ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Е. А. Бакланов

Новосибирск, 2023

1. Основные понятия теории вероятностей
   - Случайные величины. Математическое ожидание
   - Теоремы о предельном переходе под знаком математического ожидания
   - Основные неравенства

Пусть $(\Omega, \mathcal{F})$ — *измеримое пространство*, т. е. непустое множество Ω, состоящее из элементарных исходов ω, с выделенной на нём системой $\mathcal{F}$ его подмножеств (*событий*), образующих σ-алгебру.

Напомним, что система $\mathcal{F}$ подмножеств пространства Ω образует *алгебру*, если $\Omega \in \mathcal{F}$, $\overline{A} \equiv \Omega \setminus A \in \mathcal{F}$ и $A \cup B \in \mathcal{F}$ для любых $A, B \in \mathcal{F}$.

Алгебра $\mathcal{F}$ образует σ-*алгебру*, если вместе с каждой последовательностью событий $A_1, A_2, \ldots,$ принадлежащих $\mathcal{F}$, объединение $\bigcup\limits_{i=1}^{\infty} A_i$ также принадлежит $\mathcal{F}$.

Функция $\mathsf{P}(A)$, определённая на множествах A из σ-алгебры $\mathcal{F}$, называется *вероятностной мерой* или *вероятностью*, если она обладает следующими свойствами:

(i) $\mathsf{P}(A) \geqslant 0$ для любого события $A \in \mathcal{F}$ (неотрицательность);

(ii) $\mathsf{P}(\Omega) = 1$ (нормированность);

(iii) $\mathsf{P}\Big(\bigcup\limits_{i=1}^{\infty} A_i\Big) = \sum\limits_{i=1}^{\infty} \mathsf{P}(A_i)$ для любого счётного набора попарно несовместных событий $A_i \in \mathcal{F}$ (σ-аддитивность).

Тройка $(\Omega, \mathcal{F}, \mathsf{P})$ называется *вероятностным пространством*.

Пусть $\mathcal{A}$ — некоторый класс множеств из Ω. Минимальная σ-алгебра, содержащая $\mathcal{A}$, называется *σ-алгеброй, порождённой классом* $\mathcal{A}$, и обозначается $\sigma(\mathcal{A})$.

Если $\mathcal{A}$ — класс открытых множеств (топологического или метрического пространства), то $\sigma(\mathcal{A})$ называется *борелевской σ-алгеброй*. Элементы борелевской σ-алгебры называют *борелевскими множествами*. Борелевскую σ-алгебру в $\mathbb{R}$ будем обозначать $\mathfrak{B}(\mathbb{R})$.

Функция $g\colon \mathbb{R} \to \mathbb{R}$ называется *борелевской*, если прообраз каждого борелевского множества есть снова борелевское множество: для любого $B \in \mathfrak{B}(\mathbb{R})$

$$g^{-1}(B) \equiv \{x\colon g(x) \in B\} \in \mathfrak{B}(\mathbb{R}).$$

Действительная функция $X = X(\omega)$, определённая на измеримом пространстве $(\Omega, \mathcal{F})$, называется *$\mathcal{F}$-измеримой функцией* или *случайной величиной*, если для любого $B \in \mathfrak{B}(\mathbb{R})$

$$X^{-1}(B) = \{\omega\colon X(\omega) \in B\} \in \mathcal{F}.$$

Простейшим примером случайной величины является *индикатор* $I_A(\omega)$ измеримого множества $A \in \mathcal{F}$:

$$I_A = I_A(\omega) = \begin{cases} 1, & \omega \in A; \\ 0, & \omega \notin A. \end{cases}$$

Случайная величина X называется *простой*, если она принимает конечное число значений. Любая простая случайная величина представима в виде

$$X(\omega) = \sum_{i=1}^{n} a_i I_{A_i}(\omega),$$

где $a_i \in \mathbb{R}$, события $A_i \in \mathcal{F}$ попарно несовместны и $\bigcup\limits_{i=1}^{n} A_i = \Omega$.

Пусть $X = X(\omega)$ — случайная величина. Рассмотрим множества из $\mathcal{F}$ вида $\{\omega\colon X(\omega) \in B\}$, $B \in \mathfrak{B}(\mathbb{R})$. Они образуют σ-алгебру, называемую *σ-алгеброй, порождённой случайной величиной* X. Будем её обозначать $\sigma(X)$.

Вероятностная мера P_X на $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, заданная соотношением

$$P_X(B) = \mathsf{P}\{\omega\colon X(\omega) \in B\}, \quad B \in \mathfrak{B}(\mathbb{R}),$$

называется *распределением вероятностей* случайной величины X на $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$.

Функция

$$F_X(x) = \mathsf{P}\{\omega\colon X(\omega) < x\}, \quad x \in \mathbb{R},$$

называется *функцией распределения* случайной величины X.

Пусть $F(x)$ — некоторая функция распределения на числовой прямой $\mathbb{R}$. Тогда на $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ существует единственная вероятностная мера P такая, что $P[a, b) = F(b) - F(a)$ для любых $-\infty < a < b \leqslant \infty$.

Таким образом, между вероятностными мерами P на $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ и функциями распределения F на числовой прямой $\mathbb{R}$ существует взаимно однозначное соответствие.

Мера P, построенная по функции F, называется вероятностной *мерой Лебега — Стилтьеса*, соответствующей функции распределения F.

Пусть $F(x) = x$, $0 \leqslant x \leqslant 1$. В этом случае соответствующую вероятностную меру называют *мерой Лебега* на отрезке $[0, 1]$. Ясно, что мера Лебега любого из интервалов $[a, b)$, $[a, b]$, $[a, b)$, (a, b) равна его длине $b - a$.

Мерой Лебега — Стилтьеса на $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ назвается любая *мера* (т. е. неотрицательная и σ-аддитивная функция) μ такая, что для любого ограниченного интервала I его мера $\mu(I) < \infty$.

Обобщённой функцией распределения на $\mathbb{R}$ называется любая неубывающая и непрерывная слева функция со значениями в $(-\infty, +\infty)$.

Соотношение

$$\mu[a, b) = F(b) - F(a), \quad a < b,$$

устанавливает взаимно однозначное соответствие между мерами Лебега — Стилтьеса μ и обобщёнными функциями распределения F.

Пусть $F(x) = x$, $x \in \mathbb{R}$. Отвечающая этой обобщённой функции распределения мера μ называется *мерой Лебега* на $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$.

Пусть $(\Omega, \mathcal{F}, \mathsf{P})$ — вероятностное пространство и
$X = X(\omega)$ — неотрицательная случайная величина.
Её *математическое ожидание*, обозначаемое $\mathsf{E}X$, есть *интеграл Лебега* $\int\limits_\Omega X(\omega)\mathsf{P}(d\omega)$, по определению равный

$$\lim_{n\to\infty}\left(\sum_{k=1}^{n2^n}\frac{k-1}{2^n}\mathsf{P}\left(\frac{k-1}{2^n}\leqslant X<\frac{k}{2^n}\right)+n\mathsf{P}(X\geqslant n)\right).$$

Так как $X \geqslant 0$, то интеграл $\int X\,d\mathsf{P} = \int\limits_\Omega X(\omega)\mathsf{P}(d\omega)$ определён, хотя, быть может, и принимает значение $+\infty$.

Пусть $X = X(\omega)$ — произвольная случайная величина и $X^+ = \max\{X, 0\}$, $X^- = -\min\{X, 0\}$. Тогда математическое ожидание $\mathsf{E}X$ определяется только в том случае, когда по крайней мере одна из величин $\mathsf{E}X^+$ или $\mathsf{E}X^-$ конечна, и полагается равным $\mathsf{E}X = \mathsf{E}X^+ - \mathsf{E}X^-$.

Случайная величина $X = X(\omega)$ называется *интегрируемой*, если $\mathsf{E}|X| = \mathsf{E}X^+ + \mathsf{E}X^- < \infty$.

Если $\mathsf{E}X$ определено, то определены также математические ожидания $\mathsf{E}(XI_A)$ для любого $A \in \mathcal{F}$.

Для $\mathsf{E}(XI_A)$ или, что то же, $\int\limits_\Omega XI_A\,d\mathsf{P}$, часто используются обозначения $\mathsf{E}(X;A)$ и $\int\limits_A X\,d\mathsf{P}$.

Интеграл $\int\limits_A X\,d\mathsf{P}$ называется *интегралом Лебега от* X *по множеству* A *(по мере* P*)*.

Пусть μ — произвольная мера, заданная на измеримом пространстве $(\Omega, \mathcal{F})$ и принимающая, быть может, значение $+\infty$, а $X = X(\omega)$ — $\mathcal{F}$-измеримая функция со значениями в $\overline{\mathbb{R}} = [-\infty, +\infty]$ (*расширенная* случайная величина).

В этом случае интеграл Лебега $\int\limits_{\Omega} X(\omega)\,\mu(d\omega)$ определяется точно так же: сначала для неотрицательных X по формуле

$$\int\limits_{\Omega} X(\omega)\,\mu(d\omega) =$$

$$= \lim_{n\to\infty} \left(\sum_{k=1}^{n2^n} \frac{k-1}{2^n}\,\mu\left(\omega\colon \frac{k-1}{2^n} \leqslant X(\omega) < \frac{k}{2^n}\right) + n\mu(\omega\colon X(\omega) \geqslant n) \right),$$

затем для произвольных X по формуле

$$\int\limits_{\Omega} X(\omega)\,\mu(d\omega) = \int\limits_{\Omega} X^{+}(\omega)\,\mu(d\omega) - \int\limits_{\Omega} X^{-}(\omega)\,\mu(d\omega),$$

если только не возникает неопределённость вида $\infty - \infty$.

В случае, когда $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, а μ — мера Лебега, интеграл $\int\limits_{\mathbb{R}} X(x)\,\mu(dx)$ обозначают $\int\limits_{\mathbb{R}} X(x)\,dx$.

Если же мера μ соответствует некоторой обобщённой функции распределения F, то интеграл $\int\limits_{\mathbb{R}} X(x)\,\mu(dx)$ называют *интегралом Лебега — Стилтьеса* (по мере, соответствующей функции распределения F) и обозначают $\int\limits_{\mathbb{R}} X(x)\,dF(x)$.

Пусть $X = X(\omega)$ — случайная величина с функцией распределения F, g — борелевская функция. Если существует один из интегралов $\int\limits_A g(x)\,dF(x)$ или $\int\limits_{X^{-1}(A)} g(X(\omega))\,\mathsf{P}(d\omega)$, то существует и другой, и они совпадают. В частности, при $A = \mathbb{R}$ получаем:

$$\mathsf{E}g(X) = \int\limits_{\Omega} g(X(\omega))\,\mathsf{P}(d\omega) = \int\limits_{\mathbb{R}} g(x)\,dF(x).$$

События $A_1, \ldots, A_n$, $n \geqslant 2$, называются *независимыми* (в совокупности), если для любого набора индексов $1 \leqslant i_1 < \ldots < i_k \leqslant n$

$$\mathsf{P}(A_{i_1} \cap \ldots \cap A_{i_k}) = \mathsf{P}(A_{i_1}) \ldots \mathsf{P}(A_{i_k}).$$

События, составляющие бесконечное семейство, называются *независимыми*, если для каждого $n \geqslant 2$ любые n из этих событий независимы.

Классы событий $\mathcal{A}_1$, $\mathcal{A}_2, \ldots$ называются *независимыми*, если любые события $A_1, A_2, \ldots$ такие, что $A_k \in \mathcal{A}_k$, являются независимыми.

Случайные величины $X_1, X_2, \ldots$, заданные на одном вероятностном пространстве, называются *независимыми*, если независимы порождённые ими σ-алгебры.

Теоремы о предельном переходе под знаком математического ожидания

В этом параграфе будут рассмотрены основные теоремы о предельном переходе под знаком математического ожидания (интеграла Лебега).

Лемма 1.1

Для произвольной неотрицательной случайной величины $X \geqslant 0$ *существует последовательность простых случайных величин* $\{X_n\}_{n\geqslant 1}$ *такая, что* $X_n \uparrow X$, *т. е.* $X_n(\omega) \uparrow X(\omega)$ *для всех* $\omega \in \Omega$.

Доказательство

Для каждого $n = 1, 2, \ldots$ положим

$$X_n(\omega) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I\left\{\omega\colon \frac{k-1}{2^n} \leqslant X(\omega) < \frac{k}{2^n}\right\} + nI(\omega\colon X(\omega) \geqslant n).$$

Непосредственно проверяется, что $X_n(\omega) \uparrow X(\omega)$ для всех $\omega \in \Omega$. $\square$

Теоремы о предельном переходе под знаком математического ожидания

Замечание

Лемма 1.1 является ключевой при построении интеграла Лебега.Так как $X_n \leqslant X_{n+1}$, то $\mathsf{E}X_n \leqslant \mathsf{E}X_{n+1}$ и, следовательно, существует предел $\lim\limits_{n\to\infty} \mathsf{E}X_n$ (который может принимать и значение $+\infty$), который и называется интегралом Лебега от неотрицательной случайной величины X.

Теоремы о предельном переходе под знаком математического ожидания

Теорема 1.1 (теорема о монотонной сходимости)

Пусть $\{X_n\}_{n\geqslant 1}$ *— последовательность неотрицательных случайных величин. Если* $X_n \uparrow X$*, то* $\mathsf{E}X_n \uparrow \mathsf{E}X$.

Теоремы о предельном переходе под знаком математического ожидания

Доказательство

Построим для каждого $k \geqslant 1$ последовательность простых случайных величин $\{X_{k,n}\}_{n \geqslant 1}$ таких, что $X_{k,n} \uparrow X_k$ при $n \to \infty$. Обозначим $Y_n = \max\limits_{1 \leqslant k \leqslant n} X_{k,n}$. Тогда

$$Y_{n-1} \leqslant Y_n \leqslant \max_{1 \leqslant k \leqslant n} X_k = X_n.$$

Пусть $Y = \lim\limits_{n \to \infty} Y_n$. Так как $X_{k,n} \leqslant Y_n \leqslant X_n$ для $1 \leqslant k \leqslant n$, то, переходя к пределу при $n \to \infty$, получим, что $X_k \leqslant Y \leqslant X$ для любого $k \geqslant 1$, а значит, $Y = X$.

Доказательство

Случайные величины Y_n простые и $Y_n \uparrow Y$. Поэтому

$$\mathsf{E}X = \mathsf{E}Y = \lim_{n\to\infty} \mathsf{E}Y_n \leqslant \lim_{n\to\infty} \mathsf{E}X_n.$$

С другой стороны, поскольку $X_n \leqslant X_{n+1} \leqslant X$, то

$$\lim_{n\to\infty} \mathsf{E}X_n \leqslant \mathsf{E}X.$$

Таким образом, $\lim\limits_{n\to\infty} \mathsf{E}X_n = \mathsf{E}X$. $\square$

Следствие 1.1

Пусть $\{X_n\}_{n\geqslant 1}$ *— последовательность неотрицательных случайных величин. Если* $\mathsf{E}X_1 < \infty$ *и* $X_n \downarrow X$*, то* $\mathsf{E}X_n \downarrow \mathsf{E}X$*.*

Доказательство

Достаточно заметить, что $0 \leqslant X_1 - X_n \uparrow X_1 - X$, и воспользоваться теоремой о монотонной сходимости. $\square$

Теоремы о предельном переходе под знаком математического ожидания

Следствие 1.2

Пусть $\{X_n\}_{n\geqslant 1}$ *— последовательность неотрицательных случайных величин. Тогда*

$$\mathsf{E}\left(\sum_{n=1}^{\infty} X_n\right) = \sum_{n=1}^{\infty} \mathsf{E}X_n.$$

Доказательство

Заметим, что $0 \leqslant Z_n \equiv \sum\limits_{i=1}^{n} X_i \uparrow \sum\limits_{i=1}^{\infty} X_i \equiv Z$. По теореме о монотонной сходимости $\mathsf{E}Z_n \uparrow \mathsf{E}Z$. С другой стороны, $\mathsf{E}Z_n = \sum\limits_{i=1}^{n} \mathsf{E}X_i \uparrow \sum\limits_{i=1}^{\infty} \mathsf{E}X_i$.

Значит, $\mathsf{E}Z = \sum\limits_{i=1}^{\infty} \mathsf{E}X_i$. $\square$

Следствие 1.3

Если $\mathsf{E}|X| < \infty$, *то* $\mathsf{E}(|X|I_A) \to 0$ *при* $\mathsf{P}(A) \to 0$.

Доказательство

Положим $X_n = |X|I(|X| \leqslant n)$. Тогда $X_n \uparrow |X|$ и, в силу теоремы о монотонной сходимости, $\mathsf{E}X_n \uparrow \mathsf{E}|X| < \infty$. Следовательно,

$$\mathsf{E}|X|I(|X| > n) = \mathsf{E}|X| - \mathsf{E}|X|I(|X| \leqslant n) \to 0,$$

т. е. для любого $\varepsilon > 0$ найдётся n_0 такое, что $\mathsf{E}|X|I(|X| > n) \leqslant \varepsilon/2$ при $n \geqslant n_0$.

Теоремы о предельном переходе под знаком математического ожидания

Доказательство

Если $\mathsf{P}(A) \leqslant \varepsilon/(2n_0)$, то

$$\begin{aligned}
\mathsf{E}(|X|I_A) &= \mathsf{E}(|X|I_A I(|X| > n_0)) + \mathsf{E}(|X|I_A I(|X| \leqslant n_0)) \leqslant \\
&\leqslant \mathsf{E}|X|I(|X| > n_0) + n_0\mathsf{P}(A) \leqslant \\
&\leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}$$

$\square$

Лемма 1.2 (Фату)

(i) *Пусть* $\{X_n\}_{n \geqslant 1}$ *— последовательность неотрицательных случайных величин. Тогда*

$$\mathsf{E}(\liminf_{n \to \infty} X_n) \leqslant \liminf_{n \to \infty} \mathsf{E} X_n.$$

(ii) *Пусть* $|X_n| \leqslant Y$ *для всех* $n \geqslant 1$ *и* $\mathsf{E} Y < \infty$. *Тогда*

$$\mathsf{E}(\liminf_{n \to \infty} X_n) \leqslant \liminf_{n \to \infty} \mathsf{E} X_n \leqslant \limsup_{n \to \infty} \mathsf{E} X_n \leqslant \mathsf{E}(\limsup_{n \to \infty} X_n).$$

Теоремы о предельном переходе под знаком математического ожидания

Доказательство

Положим $Y_n = \inf\limits_{k \geqslant n} X_k$. Тогда $Y_n \uparrow \liminf\limits_{n\to\infty} X_n$. По теореме о монотонной сходимости

$$\lim_{n\to\infty} \mathsf{E}Y_n = \mathsf{E}(\lim_{n\to\infty} Y_n) = \mathsf{E}(\liminf_{n\to\infty} X_n).$$

Так как $Y_n \leqslant X_n$, то

$$\lim_{n\to\infty} \mathsf{E}Y_n = \liminf_{n\to\infty} \mathsf{E}Y_n \leqslant \liminf_{n\to\infty} \mathsf{E}X_n.$$

Теоремы о предельном переходе под знаком математического ожидания

Доказательство

Так как $Y - X_n \geqslant 0$, $Y + X_n \geqslant 0$ и

$$\liminf_{n\to\infty}(Y - X_n) = Y - \limsup_{n\to\infty} X_n, \quad \liminf_{n\to\infty}(Y + X_n) = Y + \liminf_{n\to\infty} X_n,$$

то второе утверждение следует из первого. $\square$

Теоремы о предельном переходе под знаком математического ожидания

Теорема 1.2 (теорема Лебега о мажорируемой сходимости)

Пусть $\{X_n\}_{n\geqslant 1}$ *— последовательность случайных величин таких, что* $|X_n| \leqslant Y$, $\mathsf{E}Y < \infty$ *и* $X_n \to X$. *Тогда* $\mathsf{E}|X| < \infty$, $\mathsf{E}X_n \to \mathsf{E}X$ *и* $\mathsf{E}|X_n - X| \to 0$.

Теоремы о предельном переходе под знаком математического ожидания

Доказательство

По условию $\liminf\limits_{n\to\infty} X_n = \limsup\limits_{n\to\infty} X_n = X$. Поэтому в силу леммы Фату

$$\mathsf{E}X = \mathsf{E}(\liminf_{n\to\infty} X_n) \leqslant \liminf_{n\to\infty} \mathsf{E}X_n \leqslant \limsup_{n\to\infty} \mathsf{E}X_n \leqslant \mathsf{E}(\limsup_{n\to\infty} X_n) = \mathsf{E}X,$$

т. е. $\lim\limits_{n\to\infty} \mathsf{E}X_n = \mathsf{E}X$. Очевидно также, что $|X| \leqslant Y$. Поэтому $\mathsf{E}|X| < \infty$.

Сходимость $\mathsf{E}|X_n - X| \to 0$ доказывается точно так же, если только заметить, что $|X_n - X| \leqslant 2Y$. $\square$

Следствие 1.4

Пусть $X_n \to X$, $|X_n| \leqslant Y$, $\mathsf{E}Y^p < \infty$ *для некоторого* $p > 0$. *Тогда* $\mathsf{E}|X|^p < \infty$ *и* $\mathsf{E}|X_n - X|^p \to 0$.

Доказательство

Достаточно заметить, что $|X| \leqslant Y$ и

$$|X_n - X|^p \leqslant (|X_n| + |X|)^p \leqslant (2Y)^p. \qquad \square$$

Приведённые в этом параграфе неравенства (для математических ожиданий) находят широкое применение и в теории вероятностей, и в математическом анализе. Сначала мы получим ряд неравенств, называемых *неравенствами Чебышёва*.

Теорема 1.3 (обобщённое неравенство Чебышёва)

Пусть g *— неотрицательная и неубывающая функция такая, что* $\mathsf{E}g(X) < \infty$. *Тогда для всех* x *таких, что* $g(x) > 0$

$$\mathsf{P}(X \geqslant x) \leqslant \frac{\mathsf{E}g(X)}{g(x)}. \tag{1.1}$$

Доказательство

Заметим, что $I_A + I_{\overline{A}} = 1$ для любого события A. Тогда

$$\begin{aligned} \mathsf{E}g(X) &= \mathsf{E}\big(g(X)I(X \geqslant x)\big) + \mathsf{E}\big(g(X)I(X < x)\big) \geqslant \\ &\geqslant \mathsf{E}\big(g(X)I(X \geqslant x)\big) \geqslant g(x)\mathsf{E}I(X \geqslant x) = \\ &= g(x)\mathsf{P}(X \geqslant x), \end{aligned}$$

где $I(A)$ также обозначает индикатор события A. □

Выбирая подходящие функции g, из неравенства (1.1) получаем следующие известные неравенства.

Неравенство Маркова

Пусть $\mathsf{E}|X|^p < \infty$ *для некоторого* $p > 0$. *Тогда для всех* $x > 0$

$$\mathsf{P}(|X| \geqslant x) \leqslant \frac{\mathsf{E}|X|^p}{x^p}. \tag{1.2}$$

Неравенство Чебышёва

Пусть $\mathsf{D}X < \infty$. *Тогда для всех* $x > 0$

$$\mathsf{P}(|X - \mathsf{E}X| \geqslant x) \leqslant \frac{\mathsf{D}X}{x^2}. \tag{1.3}$$

Основные неравенства

Экспоненциальное неравенство Чебышёва

Пусть $\mathsf{E}e^{\lambda X} < \infty$ *для некоторого* $\lambda > 0$. *Тогда для всех* $x \in \mathbb{R}$

$$\mathsf{P}(X \geqslant x) \leqslant \frac{\mathsf{E}e^{\lambda X}}{e^{\lambda x}}. \tag{1.4}$$

Лемма 1.3 (неравенство Кантелли)

Пусть $\mathsf{D}X < \infty$. *Тогда для всех* $x > 0$

$$\mathsf{P}(X - \mathsf{E}X \geqslant x) \leqslant \frac{\mathsf{D}X}{x^2 + \mathsf{D}X}.$$

Доказательство

Пусть t — произвольное положительное число. Тогда

$$\begin{aligned}
\mathsf{P}(X - \mathsf{E}X \geqslant x) &= \mathsf{P}(X - \mathsf{E}X + t \geqslant x + t) \leqslant \\
&\leqslant \mathsf{P}(|X - \mathsf{E}X + t| \geqslant x + t) \leqslant \\
&\leqslant \frac{\mathsf{E}(X - \mathsf{E}X + t)^2}{(x+t)^2} = \frac{\mathsf{D}X + t^2}{(x+t)^2}.
\end{aligned}$$

В силу произвольности t получаем:

$$\mathsf{P}(X - \mathsf{E}X \geqslant x) \leqslant \inf_{t>0} \frac{\mathsf{D}X + t^2}{(x+t)^2} = \frac{\mathsf{D}X}{x^2 + \mathsf{D}X}.$$

□

Определение

Функция $g = g(x)$ называется *выпуклой*, если для любых $x, y \in \mathbb{R}$ и для любого $\alpha \in [0, 1]$

$$\alpha g(x) + (1 - \alpha)g(y) \geqslant g(\alpha x + (1 - \alpha)y). \tag{1.5}$$

Докажем одно полезное свойство выпуклых функций.

Лемма 1.4

Пусть g — выпуклая функция. Тогда для каждого $z \in \mathbb{R}$ найдётся число $c = c(z)$ такое, что для всех $x \in \mathbb{R}$

$$g(x) \geqslant g(z) + c(x - z).$$

Основные неравенства

Доказательство

Пусть $x < z < y$. Тогда $z = \alpha x + (1-\alpha)y$, где $\alpha = \dfrac{y-z}{y-x}$. В силу выпуклости функции g имеем:

$$\alpha g(z) + (1-\alpha)g(z) = g(z) = g(\alpha x + (1-\alpha)y) \leqslant \alpha g(x) + (1-\alpha)g(y).$$

Отсюда получаем:

$$\begin{aligned} \frac{y-z}{y-x}(g(z)-g(x)) &= \alpha(g(z)-g(x)) \leqslant (1-\alpha)(g(y)-g(z)) = \\ &= \frac{z-x}{y-x}(g(y)-g(z)), \end{aligned}$$

и, следовательно,

$$\frac{g(z)-g(x)}{z-x} \leqslant \frac{g(y)-g(z)}{y-z}.$$

Доказательство

Зафиксируем точку $z \in \mathbb{R}$. Тогда найдётся число $c \in \mathbb{R}$ такое, что для всех $x < z$ и всех $y > z$

$$\frac{g(z) - g(x)}{z - x} \leqslant c \leqslant \frac{g(y) - g(z)}{y - z}.$$

Отсюда следует, что $g(x) \geqslant g(z) + c(x - z)$ при $x < z$ и $g(y) \geqslant g(z) + c(y - z)$ при $y > z$, т. е. $g(x) \geqslant g(z) + c(x - z)$ для всех $x \in \mathbb{R}$. $\square$

Теорема 1.4 (неравенство Йенсена)

Пусть $\mathsf{E}|X| < \infty$ *и пусть* g *— выпуклая функция. Тогда*

$$g(\mathsf{E}X) \leqslant \mathsf{E}g(X). \tag{1.6}$$

Основные неравенства

Доказательство

Так как g — выпуклая функция, то в силу леммы 1.4 для каждого $y \in \mathbb{R}$ найдётся число $c(y)$ такое, что для всех $x \in \mathbb{R}$

$$g(x) \geqslant g(y) + c(y)(x - y).$$

Положив здесь $x = X$, $y = \mathsf{E}X$ и взяв математическое ожидание обеих частей этого неравенства, получим:

$$\mathsf{E}g(X) \geqslant g(\mathsf{E}X).$$

$\square$

Из неравенства Йенсена выводится ряд полезных неравенств. Например, получим следующее неравенство, связывающее моменты разных порядков.

Теорема 1.5 (неравенство Ляпунова)

Если $0 < s < t$, *то*

$$(\mathsf{E}|X|^s)^{1/s} \leqslant (\mathsf{E}|X|^t)^{1/t}. \tag{1.7}$$

Доказательство

Так как $r = t/s > 1$, то функция $g(x) = |x|^r$ — выпуклая. Из неравенства Йенсена для $Y = |X|^s$ находим, что $|\mathsf{E}Y|^r \leqslant \mathsf{E}|Y|^r$, т. е. $(\mathsf{E}|X|^s)^{t/s} \leqslant \mathsf{E}|X|^t$, что и доказывает (1.7). □

Основные неравенства

Из неравенства Ляпунова вытекает следующая цепочка неравенств между абсолютными моментами:

$$\mathsf{E}|X| \leqslant (\mathsf{E}|X|^2)^{1/2} \leqslant \ldots \leqslant (\mathsf{E}|X|^m)^{1/m} \leqslant \ldots \qquad (1.8)$$

Теорема 1.6 (неравенство Гёльдера)

Пусть $1 < p < \infty$ *и* $1/p + 1/q = 1$. *Тогда*

$$\mathsf{E}|XY| \leqslant (\mathsf{E}|X|^p)^{1/p}(\mathsf{E}|Y|^q)^{1/q}. \tag{1.9}$$

Основные неравенства

Доказательство

Если $\mathsf{E}|X|^p = 0$ или $\mathsf{E}|Y|^q = 0$, то неравенство (1.9) очевидно. Пусть $\mathsf{E}|X|^p > 0$, $\mathsf{E}|Y|^q > 0$. Положим

$$\widetilde{X} = \frac{|X|}{(\mathsf{E}|X|^p)^{1/p}}, \quad \widetilde{Y} = \frac{|Y|}{(\mathsf{E}|Y|^q)^{1/q}}.$$

Воспользуемся *неравенством Юнга*

$$xy \leqslant \frac{x^p}{p} + \frac{y^q}{q}, \quad x, y > 0, \quad 1/p + 1/q = 1, \tag{1.10}$$

вытекающим непосредственно из выпуклости функции $g(x) = -\ln x$:

$$xy = e^{\ln(xy)} = \exp\left\{\frac{\ln x^p}{p} + \frac{\ln y^q}{q}\right\} \leqslant \exp\left\{\ln\left(\frac{x^p}{p} + \frac{y^q}{q}\right)\right\} = \frac{x^p}{p} + \frac{y^q}{q}.$$

Доказательство

Из (1.10) получаем

$$\mathsf{E}(\widetilde{X}\widetilde{Y}) \leqslant \frac{1}{p}\,\mathsf{E}\widetilde{X}^p + \frac{1}{q}\,\mathsf{E}\widetilde{Y}^q = 1/p + 1/q = 1,$$

что и доказывает (1.9). $\square$

Основные неравенства

Частным случаем неравенства Гёльдера при $p = q = 2$ является *неравенство Коши — Буняковского*:

$$\mathsf{E}|XY| \leqslant \sqrt{\mathsf{E}X^2}\,\sqrt{\mathsf{E}Y^2}.$$

Основные неравенства

Для доказательства ещё одного известного неравенства — *неравенства Минковского* — нам понадобится так называемое C_r-*неравенство*.

Лемма 1.5

Пусть $r > 0$. *Тогда*

$$|x + y|^r \leqslant C_r(|x|^r + |y|^r), \tag{1.11}$$

где $C_r = \max\{1, 2^{r-1}\}$.

Основные неравенства

Доказательство

Без ограничения общности можно считать, что $x, y > 0$. Пусть $0 < r \leqslant 1$. Так как в этом случае $x^{1/r} \leqslant x$ для всех $0 < x \leqslant 1$, то

$$\left(\frac{x^r}{x^r + y^r}\right)^{1/r} + \left(\frac{y^r}{x^r + y^r}\right)^{1/r} \leqslant \frac{x^r}{x^r + y^r} + \frac{y^r}{x^r + y^r} = 1.$$

Отсюда находим, что $(x + y) \leqslant (x^r + y^r)^{1/r}$, т. е. $(x + y)^r \leqslant x^r + y^r$. Пусть теперь $r \geqslant 1$. Тогда функция $|x|^r$ выпукла и, следовательно,

$$\left(\frac{x + y}{2}\right)^r \leqslant \frac{x^r + y^r}{2},$$

что и даёт неравенство (1.11). □

Замечание

C_r-неравенство нетрудно обобщить на случай n слагаемых:

$$\left|\sum_{i=1}^{n} x_i\right|^r \leqslant \max\{1, n^{r-1}\} \sum_{i=1}^{n} |x_i|^r, \quad r > 0. \tag{1.12}$$

Сначала покажем, что функция $g = g(x)$ выпукла тогда и только тогда, когда для любых $x_i \in \mathbb{R}$ и для любых $\alpha_i \geqslant 0$, $1 \leqslant i \leqslant n$, $\alpha_1 + \ldots + \alpha_n = 1$,

$$g(\alpha_1 x_1 + \ldots + \alpha_n x_n) \leqslant \alpha_1 g(x_1) + \ldots + \alpha_n g(x_n). \tag{1.13}$$

Очевидно, из (1.13) следует (1.5). Пусть теперь выполнено (1.5) и пусть неравенство (1.13) справедливо для некоторого $n > 2$. Тогда, учитывая равенство $1 - \alpha_{n+1} = \alpha_1 + \ldots + \alpha_n$, получаем:

$$\begin{aligned}
& g(\alpha_1 x_1 + \ldots + \alpha_{n+1} x_{n+1}) \leqslant \\
\leqslant\ & \alpha_{n+1} g(x_{n+1}) + (1 - \alpha_{n+1})\, g\left(\frac{\alpha_1 x_1 + \ldots + \alpha_n x_n}{1 - \alpha_{n+1}}\right) \leqslant \\
\leqslant\ & \alpha_{n+1} g(x_{n+1}) + (1 - \alpha_{n+1}) \sum_{i=1}^{n} \frac{\alpha_i}{1 - \alpha_{n+1}}\, g(x_i) = \\
=\ & \alpha_1 g(x_1) + \ldots + \alpha_{n+1} g(x_{n+1}).
\end{aligned}$$

Основные неравенства

Из (1.13), в частности, получаем следующее неравенство для выпуклых функций:

$$g\left(\frac{x_1 + \ldots + x_n}{n}\right) \leqslant \frac{g(x_1) + \ldots + g(x_n)}{n}.$$

Теперь неравенство (1.12) доказывается точно так же, как и (1.11). Действительно, если $0 < r \leqslant 1$, то

$$\begin{aligned}
&\frac{|x_1| + \ldots + |x_n|}{(|x_1|^r + \ldots + |x_n|^r)^{1/r}} = \\
= \quad &\left(\frac{|x_1|^r}{|x_1|^r + \ldots + |x_n|^r}\right)^{1/r} + \ldots + \left(\frac{|x_n|^r}{|x_1|^r + \ldots + |x_n|^r}\right)^{1/r} \leqslant \\
\leqslant \quad &\frac{|x_1|^r}{|x_1|^r + \ldots + |x_n|^r} + \ldots + \frac{|x_n|^r}{|x_1|^r + \ldots + |x_n|^r} = 1.
\end{aligned}$$

Если $r \geqslant 1$, то из выпуклости функции $|x|^r$ и неравенства (1.13) получаем:

$$\left|\frac{x_1 + \ldots + x_n}{n}\right|^r \leqslant \frac{|x_1|^r + \ldots + |x_n|^r}{n}.$$

Итак, неравенство (1.12) доказано. □

Основные неравенства

Теорема 1.7 (неравенство Минковского)

Если $1 \leqslant p < \infty$, *то*

$$(\mathsf{E}|X+Y|^p)^{1/p} \leqslant (\mathsf{E}|X|^p)^{1/p} + (\mathsf{E}|Y|^p)^{1/p}. \quad (1.14)$$

Основные неравенства

Доказательство

Из неравенства (1.11) следует, что $\mathsf{E}|X+Y|^p < \infty$, если $\mathsf{E}|X|^p < \infty$ и $\mathsf{E}|Y|^p < \infty$.
Если $p = 1$, то неравенство (1.14) следует из (1.11).
Пусть $p > 1$. Тогда

$$|X+Y|^p = |X+Y||X+Y|^{p-1} \leqslant |X||X+Y|^{p-1} + |Y||X+Y|^{p-1}. \quad (1.15)$$

Пусть $q = p/(p-1) > 1$. Тогда

$$\mathsf{E}(|X+Y|^{p-1})^q = \mathsf{E}|X+Y|^p < \infty,$$

и, значит, в силу неравенства Гёльдера

$$\begin{aligned} \mathsf{E}(|X||X+Y|^{p-1}) &\leqslant (\mathsf{E}|X|^p)^{1/p} \left(\mathsf{E}|X+Y|^{(p-1)q}\right)^{1/q} = \\ &= (\mathsf{E}|X|^p)^{1/p} (\mathsf{E}|X+Y|^p)^{1/q}. \end{aligned}$$

Основные неравенства

Доказательство

Точно так же и

$$\mathsf{E}(|Y||X+Y|^{p-1}) \leqslant (\mathsf{E}|Y|^p)^{1/p} (\mathsf{E}|X+Y|^p)^{1/q} .$$

Поэтому в силу (1.15)

$$\mathsf{E}|X+Y|^p \leqslant (\mathsf{E}|X+Y|^p)^{1/q} \left((\mathsf{E}|X|^p)^{1/p} + (\mathsf{E}|Y|^p)^{1/p} \right) . \tag{1.16}$$

Если $\mathsf{E}|X+Y|^p = 0$, то неравенство (1.14) очевидно. Если $\mathsf{E}|X+Y|^p > 0$, то из (1.16) получаем:

$$(\mathsf{E}|X+Y|^p)^{1-1/q} \leqslant (\mathsf{E}|X|^p)^{1/p} + (\mathsf{E}|Y|^p)^{1/p},$$

что и доказывает неравенство (1.14), поскольку $1 - 1/q = 1/p$. $\square$