V. B. Nevzorov, The rate of convergence of order statistics for nonidentically distributed random variables to the normal law, *Zap. Nauchn. Sem. LOMI*, 1983, Volume 130, 137–146

В.Б.Невзоров

СКОРОСТЬ СХОДИМОСТИ К НОРМАЛЬНОМУ ЗАКОНУ ПОРЯДКОВЫХ СТАТИСТИК ДЛЯ НЕОДИНАКОВО РАСПРЕДЕЛЕННЫХ ВЕЛИЧИН

I. Введение. Пусть $X_{1,n}\leqslant X_{2,n}\leqslant\cdots\leqslant X_{n,n}$ – вариационный ряд, построенный по выборке $X_1, X_2, \ldots, X_n$. В последнее время большое внимание уделяется изучению асимптотических свойств порядковых статистик $X_{k,n}$ и их линейных комбинаций. В случае, когда вариационный ряд построен по выборке из одной генеральной совокупности, т.е. когда $X_1, X_2, \ldots, X_n$ представляют независимые одинаково распределенные случайные величины, получены оценки типа Берри–Эссеена, асимптотические разложения и доказаны некоторые другие предельные теоремы для порядковых статистик и их линейных комбинаций ([I] – [IO] и др.). Менее изучены ситуации, когда выборка $X_1, X_2, \ldots, X_n$ представляет смесь двух или нескольких выборок (например, присутствуют так называемые аномальные наблюдения). В работе [I2] рассматривались выборочные квантили $U_{\tau,N}$, построенные по общей выборке объема $N=n_1+n_2+\cdots+n_k$, представляющей смесь k выборок $X_{j1}, X_{j2}, \ldots, X_{jn_j}$ $(1\leqslant j\leqslant k)$, которым соответствуют различные функции распределения $F_1(x), F_2(x), \ldots, F_k(x)$. В этом случае при некоторых ограничениях на функции распределения доказана асимптотическая нормальность выборочных квантилей $U_{\tau,N}$. В ситуации, когда объемы всех выборок одинаковы ($n_1=n_2=\cdots=n_k=n$ и $N=nk$), были построены асимптотические разложения функций распределения величин $U_{\tau,N}$ ([I2]) и получены теоремы для вероятностей больших уклонений этих величин ([II]). Мы приведем здесь оценки скорости сходимости функций распределения порядковых статистик к нормальному закону в более общей ситуации.

2. Результаты. Рассмотрим последовательность независимых случайных величин $X_1, X_2, \ldots$ с функциями распределения $F_1(x), F_2(x), \ldots$ и плотностями $f_1(x), f_2(x), \ldots$ такими, что

$$\sup_{f_i(x)>0} |f_i'(x)|\leqslant M \quad (i=1,2,\ldots), \tag{I}$$

где $0<M<\infty$ – некоторая константа. Обозначим

$$H(x)=\frac{1}{n}\sum_{i=1}^{n} F_i(x), \quad h(x)=\frac{1}{n}\sum_{i=1}^{n} f_i(x) \tag{2}$$

(точнее было бы обозначить эти суммы через $H_n(x)$ и $h_n(x)$, но мы опускаем индекс для простоты обозначений). Выберем точку $\alpha_{k,n}$ так, чтобы

$$H(\alpha_{k,n})=\frac{k}{n} \quad (1\leqslant k\leqslant n). \tag{3}$$

Пусть $\mathcal{A}_0=\mathcal{A}(n)=\sum_{i=1}^{n} F_i(\alpha_{k,n})(1-F_i(\alpha_{k,n}))$ и $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left\{-\frac{t^2}{2}\right\}dt$.

Приведем основной результат.

ТЕОРЕМА I. Если случайные величины $X_1, X_2, \ldots,$ удовлетворяют условию (I), то для порядковых статистик $X_{k,n}$ $(1\leqslant k\leqslant n)$ справедлива оценка

$$\Delta_n=\sup_{x}\left|P\left\{\frac{(X_{k,n}-\alpha_{k,n})nh(\alpha_{k,n})}{\sqrt{\mathcal{A}(n)}}<x\right\}-\Phi(x)\right|\leqslant C\left(\frac{1}{\sqrt{\mathcal{A}(n)}}+\frac{M\sqrt{\mathcal{A}(n)}}{nh^2(\alpha_{k,n})}\right), \tag{4}$$

где C - некоторая абсолютная константа (здесь и далее мы будем использовать одно и то же обозначение C для различных абсолютных констант).

СЛЕДСТВИЕ I. Если случайные величины $X_1, X_2, \ldots$ одинаково распределены, то из теоремы I следует теорема 3 работы [3].

СЛЕДСТВИЕ 2. Можно получить следующий результат.

ТЕОРЕМА 2. Если $X_1, X_2, \ldots, X_n$ - независимые симметричные случайные величины, удовлетворяющие условию (I), и $X_{(med)}$ - соответствующая выборочная медиана, то

$$\sup_{x}\left|P\left\{2h(0)\sqrt{n}\,X_{(med)}<x\right\}-\Phi(x)\right|\leqslant\frac{C}{\sqrt{n}}\max\left\{1, \frac{M}{h^2(0)}\right\}.$$

Отметим, что теорема 2 при нечетных n непосредственно следует из теоремы I, а при четных n для ее доказательства требуются некоторые простые дополнительные выкладки.

ЗАМЕЧАНИЕ I. Методы данной работы и работы и работы [5] можно использовать для доказательства многомерного аналога теоремы I, т.е. для получения оценки близости вектора $Z=(Z_1, Z_2, \ldots, Z_p)$, где $Z_i=\sqrt{n}(X_{l_i,n}-\alpha_{l_i,n})$ и $0<l_1<l_2<\cdots<l_p<n$ - целые числа, и p-мерного нормального вектора $\xi=(\xi_1,\xi_2,\ldots,\xi_p)$, компоненты которого имеют нулевые математические ожидания и ковариации

$$\Lambda_{ij}=\operatorname{cov}(\xi_i,\xi_j)=\frac{1}{nh(\alpha_{l_i,n})h(\alpha_{l_j,n})}\sum_{t=1}^{n}F_t(\alpha_{l_i,n})(1-F_t(\alpha_{l_j,n}))_{(i\leqslant j)}.$$

3. Обсуждение результатов.

ЗАМЕЧАНИЕ 2. Утверждение теоремы I останется справедливым, если условие (I) заменить условием

$$\max_{1\le i\le n}\ \sup_{f_i(x)>0} |f_i'(x)| \le M(n), \qquad (I')$$

т.е. разрешить величине M меняться с ростом n. Отметим, что в соотношениях (I) и (I′) супремум берется по области $\{x: f(x)>0\}$, а не по всем x, чтобы не исключать из рассмотрения такие, например, распределения, как равномерное или экспоненциальное, у которых могли бы возникнуть трудности с определением производной $f'(x)$ в крайних точках носителя.

ЗАМЕЧАНИЕ 3. Мы не требуем однозначности определения величины $\alpha_{k,n}$ в соотношении (3), так как в ситуациях, когда этой однозначности нет, $h(\alpha_{k,n})=0$ при любом выборе $\alpha_{k,n}$, удовлетворяющем соотношению (3), и неравенство (4) становится тривиальным. В теореме I не приводятся также условия, исключающие равенство $\mathcal{A}(n)=0$ (оно возможно при $k=n$), так как и в этой ситуации неравенство (4) тривиально выполняется. Отметим, что поскольку $\mathcal{A}(n)=\sum_{i=1}^{n} F_i(\alpha_{k,n})(1-F_i(\alpha_{k,n}))\le \min\{\sum_{i=1}^{n} F_i(\alpha_{k,n}),$ $\sum_{i=1}^{n}(1-F_i(\alpha_{k,n}))\}=n\min\{H(\alpha_{k,n}),1-H(\alpha_{k,n})\}=\min\{k,n-k\}$, то правая часть соотношения (4) не меньше, чем $C\left(\frac{1}{\sqrt{k}}+\frac{1}{\sqrt{n-k}}\right)$. Поэтому для крайних членов вариационного ряда, т.е. таких $X_{k,n}$, у которых при $n\to\infty$ фиксировано k или $n-k$, оценка (4) является формальной – она не стремится к нулю с ростом n.

ЗАМЕЧАНИЕ 4. Условия, накладываемые на случайные величины в теореме I, могут быть ослаблены. При некотором $\varepsilon>0$ рассмотрим область $\Delta_{n,\varepsilon}=(\alpha_{k,n}-\varepsilon,\alpha_{k,n}+\varepsilon)$. Потребуем, чтобы в этой области существовали производные $F_i'(x)$ и $F_i''(x)$ $(1\le i\le n)$ и выполнялось условие

$$\sup_{x\in\Delta_{n,\varepsilon}} |F_i''(x)|\le M \quad (1\le i\le n) \qquad (5)$$

В этой ситуации обозначим $h(\alpha_{k,n})=\frac{1}{n}\sum_{i=1}^{n} F_i'(\alpha_{k,n})$. При выполнении этих условий вместо условия (I) соотношение (4) будет иметь место с дополнительным слагаемым в правой части, имеющим вид $C\exp\{-\varepsilon^2 n h^2(\alpha_{k,n})\}$. Отметим, что этим слагаемым можно пренебречь, если, например, $\varepsilon>\ln^{1/2} n/\sqrt{n}\,h(\alpha_{k,n})$. Мы в этом случае можем также записать оценку, несколько ухудшив ее, но зато сделав проще, в следующем виде:

$$\Delta_n \leqslant C\left(\frac{1}{\sqrt{A(n)}} + \frac{M\sqrt{A(n)}}{n h^2(\alpha_{k,n})}\left(1+\frac{1}{\varepsilon\sqrt{M}}\right)\right). \qquad (4')$$

Видим, что если $\varepsilon \geqslant \frac{C}{\sqrt{M}}$, то оценки (4) и (4′) имеют одинаковый порядок. Для того, чтобы использовать соотношение (4′) при $n \to \infty$, мы должны учитывать, что с ростом n точка $\alpha_{k,n}$ может меняться, и в этом случае требуется, чтобы условия на величины $X_1, X_2, \ldots$ выполнялись в области $A_\varepsilon = \bigcup_n \Delta_{n\varepsilon}$. В некоторых ситуациях условия упрощаются.

ПРИМЕР. Пусть функции распределения $F_1(x), F_2(x), \ldots$ таковы, что $F_i(x_q) = q\ (0<q<1)$ при некотором x_q. Такая ситуация может возникнуть, например, при "усечении" элементов выборки на разных уровнях, когда наблюдаются величины $Z_k = \max\{C_k, X_k\}$, где $C_k < x_q\ (k=1,2,\ldots)$ – некоторые константы. Потребуем, чтобы в некоторой окрестности точки x_q функции $F_i(x)$ имели бы равномерно ограниченную вторую производную:

$$\sup_{|x_q - x|<\varepsilon} |F_i''(x)| \leqslant M \qquad (i=1,2,\ldots),$$

и определим $X_{(q)} = X_{[qn]+1,n}$ – выборочную квантиль порядка q. Тогда, несколько изменив доказательство теоремы I, можно получить оценку

$$\sup_x \left| P\left\{ \frac{(X_{(q)} - x_q)\sqrt{n}\, h(x_q)}{\sqrt{q(1-q)}} < x \right\} - \Phi(x) \right| \leqslant C(\varepsilon\sqrt{M}) \frac{1}{\sqrt{n}} \max\left\{ \frac{1}{\sqrt{q(1-q)}}; \frac{M}{h^2(x_q)} \right\}, \qquad (10)$$

где $h(x_q) = \frac{1}{n}\sum_{i=1}^{n} F_i'(x_q)$, а константа $C(\varepsilon\sqrt{M})$ зависит только от произведения $\varepsilon\sqrt{M}$. Этот результат при $q=\frac{1}{2}$ усиливает теорему 2.

4. ДОКАЗАТЕЛЬСТВО теоремы I. По последовательности случайных величин $X_1, X_2, \ldots$ построим, соответственно, величины

$$\xi_i = \xi_i^{(x)} = \begin{cases} 1, & \text{если } X_i < x, \\ 0, & \text{если } X_i \geqslant x \end{cases} \qquad (i=1,2,\ldots)$$

и пусть $S_n = \xi_1 + \xi_2 + \cdots + \xi_n$. Тогда

$$P\{X_{k,n} < x\} = P\{S_n \geqslant k\} = P\left\{ \frac{S_n - ES_n}{\sqrt{\mathcal{D}S_n}} \geqslant \frac{k - ES_n}{\sqrt{\mathcal{D}S_n}} \right\}. \qquad (6)$$

Отметим, что $ES_n = F_1(x) + \cdots + F_n(x)$ и $\mathcal{D}S_n = \sum_{i=1}^{n} F_i(x)(1-F_i(x)) \leqslant \frac{n}{4}$.

Используя неравномерную оценку Бикялиса ([I3]) в центральной предельной теореме, получаем, что

$$P\{X_{k,n} < x\} = 1 - \Phi\left(\frac{k - ES_n}{\sqrt{\mathcal{D}S_n}}\right) + \rho_n ,$$

где

$$|\rho_n| < C/\sqrt{\mathcal{D}S_n}\left(1 + \left|\frac{ES_n - k}{\sqrt{\mathcal{D}S_n}}\right|^3\right) \leqslant \min\left\{\frac{C}{\sqrt{\mathcal{D}S_n}}; \frac{C\mathcal{D}S_n}{|ES_n - k|^3}\right\}. \quad (7)$$

Обозначив $V = x - \alpha_{k,n}$, имеем $H(x) - \frac{k}{n} = H(V + \alpha_{k,n}) - H(\alpha_{k,n}) = Vh(\alpha_{k,n}) + \frac{V^2}{2}h'(\delta)$, где δ – некоторая точка, находящаяся между $\alpha_{k,n}$ и $\alpha_{k,n} + V$
Отсюда получаем, что

$$H(x) - \frac{k}{n} = Vh(\alpha_{k,n})\left(1 + \frac{M\Theta V}{2h(\alpha_{k,n})}\right), \quad (8)$$

где $|\Theta| \leqslant 1$. Аналогично получаем, что

$$\mathcal{D}S_n = \sum_{i=1}^{n} F_i(\alpha_{k,n} + V)(1 - F_i(\alpha_{k,n} + V)) = A_0 + A_1 V + A_2 V^2 + A_3 V^3 + A_4 V^4, \quad (9)$$

где

$$A_0 = A(n) = \sum_{i=1}^{n} F_i(\alpha_{k,n})(1 - F_i(\alpha_{k,n})), \quad A_1 = \sum_{i=1}^{n}(1 - 2F_i(\alpha_{k,n}))f_i(\alpha_{k,n}),$$

$$A_2 = \frac{1}{2}\sum_{i=1}^{n} f_i'(\delta_i)(1 - 2F_i(\alpha_{k,n})) - \sum_{i=1}^{n} f_i^2(\alpha_{k,n}),$$

$$A_3 = -\sum_{i=1}^{n} f_i(\alpha_{k,n}) f_i'(\delta_i), \quad A_4 = -\frac{1}{4}\sum_{i=1}^{n}[f_i'(\delta_i)]^2,$$

и δ_i находятся между $\alpha_{k,n}$ и $\alpha_{k,n} + V$

Из соотношений (8) и (9) следует, что

$$\frac{ES_n - k}{\sqrt{\mathcal{D}S_n}} = \frac{z(1 + B_0 z)}{\sqrt{1 + B_1 z + B_2 z^2 + B_3 z^3 + B_4 z^4}}, \quad (10)$$

где $z = Vh(\alpha_{k,n})n/\sqrt{A_0}$ и

$$B_0 = \frac{M\sqrt{A_0}}{2nh^2(\alpha_{k,n})}, \quad B_1 = \frac{A_1}{\sqrt{A_0}} \cdot \frac{1}{nh(\alpha_{k,n})}, \quad B_2 = \frac{A_2}{h^2(\alpha_{k,n})n^2},$$

$$B_3 = \frac{A_3\sqrt{A_0}}{h^3(\alpha_{k,n})n^3}, \quad B_4 = \frac{A_4 A_0}{h^4(\alpha_{k,n})n^4}.$$

Видим, что для оценки разности

$$|P\{(X_{k,n}-\alpha_{k,n})nh(\alpha_{k,n})/\sqrt{A_0}<z\}-\Phi(z)|=|P\left\{\frac{S_n-ES_n}{\sqrt{\mathcal{D}S_n}}\geqslant\frac{k-ES_n}{\sqrt{\mathcal{D}S_n}}\right\}-\Phi(z)|=$$

$$=\left|\left[P\left\{\frac{S_n-ES_n}{\sqrt{\mathcal{D}S_n}}\geqslant\frac{k-ES_n}{\sqrt{\mathcal{D}S_n}}\right\}-\Phi\left(\frac{ES_n-k}{\sqrt{\mathcal{D}S_n}}\right)\right]+\left[\Phi\left(\frac{ES_n-k}{\sqrt{\mathcal{D}S_n}}\right)-\Phi(z)\right]\right| \quad (11)$$

нужно оценить величину ρ_n и разность

$$R_n=\left|\Phi\left(\frac{z(1+B_0z)}{\sqrt{1+B_1z+B_2z^2+B_3z^3+B_4z^4}}\right)-\Phi(z)\right|.$$

Рассмотрим различные ситуации.

1) Пусть $z\geqslant\frac{1}{8B_0}$. Тогда $x-\alpha_{k,n}=v\geqslant\frac{h(\alpha_{k,n})}{4M}$. Если $|t-\alpha_{k,n}|\leqslant\frac{h(\alpha_{k,n})}{4M}$, то $|h(t)-h(\alpha_{k,n})|\leqslant M|t-\alpha_{k,n}|\leqslant\frac{h(\alpha_{k,n})}{4}$ и $h(t)\geqslant\frac{3}{4}h(\alpha_{k,n})$. Отсюда следует, что

$$H(x)-\frac{k}{n}\geqslant H\left(\alpha_{k,n}+\frac{1}{4M}h(\alpha_{k,n})\right)-H(\alpha_{k,n})\geqslant\frac{3}{16}\frac{h^2(\alpha_{k,n})}{M}$$

В этой ситуации

$$1-\Phi(z)\leqslant 1-\Phi\left(\frac{1}{8B_0}\right)\leqslant CB_0\leqslant C\frac{M\sqrt{A_0}}{nh^2(\alpha_{k,n})} \quad (12)$$

и

$$1-P\left\{\frac{S_n-ES_n}{\sqrt{\mathcal{D}S_n}}\geqslant\frac{k-ES_n}{\sqrt{\mathcal{D}S_n}}\right\}=P\{S_n-ES_n\leqslant-n[H(x)-H(\alpha_{k,n})]\}\leqslant\frac{\mathcal{D}S_n}{n^2(H(x)-H(\alpha_{k,n}))^2} \quad (13)$$

Для оценки правой части соотношения (13) нам потребуется следующий результат.

ЛЕММА. Справедлива оценка $|\mathcal{D}S_n-A_0|\leqslant n|H(x)-H(\alpha_{k,n})|$.

Действительно,

$$|\mathcal{D}S_n-A_0|=\left|\sum_{i=1}^{n}(F_i(x)-F_i(\alpha_{k,n}))-\sum_{i=1}^{n}(F_i^2(x)-F_i^2(\alpha_{k,n}))\right|=$$

$$=\left|n(H(x)-H(\alpha_{k,n}))-\sum_{i=1}^{n}(F_i(x)-F_i(\alpha_{k,n}))(F_i(x)+F_i(\alpha_{k,n}))\right|.$$

Учитывая, что все слагаемые $F_i(x)-F_i(\alpha_{k,n})$ имеют один и тот же знак и $0\leqslant F_i(x)+F_i(\alpha_{k,n})\leqslant 2$, получаем оценку $|\mathcal{D}S_n-A_0|\leqslant n|H(x)-H(\alpha_{k,n})|$.

Из леммы и соотношения (13) получаем, что

$$1-P\left\{\frac{S_n-ES_n}{\sqrt{\mathcal{D}S_n}}\geqslant\frac{k-ES_n}{\sqrt{\mathcal{D}S_n}}\right\}\leqslant C\left[\frac{A_0M^2}{n^2h^4(\alpha_{k,n})}+\frac{M}{nh^2(\alpha_{k,n})}\right] \quad (14)$$

Мы можем считать, что $A_0 > 1$ и $\frac{\sqrt{A_0}M}{nh^2(\alpha_{k,n})} < 1$, так как иначе неравенство (4) тривиально выполняется. Но тогда из соотношений (II),(I2) и (I4) следует, что в нашей ситуации справедливо утверждение теоремы I.

2) В случае $z < -\frac{1}{8B_0}$ предыдущие доказательства требуют лишь незначительных изменений.

3) Рассмотрим величину $R_n = \left|\Phi\left(\frac{z(1+B_0 z)}{\sqrt{1+B_1 z+B_2 z^2+B_3 z^3+B_4 z^4}}\right) - \Phi(z)\right|$ в случае, когда $|B_0 z| \leq \frac{1}{8}$ и $-\frac{1}{2} < C(z) = B_1 z + B_2 z^2 + B_3 z^3 + B_4 z^4 < 1$. Имеем

$$R_n \leq \frac{1}{\sqrt{2\pi}}\left|\frac{z(1+B_0 z)}{\sqrt{1+C(z)}} - z\right| e^{-\frac{z^2}{8}} \leq \frac{|z|}{\sqrt{\pi}} e^{-\frac{z^2}{8}} \left|1+B_0 z - \sqrt{1+C(z)}\right| \leq$$

$$\leq C|z| e^{-\frac{z^2}{2}} \left(|B_0 z| + |C(z)|\right) \leq C\left(B_0 + |B_1| + |B_2| + |B_3| + |B_4|\right). \tag{I5}$$

Оставим оценку для R_n в виде (I5) и перейдем к другим случаям.

4) Пусть $|B_0 z| \leq \frac{1}{8}$ и $C(z) > 1$. Заметим, что

$$1 < C(z) = B_1 z + B_2 z^2 + B_3 z^3 + B_4 z^4 \leq B_1 z + B_2' z^2, \tag{I6}$$

где

$$B_2' = \sum_{i=1}^{n} \frac{f_i'(\delta_i)(1-2F_i(\alpha_{k,n}))}{2h^2(\alpha_{k,n})n^2},$$

так как

$$(B_2 - B_2')z^2 + B_3 z^3 + B_4 z^4 = -\frac{z^2}{h^2(\alpha_{k,n})n^2}\sum_{i=1}^{n}\left[f_i(\alpha_{k,n}) + \frac{z}{2}\frac{f_i'(\delta_i)\sqrt{A_0}}{h(\alpha_{k,n})n}\right]^2 \leq 0.$$

Соотношение (I6) приводит к одной из следующих ситуаций: либо $B_1 z \leq B_2' z^2$ и $B_2' z^2 \geq \frac{1}{2}$, либо $B_1 z \geq B_2' z^2$ и $B_1 z \geq \frac{1}{2}$.

а) Пусть $B_1 z \leq B_2' z^2$ и $B_2' z^2 \geq \frac{1}{2}$. Тогда $|z| \geq \frac{1}{\sqrt{2B_2'}}$ и $1 + C(z) \leq 4B_2' z^2$ (в этом случае $B_2' > 0$). Получаем, что

$$\left|\frac{z(1+B_0 z)}{\sqrt{1+C(z)}}\right| \geq \frac{|z|(1-|B_0 z|)}{2\sqrt{B_2' z^2}} = \frac{1-|B_0 z|}{2\sqrt{B_2'}} \geq \frac{7}{16}\frac{1}{\sqrt{B_2'}}.$$

Если $z > 0$, то $1 - \Phi(z) \leq 1 - \Phi\left(\frac{1}{8B_0}\right) \leq CB_0$,

$1 - \Phi\left(\frac{z(1+B_0 z)}{\sqrt{1+C(z)}}\right) \leq CB_2'$ и

$$R_n \leqslant (1-\Phi(z)) + \left(1-\Phi\left(\frac{z(1+B_0 z)}{\sqrt{1+C(z)}}\right)\right) \leqslant C(B_0 + |B_2'|). \qquad (17)$$

Если же $z<0$, то $\Phi(z) \leqslant \Phi\left(-\frac{1}{8B_0}\right) \leqslant CB_0$,

$\Phi\left(\frac{z(1+B_0 z)}{\sqrt{1+C(z)}}\right) \leqslant \Phi\left(-\frac{7}{16}\frac{1}{\sqrt{B_2'}}\right) \leqslant CB_2'$ и вновь

$$R_n \leqslant \Phi(z) + \Phi\left(\frac{z(1+B_0 z)}{\sqrt{1+C(z)}}\right) \leqslant C(B_0 + |B_2'|). \qquad (18)$$

Из соотношений (17) и (18) следует, что в этой ситуации

$$R_n \leqslant C(B_0 + |B_2'|). \qquad (19)$$

б) Если $B_1 z \geqslant B_2' z^2$ и $B_1 z \geqslant \frac{1}{2}$, то аналогичные рассуждения приводят к неравенству

$$R_n \leqslant C|B_1|. \qquad (20)$$

5) Пусть теперь $|B_0 z| < \frac{1}{8}$ и $C(z) < -\frac{1}{2}$. В этом случае

$$\mathfrak{D} S_n = A_0(1+C(z)) \leqslant \frac{A_0}{2}. \qquad (21)$$

Условие $C(z) < -\frac{1}{2}$ выполняется, если имеет место хотя бы одно из следующих неравенств:

$$B_1 z \leqslant -\frac{1}{8}; \; B_2 z^2 \leqslant -\frac{1}{8}; \; B_3 z^3 \leqslant -\frac{1}{8}; \; B_4 z^4 \leqslant -\frac{1}{8}; \qquad (22)$$

и нетрудно убедиться, что тогда $|v| \geqslant CT$, где

$$T = \min\left\{\frac{A_0}{|A_1|}, \frac{\sqrt{A_0}}{\sqrt{|A_2|}}, \frac{A_0^{1/3}}{|A_3|^{1/3}}, \frac{A_0^{1/4}}{|A_4|^{1/4}}\right\}. \qquad (23)$$

Из условия $|B_0 z| < \frac{1}{8}$ следует, что $|v| < \frac{h(\alpha_{k,n})}{4M}$ и $\left|\frac{Mv}{2h(\alpha_{k,n})}\right| \leqslant \frac{1}{8}$

Оценим величину $\left|H(x) - \frac{k}{n}\right|$. Имеем

$$\left|H(x) - \frac{k}{n}\right| = |v| h(\alpha_{k,n})\left(1 + \frac{M\Theta v}{2h(\alpha_{k,n})}\right) \geqslant CTh(\alpha_{k,n}). \qquad (24)$$

Используя неравенства (21) и (24), получаем в случае, когда $\frac{k}{n} - H(x) > 0$, оценку

$$|P\left\{\frac{S_n - ES_n}{\sqrt{\mathfrak{D}S_n}} \geqslant \frac{k - ES_n}{\sqrt{\mathfrak{D}S_n}}\right\} = P\left\{S_n - ES_n \geqslant n\left(\frac{k}{n} - H(x)\right)\right\} \leqslant \frac{\mathfrak{D}S_n}{n^2\left(H(x) - \frac{k}{n}\right)^2} \leqslant$$

$$\leqslant CA_0 / n^2 T^2 h^2(\alpha_{k,n}) \leqslant C(B_1^2 + |B_2| + |B_3|^{2/3} + |B_4|^{1/2}). \qquad (25)$$

Если $k/n - H(x) > 0$, то $v < 0$ и $z < 0$. Из соотношений (22) следует, что $z < -C\min\{|B_1|^{-1}, |B_2|^{-1/2}, |B_3|^{-1/3}, |B_4|^{-1/4}\}$ и

$$\Phi(z) \leq C(B_1^2 + |B_2| + |B_3|^{2/3} + |B_4|^{1/2}) . \tag{26}$$

Из неравенств (25) и (26) получаем, что

$$R_n \leq C(B_1^2 + |B_2| + |B_3|^{2/3} + |B_4|^{1/2}) . \tag{27}$$

Аналогично выводим неравенство (27) и в случае, когда $\frac{k}{n} - H(x) < 0$. Таким образом, из неравенств (I5),(I9),(20) и (27) следует, что если $|B_0 z| \leq \frac{1}{8}$, то

$$R_n \leq C(B_0 + |B_1| + |B_2| + |B_3| + |B_4| + |B_2'| + B_1^2 + |B_3|^{2/3} + |B_4|^{1/2}) . \tag{28}$$

Оценивая величины, входящие в правую часть неравенства (28), получаем, что

$$R_n \leq C\left[\frac{1}{A_0} + \frac{M\sqrt{A_0}}{nh^2(\alpha_{k,n})} + \frac{\sum_{i=1}^{n} f_i^2(\alpha_{k,n})}{n^2h^2(\alpha_{k,n})}\right] . \tag{29}$$

Нетрудно убедиться, что в условиях теоремы I справедливо неравенство $\sum_{i=1}^{n} f_i^2(\alpha_{k,n}) \leq 2Mn$,
и отсюда следует, что

$$R_n \leq C\left[\frac{1}{A_0} + \frac{M\sqrt{A_0}}{nh^2(\alpha_{k,n})}\right] . \tag{30}$$

Осталось оценить величину ρ_n в случае, когда $|B_0 z| \leq \frac{1}{8}$. Справедливо соотношение (7):

$$|\rho_n| \leq C\min\left\{\frac{1}{\sqrt{\mathcal{D}S_n}} ; \frac{\mathcal{D}S_n}{|ES_n - k|^3}\right\} .$$

Если $\mathcal{D}S_n > \frac{A_0}{2}$, то сразу из неравенства (7) получаем, что

$$|\rho_n| \leq C/\sqrt{A_0} . \tag{31}$$

Будем считать, что $\mathcal{D}S_n \leq A_0/2$. В этом случае $|v| \geq CT$, где T определяется соотношением (23), и справедливо неравенство (24), с помощью которого получаем оценку

$$|\rho_n| \leq C\frac{A_0}{T^3h^3(\alpha_{k,n})n^3} . \tag{32}$$

Из неравенств (31) и (32) следует, что

$$|\rho_n| \leq C\max\left\{\frac{1}{\sqrt{A_0}} + \frac{M^{3/2}}{n^{3/2}h^3(\alpha_{k,n})\sqrt{A_0}}\right\} . \tag{33}$$

Соотношения из (30) и (33) дают нам оценку

$$\Delta_n \leq |\rho_n| + R_n \leq C\max\left\{\frac{1}{\sqrt{A_0}} + \frac{M\sqrt{A_0}}{nh^2(\alpha_{k,n})}\right\} , \tag{34}$$

если учесть, что в условиях, когда

$$\frac{M^{3/2}}{n^{3/2}h^3(\alpha_{k,n})\sqrt{A_0}} \geq \frac{M\sqrt{A_0}}{nh^2(\alpha_{k,n})} ,$$

имеет место неравенство $M\sqrt{A_0}/nh^2(\alpha_{k,n}) \geq A_0^3$, и соотношение (34) тривиально выполняется. Теорема I доказана.

Литература

1. B o o k S.A. Large deviation probabilities for order statistics. – Nav.res.log.quart., 1971, vol.18, N 4, p.521-523.
2. R e i s s P.D. On the accuracy of the normal approximation for quantils. – Ann.Prob., 1974, vol.2, p.741-744.
3. E g o r o v V.A., N e v z o r o v V.B. Limit theorems for linear combinations of order statistics. – Lecture Notes in Mathematics, 1976, vol.550, p.63-79.
4. B j e r v e S. Error bounds for linear combinations of order statistics. – Ann.Stat., 1977, vol.5, p.357-369.
5. Е г о р о в В.А., Н е в з о р о в В.Б. О скорости сходимости совместного распределения порядковых статистик к нормальному закону. – Зап.науч.семин.ЛОМИ, 1977, т.72, с.75-83.
6. Н е в з о р о в В.Б. Большие уклонения для выборочных квантилей. – Зап.науч.семин.ЛОМИ, 1980, т.97, с.144-150.
7. van Z w e t W.R. A strong law for linear functions of order statistics. – Ann.Prob., 1980, vol.8, p.986-990.
8. H e l m e r s R. Edgeworth expancions for linear combinations of order statistics with smooth weight functions. – Ann. Stat., 1980, vol.8, p.1361-1374.
9. H e l m e r s R. A Berry-Esseen theorem for linear combinations of order statictics. – Ann.Prob., 1981, vol.9, p.342-347
10. V e r a v e r b e k e N. Large deviations for linear combinations of order statistictics.– Тезисы докладов III международной Вильнюсской конференции по теории вероятностей и математической статистике, 1981, т.3, с.343-344.
11. C h a n d a K.C. Probabilities of large deviations of sample quantils. – Calcutta Stat.Ass.Bull., 1978, vol.27, p.15-22.
12. C h a n d a K.C. Asymptotic expansion of quantiles computed from mixed samples. – Calcutta Stat.Ass.Bull., 1979, vol.28, p.37-46.
13. Б и к я л и с А. Оценки остаточного члена в центральной предельной теореме. – Лит.мат.сб., 1966, т.6, с.323-346.