

УДК 519.521

Метод Стейна и характеристические функции

А. Н. Тихомиров

В работе дается обзор различных приложений метода описания аппроксимирующих распределений с помощью дифференциальных уравнений для характеристических функций и, в частности, применений этого описания к оценке близости распределений. Эта идея впервые была предложена автором в 1976 г. В дальнейшем такой подход, в последнее время называемый некоторыми авторами методом Стейна–Тихомирова (см., например, работы П. Айхельсбахера, В. Редносса, Й. Сунклодаса, Ш. К. Форманова), получил широкое развитие.

Библиография: 78 названий.

Ключевые слова: метод Стейна, метод Стейна–Тихомирова, слабая зависимость, нормальная аппроксимация.

DOI: <https://doi.org/10.4213/rm10242>

СОДЕРЖАНИЕ

1. Введение.....	122
2. Краткое описание метода Стейна	125
2.1. Аппроксимация нормальным законом распределения	126
2.2. Оценка точности аппроксимации сумм индикаторных функций распределением Пуассона в метрике полной вариации.....	131
3. Обзор литературы по методу Стейна	133
4. Характеристические функции и метод Стейна.....	137
5. Аппроксимация нормальным законом.....	137
5.1. Квадратичные формы.....	142
5.2. Аппроксимация вероятности попадания в шар суммы гильбертовозначных случайных величин.....	147
5.3. Применение метода к оценке точности аппроксимации распределения сумм мультииндексированных случайных величин. Случайные поля	151

Разделы 1, 3, 4, 5 и 7 настоящей работы выполнены в рамках проекта Коми научного центра УрО РАН № 125031203621-2 “Асимптотические и неасимптотические проблемы в математических моделях сложных систем, включая модели большой размерности”, а разделы 2, 6 выполнены в рамках Программы фундаментальных исследований НИУ ВШЭ.

6. Развитие метода характеристических функций	153
6.1. Вариант А. Роллина	153
6.2. Вариант П. Айхельсбахера и Б. Редносса	156
6.3. Подход Ш. К. Форманова	157
6.4. Новый подход к описанию метода Стейна–Тихомирова	158
6.5. Оценка скорости сходимости к сопровождающему распределению Пуассона	159
7. Подход В. Бенткуса и применение этого подхода к случайным матрицам	161
Список литературы	168

1. Введение

Л. Чен (L. H. Y. Chen) в своей статье [24] пишет, что метод Стейна родился в середине шестидесятых годов прошлого столетия, за несколько лет до выхода статьи Стейна 1972 г. [68]. Однако для большинства исследователей этот метод стал известен именно благодаря упомянутой работе. В 1960-е годы одной из самых популярных тем в теории вероятностей были предельные теоремы для слабо зависимых величин. В качестве условий слабой зависимости рассматривались различные условия перемешивания, в частности условие m -зависимости, условие сильного перемешивания или условие равномерно сильного перемешивания (см., например, [51]). Одним из основных методов доказательства предельных теорем для слабо зависимых величин являлся метод Бернштейна, или метод секционирования. Суть этого метода проста. Допустим, нам надо аппроксимировать распределение случайной величины $\Psi_n = \Psi(X_1, X_2, \dots, X_n)$, которая является функцией от последовательности случайных величин $X_1, X_2, \dots$, удовлетворяющей одному из условий слабой зависимости. Представим последовательность случайных величин в виде длинных и коротких серий

$$Y_1 = (X_1, \dots, X_l), \quad Y_2 = (X_{l+1}, \dots, X_{l+m}), \quad Y_3 = (X_{m+l+1}, \dots, X_{2l+m}), \quad \dots$$

“Расстояние” между $Y_1, Y_3, \dots$ равно $m+1$. Это означает, что можно выбрать m достаточно большим, чтобы эти величины были “почти” независимы (например, в случае m -зависимых последовательностей они будут просто независимыми). С другой стороны, мы можем выбрать l достаточно большим, а m достаточно малым по сравнению с l так, чтобы вклад коротких серий в предельное распределение Ψ_n был пренебрежимо мал. Иными словами, мы можем аппроксимировать распределение случайных величин Ψ_n распределением случайных величин $\tilde{\Psi}_n = \Psi_n(Y_1, Y_3, \dots)$. В свою очередь, распределение случайных величин $\tilde{\Psi}_n$ мы аппроксимируем распределением случайных величин $\hat{\Psi} = \Psi(\hat{Y}_1, \hat{Y}_3, \dots)$, где $\hat{Y}_1, \hat{Y}_3, \dots$ – независимые копии $Y_1, Y_3, \dots$. С применением метода Бернштейна были получены фундаментальные результаты о применимости центральной предельной теоремы для слабо зависимых величин (см., например, [51]). Однако оценок скорости сходимости в центральной предельной теореме для случайных величин с перемешиванием (как, впрочем, и для m -зависимых величин), близких к оценкам для независимых величин,

методом Бернштейна получить не удавалось. Наилучший результат, известный к 1969 г., – это результат В. Филиппа (W. Philipp) [59]. Им была получена оценка остаточного члена в центральной предельной теореме для сумм случайных последовательностей, удовлетворяющих условию перемешивания по Ибрагимову с экспоненциально быстро убывающим коэффициентом, в метрике Колмогорова, а именно оценка порядка $O(n^{-1/4})$.

В 1972 г. вышла статья Ч. Стейна (Ch. Stein) “A bound for the error in the normal approximation to the distribution of a sum of dependent random variables” [68]. В этой статье впервые была получена оптимальная по порядку зависимости от числа слагаемых n (с точностью до логарифмического множителя, т.е. отличающаяся от нижней оценки на логарифмический множитель) верхняя оценка скорости сходимости в центральной предельной теореме для случайных величин, удовлетворяющих условию равномерно сильного перемешивания с экспоненциально быстро убывающим коэффициентом равномерно сильного перемешивания, порядка $O(n^{-1/2} \log^8 n)$. Главный результат упомянутой статьи – принципиально новый подход к оценке близости к нормальному распределению. Этот подход известен теперь как метод Стейна. Методу Стейна посвящена обширная литература, включая обзорные статьи и монографии. Мы приведем здесь лишь небольшую часть из огромного списка работ, в той или иной мере посвященных методу Стейна. Популярность метода Стейна связана с его простотой и универсальностью. Если работа Стейна 1972 г., о которой говорилось выше, была посвящена аппроксимации нормальным распределением, то уже через три года была опубликована статья Л. Чена об аппроксимации распределением Пуассона [22]. В дальнейшем список аппроксимирующих распределений был расширен, и сейчас он включает показательное распределение, биномиальное распределение и многие, многие другие.

В настоящем обзоре мы дадим краткое описание основных идей метода Стейна и приведем примеры его применения, иллюстрирующие эти идеи. В разделе 3 мы дадим краткий обзор основной литературы по методу Стейна. Однако ключевой упор в обзоре сделан на симбиоз идей Стейна и метода характеристических функций. Этот подход был впервые предложен автором в 1976 г. в работе [73]. Дело в том, что Стейн и большинство его последователей не задействовали аппарат преобразования Фурье, а применяли так называемые прямые методы, т.е. использовали непосредственно характеризацию аппроксимирующих распределений (плотностей). Более того, в работе 1972 г. Стейн писал: “The abstract normal approximation theorem of Section 2 is elementary in the sense that it uses only the basic properties of conditional expectation and the elements of analysis, including the solution of a first order linear differential equation. It is also direct, in the sense that the expectation of a fairly arbitrary function of the random variable W in question is approximated directly without going through the characteristic function. Because of the clumsiness of the technique used at this point, it is likely that slightly better results could be obtained by using the basic identity, Lemma 2.1, to approximate the characteristic function of W and, by standard procedures, to go from this to an approximation for the distribution of W . However, I believe that, in the long run, better results will be obtained by direct methods”.

Стейн отдавал предпочтение прямым методам и высказывал предположение, что прямые методы дадут лучшие результаты, нежели характеристические функции. Однако, как показывает время, во многих задачах использование аппарата характеристических функций имеет преимущество перед прямыми методами. Обсуждению вопросов применения характеристических функций в методе Стейна посвящены разделы 4–7. В разделе 5 мы рассмотрим задачи нормальной аппроксимации распределений сумм случайных величин, оценки скорости сходимости в центральной предельной теореме для сумм случайных величин (как зависимых, так и независимых). В частности, в п. 5.1 выводится оценка скорости сходимости в центральной предельной теореме для квадратичных форм от независимых случайных величин. В п. 5.2 рассматривается задача оценки точности аппроксимации вероятности попадания в шар сумм гильбертовозначных случайных величин. Пункт 5.3 посвящен применениям метода характеристических функций к случайным полям. Наконец, в пп. 6.1–6.4 и в разделе 7 обсуждаются варианты развития метода характеристических функций различными авторами. В п. 6.5 рассмотрено применение характеристических функций к задаче аппроксимации сопровождающим пуассоновским распределением.

Список обозначений. Греческими буквами ξ, η, ζ мы, как правило, будем обозначать случайные величины. Также для обозначения случайных величин будем использовать прописные буквы латинского алфавита X, Y, W, Z .

Прописными латинским буквами (с нижним индексом или без него) будем обозначать функции распределения, $F(x), G(x)$. Прописные латинские буквы **A, B, C**..., выделенные жирным шрифтом, используются для обозначения операторов и матриц. Символ $\mathbb{Z}$ будет обозначать множество целых чисел. Символами $\mathbb{R}$ и $\mathbb{C}$ обозначаем вещественную прямую и комплексную плоскость соответственно. Символами $\mathbb{R}^q$ и $\mathbb{C}^q$ обозначаем q -мерные вещественное и комплексное векторные пространства. Строчными латинскими буквами, выделенными жирным шрифтом, будем обозначать векторы в $\mathbb{R}^q, \mathbb{C}^q$ для $q > 1$.

Символом $\varphi_{\text{нор}}(x)$ будем обозначать плотность стандартного нормального закона:

$$\varphi_{\text{нор}}(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\}.$$

Символом $\Phi_{\text{нор}}(x)$ обозначается функция распределения стандартного нормального закона:

$$\Phi_{\text{нор}}(x) = \int_{-\infty}^x \varphi_{\text{нор}}(y) dy.$$

Далее $f_{\text{нор}}(t)$ будет обозначать характеристическую функцию стандартного нормального закона:

$$f_{\text{нор}}(t) = \exp\left\{-\frac{t^2}{2}\right\}.$$

Символом $\mathcal{W}_1(W, z)$ обозначим расстояние Канторовича–Вассерштейна между случайными величинами W и Z , а символом $\rho(F, G)$ – расстояние Колмогорова между функциями распределения F и G . Пусть $d_{\text{TV}}(F, G)$ обозначает

расстояние полной вариации между функциями распределения F и G . Определения см. ниже, в начале раздела 2. Прописными латинскими буквами, выделенными каллиграфическим шрифтом, $\mathcal{F}, \mathcal{H}, \dots$ будем обозначать множества (классы) функций с определенными свойствами. Символом $\theta(t)$ с индексом или без него будем обозначать произвольную функцию такую, что $|\theta(t)| \leq 1$. Буквами C, c будем обозначать абсолютные постоянные, значения которых не обязательно одни и те же при разных вхождениях.

2. Краткое описание метода Стейна

Основные идеи метода Стейна можно представить следующим образом. Предположим, что мы оцениваем близость функции распределения случайной величины W , $F(x) = \Pr\{W < x\}$, к функции распределения случайной величины Z , $G(x) = \Pr\{Z < x\}$, в некоторой метрике

$$d(W, Z) = \sup_{h \in \mathcal{H}} |Eh(W) - Eh(Z)|,$$

где $\mathcal{H}$ – некоторое пространство функций, определяющее метрику. Наиболее часто рассматривают следующие метрики:

- метрика Канторовича–Вассерштейна

$$\mathcal{W}_1(W, Z) = \sup_{h \in \mathcal{H}} |Eh(Z) - Eh(W)|,$$

где $\mathcal{H}$ – пространство 1-липшицевых функций, т. е.

$$\mathcal{H} = \{h: |h(x) - h(y)| \leq |x - y|\};$$

- метрика Колмогорова

$$\rho(W, Z) = \sup_{h \in \mathcal{H}} |Eh(Z) - Eh(W)|,$$

где $\mathcal{H}$ – пространство индикаторов интервалов $(-\infty, y)$, $y \in \mathbb{R}$, т. е.

$$\mathcal{H} = \left\{ h_y(x), y \in \mathbb{R}: h_y(x) = \begin{cases} 1, & \text{если } x < y \\ 0 & \text{в противном случае} \end{cases} \right\};$$

- метрика полной вариации

$$d_{\text{TV}}(W, Z) = \sup_{h \in \mathcal{H}} |Eh(W) - Eh(Z)|,$$

где $\mathcal{H}$ – пространство индикаторов всех событий.

Пусть выбрана метрика с соответствующим пространством $\mathcal{H}$. Для случайной величины Z определим оператор $\mathbf{A}$ и пространство функций $\mathcal{F} = \mathcal{F}(\mathbf{A})$ такое, что случайная величина Z имеет функцию распределения $G(x)$ тогда и только тогда, когда $\mathbf{E}\mathbf{A}f(Z) = 0$ для всех $f \in \mathcal{F}$. Далее, для каждой функции $h \in \mathcal{H}$ найдем такую функцию $f_h \in \mathcal{F}$, что

$$\mathbf{A}f_h(x) = h(x) - Eh(Z). \quad (1)$$

Это равенство называют уравнением Стейна. Полагая $x = W$ и беря математическое ожидание от обеих частей уравнения Стейна, мы получим

$$\mathbf{E}\mathbf{A}f(W) = \mathbf{E}h(W) - \mathbf{E}h(Z).$$

Введем обозначение

$$\mathcal{F}(\mathcal{H}) = \{f \in \mathcal{F} : \text{существует } h \in \mathcal{H} \text{ такая, что } f = f_h\}.$$

Нетрудно видеть, что в наших обозначениях

$$d(F, G) = d(X, Z) \leq \sup_{f \in \mathcal{F}(\mathcal{H})} |\mathbf{E}\mathbf{A}f(W)|.$$

Таким образом, задача оценки расстояния $d(F, G)$ между функциями распределения $F(x)$ и $G(x)$ сводится к вычислению $\mathbf{E}\mathbf{A}f(W)$.

2.1. Аппроксимация нормальным законом распределения. В этом пункте мы рассмотрим применение метода Стейна к задаче аппроксимации нормальным законом распределения. Прежде всего дадим несколько общих определений.

ОПРЕДЕЛЕНИЕ 1. Пусть $p(x)$ – плотность распределения вероятностей с носителем $\text{supp}\{p\} = (a, b)$. *Классом Стейна* называется множество функций $\mathcal{F}_p$ таких, что

- (а) отображение $x \rightarrow f(x)p(x)$ дифференцируемо;
- (б) имеют место равенства

$$\lim_{x \rightarrow a+} f(x)p(x) = \lim_{x \rightarrow b-} f(x)p(x) = 0.$$

ОПРЕДЕЛЕНИЕ 2. *Оператор Стейна* $\mathbf{A}_p$ определяется соотношением

$$f \rightarrow \mathbf{A}_p f = \begin{cases} \frac{(fp)'}{p}, & \text{если } x \in (a, b); \\ 0, & \text{если } x \notin (a, b). \end{cases}$$

Мы можем записать оператор $\mathbf{A}_p$ в виде

$$\mathbf{A}_p f(x) = f'(x) + \frac{p'(x)}{p(x)} f(x). \quad (2)$$

Определим класс Стейна и оператор Стейна для нормального закона. В качестве класса Стейна $\mathcal{F}_p$ рассмотрим множество непрерывно дифференцируемых ограниченных функций. Нетрудно видеть, что для любого $f \in \mathcal{F}_{\varphi_{\text{nor}}}$ функция $f\varphi_{\text{nor}}$ дифференцируема и

$$\lim_{x \rightarrow -\infty} f(x)\varphi_{\text{nor}}(x) = \lim_{x \rightarrow \infty} f(x)\varphi_{\text{nor}}(x) = 0.$$

В силу равенства (2) имеем

$$\mathbf{A}_{\varphi_{\text{nor}}} f(x) = f'(x) - xf(x). \quad (3)$$

Уравнение Стейна (1) примет вид

$$f'(x) - xf(x) = h(x) - \mathbf{E}h(Z), \quad (4)$$

где $h \in \mathcal{H}$.

Решая уравнение Стейна относительно функции f , мы получим

$$f_h(x) = \exp\left\{\frac{x^2}{2}\right\} \int_{-\infty}^x \exp\left\{-\frac{u^2}{2}\right\} (h(u) - \mathbf{E}h(Z)) du. \quad (5)$$

Заметим, что

$$f_h(x)\varphi_{\text{nor}}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{u^2}{2}\right\} (h(u) - \mathbf{E}h(Z)) du.$$

Очевидно, что

$$\lim_{x \rightarrow -\infty} f_h(x)\varphi_{\text{nor}}(x) = 0. \quad (6)$$

Для функций $h \in \mathcal{H}$ справедливо равенство

$$\mathbf{E}h(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\{-u^2\} h(u) du.$$

Нетрудно видеть, что

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{u^2}{2}\right\} (h(x) - \mathbf{E}h(Z)) du = \mathbf{E}h(Z) - \mathbf{E}h(Z) = 0. \quad (7)$$

Следовательно,

$$\lim_{x \rightarrow \infty} f_h(x)\varphi_{\text{nor}}(x) = 0.$$

Отсюда следует, что функция $f_h(x)$ принадлежит классу Стейна и

$$\sup_x |F(x) - \Phi(x)| \leqslant \sup_{h \in \mathcal{H}} |\mathbf{E}[f'_h(W) - Wf_h(W)]|, \quad (8)$$

где, напомним, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{u^2}{2}\right\} du$.

Можно выделить две части метода Стейна. Первая – это выбор оператора Стейна и класса Стейна, сведение задачи оценки близости функций распределения к неравенству Стейна (8). Вторая – это вычисление $\mathbf{E}\mathbf{A}f_h(W)$. Вторая часть зависит от структуры случайной величины W .

Рассмотрим в качестве примера классическую задачу оценки точности нормальной аппроксимации функции распределения нормированной суммы независимых одинаково распределенных случайных величин.

Оценка точности аппроксимации функции распределения суммы независимых одинаково распределенных случайных величин нормальным законом в метрике Канторовича–Вассерштейна. Предположим, что $X_1, X_2, \dots$ – последовательность независимых одинаково распределенных случайных величин с $\mathbf{E}X_1 = 0$ и $\mathbf{E}X_1^2 = 1$. Положим

$$W = \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j.$$

Мы оценим близость функции распределения $F_n(x) = \Pr\{W < x\}$ к $\Phi(x)$ в метрике Канторовича–Вассерштейна. Пространство $\mathcal{H}$ – пространство 1-липшицевых функций.

ТЕОРЕМА 1. Пусть

$$\mathbb{E}|X_1|^3 = \beta_3 < \infty.$$

Тогда существует постоянная $C > 0$ такая, что

$$\mathcal{W}_1(W, Z) \leq C \frac{\beta_3}{\sqrt{n}}.$$

Константа C может быть выбрана равной

$$C = 14 + 4 \sqrt{\frac{e\pi}{2}}.$$

ДОКАЗАТЕЛЬСТВО. Начнем с двух вспомогательных утверждений.

ЛЕММА 1. Существует постоянная $C_1 > 0$ такая, что для любой функции $h \in \mathcal{H}$

$$|f_h(x)| \leq C_1.$$

Можно положить $C_1 = 2 + \sqrt{e\pi/2}$.

ДОКАЗАТЕЛЬСТВО ЛЕММЫ 1. Не умаляя общности, можно считать, что $h(0) = 0$ для любого $h \in \mathcal{H}$. Поскольку $h(x)$ – это 1-липшицева функция, то

$$|h(x)| \leq |x|. \quad (9)$$

Заметим, что для любой функции $h \in \mathcal{H}$

$$\int_{-\infty}^{\infty} \exp\left\{-\frac{u^2}{2}\right\} (h(u) - \mathbb{E}h(Z)) du = 0. \quad (10)$$

Рассмотрим случай $x > 0$. В силу последнего равенства мы можем записать $f(x)$ в виде

$$f(x) = -\exp\left\{\frac{x^2}{2}\right\} \int_x^{\infty} (h(u) - \mathbb{E}h(Z)) \exp\left\{-\frac{u^2}{2}\right\} du. \quad (11)$$

Поскольку Z – стандартная нормально распределенная случайная величина, имеем $\mathbb{E}|h(Z)| \leq \mathbb{E}|Z| \leq 1$. В силу этого выполняется неравенство

$$|f(x)| \leq \exp\left\{\frac{x^2}{2}\right\} \int_x^{\infty} (u+1) \exp\{-u^2\} du. \quad (12)$$

Далее,

$$\exp\left\{\frac{x^2}{2}\right\} \int_x^{\infty} \exp\left\{-\frac{u^2}{2}\right\} du \leq \begin{cases} \sqrt{\frac{\pi e}{2}}, & \text{если } x \leq 1, \\ 1, & \text{если } x \geq 1. \end{cases} \quad (13)$$

Нетрудно видеть, что

$$\exp\left\{\frac{x^2}{2}\right\} \int_x^\infty u \exp\left\{-\frac{u^2}{2}\right\} du \leq 1. \quad (14)$$

Объединяя неравенства (12)–(14), получим

$$|f(x)| \leq 1 + \sqrt{\frac{\pi e}{2}}. \quad (15)$$

Для $x < 0$ воспользуемся равенством

$$f(x) = -\exp\left\{\frac{x^2}{2}\right\} \int_{-x}^\infty [h(-u) - \mathbb{E}h(Z)] \exp\left\{-\frac{u^2}{2}\right\} du,$$

откуда следует оценка (15). Лемма доказана.

ЛЕММА 2. *Существует постоянная $C_2 > 0$ такая, что для всех $h \in \mathcal{H}$ и для всех x имеет место неравенство*

$$|f'_h(x)| \leq C_2.$$

Константу C_2 можно выбрать равной

$$C_2 = 4 + \sqrt{\frac{e\pi}{2}}.$$

ЗАМЕЧАНИЕ 1. Производные ν -го порядка функция f_h зависят от свойств производной $h^{(\nu-1)}$ функции h порядка $\nu - 1$. Поскольку мы предполагаем лишь липшицевость функции h , то уместно говорить только о производной первого порядка.

ДОКАЗАТЕЛЬСТВО ЛЕММЫ 2. Воспользуемся равенством (4). Для $x > 0$ мы имеем представление

$$\begin{aligned} f'_h(x) &= -x \exp\left\{\frac{x^2}{2}\right\} \int_x^\infty [h(u) - h(Z)] \exp\left\{-\frac{u^2}{2}\right\} du + h(x) - h(Z) \\ &= -x \exp\left\{\frac{x^2}{2}\right\} \int_x^\infty [h(u) - h(x)] \exp\left\{-\frac{u^2}{2}\right\} du \\ &\quad + [h(x) - \mathbb{E}h(Z)] \left(1 - x \exp\left\{\frac{x^2}{2}\right\} \int_x^\infty \exp\left\{-\frac{u^2}{2}\right\} du\right). \end{aligned} \quad (16)$$

Используя свойство 1-липшицевости функции h , мы получим

$$\begin{aligned} |f'_h(x)| &\leq x \exp\left\{\frac{x^2}{2}\right\} \int_x^\infty (u - x) \exp\left\{-\frac{u^2}{2}\right\} du \\ &\quad + (x + 1) \left|1 - x \exp\left\{\frac{x^2}{2}\right\} \int_x^\infty \exp\left\{-\frac{u^2}{2}\right\} du\right|. \end{aligned} \quad (17)$$

Нетрудно видеть, что

$$\begin{aligned} x \exp\left\{\frac{x^2}{2}\right\} \int_x^\infty (u-x) \exp\left\{-\frac{u^2}{2}\right\} du \\ \leq \int_x^\infty x(u-x) \exp\{-x(u-x)\} du \leq \frac{1}{x} \end{aligned} \quad (18)$$

и

$$\begin{aligned} 0 \leq 1 - x \exp\left\{\frac{x^2}{2}\right\} \int_x^\infty \exp\left\{-\frac{u^2}{2}\right\} du \\ = \exp\left\{\frac{x^2}{2}\right\} \int_x^\infty (u-x) \exp\left\{-\frac{u^2}{2}\right\} du \leq \frac{1}{x}. \end{aligned} \quad (19)$$

Объединяя последние два неравенства и справедливое при $x > 1$ неравенство

$$|f'_h(x)| \leq 2, \quad (20)$$

получим, что для $0 < x \leq 1$

$$|f'_h(x)| \leq C_1 + 2.$$

В случае $x < 0$ доказательство получается заменой x на $-x$. Лемма доказана.

Теперь мы можем перейти к вычислению $\mathbf{E}\mathbf{A}f_h(W)$. Для упрощения записи мы будем опускать в дальнейшем индекс h в обозначении f_h и писать просто f . Введем в рассмотрение случайную величину

$$W^{(j)} = \frac{1}{\sqrt{n}} \sum_{k \neq j} X_k.$$

Заметим, что $W - W^{(j)} = X_j/\sqrt{n}$, а X_j и $W^{(j)}$ независимы. Кроме того, мы будем использовать равенство

$$\begin{aligned} f(W) - f(W^{(j)}) &= \int_0^1 X_j f'(W^{(j)} + \tau X_j) d\tau \\ &= X_j f'(W^{(j)}) + \int_0^1 X_j (f'(W^{(j)} + \tau X_j) - f'(W^{(j)})) d\tau. \end{aligned} \quad (21)$$

По определению оператора Стейна имеем

$$\mathbf{E}\mathbf{A}f(W) = \mathbf{E}[f'(W) - Wf(W)] = \mathbf{E}f'(W) - \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbf{E}X_j f(W). \quad (22)$$

Используя равенство (21), мы получим

$$\begin{aligned} \mathbf{E}\mathbf{A}f(W) &= \frac{1}{n} \sum_{j=1}^n \mathbf{E}|X_j|^2 \mathbf{E}(f'(W) - f'(W^{(j)})) \\ &\quad + \frac{1}{n} \sum_{j=1}^n \mathbf{E}X_j^2 \int_0^1 (f'(W^{(j)} + \tau X_j) - f'(W^{(j)})) d\tau \end{aligned} \quad (23)$$

(в первой сумме мы воспользовались тем, что $E|X_1|^2 = 1$). В силу равенства (1) имеем

$$|f'(x) - f'(y)| \leq |x - y| |f(x)| + |y| |f(x) - f(y)| + |h(x) - h(y)|. \quad (24)$$

Применяя леммы 1 и 2, приходим к неравенству

$$|f'(x) - f'(y)| \leq C_1|x - y| + C_2|y| |x - y| + |x - y| = (C_1 + C_2|y| + 1)|x - y|. \quad (25)$$

Полагая $x = W$, $y = W^{(j)}$ и используя последнее неравенство, получим

$$\begin{aligned} |E\mathbf{A}f(W)| &\leq \frac{1}{n^{3/2}} \sum_{j=1}^n (E|X_j|^2 E|X_j| + E|X_j|^3) (C_1 + C_2 E|W^{(j)}| + 1) \\ &\quad + 2 \frac{C_1 + C_2 + 1}{\sqrt{n}} E|X_1|^3. \end{aligned} \quad (26)$$

Последнее неравенство влечет, что для $C = 2(C_1 + C_2 + 1)$

$$\mathcal{W}_1(W, Z) \leq \frac{C}{\sqrt{n}} E|X_1|^3. \quad (27)$$

Теорема 1 доказана.

Приведенные рассуждения без труда обобщаются на суммы слабо зависимых величин. Детали можно найти, например, в работе Стейна [68].

2.2. Оценка точности аппроксимации сумм индикаторных функций распределением Пуассона в метрике полной вариации. Пусть $X_1^{(n)}, X_2^{(n)}, \dots$ – схема серий независимых в каждой серии бернуллиевских случайных величин с $E X_j^{(n)} = p_j^{(n)}$. Положим

$$\lambda = \sum_{j=1}^n p_j^{(n)}.$$

Нас будет интересовать оценка близости распределений случайной величины $W = \sum_{j=1}^n X_j^{(n)}$ и распределения пуассоновской случайной величины Z с параметром λ .

В 1975 г. Л. Чен опубликовал статью “Poisson approximation for dependent trials” [22]. Чен предложил модификацию метода Стейна для оценки близости к распределению Пуассона. Мы приведем иллюстрацию применения метода Стейна для оценки близости к распределению Пуассона в метрике полной вариации. Пусть $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ – множество неотрицательных целых чисел. Распределение Пуассона задается функцией

$$p(x) = \frac{\lambda^x}{x!} \exp\{-\lambda\}, \quad x \in \mathbb{Z}_+.$$

Оператор Стейна для распределения Пуассона, предложенный Ченом, задается равенством

$$\mathbf{A}f(x) = xf(x) - \lambda f(x + 1). \quad (28)$$

Уравнение Стейна принимает вид

$$xf(x) - \lambda f(x+1) = h(x) - \mathbf{E}h(Z), \quad (29)$$

где $\mathbf{E}h(Z) = \sum_{x=0}^{\infty} h(x)p(x)$. Решение уравнения (29) для $x \geq 1$ имеет вид

$$\begin{aligned} f_h(x) &= -(x-1)! \lambda^{-x} \sum_{k=0}^{x-1} [h(k) - \mathbf{E}h(Z)] \frac{\lambda^k}{k!} \\ &= (x-1)! \lambda^{-x} \sum_{k=x}^{\infty} [h(k) - \mathbf{E}h(Z)] \frac{\lambda^k}{k!}. \end{aligned} \quad (30)$$

Мы будем оценивать расстояние полной вариации. Определяющее метрику пространство расширим до пространства функций таких, что $|h(x)| \leq 1$. В этом случае, как показал Чен в [22; леммы 3.1, 3.2 и 3.5], при $x \geq 1$

$$|f_h(x+1) - f_h(x)| \leq \lambda^{-1}(2 + |x - \lambda| \min\{1, \lambda^{-1/2}\}). \quad (31)$$

Далее, пользуясь тем, что случайные величины принимают только значения 0 или 1, получим

$$\mathbf{E}W f_h(W) = \sum_{j=1}^n \mathbf{E}X_j^{(n)} f_h(W) = \sum_{j=1}^n p_j^{(n)} \mathbf{E}f(W^{(j)} + 1), \quad (32)$$

где $W^{(j)} = W - X_j^{(n)}$.

Поскольку $\lambda = \sum_{j=1}^n p_j^{(n)}$, можно записать

$$\mathbf{E}A f_h(W) = \sum_{j=1}^n p_j^{(n)} \mathbf{E}(f(W^{(j)} + 1) - f(W + 1)). \quad (33)$$

Продолжая последнее равенство, мы приходим к следующему представлению:

$$\mathbf{E}A f_h(W) = \sum_{j=1}^n (p_j^{(n)})^2 \mathbf{E}(f(W^{(j)} + 1) - f(W^{(j)} + 2)). \quad (34)$$

Поскольку $W^{(j)} + 1 \geq 1$, мы можем воспользоваться оценкой (31). Получим

$$|\mathbf{E}A f_h(W)| \leq \lambda^{-1} \sum_{j=1}^n (p_j^{(n)})^2 (2 + \mathbf{E}|W^{(j)} + 1 - \lambda| \min\{1, \lambda^{-1/2}\}). \quad (35)$$

Очевидно, что $\mathbf{E}|W^{(j)} + 1 - \lambda| \leq 1 + \sqrt{\lambda}$. С учетом этой оценки неравенство (35) влечет, что

$$d_{\text{TV}}(W, Z) \leq 4\lambda^{-1} \sum_{j=1}^n (p_j^{(n)})^2. \quad (36)$$

Так же, как и в случае нормальной аппроксимации, приведенные рассуждения обобщаются на случай слабо зависимых величин. Детали можно найти в работе Чена [22].

ЗАМЕЧАНИЕ 2. Для достаточно большого класса распределений можно привести оператор Стейна. Например, для показательного распределения

$$p(x) = \begin{cases} \alpha \exp\{-\alpha x\}, & \text{если } x > 0, \\ 0, & \text{если } x < 0, \end{cases}$$

оператор Стейна задается формулой

$$\mathbf{A}f(x) = f'(x) - \alpha f(x).$$

Для биномиального распределения $\text{Bi}(n, p)$ с параметрами n, p оператор Стейна задается формулой

$$\mathbf{A}f(x) = \frac{1}{1-p}(np-x)f(x+1) - xf(x). \quad (37)$$

Пространство функций $\mathcal{H}$ определяется метрикой. А вот класс функций Стейна определяется и оператором, и пространством $\mathcal{H}$, т.е. метрикой. Как видно из приведенных выше примеров, наиболее сложная часть метода – это исследование свойств функций $f \in \mathcal{F}(\mathbf{A}, \mathcal{H})$, где $\mathcal{F}(\mathbf{A}, \mathcal{H})$ – множество функций $f \in \mathcal{F}(\mathbf{A})$ таких, что существуют $h \in \mathcal{H}$, для которых

$$\mathbf{A}f(x) = h(x) - \mathbf{E}h(Z). \quad (38)$$

3. Обзор литературы по методу Стейна

Описанию метода Стейна посвящена обширная литература. Метод Стейна применялся многими авторами при оценивании скорости сходимости в различных предельных теоремах – см., например, [17]–[19], [67], [71], [43].

Первая монография по методу Стейна написана самим Стейном в 1986 г. [69]. В этой книге Стейн приводит описание метода оценки точности аппроксимации математического ожидания функции от некоторой случайной величины математическим ожиданием другой случайной величины. В случае индикаторных функций задача эквивалентна оценке точности аппроксимации функций распределения. В монографии приводится не только применение подхода к аппроксимации нормальным законом, но и, для дискретных случайных величин, аппроксимация биномиальным законом распределения числа единиц в двоичном представлении случайного целого числа, аппроксимация пуассоновским распределением распределения сумм бернуллиевских величин, задача вычисления математического ожидания числа латинских прямоугольников.

В 1992 г. вышла в свет монография А. Барбура (A. D. Barbour), Л. Холста (L. Holst) и С. Янсона (S. Janson) “Poisson approximation” [10]. В ней подробно изложен метод Стейна–Чена пуассоновской аппроксимации и приведены многочисленные примеры применения этого метода.

В 1998 г. вышла обзорная статья Л. Чена [23] о методе Стейна.

В 2004 г. издан сборник под редакцией П. Дьякониса и С. Холмса “Stein’s method: expository lectures and applications” [32], в котором представлены статьи Ч. Стейна, П. Дьякониса (P. Diaconis), С. Холмса (S. Holmes), Г. Райнерта (G. Reinert) и дается дальнейшее развитие метода Стейна. В частности,

рассматривается применение метода Стейна для процесса рождения и гибели (Холмс), для цепей Маркова (Дьяконис), применение метода Стейна для бутстреп-процедур (Райнерт и Холмс).

В 2005 г. опубликован сборник под редакцией А. Барбура и Л. Чена “An introduction to Stein’s method” [8]. В сборнике приводятся обзорные статьи Ки-Ман Шао и Л. Чена о нормальной аппроксимации методом Стейна, статья Т. Эрхардсона (Т. Erhardsson) об аппроксимации методом Стейна пуассоновским распределением и сопровождающим пуассоновским распределением, а также статья Г. Райнерт о трех общих подходах к методу Стейна.

В 2010 г. вышла в свет книга Л. Чена, Л. Гольдштейна (L. Goldstein) и К.-М. Шао (Q.-M. Shao) “Normal approximation by Stein’s method” [25], в которой приводится история метода Стейна, его обоснование и различные приложения. Последние главы книги посвящены методу Стейна и исчислению Маллявэна.

В 2012 г. издана книга И. Нурдина (I. Nourdin) и Дж. Пеккати (G. Peccati) “Normal approximations with Malliavin calculus: from Stein’s method to universality” [57]. В книге расширяется сфера применения метода Стейна. Рассматриваются задачи, связанные с винеровским хаосом, случайными процессами, стохастическим анализом. Приводится известная теорема четырех моментов для винеровского хаоса.

В 2019 г. вышла в свет книга Б. Аппаса (B. Arras) и К. Удре (C. Houdré) “On Stein’s method for infinitely divisible laws with finite first moment” [1]. В этой книге основное внимание уделяется оценкам скорости сходимости в предельных теоремах, когда предельный закон является безгранично делимым и имеет конечный первый момент. Представлены и разработаны два метода для получения таких оценок. Первый метод основан на характеристических функциях и тождествах типа Стейна. Второй метод представляет собой общий подход метода Стейна для одномерных саморазложимых законов с конечным первым моментом.

Следует упомянуть также некоторые обзорные статьи по методу Стейна. Одна из первых обзорных работ по методу Стейна на русском языке – это статья Й. Сунклодаса [70] в сборнике “Итоги науки и техники” 1991 г. Автор делает основной упор на обзор результатов по скорости сходимости в центральной предельной теореме для слабо зависимых величин (включая m -зависимые). Сунклодас разделяет результаты, полученные прямыми методами (метод Стейна), и результаты, полученные с применением уравнения для характеристических функций, называя это методом Тихомирова [70].

Обзорная статья В. Ротаря и Й. Ринотта (Y. Rinott) 2000 г. [62] посвящена применению метода Стейна к нормальной аппроксимации. В статье приводится элементарное обоснование метода Стейна и дается ряд примеров применения этого метода. В частности, рассматриваются перестановочные величины и локально зависимые величины.

В обзорной статье Н. Росса (N. Ross) [65] 2011 г. обсуждаются основные концепции метода Стейна для аппроксимации нормальным распределением, распределением Пуассона, экспоненциальным и геометрическим распределениями, а также связь метода Стейна с неравенствами концентрации меры.

Первые работы с применением метода Стейна были посвящены либо нормальной аппроксимации, либо пуассоновской аппроксимации для зависимых величин.

В этой связи отметим обзорную статью З. Чаттерджи (S. Chatterjee) [20], в которой дается краткое описание исторического развития метода Стейна, а также приводятся некоторые современные достижения этого метода.

Обзорная статья Л. Чена 2021 г. [24] интересна с точки зрения истории возникновения метода Стейна. Автор делится своими личными впечатлениями и воспоминаниями о работе с Чарльзом Стейном. Чен пишет, что Ч. Стейн пришел к своему характеристическому уравнению для нормальной плотности исходя из уравнения для характеристической функции нормального закона, которое Стейн использовал для доказательства комбинаторной центральной предельной теоремы. В то же время в обзоре подчеркивается, что яркой особенностью метода Стейна является то, что он не работает с аналитическими методами преобразования Фурье (характеристическими функциями) и что в основе метода лежит решение дифференциального уравнения первого порядка. Отмечается также, что применения метода Стейна, начавшиеся с нормальной и пуассоновской аппроксимаций, распространяются на все более широкие классы распределений.

Различным задачам с применением метода Стейна посвящены сотни работ, и перечислить все просто невозможно. Тем не менее мы отметим некоторые из них, в основном касающиеся нормальной аппроксимации, пуассоновской аппроксимации, распространения метода Стейна на другие распределения, многомерных вариантов метода Стейна. Одно из первых направлений в развитии метода Стейна – пуассоновская аппроксимация. Л. Чен предложил модификацию метода Стейна для аппроксимации суммы зависимых индикаторов пуассоновским распределением (см. [22]). В 1977 г. был опубликован результат, полученный совместно П. Дьяконисом и Ч. Стейном (см. [30; теорема 8]), – оценка точности нормальной аппроксимации распределения числа единиц в двоичном представлении случайного целого числа. В 1978 г. в работе Л. Чена и С.-Т. Хо (S.-T. Ho) [50] методом Стейна была получена оценка точности нормальной аппроксимации сумм независимых двухиндексных случайных величин по случайным перестановкам. В работе Барбура 1982 г. [3] рассматривается применение метода Стейна к решению задач, связанных со случайными графами. При этом рассматривается как пуассоновская, так и нормальная аппроксимация. В работах [4]–[6] метод Стейна используется для вывода асимптотических разложений математического ожидания гладких функций от сумм независимых случайных величин. В работе [7] метод Стейна развит в контексте функциональной аппроксимации винеровским процессом и другими гауссовскими процессами. В 1991 г. Ф. Гётце (F. Götze) [43] развил технику Стейна для получения оценок точности аппроксимации многомерным нормальным распределением. В 1992 г. А. Барбур, Л. Чен и В.-Л. Ло (W.-L. Loh) распространили метод Стейна на сопровождающее пуассоновское распределение (см. [9]). Продолжение этой тематики можно найти в работе [64]. В статье [61] 1997 г. авторы приводят дальнейшее развитие метода Стейна нормальной аппроксимации, применяя его к различным типам зависимости, включая функции от цепей

Маркова и U -статистики. Тема многомерной нормальной аппроксимации и метод Стейна с приложением к задачам математической статистики получили дальнейшее развитие в работах [42], [41], [53]. К.-М. Шао и Л. Чен в работе [26] сделали основной упор на развитие метода Стейна для доказательства неравномерных оценок в нормальной аппроксимации сумм неодинаково распределенных случайных величин. В работе П. Дьякониса [31] рассматриваются примеры, демонстрирующие применение метода Стейна к цепям Маркова. Приводятся оценки скорости сходимости к стационарному распределению для некоторых простых цепей Маркова. Чен и Шао в работе 2004 г. [27] рассмотрели применение метода Стейна к локально зависимым величинам (обобщение m -зависимости). Они получили методом Стейна равномерную и неравномерную оценки типа Берри–Эссеена для нормальной аппроксимации сумм локально зависимых величин. В работах Ф. Гётце и автора [47] и [48] метод Стейна применялся к доказательству полукругового закона для спектра случайных матриц с мартингальной зависимостью элементов. М. Райс (M. Raic) в [60; теорема 2.2] использовал метод Стейна для оценки вероятностей больших отклонений для сумм независимых случайных величин. Он применил модифицированное уравнение Стейна с учетом преобразования Крамера. В работе [36] этот метод развит для дискретных гиббсовских мер. В работах [56], [55], [34], [37] авторы объединили метод исчисления Маллявэна и метод Стейна, получив оценки типа Берри–Эссеена для гауссовской аппроксимации и гамма-аппроксимации случайных величин в винеровском хаосе гауссовских процессов, а также для аппроксимации функционалов от гауссовских полей. Работа [33] посвящена развитию метода Стейна для перестановочных случайных величин. В работе [21] метод Стейна развивается для оценки точности негауссовской аппроксимации в модели Кюри–Вайсса. В работе [58] авторы применяют метод Стейна для оценки скорости сходимости распределения числа шаров заданного цвета в урновой модели Пойя к классу обобщенных гамма-распределений, который включает в себя рэлеевское, полунормальное и гамма-распределения. В работе [40] рассматривается применение метода Стейна к аппроксимации χ^2 -распределением и критерию Неймана–Пирсона. К.-М. Шао и З. С. Жанг (Z. S. Zhang) в работе [66] рассматривали как гауссовскую, так и негауссовскую аппроксимацию в случае неограниченных перестановочных случайных величин. В работе [54] дается общее описание метода Стейна для оценки точности аппроксимации произвольным многомерным непрерывным распределением. Рассматриваются оценки в метрике Вассерштейна. В работе [28] помощью метода Стейна даются оценки аппроксимации α -устойчивым законом с $\alpha \in (0, 2)$ в метрике типа Вассерштейна. В работе [16] разрабатывается новый метод, основанный на методе Стейна, для сравнения стационарных распределений неприводимой цепи Маркова. В недавней работе [35] П. Айхельсбахер (P. Eichelsbacher) и Б. Редносс (B. Rednoß) использовали метод Стейна и характеристические функции для анализа асимптотической нормальности числа подграфов в случайном графе Эрдёша–Реньи.

4. Характеристические функции и метод Стейна

В дальнейшем мы уделяем основное внимание применению характеристических функций.

Если перейти к характеристическим функциям, то новая характеристика запишется так. Мы рассматриваем такой оператор, что для характеристической функции $\varphi(t) = \mathbb{E} \exp\{it\xi\}$ случайной величины ξ равенство

$$\mathbf{A}\varphi(t) = 0 \quad \text{при любом } t \in \mathbb{R} \quad (39)$$

выполняется тогда и только тогда, когда $\varphi(t) = \mathbb{E} \exp\{itZ\}$, т. е. распределение случайной величины ξ совпадает с распределением случайной величины Z . Если рассмотреть действие этого оператора на характеристическую функцию $\psi(t) = \mathbb{E} \exp\{itW\}$, то мы можем записать

$$\mathbf{A}\psi(t) = \mathbf{A}(\psi(t) - \varphi(t)) =: \varepsilon(t). \quad (40)$$

Решая уравнение

$$\mathbf{A}g(t) = \varepsilon(t), \quad \varepsilon(0) = 0, \quad (41)$$

мы получим оценку разности $g(t) = \psi(t) - \varphi(t)$ характеристических функций $\varphi(t)$ и $\psi(t)$.

Существует обширная литература, отражающая возможности оценивать различные метрики в классе функций распределения через характеристические функции. Можно, например, ознакомиться с замечательным обзором по этой тематике С. Г. Бобкова [13].

5. Аппроксимация нормальным законом

Пусть случайная величина Z распределена по стандартному нормальному закону. Очевидно, что характеристическая функция $f_{\text{nor}}(t)$ удовлетворяет уравнению

$$f'(t) = -tf(t), \quad f(0) = 1.$$

Оператор $\mathbf{A}$ определяется равенством

$$\mathbf{A}f(t) = f'(t) + tf(t).$$

Мы можем сравнить этот оператор с усредненным оператором Стейна

$$\mathbf{A}_{\varphi_{\text{nor}}} \varkappa(x) = \varkappa'(x) + x\varkappa(x)$$

в классе функций $\mathcal{F} = \{\varkappa_t(x) = \exp\{itx\}, t \in \mathbb{R}\}$:

$$\mathbf{E} \mathbf{A}_{\varphi_{\text{nor}}} \varkappa_t(W) = \mathbb{E}(it\varkappa_t(x) - W\varkappa_t(W)) = i(f(t) + tf'(t)).$$

Вычисляя значение оператора $\mathbf{A}$ на функции $f(t) = \mathbb{E} \exp\{itW\}$ и учитывая, что $\mathbf{A}f_{\text{nor}}(t) = 0$, можем записать

$$\mathbf{A}f(t) = \mathbf{A}(f(t) - f_{\text{nor}}(t)) =: \varepsilon(t).$$

Полагая $g(t) = f(t) - f_{\text{нор}}(t)$ и решая уравнение

$$\mathbf{A}g(t) = \varepsilon(t), \quad g(t) = 0, \quad (42)$$

мы получим равенство

$$g(t) = f(t) - f_{\text{нор}}(t) = \exp\left\{-\frac{t^2}{2}\right\} \int_0^t \exp\left\{\frac{u^2}{2}\right\} \varepsilon(u) du. \quad (43)$$

Докажем следующую лемму.

ЛЕММА 3. Пусть характеристическая функция $f(t) = \mathbf{E} \exp\{itW\}$ удовлетворяет соотношению

$$f'(t) = (-t + \theta_1(t)A|t|^2)f(t) + \theta_2(t)B^k|t|^k, \quad f(0) = 1, \quad (44)$$

где $|\theta_j(t)| \leq 1$ для $j = 1, 2$, а A и B — некоторые положительные величины, не зависящие от t , и $k \geq 1$. Тогда существуют абсолютные постоянные $C_1, C_2 > 0$ такие, что при $A > 0$

$$\sup_x |\Pr\{W < x\} - \Phi(x)| \leq \begin{cases} C_1 \left(A + \frac{B^{k-1}}{A^{k-1}} B \right), & \text{если } k > 1, \\ C_1(A + B|\log A|), & \text{если } k = 1, \end{cases} \quad (45)$$

и при $A = 0$

$$\sup_x |\Pr\{W < x\} - \Phi(x)| \leq \begin{cases} CB, & \text{если } k > 1, \\ CB|\log B|, & \text{если } k = 1. \end{cases} \quad (46)$$

ДОКАЗАТЕЛЬСТВО. Решение уравнения (44) можно записать в виде

$$\begin{aligned} f(t) = & \exp\left\{-\frac{t^2}{2} + A \int_0^t \theta_1(u)u^2 du\right\} \\ & \times \left(1 + B^k \int_0^t \theta_2(u)u^k \exp\left\{\frac{u^2}{2} - A \int_0^u v^2 \theta_1(v) dv\right\} du\right). \end{aligned} \quad (47)$$

Простые вычисления показывают, что для любых $0 \leq u \leq t \leq 1/(2A)$

$$\begin{aligned} -\frac{t^2}{2} + \frac{u^2}{2} + A \int_u^t \theta_1(v)v^2 dv & \leq -\frac{t^2}{2} + \frac{u^2}{2} + \frac{1}{3}A(t^3 - u^3) \\ & \leq -\frac{t^2}{4} + \frac{u^2}{4}. \end{aligned} \quad (48)$$

Равенство (47) влечет

$$\begin{aligned} \left| f(t) - \exp\left\{-\frac{t^2}{2}\right\} \right| & \leq \exp\left\{-\frac{t^2}{2}\right\} \left| \exp\left\{A \int_0^t \theta_1(u)u du\right\} - 1 \right| \\ & \quad + B^k \left| \int_0^t \theta_2(u)u^k \exp\left\{-\frac{t^2}{2} + \frac{u^2}{2} + A \int_u^t \theta_1(v) dv\right\} du \right|. \end{aligned} \quad (49)$$

Продолжая это неравенство, с учетом оценки (48) получим, что для $|t| \leq 1/(2A)$

$$\left| f(t) - \exp\left\{-\frac{t^2}{2}\right\} \right| \leq |A| |t| \exp\left\{-\frac{t^2}{4}\right\} + B^k \min\{|t|^{k-1}, |t|^k\}. \quad (50)$$

Нетрудно видеть, что это неравенство и неравенство Эссеена (см. [13; теорема 2.1]) влекут

$$\sup_x |\Pr\{W < x\} - \Phi(x)| \leq \begin{cases} C_1 \left(A + \frac{B^k}{A^{k-1}} \right), & \text{если } k > 1, \\ C_1 (A + B |\log A|), & \text{если } k = 1. \end{cases} \quad (51)$$

Случай $A = 0$ рассматривается аналогично. Лемма доказана.

В качестве примера применения леммы 3 рассмотрим оценку точности нормальной аппроксимации сумм m -зависимых случайных величин при условии, что последовательность случайных величин стационарна.

ОПРЕДЕЛЕНИЕ 3. Будем говорить, что последовательность случайных величин $X_1, X_2, \dots$ удовлетворяет условию m -зависимости, если для любых $1 \leq a < b < c \leq d$ случайные величины X_j , $j \in [a, b]$, и X_i , $i \in [c, d]$, независимы, как только $c - b > m$.

Мы будем считать, что

$$\mathbb{E}X_1 = 0, \quad \mathbb{E}X_1^2 = 1 \quad \text{и} \quad \mathbb{E}|X_1|^4 < \infty.$$

Положим

$$W = \frac{1}{\sqrt{n} \sigma_n} \sum_{j=1}^n X_j, \quad \text{где } \sigma_n^2 = \frac{1}{n} \mathbb{E} \left(\sum_{j=1}^n X_j \right)^2.$$

Введем в рассмотрение величины

$$\Delta_j = \frac{1}{\sqrt{n} \sigma_n} \sum_{k: |k-j| \leq m} X_k \quad \text{и} \quad W^{(j)} = W - \Delta_j.$$

Заметим, что величины X_j и $W^{(j)}$ независимы при любом $j = 1, \dots, n$. В этих обозначениях мы можем записать

$$\begin{aligned} \mathbf{A}f(t) &= i\mathbb{E}W \exp\{itW\} + tf(t)(t) \\ &= i \sum_{j=1}^n \frac{1}{\sqrt{n} \sigma_n} \mathbb{E}X_j \exp\{itW\} + t \mathbb{E} \exp\{itW\} \\ &= \frac{i}{\sqrt{n} \sigma_n} \sum_{j=1}^n \mathbb{E}X_j (\exp\{itW\} - \exp\{itW^{(j)}\}) + t \mathbb{E} \exp\{itW\} \\ &= \frac{i}{\sqrt{n} \sigma_n} \mathbb{E} \left(\sum_{j=1}^n X_j (1 - \exp\{-it\Delta_j\}) \right) \exp\{itW\} + t \mathbb{E} \exp\{itW\} \\ &= \frac{i}{\sqrt{n} \sigma_n} \mathbb{E} \left(\sum_{j=1}^n X_j \left(1 - \exp\{-it\Delta_j\} - \frac{1}{\sqrt{n} \sigma_n} it\Delta_j \right) \right) \exp\{itW\} \\ &\quad - t \mathbb{E} \left(\frac{1}{n\sigma_n^2} \sum_{j=1}^n (X_j \Delta_j - 1) \right) \exp\{itW\} =: A_1 + A_2. \end{aligned} \quad (52)$$

Заметим, что величины $\xi_j = X_j \Delta_j$ и $\eta_j = X_j(1 - \exp\{-it\Delta_j\})$ для $j = 1, 2, \dots$ являются $2m$ -зависимыми. Пусть $h = \lfloor n/(2m) \rfloor$, где $\lfloor a \rfloor$ обозначает целую часть числа a . Положим

$$\gamma_l = \sum_{j=2(l-1)m}^{2lm} \eta_j \quad \text{и} \quad \beta_l = \sum_{j=2(l-1)m}^{2lm} \xi_j, \quad \text{если } l = 1, \dots, h,$$

и

$$\gamma_{h+1} = \sum_{j=2mh+1}^n \eta_j, \quad \beta_{h+1} = \sum_{j=2mh+1}^n \xi_j.$$

Отсюда нетрудно получить, что

$$\mathbb{E} \left| \sum_{j=1}^n X_j(1 - \exp\{-it\Delta_j\}) \right|^2 \leq \sum_{l=1}^{h/2} \mathbb{E} |\eta_{2l+1}|^2 + \sum_{l=1}^{h/2} \mathbb{E} |\eta_{2l}|^2. \quad (53)$$

Очевидно, что

$$|\eta_j| \leq \frac{1}{n\sigma_n^2} |t| |X_j| |\Delta_j| \leq \frac{1}{2n\sigma_n^2} m |t|^2 \left(m |X_j|^2 + \sum_{l: |l-j| \leq m} |X_l|^2 \right), \quad (54)$$

откуда следует неравенство

$$\mathbb{E} \left| \frac{1}{\sqrt{n} \sigma_n} \sum_{j=1}^n X_j(1 - \exp\{-it\Delta_j\}) \right|^2 \leq C \frac{m^2}{n} |t|^2 \mathbb{E} |X_1|^4. \quad (55)$$

Применяя последнюю оценку, мы получим

$$\begin{aligned} & \left| A_1 - \mathbb{E} \left(\frac{1}{\sqrt{n} \sigma_n} \sum_{j=1}^n X_j(1 - \exp\{-it\Delta_j\}) \right) \mathbb{E} \exp\{itW\} \right| \\ & \leq \mathbb{E}^{1/2} \left| \frac{1}{\sqrt{n} \sigma_n} \sum_{j=1}^n X_j(1 - \exp\{-it\Delta_j\}) - \mathbb{E} \frac{1}{\sqrt{n} \sigma_n} \sum_{j=1}^n X_j(1 - \exp\{-it\Delta_j\}) \right|^2 \\ & \leq C \frac{m}{\sqrt{n}} |t| \mathbb{E}^{1/2} |X_1|^4. \end{aligned} \quad (56)$$

Далее, учитывая, что

$$\frac{1}{n\sigma_n^2} \mathbb{E} \left(\sum_{j=1}^n X_j \right)^2 = \frac{1}{\sqrt{n} \sigma_n} \sum_{j=1}^n \mathbb{E} X_j (\Delta_j + W^{(j)}) = \frac{1}{\sqrt{n} \sigma_n} \sum_{j=1}^n \mathbb{E} X_j \Delta_j = 1,$$

мы получим

$$\begin{aligned} & i \mathbb{E} \frac{1}{\sqrt{n} \sigma_n} \left(\sum_{j=1}^n X_j(1 - \exp\{-it\Delta_j\}) \right) \mathbb{E} \exp\{itW\} + t \mathbb{E} \exp\{itW\} \\ & = i \mathbb{E} \left(\frac{1}{\sqrt{n} \sigma_n} \sum_{j=1}^n \mathbb{E} X_j(1 - \exp\{-it\Delta_j\} - it\Delta_j) \right) \mathbb{E} \exp\{itW\}. \end{aligned} \quad (57)$$

Неравенство $|\exp\{ix\} - 1 - ix| \leq x^2/2$ влечет

$$\left| i \mathbb{E} \frac{1}{\sqrt{n} \sigma_n} \left(\sum_{j=1}^n X_j (1 - \exp\{-it\Delta_j\}) \right) \mathbb{E} \exp\{itW\} + t \mathbb{E} \exp\{itW\} \right| \leq \frac{t^2}{\sqrt{n} \sigma_n} \sum_{j=1}^n \mathbb{E} |X_j| |\Delta_j|^2. \quad (58)$$

Объединяя полученные неравенства, мы придем к оценке

$$|\mathbf{A}f(t)| \leq C \left(\frac{m|t|^2}{\sqrt{n}} \mathbb{E} |X_1|^3 |\psi(t)| + \frac{|t|m^2}{\sqrt{n}} \mathbb{E}^{1/2} |X_1|^4 \right). \quad (59)$$

Мы можем переписать последнее неравенство в виде

$$\mathbf{A}f(t) = C \frac{m|t|^2}{\sqrt{n}} \mathbb{E} |X_1|^3 \theta_1(t) \psi(t) + \frac{C|t|m^2}{\sqrt{n}} \mathbb{E}^{1/2} |X_1|^4 \theta_2(t). \quad (60)$$

Продолжая приведенное равенство, мы получим следующее:

$$\begin{aligned} g'(t) + \left(t - C \frac{m|t|^2}{\sqrt{n}} \mathbb{E} |X_1|^3 \theta_1(t) \right) g(t) &= C \frac{m|t|^2}{\sqrt{n}} \mathbb{E} |X_1|^3 \theta_1(t) \varphi(t) \\ &\quad + \frac{C|t|m^2}{\sqrt{n}} \mathbb{E}^{1/2} |X_1|^4 \theta_2(t), \quad g(0) = 0. \end{aligned} \quad (61)$$

Решение этого уравнения представляется в виде

$$\begin{aligned} g(t) &= \exp \left\{ -\frac{t^2}{2} - \int_0^t C \frac{m}{\sqrt{n}} \mathbb{E} |X_1|^3 u^2 \theta_1(u) du \right\} \\ &\quad \times \int_0^t \exp \left\{ \frac{u^2}{2} + \int_0^u C \frac{m}{\sqrt{n}} \mathbb{E} |X_1|^3 \theta_1(v) v^2 dv \right\} \frac{C|u|m^2}{\sqrt{n}} \mathbb{E}^{1/2} |X_1|^4 \theta_2(u) du. \end{aligned} \quad (62)$$

Применяя теперь лемму 3 с

$$A = C \frac{m}{\sqrt{n}} \mathbb{E} |X_1|^3, \quad B = \frac{Cm^2}{\sqrt{n}} \mathbb{E}^{1/2} |X_1|^4 \quad \text{и} \quad k = 1,$$

получим

$$\begin{aligned} \rho(W, Z) &= \sup_x |\Pr\{W < x\} - \Pr\{Z < x\}| \leq C \left(\frac{1}{T} + \int_0^T |\psi(t) - \varphi(t)| \frac{dt}{t} \right) \\ &\leq \frac{Cm}{\sqrt{n}} \mathbb{E} |X_1|^3 + \frac{Cm^2 \log n}{\sqrt{n}} \mathbb{E}^{1/2} |X_1|^4. \end{aligned} \quad (63)$$

Приведенная оценка оптимальна по порядку зависимости от числа слагаемых n (с точностью до логарифма). Можно без особых усилий улучшить эту оценку и по зависимости от моментов (снизить условия до $\mathbb{E} |X_1|^3 < \infty$), и по зависимости от n (избавиться от логарифмического множителя). Подробности имеются в [73], [74]. Дальнейшее развитие метода и его приложений можно найти также в работах [75], [76], [12].

5.1. Квадратичные формы. В качестве еще одного применения метода, описанного выше в этом разделе (лемма 3), мы приведем доказательство центральной предельной теоремы для квадратичных форм. Рассмотрим последовательность независимых одинаково распределенных случайных величин $X_1, X_2, \dots$ таких, что

$$\mathbb{E}X_1 = 0, \quad \mathbb{E}X_1^2 = 1, \quad \mathbb{E}|X_1|^3 = \beta_3.$$

Пусть $\mathbf{B} = (b_{ij})_{i,j=1}^n$ – симметричная матрица с $b_{jj} = 0$ для $j = 1, \dots, n$. Будем считать, что

$$\|\mathbf{B}\|_2^2 = \sum_{j,k=1}^n |b_{jk}|^2 = \frac{1}{2}.$$

Символом $\|\mathbf{B}\|$ обозначим операторную норму матрицы $\mathbf{B}$:

$$\|\mathbf{B}\| = \sup_{\mathbf{x} \in \mathbb{R}^n: \|\mathbf{x}\|=1} \langle \mathbf{x}, \mathbf{B}\mathbf{x} \rangle,$$

где $\langle \cdot, \cdot \rangle$ – скалярное произведение в $\mathbb{R}^n$.

Рассмотрим квадратичную форму

$$Q = \sum_{j,k=1}^n b_{jk} X_j X_k. \quad (64)$$

Нетрудно видеть, что

$$\mathbb{E}Q^2 = 2\|\mathbf{B}\|_2^2 = 1. \quad (65)$$

ТЕОРЕМА 2. Пусть $\lim_{n \rightarrow \infty} \|\mathbf{B}\| = 0$. Тогда

$$\lim \Pr\{Q < x\} = \Phi(x), \quad (66)$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{u^2}{2}\right\} du$. Более того,

$$\sup_x |\Pr\{Q < x\} - \Phi(x)| \leq C \|\mathbf{B}\|^{1/2} \beta_3^2. \quad (67)$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим характеристическую функцию

$$f(t) = \mathbb{E} \exp\{itQ\}.$$

Положим $\mathbf{X} = (X_1, \dots, X_n)^\top$ и

$$\mathbf{X}^{(j)} = (X_1, \dots, X_{j-1}, 0, X_{j+1}, \dots, X_n)^\top, \quad \mathbf{X}_j = (0, \dots, 0, X_j, 0, \dots, 0)^\top$$

для $j = 1, \dots, n$.

Во введенных обозначениях мы можем записать значение оператора $\mathbf{A}$ в следующем виде:

$$\begin{aligned} f'(t) + tf(t) &= i\mathbb{E}\langle \mathbf{X}, \mathbf{B}\mathbf{X} \rangle \exp\{itQ\} + t\mathbb{E} \exp\{itQ\} \\ &= i \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{B}\mathbf{X} \rangle \exp\{itQ\} + t\mathbb{E} \exp\{itQ\}. \end{aligned} \quad (68)$$

Положим $Q^{(j)} = \langle \mathbf{X}^{(j)}, \mathbf{B}\mathbf{X}^{(j)} \rangle$. Заметим, что

$$\Delta Q_j = Q - Q^{(j)} = 2\langle \mathbf{X}_j, \mathbf{B}\mathbf{X}^{(j)} \rangle, \quad (69)$$

поскольку $\langle \mathbf{X}_j, \mathbf{B}\mathbf{X}_j \rangle = b_{jj} X_j^2 = 0$.

Введем в рассмотрение резольвентную матрицу $\mathbf{R}_t = (\mathbf{I} - 2it\mathbf{B})^{-1}$, где $\mathbf{I}$ – единичная матрица. Рассмотрим выражение

$$\mathbb{E}\langle \mathbf{X}, \mathbf{R}_t \mathbf{B}\mathbf{X} \rangle \exp\{itQ\} = \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X} \rangle \exp\{itQ\}. \quad (70)$$

Учитывая независимость X_j и $Q^{(j)}$ от $\mathbf{X}^{(j)}$, получаем

$$\begin{aligned} \mathbb{E}\langle \mathbf{X}, \mathbf{R}_t \mathbf{B}\mathbf{X} \rangle \exp\{itQ\} &= \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X}^{(j)} \rangle (\exp\{it\Delta Q_j\} - 1) \exp\{itQ^{(j)}\} \\ &\quad + \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X}_j \rangle \exp\{itQ\}. \end{aligned} \quad (71)$$

Продолжая это равенство, мы приходим к следующему представлению:

$$\begin{aligned} \mathbb{E}\langle \mathbf{X}, \mathbf{R}_t \mathbf{B}\mathbf{X} \rangle \exp\{itQ\} &= it \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X}^{(j)} \rangle \Delta Q_j \exp\{itQ^{(j)}\} \\ &\quad + \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X}_j \rangle \exp\{itQ^{(j)}\} + \varepsilon_1(t) + \varepsilon_2(t), \end{aligned} \quad (72)$$

где

$$\begin{aligned} \varepsilon_1(t) &= \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X}^{(j)} \rangle (\exp\{it\Delta Q_j\} - 1 - it\Delta Q_j) \exp\{itQ^{(j)}\}, \\ \varepsilon_2(t) &= \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X}_j \rangle (\exp\{it\Delta Q_j\} - 1) \exp\{itQ^{(j)}\}. \end{aligned} \quad (73)$$

Далее,

$$\begin{aligned} it \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X}^{(j)} \rangle \Delta Q_j \exp\{itQ^{(j)}\} \\ = 2it \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X}^{(j)} \rangle \langle \mathbf{X}_j, \mathbf{B}\mathbf{X}^{(j)} \rangle \exp\{itQ^{(j)}\}. \end{aligned} \quad (74)$$

Обозначим символом $\mathbf{B}^{(j)}$ матрицу, у которой только один элемент, $b_{jj}^{(j)}$, отличен от нуля: $b_{jj}^{(j)} = 1$. В этих обозначениях, принимая во внимание независимость величин $\mathbf{X}_j$ и $\mathbf{X}^{(j)}$ от $Q^{(j)}$, мы получим

$$\begin{aligned} it \sum_{j=1}^n \mathbb{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B}\mathbf{X}^{(j)} \rangle \Delta Q_j \exp\{itQ^{(j)}\} \\ = 2it \sum_{j=1}^n \mathbb{E}\langle \mathbf{B}^{(j)} \mathbf{B}\mathbf{X}^{(j)}, \mathbf{R}_t \mathbf{B}\mathbf{X}^{(j)} \rangle \exp\{itQ^{(j)}\}. \end{aligned} \quad (75)$$

Продолжая последнее равенство, нетрудно получить

$$\begin{aligned}
 & it \sum_{j=1}^n \mathbb{E} \langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle \Delta Q_j \exp\{itQ^{(j)}\} \\
 &= 2it \sum_{j=1}^n \mathbb{E} \langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}, \mathbf{R}_t \mathbf{B} \mathbf{X} \rangle \exp\{itQ\} \\
 &\quad + 2it \sum_{j=1}^n \mathbb{E} \langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle \exp\{itQ^{(j)}\} + \varepsilon_3(t) + \varepsilon_4(t), \tag{76}
 \end{aligned}$$

где

$$\begin{aligned}
 \varepsilon_3(t) &= 4it \sum_{j=1}^n \mathbb{E} \langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle \exp\{itQ^{(j)}\}, \\
 \varepsilon_4(t) &= 2it \sum_{j=1}^n \mathbb{E} \langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}, \mathbf{R}_t \mathbf{B} \mathbf{X} \rangle (1 - \exp\{it\Delta Q^{(j)}\}) \exp\{itQ^{(j)}\}.
 \end{aligned}$$

Поскольку $\sum_{j=1}^n \mathbf{B}_j = \mathbf{I}$, то, принимая во внимание симметричность матрицы $\mathbf{B}$, мы можем переписать представление (76) в виде

$$\begin{aligned}
 & it \sum_{j=1}^n \mathbb{E} \langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle \Delta Q_j \exp\{itQ^{(j)}\} \\
 &= 2it \mathbb{E} \langle \mathbf{X}, \mathbf{R}_t \mathbf{B}^2 \mathbf{X} \rangle \exp\{itQ\} + \varepsilon_3(t) + \varepsilon_4(t). \tag{77}
 \end{aligned}$$

Аналогично,

$$\begin{aligned}
 & it \sum_{j=1}^n \mathbb{E} \langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle \Delta Q_j \exp\{itQ^{(j)}\} \\
 &= 2it \sum_{j=1}^n \mathbb{E} \langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle \langle \mathbf{X}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle \exp\{itQ^{(j)}\} \\
 &= 2it \sum_{j=1}^n \mathbb{E} \langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle \exp\{itQ^{(j)}\} = 0. \tag{78}
 \end{aligned}$$

Нетрудно видеть, что

$$\varepsilon_3(t) = 0. \tag{79}$$

Далее, рассмотрим величину

$$\varkappa(t) = \sum_{j=1}^n \mathbb{E} \langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle \exp\{itQ^{(j)}\} = \sum_{j=1}^n \mathbb{E} \langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle \mathbb{E} \exp\{itQ^{(j)}\}. \tag{80}$$

Продолжая последнее равенство, мы придем к следующему выражению:

$$\varkappa(t) = \sum_{j=1}^n \mathbb{E} \langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle \mathbb{E} \exp\{itQ\} + \varepsilon_5(t), \tag{81}$$

где

$$\varepsilon_5(t) = \sum_{j=1}^n \mathbf{E}\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle \mathbf{E}(1 - \exp\{it\Delta Q_j\}) \exp\{itQ^{(j)}\}. \quad (82)$$

Пусть $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ – собственные числа матрицы $\mathbf{B}$. Тогда

$$\|\mathbf{B}\|_2^2 = 1 = \sum_{j=1}^n \lambda_j^2, \quad \|\mathbf{B}\| = \max_{1 \leq j \leq n} |\lambda_j|.$$

Во введенных обозначениях мы имеем

$$\mathrm{Tr} \mathbf{R}_t \mathbf{B} = \sum_{j=1}^n \frac{\lambda_j}{1 - 2it\lambda_j} = \sum_{j=1}^n \lambda_j + 2it \sum_{j=1}^n \lambda_j^2 + (2it)^2 \sum_{j=1}^n \frac{\lambda_j^3}{1 - 2it\lambda_j}. \quad (83)$$

Заметим, что

$$\begin{aligned} \sum_{j=1}^n \lambda_j &= 0, \quad \sum_{j=1}^n \lambda_j^2 = \frac{1}{2}, \\ 2t^2 \left| \sum_{j=1}^n \frac{\lambda_j^3}{1 - 2it\lambda_j} \right| &\leq 2t^2 \max_{1 \leq j \leq n} |\lambda_j| = 2|t|^2 \|\mathbf{B}\|. \end{aligned} \quad (84)$$

Обозначив символом $\theta(t)$ некоторую функцию такую, что $|\theta(t)| \leq 1$, мы можем записать

$$\mathrm{Tr} \mathbf{R}_t \mathbf{B} = 2it + 4\theta(t)|t|^2 \|\mathbf{B}\|. \quad (85)$$

Объединяя равенства (70), (72), (76), (77), (81) и (85) с учетом (79), мы получим

$$\begin{aligned} \mathbf{E}\langle \mathbf{X}, \mathbf{R}_t \mathbf{B} \mathbf{X} \rangle \exp\{itQ\} &= 2it\mathbf{E}\langle \mathbf{X}, \mathbf{R}_t \mathbf{B}^2 \mathbf{X} \rangle \exp\{itQ\} + 2it\mathbf{E} \exp\{itQ\} \\ &\quad + \varepsilon_1(t) + \varepsilon_2(t) + \varepsilon_4(t) + 4\theta(t)|t|^2 \|\mathbf{B}\| \mathbf{E} \exp\{itQ\}. \end{aligned} \quad (86)$$

Переносим первое слагаемое из правой части последнего равенства и учитывая, что $\mathbf{R}_t(\mathbf{I} - 2it\mathbf{B}) = \mathbf{I}$, мы приходим к равенству

$$\begin{aligned} \mathbf{E}\langle \mathbf{X}, \mathbf{B} \mathbf{X} \rangle \exp\{itQ\} &= it\mathbf{E} \exp\{itQ\} + \varepsilon_1(t) + \varepsilon_2(t) + \varepsilon_4(t) \\ &\quad + 4\theta(t)|t|^2 \|\mathbf{B}\| \mathbf{E} \exp\{itQ\}. \end{aligned} \quad (87)$$

Для завершения доказательства достаточно оценить $\varepsilon_1(t)$, $\varepsilon_2(t)$, $\varepsilon_4(t)$. Обозначим $\mathbf{e}_j = (0, \dots, 0, 1, 0, \dots, 0)^\top$ ортонормированные векторы в $\mathbb{R}^n$. Оценим $\varepsilon_1(t)$. Применяя элементарные неравенства, мы получим

$$\begin{aligned} |\varepsilon_1(t)| &\leq 4t^2 \sum_{j=1}^n \mathbf{E}|\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle| |\langle \mathbf{X}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle|^2 \\ &= 4t^2 \sum_{j=1}^n \mathbf{E}|X_j|^3 \mathbf{E}|\langle \mathbf{e}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle| |\langle \mathbf{e}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle|^2 \\ &\leq 4t^2 \sum_{j=1}^n \mathbf{E}|X_j|^3 \mathbf{E}^{1/3} |\langle \mathbf{e}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle|^3 \mathbf{E}^{2/3} |\langle \mathbf{e}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle|^3 \\ &\leq 4t^2 \beta_3 \left(\sum_{j=1}^n \mathbf{E}|\langle \mathbf{e}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle|^3 \right)^{1/3} \left(\sum_{j=1}^n \mathbf{E}|\langle \mathbf{e}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle|^3 \right)^{2/3}. \end{aligned} \quad (88)$$

Используя неравенство Розенталя, нетрудно получить оценку

$$|\varepsilon_1| \leq C\beta_3^2 |t|^2 \|\mathbf{B}\|. \quad (89)$$

Далее,

$$\begin{aligned} |\varepsilon_2(t)| &\leq 2|t| \sum_{j=1}^n \mathbb{E} |\langle \mathbf{X}_j, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle| |\langle \mathbf{X}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle| \\ &\leq \beta_3 |t| \sum_{j=1}^n |\langle \mathbf{e}_j, \mathbf{R}_t \mathbf{B} \mathbf{e}_j \rangle| \mathbb{E} |\langle \mathbf{e}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle|. \end{aligned} \quad (90)$$

В силу резольвентного равенства $\mathbf{R}_t = \mathbf{I} + 2it\mathbf{R}_t\mathbf{B}$ мы имеем

$$\begin{aligned} \sum_{j=1}^n |[\mathbf{R}_t \mathbf{B}]_{jj}| &= \sum_{j=1}^n |[\mathbf{B}]_{jj} + [\mathbf{R}_t \mathbf{B}^2]_{jj}| = \sum_{j=1}^n |[\mathbf{R}_t \mathbf{B}^2]_{jj}| \\ &\leq \sum_{j=1}^n \frac{|\lambda_j^2|}{|1 - 2it\lambda_j|} \leq \|\mathbf{B}\|^2. \end{aligned} \quad (91)$$

Кроме того,

$$\mathbb{E} |\langle \mathbf{e}_k, \mathbf{B} \mathbf{X}^{(k)} \rangle| \leq \mathbb{E}^{1/2} \left| \sum_{j=1}^n [\mathbf{B}]_{jk} X_k \right|^2 \leq \left(\sum_{j=1}^n |[\mathbf{B}]_{jk}|^2 \right)^{1/2} \leq \|\mathbf{B}\|. \quad (92)$$

Отсюда следует, что

$$|\varepsilon_2(t)| \leq C|t|\beta_3 \|\mathbf{B}\|. \quad (93)$$

Наконец, оценим $\varepsilon_4(t)$. Мы имеем

$$|\varepsilon_4(t)| \leq 2|t| \sum_{j=1}^n \mathbb{E} |\langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}, \mathbf{R}_t \mathbf{B} \mathbf{X} \rangle| |\Delta Q^{(j)}|. \quad (94)$$

Используя равенство $\mathbf{B}^{(j)} \mathbf{B} \mathbf{X}_j = 0$, получим

$$\begin{aligned} |\varepsilon_4(t)| &\leq 2|t| \sum_{j=1}^n (\mathbb{E} |\langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}^{(j)}, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle| |\Delta Q^{(j)}| \\ &\quad + \mathbb{E} |\langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}^{(j)}, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle| |\Delta Q^{(j)}|) \\ &\leq 4|t|^2 \sum_{j=1}^n (\mathbb{E} |\langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}^{(j)}, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle| |\langle \mathbf{X}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle| \\ &\quad + \mathbb{E} |\langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}^{(j)}, \mathbf{R}_t \mathbf{B} \mathbf{X}_j \rangle| |\langle \mathbf{X}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle|). \end{aligned} \quad (95)$$

Заметим, что

$$\begin{aligned} &\mathbb{E} |\langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}^{(j)}, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle| |\langle \mathbf{X}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle| \\ &\leq \mathbb{E}^{2/3} |\langle \mathbf{B}^{(j)} \mathbf{B} \mathbf{X}^{(j)}, \mathbf{R}_t \mathbf{B} \mathbf{X}^{(j)} \rangle|^{3/2} \mathbb{E}^{1/3} |\langle \mathbf{e}_j, \mathbf{B} \mathbf{X}^{(j)} \rangle|^3 \\ &\leq \mathbb{E}^{2/3} \left| \sum_{k=1}^n [\mathbf{B}]_{jk} X_k \right|^3 \mathbb{E}^{1/3} \left| \sum_{k=1}^n [\mathbf{R}_t \mathbf{B}]_{jk} X_k \right|^3 \\ &\leq \beta_3 \left(\sum_{k=1}^n |[\mathbf{B}]_{jk}|^2 \right) \|\mathbf{B}\|. \end{aligned} \quad (96)$$

Из последнего неравенства следует, что

$$|\varepsilon_4(t)| \leq C|t|^2\beta_3\|\mathbf{B}\|. \quad (97)$$

Равенство (87) и неравенства (89), (93) и (97) влекут представление

$$f'_n(t) + tf_n(t) = \theta_1(t)C|t|^2\|\mathbf{B}\|\beta_3 + \theta_2(t)t^2\beta_3^2\|\mathbf{B}\|f(t). \quad (98)$$

Применяя лемму 3, мы получим

$$\sup_x |\Pr\{Q < x\} - \Phi(x)| \leq C\|\mathbf{B}\|^{1/2}\mu_3. \quad (99)$$

Теорема 2 доказана.

Представленный результат неоптимален. Это прежде всего иллюстрация метода. Оптимальный результат, а именно неравенство

$$\sup_x |\Pr\{Q < x\} - \Phi(x)| \leq C\|\mathbf{B}\|,$$

получен в работе [46].

5.2. Аппроксимация вероятности попадания в шар суммы гильбертовозначных случайных величин. Пусть $X_1, X_2, \dots$ – стационарная в узком смысле последовательность m -зависимых случайных величин со значениями в сепарабельном гильбертовом пространстве $\mathbb{H}$. Будем обозначать символами $\langle \cdot, \cdot \rangle$ и $\|\cdot\|$ соответственно скалярное произведение и норму в $\mathbb{H}$. Мы будем считать, что $\mathbb{E}X_1 = 0$, $\mathbb{E}\|X_1\|^2 < \infty$ и определен ковариационный оператор $\mathbf{B}$ случайной величины X_1 . Будем предполагать, что случайные величины X_i, X_j , $i \neq j$, некоррелированы, т. е.

$$\mathbb{E}\langle X_i, h \rangle \langle X_j, g \rangle = 0 \quad \text{при } h, g \in \mathbb{H}, i \neq j. \quad (100)$$

Рассмотрим сумму

$$S_n = (n \operatorname{Tr} \mathbf{B})^{-1/2} \sum_{j=1}^n X_j$$

и гауссовский вектор Y , принимающий значения в $\mathbb{H}$, с $\mathbb{E}Y = 0$ и ковариационным оператором $\mathbf{B}/\operatorname{Tr} \mathbf{B}$. Нас будет интересовать оценка величины

$$\Delta_n = \sup_{x>0} |\Pr\{\|S_n\| < x\} - \Pr\{\|Y\| < x\}| \quad (101)$$

для последовательности $\{X_j\}_{j=1}^\infty$ случайных величин, являющихся m -зависимыми. Мы приведем доказательство следующей теоремы.

ТЕОРЕМА 3. Пусть выполнено условие (100). Предположим, что существует постоянная $M > 0$ такая, что

$$\mathbb{E}\|X_1\|^4 < \infty. \quad (102)$$

Тогда найдется постоянная C , значение которой определяется распределением собственных чисел оператора $\mathbf{B}$, такая, что

$$\Delta \leq Cmn^{-1/2}. \quad (103)$$

ДОКАЗАТЕЛЬСТВО теоремы основано на исследовании поведения характеристической функции $f_n(t) = \mathbb{E} \exp\{it\|S_n\|\}$ методом, описанным в предыдущем разделе. Пусть $f(t) = \mathbb{E}\|Y\|^2$ – характеристическая функция случайной величины $\|Y\|^2$. Пусть $\lambda_1 \geq \lambda_2 \geq \dots$ – собственные числа оператора $\mathbf{B}/\text{Tr } \mathbf{B}$. Известно, что

$$f(t) = \prod_{j=1}^{\infty} (1 - 2it\lambda_j)^{-1/2}. \quad (104)$$

Введем в рассмотрение резольвентный оператор $\mathbf{R}_t = (\mathbf{I} - 2it\mathbf{B})^{-1}$. Нетрудно видеть, что

$$f'(t) = i \left(\sum_{j=1}^{\infty} \frac{\lambda_j}{1 - 2it\lambda_j} \right) f(t) = i(\text{Tr } \mathbf{R}_t \mathbf{B}) f(t). \quad (105)$$

Оператор $\mathbf{A}$, действующий на множестве характеристических функций, мы можем определить следующим образом:

$$\mathbf{A}\psi(t) = \psi'(t) - i(\text{Tr } \mathbf{R}_t \mathbf{B})\psi(t). \quad (106)$$

Чтобы оценить действие оператора $\mathbf{A}$ на характеристическую функцию $f_n(t)$, рассмотрим величину

$$\psi_n(t) = \mathbb{E}\langle S_n, \mathbf{R}_t S_n \rangle \exp\{it\|S_n\|^2\}. \quad (107)$$

Для любого $j = 1, \dots, n$ определим величины

$$\eta_j = (n \text{Tr } \mathbf{B})^{-1/2} \sum_{k: |k-j| \leq m} X_k, \quad S_n^{(j)} = S_n - \eta_j. \quad (108)$$

Не умаляя общности, будем считать в дальнейшем, что $\text{Tr } \mathbf{B} = 1$. Мы можем представить функцию $\psi_n(t)$ в виде

$$\begin{aligned} \psi_n(t) &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}\langle X_j, \mathbf{R}_t S_n^{(j)} \rangle \exp\{it\|S_n^{(j)}\|^2\} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}\langle X_j, \mathbf{R}_t \eta_j \rangle \exp\{it\|S_n\|^2\} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}\langle X_j, \mathbf{R}_t S_n^{(j)} \rangle (\exp\{it(\|S_n\|^2 - \|S_n^{(j)}\|^2)\} - 1) \exp\{it\|S_n^{(j)}\|^2\}. \end{aligned} \quad (109)$$

Поскольку случайные величины X_j и $S_n^{(j)}$ независимы, первое слагаемое в правой части последнего равенства равно нулю. Далее, в силу условия (100)

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}\langle X_j, \mathbf{R}_t \eta_j \rangle = \frac{1}{n} \sum_{j=1}^n \mathbb{E}\langle X_j, \mathbf{R}_t \mathbf{X}_j \rangle = \text{Tr } \mathbf{R}_t \mathbf{B}. \quad (110)$$

Учитывая сделанные замечания, мы можем привести равенство (109) к виду

$$\begin{aligned} \psi_n(t) = & \operatorname{Tr} \mathbf{R}_t \mathbf{B} f_n(t) + \theta(t) D^{1/2} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n \langle X_j, \mathbf{R}_t \eta_j \rangle \right) \\ & + \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E} \langle X_j, \mathbf{R}_t S_n^{(j)} \rangle (\exp\{it(\|S_n\|^2 - \|S_n^{(j)}\|^2)\} - 1) \exp\{it\|S_n^{(j)}\|^2\}. \end{aligned} \quad (111)$$

Заметим, что

$$\|S_n\|^2 - \|S_n^{(j)}\|^2 = 2it \langle X_j, S_n^{(j)} \rangle + it \langle X_j, X_j \rangle. \quad (112)$$

Продолжая равенство (111), мы придем к следующему:

$$\begin{aligned} \psi_n(t) = & \operatorname{Tr} \mathbf{R}_t \mathbf{B} \psi_n(t) + \theta(t) D^{1/2} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n \langle X_j, \mathbf{R}_t \eta_j \rangle \right) \\ & + 2it \frac{1}{n} \sum_{j=1}^n \mathbb{E} \langle X_j, \mathbf{R}_t S_n^{(j)} \rangle \langle X_j, S_n^{(j)} \rangle \exp\{it\|S_n^{(j)}\|^2\} \\ & + it \frac{1}{n^{3/2}} \sum_{j=1}^n \mathbb{E} \langle X_j, \mathbf{R}_t S_n^{(j)} \rangle \langle X_j, X_j \rangle \exp\{it\|S_n^{(j)}\|^2\} \\ & + \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E} \langle X_j, \mathbf{R}_t S_n^{(j)} \rangle (\exp\{it(\|S_n\|^2 - \|S_n^{(j)}\|^2)\} - 1 \\ & - it(\|S_n\|^2 - \|S_n^{(j)}\|^2)) \exp\{it\|S_n^{(j)}\|^2\}. \end{aligned} \quad (113)$$

Так как случайные величины $\xi_j = X_j \eta_j$, $j = 1, \dots, n$, являются $2m$ -зависимыми и $\|\mathbf{R}_t\| \leq 1$, мы получим

$$D^{1/2} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n \langle X_j, \mathbf{R}_t \eta_j \rangle \right) \leq \frac{Cm}{\sqrt{n}} \mathbb{E}^{1/2} \|X_1\|^4. \quad (114)$$

Далее,

$$\begin{aligned} & 2it \frac{1}{n} \sum_{j=1}^n \mathbb{E} \langle X_j, \mathbf{R}_t S_n^{(j)} \rangle \langle X_j, S_n^{(j)} \rangle \exp\{it\|S_n^{(j)}\|^2\} \\ & = 2it \frac{1}{n} \sum_{j=1}^n \mathbb{E} \langle S_n^{(j)}, \mathbf{R}_t \mathbf{B} S_n^{(j)} \rangle \exp\{it\|S_n^{(j)}\|^2\} \\ & = 2it \mathbb{E} \langle S_n, \mathbf{R}_t \mathbf{B} S_n \rangle \exp\{it\|S_n\|^2\} \\ & \quad - 4it \frac{1}{n\sqrt{n}} \sum_{j=1}^n \mathbb{E} \langle X_j, \mathbf{R}_t \mathbf{B} S_n^{(j)} \rangle \exp\{\|S_n^{(j)}\|^2\} \\ & \quad - 2it \frac{1}{n\sqrt{n}} \sum_{j=1}^n \mathbb{E} \langle X_j, \mathbf{R}_t \mathbf{B} X_j \rangle \exp\{it\|S_n^{(j)}\|^2\} \\ & \quad - 2it \frac{1}{n} \sum_{j=1}^n \mathbb{E} \langle S_n, \mathbf{R}_t \mathbf{B} S_n \rangle (\exp\{it(\|S_n\|^2 - \|S_n^{(j)}\|^2)\} - 1) \exp\{\|S_n^{(j)}\|^2\}. \end{aligned} \quad (115)$$

Очевидно, что в силу независимости x_j и $S_n^{(j)}$ справедливо равенство

$$4it \frac{1}{n\sqrt{n}} \sum_{j=1}^n \mathbb{E} \langle X_j, \mathbf{R}_t \mathbf{B} S_n^{(j)} \rangle \exp\{\|S_n^{(j)}\|^2\} = 0. \quad (116)$$

Далее,

$$\left| 2it \frac{1}{n\sqrt{n}} \sum_{j=1}^n \mathbb{E} \langle X_j, \mathbf{R}_t \mathbf{B} X_j \rangle \exp\{it\|S_n^{(j)}\|^2\} \right| \leq \frac{2|t|}{\sqrt{n}} \mathbb{E} \|X_1\|^4. \quad (117)$$

Наконец,

$$\begin{aligned} & \left| 2it \frac{1}{n} \sum_{j=1}^n \mathbb{E} \langle S_n, \mathbf{R}_t \mathbf{B} S_n \rangle (\exp\{it(\|S_n\|^2 - \|S_n^{(j)}\|^2)\} - 1) \exp\{\|S_n^{(j)}\|^2\} \right| \\ & \leq C|t| \frac{1}{n} \sum_{j=1}^n \mathbb{E} |\langle S_n, \mathbf{R}_t \mathbf{B} S_n \rangle| \left(\frac{1}{\sqrt{n}} |\langle X_j, S_n^{(j)} \rangle| + \frac{1}{n} |\langle X_j, X_j \rangle| \right) \\ & \quad + C|t| \frac{1}{n\sqrt{n}} \sum_{j=1}^n \mathbb{E}^{1/2} |\langle S_n, \mathbf{R}_t \mathbf{B} S_n \rangle|^2 \mathbb{E}^{1/2} |\langle X_j, S_n^{(j)} \rangle|^2 \\ & \quad + C|t| \frac{1}{n^2} \sum_{j=1}^n \mathbb{E}^{1/2} |\langle S_n, \mathbf{R}_t \mathbf{B} S_n \rangle|^2 \mathbb{E}^{1/2} |\langle X_j, X_j \rangle|^2 \leq \frac{C|t|}{\sqrt{n}} m^2 \mathbb{E} \|X_1\|^4. \end{aligned} \quad (118)$$

Объединяя соотношения (109)–(118), мы получим

$$\begin{aligned} \psi_n(t) &= \text{Tr } \mathbf{R}_t \mathbf{B} \mathbb{E} \exp\{it\|S_n\|^2\} + 2it \mathbb{E} \langle S_n, \mathbf{R}_t \mathbf{B} S_n \rangle \exp\{it\|S_n\|^2\} \\ & \quad + \theta_1(t) \frac{C(1+|t|)}{\sqrt{n}} \mathbb{E} \|X_1\|^4. \end{aligned} \quad (119)$$

Принимая во внимание тот факт, что $(\mathbf{I} - 2it\mathbf{B})\mathbf{R}_t = \mathbf{I}$, мы придем к равенству

$$\mathbb{E} \langle S_n, S_n \rangle \exp\{it\|S_n\|^2\} = \text{Tr } \mathbf{R}_t \mathbf{B} \mathbb{E} \exp\{it\|S_n\|^2\} + \theta_1(t) \frac{Cm(1+|t|)}{\sqrt{n}} \mathbb{E} \|X_1\|^4. \quad (120)$$

Последнее равенство влечет

$$f'_n(t) - i \text{Tr } \mathbf{R}_t \mathbf{B} f_n(t) = \theta_1(t) \frac{C(1+|t|)}{\sqrt{n}} \mathbb{E} \|X_1\|^4. \quad (121)$$

Интегрируя последнее уравнение с начальным условием $f_n(0) = 1$, мы получим

$$f_n(t) = \prod_{j=1}^{\infty} \frac{1}{\sqrt{1-2it\lambda_j}} + \frac{Cm}{\sqrt{n}} \mathbb{E} \|X_1\|^4 \theta_1(t) \int_0^{|t|} (1+u) \prod_{j=1}^{\infty} \left| \frac{1-2iu\lambda_j}{1-2it\lambda_j} \right|^{1/2} du. \quad (122)$$

Неравенство Эссеена (см., например, [13; теорема 2.1]) и последнее представление влекут оценку

$$\sup_x |\Pr\{\|S_n\|^2 < x\} - \Pr\{\|Y\|^2 < x\}| \leq Cn^{-1/6} m \mathbb{E} \|X_1\|^4. \quad (123)$$

Теорема 3 доказана.

Различные уточнения оценки точности нормальной аппроксимации вероятности попадания в шар сумм гильбертовозначных слабо зависимых величин можно найти, например, в работах [78], [75], [12].

5.3. Применение метода к оценке точности аппроксимации распределения сумм мультииндексированных случайных величин. Случайные поля. В этом пункте мы остановимся на применении метода характеристических функций к суммированию случайных полей. Одними из первых этот метод для случайных полей применили авторы работы [49]. Рассмотрим случайное поле $X = \{X_j, j \in \mathbb{Z}^d\}$ такое, что $\mathbb{E}X_j = 0$ и $\mathbb{E}X_j^2 < \infty$. Пусть $\mathbb{A}_n \subset \mathbb{Z}^d$ – последовательность конечных подмножеств целочисленной решетки $\mathbb{Z}^d$. Определим S_n и σ_n следующим образом:

$$S_n = \sum_{j \in \mathbb{A}_n} X_j, \quad \sigma_n^2 = \mathbb{E}S_n^2. \quad (124)$$

Пусть $|\mathbb{A}_n|$ обозначает число элементов множества $\mathbb{A}_n$. Будем предполагать, что

$$\liminf_{n \rightarrow \infty} \sigma_n^2 |\mathbb{A}_n|^{-1} = \alpha > 0. \quad (125)$$

Пусть, кроме того, для некоторого $\delta > 0$

$$\sup_{j \in \mathbb{Z}^d} \mathbb{E}|X_j|^{2+\delta} =: \|X\|_{2+\delta}^{2+\delta} < \infty. \quad (126)$$

Положим

$$\Delta_n = \sup_x \left| \Pr \left\{ \frac{S_n}{\sigma_n} < x \right\} - \Phi(x) \right|,$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{u^2}{2}\right\} du$.

Результаты, полученные в работе [49], следующие.

ТЕОРЕМА 4. Пусть X – поле m -зависимых величин над $\mathbb{Z}^d$, для которого выполнены условия

$$\mathbb{E}X_j = 0, \quad \mathbb{E}X_j^2 < \infty \quad (127)$$

и условия (125), (126). Тогда

$$\Delta_n = \begin{cases} O(\sigma_n^{-\delta}), & \text{если } 0 < \delta < 1; \\ O((\log \sigma_n)^{(d-1)/2} \sigma_n^{-1}), & \text{если } \delta > 1. \end{cases} \quad (128)$$

Для любого подмножества индексов $\mathbb{A}$ символом $\mathfrak{M}_{\mathbb{A}}$ обозначим σ -алгебру, порожденную случайными величинами $X_j, j \in \mathbb{A}$. Для любых двух подмножеств $\mathbb{A}, \mathbb{B} \subset \mathbb{Z}^d$ индексов обозначим символом $d(\mathbb{A}, \mathbb{B})$ евклидово расстояние между этими множествами.

Условие сильного перемешивания для полей определим следующим образом: существует функция $\alpha(n)$ такая, что $\lim_{n \rightarrow \infty} \alpha(n) = 0$ и для любых $\mathbb{A}, \mathbb{B} \subset \mathbb{Z}^d$ выполнено неравенство

$$\sup_{A \in \mathfrak{M}_{\mathbb{A}}, B \in \mathfrak{M}_{\mathbb{B}}} |\Pr\{AB\} - \Pr\{A\} \Pr\{B\}| \leq \alpha(d(\mathbb{A}, \mathbb{B})). \quad (129)$$

Для полей с сильным перемешиванием была получена следующая оценка.

ТЕОРЕМА 5. Пусть X – центрированное случайное поле над $\mathbb{Z}^d$, удовлетворяющее условию сильного перемешивания с коэффициентом $\alpha(n)$, убывающим экспоненциально быстро, т. е. таким, что

$$\alpha(n) \leq c_1 \exp\{-c_2 n\} \quad \text{для всех } n \geq 1, \quad (130)$$

где $c_1, c_2 > 0$ – постоянные. Пусть выполнено условие (126). Тогда

$$\Delta_n = O((\log \sigma_n)^{d(1+\delta) \wedge 2} \sigma_n^{-\delta \wedge 1}) \quad (131)$$

(здесь и далее $a \wedge b = \min\{a, b\}$).

Доказательство этих теорем основано на методе Стейна–Тихомирова. В случае m -зависимых величин для $l \geq 1$ обозначим символом $\mathbb{B}(j, lm)$ шар радиуса lm с центром в точке $j \in \mathbb{Z}^d$. Пусть $\overline{\mathbb{B}(j, lm)}$ обозначает дополнение шара $\mathbb{B}(j, lm)$. Введем в рассмотрение величины

$$\begin{aligned} S_j^{(l)} &= \frac{1}{\sigma_n} \sum_{k \in \mathbb{A}_n \cap \mathbb{B}(j, lm)} X_k, & S_j^{(0)} &= \frac{1}{\sigma_n} \sum_{k \in \mathbb{A}_n} X_k, \\ \xi_j^{(l)} &= \exp\{it(S_j^{(l-1)} - S_j^{(l)})\} - 1, & l &\geq 1. \end{aligned} \quad (132)$$

Для производной $f'_n(t)$ характеристической функции $f_n(t) = \mathbb{E} \exp\{itS_n\}$ во введенных обозначениях справедливо представление

$$\begin{aligned} f'_n(t) &= \frac{i}{n} \left[\sum_{j \in \mathbb{A}_n} \mathbb{E}(X_j \exp\{itS_j^{(1)}\}) + \sum_{r=2}^q \sum_{j \in \mathbb{A}_n} \mathbb{E}\left(X_j \prod_{l=1}^{r-1} \xi_j^{(l)} \exp\{itS_j^{(r)}\}\right) \right. \\ &\quad \left. + \sum_{j \in \mathbb{A}_n} \mathbb{E}\left(X_j \prod_{l=1}^q \xi_j^{(l)} \exp\{itS_j^{(q)}\}\right) \right]. \end{aligned} \quad (133)$$

Обозначим $a_j(r) := \mathbb{E}\left(\prod_{l=1}^{r-1} \xi_j^{(l)}\right)$, $j = 1, \dots, n$, $2 \leq r \leq q$. Положим

$$\begin{aligned} T_0 &= \sum_{j \in \mathbb{A}_n} \mathbb{E}(X_j \exp\{itS_j^{(1)}\}), \\ T_1 &= \frac{i}{\sigma_n} \sum_{j \in \mathbb{A}_n} a_j(r) f_n(t), \\ T_2(r) &= \frac{i}{\sigma_n} \sum_{j \in \mathbb{A}_n} a_j(r) (\mathbb{E} \exp\{itS_j^{(r)}\} - f_n(t)), \\ T_3(r) &= \frac{i}{\sigma_n} \sum_{j \in \mathbb{A}_n} \mathbb{E}\left(X_j \prod_{l=1}^{r-1} \xi_j^{(l)} (\exp\{itS_j^{(r)}\} - \mathbb{E} \exp\{itS_j^{(r)}\})\right), \\ R_k &= \frac{i}{\sigma_n} \sum_{j \in \mathbb{A}_n} \mathbb{E}\left(X_j \prod_{l=1}^q \xi_j^{(l)} \exp\{itS_j^{(q)}\}\right). \end{aligned} \quad (134)$$

В силу m -зависимости случайные величины X_j и $S_j^{(1)}$ независимы. Следовательно, $T_0 = 0$. Аналогично, $T_3(r) = 0$. Таким образом,

$$f'_n(t) = \sum_{r=2}^q (T_1(r) + T_2(r)) + R_q. \quad (135)$$

Технически несложные, но достаточно объемные вычисления приводят в обозначениях работы [49] к выражению вида

$$f'_n(t) = -tf_n(t) + a_{n2}|t|^{(1+\delta)\wedge 2}\theta_{n2}(t)f_n(t) + b_{n0}\widehat{\theta}_{n0}(t) + b_{n2}(t)|t|^{(1+\delta)\wedge 2}, \quad (136)$$

где

$$a_{n2} = \sigma_n^{-(d\wedge 1)}, \quad b_{n0} \leq \sigma_n^{-1} \quad \text{и} \quad b_{n2} = \sigma_n^{-(1+\delta)\wedge 2}.$$

Отсюда с применением леммы 3 нетрудно получить требуемый результат.

Центральная предельная теорема для случайных полей с сильным перемешиванием рассматривалась в работе Э. Больтхаузена (E. Bolthausen) 1982 г. [15]. Оценкам скорости сходимости в центральной предельной теореме для случайных полей с различными условиями перемешивания посвящена обширная литература. См., например, обзор [70]. В работах Х. Такахата [71], [72] были получены оценки, улучшающие результаты работы [49] с применением метода Стейна (без характеристических функций). В работах А. В. Булинского рассматривались оценки скорости сходимости в центральной предельной теореме для случайных полей ассоциированных величин. Он использовал как метод Стейна с прямыми вычислениями, так и метод с характеристическими функциями (см. [18]). В работе [75] приведены оценки в центральной предельной теореме для векторнозначных случайных полей с перемешиванием.

6. Развитие метода характеристических функций

Пусть Z – случайная величина, распределенная по стандартному нормальному закону.

6.1. Вариант А. Роллина. В работе [63] был предложен следующий вариант обобщения леммы 3.

ЛЕММА 4 (A. Röllin). *Пусть W – случайная величина такая, что $E|W| < \infty$. Предположим, что ее характеристическая функция $\varphi(t) = E \exp\{itW\}$ удовлетворяет дифференциальному уравнению*

$$\varphi'(t) = -t(1 + a(t))\varphi(t) + b(t), \quad t \in \mathbb{R}, \quad (137)$$

где $a(t)$ и $b(t)$ – функции (возможно, с комплексными значениями), удовлетворяющие неравенствам

$$|a(t)| \leq A_0 + A_1|t|, \quad |b(t)| \leq B_0 + B_1|t| + B_2t^2, \quad (138)$$

с некоторыми неотрицательными константами $A_0 < 1/2$, A_1 , B_0 , B_1 и B_2 . Тогда для любого $t \geq 2A_1/(1 - 2A_0)$

$$\sup_x |\Pr\{W < x\} - \Pr\{Z < x\}| \leq \frac{2}{\pi} A_0 + \frac{4}{2\sqrt{3}} A_1 |t| + \frac{2}{\pi} B_1 \left(1 + 2 \log_+ \frac{1}{2t}\right) + \frac{4}{\pi} \frac{B_2}{t} + \frac{24t}{\pi\sqrt{2\pi}}. \quad (139)$$

ДОКАЗАТЕЛЬСТВО. Мы приведем доказательство, предложенное А. Роллин-ном (A. Röllin) в [63; лемма 2.3]. Введем обозначения

$$\tilde{a}(t) = -t - a(t)t \quad \text{и} \quad \Delta(t) = \int_0^t ua(u) du.$$

Запишем решение уравнения (137) в виде

$$\begin{aligned} \varphi(t) &= \exp\left\{-\frac{t^2}{2} - \Delta(t)\right\} \left(1 + \int_0^t \exp\left\{\frac{u^2}{2} + \Delta(u)\right\} b(u) du\right) \\ &= \exp\left\{-\frac{t^2}{2}\right\} + R_1(t) + R_2(t), \end{aligned} \quad (140)$$

где

$$\begin{aligned} R_1(t) &:= \exp\left\{-\frac{t^2}{2} - \Delta(t)\right\} - \exp\left\{-\frac{t^2}{2}\right\}, \\ R_2(t) &:= \exp\left\{-\frac{t^2}{2} - \Delta(t)\right\} \int_0^t \exp\left\{\frac{u^2}{2} + \Delta(u)\right\} b(u) du. \end{aligned}$$

Далее, используя равенство $\exp\{x + h\} - \exp\{x\} = h \int_0^1 \exp\{x + sh\} ds$, мы запишем

$$R_1(t) = -\Delta(t) \int_0^1 \exp\left\{-\frac{t^2}{2} - s\Delta(t)\right\} ds. \quad (141)$$

Используя оценку для функции $a(t)$, получаем

$$|\Delta(t)| \leq \left| \int_0^t ua(u) du \right| \leq \frac{A_0}{2} t^2 + \frac{A_1}{3} |t|^3.$$

Поскольку неравенство $A_0/2 + A_1|t|/3 \leq 1/4$ эквивалентно тому, что

$$-\frac{t^2}{2} - s\Delta(t) \leq -\frac{t^2}{4}, \quad (142)$$

мы получим

$$\begin{aligned} |R_1(t)| &= \left| \Delta(t) \int_0^1 \exp\left\{-\frac{t^2}{2} - s\Delta(t)\right\} ds \right| \\ &\leq \left(\frac{A_0}{2} t^2 + \frac{A_1}{3} |t|^3 \right) \exp\left\{-\frac{t^2}{4}\right\} \quad \text{для } |t| \leq T_1. \end{aligned} \quad (143)$$

Остаточный член $R_2(t)$ можно представить в виде

$$R_2(t) = \int_0^t \exp\left\{-\left(\frac{t^2}{2} - \frac{u^2}{2}\right) - (\Delta(u) - \Delta(t))\right\} b(u) du. \quad (144)$$

Нетрудно видеть, что для $0 \leq u \leq t$ и $t \leq u \leq 0$

$$|\Delta(t) - \Delta(u)| = \left| \int_u^t va(v) dv \right| \leq (A_0 + A_1|t|) \left| \int_u^t v dv \right| \leq \frac{A_0 + A_1|t|}{2} (t^2 - u^2). \quad (145)$$

Поскольку для t таких, что $|t| \leq (1 - 2A_0)/(2A_1) =: T_2$, имеет место неравенство $(A_0 + A_1|t|)/2 \leq 1/4$, мы получим

$$|R_2(t)| \leq \exp\left\{-\frac{t^2}{4}\right\} \int_0^{|t|} |b(u)| \exp\left\{\frac{u^2}{4}\right\} du \quad \text{для } |t| \leq T_2. \quad (146)$$

Отсюда нетрудно получить с учетом оценки $b(t)$, что

$$\begin{aligned} |R_2(t)| &\leq B_0 \exp\left\{-\frac{t^2}{4}\right\} \int_0^{|t|} \exp\left\{\frac{u^2}{4}\right\} du + B_1 \exp\left\{-\frac{t^2}{4}\right\} \int_0^{|t|} u \exp\left\{\frac{u^2}{4}\right\} du \\ &\quad + B_2 \exp\left\{-\frac{t^2}{4}\right\} \int_0^{|t|} u^2 \exp\left\{\frac{u^2}{4}\right\} du \\ &\leq 2B_0|F(T_2)| + 2B_1\left(1 - \exp\left\{-\frac{t^2}{4}\right\}\right) + 2B_2|t|\left(1 - \exp\left\{-\frac{t^2}{4}\right\}\right) \\ &\leq 2B_0|F(T_2)| + 2(B_1 + B_2|t|) \min\left\{1, \frac{t^2}{4}\right\}, \end{aligned} \quad (147)$$

где

$$F(t) = \exp\{-t^2\} \int_0^{u^2} du$$

– функция Доусона. Используя неравенство Эссеена [13; теорема 2.1], отсюда несложно получить требуемую оценку. Лемма 4 доказана.

Роллин использовал свой вариант леммы 3 для оценки скорости сходимости распределения случайной величины $W = \frac{T - \mathbb{E}T}{\sqrt{\mathbb{D}T}}$ – централизованного и нормированного числа треугольников T в случайном графе Эрдёша–Реньи $G(n, p)$, с вероятностью выбора ребра p . Сформулируем результат, полученный А. Роллином с применением леммы 4.

ТЕОРЕМА 6. Пусть $W = \frac{T - \mathbb{E}T}{\sqrt{\mathbb{D}T}}$, а Z – стандартная нормальная случайная величина. Существует универсальная постоянная $C > 0$ такая, что для любого $0 < p < 1$

$$\sup_x |\Pr\{W < x\} - \Pr\{Z < x\}| \leq \begin{cases} \frac{C}{n(1-p)^{1/2}}, & \text{если } \frac{1}{2} < p < 1, \\ \frac{C}{np^{1/2}}, & \text{если } n^{-1/2} < p \leq \frac{1}{2}, \\ \frac{C}{n^{3/2}p^{3/2}}, & \text{если } 0 < p \leq n^{-1/2}. \end{cases} \quad (148)$$

Подробному доказательству этого результата посвящена работа [63]. Лемма 6 является ключевым моментом этого доказательства.

6.2. Вариант П. Айхельсбахера и Б. Редносса. П. Айхельсбахер (P. Eichelsbacher) и Б. Редносс (B. Rednoß) предложили следующую версию леммы 3.

ЛЕММА 5. Пусть $\varphi(t) = \mathbb{E} \exp\{itW\}$ – характеристическая функция случайной величины W такой, что $\mathbb{E}|W| < \infty$. Пусть Z – случайная величина, имеющая стандартное нормальное распределение. Пусть $\alpha, \beta: \mathbb{R} \rightarrow \mathbb{C}$ – непрерывные ограниченные функции такие, что $\|\alpha\|_\infty < A < \infty$, $\|\beta\|_\infty \leq B < \infty$, где $A > 0$, $B \geq 0$, и

$$\varphi'(t) + t\varphi(t) = t^2(\alpha(t)\varphi(t) + \beta(t)) \quad \text{для всех } t \in \mathbb{R}. \quad (149)$$

Тогда

$$\sup_x |\Pr\{W < x\} - \Pr\{Z < x\}| \leq \frac{4}{3\sqrt{\pi}} + \frac{24\sqrt{2}}{\sqrt{\pi^3}} A + \frac{2}{\pi} \frac{B}{A}. \quad (150)$$

ДОКАЗАТЕЛЬСТВО. Требуемое утверждение немедленно следует из леммы 4, если положить

$$a(t) = t\alpha(t), \quad b(t) = \beta(t), \quad t \in \mathbb{R},$$

и

$$A_1 = A, \quad B_2 = B, \quad A_0 = B_0 = B_1 = 0, \quad t = 2A.$$

Лемма доказана.

Сформулируем результат, который был получен Айхельсбахером и Редноссом с применением леммы 5. Введем в рассмотрение следующую характеристику графа G_0 :

$$\Psi_{\min} := \Psi_{\min}(G_0) = \min\{n^{\nu_H} p^{e_H} : H \subset G_0, e_H \geq 1\}, \quad (151)$$

где $H \subset G_0$ означает подграф графа G_0 , ν_H – число вершин подграфа H , а e_H – число ребер подграфа H . Символом T будем обозначать число подграфов графа G_0 . Положим

$$W = \frac{T - \mathbb{E}T}{\sqrt{\mathbb{D}T}}.$$

Случайная величина Z , как и ранее, распределена по стандартному нормальному закону. Основной результат Айхельсбахера и Редносса заключается в следующем.

ТЕОРЕМА 7. Пусть T – число подграфов графа Эрдёша–Реньи $G(n, p)$, изоморфных графу G_0 и имеющих по крайней мере одно ребро. Пусть $p_0 \in (0, 1)$. Тогда существует константа C_{G_0, p_0} , зависящая только от графа G_0 и p_0 , такая, что

$$\sup_x |\Pr\{W < x\} - \Pr\{Z < x\}| \leq C_{G_0, p_0} \times \begin{cases} \frac{1}{n\sqrt{1-p}}, & \text{если } p > p_0, \\ \frac{1}{\sqrt{\Psi_{\min}}}, & \text{если } p \leq p_0. \end{cases} \quad (152)$$

6.3. Подход Ш. К. Форманова. Форманов [38], [39] предложил рассматривать предельное соотношение

$$\lim_{n \rightarrow \infty} \mathbf{A}(\mathbb{E} \exp\{itW_n\}) = 0 \quad (153)$$

как модификацию метода, описанного в предыдущих пунктах, для доказательства центральной предельной теоремы для последовательности случайных величин W_n . Пусть $\mathcal{F}_N$ – класс характеристических функций таких, что

$$f'(0) = 0, \quad f''(0) = -\sigma^2 > -\infty.$$

Форманов отметил одно из основных свойств оператора $\mathbf{A}$.

ЛЕММА 6. Пусть характеристические функции f, g принадлежат классу $\mathcal{F}_N$. Тогда

$$\mathbf{A}(fg) = f\mathbf{A}(g) + g\mathbf{A}(f). \quad (154)$$

ДОКАЗАТЕЛЬСТВО практически очевидно, но мы воспроизведем его здесь для удобства читателей. Пусть $\sigma_1^2 = -f''(0)$, $\sigma_2^2 = -g''(0)$. По определению оператора $\mathbf{A}$

$$\begin{aligned} \mathbf{A}(fg) &= (fg)' + t(-(fg)''_{t=0})fg = f'g + g'f + t(\sigma_1^2 + \sigma_2^2)fg \\ &= (f' + \sigma_1^2 t)fg + f(g' + \sigma_2^2 tg) = f\mathbf{A}(g) + g\mathbf{A}(f), \end{aligned} \quad (155)$$

что и доказывает лемму.

Чтобы сформулировать основной результат Форманова, введем следующие обозначения: $X_{n1}, X_{n2}, \dots$ – последовательность независимых случайных величин, образующих схему серий. Рассмотрим величину

$$S_n = X_{n1} + \dots + X_{nk_n},$$

причем k_n может быть как конечным, так и бесконечным. Следуя В. М. Золотарёву [77], предельные теоремы для распределения S_n без условия бесконечной малости называют неклассическими. Будем предполагать, что

$$\mathbb{E}X_{nj} = 0, \quad \sigma_{nj}^2 = \mathbb{E}X_{nj}^2, \quad j = 1, 2, \dots, \quad \sum_{j=1}^{k_n} \sigma_{nj}^2 = 1. \quad (156)$$

Пусть $F_n(x) = \Pr\{S_n < x\}$ – функция распределения суммы S_n , $f_{nj}(t) = \mathbb{E} \exp\{itX_{nj}\}$ – характеристическая функция случайной величины X_{nj} , а $\Phi(x)$ – функция распределения стандартного нормального закона. Результат Форманова состоит в следующем.

ТЕОРЕМА 8. Пусть выполнено условие (156). Тогда

$$\sup_x |F_n(x) - \Phi(x)| \rightarrow 0 \quad \text{при } n \rightarrow \infty \quad (157)$$

в том и только том случае, когда при любом t

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{k_n} |\mathbf{A}(f_{nj})(t)| = 0. \quad (158)$$

Доказательство этой теоремы не представляет трудностей, и читатель может найти его в работе [39].

6.4. Новый подход к описанию метода Стейна–Тихомирова. В настоящем пункте мы представим подход, предложенный в работе [2]. Рассмотрим последовательность случайных величин X_n с характеристическими функциями $\varphi_n(t) = \mathbb{E} \exp\{itX_n\}$, и пусть X_∞ – случайная величина, имеющая характеристическую функцию $\varphi_\infty(t) = \mathbb{E} \exp\{itX_\infty\}$. Оценивается разность характеристических функций

$$\Delta_n(t) = \varphi_n(t) - \varphi_\infty(t) \quad (159)$$

или, что эквивалентно, $|\Delta_n(t)| = |\varphi_n(t) - \varphi_\infty(t)|$. Оценка близости к нулю величины $\Delta_n(t)$ влечет, как уже отмечалось в п. 6.1, оценку близости к нулю в определенной метрике распределений случайных величин X_n и X_∞ . В случае нормального распределения X_∞ мы можем воспользоваться неравенством Эссеена [13; теорема 2.1]. В общем случае можно записать, что

$$\Delta_n(t) = \varphi_n(t) - \varphi_\infty(t) = (\mathcal{D}_\infty \circ L_\infty)\varphi_n(t) \quad (160)$$

с дифференциальным оператором L_∞ , определенным соотношением

$$\varphi \rightarrow L_\infty\varphi(t) = \varphi'(t) - \frac{\varphi'_\infty(t)}{\varphi_\infty(t)} \varphi(t), \quad (161)$$

и интегральным оператором

$$\mathcal{D}_\infty f(t) = \varphi_\infty(t) \int_0^t \frac{1}{\varphi_\infty(u)} f(u) du. \quad (162)$$

В случае аппроксимации нормальным законом дифференциальный оператор имеет вид

$$L_\infty\varphi(t) = \varphi'(t) + t\varphi(t), \quad (163)$$

а интегральный оператор – вид

$$\mathcal{D}_\infty\varphi(t) = \exp\left\{-\frac{t^2}{2}\right\} \int_0^t \exp\left\{\frac{u^2}{2}\right\} du. \quad (164)$$

Авторы обзора [2] предложили следующий вариант описания метода.

- Шаг 1: Вычисление оператора $L_\infty(\varphi_n - \varphi_\infty)(t)$. Используя свойства аппроксимирующей характеристической функции, получаем оценку вида

$$|L_\infty(\varphi_n - \varphi_\infty)(t)| \leq \Delta(F_n, F_\infty)\psi(t), \quad (165)$$

где $\Delta(F_n, F_\infty) > 0$ – некоторая подходящая вычислимая величина, $\psi(t) \geq 0$ – точно известная функция.

- Шаг 2: Используя выведенную на первом шаге оценку, получаем

$$|\varphi_n(t) - \varphi_\infty(t)| \leq \Delta(F_n, F_\infty)|\varphi_\infty(t)| \int_0^t \frac{1}{|\varphi_\infty(u)|} \psi(u) du =: \mathcal{D}_\infty\psi(t). \quad (166)$$

Затем – изучение свойств дифференциального оператора $\mathcal{D}_\infty\psi(t)$.

• *Шаг 3:* Используя информацию, полученную на шаге 1 и шаге 2, а также неравенства, позволяющие оценивать различные метрики в пространстве распределений через их характеристические функции (см., например, [13]), переходим к непосредственным оценкам близости распределений.

Например, в случае аппроксимации вероятности попадания в шар суммы гильбертовозначных векторов с ковариационным оператором $\mathbf{B}$, рассмотренном в п. 5.2, мы можем указать выражения для дифференциального оператора:

$$L_\infty \varphi(t) = \varphi'(t) - i \operatorname{Tr} \mathbf{R}_t \mathbf{B} \varphi(t), \quad (167)$$

и интегрального оператора:

$$\mathcal{D}_\infty \psi(t) = \frac{1}{\operatorname{Tr} \mathbf{R}_t \mathbf{B}} \int_0^t \psi(u) \operatorname{Tr} \mathbf{R}_u \mathbf{B} du. \quad (168)$$

6.5. Оценка скорости сходимости к сопровождающему распределению Пуассона. Мы рассмотрим аппроксимацию сопровождающим распределением Пуассона (compound Poisson distribution). Модель, которую мы будем изучать в этом пункте, описана в работе [29]. Пусть $\varphi_\infty(t)$ – характеристическая функция сопровождающего закона Пуассона с интенсивностью μ и характеристической функцией скачков, равной $\varkappa(t)$. Функцию $\varphi_\infty(t)$ можно представить в виде

$$\varphi_\infty(t) = \exp\{\mu(\varkappa(t) - 1)\}, \quad (169)$$

или

$$\varphi'_\infty(t) = \mu \varphi'(t) \varphi_\infty(t), \quad \varphi_\infty(0) = 1. \quad (170)$$

В терминах из предыдущего пункта дифференциальный оператор имеет вид

$$L_\infty \varphi(t) = \varphi'(t) - \mu \varkappa'(t) \varphi(t), \quad (171)$$

а интегральный оператор – вид

$$\mathcal{D}_\infty \varphi(t) = \exp\{\mu(\varkappa(t) - 1)\} \int_0^t \exp\{-\mu(\varkappa(u) - 1)\} \varphi(u) du. \quad (172)$$

Чтобы описать последовательность случайных величин W_n , рассмотрим треугольный массив случайных величин Y_{nj} , $n \geq 1$, $j = 1, \dots, n$. Пусть случайные величины Y_{nj} согласованы с последовательностью σ -алгебр $\mathcal{F}_j$, $j = 1, 2, \dots$. Относительно случайных величин Y_{nj} будем предполагать выполнение следующих условий.

УСЛОВИЕ (A1). Имеют место оценки

$$\Pr\{Y_{nj} \neq 0\} \leq \frac{C}{n}, \quad \mathbb{E}|Y_{nj}| \leq \frac{C_1}{n}, \quad j = 1, \dots, n, \quad (173)$$

и

$$\mathbb{E}|T_{nj}| \mathbb{I}\{Y_{ni} \neq 0\} \leq \left(\frac{C}{n}\right)^2, \quad i \neq j. \quad (174)$$

УСЛОВИЕ (A2). Существует $l \geq 1$ такое, что

$$|E(Y_{nj} \mid \mathcal{F}_i) - EY_{nj}| \leq \frac{C_1}{n} \alpha(j-l), \quad i \leq j-l, \quad (175)$$

где $\alpha(n)$ – убывающая последовательность такая, что $\lim_{n \rightarrow \infty} \alpha(n) = 0$.

УСЛОВИЕ (A3). Для любой измеримой функции $v(x)$, $x \in \mathbb{R}^{n-j+1}$, такой, что $|v(x)| \leq 1$, справедливо неравенство

$$|E(v(Y_{nj}, \dots, Y_{nn}) \mid \mathcal{F}_i) - Ev(Y_{nj}, \dots, Y_{nn})| \leq \alpha(j-i), \quad i < j. \quad (176)$$

УСЛОВИЕ (A4). Пусть $\varphi_{nj}(t) = E \exp\{itY_{nj}\}$ – характеристическая функция случайной величины Y_{nj} . Существуют характеристическая функция $\varphi(t)$ и положительная постоянная $C_2 > 0$ такая, что

$$|\varphi'_{nj}(t) - \varphi'(t)| \leq C_2 n^{-2}, \quad t \in \mathbb{R}. \quad (177)$$

УСЛОВИЕ (A5). Существуют постоянная $C_3 > 0$ и число $r \in (0, 1)$ такие, что для всех $k \geq 1$

$$\alpha(k) \leq C_3 r^k. \quad (178)$$

Положим $W_n = \sum_{j=1}^n Y_{n,j}$. Пусть Z – случайная величина, имеющая сопрягающее распределение Пуассона с интенсивностью μ и скачками с характеристической функцией $\varphi(t)$, т. е. распределение с характеристической функцией

$$\psi(t) = E \exp\{itS\} = \exp\{\mu(\varphi(t) - 1)\}.$$

Определим расстояние Леви между любыми двумя функциями распределения $F(x) = \Pr\{\xi < x\}$ и $G(x) = \Pr\{\eta < x\}$ случайных величин ξ и η равенством

$$L(F, G) = L(\xi, \eta) = \inf\{h > 0: G(x-h) - h < F(x) < G(x+h) + h\}. \quad (179)$$

Следующий результат получен Ф. Клебанером с применением метода характеристических функций.

ТЕОРЕМА 9. Пусть выполнены условия (A1)–(A5). Тогда найдется постоянная $C > 0$ такая, что

$$L(W_n, Z) \leq C n^{-1/2} \log n. \quad (180)$$

ДОКАЗАТЕЛЬСТВО основано на выводе для разности характеристических функций $\Delta_n(t) = \psi_n(t) - \psi(t)$ уравнения вида

$$\Delta'_n(t) = \mu \varphi'(t) \Delta_n(t) + r_n(t). \quad (181)$$

В обозначениях раздела 6 дифференциальный и интегральный операторы имеют вид

$$L_{\infty}(f)(t) = \mu\varphi'(t)f(t) - f'(t)$$

и

$$\mathcal{D}_{\infty}(r_n)(t) = \int_0^t \exp\{\mu(\varphi(t) - \varphi(s))\} r_n(s) ds$$

соответственно. В условиях теоремы справедлива оценка

$$|r_n(t)| = |L_{\infty}(\psi_n)(t)| \leq C_2 \frac{1}{n} + 3C_1 \alpha(b \log n) + 8C_1^2 b \frac{\log n}{n}, \quad (182)$$

где b – некоторая постоянная такая, что $b \log n$ – положительное целое число (доказательство оценки (182) приведено в [29; с. 8–10]). Поскольку $|\varphi(t)| \leq 1$, отсюда немедленно следует, что

$$|\Delta_n(t)| \leq |\mathcal{D}_{\infty} r_n(t)| \leq C|t| \left(C_2 \frac{1}{n} + 3C_1 \alpha(b \log n) + 8C_1^2 b \frac{\log n}{n} \right). \quad (183)$$

Теперь, используя неравенство Золотарёва (см., например, [13]), нетрудно получить требуемую оценку. Теорема доказана.

7. Подход В. Бенткуса и применение этого подхода к случайным матрицам

Мы рассмотрим задачу аппроксимации распределения функций от независимых случайных величин $X_1, \dots, X_n$, имеющих распределение $F(x)$, с помощью функций от независимых случайных величин $Y_1, Y_2, \dots, Y_n$ (не зависящих от $(X_1, X_2, \dots)$) с распределением $G(x)$. Аппроксимируемые величины:

$$W_n = H_n(X_1, \dots, X_n),$$

аппроксимирующая величина: $Z_n = H_n(Y_1, \dots, Y_n)$. Например, аппроксимация распределений сумм случайных величин нормальным законом может быть описана предлагаемой схемой. Функция H_n имеет вид

$$H_n(x_1, \dots, x_n) = \frac{1}{\sqrt{n}} \sum_{j=1}^n x_j.$$

Если в качестве случайных величин $Y_1, \dots, Y_n$ рассматривать независимые стандартные нормальные величины, то распределение случайной величины $Z_n = H_n(Y_1, \dots, Y_n)$ не зависит от n и совпадает со стандартным нормальным законом.

Введем в рассмотрение случайные величины

$$Z_j(\alpha) = X_j \cos \alpha + Y_j \sin \alpha, \quad j = 1, 2, \dots, \quad \alpha \in \left[0, \frac{\pi}{2}\right].$$

Пусть

$$\varphi_n(t, \alpha) = \mathbb{E} \exp\{itH_n(\alpha)\}.$$

Заметим, что

$$\varphi_n(t, 0) = \varphi_n(t) = \mathbb{E} \exp\{itW_n\}, \quad \varphi_n\left(t, \frac{\pi}{2}\right) = \psi_n(t) = \mathbb{E} \exp\{itZ_n\}.$$

Таким образом, мы можем записать

$$\psi_n(t) - \varphi_n(t) = \int_0^t \frac{\partial \varphi_n(t, \alpha)}{\partial \alpha} d\alpha. \quad (184)$$

При таком подходе не требуется характеристизация аппроксимирующего распределения. Остается только проблема оценки функции $\frac{\partial \varphi_n(t, \alpha)}{\partial \alpha}$. В случае сумм случайных величин имеем

$$\frac{\partial \varphi_n(t, \alpha)}{\partial \alpha} = \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}(-X_j \sin \alpha + Y_j \cos \alpha) \exp\{itH_n(\alpha)\}. \quad (185)$$

Выражения типа $\sum_{j=1}^n \mathbb{E} X_j f(X_1, \dots, X_n)$ ассоциируются с методом Стейна. Для сумм случайных величин, принимающих значения в различных пространствах, этот подход подробно изложен в работе В. Бенткуса [11].

В настоящем разделе мы приведем пример применения указанного выше подхода в теории случайных матриц. Пусть

$$\mathbf{W}_n = \frac{1}{\sqrt{n}} (X_{ij})_{i,j=1}^n$$

– случайная симметричная матрица ($\mathbf{W}_n = \mathbf{W}_n^\top$, где $(\cdot)^\top$ означает транспонирование). Будем предполагать, что элементы матрицы – это независимые (с точностью до симметрии) случайные величины с $\mathbb{E} X_{ij} = 0$ и $\mathbb{E} X_{ij}^2 = \sigma_{ij}^2 < \infty$. Кроме того, рассмотрим симметричную матрицу

$$\mathbf{Z}_n = \frac{1}{\sqrt{n}} (Y_{jk})_{j,k=1}^n,$$

где Y_{jk} – независимые (с точностью до симметрии) и не зависящие от X_{ij} , $i, j = 1, \dots, n$, случайные величины с $\mathbb{E} Y_{jk} = 0$ и $\mathbb{E} Y_{jk}^2 = \sigma_{jk}^2 < \infty$. Обозначим

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

собственные числа матрицы $\mathbf{W}_n$, упорядоченные по убыванию, и

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$$

собственные числа матрицы $\mathbf{Z}_n$. Одним из основных объектов исследования в теории случайных матриц является эмпирическая функция распределения, определяемая для введенных выше матриц $\mathbf{W}_n$ и $\mathbf{Z}_n$ равенствами

$$F_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbb{I}\{\lambda_j < x\}, \quad G_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbb{I}\{\mu_j < x\}, \quad (186)$$

где $\mathbb{I}\{A\}$ обозначает индикатор события A . Например, мы можем оценить близость ожидаемых спектральных функций распределения матриц $\mathbf{W}_n$ и $\mathbf{Z}_n$, а именно величину

$$\Delta_n := \sup_x |\mathbb{E}F_n(x) - \mathbb{E}G_n(x)|. \quad (187)$$

В качестве одного из основных инструментов оценки величин типа Δ_n в теории случайных матриц используется преобразование Стилтеса. Для любой случайной величины ξ преобразование Стилтеса определяется равенством

$$s(z) := \mathbb{E} \frac{1}{\xi - z}, \quad (188)$$

где $z = u + iv$ – любые комплексные числа с положительной мнимой частью $v > 0$. Это преобразование позволяет, в частности, оценивать близость функций распределения. Одно из неравенств было получено, например, в [14; предложение 2.1].

ЛЕММА 7. Пусть F и G – функции распределения случайных величин ξ и η . Пусть α, β выбраны так, что

$$G(\alpha) \leq v_0 \quad \text{и} \quad 1 - G(\beta) \leq v_0.$$

Тогда существует абсолютная постоянная $C > 0$ такая, что для любых $v_1 > v_0$ расстояние Леви $L(F, G)$ между функциями распределения F и G допускает оценку

$$\begin{aligned} L(F, G) \leq & v_0 + v_0 \sup_x \operatorname{Im} S_G(x + iv_0) + \int_{-\infty}^{\infty} |S_F(u + iv_1) - S_G(u + iv_1)| du \\ & + \sup_{x \in [\alpha - 2v_0, \beta + 2v_0]} \int_{v_0}^{v_1} |(S_F(u + iv) - S_G(u + iv))| dv, \end{aligned} \quad (189)$$

где

$$S_F(z) = \int_{-\infty}^{\infty} \frac{1}{x - z} dF(x), \quad S_G(z) = \int_{-\infty}^{\infty} \frac{1}{x - z} dG(x). \quad (190)$$

Преобразование Стилтеса для эмпирических спектральных функций распределения случайных матриц легко выражается через алгебраические характеристики этих матриц. Введем в рассмотрение резольвентные матрицы матриц $\mathbf{W}_n$ и $\mathbf{Z}_n$:

$$\mathbf{R}_W(z) = (\mathbf{W}_n - z\mathbf{I})^{-1} \quad \text{и} \quad \mathbf{R}_Z(z) = (\mathbf{Z}_n - z\mathbf{I})^{-1}.$$

Здесь и далее $\mathbf{I}$ обозначает единичную матрицу соответствующей размерности. Имеют место представления

$$\begin{aligned} S_W(z) &= \int_{-\infty}^{\infty} \frac{1}{x - z} dF_n(x) = \frac{1}{n} \sum_{j=1}^n \frac{1}{\lambda_j - z} = \frac{1}{n} \operatorname{Tr} \mathbf{R}_W(z), \\ S_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{x - z} dG_n(x) = \frac{1}{n} \sum_{j=1}^n \frac{1}{\mu_j - z} = \frac{1}{n} \operatorname{Tr} \mathbf{R}_Z(z). \end{aligned} \quad (191)$$

Введем в рассмотрение матрицы

$$\mathbf{W}_n(\alpha) = \frac{1}{\sqrt{n}}(Z_{jk}(\alpha))_{j,k=1}^n = \frac{1}{\sqrt{n}}(X_{jk} \cos \alpha + Y_{jk} \sin \alpha)_{j,k=1}^n, \quad \alpha \in \left[0, \frac{\pi}{2}\right].$$

Мы докажем следующую теорему.

ТЕОРЕМА 10. Пусть случайные величины X_{jk} и Y_{jk} , $j, k = 1, 2, \dots$, независимы и удовлетворяют следующим условиям:

- (a) $\mathbf{E}X_{jk} = \mathbf{E}Y_{jk} = 0$;
- (b) $\mathbf{E}X_{jk}^2 = \mathbf{E}Y_{jk}^2 = \sigma_{jk}^2 < \infty$;
- (c) существует постоянная $\mu_3 > 0$ такая, что

$$\sup_{j,k \geq 1} \max\{\mathbf{E}|X_{jk}|^3, \mathbf{E}|Y_{jk}|^3\} \leq \mu_3 < \infty.$$

Тогда

$$L(\mathbf{E}F_n, \mathbf{E}G_n) \rightarrow 0 \quad \text{при } n \rightarrow \infty. \quad (192)$$

Если существует постоянная C_0 такая, что $\sup_{u \in \mathbb{R}} \operatorname{Im} S_G(u + iv) \leq C_0$, то найдется постоянная $C > 0$ такая, что

$$L(\mathbf{E}F_n, \mathbf{E}G_n) \leq C\mu_3 n^{-1/8}. \quad (193)$$

ЗАМЕЧАНИЕ 3. Мы, разумеется, не задавались целью получить оптимальную по n оценку. Примечательно в этой теореме то, что при любом профиле дисперсий (мы не налагаем никаких ограничений на σ_{jk}^2) верно следующее: если у элементов матриц совпадают первые два момента и третьи моменты ограничены, то предельные ожидаемые спектральные распределения этих матриц совпадают.

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 10. Положим

$$\mathbf{R}(z, \alpha) = (\mathbf{W}_n(\alpha) - z\mathbf{I})^{-1}, \quad S(z, \alpha) = \frac{1}{n} \operatorname{Tr} \mathbf{R}(z, \alpha). \quad (194)$$

Нетрудно видеть, что

$$S_Z(z) - S_W(z) = \int_0^{\pi/2} \frac{\partial S(z, \alpha)}{\partial \alpha} d\alpha. \quad (195)$$

Далее,

$$\frac{\partial S(z, \alpha)}{\partial \alpha} = \frac{1}{n\sqrt{n}} \sum_{j=1}^n \frac{\partial R_{jj}(z, \alpha)}{\partial \alpha} = \frac{1}{n} \sum_{j=1}^n \sum_{k,l=1}^n \frac{\partial R_{jj}(z, \alpha)}{\partial Z_{kl}} \frac{\partial Z_{kl}}{\partial \alpha} \quad (196)$$

($Z_{kl} = Z_{kl}(\alpha)$). Используя следующие формулы дифференцирования элементов обратной матрицы (здесь A_{ij}^{-1} – элементы матрицы, обратной к $\mathbf{A} = (A_{ij})$):

$$\frac{\partial A_{jj}^{-1}}{\partial A_{kl}} = -A_{jk}^{-1} A_{jl}^{-1}, \quad (197)$$

для резольвентной матрицы, учитывая ее симметрию, мы получим

$$\frac{\partial R_{jj}(z, \alpha)}{\partial Z_{kl}} = -R_{jl}(z, \alpha)R_{jk}(z, \alpha)(2 - \delta_{kl}), \quad (198)$$

где δ_{kl} – символ Кронекера. (См., например, [52].) Подставляя равенство (198) в равенство (196), приходим к соотношению

$$\begin{aligned} \frac{\partial S(z, \alpha)}{\partial \alpha} &= \frac{1}{n\sqrt{n}} \sum_{j=1}^n \sum_{k,l=1}^n (-X_{kl} \sin \alpha + Y_{kl} \cos \alpha) \\ &\quad \times R_{jk}(z, \alpha)R_{jl}(z, \alpha)(2 - \delta_{kl}). \end{aligned} \quad (199)$$

Далее, введем в рассмотрение ортонормированные векторы в $\mathbb{R}^n$:

$$\mathbf{e}_j = (0, \dots, 0, 1, 0, \dots, 0)^\top \quad \text{с единицей на } j\text{-м месте.}$$

Пусть $\mathbf{W}_n^{(kl)}(\alpha)$ обозначает матрицу, полученную из матрицы $\mathbf{W}_n(\alpha)$ заменой элемента $Z_{kl}(\alpha)$ нулем. Соответствующие резольвентные матрицы обозначим

$$\mathbf{R}^{(kl)}(z, \alpha) = (\mathbf{W}_n^{(kl)}(\alpha) - z\mathbf{I})^{-1}.$$

Используя определение резольвенты, нетрудно вывести следующее представление:

$$\mathbf{R}(z, \alpha) = \mathbf{R}^{(kl)}(z, \alpha) - \frac{1}{\sqrt{n}} Z_{kl}(\alpha) \mathbf{R}^{(kl)}(z, \alpha) (\mathbf{e}_k \mathbf{e}_l^\top + \mathbf{e}_l \mathbf{e}_k^\top) \mathbf{R}(z, \alpha). \quad (200)$$

Последнее равенство влечет

$$\begin{aligned} R_{jk}(z, \alpha) &= R_{jk}^{(kl)}(z, \alpha) - \frac{1}{\sqrt{n}} Z_{kl}(\alpha) R_{jk}^{(kl)}(z, \alpha) R_{lk}(z, \alpha) \\ &\quad - \frac{1}{\sqrt{n}} Z_{kl}(\alpha) R_{jl}^{(kl)}(z, \alpha) R_{kk}(z, \alpha), \\ R_{jl}(z, \alpha) &= R_{jl}^{(kl)}(z, \alpha) - \frac{1}{\sqrt{n}} Z_{kl}(\alpha) R_{jk}^{(kl)}(z, \alpha) R_{ll}(z, \alpha) \\ &\quad - \frac{1}{\sqrt{n}} Z_{kl}(\alpha) R_{jl}^{(kl)}(z, \alpha) R_{kl}(z, \alpha). \end{aligned} \quad (201)$$

Подставляя эти равенства в (199), получим

$$\frac{\partial S(z, \alpha)}{\partial \alpha} =: S_1 + S_2 + S_3 + S_4, \quad (202)$$

где

$$\begin{aligned}
 S_1 &= \frac{1}{n\sqrt{n}} \sum_{j=1}^n \sum_{k,l=1}^n (-X_{kl} \sin \alpha + Y_{kl} \cos \alpha) R_{jl}^{(kl)}(z, \alpha) R_{jk}^{(kl)}(z, \alpha), \\
 S_2 &= \frac{1}{n^2} \sum_{j=1}^n \sum_{k,l=1}^n (-X_{kl} \sin \alpha + Y_{kl} \cos \alpha) Z_{jk}(\alpha) \\
 &\quad \times \left((R_{jk}^{(kl)}(z, \alpha) R_{lk}^{(kl)}(z, \alpha) + R_{jl}^{(kl)}(z, \alpha) R_{kk}^{(kl)}(z, \alpha)) R_{jl}^{(kl)}(z, \alpha) \right. \\
 &\quad \left. + (R_{jk}^{(kl)}(z, \alpha) R_{lk}^{(kl)}(z, \alpha) + R_{jl}^{(kl)}(z, \alpha) R_{kl}^{(kl)}(z, \alpha)) R_{jk}^{(kl)}(z, \alpha) \right), \\
 S_3 &= \frac{1}{n^2\sqrt{n}} \sum_{j=1}^n \sum_{k,l=1}^n (-X_{kl} \sin \alpha + Y_{kl} \cos \alpha) Z_{kl}^2(\alpha) \\
 &\quad \times \left((R_{jk}^{(kl)}(z, \alpha))^2 R_{lk}(z, \alpha) R_{ll}(z, \alpha) + R_{jl}^{(kl)}(z, \alpha) R_{jk}^{kl}(z, \alpha) R_{kk}(z, \alpha) R_{ll}(z, \alpha) \right. \\
 &\quad \left. + (R_{jl}^{(kl)}(z, \alpha))^2 R_{kl}(z, \alpha) R_{kk}(z, \alpha) + R_{jk}^{(kl)}(z, \alpha) R_{jl}^{(kl)}(z, \alpha) (R_{lk}(z, \alpha))^2 \right)
 \end{aligned}$$

и (для краткости мы опускаем аргументы (z, α) у элементов матриц $\mathbf{R}(z, \alpha)$, $\mathbf{R}^{(kl)}(z, \alpha)$)

$$\begin{aligned}
 S_4 &= \frac{1}{n^2\sqrt{n}} \sum_{j=1}^n \sum_{k,l=1}^n (-X_{kl} \sin \alpha + Y_{kl} \cos \alpha) Z_{kl}^2(\alpha) \\
 &\quad \times \left((R_{jk}^{(kl)})^2 R_{lk}^{(kl)} R_{ll} + (R_{jk}^{(kl)})^2 R_{ll}^{(kl)} R_{kl} + R_{jk}^{(kl)} R_{jl}^{(kl)} R_{kk}^{(kl)} R_{ll} \right. \\
 &\quad \left. + R_{jk}^{(kl)} R_{jl}^{(kl)} R_{kl}^{(kl)} R_{kl} + R_{jl}^{(kl)} R_{jk}^{(kl)} R_{lk}^{(kl)} R_{ll} \right. \\
 &\quad \left. + R_{jl}^{(kl)} R_{jk}^{(kl)} R_{lk}^{(kl)} R_{kl} + (R_{jl}^{(kl)})^2 R_{kk}^{(kl)} R_{ll} + (R_{jl}^{(kl)})^2 R_{kl}^{(kl)} R_{kk} \right). \quad (203)
 \end{aligned}$$

Поскольку случайные величины X_{kl} , Y_{kl} и матрица $\mathbf{R}^{(kl)}(z, \alpha)$ независимы, мы имеем

$$ES_1 = 0. \quad (204)$$

Далее, заметим, что

$$E(-X_{kl} \sin \alpha + Y_{kl} \cos \alpha) Z_{kl} = (-EX_{kl}^2 + EY_{kl}^2) \cos \alpha \sin \alpha = 0. \quad (205)$$

Отсюда в силу независимости случайных величин X_{kl} , Y_{kl} , $Z_{kl}(\alpha)$ и матрицы $\mathbf{R}^{(kl)}(z, \alpha)$ мы получим

$$ES_2 = 0. \quad (206)$$

Для оценки величин $E|S_3|$ и $E|S_4|$ воспользуемся простым неравенством

$$\max\{\|\mathbf{R}(z, \alpha)\|, \|\mathbf{R}^{(kl)}(z, \alpha)\|, k, l = 1, \dots, n\} \leq \frac{1}{v}, \quad z = u + iv,$$

которое влечет также, что

$$\sum_{l=1}^n |R_{jl}(z, \alpha)|^2 \leq \frac{1}{v^2}, \quad \sum_{j=1}^n |R_{jl}^{(kl)}(z, \alpha)|^2 \leq \frac{1}{v^2} \quad (207)$$

для любых $j = 1, \dots, n$ и $k, l = 1, \dots, n$. Используя приведенные неравенства, нетрудно получить оценку

$$\max\{E|S_3|, E|S_4|\} \leq C\mu_3 \frac{1}{v^4\sqrt{n}}. \quad (208)$$

А именно, все слагаемые, возникающие в правой части неравенств для указанного максимума, вытекающие из определения величин S_3, S_4 , оцениваются одинаково. Например,

$$\begin{aligned} & \frac{1}{n^2\sqrt{n}} \sum_{j=1}^n \sum_{k,l=1}^n E|-X_{kl} \sin \alpha + Y_{kl} \cos \alpha| |Z_{kl}|^2 |R_{jk}^{(kl)}(z, \alpha)|^2 R_{lk}(z, \alpha) R_{ll}(z, \alpha)| \\ & \leq \frac{1}{n^2\sqrt{n}} \sum_{k,l=1}^n E(|X_{lk}|^3 + E|Y_{kl}|^3) \left(\sum_{j=1}^n |R_{jk}^{(kl)}(z, \alpha)|^2 \right) |R_{lk}(z, \alpha) R_{ll}(z, \alpha)| \\ & \leq \frac{1}{v^4 n^2 \sqrt{n}} \sum_{k,l=1}^n E(|X_{lk}|^3 + E|Y_{kl}|^3) \leq \frac{C\mu_3}{v^4\sqrt{n}}. \end{aligned} \quad (209)$$

Аналогично,

$$\begin{aligned} & \frac{1}{n^2\sqrt{n}} \sum_{j=1}^n \sum_{k,l=1}^n E|-X_{kl} \sin \alpha + Y_{kl} \cos \alpha| |Z_{kl}|^2 |R_{jk}^{(kl)} R_{jl}^{(kl)} R_{kk}^{(kl)} R_{ll}| \\ & \leq \frac{1}{n^2\sqrt{n}} \sum_{k,l=1}^n E(|X_{lk}|^3 + E|Y_{kl}|^3) \left(\sum_{j=1}^n |R_{jk}^{(kl)}(z, \alpha) R_{jl}^{(kl)}(z, \alpha)| \right) \\ & \quad \times |R_{lk}(z, \alpha) R_{ll}(z, \alpha)| \\ & \leq \frac{1}{n^2\sqrt{n}} \sum_{k,l=1}^n E(|X_{lk}|^3 + E|Y_{kl}|^3) \left(\sum_{j=1}^n |R_{jk}^{(kl)}(z, \alpha)|^2 \right)^{1/2} \\ & \quad \times \left(\sum_{j=1}^n |R_{jl}^{(kl)}(z, \alpha)|^2 \right)^{1/2} |R_{lk}(z, \alpha) R_{ll}(z, \alpha)| \\ & \leq \frac{C\mu_3}{v^4\sqrt{n}}. \end{aligned} \quad (210)$$

Таким образом, мы получим неравенство

$$|S_W(z) - S_Z(z)| \leq \frac{C\mu_3}{v^4\sqrt{n}}. \quad (211)$$

Соотношение (192) следует теперь из свойств преобразования Стилтеса, а неравенство (193) – из леммы 7 с $v_1 = 1, v_0 = n^{-1/8}$. Теорема 10 доказана.

Развитие метода Стейна в теории случайных матриц с использованием преобразования Стилтеса можно найти также в работах [45], [48], [44].

Список литературы

- [1] B. Arras, C. Houdré, *On Stein's method for infinitely divisible laws with finite first moment*, SpringerBriefs Probab. Math. Stat., Springer, Cham, 2019, xi+104 pp.
- [2] B. Arras, G. Mijoule, G. Poly, Y. Swan, *A new approach to the Stein–Tikhomirov method: with applications to the second Wiener chaos and Dickman convergence*, 2017 (v1 – 2016), 45 pp., arXiv: 1605.06819.
- [3] A. D. Barbour, “Poisson convergence and random graphs”, *Math. Proc. Cambridge Philos. Soc.*, **92**:2 (1982), 349–359.
- [4] A. D. Barbour, “Asymptotic expansions based on smooth functions in the central limit theorem”, *Probab. Theory Relat. Fields*, **72**:2 (1986), 289–303.
- [5] A. D. Barbour, “Asymptotic expansions in the Poisson limit theorem”, *Ann. Probab.*, **15**:2 (1987), 748–766.
- [6] A. D. Barbour, “Stein's method and Poisson process convergence”, *J. Appl. Probab.*, **25**:A (1988), 175–184.
- [7] A. D. Barbour, “Stein's method for diffusion approximations”, *Probab. Theory Related Fields*, **84**:3 (1990), 297–322.
- [8] A. D. Barbour, L. H. Y. Chen (eds.), *An introduction to Stein's method*, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., **4**, Singapore Univ. Press, Singapore; World Sci. Publ., Hackensack, NJ, 2005, xii+225 pp.
- [9] A. D. Barbour, L. H. Y. Chen, Wei-Liem Loh, “Compound Poisson approximation for nonnegative random variables via Stein's method”, *Ann. Probab.*, **20**:4 (1992), 1843–1866.
- [10] A. D. Barbour, L. Holst, S. Janson, *Poisson approximation*, Oxford Stud. Probab., **2**, The Clarendon Press, Oxford Univ. Press, New York, 1992, x+277 pp.
- [11] В. Бенткус, “Новый подход к аппроксимациям в теории вероятностей и теории операторов”, *Liet. Mat. Rink.*, **43**:4 (2003), 444–470; англ. пер.: V. Bentkus, “A new method for approximations in probability and operator theories”, *Lithuanian Math. J.*, **43**:4 (2003), 367–388.
- [12] V. Bentkus, F. Götze, A. Tikhomirov, “Berry–Esseen bounds for statistics of weakly dependent samples”, *Bernoulli*, **3**:3 (1997), 329–349.
- [13] С. Г. Бобков, “Близость вероятностных распределений в терминах преобразований Фурье–Стилтьеса”, *УМН*, **71**:6(432) (2016), 37–98; англ. пер.: S. G. Bobkov, “Proximity of probability distributions in terms of Fourier–Stieltjes transforms”, *Russian Math. Surveys*, **71**:6 (2016), 1021–1079.
- [14] S. G. Bobkov, F. Götze, A. N. Tikhomirov, “On concentration of empirical measures and convergence to the semi-circle law”, *J. Theoret. Probab.*, **23**:3 (2010), 792–823.
- [15] E. Bolthausen, “On the central limit theorem for stationary mixing random fields”, *Ann. Probab.*, **10**:4 (1982), 1047–1050.
- [16] G. Bresler, D. Nagaraj, “Stein's method for stationary distributions of Markov chains and application to Ising models”, *Ann. Appl. Probab.*, **29**:5 (2019), 3230–3265.
- [17] A. V. Bulinski, P. Doukhan, “Vitesse de convergence dans le théorème de limite centrale pour les champs mélangeants satisfaisant des hypothèses de moments faibles”, *C. R. Acad. Sci. Paris Sér. I Math.*, **311**:12 (1990), 801–805.
- [18] A. Bulinski, A. Shashkin, *Limit theorems for associated random fields and related systems*, Adv. Ser. Stat. Sci. Appl. Probab., **10**, World Sci. Publ., Hackensack, NJ, 2007, xii+436 pp.
- [19] A. Bulinski, N. Slepov, “Sharp estimates for proximity of geometric and related sums distributions to limit laws”, *Mathematics*, **10**:24 (2022), 4747, 37 pp.

- [20] S. Chatterjee, “A short survey of Stein’s method”, *Proceedings of the international congress of mathematicians (ICM 2014, Seoul)*, v.IV, Kyung Moon Sa, Seoul, 2014, 1–24.
- [21] S. Chatterjee, Qi-Man Shao, “Nonnormal approximation by Stein’s method of exchangeable pairs with application to the Curie–Weiss model”, *Ann. Appl. Probab.*, **21**:2 (2011), 464–483.
- [22] L. H. Y. Chen, “Poisson approximation for dependent trials”, *Ann. Probab.*, **3**:3 (1975), 534–545.
- [23] L. H. Y. Chen, “Stein’s method: some perspectives with applications”, *Probability towards 2000*, Lect. Notes Stat., **128**, Springer-Verlag, New York, 1998, 97–122.
- [24] L. H. Y. Chen, “Stein’s method of normal approximation: some recollections and reflections”, *Ann. Statist.*, **49**:4 (2021), 1850–1863.
- [25] L. H. Y. Chen, L. Goldstein, Qi-Man Shao, *Normal approximation by Stein’s method*, Probab. Appl. (N. Y.), Springer, Heidelberg, 2011, xii+405 pp.
- [26] L. H. Y. Chen, Qi-Man Shao, “A non-uniform Berry–Esseen bound via Stein’s method”, *Probab. Theory Related Fields*, **120**:2 (2001), 236–254.
- [27] L. H. Y. Chen, Qi-Man Shao, “Normal approximation under local dependence”, *Ann. Probab.*, **32**:3A (2004), 1985–2028.
- [28] Peng Chen, I. Nourdin, Lihu Xu, Xiaochuan Yang, “Multivariate stable approximation by Stein’s method”, *J. Theoret. Probab.*, **37**:1 (2024), 446–488.
- [29] P. Chigansky, F. C. Klebaner, “Compound Poisson approximation for triangular arrays with applications to threshold estimation”, *Electron. Commun. Probab.*, **17** (2012), 29, 10 pp.
- [30] P. Diaconis, “The distribution of leading digits and uniform distribution mod 1”, *Ann. Probab.*, **5**:1 (1977), 72–81.
- [31] P. Diaconis, “Stein’s method for Markov chains: first examples”, *Stein’s method: expository lectures and applications*, IMS Lecture Notes Monogr. Ser., **46**, Inst. Math. Statist., Beachwood, OH, 2004, 27–43.
- [32] P. Diaconis, S. Holmes (eds.), *Stein’s method: expository lectures and applications*, IMS Lecture Notes Monogr. Ser., **46**, Inst. Math. Statist., Beachwood, OH, 2004, vi+139 pp.
- [33] P. Eichelsbacher, M. Löwe, “Stein’s method for dependent random variables occurring in statistical mechanics”, *Electron. J. Probab.*, **15** (2010), 962–988, 30.
- [34] P. Eichelsbacher, B. Martschink, “On rates of convergence in the Curie–Weiss–Potts model with an external field”, *Ann. Inst. Henri Poincaré Probab. Stat.*, **51**:1 (2015), 252–282.
- [35] P. Eichelsbacher, B. Rednoß, “Kolmogorov bounds for decomposable random variables and subgraph counting by the Stein–Tikhomirov method”, *Bernoulli*, **29**:3 (2023), 1821–1848.
- [36] P. Eichelsbacher, G. Reinert, “Stein’s method for discrete Gibbs measures”, *Ann. Appl. Probab.*, **18**:4 (2008), 1588–1618.
- [37] P. Eichelsbacher, C. Thäle, “Malliavin–Stein method for variance-gamma approximation on Wiener space”, *Electron. J. Probab.*, **20** (2015), 123, 28 pp.
- [38] Ш. К. Форманов, “Метод Стейна–Тихомирова и неклассическая центральная предельная теорема”, *Докл. РАН*, **376**:4 (2001), 458–460; англ. пер.: Sh. K. Formanov, “The Stein–Tikhomirov method and a nonclassical central limit theorem”, *Dokl. Math.*, **63**:1 (2001), 75–77.

- [39] Ш. К. Форманов, “О методе Стейна–Тихомирова и его приложениях в неклассических предельных теоремах”, *Дискрет. матем.*, **19**:1 (2007), 27–39; англ. пер.: Sh. K. Formanov, “On the Stein–Tikhomirov method and its applications in nonclassical limit theorems”, *Discrete Math. Appl.*, **17**:1 (2007), 23–36.
- [40] R. E. Gaunt, A. M. Pickett, G. Reinert, “Chi-square approximation by Stein’s method with application to Pearson’s statistic”, *Ann. Appl. Probab.*, **27**:2 (2017), 720–756.
- [41] L. Goldstein, G. Reinert, “Stein’s method and the zero bias transformation with application to simple random sampling”, *Ann. Appl. Probab.*, **7**:4 (1997), 935–952.
- [42] L. Goldstein, Y. Rinott, “Multivariate normal approximations by Stein’s method and size bias couplings”, *J. Appl. Probab.*, **33**:1 (1996), 1–17.
- [43] F. Götze, “On the rate of convergence in the multivariate CLT”, *Ann. Probab.*, **19**:2 (1991), 724–739.
- [44] F. Götze, H. Kösters, A. N. Tikhomirov, “Asymptotic spectra of matrix-valued functions of independent random matrices and free probability”, *Random Matrices Theory Appl.*, **4**:2 (2015), 1550005, 85 pp.
- [45] Ф. Гётце, А. А. Наумов, А. Н. Тихомиров, “Предельные теоремы для двух классов случайных матриц с зависимыми элементами”, *Теория вероятн. и ее примен.*, **59**:1 (2014), 61–80; англ. пер.: F. Götze, A. A. Naumov, A. N. Tikhomirov, “Limit theorems for two classes of random matrices with dependent entries”, *Theory Probab. Appl.*, **59**:1 (2015), 23–39.
- [46] F. Götze, A. N. Tikhomirov, “Asymptotic distribution of quadratic forms”, *Ann. Probab.*, **27**:2 (1999), 1072–1098.
- [47] F. Götze, A. N. Tikhomirov, “Limit theorems for spectra of random matrices with martingale structure”, *Stein’s method and applications*, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., **5**, Singapore Univ. Press, Singapore; World Sci. Publ., River Edge, NJ, 2005, 181–193.
- [48] F. Götze, A. N. Tikhomirov, “Limit theorems for spectra of random matrices with martingale structure”, *Теория вероятн. и ее примен.*, **51**:1 (2006), 171–192; *Theory Probab. Appl.*, **51**:1 (2007), 42–64.
- [49] X. Guyon, S. Richardson, “Vitesse de convergence du théorème de la limite centrale pour des champs faiblement dépendants”, *Z. Wahrsch. Verw. Gebiete*, **66**:2 (1984), 297–314.
- [50] Soo-Thong Ho, L. H. Y. Chen, “An L_p bound for the remainder in a combinatorial central limit theorem”, *Ann. Probab.*, **6**:2 (1978), 231–249.
- [51] И. А. Ибрагимов, Ю. В. Линник, *Независимые и стационарно связанные величины*, Наука, М., 1965, 524 с.; англ. пер.: I. A. Ibragimov, Yu. V. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publ., Groningen, 1971, 443 pp.
- [52] A. M. Khorunzhy, B. A. Khoruzhenko, L. A. Pastur, “Asymptotic properties of large random matrices with independent entries”, *J. Math. Phys.*, **37**:10 (1996), 5033–5060.
- [53] E. Meckes, “On Stein’s method for multivariate normal approximation”, *High dimensional probability V: the Luminy volume*, Inst. Math. Stat. (IMS) Collect., **5**, Inst. Math. Statist., Beachwood, OH, 2009, 153–178.
- [54] G. Mijoule, M. Raič, G. Reinert, Y. Swan, “Stein’s density method for multivariate continuous distributions”, *Electron. J. Probab.*, **28** (2023), 59, 40 pp.
- [55] I. Nourdin, G. Peccati, “Stein’s method and exact Berry–Esseen asymptotics for functionals of Gaussian fields”, *Ann. Probab.*, **37**:6 (2009), 2231–2261.
- [56] I. Nourdin, G. Peccati, “Stein’s method on Wiener chaos”, *Probab. Theory Related Fields*, **145**:1-2 (2009), 75–118.

- [57] I. Nourdin, G. Peccati, *Normal approximations with Malliavin calculus. From Stein's method to universality*, Cambridge Tracts in Math., **192**, Cambridge Univ. Press, Cambridge, 2012, xiv+239 pp.
- [58] E. A. Peköz, A. Röllin, N. Ross, "Generalized gamma approximation with rates for urns, walks and trees", *Ann. Probab.*, **44**:3 (2016), 1776–1816.
- [59] W. Philipp, "The remainder in the central limit theorem for mixing stochastic processes", *Ann. Math. Statist.*, **40**:2 (1969), 601–609.
- [60] M. Raič, "CLT-related large deviation bounds based on Stein's method", *Adv. in Appl. Probab.*, **39**:3 (2007), 731–752.
- [61] Y. Rinott, V. Rotar, "On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U -statistics", *Ann. Appl. Probab.*, **7**:4 (1997), 1080–1105.
- [62] Y. Rinot, V. Rotar, "Normal approximations by Stein's method", *Decis. Econ. Finance*, **23**:1 (2000), 15–29.
- [63] A. Röllin, "Kolmogorov bounds for the normal approximation of the number of triangles in the Erdős–Rényi random graph", *Probab. Engrg. Inform. Sci.*, **36**:3 (2022), 747–773.
- [64] M. Roos, "Stein's method for compound Poisson approximation: the local approach", *Ann. Appl. Probab.*, **4**:4 (1994), 1177–1187.
- [65] N. Ross, "Fundamentals of Stein's method", *Probab. Surv.*, **8** (2011), 210–293.
- [66] Qi-Man Shao, Zhuo-Song Zhang, "Berry–Esseen bounds of normal and nonnormal approximation for unbounded exchangeable pairs", *Ann. Probab.*, **47**:1 (2019), 61–108.
- [67] И. Г. Шевцова, "Об абсолютных константах в неравенстве Берри–Эссеена и его структурных и неравномерных уточнениях", *Информ. и ее примен.*, **7**:1 (2013), 124–125.
- [68] C. Stein, "A bound for the error in the normal approximation to the distribution of a sum of dependent random variables", *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability*, v. II: *Probability theory* (Univ. California, Berkeley, CA, 1970/1971), Univ. California Press, Berkeley, CA, 1972, 583–602.
- [69] C. Stein, *Approximate computation of expectations*, IMS Lecture Notes Monogr. Ser., **7**, Inst. Math. Statist., Hayward, CA, 1986, iv+164 pp.
- [70] Й. Сунклодас, "Аппроксимация распределений сумм слабо зависимых случайных величин нормальным распределением", *Теория вероятностей – 6. Предельные теоремы теории вероятностей*, Итоги науки и техн. Сер. Соврем. пробл. матем. Фундам. направления, **81**, ВИНТИ, М., 1991, 140–199.
- [71] H. Takahata, " L_∞ -bound for asymptotic normality of weakly dependent summands using Stein's result", *Ann. Probab.*, **9**:4 (1981), 676–683.
- [72] H. Takahata, "On the rates in the central limit theorem for weakly dependent random fields", *Z. Wahrsch. Verw. Gebiete*, **64**:4 (1983), 445–456.
- [73] А. Н. Тихомиров, "О скорости сходимости в центральной предельной теореме для слабо зависимых величин", *Вестн. Ленингр. ун-та. Сер. матем., мех., астрон.*, 1976, № 7(2), 158–159.
- [74] А. Н. Тихомиров, "О скорости сходимости в центральной предельной теореме для слабо зависимых величин", *Теория вероятн. и ее примен.*, **25**:4 (1980), 800–818; англ. пер.: A. N. Tikhomirov, "On the convergence rate in the central limit theorem for weakly dependent random variables", *Theory Probab. Appl.*, **25**:4 (1981), 790–809.
- [75] А. Н. Тихомиров, "О нормальной аппроксимации сумм векторных случайных полей с перемешиванием", *Докл. АН СССР*, **272**:2 (1983), 312–314; англ. пер.: A. N. Tikhomirov, "On normal approximation of sums of vector-valued random fields with mixing", *Soviet Math. Dokl.*, **28**:2 (1983), 396–397.

- [76] А. Н. Тихомиров, “О распределении максимальной суммы слабо зависимых величин”, *Теория вероятн. и ее примен.*, **31**:4 (1986), 829–834; англ. пер.: A. N. Tikhomirov, “On the distribution of the maximal sum of weakly dependent variables”, *Theory Probab. Appl.*, **31**:4 (1986), 733–738.
- [77] В. М. Золотарев, *Современная теория суммирования независимых случайных величин*, Наука, М., 1986, 416 с.
- [78] А. Н. Тихомиров, “О точности нормальной аппроксимации вероятности попадания в шар сумм слабо зависимых гильбертовозначных случайных величин. I”, *Теория вероятн. и ее примен.*, **36**:4 (1991), 699–710; англ. пер.: A. N. Tikhomirov, “On the accuracy of normal approximation of the probability of hitting a ball of sums of weakly dependent Hilbert space valued random variables. I”, *Theory Probab. Appl.*, **36**:4 (1991), 738–751.

Александр Николаевич Тихомиров
(Aleksandr N. Tikhomirov)

Физико-математический институт,
Коми научный центр Уральского отделения
Российской академии наук;
Национальный исследовательский университет
“Высшая школа экономики”
E-mail: tikhomirov@ipm.komisc.ru

Поступила в редакцию
03.03.2025