ЧАСТЬ II

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ НЕКОТОРЫХ ТИПОВ СЛУЧАЙНЫХ ПРОЦЕССОВ

О СКОРОСТИ СХОДИМОСТИ В ЦЕНТРАЛЬНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМЕ ДЛЯ ЭМПИРИЧЕСКИХ МЕР

И. С. БОРИСОВ

1. Введение и формулировка основных результатов

Пусть $\{X_k\}$ — последовательность независимых одинаково распределенных случайных величин (с. в.) со значениями в произвольном измеримом пространстве $(\mathfrak{X}, \mathscr{A})$ и распределением $P(\cdot)$ на $\mathscr{A}$. Обозначим через $P_n(\cdot)$ эмпирическое распределение, построенное по выборке $X_1, X_2, \ldots, X_n$:

$$P_n(A) = n^{-1} \sum_{i=1}^{n} I_{X_i}(A),$$

где $I_{X_i}(A) = 1$, если $X_i \in A$, и $I_{X_i}(A) = 0$ в противном случае.

Рассмотрим последовательность так называемых эмпирических мер:

$$Q_n(\cdot) = n^{1/2}(P_n(\cdot) - P(\cdot)), \quad n = 1, 2, \ldots, \tag{1}$$

которые нам будет удобно интерпретировать как случайные процессы, заданные на параметрическом множестве $\mathscr{A}$. В силу многомерной центральной предельной теоремы можно утверждать, что конечномерные распределения $Q_n(\cdot)$ при $n \to \infty$ слабо сходятся к соответствующим конечномерным распределениям гауссовского случайного процесса $G(A)$, $A \in \mathscr{A}$, с нулевым средним и ковариацией

$$\mathbf{E}G(A)G(B) = P(A \cap B) - P(A)P(B). \tag{2}$$

Значительный интерес представляет более сильный вид сходимости Q_n к G, из которого следует слабая сходимость распределений для достаточно широкого класса функционалов от указанных процессов.

Пусть $\mathscr{B}$ — произвольный подкласс множеств из $\mathscr{A}$. Будем говорить, что для последовательности $\{Q_n\}$ имеет место центральная предельная теорема равномерно по $\mathscr{B}$, если на некотором вероятностном пространстве можно построить случайные процессы $\{\tilde{Q}_n\}$, а также гауссовский процесс G, для которых

$$\tilde{Q}_n \overset{\mathscr{D}}{=} Q_n, \quad n = 1, 2, \ldots, \tag{3}$$

и, кроме того, при $n \to \infty$

$$\sup_{A \in \mathscr{B}} |\tilde{Q}_n(A) - G(A)| \to 0 \tag{4}$$

с вероятностью 1, где символ „$\overset{\mathscr{D}}{=}$“ в (3) и далее обозначает равенство конечномерных распределений.

Отметим, что в данной формулировке не требуется, чтобы супремум в (4) был измеримым отображением, поскольку сходимость с вероятностью 1 может быть определена для последовательностей произвольных функций от элементарного исхода.

Замечание. Аналогичное определение приведено в [1]. Однако запись „$\overset{\mathscr{D}}{=}$“ там означает совпадение распределений на σ-алгебре, порожденной шарами в равномерной топологии некоторого функционального пространства. При этом, разумеется, необходимо требовать, чтобы процессы Q_n были измеримыми относительно указанной σ-алгебры. Отметим также, что при достаточно широких предположениях относительно $\mathscr{B}$ и P эти два определения эквивалентны (см. также замечание к теореме 2).

Обозначим

$$d(A, B) = P(A \setminus B) + P(B \setminus A).$$

Пару $(\mathscr{A}, d)$ можно рассматривать как метрическое пространство с псевдометрикой $d(\cdot)$. Всюду в дальнейшем предполагается, что $\mathscr{B}$ — вполне ограниченное множество в $(\mathscr{A}, d)$.

Обозначим через $N_1(\varepsilon) = N_1(\varepsilon, \mathscr{B}, P)$ минимальное число элементов так называемой ε-сети с вложением (см. [2]) для $\mathscr{B}$, т. е. $N_1(\varepsilon)$ — наименьшее натуральное число, для которого найдутся такие множества A_1, $A_2, \ldots, A_{N_1(\varepsilon)} \in \mathscr{A}$, что для любого $B \in \mathscr{B}$ можно подобрать пару множеств A_{i_1} и A_{i_2}, $i_1, i_2 \leqslant N_1(\varepsilon)$, со следующими свойствами:

$$A_{i_1} \subseteq B \subseteq A_{i_2}, \quad d\left(A_{i_1}, A_{i_2}\right) \leqslant \varepsilon.$$

Величину $H_1(\varepsilon) = \log N_1(\varepsilon)$ назовем, следуя Р. Дадли (см. [2]), метрической энтропией с вложением.

В [2] показано, что для выполнения центральной предельной теоремы равномерно по $\mathscr{B}$ достаточно, чтобы

$$\sum_{k=1}^{\infty} [2^{-k} H_1(2^{-k})]^{1/2} < \infty. \tag{5}$$

Там же установлено, что условие (5), вообще говоря, не является необходимым. Однако в [3] приведен пример, в котором соотношение (5) все же необходимо для выполнения центральной предельной теоремы равномерно по $\mathscr{B}$. Иными словами, условие (5) в известном смысле неулучшаемо.

Цель настоящей работы — получить при выполнении условия (5) оценки скорости сходимости в (4).

Теорема 1. *Пусть выполнено условие* (5). *Тогда на некотором вероятностном пространстве можно построить d-сепарабельные случайные процессы $\tilde{Q}_n(\cdot)$, $n = 1, 2, \ldots$, и $G(\cdot)$, для которых при любых $x \geqslant 0$, $r \geqslant 0$, $M = 0, 1, \ldots$ и $y \geqslant 2r + 5$ справедливо неравенство*

$$\mathbf{P}\Big\{\sup_{A \in \mathscr{B}} |\tilde{Q}_n(A) - G(A)| \geqslant C n^{-1/2} \gamma^*(2^{-M}, \mathscr{B}, P)(\log n + x) +$$
$$+ (4y+1)\Big(\sum_{k>M} [2^{-k} H_1(2^{-k})]^{1/2} + 5 \cdot 2^{-M/2}\Big)\Big\} \leqslant K e^{-\lambda x} + N_1(2^{-M})^{1-r} \times$$
$$\times [R_1(y, r) + I(n, M, H_1) R_2(y, r)] + N_1(2^{-M})^3 (1 - I(n, M, H_1)) \times$$
$$\times \exp\Big\{- y n^{1/2} \sum_{k>M} [2^{-k} H_1(2^{-k})]^{1/2}\Big\}, \tag{6}$$

где $I(n, M, H_1) = \begin{cases} 1, \text{ если } \max(1, H_1(2^{-M-1})) < n 2^{-M+1}, \\ 0 \text{ в противном случае}; \end{cases}$

$$R_1(y, r) = 3\left[\exp\left\{\frac{(y-r-2)^2 - (r+3)^2 + 1}{2(1+y)}\right\} - 1\right]^{-1};$$

$$R_2(y, r) = (2\pi)^{-1/2} \int_y^{\infty} [\exp\{(t^2 - 5)/2 - r\} - 1]^{-1} dt;$$

C, K, λ — абсолютные положительные постоянные; $\gamma^*(\cdot)$ — некоторый функционал (подробнее см. раздел 2), удовлетворяющий неравенству

$$\gamma^*(\varepsilon, \mathscr{B}, P) \leqslant N_1(\varepsilon). \tag{7}$$

Замечание 1. Случайные процессы, сепарабельные относительно псевдометрики $d(\cdot)$, обладают всеми свойствами обычных сепарабельных (относительно метрик) процессов (см. [4]). В частности, супремум в (6) будет случайной величиной.

Замечание 2. Соотношение (6) останется в силе, если функцию $N_1(\cdot)$ заменить на любую ее мажоранту, для которой имеет место (5). Такое видоизменение часто бывает более удобным для получения оценок скорости сходимости в (4).

Сформулированная теорема допускает следующее обобщение.

Теорема 2. *Пусть выполнено условие* (5). *Тогда на некотором вероятностном пространстве можно построить последовательность серий с. в.* $X_{1n}, X_{2n}, \ldots, X_{nn}$, $n=1, 2, \ldots$, *совпадающих по распределению при каждом* n *с выборкой* $X_1, X_2, \ldots, X_n$, *а также* d*-сепарабельный гауссовский процесс* $G(\cdot)$, *для которых соотношение* (6) *будет иметь место при замене супремума на некоторую с. в.* Λ_n, *удовлетворяющую с вероятностью* 1 *неравенству*

$$\sup_{A\in\mathscr{B}} |\tilde{Q}_n(A)-G(A)| \leqslant \Lambda_n,$$

где случайный процесс $\tilde{Q}_n(\cdot)$ *определен в* (1) *для выборки* $X_{1n}, \ldots, X_{nn}$.

Замечание. Процессы $Q_n(\cdot)$ и $\tilde{Q}_n(\cdot)$, определенные в (1), могут уже не быть d-сепарабельными. Поэтому супремум в (4) может не быть случайной величиной.

Если $\mathscr{B}$ является «эмпирически измеримым для P» (см. [1]), то запись „$\overset{\mathscr{D}}{=}$" в определении центральной предельной теоремы можно понимать так же, как и в [1], т. е. как совпадение вероятностных распределений $Q_n(\cdot)$ и $\tilde{Q}_n(\cdot)$ на σ-алгебре, порожденной шарами в равномерной топологии некоторого функционального пространства. Однако, как показывает теорема 2, для доказательства центральной предельной теоремы при выполнении условия (5) нет никакой необходимости в привлечении тех или иных условий измеримости. При этом класс функционалов от $Q_n(\cdot)$, сходящихся по распределению к соответствующим функционалам от $G(\cdot)$, становится шире.

Положим в (6) $r=1$, $x=2\lambda^{-1}\log n$, $y=3\log n$. Тогда правая часть (6), очевидно, будет иметь порядок малости $O(n^{-2})$. Пусть, кроме того,

$$M\equiv M(n)=\max\left\{M\geqslant 0\colon \sum_{k>M}[2^{-k}H_1(2^{-k})]^{1/2}\geqslant N_1(2^{-M})\,n^{-1/2}\right\}.$$

Из леммы Бореля-Кантелли следует, что на соответствующем вероятностном пространстве при $n\to\infty$

$$\sup_{A\in\mathscr{B}}|\tilde{Q}_n(A)-G(A)|=O\left(\log n\sum_{k>M(n)}[2^{-k}H_1(2^{-k})]^{1/2}\right) \tag{8}$$

с вероятностью 1, причем постоянная в определении символа $O(\cdot)$ абсолютна.

Приведем еще одну оценку, которая в ряде случаев точнее (8). Положим в (6) $x=2\lambda^{-1}\log n$, $y=6+2r$, $r=1+2\log n/\log N_1(2^{-M})$, а в качестве $M=\tilde{M}(n)$ возьмем наименьшее целое неотрицательное число, минимизирующее на множестве $\{M: \max(1,\ H_1(2^{-M-1}))<n2^{1-M}\}$ выражение

$$\Delta_n(M)=\max\left\{\left(1+\frac{\log n}{H_1(2^{-M})}\right)\sum_{k>M}[2^{-k}H_1(2^{-k})]^{1/2},\ \frac{\gamma^*(2^{-M},\mathscr{B},P)\log n}{\sqrt{n}}\right\}.$$

Нетрудно видеть, что в этом случае из (6) следует неравенство

$$\mathbf{P}\left\{\sup_{A\in\mathscr{B}}|\tilde{Q}_n(A)-G(A)|>C_0\Delta_n(\tilde{M}(n))\right\}\leqslant Cn^{-2}, \tag{9}$$

где C_0, C — абсолютные положительные постоянные. Отсюда так же, как и в (8), можно получить оценку скорости сходимости в (4).

В качестве примера рассмотрим случай, когда $\mathfrak{X}=R^m$, $\mathscr{A}$ — σ-алгебра борелевских подмножеств в R^m, $\mathscr{B}=\{\{z\in R^m\colon z<t\};\ t\in R^m\}$ — мно-

жество открытых «углов» в R^m (соотношение $z < t$ понимается в смысле обычного частичного порядка в R^m). Иными словами, в этом примере речь идет об аппроксимации эмпирических функций распределения в R^m. В [5] показано, что для любого распределения P в R^m

$$H_1(2^{-k}) \leqslant \widehat{H}_1(2^{-k}) \equiv mk \log 2, \quad \gamma^*(2^{-M}, \mathscr{B}, P) \leqslant 2^{(m-1)M}. \tag{10}$$

Принимая во внимание замечание 2 к теореме 1, можно заменить в выражении для $\Delta_n(M)$ функцию $H_1(\cdot)$ на $\widehat{H}_1(\cdot)$.

Положим

$$M \equiv \widehat{M}(n) = [(2m-1)^{-1}(\log_2 n - \log_2 \log_2 n)] - 1,$$

где $[\cdot]$ — целая часть числа. Если $n \geqslant 2^{3(2m-1)}$, то $\widehat{M}(n) > 0$. Кроме того, легко видеть, что $m(M+1) < n2^{1-M}$. Поэтому

$$\inf_{\substack{M=0,1,\ldots \\ \max(1,\widehat{H}_1(2^{-M-1}))<n2^{1-M}}} \Delta_n(M) \leqslant \Delta_n(\widehat{M}(n)).$$

Далее, из (10) следует, что при $n \geqslant 2^{3(2m-1)}$

$$\widehat{H}^{-1}(2^{-M}) \sum_{k>M} [2^{-k}\widehat{H}_1(2^{-k})]^{1/2} \leqslant C_1 (mM2^M)^{-1/2} \leqslant C_2 n^{-\frac{1}{2(2m-1)}},$$

$$n^{-1/2}\gamma^*(2^{-M}, \mathscr{B}, P) \leqslant C_3 n^{-\frac{1}{2(2m-1)}},$$

где C_1, C_2, C_3 — абсолютные положительные постоянные. Из полученных неравенств и из (9) окончательно получаем, что при $n \geqslant 2^{3(2m-1)}$

$$\mathbf{P}\left(\sup_{A \in \mathscr{B}} |\widetilde{Q}_n(A) - G(A)| > C_0 \max(C_2, C_3)\, n^{-\frac{1}{2(2m-1)}} \log n\right) \leqslant Cn^{-2}. \tag{11}$$

Соотношение (11) усиливает один из результатов в [6], а также результат автора [5] (в смысле явной зависимости постоянных от размерности R^m).

В качестве следствия приведенной теоремы можно получить степенные оценки скорости сходимости в (4), по существу, для любого класса $\mathscr{B}$, который вкладывается в конечномерное евклидово пространство (например, для множества всех шаров или эллипсоидов в R^m и т. д.).

Рассмотрим еще один пример, с которым, в частности, будут связаны рассуждения раздела 3 настоящей статьи, касающиеся неулучшаемости полученных оценок скорости сходимости в (4). Пусть $P = \{p_i\}$ — произвольное дискретное распределение в $(\mathfrak{X}, \mathscr{A})$. Можно считать, что $\mathfrak{X} = \{a_1, a_2, \ldots\}$, где a_i — атомы распределения P, и $\mathscr{A}$ — всевозможные подмножества $\mathfrak{X}$. Положим $\mathscr{B} = \mathscr{A}$ и обозначим $\varepsilon_m = \sum_{i>m} p_i$. Пусть атомы $\{a_i\}$ занумерованы так, что последовательность $\{p_i\}$ монотонно убывает. Тогда легко видеть, что совокупность подмножеств $\left(a_0 \overset{\text{def}}{=} \varnothing\right)$

$$\{\{a_{i_1}, \ldots, a_{i_m}\}, \{a_{i_1}, \ldots, a_{i_m}\} \cup \{a_i;\ \ i > m\};\ \ 0 \leqslant i_1 \leqslant \ldots \leqslant i_m \leqslant m\}$$

образует ε_m-сеть с вложением для $\mathscr{A}$, т. е. $H_1(\varepsilon_m) \leqslant \widehat{H}_1(\varepsilon_m) \equiv m+1$. В [3] показано, что в нашем случае существует абсолютная постоянная c_0, для которой при любом натуральном m

$$\int_0^{\sqrt{\varepsilon_m}} H_1^{1/2}(x^2)\, dx \leqslant c_0 \sum_{i>m} \sqrt{p_i}.$$

Заметим, что именно из этого соотношения следует необходимость условия (5) для выполнения центральной предельной теоремы равномерно по $\mathscr{B} = \mathscr{A}$ в указанном классе распределений, поскольку сходимость ряда $\Sigma\sqrt{p_i}$ является в данном случае необходимым и достаточным условием для выполнения (4) (см. [3]).

Далее, принимая во внимание замечание, сделанное после доказательства леммы 3 в разделе 2 настоящей статьи, заключаем, что $\gamma^*(\varepsilon_m, \mathscr{A}, P) \leqslant m$. Обозначим $M(m) = \lfloor -\log_2 \varepsilon_m \rfloor - 1$. Будем считать, что $\varepsilon_m \leqslant 1/2$. Поскольку $\varepsilon_m \leqslant 2^{-M(m)-1}$, очевидно,

$$\Delta_n(\widetilde{M}(n)) \leqslant \widetilde{C} \min_{m:m+1\leqslant 4n\varepsilon_m} \widehat{\Delta}_n(m) \equiv$$

$$\equiv \widetilde{C} \min_{m:m+1\leqslant 4n\varepsilon_m} \max\left\{\left(1 + m^{-1}\log n\right) \sum_{i>m} \sqrt{p_i},\ mn^{-1/2}\log n\right\},$$

где $\widetilde{C}$ — некоторая абсолютная постоянная. Обозначая через $m_0(n)$ целое неотрицательное число, минимизирующее на множестве $\{m \geqslant 1 : m+1 \leqslant 4n\varepsilon_m\}$ функцию $\widehat{\Delta}_n(m)$, получаем из (9)

$$\mathbf{P}\left(\sup_{A\in\mathscr{A}} \left|\widetilde{Q}_n(A) - G(A)\right| > \widetilde{C}C_0\widehat{\Delta}_n(m_0(n))\right) \leqslant Cn^{-2}.$$

2. Доказательство теорем

Схема доказательства состоит в следующем. Сначала в $(\mathscr{A}, d)$ будет выделена конечная 2^{-M}-сеть $\{A_i\}$ с вложением для $\mathscr{B}$. Затем на одном вероятностном пространстве мы построим совокупности с. в. $\{Q_n(A_i)\}$ и $\{G(A_i)\}$. При этом задача построения сведется к аналогичной проблеме для совокупностей с. в. $\{\overline{Q}_n(A_i)\}$ и $\{\overline{G}(A_i)\}$, где $\overline{Q}_n$ и $\overline{G}$ определены в (1) и (2) для некоторого дискретного распределения $\overline{P}$ (леммы 2, 3). В свою очередь, для $\overline{Q}_n$ и $\overline{G}$ имеют место специальные представления (лемма 1), позволяющие достаточно просто конструировать на одном вероятностном пространстве случайные процессы $\overline{Q}_n$ и $\overline{G}$ с достаточно близкими траекториями. После того, как совокупности с. в. $\{Q_n(A_i)\}$ и $\{G(A_i)\}$ будут заданы на одном вероятностном пространстве, можно на расширенном вероятностном пространстве доопределить тем или иным способом значения процессов $Q_n(A)$ и $G(A)$ и в других точках $A \in \mathscr{A}$. Для завершения доказательства необходимо будет оценить выбросы приращений случайных процессов Q_n и G в окрестностях узлов 2^{-M}-сети (леммы 4, 5).

Итак, приступим к доказательству основных вспомогательных утверждений.

Лемма 1. *Пусть* $\overline{P} = \{p_k;\ k = 1, 2, \ldots, m\}$ — *дискретное распределение в* $(\mathfrak{X}, \mathscr{A})$. *Тогда имеют место соотношения*

$$\overline{Q}_n(\cdot) \overset{\mathscr{D}}{=} \sum_{i=1}^{m} I_{a_i}(\cdot)\left[V_n(F(i+1)) - V_n(F(i))\right], \tag{12}$$

$$\overline{G}(\cdot) \overset{\mathscr{D}}{=} \sum_{i=1}^{m} I_{a_i}(\cdot)\left[\overset{\circ}{w}(F(i+1)) - \overset{\circ}{w}(F(i))\right], \tag{13}$$

где $\{a_i;\ i = 1, 2, \ldots, m\}$ — *атомы распределения* $\overline{P}$; $V_n(\cdot)$ — *эмпирический процесс на* $[0, 1]$, *построенный по выборке объема* n *из равномерного на* $[0, 1]$ *распределения*; $\overset{\circ}{w}(\cdot)$ — *«броуновский мост» на* $[0, 1]$, $F(t) = \sum_{i<t} p_i$; *процессы* $\overline{Q}_n$ *и* $\overline{G}$ *определены в* (1) *и* (2) *для распределения* $\overline{P}$.

Доказательство этой леммы по существу содержится в [5], и поэтому мы его опускаем.

Далее, зададим случайные процессы $V_n(\cdot)$ и $\overset{\circ}{w}(\cdot)$ на одном вероятностном пространстве методом работы [7]. Тогда из (12) и (13) следует, что для любого m-точечного распределения в $(\mathfrak{X}, \mathscr{A})$ равномерно по любому подклассу $\mathscr{B} \subseteq \mathscr{A}$ при $n \to \infty$

$$|\overline{Q}_n(\cdot) - \overline{G}(\cdot)| = O(mn^{-1/2}\log n) \tag{14}$$

с вероятностью 1. Здесь постоянная, входящая в определение $O(\cdot)$, абсолютна. Нетрудно видеть, что множитель m в (14) может быть существенно уменьшен в зависимости от свойств класса $\mathscr{B}$ и порядка нумерации атомов $\{a_i\}$.

В дальнейшем множество вида $\{a_i;\ k\leqslant i\leqslant l\}$ будем называть цепочкой атомов. Ясно, что для любого фиксированного $A\in\mathscr{B}$ близость $\widetilde{Q}_n(A)$ и $G(A)$, определенных на одном вероятностном пространстве указанным методом, зависит не от числа атомов m, а от числа цепочек атомов (обозначим его $\gamma(A)$) дискретного распределения, содержащихся в A. Очевидно, что $\gamma(A)$ существенно зависит от способа нумерации атомов распределения. В связи с этим замечанием введем в рассмотрение величину

$$\gamma_0(\mathscr{B},\overline{P})=\min\max_{A\in\mathscr{B}}\gamma(A),$$

где минимум берется по всевозможным нумерациям атомов $\{a_i\}$. Теперь оценку (14) можно уточнить, заменив в (14) множитель m на величину $\gamma_0(\cdot)$. Например, пусть $\mathfrak{X}=R^k$, $\mathscr{A}$ — борелевская σ-алгебра, $\mathscr{B}=\{\{z\in R^k: z<t\};\ t\in R^k\}$ — множество открытых «углов» в R^k. Предположим, что атомы $\{a_i\}$ дискретного распределения $\overline{P}$ занумерованы так, что $a_i\leqslant a_{i+1}$ для любого $i<m$. Тогда, очевидно, $\gamma(\cdot)=1$, и предложенный метод позволяет получить оценку

$$\sup_{A\in\mathscr{B}}\left|\overline{Q}_n(A)-\overline{G}(A)\right|=O\left(n^{-1/2}\log n\right)$$

с вероятностью 1.

Приведем более содержательный и важный для приложения пример. Пусть $\overline{P}$ — произвольное дискретное распределение в R^k, сосредоточенное на решетке:

$$R(N)=\{(t_{i_1}^{(1)},\ldots,t_{i_k}^{(k)});\quad i_s=1,2,\ldots,N,\ s\leqslant k\},$$

где $t_i^{(l)}$ — любые числа. Пусть $\mathscr{B}$ — множество открытых «углов», рассмотренное в предыдущем примере. Упорядочивая точки множества $R(N)$ по лексикографическому принципу, можно показать (см. [5]), что

$$\gamma_0(\mathscr{B},\overline{P})\leqslant N^{k-1}.$$

Лемма 2. *Пусть $\mathscr{B}_N=\{A_1,\ldots,A_N\}$ — произвольный конечный набор множеств из $\mathscr{A}$. Тогда существует дискретное распределение $\overline{P}$, имеющее не более 2^N атомов, и последовательность независимых с. в. $\{\overline{X}_k\}$, распределенных по закону $\overline{P}$ и заданных на том же вероятностном пространстве, что и $\{X_k\}$, для которых $\overline{Q}_n(A_i)=Q_n(A_i)$ при любом $A_i\in\mathscr{B}_N$, где $\overline{Q}_n(\cdot)$ — эмпирический процесс, построенный по выборке $\overline{X}_1,\ldots,\overline{X}_n$.*

Доказательство. Введем следующие обозначения:

$$\begin{aligned}
B_0&=\mathfrak{X}\setminus\bigcup_{i\leqslant N}A_i,\\
B_k&=A_k\setminus\bigcup_{\substack{i\leqslant N,\\ i\neq k}}A_i,\quad k=1,2,\ldots,N,\\
B_{k_1k_2}&=A_{k_1}\cap A_{k_2}\setminus\bigcup_{\substack{i\leqslant N,\\ i\neq k_1,k_2}}A_i,\ k_1<k_2,\ k_l\leqslant N,\\
B_{k_1k_2k_3}&=A_{k_1}\cap A_{k_2}\cap A_{k_3}\setminus\bigcup_{\substack{i\leqslant N,\\ i\neq k_1,k_2,k_3}}A_i,\ k_1<k_2<k_3,\ k_l\leqslant N,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
B_{k_1\ldots k_{N-1}}&=\bigcap_{l\leqslant N-1}A_{k_l}\setminus\bigcup_{\substack{i\leqslant N,\\ i\neq k_1,\ldots,k_{N-1}}}A_i,\ k_1<\ldots<k_{N-1},\ k_l\leqslant N,\\
B_{k_1\ldots k_N}&=\bigcap_{i\leqslant N}A_i.
\end{aligned}\tag{15}$$

Нетрудно видеть, что таким образом введенная совокупность $\{B.\}$ разбивает $\mathfrak{X}$ на непересекающиеся подмножества. Без ограничения общности считаем, что $\mathcal{B}_N \neq \{\mathfrak{X}\}$. Так что семейство (15) содержит, по крайней мере, два непустых множества.

Далее, из каждого непустого множества $B_{k_1 \dots k_l}$ выберем по одной точке $b_{k_1 \dots k_l}$. Последовательность $\{\overline{X}_k\}$ определим следующим образом: если $X_k \in B_{k_1 \dots k_l}$, то $\overline{X}_k = b_{k_1 \dots k_l}$. Из (15) следует, что число атомов этого дискретного распределения не превосходит величины 2^N.

Покажем, что любое множество $A_i \in \mathcal{B}_N$ представимо в виде объединения множеств из $\{B.\}$. В самом деле, выберем из системы множеств (15) все множества, у которых одна из координат мультииндекса равна i. Очевидно, объединение таких множеств содержится в A_i. С другой стороны, для любой точки $a \in A_i$ выберем из $\mathcal{B}_N \backslash \{A_i\}$ все множества A_{k_1}, $A_{k_2}, \dots, A_{k_s}$, содержащие точку a. Понятно, что

$$a \in A_i \bigcap_{j \leqslant s} A_{k_j} \diagdown \bigcup_{\substack{l \leqslant N, \\ l \neq i, k_1, \dots, k_s}} A_l \equiv B_{\cdot i}\,.$$

Следовательно, объединение множеств из $\{B.\}$, у которых одна из координат мультииндекса равна i, совпадает с множеством A_i.

Далее, используя аддитивность вероятностных мер, заключаем, что для распределения $\overline{P}$ с. в. $\overline{X}_1$ и соответствующего эмпирического распределения имеют место равенства

$$\overline{P}(A_i) = P(A_i), \quad \overline{P}_n(A_i) = P_n(A_i), \quad A_i \in \mathcal{B}_N.$$

Лемма доказана.

Лемма 3. *Пусть* $\mathcal{B}_N = \{A_1, \dots, A_N\}$ *и* $\overline{P}$ — *дискретное распределение, построенное в лемме* 2. *Тогда*

$$\gamma_0(\mathcal{B}_N, \overline{P}) \leqslant N.$$

Доказательство. Прежде всего, разобьем совокупность множеств $\{B.\}$, определенных в (15), на $N+1$ непересекающихся классов:

$$\{B.\} = \bigcup_{k=1}^{N} \{B_k.\} \cup \{B_0\},$$

где $B_k.$ — множество вида (15), у которого первая координата мультииндекса равна k. Для простоты рассуждений будем считать, что все множества $B.$ не пусты.

Рассмотрим отдельно класс $\{B_1.\}$. Его можно представить следующим образом:

$$\{B_1.\} = \{B_1\} \cup \bigcup_{l=2}^{N} \{B_{1l}.\}.$$

Теперь сопоставим множеству B_1 номер «1» (точнее, занумеруем атом $b_1 \in B_1$ распределения $\overline{P}$). Далее, элементы класса $\{B_{12}.\}$ занумеруем в произвольном порядке числами от 2 до $m_1 = 2^{N-2}$ (без пропусков); элементы класса $\{B_{13}.\}$ — числами от m_1+1 до $m_1 + m_2 = 2^{N-2} + 2^{N-3}$ и т. д. до тех пор, пока не дойдем до класса $\{B_{1N}\}$, состоящего из одного множества B_{1N}, которому присвоим номер $\sum_{i=1}^{N-1} m_i + 1 = \sum_{i=0}^{N-2} 2^i + 1$.

Таким образом, мы занумеровали все элементы подкласса $\{B_1.\}$, образующего разбиение множества A_1. Тем самым, внутри A_1 образовалась одна цепочка атомов. В то же время нумерация без пропусков элементов подкласса $\{B_{1k}.\}$ образует также одну цепочку атомов внутри A_k.

Следующий этап нумерации состоит в проведении аналогичной процедуры с классом $\{B_2.\}$. Именно

$$\{B_2.\} = \{B_2\} \cup \bigcup_{l=3}^{N} \{B_{2l}.\}.$$

Множеству B_2 присваиваем номер $\sum_{i=1}^{N-1} m_i + 2$. А затем так же, как на предыдущем этапе, осуществляем сквозную нумерацию подклассов $\{B_{23\cdot}\}$, $\{B_{24\cdot}\}$, ..., $\{B_{2N}\}$ (внутри каждого подкласса $\{B_{2k\cdot}\}$ нумерация элементов без пропусков!). Тем самым, внутри A_2 образовалась вторая цепочка атомов, и нумерация элементов разбиения A_2 закончена. Кроме того, внутри каждого из множеств A_3, A_4, ..., A_N построено также по второй цепочке атомов.

Продолжая эту процедуру, можно таким образом построить внутри каждого множества A_k не более k цепочек атомов. Лемма доказана.

Замечание. В реальных задачах, представляющих интерес для приложений, величина $\gamma_0(\mathscr{B}_n, \bar{P})$, как правило, существенно меньше N (несмотря на то, что оценка в лемме 3 неулучшаема). Например, пусть A_1, ..., A_N — суть всевозможные объединения (включая пустое множество) непересекающихся множеств B_1, ..., B_m. Очевидно, в этом случае $N = 2^m$. В то же время совокупность $\{B_i;\ i \leqslant m\}$ совпадает с системой множеств $\{B.\}$, рассмотренных в леммах 2, 3. Иными словами, $\{B_i;\ i \leqslant m\}$ — минимальная система порождающих для $\{A_i;\ i \leqslant N\}$. Так что для любого распределения P имеет место оценка $\gamma_0(\mathscr{B}_N, \bar{P}) < m = \log_2 N$, где распределение $\bar{P}$, соответствующее P и классу $\{A_i;\ i \leqslant N\}$, построено в лемме 2.

В дальнейшем символом $A_{M,i}$ будем обозначать элемент ε-сети с вложением при $\varepsilon = 2^{-M}$. Обозначим также

$$\mathscr{B}_{(M,i)} = \{A \in \mathscr{B}:\ d(A, A_{M,i}) \leqslant 2^{-M}\}, \quad i = 1, 2, \ldots, N_1(2^{-M}).$$

В следующих двух леммах случайные процессы $Q_n(\cdot)$ и $G(\cdot)$ предполагаются d-сепарабельными. Замечание, связанное с этим допущением, будет сделано при доказательстве теоремы 1.

Лемма 4. *Для любых $r \geqslant 0$, $M = 0, 1, \ldots$ и $y \geqslant 2r + 5$ справедливо неравенство*

$$\mathbf{P}\Big(\sup_{A \in \mathscr{B}(M,i)} |Q_n(A) - Q_n(A_{M,i})| \geqslant (4y+1)\Big[\sum_{k>M}[2^{-k}H_1(2^{-k})]^{1/2} + 5\cdot 2^{-M/2}\Big]\Big) \leqslant$$
$$\leqslant I(n, M, H_1)\, N_1(2^{-M})^{-r} R_1(y, r) + (1 - I(n, M, H_1))\, N_1(2^{-M})^2 \times$$
$$\times \exp\Big\{-yn^{1/2} \sum_{k>M} [2^{-k}H_1(2^{-k})]^{1/2}\Big\}, \qquad (16)$$

где функции $R_1(y, r)$ и $I(n, M, H_1)$ определены в теореме 1.

Доказательство. Обозначим через $l^- = l^-(k, A)$ и $l^+ = l^+(k, A)$ соответственно номера внутренней и внешней окрестностей 2^{-k}-сети для множества $A \in \mathscr{B}$, т. е. $A_{k,l^-} \subseteq A \subseteq A_{k,l^+}$ и $d(A_{k,l^-}, A_{k,l^+}) \leqslant 2^{-k}$. Отметим, что $l^{\pm}(\cdot)$ определяются, вообще говоря, не единственным образом. Так что $l^{\pm}(\cdot)$ — это любая пара натуральных чисел с отмеченным свойством.

Далее, введем следующее обозначение:

$$N^* = \min\{k = 0, 1, \ldots : \max(k, H_1(2^{-M-k})) \geqslant 4n2^{-M-k}\} - 1. \qquad (17)$$

Предположим сначала, что $N^* \geqslant 1$. Разобьем совокупность $\mathscr{B}(M, i)$ на непересекающиеся подклассы, для каждого из которых найдется пара $l^{\pm}(M + N^*, \cdot)$, единая для всех элементов из данного подкласса. Очевидно, число таких подклассов не превосходит величины

$$N_1(2^{-M-N^*})(N_1(2^{-M-N^*}) - 1)/2. \qquad (18)$$

Легко видеть, что

$$\sup_{A \in \mathscr{B}(M,i)} |Q_n(A) - Q_n(A_{M,i})| \leqslant \max_{m:\, l^-(M+N^*, A) = m,\, A \in \mathscr{B}(M,i)} |Q_n(A_{M+N^*,m}) -$$
$$- Q_n(A_{M,i})| + \max_{l^{\pm}(M+N^*,\cdot)} \sup_{A \in \mathscr{B}(M,i):\, A_{M+N^*,l^-} \subseteq A \subseteq A_{M+N^*,l^+}} |Q_n(A) -$$
$$- Q_n(A_{M+N^*,l^-})|. \qquad (19)$$

Пусть $A_{M+N^*,m}$ — произвольное подмножество из совокупности, по которой берется максимум в первом слагаемом правой части (19). Очевидно, найдется множество $A\in\mathscr{B}(M, i)$, для которого

$$d(A_{M+N^*,m}, A)\leqslant 2^{-M-N^*}.$$

Теперь по данному A выберем множества $\{A^{(m)}_{M+j,l(j)};\ j=1, 2, \ldots, N^*-1\}$, для которых

$$d\left(A^{(m)}_{M+j,l(j)}, A\right)\leqslant 2^{-M-j}.$$

Положим также $A^{(m)}_{M+N^*,\,l(0)}=A_{M+N^*,m}$, $A^{(m)}_{M,\,l(N^*)}=A_{M,\,i}$. Из неравенства треугольника для $d(\cdot)$ следует, что при любом $j=1, 2, \ldots, N^*$

$$d\left(A^{(m)}_{M+j,l(j)}, A^{(m)}_{M+j-1,l(j-1)}\right)\leqslant 2^{-M-j+2}. \tag{20}$$

Оценим сверху первое и второе слагаемые правой части (19). Имеем

$$\max_{m:\,l^-(M+N^*,A)=m,\,A\in\mathscr{B}(M,i)} |Q_n(A_{M+N^*,m})-Q_n(A_{M,i})|\leqslant$$

$$\leqslant\sum_{j=1}^{N^*}\max_{\substack{m:\,l^-(\cdot)=m,\\ A\in\mathscr{B}(M,i)}}\left|Q_n\left(A^{(m)}_{M+j,l(j)}\right)-Q_n\left(A^{(m)}_{M+j-1,l(j-1)}\right)\right|\leqslant$$

$$\leqslant\sum_{k=1}^{N^*}\max_{m,l:\,d(A_{M+k,m},A_{M+k-1,l})\leqslant 2^{-M-k+2}}|Q_n(A_{M+k,m})-Q_n(A_{M+k-1,l})|\equiv J_1. \tag{21}$$

Оценка для второго слагаемого правой части (19) такова:

$$\max_{l^{\pm}(M+N^*,\cdot)}\ \sup_{A\in\mathscr{B}(M,i):\,A_{M+N^*,l^-}\subseteq A\subseteq A_{M+N^*,l^+}}\left|Q_n(A)-Q_n\left(A_{M+N^*,l^-}\right)\right|=$$

$$=\max_{l^{\pm}(M+N^*,\cdot)}\ \sup_{\cdots}\max\left\{Q_n(A)-Q_n\left(A_{M+N^*,l^-}\right), Q_n\left(A_{M+N^*,l^-}\right)-Q_n(A)\right\}\leqslant$$

$$\leqslant\max_{\substack{k,j:\,A_{M+N^*,k}\subseteq A_{M+N^*,j}\\ d(A_{M+N^*,k},A_{M+N^*,j})\leqslant 2^{-M-N^*}}}\max\{\sqrt{n}\,[P_n(A_{M+N^*,j})-P(A_{M+N^*,k})]-$$

$$-Q_n(A_{M+N^*,k}),\ \sqrt{n}\,[P(A_{M+N^*,j})-P(A_{M+N^*,k})]\}\leqslant$$

$$\leqslant\max_{\substack{k,j:\,A_{M+N^*,k}\subseteq A_{M+N^*,j}\\ d(A_{M+N^*,k},A_{M+N^*,j})\leqslant 2^{-M-N^*}}}|Q_n(A_{M+N^*,j})-Q_n(A_{M+N^*,k})|+$$

$$+\sqrt{n}\,2^{-M-N^*}\equiv J_2+\sqrt{n}\,2^{-M-N^*}. \tag{22}$$

Оценим вероятности больших уклонений с. в. J_1 и J_2. Для любых положительных $y_1, \ldots, y_{N^*}$ из (21) имеем

$$\mathbf{P}\left(J_1>\sum_{k=1}^{N^*}y_k 2^{-(M+k-2)/2}\right)\leqslant\sum_{k=1}^{N^*}N_1\left(2^{-M-k}\right)^2\times$$

$$\times\sup_{A,B:\,d(A,B)\leqslant 2^{-M-k+2}}\mathbf{P}\left(|Q_n(A)-Q_n(B)|>y_k 2^{-(M+k-2)/2}\right). \tag{23}$$

Величина $Q_n(A)-Q_n(B)$ есть нормированная сумма независимых одинаково распределенных с. в. с нулевым средним. Так что для оценки вероятностей, стоящих в правой части (23), можно воспользоваться неравенством С. Н. Бернштейна, которое применительно к нашей задаче выглядит следующим образом. Пусть $\xi_1, \ldots, \xi_n$ — независимые одинаково распределенные с. в. с нулевым средним. Пусть, кроме того, $|\xi_i|\leqslant 1$ с вероятностью 1 и $\mathbf{D}\xi_i\leqslant\sigma$. Тогда для любого $y\geqslant 0$ (см. [8])

$$\mathbf{P}\left(n^{-1/2}\left|\sum_{i=1}^{n}\xi_i\right|>y\sqrt{\sigma}\right)\leqslant 2\exp\left\{-\frac{y^2}{2\left(1+y/\sqrt{n\sigma}\right)}\right\}. \tag{24}$$

Теперь отметим, что для любых $A, B \in \mathscr{A}$

$$\mathbf{D}(Q_n(A) - Q_n(B)) \leqslant d(A, B). \tag{25}$$

Поэтому из (20), (23)—(25) при любых $y_k \geqslant 0$ следует неравенство

$$\mathbf{P}\left(J_1 > \sum_{k=1}^{N^*} y_k 2^{-(M+k-2)/2}\right) \leqslant 2 \sum_{k=1}^{N^*} N_1(2^{-M-k}) \times$$
$$\times \exp\left\{-\frac{y_k^2}{2\left(1 + y_k/(n2^{-M-k+2})^{1/2}\right)}\right\} = 2 \sum_{k=1}^{N^*} \exp\left\{2H_1(2^{-M-k}) - \right.$$
$$\left. - \frac{y_k^2}{2\left(1 + y_k/(n2^{-M-k+2})^{1/2}\right)}\right\}.$$

Положим

$$y_k = y\left[\max\left(H_1(2^{-M-k}), k\right)\right]^{1/2}. \tag{26}$$

Тогда для любого $k \leqslant N^*$

$$y_k \leqslant y(n2^{-M-k+2})^{1/2}$$

и, стало быть,

$$\mathbf{P}\left(J_1 > \sum_{k=1}^{N^*} y_k 2^{-(M+k-2)/2}\right) \leqslant 2 \sum_{k=1}^{N^*} \exp\left\{-\left[\frac{y^2}{2(1+y)} - 2\right] \max\left(H_1(2^{-M-k}), k\right)\right\} =$$
$$= 2 \sum_{k=1}^{N^*} \exp\left\{-\left[\frac{(y-r-2)^2 - (r+3)^2 + 1}{2(1+y)} + r\right] \max\left(H_1(2^{-M-k}), k\right)\right\} \leqslant$$
$$\leqslant 2N_1(2^{-M})^{-r} \sum_{k=1}^{N^*} \exp\left\{-\frac{(y-r-2)^2 - (r+3)^2 + 1}{2(1+y)} \max\left(H_1(2^{-M-k}), k\right)\right\}, \tag{27}$$

где r — произвольное неотрицательное число. Очевидно, при $y \geqslant 2r+5$ аргументы у экспонент, стоящих под знаком суммы в правой части (27), становятся отрицательными. Поэтому при $y \geqslant 2r+5$ величины $\max(H_1(2^{-M-k}), k)$ в (27) можно заменить на k. Таким образом,

$$\mathbf{P}\left(J_1 > \sum_{k=1}^{N^*} y_k 2^{-(M+k-2)/2}\right) \leqslant$$
$$\leqslant 2N_1(2^{-M})^{-r}\left[\exp\left\{\frac{(y-r-2)^2 - (r+3)^2 + 1}{2(1+y)}\right\} - 1\right]^{-1}. \tag{28}$$

Далее, из (26) получаем

$$\sum_{k=1}^{N^*} y_k 2^{-(M+k-2)/2} \leqslant 2y \sum_{k=M+1}^{N^*} \left[2^{-k} H_1(2^{-k})\right]^{1/2} + y 2^{1-M/2} \sum_{k=1}^{N^*} (k2^{-k})^{1/2}. \tag{29}$$

Рассмотрим правую часть (22). Принимая во внимание соотношения (18) и (25), с помощью неравенства С. Н. Бернштейна (24) получаем

$$\mathbf{P}(J_2 > z\sqrt{n} 2^{-M-N^*}) \leqslant \exp\left\{2H_1(2^{-M-N^*}) - \frac{z^2 n 2^{-M-N^*}}{2(1+z)}\right\}.$$

Из (17) следует

$$n2^{-M-N^*} \geqslant 4^{-1} \max\left(H_1(2^{-M-N^*}), N^*\right),$$

стало быть,

$$\mathbf{P}(J_2 > z\sqrt{n} 2^{-M-N^*}) \leqslant \exp\left\{-\left[\frac{z^2}{8(1+z)} - 2\right] \max\left(H_1(2^{-M-N^*}), N^*\right)\right\}.$$

Теперь положим $z=4y$. Тогда нетрудно проверить, что при $y\geqslant 5$

$$\frac{z^2}{8(1+z)}>\frac{y^2}{2(1+y)}.$$

Поэтому (см. вывод (27)) при $y\geqslant 2r+5$

$$\mathbf{P}\left(J_2>z\sqrt{n}2^{-M-N^*}\right)\leqslant\exp\left\{-\left[\frac{y^2}{2(1+y)}-2\right]\max\left(H_1\left(2^{-M-N^*}\right),N^*\right)\right\}\leqslant$$
$$\leqslant N_1\left(2^{-M}\right)\exp\left\{-\frac{(y-r-2)^2-(r+3)^2+1}{2(1+y)}\cdot N^*\right\},\qquad(30)$$

где r — произвольное неотрицательное число.

Далее, из (17) получаем

$$\sqrt{n}2^{-M-N^*}\leqslant\left[\max\left(H_1\left(2^{-M-N^*-1}\right),N^*+1\right)\right]^{1/2}2^{-(M+N^*+1)/2}\leqslant$$
$$\leqslant\left[H_1\left(2^{-M-N^*-1}\right)2^{-M-N^*-1}\right]^{1/2}+\left[(N^*+1)\,2^{-N^*-1}\right]^{1/2}2^{-M/2}.\qquad(31)$$

Таким образом, из (22), (30) и (31) следует, что при $y\geqslant 2r+5$

$$\mathbf{P}\Bigg(\max_{l^{\pm}(M+N^*,\cdot)}\ \sup_{\substack{A\in\mathscr{B}(M,i):\\ A_{M+N^*,l^-}\subseteq A\subseteq A_{M+N^*,l^+}}}\left|Q_n(A)-Q_n\left(A_{M+N^*,l^-}\right)\right|>$$
$$>(4y+1)\left\{\left[H_1\left(2^{-M-N^*-1}\right)2^{-M-N^*-1}\right]^{1/2}+\left[(N^*+1)\,2^{-N^*-1}\right]^{1/2}2^{-M/2}\right\}\Bigg)\leqslant$$
$$\leqslant N_1\left(2^{-M}\right)^{-r}\exp\left\{-\frac{(y-r-2)^2-(r+3)^2+1}{2(1+y)}\cdot N^*\right\}.\qquad(32)$$

Утверждение леммы следует из (19), (21), (28), (29), (32) и неравенства $\sum\limits_{k\geqslant 1}\left(k2^{-k}\right)^{1/2}<5$.

Пусть теперь $N^*<1$ (т. е. $n2^{-M+1}\leqslant\max\left(H_1(2^{-M-1}),\ 1\right)$). Тогда так же, как в (19) и (22), получаем

$$\sup_{A\in\mathscr{B}(M,i)}\left|Q_n(A)-Q_n\left(\hat{A}_{M,i}\right)\right|\leqslant$$
$$\leqslant\max_{k,j:A_{M,k}\subseteq A_{M,j},d(A_{M,k},A_{M,j})<2^{-M}}\left|Q_n\left(A_{M,k}\right)-Q_n\left(A_{M,j}\right)\right|+$$
$$+\sqrt{n}2^{-M}\equiv J_3+\sqrt{n}2^{-M}.\qquad(33)$$

Далее, при любом $z\geqslant 1$ из неравенства С. Н. Бернштейна (24) следует, что

$$\mathbf{P}\left(J_3>z\sum_{k>M}\left[2^{-k}H_1\left(2^{-k}\right)\right]^{1/2}\right)\leqslant$$
$$\leqslant N_1\left(2^{-M}\right)^2\exp\left\{-\frac{z^2 2^M\left(\sum\limits_{k>M}\left[2^{-k}H_1\left(2^{-k}\right)\right]^{1/2}\right)^2}{2\left(1+z2^Mn^{-1/2}\sum\limits_{k>M}\left[2^{-k}H_1\left(2^{-k}\right)\right]^{1/2}\right)}\right\}\leqslant$$
$$\leqslant N_1\left(2^{-M}\right)^2\exp\left\{-4^{-1}zn^{1/2}\sum_{k>M}\left[2^{-k}H_1\left(2^{-k}\right)\right]^{1/2}\right\}.\qquad(34)$$

Кроме того, если $N^*<1$, то

$$\sqrt{n}2^{-M}<\left[H_1(2^{-M-1})2^{-M-1}\right]^{1/2}+2^{-M/2}.\qquad(35)$$

Полагая $z=4y$, из (33)—(35) получаем требуемое. Лемма доказана.

Замечание. Специфика метрической энтропии с вложением используется по существу лишь при доказательстве леммы 4. Можно привлекать и другие соображения для получения неравенств вида (16). Однако это выходит за рамки данного исследования.

Лемма 5. *При выполнении условия* (5) *для любых* $r \geqslant 0$, $M = 0, 1, \ldots$ *и* $y \geqslant (2r+5)^{1/2}$

$$\mathbf{P}\left(\sup_{A \in \mathscr{B}(M,i)} |G(A) - G(A_{M,i})| > 2y\left(\sum_{k>M}[2^{-k}H_1(2^{-k})]^{1/2} + 5\cdot 2^{-M/2}\right)\right) \leqslant$$
$$\leqslant N_1(2^{-M})^{-r} R_2(y, r),$$

где функция $R_2(y, r)$ *определена в теореме* 1.

Доказательство. В основе доказательства леммы лежат хорошо известные приемы оценки максимального выброса траекторий сепарабельного гауссовского случайного процесса с конечномерным параметрическим множеством (см., например, [4]). Наши рассуждения в известной степени аналогичны рассуждениям, имевшим место при доказательстве леммы 4.

Итак, для любого d-сепарабельного случайного процесса на $\mathscr{A}$ с вероятностью 1 справедливо неравенство

$$\sup_{A \in \mathscr{B}(M,i)} |G(A) - G(A_{M,i})| \leqslant$$
$$\leqslant \sup_{A \in \mathscr{B}(M,i)} \sup_{k=1,2,\ldots} \left|G\left(A_{M+k,l^-(M+k,A)}\right) - G(A_{M,i})\right| \equiv J_4, \quad (36)$$

где, как и прежде, $l^-(M+k, A)$ определяется из условия

$$d\left(A_{M+k,l^-(\cdot)}, A\right) \leqslant 2^{-M-k}.$$

Очевидно,

$$J_4 \leqslant \sum_{k=1}^{\infty} \sup_{A \in \mathscr{B}(M,i)} \left|G(A_{M+k,l(M+k,A)}) - G\left(A_{M+k-1,l^-(M+k-1,A)}\right)\right|,$$

где для удобства обозначений полагаем $A_{M,l^-(M,A)} = A_{M,i}$. Так как подкласс $\{A_{M+k,l^-(M+k,A)};\ A \in \mathscr{B}(M, i)\}$ конечен, то с учетом (20) можно еще раз усилить предыдущее неравенство:

$$J_4 \leqslant \sum_{k=1}^{\infty} \max_{k,j:d\left(A_{M+k,l}, A_{M+k-1,j}\right) \leqslant 2^{-M-k+2}} |G(A_{M+k,l}) - G(A_{M+k-1,j})|. \quad (37)$$

Следовательно, в силу (24), (25) и (37) при любых $y_k \geqslant 0$

$$\mathbf{P}\left(J_4 > \sum_{k=1}^{\infty} y_k 2^{-(M+k-2)/2}\right) \leqslant \sum_{k=1}^{\infty} N_1(2^{-M-k})^2\, \overline{\Phi}(y_k), \quad (38)$$

где $\overline{\Phi}(y) = (2\pi)^{-1/2} \int\limits_y^{\infty} e^{-t^2/2} dt.$

Далее, как и при доказательстве леммы 4, положим в (38)

$$y_k = y[\max(H_1(2^{-M-k}),\ k)]^{1/2}.$$

Тогда для любого $r \geqslant 0$

$$\mathbf{P}\left(J_4 > \sum_{k=1}^{\infty} y_k 2^{-(M+k-2)/2}\right) \leqslant \sum_{k=1}^{\infty} [\max(H_1(2^{-M-k}), k)/2\pi]^{1/2} \times$$
$$\times \int\limits_y^{\infty} \exp\{-(t^2/2 - 2)\max(H_1(2^{-M-k}), k)\}dt \leqslant N_1(2^{-M})^{-r}(2\pi)^{-1/2} \times$$
$$\times \sum_{k=1}^{\infty} \int\limits_y^{\infty} \exp\{-((t^2 - 2r - 5)/2)\max(H_1(2^{-M-k}), k)\,dt. \quad (39)$$

Если $y \geqslant (2r+5)^{1/2}$, то, заменив величину $\max(H_1(2^{-M-k}),\ k)$ на k в правой части (39), окончательно получаем

$$\mathbf{P}\left(J_4 > \sum_{k=1}^{\infty} y_k 2^{-(M+k-2)/2}\right) \leqslant$$

$$\leqslant N_1 (2^{-M})^{-r} (2\pi)^{-1/2} \int_y^{\infty} [\exp\{(t^2 - 2r - 5)/2\} - 1]^{-1}\, dt. \quad (40)$$

Утверждение леммы следует из (36), (40) и из неравенства

$$\sum_{k=1}^{\infty} y_k 2^{-(M+k-2)/2} \leqslant 2y \left(\sum_{k>M} [2^{-k} H_1(2^{-k})]^{1/2} + 5 \cdot 2^{-M/2}\right).$$

Доказательство теоремы 1. Пусть M — произвольное целое неотрицательное число. С помощью лемм 1—3 методом квантильных преобразований (см. [7]) можно так построить на траектории броуновского моста $\mathring{w}(\cdot)$ совокупности с. в. $\{\tilde{Q}_n(A_{M,i});\ i \leqslant N_1(2^{-M})\}$, $n = 1, 2, \ldots$, и $\{G(A_{M,i});\ i \leqslant N_1(2^{-M})\}$, что для любого n

$$\mathbf{P}\left(\max_{i \leqslant N_1(2^{-M})} |\tilde{Q}_n(A_{M,i}) - G(A_{M,i})| \geqslant\right.$$

$$\left.\geqslant C\gamma^*(2^{-M}, \mathscr{B}, P)(\log n + x)\, n^{-1/2}\right) \leqslant K e^{-\lambda x}, \quad (41)$$

где C, k, λ — абсолютные положительные постоянные, $\gamma^*(2^{-M}, \mathscr{B}, P) = \gamma_0(\mathscr{B}_N, \bar{P})$, а $\bar{P}$ определено в лемме 2 для семейства множеств $\mathscr{B}_N = \{A_{M,i};\ i \leqslant N_1(2^{-M})\}$.

Из леммы 3 следует, что величина $\gamma^*(\varepsilon, \mathscr{B}, P)$ удовлетворяет неравенству (7).

Далее, применяя теорему А. Н. Колмогорова для согласованных условных распределений, расширив соответствующим образом уже построенное вероятностное пространство (на котором задан $\mathring{w}(\cdot)$), доопределим значения $\tilde{Q}_n(A)$, $n = 1, 2, \ldots$, и $G(A)$ во всех оставшихся «точках» $A \in \mathscr{A}$. При этом $\tilde{Q}_1(\cdot)$, $\tilde{Q}_2(\cdot)$, ..., $G(\cdot)$ строятся как условно независимые относительно σ-алгебры, порожденной $\mathring{w}(\cdot)$, случайные процессы. Без ограничения общности рассуждений можно считать эти случайные процессы d-сепарабельными (см. [4]). В связи с этим необходимо отметить, что при выводе оценок (22) и (33) в лемме 4 использовался явный вид процесса $Q_n(\cdot)$, который предполагался d-сепарабельным. Однако, если принять во внимание процедуру построения d-сепарабельной модификации процесса $\tilde{Q}_n(\cdot)$ (см. [4]), нетрудно понять, что соответствующие неравенства имеют место и для сепарабелизированного случайного процесса $\tilde{Q}_n(\cdot)$.

Для построенных на одном вероятностном пространстве случайных процессов $\tilde{Q}_n(\cdot)$ и $G(\cdot)$ справедливо неравенство

$$\sup_{A \in \mathscr{B}} |\tilde{Q}_n(A) - G(A)| \leqslant \max_{i \leqslant N_1(2^{-M})} |\tilde{Q}_n(A_{M,i}) - G(A_{M,i})| +$$

$$+ \max_{i \leqslant N_1(2^{-M})} \sup_{A \in \mathscr{B}(M,i)} |\tilde{Q}_n(A) - \tilde{Q}_n(A_{M,i})| +$$

$$+ \max_{i \leqslant N_1(2^{-M})} \sup_{A \in \mathscr{B}(M,i)} |G(A) - G(A_{M,i})|. \quad (42)$$

Утверждение теоремы следует из (41), (42) и лемм 4, 5.

Доказательство теоремы 2. Отличие от предыдущего доказательства состоит лишь в построении случайных процессов $\tilde{Q}_n(\cdot)$.

Обозначим через $\nu_n(A)$ число точек из выборки X_1, X_2, ..., X_n, попавших в множество $A \in \mathscr{A}$. Теперь для каждого n на некотором вероятностном пространстве зададим совокупность независимых случайных векторов, занумерованных мультииндексом $(n_1, \ldots, n_s)$:

$$\left\{(X_{1n}(n_1, \ldots, n_s), \ldots, X_{nn}(n_1, \ldots, n_s));\ n_j \geqslant 0,\ \sum_{j=1}^{s} n_j = n\right\}, \quad (43)$$

где $s = N_1(2^{-M})$, а распределение вектора при фиксированном мультииндексе $(n_1, \ldots, n_s)$ совпадает с условным распределением выборки $X_1, \ldots, X_n$ при условии $\bigcap_{i \leqslant s} \{\nu_n(A_{M,i}) = n_i\}$. При этом если событие $\bigcap_{i \leqslant s} \{\nu_n(A_{M,i}) = n_i\}$ имеет нулевую вероятность, то распределение соответствующего вектора $(X_{1n}(n_1, \ldots, n_s), X_{2n}(n_1, \ldots, n_s), \ldots, X_{nn}(n_1, \ldots, n_s))$ можно считать произвольным. При различных n семейства случайных векторов (43) предполагаются независимыми в совокупности.

Далее, на траектории броуновского моста $\overset{\circ}{w}(\cdot)$, не зависящего от указанного множества случайных векторов, построим так же, как и в теореме 1, совокупности с. в. $\{\tilde{Q}_n(A_{M,i});\ i \leqslant s\}$, $n = 1, 2, \ldots$, и $\{G(A_{M,i});\ i \leqslant s\}$. Обозначим

$$X_{kn} = X_{kn}(\tilde{\nu}_n(A_{M,1}), \ldots, \tilde{\nu}_n(A_{M,s})),$$

где $\tilde{\nu}(A_{M,i}) = n^{1/2}\tilde{Q}_n(A_{M,i}) + nP(A_{M,i})$ — копия с. в. $\nu_n(A_{M,i})$. Очевидно, случайный вектор $(X_{1n}, \ldots, X_{nn})$ совпадает по распределению с выборкой $X_1, \ldots, X_n$ и число точек из выборки $X_{1n}, \ldots, X_{nn}$, попавших в множество $A_{M,i}$, с вероятностью 1 равно $\tilde{\nu}_n(A_{M,i})$.

Теперь построим эмпирическую меру $\tilde{Q}_n(\cdot)$ по выборке $X_{1n}, \ldots, X_{nn}$. Очевидно, распределения любых измеримых функционалов от $Q_n(\cdot)$ и $\tilde{Q}_n(\cdot)$ совпадают. Так что запись „$\overset{\mathscr{D}}{=}$“ в этом случае можно понимать в значительно более сильном смысле, чем в (3).

Отметим также, что случайный процесс $G(\cdot)$ в этом случае строится так же, как и в предыдущей теореме, т. е. с помощью теоремы А. Н. Колмогорова о согласованных распределениях. При этом можно считать процесс $G(\cdot)$ условно (относительно σ-алгебры, порожденной с. в. $\{G(A_{M,i});\ i \leqslant s\}$) не зависящим от $\{\tilde{Q}_n(\cdot);\ n = 1, 2, \ldots\}$.

Существование измеримой оценки Λ_n следует из неравенства (42) и неравенств, полученных при доказательстве леммы 4. Теорема 2 доказана.

3. Нижние оценки скорости сходимости

В этом разделе речь пойдет о точности оценок вида (9). Будет показано, что полученные оценки скорости сходимости в центральной предельной теореме для эмпирических мер в известном смысле неулучшаемы.

Рассмотрим сначала случай, когда $\mathfrak{X} = R$, $\mathscr{B} = \{(-\infty, t);\ t \in R\}$. Из (11) при $m = 1$ следует, что указанная оценка скорости сходимости с вероятностью 1 имеет порядок малости $O(n^{-1/2}\log n)$, который, как показано в [7], является максимально возможным (т. е. при любом способе построения случайных процессов Q_n и G на одном вероятностном пространстве порядок близости их траекторий не может быть больше, чем $O(n^{-1/2}\log n)$). Отметим, что в этом случае $H_1(\varepsilon) \sim |\log \varepsilon|$ при $\varepsilon \to 0$, т. е. в данном примере метрическая энтропия с вложением при $\varepsilon \to 0$ растет достаточно медленно. В то же время, если $H_1(\varepsilon)$ при $\varepsilon \to 0$ растет, скажем, степенным образом, то величина $\Delta_n(\tilde{M}(n))$ в (9) будет убывать значительно медленнее $n^{-1/2}\log n$. Скорость этого убывания в зависимости от $H_1(\cdot)$ может быть сколь угодно малой. Возникает вопрос, по существу ли в рассматриваемой общности такое ухудшение? Положительный ответ на поставленный вопрос дает

Теорема 3. *Существуют распределение P и класс $\mathscr{B} \subseteq \mathscr{A}$, для которых:*

1) $H_1(\varepsilon) \sim \psi(\varepsilon)\varepsilon^{-1}$

при $\varepsilon \to 0$, *где* $\psi(\cdot)$ — *положительная функция, удовлетворяющая соотношениям* $\lim\limits_{\varepsilon\to 0}\psi(\varepsilon)=0$, $\lim\limits_{\varepsilon\to 0}\psi(\varepsilon)\varepsilon^{-\alpha}=\infty$ *для всякого* $\alpha>0$, $\sum\limits_{k\geqslant 1}\psi(2^{-k})<$ $<\infty$;

2) *при любом методе задания случайных процессов* $\{Q_n(\cdot)\}$ *и* $G(\cdot)$ *на одном вероятностном пространстве*

$$\sup_{A\in\mathscr{B}} |Q_n(A)-G(A)| \geqslant c\Delta_n(\widetilde{M}(n)) \tag{44}$$

с вероятностью 1 *при всех достаточно больших* n, *где* c — *абсолютная положительная постоянная.*

Доказательство. Пусть $\overline{P}=\{p_i\}$ — произвольное дискретное распределение в $\mathfrak{X}$ со счетным числом атомов $a_1, a_2, \ldots$, причем $\{a_i\}$ занумерованы так, что последовательность соответствующих им масс $\{p_i\}$ монотонно убывает. Можно считать, что процессы $\overline{Q}_n(\cdot)$ и $\overline{G}(\cdot)$, соответствующие данному $\overline{P}$, определены по формулам (12) и (13) при $m=\infty$ (подробнее см. [3]). Допустим, что случайные процессы $V_n(\cdot)$ и $\overset{\circ}{w}(\cdot)$ в формулах (12) и (13), а тем самым и процессы $\overline{Q}_n(\cdot)$ и $\overline{G}(\cdot)$ уже заданы каким-нибудь образом на одном вероятностном пространстве.

Положим $\mathscr{B}=\mathscr{A}$ и будем считать, что

$$\sum_{i\geqslant 1}\sqrt{p_i}<\infty. \tag{45}$$

В этом случае величина $\sup\limits_{A\in\mathscr{A}}|\overline{Q}_n(A)-\overline{G}(A)|$ есть не что иное, как расстояние по вариации между зарядами $\overline{Q}_n$ и $\overline{G}$, которые абсолютно непрерывны относительно считающей меры в точках $\{a_i\}$, т. е.

$$\sup_{A\in\mathscr{A}}|\overline{Q}_n(A)-\overline{G}(A)|\geqslant\frac{1}{2}\sum_{i>m}|\Delta_iV_n-\Delta_i\overset{\circ}{w}|\geqslant$$

$$\geqslant\frac{1}{2}\sum_{i>m}|\Delta_i\overset{\circ}{w}|-\frac{1}{2}\sum_{i>m}|\Delta_iV_n|\equiv\zeta, \tag{46}$$

где $\Delta_i g=g(F(i+1))-g(F(i))$, $F(i)=\sum\limits_{k<i}p_k$; m — произвольное натуральное число, которое далее будет выбрано. Отметим, что при выполнении (45) с. в. ζ будет собственной (см. [3]).

Нам понадобятся оценки сверху для вероятностей уклонений каждой из двух сумм правой части (46).

Лемма 6. *Пусть выполнено условие* (45). *Тогда для любых положительных* $x\leqslant\sqrt{n}\sum\limits_{i>m}p_i$ *и* z *имеют место неравенства*

$$\mathbf{P}\left(\sum_{i>m}|\Delta_iV_n|>x+2\sqrt{n}\sum_{i>m}p_i\right)\leqslant\exp\left\{-\frac{x^2}{4\sum\limits_{i>m}p_i}\right\}, \tag{47}$$

$$\mathbf{P}\left(\left|\sum_{i>m}|\Delta_i\overset{\circ}{w}|-\sqrt{\frac{2}{\pi}}\sum_{i>m}\sqrt{p_i}\right|>z\right)\leqslant 4\exp\left\{-\frac{z^2}{8\sum\limits_{i>m}p_i}\right\}. \tag{48}$$

Доказательство. Соотношение (47) сразу следует из неравенства С. Н. Бернштейна (24), поскольку

$$\sum_{i>m}|\Delta_iV_n|\leqslant\frac{1}{\sqrt{n}}\sum_{j=1}^{n}I_{\xi_j}([F(m+1),1])+\sqrt{n}\sum_{i>m}p_i,$$

где $\xi_1, \ldots, \xi_n$ — независимые с. в. с равномерным распределением на отрезке $[0, 1]$.

Для доказательства (48) воспользуемся известным представлением $\overset{\circ}{w}(t)=w(t)-tw(1)$, где $w(t)$ — стандартный винеровский процесс.

Тогда

$$\mathbf{P}\left(\left|\sum_{i>m}|\Delta_i \mathring{w}|-\sqrt{\frac{2}{\pi}}\sum_{i>m}\sqrt{p_i}\right|>z\right)\leqslant$$

$$\leqslant \mathbf{P}\left(\left|\sum_{i>m}|\Delta_i w|-\sqrt{\frac{2}{\pi}}\sum_{i>m}\sqrt{p_i}\right|>z/2\right)+\mathbf{P}\left(|w(1)|\sum_{i>m}p_i>z/2\right)\leqslant$$

$$\leqslant \mathbf{P}\left(\sum_{i>m}|\Delta_i w|<\sqrt{\frac{2}{\pi}}\sum_{i>m}\sqrt{p_i}-z/2\right)+$$

$$+\mathbf{P}\left(\sum_{i>m}|\Delta_i w|>\sqrt{\frac{2}{\pi}}\sum_{i>m}\sqrt{p_i}+z/2\right)+2\exp\left\{-\frac{z^2}{8\left(\sum\limits_{i>m}p_i\right)^2}\right\}. \quad (49)$$

Оценим первое слагаемое правой части (49). Применяя экспоненциальное неравенство Чебышева и учитывая независимость с. в. $\{\Delta_i w\}$, получаем для любого $t>0$

$$\mathbf{P}\left(\sum_{i>m}|\Delta_i w|<\sqrt{\frac{2}{\pi}}\sum_{i>m}\sqrt{p_i}-z/2\right)\leqslant$$

$$\leqslant \exp\left\{t\sqrt{\frac{2}{\pi}}\sum_{i>m}\sqrt{p_i}-tz/2\right\}\prod_{i>m}\mathbf{E}\exp\{-t|\Delta_i w|\}. \quad (50)$$

Отметим, что

$$\mathbf{E}|\Delta_i w|=\sqrt{2p_i/\pi},\quad \mathbf{E}(\Delta_i w)^2=p_i.$$

Используя элементарные неравенства

$$e^{-x}\leqslant 1-x+x^2/2,\quad x\geqslant 0,$$
$$1+y\leqslant e^y,\quad y\in R,$$

легко устанавливаем, что

$$\prod_{i>m}\mathbf{E}\exp\{-t|\Delta_i w|\}\leqslant\exp\left\{-t\sqrt{\frac{2}{\pi}}\sum_{i>m}\sqrt{p_i}+\frac{t^2}{2}\sum_{i>m}p_i\right\}.$$

Подставляя полученную оценку в (50) и полагая $t=z\left(2\sum\limits_{i>m}p_i\right)^{-1}$, окончательно получаем

$$\mathbf{P}\left(\sum_{i>m}|\Delta_i w|<\sqrt{\frac{2}{\pi}}\sum_{i>m}\sqrt{p_i}-z/2\right)\leqslant\exp\left\{-\frac{z^2}{8\sum\limits_{i>m}p_i}\right\}.$$

Аналогичным образом можно показать, что такая же оценка справедлива и для второго слагаемого правой части (49). Лемма доказана.

Далее, положим в (47) и (48) $x=z=\sqrt{n}\sum\limits_{i>m}p_i$. Тогда из (46) следует, что на множестве элементарных исходов, вероятность которого не меньше величины $1-5\exp\left\{-n\sum\limits_{i>m}p_i/8\right\}$, справедливо неравенство

$$\zeta\geqslant\frac{1}{\sqrt{2\pi}}\sum_{i>m}\sqrt{p_i}-2\sqrt{n}\sum_{i>m}p_i. \quad (51)$$

Покажем, что в достаточно широком классе дискретных распределений при надлежащем выборе m правая часть (51) по порядку малости совпадает с верхней оценкой, полученной в первом разделе статьи. Заодно будет установлено, что скорость сходимости в центральной предельной теореме может быть сколь угодно малой.

Из условия (45) и монотонности $\{p_i\}$ следует представление

$$p_i=\varphi^2(i)i^{-2}, \quad (52)$$

где $\lim\limits_{i\to\infty}\varphi(i)=0$. Без ограничения общности можно считать, что $\varphi(x)$ — положительная и непрерывно дифференцируемая на $(0,\ \infty)$ функция. Рассмотрим класс дискретных распределений, для которых функция $\varphi(x)$ в (52) при всех $x>0$ удовлетворяет неравенству

$$|\varphi'(x)|\leqslant\gamma(x)\varphi(x)x^{-1}, \tag{53}$$

где $\lim\limits_{x\to\infty}\gamma(x)=0$. Для любого $\varepsilon\in(0,\ 1/2)$ обозначим через $m(\varepsilon,\ \gamma)$ наименьшее натуральное число, определяемое условием $\gamma(x)\leqslant\varepsilon$ при всех $x\geqslant m(\varepsilon,\ \gamma)$. Так как для любого натурального m

$$\int\limits_{m+1}^{\infty}x^{-2}\varphi^2(x)\,dx\leqslant\sum_{i>m}p_i\leqslant\int\limits_{m}^{\infty}x^{-2}\varphi^2(x)\,dx,$$

то, имея в виду равенство

$$\int\limits_{m}^{\infty}x^{-2}\varphi^2(x)\,dx=m^{-1}\varphi^2(m)+\int\limits_{m}^{\infty}x^{-1}\varphi'(x)\,dx$$

и учитывая (53), получаем, что при всех $m\geqslant m(\varepsilon,\ \gamma)$

$$\frac{\varphi^2(m+1)}{(1+\varepsilon)(m+1)}\leqslant\sum_{i>m}p_i\leqslant\frac{\varphi^2(m)}{(1-\varepsilon)m}. \tag{54}$$

Отсюда, в свою очередь, следуют соотношения

$$\frac{\varphi^2(m)}{m}\leqslant\frac{m(1+\varepsilon)}{m-1-\varepsilon}\sum_{i>m}p_i, \tag{55}$$

$$\sum_{i>m}p_i\leqslant\frac{(1-\varepsilon)(m+1)+1}{(1-\varepsilon)(m+1)}\cdot\frac{\varphi^2(m+1)}{m+1}. \tag{56}$$

Обозначим

$$m(n)=\max\{i:np_i\geqslant 1\},$$
$$m^*(n)=\max\{m(n),\ m(\varepsilon,\ \gamma)+1\}.$$

Положим в (51) $m=m^*(n)$. Тогда из (56) и определения $m^*(n)$ легко получаем

$$\sqrt{n}\sum_{i>m^*(n)}p_i\leqslant C\varphi(m^*(n)+1).$$

Далее, с помощью правила Лопиталя и (53) убеждаемся в том, что при $m\to\infty$

$$\varphi(m)\left[\int\limits_{m}^{\infty}x^{-1}\varphi(x)\,dx\right]^{-1}\to 0.$$

Поскольку $\lim\limits_{n\to\infty}m^*(n)=\infty$ и

$$\sum_{i>m}\sqrt{p_k}\geqslant\int\limits_{m+1}^{\infty}x^{-1}\varphi(x)\,dx, \tag{57}$$

то для любого сколь угодно малого $\alpha>0$ и всех $n\geqslant n_0(\alpha,\ \varphi)$ справедливо неравенство

$$\sqrt{n}\sum_{i>m^*(n)}p_i\leqslant\alpha\sum_{i>m^*(n)}\sqrt{p_i}. \tag{58}$$

Таким образом, из (51), (57), (58) и определения $m^*(n)$ следует, что с вероятностью, не меньшей $1-5\exp\{-m^*(n)/8\}$, при $n\geqslant n_0(\alpha,\ \varphi)$ имеет место неравенство

$$\zeta \geqslant \left[(2\pi)^{-1/2} - 2\alpha\right] \int\limits_{m^*(n)+1}^{\infty} x^{-1}\varphi(x)\,dx, \tag{59}$$

где, конечно, предполагается, что $\alpha < (2\sqrt{2\pi})^{-1}$.

Теперь рассмотрим верхнюю оценку скорости сходимости. В первом разделе статьи показано, что для любого дискретного распределения (с монотонным убыванием масс $\{p_i\}$)

$$\Delta_n(\widetilde{M}(n)) \leqslant \delta(n) \equiv$$
$$\equiv \widetilde{C} \min_{m:\, m \leqslant 4n \sum\limits_{i>m} p_i} \max\left(\left[1 + m^{-1}\log n\right] \sum_{i>m} \sqrt{p_i},\ mn^{-1/2}\log n\right).$$

Заметим, что из (55) и определения $m^*(n)$ следуют оценки

$$m^*(n)n^{-1/2} \leqslant \varphi(m^*(n)),$$
$$m^*(n) \leqslant \frac{nm^*(n)(1+\varepsilon)}{m^*(n) - 1 - \varepsilon} \sum_{i>m^*(n)} p_i < 3n \sum_{i>m^*(n)} p_i.$$

Иными словами,

$$\delta(n) \leqslant \widetilde{C} \max\left(\left[1 + \frac{\log n}{m^*(n)}\right] \int\limits_{m^*(n)}^{\infty} x^{-1}\varphi(x)\,dx,\quad \varphi(m^*(n))\log n\right). \tag{60}$$

Теперь уже нетрудно построить пример распределения, для которого порядки малости верхней и нижней оценок совпадают. Однако оценку (60) можно еще конкретизировать, несколько сузив класс дискретных распределений, определяемый условием (53). Именно, дополнительно к (53) потребуем, чтобы функция $\varphi(x)$ в (52) монотонно убывала при $x \to \infty$. В этом случае справедливо представление

$$\varphi(x) = \beta(x)(\log x)^{-1}, \quad x > 1, \tag{61}$$

где $\lim\limits_{x\to\infty} \beta(x) = 0$ и, кроме того, условие (53) означает, что $\varphi(x)$ и $\beta(x)$ — медленно меняющиеся при $x \to \infty$ функции, т. е. для любого $t > 0$

$$\lim_{n\to\infty} m^*(n)\, n^{-(1/2-t)} = \infty. \tag{62}$$

Очевидно, при $n \geqslant n_0(\varphi)$ из (60) с учетом (61) и (62) следует оценка

$$\delta(n) \leqslant 2\widetilde{C} \max\left(\int\limits_{m^*(n)}^{\infty} \beta(x)(x\log x)^{-1}dx,\ \beta(m^*(n))\right), \tag{63}$$

по существу совпадающая с оценкой, полученной в (9). В то же время правая часть (59) является нижней границей, за которую величина $\sup\limits_{A\in\mathcal{A}} |\widetilde{Q}_n(A) - \overline{G}(A)|$, начиная с некоторого n, не выходит с вероятностью 1 (это вытекает из леммы Бореля-Кантелли, так как $\exp\{-m^*(n)/8\} \leqslant \exp\{-n^{1/4}/8\}$ при всех достаточно больших n).

В качестве примера рассмотрим дискретное распределение с массами

$$p_k = C\left[k\left[\log^{(r)}(k+d)\right]^{\varepsilon} \prod_{1\leqslant i\leqslant r} \log^{(i)}(k+d)\right]^{-2}, \tag{64}$$

где $\log^{(i)}(\cdot)$ — i-кратный логарифм, $\log^{(r)} d > 0$, $\varepsilon > 0$, C — постоянная нормировки. Легко видеть, что в этом случае выполнено условие (45) и, кроме того,

$$m^*(n) \underset{n\to\infty}{\sim} c_1 \sqrt{n} \left(\log^{(r)} n\right)^{\varepsilon} \prod_{1\leqslant i\leqslant r} \log^{(i)} n,$$
$$\varphi(x)\log x \equiv \beta(x) = c_2\left[\left[\log^{(r)}(x+d)\right]^{\varepsilon} \prod_{2\leqslant i\leqslant r} \log^{(i)}(x+d)\right]^{-1}, \tag{65}$$

$$\int_m^\infty \frac{\beta(x)}{x \log x} dx \underset{m\to\infty}{\sim} c_3 \left(\log^{(r)} m\right)^{-\varepsilon}, \tag{66}$$

где, по определению, $\prod_{h \leq i \leq m} = 1$ при $k > m$ (случай $r = 1$).

Поскольку при $n \to \infty$ порядок малости в (65) не меньше, чем в (66), то для распределения (64) верхняя и нижняя оценки имеют один и тот же порядок $(\log^{(r)} n)^{-\varepsilon}$. Кроме того, из результатов первого раздела и соотношения (54) следует, что в этих условиях

$$H_1(\varepsilon) \underset{\varepsilon\to 0}{\sim} \psi(\varepsilon)\varepsilon^{-1},$$

где $\psi(\varepsilon) = c_4\varphi(1/\varepsilon)$. Неравенство (44) следует из (59), (63)—(66). Теорема доказана.

Замечание. Из определения $m^*(n)$, (59) и (63) следует, что если функция $\beta(x)$ в (61) медленно меняется при $x \to \infty$ (т. е. энтропия $H_1(\varepsilon)$ близка к критической), то оценка скорости сходимости в (4) при $\mathscr{B} = \mathscr{A}$ для любого дискретного распределения по существу определяется величиной интеграла $\int_{\sqrt{n}}^\infty \beta(x)(x \log x)^{-1} dx$, от которого, кроме существования, ничего не требуется. Так что скорость сходимости в (4) в рассматриваемой общности может быть сколь угодно малой.

ЛИТЕРАТУРА

1. **Dudley R. M.** Limit theorems for empirical measures.— Second Vilnius Conf. on Probab. Theory and Math. Statist., Abstracts of Communic., Vilnius, 1977, v. 3, p. 40—41.
2. **Dudley R. M.** Central limit theorems for empirical measures.— Ann. Probab., 1978, v. 6, N 6, p. 899—929.
3. **Борисов И. С.** К вопросу о точности приближения в центральной предельной теореме для эмпирических мер.— Сиб. мат. журн., 1983, т. 24, № 6, с. 14—25.
4. **Ферник К.** Регулярность траекторий гауссовских случайных функций.— В кн.: Случайные процессы. Выборочные функции и пересечения. М.: Мир, 1978, с. 63—132.
5. **Борисов И. С.** Аппроксимация эмпирических полей, построенных по векторным наблюдениям с зависимыми координатами.— Сиб. мат. журн., 1982, т. 23, № 5, с. 31—41.
6. **Philipp W., Pinzur L.** Almost sure approximation theorems for the multivariate empirical process.— Z. Wahrscheinlichkeitstheorie verw. Geb., 1980, B. 54, N 1, S. 1—13.
7. **Komlos J., Major P., Tusnady G.** An approximation of partial sums of independent RV's and the sample DF.— Z. Wahrscheinlichkeitstheorie verw. Geb., 1975, B. 32, N 1/2, S. 111—133.
8. **Yurinskii V. V.** Exponential inequalities for sums of random vectors.— J. Multivar. Anal., 1976, v. 6, N 4, p. 473—499.

О РАСПРЕДЕЛЕНИИ ВРЕМЕНИ ПРЕБЫВАНИЯ НА ПОЛУОСИ И ПОЛОЖЕНИЯ В ПОСЛЕДНИЙ МОМЕНТ ВРЕМЕНИ ПРОЦЕССА С НЕЗАВИСИМЫМИ ПРИРАЩЕНИЯМИ, УПРАВЛЯЕМОГО ЦЕПЬЮ МАРКОВА

В. С. ЛУГАВОВ

1. Введение

Рассматривается однородный по времени и аддитивный по первой компоненте двумерный марковский процесс

$$\mathscr{L}_t = \{\xi_t, \varkappa_t;\ t \geq 0\}, \quad \xi_0 = 0.$$

Вторая координата $\{\varkappa_t;\ t \geq 0\}$ есть однородный марковский процесс, принимающий значения $1, 2, \ldots, N$, а первая координата $\{\xi_t;\ t \geq 0\}$