

ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Е. А. Бакланов

Новосибирск, 2023

1 Основные понятия теории вероятностей

2 Вероятностные неравенства

- Неравенства для распределения максимума сумм независимых случайных величин

Неравенства для распределения максимума сумм независимых случайных величин

Пусть $\{X_n\}_{n \geq 1}$ — последовательность независимых случайных величин, $S_k = X_1 + \dots + X_k$, $\bar{S}_n = \max_{1 \leq k \leq n} S_k$.

В настоящем параграфе будут получены оценки для распределения \bar{S}_n , которые играют важное значение при изучении предельного поведения сумм независимых случайных величин и случайных процессов.

Неравенства для распределения максимума сумм независимых случайных величин

Лемма 2.7

Пусть на одном вероятностном пространстве (Ω, \mathcal{F}, P) заданы событие A и случайная величина τ , принимающая только положительные целые значения ($A \in \mathcal{F}$ и $\{\tau = k\} \in \mathcal{F}, k \geq 1$). Тогда для всех $n \geq 1$ имеет место неравенство

$$P(\tau \leq n) \leq \frac{P(A)}{\min_{1 \leq k \leq n} P(A \mid \tau = k)}. \quad (2.21)$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

В силу формулы полной вероятности

$$\begin{aligned} P(A) &= \sum_{k=1}^{\infty} P(A \mid \tau = k)P(\tau = k) \geq \sum_{k=1}^n P(A \mid \tau = k)P(\tau = k) \geq \\ &\geq \min_{k \leq n} P(A \mid \tau = k)P(\tau \leq n), \end{aligned}$$

откуда и следует неравенство (2.21). □

Неравенства для распределения максимума сумм независимых случайных величин

Теорема 2.9

Для всех $x, y \in \mathbb{R}$ имеет место неравенство

$$P(\bar{S}_n \geq x) \leq \frac{P(S_n \geq x - y)}{\min_{1 \leq k \leq n} P(S_n - S_k \geq -y)}. \quad (2.22)$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

В лемме 2.7 положим $\tau = \min\{k \geq 1 : S_k \geq x\}$, $A = \{S_n \geq x - y\}$ и заметим, что $\{\tau \leq n\} = \{\bar{S}_n \geq x\}$ и $S_k \geq x$ на множестве $\{\tau = k\}$.

Тогда

$$\begin{aligned} P(A \mid \tau = k) &= P(S_n \geq x - y \mid \tau = k) = \\ &= P(S_k + S_n - S_k \geq x - y \mid \tau = k) \geq \\ &\geq P(x + S_n - S_k \geq x - y \mid \tau = k) = \\ &= P(S_n - S_k \geq -y \mid \tau = k). \end{aligned} \tag{2.23}$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Нетрудно заметить, что разность $S_n - S_k = X_{k+1} + \dots + X_n$ измерима относительно σ -алгебры $\sigma(X_{k+1}, \dots, X_n)$. С другой стороны,

$$\{\tau = k\} = \{S_1 < x, \dots, S_{k-1} < x, S_k \geq x\} \in \sigma(X_1, \dots, X_k).$$

Поэтому события $\{S_n - S_k \geq -y\}$ и $\{\tau = k\}$ независимы и, следовательно,

$$P(S_n - S_k \geq -y \mid \tau = k) = P(S_n - S_k \geq -y).$$

Тогда для всех $k \leq n$ из (2.23) получаем:

$$P(A \mid \tau = k) \geq P(S_n - S_k \geq -y \mid \tau = k) = P(S_n - S_k \geq -y).$$

Теперь из (2.21) получаем неравенство (2.22). □

Неравенства для распределения максимума сумм независимых случайных величин

Следствие 2.8 (неравенство Леви — Колмогорова)

Пусть $EX_j = 0$, $DX_j < \infty$, $j \leq n$. Тогда для всех $x \in \mathbb{R}$

$$P(\bar{S}_n \geq x) \leq 2P(S_n \geq x - \sqrt{2DS_n}). \quad (2.24)$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

В теореме 2.9 положим $y = \sqrt{2DS_n}$ и покажем, что для всех $k \leq n$

$$P(S_n - S_k \geq -y) \geq \frac{1}{2}.$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Действительно,

$$\begin{aligned} P(S_n - S_k < -y) &= P(S_k - S_n > y) \leq P(|S_k - S_n| > y) \leq \\ &\leq \frac{D(S_k - S_n)}{y^2} = \frac{1}{y^2} D\left(\sum_{j=k+1}^n X_j\right) = \\ &= \frac{1}{y^2} \sum_{j=k+1}^n DX_j \leq \frac{1}{y^2} \sum_{j=1}^n DX_j = \frac{DS_n}{y^2} = \frac{1}{2}, \end{aligned}$$

откуда следует неравенство (2.24). □

Неравенства для распределения максимума сумм независимых случайных величин

Определение

Говорят, что две случайные величины *одинаково распределены*, и пишут $X \stackrel{d}{=} Y$, если $P(X \in B) = P(Y \in B)$ для любого $B \in \mathfrak{B}(\mathbb{R})$.

Заметим, что эти случайные величины могут быть заданы на разных вероятностных пространствах.

Определение

Говорят, что случайная величина X имеет *симметричное распределение*, если $X \stackrel{d}{=} -X$.

Неравенства для распределения максимума сумм независимых случайных величин

Отметим некоторые свойства симметричных случайных величин.

1. Если случайная величина X имеет симметричное распределение и её математическое ожидание существует, то $EX = 0$.
2. Если случайная величина X имеет симметричное распределение, то $P(X \geq 0) \geq 1/2$.

Действительно,

$$\begin{aligned} 2P(X \geq 0) &= P(X \geq 0) + P(-X \geq 0) = P(X \geq 0) + P(X \leq 0) = \\ &= 1 + P(X = 0) \geq 1. \end{aligned}$$

Неравенства для распределения максимума сумм независимых случайных величин

3. Если X_1, \dots, X_n — независимые случайные величины, имеющие симметричное распределение, то $S_n = X_1 + \dots + X_n$ имеет симметричное распределение.

Для доказательства этого свойства достаточно рассмотреть две независимые симметричные случайные величины X_1 и X_2 и воспользоваться формулой свёртки. Пусть F и G — функции распределения X_1 и X_2 соответственно. В силу симметричности имеем $F(x) = P(-X_1 < x)$ и $G(x) = P(-X_2 < x)$. Тогда

$$\begin{aligned} P(-(X_1 + X_2) < x) &= P(-X_1 + (-X_2) < x) = \int_{-\infty}^{+\infty} F(x - t) dG(t) = \\ &= P(X_1 + X_2 < x). \end{aligned}$$

Неравенства для распределения максимума сумм независимых случайных величин

Следствие 2.9 (неравенство Леви)

Пусть X_1, \dots, X_n — независимые случайные величины, имеющие симметричное распределение. Тогда для всех $x \in \mathbb{R}$

$$P(\bar{S}_n \geq x) \leq 2P(S_n \geq x). \quad (2.25)$$

Доказательство

Это неравенство есть непосредственное следствие теоремы 2.9. Действительно, положив $y = 0$ в (2.22) и воспользовавшись свойствами симметричного распределения, получаем:

$$P(S_n - S_k \geq 0) = P(X_{k+1} + \dots + X_n \geq 0) \geq 1/2.$$

Следовательно, $\min_{1 \leq k \leq n} P(S_n - S_k \geq 0) \geq 1/2$. Отсюда следует (2.25). □

Неравенства для распределения максимума сумм независимых случайных величин

Замечание

Применяя неравенство Леви (2.25) к случайным величинам $-X_j$, получаем:

$$P\left(\max_{1 \leq k \leq n} (-S_k) \geq x\right) \leq 2P(-S_n \geq x),$$

и, следовательно, для всех $x > 0$

$$\begin{aligned} P\left(\max_{1 \leq k \leq n} |S_k| \geq x\right) &\leq P\left(\max_{1 \leq k \leq n} S_k \geq x\right) + P\left(\max_{1 \leq k \leq n} (-S_k) \geq x\right) \leq \\ &\leq 2P(S_n \geq x) + 2P(-S_n \geq x) = \\ &= 2P(|S_n| \geq x). \end{aligned} \tag{2.26}$$

Неравенства для распределения максимума сумм независимых случайных величин

Определение

Медианой случайной величины X называется любое число $\text{med}(X)$, удовлетворяющее неравенствам

$$P(X \geq \text{med}(X)) \geq 1/2 \quad \text{и} \quad P(X \leq \text{med}(X)) \geq 1/2.$$

Неравенства для распределения максимума сумм независимых случайных величин

Медиана существует у любой случайной величины, но может быть не единственна.

Пусть случайная величина X имеет функцию распределения F . Если уравнение $F(t) = 1/2$ имеет единственное решение, то это решение и является медианой.

Если уравнение $F(t) = 1/2$ имеет бесконечно много решений, то существует отрезок I такой, что $F(t) = 1/2$ для всех $t \in I$. Любая точка этого отрезка является медианой.

Если же уравнение $F(t) = 1/2$ не имеет решений, то найдётся точка t_0 такая, что $F(t_0) < 1/2$ и $F(t_0 + 0) \geq 1/2$. Эта точка t_0 и будет медианой.

Неравенства для распределения максимума сумм независимых случайных величин

Отметим также, что

$$F(\text{med}(X)) \leq \frac{1}{2} \leq F(\text{med}(X) + 0)$$

и $\text{med}(-X) = -\text{med}(X)$. В частности, если X имеет симметричное распределение, то 0 является медианой X .

Неравенства для распределения максимума сумм независимых случайных величин

Следствие 2.10

Для всех $x \in \mathbb{R}$

$$P(\bar{S}_n > x) \leq 2P(S_n > x - \max_{1 \leq k \leq n} \text{med}(S_k - S_n)).$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

В теореме 2.9 положим $y = \max_{1 \leq k \leq n} \text{med}(S_k - S_n)$ и покажем, что для всех $k \leq n$

$$P(S_n - S_k > -y) \geq 1/2.$$

Действительно, $y \geq \text{med}(S_k - S_n) = -\text{med}(S_n - S_k)$, т. е. $\text{med}(S_n - S_k) \geq -y$ для всех $k \leq n$. Тогда по определению медианы

$$P(S_n - S_k > -y) \geq P(S_n - S_k > \text{med}(S_n - S_k)) \geq 1/2.$$



Неравенства для распределения максимума сумм независимых случайных величин

Лемма 2.8

Пусть случайные величины X и τ заданы на одном вероятностном пространстве, $X > 0$, $\tau \in \mathbb{N}$. Тогда для любого $n \geq 1$

$$P(\tau \leq n) \leq \frac{EX}{\min_{1 \leq k \leq n} E(X|\tau = k)}.$$

Замечание

Лемма 2.7 есть частный случай леммы 2.8 при $X = I_A$.

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Так как $\sum_{k=1}^{\infty} I(\tau = k) = 1$ и

$$E(X | \tau = k) = \frac{E(X; \tau = k)}{P(\tau = k)} = \frac{E(XI(\tau = k))}{P(\tau = k)},$$

то

$$\begin{aligned} EX &= \sum_{k=1}^{\infty} E(XI(\tau = k)) = \sum_{k=1}^{\infty} P(\tau = k)E(X | \tau = k) \geq \\ &\geq \sum_{k=1}^n P(\tau = k)E(X | \tau = k) \geq P(\tau \leq n) \min_{1 \leq k \leq n} E(X | \tau = k). \end{aligned}$$



Неравенства для распределения максимума сумм независимых случайных величин

Теорема 2.10

Пусть g — неотрицательная и неубывающая функция. Тогда для всех $x \in \mathbb{R}$

$$P(\bar{S}_n \geq x) \leq \frac{Eg(S_n)}{\min_{1 \leq k \leq n} Eg(S_n - S_k + x)}. \quad (2.27)$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Положим в лемме 2.8 $X = g(S_n) \geq 0$, $\tau = \min\{k \geq 1 : S_k \geq x\}$ и $\tau = \infty$, если $S_k \leq x$ для всех $k \geq 1$.

Заметим, что $\{\tau = k\} = \{S_1 < x, \dots, S_{k-1} < x, S_k \geq x\}$ и $\{\tau \leq n\} = \{\bar{S}_n \geq x\}$. Так как $S_k \geq x$ на множестве $\{\tau = k\}$, для всех $k \leq n$ получаем:

$$\begin{aligned} E(g(S_n) \mid \tau = k) &= E(g(S_n - S_k + S_k) \mid \tau = k) \geq \\ &\geq E(g(S_n - S_k + x) \mid \tau = k) = E g(S_n - S_k + x). \end{aligned}$$

Последнее равенство имеет место в силу независимости случайных величин $S_n - S_k$ и $I(\tau = k)$. □

Неравенства для распределения максимума сумм независимых случайных величин

Следующее неравенство есть аналог обобщённого неравенства Чебышёва для максимума сумм независимых случайных величин.

Следствие 2.11

Пусть $EX_j = 0$, $j \leq n$, и пусть g — неотрицательная, неубывающая и выпуклая функция. Тогда для всех x таких, что $g(x) > 0$

$$P(\bar{S}_n \geq x) \leq \frac{Eg(S_n)}{g(x)}. \quad (2.28)$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

В силу выпуклости g из неравенства Йенсена (1.6) получаем:

$$Eg(S_n - S_k + x) \geq g(E(S_n - S_k + x)) = g(x).$$

Теперь из (2.27) следует неравенство (2.28). □

Неравенства для распределения максимума сумм независимых случайных величин

Следствие 2.12 (неравенство Колмогорова)

Пусть $EX_j = 0$, $DX_j < \infty$, $j \leq n$. Тогда для всех $x > 0$

$$P\left(\max_{1 \leq k \leq n} |S_k| \geq x\right) \leq \frac{DS_n}{x^2}. \quad (2.29)$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Рассмотрим функцию

$$g(x) = \begin{cases} x^2, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Эта функция удовлетворяет условиям следствия 2.11. Заметим, что $g(x) + g(-x) = x^2$ для всех $x \in \mathbb{R}$.

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Применим неравенство (2.28) к случайным величинам X_j и $-X_j$:

$$\begin{aligned} P\left(\max_{1 \leq k \leq n} |S_k| \geq x\right) &= P\left(\left\{\max_{1 \leq k \leq n} S_k \geq x\right\} \cup \left\{\max_{1 \leq k \leq n} (-S_k) \geq x\right\}\right) \leq \\ &\leq P\left(\max_{1 \leq k \leq n} S_k \geq x\right) + P\left(\max_{1 \leq k \leq n} (-S_k) \geq x\right) \leq \\ &\leq \frac{Eg(S_n)}{x^2} + \frac{Eg(-S_n)}{x^2} = \frac{E(g(S_n) + g(-S_n))}{x^2} = \\ &= \frac{ES_n^2}{x^2} = \frac{DS_n}{x^2}. \end{aligned}$$



Неравенства для распределения максимума сумм независимых случайных величин

Обобщением неравенства Колмогорова (2.29) является следующее неравенство Хайека — Реньи.

Следствие 2.13 (неравенство Хайека — Реньи)

Пусть $EX_j = 0$, $EX_j^2 < \infty$, $j \leq n$. Пусть также $0 < c_n \leq c_{n-1} \leq \dots \leq c_1$. Тогда для всех $x > 0$ и $m < n$

$$P\left(\max_{m \leq k \leq n} c_k |S_k| \geq x\right) \leq \frac{1}{x^2} \left(c_m^2 \sum_{k=1}^m EX_k^2 + \sum_{k=m+1}^n c_k^2 EX_k^2 \right). \quad (2.30)$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Положим в лемме 2.8

$$X = \sum_{j=m}^{n-1} (c_j^2 - c_{j+1}^2) S_j^2 + c_n^2 S_n^2, \quad \tau = \min\{j \geq m : c_j |S_j| \geq x\}.$$

Нетрудно проверить, что

$$EX = c_m^2 \sum_{k=1}^m EX_k^2 + \sum_{k=m+1}^n c_k^2 EX_k^2, \quad \{\tau \leq n\} = \left\{ \max_{m \leq k \leq n} c_k |S_k| \geq x \right\}.$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Так как $S_k^2 \geq x^2/c_k^2$ на множестве $\{\tau = k\}$, то $E(S_k^2 | \tau = k) \geq x^2/c_k^2$.
Если $j > k$, то X_j и $S_k I(\tau = k)$ независимы и, следовательно,

$$E(X_j S_k | \tau = k) = EX_j E(S_k | \tau = k) = 0.$$

Таким образом, для всех $j > k$ имеем

$$\begin{aligned} E(S_j^2 | \tau = k) &= E((S_j - S_k)^2 | \tau = k) + E(S_k^2 | \tau = k) \geq \\ &\geq E(S_k^2 | \tau = k) \geq \frac{x^2}{c_k^2}. \end{aligned}$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Далее, для $k \geq m$

$$\begin{aligned} E(X \mid \tau = k) &= \sum_{j=m}^{n-1} (c_j^2 - c_{j+1}^2) E(S_j^2 \mid \tau = k) + c_n^2 E(S_n^2 \mid \tau = k) \geq \\ &\geq \sum_{j=k}^{n-1} (c_j^2 - c_{j+1}^2) E(S_j^2 \mid \tau = k) + c_n^2 E(S_n^2 \mid \tau = k) \geq \\ &\geq \frac{x^2}{c_k^2} \left(\sum_{j=k}^{n-1} (c_j^2 - c_{j+1}^2) + c_n^2 \right) = x^2. \end{aligned}$$

Отсюда и из леммы 2.8 вытекает неравенство (2.30). □

Неравенства для распределения максимума сумм независимых случайных величин

В заключение этого параграфа приведём ещё одно неравенство Колмогорова — оценку снизу для распределения максимума сумм ограниченных случайных величин.

Теорема 2.11

Пусть $EX_j = 0$ и пусть $P(|X_j| \leq c) = 1$ для некоторого $c > 0$, $j \leq n$. Тогда для всех $x > 0$

$$P\left(\max_{1 \leq k \leq n} |S_k| \geq x\right) \geq 1 - \frac{(x+c)^2}{B_n^2}, \quad (2.31)$$

где $B_n^2 = DS_n = \sum_{j=1}^n DX_j$.

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Положим $\tau = \min\{k \geq 1 : |S_k| \geq x\}$, $A = \{\max_{1 \leq k \leq n} |S_k| \geq x\}$. Тогда $A = \{\tau \leq n\}$ и

$$E(S_n^2; \bar{A}) \leq x^2 E(1; \bar{A}) = x^2 P(\bar{A}) = x^2 - x^2 P(A),$$

$$E(S_n^2; A) = ES_n^2 - E(S_n^2; \bar{A}) \geq B_n^2 - x^2 + x^2 P(A).$$

Далее,

$$\begin{aligned} E(S_n^2 \mid \tau = k) &= E((S_n - S_k + S_k)^2 \mid \tau = k) = \\ &= E((S_n - S_k)^2 \mid \tau = k) + E(S_k^2 \mid \tau = k) + \\ &+ 2E(S_k(S_n - S_k) \mid \tau = k). \end{aligned}$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Так как на множестве $\{\tau = k\}$

$$|S_k| = |S_{k-1} + X_k| \leq |S_{k-1}| + |X_k| \leq |S_{k-1}| + c \leq x + c,$$

то $E(S_k^2 | \tau = k) \leq (x + c)^2$. Так как случайные величины $S_k I(\tau = k)$ и $S_n - S_k$ независимы, то $E(S_k(S_n - S_k) | \tau = k) = 0$ и

$$E((S_n - S_k)^2 | \tau = k) = E(S_n - S_k)^2 = D(S_n - S_k) = \sum_{j=k+1}^n DX_j \leq B_n^2,$$

и, следовательно, для всех $k \leq n$

$$E(S_n^2 | \tau = k) \leq (x + c)^2 + B_n^2.$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Таким образом,

$$\begin{aligned} E(S_n^2; A) &= E(S_n^2; \tau \leq n) = \sum_{k=1}^n E(S_n^2; \tau = k) = \\ &= \sum_{k=1}^n P(\tau = k) E(S_n^2 \mid \tau = k) \leq \\ &\leq ((x + c)^2 + B_n^2) P(A). \end{aligned}$$

Неравенства для распределения максимума сумм независимых случайных величин

Доказательство

Отсюда получаем:

$$((x + c)^2 + B_n^2)P(A) \geq E(S_n^2; A) \geq B_n^2 - x^2 + x^2P(A)$$

и

$$P(A) \geq \frac{B_n^2 - x^2}{(x + c)^2 + B_n^2 - x^2} = 1 - \frac{(x + c)^2}{B_n^2 + (x + c)^2 - x^2} \geq 1 - \frac{(x + c)^2}{B_n^2}.$$



Неравенства для распределения максимума сумм независимых случайных величин

Замечание

В неравенствах Колмогорова, приведённых выше, мы получили оценки для распределений максимума частичных сумм независимых случайных величин через дисперсии указанных сумм. При изучении предельного поведения частичных сумм, в законах больших чисел в частности, оказываются полезными обратные неравенства — оценки дисперсии сумм независимых случайных величин через распределения их максимумов.

Неравенства для распределения максимума сумм независимых случайных величин

Замечание

Одно из таких неравенств можно получить из неравенства (2.31). Если в доказательстве теоремы 2.11 опустить последнее неравенство, то мы приходим к следующей оценке:

$$\begin{aligned} P\left(\max_{1 \leq k \leq n} |S_k| < x\right) \sum_{j=1}^n DX_j &\leq x^2 P\left(\max_{1 \leq k \leq n} |S_k| < x\right) + \\ &+ (x+c)^2 P\left(\max_{1 \leq k \leq n} |S_k| \geq x\right). \end{aligned}$$

Неравенства для распределения максимума сумм независимых случайных величин

Следствие 2.14

Пусть $EX_j = 0$ и пусть $P(|X_j| \leq c) = 1$ для некоторого $c > 0$, $j \leq n$.
Тогда для всех $x > 0$

$$\sum_{j=1}^n DX_j \leq x^2 + (x + c)^2 \frac{P(\max_{1 \leq k \leq n} |S_k| \geq x)}{1 - P(\max_{1 \leq k \leq n} |S_k| \geq x)}. \quad (2.32)$$

В частности, если $P(\max_{1 \leq k \leq n} |S_k| \geq x) \leq \delta$ для некоторого $0 < \delta < 1$, то

$$\sum_{j=1}^n DX_j \leq x^2 + (x + c)^2 \frac{\delta}{1 - \delta}.$$