

Числовые характеристики случайных величин

октябрь–декабрь 2023 г.

Курс лекций «Теория вероятностей»

- Математическое ожидание (среднее)
- Дисперсия
- Моменты
- Медиана

Математическое ожидание

Другое название: среднее значение (expected value, mean value)

Пусть для начала ξ — дискретная случайная величина, заданная на дискретном вероятностном пространстве $\Omega = \{\omega_1, \dots, \omega_n, \dots\}$.

Определение

Число

$$E(\xi) = \sum_i \xi(\omega_i)P\{\omega_i\}$$

называется математическим ожиданием случайной величины X при условии абсолютной сходимости ряда.

Математическое ожидание, II

Как известно, распределение дискретной случайной величины можно задавать таблицей

ξ	x_1	x_2	$\dots$	x_k	$\dots$
$P(\xi)$	p_1	p_2	$\dots$	p_k	$\dots$

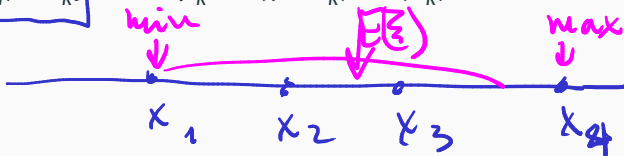
Обозначим $B_k = \{\omega_i : \xi(\omega_i) = x_k\}$ Тогда $p_k = P(\xi = x_k) = P(B_k)$, $k = 1, 2, \dots$

Предложение

$$E(\xi) = \sum_k x_k p_k.$$

Доказательство.

Имеем: $E(\xi) = \sum_i \xi(\omega_i) P\{\omega_i\} = \sum_k \sum_{\omega_i \in B_k} \xi(\omega_i) P\{\omega_i\} = \sum_k x_k \sum_{\omega_i \in B_k} P\{\omega_i\} = \sum_k x_k p_k$,
как и следовало ожидать. □



Математическое ожидание: свойства

Пусть на этом же вероятностном пространстве $\Omega = \{\omega_1, \dots, \omega_n, \dots\}$ задана другая случайная величина η .

Предложение

$$E(\xi + \eta) = E(\xi) + E(\eta).$$

Доказательство.

$$E(\xi + \eta) = \sum_i (\xi + \eta)(\omega_i) P\{\omega_i\} = \sum_i (\xi(\omega_i) + \eta(\omega_i)) P\{\omega_i\} = E(\xi) + E(\eta).$$

□

Аналогично доказывается, что $E(a\xi) = aE(\xi)$ для любой константы a .

Математическое ожидание: упражнение

$$EX = \sum_{k=1}^{\infty} x_k p_k$$

↓
0

Упражнение

Пусть $X \geq 0$ — неотрицательная (дискретная) случайная величина, и $E(X) = 0$.
Докажите, что тогда $P(X = 0) = 1$.

В теории вероятностей принято обозначение: $P(X = 0) = 1 \iff X = 0$ почти наверное (п.н., a.s., almost surely).

Пусть $P(X=0) < 1 \Rightarrow \exists x_i > 0 : p_i > 0 \Rightarrow$
 $\downarrow$
 $P(X > 0) > 0 \quad EX \geq x_i p_i > 0$. Противоречие

Математическое ожидание: замена переменной

$$\eta = g(\xi) \Rightarrow y_1, y_2, \dots, g_1, g_2, \dots, g_i = P(g(\xi) = y_i)$$

Аналогично предыдущему предложению доказывается

$$\sum y_i g_i$$

Предложение

Пусть ξ — случайная величина с заданной таблицей распределения

$\{x_k, p_k, k = 1, 2, \dots\}$ и $g(x)$ — числовая функция числового аргумента.

Обозначим $\eta = g(\xi)$. Тогда

$$E(\eta) = E(g(\xi)) = \sum_k g(x_k) p_k$$

$$E(\eta) = \sum_i i p g^{i-1}$$

Этот результат называют теоремой о замене переменных.

x -геометр. с.б. $x_i = i, p_i = p q^{i-1}, i = 1, 2, \dots$
 $g(x) = e^x \Rightarrow E(g(\xi)) = \sum e^i p q^{i-1}$

Математическое ожидание: общий случай

Пусть X — произвольная случайная величина.

В общей теории средним значением или математическим ожиданием случайной величины X называется интеграл Лебега

$$E(X) = \int_{\Omega} X(\omega)P(d\omega).$$

Доказывается, что для любой ограниченной случайной величины X (измеримой функции) на $(\Omega, \mathcal{F}, P)$ можно корректно определить такой интеграл, причем будут выполняться привычные свойства интеграла — линейность и неотрицательность.

Математическое ожидание: непрерывный случай

Можно показать, что для непрерывной случайной величины X с плотностью $f_X(x)$



$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

$$\sum x_i f_X(x_i) \Delta_i$$

и справедлива теорема о замене переменных. А именно, справедливо

Предложение

Пусть X — случайная величина, имеющая (для простоты) плотность распределения $f_X(x)$ и пусть $h: \mathbb{R} \rightarrow \mathbb{R}$ — измеримая функция. Обозначим

$Y = h(X)$ Тогда

$$E(Y) = \int_{-\infty}^{+\infty} h(x) f_X(x) dx$$

(теорема о замене переменных в интеграле Лебега).

Математическое ожидание: свойства

1. Линейность: $E(aX + bY) = aE(X) + bE(Y)$.

2. Положительность:

- если $X \geq 0$ (п.н.), то $E(X) \geq 0$;
- если $X \geq 0$ (п.н.) и $E(X) = 0$, то $X = 0$ (п.н.).

Строгое доказательство в общем случае следует из свойств интеграла Лебега.



• если $X \geq 0$ (п.н.) и $P(X > 0) > 0$, то $E(X) > 0$

Мат. ожидание: константа

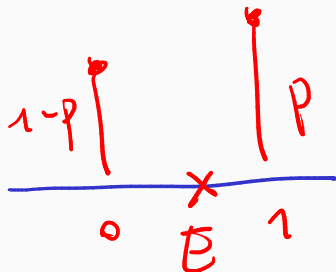
$$E(m) = m \cdot 1 = m$$

m	m	— расст.
$P(m)$	1	

Математическое ожидание, примеры: бернуллиевская величина

Пусть ε — бернуллиевская случайная величина с параметром p . Тогда

$$E(\varepsilon) = 0 \cdot (1 - p) + 1 \cdot p = p.$$



Математическое ожидание, примеры: биномиальная величина

Пусть X — биномиальная случайная величина с параметрами n, p . По определению ($q = 1 - p$)

$$\begin{aligned} E(X) &= \sum_{k=0}^n k C_n^k p^k q^{n-k} = \sum_{k=1}^n k \frac{n!}{k!(n-k)!} p^k q^{n-k} = np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} q^{n-k} \\ &= |m = k-1| = np \sum_{m=0}^{n-1} \frac{(n-1)!}{m!(n-1-m)!} p^m q^{(n-1)-m} = np(p+q)^{n-1} = np. \end{aligned}$$

Результат согласуется с интуицией и здравым смыслом. Этот результат можно получить без рутинных вычислений, если воспользоваться представлением

$$X = \sum_{i=1}^n \varepsilon_i,$$

где ε_i — бернуллиевские случайные величины, и свойством линейности математического ожидания: $E(X) = \sum_{i=1}^n E(\varepsilon_i) = np$.

Математическое ожидание, примеры: пуассоновская величина

Пусть Π — пуассоновская случайная величина с параметром $\lambda > 0$. Тогда

$$E(\Pi) = e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda.$$

Π - пуасс. величина со средним λ .

Математическое ожидание, примеры: геометрическая величина

Число исп. в схеме Бернулли до 1-го успеха (вкл.)

Пусть X — геометрическая случайная величина с параметром p ,

$P(X = k) = pq^{k-1}$, $q = 1 - p$, $k = 1, 2, \dots$. Тогда

$$E(X) = p \sum_{k=1}^{\infty} kq^{k-1} = p(1 + 2q + 3q^2 + \dots) = p(q + q^2 + q^3 + \dots)' = p \left(\frac{q}{1-q} \right)' = p \frac{1}{(1-q)^2} = \frac{1}{p}.$$

Один из переходов (почленное дифференцирование ряда) вытекает из свойств степенных рядов.

Полученный результат, несомненно, соответствует интуиции.

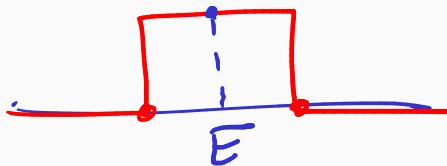
Математическое ожидание, примеры: равномерная величина

Пусть X — дискретная равномерная случайная величина со значениями $x_1, x_2, \dots, x_n$. Тогда

$$E(X) = \sum_{k=1}^n x_k \frac{1}{n} = \frac{x_1 + \dots + x_n}{n} = \bar{x}$$

— среднее арифметическое.

Математическое ожидание, примеры: равномерная величина



Пусть X — равномерная на $[a, b]$. Тогда

$$E(X) = \frac{1}{b-a} \int_a^b x dx = \frac{1}{b-a} \cdot \frac{x^2}{2} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2},$$

как и следовало ожидать.

плотность.

Математическое ожидание, примеры: показательная величина

Часто говорят. "Показ. расстр. со средним θ "
 $\Rightarrow$ плотность $f(x, \theta) = \frac{1}{\theta} e^{-x/\theta} I(x \geq 0)$

Пусть X — показательная с параметром $\lambda > 0$. Интегрированием по частям получаем, что

$$E(X) = \int_0^{+\infty} x \lambda e^{-\lambda x} dx = \int_0^{+\infty} x d(-e^{-\lambda x}) = (-x e^{-\lambda x}) \Big|_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx = (-e^{-\lambda x} / \lambda) \Big|_0^{+\infty} = \frac{1}{\lambda}.$$

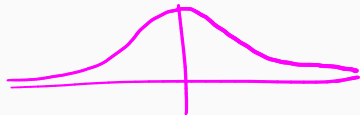
x — время жизни элемента в секундах,
 $\Rightarrow \lambda$ размерность $\frac{1}{\text{сек.}}$

Математическое ожидание, примеры: стандартная нормальная величина

Напоминание. Случайная величина $Z \sim N(0, 1)$ называется стандартной нормальной случайной величиной. Ее плотность распределения есть

$$\varphi(-x) = \varphi(x)$$

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



а функция распределения $\Phi(x) = \int_{-\infty}^x \varphi(y) dy$. Эту функцию называют функцией Лапласа.

Функция $\varphi(x)$ четная, поэтому $E(Z) = \int_{-\infty}^{+\infty} \underbrace{x\varphi(x)}_{\text{нечетное}} dx = 0$ как интеграл нечетной функции (интеграл существует, так как $x e^{-x^2/2}$ стремится к нулю достаточно быстро при $x \rightarrow \pm\infty$).

Математическое ожидание, примеры: нормальная величина

Пусть $Z \sim N(0, 1)$, и пусть $m \in \mathbb{R}$, $\sigma > 0$ — два числа. Определим новую случайную величину $X = m + \sigma Z$.

Предложение

$X \sim N(m, \sigma^2)$

Доказательство.

Имеем: $F_X(x) = P(X \leq x) = P(m + \sigma Z \leq x) = P\left(Z \leq \frac{x - m}{\sigma}\right) = \Phi\left(\frac{x - m}{\sigma}\right)$. Поэтому

$$f_X(x) = F'_X(x) = \frac{1}{\sigma} \Phi'\left(\frac{x - m}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - m)^2}{2\sigma^2}\right).$$

А это и означает, что $X \sim N(m, \sigma^2)$.



Математическое ожидание, примеры: нормальная величина, II

Точно так же доказывается в каком-то смысле обратное утверждение.

Предложение

Если $X \sim N(m, \sigma^2)$, то $\frac{X - m}{\sigma} \sim N(0, 1)$.

Отсюда следует, что $E(X) = E(m + \sigma Z) = m$ — почти очевидный факт.

$$\begin{array}{c} \parallel \\ E(m) + \sigma E(Z) \\ \parallel \qquad \qquad \parallel \\ m \qquad \qquad 0 \end{array}$$

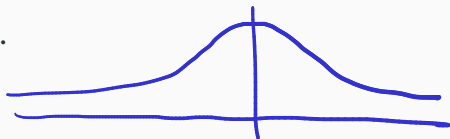
Математическое ожидание: пример, когда не существует

Рассмотрим распределение Коши (оно же распределение Стьюдента $t(1)$):

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

Для него

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



не существует (функция нечетная, но интеграл модуля не существует).

Математическое ожидание: центр распределения

Среднее значение — это число, претендующее на роль центра распределения случайной величины X . Это оправдывается равенством $E(X - m) = 0$.

Если распределение случайной величины X симметричное относительно μ (для непрерывной случайной величины это означает $f(x - \mu) \equiv f(\mu - x)$), то $E(X) = \mu$, если существует.



Дисперсия (variance)

Неформально дисперсия является числовым измерителем интенсивности «колебания» случайной величины X вокруг центра m .

Определение

Число $\text{Var}(X) = E((X - m)^2)$, где $m = E(X)$, называется дисперсией случайной величины X .

Обозначение: $\text{Var}(X) = \sigma_X^2$.

Из теоремы о замене переменных получаем, что

- $\text{Var}(X) = \sum_k (x_k - m)^2 p_k$ для дискретной случайной величины X
- $\text{Var}(X) = \int_{-\infty}^{+\infty} (x - m)^2 f_X(x) dx$ для непрерывной случайной величины X

Определение

Число $\sigma_X = \sqrt{\text{Var}(X)} = \sqrt{\sigma_X^2}$ называется стандартным отклонением случайной величины X .

Если X размерная (сек)
то $\text{Var}(X) \sim \text{сек}^2$

разм. та же, что и X

Дисперсия: свойства

1. $\text{Var}(X) \geq 0$; $\text{Var}(X) = 0 \iff X \equiv m = \text{const}$ (п.н.)
2. $\text{Var}(aX + b) = a^2 \text{Var}(X)$
3. $\text{Var}(X) = E((X - m)^2) = E(X^2 - 2mX + m^2) = E(X^2) - 2mE(X) + m^2 = E(X^2) - m^2 = E(X^2) - (E(X))^2$

Дисперсия суммы

Примеры:

- $Y = X$: $\text{Var}(X + Y) = \text{Var}(2X) = 4 \text{Var}(X) = 2(\text{Var}(X) + \text{Var}(Y)) > \text{Var}(X) + \text{Var}(Y)$
- $Y = -X$: $\text{Var}(X + Y) = \text{Var}(X - X) = 0 < \text{Var}(X) + \text{Var}(Y)$

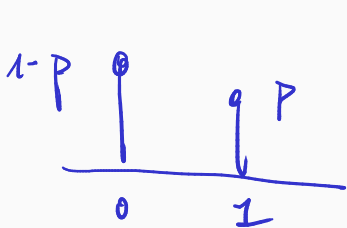
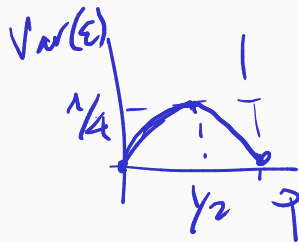
Замечание

В отличие от математического ожидания равенство $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ в общем случае не выполняется. Позже мы выделим класс случайных величин, для которых это равенство выполняется. (Спойлер: независимые, независимые в среднем, некоррелированные)

Дисперсия, примеры: бернуллиевская величина

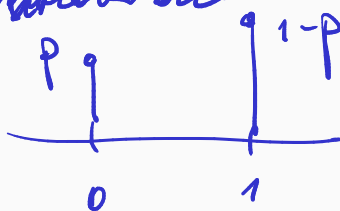
Заметим, что $\varepsilon^2 = \varepsilon$, откуда следует:

$$\text{Var}(\varepsilon) = E(\varepsilon^2) - (E(\varepsilon))^2 = p - p^2 = p(1 - p) = pq$$



var. ~~ошибка~~ *ошибка* ~~base~~

↪



$$\begin{aligned} \text{Var}(1 - \varepsilon) &= \\ &= \text{Var}(-\varepsilon) = (-1)^2 \text{Var} \varepsilon \\ &= \text{Var} \varepsilon_{26} \end{aligned}$$

Дисперсия, примеры: биномиальная величина

Биномиальная $X \sim B(n, p)$: мы показали, что $E(X) = \sum_{k=0}^n k C_n^k p^k q^{n-k} = np$.

С помощью аналогичных рутинных вычислений можно показать, что $\text{Var}(X) = npq$.

Напомним, что, первое равенство есть следствие представления $X = \sum_{i=1}^n \varepsilon_i$, где ε_i — независимые бернуллиевские случайные величины, и линейности математического ожидания.

Выражение для дисперсии тоже можно вывести из этого представления, но при этом надо пользоваться фактом, который мы установим позже: дисперсия суммы независимых случайных величин равна сумме их дисперсий.

$$\sigma_X = \sqrt{npq} \sim \sqrt{n}$$

Дисперсия, примеры: пуассоновская величина

$$E(P^2) = e^{-\lambda} \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} = \lambda^2 + \lambda k = \lambda^2 + \lambda$$

Пусть P — пуассоновская случайная величина с параметром $\lambda > 0$. Доказали, что

$$E(P) = \lambda.$$

← безразм. параметр.

Аналогично можно доказать, что

$$\text{Var}(P) = \lambda.$$

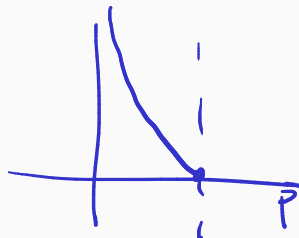
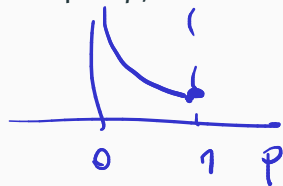
Дисперсия, примеры: геометрическая величина

Пусть X — геометрическая случайная величина с параметром p ,
 $P(X = k) = pq^{k-1}$, $q = 1 - p$, $k = 1, 2, \dots$. Доказали, что

$$E(X) = \frac{1}{p}.$$

Можно показать, что

$$\text{Var}(X) = \frac{1-p}{p^2}.$$



Дисперсия, примеры: равномерная величина

Пусть X — дискретная равномерная случайная величина со значениями $x_1, x_2, \dots, x_n$. Доказали, что $E(X) = \bar{x}$.

Аналогично получаем, что $E(X^2) = \frac{x_1^2 + \dots + x_n^2}{n} = \overline{x^2}$, откуда

$$\text{Var}(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \overline{x^2} - (\bar{x})^2.$$

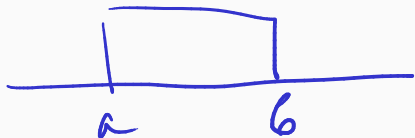
Дисперсия, примеры: равномерная величина

Пусть X — равномерная на $[a, b]$. Доказали, что $E(X) = \frac{a+b}{2}$.

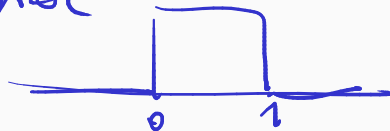
$$E(X^2) = \frac{1}{b-a} \int_a^b x^2 dx = \frac{1}{b-a} \cdot \frac{x^3}{3} \Big|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3},$$

и дисперсия

$$\text{Var}(X) = \frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$



расстояние



Дисперсия, примеры: показательная величина

Пусть X — показательная с параметром $\lambda > 0$. Доказали, что $E(X) = 1/\lambda$.

$$E(X^2) = \lambda \int_0^{+\infty} x^2 e^{-\lambda x} dx = \int_0^{+\infty} x^2 d(-e^{-\lambda x}) = (-x^2 e^{-\lambda x}) \Big|_0^{+\infty} + 2 \int_0^{+\infty} x e^{-\lambda x} dx = \frac{2}{\lambda^2}.$$

Значит, $\text{Var}(X) = 2/\lambda^2 - (1/\lambda)^2 = 1/\lambda^2$.

$$\sigma_x = \frac{1}{\lambda}$$

Дисперсия, примеры: нормальная величина

Нормальная $X \sim N(m, \sigma^2)$. Доказали, что $X = m + \sigma Z$, $Z \sim N(0, 1)$ и $E(X) = m$.

По свойству дисперсии $\text{Var}(X) = \sigma^2 \text{Var}(Z)$.

Заметим, что $d(e^{-x^2/2}) = -xe^{-x^2/2}dx$.

Имеем:

$$\begin{aligned}\text{Var}(Z) = E(Z^2) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x d(-e^{-x^2/2}) dx \\ &= -\frac{1}{\sqrt{2\pi}} x e^{-x^2/2} \Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2} dx = 1\end{aligned}$$

Значит, $\text{Var}(X) = \sigma^2$.

«Нормальная величина с параметрами m и σ^2 » эквивалентно «нормальная величина с матожиданием (со средним) m и дисперсией σ^2 ».

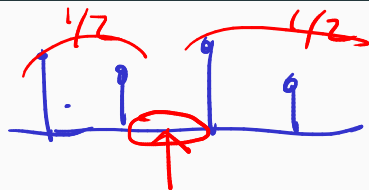
Определение

1. Моментом степени k (k -м моментом) случайной величины X называется число $m_k = E(X^k)$.
2. Число $n_k = E((X - m)^k)$ называется k -м центральным моментом.
3. Число $\mu_k = E(|X|^k)$ называется k -м абсолютным моментом.
4. Число $\nu_k = E(|X - m|^k)$ называется k -м абсолютным центральным моментом.

$$Z \sim N(0, 1)$$

$$E(Z^4) = 3$$

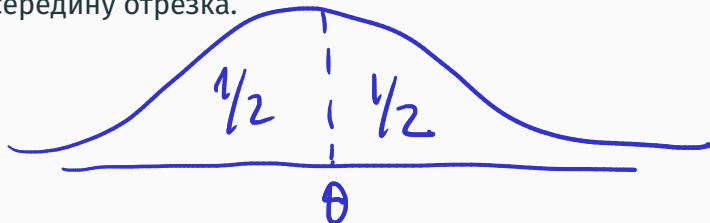
$$E(Z^3) = 0$$



Медиана случайной величины определяется так:

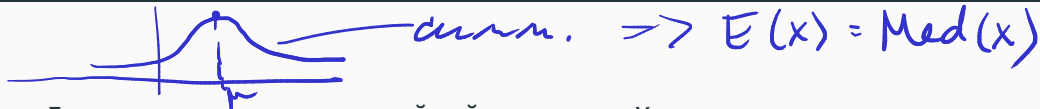
$$\text{Med}(X) = \{\theta : P(X \leq \theta) \geq 1/2, P(X \geq \theta) \geq 1/2\}.$$

Для дискретных величин медиан может быть несколько: для бернуллиевской с $p = 1/2$ весь отрезок $[0, 1]$. В таких случаях берут минимум, максимум или середину отрезка.

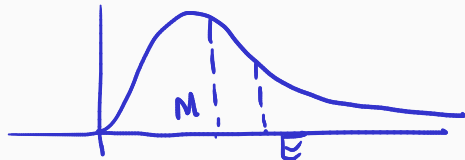


1. Нелинейность: вообще говоря, $\text{Med}(X + Y) \neq \text{Med}(X) + \text{Med}(Y)$
2. $\text{Med}(X + b) = \text{Med}(X) + b$
3. $\text{Med}(aX) = a \text{Med}(X)$ (если рассматривать, скажем, середину отрезка из определения)

Медиана: симметрия



- Если распределение случайной величины X симметричное относительно μ , то медиана $\text{Med}(X) = \mu$ (и равна $E(X)$).
- Если распределение «перекошено» в сторону малых значений (например, показательное), то $\text{Med}(X) < E(X)$. В этом случае также $E(X - m)^3 > 0$.
- Если распределение «перекошено» в сторону больших значений (например, бернуллиевское с $p > 1/2$), то $\text{Med}(X) > E(X)$. В этом случае также $E(X - m)^3 < 0$.



$$\text{Skew} = \frac{E((X - m)^3)}{(\text{Var}(X))^{3/2}}$$

Медиана: пример



Пусть X — показательная случайная величина с параметром $\lambda > 0$. Тогда $F(x) = 1 - e^{-\lambda x}$ для $x \geq 0$ и $\text{Med}(X)$ — решение уравнения $F(x) = 1/2 \iff e^{-\lambda x} = 1/2 \iff \lambda x = \ln 2 \iff x = \ln 2 / \lambda < 1/\lambda$.

Оптимизационное свойство матожидания

$L(x, c)$ — потери
 $E(L(x, c))$ — ожидаемые

«Угадываем» значение случайной величины X . c — догадка, $X - c$ — ошибка.
Функция потерь $E((X - c)^2)$ (ожидаемые потери). Чему должно быть равно c , чтобы ожидаемые потери были наименьшими?

Ответ: $c = E(X)$.

MSE

Оптимизационное свойство медианы

MAE

Рассмотрим ту же задачу, но с ожидаемыми потерями $E|X - c|$. Тогда оптимальное значение — медиана $\text{Med}(X)$.

- [1] Sh. Ross. *A First Course in Probability*. 9th ed. London: Pearson, 2014.
- [2] Б. В. Гнеденко. *Курс теории вероятностей*. М.: Издательство УРСС, 2005.
- [3] С. В. Головань, П. К. Катышев, А. А. Пересецкий. *Сборник задач по курсу теории вероятностей*. М.: Издательский дом «Дело» РАНХиГС, 2021.

Для заметок