Лекция 7-8. Числовые характеристики случайных величин

Велищанский Михаил Александрович

Московский Государственный Технический Университет имени Н.Э. Баумана

Москва, 2023

Определение

Математическим ожиданием (средним значением) $\mathsf{M}\,X$ дискретной случайной величины X *называют сумму произведений значений x_i случайной величины и вероятностей $p_i = \mathsf{P}\{X = x_i\}$, с которыми случайная величина принимает эти значения:*

$$\mathsf{M}\,X = \sum_i x_i p_i.$$

При этом, если множество возможных значений случайной величины X счетно, предполагается, что

$$\sum_{i=1}^{\infty} |x_i| p_i < +\infty,$$

т.е. ряд, определяющий математическое ожидание, сходится абсолютно; в противном случае говорят, что математическое ожидание случайной величины X не существует.

Замечание

Пусть на прямой расположена система материальных точек с массами p_i $\left(\sum\limits_i p_i = 1\right)$ *и пусть* x_i *— координата* i*-й точки.*

Тогда центр масс системы будет иметь координату

$$\overline{X} = \frac{\sum\limits_i x_i p_i}{\sum\limits_i p_i} = \frac{\sum\limits_i x_i p_i}{1} = \sum_i x_i p_i,$$

совпадающую с математическим ожиданием $\mathbf{M}\, X$ *случайной величины* X*.*

Пример

Пусть X *— число выпавших очков при подбрасывании игральной кости. Так как* $p_i = \mathbf{P}\{X = i\} = \dfrac{1}{6}$, $i = \overline{1, 6}$, *то*

$$\mathbf{M}\, X = \sum_{i=1}^{6} i \frac{1}{6} = 7/2 = 3.5.$$

Определение

Математическим ожиданием (средним значением) $\mathsf{M}\,X$ **непрерывной случайной величины** X *называют интеграл*

$$\mathsf{M}\,X = \int\limits_{-\infty}^{+\infty} xf(x)\,dx.$$

При этом предполагается, что $\int\limits_{-\infty}^{+\infty} |x|f(x)\,dx < +\infty$, *т.е. несобственный интеграл, определяющий математическое ожидание, сходится абсолютно.*

Замечание

Так же как и в дискретном случае, математическое ожидание непрерывной случайной величины можно интерпретировать как центр масс стержня, плотность массы которого в точке x *равна* $f(x)$.

Пример

Пусть случайная величина X имеет распределение Пуассона. Тогда

$$\mathbf{M}\,X = \sum_{i=0}^{\infty} i\frac{\lambda^i}{i!}e^{-\lambda} = \lambda\sum_{i=1}^{\infty}\frac{\lambda^{i-1}}{(i-1)!}e^{-\lambda} = \lambda\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}e^{-\lambda} = \lambda e^{\lambda}e^{-\lambda} = \lambda.$$

Пример

Найдем математическое ожидание случайной величины X, имеющей геометрическое распределение:

$$\mathbf{M}\,X = \sum_{i=0}^{\infty} ipq^i = pq\sum_{i=0}^{\infty} iq^{i-1} = pq\Big(\sum_{i=0}^{\infty} q^i\Big)'_q =$$
$$= pq\left(\frac{1}{1-q}\right)' = \frac{pq}{(1-q)^2} = \frac{pq}{p^2} = \frac{q}{p}.$$

Пример

Пусть непрерывная случайная величина X распределена равномерно на $(a,\ b)$. Тогда

$$\mathbf{M}\,X = \int\limits_{-\infty}^{+\infty} xf(x)dx = \int\limits_{a}^{b} \frac{x}{b-a}dx = \frac{1}{b-a}\cdot\frac{x^2}{2}\Big|_a^b = \frac{b+a}{2}.$$

Пример

Пусть непрерывная случайная величина X распределена по экспоненциальному закону с параметром λ. Тогда ее математическое ожидание равно

$$\mathbf{M}\,X = \int\limits_{-\infty}^{+\infty} x\,f(x)\,dx = \int\limits_{0}^{+\infty} \lambda x e^{-\lambda x}\,dx =$$

$$= -\left(xe^{-\lambda x}\Big|_0^{+\infty} - \int\limits_{0}^{+\infty} e^{-\lambda x}\,dx \right) = -\frac{1}{\lambda}e^{-\lambda x}\Big|_0^{+\infty} = \frac{1}{\lambda}.$$

Пример

Пусть непрерывная случайная величина X распределена по нормальному закону с параметрами m и σ, тогда:

$$\mathsf{M}\,X = \int\limits_{-\infty}^{+\infty} x\varphi_{m,\sigma}(x)\,dx = \int\limits_{-\infty}^{+\infty} \frac{x}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/(2\sigma^2)} dx =$$

$$= \left| y = \frac{(x-m)}{\sigma} \right| = \int\limits_{-\infty}^{+\infty} \frac{\sigma y + m}{\sqrt{2\pi}} e^{-y^2/2} dy = \int\limits_{-\infty}^{+\infty} \frac{\sigma y}{\sqrt{2\pi}} e^{-y^2/2} dy +$$

$$+ m \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = \frac{\sigma}{\sqrt{2\pi}} \underbrace{\int\limits_{-\infty}^{+\infty} y e^{-y^2/2} dy}_{=0} + m \underbrace{\int\limits_{-\infty}^{+\infty} \varphi(y)\,dy}_{=1} = m.$$

Таким образом, $\mathsf{M}\,X = m$, т.е. параметр m имеет смысл математического ожидания случайной величины X.

Найдем математическое ожидание функции случайной величины (*случайного вектора*).
Пусть $Y = Y(X)$ является функцией от случайной величины X.
Рассмотрим сначала *дискретную случайную величину* X, принимающую значения $x_1, \ldots, x_n, \ldots$ с вероятностями $p_n = \mathbf{P}\{X = x_n\}$, $n = 1, 2, \ldots$.
Тогда случайная величина $Y = Y(X)$ принимает значения $Y(x_n)$ с вероятностями $p_n = \mathbf{P}\{X = x_n\}$, $n = 1, 2, \ldots$, и ее математическое ожидание имеет вид

$$\mathbf{M}\,Y = \mathbf{M}\,Y(X) = \sum_{i=1}^{\infty} Y(x_i)p_i,$$

при этом требуется выполнение условия $\sum\limits_{i=1}^{\infty} |Y(x_i)|p_i < +\infty$.

Математическое ожидание функции от случайной величины

Для *непрерывной случайной величины* X, имеющей плотность распределения $f(x)$, математическое ожидание случайной величины $Y = Y(X)$ можно найти, используя аналогичную дискретному случаю формулу

$$\mathbf{M} = \mathbf{M}\,Y(X) = \int\limits_{-\infty}^{+\infty} Y(x) f(x) dx\,,$$

при этом необходимо чтобы $\int\limits_{-\infty}^{+\infty} |Y(x)| f(x) dx < +\infty.$

Аналогичным образом определяется математическое ожидание функции от *многомерной случайной величины.*

Математическое ожидание функции от случайного вектора

Математическое ожидание $\mathbf{M}\,Y$ функции $Y = Y(X_1, X_2)$ от *дискретной* двумерной случайной величины $(X_1,\, X_2)$ можно найти, воспользовавшись формулой

$$\mathbf{M}\,Y = \mathbf{M}\,Y(X_1, X_2) = \sum_{i,j} Y(x_{1i}, x_{2j})p_{ij},$$

где $p_{ij} = \mathbf{P}\{X_1 = x_{1i},\, X_2 = x_{2j}\}$.

Математическое ожидание $\mathbf{M}\,Y$ функции $Y = Y(X_1, X_2)$ от двумерной *непрерывной* случайной величины $(X_1,\, X_2)$ определяется формулой

$$\mathbf{M}\,Y = \mathbf{M}\,Y(X_1, X_2) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} Y(x_1, x_2) f_{X_1,X_2}(x_1, x_2)\, dx_1 dx_2,$$

где $f_{X_1,X_2}(x_1, x_2)$ — совместная плотность распределения случайных величин X_1 и X_2.

Теорема

Математическое ожидание удовлетворяет следующим свойствам:

1. *Если случайная величина X принимает всего одно значение C с вероятностью единица (т.е., по сути дела, не является случайной величиной), то* $\mathbf{M}\, C = C$.
2. $\mathbf{M}(aX + b) = a\,\mathbf{M}\, X + b$, *где a, b — постоянные.*
3. $\mathbf{M}(X_1 + X_2) = \mathbf{M}\, X_1 + \mathbf{M}\, X_2$.
4. $\mathbf{M}(X_1 X_2) = \mathbf{M}\, X_1 \cdot \mathbf{M}\, X_2$ *для независимых случайных величин X_1 и X_2.*

Доказательство.

1. Если случайная величина X принимает всего одно значение C с вероятностью единица, то, согласно определению, имеем $\mathbf{M}\, C = C \cdot 1 = C$.

доказательство (продолжение).

Доказательство свойств 2–4 проведем для непрерывных случайных величин (для дискретных — аналогично).

2. Докажем, что $\mathbf{M}(aX+b) = a\,\mathbf{M}\,X + b$, где a, b — постоянные.
   Найдем математическое ожидание случайной величины $Y = aX + b$ ($Y = Y(x) = ax + b$):
   $$\mathbf{M}\,Y = \mathbf{M}(aX+b) = \int\limits_{-\infty}^{+\infty} (ax+b) f_X(x)\,dx =$$
   $$= a\int\limits_{-\infty}^{+\infty} x f_X(x)\,dx + b\int\limits_{-\infty}^{+\infty} f_X(x)\,dx =$$
   $$= a\,\mathbf{M}\,X + b\cdot 1 = a\,\mathbf{M}\,X + b.$$

Свойства математического ожидания случайной величины

доказательство (продолжение).

3. Докажем, что $\mathsf{M}(X_1 + X_2) = \mathsf{M}\, X_1 + \mathsf{M}\, X_2$.
   Пусть теперь $Y = X_1 + X_2$ $(Y = Y(x_1, x_2) = x_1 + x_2)$. Тогда

$$\mathsf{M}\, Y = \mathsf{M}(X_1 + X_2) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x_1 + x_2) f_{X_1,X_2}(x_1, x_2)\, dx_1 dx_2 =$$

$$= \int\limits_{-\infty}^{+\infty} x_1 \left(\int\limits_{-\infty}^{+\infty} f_{X_1,X_2}(x_1, x_2)\, dx_2 \right) dx_1 +$$

$$+ \int\limits_{-\infty}^{+\infty} x_2 \left(\int\limits_{-\infty}^{+\infty} f_{X_1,X_2}(x_1, x_2)\, dx_1 \right) dx_2 =$$

$$= \int\limits_{-\infty}^{+\infty} x_1 f_{X_1}(x_1) dx_1 + \int\limits_{-\infty}^{+\infty} x_2 f_{X_2}(x_2) dx_2 = \mathsf{M}\, X_1 + \mathsf{M}\, X_2.$$

□

доказательство (продолжение).

4. Докажем, что $\mathbf{M}(X_1X_2) = \mathbf{M}\,X_1 \cdot \mathbf{M}\,X_2$ для *независимых случайных величин* X_1 и X_2.
   Так как X_1 и X_2 независимые случайные величины, то воспользовавшись формулой для математического ожидания функции случайного вектора $(Y = X_1X_2)$ имеем:

$$\mathbf{M}\,Y = \mathbf{M}(X_1X_2) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} x_1x_2 f_{X_1,X_2}(x_1,x_2)\,dx_1\,dx_2 =$$

$$= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} x_1x_2 f_{X_1}(x_1) f_{X_2}(x_2)\,dx_1\,dx_2 =$$

$$= \left(\int\limits_{-\infty}^{+\infty} x_1 f_{X_1}(x_1)\,dx_1\right)\left(\int\limits_{-\infty}^{+\infty} x_2 f_{X_2}(x_2)\,dx_2\right) =$$

$$= \mathbf{M}\,X_1\,\mathbf{M}\,X_2.$$

Замечание

1. *В данной теореме предполагается, что математические ожидания случайных величин* X, X_1 *и* X_2 *существуют, хотя математическое ожидание суммы случайных величин может существовать даже тогда, когда математические ожидания обоих слагаемых не существует, поскольку* $\mathbf{M}(X - X) = \mathbf{M}\,0 = 0$, *даже если* $\mathbf{M}\,X$ *не существует.*
2. *Свойство 3 можно обобщить на произвольное число слагаемых, т.е.*

$$\mathbf{M}(X_1 + \ldots + X_n) = \mathbf{M}\,X_1 + \ldots + \mathbf{M}\,X_n.$$

3. *Свойство 4 также допускает обобщение на произведение конечного числа (независимых в совокупности) случайных величин:*

$$\mathbf{M}(X_1 \cdot X_2 \ldots X_n) = \mathbf{M}\,X_1 \cdot \mathbf{M}\,X_2 \cdot \ldots \cdot \mathbf{M}\,X_n.$$

Дисперсия случайной величины

Две случайные величины могут иметь одинаковые математические ожидания, но их возможные значения будут по-разному рассеиваться вокруг этого значения.

Например, средний балл на экзамене в двух группах равен 4, но в первой группе почти все студенты получили 4, а во второй группе 4 нет вообще, а основные оценки 3 и 5.

Поэтому, наряду с математическим ожиданием, вводят число, характеризующее "разброс" случайной величины относительно своего математического ожидания. Такой характеристикой обычно служит *дисперсия*.

Определение

Дисперсией $\mathbf{D}\,X$ случайной величины X *называют математическое ожидание квадрата отклонения случайной величины* X *от ее среднего значения, т.е.* $\mathbf{D}\,X = \mathbf{M}(X - \mathbf{M}\,X)^2$.

Полагая $Y(x) = (x - \mathbf{M}\,X)^2$ и используя формулы для вычисления математического ожидания функции от СВ, получаем:

$$\mathbf{D}\,X = \sum_i (x_i - \mathbf{M}\,X)^2 p_i \text{ и } \mathbf{D}\,X = \int\limits_{-\infty}^{+\infty} (x - \mathbf{M}\,X)^2 f(x)\,dx.$$

Замечание

Из определения непосредственно следует, что дисперсия любой случайной величины является неотрицательным числом.

Механический смысл дисперсии — центральный (относительно центра масс) момент инерции массы, распределенной на оси с линейной плотностью $f(x)$.
Дисперсия $\mathbf{D}\,X$ представляет собой второй момент *центрированной* (имеющей нулевое математическое ожидание) случайной величины $\overset{o}{X} = X - \mathbf{M}\,X$.
Поэтому иногда дисперсию называют **вторым центральным моментом** случайной величины.

Наряду с понятием второго центрального момента можно ввести понятие *второго начального момента*.

Определение

Вторым начальным моментом m_2 **случайной величины** X *называют математическое ожидание квадрата случайной величины* X, *т.е.* $m_2 = \mathbf{M}\,X^2$.

В литературе часто встречаются понятия моментов высших порядков.

Определение

Моментом k-го порядка m_k **(k-м моментом)** *случайной величины* X *называют математическое ожидание* k*-й степени случайной величины* X, *т.е.* $m_k = \mathsf{M}\, X^k$.

Иногда k-й момент называют также начальным моментом k-го порядка.

Определение

Центральным моментом k-го порядка $\overset{o}{m}_k$ **(k-м центральным моментом)** *случайной величины* X *называют математическое ожидание* k*-й степени случайной величины* $\overset{o}{X} = X - \mathsf{M}\, X$ *т.е.* $\overset{o}{m}_k = \mathsf{M}(X - \mathsf{M}\, X)^k$.

Свойства дисперсии

Теорема

Дисперсия удовлетворяет следующим свойствам:

1. *Если случайная величина* X *принимает всего одно значение* C *с вероятностью единица, то* $\mathbf{D}\, C = 0$.
2. $\mathbf{D}(aX + b) = a^2\, \mathbf{D}\, X$.
3. $\mathbf{D}\, X = \mathbf{M}\, X^2 - (\mathbf{M}\, X)^2$.
4. $\mathbf{D}(X + Y) = \mathbf{D}\, X + \mathbf{D}\, Y$ *для независимых случайных величин* X *и* Y.

Доказательство.

1. Если случайная величина X с вероятностью единица принимает всего одно значение C, то в силу свойства 1 математического ожидания $(\mathbf{M}\, X = C)$ получаем

$$\mathbf{D}\, X = \mathbf{M}(X - C)^2 = (C - C)^2 \cdot 1 = 0.$$

доказательство (продолжение).

2. Определим дисперсию случайной величины $Y = aX + b$. Используя свойство 2 математического ожидания, имеем

$$\mathbf{D}\,Y = \mathbf{M}(Y - \mathbf{M}\,Y)^2 = \mathbf{M}\big(aX + b - \mathbf{M}(aX + b)\big)^2 =$$
$$= \mathbf{M}(aX + b - aMX - b)^2 = \mathbf{M}\big(a(X - \mathbf{M}\,X)\big)^2 =$$
$$= \mathbf{M}\big(a^2(X - \mathbf{M}\,X)^2\big) = a^2\,\mathbf{M}(X - \mathbf{M}\,X)^2 = a^2\,\mathbf{D}\,X.$$

3. Согласно свойствам 2 и 3 математического ожидания, получаем

$$\mathbf{D}\,X = \mathbf{M}(X - \mathbf{M}\,X)^2 = \mathbf{M}\big(X^2 - 2XMX + (\mathbf{M}\,X)^2\big) =$$
$$= \mathbf{M}\,X^2 - 2(\mathbf{M}\,X)^2 + (\mathbf{M}\,X)^2 = \mathbf{M}\,X^2 - (\mathbf{M}\,X)^2.$$

□

доказательство (продолжение).

4. Пусть X и Y — независимые случайные величины. Тогда, используя независимость случайных величин $\overset{o}{X} = X - \mathsf{M}\,X$ и $\overset{o}{Y} = Y - \mathsf{M}\,Y$, а также свойства 2–4 математического ожидания, получаем

$$\mathsf{D}(X+Y) = \mathsf{M}\big(X+Y-\mathsf{M}(X+Y)\big)^2 =$$
$$= \mathsf{M}\big((X-\mathsf{M}\,X)+(Y-\mathsf{M}\,Y)\big)^2 = \mathsf{M}(X-\mathsf{M}\,X)^2 +$$
$$+ 2\,\mathsf{M}\big((X-\mathsf{M}\,X)(Y-\mathsf{M}\,Y)\big) + \mathsf{M}(Y-\mathsf{M}\,Y)^2 =$$
$$= \mathsf{D}\,X + 2(\mathsf{M}\overset{o}{X}\cdot\mathsf{M}\overset{o}{Y}) + \mathsf{D}\,Y = \mathsf{D}\,X + \mathsf{D}\,Y,$$

поскольку $\mathsf{M}\overset{o}{X} = 0$ и $\mathsf{M}\overset{o}{Y} = 0$.

□

Свойства дисперсии

Замечание

Можно показать, что справедливо и свойство, обратное свойству 1, т.е. имеет место утверждение: дисперсия случайной величины X равна нулю тогда и только тогда, когда X с вероятностью 1 принимает всего одно значение.

Замечание

Очевидно, что свойство 4 справедливо для суммы не только двух, но и любого числа n попарно независимых случайных величин $\mathbf{D}(X_1 + \ldots + X_n) = \mathbf{D}\,X_1 + \ldots + \mathbf{D}\,X_n$.

Замечание

Позднее будет приведена формула для дисперсии суммы любых (а не только независимых) слагаемых.

На практике на ряду с дисперсией $\mathbf{D}\,X$ случайной величины X часто используют величину $\sigma = \sqrt{\mathbf{D}\,X}$, которую называют **средним квадратичным отклонением** случайной величины X.

Пример

Пусть случайная величина X имеет распределение Пуассона ($\mathbf{M}\,X = \lambda$). Необходимо найти дисперсию $\mathbf{D}\,X$. Для начала найдем второй момент

$$\mathbf{M}\,X^2 = \sum_{i=0}^{\infty} i^2 \frac{\lambda^i}{i!} e^{-\lambda} = \lambda \sum_{i=1}^{\infty} i \frac{\lambda^{i-1}}{(i-1)!} e^{-\lambda} = \lambda \sum_{j=0}^{\infty} (j+1) \frac{\lambda^j}{j!} e^{-\lambda} =$$

$$= \lambda \left(\sum_{j=0}^{\infty} j \frac{\lambda^j}{j!} e^{-\lambda} + \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} e^{-\lambda} \right) = \lambda(\mathbf{M}\,X + 1) = \lambda^2 + \lambda.$$

Тогда, воспользовавшись свойством 3 дисперсии, имеем

$$\mathbf{D}\,X = \mathbf{M}\,X^2 - (\mathbf{M}\,X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Пример

Пусть непрерывная случайная величина X распределена равномерно на (a, b). Тогда

$$\mathbf{D}\,X = \int\limits_a^b \left(x - \frac{b+a}{2}\right)^2 \frac{1}{b-a}\,dx = \frac{1}{3(b-a)}\left(\left(b - \frac{b+a}{2}\right)^3 - \right.$$

$$\left. - \left(a - \frac{b+a}{2}\right)^3\right) = \frac{(b-a)^3}{12(b-a)} = \frac{(b-a)^2}{12}.$$

Пример

Пусть $X \sim Exp(\lambda)$. Тогда

$$\mathbf{M}\,X^2 = \int\limits_{-\infty}^{+\infty} x^2 f(x)\,dx = \int\limits_0^{+\infty} \lambda x^2 e^{-\lambda x}\,dx = -\left(x^2 e^{-\lambda x}\Big|_0^{+\infty} - \right.$$

$$\left. - \int\limits_0^{+\infty} 2xe^{-\lambda x}\,dx \right) = \ldots = -\frac{2}{\lambda^2} e^{-\lambda x}\Big|_0^{+\infty} = \frac{2}{\lambda^2} \Rightarrow$$

$$\Rightarrow \mathbf{D}\,X = MX^2 - (MX)^2 = 1/\lambda^2.$$

Пример

Найдем дисперсию случайной величины X, распределенной по нормальному закону с параметрами m и σ

$$\mathbf{D}\,X = \int_{-\infty}^{+\infty} (x-m)^2 \varphi_{m,\sigma}(x)\,dx = \int_{-\infty}^{+\infty} \frac{(x-m)^2}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-m)^2}{2\sigma^2}}\,dx =$$

$$= \left| y = \frac{(x-m)}{\sigma} \right| = \sigma^2 \int_{-\infty}^{+\infty} \frac{y^2}{\sqrt{2\pi}} e^{-y^2/2} dy = \left| \begin{array}{c} \text{по частям} \\ u = y/\sqrt{2\pi} \\ dv = ye^{-y^2/2}dy \end{array} \right| =$$

$$\ldots = \sigma^2 \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\,dy = \sigma^2 \int_{-\infty}^{+\infty} \varphi(y)\,dy = \sigma^2.$$

Таким образом, дисперсия нормально распределенной случайной величины совпадает с квадратом второго параметра.

Пусть $(X_1; X_2)$ — *двумерный случайный вектор.*

Определение

Ковариацией (корреляционным моментом) $\mathrm{cov}(X_1, X_2)$ *случайных величин* X_1 *и* X_2 *называют математическое ожидание произведения случайных величин* $\overset{o}{X}_1 = X_1 - \mathsf{M}\,X_1$ *и* $\overset{o}{X}_2 = X_2 - \mathsf{M}\,X_2$*:*

$$\mathrm{cov}(X_1, X_2) = \mathsf{M}(\overset{o}{X}_1 \overset{o}{X}_2) = \mathsf{M}\big((X_1 - \mathsf{M}\,X_1)(X_2 - \mathsf{M}\,X_2)\big).$$

Для дискретных случайных величин X_1 и X_2

$$\mathrm{cov}(X_1, X_2) = \sum_{i,j} (x_i - \mathsf{M}\,X_1)(y_j - \mathsf{M}\,X_2) p_{ij},$$

для непрерывных случайных величин X_1 и X_2

$$\mathrm{cov}(X_1, X_2) = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} (x_1 - \mathsf{M}\,X_1)(x_2 - \mathsf{M}\,X_2) f_{X_1,X_2}(x_1, x_2)\, dx_1\, dx_2.$$

Поскольку

$$\begin{aligned}
&\mathbf{D}(X+Y) = \mathbf{M}\big((X+Y) - \mathbf{M}(X+Y)\big)^2 = \\
&= \mathbf{M}(X - \mathbf{M}\,X)^2 + 2\,\mathbf{M}\big((X - MX)(Y - MY)\big) + \mathbf{M}(Y - MY)^2 = \\
&= \mathbf{D}\,X + \mathbf{D}\,Y + 2\,\mathrm{cov}(X, Y),
\end{aligned}$$

то можно ввести свойство 5 для дисперсии суммы случайных величин, справедливое для любых случайных величин X и Y:

$$\mathbf{D}(X+Y) = \mathbf{D}\,X + \mathbf{D}\,Y + 2\,\mathrm{cov}(X, Y).$$

Свойство 5 дисперсии допускает обобщение на произвольное число слагаемых:

$$\mathbf{D}(X_1 + \ldots + X_n) = \mathbf{D}\,X_1 + \ldots + \mathbf{D}\,X_n + 2 \sum_{1 \leqslant i < j \leqslant n} \mathrm{cov}(X_i, X_j).$$

Свойства ковариации случайных величин

Теорема

Ковариация имеет следующие свойства

1. $\operatorname{cov}(X, X) = \mathbf{D}\, X$.
2. $\operatorname{cov}(X_1, X_2) = 0$ *для независимых случайных величин* X_1 *и* X_2.
3. *Если* $Y_1 = a_1 X_1 + b_1$, *и* $Y_2 = a_2 X_2 + b_2$, *то* $\operatorname{cov}(Y_1, Y_2) = a_1 a_2 \operatorname{cov}(X_1, X_2)$.
4. $-\sqrt{\mathbf{D}\, X_1\, \mathbf{D}\, X_2} \leqslant \operatorname{cov}(X_1, X_2) \leqslant \sqrt{\mathbf{D}\, X_1\, \mathbf{D}\, X_2}$.
5. $|\operatorname{cov}(X_1, X_2)| = \sqrt{\mathbf{D}\, X_1\, \mathbf{D}\, X_2}$ *тогда и только тогда, когда случайные величины* X_1 *и* X_2 *связаны линейной зависимостью, т.е. существуют такие числа* a *и* b, *при которых* $X_2 = aX_1 - b$.
6. $\operatorname{cov}(X_1, X_2) = \mathbf{M}(X_1 X_2) - MX_1 MX_2$.

Доказательство.

1. По определению ковариации и дисперсии имеем: $\operatorname{cov}(X, X) = \mathbf{M}(X - MX)^2 = \mathbf{D}\, X$.
2. Если случайные величины X_1 и X_2 являются независимыми (и имеют математические ожидания), то

$$\begin{aligned}\operatorname{cov}(X_1, X_2) = \mathbf{M}\big((X_1 - MX_1)(X_2 - MX_2)\big) =\\ = \big(\mathbf{M}(X_1 - MX_1)\big)\big(\mathbf{M}(X_2 - MX_2)\big) = 0.\end{aligned}$$

3. Пусть $Y_1 = a_1X_1 + b_1$, $Y_2 = a_2X_2 + b_2$.
Тогда

$$\begin{aligned}&\operatorname{cov}(Y_1, Y_2) = \mathbf{M}\big((Y_1 - MY_1)(Y_2 - MY_2)\big) =\\ &= \mathbf{M}\big((a_1X_1 + b_1 - a_1MX_1 - b_1)(a_2X_2 + b_2 - a_2MX_2 - b_2)\big) =\\ &\qquad = \mathbf{M}\big(a_1a_2(X_1 - MX_1)(X_2 - MX_2)\big) = a_1a_2 \operatorname{cov}(X_1, X_2).\end{aligned}$$

□

доказательство (продолжение).

4. Рассмотрим дисперсию случайной величины
$Y_x = xX_1 - X_2$, где x — произвольное число.
В силу свойств дисперсии и свойства 3 ковариации

$$\mathbf{D}\,Y_x = \mathbf{D}(xX_1) + 2\operatorname{cov}(xX_1, -X_2) + \mathbf{D}(-X_2) = \\ = x^2\,\mathbf{D}\,X_1 - 2x\operatorname{cov}(X_1, X_2) + \mathbf{D}\,X_2.$$

Т.к. $\mathbf{D}\,Y_x \geqslant 0$ то дискриминант

$$\mathbf{D} = \left(-2\operatorname{cov}(X_1, X_2)\right)^2 - 4\,\mathbf{D}\,X_1\,\mathbf{D}\,X_2 \leqslant 0.$$

Следовательно $\left(2\operatorname{cov}(X_1, X_2)\right)^2 \leqslant 4\,\mathbf{D}\,X_1\,\mathbf{D}\,X_2 \Rightarrow$
$|\operatorname{cov}(X_1, X_2)| \leqslant \sqrt{\mathbf{D}\,X_1\,\mathbf{D}\,X_2}$.

$\square$

доказательство (продолжение).

5. Пусть выполнено равенство $|\operatorname{cov}(X_1, X_2)| = \sqrt{\mathsf{D}\, X_1\, \mathsf{D}\, X_2}$.
Значит, дискриминант $\mathsf{D} = \big(2\operatorname{cov}(X_1, X_2)\big)^2 - 4\,\mathsf{D}\, X_1\, \mathsf{D}\, X_2$ равен нулю, и уравнение $\mathsf{D}\, Y_x = 0$ имеет решение, которое обозначим a.
Тогда случайная величина $Y_a = aX_1 - X_2$ принимает всего одно значение (допустим, b), и, следовательно, $X_2 = aX_1 - b$.

Пусть теперь выполнено $X_2 = aX_1 - b$. Тогда в соответствии со свойством 1 дисперсии $\mathsf{D}\, Y_a = 0$, а значит, дискриминант является неотрицательным.
Поскольку при доказательстве утверждения 4 было показано, что этот дискриминант неположителен, то он равен нулю, и значит $|\operatorname{cov}(X_1, X_2)| = \sqrt{\mathsf{D}\, X_1\, \mathsf{D}\, X_2}$.

□

доказательство (продолжение).

6.

$$\begin{aligned}\operatorname{cov}(X_1, X_2) = \mathbf{M}\big((X_1 - \mathbf{M}\,X_1)(X_2 - \mathbf{M}\,X_2)\big) = \\ = \mathbf{M}\big(X_1X_2 - X_2\,\mathbf{M}\,X_1 - X_1\,\mathbf{M}\,X_2 + \mathbf{M}\,X_1\,\mathbf{M}\,X_2\big) = \\ = \mathbf{M}(X_1X_2) - \mathbf{M}(X_2\,\mathbf{M}\,X_1) - \mathbf{M}(X_1\,\mathbf{M}\,X_2) + \mathbf{M}\,X_1\,\mathbf{M}\,X_2 = \\ = \mathbf{M}(X_1X_2) - \mathbf{M}\,X_1\,\mathbf{M}\,X_2.\end{aligned}$$

□

Замечание

Если случайные величины связаны линейной зависимостью $X_2 = aX_1 - b,\ a \neq 0$, *то в соответствии со свойствами 3 и 1* $\operatorname{cov}(X_1, X_2) = a\operatorname{cov}(X_1, X_1) = a\,\mathbf{D}\,X_1$. *Поэтому свойство 5 допускает следующее уточнение:*

$$\operatorname{cov}(X_1, X_2) = \sqrt{\mathbf{D}\,X_1\,\mathbf{D}\,X_2} \qquad \text{при } a > 0;$$

$$\operatorname{cov}(X_1, X_2) = -\sqrt{\mathbf{D}\,X_1\,\mathbf{D}\,X_2} \qquad \text{при } a < 0.$$

Замечание

Имеет место еще одно полезное при расчетах свойство дисперсии
$\mathbf{D}(a_1X_1 + a_2X_2 + b) = a_1^2\,\mathbf{D}\,X_1 + a_2^2\,\mathbf{D}\,X_2 + 2a_1a_2\,\mathrm{cov}(X_1, X_2).$

Замечание

Как следует из свойства 2, ковариация независимых случайных величин равна нулю. Однако обратное, вообще говоря, неверно. Существуют зависимые и даже функционально зависимые случайные величины, ковариация которых равна нулю.

Определение

Случайные величины X и Y называют **некоррелированными**, *если их ковариация равна нулю, т.е.* $\mathrm{cov}(X, Y) = 0$.

Таким образом из *некоррелированности* случайных величин не следует их *независимость*.

Матрицей ковариаций случайного вектора

Рассмотрим теперь n-мерный случайный вектор $\vec{X} = (X_1, \ldots, X_n)$.

Определение

Матрицей ковариаций (ковариационной матрицей) *случайного вектора* $\vec{X}$ *называют матрицу* $\Sigma = (\sigma_{ij}) = (\operatorname{cov}(X_i, X_j))$, *состоящую из ковариаций случайных величин* X_i *и* X_j.

Теорема (б/д)

Ковариационная матрица обладает следующими свойствами.

1. *Матрица ковариаций* Σ *является симметрической.*
2. *Если случайный вектор* $\vec{Y} = \vec{X}B + \vec{c}$, *то* $\Sigma_{\vec{Y}} = B^T \Sigma_{\vec{X}} B$, *где* $\Sigma_{\vec{Y}}$ *и* $\Sigma_{\vec{X}}$ *матрицы ковариаций случайных векторов* $\vec{Y}$ *и* $\vec{X}$ *соответственно.*
3. *Матрица ковариаций* Σ *является неотрицательно определенной.*

Определение

Коэффициентом корреляции *случайных величин* X *и* Y *называют число* $\rho = \rho(X, Y)$, *определяемое равенством (предполагается, что* $\mathsf{D}\,X > 0$ *и* $\mathsf{D}\,Y > 0$*)* $\rho = \dfrac{\operatorname{cov}(X, Y)}{\sqrt{\mathsf{D}\,X \cdot \mathsf{D}\,Y}}$.

Теорема (б/д)

Коэффициент корреляции имеет следующие свойства.

1. $\rho(X, X) = 1$.
2. *Если случайные величины* X *и* Y *являются независимыми (и существуют* $\mathsf{D}\,X > 0$ *и* $\mathsf{D}\,Y > 0$*), то* $\rho(X, Y) = 0$.
3. $\rho(a_1X_1 + b_1,\, a_2X_2 + b_2) = \pm\rho(X_1, X_2)$. *При этом знак плюс нужно брать в том случае, когда* a_1 *и* a_2 *имеют одинаковые знаки, и минус — в противном случае.*
4. $-1 \leqslant \rho(X,\, Y) \leqslant 1$.
5. $|\rho(X,\, Y)| = 1$ *тогда и только тогда, когда случайные величины* X *и* Y *связаны линейной зависимостью.*

Коэффициент корреляции случайных величин

Коэффициент корреляции (также как и ковариация) отражают "степень линейной близости" случайных величин. При $\rho > 0$ говорят о *положительной* корреляционной зависимости X и Y, при $\rho < 0$ — об *отрицательной*.

Однако, коэффициент корреляции (ковариация) может не улавливать "степень нелинейной близости" случайных величин. Для этой цели служат другие характеристики.

По аналогии с ковариационной матрицей для случайного вектора $\vec{X} = (X_1, \ldots, X_n)$ можно ввести корреляционную матрицу.

Определение

Корреляционной (нормированной ковариационной) матрицей *случайного вектора* $\vec{X}$ *называют матрицу* $P = (\rho_{ij}) = (\rho(X_i, X_j))$, *состоящую из коэффициентов корреляций случайных величин* X_i *и* X_j.

Рассмотрим лишь одну дополнительную характеристику, широко используемую, в частности, в математической статистике.

Определение

Квантилью уровня α, или α-квантилью, *$(0 < \alpha < 1)$ случайной величины X (распределения случайной величины X) называют число Q_α, удовлетворяющее неравенствам*

$$\mathsf{P}\{X < Q_\alpha\} \leqslant \alpha \quad \text{и} \quad \mathsf{P}\{X > Q_\alpha\} \leqslant 1 - \alpha.$$

$1/2$-*квантиль называют также медианой M случайной величины X.*

Квантиль

Для *непрерывной случайной величины* X α-квантиль Q_α является решением уравнения

$$F(Q_\alpha) = \alpha,$$

где $F(x)$ — *функция распределения* случайной величины X.
Таким образом, для непрерывной случайной величины X квантиль Q_α — это такое число, меньше которого X принимает значение с *вероятностью* α.
Если известна *плотность распределения* $f(x)$ случайной величины X, то, учитывая связь между функцией распределения и плотностью распределения, уравнение для определения α-квантили можно записать в виде

$$\int_{-\infty}^{Q_\alpha} f(x)\,dx = \alpha.$$