О СКОРОСТИ СХОДИМОСТИ В ТЕОРЕМЕ ЛЯПУНОВА

И. С. Тюрин

1. Введение. Рассмотрим независимые (действительные) случайные величины (с.в.) $X_1, \ldots, X_n$ с нулевыми средними, дисперсиями равными $\sigma_1^2, \ldots, \sigma_n^2$ и конечными третьими абсолютными моментами $\beta_1, \ldots, \beta_n$. Обозначим

$$\sigma^2 = \sigma^2(n) := \sum_{j=1}^{n} \sigma_j^2, \quad \varepsilon_n := \frac{1}{\sigma^3} \sum_{j=1}^{n} \beta_j.$$

Согласно теореме Ляпунова $S_n := (X_1 + \ldots + X_n)/\sigma(n)$ слабо сходится к стандартной нормальной с.в. при $\varepsilon_n \to 0$. С теоретической и практической точки зрения немалый интерес представляет оценка скорости сходимости в этой теореме. Как известно [1, 2], существует такая наименьшая постоянная C, что для расстояния Колмогорова между S_n и стандартной гауссовской величиной N справедливо неравенство

$$\rho(S_n, N) := \sup_{x \in \mathbb{R}} \bigl| \mathsf{P}(S_n \leqslant x) - \mathsf{P}(N \leqslant x) \bigr| \leqslant C\varepsilon_n, \ \ n \in \mathbb{N}. \tag{1}$$

Оценке этой константы посвящено множество работ. Эссеен [2] показал, что $C \leqslant 7{,}5$. Бергстрем [3] получил оценку $C \leqslant 4{,}8$. Такано [4] установил, что в случае независимых одинаково распределенных (н.о.р.) слагаемых $C \leqslant 2{,}031$. В. М. Золотарев [5, 6, 7, 8] получил новое неравенство, позволяющее оценивать близость двух сумм независимых с. в. и с его помощью показал последовательно справедливость оценок $C \leqslant 1{,}322$ и $C \leqslant 0{,}9051$, а для случая н.о.р. величин $C \leqslant 1{,}301$ и $C \leqslant 0{,}8197$. Предложенный метод получил развитие в работах ван Бика [9] и И. С. Шиганова [10], которые доказали, соответственно, оценки $C \leqslant 0{,}7975$ и $C \leqslant 0{,}7915$. Для сумм с.в., имеющих одинаковое распределение, И. С. Шиганов получил оценку $C \leqslant 0{,}7655$, которую в 2006 году улучшила И. Г. Шевцова [11], показав, что в этом случае $C \leqslant 0{,}7056$.

Отдельного упоминания заслуживает родственная задача отыскания асимптотически правильных постоянных в теореме Ляпунова. Как показал Эссеен [12], если все с.в. X_j, $j = 1, 2, \ldots$ имеют одно и то же распределение, то

$$\limsup_{n \to \infty} \frac{\rho(S_n, N)}{\varepsilon_n} \leqslant C_1 := \frac{\sqrt{10} + 3}{6\sqrt{2\pi}} = 0{,}409\ldots, \tag{2}$$

причем константа в правой части этого неравенства не может быть уменьшена (откуда, в частности, следует нижняя оценка $C \geqslant C_1$). Последний результат был развит Б. А. Рогозиным [13], который установил, что при тех же предположениях

$$\limsup_{n \to \infty} \frac{\rho(S_n, \mathcal{N})}{\varepsilon_n} \leqslant C_2 := \frac{1}{\sqrt{2\pi}}, \tag{3}$$

где $\rho(S_n, \mathcal{N}) := \inf_{G \in \mathcal{N}} \rho(S_n, G)$, $\mathcal{N}$ – множество всех нормальных с.в.

Г. П. Чистяков [14, 15, 16] обобщил (2) и (3) на случай неодинаково распределенных слагаемых. Он доказал, что

$$\rho(S_n, N) \leqslant C_1\varepsilon_n + r_1(\varepsilon_n), \qquad \rho(S_n, \mathcal{N}) \leqslant C_2\varepsilon_n + r_2(\varepsilon_n),$$

где $r_1(\varepsilon_n), r_2(\varepsilon_n)$ есть $o(\varepsilon_n)$ при $\varepsilon_n \to 0$.

Известны также оценки скорости сходимости в теореме Ляпунова при наличии моментов порядка $2+\delta$ (см. [17, 18]).

Аналоги (1) существуют и для других вероятностных метрик, например, ζ_r (где $r = 1, 2, 3$), о которых подробнее будет сказано в разделе 2. Оценки в этих метриках могут быть получены естественным образом при использовании метода Стейна. При доказательстве упомянутых оценок, в частности, используется так называемое нулевое смещение вероятностного распределения (см. [19]).

Для расстояния в метриках ζ_r $(r = 1, 2, 3)$ известны следующие оценки (см. [20])

$$\zeta_1(S_n, N) \leqslant 3\varepsilon_n, \qquad \zeta_2(S_n, N) \leqslant \frac{3\sqrt{2\pi}}{8}\varepsilon_n, \qquad \zeta_3(S_n, N) \leqslant \frac{1}{2}\varepsilon_n. \tag{4}$$

В [21] рассмотрена задача об отыскании наименьшей верхней грани $\mathsf{E} f(X_1, \ldots, X_n)$ по множеству наборов независимых простых с.в., удовлетворяющих m ограничениям вида $\mathsf{E}\, g_{ij}(X_j) = c_{ij}, j = 1, \ldots, n$. А именно, установлено, что в этом случае можно ограничиться рассмотрением с.в., принимающих не более $m+1$ значений. В настоящей работе результаты [21] обобщены на случай произвольного квазивыпуклого функционала, заданного на множестве вероятностных распределений.

При помощи полученных результатов установлена неулучшаемая оценка близости в средней метрике вероятностного распределения и его преобразования нулевого смещения. Последнее применено к оценке точности гауссовской аппроксимации распределений нормированных сумм независимых слагаемых. Установлено, что значения констант в (4) можно уменьшить в 3 раза. При этом, наша оценка, полученная для метрики ζ_3, оптимальна. Кроме того, выведены новые оценки близости характеристических функций нормированной суммы и стандартной нормальной с.в. Это позволило доказать, что $C \leqslant 0{,}6379$, а в случае н.о.р. слагаемых $C \leqslant 0{,}5894$.

2. Обозначения и формулировки результатов. Пусть (S, d) – некоторое метрическое пространство, Q – множество зарядов ограниченной вариации на σ-алгебре $\mathcal{B}(S)$ борелевских подмножеств S с операциями умножения на скаляр и сложения, определенными следующим образом: пусть $\mu, \mu_1, \mu_2 \in Q,\ c \in \mathbb{R}$, положим для каждого $A \in \mathcal{B}(S)$

$$(c\mu)(A) := c \cdot \mu(A), \quad (\mu_1 + \mu_2)(A) := \mu_1(A) + \mu_2(A).$$

Легко видеть, что Q замкнуто относительно этих операций и представляет собой линейное пространство, а множество D дискретных вероятностных мер, сосредоточенных в конечном числе точек, есть выпуклое подмножество Q. Последнее означает, что $\alpha\mu_1 + (1-\alpha)\mu_2 \in D$ для произвольных $\mu_1, \mu_2 \in D$ и $\alpha \in (0, 1)$.

Рассмотрим множество наборов, состоящих из n независимых с.в. $X_1, \ldots, X_n$. Тогда

$$\mathsf{E} f(X_1, \ldots, X_n) = \int\limits_{\mathbb{R}^n} f dP_{X_1} \ldots dP_{X_n},$$

где $f : \mathbb{R}^n \to \mathbb{R}$, $P_{X_1}, \ldots, P_{X_n}$ – распределения с.в. $X_1, \ldots, X_n$. Таким образом, $\mathsf{E} f(X_1, \ldots, X_n)$ можно рассматривать как функцию на множестве мер, линейную по каждому из n своих аргументов.

Функция $g : G \to \mathbb{R}$, где G – выпуклое множество, называется квазивыпуклой, если для любых $x, y \in G$ и $\alpha \in (0, 1)$ имеем

$$g(\alpha x + (1-\alpha)y) \leqslant \max\{g(x), g(y)\}.$$

Всюду далее будем предполагать, что заданы $h_1, \ldots, h_m$ – некоторые вещественнозначные функции на S. Рассмотрим множество

$$K := \{\mu \in D : \langle h_i, \mu\rangle = 0,\ i = 1, \ldots, m\},\ \text{где}\ \langle f, \mu\rangle := \int\limits_S f d\mu.$$

Нетрудно видеть, что K – выпуклое множество. Кроме того, обозначим K_j – множество мер $\mu \in K$, сосредоточенных на не более чем j точках ($j \in \mathbb{N}$).

Теорема 1. *Для произвольной квазивыпуклой функции* $g : K \to \mathbb{R}$ *имеем*

$$\sup_{\mu \in K} g(\mu) = \sup_{\mu \in K_{m+1}} g(\mu).$$

В последнем выражении супремум по пустому множеству считаем равным нулю.

Теорема 2. *Пусть* f *– неотрицательная функция на* S, V *– векторное пространство с нормой* $\|\cdot\|$, $A : K \to V$ *– такое отображение, что*

$$A(\alpha\mu + (1-\alpha)\nu) = \alpha A\mu + (1-\alpha)A\nu \tag{5}$$

для любых $\mu, \nu \in K, \alpha \in (0, 1)$. *Тогда наименьшее* γ, *при котором неравенство*

$$\|A\mu\| \leqslant \gamma\langle f, \mu\rangle \tag{6}$$

выполнено для любой меры $\mu \in K$, *совпадает с наименьшим* γ, *при котором* (6) *верно для произвольной меры* $\mu \in K_{m+1}$.

Пусть W – с.в. с нулевым средним и дисперсией $\sigma^2 > 0$. Говорят, что с.в. W^* имеет распределение W-нулевого смещения, если

$$\mathsf{E}Wf(W) = \sigma^2\mathsf{E}f'(W^*) \tag{7}$$

для каждой дифференцируемой функции $f : \mathbb{R} \to \mathbb{R}$, для которой левая часть (7) определена. Известно (см. [19]), что W^* существует для любой описанной выше W и имеет плотность

$$p(w) = \begin{cases} \sigma^{-2}\mathsf{E}\left(W \cdot \mathbf{1}\{W > w\}\right), & \text{если } w \geqslant 0; \\ \sigma^{-2}\mathsf{E}\left(-W \cdot \mathbf{1}\{W < w\}\right), & \text{если } w < 0. \end{cases} \tag{8}$$

Для произвольной функции $f \in C^{(r-1)}(\mathbb{R})$, где $r \in \mathbb{N}$, положим

$$M_r(f) := \sup_{x \neq y}\left|\frac{f^{(r-1)}(x) - f^{(r-1)}(y)}{x - y}\right|.$$

Как обычно, $C^{(0)}(\mathbb{R}) := C(\mathbb{R})$. Если $f \notin C^{(r-1)}(\mathbb{R})$, то считаем $M_r(f) = \infty$. Положим

$$\zeta_r(X, Y) := \sup\{|\mathsf{E}f(X) - \mathsf{E}f(Y)| : f \in \mathcal{F}_r\},\ \ r = 1, 2, \ldots,$$

где $\mathcal{F}_r$ – множество вещественных ограниченных функций с $M_r(f) \leqslant 1$.

Заметим, что ζ_1 имеет другие представления. Это так называемая средняя метрика

$$\varkappa_1(X, Y) := \int_{-\infty}^{\infty}\big|\mathsf{P}(X \leqslant x) - \mathsf{P}(Y \leqslant x)\big|dx,$$

а также минимальная L_1-метрика

$$l_1(X,Y) := \inf\left\{ \mathsf{E}\,|\widetilde{X}-\widetilde{Y}| : Law(\widetilde{X}) = Law(X), Law(\widetilde{Y}) = Law(Y)\right\}.$$

По этому поводу см. [22, с. 56].

Теорема 3. *Пусть W центрирована, имеет единичную дисперсию и конечный третий абсолютный момент. Тогда*

$$\zeta_1(W,W^*) \leqslant \frac{1}{2}\mathsf{E}|W|^3, \tag{9}$$

причем равенство достигается, когда W принимает ровно два значения.

Следствие 1. *Пусть с.в. S_n^* имеет распределение S_n-нулевого смещения. Тогда*

$$\zeta_1(S_n,S_n^*) \leqslant \frac{1}{2}\varepsilon_n.$$

Теорема 4. *Справедливы следующие неравенства:*

$$\zeta_1(S_n,N) \leqslant 2\zeta_1(S_n,S_n^*) \leqslant \varepsilon_n, \qquad \zeta_2(S_n,N) \leqslant \frac{\sqrt{2\pi}}{4}\zeta_1(S_n,S_n^*) \leqslant \frac{\sqrt{2\pi}}{8}\varepsilon_n, \tag{10}$$

$$\zeta_3(S_n,N) \leqslant \frac{1}{3}\zeta_1(S_n,S_n^*) \leqslant \frac{1}{6}\varepsilon_n. \tag{11}$$

Последнее двойное неравенство оптимально, а именно, для любого $\delta > 0$ существует такая последовательность н.о.р. с.в. $X_1, X_2, \ldots$, что

$$\frac{\zeta_3(S_n,N)}{\varepsilon_n} \geqslant \frac{1}{6} - \delta, \quad n = 1, 2, \ldots$$

Для $\gamma > 0$ и $t \in \mathbb{R}$ положим

$$b(t,\gamma) := \begin{cases} -t^2 + 2\gamma a|t|^3, & \text{если } \gamma|t| < M; \\ -2\left(\frac{1}{\gamma}\right)^2(1-\cos\gamma t), & \text{если } M \leqslant \gamma|t| \leqslant 2\pi; \\ 0, & \text{если } \gamma|t| > 2\pi. \end{cases}$$

Здесь

$$a := \max_{x>0}\{(\cos(x) - 1 + x^2/2)/x^3\} \approx 0{,}099162,$$

M – точка, в которой достигается этот максимум, $M \approx 3{,}995896$.

Обозначим $f_{S_n}(t) := \mathsf{E}e^{itS_n}$, $\varphi(t) := \exp(-t^2/2)$, $\delta_n(t) := |f_{S_n}(t) - \varphi(t)|$, $t \in \mathbb{R}$.

Теорема 5. *Для произвольного $t \in \mathbb{R}$ имеем*

$$\delta_n(t) \leqslant \delta^{(1)}(\varepsilon_n, t) := \exp\left(\frac{1}{2}b(t, 2\varepsilon_n)\right) + \varphi(t), \tag{12}$$

$$\delta_n(t) \leqslant \delta^{(2)}(\varepsilon_n, t) := \varepsilon_n\varphi(t)\int_0^t \frac{s^2}{2}\exp\left(\frac{s^2}{2}\right)ds. \tag{13}$$

Обозначим $A := \varepsilon_n^{-1/3}/6a$. *Для всех* $t \in \mathbb{R}$ *верна оценка*

$$\delta_n(t) \leqslant \delta^{(3)}(\varepsilon_n, t) := \begin{cases} \varepsilon_n \varphi(t) \int\limits_0^t \frac{s^2}{2} \exp\left(\frac{s^2 \varepsilon_n^{2/3}}{2}\right) ds, & |t| \leqslant A; \\ \varepsilon_n \varphi(t) \left(\int\limits_0^A \frac{s^2}{2} \exp\left(\frac{s^2 \varepsilon_n^{2/3}}{2}\right) ds + \int\limits_A^t \frac{s^2}{2l} \exp\left(2a\varepsilon_n s^3\right) ds \right), & |t| > A. \end{cases} \tag{14}$$

где

$$l := \inf_{t \geqslant 0} \left\{ \exp\left(-t^2/2 + 2at^3\right)\right\} \approx 0{,}624489.$$

Для случая н.о.р. величин оценки могут быть несколько улучшены. Обозначим $\tau_n := \frac{1}{\sigma^3} \sum_{j=1}^n \sigma_j^3$. Пусть $X_1, X_2, \ldots$ – центрированные н.о.р. с.в. с дисперсиями 1 и конечными третьими абсолютными моментами β. Тогда $\varepsilon_n = \beta/\sqrt{n}$, $\tau_n = 1/\sqrt{n}$.

Теорема 6. *Для описанной последовательности с.в. и произвольного* $t \in \mathbb{R}$ *справедливы неравенства*

$$\delta_n(t) \leqslant \delta^{(4)}(n, \varepsilon_n, t) := \left(1 + \frac{b(t, \varepsilon_n + \tau_n)}{n}\right)^{\frac{n}{2}} + \varphi(t), \tag{15}$$

$$\delta_n(t) \leqslant \delta^{(5)}(n, \varepsilon_n, t) := \varepsilon_n \varphi(t) \int\limits_0^t \left(1 + \frac{b(s, \varepsilon_n + \tau_n)}{n}\right)^{\frac{n-1}{2}} \frac{s^2}{2} \exp\left(\frac{s^2}{2}\right) ds. \tag{16}$$

Пусть $m \in \mathbb{N}$ *и* $n \geqslant m$. *Тогда*

$$\delta_n(t) \leqslant \delta^{(6)}(m, \varepsilon_n, t) := \exp\left(\frac{1}{2} b(t, \varepsilon_n + \tau_m)\right) + \varphi(t), \tag{17}$$

$$\delta_n(t) \leqslant \delta^{(7)}(m, \varepsilon_n, t) := \varepsilon_n \varphi(t) \int\limits_0^t \exp\left(\frac{m-1}{2m} b(s, \varepsilon_n + \tau_m) + \frac{s^2}{2}\right) \frac{s^2}{2} ds. \tag{18}$$

Оценки (12)-(18) позволили установить следующий результат.

Теорема 7. *Неравенство (1) имеет место с константой* $C \leqslant 0{,}6379$, *а в случае одинаково распределенных слагаемых* $C \leqslant 0{,}5894$.

3. Доказательства основных результатов.

Доказательство теоремы 1. Если $K = \varnothing$, то $K_{m+1} = \varnothing$, и утверждение теоремы выполнено. Далее предполагаем, что множество K непусто.

Последовательность множеств $K_1, K_2, \ldots$ возрастает к множеству K, следовательно,

$$\sup_{\mu \in K} g(\mu) = \sup_{j \geqslant 1} \sup_{\mu \in K_j} g(\mu) = \sup_{j \geqslant m+1} \sup_{\mu \in K_j} g(\mu).$$

Поэтому достаточно установить, что

$$\sup_{\mu \in K_{m+1}} g(\mu) \geqslant \sup_{\mu \in K_{m+2}} g(\mu) \geqslant \sup_{\mu \in K_{m+3}} g(\mu) \geqslant \ldots$$

Для этого возьмем произвольную меру $\mu \in K_j$, где $j > m+1$, и покажем, что существует $\mu' \in K_{j-1}$, такая что $g(\mu') \geqslant g(\mu)$.

Пусть μ сосредоточена в точках $s_1, \ldots, s_j \in S$ и

$$\mu(s_i) = \mu_i \geqslant 0, \;\; i = 1, \ldots, j. \tag{19}$$

Вектор $\bar{\mu} = (\mu_1, \ldots, \mu_j)$ задает вероятностное распределение, поэтому

$$\mu_1 + \ldots + \mu_j = 1. \tag{20}$$

Кроме того, выполнены условия $\langle h_i, \mu \rangle = 0, \ i = 1, \ldots, m,$ и, следовательно,

$$\mu_1 \cdot h_i(s_1) + \ldots + \mu_j \cdot h_i(s_j) = 0, \quad i = 1, \ldots, m. \tag{21}$$

Обратно, всякий вектор с неотрицательными координатами, удовлетворяющий системе линейных уравнений (20) и (21) задает согласно формулам (19) элемент множества K_j, а если одна из его координат равна нулю, то элемент K_{j-1}. Имеем $m+1$ уравнение и как минимум $m+2$ неизвестных, следовательно, существует ненулевое решение $\bar{\nu} = (\nu_1, \ldots, \nu_j)$ соответствующей однородной системы. Поскольку сумма координат этого вектора равна нулю, но сам вектор ненулевой, то у него есть положительная и отрицательная координаты. Поэтому существуют наименьшие $\alpha \geqslant 0$ и $\beta \geqslant 0$, что соответственно одна из координат вектора $\bar{\mu_*} = \bar{\mu} - \alpha\bar{\nu}$ равна нулю, и некоторая координата $\bar{\mu^*} = \bar{\mu} + \beta\bar{\nu}$ равна нулю. Если $\alpha = 0$, то $\mu \in K_{j-1}$, иначе

$$\bar{\mu} = \frac{\beta}{\alpha+\beta}\bar{\mu_*} + \frac{\alpha}{\alpha+\beta}\bar{\mu^*},$$

и в силу квазивыпуклости $g(\mu) \leqslant \max\{g(\mu_*), g(\mu^*)\}$, где μ_*, μ^* – распределения, отвечающие $\bar{\mu_*}$ и $\bar{\mu^*}$. Таким образом $g(\mu) \leqslant g(\mu_*)$ либо $g(\mu) \leqslant g(\mu^*)$. Но μ_* и $\mu^* \in K_{j-1}$. $\square$

Доказательство теоремы 2. В силу теоремы 1 достаточно показать, что при каждом фиксированном значении γ функция

$$g(\mu) := \|A\mu\| - \gamma\langle f, \mu\rangle$$

является квазивыпуклой. Пусть $\alpha + \beta = 1$. Согласно свойствам нормы

$$\begin{aligned}\|A(\alpha\mu + \beta\nu)\| - \gamma\langle f, \alpha\mu + \beta\nu\rangle = \|\alpha A\mu + \beta A\nu\| - \gamma\langle f, \alpha\mu\rangle - \gamma\langle f, \beta\nu\rangle \leqslant \\ \leqslant \|\alpha A\mu\| + \|\beta A\nu\| - \gamma\langle f, \alpha\mu\rangle - \gamma\langle f, \beta\nu\rangle = \alpha g(\mu) + \beta g(\nu) \leqslant \max\{g(\mu), g(\nu)\}. \ \square\end{aligned}$$

Доказательство теоремы 3. Вначале проверим, что можно ограничиться рассмотрением простых с.в. W. Достаточно показать, что для любой с.в. W, удовлетворяющей условиям теоремы, существует последовательность $(W_n)_{n\geqslant 1}$ простых с.в. с нулевым средним и единичной дисперсией, такая что

$$\zeta_1(W_n, W_n^*) \to \zeta_1(W, W^*) \ \text{ и } \ \mathsf{E}|W_n|^3 \to \mathsf{E}|W|^3, \quad n \to \infty. \tag{22}$$

Считаем, что с.в. W задана на вероятностном пространстве $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathsf{P} = P_W)$. Построим последовательность простых с.в. $(W_n')_{n\geqslant 1}$, сходящуюся к W по норме L_3. Положим

$$W_n := \frac{W_n' - \mathsf{E}W_n'}{\sqrt{\operatorname{Var} W_n'}}.$$

Легко видеть, что W_n тоже сходится к W по норме L_3. Поэтому второе из условий (22), очевидно, выполнено. Остается показать справедливость первого из этих условий.

Из неравенства треугольника для метрики ζ_1 нетрудно получить

$$|\zeta_1(W, W^*) - \zeta_1(W_n, W_n^*)| \leqslant \zeta_1(W, W_n) + \zeta_1(W^*, W_n^*). \tag{23}$$

Первое слагаемое в правой части (23) стремится к нулю, поскольку

$$\zeta_1(W,W_n)=l_1(W,W_n)\leqslant \mathsf{E}\,|W-W_n|\leqslant \left(\mathsf{E}\,|W-W_n|^3\right)^{\frac{1}{3}}.$$

Оценим второе слагаемое. Для функции $f\in\mathcal{F}_1$ положим $F(x):=\int_0^x f(u)du$. Тогда

$$\mathsf{E}f(W^*)-\mathsf{E}f(W_n^*)=\mathsf{E}WF(W)-\mathsf{E}W_nF(W_n). \tag{24}$$

Поскольку при замене функции $f(x)$ на $f(x)-f(0)$ разность математических ожиданий в левой части (24) не изменится, можем считать без ограничения общности, что $f(0)=0$. Тогда $|f(x)-f(y)|\leqslant|x-y|$ влечет $|f(x)|\leqslant|x|$ и, как следствие, $|F(x)|\leqslant|x|^2$, $|xf(x)|\leqslant|x|^2$. По теореме о конечном приращении

$$WF(W)-W_nF(W_n)=(W-W_n)\cdot\{xF(x)\}'\,|_{x=\xi}=(W-W_n)\{F(\xi)+\xi f(\xi)\},$$

где ξ – число между W и W_n. Последнее, с учетом того, что

$$|F(\xi)+\xi f(\xi)|\leqslant 2|\xi|^2\leqslant 2(|W|+|W_n|)^2,$$

дает оценку

$$|\mathsf{E}\{WF(W)-W_nF(W_n)\}|\leqslant 2\mathsf{E}|W-W_n|\,(|W|+|W_n|)^2.$$

Наконец, из неравенства Гельдера

$$\mathsf{E}|W-W_n|\,(|W|+|W_n|)^2\leqslant\left(\mathsf{E}|W-W_n|^3\right)^{\frac{1}{3}}\left(\mathsf{E}\,(|W|+|W_n|)^3\right)^{\frac{2}{3}}.$$

Таким образом к нулю стремится и второе слагаемое (23), поскольку $(\mathsf{E}|W-W_n|^3)^{\frac{1}{3}}\to 0$ по определению сходимости в L_3.

Итак, достаточно рассматривать только простые с.в. Пусть A_1 – отображение, сопоставляющее распределению P_X с.в. X распределение ее нулевого смещения P_X^*. Кроме того, рассмотрим линейный оператор A_2, сопоставляющий заряду ν его функцию распределения (ф.р.) $G_\nu(x):=\nu\left((-\infty,x]\right)$. Нетрудно видеть, что

$$(\alpha P_{W_1}+(1-\alpha)P_{W_2})^*=\alpha P_{W_1}^*+(1-\alpha)P_{W_2}^*,$$

поэтому отображение $A_2-A_2A_1$ удовлетворяет (5). Если положить $h_1(x)=x$, $h_2(x)=x^2-1$ и в теореме 2 взять $f(x)=|x|^3$, $A=A_2-A_2A_1$, V – пространство интегрируемых (по мере Лебега) на вещественной прямой функций с нормой

$$\|G\|=\int_{-\infty}^{\infty}|G(x)|dx,$$

то получим, что достаточно рассматривать только простые с.в., принимающие не более трех значений. Ф.р. простой с.в. W имеет ступенчатый вид. По формуле (8) легко получить ф.р. W^*. Зная обе эти функции, не составляет труда найти явное выражение расстояния в средней метрике $\varkappa_1(W,W^*)$.

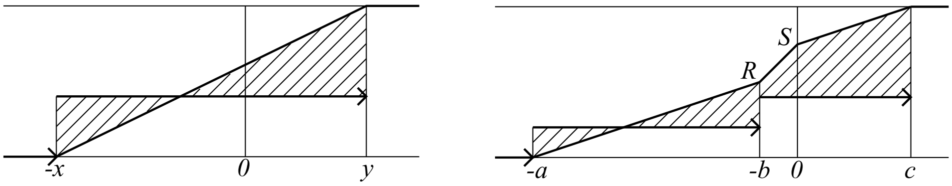

Рис. 1

Пусть W принимает ровно два значения $-x$ и y с вероятностями p и q. Тогда ее ф.р. кусочно-постоянна и имеет два скачка в точках $-x$ и y равных, соответственно, p и q. Так как W центрирована, имеем $px = qy$. Откуда, согласно (8), следует, что W^* распределена равномерно на отрезке $[-x, y]$. Следовательно, на $[-x, y]$ ее ф.р. линейна, а график представляет собой отрезок, соединяющий точки $(-x, 0)$ и $(y, 1)$. По определению $\varkappa_1(W, W^*)$ равна площади фигуры, заключенной между ф.р. (в рассматриваемом случае это объединение двух треугольников, см. рис. 1 слева).

Из того, что $\mathsf{E}\, W = 0$ и $\mathsf{E}\, W^2 = 1$ следует, что $x = \sqrt{q/p},\ \ y = \sqrt{p/q}$. Отсюда

$$\mathsf{E}\,|W|^3 = px^3 + qy^3 = q\sqrt{\frac{q}{p}} + p\sqrt{\frac{p}{q}}, \tag{25}$$

Найдем площадь фигуры, заключенной между ф.р. с.в. W и W^*. Плотность W^* равна, $px = qy = \sqrt{pq}$. Следовательно, ф.р. этой с.в. на отрезке [-x,y] имеет коэффициент наклона $\sqrt{pq}$. Длина вертикального катета первого треугольника равна p, а второго q. Отсюда сумма площадей этих треугольников равна

$$\frac{1}{2}p^2\frac{1}{\sqrt{pq}} + \frac{1}{2}q^2\frac{1}{\sqrt{pq}} = \frac{1}{2}p\sqrt{\frac{p}{q}} + \frac{1}{2}q\sqrt{\frac{q}{p}} = \frac{1}{2}\,\mathsf{E}\,|W|^3.$$

Итак, если W принимает ровно два значения, в (9) имеем равенство.

Рассмотрим случай, когда W принимает три значения. Считаем без ограничения общности, что два из них ($-a < -b$) не превосходят нуля, а одно (c) положительно. Как и ранее ф.р. с.в. W кусочно постоянна, а ф.р. W^* кусочно-линейна, однако форма фигуры, заключенной между ними, имеет более сложный вид (см. рис. 1 справа). Обозначим R значение ф.р. с.в. W^* в точке $-b$, S – ее значение в точке 0. Пусть W принимает значения $-a, -b, c$ соответственно с вероятностями p, q, r. Тогда в силу моментных ограничений

$$\begin{cases} p + q + r = 1, \\ -pa - qb + rc = 0, \\ pa^2 + qb^2 + rc^2 = 1. \end{cases}$$

Это система линейных уравнений относительно p, q, r. Воспользовавшись правилом Крамера, получим

$$p = (1 - bc)(c + b)/\Delta, \qquad q = (ac - 1)(a + c)/\Delta, \qquad r = (ab + 1)(a - b)/\Delta,$$

где $\Delta = (a+c)(b+c)(a-b)$. Таким образом, всякая с.в. с нулевым средним и дисперсией 1, принимающая три значения, определяется этими тремя значениями однозначно. Легко видеть, что p, q, r неотрицательны тогда и только тогда, когда

$$ac \geqslant 1, \qquad bc \leqslant 1. \tag{26}$$

Другими словами, с.в. W, принимающая значения $-a, -b, c$ существует в том и только том случае, когда выполнено (26). Наша цель – доказать, что функция

$$g(a, b, c) := \varkappa_1(W, W^*) - \frac{1}{2}\,\mathsf{E}\,|W|^3 \tag{27}$$

не превосходит нуля. Ее явное выражение через переменные a, b, c зависит от взаимного расположения ф.р. с.в. W и W^*. Имеется 5 случаев:

I. $R \leqslant p$, $S \leqslant p+q$, или, эквивалентно, $a(a-b) \leqslant 1$, $c \geqslant 1$. В этом случае

$$g(a,b,c) = \frac{r}{c} - pa^3 - qb^3.$$

II. $R \leqslant p$, $S \geqslant p+q \Leftrightarrow a(a-b) \leqslant 1$, $c \leqslant 1$.

$$g(a,b,c) = \frac{r}{c} - pa^3 - qb^3.$$

III. $R \geqslant p$, $S \leqslant p+q \Leftrightarrow a(a-b) \geqslant 1$, $c \geqslant 1$ (последнее влечет $c(b+c) \geqslant 1$).

$$g(a,b,c) = pa\left\{a - b - \frac{1}{a}\right\}^2 + \frac{r}{c} - pa^3 - qb^3.$$

IV. $p \leqslant R \leqslant p+q$, $S \geqslant p+q \Leftrightarrow a(a-b) \geqslant 1$, $c(b+c) \geqslant 1$, $c \leqslant 1$.

$$g(a,b,c) = pa\left\{a - b - \frac{1}{a}\right\}^2 + \frac{r}{c} - pa^3 - qb^3.$$

V. $R \geqslant p+q \Leftrightarrow c(b+c) \leqslant 1$.

$$g(a,b,c) = \frac{p}{a} - rc^3.$$

Заметим, что в каждом из случаев g – это одна и та же функция, заданная соотношением (27). Поэтому, если набор значений a, b, c удовлетворяет ограничениям сразу двух случаев, для функции g можно использовать как выражение, соответствующее первому из них, так и выражение, соответствующее второму.

Как видно, в случаях I, II, а также в случаях III, IV функция g имеет одно и то же представление. I и II объединим в случай A, III и IV – в случай B.

A. $a(a-b) \leqslant 1$.

$$g(a,b,c) = \frac{r}{c} - pa^3 - qb^3.$$

B. $a(a-b) \geqslant 1$, $c(b+c) \geqslant 1$

$$g(a,b,c) = pa\left\{a - b - \frac{1}{a}\right\}^2 + \frac{r}{c} - pa^3 - qb^3.$$

Покажем, что в каждом из случаев A, B, V функция g не превосходит нуля.

Случай A.

$$g(a,b,c) = \frac{r}{c} - pa^3 - qb^3 = \frac{(1+ab)(a-b)}{\Delta c} - \frac{(1-bc)(b+c)a^3}{\Delta} - \frac{(ac-1)(a+c)b^3}{\Delta} =$$
$$= \frac{(a-b)(ac-1)(1-bc)}{\Delta c}(-1 - bc - ac - ab) \leqslant 0$$

в силу (26), неотрицательности чисел a, b, c и того, что $a > b$.

Случай V.

$$g(a,b,c) = \frac{p}{a} - rc^3 = \frac{(1-bc)(b+c)}{\Delta a} - \frac{(1+ab)(a-b)c^3}{\Delta} =$$
$$= \frac{(ac-1)}{\Delta a}\left\{-\left[(ab+1)(a-b)\right]c^2 - \left[1 + b(a-b)\right]c - b\right\}.$$

Поскольку $ac \geqslant 1$ достаточно доказать, что выражение, стоящее внутри фигурных скобок, не превосходит нуля. Рассмотрим это выражение как функцию от переменной c, зафиксировав остальные. При $c = 0$ эта функция равна $-b \leqslant 0$. Осталось заметить, что она убывает по c, поскольку коэффициенты при c и c^2 отрицательны, и, как следствие, не превосходит нуля при любых положительных значениях c.
Случай В.

$$g(a,b,c) = pa\left\{a - b - \frac{1}{a}\right\}^2 + \frac{r}{c} - pa^3 - qb^3 = \frac{(1-bc)(b+c)a}{\Delta}\left\{a - b - \frac{1}{a}\right\}^2 + \\ + \frac{(1+ab)(a-b)}{\Delta c} - \frac{(1-bc)(b+c)a^3}{\Delta} - \frac{(ac-1)(a+c)b^3}{\Delta} = \frac{1-bc}{\Delta ac}\left\{k_2c^2 + k_1c + k_0\right\},$$

где $k_2 = 1 - 2a(a-b)(1+ab)$, $k_1 = b\left[1 - a(a-b)(1+ab)\right]$, $k_0 = a(a-b)(1+ab)$. Допустим, что $g(a,b,c) > 0$. В силу условия $a(a-b) \geqslant 1$ имеем

$$1 - a(a-b)(1+ab) \leqslant 1 - (1+ab) = -ab \leqslant 0,$$
$$1 - 2a(a-b)(1+ab) \leqslant 1 - 2(1+ab) = -1 - 2ab < 0.$$

Значит k_2 и k_1 не превосходят нуля и, следовательно, $k_2c^2 + k_1c + k_0$ убывает по c. Поэтому, если при фиксированных a и b уменьшить значение переменной c, g останется положительной. Величину c снизу ограничивают два условия задачи:

$$ac \geqslant 1 \ \text{ и } \ c(b+c) \geqslant 1.$$

Первое из этих условий избыточно, оно следует из двух других:

$$c(b+c) \geqslant 1 \ \text{ и } \ a(a-b) \geqslant 1.$$

Действительно, пусть $ac < 1$. Тогда

$$1 \leqslant c(b+c) < \frac{1}{a}\left(b + \frac{1}{a}\right) \Rightarrow a^2 < ab + 1 \Rightarrow a(a-b) < 1.$$

Значит можем уменьшить c до такого значения c_*, что $c_*(b + c_*) = 1$. При этом g останется положительной. Но ситуация, когда $c(b+c) = 1$ удовлетворяет ограничениям случая V, для которого было доказано, что $g \leqslant 0$. Пришли к противоречию. $\square$

Доказательство следствия 1. Без ограничения общности считаем $\sigma = 1$. Пусть I – случайный индекс, принимающий значения $1, \ldots, n$ с вероятностями $\sigma_1^2, \ldots, \sigma_n^2$, не зависящий от $X_1, \ldots, X_n$. Построим на расширенном вероятностном пространстве

$$S_i' := \sum_{j \neq i} X_j + X_i^*, \quad i = 1, \ldots, n,$$

где X_i^* имеет распределение X_i-нулевого смещения и не зависит от $I, X_1, \ldots, X_n$, $i = 1, \ldots, n$. Тогда $S_n^* = S_I'$ (см. [19]). Поэтому для произвольной функции $f \in \mathcal{F}_1$ имеем

$$\mathsf{E}f(S_n^*) = \mathsf{E}f(S_I') = \sum_{k=1}^n \mathsf{E}f(S_I')\mathbf{1}\{I = k\} = \sum_{k=1}^n \mathsf{E}f(S_k')\mathbf{1}\{I = k\} = \sum_{k=1}^n \sigma_k^2 \mathsf{E}f(S_k'). \qquad (28)$$

Следовательно,

$$|\mathsf{E}f(S_n) - \mathsf{E}f(S_n^*)| = \left|\sum_{k=1}^n \sigma_k^2 \mathsf{E}f(S_n) - \sum_{k=1}^n \sigma_k^2 \mathsf{E}f(S_k')\right| \leqslant \sum_{k=1}^n \sigma_k^2 \,|\mathsf{E}f(S_n) - \mathsf{E}f(S_k')| \leqslant$$
$$\leqslant \sum_{k=1}^n \sigma_k^2 \zeta_1(X_k, X_k^*) = \sum_{k=1}^n \sigma_k^3 \zeta_1\left(\frac{X_k}{\sigma_k}, \frac{X_k^*}{\sigma_k}\right) \leqslant \sum_{k=1}^n \sigma_k^3 \cdot \frac{1}{2}\,\mathsf{E}\left|\frac{X_k}{\sigma_k}\right|^3 = \frac{1}{2}\sum_{k=1}^n \beta_k = \frac{1}{2}\varepsilon_n.$$

Здесь было использовано утверждение теоремы 3 для с.в. $\frac{1}{\sigma_k}X_k$, а также свойство однородности метрики ζ_1 (т. е. $\zeta_1(cX, cY) = c\zeta_1(X,Y)$) и $(\alpha X)^* \stackrel{D}{=} \alpha X^*$. $\square$

Доказательство теоремы 4. Нетрудно видеть, что для непрерывной $f : \mathbb{R} \to \mathbb{R}$

$$h(w) := e^{\frac{w^2}{2}} \int\limits_{-\infty}^{w} (f(w) - \mathsf{E}f(N)) e^{-\frac{x^2}{2}}\, dx$$

удовлетворяет уравнению Стейна

$$h'(w) - wh(w) = f(w) - \mathsf{E}f(N).$$

Поэтому

$$|\mathsf{E}f(S_n) - \mathsf{E}f(N)| = |\mathsf{E}h'(S_n) - \mathsf{E}S_n h(S_n)| = |\mathsf{E}h'(S_n) - \mathsf{E}h'(S_n^*)| \leqslant M_2(h)\zeta_1(S_n, S_n^*).$$

Как показано в [20],

$$M_2(h) \leqslant \min\{2M_1(f), \frac{\sqrt{2\pi}}{4}M_2(f), \frac{1}{3}M_3(f)\}.$$

Принимая во внимание то, что $M_r(f) \leqslant 1$ в определении метрики ζ_r, имеем

$$\zeta_1(S_n, N) \leqslant 2\zeta_1(S_n, S_n^*), \quad \zeta_2(S_n, N) \leqslant \frac{\sqrt{2\pi}}{4}\zeta_1(S_n, S_n^*), \quad \zeta_3(S_n, N) \leqslant \frac{1}{3}\zeta_1(S_n, S_n^*). \tag{29}$$

Оценки, содержащие ε_n, получаются применением следствия 1 к (29).

Докажем оптимальность (11). Положим $f(x) := x^3/6$. Тогда $M_3(f) = 1$ и $\mathsf{E}\, f(N) = 0$, поскольку с.в. N симметрична, а функция f – нечетна. Пусть $X_1, X_2, \ldots$ – н.о.р. величины с нулевыми средними и дисперсиями равными 1. Тогда

$$\mathsf{E}\, S_n^3 = \frac{\mathsf{E}\, X_1^3}{\sqrt{n}}, \quad \text{а} \quad \varepsilon_n = \frac{\mathsf{E}\,|X_1|^3}{\sqrt{n}}.$$

Следовательно,

$$\frac{\zeta_3(S_n, N)}{\varepsilon_n} \geqslant |\,\mathsf{E}\, f(S_n) - \mathsf{E}\, f(N)| = \frac{1}{6}\frac{|\,\mathsf{E}\, X_1^3|}{\mathsf{E}\,|X_1|^3}.$$

Остается доказать, что $|\,\mathsf{E}\, X_1^3|/\,\mathsf{E}\,|X_1|^3$ может быть сколь угодно близко к единице. Согласно (25) третий абсолютный момент центрированной с.в. X_1 с дисперсией 1, принимающей два значения $x = -\sqrt{q/p}$, $y = \sqrt{p/q}$ соответственно с вероятностями p и q, равен

$$\mathsf{E}\,|X_1|^3 = q\sqrt{\frac{q}{p}} + p\sqrt{\frac{p}{q}}.$$

нетрудно видеть, что третий момент этой с.в. равен

$$\mathsf{E}\,X_1^3 = -q\sqrt{\frac{q}{p}} + p\sqrt{\frac{p}{q}}.$$

При $p \to 1$ имеем $|\,\mathsf{E}\,X_1|/\,\mathsf{E}\,|X_1|^3 \to 1$. $\square$

Лемма 1 ([23]). *Пусть W – центрированная с.в. с дисперсией 1 и конечным третьим абсолютным моментом β. Обозначим $f(t) := \mathsf{E}e^{itW}, t \in \mathbb{R}$. Тогда при всех $t \in \mathbb{R}$*

$$|f(t)|^2 \leqslant 1 + b(t, \beta + 1), \tag{30}$$

кроме того,

$$|f_{S_n}(t)|^2 \leqslant \exp\left(b(t, \varepsilon_n + \tau_n)\right). \tag{31}$$

Лемма 2. *Функция $b(t, \gamma)$ не убывает по γ при каждом $t \in \mathbb{R}$.*

Доказательство. Проверяется непосредственно вычислением производной. $\square$

Следствие 2. *Для всех $t \in \mathbb{R}$ и $n \in \mathbb{N}$*

$$|f_{S_n}(t)|^2 \leqslant \exp\left(b(t, 2\varepsilon_n)\right). \tag{32}$$

Доказательство. По неравенству Ляпунова имеем $\sigma_j^3 \leqslant \beta_j, j = 1, \dots, n$. Поэтому $\tau_n \leqslant \varepsilon_n$. Теперь (32) следует из (31) и леммы 2. $\square$

Лемма 3. *При выполнении условий леммы* 1

$$|f(t) - \varphi(t)| \leqslant \varphi(t) \int\limits_0^t |f(s) - f^*(s)| s \exp\left(\frac{s^2}{2}\right) ds, \quad t \in \mathbb{R}, \tag{33}$$

где $f^*(t)$ – характеристическая функция с.в. W^*, имеющей распределение W-нулевого смещения.

Доказательство. В силу определения распределения W-нулевого смещения

$$f'(t) = i\,\mathsf{E}\,W\cos(tW) - \mathsf{E}\,W\sin(tW) = -it\,\mathsf{E}\sin(tW^*) - t\,\mathsf{E}\cos(tW^*) = -t\,\mathsf{E}\,e^{itW^*}. \tag{34}$$

Рассмотрим функцию $\psi(t) := \frac{f(t)}{\varphi(t)}$. Заметим, что $\psi(0) = 1$. С учетом (34) Имеем

$$\psi'(t) = \frac{d}{dt}\left(f(t)e^{\frac{1}{2}t^2}\right) = f'(t)e^{\frac{1}{2}t^2} + tf(t)e^{\frac{1}{2}t^2} = \{f(t) - f^*(t)\}te^{\frac{1}{2}t^2}.$$

Тогда

$$\left|\frac{f(t)}{\varphi(t)} - 1\right| = |\psi(t) - \psi(0)| \leqslant \int_0^t |\psi'(s)| ds = \int_0^t |f(s) - f^*(s)| s e^{\frac{1}{2}s^2} ds. \ \square$$

Лемма 4. *Для любых с.в. X и Y имеем*

$$\left|\mathsf{E}\,e^{itX} - \mathsf{E}\,e^{itY}\right| \leqslant t\zeta_1(X, Y).$$

Доказательство. Хорошо известно, что для любых $t, x, y \in \mathbb{R}$ выполнено неравенство

$$\left|e^{itx} - e^{ity}\right| \leqslant t|x - y|.$$

Поэтому для любых $\widetilde{X}, \widetilde{Y}$, определенных на одном вероятностном пространстве, таких что $Law(\widetilde{X}) = Law(X)$ и $Law(\widetilde{Y}) = Law(Y)$, имеем

$$\left|\mathsf{E}\, e^{itX} - \mathsf{E}\, e^{itY}\right| = \left|\mathsf{E}\, e^{it\widetilde{X}} - \mathsf{E}\, e^{it\widetilde{Y}}\right| \leqslant \mathsf{E}\left|e^{it\widetilde{X}} - e^{it\widetilde{Y}}\right| \leqslant t\,\mathsf{E}\,|\widetilde{X} - \widetilde{Y}|. \tag{35}$$

Переходя в (35) к нижней грани по всевозможным $\widetilde{X}, \widetilde{Y}$, получим

$$\left|\mathsf{E}\, e^{itX} - \mathsf{E}\, e^{itY}\right| \leqslant t l_1(X,Y) = t\zeta_1(X,Y).\ \square$$

Доказательство теоремы 5. С учетом (32) неравенство (12) тривиально:

$$|f_{S_n}(t) - \varphi(t)| \leqslant |f_{S_n}(t)| + |\varphi(t)| \leqslant \exp\left(\frac{1}{2} b(t, 2\varepsilon_n)\right) + \varphi(t).$$

Не ограничивая общности, считаем, что $\sigma = 1$. Обозначим $f_j(t) := \mathsf{E}\, e^{itX_j}$, $f_j^*(t) := \mathsf{E}\, e^{itX_j^*}$, $j = 1, \ldots, n$. Положим в (33) $W := S_n$. Воспользовавшись леммой 4 и следствием 1, получим

$$|f(s) - f^*(s)| \leqslant s\zeta_1(S_n, S_n^*) \leqslant \frac{\varepsilon_n s}{2}.$$

Подставив последнее выражение в (33), приходим к (13).

Согласно (28)

$$f^*(s) = \sum_{m=1}^n \sigma_m^2 \mathsf{E} e^{isS'_m} = \sum_{m=1}^n \sigma_m^2 f_m^*(s) \prod_{j \neq m} f_j(s).$$

Следовательно,

$$|f(s) - f^*(s)| = \left| \sum_{i=1}^n \sigma_i^2 \prod_j f_j(s) - \sum_{i=1}^n \sigma_i^2 f_i^*(s) \prod_{j\neq i} f_j(s) \right| =$$

$$= \left| \sum_{i=1}^n \sigma_i^2 \left\{ f_i(s) - f_i^*(s) \right\} \prod_{j \neq i} f_j(s) \right|. \tag{36}$$

По лемме 4 и теореме 3 имеем

$$|f_j(s) - f_j^*(s)| \leqslant s\zeta_1(X_j, X_j^*) = s\sigma_j \zeta_1\left(\frac{X_j}{\sigma_j}, \frac{X_j^*}{\sigma_j}\right) \leqslant \frac{\beta_j s}{2\sigma_j^2}. \tag{37}$$

Из (30) следует, что $|f_j(s)| \leqslant \exp(-\sigma_j^2 s^2/2 + 2\beta_j a|s|^3)$ для всех вещественных s. Поэтому

$$|f(s) - f^*(s)| \leqslant \left| \sum_{i=1}^n \beta_i \frac{s}{2} \prod_{j \neq i} \exp\left(-\frac{\sigma_j^2 s^2}{2} + 2\beta_j a s^3\right) \right| =$$

$$= \left| \sum_{j=1}^n \beta_j \frac{s}{2} \exp\left(\frac{\sigma_j^2 s^2}{2} - 2\beta_j a s^3\right) \right| \exp\left(-\frac{s^2}{2} + 2\varepsilon_n a s^3\right). \tag{38}$$

Поскольку $\sigma_j^3 \leqslant \beta_j \leqslant \varepsilon_n$ для $j = 1, \ldots, n$, имеем для каждого такого j

$$\exp\left(\frac{\sigma_j^2 t^2}{2} - 2\beta_j a t^3\right) \leqslant \exp\left(\frac{\sigma_j^2 t^2}{2} - 2\sigma_j^3 a t^3\right) = \exp\left(\frac{s^2}{2} - 2as^3\right)\Big|_{s=t\sigma_j} \leqslant$$
$$\leqslant \sup\left\{\exp\left(\frac{s^2}{2} - 2as^3\right) : s \in [0, t\varepsilon_n^{1/3}]\right\}. \quad (39)$$

Функция $\exp\left(\frac{s^2}{2} - 2as^3\right)$ возрастает на отрезке $[0, 1/6a]$, в точке $1/6a$ – глобальный максимум равный $1/l$. Поэтому

$$\sup\left\{\exp\left(\frac{s^2}{2} - 2as^3\right) : s \in [0, t\varepsilon_n^{1/3}]\right\} = \begin{cases} \exp\left(\frac{t^2\varepsilon_n^{2/3}}{2} - 2a\varepsilon_n t^3\right) & \text{если } t\varepsilon_n^{1/3} \leqslant 1/6a; \\ 1/l, & \text{иначе.} \end{cases} \quad (40)$$

Объединяя (38), (39) и (40), имеем при $t\varepsilon_n^{1/3} \leqslant 1/6a$

$$|f(s) - f^*(s)| \leqslant \frac{\varepsilon_n s}{2} \exp\left(\frac{s^2 \varepsilon_n^{2/3}}{2} - \frac{s^2}{2}\right),$$

а при $t\varepsilon_n^{1/3} > 1/6a$

$$|f(s) - f^*(s)| \leqslant \frac{\varepsilon_n s}{2l} \exp\left(-\frac{s^2}{2} + 2a\varepsilon_n s^3\right).$$

Подставляя полученные выражения в (33), получим требуемые оценки.

Доказательство теоремы 6. Докажем вначале (15). Пусть $f_1(t) := \mathsf{E}\, e^{itX_1}$. Имеем

$$|f_{S_n}(t) - \varphi(t)| \leqslant |f_{S_n}(t)| + |\varphi(t)| = \left|f_1\left(\frac{t}{\sqrt{n}}\right)\right|^n + |\varphi(t)|. \quad (41)$$

Из леммы 1 следует, что

$$\left|f_1\left(\frac{t}{\sqrt{n}}\right)\right| \leqslant \sqrt{1 + b\left(\frac{t}{\sqrt{n}}, \beta + 1\right)} = \sqrt{1 + \frac{1}{n} b\left(t, \frac{\beta+1}{\sqrt{n}}\right)} = \sqrt{1 + \frac{1}{n} b\,(t, \varepsilon_n + \tau_n)}. \quad (42)$$

Подставив последнее выражение в (41) получим требуемое утверждение. Чтобы установить (17), заметим, что $1 + x \leqslant e^x$ при всех вещественных x. Применив это неравенство к (15) имеем

$$|f_{S_n}(t) - \varphi(t)| \leqslant \exp\left(\frac{1}{2} b(t, \varepsilon_n + \tau_n)\right) + \varphi(t). \quad (43)$$

Остается заметить, что последовательность $(\tau_m)_{m \geqslant 1}$ убывает, откуда следует (17).

Положим $W := S_n$ в (33). Воспользовавшись (36) для с.в. $\frac{1}{\sqrt{n}} X_1, \ldots, \frac{1}{\sqrt{n}} X_n$, получим

$$|f(s) - f^*(s)| = \left|f_1\left(\frac{s}{\sqrt{n}}\right)\right|^{n-1} \cdot \left|f_1\left(\frac{s}{\sqrt{n}}\right) - f_1^*\left(\frac{s}{\sqrt{n}}\right)\right|.$$

Первый из множителей оцениваем при помощи (42), второй – при помощи (37). Имеем

$$|f(s) - f^*(s)| \leqslant \left(1 + \frac{b(s, \varepsilon_n + \tau_n)}{n}\right)^{\frac{n-1}{2}} \frac{\varepsilon_n s}{2}. \quad (44)$$

Подставив полученное выражение в (33), получим (16). Для того, чтобы установить (18), применим к первому множителю, стоящему в правой части (44), неравенство $1+x \leqslant e^x$ и заметим, что $(\tau_m)_{m\geqslant 1}$ убывает. $\square$

Доказательство теоремы 7. Поскольку $\rho(S_n,N) \leqslant 1$, то при $\varepsilon_n \geqslant 1/C$ утверждение теоремы тривиально, поэтому достаточно рассматривать $\varepsilon_n \leqslant 1/0{,}6379$ в общем случае, а для н.о.р. величин $\varepsilon_n \leqslant 1/0{,}5894$. Кроме того, в [24] показано, что при $\varepsilon_n \leqslant 1/7$

$$\rho(S_n,N) \leqslant 0{,}527\varepsilon_n\lambda_n^{3/2},$$

а при $\varepsilon_n \leqslant 0{,}1$

$$\rho(S_n,N) \leqslant 0{,}5151\varepsilon_n\lambda_n^{3/2},$$

где $\lambda_n := \sigma^2(n)/(\sigma^2(n)-\sigma_m)$, $\sigma_m := \max\limits_{j=1,\ldots,n}\sigma_j$. Для н.о.р. слагаемых $\varepsilon_n = \beta_1/(\sigma_1^3\sqrt{n}) \geqslant 1/\sqrt{n}$. Значит, $n \geqslant \lceil 1/\varepsilon_n^2\rceil$. Поэтому при $\varepsilon_n \leqslant 1/7$ имеем $\lambda_n \leqslant 49/48$ и, следовательно,

$$\rho(S_n,N) < 0{,}5436\varepsilon_n.$$

В общем случае без ограничения общности считаем $\sigma=1$, тогда $\sigma_m^3 \leqslant \beta_m \leqslant \sum_{j=1}^n \beta_j = \varepsilon_n$. Таким образом, при $\varepsilon_n \leqslant 0{,}0484$ выполнено

$$\rho(S_n,N) \leqslant 0{,}5151\varepsilon_n\left(\frac{1}{1-\varepsilon_n^{2/3}}\right)^{\frac{3}{2}} < 0{,}5151\varepsilon_n\left(\frac{1}{1-0{,}0484^{2/3}}\right)^{\frac{3}{2}} < 0{,}6379\varepsilon_n.$$

Следовательно, в общем случае можно ограничиться рассмотрением ε_n из отрезка $I_1 = [0{,}0484; 1/0{,}6379]$, а в случае н.о.р. величин – из $I_2 = [1/7; 1/0{,}5894]$. В основе доказательства теоремы для этих значений ε_n лежит неравенство, полученное В. М. Золотаревым и впоследствии улучшенное ван Биком.

Пусть $p(u)$ – непрерывная симметричная функция из $L_1(\mathbb{R})$, такая что ее преобразование Фурье $\widehat{p}(t) := \int_{-\infty}^{\infty} e^{itu}p(u)du \in L_1(\mathbb{R})$,

$$\|p\| := \int\limits_{-\infty}^{\infty}|p(u)|du, \ \ v(x) := \int\limits_0^x \max\{p(u),0\}du, \ \ q(y) := \frac{1}{\sqrt{2\pi}}\int\limits_0^{\infty}\left|\widehat{p}\left(\frac{t}{y}\right)\right|\frac{\delta_n(t)}{t}dt,$$

и существует χ_p – единственный положительный корень уравнения $4v(x) = \|p\|$. Тогда согласно неравенству ван Бика (см., напр., [9])

$$\frac{\rho(S_n,N)}{\varepsilon_n} \leqslant \sqrt{\frac{2}{\pi}}\cdot\frac{v(x)\,x/y+q(y)}{\varepsilon_n(4v(x)-\|p\|)}, \quad \forall\, y>0, x>\chi_p. \tag{45}$$

Обозначим

$$p_1(u) := \frac{\sin u}{2\pi u(1-u^2/\pi^2)}, \quad \widehat{p}_1(t) := \cos^2\left(\frac{\pi t}{2}\right)\mathbf{1}(|t|\leqslant 1),$$

где $\mathbf{1}(\cdot)$ – индикаторная функция. Рассмотрим семейство функций

$$p_{a_1,a_2,a_3}(u) := 0{,}5(p_1(u-a_1)+p_1(u-a_1)+a_3(p_1(u-a_2)+p_1(u-a_2))), \tag{46}$$

которым соответствуют преобразования Фурье

$$\widehat{p}_{a_1,a_2,a_3}(t) := \widehat{p}_1(t)(\cos(a_1t)+a_3\cos(a_2t)), \quad a_1,a_2\in\mathbb{R}, a_3>-1.$$

Зафиксируем некоторое $\varepsilon > 0$. Рассмотрим всевозможные наборы с.в. $X_1, \dots, X_n$, $n \in \mathbb{N}$, такие что $\varepsilon_n = \varepsilon$. Из (45) вытекает, что наименьшая верхняя грань $\rho(S_n, N)/\varepsilon_n$ по всем указанным наборам не превосходит при любых значениях параметров x, y, a_1, a_2, a_3 величины $D(\varepsilon, x, y, a_1, a_2, a_3)$, которая получается из правой части (45), если величину $\delta_n(t)$, фигурирующую в определении $q(y)$, заменить ее оценкой $\Delta(\varepsilon, t) := \min\{\delta^{(1)}(\varepsilon, t), \delta^{(2)}(\varepsilon, t), \delta^{(3)}(\varepsilon, t)\}$ и подставить $p := p_{a_1,a_2,a_3}$ из (46). Для $\varepsilon \in I_1$ будем подбирать значения x, y, a_1, a_2, a_3 так, чтобы $D(\varepsilon, x, y, a_1, a_2, a_3)$ оказалась как можно меньше, после чего оценка для C получится взятием верхней грани по всем $\varepsilon \in I_1$. Задачу подбора x, y, a_1, a_2, a_3 можно решать численно при помощи компьютера. В данной работе для этой цели был использован метод сопряженных градиентов, соответствующая процедура взята из библиотеки GNU Scientific Library версии 1.11. Вычисления были проведены для 440 значений ε, для остальных использовано неравенство

$$D(\varepsilon^{(1)}, x, y, a_1, a_2, a_3) \leqslant \frac{\varepsilon^{(2)}}{\varepsilon^{(1)}} D(\varepsilon^{(2)}, x, y, a_1, a_2, a_3), \ \ \varepsilon^{(1)} < \varepsilon^{(2)}, \tag{47}$$

которое следует из того, что каждая из функций $\delta^{(1)}(\varepsilon, t), \delta^{(2)}(\varepsilon, t), \delta^{(3)}(\varepsilon, t)$ не убывает по ε. Экстремальное значение $D(\varepsilon, x, y, a_1, a_2, a_3) = 0{,}637733$ возникло при

$$\varepsilon = 0{,}4767, \ \ x = 7{,}5916, \ \ y = 17{,}9896, \ \ a_1 = 2{,}2371, \ \ a_2 = 5{,}3685, \ \ a_3 = 0{,}5062.$$

В случае н.о.р. с.в. использование неравномерных по n оценок позволило добиться лучших результатов. Обозначим $J(\varepsilon, n)$ наименьшую верхнюю грань $\rho(S_n, N)/\varepsilon_n$ по всем наборам н.о.р. с.в. $X_1, \dots, X_n$ с $\varepsilon_n = \varepsilon$. Тогда $C = \sup\{J(\varepsilon, n) : \varepsilon > 0, n \in \mathbb{N}\}$. Достаточно показать, что $J(\varepsilon, n) \leqslant 0{,}5894$ при $\varepsilon \in I_2, n \geqslant n_0(\varepsilon) := \lceil 1/\varepsilon^2 \rceil$. Зафиксируем ε, после чего для нескольких первых значений $n = n_0, n_0+1, \dots, n_0+k$ получим оценку для $J(\varepsilon, n)$ следующим образом. Заменим величину $\delta_n(t)$, фигурирующую в определении $q(y)$, оценкой $\Delta_1(n, \varepsilon, t) := \min\{\delta^{(4)}(n, \varepsilon, t), \delta^{(5)}(n, \varepsilon, t)\}$ и возьмем $p = p_{a_1,a_2,a_3}$, после чего в правой части (45) возникнет функция $J_1(\varepsilon, n, x, y, a_1, a_2, a_3)$, при каждом допустимом наборе параметров x, y, a_1, a_2, a_3, представляющая собой оценку величины $J(\varepsilon, n)$. Как и ранее, при помощи компьютера подберем x, y, a_1, a_2, a_3 так, чтобы $J_1(\varepsilon, n, x, y, a_1, a_2, a_3)$ оказалась как можно меньше.

Для $n > n_0 + k$ величины $J(\varepsilon, n)$ оценим равномерно. Вместо $\delta_n(t)$ подставим $\Delta_2(n, \varepsilon, t) := \min\{\delta^{(6)}(n_0+k+1, \varepsilon, t), \delta^{(5)}(n_0+k+1, \varepsilon, t)\}$. В правой части (45) возникнет функция $J_2(\varepsilon, x, y, a_1, a_2, a_3) \geqslant \sup\limits_{n > n_0+k} J(\varepsilon, n)$. Подобрав должным образом x, y, a_1, a_2, a_3, найдем искомую оценку. Обозначим $J(\varepsilon)$ максимум из $k{+}2$ полученных оценок. $J(\varepsilon)$ оценивает сверху величину $\sup\limits_n J(\varepsilon, n)$. Поскольку при каждом фиксированном n функции $\delta^{(k)}(n, \varepsilon, t)$, $k = 4, 5, 6, 7$, не убывают по ε, то $J(\varepsilon)$ обладает аналогичным (47) свойством, а именно,

$$J(\varepsilon^{(1)}) \leqslant \frac{\varepsilon^{(2)}}{\varepsilon^{(1)}} J(\varepsilon^{(2)}), \ \ \varepsilon^{(1)} < \varepsilon^{(2)},$$

что позволило проводить вычисления лишь для конечного числа значений ε.

Для случая н.о.р. величин экстремальное значение $J(\varepsilon) = J_1(\varepsilon, n, x, y, a_1, a_2, a_3) = 0{,}589327$ было достигнуто при $\varepsilon = 0{,}14286, n = 49$, при этом

$$x = 3{,}8570, \ \ y = 28{,}6444, \ \ a_1 = 1{,}8827, \ \ a_2 = 2{,}8276, \ \ a_3 = -0{,}4959.$$

Итак, утверждение теоремы справедливо во всех рассмотренных случаях. $\square$

Автор считает своим приятным долгом выразить искреннюю признательность профессору А. В. Булинскому за внимание к работе и ценные указания.

Список литературы

[1] A. C. Berry (1941), The accuracy of the Gaussian approximation to the sum of independent variates. Trans. Amer. Math. Soc., vol. 49, N 1, p. 122-139.

[2] C.-G. Esseen (1942), On the Liapunoff limit of error in the theory of probability. Ark. Mat. Astron. Fys., vol. A28, N 9, p. 1-19.

[3] H. Bergström (1949), On the central limit theorem in the case of not equally distributed random variables. Skand. Aktuarietidskr., vol. 33, p. 37-62.

[4] K. Takano (1951), A remark to a result of A. C. Berry. Res. Mem. Inst. Math., vol. 9, N 6, p. 4.08-4.15.

[5] В. М. Золотарев (1965), О близости распределений двух сумм независимых случайных величин, Теор. вероятн. примен., т. 10, вып. 3, с. 519-526.

[6] В. М. Золотарев (1966), Абсолютная оценка остаточного члена в центральной предельной теореме. Теор. вероятн. примен., т. 11, вып. 1, с. 108-119.

[7] V. M. Zolotarev (1967), A sharpening of the inequality of Berry-Esseen. Z. Wahrsch. Verw. Geb., vol. 8, p. 332-342.

[8] В. М. Золотарев (1967), Некоторые неравенства теории вероятностей и их применение к уточнению теоремы А. М. Ляпунова. Докл. Акад. наук СССР, т. 177, вып. 3, с. 501-504.

[9] P. van Beek (1972), An application of Fourier methods to the problem of sharpening the Berry-Esseen inequality. Z. Wahrsh. Verw. Geb., vol. 23, p. 187-196.

[10] И. С. Шиганов (1982), Об уточнении верхней константы в остаточном члене центральной предельной теоремы. Проблемы устойчивости стохастических моделей. Труды ВНИИСИ, с. 109-115.

[11] И. Г. Шевцова (2006), Уточнение верхней оценки абсолютной постоянной в неравенстве Берри-Эссеена. Теор. вероятн. примен., т. 51, вып. 3, с. 622-626.

[12] C.-G. Esseen (1956), A moment inequality with an application to the central limit theorem, Scand. Aktuarietidskr., vol. 39, p. 160-170.

[13] Б. А. Рогозин (1960). Одно замечание к работе Эссеена "Моментное неравенство с применением к центральной предельной теореме". Теор. вероятн. примен., т. 5, вып. 1, с. 125-128.

[14] Г. П. Чистяков (2001) Новое асимптотическое разложение и асимптотически наилучшие постоянные в теореме Ляпунова. I. Теор. вероятн. примен., т. 46, вып. 2, с. 326-344.

[15] Г. П. Чистяков (2001) Новое асимптотическое разложение и асимптотически наилучшие постоянные в теореме Ляпунова. II. Теор. вероятн. примен., т. 46, вып. 3, с. 573-579.

[16] Г. П. Чистяков (2002) Новое асимптотическое разложение и асимптотически наилучшие постоянные в теореме Ляпунова. III. Теор. вероятн. примен., т. 47, вып. 3, с. 475-497.

[17] В. Ю. Королев, И. Г. Шевцова (2005). О точности нормальной аппроксимации. I. Теор. вероятн. примен., т. 50, вып. 2, с. 353-366.

[18] В. Ю. Королев, И. Г. Шевцова (2005). О точности нормальной аппроксимации. II. Теор. вероятн. примен., т. 50, вып. 3, с. 555-564.

[19] L. Goldstein, G. Reinert (1997), Stein's method and the zero bias transformation with application to simple random sampling. Ann. Appl. Prob., vol. 7, N 4, p. 935-952.

[20] M. Raic (2003), Normal approximation by Stein's method. Proceedings of the Seventh Young Statisticians Meeting. Metodoloski zvezki, vol. 21, p. 71-97.

[21] W. Hoeffding (1955), The extrema of the expected value of a function of independent random variables. Ann. Math. Statist., vol. 26, N 2, p. 268-275.

[22] Золотарев В. М. Современная теория суммирования независимых случайных величин. М.: Наука, 1986.

[23] H. Prawitz (1974), Weitere Ungleichungen fur den absoluten Betrag einer charakteristischen Function. Scand. Actuarial J., p. 21-28.

[24] H. Prawitz (1975), On the remainder in the central limit theorem. Scand. Actuarial J., p. 145-156.