Записки научных
семинаров ПОМИ
Том 320, 2004 г.

А. И. Мартикайнен

ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА ПОЧТИ НАВЕРНОЕ БЕЗ ЛОГАРИФМИЧЕСКОГО СУММИРОВАНИЯ

1. Введение. Пусть $\{X_n\}$ – последовательность независимых случайных величин с нулевыми математическими ожиданиями и конечными дисперсиями. Обозначим $S_n = X_1 + \dots + X_n$ и предположим, что $B_n^2 = EX_1^2 + \dots + EX_n^2 \to \infty$. Предположим, что выполнено условие Линдеберга и, следовательно, справедлива центральная предельная теорема (ЦПТ). Если дополнительно

$$B_l/B_k \geqslant Kl^a/k^a \qquad 1 \leqslant k \leqslant l \tag{1}$$

для некоторых постоянных $K > 0$, $a > 0$, то справедлива ЦПТ почти наверное:

$$\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}I(S_k < xB_k) \to \Phi(x) \ \text{ п.н.}, \tag{2}$$

равномерно по $x \in R$. Здесь и далее $I(\cdot)$ – индикаторная функция события в скобках, $\Phi(x)$ – стандартная нормальная функция распределения. Этот результат был получен Беркесом и Делингом в 1993 году (при дополнительном условии одинаковой распределенности X_n и дополнительных моментных условиях его доказали в 1988 независимо Бросамлер и Шатте, оказавшиеся пионерами этого направления). Условие (1) влечет как мимимум степенную скорость роста B_n. Если его опустить, ЦПТ почти наверное может не выполняться (см. снова [8]). Однако, его можно опустить, если разрешить системе весов зависеть от B_n. Атлах в 1993 году показал, что при выполнении условия Линдеберга имеет место следующая форма ЦПТ почти наверное:

$$\frac{1}{\log B_n^2}\sum_{k=1}^{n}\frac{EX_{k+1}^2}{B_k^2}I(S_k < xB_k) \to \Phi(x) \ \text{ п.н.}, \tag{3}$$

Работа выполнена при поддержке грантов E02-1.0-56 Минобразования РФ, 02-01-00779 РФФИ и НШ-2258.2003.1.

равномерно по $x \in R$. К настоящему моменту имеются десятки публикаций, исследующих ЦПТ почти наверное в различных направлениях. Практически каждая предельная теорема, формулируемая на языке слабой сходимости, имеет свой аналог "почти наверное" (см., например, Беркес и Чаки (2001)).

Исследовалась и скорость сходимости в (2). Понятно, что она не может быть лучше логарифмической – первый индикатор в (2) имеет логарифмически убывающий вес. Отметим, что попытки заменить систему весов предпринимались многократно. Однако, в наших предположениях выбор логарифмического метода суммирования или близкого к нему метода (3) оказывается обоснованным: лучшая скорость сходимости, по-видимому, не достигается ни на каком другом методе суммирования. Отметим также, что при любом x обычные средние арифметические индикаторов $I(S_k < xB_k)$ при сходимости с вероятностью 1 оказываются неустойчивыми, имея верхний предел равный 1 и нижний предел равный 0.

В 2002 году при условии одинаковой распределенности независимых случайных величин X_n в [9] была доказана невозможность значительного улучшения скорости сходимости в ЦПТ почти наверное: если $EX_1 = 0$ и $EX_1^2 = 1$, то из соотношения

$$b_n(H_n(I(S_1 < x), \dots, I(S_n < x\sqrt{n})) - \Phi(x)) \to 0$$
$$\text{п.н. равномерно по } x$$

для некоторой последовательности постоянных b_n и некоторой последовательности функций H_n (H_n – произвольная функция n вещественных аргументов) вытекает оценка $b_n = o(\log n \log\log n)$. В [9] была предложена иная форма ЦПТ почти наверное, обеспечивающая степенную скорость сходимости. Разобьем множество индексов $\{1, \dots, n\}$ на $l_n = [\sqrt{n \log n}]$ последовательных непересекающихся интервалов $\Delta_{n,1}, \dots, \Delta_{n,l_n}$ одной и той же длины $k_n = [n/l_n]$ (возможно, n не делится на l_n нацело, в этом случае отбросим несколько последних индексов). С помошью сумм $S_{n,j} = \sum_{i \in \Delta_{n,j}} X_i$ построим статистику

$$T_n(x) = \frac{1}{l_n}(I(S_{n,1} < x\sqrt{k_n}) + \dots + I(S_{n,l_n} < x\sqrt{k_n})).$$

В [9] было показано, что для любой числовой последовательности x_n,

$$T_n(x_n) - \Phi(x_n) = O(n^{-1/4}(\log n)^{1/4}) \text{ п.н.}, \qquad (4)$$

если момент третьего порядка X_1 конечен.

В п. 2 мы покажем (в теореме 1), что ЦПТ почти наверное, основанная на индикаторах $P(S_k < xB_k)$, не может иметь степенной скорости сходимости при суммировании независимых случайных величин с нулевыми средними и конечными дисперсиями (при некоторых дополнительных условиях). В теореме 2 мы предположим, что S_k образуют мартингал. При дополнительных предположениях (более жестких, чем в теореме 1) будет установлено, что скорость сходимости в ЦПТ почти наверное не может быть лучше логарифмической, как и в случае сумм независимых одинаково распределенных слагаемых. В п. 3 мы покажем, что предложенный в [9] вариант ЦПТ почти наверное может быть доказан равномерно по x (то есть (4) выполняется равномерно по $x \in R$), что использованная в [9] техника переносится на случай независимых неодинаково распределенных случайных величин, и даже на схему серий с сохранением степенной скорости сходимости (теорема 3).

2. О невозможности степенной скорости сходимости в традиционной ЦПТ почти наверное.

Теорема 1. *Пусть $\{X_n\}$ – последовательность независимых случайных величин, $EX_n = 0$, $EX_n^2 < \infty$, $EX_1^2 > 0$, $B_n^2 = \sum_{k=1}^n EX_k^2 \to \infty$ и при каждом $\varepsilon > 0$ выполнено условие*

$$\sum_{k=1}^n P(X_k > \varepsilon B_k \sqrt{\log B_n}) \to 0. \tag{5}$$

Предположим, что для некоторой последовательности функций $H_n : \{0,1\}^n \to R$ и для некоторой последовательности постоянных $\{b_n\}$ выполнено с вероятностью 1 равномерно по $x \in R$ соотношение

$$b_n(H_n(I_1(x), \dots, I_n(x)) - \Phi(x)) \to 0, \tag{6}$$

где

$$I_k(x) = I(X_1 + \dots + X_k < xB_k). \tag{7}$$

Тогда $b_n = o(B_n^\delta)$ для любого $\delta > 0$.

Условие (5), конечно, значительно слабее условия Линдеберга. Непосредственно из неравенства Чебышева вытекает оценка

$$\sum_{k=1}^n P(X_k > \varepsilon B_k \sqrt{\log B_n}) \leqslant \frac{1}{\varepsilon^2 \log B_n} \sum_{k=1}^n \frac{EX_k^2}{B_k^2} \leqslant \frac{2}{\varepsilon^2} + o(1).$$

Так что требуются лишь небольшие предположения, чтобы вместо ограниченности правой части иметь ее стремление к 0. При более жестких дополнительных условиях можно доказать, что b_n растет не быстрее логарифмической функции. Следующий результат можно доказать, перенося технику из [9] на мартингальный случай.

Теорема 2. *Пусть* $\{X_n, \mathcal{F}_n\}_{n\geqslant 0}$ – *последовательность мартингал-разностей с конечными вторыми моментами,* $X_0 = 0$, $\mathcal{F}_0 = \{\phi, \Omega\}$. *Предположим, что*

$$B_n^2 = \sum_{k=1}^{n} E(X_k|\mathcal{F}_{k-1}) \to \infty \quad \text{п.н.},$$

и для некоторой положительной предсказуемой последовательности $\{\varepsilon_n, \mathcal{F}_n\}$ *выполнены условия* $\varepsilon_n \searrow 0$ *п.н.,*

$$|X_n| \leqslant \varepsilon_n B_n/(\log\log B_n)^{3/2} \ \text{п.н.}, \quad n \geqslant 1 \tag{8}$$

Если (6) *имеет место с вероятностью* 1 *равномерно по* $x \in R$ *с* $I_k(x)$, *определенными в* (7), *то* $b_n = o(\log B_n \log\log B_n)$.

Вопрос о скорости сходимости в ЦПТ почти наверное при нарушении условия (5) остается открытым.

Доказательство этих двух теорем предварим несколькими леммами.

Лемма 1. *Пусть* $H_n : \{0,1\}^n \to R$ – *последовательность функций,* $\{Y_{n,k}, 1 \leqslant k \leqslant n, n \geqslant 1\}$, – *случайные величины. Предположим, что числовая последовательность* $x_n \to \infty$ *выбрана так, что*

$$\liminf P(\max_k Y_{n,k} < x_n) > 0.$$

Если

$$b_n(H_n(I(Y_{n,1} < x_n), \ldots, I(Y_{n,n} < x_n)) - \Phi(x_n)) \xrightarrow{P} 0, \tag{9}$$

и, либо $H_n(1,\ldots,1) = 1$ *при всех* n, *либо* (9) *выполнено дополнительно с заменой* x_n *на некоторую последовательность постоянных* $x_n' \geqslant x_n + x_n^{-1}$, *то* $b_n = o(x_n \exp(x_n^2/2))$.

Доказательство (см. также [9]). Положим $J_n(x_n) = I(\max_k Y_{n,k} < x_n)$. При каждом $\varepsilon > 0$ величины

$$\begin{aligned} P(J_n(x_n) = 1, |b_n(H_n(I(Y_{n,1} < x_n), \ldots, I(Y_{n,n} < x_n)) - \Phi(x_n))| > \varepsilon) \\ = I(|b_n(H_n(1,\ldots,1) - \Phi(x_n))| > \varepsilon) P(J_n(x_n) = 1) \end{aligned}$$

стремятся к нулю. Поскольку вероятности в правой части равенства отделены от нуля при больших n, индикаторы в правой части обращаются в 0 при больших n. Таким образом,

$$b_n(H_n(1,\ldots,1)-\Phi(x_n))\to 0. \qquad (10)$$

Ясно, что при выполнении (9) с x'_n вместо x_n, имеет место и (10) с x'_n вместо x_n. Учитывая, что $x_n\to\infty$ и

$$1-\Phi(x_n)\sim\frac{1}{\sqrt{2\pi}x_n}e^{-x_n^2/2}\geqslant(e^{-1}+o(1))(1-\Phi(x'_n)),$$

получаем утверждение леммы.

В леммах 2 и 4 предполагается, что независимые случайные величины $X_1,X_2,\ldots,$ имеют нулевые средние и конечные вторые моменты, S_n и B_n определены в начале статьи.

Лемма 2. *Предположим, что* $0<b_1\leqslant\cdots\leqslant b_n$ *-постоянные. Тогда при каждом* $x>0$ *справедливо неравенство*

$$P(\max_{1\leqslant k\leqslant n}|S_k|/b_k>x)\leqslant\frac{1}{x^4}\Big(\sum_{k=1}^{n}\frac{EX_k^4}{b_k^4}+2\sum_{k=2}^{n}\frac{EX_k^2B_{k-1}^2}{b_k^4}\Big).$$

Эта лемма является переформулировкой одного неравенства В. М. Золотарева см. [1, с. 76]). Согласно этому неравенству

$$P(\max_{1\leqslant k\leqslant n}|S_k|/b_k>x)\leqslant\frac{1}{x^4}\Big(\frac{ES_n^4}{b_n^4}+\sum_{k=1}^{n-1}\Big(\frac{1}{b_k^4}-\frac{1}{b_{k+1}^4}\Big)ES_k^4\Big).$$

Правая часть может быть переписана в таком виде

$$\frac{1}{x^4}\sum_{k=1}^{n}\frac{1}{b_k^4}(ES_k^4-ES_{k-1}^4),$$

где $S_0=0$. Замечая, что $ES_k^4=\sum_{j=1}^{k}EX_j^4+2\sum_{j=2}^{k}EX_j^2B_{j-1}^2$ при $k>1$, получаем утверждение леммы.

Лемма 3. *Пусть* $A_{n,k}$ *и* $B_{n,k}$ *-два набора событий* $1\leqslant n<\infty$, $1\leqslant k<k_n\leqslant\infty$. *Предположим, что при всех* n, k *и* $l\neq k$ *независимы события* $A_{n,k}$ *и* $B_{n,k}$, *а также события* $A_{n,k}$ *и* $A_{n,l}$. *Если* $P(B_{n,k})\geqslant C>0$, $P(A_{n,k})\to 0$ *равномерно по* k *и* $P(\cup_k A_{n,k}B_{n,k})\to 0$, *то*

$$P(\cup_k A_{n,k}B_{n,k})\geqslant(C+o(1))\sum_k P(A_{n,k}).$$

Похожая лемма доказана, например, в [2, лемма 1]. Доказательство нашей леммы аналогично.

Обозначим $T_n = \max_{1\leqslant k\leqslant n} S_k/B_k$, $M_n = \max_{1\leqslant k\leqslant n} X_k/B_k$. Напомним, что $B_1 > 0$ и $B_n \to \infty$.

Лемма 4. *Пусть $\{a_n\}$ – последовательность положительных чисел, такая, что*

$$\liminf \frac{a_n^2}{\log B_n} > 0.$$

Тогда эквивалентны утверждения

(A) $P(T_n > \varepsilon a_n) \to 0$ *при каждом* $\varepsilon > 0$;

(B) $P(M_n > \varepsilon a_n) \to 0$ *при каждом* $\varepsilon > 0$;

(C) $\sum_{k=1}^n P(X_k > \varepsilon a_n B_k) \to 0$ *при каждом* $\varepsilon > 0$.

Доказательство. Эквивалентность (B) и (C) хорошо известна. Для доказательства импликации (A)$\Rightarrow$ (C) применим к событиям $A_{n,k} = \{X_k > 2\varepsilon B_k a_n\}$ и $B_{n,k} = \{S_{k-1} \geqslant -\varepsilon B_k a_n\}, 1 \leqslant k \leqslant n$, лемму 3 (в силу неравенства Чебышева $P(B_{n,k}) \to 1$ равномерно по k).

В заключение докажем, что из (C) следует (A). Обозначим $Y_{n,k} = X_k I(|X_k| \leqslant \delta B_k a_n)$, $R_{n,k} = Y_{n,1} + \cdots + Y_{n,k}$. Имеем при $B_0 = 0, \delta > 0$,

$$\max_{1\leqslant k\leqslant n} \frac{E|R_{n,k}|}{B_k} \leqslant \max_{1\leqslant k\leqslant n} \frac{1}{B_k} \sum_{i=1}^k E|X_i| I(|X_i| > \delta B_i a_n)$$

$$\leqslant \max_{1\leqslant k\leqslant n} \frac{1}{a_n B_k \delta} \sum_{i=1}^k \frac{EX_i^2}{B_i} \leqslant \max_{1\leqslant k\leqslant n} \frac{1}{a_n B_k \delta} \sum_{i=1}^k \int\limits_{B_{i-1}^2}^{B_i^2} \frac{dx}{\sqrt{x}} = \frac{1}{a_n \delta}.$$

Отсюда с помощью леммы 2 при достаточно больших n получаем

$$\begin{aligned} P(T_n > 2\varepsilon a_n) &\leqslant P\Big(\max_{1\leqslant k\leqslant n} \frac{R_{n,k}}{B_k} > 2\varepsilon a_n\Big) + \sum_{k=1}^n P(X_k > \delta a_n B_k) \\ &\leqslant P\Big(\max_{1\leqslant k\leqslant n} \frac{R_{n,k} - ER_{n,k}}{B_k} > \varepsilon a_n\Big) + o(1) \\ &\leqslant \frac{1}{\varepsilon^4 a_n^4}\Big(\sum_{k=1}^n \frac{E(Y_{n,k} - EY_{n,k})^4}{B_k^4} + 2\sum_{k=2}^n \frac{EX_k^2}{B_k^2}\Big) + o(1) \\ &\leqslant \frac{4a_n^2\delta^2 + 2}{\varepsilon^4 a_n^4} \sum_{k=1}^n \frac{EX_k^2}{B_k^2} + o(1) \end{aligned}$$

$$\leqslant \frac{4a_n^2\delta^2+2}{\varepsilon^4 a_n^4}(1+\log B_n^2/B_1^2)+o(1).$$

Левая часть этих неравенств от δ не зависит. Устремляя в правой части δ к нулю, получаем оценку

$$P(T_n > \varepsilon a_n) = o\Big(\frac{\log B_n}{a_n^2}\Big) + o(1),$$

откуда следует (A).

Доказательство теоремы 1 сводится к применению лемм 1 и 4 с $x_n = \sqrt{2\delta \log B_n}$.

Лемма 5. *В условиях теоремы 2 при каждом вещественном* x

$$P(L_nT_n < x + L_n^2 + \log L_n - 0.5\log(8\pi)) \to \exp(-e^{-x}).$$

Здесь $L_n = (2\log\log(B_n^2 \vee 10))^{1/2}$.

Эта лемма представляет собой один из вариантов теоремы Дарлинга–Эрдеша и была доказана в [5].

Доказательство теоремы 2. Пользуясь леммой 5 с

$$x=0, \quad x_n = L_n + \frac{\log L_n}{L_n} - \frac{\log 8\pi}{2L_n},$$

получаем $P(T_n < x_n) \to e^{-1}$. Теорема 2 вытекает теперь из леммы 1, поскольку

$$x_n \exp(x_n^2/2) \sim \frac{1}{\sqrt{8\pi}} L_n^2 \exp(L_n^2/2) \sim \sqrt{\frac{2}{\pi}} \log B_n \log\log B_n.$$

3. ЦПТ почти наверное в схеме серий. Пусть $X_{n,1},\dots,X_{n,n}$ – последовательность серий независимых в каждой серии случайных величин с $EX_{n,j}=0$ и $EX_{n,1}^2+\dots+EX_{n,n}^2=1$. Разобьем множество $\{1,\dots,n\}$ на непересекающиеся подмножества (не обязательно интервалы) $\Delta_{n,1},\dots\Delta_{n,l_n}$. Здесь $l_n\in\{1,\dots,n\}$. Обозначим

$$S_{n,j} = \sum_{i\in\Delta_{n,j}} X_{n,i}, \quad B_{n,j}^2 = ES_{n,j}^2.$$

Предположим, что $l_nB_{n,j}^2 \to 1$ равномерно по $j=1,\dots,n$, то есть разбиения выбираются так, чтобы выровнять суммарные дисперсии. Из дальнейшего видно, что содержательный результат получается при $l_n \to \infty$. При этом нужное разбиение может не существовать, если серия случайных величин вырождается, скажем, к одному (двум) заметным слагаемым. Нужное разбиение заведомо существует, если выполнено условие "бесконечной малости" $\max_j EX^2_{n,j} = o(1/l_n)$.

Теорема 3. *С вероятностью* 1 *равномерно по всем* x

$$\frac{1}{l_n}\sum_{j=1}^{l_n} I(S_{n,j} < xB_{n,j}) - \Phi(x) = O\Big(\sqrt{\frac{\log n}{l_n}} + \sqrt{l_n}\sum_{i=1}^{n} E|X_{n,i}|^3\Big). \quad (11)$$

Оценка (4) получается как следствие при $X_{n,i} = X_i/\sqrt{n}$.

Доказательство. Оценим вначале отклонение случайной величины

$$T_n(x) = \frac{1}{l_n}\sum_{j=1}^{l_n} I(S_{n,j} < xB_{n,j})$$

от ее математического ожидания. Для этого воспользуемся неравенством Хефдинга (см., например, [1, с. 76]: если $Y_1,\dots,Y_l$ – независимые случайные величины, $0 \leqslant Y_j \leqslant 1$ то

$$P(Y_1+\dots+Y_l - EY_1 - \dots - EY_l \geqslant ly) \leqslant \exp\left(-2ly^2\right), \quad y>0.$$

Имеем

$$P(|T_n(x) - ET_n(x)| \geqslant y) \leqslant 2\exp(-2l_n y^2). \quad (12)$$

Для тех n, для которых $r_n = \sqrt{\log n/l_n} \geqslant 1$, оценка (11) тривиальна. Для остальных $n>1$ (обозначим их множество A и предположим, что A бесконечно) обозначим $x_{n,j}$ решение уравнения $\Phi(x_{n,j}) = jr_n$ для тех целых j, для которых $0 < jr_n < 1$. Заменяя в (12) y на r_n, x на $x_{n,j}$ и суммируя по j, получаем при $n \in A$

$$\sum_j P(|T_n(x_{n,j}) - ET_n(x_{n,j})| \geqslant r_n) \leqslant 2r_n^{-1}\exp(-2l_n r_n^2) \leqslant 2n^{-1.5}.$$

Мы использовали здесь неравенство $r_n^{-1} \leqslant \sqrt{n}$, вытекающее из определения r_n и условия $l_n \leqslant n$. Суммируя эту оценку по $n \in A$, получаем с помощью леммы Бореля–Кантелли

$$P(\sup_j |T_n(x_{n,j}) - ET_n(x_{n,j})| \geqslant r_n \quad \text{б. ч. в А}) = 0. \quad (13)$$

Для оценки отклонения $ET_n(x)$ от $\Phi(x)$ воспользуемся неравенством Эссеена [1, с. 139]. При всех достаточно больших n и $j = 1, \ldots, l_n$ имеем

$$\sup_x |P(S_{n,j} < xB_{n,j}) - \Phi(x)| \leqslant 0.8 \sum_{i\in\Delta_{n,j}} E|X_{n,i}|^3/B_{n,j}^3$$
$$\leqslant l_n^{3/2} \sum_{i\in\Delta_{n,j}} E|X_{n,i}|^3.$$

После усреднения по j получаем при достаточно больших n

$$\sup_x |ET_n(x) - \Phi(x)| \leqslant \sqrt{l_n} \sum_{i=1}^n E|X_{n,i}|^3.$$

В сочетании с (13) это приводит при $n \in A$, $n \to \infty$, к оценке

$$\sup_j |T_n(x_{n,j}) - \Phi(x_{n,j})| = O\Big(\sqrt{\frac{\log n}{l_n}} + \sqrt{l_n} \sum_{i=1}^n E|X_{n,i}|^3\Big) \quad \text{п.н.}$$

Учитывая выбор $x_{n,j}$ и монотонность случайных величин $T_n(x)$ по x, получаем утверждение теоремы.

Литература

1. В. В. Петров, *Суммы независимых случайных величин.* М. (1972).
2. А. И. Мартикайнен, *Об усиленном заорне больших чисел для управляемых сумм.* — Вестник ЛГУ, вып. 1 (1977), 56–65.
3. G. A. Brosamler, *An almost everywhere central limit theorem.* — Math. Proc. Cambridge Philos. Soc. **104** (1988), 561–574.
4. P. Schatte, *On the central limit theorem with almost sure convergence.* — Math. Nachr. **137** (1988), 249–256.
5. U. Einmahl, D. M. Mason, *Darling–Erdös theorems for martingales.* — J. Theor. Probab. **2** (1989), 437–460.
6. M. Atlagh, *Théorème central limite prescue sûr et loi du logarithme itéré pour des sommes de variables aléatoires indépendantes.* — C. R. Acad. Paris, Sér. I **316** (1993), 929–933.
7. I. Berkes, H. Dehling, *Some limit theorems in log density.* — Ann. Probab. **21** (1993), 1640–1670.
8. I. Berkes, E. Csáki, *A universal result in almost sure central limit theorem.* — Stoch. Proc. Appl. **94** (2001), 105–134.
9. A. I. Martikainen, *On rate of convergence in almost sure central limit theorem.* Teor. Veroyatn. Mat. Statist. **66** (2002), 77–81.

Martikainen A. I. Almost sure central limit theorem without logarithmic sums.

We investigate possible rates of convergence in almost sure CLT for sums of independent random variables and martingales.

С.-Петербургский
государственный университет

Поступило 15 декабря 2003 г.