УДК 519.21
ББК 22.171
Л64

Р е ц е н з е н т ы:
докт. физ.-мат. наук, проф. Я.Ю. Никитин (С.-Петерб. гос. ун-т),
докт. физ.-мат. наук, проф. А.Н. Бородин (ПОМИ РАН)

Печатается по постановлению
Редакционно-издательского совета
С.-Петербургского государственного университета

Лифшиц М.А.
Л64 Лекции по сильной аппроксимации. Учебно-метод. пособие. — СПб., 2007. — 32 с.

Пособие посвящено оригинальному типу предельных теорем теории вероятностей – результатам о сильной аппроксимации сумм независимых случайных величин суммами аналогичных нормально распределенных величин. Методы исследования, основанные на сильной аппроксимации, существенно повлияли на центральные разделы вероятностной теории и ее статистические приложения. Наряду с классическими результатами рассматривается аппроксимация не одинаково распределенных величин и многомерных векторов.

Предназначено для студентов старших курсов и аспирантов математических специальностей.

ББК 22.171

1. Введение

Центральная предельная теорема – это утверждение о близости распределения сумм независимых или слабо зависимых случайных величин к нормальному распределению. Функциональная центральная предельная теорема идет дальше и утверждает близость распределения (в подходящем функциональном пространстве) процесса последовательных сумм случайных величин и распределения винеровского процесса. В обоих случаях речь идет не о близости двух стохастических объектов, а лишь о близости их распределений.

Тот же самый эффект нормализации распределений сумм может быть выражен по иному – путем построения двух взаимно близких объектов на одном вероятностном пространстве. Пусть, например, $X = \{X_1, \dots, X_j, \dots\}$ – последовательность независимых случайных величин, а $Y = \{Y_1, \dots, Y_j, \dots\}$ – соответствующая последовательность гауссовских величин (имеется в виду, что Y_j имеет те же математическое ожидание и дисперсию, что и X_j). Требуется построить на некотором вероятностном пространстве последовательности $\tilde{X} = \{\tilde{X}_1, \dots, \tilde{X}_j, \dots\}$ и $\tilde{Y} = \{\tilde{Y}_1, \dots, \tilde{Y}_j, \dots\}$, равнораспределенные с X и Y соответственно таким образом, чтобы величина

$$\Delta_n(\tilde{X}, \tilde{Y}) = \max_{1 \le k \le n} \left| \sum_{j=1}^{k} \tilde{X}_j - \sum_{j=1}^{k} \tilde{Y}_j \right| \tag{1.1}$$

была как можно меньше. При этом малость понимается двояко:

а) в терминах неравенств: доказывается, что вероятность того, что $\Delta_n(\tilde{X}, \tilde{Y})$ велико, мала;

б) в терминах сходимости почти наверное: доказывается, что при $n \to \infty$ с вероятностью единица $\Delta_n(\tilde{X}, \tilde{Y})$ растет не быстрее, чем некоторая известная функция, определяемая моментными характеристиками последовательности X.

Утверждения типа а) являются фундаментальными, а утверждения типа б) легко выводятся из них и часто используются при решении различных асимптотических задач теории вероятностей и статистики. Поскольку в них речь идет о сближении двух последовательностей с вероятностью единица, то по аналогии с усиленным законом больших чисел вся тематика получила название *сильной аппроксимации.*

Доказательства утверждений типа а) и б), как правило, трудны, но зато замечательно большое число полезных следствий и результатов может быть получено из них с большой легкостью. Поэтому инструментарий сильной аппроксимации должен быть под рукой у каждого, кто занимается теоретическими исследованиями, связанными так или иначе с суммами случайных величин.

Первый достаточно точный метод сильной аппроксимации (так называемый метод вложения в винеровский процесс) предложил А.В.Скороход. После этого некоторое время считалось, что его оценки оптимальны. В конечном счете, однако, выяснилось, что для суммирования независимых величин можно получить лучший результат (впрочем, в более широком классе мартингальных последовательностей оценки Скорохода действительно являются оптимальными). Оптимальную оценку скорости сильной аппроксимации сумм независимых одинаково распределенных величин нашли венгерские математики Я. Комлош, П. Майор и Г. Тушнади, предложившие в работах [10, 11] метод диадической аппроксимации (называемый чаще по именам авторов КМТ-конструкцией или венгерской конструкцией). Их результаты представлены ниже. Дальнейший прогресс связан, прежде всего, с работами А.И. Саханенко (суммы независимых разнораспределенных величин), У. Айнмаля (U. Einmahl) и А.Ю. Зайцева (суммы независимых векторов).

Содержание этих лекций отчасти представляет собой упрощенный пересказ того, что в разное время объяснял мне А.Ю. Зайцев. Без его дружеской помощи я, разумеется, не смог бы ничего сделать.

2. Теорема Комлоша, Майора и Тушнади

Следующая теорема представляет собой оптимальную оценку в сильной аппроксимации сумм независимых одинаково распределенных случайных величин.

Теорема 2.1. *(КМТ-неравенство.) Пусть* $X = \{X_j\}$ *– последовательность независимых одинаково распределенных случайных величин, имеющих конечный экспоненциальный момент, т.е.*

при некотором $z > 0$ $\mathbf{E}\exp\{z|X_j|\} < \infty$. *Тогда существуют такие положительные постоянные* C_1, C_2, *зависящие от общего распределения* F *величин* X_j, *что для любого* n *можно построить на некотором вероятностном пространстве последовательность* $\tilde{X} = \{\tilde{X}_1, \ldots, \tilde{X}_n\}$, *равнораспределенную с последовательностью* $\{X_1, \ldots, X_n\}$, *и последовательность независимых гауссовских величин* $\{\tilde{Y}_1, \ldots, \tilde{Y}_n\}$, *имеющих те же математические ожидания и дисперсии, таким образом, чтобы для разности* $\Delta_n(\tilde{X}, \tilde{Y})$, *определенной в* (1.1), *было выполнено*

$$\mathbf{E}\exp\left\{C_1\Delta_n(\tilde{X}, \tilde{Y})\right\} \le 1 + C_2 n^{1/2}. \tag{2.1}$$

Следствие 2.2. В условиях теоремы 2.1 мы имеем для любого $x > 0$

$$\mathbf{P}\left\{\Delta_n(\tilde{X}, \tilde{Y}) \ge x\right\} \le \exp\{-C_1 x\}\,(1 + C_2 n^{1/2}). \tag{2.2}$$

Это неравенство мгновенно получается из (2.1) применением экспоненциального неравенства Чебышева. Именно в такой форме чаще всего и цитируется КМТ-неравенство. Однако формулировка (2.1) (предложенная Саханенко) выглядит более естественной, так как она не содержит лишнего параметра x. Зависимость параметров C_1, C_2 от свойств распределения F мы обсудим ниже, когда речь пойдет о суммах разнораспределенных величин.

Обратите внимание, что в теореме 2.1 речь идет о построении *конечной* последовательности величин. Тем не менее, из теоремы 2.1 сравнительно легко вывести следующий результат о сильной аппроксимации бесконечной последовательности.

Следствие 2.3. В условиях теоремы 2.1 можно построить на некотором вероятностном пространстве последовательность $\tilde{X} = \{\tilde{X}_j,\}$, равнораспределенную с последовательностью X, а также последовательность независимых гауссовских величин $\{\tilde{Y}_j\}$, имеющих те же математические ожидания и дисперсии, таким образом, чтобы с вероятностью единица было выполнено

$$\sum_{j=1}^{n}\tilde{X}_j-\sum_{j=1}^{n}\tilde{Y}_j=O(\log n). \tag{2.3}$$

Доказательство следствия. Положим $n_m=2^{2^m}$ для всех $m=1,2,\dots$. Пусть $N_m=\sum_{k\le m}n_k$, а также $N_0=0$. Для каждого m согласно следствию 2.2 построим на некотором вероятностном пространстве Ω_m последовательности случайных величин $\tilde{X}^{(m)}=\{\tilde{X}_j^{(m)}\}$ и $\{\tilde{Y}_j^{(m)}\}$, $1\le j\le n_m$, которые удовлетворяют оценке (2.2), т.е.

$$\mathbf{P}\left\{\Delta_{n_m}(\tilde{X}^{(m)},\tilde{Y}^{(m)})\ge x\right\}\le\exp\{-C_1x\}(1+C_2n_m^{1/2}).$$

В частности, ряд из вероятностей

$$\begin{aligned}&\sum_m\mathbf{P}\left\{\Delta_{n_m}(\tilde{X}^{(m)},\tilde{Y}^{(m)})\ge A\log n_m\right\}\\ \le\quad&\sum_m\exp\{-C_1A\log n_m\}(1+C_2n_m^{1/2})\end{aligned}$$

будет сходиться, если $C_1A>1/2$.

Перенесем построенное на единое вероятностное пространство, положив $\Omega=\prod_m\Omega_m$ и определив для $\omega=(\omega_1,\omega_2,\dots)$ случайные величины $\tilde{X}_j,\tilde{Y}_j,\ N_{m-1}<j\le N_m$ по формуле

$$\tilde{X}_j(\omega)=\tilde{X}^{(m)}_{j-N_{m-1}}(\omega_m),\qquad\tilde{Y}_j(\omega)=\tilde{Y}^{(m)}_{j-N_{m-1}}(\omega_m).$$

Заметим, что при каждом m верна оценка

$$\Delta_{N_m}(\tilde{X},\tilde{Y})\le\sum_{k=1}^{m}\Delta_{n_k}(\tilde{X}^{(m)},\tilde{Y}^{(m)}).$$

С учетом леммы Бореля–Кантелли и сходимости вышеуказанных рядов имеем почти наверное

$$\Delta_{N_m}(\tilde{X},\tilde{Y})=O\left(\sum_{k=1}^{m}A\log n_k\right)=O(2^m)=O(\log N_m).$$

Поскольку последовательность $\Delta_n(\tilde{X},\tilde{Y})$ является неубывающей и при любом $n\in(N_{m-1},N_m]$ верно $2\log n\ge\log N_m$, то мы можем перейти от подпоследовательности $\{\Delta_{N_m}\}$ ко всей последовательности $\{\Delta_n\}$ и получить оценку (2.3), что и требовалось.

Нетрудно сообразить, что оценка (2.3) является оптимальной. В самом деле, рассмотрим последовательность величин с симметричными экспоненциальными распределениями, т.е. такую, для которой $\mathbf{P}\{|X_j| \geq r\} = \exp\{-r\}$. В этом случае для любого $a > 0$ ряд

$$\sum_j \mathbf{P}\{|X_j| \geq a \log j\} = \sum_j j^{-a}$$

сходится тогда и только тогда, когда $a > 1$, и по лемме Бореля–Кантелли имеем

$$\limsup_{j\to\infty} |X_j| / \log j = 1$$

почти наверное. В то же время хорошо известно, что для гауссовских величин $\tilde{Y}_j = O(\sqrt{\log j}) = o(\log j)$. Поэтому при любом построении $\tilde{X}$ и $\tilde{Y}$

$$\limsup_{j\to\infty} \Delta_n(\tilde{X}, \tilde{Y}) / \log n \geq 1/2$$

с вероятностью единица и видно, что (2.3) невозможно улучшить.

На самом деле неулучшаемость (2.3) верна в следующем еще более сильном и даже удивительном смысле: если в (2.3) можно заменить $O(\log j)$ на $o(\log j)$, то, как показал Бартфаи [5], величины X_j имеют нормальные распределения!

Сила КМТ-оценки (2.3) станет еще более очевидной, если сказать, что предшествовавшая ее появлению техника Скорохода давала аналогичные асимптотики с границей всего лишь $O(n^{1/4})$.

С первого взгляда может показаться, что конечность экспоненциального момента, которая требуется в теореме 2.1, делает этот результат недостаточно широко применимым. На самом же деле несложные манипуляции с урезанием случайных величин позволяют вывести из теоремы 2.1 оценки того же типа для последовательностей величин, подчиняющихся гораздо более слабым моментным ограничениям.

Теорема 2.2. *Пусть $X = \{X_1, \ldots, X_j, \ldots\}$ – последовательность независимых одинаково распределенных случайных величин, имеющих конечный момент порядка $p > 2$, т.е. $\mathbf{E}|X_j|^p < \infty$. Тогда можно таким образом построить на некотором вероятностном пространстве последовательность $\tilde{X} = \{\tilde{X}_1, \ldots, \tilde{X}_j, \ldots\}$, равнораспределенную с X, и последовательность независимых гауссовских величин $\{\tilde{Y}_1, \ldots, \tilde{Y}_j, \ldots\}$, имеющих те же математические* *ожидания и дисперсии, чтобы с вероятностью единица было выполнено*

$$\sum_{j=1}^{n} \tilde{X}_j - \sum_{j=1}^{n} \tilde{Y}_j = o(n^{1/p}). \tag{2.5}$$

Эту теорему мы выведем из более общих результатов в следующем разделе.

Особый интерес представляет теорема Майора [12], которая рассматривает случай $p = 2$, наиболее естественный с точки зрения условий центральной предельной теоремы.

Теорема 2.3. *Пусть $X = \{X_j\}$ последовательность независимых одинаково распределенных случайных величин с нулевыми средними и единичными дисперсиями. Тогда можно таким образом построить на некотором вероятностном пространстве последовательность $\tilde{X} = \{\tilde{X}_j\}$, равнораспределенную с X, и последовательность независимых гауссовских величин $\{\tilde{Y}_j\}$ с нулевыми средними и дисперсиями*

$$\mathbf{E}\tilde{Y}_j^2 = \mathbf{Var}(X_j \mathbf{1}_{|X_j| \le n}) \to 1,$$

чтобы с вероятностью единица было выполнено

$$\sum_{j=1}^{n} \tilde{X}_j - \sum_{j=1}^{n} \tilde{Y}_j = o(n^{1/2}). \tag{2.6}$$

Разделив (2.6) на $n^{1/2}$ и рассмотрев распределения нормированных сумм, в качестве тривиального следствия получим центральную предельную теорему Леви.

Интересно, что без изменения дисперсий можно получить лишь

$$\sum_{j=1}^{n} \tilde{X}_j - \sum_{j=1}^{n} \tilde{Y}_j = o((n \, \log\log n)^{1/2})$$

(Штрассен, [14]), и в общем случае эта оценка неулучшаема.

Для математической статистики большое значение имеет также и другой результат Комлоша, Майора и Тушнади, аналогичный теореме 2.1, но относящийся к аппроксимации эмпирической функции распределения.

3. Сильная аппроксимация не одинаково распределенных величин

В этом разделе мы рассмотрим аналоги КМТ-результатов о сильной аппроксимации сумм независимых величин, имеющих *различные* распределения. Интуитивно ясно, что нужно потребовать определенную равномерную ограниченность распределений слагаемых (или, по крайней мере, их равномерную близость к классу гауссовских распределений). При этом равномерную ограниченность можно понимать различными способами – в терминах экспоненциальных моментов (Саханенко), в терминах обычных моментов (условия Бернштейна), в терминах параметров характеристических функций (Зайцев). Мы будем двигаться от простого к сложному, от частного к общему, и следить за взаимосвязью возникающих условий.

Параметр Крамера. Говорят, что случайная величина X удовлетворяет условию Крамера, если она имеет конечный экспоненциальный момент, т.е. при некотором $h > 0$

$$\mathbf{E}\exp\{h|X|\} < \infty.$$

Соответственно можно ввести *параметр Крамера*

$$h(X) = \sup\{h : \ \mathbf{E}\exp\{h|X|\} < \infty\}$$

как характеристику концентрации распределения X. В дальнейшем мы увидим, что наиболее интересные результаты о сильной аппроксимации достигаются при выполнении условия Крамера, но параметр Крамера не подходит для получения количественных оценок.

Параметр Саханенко. Пусть X случайная величина, имеющая нулевое ожидание и удовлетворяющая условию Крамера. Следуя [4], определим *параметр Саханенко* соотношением

$$\lambda(X) = \sup\left\{\lambda : \ \lambda\mathbf{E}|X|^3\exp\{\lambda|X|\} \le \mathbf{E}X^2\right\}.$$

Выражение $\mathbf{E}|X|^3\exp\{\lambda|X|\}$, стоящее под знаком супремума, конечно при $\lambda < h(X)$, так что $0 < \lambda(X) < \infty$ (кроме вырожденного случая $X = 0$, когда $\lambda(X) = \infty$). Разумеется, $\lambda(X)$ зависит только от распределения величины X. Отметим также, что функционал $\lambda(\cdot)$ является однородным степени -1, т.е. $\lambda(cX) = |c|^{-1}\lambda(X)$.

Если величина X ограничена, то параметр $\lambda(X)$ нетрудно оценить. Пусть $|X| \le a$. Тогда при всех $x \in (0, a)$ запишем неравенство

$$\lambda x^3 e^{\lambda x} = \lambda x e^{\lambda x} x^2 \le \lambda a e^{\lambda a} x^2.$$

Полагая $\lambda = 1/2a$, находим

$$\lambda\, \mathbf{E}|X|^3 e^{\lambda X} \le \sqrt{e/4}\, \mathbf{E}X^2 < \mathbf{E}X^2.$$

Поэтому $\lambda(X) \ge 1/2a$, т.е. $\lambda(X)^{-1} \le 2a$.

Через параметр $\lambda(X)$ можно оценить дисперсию величины X. В самом деле

$$(\mathbf{E}X^2)^{3/2} \le \mathbf{E}|X|^3 \le \mathbf{E}|X|^3 e^{\lambda|X|} \le \lambda^{-1}\mathbf{E}X^2.$$

Отсюда

$$\mathbf{Var}X = \mathbf{E}X^2 \le \lambda(X)^{-2}. \tag{3.1}$$

Обсудим теперь связь параметра Саханенко с другими, более известными характеристиками распределения, такими как параметры Бернштейна и Крамера.

Параметр Бернштейна $b(X)$ определяется для величин с нулевым математическим ожиданием и выражается в терминах моментов величины X:

$$b(X) = \inf\{\tau : |\mathbf{E}X^m| \le \frac{m!}{2}\tau^{m-2}\mathbf{Var}(X), \quad m = 3, 4, \dots\}.$$

Например, если $|X| \le a$, то $b(X) \le a$.

Параметр Бернштейна напрямую контролирует величину четных моментов, однако и нечетные абсолютные моменты также допускают факториальные оценки. Обозначим $\tau = b(X)$. По неравенству Гельдера для любого нечетного $m \ge 3$ верно

$$\begin{aligned} \mathbf{E}|X|^m &= \mathbf{E}(|X|^{(m-1)/2}\, |X|^{(m+1)/2}) \le [\mathbf{E}X^{m-1}]^{1/2}\, [\mathbf{E}X^{m+1}]^{1/2} \\ &\le \left[\frac{(m-1)!}{2}\tau^{m-3}\mathbf{Var}(X) \cdot \frac{(m+1)!}{2}\tau^{m-1}\mathbf{Var}(X)\right]^{1/2} \\ &= \frac{m!}{2}\tau^{m-2}\mathbf{Var}(X)[(m+1)/m]^{1/2} \le \frac{m!}{\sqrt{3}}\tau^{m-2}\mathbf{Var}(X). \end{aligned}$$

Поэтому для любого u

$$\mathbf{E}|X|^3 e^{|uX|} = \sum_{m=0}^{\infty} \mathbf{E}|X|^{m+3}|u|^m/m!$$

$$\le \sum_{m=0}^{\infty} \mathbf{Var}(X)\tau^{m+1}|u|^m \frac{(m+3)!}{\sqrt{3}\, m!}$$

$$
\begin{aligned}
&= \frac{\mathbf{Var}(X)\tau}{\sqrt{3}} \sum_{m=0}^{\infty} (m+3)(m+2)(m+1)(\tau|u|)^m \\
&= \frac{\mathbf{Var}(X)x}{\sqrt{3}|u|} \left[6(1-x)^{-1} + 18x(1-x)^{-2} \right. \\
&\quad \left. +18x^2(1-x)^{-3} + 6x^3(1-x)^{-4}\right],
\end{aligned}
$$

где $x = \tau|u|$. Полагая $u = (7\tau)^{-1}, x = 1/7$, получаем

$$
\mathbf{E}|X|^3 e^{|uX|} \le 0.92|u|^{-1}\mathbf{Var}(X),
$$

т.е. $\lambda(X) \ge (7\tau)^{-1}$ или $\lambda(X)^{-1} \le 7b(X)$.

Неравенство в обратную сторону почти очевидно. Если $\lambda = \lambda(X)$, то для любого $m \ge 3$

$$
\frac{\mathbf{E}|X|^3(\lambda|X|)^{m-3}}{(m-3)!} \le \mathbf{E}|X|^3 e^{|\lambda X|} \le \lambda^{-1}\mathbf{Var}(X),
$$

откуда

$$
\mathbf{E}|X|^m \le (m-3)!\lambda^{2-m}\mathbf{Var}(X) \le \frac{m!}{2}\lambda^{2-m}\mathbf{Var}(X)
$$

и

$$
b(X) \le \lambda(X)^{-1}.
$$

Что касается связи между параметрами Бернштейна (или эквивалентного ему параметра Саханенко) и Крамера, то почти очевидным является неравенство

$$
b(X) \ge h(X)^{-1}.
$$

Действительно, если $h\, b(X) < 1$, то

$$
\begin{aligned}
\mathbf{E}e^{h|X|} &\le \sum_{m\ge 0} \frac{h^m|\mathbf{E}X^m|}{m!} \\
&\le 1 + h\,\mathbf{E}|X| + \frac{h^2\mathbf{E}X^2}{2} + \sum_{m\ge 3} \frac{h^m b(X)^{m-2}\mathbf{Var}X}{2} < \infty.
\end{aligned}
$$

В обратную сторону можно лишь показать, что

$$
b(X) \le 2 \inf_{h>0} \frac{\mathbf{E}e^{h|X|}}{h^3\mathbf{Var}X},
$$

т.е. показатели Бернштейна и Саханенко конечны тогда и только тогда, когда выполнено условие Крамера. Однако следующий пример показывает, что через параметр Крамера оценить параметр

Бернштейна нельзя. Пусть $R \geq 0$ и случайная величина X_R имеет плотность $p(x) = \frac{1}{2}\, e^{R-|x|} \mathbf{1}_{[R,\infty)}(|x|)$. Тогда $h(X_R) = 1$, но

$$b(X_R) \geq \frac{\mathbf{E} X_R^4}{12\, \mathbf{E} X_R^2} \sim \frac{R^2}{12} \to \infty \quad (R \to \infty).$$

Следующая теорема представляет собой оптимальную оценку в сильной аппроксимации сумм независимых случайных величин в терминах параметра Саханенко.

Теорема 3.1. *(Экспоненциальное неравенство Саханенко.) Пусть $X = \{X_j\}$ – последовательность независимых случайных величин, имеющих конечные экспоненциальные моменты. Тогда для любого натурального n можно построить на некотором вероятностном пространстве последовательность $\tilde{X} = \{\tilde{X}_1, \ldots, \tilde{X}_n\}$, равнораспределенную с $\{X_1, \ldots, X_n\}$, а также последовательность независимых гауссовских величин $\{\tilde{Y}_1, \ldots, \tilde{Y}_n\}$, имеющих те же математические ожидания и дисперсии, таким образом, чтобы для разности $\Delta_n(\tilde{X}, \tilde{Y})$, определенной в (1.1), было выполнено*

$$\mathbf{E} \exp\left\{ C_3 \lambda \Delta_n(\tilde{X}, \tilde{Y}) \right\} \leq 1 + \lambda B_n, \tag{3.2}$$

где C_3 некоторая абсолютная постоянная, $B_n^2 = \sum_{j=1}^n \mathbf{Var} X_j$ и $\lambda = \inf_{j \leq n} \lambda(X_j - \mathbf{E} X_j)$.

Замечание 3.2. Неравенство КМТ (2.1) очевидным образом следует из (3.2), причем в нем можно положить $C_1 = C_3 \lambda(X_1)$, $C_2 = \lambda(X_1)(\mathbf{Var} X_1)^{1/2}$.

Явная зависимость всех параметров неравенства (3.2) от распределений слагаемых делает возможным его применение к срезкам случайных величин, не имеющих экспоненциальных моментов. В результате получаются неравенства типа (3.2), хотя и более скромные. Например, для величин, имеющих моменты порядка $p > 2$, верен следующий аналог теоремы 3.1.

Теорема 3.2. *(Степенное неравенство Саханенко.) Пусть $X = \{X_j\}$ – последовательность независимых случайных величин, имеющих конечные моменты порядка $p > 2$. Тогда для любого натурального n можно построить на некотором вероятностном пространстве последовательность $\tilde{X} = \{\tilde{X}_1, \ldots, \tilde{X}_n\}$, равнораспределенную с $\{X_1, \ldots, X_n\}$, и последовательность независимых* *гауссовских величин* $\{\tilde{Y}_1, \ldots, \tilde{Y}_n\}$, *имеющих те же математические ожидания и дисперсии, таким образом, чтобы для разности* $\Delta_n(\tilde{X}, \tilde{Y})$, *определенной в* (1.1), *было выполнено*

$$\mathbf{E}\Delta_n(\tilde{X}, \tilde{Y})^p \leq C(p) \sum_{j=1}^{n} \mathbf{E}|X_j - \mathbf{E}X_j|^p, \tag{3.3}$$

где $C(p)$ *– некоторая постоянная, зависящая только от* p.

Предложение 3.3. *Существует такая константа* $C'(p)$, *что построение из предыдущей теоремы возможно провести так, чтобы неравенство* (3.3) *выполнялось для всех* n *одновременно, хотя и с заменой* $C(p)$ *на* $C'(p)$.

Доказательство. Пусть имеется некоторое разбиение натурального ряда на конечные блоки $\{\mathcal{N}_m\}$, причем

$$\sum_{j \in \mathcal{N}_m} \mathbf{E}|X_j - \mathbf{E}X_j|^p \leq T_m.$$

Тогда по теореме 3.2 можно так построить для каждого блока аппроксимирующие гауссовские последовательности $\tilde{Y}_j,\ j \in \mathcal{N}_m\}$, чтобы соответствующие отклонения Δ_m удовлетворяли оценкам

$$\mathbf{E}\Delta_m^p \leq C(p) T_m.$$

Не ограничивая общности, можно считать, что все эти конструкции реализованы на общем вероятностном пространстве. Тогда верны следующие неравенства:

$$\mathbf{E}\Delta_m \leq C(p)^{1/p} T_m^{1/p}, \qquad \mathbf{E}\Delta_m^2 \leq C(p)^{2/p} T_m^{2/p}.$$

Полагая далее $\tilde{\Delta}_m = \Delta_m - \mathbf{E}\Delta_m$, получаем аналогичные соотношения

$$\mathbf{E}|\tilde{\Delta}_m|^p \leq \mathbf{E}\Delta_m^p + (\mathbf{E}\Delta_m)^p \leq 2\mathbf{E}\Delta_m^p \leq 2\ C(p) T_m,$$

$$\mathbf{E}\tilde{\Delta}_m^2 \leq \mathbf{E}\Delta_m^2 \leq C(p)^{2/p} T_m^{2/p}.$$

Пусть $S = \sum_m \Delta_m,\ \tilde{S} = S - \mathbf{E}S$. Тогда неравенство Розенталя [3, теорема 3.5.19] дает

$$
\begin{aligned}
\mathbf{E}|\tilde{S}|^p &\le c_R(p)\left(\sum_m \mathbf{E}|\tilde{\Delta}_m|^p + \left(\sum_m \mathbf{E}\tilde{\Delta}_m^2\right)^{p/2}\right) \\
&\le c_R(p)C(p)\left(2\sum_m T_m + \left(\sum_m T_m^{2/p}\right)^{p/2}\right)
\end{aligned}
$$

с некоторой постоянной $c_R(p)$, зависящей только от p. Имеем также

$$
\mathbf{E}S = \sum_m \mathbf{E}\Delta_m \le C(p)^{1/p}\sum_m T_m^{1/p}.
$$

Поэтому

$$
\mathbf{E}S^p = \mathbf{E}(\tilde{S} + \mathbf{E}S)^p \le 2^p\ [\mathbf{E}|\tilde{S}|^p + (\mathbf{E}S)^p] \le 2^p C(p) Q, \tag{3.4}
$$

где

$$
Q = \left(2c_R(p)\sum_m T_m + c_R(p)\left(\sum_m T_m^{2/p}\right)^{p/2} + \left(\sum_m T_m^{1/p}\right)^{p}\right).
$$

Теперь докажем собственно предложение 3.3. Для краткости записи считаем $\mathbf{E}X_j = 0$. Рассмотрим два случая.

а) Пусть $T = \sum_j \mathbf{E}|X_j|^p = \infty$. В этом случае мы построим блоки $\{\mathcal{N}_m, -\infty < m < \infty\}$ по формуле

$$
\mathcal{N}_m = \left\{ n : 2^{m-1} < \sum_{j\le n} \mathbf{E}|X_j|^p \le 2^m \right\}.
$$

Соответственно можно положить $T_m = 2^m$. Для всех M и всех $n \in \mathcal{N}_M$ верно

$$
\Delta_n(\tilde{X}, \tilde{Y}) \le \sum_{m\le M} \Delta_m.
$$

Подставляя в (3.4) конкретные значения T_m, получаем

$$
\begin{aligned}
&\mathbf{E}\Delta_n(\tilde{X},\tilde{Y})^p \le \mathbf{E}\left(\sum_{m\le M}\Delta_m\right)^p \\
\le\ & 2^{p+1+M}C(p)\left(2c_R(p) + c_R(p)\left(2^{2/p}-1\right)^{-p/2} + \left(2^{1/p}-1\right)^{-p}\right),
\end{aligned}
$$

в то время как

$$\sum_{j\le n} \mathbf{E}|X_j|^p > 2^{M-1}$$

и (3.3) проверено.

б) Пусть $T = \sum_j \mathbf{E}|X_j|^p < \infty$. Здесь простая конструкция из предыдущего пункта не проходит, так как в один из классов попадет бесконечное число индексов. Построим два типа блоков. По типу предыдущего пункта, положим для $m < 0$

$$\mathcal{N}_m = \left\{ n : 2^{m-1}T < \sum_{j\le n} \mathbf{E}|X_j|^p \le 2^m T \right\}$$

и $T_m = 2^m T$.

Для $m > 0$ выберем

$$\mathcal{N}_m = \left\{ n : 2^{-m}T < \sum_{j\ge n} \mathbf{E}|X_j|^p \le 2^{1-m} T \right\}$$

и $T_m = 2^{-m} T$.

При таком построении останется непокрытым единственный индекс

$$n_0 = \inf\{n : \sum_{j\ge n} \mathbf{E}|X_j|^p > T/2\}.$$

Из него образуем класс $\mathcal{N}_0$ и положим $T_0 = T$.

Для индексов $n \in \cup_{m<0}\mathcal{N}_m$ проходят оценки из п. а). Для индексов $n \in \cup_{m\ge 0}\mathcal{N}_m$ в силу (3.4) верно

$$\mathbf{E}\Delta_n(\tilde{X},\tilde{Y})^p \le \mathbf{E}\left(\sum_{m=-\infty}^{\infty} \Delta_m\right)^p \le C'(p)T,$$

в то время как

$$\sum_{j\le n} \mathbf{E}|X_j|^p > T/2$$

и (3.3) проверено.

Сейчас мы покажем как можно пользоваться оценкой (3.3) для получения достаточно произвольных скоростей сходимости. Сначала сформулируем результат для величин с заданными оценками моментов степенного порядка.

Теорема 3.4. *(К. Шао,* [13].*) Пусть $X = \{X_j\}$ – последовательность независимых случайных величин с нулевыми средними,* *а $H_j \nearrow \infty$ – положительная последовательность и при некотором $p > 2$*

$$\sum_{j=1}^{\infty} \frac{\mathbf{E}|X_j|^p}{H_j^p} < \infty. \tag{3.5}$$

Тогда можно таким образом построить на некотором вероятностном пространстве последовательность $\tilde{X} = \{\tilde{X}_j\}$, равнораспределенную с X, и последовательность независимых гауссовских величин $\tilde{Y} = \{\tilde{Y}_j\}$, имеющих $\mathbf{E}\tilde{Y}_j = 0$ и $\mathbf{Var}\tilde{Y}_j = \mathbf{Var}X_j$, чтобы для разности $\Delta_n(\tilde{X}, \tilde{Y})$, определенной в (1.1), было почти наверное выполнено

$$\Delta_n(\tilde{X}, \tilde{Y}) = o(H_n).$$

Следующая теорема показывает как можно получать оценки типа КМТ на примере сумм случайных величин имеющих равномерно ограниченные хвосты. В случае одинаково распределенных слагаемых из нее мгновенно следует теорема 2.3 и, после небольшого дополнительного вычисления, теорема 2.2.

Теорема 3.5. *(К.Шао, У.Айнмаль.) Пусть заданы положительная случайная величина Z, последовательность независимых случайных величин $X = \{X_j\}$ и функция $G : \mathbf{R}_+ \to \mathbf{R}_+$, удовлетворяющие следующим условиям:*

а) при некотором $\alpha > 1$ функция $x \to G(x)/x^\alpha$ не убывает;

б) при некотором $q > 0$ функция $x \to G(x)/x^q$ не возрастает;

в) для некоторого $c > 0$ при всех $r \geq 0$ верно

$$\sup_j \mathbf{P}\{|X_j| \geq r\} \leq c\, \mathbf{P}\{Z \geq r\};$$

г) $\mathbf{E}G(Z) < \infty$.

Тогда можно таким образом построить на некотором вероятностном пространстве последовательность $\tilde{X} = \{\tilde{X}_j\}$, равнораспределенную с X, и последовательность независимых гауссовских величин $\tilde{Y} = \{\tilde{Y}_j\}$, имеющих математические ожидания $\mathbf{E}X_j$ и дисперсии

$$\mathbf{Var}\,\tilde{Y}_j^2 = \mathbf{Var}(X_j \mathbf{1}_{G(|X_j|) \leq j}),$$

чтобы почти наверное

$$\sum_{j=1}^{n} \tilde{X}_j - \sum_{j=1}^{n} \tilde{Y}_j = o(G^{-1}(n)). \tag{3.6}$$

Доказательство теоремы 3.4. Сначала немного уменьшим H_j: выберем такую возрастающую последовательность H'_j, чтобы $H'_j = o(H_j)$, но по-прежнему

$$\sum_{j=1}^{\infty} \frac{\mathbf{E}|X_j|^p}{(H'_j)^p} < \infty.$$

Положим $X'_j = X_j/H'_j$ и построим на одном вероятностном пространстве последовательность $\tilde{X}'$ и соответствующую гауссовскую последовательность $\tilde{Y}'$ таким образом, чтобы для любого n выполнялось (3.3), т.е.

$$\mathbf{E}\Delta_n(\tilde{X}', \tilde{Y}')^p \le C(p) \sum_{j=1}^{n} \mathbf{E}|X'_j|^p = C(p) \sum_{j=1}^{n} \frac{\mathbf{E}|X_j|^p}{(H'_j)^p}. \tag{3.7}$$

Поскольку правая часть (3.7) равномерно по n ограничена, то мы имеем

$$\mathbf{E}\Delta_\infty(\tilde{X}', \tilde{Y}')^p \le C(p) \sum_{j=1}^{\infty} \frac{\mathbf{E}|X_j|^p}{(H'_j)^p} < \infty,$$

где $\Delta_\infty(\tilde{X}', \tilde{Y}') = \sup_n \Delta_n(\tilde{X}', \tilde{Y}')$. Поэтому последовательность

$$S'_n = \sum_{j=1}^{n} \tilde{X}'_j - \sum_{j=1}^{n} \tilde{Y}'_j\,,$$

для которой $|S'_n| \le \Delta_n(\tilde{X}', \tilde{Y}')$, будет почти наверное ограничена.

Теперь на том же вероятностном пространстве вернем величинам правильный масштаб и положим $\tilde{X}_j = H'_j \tilde{X}'_j, \tilde{Y}_j = H'_j \tilde{Y}'_j$. Для разностей имеем

$$\begin{aligned} \sum_{j=1}^{k} \tilde{X}_j - \sum_{j=1}^{k} \tilde{Y}_j &= \sum_{j=1}^{k} H'_j(\tilde{X}'_j - \tilde{Y}'_j) \\ = \quad \sum_{j=1}^{k} H'_j(S'_j - S'_{j-1}) &= H'_k S'_k - \sum_{j=1}^{k-1} S'_j(H'_{j+1} - H'_j). \end{aligned}$$

Следовательно,

$$\left|\sum_{j=1}^{k}\tilde{X}_j - \sum_{j=1}^{k}\tilde{Y}_j\right| \le 2H'_k \sup_{1\le j<\infty}|S'_j| = O(H'_k) = o(H_k).$$

Поскольку $\tilde{X}_j$ и $\tilde{Y}_j$ имеют нужные распределения и обладают требуемой независимостью, теорема 3.4 доказана.

Доказательство теоремы 3.5. Будем считать, что $\mathbf{E}X_j = 0$. Положим $H_j = G^{-1}(j)$ и разобьем X_j на три части:

$$X_j = \left(X_j\mathbf{1}_{|X_j|\le H_j} - \mathbf{E}[X_j\mathbf{1}_{|X_j|\le H_j}]\right) + X_j\mathbf{1}_{|X_j|>H_j} - \mathbf{E}[X_j\mathbf{1}_{|X_j|>H_j}]. \tag{3.8}$$

Аппроксимации (с помощью теоремы 3.4) заслуживает только первое слагаемое. Проверим выполнение условия (3.5) применительно к величинам $X_j\mathbf{1}_{|X_j|\le H_j} - \mathbf{E}[X_j\mathbf{1}_{|X_j|\le H_j}]$. Убедимся, что для любого $p > q$

$$\sum_{j=1}^{\infty}\frac{\mathbf{E}|X_j|^p\mathbf{1}_{|X_j|\le H_j}}{H_j^p} < \infty. \tag{3.9}$$

По формуле интегрирования по частям и условию в) при всех $u > 0$ и j

$$\begin{aligned} \mathbf{E}|X_j|^p\mathbf{1}_{|X_j|\le u} &= \int_0^u px^{p-1}\mathbf{P}\{|X_j| \ge v\}dv - u^p\mathbf{P}\{|X_j > u\} \\ \le \quad & c\int_0^u px^{p-1}\mathbf{P}\{Z \ge v\}dv = c\ \mathbf{E}Z^p\mathbf{1}_{Z\le u} + c\ u^p\mathbf{P}\{Z > u\}. \end{aligned}$$

Применяя это к $u = H_j$, мы придем к необходимости оценить ряды

$$\sum_{j=1}^{\infty}\frac{\mathbf{E}Z^p\mathbf{1}_{Z\le H_j}}{H_j^p}\ , \qquad \sum_{j=1}^{\infty}\mathbf{P}\{Z > H_j\}.$$

Для второго ряда по условию г)

$$\sum_{j=1}^{\infty}\mathbf{P}\{Z \ge H_j\} = \mathbf{E}\sum_{j=1}^{[G(Z)]} 1 \le \mathbf{E}G(Z) < \infty. \tag{3.10}$$

Для первого ряда

$$\sum_{j=1}^{\infty} \frac{\mathbf{E} Z^p \mathbf{1}_{Z \le H_j}}{H_j^p} \le \mathbf{E} Z^p \sum_{j=N(Z)}^{\infty} H_j^{-p},$$

где

$$N = N(Z) = \begin{cases} [G(Z)] + 1, & \text{если } G(Z) \text{ не целое,} \\ G(Z), & \text{если } G(Z) \text{ целое.} \end{cases}$$

Далее, из условия б) теоремы следует, что функция

$$y \to G^{-1}(y) y^{-1/q}$$

не убывает. Соответственно при любом $j \ge N$ имеем

$$\frac{H_j}{j^{1/q}} = \frac{G^{-1}(j)}{j^{1/q}} \ge \frac{G^{-1}(N)}{N^{1/q}},$$

$$H_j \ge \frac{j^{1/q} G^{-1}(N)}{N^{1/q}} \ ,$$

откуда

$$\begin{aligned} \sum_{j=N}^{\infty} H_j^{-p} &\le G^{-1}(N)^{-p} N^{p/q} \sum_{j=N}^{\infty} j^{-p/q} \\ \le \quad & c(p,q) G^{-1}(N)^{-p} N \le c(p,q) Z^{-p} N. \end{aligned}$$

(Отметим, что в последнем переходе было использовано неравенство $G^{-1}(N(Z)) \ge Z$.) Следовательно,

$$\sum_{j=1}^{\infty} \frac{\mathbf{E} Z^p \mathbf{1}_{Z \le H_j}}{H_j^p} \le c(p,q) \mathbf{E} N(Z) \le c(p,q)(\mathbf{E} G(Z) + 1) < \infty.$$

Таким образом, условие (3.9) проверено. Учитывая, что для любой случайной величины V по неравенству Йенсена

$$\begin{aligned} \mathbf{E}|V - \mathbf{E}V|^p &\le 2^p \mathbf{E} \max\{|V|^p; |\mathbf{E}V|^p\} \\ \le \quad & 2^p (\mathbf{E}|V|^p + |\mathbf{E}V|^p) \le 2^{p+1} \mathbf{E}|V|^p, \end{aligned}$$

мы видим, что первые слагаемые в (3.8) подчиняются условию теоремы 3.4 и тем самым для их сумм имеется гауссовская аппроксимация нужного порядка $o(H_n) = o(G^{-1}(n))$.

Теперь покажем, что второе и третье слагаемое в (3.8) никакой роли не играют. В самом деле, для второго слагаемого по (3.10) мы имеем

$$\sum_j \mathbf{P}\{|X_j| > H_j\} \le c \sum_j \mathbf{P}\{Z \ge H_j\} < \infty.$$

Поэтому по лемме Бореля–Кантелли второе слагаемое, начиная с некоторого номера j, просто обращается в нуль.

Для оценки третьего слагаемого воспользуемся известной леммой Кронекера. Она утверждает, что для произвольной последовательности x_j и произвольной положительной последовательности $c_n \nearrow \infty$ из $\sum_j \frac{x_j}{c_j} < \infty$ следует $\frac{1}{c_n}\sum_{j\le n} x_j \to 0$. Поэтому для проверки соотношения

$$\frac{1}{H_n}\sum_{j\le n} \mathbf{E}[X_j \mathbf{1}_{|X_j|>j}] \to 0,$$

уничтожающего третье слагаемое, достаточно проверить конечность суммы

$$\sum_{j=1}^{\infty} \frac{\mathbf{E}|X_j|\mathbf{1}_{|X_j|>H_j}}{H_j}.$$

Поскольку для любых $u > 0$ и j

$$\begin{aligned}
\mathbf{E}|X_j|\mathbf{1}_{|X_j|>u} &= u\,\mathbf{P}\{|X_j| > u\} + \int_u^\infty \mathbf{P}\{|X_j| > v\}dv \\
&\le c\,u\,\mathbf{P}\{Z > u\} + c\int_u^\infty \mathbf{P}\{Z > v\}dv \\
&= c\,\mathbf{E}Z\mathbf{1}_{Z>u}\ ,
\end{aligned}$$

то дело сводится к математическому ожиданию

$$\mathbf{E}\left(Z \sum_{j<G(Z)} H_j^{-1}\right).$$

Воспользуемся условием в) теоремы. Оно гарантирует, что функция $x \to G^{-1}(x)x^{-1/\alpha}$ не возрастает. Поэтому для любого $j \le m$, где $m = m(Z) = [G(Z)] + 1$, имеем

$$\frac{H_j}{j^{1/a}} = \frac{G^{-1}(j)}{j^{1/a}} \ge \frac{G^{-1}(m)}{m^{1/a}} = \frac{H_m}{m^{1/a}},$$

$$H_j \ge \frac{H_m j^{1/a}}{m^{1/a}}$$

и так как $\alpha > 1$, то

$$\begin{aligned}\sum_{j\le m} H_j^{-1} &\le m^{1/\alpha} H_m^{-1} \sum_{j\le m} j^{-1/\alpha}\\ &\le \quad m^{1/\alpha} H_m^{-1} c(\alpha) m^{1-1/\alpha} = c(\alpha)\ m\ H_m^{-1}.\end{aligned}$$

С учетом неравенства $Z \le H_{m(Z)}$,

$$\begin{aligned}&\mathbf{E}\left(Z \sum_{j<G(Z)} H_j^{-1}\right) \le \mathbf{E}\left(Z \sum_{j\le m(Z)} H_j^{-1}\right)\\ \le \quad & c(\alpha)\mathbf{E}\left(Z\ m(Z) H_{m(Z)}^{-1}\right)\\ \le \quad & c(\alpha)\mathbf{E}m(Z) \le c(\alpha)(\mathbf{E}G(Z)+1) < \infty.\end{aligned}$$

Этого достаточно для применения леммы Кронекера и уничтожения третьей части в (3.8).

Доказательство теоремы 2.2. Не ограничивая общности, можно считать, что наши одинаково распределенные величины X_j имеют единичные дисперсии и нулевые средние. Применяя теорему 3.5 с $G(x) = x^p$, мы построим гауссовские величины $\tilde{Y}_j$, дающие нужный порядок аппроксимации $o(n^{1/p})$, но имеющие не единичные дисперсии. Исправляя этот недостаток, полагаем

$$Y_j = (\mathbf{Var}\tilde{Y}_j)^{-1/2}\tilde{Y}_j.$$

Достаточно проверить, что разности $Z_j = Y_j - \tilde{Y}_j$ почти наверное удовлетворяют оценке

$$n^{-1/p}\sum_{j=1}^{n} Z_j \to 0.$$

По усиленному закону больших чисел Колмогорова, для этого достаточно убедиться в сходимости ряда из дисперсий

$$\sum_{j=1}^{\infty} j^{-2/p}\mathbf{Var}(Z_j). \tag{3.11}$$

Положим $v_j = \mathbf{Var}Y_j$ и запишем тождество

$$Z_j = (v_j^{-1/2} - 1)\tilde{Y}_j = \frac{1 - v_j}{v_j^{1/2}(v_j^{-1/2} + 1)}\tilde{Y}_j.$$

По построению из доказательства теоремы 3.2 мы имеем $H_j = j^{1/p}$, а также

$$\begin{aligned}1-v_j &= \mathbf{E}X_j^2-\mathbf{E}X_j^2\mathbf{1}_{|X_j|\le H_j}+(\mathbf{E}X_j\mathbf{1}_{|X_j|\le H_j})^2\\ &= \mathbf{E}X_j^2\mathbf{1}_{|X_j|>H_j}+(\mathbf{E}X_j\mathbf{1}_{|X_j|>H_j})^2\\ &\le 2\mathbf{E}X_j^2\mathbf{1}_{|X_j|>H_j}=2\sigma_j^2\searrow 0.\end{aligned}$$

Более того,

$$\begin{aligned}\sum_{j=1}^{\infty}\sigma_j^2\ j^{-2/p} &= \mathbf{E}\left(X_1^2\sum_{j=1}^{\infty}\mathbf{1}_{|X_1|>j^{1/p}}j^{-2/p}\right)\\ &\le c(p)\mathbf{E}\left(X_1^2[|X_1|^p]^{1-2/p}\right)\\ &\le c(p)\ \mathbf{E}|X_1|^p<\infty.\end{aligned}$$

Учитывая, что $\{\sigma_j^2\}$ монотонная последовательность, находим отсюда

$$\sigma_n^2\le c(p)\ \mathbf{E}|X_1|^p\left(\sum_{j=1}^{n}j^{-2/p}\right)^{-1}\le c\ n^{\frac{2-p}{p}}.$$

Окончательно, пользуясь неравенством $p>2$, имеем

$$\sum_{j=1}^{\infty}j^{-2/p}\mathbf{Var}(Z_j)\le\sum_{j=1}^{\infty}\frac{(cj^{\frac{2-p}{p}})^2}{(1+o(1))j^{2/p}}\le c\ \sum_{j=1}^{\infty}j^{\frac{2-2p}{p}}<\infty,$$

и (3.11) проверено. Теорема доказана.

Параметр Зайцева. Пусть X случайная величина. Рассмотрим ее комплексные экспоненциальные моменты

$$\Lambda(u)=\mathbf{E}\exp\{uX\},\qquad u\in\mathbf{C},$$

и определим, следуя работе [1], *параметр Зайцева* соотношением

$$\tau_z(X)=\inf\{\tau:|(\log\Lambda)'''(u)|\le\tau\mathbf{Var}(X),\quad\forall u:\ |u|\le\tau^{-1}\}.$$

Здесь имеется в виду, что функция $\Lambda(\cdot)$ определена и дифференцируема во всех точках упомянутого круга $\{u:|u|\le\tau^{-1}\}$. Разумеется, если $\tau_z(X)\ge\tau$, то конечны и вещественные экспоненциальные моменты $\mathbf{E}\exp\{\lambda X\},\ \lambda\in(-\tau^{-1},\tau^{-1})$). Параметр $\tau_z(X)$ характеризует близость распределения X к классу нормальных распределений. Очевидно, что $\tau_z(X)=0$ эквивалентно нормальности распределения X (в самом деле, условие $(\log\Lambda)'''(\cdot)=0$ означает, что $\log\Lambda$ есть многочлен второго порядка). В отличие от параметра Саханенко, малость параметра $\tau_z(X)$ не означает малость случайной величины X (которая может быть сама по себе достаточно большой, но близкой по распределению к некоторой нормальной величине).

Функционал $\tau_z(X)$ является однородным степени 1, т.е. верно $\tau_z(cX) = |c|\tau_z(X)$.

Выполнено также замечательное свойство согласованности с операцией суммирования независимых величин (свертки распределений): для любых независимых величин X, Y верно $\tau_z(X+Y) \leq \max\{\tau_z(X), \tau_z(Y)\}$.

Следующая теорема представляет собой оптимальную оценку в сильной аппроксимации сумм независимых случайных величин в терминах параметра Зайцева.

Теорема 3.6. *(Одномерное неравенство Зайцева.) Пусть* $X = \{X_j\}$ *– последовательность независимых случайных величин, имеющих нулевые средние и единичные дисперсии. Тогда для любого натурального* n *можно таким образом построить на некотором вероятностном пространстве последовательность* $\tilde{X} = \{\tilde{X}_j\}_{j\leq n}$*, равнораспределенную с* $\{X_j\}_{j\leq n}$*, и последовательность независимых гауссовских величин* $\{\tilde{Y}_j\}_{j\leq n}$ *с теми же средними и единичными дисперсиями, чтобы для разности* $\Delta_n(\tilde{X}, \tilde{Y})$*, определенной в* (1.1)*, было выполнено*

$$\mathbf{E}\exp\left\{C_1\tau^{-1}\Delta_n(\tilde{X}, \tilde{Y})\right\} \leq \exp\{C_2\log^+(n/\tau^2)\}, \tag{3.12}$$

где C_1, C_2 *– некоторые абсолютные постоянные, и*

$$\log^+ v = \max\{1, \log v\}, \qquad \tau = \tau_n = \sup_{j\leq n}\tau_z(X_j - \mathbf{E}X_j)\ .$$

Отсюда, конечно, вытекает экспоненциальное неравенство – для любого $x > 0$

$$\mathbf{P}\left\{\Delta_n(\tilde{X}, \tilde{Y}) \geq C_2\tau\log^+(n/\tau^2)/C_1 + x\right\} \leq \exp\{-C_1x/\tau\} \tag{3.13}$$

и логарифмическая аппроксимация КМТ-типа (2.3) для бесконечной последовательности, если

$$\sup_{1\leq j<\infty}\tau_z(X_j - \mathbf{E}X_j) < \infty.$$

Связь параметров Зайцева и Саханенко. Оценим параметр Зайцева через параметр Саханенко. Пусть X – вещественная случайная величина и $\lambda = \lambda(X) < \infty$. Положим $D = \mathbf{Var}(X)$ и

$\Lambda(u) = \mathbf{E}e^{uX}$. Мы уже видели, что $D \le \lambda(X)^{-2}$. Для оценки параметра $\tau_z(X)$ необходимо оценить величину

$$(\log\Lambda)'''(u) = \frac{\Lambda'''}{\Lambda}(u) - \frac{3\Lambda''\Lambda'}{\Lambda^2}(u) + \frac{2(\Lambda')^3}{\Lambda^3}(u) \qquad (3.14)$$

для не слишком больших $u \in \mathbf{C}$. Пусть $|u| \le c\lambda$. Оценим Λ и производные $\Lambda', \Lambda'', \Lambda'''$, появляющиеся в (3.14).

1) Из разложения

$$\Lambda(u) = \mathbf{E}e^{uX} = \mathbf{E}(1 + uX + u^2X^2/2 \pm (|uX|^3/6 + ...))$$

следует

$$\begin{aligned} |\Lambda(u)| &\ge 1 - |u|^2D/2 - |u|^3\mathbf{E}(|X|^3e^{|uX|})/6 \\ &\ge 1 - |u|^2D/2 - |u|^3\lambda^{-1}D/6 \\ &\ge 1 - (c^2/2 + c^3/6) = c_0. \end{aligned}$$

2) Из разложения

$$\Lambda'(u) = \mathbf{E}Xe^{uX} = \mathbf{E}(X + uX^2 \pm |X|(|uX|^2/2 + ...))$$

следует

$$\begin{aligned} |\Lambda'(u)| &\le D|u| + |u|^2\mathbf{E}|X|^3e^{|uX|}/2 \\ &\le D|u| + |u|^2\lambda^{-1}D/2 \le D|u|(1 + c/2). \end{aligned}$$

3) Из разложения

$$\Lambda''(u) = \mathbf{E}X^2e^{uX} = \mathbf{E}(X^2 \pm X^2(|uX| + ...))$$

следует

$$|\Lambda''(u)| \le D + |u|\mathbf{E}|X|^3e^{|uX|} \le D + |u|\lambda^{-1}D \le D(1 + c).$$

4) Из представления $\Lambda'''(u) = \mathbf{E}X^3e^{uX}$ следует

$$|\Lambda'''(u)| \le \mathbf{E}|X|^3e^{|uX|} \le \lambda^{-1}D.$$

Собирая все оценки вместе, находим

$$\begin{aligned} &|(\log\Lambda)'''(u)| \\ \le\ & \frac{\lambda^{-1}D}{c_0} + \frac{3D(1+c)D|u|(1+c/2)}{c_0^2} + \frac{2(D|u|(1+c/2))^3}{c_0^3} \\ \le\ & \lambda^{-1}D\left[\frac{1}{c_0} + \frac{3c(1+c)(1+c/2)}{c_0^2} + \frac{2c^3(1+c/2)^3}{c_0^3}\right]. \end{aligned}$$

Полагая здесь $c = 1/3$, получаем

$$|(\log \Lambda)'''(u)| \leq 3\lambda^{-1} D = c^{-1}\lambda^{-1} D,$$

откуда следует $\tau_z(X) \leq c^{-1}\lambda^{-1} = 3\lambda(X)^{-1}$. Таким образом, для любой случайной величины X верно

$$\tau_z(X) \leq 3\lambda^{-1}(X).$$

Оценка параметра Саханенко (или Бернштейна) через параметр Зайцева требует дополнительного ограничения на дисперсию. Предположим, что $\tau_z(X) \leq 1$, $\mathbf{E}X = 0$ и $D = \mathbf{Var}(X) \leq 1$. Тогда для $u : |u| \leq 1$ мы имеем $|(\log \Lambda)'''(u)| \leq D$, а также начальные условия $(\log \Lambda)''(0) = D$, $(\log \Lambda)'(0) = (\log \Lambda)(0) = 0$. Интегрирование дает

$$|(\log \Lambda)'(u)| \leq D(|u| + |u|^2/2)$$

и

$$|(\log \Lambda)(u)| \leq D(|u|^2/2 + |u|^3/6).$$

Отсюда моментное неравенство для вещественных u :

$$\begin{aligned}
|\mathbf{E}X^2(e^{uX} - 1)| &= |\Lambda''(u) - \Lambda''(0)| \leq \int_0^u |\Lambda'(v)| dv \\
&= \int_0^u |(\log \Lambda)'(v)\Lambda(v)| dv \\
&\leq \int_0^u D(|v| + |v|^2/2) \exp\{D(|u|^2/2 + |u|^3/6)\} dv \\
&\leq D(|u|^2/2 + |u|^3/6) \exp\{2/3\} \\
&\leq \frac{2}{3} \exp\{2/3\} D = c\ D.
\end{aligned}$$

Полагая $u = \pm 1$, после осреднения получаем

$$c\ D \geq \left| \mathbf{E}X^2 \left(\frac{e^X + e^{-X}}{2} - 1 \right) \right| = \mathbf{E} \sum_{m=1}^{\infty} \frac{X^{2m+2}}{(2m)!}.$$

Для любого четного момента

$$\mathbf{E}X^{2m+2} \leq c(2m)! D.$$

Для нечетных абсолютных моментов по неравенству Гельдера

$$\begin{aligned}
\mathbf{E}X^{2m+1} &\leq [\mathbf{E}X^{2m}]^{1/2}\ [\mathbf{E}X^{2m+2}]^{1/2} \\
&\leq c\ D(2m-1)! \sqrt{2m/(2m-1)} \leq \sqrt{2} c D(2m-1)!.
\end{aligned}$$

Поскольку при всех $m \geq 3$ верно $\sqrt{2}c \leq (2^{3/2}c)^{m-2}/2$, то $b(X) \leq 2^{3/2}c < 4$. В силу однородности показателей Зайцева и Бернштейна имеем для всех случайных величин с нулевым средним

$$b(X) \leq 4\max\{\tau_z(X), \sqrt{\mathbf{Var}(X)}\}.$$

Требуемая оценка получена.

4. Сильная аппроксимация сумм многомерных величин

В этом разделе мы рассматриваем ту же задачу о сильной аппроксимации сумм, предполагая, что величины X_j принимают значения в $\mathbf{R}^d$. Соответственно нам потребуются такие понятия как евклидова норма $\|\cdot\|$, скалярное произведение $(\cdot,\cdot)$ в $\mathbf{R}^d$ и $\mathbf{C}^d$; математическое ожидание $\mathbf{E}X \in \mathbf{R}^d$ и ковариационный оператор $D = \mathbf{Cov}X : \mathbf{R}^d \to \mathbf{R}^d$ для $\mathbf{R}^d$-значной случайной величины X, определяемые формулами

$$(\mathbf{E}X, v) = \mathbf{E}(X, v),$$

$$(Dv, w) = \mathbf{Cov}((X,v),(X,w)) = \mathbf{E}(X - \mathbf{E}X, v)(X - \mathbf{E}X, w).$$

Символом $\partial_v f$ обозначаем частную производную функции f в направлении $v \in \mathbf{R}^d$. Аналогично $\partial_v^2 f$ обозначает частную производную второго порядка.

Результаты, которые удается получить для многомерного случая, во многом аналогичны одномерным, но приходится учитывать новый мешающий фактор – возможное вырождение ковариационных операторов. Это не так важно для случая одинаково распределенных слагаемых, где линейной заменой можно привести ситуацию к сложению векторов с единичными ковариациями, но может сильно испортить дело, если мы складываем величины, чьи ковариации вырождаются в разных направлениях.

Определим многомерные аналоги условия Крамера, а также параметров Бернштейна, Саханенко и Зайцева. Случайная величина $X \in \mathbf{R}^d$ удовлетворяет условию Крамера, если она имеет конечный экспоненциальный момент, т.е. при некотором $h > 0$

$$\mathbf{E}\exp\{h||X||\} < \infty.$$

Как и в одномерном случае, наиболее интересные результаты получаются при выполнении этого условия.

Параметр Саханенко. Пусть $X \in \mathbf{R}^d$ случайная величина, имеющая нулевое ожидание и удовлетворяющая условию Крамера.

Определим *параметр Саханенко* соотношением

$$\lambda(X) = \sup\big\{\lambda : \ \lambda\, \mathbf{E}\left((X,v)^2|(X,w)|\exp\{\lambda|(X,w)|\}\right) \le \mathbf{E}(X,w)^2,$$
$$\forall v,w \in \mathbf{R}^d : \ \|v\| = \|w\| = 1\big\}.$$

Параметр Бернштейна определяется для величин с нулевым математическим ожиданием и выражается в терминах моментов величины X:

$$b(X) = \inf\{\tau : |\mathbf{E}(X,v)^2(X,w)^{m-2}| \le \frac{m!}{2}\tau^{m-2}\mathbf{E}(X,v)^2,$$
$$\forall v,w \in \mathbf{R}^d,\ ||w|| = 1, \quad \forall m = 3,4,\dots\}.$$

Параметр Зайцева. Пусть $X \in \mathbf{R}^d$ случайная величина. Рассмотрим ее комплексные экспоненциальные моменты

$$\Lambda(u) = \mathbf{E}\exp\{(u,X)\}, \qquad u \in \mathbf{C}^d,$$

и определим *параметр Зайцева* соотношением

$$\tau_z(X) = \inf\big\{\tau : \partial_w\partial_v^2(\log\Lambda)(u)| \le \tau(\mathbf{Cov}X\, v, v),$$
$$\forall u \in \mathbf{C}^d, v,w \in \mathbf{R}^d : \ |u| \le \tau^{-1}, \|w\| = \|v\| = 1\big\}.$$

Соотношения между тремя показателями такие же, как в одномерном случае. Параметры Бернштейна и Саханенко эквивалентны, т.е.

$$[7\ \lambda(X)]^{-1} \le b(X) \le \lambda(X)^{-1}.$$

Параметр Зайцева можно оценить через параметр Бернштейна или Саханенко – с некоторой абсолютной постоянной c верно

$$\tau_z(X) \le c\ b(X).$$

Для обратной оценки приходится привлекать величину B^2 максимального собственного числа оператора $\mathbf{Cov}(X)$ (в одномерном случае здесь использовалась дисперсия). Имеем для величин с нулевым средним и некоторой абсолютной постоянной c:

$$b(X) \le c\ \max\{\tau_z(X), B\}.$$

Теорема 4.1. *(Многомерное неравенство Зайцева.) Пусть заданы $\alpha > 0$, $D > 0$, а $X = \{X_j\}$ – последовательность независимых случайных величин, имеющих нулевые средние и ковариационные* *операторы D_j , удовлетворяющие условию равномерной невырожденности*

$$\beta_1||v||^2 \le D^2(D_j v, v) \le \beta_2||v||^2, \ \forall v \in \mathbf{R}^d. \tag{4.1}$$

Тогда для любого натурального n можно таким образом построить на некотором вероятностном пространстве последовательность $\tilde{X} = \{\tilde{X}_1, \ldots, \tilde{X}_n\}$, равнораспределенную с $\{X_1, \ldots, X_n\}$, и последовательность независимых гауссовских случайных величин $\{\tilde{Y}_1, \ldots, \tilde{Y}_n\}$ с нулевыми средними и ковариационными операторами D_j , чтобы для разности

$$\Delta_n(\tilde{X}, \tilde{Y}) = \max_{1 \le k \le n} \left\| \sum_{j=1}^{k} \tilde{X}_j - \sum_{j=1}^{k} \tilde{Y}_j \right\|$$

было выполнено

$$\mathbf{E} \exp \left\{ \frac{C_1 D \Delta_n(\tilde{X}, \tilde{Y})}{\tau d^{9/2} \log^+ d} \right\} \le \exp\{C_2 d^{3+\alpha} \log^+(n/\tau^2)\}, \tag{4.2}$$

где C_1, C_2 – некоторые постоянные, зависящие от α, β_1, β_2 , и

$$\tau = \max\{1, \tau_z(X_1), \ldots, \tau_z(X_n)\}.$$

Следствие 4.2. В условиях теоремы верно экспоненциальное неравенство – для любого $x > 0$

$$\mathbf{P} \left\{ C_1 \Delta_n(\tilde{X}, \tilde{Y}) \ge C_2 \tau d^{15/2+\alpha} \log^+ d \log^+(n/\tau^2) + x \right\}$$
$$\le \exp \left\{ - \frac{x}{\tau d^{9/2} \log^+ d} \right\}.$$

Следствие 4.3. Если ковариационные операторы равномерно ограничены в смысле (4.1) и

$$\sup_{1 \le j < \infty} \tau_z(X_j) < \infty,$$

то для бесконечной последовательности верна аппроксимация КМТ-типа

$$\sum_{j=1}^{n} \tilde{X}_j - \sum_{j=1}^{n} \tilde{Y}_j = O(\log n).$$

В частности, это верно для сумм независимых одинаково распределенных $\mathbf{R}^d$-значных случайных величин, удовлетворяющих условию Крамера.

Замечание 4.4. Ввиду отмеченных выше соотношений между параметрами Бернштейна, Зайцева и Саханенко, в формулировке теоремы можно заменить $\tau_z(X_j)$ на $b(X_j)$ или $\lambda(X_j)^{-1}$.

Замечание 4.5. Определенное неудобство в применении теоремы может выбрать равномерность оценки (4.1). В этом отношении представляет интерес следующее обобщение теоремы 4.1 (собственно говоря, оно и представляет собой настоящее неравенство Зайцева). Отрезок натурального ряда $[1..n]$ разбивается на последовательные блоки $\mathcal{N}_1, \ldots, \mathcal{N}_l$ и вместо (4.1) ограничение накладывается на ковариации блочных сумм, т.е.

$$\beta_1 ||v||^2 \leq D^2 \sum_{j \in \mathcal{N}_k} (D_j v, v) \leq \beta_2 ||v||^2, \ \forall k \leq l, \quad \forall v \in \mathbf{R}^d.$$

Тогда верно (4.2) с заменой n на l в правой части.

Замечание 4.6. В некоторых случаях можно несколько улучшить зависимость постоянных от размерности. Например, если все величины X_j имеют единичные ковариационные операторы, то, согласно работе [15], вместо (4.2) можно записать

$$\mathbf{E} \exp \left\{ \frac{C_1 D \Delta_n(\tilde{X}, \tilde{Y})}{\tau d^3 \log^+ d} \right\} \leq \exp\{C_2 d^{9/4+\alpha} \log^+ n\},$$

где C_1, C_2 – некоторые постоянные, зависящие от α. С другой стороны, если моменты третьего порядка всех слагаемых равны нулю, то можно выбросить $\log^+ d$ в знаменателе левой части (4.2).

Замечание 4.7. Ограничение, присутствующее в теореме 4.1 не позволяет использовать возможную малость параметров $\tau_z(X_j)$, т.е. близость распределений к гауссовским. Результаты без этого ограничения (но только для достаточно гладких распределений с единичными ковариационными операторами) см. в работе [9].

Литература

1. *Зайцев А.Ю.* Оценки расстояния Леви-Прохорова в многомерной центральной предельной теореме для случайных векторов с конечными экспоненциальными моментами // Теория вероятностей и ее применения. 1986. Т. 31. С. 246–265.

2. *Зайцев А.Ю.* Multidimensional version of a result of Sakhanenko in the invariance principle for vectors with finite exponential moments // Теория вероятностей и ее применения. 2000. Т. 45. С. 718–738.

3. *Петров В.В.* Предельные теоремы для сумм независимых случайных величин. М., 1987.

4. *Саханенко А.И.* Скорость сходимости в принципе инвариантности для разнораспределенных величин с экспоненциальными моментами // Труды Института Математики СО АН СССР. 1984. Т. 3. С. 4–49.

5. *Bártfai P.* Die Bestimmung der zu einem wiederkehrenden Prozess gehörenden Verteilungfunktion aus den mit Fehlern behafteten Daten einer einzigen Realisation// Studia Sci. Math. Hungar. 1966. Vol. 1. P. 161–168.

6. *Csörgő M., Révész P.* Strong Approximations in Probability and Statistics. New York, 1981.

7. *Einmahl U.* Strong invariance principles for partial sums of independent random vectors // Ann. Probab. 1987. Vol. 15. P. 1419–1440.

8. *Einmahl U.* Extensions of results of Komlós, Major, and Tusnády to the multivariate case // J. Multivar. Anal. 1986. Vol. 28. P. 20–68.

9. *Götze F., Zaitsev A.Yu.* Hungarian construction for vectors with almost Gaussian smooth distributions // Asymptotic Methods in Probability and Statistics with Applications. Boston, 2001. P. 101–132.

10. *Komlós J., Major P., Tusnády G.* An approximation of partial sums of independent RV'-s and the sample DF.I // Z. Wahrscheinlichkeitstheor. verw. Geb. 1975. Vol. 32. P. 111–131.

11. *Komlós J., Major P., Tusnády G.* An approximation of partial sums of independent RV'-s and the sample DF.II // Z. Wahrscheinlichkeitstheor. verw. Geb. 1976. Vol. 34, P. 34–58.

12. *Major P.* An improvement of Strassen's invariance principle. // Ann. Probab. 1979. Vol. 7. P. 55–61.

13. *Shao Q.* Strong approximation theorems for independent random variables and their applications // J. Multivar. Anal. 1995. Vol. 52. P. 107–130.

14. *Strassen V.* An invariance principle for the law of iterated logarithm // Z. Wahrscheinlichkeitstheor. verw. Geb. 1964. Vol. 3. P. 211–226.

15. *Zaitsev A.Yu.* Multidimensional version of the results of Komlós, Major, and Tusnády for vectors with finite exponential moments // ESAIM: Probability and Statistics. 1998. Vol. 2. P. 41–108.

Оглавление

У ч е б н о е   и з д а н и е

Михаил Анатольевич Лифшиц

ЛЕКЦИИ ПО СИЛЬНОЙ АППРОКСИМАЦИИ

Учебно-методическое пособие

Зав. редакцией Г.И. Чередниченко
Редактор Ф.С. Бастиан
Техн. редактор Л.Н. Иванова
Обложка А.В. Калининой
Компьютерная верстка автора

Подписано в печать с оригинала-макета 15.01.2007.
Ф-т 60х84/16. Усл. печ. л. 1,86. Уч.-изд. л. 1,6. Тираж 100 экз.

Заказ №

РОПИ С.-Петербургского государственного университета.

199034, С.-Петербург, Университетская наб., 7/9.

Отпечатано в отделе оперативной полиграфии НИИХ СПбГУ
с оригинала-макета заказчика.

198504, С.-Петербург, Старый Петергоф. Университетский пр. 26.

Предназначено для учебного процесса.