Math-Net.Ru

Общероссийский математический портал

И. Ринотт, В. И. Ротарь, Некоторые оценки скорости сходимости в ЦПТ для мартингалов. I, *Теория вероятн. и ее примен.*, 1998, том 43, выпуск 4, 692–710

DOI: <https://doi.org/10.4213/tvp2016>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.147.182.235

26 апреля 2023 г., 06:26:07

© 1998 г.

РИНОТТ И. *, РОТАРЬ В. И. **

НЕКОТОРЫЕ ОЦЕНКИ СКОРОСТИ СХОДИМОСТИ В ЦПТ ДЛЯ МАРТИНГАЛОВ. I¹⁾

Настоящая статья касается точности нормальной аппроксимации распределений случайных величин $S_n = \sum_1^n X_m$, где X_m — мартингал-разности. Известно, что в общем случае, даже если третьи моменты слагаемых конечны, точность аппроксимации не может иметь порядок лучший, чем $O(n^{-1/8})$. Если условные дисперсии $E\{X_m^2 \mid X_1, \dots, X_{m-1}\} = EX_m^2$, то скорость сходимости имеет порядок $O(n^{-1/4})$, в то время как при дополнительном условии независимости слагаемых точность аппроксимации имеет порядок $O(n^{-1/2})$. Настоящая статья представляет попытку объединить упомянутые выше случаи в одной оценке, а также рассмотреть ряд промежуточных ситуаций. Оценка дана в терминах определенных характеристик зависимости между слагаемыми, отражающих влияние различных факторов на скорость сходимости.

Ключевые слова и фразы: центральная предельная теорема, мартингалы, скорость сходимости.

1. Введение и основные результаты

1.1. Введение и цели работы. Пусть случайные величины (с.в.) $X_1, X_2, \dots$ образуют последовательность мартингал-разностей, точнее,

$$E\{X_m \mid \mathcal{F}^{m-1}\} = 0, \quad (1.1)$$

где $\mathcal{F}^m$ — σ -алгебра, порожденная с.в. $X_1, \dots, X_m$. Пусть $S_0 = 0$, $S_n = \sum_{m=1}^n X_m$ при $n \geq 1$. Положим $\sigma_m^2 = E\{X_m^2 \mid \mathcal{F}^{m-1}\}$, $\bar{\sigma}_m^2 = EX_m^2$ и $B_n^2 = \sum_{m=1}^n \bar{\sigma}_m^2$. Ниже мы будем также всегда полагать $E|X_m^3| < \infty$ для всех m .

Настоящая статья посвящена оценкам величины $\delta_n = \sup_x |\mathbf{P}\{S_n \leq x\} - \Phi(x/B_n)|$, где Φ — функция стандартного нормального распределе-

*Department of Mathematics, UCSD, La Jolla, CA 92093; e-mail: yrinott@ucsd.edu

**Центральный экономико-математический институт РАН, Нахимовский просп., 47, 117428 Москва, Россия; e-mail: rotar@cemi.rssi.ru

¹⁾Работа выполнена при поддержке гранта NSF #DMS 95-04616; второй автор пользовался также поддержкой Российского фонда фундаментальных исследований, грант № 98-0100358.

ния. Эта задача рассматривалась во многих работах; см., например, [2], [6], [8]–[10], [12]–[15], [17], [18], [20] и библиографию в этих работах. Одни из первых общих оценок были получены в [12] и [18], причем оценка в [18] была получена для семимартингалов и дана в терминах «первичных» характеристик процесса. Из этой работы следует, в частности (см. также [2], [8], [9]), что в общем случае нельзя надеяться на скорость сходимости порядка лучшего, чем $O(n^{-1/8})$. (Говоря так, мы, конечно, имеем в виду случай, когда значения слагаемых имеют один порядок; точные формулировки будут даны позже.)

Чтобы прийти к лучшей скорости, следует наложить дополнительные условия на структуру зависимости слагаемых, в частности, потребовать, чтобы условные дисперсии σ_m^2 были близки, в определенном смысле, к безусловным $\bar{\sigma}_m^2$.

Существенное продвижение в этом направлении было осуществлено в [2], где весьма точные оценки были получены при тех или иных условиях на совместное распределение слагаемых. В частности, там было показано, что, если $B_n^{-2} \sum_{m=1}^n \sigma_m^2 = 1$ и случайные слагаемые ограничены, скорость имеет порядок $O(\ln n / \sqrt{n})$, а также установлены некоторые условия, при которых скорость имеет тот же порядок, что и в случае независимости, т.е. $O(n^{-1/2})$. Работа [2] содержит также ряд примеров, указывающих на наилучшую возможную точность нормальной аппроксимации при различных типах зависимости. В частности, в [2] показано, что в случае, когда $\sigma_m^2 = \bar{\sigma}_m^2$, без дополнительных условий можно надеяться только на скорость порядка $O(n^{-1/4})$.

Оценки работы [16], полученные для частного случая $\sigma_m^2 \equiv 1$, включали некоторые специфические характеристики зависимости, позволившие рассмотреть одновременно как крайние случаи: наихудший со скоростью $O(n^{-1/4})$ и наилучший, т.е. случай независимых слагаемых, со скоростью $O(n^{-1/2})$, так и некоторые «промежуточные» случаи. Там же было показано, каким образом порядок наибольшего конечного момента слагаемых влияет на точность аппроксимации. Настоящая статья использует некоторые модификации характеристик из [16]; мы обсудим их в следующем разделе.

Случай конечности моментов высоких порядков в общей ситуации был рассмотрен в [9]. Неравномерные оценки того же типа были получены в [10].

Настоящая статья представляет попытку рассмотреть указанные выше крайние случаи и ряд промежуточных одновременно, в одной оценке. Понятно, что, чтобы осуществить это, надо ввести в рассмотрение характеристики зависимости, отражающие влияние на скорость сходимости, по крайней мере, следующих факторов:

- близости $\sum_{m=1}^n \sigma_m^2 / B_n^2$ к единице;
- близости условных дисперсий к безусловным (оба аспекта в той

или иной степени учитывались во всех работах на эту тему);

— более «тонких» особенностей зависимости, которые могли бы привести к точности аппроксимации лучшей, чем $O(n^{-1/4})$.

Характеристики, предлагаемые в настоящей работе, по мнению авторов, достаточно естественны, однако их вычисление в конкретных ситуациях может потребовать дополнительной работы по учету тех или иных особенностей зависимости. Ниже будет приведен ряд примеров.

Хочется также подчеркнуть, что оценки, полученные ниже, совсем не претендуют на то, что из них «все следует», — мы рассмотрим, в частности, ситуации, когда изучение конкретных случаев приводит к оценкам, лучшим, чем те, что следуют из общих утверждений. Цель настоящей работы проще: исследовать возможность единого подхода к различным типам зависимости и попытаться понять, какие характеристики зависимости могут быть ответственны за скорость сходимости.

1.2. Основной результат и обсуждения. Ниже одним и тем же символом C обозначаются абсолютные постоянные, вообще говоря, различные в различных формулах.

Вместе со с.в. X_j мы будем рассматривать нормированные с.в. $Y_j = X_j/\sigma_j$ и предполагать, что $E|X_j|^3 < \infty$, $E|Y_j|^3 < \infty$. Соотношение между указанными моментами может быть произвольным.

Пусть

$$\Gamma_{1n} := E \left| 1 - B_n^{-2} \sum_{m=1}^n \sigma_m^2 \right|.$$

Эта естественная характеристика использовалась во многих работах; см., например, [11], [18], [2].

Далее, для $j, m = 1, 2, \dots$, $j \geq m$, положим $\sigma_{jm}^2 = E\{X_j^2 | \mathcal{F}^{m-1}\}$. В частности, $\sigma_{mm}^2 = \sigma_m^2$. При фиксированном n и $m = 1, \dots, n-1$ положим $V_{mn}^2 = \sum_{j=m+1}^n \sigma_{jm}^2$ и $\bar{V}_{mn}^2 = \sum_{j=m+1}^n \bar{\sigma}_j^2$. Мы будем описывать степень близости условных дисперсий к их средним в терминах с.в.

$$\alpha_m = \frac{\sigma_m^2 - \bar{\sigma}_m^2}{\sigma_m^2 + \bar{\sigma}_m^2} \quad \text{и} \quad \beta_{mn} = \frac{V_{mn}^2 - \bar{V}_{mn}^2}{V_{mn}^2 + \bar{V}_{mn}^2} \quad \text{для } m = 1, \dots, n-1,$$

полагая $\beta_{nn} = 0$.

Если знаменатели в последнем определении равны нулю, полагаем $\alpha_m = \beta_{mn} = 0$.

Заметим, что α_m, β_{mn} измеримы относительно $\mathcal{F}^{m-1}$ и $-1 \leq \alpha_m, \beta_{mn} \leq 1$. В качестве соответствующей детерминированной характеристики рассмотрим величину

$$\Gamma_{2n} := \frac{1}{B_n^3} \sum_{m=1}^n E(|\alpha_m| + |\beta_{mn}|) (|X_m^3| + \bar{\sigma}_m^3 |Y_m^3|).$$

Ясно, что если $\sigma_m^2 = \bar{\sigma}_m^2$ для всех $m \leq n$, то $\Gamma_{1n} = \Gamma_{2n} = 0$. Грубой оценкой для Γ_{2n} может служить неравенство

$$\Gamma_{2n} \leq \frac{2}{B_n^3} \sum_{m=1}^n \bar{\sigma}_m^3 (\mu_m + \nu_m), \quad (1.2)$$

где $\mu_m := E|X_m^3|/\bar{\sigma}_m^3$, $\nu_m := E|Y_m^3|$.

Обратимся к характеристикам, отражающим более «тонкие» особенности зависимости. Для любой с.в. Y положим

$$\Delta_m(Y) := \sup_z \left| P\{S_m \leq z \mid Y\} - P\{S_m \leq z\} \right|.$$

(Выбирается регулярный вариант условного распределения и, так как достаточно рассмотреть $\sup$ только по множеству рациональных z , величина $\Delta_m(Y)$ измерима.)

Пусть $r_m = E\{|Y_m^3| \Delta_{m-1}(Y_m)\}$, иначе $r_m = \int_{-\infty}^{\infty} |y|^3 \sup_z |P\{S_{m-1} \leq z \mid Y_m = y\} - P\{S_{m-1} \leq z\}| P\{Y_m \in dy\}$. Очевидно, что $r_m \leq \nu_m$ и $r_m \equiv 0$, если с.в. $X_1, \dots, X_m$ независимы.

То, что в определении r_m фигурирует не само слагаемое X_m , а его «нормализация» Y_m , может улучшить характеристику: деля X_m на условное среднеквадратичное отклонение, можно «уменьшить влияние» S_{m-1} . Пусть, например, $X_m = h(S_{m-1})Z_m$, где h — некоторая положительная функция, и Z_m не зависит от S_{m-1} , $EZ_m = 0$. Тогда, как легко видеть, $r_m = 0$, в то время как $\hat{r}_m := E\{|X_m^3| \Delta_{m-1}(X_m)\}$ может и не равняться нулю. С другой стороны, последняя характеристика может оказаться более легкой для вычислений. В разделе 4 мы коснемся оценок, данных в терминах $\hat{r}_m$.

Вводя в рассмотрение характеристики r_m , мы исходим из предположения, что при не слишком обременительных условиях эти характеристики могут быть малы при больших m : «сумма большого числа слагаемых S_{m-1} слабо зависит от отдельно взятой с.в. X_m ». В следующем разделе будут рассмотрены примеры, подтверждающие эту гипотезу, и некий общий подход к вычислению r_m .²⁾

В частном случае $\sigma_j^2 = \bar{\sigma}_j^2$ характеристики, близкие по смыслу к r_m , рассматривались ранее в [16]. Отличие в том, что в характеристиках из [16] фигурирует условное распределение с.в. X_m при заданной S_{m-1} , что

²⁾Убедиться в правдоподобности указанного предположения можно, вычислив сперва более простую характеристику, именно, корреляцию $\text{Corr}\{S_{m-1}, g(X_m)\}$, где g — некоторая функция. Пусть, например, сумма $\sum_{j=1}^{m-1} |\text{Corr}\{X_j, g(X_m)\}|$ ограничена равномерно по m (что верно, скажем, при достаточно слабых условиях типа перемешивания). Предположим также, что все $\bar{\sigma}_m^2$ имеют один порядок (скажем, в смысле условия (1.4) ниже), и то же верно для дисперсий величин $g(X_m)$. Тогда, как легко подсчитать, $\text{Corr}\{S_{m-1}, g(X_m)\} = O(1/\sqrt{m})$. Конечно, вычислить r_m существенно труднее.

грубее. Например, если значения X_m определяются тем, приняли предшествующие слагаемые рациональные или иррациональные значения, то значение S_{m-1} содержит существенную информацию о возможных значениях X_m и «влияние» S_{m-1} на X_m может и не уменьшаться при росте m .

Последний шаг в подготовке к формулировке результата касается того, что мартингальный тип зависимости связан с порядком, в котором рассматриваются слагаемые, и не безразлично, стоят слагаемые с тем или иным свойством в начале или в конце последовательности $X_1, \dots, X_n$. До некоторой степени это можно учесть, вводя для любой последовательности чисел $a_1, a_2, \dots$ операцию

$$a_m^* = \frac{1}{B_m^2} \sum_{j=1}^m \bar{\sigma}_j^2 \max_{j \leq k \leq m} a_k. \quad (1.3)$$

Если $a_m \rightarrow 0$ при $m \rightarrow \infty$ (что, как правило, верно для характеристик, рассматриваемых ниже), то операция (1.3) изменяет последовательность a_m не существенно. Действительно, если a_m убывают, величина a_m^* равна взвешенному среднему $B_m^{-2} \sum_{j=1}^m \bar{\sigma}_j^2 a_j$. Более того, если $\bar{\sigma}_j^2$ имеют один порядок роста, скажем,

$$c^{-1} \leq \bar{\sigma}_j^2 \leq c \quad (1.4)$$

для всех j и некоторого положительного абсолютного c , и если a_m убывают «регулярно и не слишком быстро» (например, если $a_m = h(m)/m^p$, где $0 < p < 1$ и h — медленно меняющаяся функция), то a_m^* имеет тот же порядок, что и a_m (это следует из того, что $\int_0^b h(x) x^{-p} dx \sim h(b) b^{1-p}/(1-p)$ для $p < 1$ и медленно меняющейся h ; см., например, [25, с. 297]).

Теорема 1. *Существует абсолютная постоянная C такая, что*

$$\delta_n \leq C \left\{ \frac{(\nu_n \bar{\sigma}_n)^*}{B_n} + \sqrt{\frac{(r_n \bar{\sigma}_n)^*}{B_n}} + (\Gamma_{1n}^*)^{1/3} + (\Gamma_{2n}^*)^{1/4} \right\}. \quad (1.5)$$

Перейдем к обсуждениям.

1.2.1. Крайние случаи. Положим $\mu = \sup_m \mu_m$. Пусть сначала с.в. X_m независимы. В этом случае все слагаемые в правой части (1.5), за исключением первого, равны нулю, и $\mu_m = \nu_m$. Следовательно, $\delta_n \leq C(\mu_n \bar{\sigma}_n)^*/B_n \leq C\mu \bar{\sigma}_n^*/B_n$. Так как в случае независимости порядок слагаемых несуществен, предположим, что $\bar{\sigma}_m$ не возрастают с ростом m . Тогда $\bar{\sigma}_n^* = \sum_1^n \sigma_m^3/B_n^2$, и (1.5) влечет оценку $\delta_n \leq C\mu \sum_1^n \sigma_m^3/B_n^3$, которая мало отличается от оценки Берри–Эссеена. В частности, если выполнено (1.4), то $\delta_n \leq C\mu/\sqrt{n}$, где C — постоянная. (Конечно, C зависит от c из (1.4); мы не выражаем эту зависимость явно ни в этом

примере, ни в более сложных ниже, где также используется (1.4).) Мы рассмотрели наилучший случай. Перейдем к наихудшему.

Пусть $\nu = \sup_m \nu_m$. Ввиду (1.2) $\Gamma_{2m} \leq 2(\mu + \nu) \sum_1^m \bar{\sigma}_j^3 / B_m^3$. Если верно (1.4), $\Gamma_{2m} \leq C(\mu + \nu) / \sqrt{m}$, и, осуществляя операцию (1.3), получаем, что $\Gamma_{2n}^* \leq Cn^{-1} \sum_{m=1}^n (\mu + \nu) / \sqrt{m} \leq C(\mu + \nu) / \sqrt{n}$. В случае (1.4) $(\nu_n \bar{\sigma}_n)^* \leq C\nu$, $(r_n \bar{\sigma}_n)^* \leq C\nu$, $B_n^2 \geq Cn$. Следовательно, $\delta_n \leq C\{\nu n^{-1/2} + \nu^{1/2} n^{-1/4} + \Gamma_{1n}^*\}^{1/3} + (\mu + \nu)^{1/4} n^{-1/8}$, и оценка имеет порядок $(\Gamma_{1n}^*)^{1/3} + n^{-1/8}$. (Члена с Γ_{1n}^* в такого рода оценках, конечно, не избежать.) Обратимся к промежуточным случаям.

1.2.2. Влияние условных дисперсий. Возможность близости σ_m^2 к $\bar{\sigma}_m^2$ отражена в двух последних членах (1.5). Как уже отмечалось, член с Γ_{1n} присутствует во многих полученных ранее оценках; см., например, [12], [2], [18]. Степень при Γ_{1n} в (1.5) та же, что и в [12], [2], [18]. Так как в нетривиальных ситуациях $\Gamma_{1m} \rightarrow 0$ при $m \rightarrow \infty$, переход от Γ_{1n} к Γ_{1n}^* не меняет оценку по существу. Отметим, что в [12], [2] были также получены оценки, использующие характеристику $\|1 - B_m^{-2} \sum_{j=1}^m \sigma_j^2\|_\infty$. Как показано в [12], [2], если указанная характеристика конечна, ее использование приводит к лучшей оценке.

Обратимся к Γ_{2n}^* . Для того чтобы правая часть (1.5) имела лучший порядок, чем $O(n^{-1/8})$, необходимо, чтобы Γ_{2n}^* имела порядок лучший, чем $O(n^{-1/2})$. Последнее возможно, только если σ_m^2 в определенном смысле близки к $\bar{\sigma}_m^2$. Рассмотрим в качестве примера случай, когда указанная близость имеет место лишь при больших m . Предполагая, что $\|\sigma_j\|_\infty < \infty$ при всех j , положим $e_j = \sup_{k \geq j} \|1 - \sigma_k^2 / \bar{\sigma}_k^2\|_\infty$. (Как обычно, $\|\cdot\|_p$ — норма в L_p .) Тогда, как легко видеть, $\|\alpha_j\|_\infty \leq e_j$, $\|\beta_{jm}\|_\infty \leq e_j$ при всех j и m . Следовательно, $\Gamma_{2m} \leq 2 \sum_1^m e_j \bar{\sigma}_j^3 (\mu_j + \nu_j) / B_m^3 \leq 2(\mu + \nu) \sum_1^m e_j \bar{\sigma}_j^3 / B_m^3$, и мы можем провести операцию (1.3).

Пусть, например, выполнено (1.4) и $e_j = O(j^{-s})$. В [2] показано, что при $0 \leq s < \frac{1}{2}$ (и некоторых других ограничениях — мы не приводим их, так как сейчас нас интересует лишь влияние σ_m) скорость сходимости имеет порядок $O(n^{-s})$. В нашем случае, если $0 \leq s < 1$, величина $\Gamma_{2m} = O(m^{-s-1/2})$, и при $0 \leq s < \frac{1}{2}$, как легко подсчитать, последний член в (1.5) $(\Gamma_{2n}^*)^{1/4} = O(n^{-1/8-(s/4)})$. При $s > \frac{1}{6}$ это хуже, чем скорость в [2], при $s < \frac{1}{6}$ — лучше. При $s = \frac{1}{2}$ имеем $(\Gamma_{2n}^*)^{1/4} = O([\ln n / n]^{1/4})$, а при $s > \frac{1}{2}$ верно $(\Gamma_{2n}^*)^{1/4} = O(n^{-1/4})$.

Если $\|\sigma_m\|_\infty = \infty$, но X_j имеют моменты более высокого порядка, чем три, можно поступать аналогично, отталкиваясь от неравенства $E(|\alpha_j| + |\beta_{jm}|) (|X_j^3| + \bar{\sigma}_j^3 |Y_j^3|) \leq \|(|\alpha_j| + |\beta_{jm}|)\|_q \| |X_j^3| + \bar{\sigma}_j^3 |Y_j^3| \|_p$, где $p^{-1} + q^{-1} = 1$. Если конечны только третьи моменты, с.в. $|\alpha_j| + |\beta_{jm}|$ и $|X_j^3| + \bar{\sigma}_j^3 |Y_j^3|$ рассматривать отдельно нельзя.

Заметим также, что Γ_{1m} , конечно, тоже можно оценивать в терми-

нах e_j , но это было бы слишком грубо — сходимость $B_m^{-2} \sum_{j=1}^m \sigma_j^2$ к 1 суть действие закона больших чисел, который проявляется не за счет близости отдельных случайных слагаемых к их средним. В разделе 4 мы коснемся возможности улучшения оценок за счет дополнительных условий на σ_j^2 , а сейчас рассмотрим

1.2.3. Основные члены в (1.5). Чтобы не смешивать их влияние с влиянием условных дисперсий, положим для простоты $\sigma_m^2 = \bar{\sigma}_m^2$ для всех $m \leq n$. Тогда $\Gamma_{1n}^* = \Gamma_{2n}^* = 0$, $\nu_m = \mu_m$ и

$$\delta_n \leq C \left\{ \frac{(\mu_n \bar{\sigma}_n)^*}{B_n} + \sqrt{\frac{(r_n \bar{\sigma}_n)^*}{B_n}} \right\}. \quad (1.6)$$

Пусть снова верно (1.4). Используя грубую оценку $r_n \leq \mu$, получаем, что $\delta_n \leq C\mu/n^{1/4}$. Как показано в [2], эта оценка точна.

С другой стороны, при не слишком сильных дополнительных условиях r_m может быть мало при больших m , что улучшает скорость сходимости. Предположим, например, что $r_m \leq CLm^{-1/2}$, где L — некоторая не зависящая от m характеристика распределений слагаемых (в следующем разделе будут рассмотрены ситуации, когда это верно, и получены явные выражения для L). Тогда, как легко видеть, $r_n^* \leq CLn^{-1} \sum_1^n m^{-1/2} \leq CLn^{-1/2}$ и $\delta_n = O(n^{-1/2})$, если верно (1.4).

Рассмотрим более общий случай, когда $r_m = O(h(m)m^{-p})$, где h — медленно меняющаяся функция. Достаточно рассмотреть $p \leq \frac{1}{2}$, и $h(m) \geq 1$, если $p = \frac{1}{2}$. (Если $r_m = o(m^{-1/2})$, второй член в (1.6) оказывается меньше, чем первый.) Тогда, как нетрудно проверить, из (1.4) и (1.6) следует, что

$$\delta_n \leq O(h^{1/2}(n)n^{-p/2-1/4}).$$

В этой статье мы ограничимся примерами ситуаций, когда $r_m \rightarrow 0$ при $m \rightarrow \infty$. Хотя ниже для ряда случаев будут получены конкретные оценки δ_n , цель следующего раздела состоит не столько в этих оценках, сколько в том, чтобы показать, что оценивание r_m возможно. Основная идея раздела состоит в том, чтобы получать оценки r_m не прямыми вычислениями, что представляется очень непростым, а в терминах δ_m . Хотя δ_m неизвестно и является как раз тем, ради чего проводятся все вычисления, указанный подход не ведет к «порочному кругу», если проводить вычисления обеих характеристик — r_m и δ_m — одновременно.

2. Примеры оценивания r_m

В этом разделе мы вновь ограничимся случаем $\sigma_m^2 = \bar{\sigma}_m^2$ для всех m . Чтобы упростить вычисления, будем также предполагать, что выполнено (1.4).

Прежде всего, несколько обобщим определение δ_n . Пусть $\bar{j}^{(n)} = \{j_1, \dots, j_n\}$, где $j_1 < \dots < j_n$ — целые. Положим $S_{\bar{j}^{(n)}} = \sum_{k=1}^n X_{j_k}$, $B_{\bar{j}^{(n)}}^2 = \sum_{k=1}^n \sigma_{j_k}^2$, и $\tilde{\delta}_n = \sup_{\bar{j}^{(n)}} \sup_z |\mathbf{P}\{S_{\bar{j}^{(n)}} < z\} - \Phi(z/B_{\bar{j}^{(n)}})|$, где первый $\sup$ берется по всем выборкам $\bar{j}^{(n)}$ размера n . Если использовать характеристики, относящиеся ко всей последовательности $X_1, X_2, \dots$, такая модификация не приведет ни к каким противоречиям, так как любые $X_{j_1}, \dots, X_{j_n}$ образуют последовательность мартингал-разностей.

Чтобы пояснить последующие выкладки, рассмотрим сначала

2.1. Простой пример. Пусть выполнено (1.4) все $|X_m|$ не превосходят некоторой постоянной $H \geq 1$ и для каждого m существует множество индексов $\mathcal{W}_m$ такое, что X_m не зависит от совокупности $\{X_j: j \notin \mathcal{W}_m\}$. Пусть $k = \max_m |\mathcal{W}_m| < \infty$, где $|W|$ — число точек в W . Такую структуру зависимости называют локальной, а множества $\mathcal{W}_m$ — множествами зависимости или окрестностями зависимости (в англоязычной литературе — *dependency sets* или *dependency neighborhoods*; см., например, [27], [23], [24]). Простейший пример локальной зависимости — k -зависимые с.в.

Полагая $m > 2k$, исключим из суммы $S_m = \sum_1^m X_j$ произвольные k слагаемых, включая, однако, все те, чьи индексы принадлежат $\mathcal{W}_{m+1}$. Обозначим оставшуюся сумму $m - k$ слагаемых $\tilde{S}_{m-k}$. Ясно, что $\tilde{S}_{m-k}$ и X_{m+1} независимы. Пусть $\tilde{B}_{m-k}^2 = \mathbf{E}\tilde{S}_{m-k}^2$. Тогда

$$\begin{aligned} & \mathbf{P}\{S_m < z \mid X_{m+1}\} - \mathbf{P}\{S_m < z\} \\ & \leq \mathbf{P}\{\tilde{S}_{m-k} < z + kH \mid X_{m+1}\} - \mathbf{P}\{S_m < z\} \\ & = \mathbf{P}\{\tilde{S}_{m-k} < z + kH\} - \mathbf{P}\{S_m < z\} \\ & \leq \tilde{\delta}_{m-k} + \tilde{\delta}_m + \Phi\left(\frac{z + kH}{\tilde{B}_{m-k}}\right) - \Phi\left(\frac{z}{B_m}\right) \\ & \leq \tilde{\delta}_{m-k} + \tilde{\delta}_m + \frac{B_m - \tilde{B}_{m-k} + kH}{\tilde{B}_{m-k}} \leq \tilde{\delta}_{m-k} + \tilde{\delta}_m + \frac{CkH}{\sqrt{m}}. \end{aligned}$$

На последнем шаге мы использовали неравенство $m > 2k$ и оценку (1.4). Нижняя оценка получается аналогично, так что $\Delta_m(X_{m+1}) := \sup_z |\mathbf{P}\{S_m < z \mid X_{m+1}\} - \mathbf{P}\{S_m < z\}| \leq \tilde{\delta}_{m-k} + \tilde{\delta}_m + CkH/\sqrt{m}$, и, следовательно,

$$r_{m+1} \leq C(\mu(\tilde{\delta}_{m-k} + \tilde{\delta}_m + kHm^{-1/2})). \quad (2.1)$$

С другой стороны, из (1.6) и (1.4) следует, что

$$\delta_n \leq C\{\mu n^{-1/2} + (r_n^*)^{1/2} n^{-1/4}\}. \quad (2.2)$$

Так как индекс в левой части (2.1) больше индексов в правой, можно применить индукцию. Положив $K = \mu + \sqrt{\mu k H}$, докажем, что

$\tilde{\delta}_m \leq \bar{C}K/\sqrt{m}$ для некоторой постоянной $\bar{C} \geq 1$. При $m \leq 2k$ справедливость оценки (при $C = 2$) очевидна. Предположим, что она верна при $m \leq n - 1$. Тогда из (2.1) следует, что $r_m \leq C\mu(\bar{C}K + kH)m^{-1/2}$, если $2k < m \leq n$. Осуществим операцию (1.3), используя (1.4) и при $m \leq 2k$ простое неравенство $r_m \leq \mu$. Учитывая также, что $H \geq 1$, нетрудно получить, что (1.3) влечет оценку $r_n^* \leq C(\bar{C}\mu K + \mu kH)n^{-1/2}$. Отсюда и из (2.2) получаем, что $\delta_n \leq C\{\mu + \sqrt{\bar{C}}\sqrt{\mu K} + \sqrt{\mu kH}\}n^{-1/2} \leq C\sqrt{\bar{C}}Kn^{-1/2}$, поскольку $\mu \leq K$. Очевидно, что форма полученной оценки позволяет распространить ее и на величины $\tilde{\delta}_n$. Чтобы завершить индукцию, осталось положить $\bar{C} \geq C^2$. Отметим, что, проводя означенные вычисления, мы получили оценку и для r_n .

Прежде чем перейти к более сложным примерам, сформулируем несколько общих утверждений. По существу, они представляют неравенства типа сглаживания для условных распределений и могут быть интересны сами по себе.

2.2. Сглаживание для условных распределений. Ниже символ $\mathcal{F}_X$ обозначает σ -алгебру, порожденную с.в. X . Для произвольных σ -алгебры $\mathcal{F}$ и с.в. X определим величину

$$\Delta(X | \mathcal{F}) := \sup_x |\mathbf{P}\{X < x | \mathcal{F}\} - \mathbf{P}\{X < x\}|.$$

Положим также

$$\delta(X) = \sup_z \left| \mathbf{P}\left\{X < \mathbf{E}X + z\sqrt{\mathbf{D}\{X\}}\right\} - \Phi(z) \right|.$$

Пусть теперь U, V — некоторые с.в. с нулевыми средними и единичными дисперсиями и $S = \varepsilon U + \sqrt{1 - \varepsilon^2}V$, где фиксированное $\varepsilon \in [0, 1]$. Некоррелированность U и V не предполагается.

Лемма 1. Пусть σ -алгебра $\mathcal{F}' \supseteq \mathcal{F}_U$. Тогда для любой σ -алгебры $\mathcal{F}$, числа $q \geq 2$ и $\varepsilon \in [0, 1]$

$$\begin{aligned} \|\Delta(S | \mathcal{F})\|_q &\leq C \left\{ \|\Delta(V | \mathcal{F})\|_q + \delta(S) + \delta(V) + \varepsilon\|U\|_q \right. \\ &\quad \left. + (\varepsilon\|U\|_q)^{1-1/q} \left(\|\Delta(V | \mathcal{F}' \vee \mathcal{F})\|_q \right)^{1/q} \right\}. \end{aligned} \quad (2.3)$$

При доказательстве неравенства (2.3) мы покажем, что оно остается верным и при всех $q \geq 1$, если к правой части добавить $\varepsilon^2 + |\tau|\varepsilon$, где $\tau = \mathbf{E}UV$. При $q \geq 2$ указанным членом можно пренебречь, так как он не больше, чем четвертый член в (2.3), поскольку $\|U\|_2 = 1$. То же касается и последующих утверждений.

Далее, пусть $\|X\|_{q|\mathcal{F}} := (\mathbf{E}\{|X|^q | \mathcal{F}\})^{1/q}$. Обобщение (2.3) представляет

Лемма 2. Для любого $q \geq 2$, $p^{-1} + q^{-1} = 1$, $\varepsilon \in [0, 1]$ и $t \geq 1$

$$\|\Delta(S | \mathcal{F})\|_q \leq C \left\{ \|\Delta(V | \mathcal{F})\|_q + \delta(S) + \delta(V) + \varepsilon \left\| (\|U\|_{t|\mathcal{F}}) \right\|_q + \left[\varepsilon \left\| (\|U\|_{t|\mathcal{F}}) \right\|_q \right]^{t/(t+p)} \left(\|\Delta(V | \mathcal{F}' \vee \mathcal{F})\|_q \right)^{p/(t+p)} \right\}. \quad (2.4)$$

Ясно, что $\|(\|U\|_{t|\mathcal{F}})\|_q \leq \|U\|_q$, если $t \leq q$, причем при $t = q$ имеет место равенство. Следовательно, (2.4) влечет (2.3).

Преимущество (2.4) — в том, что $\|(\|U\|_{t|\mathcal{F}})\|_q$ может быть конечным, в то время как $\|U\|_q = \infty$. Предположим, например, что $\| \|U\|_{t|\mathcal{F}} \|_\infty < \infty$ для некоторого $t > 1$. Тогда в (2.4) можно положить $q = \infty$, что влечет оценку

$$\|\Delta(S | \mathcal{F})\|_\infty \leq C \left\{ \|\Delta(V | \mathcal{F})\|_\infty + \delta(S) + \delta(V) + \varepsilon \left\| (\|U\|_{t|\mathcal{F}}) \right\|_\infty + \left[\varepsilon \left\| (\|U\|_{t|\mathcal{F}}) \right\|_\infty \right]^{t/(t+1)} \left(\|\Delta(V | \mathcal{F}' \vee \mathcal{F})\|_\infty \right)^{1/(t+1)} \right\}. \quad (2.5)$$

При дополнительных условиях, имеющих характер марковских, член $\Delta(V | \mathcal{F})$ можно рассмотреть отдельно. Предположим, что для некоторой σ -алгебры $\mathcal{A}$

$$\mathbf{P}\{B | \mathcal{A} \vee \mathcal{F}\} = \mathbf{P}\{B | \mathcal{A}\} \quad \text{для любого } B \in \mathcal{F}_V. \quad (2.6)$$

Ниже нам будет удобно использовать для условных распределений запись типа $\mathbf{P}\{B | \mathcal{F}\}[\omega]$ (представляя элементарный исход ω , от которого зависит $\mathbf{P}\{B | \mathcal{F}\}$, в явном виде). Заметим также, что в регулярном варианте условной функции распределения типа $\mathbf{P}\{V < y | \mathcal{F}\}$, где V — с.в., в качестве аргумента y может выступать любая с.в., измеримая относительно $\mathcal{F}$; см., например, [7].

Обозначим $G_{\mathcal{A}|\mathcal{F}}(B | \omega)$ сужение условного распределения $\mathbf{P}\{B | \mathcal{F}\}[\omega]$ на $B \in \mathcal{A}$. Пусть $L_{\mathcal{A}|\mathcal{F}}(B | \omega', \omega'') = G_{\mathcal{A}|\mathcal{F}}(B | \omega') - G_{\mathcal{A}|\mathcal{F}}(B | \omega'')$, и $\rho_{\mathcal{A}|\mathcal{F}} = \rho_{\mathcal{A}|\mathcal{F}}(\omega', \omega'') = \frac{1}{2} \int |L_{\mathcal{A}|\mathcal{F}}(d\omega | \omega', \omega'')|$, где $|L_{\mathcal{A}|\mathcal{F}}(d\omega | \cdot)|$ обозначает вариацию. Заметим, что с.в. $\rho_{\mathcal{A}|\mathcal{F}}$ измерима относительно $\mathcal{F} \times \mathcal{F}$. При $q \geq 1$ рассмотрим норму

$$\|\rho_{\mathcal{A}|\mathcal{F}}\|_q := \left[\int \int \rho_{\mathcal{A}|\mathcal{F}}^q(\omega', \omega'') \mathbf{P}\{d\omega'\} \mathbf{P}\{d\omega''\} \right]^{1/q}.$$

Лемма 3. Для любых $q \geq 1$ и $s \in [0, 1]$

$$\|\Delta(V | \mathcal{F})\|_q \leq 4 \|\Delta(V | \mathcal{A})\|_q^s \cdot (\|\rho_{\mathcal{A}|\mathcal{F}}\|_{(q-1)/(1-s)})^{(q-1)/q}.$$

При $\|\rho_{\mathcal{A}|\mathcal{F}}\|_\infty < \infty$ можно положить $s = 1$ и $q = \infty$, что влечет оценку

$$\|\Delta(V | \mathcal{F})\|_\infty \leq 4 \|\Delta(V | \mathcal{A})\|_\infty \cdot \|\rho_{\mathcal{A}|\mathcal{F}}\|_\infty. \quad (2.7)$$

2.3. Дальнейшие примеры. При фиксированном m положим $S_m = \tilde{S}_{m-k} + Z_k$, где Z_k — сумма k с.в., выбранных каким-то образом из последовательности $X_1, \dots, X_m$, и $\tilde{S}_{m-k}$ — остаточная сумма. Пусть $U = Z_k / \sqrt{\mathbf{E}Z_k^2}$, $V = \tilde{S}_{m-k} / \tilde{B}_{m-k}$, где вновь $\tilde{B}_{m-k}^2 = \mathbf{E}\tilde{S}_{m-k}^2$. Тогда $(S_m/B_m) = S := \varepsilon U + \sqrt{1 - \varepsilon^2}V$, где $\varepsilon^2 = \mathbf{E}Z_k^2/B_m^2$. Положим $\psi(q) = \sup_m \|X_m\|_q$.

При условии (1.4) $\varepsilon^2 \leq Ck/m$, $\psi(q) \geq \psi(2) \geq C$ для $q \geq 2$, и $\|U\|_q \leq C\sqrt{k}\psi(q)$. (Заметим, что, используя мартингалльное свойство и неравенство Буркхольдера, можно исключить k в правой части последнего неравенства, но в этом случае C зависело бы от q , что для нас нежелательно.)

Учитывая, что $\|\Delta(S_m | \mathcal{F})\|_q = \|\Delta(S_m/B_m | \mathcal{F})\|_q$, и применяя (2.3), получаем, что для любой $\mathcal{F}$ и любой $\mathcal{F}' \supseteq \mathcal{F}_{Z_k}$

$$\begin{aligned} \|\Delta(S_m | \mathcal{F})\|_q &\leq C \left\{ \|\Delta(\tilde{S}_{m-k} | \mathcal{F})\|_q + \delta_m + \tilde{\delta}_{m-k} + k\psi(q) m^{-1/2} \right. \\ &\quad \left. + \left(k\psi(q) m^{-1/2} \right)^{1-1/q} \left(\|\Delta(\tilde{S}_{m-k} | \mathcal{F}' \vee \mathcal{F})\|_q \right)^{1/q} \right\}. \end{aligned} \quad (2.8)$$

Пусть теперь $\mathcal{F}_m = \mathcal{F}_{X_m}$. Так как мы предположили $\sigma_m = \bar{\sigma}_m$, по определению,

$$\begin{aligned} r_{m+1} &\leq \bar{\sigma}_m^{-3} \|X_{m+1}\|_p \|\Delta(S_m | \mathcal{F}_{m+1})\|_q \\ &\leq \bar{\sigma}_m^{-3} \psi^3(3p) \|\Delta(S_m | \mathcal{F}_{m+1})\|_q \end{aligned} \quad (2.9)$$

при $p^{-1} + q^{-1} = 1$. Сравнивая порядки норм в (2.9) и (2.8), мы видим, что плата за грубое оценивание в (2.9) состоит в том, что теперь мы должны предполагать конечность четвертых моментов $\mathbf{E}X_m^4$ (так как если $3p = q$, то $q = 4$). Поэтому будем рассматривать только $q \geq 4$, что вместе с (1.4) влечет оценку

$$r_{m+1} \leq C\psi^3(q) \|\Delta(S_m | \mathcal{F}_{m+1})\|_q. \quad (2.10)$$

Прежде всего, вернемся к примеру п. 2.1.

2.3.1. Локальная зависимость. Случай ограниченных слагаемых, затронутый в п. 2.1 для примера, рассматривался ранее в [23] и [24], где скорость порядка $O(n^{-1/2})$ была установлена без каких-либо требований на первые два условных момента. Ниже, в рамках условий (1.1) и $\sigma_m = \bar{\sigma}_m$, мы увяжем скорость сходимости с порядком наибольшего конечного момента, точнее, выведем из (2.8) существование постоянной $\bar{C}$ такой, что для любого $q \geq 4$ и всех m

$$\tilde{\delta}_m \leq \bar{C} k^{1/2} \psi^3(q) m^{-1/2+1/4q}. \quad (2.11)$$

(Определение $\tilde{\delta}_m$ см. в п. 2.1.) В случае ограниченных слагаемых (2.11) влечет оценку

$$\bar{C} k^{1/2} \psi^3(\infty) m^{-1/2}.$$

Доказательство просто. Пусть $\tilde{S}_{m-k}$ то же, что и в п. 2.1, $\mathcal{F} = \mathcal{F}_{m+1}$, $\mathcal{F}' \supseteq \mathcal{F}_U$. Первое слагаемое в (2.8) равно нулю. Второй сомножитель в последнем члене (2.8) оценим грубо — единицей. После этого последний член становится больше, чем четвертый, так как этот член можно считать меньшим единицы. Вспомним также, что ввиду (1.4), $\psi(q) \geq C > 0$. Таким образом,

$$\begin{aligned} \|\Delta(S_m | \mathcal{F}_{m+1})\|_q &\leq C \left\{ \delta_m + \tilde{\delta}_{m-k} + (k\psi(q) m^{-1/2})^{1-1/q} \right\} \\ &\leq C \left\{ \delta_m + \tilde{\delta}_{m-k} + k\psi(q) m^{-1/2+1/2q} \right\}. \end{aligned} \quad (2.12)$$

Осталось применить индукцию. Так как $\psi \geq C$, (2.11) очевидно при $m \leq 2k$. Предположим, что (2.11) верно для некоторого $\bar{C} \geq 1$ и $m \leq n-1$. Ввиду (2.10) и (2.12) $r_m \leq C\bar{C}k\psi^6(q) m^{-1/2+1/2q}$ при $m \leq n$. (Выводя последнюю оценку, надо учесть, что при $m \leq 2k$ она тривиальна.) Отсюда и из (1.4) следует, что $r_n^* \leq C\bar{C}k\psi^6(q) n^{-1/2+1/2q}$. Подставляя последнее в (1.6), легко получить, что $\delta_n \leq C\bar{C}^{-1/2} k^{1/2} \psi^3(q) n^{-1/2+1/4q}$. Ясно, что та же оценка верна и для $\tilde{\delta}_n$, что влечет справедливость (2.11) при $m = n$.

2.3.2. Перемешивание. Пусть σ -алгебра $\mathcal{F}_m^k = \bigvee_{j=m}^k \mathcal{F}_j$ (т.е. $\mathcal{F}_m^k$ порождена $(X_m, \dots, X_k)$), и $\mathcal{F}^m = \mathcal{F}_1^m$. Возвращаясь к (2.8), положим $\tilde{S}_{m-k} = S_{m-k}$, $\mathcal{F} = \mathcal{F}_{m+1}^\infty$, $\mathcal{F}' = \mathcal{F}_{m-k+1}^m$. Обозначим $\|\Delta(S_m | \mathcal{F}_{m+1}^\infty)\|_q$ символом $\chi_m(q)$ и положим $d(k) = d(k; q) = \sup_m \|\Delta(S_{m-k} | \mathcal{F}_{m+1}^\infty)\|_q$. Отметим, что $d(k; q)$ — по существу, характеристика перемешивания, причем она может оказаться слабее обычных характеристик такого типа, так как в ней фигурируют не меры (условные и безусловные), определенные на всей σ -алгебре $\mathcal{F}^{m-k}$, а лишь распределения суммы S_{m-k} . (Описание различных типов перемешивания можно найти, например, в [4], обзор предельных теорем для последовательностей с перемешиванием — в [22].)

Из (2.8) следует, что

$$\begin{aligned} \chi_m(q) &\leq C \left\{ d(k; q) + \delta_m + \delta_{m-k} + k\psi(q) m^{-1/2} \right. \\ &\quad \left. + (k\psi(q) m^{-1/2})^{1-1/q} (\chi_{m-k}(q))^{1/q} \right\}. \end{aligned} \quad (2.13)$$

Чтобы найти оптимальную оценку, надо минимизировать правую часть (2.13) по k . Мы поступим проще. Прежде всего заметим, что теперь нам придется проводить индукцию, вычисляя δ_m и χ_m одновременно. В этом случае следует принимать во внимание оба сомножителя в последнем члене из (2.13) (ранее второй сомножитель заменялся просто

единицей). Далее, поскольку степени при указанных сомножителях в сумме равны единице, значение q , как можно убедиться, проводя конкретные вычисления, мало влияет на вид получаемой оценки. Поэтому положим $q = 4$, $\psi = \psi(4)$, $d(k) = d(k; 4)$.

Предположим далее, что существует $k = k(m)$ такое, что $d(k(m)) = m^{-1/2}$. Чтобы применить впоследствии (1.3), предположим также, что $k(m)$ — возрастающая регулярно меняющаяся функция. Чтобы осуществить индукцию, достаточно следовать тем же аргументам, что использовались ранее. При этом, применяя (2.9), надо учесть, что $\|\Delta(X | \mathcal{F})\|_q \leq \|\Delta(X | \mathcal{F}')\|_q$ для любых X и $\mathcal{F} \subseteq \mathcal{F}'$. Окончательно, это приведет к оценкам

$$\delta_n \leq \frac{C\psi^3 k(n)}{\sqrt{n}}, \quad \chi_n(4) \leq \frac{C\psi k(n)}{\sqrt{n}}. \quad (2.14)$$

Для k -зависимых с.в. $k(n) = k$ числители в (2.14) не зависят от n .

Отметим теперь, что условие конечности моментов четвертого или более высоких порядков может быть заменено на определенное требование к условным моментам с.в. X_m — обстоятельство, видимо, впервые замеченное в [2]. Для этого вместо неравенства (2.3) следует использовать неравенство (2.4) или (2.5). Предположим, например, что $\bar{\psi}(t) := \sup_m \max_{j < m} \|(\mathbf{E}\{|X_j|^t | X_m\})^{1/t}\|_\infty < \infty$ при некотором $t > 1$. Тогда, как нетрудно вывести непосредственно из (2.5),

$$\begin{aligned} \chi_m(\infty) \leq C \Big\{ d(k; \infty) + \delta_m + \delta_{m-k} + k\bar{\psi}(t) m^{-1/2} \\ + [k\bar{\psi}(t) m^{-1/2}]^{t/(t+1)} (\chi_{m-k}(\infty))^{1/(t+1)} \Big\}. \end{aligned} \quad (2.15)$$

Сравнивая последнее неравенство с (2.13) и отмечая, что степени при последних членах в обоих неравенствах не играют существенной роли: важно лишь, что в сумме они дают единицу, приходим к (2.14) с заменой $\psi = \psi(4)$ на $\bar{\psi}(t)$.

2.3.3. Зависимости, сходные с марковской. При дополнительных условиях член $k(n)$ в (2.14) можно уменьшить или даже убрать. Будем описывать структуру зависимости в терминах σ -алгебр $\mathcal{A}_m \subset \mathcal{F}_m^\infty$ таких, что для любых m и x

$$\mathbf{P}\{S_{m-1} < x | \mathcal{F}_m^\infty\} = \mathbf{P}\{S_{m-1} < x | \mathcal{A}_m\}. \quad (2.16)$$

Типичным примером может служить марковская цепь — в этом случае $\mathcal{A}_m = \mathcal{F}_m$, однако, на самом деле, мы можем рассматривать σ -алгебры, содержащие события, определяемые значениями всей последовательности $(X_m, X_{m+1}, \dots)$. Конечно, формально такая схема может быть сведена к марковской, но для такой «обобщенной» марковской цепи может оказаться трудным получить хорошую оценку скорости сходимости

(скажем, это будет цепь со значениями в пространстве высокой размерности). В вычислениях, проводимых ниже, конкретный вид $\mathcal{A}_m$ не играет роли, так что мы рассматриваем случай более общий, чем марковский, по существу.

Обозначим $G_{mk}(B | \omega)$ сужение условного распределения $P\{B | \mathcal{F}_{m+k}^\infty\}[\omega]$ на $\mathcal{A}_m$. Скажем, в случае марковской цепи и $\mathcal{A}_m = \mathcal{F}_m$ это соответствует распределению $P\{X_m \in dz | \mathcal{F}_{m+k}^\infty\} = P\{X_m \in dz | X_{m+k}\}$.

Пусть $L_{mk}(B | \omega', \omega'') = G_{mk}(B | \omega') - G_{mk}(B | \omega'')$, $\rho_{km}(\omega', \omega'') = \frac{1}{2} \int |L_{mk}(d\omega | \omega', \omega'')|$ и $\rho(k; q) = \sup_m [\int \int \rho_{km}^q(\omega', \omega'') P\{d\omega''\} P\{d\omega'\}]^{1/q}$.

В случае $\mathcal{A}_m = \mathcal{F}_m$

$$\rho(k; q) = \sup_m \left(\int \int \left[\frac{1}{2} \int |L_{mk}(dz; x, y)| \right]^q F_{m+k}(dx) F_{m+k}(dy) \right)^{1/q},$$

где $L_{mk}(dz; x, y) = P\{X_m \in dz | X_{m+k} = x\} - P\{X_m \in dz | X_{m+k} = y\}$ и F_m — распределение X_m .

Из леммы 3 и (2.16) при $V = S_{m-k}$, $\mathcal{F} = \mathcal{F}_{m+1}^\infty$ и $\mathcal{A} = \mathcal{A}_{m-k+1}$ следует, что

$$\|\Delta(S_{m-k} | \mathcal{F}_{m+1}^\infty)\|_q \leq 4 \|\Delta(S_{m-k} | \mathcal{F}_{m-k+1}^\infty)\|_q^s \cdot \left[\rho\left(k; \frac{q-1}{1-s}\right) \right]^{1-1/q},$$

где $s \leq 1$. Следовательно, вместо (2.13) мы имеем

$$\begin{aligned} \chi_m(q) \leq C \left\{ 4\chi_{m-k}^s(q) \left[\rho\left(k; \frac{q-1}{1-s}\right) \right]^{1-1/q} + \delta_m + \delta_{m-k} \right. \\ \left. + k\psi(q) m^{-1/2} + \left(k\psi(q) m^{-1/2} \right)^{1-1/q} (\chi_{m-k}(q))^{1/q} \right\}, \quad (2.17) \end{aligned}$$

и мы можем провести индукцию более аккуратно.

Ограничимся случаем $\rho(k; \infty) \rightarrow 0$ при $k \rightarrow \infty$. Это сильное требование; оно выполнено, например, если справедливо условие Деблина. В рассматриваемом случае можно положить в (2.17) $s = 1$ и $q = 4$. Вводя обозначения $\chi_m = \chi_m(4)$, $\psi = \psi(4)$, имеем

$$\begin{aligned} \chi_m \leq C \left\{ 4\chi_{m-k} [\rho(k; \infty)]^{3/4} + \delta_m + \delta_{m-k} \right. \\ \left. + k\psi m^{-1/2} + (k\psi m^{-1/2})^{3/4} (\chi_{m-k})^{1/4} \right\}. \end{aligned}$$

Пусть k_0 таково, что $\rho(k_0; \infty) \leq \frac{1}{8^{4/3}}$. Полагая $k = k_0$, получаем

$$\begin{aligned} \chi_m \leq C \left\{ \frac{1}{2} \chi_{m-k_0} + \delta_m + \delta_{m-k_0} + k_0 \psi m^{-1/2} \right. \\ \left. + \left(k_0 \psi m^{-1/2} \right)^{3/4} (\chi_{m-k_0})^{1/4} \right\}. \quad (2.18) \end{aligned}$$

Проводя вычисления, аналогичные тем, что уже проводились выше, легко проверить, что (2.18) влечет неравенство $\delta_n \leq C\psi^3 k_0/\sqrt{n}$.

В чисто марковском случае последняя оценка имеет тот же порядок, что и оценки в [21] и [3], однако мы требовали выше конечность четвертого момента (не говоря уже о мартингальном условии и требовании к условным дисперсиям). Ввиду этого уместно подчеркнуть еще раз, что для нас было важно не столько получить эти оценки сами по себе, сколько показать, что они следуют из общих оценок, причем специфика конкретного типа зависимости учитывается лишь в самом конце вычислений.

2.4. Доказательства лемм.

Доказательство леммы 2. Пусть $\Delta(y | \mathcal{F}) = |\mathbf{P}\{S < y | \mathcal{F}\} - G(y)|$, $G(y) = \mathbf{P}\{S < y\}$, $G_V(y) = \mathbf{P}\{V < y\}$ и h обозначает измеримую относительно $\mathcal{F}$ положительную с.в., вид которой будет уточнен позже. Имеем

$$\begin{aligned}\Delta(y | \mathcal{F}) &= \mathbf{E} \left\{ \left[I(\varepsilon U + \sqrt{1 - \varepsilon^2} V < y) - G(y) \right] I(|U| > h) \mid \mathcal{F} \right\} \\ &\quad + \mathbf{E} \left\{ \left[I(\varepsilon U + \sqrt{1 - \varepsilon^2} V < y) - G(y) \right] I(|U| \leq h) \mid \mathcal{F} \right\} \\ &= J_1 + J_2.\end{aligned}$$

Далее,

$$\begin{aligned}|J_1| &= \left| \mathbf{E} \left\{ I(|U| > h) \mathbf{E} \left\{ \left[I\left(\sqrt{1 - \varepsilon^2} V < y - \varepsilon U\right) - G(y) \right] \mid \mathcal{F}' \vee \mathcal{F} \right\} \mid \mathcal{F} \right\} \right| \\ &\leq \left| \mathbf{E} \left\{ I(|U| > h) \mathbf{E} \left\{ \left[I\left(\sqrt{1 - \varepsilon^2} V < y - \varepsilon U\right) - G_V\left(\frac{y - \varepsilon U}{\sqrt{1 - \varepsilon^2}}\right) \right] \mid \mathcal{F}' \vee \mathcal{F} \right\} \mid \mathcal{F} \right\} \right| \\ &\quad + \left| \mathbf{E} \left\{ I(|U| > h) \mathbf{E} \left\{ G_V\left(\frac{y - \varepsilon U}{\sqrt{1 - \varepsilon^2}}\right) - G(y) \mid \mathcal{F}' \vee \mathcal{F} \right\} \mid \mathcal{F} \right\} \right| \\ &= |J_{11}| + |J_{12}|.\end{aligned}$$

Ясно, что $|J_{11}| \leq \mathbf{E}\{I(|U| > h)\Delta(V | \mathcal{F}' \vee \mathcal{F}) | \mathcal{F}\} \leq \|\Delta(V | \mathcal{F}' \vee \mathcal{F})\|_{q|\mathcal{F}} [\mathbf{P}\{|U| > h | \mathcal{F}\}]^{1/p} \leq \|\Delta(V | \mathcal{F}' \vee \mathcal{F})\|_{q|\mathcal{F}} [\|U\|_{t|\mathcal{F}}/h]^{t/p}$, где $q^{-1} + p^{-1} = 1$, $t > 1$.

Далее,

$$\left| G_V \left(\frac{y - \varepsilon U}{\sqrt{1 - \varepsilon^2}} \right) - G(y) \right| \leq \delta(V) + \delta(S) + \left[\Phi \left(\frac{y - \varepsilon U}{\sqrt{1 - \varepsilon^2}} \right) - \Phi \left(\frac{y - \varepsilon U}{\sigma_S} \right) \right] + \left[\Phi \left(\frac{y - \varepsilon U}{\sigma_S} \right) - \Phi \left(\frac{y}{\sigma_S} \right) \right].$$

Заметим, что $\sigma_S^2 = 1 + 2r\varepsilon\sqrt{1 - \varepsilon^2}$. Отсюда легко следует, что выражение в первых квадратных скобках $[\cdot]$ не превосходит $\min\{(\frac{1}{2}), (2\pi e)^{-1/2}(\varepsilon^2 + 2|r|\varepsilon)/(2\min\{1 - 2|r|\varepsilon, 1 - \varepsilon^2\})\} \leq \varepsilon^2 + 2|r|\varepsilon$. Выражение во вторых квадратных скобках не превосходит $\min\{1, (2\pi)^{-1/2}\varepsilon U/[1 - 2|r|\varepsilon]\} \leq \varepsilon U + (\frac{5}{2})r\varepsilon$.

Итак, $|J_{12}| \leq \delta(V) + \delta(S) + \varepsilon^2 + 5|r|\varepsilon + \varepsilon \mathbf{E}\{|U| \mid \mathcal{F}\}$.

Обратившись к J_2 , получаем

$$\begin{aligned} J_2 &\leq \mathbf{E} \left\{ \left[I \left(\sqrt{1 - \varepsilon^2} V < y + \varepsilon h \right) - G(y) \right] I(|U| \leq h) \mid \mathcal{F} \right\} \\ &= \left[\mathbf{P} \left\{ \sqrt{1 - \varepsilon^2} V < y + \varepsilon h \mid \mathcal{F} \right\} - G(y) \right] \\ &\quad - \mathbf{E} \left\{ \left[I \left(\sqrt{1 - \varepsilon^2} V < y + \varepsilon h \right) - G(y) \right] I(|U| > h) \mid \mathcal{F} \right\} = J_{21} - J_{22}. \end{aligned}$$

Поступая как и выше, получаем

$$\begin{aligned} J_{21} &= \mathbf{P} \left\{ \sqrt{1 - \varepsilon^2} V < y + \varepsilon h \mid \mathcal{F} \right\} - G_V \left(\frac{y + \varepsilon h}{\sqrt{1 - \varepsilon^2}} \right) \\ &\quad + G_V \left(\frac{y + \varepsilon h}{\sqrt{1 - \varepsilon^2}} \right) - G(y) \leq \Delta(V \mid \mathcal{F}) + \delta(V) + \delta(S) \\ &\quad + \Phi \left(\frac{y + \varepsilon h}{\sqrt{1 - \varepsilon^2}} \right) - \Phi \left(\frac{y}{\sqrt{1 - \varepsilon^2}} \right) + \Phi \left(\frac{y}{\sqrt{1 - \varepsilon^2}} \right) - \Phi \left(\frac{y}{\sigma_S} \right) \\ &\leq \Delta(V \mid \mathcal{F}) + \delta(V) + \delta(S) + \varepsilon h + 2\varepsilon^2 + \varepsilon^2 + 2|r|\varepsilon. \end{aligned}$$

Нижняя оценка для J_{21} получается аналогично с заменой h на $-h$. Легко видеть также, что J_{22} может быть рассмотрено в точности так же, как и J_1 , причем мы придем к той же оценке. Окончательно, собирая все оценки, имеем

$$\begin{aligned} \Delta(S \mid \mathcal{F}) &\leq C \left\{ \Delta(V \mid \mathcal{F}) + \delta(V) + \delta(S) + \varepsilon^2 + r\varepsilon + \varepsilon \mathbf{E}\{|U| \mid \mathcal{F}\} + \varepsilon h \right. \\ &\quad \left. + \|\Delta(V \mid \mathcal{F}' \vee \mathcal{F})\|_{q|\mathcal{F}} \left[\frac{\|U\|_{t|\mathcal{F}}}{h} \right]^{t/p} \right\}. \end{aligned}$$

Минимизируя правую часть по h (h должно быть измеримо относительно $\mathcal{F}$), нетрудно получить, что

$$\Delta(S | \mathcal{F}) \leq C \left\{ \Delta(V | \mathcal{F}) + \delta(V) + \delta(S) + \varepsilon^2 + \tau\varepsilon + \varepsilon \mathbf{E}\{|U| | \mathcal{F}\} + [\varepsilon \|U\|_{t|\mathcal{F}}]^{t/(t+p)} (\|\Delta(V | \mathcal{F}' \vee \mathcal{F})\|_{q|\mathcal{F}})^{p/(t+p)} \right\}.$$

Вычисляя теперь «неслучайные» нормы $\|\cdot\|_q$ обеих частей последнего неравенства и применяя неравенство Гёльдера, приходим к (2.4). Лемма 2 доказана.

Доказательство леммы 3. В силу (2.6)

$$\begin{aligned} \mathbf{P}\{V < y | \mathcal{F}\}[\omega'] &= \int \mathbf{P}\{V < y | \mathcal{A} \vee \mathcal{F}\}[\omega] \mathbf{P}\{d\omega | \omega'\} \\ &= \int \mathbf{P}\{V < y | \mathcal{A}\}[\omega] G_{\mathcal{A}|\mathcal{F}}(d\omega | \omega'). \end{aligned}$$

Пусть

$$\Delta(y; V | \mathcal{F})[\omega'] = \mathbf{P}\{V < y | \mathcal{F}\}[\omega'] - \mathbf{P}\{V < y\}$$

и $\mathbf{P}_{\mathcal{F}}$ — сужение $\mathbf{P}$ на $\mathcal{F}$. Легко видеть, что

$$\begin{aligned} \Delta(y; V | \mathcal{F})[\omega'] &= \int \left(\int \mathbf{P}\{V < y | \mathcal{A}\}[\omega] L_{\mathcal{A}|\mathcal{F}}(d\omega | \omega', \omega'') \right) \mathbf{P}_{\mathcal{F}}(d\omega'') \\ &= \int \left(\int [\mathbf{P}\{V < y | \mathcal{A}\}[\omega] - G_V(y)] L_{\mathcal{A}|\mathcal{F}}(d\omega | \omega', \omega'') \right) \mathbf{P}_{\mathcal{F}}(d\omega''), \end{aligned}$$

так как интеграл от $G_V(y)$ равен нулю. Итак,

$$\begin{aligned} \Delta(V | \mathcal{F}) &:= \sup_y |\Delta(y; V | \mathcal{F})| \\ &\leq \iint \Delta(V | \mathcal{A})[\omega] |L_{\mathcal{A}|\mathcal{F}}(d\omega | \omega', \omega'')| \mathbf{P}_{\mathcal{F}}(d\omega'') \\ &\leq \left(\iint \Delta^q(V | \mathcal{A})[\omega] |L_{\mathcal{A}|\mathcal{F}}(d\omega | \omega', \omega'')| \mathbf{P}_{\mathcal{F}}(d\omega'') \right)^{1/q} \\ &\quad \times \left(\iint |L_{\mathcal{A}|\mathcal{F}}(d\omega | \omega', \omega'')| \mathbf{P}_{\mathcal{F}}(d\omega'') \right)^{1/p}, \end{aligned}$$

где $p^{-1} + q^{-1} = 1$ и $|L(d\omega|\cdot)|$ обозначает вариацию.

Заменяя в первом сомножителе $|L_{\mathcal{A}|\mathcal{F}}(d\omega|\omega', \omega'')|$ на $G_{\mathcal{A}|\mathcal{F}}(d\omega|\omega') + G_{\mathcal{A}|\mathcal{F}}(d\omega|\omega'')$, мы видим, что этот сомножитель не превосходит $(\|\Delta(V | \mathcal{A})\|_{q|\mathcal{F}}^q + \|\Delta(V | \mathcal{A})\|_q^q)^{1/q}$. Используя эту оценку, тот факт, что $|\Delta(V | \mathcal{A})| \leq 1$ и $\int |L_{\mathcal{A}|\mathcal{F}}(d\omega|\cdot)| \leq 2$, и неравенство Гёльдера, получаем,

что для любых $s \in (0, 1)$ и $q \geq 2$ верно

$$\begin{aligned}
 \|\Delta(V | \mathcal{F})\|_q &\leq \left[\int \left(\|\Delta(V | \mathcal{A})\|_{q|\mathcal{F}}^q[\omega'] + \|\Delta(V | \mathcal{A})\|_q^q \right) \right. \\
 &\quad \times \left. \left(\iint |L_{\mathcal{A}|\mathcal{F}}(d\omega | \omega', \omega'')| \mathbf{P}_{\mathcal{F}}(d\omega'') \right)^{q-1} \mathbf{P}_{\mathcal{F}}(d\omega') \right]^{1/q} \\
 &\leq \left[\int \left(\|\Delta(V | \mathcal{A})\|_{q|\mathcal{F}}^q[\omega'] + \|\Delta(V | \mathcal{A})\|_q^q \right)^{1/s} \mathbf{P}_{\mathcal{F}}(d\omega') \right]^{s/q} \\
 &\quad \times \left[\int \left(\iint |L_{\mathcal{A}|\mathcal{F}}(d\omega | \omega', \omega'')| \mathbf{P}_{\mathcal{F}}(d\omega'') \right)^{(q-1)/(1-s)} \mathbf{P}_{\mathcal{F}}(d\omega') \right]^{(1-s)/q} \\
 &\leq 2^{1/q} \left(\mathbf{E} \left(\|\Delta(V | \mathcal{A})\|_{q|\mathcal{F}}^{q/s} + \|\Delta(V | \mathcal{A})\|_q^{q/s} \right) \right)^{s/q} \\
 &\quad \times 2^{(q-1)/q} \left(\iint \left(\frac{1}{2} \int |L_{\mathcal{A}|\mathcal{F}}(d\omega | \omega', \omega'')| \right)^{(q-1)/(1-s)} \right. \\
 &\quad \times \left. \mathbf{P}_{\mathcal{F}}(d\omega'') \mathbf{P}_{\mathcal{F}}(d\omega') \right)^{(1-s)/q} \\
 &\leq 2 \left(2 \|\Delta(V | \mathcal{A})\|_q^q \right)^{s/q} (\|\rho_{\mathcal{A}|\mathcal{F}}\|_{(q-1)/(1-s)})^{(q-1)/q} \\
 &\leq 4 \|\Delta(V | \mathcal{A})\|_q^s \cdot (\|\rho_{\mathcal{A}|\mathcal{F}}\|_{(q-1)/(1-s)})^{(q-1)/q}.
 \end{aligned}$$

Лемма 3 доказана.

СПИСОК ЛИТЕРАТУРЫ

1. *Bhattacharya R. N., Ranga Rao R.* Normal Approximation and Asymptotic Expansion. Melbourne: Krieger, 1986.
2. *Bolthausen E.* Exact convergence rates in some martingale central limit theorems. — Ann. Probab., 1982, v. 10, p. 672–688.
3. *Bolthausen E.* The Berry–Esseen theorem for strongly mixing Harris recurrent Markov chains. — Z. Wahrscheinlichkeitstheor. verw. Geb., 1982, B. 60, H. 3, S. 283–289.
4. *Bradley R.* Basic properties of strong mixing conditions. — Dependence in Probability and Statistics. Ed. by E. Eberlein, M. Taqqu. Boston: Birkhäuser. Progr. Probab. Statist., 1986, v. 11, p. 165–192.
5. *Dvoretzky A.* Asymptotic normality for sums of dependent random variables. — Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, v. 3. Berkeley: Univ. of California Press, 1972, p. 513–535.
6. *Erickson R. V., Quine M. P., Weber N. C.* Explicit bounds for the departure from normality of sums of dependent random variables. — Acta Math. Acad. Sci. Hungar., 1979, v. 34, p. 27–32.

7. Евстигнеев И. В. Регулярные условные математические ожидания случайных величин, зависящих от параметров. — Теория вероятн. и ее примен., 1986, т. XXXI, в. 3, с. 586–589.
8. Грамз И. Г. Нормальная аппроксимация для семимартингалов. — Успехи матем. наук, 1987, т. 42, № 6, с. 189–190.
9. Haeusler E. On the rate of convergence in the central limit theorem for martingales. — Ann. Probab., 1988, v. 16, p. 275–299.
10. Haeusler E., Joos K. A nonuniform bound on the rate of convergence in the central limit theorem for martingales. — Ann. Probab., 1989, v. 16, p. 1690–1720.
11. Hall P., Heyde C. C. Martingale Limit Theory and its Applications. New York: Academic Press, 1980, 308 p.
12. Hall P., Heyde C. C. Rates of convergence in martingale limit theorem. — Ann. Probab., 1981, v. 9, № 2, p. 395–404.
13. Heyde C. C., Brown B. M. On the departure from normality of a certain class of martingales. — Ann. Math. Statist., 1970, v. 41, p. 2161–2165.
14. Жакоб Ж., Ширяев А. Н. Предельные теоремы для случайных процессов, т. 2. М.: Физматлит, 1994.
15. Kato Y. Rates of convergence in the central limit theorem for martingale differences. — Bull. Math. Statist., 1979, p. 1–8.
16. Кирьянова Л. В., Ротарь В. И. Оценка скорости сходимости в центральной предельной теореме для мартингалов. — Теория вероятн. и ее примен., 1991, т. 36, в. 2, с. 301–312.
17. Кубилюс К. Скорость сходимости в функциональной центральной предельной теореме. — Лит. матем. сб., 1989, т. XXV, в. 1, с. 59–69.
18. Липцер Р. Ш., Ширяев А. Н. О скорости сходимости в центральной предельной теореме для мартингалов. — Теория вероятн. и ее примен., 1982, т. XXVII, в. 1, с. 3–14.
19. Липцер Р. Ш., Ширяев А. Н. Теория мартингалов. М.: Наука, 1986, 512 с.
20. Mori T. On the rate of convergence in the martingale limit theory. — Studia Sci. Math. Hungar., 1977, v. 12, № 3, p. 413–417.
21. Нагаев С. В. Уточнение предельных теорем для однородных цепей Маркова. — Теория вероятн. и ее примен., 1961, т. VI, в. 1, с. 67–86.
22. Peligrad M. Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables (a survey). — Dependence in Probability and Statistics. Ed. by E. Eberlein, M. Taqqu. Boston: Birkhäuser. Progr. Probab. Statist., 1986, v. 11, p. 193–223.
23. Rinott Y. On normal approximation rates for certain sums of dependent random variables. — J. Comput. Appl. Math., 1994, v. 55, p. 135–143.
24. Rinott Y., Rotar V. A multivariate CLT for local dependence with $n^{-1/2} \log n$ rate, and applications to multivariate graph related statistics. — J. Multivariate Anal., 1996, v. 56, p. 333–350.
25. Ротарь В. И. Теория вероятностей. М.: Высшая школа, 1992, 368 с.
26. Sazonov V. V. Normal approximation — some recent advances. — Lecture Notes in Math., 1981, v. 879, p. 1–105.
27. Stein C. Approximate Computation of Expectations. Hayward, CA: IMS, 1986, 164 p.

Поступила в редакцию
12.VIII.1997