x)dx.$$

Числовые характеристики непрерывных случайных величин

Математическим ожиданием непрерывной случайной величины X, возможные значения которой принадлежат интервалу [*a, b*], называют определенный интеграл

$$M(X) = \int_a^b x p(x) dx.$$

Если возможные значения случайной величины принадлежат всей числовой оси, то

$$M(X) = \int_{-\infty}^{\infty} x p(x) dx.$$

(предполагается, что несобственный интеграл сходится абсолютно, т.е. существует интеграл

$$\int_{-\infty}^{\infty} |x| p(x) dx.$$

Дисперсией непрерывной случайной величины называют математическое ожидание квадрата ее отклонения.

Если возможные значения случайной величины принадлежат интервалу [*a, b*], то

$$D(X)=\int_a^b [x-M(x)]^2\,p(x)dx.$$

Если возможные значения случайной величины принадлежат всей числовой оси, то

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

Замечание.

$$D(X)=\int_a^b x^2 p(x)dx-[M(x)]^2. \quad D(X)=\int_{-\infty}^{\infty} x^2 p(x)dx-[M(x)]^2.$$

Среднее квадратическое отклонение непрерывной случайной величины

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

Медианой непрерывной случайной величины называется такое ее значение *m*, при котором *F*(*m*)=0,5;другими словами,

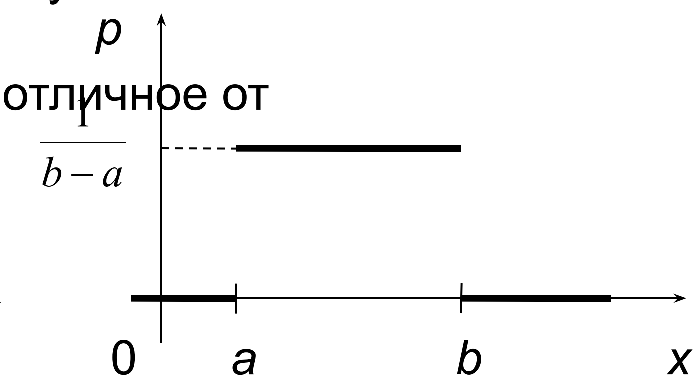
$$P(X<m)=P(X>m)=0{,}5.$$

Квантилью порядка *p* (0 < *p* < 1) называется корень уравнения *F*(*x*) = *p*.
Если случайная величина непрерывна, то *модой* распределения называют то значение аргумента, при котором плотность достигает максимума.
Модой дискретной случайной величины называют ее наиболее вероятное значение.

Равномерное распределение вероятностей

Распределение вероятностей называется *равномерным*, если на интервале, которому принадлежат все возможные значения случайной величины, плотность распределения сохраняет постоянное отличное от нуля значение:

$$p(x)=\begin{cases}0 & \text{при} \quad x\le a,\\ C & \text{при} \quad a<x\le b,\\ 0 & \text{при} \quad b<x.\end{cases}$$



Замечание.

$$p(x)=C\cdot f(x) \quad \int_a^b C\cdot f(x)dx=1 \Rightarrow C=\frac{1}{\int_a^b f(x)dx}$$

$$\int_a^b Cdx=1 \Rightarrow C=\frac{1}{\int_a^b dx} \Rightarrow C=\frac{1}{b-a} \qquad p(x)=\begin{cases}0 & \text{при} \quad x\le a,\\ \dfrac{1}{b-a} & \text{при} \quad a<x\le b,\\ 0 & \text{при} \quad b<x.\end{cases}$$

Числовые характеристики равномерно распределенной случайной величины

$$M(X)=\int_a^b xp(x)dx =\frac{1}{b-a}\int_a^b xdx = \qquad M(X)=\frac{a+b}{2}$$

$$D(X)=\int_a^b x^2 p(x)dx-[M(x)]^2.=\frac{1}{b-a}\int_a^b x^2dx-\left(\frac{a+b}{2}\right)^2= \qquad D(X)=\frac{(b-a)^2}{12}$$

Нормальное распределение

Нормальным называется распределение вероятностей непрерывной случайной величины, которое описывается плотностью

$$\varphi_0(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$$

Нормальное распределение определяется двумя параметрами: *a* и σ.

$$M(X)=\int_{-\infty}^{\infty}x\varphi_0(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}x\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}dx\quad\left|\begin{matrix}z=\dfrac{x-a}{\sigma}\\ dx=\sigma dz\end{matrix}\right|$$

$$=\frac{\sigma}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma z+a)e^{-\frac{z^2}{2}}dz=\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty}ze^{-\frac{z^2}{2}}dz+\frac{a}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz=\left|\int_{-\infty}^{\infty}ze^{-\frac{z^2}{2}}dz=0,\ \int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz=\sqrt{2\pi}\right|=a.$$

$$M(X)=a$$

$$D(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-a)^2\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}dx\quad\left|\begin{matrix}z=\dfrac{x-a}{\sigma}\\ dx=\sigma dz\end{matrix}\right|=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z\cdot ze^{-\frac{z^2}{2}}dz=\quad=\sigma^2$$

$$D(X)=\sigma^2$$

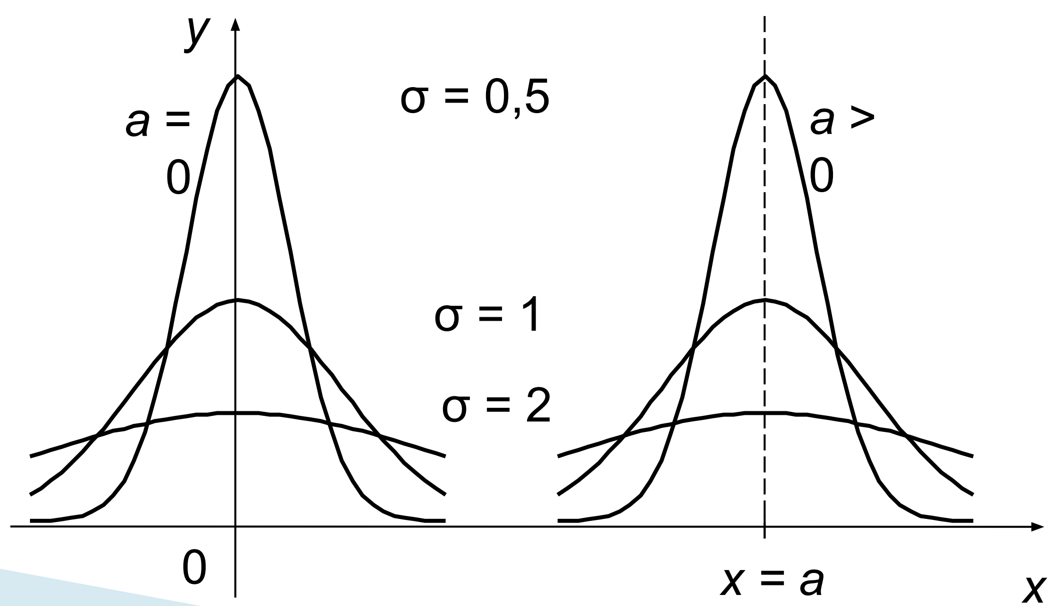
$$\sigma(X)=\sigma$$

Общим называется нормальное распределение с произвольными параметрами *a* и σ. *Нормированным* называется нормальное распределение с параметрами *a* = 0 и σ = 1. Таким образом, если *X* — нормальная величина, то *U* = (*x* – *a*)/σ — нормированная нормальная величина, причем *M*(*U*) = 0, *D*(*U*) = 1.

Плотность нормированного распределения (нормированная функция Гаусса)

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

График плотности нормального распределения называют *нормальной кривой* (*кривой Гаусса*)



1. Функция $F_0(x)$ общего нормального распределения

$$F_0(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(z-a)^2}{2\sigma^2}} dz$$

и функция $F(x)$ нормированного распределения

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz$$

связаны соотношением

$$F_0(x) = F\left(\frac{x-a}{\sigma}\right).$$

2. Вероятность попадания нормированной нормальной величины X в интервал $(0, x)$ вычисляется при помощи функции Лапласа:

$$P(0 < X < x) = \int_0^x \varphi(z)dz = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz = \Phi(x).$$

3. $\int_{-\infty}^{\infty} \varphi(x)dx = 1 \Rightarrow \int_{-\infty}^{0} \varphi(x)dx = \frac{1}{2} \quad \left(P(-\infty < X < 0) = \frac{1}{2}\right) \Rightarrow F(x) = \frac{1}{2} + \Phi(x)$

Вероятность попадания в заданный интервал нормальной случайной величины

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

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

$$\left| z = \frac{x-a}{\sigma}, \quad x = \sigma z + a, \quad dx = \sigma dz; \quad x_1 = \alpha \Rightarrow z_1 = \frac{\alpha - a}{\sigma}, \quad x_2 = \beta \Rightarrow z_2 = \frac{\beta - a}{\sigma} \right|$$

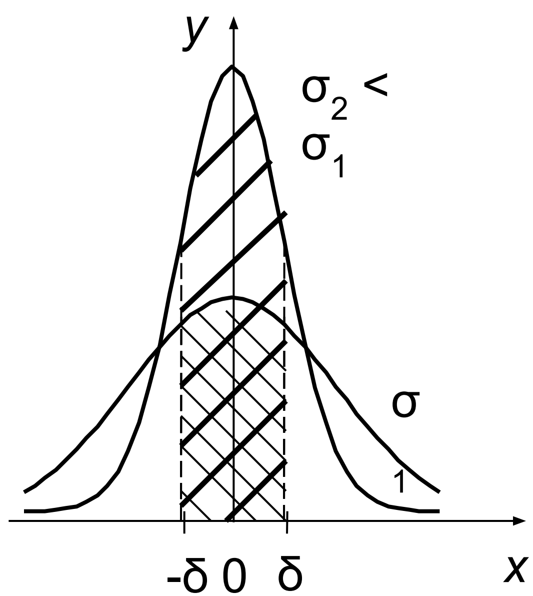
$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{z^2}{2}} \sigma dz = \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{\beta-a}{\sigma}} e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{0} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{\beta-a}{\sigma}} e^{-\frac{z^2}{2}} dz - \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{\alpha-a}{\sigma}} e^{-\frac{z^2}{2}} dz =$$

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

Вероятность заданного отклонения

$$P(|X-a|<\delta)=P(a-\delta<X<a+\delta)=\Phi\left(\frac{(a+\delta)-a}{\sigma}\right)-\Phi\left(\frac{(a-\delta)-a}{\sigma}\right)=\Phi\left(\frac{\delta}{\sigma}\right)-\Phi\left(-\frac{\delta}{\sigma}\right)=2\Phi\left(\frac{\delta}{\sigma}\right)$$

при *a* = 0 $\quad P(|X|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right)$



Правило «трех сигм»

$$P(|X-a|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right) \qquad \delta=\sigma\cdot t \qquad P(|X-a|<\sigma t)=2\Phi(t) \qquad t=3$$

$$P(|X-a|<3\sigma)=2\Phi(3)=2\cdot 0{,}49865=0{,}9973$$

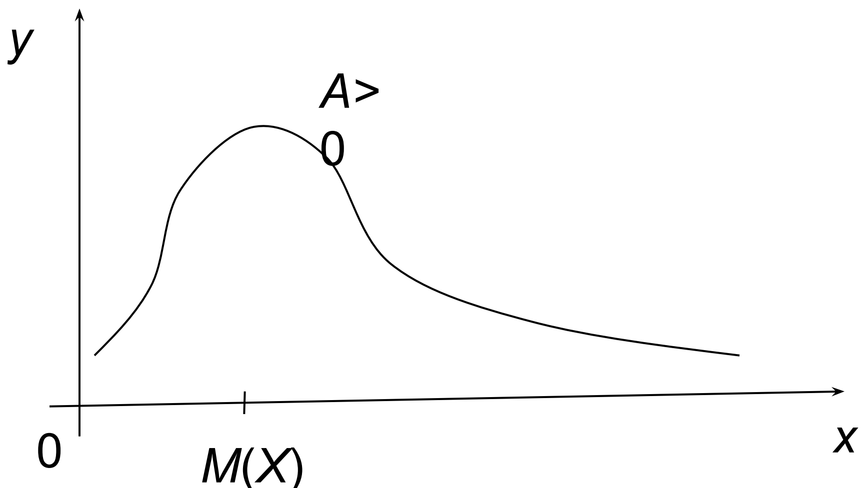
$$P(|X-a|<3\sigma)=0{,}9973 \qquad P(|X-a|>3\sigma)=0{,}0027$$

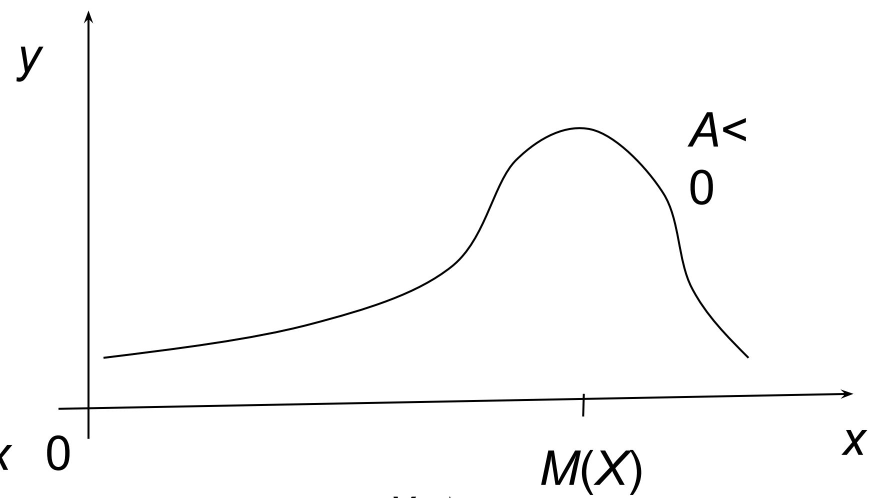
Вероятность того, что отклонение по абсолютной величине будет меньше утроенного среднего квадратического отклонения, равна 0,9973.
Другими словами, вероятность того, что абсолютная величина отклонения превысит утроенное среднее квадратическое отклонение очень мала, а именно равна 0,0027.
Это означает, что такое может произойти лишь в 0,27% случаев.
Такие события исходя из принципа невозможности маловероятных событий можно считать практически невозможными.
В этом и состоит сущность правила трех сигм:
если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения.

При изучении распределений, отличных от нормального, возникает необходимость качественно оценить это различие. С этой целью вводят специальные числовые характеристики, в частности, асимметрию и эксцесс. Для нормального распределения эти характеристики равны нулю. Поэтому небольшие значения асимметрии и эксцесса дают возможность предположить, что такое распределение близко к нормальному; большие значения указывают на значительное отклонение от нормального распределения.
Можно показать, что для симметричных распределений каждый центральный момент нечетного порядка равен нулю. Для несимметричных распределений такие моменты отличны от нуля. Поэтому центральный момент третьего порядка используется для оценки асимметрии.
Асимметрия положительна, если более пологая часть кривой распределения расположена справа от математического ожидания и отрицательна, если слева.
Для оценки «крутизны» подъема распределения по сравнению с нормальным используется характеристика, называемая эксцессом.
Если эксцесс больше нуля, то кривая такого распределения имеет более высокую и острую вершину, чем нормальная кривая, если эксцесс меньше нуля, то сраниваемая кривая имеет более низкую и плоскую вершину, чем нормальная.

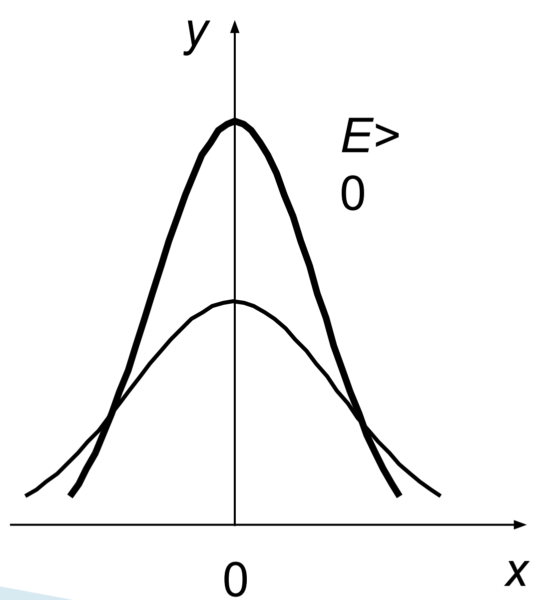
Асимметрия и эксцесс

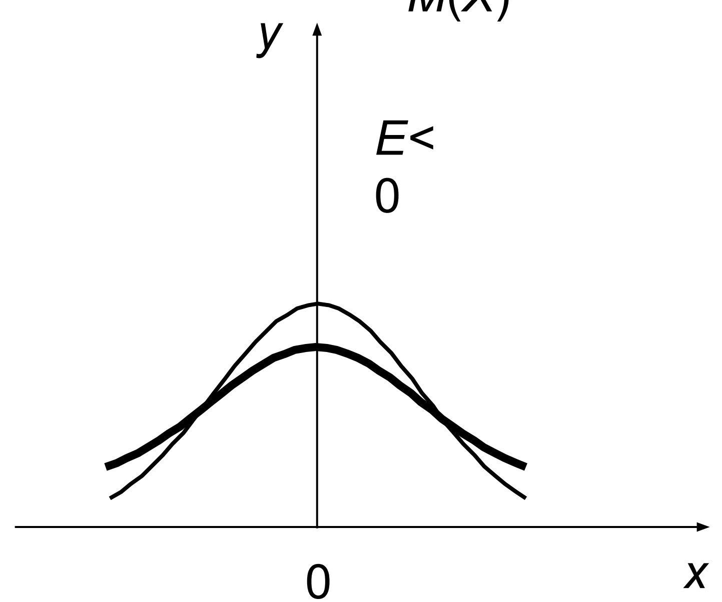
$$A = \frac{\mu_3}{\sigma^3}$$





$$E = \frac{\mu_4}{\sigma^4} - 3$$





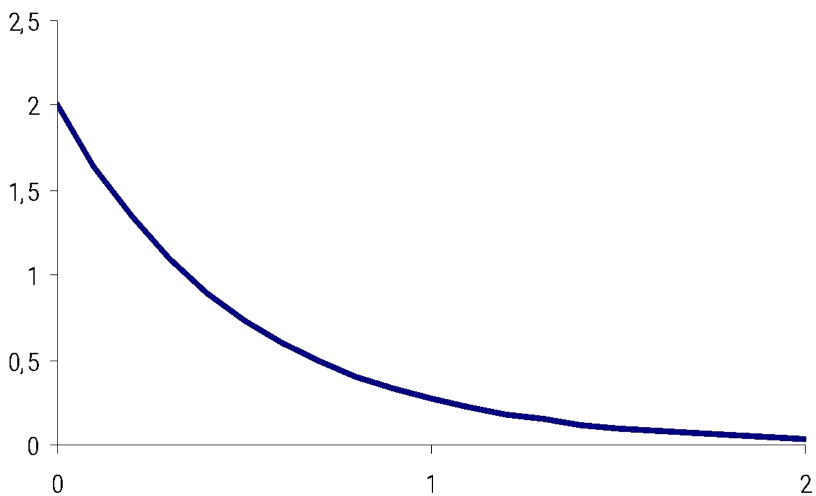
Показательное распределение

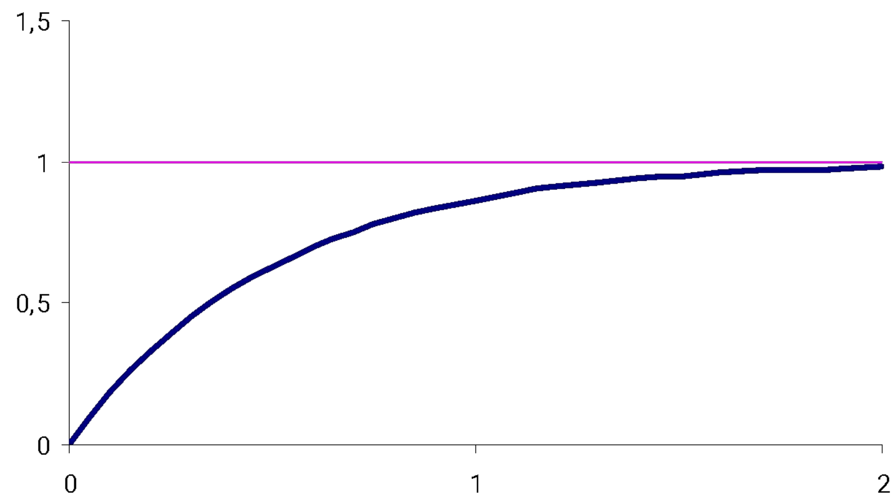
Показательным (*экспоненциальным*) называют распределение вероятностей непрерывной случайной величины *X*, которое описывается плотностью

$$p(x)=\begin{cases}0 & \text{при } x<0,\\ \lambda e^{-\lambda x} & \text{при } x\ge 0.\end{cases}$$

Функция распределения показательного закона:

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





Вероятность попадания в заданный интервал показательно распределенной случайной величины

Воспользуемся формулой $P(a < X < b) = F(b) - F(a)$.

Учитывая, что при $x \geq 0$

$$F(x) = 1 - e^{-\lambda x}, \qquad F(b) = 1 - e^{-\lambda b}, \qquad F(a) = 1 - e^{-\lambda a},$$

получаем

$$P(a < x < b) = e^{-\lambda a} - e^{-\lambda b}.$$

Числовые характеристики показательного распределения

Пусть случайная величина X распределена по показательному закону

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

Математическое ожидание

$$M(X) = \int_{-\infty}^{x} x p(x) dx = \lambda \int_{-\infty}^{x} x e^{-\lambda x} dx = \qquad M(X) = \frac{1}{\lambda}$$

Дисперсия

$$D(X) = \int_{-\infty}^{x} x^2 p(x) dx - [M(X)]^2 = \lambda \int_{-\infty}^{x} x^2 e^{-\lambda x} dx - \left(\frac{1}{\lambda}\right)^2; \quad \lambda \int_{-\infty}^{x} x^2 e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

$$D(X) = \frac{1}{\lambda^2}$$

Среднее квадратическое отклонение

$$\sigma(X) = \frac{1}{\lambda}.$$

Система двух случайных величин

Закон распределения двумерной случайной величины

Кроме одномерных случайных величин изучают случайные величины, возможные значения которой определяются двумя, тремя, …, *n* числами.

Такие величины называют соответственно *двумерными, трехмерными*, и т.д.

Двумерную случайную величину будем обозначать (*X*, *Y*).

Каждую из величин *X*, *Y* называют *составляющей* (*компонентой*) двумерной случайной величины.

Аналогично *n*-мерная случайная величина определяется как система *n* случайных величин.

Закон распределения двумерной случайной величины

Законом распределения дискретной двумерной случайной величины называют перечень возможных значений этой величины, т.е. пар (x_i, y_j) и их вероятностей $p_{ij} = p(x_i, y_j)$ $(i = 1, \ldots, n;\ j = 1, \ldots, m)$.

Обычно закон распределения двумерной дискретной случайной величины задают в виде таблицы.

Y	X						$P(Y = y_j)$
	x_1	x_2	…	x_i	…	x_n	
y_1	$p(x_1, y_1)$	$p(x_2, y_1)$	…	$p(x_i, y_1)$	…	$p(x_n, y_1)$	$P(Y = y_1) = \sum_{i=1}^{n} p(x_i, y_1)$
…	…	…	…	…	…	…	
y_j	$p(x_1, y_j)$	$p(x_2, y_j)$	…	$p(x_i, y_j)$	…	$p(x_n, y_j)$	
…	…	…	…	…	…	…	
y_m	$p(x_1, y_m)$	$p(x_2, y_m)$	…	$p(x_i, y_m)$	…	$p(x_n, y_m)$	
$P(X = x_i)$	$P(X = x_1) = \sum_{j=1}^{m} p(x_1, y_j)$	$P(X = x_2) = \sum_{j=1}^{m} p(x_2, y_j)$					$\sum\sum p(x_i, y_j) = 1$ $\sum P(X = x_i) = 1$ $\sum P(Y = y_j) = 1$

События $\{X = x_i,\ Y = y_j\}$ образуют полную группу

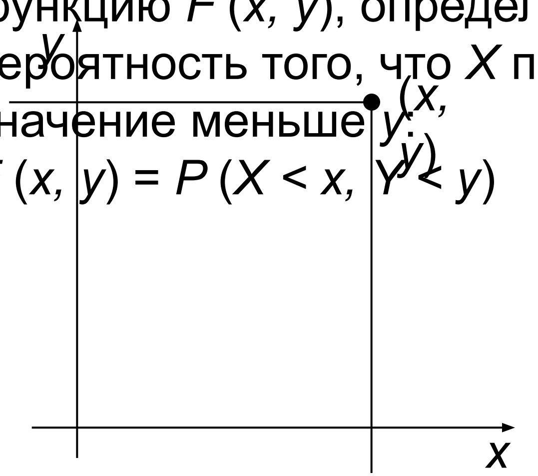
Событие $\{X = x_1\} = \left(\{X = x_1;\ Y = y_1\}\right.$ или $\{X = x_1;\ Y = y_2\}$ … или $\left.\{X = x_1;\ Y = y_m\}\right)$

Функция распределения двумерной случайной величины

Рассмотрим двумерную случайную величину (X, Y) (дискретную или непрерывную).

Функцией распределения двумерной случайной величины (X, Y) называют функцию $F(x, y)$, определяющую для каждой пары чисел (x, y) вероятность того, что X примет значение, меньшее x, Y примет значение меньше y:

$F(x, y) = P(X < x, Y < y)$



Пример. $F(x,y)=\left(\frac{1}{\pi}\text{arctg}\frac{x}{2}+\frac{1}{2}\right)\cdot\left(\frac{1}{\pi}\text{arctg}\frac{x}{3}+\frac{1}{2}\right)$

Найти $P(X < 2, Y < 3)$.

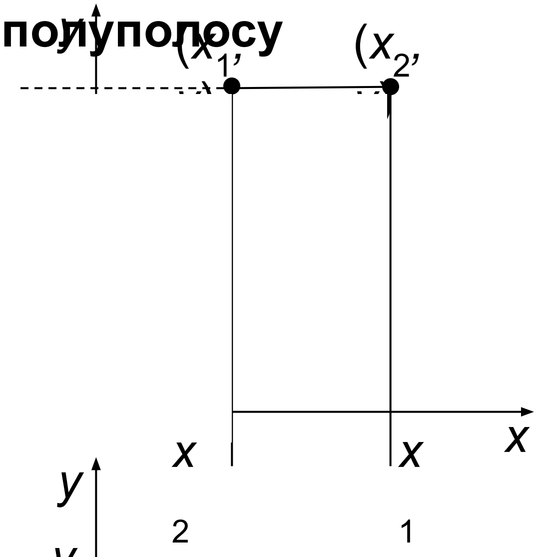
Решение.

$$P(X<2, Y<3)=F(2,3)=\left(\frac{1}{\pi}\text{arctg}\frac{2}{2}+\frac{1}{2}\right)\cdot\left(\frac{1}{\pi}\text{arctg}\frac{3}{3}+\frac{1}{2}\right)=\frac{9}{16}$$

Свойства функции распределения двумерной случайной величины

1. $0 \le F(x, y) \le 1$.
2. $F(x_2, y) \ge F(x_1, y)$, если $x_2 \ge x_1$;
   $F(x, y_2) \ge F(x, y_1)$, если $y_2 \ge y_1$.
3. $F(-\infty, y) = 0$;
   $F(x, -\infty) = 0$;
   $F(-\infty, -\infty) = 0$;
   $F(\infty, \infty) = 1$.
4. $F(x, \infty) = F_1(x)$;

Вероятность попадания случайной точки в полуполосу

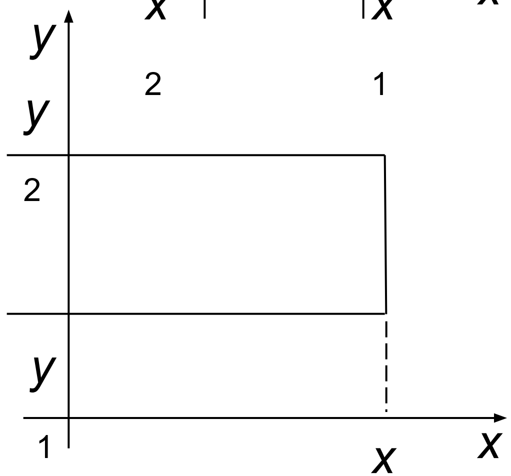


$P(x_1 < X < x_2,\ Y < y)$

Поскольку $\{X < x_2, Y < y\} = \{X < x_1, Y < y \text{ или } x_1 < X < x_2, Y < y\}$

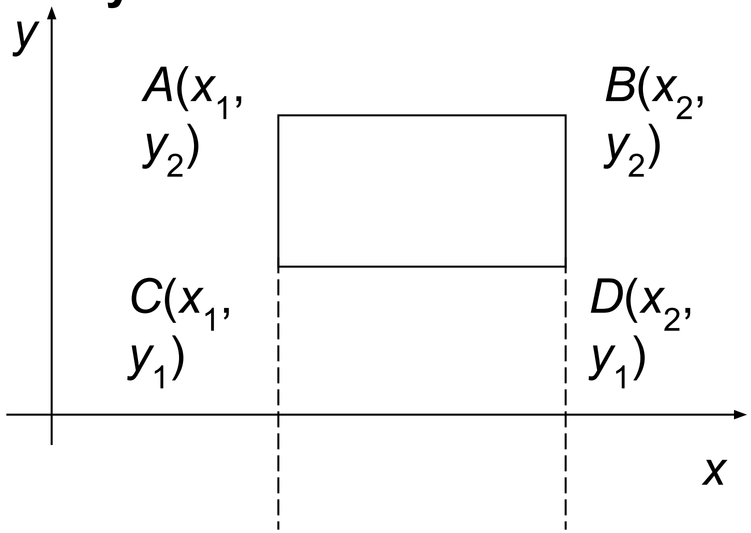
$$P(x_1 < X < x_2, Y < y) = P(X < x_2, Y < y) - P(X < x_1, Y < y)$$

$$P(x_1 < X < x_2, Y < y) = F(x_2, y) - F(x_1, y)$$



$$P(X < x, y_1 < Y < y_2) = F(x, y_2) - F(x, y_1)$$

Вероятность попадания случайной точки в прямоугольник



$$P(x_1 < X < x_2, y_1 < Y < y_2) = P(x_1 < X < x_2, Y < y_2) - P(x_1 < X < x_2, Y < y_1)$$

$$P(x_1 < X < x_2, y_1 < Y < y_2) = [F(x_2, y_2) - F(x_1, y_2)] - [F(x_2, y_1) - F(x_1, y_1)]$$

Плотность совместного распределения вероятностей двумерной случайной величины

Будем предполагать, что функция распределения *F*(*x*, *y*) непрерывна и имеет почти всюду непрерывные частные производные второго порядка. *Плотностью совместного распределения вероятностей p*(*x, y*) двумерной непрерывной случайной величины называют вторую смешанную производную от функции распределения:

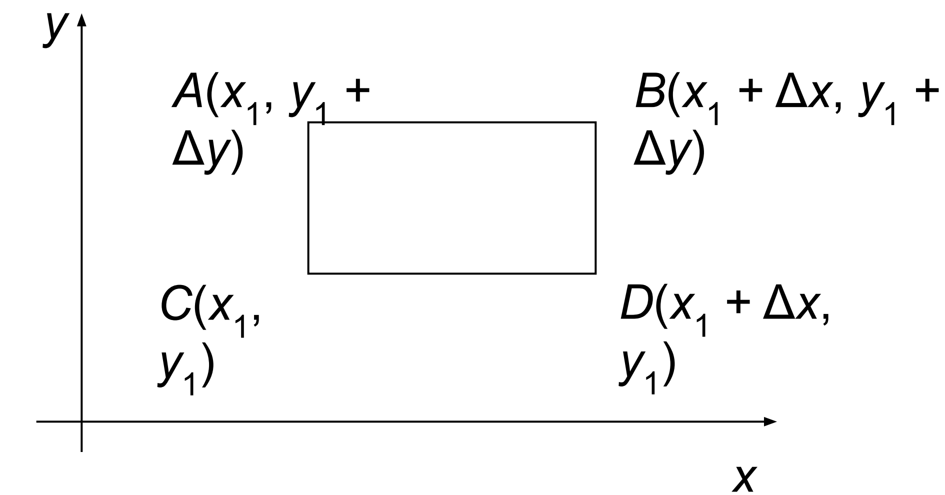
$$p(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

Зная плотность совместного распределения, можно найти функцию распределения *F*(*x*, *y*) по формуле

$$F(x, y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(x, y)\, dx\, dy$$

Вероятность попадания случайной точки в двумерную область

$$P(x_1 < X < x_2, y_1 < Y < y_2) = [F(x_2, y_2) - F(x_1, y_2)] - [F(x_2, y_1) - F(x_1, y_1)]$$

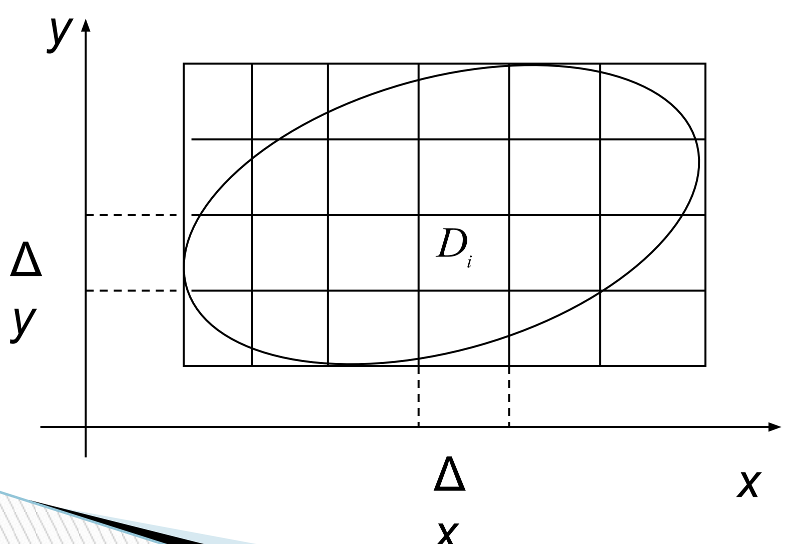


$$x_2 = x_1 + \Delta x;$$

$$y_2 = y_1 + \Delta y.$$

$$P_{ABCD} \approx F''_{xy}(\xi, \eta)\Delta x \Delta y = p(\xi, \eta)\Delta x \Delta y$$

$$x_1 < \xi < x_1 + \Delta x, y_1 < \eta < y_1 + \Delta y$$



$$P((X,Y) \subset D) \approx \sum_{i=1}^{n} \sum_{j=1}^{m} p(\xi_i, \eta_i)\Delta x \Delta y$$

$$\Delta x \to 0, \Delta y \to 0$$

$$P((X,Y) \subset D) = \iint_D p(x, y)dxdy$$

Свойства двумерной плотности вероятности

1. Двумерная плотность вероятности неотрицательна: $p(x, y) \geq 0$.
2. $\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} p(x,y)dxdy = 1$

Отыскание плотностей вероятности составляющих двумерной случайной величины

Пусть известна плотность совместного распределения вероятностей
системы двух случайных величин $p(x, y)$.
Найдем плотность распределения составляющей X.
Обозначим через $F_1(x)$ функцию распределения составляющей X.
По определению плотности распределения одномерной случайной величины

$$p_X(x) = \frac{dF_1(x)}{dx}$$

$$\left.\begin{aligned} F(x,y) &= \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} p(x,y)dxdy \\ F_1(x) &= F(x,\infty) \end{aligned}\right\} \Rightarrow F_1(x) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{\infty} p(x,y)dydx \qquad \frac{dF_1(x)}{dx} = \int\limits_{-\infty}^{\infty} p(x,y)dy$$

$$\Rightarrow p_X(x) = \int\limits_{-\infty}^{\infty} p(x,y)dy \qquad p_Y(y) = \int\limits_{-\infty}^{\infty} p(x,y)dx$$

Плотность распределения одной из составляющих равна несобственному интегралу
с бесконечными пределами от плотности совместного распределения системы,
причем переменная интегрирования соответствует другой составляющей.

Условные законы распределения составляющих системы дискретных случайных величин

Для того чтобы охарактеризовать зависимость между составляющими случайной величины,
введем понятие условного распределения.
Рассмотрим дискретную двумерную случайную величину (X, Y).
Пусть возможные значения составляющих таковы: $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m$.
Допустим, что в результате испытания величина Y приняла значение y_j: $Y = y_j$;
при этом X примет одно из возможных значений x_1, или x_2, … или x_n.
Обозначим $p(x_i \mid y_j)$ вероятность того, что случайная величина X примет значение x_i,
при условии, что $Y = y_j$. Эта вероятность, вообще говоря, не будет равна безусловной вероятности $p(x_i)$.
Условным распределением составляющей X при $Y = y_j$ называют совокупность условных вероятностей $p(x_1 \mid y_j)$, $p(x_2 \mid y_j)$, …, $p(x_n \mid y_j)$, вычисленных в предположении,
что событие $\{Y = y_j\}$ уже наступило.
Аналогично определяется условное распределение составляющей Y.
Зная закон распределения двумерной случайной величины, можно, пользуясь формулой
условной вероятности, вычислить условные законы распределения составляющих.
Например, условный закон распределения X в предположении, что событие $Y = y_j$ уже произошло,
может быть найден по формуле

$$p(x_i \mid y_j) = \frac{p(x_i, y_j)}{p(y_j)} \quad (i = 1, \dots, n)$$

Аналогично, условные законы распределения составляющей Y:

$$p(y_j \mid x_i) = \frac{p(x_i, y_j)}{p(x_i)}, \quad (j = 1, \dots, m)$$

Замечание. Сумма вероятностей условного распределения равна 1:

$$\sum_{i=1}^{n} p(x_i \mid y_j) = 1$$ при фиксированном y_j

$$\sum_{j=1}^{m} p(y_j \mid x_i) = 1$$ при фиксированном x_i

Пример. Дискретная двумерная случайная величина задана следующим законом распределения

Y	X			$P(Y = y_j)$
	x_1	x_2	x_3	
y_1	0,10	0,30	0,20	0,60
y_2	0,06	0,18	0,16	0,40

Найти условные законы распределения составляющей *X*.

Решение. Найдем закон распределения составляющей *Y*:

$$P(Y = y_1) = \sum_{i=1}^{3} p(x_i, y_1) = 0,10 + 0,30 + 0,20 = 0,60. \quad P(Y = y_2) = \sum_{i=1}^{3} p(x_i, y_2) = 0,06 + 0,18 + 0,16 = 0,40.$$

Далее, по формуле $p(x_i \mid y_j) = \dfrac{p(x_i, y_j)}{p(y_j)}$:

$$p(x_1 \mid y_1) = \frac{p(x_1, y_1)}{p(y_1)} = \frac{0,10}{0,60} = \frac{1}{6}, \quad p(x_2 \mid y_1) = \frac{0,30}{0,60} = \frac{1}{2}, \quad p(x_3 \mid y_1) = \frac{0,20}{0,60} = \frac{1}{3}, \quad \sum_{i=1}^{3} p(x_i \mid y_1) = 1;$$

$$p(x_2 \mid y_2) = \frac{p(x_2, y_2)}{p(y_2)} = \frac{0,06}{0,40} = \frac{3}{20}, \quad p(x_2 \mid y_2) = \frac{0,18}{0,40} = \frac{9}{20}, \quad p(x_3 \mid y_2) = \frac{0,16}{0,40} = \frac{2}{5}, \sum_{i=1}^{3} p(x_i \mid y_2) = 1.$$

Условные законы распределения составляющих системы непрерывных случайных величин

Пусть (X, Y) — непрерывная случайная величина.

Условной плотностью $\varphi(x \mid y)$ распределения составляющей X при данном значении $Y = y$

называют отношение плотности совместного распределения $p(x, y)$ системы (X, Y)

к плотности распределения $p_Y(y)$ составляющей Y:

$$\varphi(x \mid y) = \frac{p(x, y)}{p_Y(y)}.$$

Отличие условной плотности $\varphi(x \mid y)$ от безусловной $p_X(x)$ состоит в том, что функция $\varphi(x \mid y)$ дает распределение X при условии $Y = y$; функция $p_X(x)$ дает распределение x независимо от того, какие из возможных значений приняла составляющая Y.

Аналогично определяется условная плотность составляющей Y при данном значении $X = x$:

$$\psi(y \mid x) = \frac{p(x, y)}{p_X(x)}.$$

Если известна плотность совместного распределения $p(x, y)$, то условные плотности составляющих могут быть вычислены по формулам

$$\varphi(x \mid y) = \frac{p(x, y)}{\int_{-\infty}^{\infty} p(x, y)\,dx}, \quad \psi(y \mid x) = \frac{p(x, y)}{\int_{-\infty}^{\infty} p(x, y)\,dy}.$$

Умножая безусловный закон распределения одной из составляющих на условный закон распре- деления другой составляющей, найдем закон распределения системы случайных величин:

$$p(x, y) = p_Y(y) \cdot \varphi(x \mid y) = p_X(x) \cdot \psi(y \mid x)$$

Свойства:

$$\varphi(x \mid y) \geq 0, \quad \int_{-\infty}^{\infty} \varphi(x \mid y)\,dx = 1, \quad \psi(y \mid x) \geq 0, \quad \int_{-\infty}^{\infty} \psi(y \mid x)\,dy = 1.$$

Пример. Двумерная случайная величина (*X, Y*) задана плотностью совместного распределения

$$f(x,y)=\begin{cases}\dfrac{1}{\pi r^2} & \text{при } x^2+y^2<r^2,\\ 0 & \text{при } x^2+y^2\ge r^2.\end{cases}$$

Найти условные законы распределения составляющих.

Решение.

$$\varphi(x\mid y)=\frac{p(x,y)}{p_Y(y)}.$$

$$p_Y(y)=\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}}\frac{1}{\pi r^2}\,dx=\frac{2\sqrt{r^2-y^2}}{\pi r^2}$$

При $x^2+y^2<r^2$ $\quad |x|<\sqrt{r^2-y^2}$

$$\varphi(x\mid y)=\frac{\dfrac{1}{\pi r^2}}{\dfrac{2\sqrt{r^2-y^2}}{\pi r^2}}=\frac{1}{2\sqrt{r^2-y^2}}$$

$$\varphi(x\mid y)=\begin{cases}\dfrac{1}{2\sqrt{r^2-y^2}} & \text{при } |x|<\sqrt{r^2-y^2}\\ 0 \quad \text{при } |x|\ge\sqrt{r^2-y^2}\end{cases}$$

$$\psi(y\mid x)=\begin{cases}\dfrac{1}{2\sqrt{r^2-x^2}} & \text{при } |y|<\sqrt{r^2-x^2}\\ 0 \quad \text{при } |y|\ge\sqrt{r^2-x^2}\end{cases}$$

Условное математическое ожидание

Условным математическим ожиданием дискретной случайной величины Y при $X = x$
(x — **определенное** возможное значение X)
называют сумму произведений возможных значений Y на их условные вероятности:

$$M(Y \mid X = x) = \sum_{i=1}^{m} y_i p(y_i \mid x)$$

Для непрерывных величин

$$M(Y \mid X = x) = \int_{-\infty}^{\infty} y\psi(y \mid x)dy,$$

где $\psi(y \mid x)$ — условная плотность случайной величины Y при $X = x$.
Условное математическое ожидание $M(Y \mid x)$ есть функция от x: $M(Y \mid x) = f(x)$,
которую называют *функцией регрессии* Y на X.
Аналогично определяется условное математическое ожидание случайной величины X
и *функция регрессии* X на Y: $M(X \mid y) = \phi(y)$.

Пример. Задана двумерная случайная величина:

Y	X			
	$x_1 = 1$	$x_2 = 3$	$x_3 = 4$	$x_4 = 8$
$y_1 = 3$	0,15	0,06	0,25	0,04
$y_2 = 6$	0,30	0,10	0,03	0,07

Найти условное математическое ожидание составляющей Y при $x_1 = 1$.

Решение.

Найдем условное распределение вероятностей величины Y при $X = 1$:

$$p(x_1) = 0,15 + 0,30 = 0,45$$

$$p(y_1 \mid x_1) = \frac{p(x_1, y_1)}{p(x_1)} = \frac{0,15}{0,45} = \frac{1}{3}; \quad p(y_2 \mid x_1) = \frac{p(x_1, y_2)}{p(x_1)} = \frac{0,30}{0,45} = \frac{2}{3}.$$

$$M(Y \mid X = x_1) = \sum_{J=1}^{2} y_j p(y_j \mid x_1) = 3 \cdot \frac{1}{3} + 6 \cdot \frac{2}{3} = 5.$$

Зависимые и независимые случайные величины.

Две случайные величины X и Y называются *независимыми*, если закон распределения одной из них
не зависит от того, какие возможные значения приняла другая величина.

Условные распределения независимых величин равны их безусловным распределениям.

Теорема 1. Для того чтобы случайные величины X и Y были независимыми, необходимо и достаточно, чтобы функция распределения системы (X, Y)
была равна произведению функций распределения составляющих:

$F(x, y) = F_1(x) F_2(y)$.

Теорема 2. Для того чтобы случайные величины X и Y были независимыми, необходимо и достаточно, чтобы плотность совместного распределения системы (X, Y)
была равна произведению плотностей распределения составляющих:

$p(x, y) = p_X(x) p_Y(y)$.

Корреляционный момент. Коэффициент корреляции.

Корреляционным моментом μ_{xy} случайных величин *X* и *Y* называют математическое ожидание произведения отклонений этих величин:

$$\mu_{xy} = M\{[X - M(X)]\cdot[Y - M(Y)]\}$$

$$\mu_{xy} = M(XY) - M(X)M(Y)$$

$$\mu_{xy} = \sum_{i=1}^{n}\sum_{j=1}^{m}[x_i - M(X)][y_j - M(y)]p(x_i, y_j)$$

$$\mu_{xy} = \iint[x - M(X)][y - M(y)]p(x,y)dxdy$$

Корреляционный момент служит для характеристики связи между величинами *X* и *Y* .

Теорема 1. Корреляционный момент двух независимых случайных величин *X* и *Y* равен нулю.

Доказательство.

$$\mu_{xy} = M\{[X - M(X)]\cdot[Y - M(Y)]\} = M[X - M(X)]\cdot M[Y - M(Y)] = 0$$

Теорема 2. $|\mu_{xy}| \le \sqrt{D_x D_y}$

Коэффициентом корреляции r_{xy} случайных величин *X* и *Y* называют отношение корреляционного момента к произведению средних квадратических отклонений этих величин

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y}$$

Теорема 3. $|r_{xy}| \le 1$.

Замечание. Пусть дана случайная величина *X*.

Нормированная случайная величина

$$X' = \frac{X - M(X)}{\sigma_x}, \quad M(X') = 0,\ D(X') = 1$$

Для двух случайных величин *X* и *Y* $r_{xy} = \mu_{x'y'}$

Коррелированность и зависимость случайных величин

Две случайные величины *X* и *Y* называют *коррелированными*, если их коэффициент корреляции (или корреляционный момент) отличен от нуля; случайные величины *X* и *Y* называют *некоррелированными,* если их корреляционный момент равен 0.

Две коррелированные величины также зависимы. Обратное предположение не верно, т.е. если две величины зависимы, то они могут быть как коррелированными, так и некоррелированными.

Пример. Двумерная случайная величина (*X*, *Y*) задана плотностью распределения

$$p(x,y)=\begin{cases}\dfrac{1}{6\pi} & \text{при } \dfrac{x^2}{9}+\dfrac{y^2}{4}\le 1,\\ 0 & \text{при } \dfrac{x^2}{9}+\dfrac{y^2}{4}>1.\end{cases}$$

Доказать, что *X* и *Y* — зависимые некоррелированные величины.

Решение. Вычислим плотности распределения составляющих (внутри эллипса)

$$p_X(x)=\int_a^b p(x,y)dy=\frac{1}{6\pi}\int_{-2\sqrt{1-x^2/9}}^{2\sqrt{1-x^2/9}}dy=\frac{2}{9\pi}\sqrt{9-x^2}\qquad p_Y(y)=\frac{1}{2\pi}\sqrt{4-y^2}$$

$p_X(x)\cdot p_Y(y)\neq p(x,y)$ ⇒ *X* и *Y* — зависимые величины

$$\mu_{xy}=\iint[x-M(X)][y-M(y)]p(x,y)dxdy$$

Поскольку $p_X(x)$ и $p_Y(y)$ симметричны относительно *Ox* и *Oy*, то *M*(*X*) = *M*(*Y*) = 0.

$$\mu_{xy}=\frac{1}{6\pi}\iint_D xy\,dxdy=\frac{1}{6\pi}\int ydy\cdot\int xdx=0.$$

⇒ *X* и *Y* — некоррелированные величины

Нормальный закон распределения на плоскости

Нормальным законом распределения на плоскости называют распределение вероятностей
двумерной случайной величины, определяемое плотностью

$$f(x,y)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r_{xy}^2}}\exp\left[-\frac{1}{2\left(1-r_{xy}^2\right)}\left(\frac{(x-a_1)^2}{\sigma_x^2}+\frac{(y-a_2)^2}{\sigma_y^2}-2r_{xy}\frac{x-a_1}{\sigma_x}\frac{y-a_2}{\sigma_y}\right)\right]$$

Нормальный закон на плоскости задается пятью параметрами:

a_1, a_2 — математические ожидания;

σ_x, σ_y — средние квадратические отклонения;

r_{xy} — коэффициент корреляции величин X и Y.

Положив $r_{xy} = 0$, получим

$$f(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{(x-a_1)^2}{\sigma_x^2}+\frac{(y-a_2)^2}{\sigma_y^2}\right)\right]=$$

$$=\frac{1}{\sigma_x\sqrt{2\pi}}\exp\left(-\frac{(x-a_1)^2}{2\sigma_x^2}\right)\cdot\frac{1}{\sigma_y\sqrt{2\pi}}\exp\left(-\frac{(y-a_2)^2}{2\sigma_y^2}\right)=f_X(x)\cdot f_Y(y)$$

Таким образом, видим, что если составляющие нормально распределенной случайной величины
некоррелированы ($r_{xy} = 0$), то ее составляющие — независимы [$f(x, y) = f_X(x) \cdot f_Y(y)$].

Можно показать, что если двумерная случайная величина распределена по нормальному закону,
то и ее составляющие также распределены но нормальному закону.

Линейная регрессия. Прямые линии среднеквадратической регрессии.

Рассмотрим двумерную случайную величину (X, Y), где X, Y — **зависимые** случайные величины.

Поставим задачу представить одну из этих величин как функцию другой $Y \approx g(X)$.

Одним из способов отыскания $g(X)$ является *метод наименьших квадратов*:

$g(X)$ наилучшим образом приближает Y в смысле метода наименьших квадратов,

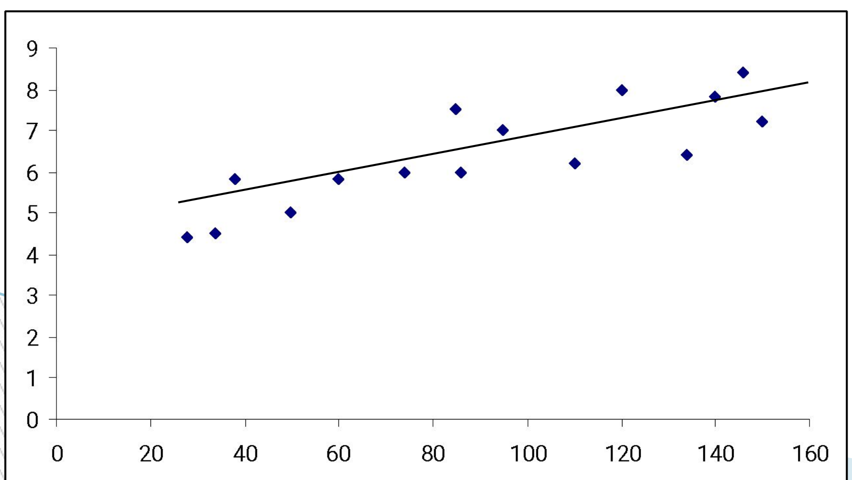
если $M[Y - g(X)]^2$ принимает наименьшее значение;

$$F(\alpha, \beta)=M[Y - \alpha X - \beta]^2 \to \min.$$

$g(X)$ называют *среднеквадратической регрессией* Y на X.

Будем искать $g(X)$ в виде $g(X) = \alpha X + \beta$, где α и β — параметры, подлежащие определению

(в этом случай $g(X)$ называют *линейной среднеквадратической регрессией* Y на X).



Линейная средняя квадратическая регрессия Y на X имеет вид

$$g(X) = m_y + r\frac{\sigma_y}{\sigma_x}(X - m_x),$$

где $m_x = M(X),\ m_y = M(Y),\ \sigma_x = \sqrt{D(X)},\ \sigma_y = \sqrt{D(Y)},\ r = \frac{\mu_{xy}}{\sigma_x \sigma_y}.$

$\alpha = r\frac{\sigma_y}{\sigma_x}$ — *коэффициент регрессии Y на X,* $\beta = m_y - r\frac{\sigma_y}{\sigma_x}m_x.$

$y - m_y = r\frac{\sigma_y}{\sigma_x}(x - m_x)$ — *прямая среднеквадратической регрессии Y на X.*

$\min F(\alpha, \beta) = \sigma_y^2(1 - r^2)$ — *остаточная дисперсия* случайной величины Y относительно X,

которая характеризует величину ошибки при замене Y линейной функцией $g(X) = \alpha X + \beta$.

При $r = \pm 1$ остаточная дисперсия равна 0.

Другими словами, при $r = \pm 1$ Y и X связаны **линейной** зависимостью.

Прямая среднеквадратической регрессии X на Y имеет вид $x - m_x = r\frac{\sigma_x}{\sigma_y}(y - m_y)$

Здесь $r\frac{\sigma_x}{\sigma_y}$ — *коэффициент регрессии X на Y,*

$\sigma_x^2(1 - r^2)$ — остаточная дисперсия случайной величины X относительно Y.

Если $r = \pm 1$, то обе прямые регрессии совпадают.

Обе прямые регрессии проходят через точку $(m_x; m_y)$, которая называется *центром совместного распределения X и Y.*