Записки научных
семинаров ПОМИ
Том 544, 2025 г.

А. Ю. Зайцев

ОЦЕНКИ ТРАНСПОРТНОГО РАССТОЯНИЯ В ЦЕНТРАЛЬНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМЕ

§1. Введение

Суммы независимых слагаемых впервые появились в теории вероятностей при рассмотрении биномиальных распределений в рамках схемы Бернулли. Были получены закон больших чисел и центральная предельная теорема Муавра–Лапласа. Было замечено, что биномиальные распределения не только хорошо аппроксимируются нормальным законом, но и убывание "хвостов" биномиальных распределений происходит похоже на убывание хвостов нормальных распределений. В качестве естественного расширения класса биномиальных распределений можно рассматривать класс распределений сумм независимых (вообще говоря, неодинаково распределенных) случайных величин, ограниченных по модулю одной и той же константой. Оцениванию хвостов таких распределений посвящено значительное количество работ, см., например, [2, 3, 4, 15, 19, 21], В настоящей работе речь пойдет не только об убывании хвостов, но и об оценивании точности аппроксимации в одномерной центральной предельной теореме.

Пусть $X_1, \ldots, X_n$ – ограниченные с вероятностью единица d-мерные независимые случайные векторы. Для простоты будем считать, что у них нулевые средние значения:

$$\mathbf{P}\{\|X_j\| \leqslant \tau\} = 1, \quad \mathbf{E}\, X_j = 0, \quad j = 1, \ldots, n. \tag{1}$$

Нас будет интересовать поведение распределения суммы $S = X_1 + \cdots + X_n$ в зависимости от ограничивающей величины $\tau > 0$.

Из неравномерной оценки Бикялиса [7] в одномерной центральной предельной теореме следует, что в одномерном случае

$$W_1(F, \Phi) \leqslant c\tau. \tag{2}$$

с абсолютной постоянной c, где W_1 – транспортное расстояние Канторовича–Рубинштейна–Васерштейна (см. обзорные статьи [9, 10]), $F = \mathcal{L}(S)$ – распределение суммы S, а $\Phi = \Phi_F$ – соответствующее

Ключевые слова: неравенства, суммы независимых случайных векторов, оценки транспортного расстояния, центральная предельная теорема.

нормальное распределение с тем же нулевым средним и той же дисперсией как у распределения F. При доказательстве неравенства (2) следует учесть, что, согласно [33],

$$W_1(F,\Phi) = \int |F(x) - \Phi(x)|\, dx, \tag{3}$$

где $F(\,\cdot\,)$ и $\Phi(\,\cdot\,)$ – соответствующие функции распределения. Кроме того, $\mathbf{E}\,|X_j|^3 \leqslant \tau\,\mathbf{E}\,X_j^2$.

Основной результат этой статьи значительно сильнее и точнее. Утверждается, что найдется такая c, что

$$W(F,\Phi) = \inf_\pi \int \exp\{|x-y|/c\tau\}\, d\pi(x,y) \leqslant c, \tag{4}$$

где инфимум берется по всем двумерным вероятностным распределениям π с маргинальными распределениями F и Φ. Результат обобщен также на распределения с достаточно медленно растущими кумулянтами из класса $\mathcal{A}_1(\tau)$, введенного в работе автора 1986 года [36]. В частных случаях мы получим некоторые результаты работы Э. Рио [23]. Обсуждается вопрос о возможности обобщения результата на многомерный случай.

Следуя работе Рио [23], определим расстояние Васерштейна, ассоциированное с функцией Орлича ψ:

$$W_\psi(G,H) = \inf\Big\{a > 0 : \inf_\pi \int \psi(|x-y|/a)\, d\pi(x,y) \leqslant 1\Big\}, \tag{5}$$

где второй инфимум берется по всем двумерным вероятностным распределениям π с маргинальными распределениями G и H.

Неравенство (4) можно переписать в виде

$$W_\psi(F,\Phi) \leqslant c\tau, \tag{6}$$

с функцией Орлича $\psi(x) = \exp\{|x|\} - 1$. Неравенство (2) тоже может быть записано в форме (6), только для функции Орлича $\psi(x) = |x|$. Неравенство (6) справедливо также для функции Орлича $\psi(x) = |x|^p$, $p \geqslant 1$. В этом случае утверждение легко выводится из (6) и превращается в оценку

$$W_p(F,\Phi) \leqslant c(p)\,\tau, \tag{7}$$

где $W_p(\,\cdot\,,\,\cdot\,)$ – стандартное p-расстояние Васерштейна. Мы учли, что $|x|^p \leqslant c(p)\,\exp\{|x|\}$.

Класс распределений сумм $S = X_1 + \cdots + X_n$, удовлетворяющих условиям (1), можно рассматривать как естественное обобщение класса биномиальных распределений, которые исторически оказались первыми распределениями сумм независимых слагаемых, изучаемыми в теории вероятностей. С. Н. Бернштейн [6] нашел менее ограничительные условия (см. определение (20) при $d = 1$), при выполнении которых хвосты распределений сумм допускают оценки, аналогичные оценкам для хвостов биномиальных распределений. В условиях неравенства Бернштейна у распределений слагаемых конечны экспоненциальные моменты, то есть выполнены условия Крамера, при которых формулируются теоремы о больших уклонениях распределений сумм независимых слагаемых. Как известно, коэффициенты возникающего в формулировках так называемого ряда Крамера–Петрова определяются по кумулянтам распределений сумм, см. [27, лемма 1.4]. Это мотивировало В. А. Статулявичуса [31] на дальнейшее расширение класса распределений, для которых справедливы результаты о больших уклонениях. Он ввел классы распределений, уже не обязательно представимых как распределения сумм большого количества независимых слагаемых, но кумулянты которых ведут себя аналогично кумулянтам таких сумм (см. (14)). В этой статье мы докажем неравенство (6) не только для распределений сумм $S = X_1 + \cdots + X_n$, удовлетворяющих условиям (1), но и для распределений из класса $\mathcal{A}_1(\tau)$, эквивалентного классу одномерных распределений, рассмотренных Статулявичусом.

Пусть $\mathcal{A}_d(\tau)$, $\tau \geqslant 0$, $d \in \mathbf{N}$, – класс d-мерных распределений, введенный в работе автора [36]. Класс $\mathcal{A}_d(\tau)$ (с фиксированным $\tau \geqslant 0$) состоит из d-мерных распределений F, для которых функция

$$\varphi(z) = \varphi(F, z) = \log \int\limits_{\mathbf{R}^d} e^{\langle z, x\rangle} F\{dx\} \quad (\varphi(0) = 0) \tag{8}$$

определена и аналитична при $\|z\|\,\tau < 1$, $z \in \mathbf{C}^d$, и

$$\left|d_u d_v^2\, \varphi(z)\right| \leqslant \|u\|\tau \left\langle \mathbb{D}\, v, v\right\rangle \tag{9}$$

для всех $u, v \in \mathbf{R}^d$ и $\|z\|\,\tau < 1$, где $\mathbb{D} = \operatorname{cov} F$ – ковариационный оператор распределения F, а $d_u\varphi$ – производная функции φ в направлении u.

Введем необходимые обозначения. Ниже символы $c, c_1, c_2, c_3, \ldots$ будут использоваться для абсолютных положительных констант. Заметим, что c может быть разным в разных (или даже в одних и тех же) формулах. Будем писать $A \ll B$, если $A \leqslant cB$. Кроме того, мы будем использовать обозначение $A \asymp B$, если $A \ll B$ и $B \ll A$. Если соответствующая константа зависит, скажем, от r, будем писать $c(r)$, $A \ll_r B$ и $A \asymp_r B$. Через $\widehat{F}(t) = \int e^{itx}\, F\{dx\}$, $t \in \mathbf{R}$, мы обозначаем характеристическую функцию одномерного распределения F.

Основной результат настоящей статьи содержится в следующих теоремах 1 и 2. В них речь идет о близости одномерных распределений.

Теорема 1. *Пусть* $F = \mathcal{L}(\xi) \in \mathcal{A}_1(\tau)$, $\tau > 0$, $\mathbf{E}\,\xi = 0$. *Тогда существует абсолютная постоянная* c_1, *такая что*

$$W(F, \Phi) = \inf_{\pi} \int \exp\{|x-y|/c_1\tau\}\, d\pi(x,y) \leqslant c_1, \tag{10}$$

где $\Phi = \Phi_F$ *– соответствующее нормальное распределение, а инфимум берется по всем двумерным вероятностными распределениям* π *с маргинальными распределениями* F *и* Φ.

Теорема 2. *Пусть* $F = \mathcal{L}(\xi) \in \mathcal{A}_1(\tau)$, $\tau > 0$, $\mathbf{E}\,\xi = 0$. *Тогда существует абсолютная постоянная* c_2, *такая что*

$$W_\psi(F, \Phi) \leqslant c_2\tau, \tag{11}$$

с функцией Орлича $\psi(x) = \exp\{|x|\} - 1$, *где* $\Phi = \Phi_F$.

Теоремы 1 и 2 эквивалентны. Если $c_1 \leqslant 2$, то из теоремы 1 следует, что

$$\inf_{\pi} \int \exp\{|x-y|/c_1\tau\}\, d\pi(x,y) \leqslant 2. \tag{12}$$

Если же $c_1 > 2$, то мы можем выбрать c_3 таким образом, чтобы $c_1^{c_3} = 2$, причем $c_3 < 1$. Тогда по неравенству Ляпунова для моментов

$$\begin{aligned}\inf_{\pi}\int\exp\{c_3|x-y|/c_1\tau\}d\pi(x,y) &\leqslant \inf_{\pi}\Big(\int\exp\{|x-y|/c_1\tau\}\, d\pi(x,y)\Big)^{c_3}\\ &\leqslant c_1^{c_3} = 2.\end{aligned} \tag{13}$$

Теперь из (12), (13) следует утверждение теоремы 2. Очевидно также, что из теоремы 2 вытекает утверждение теоремы 1.

§2. Свойства классов $\mathcal{A}_d(\tau)$

Рассмотрим элементарные свойства классов $\mathcal{A}_d(\tau)$ (см. [36, 38, 40, 41, 42]). Легко видеть, что если $\tau_1 < \tau_2$, то $\mathcal{A}_d(\tau_1) \subset \mathcal{A}_d(\tau_2)$. Кроме того, класс $\mathcal{A}_d(\tau)$ замкнут относительно операции свертки: если распределения F_1, F_2 принадлежат $\mathcal{A}_d(\tau)$, то $F_1F_2 = F_1 * F_2 \in \mathcal{A}_d(\tau)$.

Здесь и далее произведения и степени мер понимаются в смысле свертки.

Пусть $\tau \geqslant 0$, $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $y \in \mathbf{R}^m$, и $\mathbb{A} : \mathbf{R}^d \to \mathbf{R}^m$ – некоторый линейный оператор. Тогда

$$\mathcal{L}(\mathbb{A}\,\xi + y) \in \mathcal{A}_m\left(\|\mathbb{A}\|\,\tau\right), \quad \text{где} \quad \|\mathbb{A}\| = \sup_{x\in\mathbf{R}^d,\,\|x\|\leqslant 1} \|\mathbb{A}\,x\|\,.$$

В частности, для любого $a \in \mathbf{R}$

$$\mathcal{L}(a\,\xi) \in \mathcal{A}_d\left(|a|\,\tau\right).$$

Классы $\mathcal{A}_d(\tau)$ тесно связаны с другими естественно определяемыми классами многомерных распределений. Из определения $\mathcal{A}_d(\tau)$ следует, что если $\mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, то вектор ξ имеет конечные экспоненциальные моменты $\mathbf{E}\,e^{\langle h,\xi\rangle} < \infty$, при $h \in \mathbf{R}^d$, $\|h\|\,\tau < 1$. Это приводит к экспоненциальному убыванию хвостов распределений.

Условие $\mathcal{L}(\xi) \in \mathcal{A}_1(\tau)$ эквивалентно условию В. А. Статулявичуса [31], см. также [13, 26, 27], на скорость роста кумулянтов γ_m случайной величины ξ:

$$|\gamma_m| \leqslant \frac{1}{2}\,m!\,\tau^{m-2}\gamma_2, \qquad m = 3, 4, \ldots. \tag{14}$$

Эта эквивалентность означает, что если одно из этих условий выполняется с параметром τ, то второе справедливо с параметром $c\tau$, где c – некоторая положительная абсолютная константа. Заметим, впрочем, что условие $\mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$ существенно отличается от других многомерных аналогов условия Статулявичуса, рассмотренных Р. Рудзкисом [24] и Л. Саулисом [25]. В обзорной статье [27] и в монографии [26] содержится большое количество примеров распределений, удовлетворяющих условиям (14), и не являющихся распределениями сумм большого числа независимых слагаемых. Заметим также, что Статулявичус рассматривал более общие условия, при выполнении которых экспоненциальные моменты не обязательно конечны.

Другой класс распределений, обозначаемый $\widetilde{\mathcal{A}}_d(\tau)$, $\tau \geqslant 0$, упоминался в работе [42]. Он определяется аналогично $\mathcal{A}_d(\tau)$ с заменой (9) на

$$\left|\,d_v^2\,\varphi(z)\right| \leqslant 2\left\langle \mathbb{D}\,v, v\right\rangle \tag{15}$$

для всех $v \in \mathbf{R}^d$ и $\|z\|\,\tau < 1$. То, что классы $\widetilde{\mathcal{A}}_d(\tau)$ и $\mathcal{A}_d(\tau)$ также эквивалентны, легко проверяется с использованием неравенств Коши. Определение классов $\widetilde{\mathcal{A}}_d(\tau)$ в некотором смысле выглядит даже более естественно, чем определение классов $\mathcal{A}_d(\tau)$. Константу 2 в определении (15) можно заменить на любую другую константу C, $1 < C < \infty$, отделенную от 1 и от бесконечности. В результате также получатся эквивалентные классы.

Очевидно, что класс $\mathcal{A}_d\,(0)$ совпадает с классом всех d-мерных гауссовских распределений. Следующее неравенство (16) было доказано в работе автора [36] и может рассматриваться как оценка устойчивости этой характеризации:

$$\text{если } F \in \mathcal{A}_d(\tau), \text{ то } \pi\,(F,\,\Phi_F) \leqslant c\,d^2\tau\,\log^*(\tau^{-1}), \tag{16}$$

где $\pi(\,\cdot\,,\,\cdot\,)$ – расстояние Прохорова, определенное в работе [22], а Φ_F обозначает гауссовское распределение, среднее и ковариационный оператор которого такие же, как у F. Здесь $\log^* b = \max\{1, \log b\}$ при $b > 0$, а log используется для обозначения натурального логарифма. Отметим, что теоремы 1 и 2 этой статьи также могут рассматриваться как оценки устойчивости упомянутой выше характеризации гауссовских распределений в транспортнымх метриках и пока в одномерном случае.

Расстояние Прохорова между распределениями F, G может быть определено по формуле

$$\pi(F, G) = \inf\{\lambda : \pi(F, G, \lambda) \leqslant \lambda\},$$

где

$$\pi(F, G, \lambda) = \sup_Y \max\left\{F\{Y\} - G\{Y^\lambda\}, G\{Y\} - F\{Y^\lambda\}\right\}, \quad \lambda > 0,$$

а $Y^\lambda = \{y \in \mathbf{R}^d : \inf\limits_{x\in Y}\|x - y\| < \lambda\}$ – λ-окрестность борелевского множества Y (см. [11, 12]).

В работе автора [36] было также установлено, что

$$\text{если } F \in \mathcal{A}_d(\tau), \text{ то } \pi(F, \Phi_F, \lambda) \leqslant c\,d^2 \exp\Big\{-\frac{\lambda}{c\,d^2\tau}\Big\}, \quad \lambda > 0. \tag{17}$$

Здесь существенно то, что неравенство (17) доказано для всех $\tau > 0$ и для произвольного ковариационного оператора $\operatorname{cov} F$.

По теореме Штрассена–Дадли (см. работу Р. Дадли [14]) и согласно неравенству (17), для любого распределения $F \in \mathcal{A}_d(\tau)$ и любого $\lambda > 0$ можно построить на одном вероятностном пространстве случайные векторы ξ и η с $\mathcal{L}(\xi) = F$ и $\mathcal{L}(\eta) = \Phi_F$, так что

$$\mathbf{P}\{\|\xi - \eta\| > \lambda\} = \pi(F, \Phi_F, \lambda) \leqslant c\,d^2 \exp\Big\{-\frac{\lambda}{c\,d^2\tau}\Big\}. \tag{18}$$

Подчеркнем, что теорема Штрассена–Дадли гарантирует существование построения с равенством в (18) только при фиксированном λ. Пример, показывающий невозможность построения с равенством в (18) при всех λ одновременно, можно найти в обзоре [11]. Теорема Штрассена–Дадли дает возможность из оценок для $\pi(F, G, \lambda)$ автоматически получать утверждения типа неравенства (18). Первоначальное доказательство Штрассена [32] было неконструктивным. Дадли [14] дал сложное конструктивное доказательство, основанное на комбинаторных идеях. Наконец, Г. Шай [30] нашел короткое доказательство, основанное на теореме о двойственности.

Если бы равенство (18) было доказано при всех $\lambda > 0$ одновременно на одном и том же вероятностном пространстве, то из него автоматически вытекало бы утверждение теоремы 1, причем для любой размерности d, $1 \leqslant d < \infty$. Поэтому неравенство (17) дает основание рассчитывать на возможность обобщить утверждения теорем 1 и 2 на многомерный случай.

Если F – безгранично делимое распределение со спектральной мерой, сосредоточенной на шаре $\left\{x \in \mathbf{R}^d : \|x\| \leqslant \tau\right\}$, то $F \in \mathcal{A}_d(c\tau)$, где c – некоторая положительная абсолютная константа. В работе [36] можно найти и менее ограничительные условия на спектральную меру, обеспечивающие принадлежность безгранично делимого распределения классу $\mathcal{A}_d(c\tau)$.

В частности, для распределения Пуассона Π_λ с параметром $\lambda > 0$ справедливо включение $\Pi_\lambda \in \mathcal{A}_1(c)$. Из теоремы 2 следует, что

$$\sup_\lambda W_\psi(\Pi_\lambda, \Phi_{\Pi_\lambda}) \leqslant c, \quad \text{при } \psi(x) = \exp\{|x|\} - 1. \tag{19}$$

Это – утверждение следствия 2.2 работы Э. Рио [23]. Но теорема 2 содержит и более общее утверждение. В (19) мы можем заменить множество распределений Пуассона на множество всех безгранично делимых распределений со спектральными мерами Леви–Хинчина, сосредоточенными на отрезке $[-1, 1]$.

Распределения из классов $\mathcal{A}_d(\tau)$ непосредственно используются в формулировках результатов автора [40, 41, 42] об оценивании точности сильной гауссовской аппроксимации для сумм независимых случайных векторов в наиболее важном случае, когда слагаемые имеют конечные экспоненциальные моменты (см. также обзорную статью [43]). Получены многомерные аналоги одномерных результатов А. И. Саханенко [28], обобщившего и существенно уточнившего результаты Я. Комлоша, П. Майора и Г. Тушнади [18], на случай неодинаково распределенных случайных величин. Он рассмотрел следующие классы одномерных распределений:

$$\mathcal{S}_1(\tau) = \left\{\mathcal{L}(\xi):\ \mathbf{E}\,\xi = 0,\ \mathbf{E}\,|\xi|^3 \exp\left\{|\xi|/\tau\right\} \leqslant \tau\,\mathbf{E}\,|\xi|^2\right\}, \quad \tau > 0.$$

В препринте автора [35] было замечено, что классы $\mathcal{S}_1(\tau)$ эквивалентны классам распределений $\mathcal{B}_1(\tau)$, удовлетворяющих условиям неравенства Бернштейна (см. определение (20)), в том смысле, что если одно из условий принадлежности к классу выполняется с параметром τ, то второе справедливо с параметром $c\tau$, где c – некоторая положительная абсолютная константа. Результаты А. И. Саханенко [28] были сформулированы в виде оценок экспоненциальных моментов максимального отклонения построенных на одном вероятностном пространстве сумм независимых случайных величин с распределениями из $\mathcal{S}_1(\tau)$ от соответствующих сумм независимых нормально распределенных слагаемых. Внешний вид оценок Саханенко при этом почти такой же как у неравенства (10), так что из них можно вывести аналог неравенства (10) с заменой правой части на $c(1 + \sigma/\tau)$ для распределений сумм независимых случайных величин с распределениями из $\mathcal{S}_1(\tau)$. Здесь через σ^2 обозначена дисперсия рассматриваемой суммы. Для сверток распределений из $\mathcal{S}_1(\tau)$ некоторые оценки моментов экспоненциального типа в центральной предельной теореме содержатся в работе А. И. Саханенко [29].

В работах автора [35] и [37] неравенства (16) и (17) (с заменой d^2 на $d^{5/2}$) были доказаны для сверток распределений из класса $\mathcal{B}_d(\tau)$, где $\tau > 0$ и

$$\begin{aligned}\mathcal{B}_d(\tau) = \Big\{F = \mathcal{L}(\xi) : \mathbf{E}\xi = 0,\ &\Big|\mathbf{E}\,\langle\xi, v\rangle^2\,\langle\xi, u\rangle^{m-2}\Big| \\ &\leqslant \frac{1}{2} m!\,\tau^{m-2}\,\|u\|^{m-2}\ \mathbf{E}\,\langle\xi, v\rangle^2 \ \text{ при всех } u, v \in \mathbf{R}^d,\ m = 3, 4, \dots\Big\},\end{aligned} \tag{20}$$

удовлетворяющих многомерным аналогам условий неравенства Бернштейна. Условие Саханенко $\mathcal{L}(\xi) \in \mathcal{S}_1(\tau)$ эквивалентно условию $\mathcal{L}(\xi) \in \mathcal{B}_1(\tau)$. Заметим, что если $F\left\{\left\{x \in \mathbf{R}^d : \|x\| \leqslant \tau\right\}\right\} = 1$, $\mathbf{E}\,\xi = 0$, то $F \in \mathcal{B}_d(\tau)$.

Сформулируем соотношения между классами $\mathcal{A}_d(\tau)$ и $\mathcal{B}_d(\tau)$. Обозначим через σ_F^2 максимальное собственное число ковариационного оператора распределения F. Тогда
a) Если $F=\mathcal{L}(\xi)\in\mathcal{B}_d(\tau)$, то $\sigma_F^2\leqslant 12\,\tau^2$, $\mathbf{E}\,\xi=0$ и $F\in\mathcal{A}_d(c\tau)$.
b) Если $F=\mathcal{L}(\xi)\in\mathcal{A}_d(\tau)$, $\sigma_F^2\leqslant\tau^2$ и $\mathbf{E}\,\xi=0$, то $F\in\mathcal{B}_d(c\tau)$.

В частности, распределение суммы $S=X_1+\cdots+X_n$ при выполнении условий (1) принадлежат классу $\mathcal{A}_d(c\tau)$.

Таким образом, грубо говоря, $\mathcal{B}_d(\tau)$ образует подкласс распределений $F=\mathcal{L}(\xi-\mathbf{E}\,\xi)$, $\mathcal{L}(\xi)\in\mathcal{A}_d(c\tau)$, для которых $\sigma_F^2\leqslant 12\,\tau^2$. Неравенства (16) и (17) в этом случае говорят лишь о том, что оба сравниваемых распределения близки к вырожденному закону E, сосредоточенному в начале координат. Если же $F=\mathcal{L}(\xi)\in\mathcal{A}_d(\tau)$ и σ_F^2 существенно больше, чем τ^2, то $\mathcal{L}(\xi/\sigma_F)\in\mathcal{A}_d(\tau/\sigma_F)$ и неравенства (16) и (17) отражают близость распределения F к соответствующему гауссовскому закону.

Пусть $\tau\geqslant 0$, $F=\mathcal{L}(\xi)\in\mathcal{A}_d(\tau)$, $\|h\|\tau<1$, $h\in\mathbf{R}^d$. Определим распределение $\overline{F}=\overline{F}(h)$ соотношением

$$\overline{F}\{dx\}=\Big(\mathbf{E}\,e^{\langle h,\xi\rangle}\Big)^{-1}e^{\langle h,x\rangle}F\{dx\}.$$

Будем обозначать $\overline{\xi}=\overline{\xi}(h)$ – случайный вектор с распределением $\mathcal{L}(\overline{\xi}(h))=\overline{F}(h)$. Распределения $\overline{F}(h)$ иногда называют преобразованиями Крамера (или Эсшера, см. [8]). При доказательстве результатов работ [36, 37, 40, 41, 42] распределения $\overline{F}(h)$ используются при оценивании вероятностей больших уклонений, соответствуюших условным плотностям. Еще одним важным свойством классов $\mathcal{A}_d(\tau)$ является то, что $\overline{F}(h)\in\mathcal{A}_d(2\tau)$ при $\|h\|\,\tau\leqslant 1/2$, см. пункт b) леммы 1. Это дает возможность систематически применять результаты, полученные для исходных распределений, к их преобразованиям Крамера и за счет этого уточнять оценки.

А. Н. Колмогоров [16] поставил задачу оценивания точности безгранично делимой аппроксимация распределений сумм независимых случайных величин, распределения которых сосредоточены на коротких интервалах длины $\tau\leqslant 1/2$ с точностью до малой вероятности p. В частном случае, когда $p=0$, речь идет об аппроксимации распределений сумм $S=X_1+\cdots+X_n$ при $d=1$ и при выполнении условий (1). В этом случае Колмогоров [16, 17] получил оценку

$$L(F,\Phi_F)\ll\tau^{1/2}\log^{1/4}(1/\tau),\tag{21}$$

где $L(\,\cdot\,,\,\cdot\,)$ – расстояние Леви. Из сказанного выше вытекает, что $F \in \mathcal{A}_1(c\tau)$ и неравенства (16) и (17) можно рассматривать как усиления, обобщения и уточнения неравенства (21). Заметим, что формулировки результатов в работах [16, 17] отличаются от неравенства (21). Чтобы вывести из них это неравенство, требуется сравнительно элементарный дополнительный анализ.

Пусть $X, X_1, \dots, X_n$ – независимые одинаково распределенные случайные величины, такие что

$$\mathbf{E}\, X = 0 \quad \text{и} \quad \mathbf{E}\, \exp\{t|X|\} < \infty \quad \text{при некотором} \quad t > 0. \tag{22}$$

Тогда несложно убедиться в том, что пайдется такое $c(F)$, для которого $F = \mathcal{L}(X) \in \mathcal{A}_1(c(F))$, а распределение нормированной суммы

$$F_n = \mathcal{L}\bigl((X_1 + \dots + X_n)/\sqrt{n}\bigr) \in \mathcal{A}_1(c(F)/\sqrt{n}).$$

Применяя теорему 2, получаем, что

$$W_\psi(F_n, \Phi_{F_n}) \ll_F 1/\sqrt{n}, \tag{23}$$

с функцией Орлича $\psi(x) = \exp\{|x|\} - 1$. Это – утверждение теоремы 2.1 работы Э. Рио [23]. Еще раз подчеркнем, что в основных результатах настоящей статьи мы рассматриваем одномерные распределения, удовлетворяющие условию (9) на преобразования Лапласа и, вообще говоря, не представимые в виде сверток большого числа одинаковых распределений. При этом условие (9) превращается в

$$|\,\varphi'''(z)| \leqslant \tau\,\sigma^2 \quad \text{при } |z|\tau \leqslant 1. \tag{24}$$

Из сказанного выше следует, что утверждения теорем 1 и 2 справедливы для сверток одномерных распределений, сосредоточенных на отрезке $[-\tau, \tau]$ или удовлетворяющих условиям неравенства Бернштейна, а также для безгранично делимых распределений со спектральными мерами Леви–Хинчина, сосредоточенными на этом же отрезке, и для распределений, удовлетворяющих условиям Статулявичуса (14). По содержанию и по методам доказательства их можно рассматривать как просто и понятно формулируемые утверждения из теории больших уклонений.

§3. Доказательство теоремы 1

Нам понадобится следующая лемма о свойствах преобразования Крамера (она содержится в работе [36, леммы 2.1, 3.1]).

Лемма 1. *Пусть* $\tau \geqslant 0$, $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $\mathbf{E}\,\xi = 0$, $h \in \mathbf{R}^d$, $\|h\|\tau < 1$, $\mathbb{D} = \operatorname{cov} F$, $\mathbb{D}(h) = \operatorname{cov} \overline{F}(h)$. *Обозначим через* σ^2 *минимальное собственное значение оператора* $\mathbb{D}$. *Тогда*

a) *для любого* $u \in \mathbf{R}^d$ *справедливы следующие соотношения*:

$$\langle \mathbb{D}(h)\, u,\, u\rangle = \langle \mathbb{D}\, u,\, u\rangle \big(1 + \theta \|h\|\tau\big), \tag{25}$$

$$\log \mathbf{E}\, e^{i\langle h,\xi\rangle} = -\frac{1}{2}\langle \mathbb{D}\, h,\, h\rangle \Big(1 + \frac{1}{3}\,\theta \|h\|\tau\Big), \tag{26}$$

$$\log \mathbf{E}\, e^{\langle h,\xi\rangle} = \frac{1}{2}\,\langle \mathbb{D}\, h,\, h\rangle \Big(1 + \frac{1}{3}\,\theta \|h\|\tau\Big) \tag{27}$$

(*здесь и далее* θ *символизирует различные величины, не превышающие по модулю единицы*: $|\theta| \leqslant 1$);

b) *Если* $\|h\|\tau \leqslant 1/2$, *то* $\overline{F}(h) \in \mathcal{A}_d(2\tau)$;

c) *При* $x \in \Pi = \big\{x \in \mathbf{R}^d : 4.8\,\tau\sigma^{-1} \big\|\mathbb{D}^{-1/2}x\big\| \leqslant 1\big\}$ *существует такой параметр* $h = h(x) \in \mathbf{R}^d$, *что*

$$\mathbf{E}\,\overline{\xi}(h) = x, \tag{28}$$

$$\|h\|\tau \leqslant 1/2, \tag{29}$$

$$\sigma\,\|h\| \leqslant \big\|\mathbb{D}^{1/2}h\big\| \leqslant 2.4\,\big\|\mathbb{D}^{-1/2}x\big\|, \tag{30}$$

$$\big\|\mathbb{D}^{1/2}h - \mathbb{D}^{-1/2}x\big\| \leqslant 2.88\,\theta\tau\sigma^{-1}\,\big\|\mathbb{D}^{-1/2}x\big\|^2, \tag{31}$$

$$\mathbf{E}\exp\big\{\langle h,\xi\rangle - \langle h,x\rangle\big\} = \exp\Big\{-\frac{1}{2}\,\big\|\mathbb{D}^{-1/2}x\big\|^2 + 10.08\,\theta\tau\sigma^{-1}\big\|\mathbb{D}^{-1/2}x\big\|^3\Big\}. \tag{32}$$

В дальнейшем

$$\Xi(x) = e^{x^2/2}\int_x^\infty e^{-y^2/2}\,dy, \quad x > 0, \tag{33}$$

– отношение Миллса. Нам потребуется следующая лемма (см. [1, лемма 1.2 главы VI]).

Лемма 2. *Пусть* $x, \varepsilon > 0$. *Тогда*

$$0 \leqslant \Xi(x) - \Xi(x+\varepsilon) \leqslant \frac{\varepsilon}{x^2}, \tag{34}$$

$$\Xi(x) = \frac{1}{x}\Big(1 - \frac{|\theta|}{x^2}\Big). \tag{35}$$

Пусть Φ_σ – одномерное гауссовское распределение с нулевым средним и дисперсией σ^2. Тогда

$$1-\Phi_\sigma(x+\varepsilon)\leqslant(1-\Phi_\sigma(x))\,\exp\Big\{-\frac{2x\varepsilon+\varepsilon^2}{2\sigma^2}\Big\}. \tag{36}$$

Следовательно, при $x>\varepsilon$

$$1-\Phi_\sigma(x)\leqslant(1-\Phi_\sigma(x-\varepsilon))\,\exp\Big\{-\frac{2x\varepsilon-\varepsilon^2}{2\sigma^2}\Big\}. \tag{37}$$

Пусть $\rho(F,\Phi)=\sup_x\big|F(x)-\Phi(x)\big|$ – расстояние Колмогорова, равномерное расстояние между функциями распределения.

Лемма 3. *В условиях теоремы* 1

$$\rho(F,\Phi)\ll\tau/\sigma, \tag{38}$$

где $\sigma^2=\operatorname{Var}\xi$ – общая дисперсия распределений F и Φ.

Доказательство. С помощью неравенства

$$\big|e^{z_1}-e^{z_2}\big|\leqslant\big|z_1-z_2\big|\,\max\big\{\big|e^{z_1}\big|,\big|e^{z_2}\big|\big\}, \tag{39}$$

справедливого для $z_1,z_2\in\mathbb{C}$, и используя неравенство (26), получаем, что при $|t|\tau\leqslant1$

$$\big|\widehat{F}(t)-\widehat{\Phi}(t)\big|\leqslant\frac{\tau}{6}\sigma^2|t|^3\exp\Big\{-\frac{1}{3}\sigma^2t^2\Big\}. \tag{40}$$

Поэтому, используя стандартное неравенство сглаживания (см. [1, теорема 1.2 главы III]), находим, что при $T=1/\tau$

$$\rho(F,\Phi)\ll\int_0^T\bigg|\frac{\widehat{F}(t)-\widehat{\Phi}(t)}{t}\bigg|\,dt+\frac{1}{\sigma T}\ll\frac{\tau}{\sigma}. \tag{41}$$

□

Следующая лемма содержит аналог неравенства Бернштейна для распределений из класса $\mathcal{A}_1(\tau)$.

Лемма 4. *Пусть в условиях теоремы* 1 $\operatorname{Var}\xi=\sigma^2$. *Тогда*

$$\mathbf{P}\big\{\xi\geqslant x\big\}\leqslant\max\Big\{\exp\Big\{-\frac{x^2}{4\,\sigma^2}\Big\},\ \exp\Big\{-\frac{x}{4\,\tau}\Big\}\Big\},\qquad x\geqslant0. \tag{42}$$

Доказательство этой леммы почти буквально повторяет доказательство неравенства Бернштейна. Пусть $0 \leqslant h\tau \leqslant \frac{1}{2}$. Согласно (27),

$$\mathbf{E}\, e^{h\xi} \leqslant \exp\{h^2\sigma^2\}$$

и

$$\mathbf{P}\{\xi \geqslant x\} \leqslant e^{-hx}\,\mathbf{E}\, e^{h\xi} \leqslant \exp\{h^2\sigma^2 - hx\}.$$

Выберем параметр h в зависимости от x. Если $0 \leqslant x \leqslant \frac{\sigma^2}{\tau}$, возьмем $h = \frac{x}{2\sigma^2}$, и получим оценку

$$\mathbf{P}(S \geqslant x) \leqslant \exp\left\{-\frac{x^2}{4\,\sigma^2}\right\}. \tag{43}$$

А если $x > \frac{\sigma^2}{\tau}$, то возьмем $h = \frac{1}{2\tau}$ и получим

$$\mathbf{P}(S \geqslant x) \leqslant \exp\left\{\frac{\sigma^2}{4\tau^2} - \frac{x}{2\tau}\right\} \leqslant \exp\left\{-\frac{x}{4\,\tau}\right\}. \tag{44}$$

Теперь неравенство (42) следует из (43) и (44).

Доказательство теоремы 1. Не нарушая общности, будем предполагать, что функция распределения F бесконечно дифференцируема и строго возрастает. Для обоснования этого достаточно вместо распределения F рассмотреть свертку этого распределения с гауссовским распределением с нулевым средним и стремящейся к нулю положительной дисперсией, а также использовать стандартное средство доказательства теорем о сильной аппроксимации – лемму А из работы Беркеша и Филиппа [5], см., например, доказательство теоремы 3.1 из работы Э. Рио [23]. Таким образом, при таких предположениях корректно определена строго возрастающая обратная функция $F^{-1}(\,\cdot\,)$.

Пусть случайная величина η имеет распределение $\mathcal{L}(\eta) = \Phi$. Положим $\xi = F^{-1}\bigl(\Phi(\eta)\bigr)$. Ясно, что $\mathcal{L}(\xi) = F$ и $\eta = \Phi^{-1}\bigl(\Phi(\eta)\bigr)$. Это означает, что случайные величины ξ и η определяются как преобразования Смирнова равномерно распределенной на отрезке $[0,1]$ случайной величины $\Phi(\eta)$. Именно так строятся случайные величины с заданными распределениями при доказательстве равенства (3) в работе [33]. Тогда если случайная величина ξ принимает некоторое конкретное значение $x \in \mathbf{R}$, случайная величина η примет значение $\Phi^{-1}\bigl(F(x)\bigr)$.

Дальнейшие рассуждения проводятся в предположении, что $\xi = x$ и $\tau \leqslant c_4\sigma$, где $\sigma^2 = \operatorname{Var}\xi$, причем выбор c_4 будет уточняться в процессе доказательства.

Сначала мы рассмотрим случай, когда $|x| \leqslant 2\sigma$. Пусть

$$\phi(u) = \frac{1}{\sqrt{2\pi}\,\sigma}\,\exp\Bigl(-\frac{u^2}{2\,\sigma^2}\Bigr) \tag{45}$$

– плотность распределения Φ. Напомним, что согласно лемме 3

$$\rho(F,\Phi) \leqslant c_5\,\frac{\tau}{\sigma}. \tag{46}$$

Пусть $|u| \leqslant 2\sigma$, $|y| \leqslant 3\sigma$. Очевидно, что тогда

$$\begin{aligned}\bigl|\Phi(u)-\Phi(y)\bigr| \geqslant |u-y|\,\phi(3\sigma) = \frac{e^{-9/2}\,|u-y|}{\sqrt{2\pi}\,\sigma} = \frac{c_6\,|u-y|}{\sigma},\\ c_6 = \frac{e^{-9/2}}{\sqrt{2\pi}},\end{aligned} \tag{47}$$

$$\bigl|F(u)-F(y)\bigr| \geqslant \frac{c_6\,|u-y|}{\sigma} - 2c_5\,\frac{\tau}{\sigma}. \tag{48}$$

Лемма 5. *Существуют такие абсолютные положительные постоянные c_4, c_7, что при $\tau \leqslant c_4\sigma$, $|x| \leqslant 2\sigma$ справедливы неравенства*

$$\Phi(x+c_7\tau) \geqslant F(x), \quad F(x+c_7\tau) \geqslant \Phi(x). \tag{49}$$

Действительно, согласно (46)–(48), при $|x| \leqslant 2\sigma$

$$\Phi(x+c_7\tau) - F(x) \geqslant F(x+c_7\tau) - F(x) - c_5\,\frac{\tau}{\sigma} \geqslant \frac{c_6\,c_7\tau}{\sigma} - 3c_5\,\frac{\tau}{\sigma} \geqslant 0, \tag{50}$$

если мы выберем $c_7 = 3c_5/c_6$ и если $x + c_7\tau \leqslant 3\sigma$. Последнее неравенство становится очевидным, если $c_7\tau \leqslant \sigma$. Для этого достаточно выбрать $c_4 \leqslant c_7^{-1}$. Аналогично проверяется и второе неравенство в (49).

Таким образом, согласно лемме 5,

$$|\,\xi-\eta\,| < c_7\tau, \qquad \text{если} \quad |\xi| \leqslant 2\sigma. \tag{51}$$

Пусть $2\sigma \leqslant x \leqslant \sigma^2/5\tau$, а параметр $h = h(x) \in \mathbf{R}$ выбран в соответствии с пунктом c) одномерного варианта леммы 1 (условие $x \in \Pi$ которого выполнено) и таков, что $\mathbf{E}\,\overline{\xi}(h) = x$, $\|h\|\tau \leqslant 1/2$, $\mathcal{L}(\overline{\xi}(h)) = \overline{F} = \overline{F}(h)$,

$$|\sigma h - x/\sigma| \leqslant 2.88\,\tau\sigma^{-1}x^2\sigma^{-2}. \tag{52}$$

Согласно соотношению (25) леммы 1,

$$\sigma^2(h) = \operatorname{Var}\overline{\xi}(h) = \sigma^2\bigl(1+\theta\|h\|\tau\bigr). \tag{53}$$

Введем распределение $H = \Phi_{\overline{F}}$. Тогда, в соответствии с пунктом b) леммы 1, $\overline{F}(h) \in \mathcal{A}_d(2\tau)$ и по лемме 3 с учетом (53)

$$\rho(H, \overline{F}) \ll \frac{\tau}{\sigma(h)} \ll \frac{\tau}{\sigma}. \tag{54}$$

Интегрируя по частям, получаем

$$\begin{aligned} 1 - F(x) &= \mathbf{E}\, e^{h\xi} \int_x^\infty e^{-hy} \overline{F}\{dy\} \\ &= \mathbf{E}\, e^{h\xi} \biggl(\int_x^\infty h\, e^{-hy} \overline{F}(y)\, dy - e^{-hx}\, \overline{F}(x) \biggr), \end{aligned} \tag{55}$$

$$\int_x^\infty e^{-hy} H\{dy\} = \int_x^\infty h\, e^{-hy} H(y)\, dy - e^{-hx}\, H(x). \tag{56}$$

С другой стороны, легко проверяется, что

$$\int_x^\infty e^{-hy} H\{dy\} = \frac{1}{\sqrt{2\pi}} e^{-hx}\, \Xi(h\sigma(h)). \tag{57}$$

Из (54)–(57) следует, что

$$\begin{aligned} &\biggl| \int_x^\infty e^{-hy} \overline{F}\{dy\} - \frac{1}{\sqrt{2\pi}} e^{-hx}\, \Xi(h\sigma(h)) \biggr| \\ &\qquad \leqslant \biggl| \int_x^\infty h\, e^{-hy} (\overline{F}(y) - H(y))\, dy - e^{-hx} (\overline{F}(x) - H(x)) \biggr| \\ &\qquad\qquad \leqslant 2\, e^{-hx}\, \rho(H, \overline{F}) \ll e^{-hx}\, \frac{\tau}{\sigma}. \end{aligned} \tag{58}$$

Применяя неравенство (34) леммы 2, получаем

$$\Bigl| \Xi(h\sigma(h)) - \Xi(h\sigma) \Bigr| \ll \frac{h|\sigma(h) - \sigma|}{h^2\sigma^2} \ll \frac{\tau}{\sigma}. \tag{59}$$

Если $x \geqslant 2\sigma$, то $h\sigma \ll x/\sigma$ и $\Xi(h\sigma) \gg \Xi(x/\sigma) \gg \sigma/x$. Применяя еще раз неравенство (34), а также (52), получаем

$$\bigl| \Xi(h\sigma) - \Xi(x/\sigma) \bigr| \ll \frac{\bigl| h\sigma - x/\sigma \bigr|\, \sigma^2}{x^2} \ll \frac{\tau}{\sigma}. \tag{60}$$

Следовательно,

$$\left|\int\limits_x^\infty e^{-hy}\overline{F}\{dy\} - \frac{1}{\sqrt{2\pi}}e^{-hx}\,\Xi(x/\sigma)\right| \ll \frac{\tau}{\sigma}\,e^{-hx}. \tag{61}$$

Применяя полученные выше неравенства, получаем, что

$$\begin{aligned}1-F(x) &= \mathbf{E}\,e^{h\xi}\Bigl(\frac{1}{\sqrt{2\pi}}e^{-hx}\,\Xi(x/\sigma)+\int\limits_x^\infty e^{-hy}\overline{F}\{dy\}-\frac{1}{\sqrt{2\pi}}e^{-hx}\,\Xi(x/\sigma)\Bigr)\\ &= \frac{1}{\sqrt{2\pi}}\mathbf{E}\,e^{h\xi-hx}\,\Xi(x/\sigma)\Bigl(1+\theta\,c\,\frac{\tau}{\sigma}\,\frac{x}{\sigma}\Bigr)\\ &= (1-\Phi(x))\exp\Bigl\{\theta\,c_8\,\frac{\tau}{\sigma}\,\frac{x^3}{\sigma^3}\Bigr\}\\ &= \frac{1}{\sqrt{2\pi}}\,e^{-x^2/2\sigma^2}\,\Xi(x/\sigma)\exp\Bigl\{\theta\,c_8\,\frac{\tau}{\sigma}\,\frac{x^3}{\sigma^3}\Bigr\}.\end{aligned} \tag{62}$$

Лемма 6. *Существуют такие абсолютные положительные постоянные* $c_9,\dots,c_{11}$, *что*

$$1-\Phi(x+\beta(x)) \leqslant 1-F(x) \leqslant 1-\Phi(x-\beta(x)) \tag{63}$$

при $\tau/\sigma \leqslant c_9$, $2\sigma \leqslant x \leqslant z = c_{10}\sigma^2/\tau$, *где*

$$\beta(x) = c_{11}\tau\,x^2\sigma^{-2}. \tag{64}$$

Мы положим $c_{11} = 4c_8$. Тогда, выбирая достаточно малое c_{10}, мы обеспечим выполнение неравенства $\beta(x) \leqslant x$. Применяя теперь неравенства (36) и (37) при $\varepsilon = \beta(x)$, получаем

$$\begin{aligned}1-F(x) &= (1-\Phi(x))\exp\Bigl\{\theta\,c_8\,\frac{\tau}{\sigma}\,\frac{x^3}{\sigma^3}\Bigr\}\\ &\leqslant (1-\Phi(x-\beta(x)))\exp\Bigl\{-\frac{(2x-\beta(x))\beta(x)}{2\sigma^2}+c_8\,\frac{\tau}{\sigma}\,\frac{x^3}{\sigma^3}\Bigr\}\\ &\leqslant 1-\Phi(x-\beta(x)),\end{aligned} \tag{65}$$

$$\begin{aligned}1-F(x) &= (1-\Phi(x))\exp\Bigl\{\theta\,c_8\,\frac{\tau}{\sigma}\,\frac{x^3}{\sigma^3}\Bigr\}\\ &\geqslant (1-\Phi(x+\beta(x)))\exp\Bigl\{\frac{(2x+\beta(x))\beta(x)}{2\sigma^2}-c_8\,\frac{\tau}{\sigma}\,\frac{x^3}{\sigma^3}\Bigr\}\\ &\geqslant 1-\Phi(x+\beta(x)),\end{aligned} \tag{66}$$

завершая доказательство леммы.

Применяя лемму 6, получаем, что

$$|\,\xi-\eta\,| < c_{11}\tau\xi^2/\sigma^2, \qquad \text{если} \quad 2\sigma \leqslant \xi \leqslant z = c_{10}\sigma^2/\tau, \quad \tau/\sigma \leqslant c_9. \tag{67}$$

Для доказательства теоремы 1 достаточно доказать, что абсолютную постоянную c_{14} можно выбрать настолько большой, что $\mathbf{E}\,\exp\bigl\{|\,\xi-\eta\,|/c_{14}\tau\bigr\} \ll 1$.

Сначала будем считать, что $\tau/\sigma \leqslant c_9$. Ясно, что

$$\begin{aligned}\exp\bigl\{|\,\xi-\eta\,|/c_{14}\tau\bigr\} &\leqslant \exp\bigl\{|\,\xi-\eta\,|/c_{14}\tau\bigr\}\,\mathbf{1}\bigl\{|\xi| \leqslant 2\sigma\bigr\}\\ &\quad + \exp\bigl\{|\,\xi-\eta\,|/c_{14}\tau\bigr\}\,\mathbf{1}\bigl\{2\sigma \leqslant |\,\xi\,| \leqslant z\bigr\}\\ &\quad + \exp\bigl\{|\,\eta\,|/c_{14}\tau + |\,\xi\,|/c_{14}\tau\bigr\}\,\mathbf{1}\bigl\{|\xi| \geqslant z\bigr\}.\end{aligned} \tag{68}$$

Согласно (51), при $c_{14} > c_7$

$$\mathbf{E}\,\exp\bigl\{|\,\xi-\eta\,|/c_{14}\tau\bigr\}\,\mathbf{1}\bigl\{|\xi| \leqslant 2\sigma\bigr\} \leqslant e.$$

В силу леммы 4,

$$\mathbf{P}\bigl\{\xi^2 \geqslant x^2\bigr\} = \mathbf{P}\bigl\{|\xi| \geqslant x\bigr\} \leqslant 2\max\Bigl\{\exp\Bigl\{-\frac{x^2}{4\sigma^2}\Bigr\}, \exp\Bigl\{-\frac{x^2}{4c_{10}\sigma^2}\Bigr\}\Bigr\} \tag{69}$$

при $0 \leqslant x \leqslant z = c_{10}\,\sigma^2/\tau$. Положим

$$W = c_{11}\xi^2/\sigma^2.$$

Используя (69), мы получаем, что существует c_{13}, такое что

$$\mathbf{P}\bigl\{W \geqslant u\bigr\} \leqslant \mathbf{P}\Bigl\{\bigl|\xi\bigr| \geqslant \sigma\,\sqrt{u/c_{11}}\Bigr\} \leqslant 2\exp\Bigl\{-u/c_{13}\Bigr\}$$

при $0 \leqslant u \leqslant \gamma = c_{11}z^2/\sigma^2$. Положим $c_{12} = 2c_{13}$, $v = 1/c_{12}$. Интегрируя по частям, получаем

$$\begin{aligned}\mathbf{E}\exp\bigl\{v\,W\bigr\}\,\mathbf{1}\bigl\{|\,\xi\,| \leqslant z\bigr\} &\leqslant 1 + \int_0^{\gamma} v\,e^{v\,u}\,\mathbf{P}\bigl\{W \geqslant u\bigr\}\,du\\ &\leqslant 1 + \frac{2}{c_{12}}\int_0^{\gamma} e^{u/c_{12}}\,\exp\Bigl\{-u/c_{13}\Bigr\}\,du\\ &\leqslant 1 + \frac{2}{c_{12}}\int_0^{\infty} e^{-u/c_{12}}\,du \ll 1,\end{aligned}$$

Согласно (67), при $c_{14} > c_{12}$

$$\exp\bigl\{|\xi-\eta|/c_{14}\tau\bigr\}\mathbf{1}\bigl\{2\sigma\leqslant|\xi|\leqslant z\bigr\}\leqslant\exp\bigl\{c_{11}\xi^2/c_{12}\sigma^2\bigr\}\,\mathbf{1}\bigl\{2\sigma\leqslant|\xi|\leqslant z\bigr\}.$$

Следовательно,

$$\mathbf{E}\,\exp\bigl\{|\,\xi-\eta\,|/c_{14}\tau\bigr\}\,\mathbf{1}\bigl\{2\sigma\leqslant|\,\xi\,|\leqslant z\bigr\}\leqslant\mathbf{E}\exp\bigl\{v\,W\bigr\}\,\mathbf{1}\bigl\{|\,\xi\,|\leqslant z\bigr\}\ll 1.$$

По неравенству Коши–Буняковского–Шварца, при $t\in\mathbf{R}$, $z>0$ мы имеем

$$\mathbf{E}\,\exp\bigl\{t\,|\,\xi\,|+t\,|\,\eta\,|\bigr\}\,\mathbf{1}\bigl\{|\xi|\geqslant z\bigr\}$$
$$\leqslant\Bigl(\mathbf{E}\,\exp\bigl\{2t\,|\,\xi\,|\bigr\}\,\mathbf{1}\,\bigl\{|\xi|\geqslant z\bigr\}\cdot\mathbf{E}\,\exp\bigl\{2t\,|\,\eta\,|\bigr\}\,\mathbf{1}\,\bigl\{|\xi|\geqslant z\bigr\}\Bigr)^{1/2}, \quad (70)$$

а также

$$\mathbf{E}\,\exp\bigl\{2t\,|\,\xi\,|\bigr\}\,\mathbf{1}\,\bigl\{|\xi|\geqslant z\bigr\}\leqslant\Bigl(\mathbf{E}\,\exp\bigl\{4t\,|\,\xi\,|\bigr\}\cdot\mathbf{P}\,\bigl\{|\xi|\geqslant z\bigr\}\Bigr)^{1/2}. \quad (71)$$

Ясно, что

$$\exp\bigl\{4t\,|\,\xi\,|\bigr\}\leqslant\exp\bigl\{4t\,\xi\bigr\}+\exp\bigl\{-4t\,\xi\bigr\}. \quad (72)$$

Применяя лемму 4, получаем, что при $z=c_{10}\,\sigma^2/\tau$

$$\mathbf{P}\,\bigl\{|\xi|\geqslant z\bigr\}\leqslant 2\,\exp\{-c\sigma^2/\tau^2\}. \quad (73)$$

Пусть $0\leqslant|h|\tau\leqslant\frac{1}{2}$. Согласно (27),

$$\mathbf{E}\,e^{h\xi}\leqslant\exp\{h^2\sigma^2\}. \quad (74)$$

Применяя (70)–(74) при $h=4t=\pm4/c_{14}\tau$, $z=c_{10}\,\sigma^2/\tau$ и выбирая константу c_{14} достаточно большой, получаем

$$\mathbf{E}\,\exp\bigl\{2t\,|\,\xi\,|\bigr\}\,\mathbf{1}\,\bigl\{|\xi|\geqslant z\bigr\}$$
$$\leqslant\Bigl(\mathbf{E}\,\bigl(\exp\bigl\{4t\,\xi\bigr\}+\exp\bigl\{-4t\,\xi\bigr\}\bigr)\cdot\mathbf{P}\,\bigl\{|\xi|\geqslant z\bigr\}\Bigr)^{1/2}\leqslant\sqrt{2}. \quad (75)$$

Аналогично проверяется, что

$$\mathbf{E}\,\exp\bigl\{2t\,|\,\eta\,|\bigr\}\,\mathbf{1}\,\bigl\{|\xi|\geqslant z\bigr\}\leqslant\sqrt{2}. \quad (76)$$

Следовательно,

$$\mathbf{E}\,\exp\bigl\{|\,\eta\,|/c_{14}\tau+|\,\xi\,|/c_{14}\tau\bigr\}\,\mathbf{1}\bigl\{|\xi|\geqslant c_{10}\,\sigma^2/\tau\bigr\}\leqslant 2\sqrt{2}. \quad (77)$$

Пусть теперь $\tau>c_9\sigma$. Тогда

$$\mathbf{E}\exp\bigl\{|\xi-\eta|/c_{14}\tau\bigr\}\leqslant\Bigl(\mathbf{E}\exp\bigl\{2|\eta|/c_{14}\tau\bigr\}\cdot\mathbf{E}\exp\bigl\{2|\xi|/c_{14}\tau\bigr\}\Bigr)^{1/2}. \quad (78)$$

$$\mathbf{E}\,\exp\bigl\{2\,|\,\xi\,|/c_{14}\tau\bigr\}\leqslant\mathbf{E}\,\exp\bigl\{2\,\xi/c_{14}\tau\bigr\}+\mathbf{E}\,\exp\bigl\{-2\,\xi/c_{14}\tau\bigr\}. \quad (79)$$

Применяя (74) при $h = \pm 2/c_{14}\tau$, и выбирая константу c_{14} достаточно большой, получаем, что

$$\mathbf{E}\,\exp\left\{|\,\xi-\eta\,|/c_{14}\tau\right\} \ll 1, \tag{80}$$

завершая доказательство теоремы.

Список литературы

1. Т. В. Арак, А. Ю. Зайцев, *Равномерные предельные теоремы для сумм независимых случайных величин.* — Тр. МИАН СССР **174** (1986).
2. G. Bennett, *Probability inequalities for the sum of independent random variables.* — J. Amer. Statist. Assoc. **57** (1962), 33–45.
3. V. Bentkus, *A remark on Bernstein, Prokhorov, Bennett, Hoeffding, and Talagrand inequalities.* — Lith. Math. J. **42**, No. 3, 262–269 (2002).
4. V. Bentkus, *On Hoeffding's inequalities.* — Ann. Probab. **32**, No. 2 (2004), 1650–1673.
5. I. Berkes, W. Philipp, *Approximation theorems for independent and weakly dependent random vectors.* — Ann. Probab. **7**, (1979), 29–54.
6. С. Н. Бернштейн, Теория вероятностей, М., Гостехиздат, 1946.
7. А. Бикялис, *Оценки остаточного члена в центральной предельной теореме.* — Лит. матем. сб. **6**, No. 3 (1966), 323–346.
8. S. Bobkov, F. Götze, *Esscher transform and the central limit theorem.* — J. Funct. Anal. **289**, No. 5, Article ID 110999, 39 p. (2025).
9. В. И. Богачев, А. В. Колесников, *Задача Монжа–Канторовича: достижения, связи и перспективы.* — Успехи матем. наук **67**, No. 5(407) (2012), 3–110.
10. В. И. Богачев, *Задача Канторовича оптимальной транспортировки мер: новые направления исследований.* — Успехи матем. наук **77**, No. 5(467) (2022), 3–52.
11. А. А. Боровков, А. А. Могульский, А. И. Саханенко, *Предельные теоремы для случайных процессов.* Теория вероятностей – 7, Итоги науки и техн. Сер. Соврем. пробл. мат. Фундам. направления, **82**, 5–194, М., ВИНИТИ, 1995.‘
12. A. A. Borovkov, A. I. Sakhanenko, *On the rate of convergence in invariance principle. — Lect. Notes Math.* **1021** (1981), 59–66.
13. H. Döring, S. Jansen, K. Schubert, *The method of cumulants for the normal approximation.* — Probab. Surv. **19** (2022), 185–270.
14. R. M. Dudley, *Distances of probability measures and random variables.* — Ann. Math. Statist. **39**, No. 5 (1968), 1563–1572.
15. W. Hoeffding, *Probability inequalities for sums of bounded random variables.* — J. Amer. Statist. Assoc. **58** (1963), 13–30.
16. А. Н. Колмогоров, *Две равномерные предельные теоремы для сумм независимых слагаемых.* — Теория вероятн. и ее примен. **1**, No. 4 (1956), 426–436.
17. А. Н. Колмогоров, *О приближении распределений сумм независимых слагаемых неограниченно делимыми распределениями.* — Труды Москов. матем. обва **12** (1963), 437–451.

18. J. Komlós, P. Major, G. Tusnády, *An approximation of partial sums of independent RV'-s and the sample DF*, I; II. — Z. Wahrscheinlichkeitstheor. verw. Geb. **32** (1975), 111–131; **34** (1976), 34–58.
19. S. V. Nagaev, *Large deviations of sums of independent random variables.* — Ann. Probab. **7** (1979), 745–789.
20. В. В. Петров, *Предельные теоремы для сумм независимых случайных величин*, М., Наука, 1987.
21. I. Pinelis, *On the Bennett–Hoeffding inequality.* — Ann. Inst. Henri Poincaré, Probab. Statist. **50**, No. 1 (2014), 15–27.
22. Ю. В. Прохоров, *Сходимость случайных процессов и предельные теоремы теории вероятностей.* — Теория вероятн. и ее примен. **1** (1956), 177–238.
23. E. Rio, *Upper bounds for minimal distances in the central limit theorem.* — Ann. Inst. Henri Poincaré, Probab. Statist. **45**, No. 3 (2009), 802–817.
24. Р. Рудзкис, *О вероятностях больших уклонений случайных векторов.* — Лит. матем. сб. **23**, No. 1 (1983), 195–204.
25. Л. Саулис, *О больших уклонениях для случайных векторов для некоторых классов множеств, I.* — Лит. матем. сб. **23**, No. 3 (1983), 142–154.
26. Л. Саулис, В. Статулявичус, *Предельные теоремы о больших уклонениях.* Вильнюс, Мокслас, 1989.
27. Л. Саулис, В. Статулявичус, *Предельные теоремы о больших уклонениях.* — Итоги науки и техн. Сер. Соврем. пробл. матем. Фундам. направления, **81** (1991), 219–312.
28. А. И. Саханенко, *Скорость сходимости в принципе инвариантности для разнораспределенных величин с экспоненциальными моментами.* — Труды инст. матем. СО АН СССР, **3**, 4–49, Новосибирск, Наука, 1984.
29. А. И. Саханенко, *Оценки точности построений на одном вероятностном пространстве в центральной предельной теореме.* — Сиб. матем. журн. **37**, No. 4 (1996), 919–931;
30. G. Schay, *Nearest random variables with given distributions.* — Ann. Probab. **2** (1974), 163–166.
31. V. A. Statulevičius, *On large deviations.* — Z. Wahrscheinlichkeitstheor. verw. Geb. **62** (1966), 133–144.
32. V. Strassen, *The existence of probability measures with given marginals.* — Ann. Math. Statist. **36** (1965), 423–439.
33. С. С. Валландер, *Вычисление расстояния по Вассерштейну между распределениями вероятностей на прямой.* — Теория вероятн. и ее примен., 18, No. 4 (1973), 824–827.
34. V. V. Yurinskii, *Exponential inequalities for sums of random vectors.* — J. Multiv. Anal. **6** (1976), 473–499.
35. А. Ю. Зайцев, *О гауссовской аппроксимации сверток при выполнении многомерных аналогов условий неравенства Бернштейна.* — Препринт Р-9-84, Л.: ЛОМИ, 1984, 54 с.
36. А. Ю. Зайцев, *Оценки расстояния Леви–Прохорова в многомерной центральной предельной теореме для случайных величин с конечными экспоненциальными моментами.* — Теория вероятн. и ее примен. **31** (1986), 246–265.

37. A. Yu. Zaitsev, *On the Gaussian approximation of convolutions under multidimensional analogues of S. N. Bernstein inequality conditions.* — Probab. Theor. Rel. Fields **74** (1987), 535–566.
38. А. Ю. Зайцев, *О связи между двумя классами вероятностных распределений.* — В сб.: Кольца и модули. Предельные теоремы теории вероятностей. **3**, Ленинград, Изд-во ЛГУ, 1988, 153–158.
39. А. Ю. Зайцев, *Многомерный вариант второй равномерной предельной теоремы Колмогорова.* — Теория вероятн. и ее примен. **34**, No. 1 (1989), 128–151.
40. А. Ю. Зайцев, *Оценки квантилей гладких условных распределений и многомерный принцип инвариантности.* — Сибирский матем. журнал **37**, No. 4 (1996), 807–831.
41. A. Yu. Zaitsev, *Multidimensional version of the results of Komlós, Major, and Tusnády for vectors with finite exponential moments.* — ESAIM: Probab. Statist. **2** (1998), 41–108.
42. A. Yu. Zaitsev, *Multidimensional version of the results of Sakhanenko in the invariance principle for vectors with finite exponential moments*, I; II; III. — Теория вероятн. и ее примен. **45** (2000), 718–738; **46** (2001), 535–561; 744–769.
43. А. Ю. Зайцев, *Точность сильной гауссовской аппроксимации для сумм независимых случайных векторов.* — Успехи матем. наук **68**, No. 4(412) (2013), 129–172.

Zaitsev A. Yu. Transport distance estimates in the Central Limit Theorem.

Let $X_1, \dots, X_n$ be d-dimensional independent random vectors bounded with probability one. For simplicity, we assume that they have zero mean values:

$$\mathbf{P}\{\|X_j\| \leqslant \tau\} = 1, \quad \mathbf{E}\, X_j = 0, \quad j = 1, \dots, n.$$

We study the distribution behavior of the sum $S = X_1 + \dots + X_n$ as a function of the bounding value τ.

From the non-uniform Bikelis estimate in the one-dimensional central limit theorem it follows that

$$W_1(F, \Phi_\sigma) \leqslant c\tau.$$

with an absolute constant c, where W_1 is the Kantorovich–Rubinstein–Wasserstein transport distance, F is the distribution of the sum S, and Φ_σ is the corresponding normal distribution. The main result of the paper is significantly stronger and more precise. It is claimed that

$$\rho(F, \Phi_\sigma) = \inf \int \exp(|x - y|/c\tau)\, d\pi(x, y) \leqslant c,$$

where the infimum is taken over all bivariate probability distributions π with marginal distributions F and Φ_σ. The result has also been generalized to distributions with sufficiently slowly growing cumulants from the class $\mathcal{A}_1(\tau)$, introduced in the author's 1986 paper. The possibility of generalizing the result to the multivariate case is discussed.

С.-Петербургское отделение
Математического института им. В. А. Стеклова,
Фонтанка 27,
Санкт-Петербург 191023, Россия
и Санкт-Петербургский государственный университет,
Университетская наб. 7/9,
Санкт-Петербург, 199034 Россия
E-mail: zaitsev@pdmi.ras.ru

Поступило 24 октября 2025 г.