Вестник Сыктывкарского университета.
Сер.1.Вып.3.1999

УДК 519.21

Представление Бахадура выборочной квантили для ассоциированной случайной последовательности[1]

С.В.Екишева

Получено представление Бахадура для эмпирической функции распределения в окрестности выборочной квантили для выборки из ассоциированной строго стационарной случайной последовательности. На основании этого представления для выборочных квантилей доказывается асимптотическая нормальность, функциональная центральная предельная теорема, функциональный закон повторного логарифма.

Целью настоящей работы является исследование предельного поведения выборочных квантилей для выборок из ассоциированных строго стационарных последовательностей. Предельные теоремы для квантилей получены на основе так называемого представления Бахадура для эмпирического распределения в окрестности выборочной квантили, что позволяет почти наверное асимптотически выразить выборочную квантиль как сумму асимптотически нормальных ассоциированных случайных величин. Впервые такое разложение получено в [4] для последовательностей независимых одинаково распределенных случайных величин, на нем было основано изящное доказательство центральной предельной теоремы, закона повторного логарифма и других предельных свойств выборочных квантилей. Позднее разложение Бахадура и вытекающие из него предельные теоремы для выборочных квантилей были получены для m-зависимых случайных процессов, не обязательно стационарных ([14]), для стационарных авторегрессионных процессов ([8]), для стационарных последовательностей с ϕ-перемешиванием ([15]) и с α-перемешиванием ([16]).

[1]Работа выполнена при поддержке гранта РФФИ №96-01-00201

1. Формулировка результатов

Пусть $\{X_j\}_{j\in\mathbf{N}}$ – строго стационарная последовательность ассоциированных действительных случайных величин, заданных на одном вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Напомним, что конечный набор случайных величин $\xi = (\xi_1, ..., \xi_m)$ называется ассоциированным, если для любых функций $f, g : \mathbf{R}^m \to \mathbf{R}$, неубывающих по каждому аргументу и таких, что $Ef(\xi)g(\xi)$, $Ef(\xi)$, $Eg(\xi)$ конечны, выполнено неравенство

$$cov\,(f(\xi), g(\xi)) \geq 0.$$

Бесконечное множество случайных величин называется ассоциированным, если любое его конечное подмножество ассоциированно.

Пусть $F(x)$ – общая функция распределения случайных величин X_j, которую мы будем полагать абсолютно непрерывной, $f(x)$ – отвечающая ей плотность. Пусть также $0 \leq X_j \leq 1$, $j \in \mathbf{N}$.

Для $0 < p < 1$ обозначим Θ_p p-квантиль распределения $F(x)$, то есть

$$\Theta_p = \inf\{x \in [0,1] : F(x) = p\}.$$

Для любого натурального n рассмотрим вариационный ряд $X_{n,1} \leq X_{n,2} \leq ... \leq X_{n,1}$, отвечающий случайному вектору $(X_1, X_2, ..., X_n)$.
Определим выборочную p-квантиль $Z_{n,p} = Z_n$ следующим образом: $Z_n = X_{n,r}$, где $r = [np] + 1$.
Введем следующие обозначения:

$$J_n = \{t \in [0,1] : \Theta_p - n^{-\frac{5}{12}} \leq t \leq \Theta_p + n^{-\frac{5}{12}}\},\ n \in \mathbf{N},$$

$F_n(x) = \frac{1}{n}\sum\limits_{i=1}^{n} \mathbf{I}\{X_i \leq x\}$, - эмпирическая функция распределения,

$$\sigma_p^2 = cov(\mathbf{I}\{X_1 \leq \Theta_p\}, \mathbf{I}\{X_1 \leq \Theta_p\}) + 2\sum_{k=2}^{\infty} cov(\mathbf{I}\{X_1 \leq \Theta_p\}, \mathbf{I}\{X_k \leq \Theta_p\});$$

$$Y_n(t) = n^{\frac{1}{2}}(F_n(t) - F(t)),\ t \in [0,1].$$

Тогда, очевидно, $\sigma_p^2 = \lim_{n\to\infty}(nVarF_n(\Theta_p))$.

Теорема 1 (представление Бахадура) *Пусть $\{X_j\}_{j\in\mathbf{N}}$ – строго стационарная последовательность ассоциированных случайных величин, $F(x)$ – абсолютно непрерывная функция, $f(x)$ – отвечающая ей плотность, $0 \leq X_j \leq 1$, $j \in \mathbf{N}$. Пусть для некоторых $p \in (0,1)$, $\nu > 0$ выполнены условия:*

$$\sigma_p^2 > 0, \tag{1}$$

$$\sum_{n=1}^{\infty} n^7 cov(X_1, X_n) < +\infty. \tag{2}$$

Тогда для всех достаточно больших n почти наверное

$$\sup\{|(F_n(t) - F(t)) - (F_n(\Theta_p) - p)| : t \in J_n\} \le Cn^{-\frac{5}{8}} \ln n, \tag{3}$$

всюду далее C, C_i обозначают некоторые положительные константы.

Теорема 2. *Пусть условия* теоремы 1 *выполнены, $f(x)$ дифференцируема в некоторой окрестности точки Θ_p, $0 < f\prime(x) < M < +\infty$ в этой окрестности. Тогда для всех достаточно больших n почти наверное справедливо неравенство*

$$|n^{\frac{1}{2}}(Z_n - \Theta_p)f(\Theta_p) + n^{\frac{1}{2}}(F_n(\Theta_p) - p)| \le Cn^{-\frac{1}{8}} \ln n \tag{4}$$

и при $n \to \infty$

$$\frac{n^{\frac{1}{2}} f(\Theta_p)}{\sigma_p}(Z_n - \Theta_p) \to^d N(0,1). \tag{5}$$

Рассмотрим пространство $C[0,1]$ всех непрерывных на $[0,1]$ функций с равномерной метрикой и определим случайный процесс $\{W_n(t), t \in [0,1]\}$ с траекториями из этого пространства:

$$W_n(0) = 0, W_n(\tfrac{k}{n}) = \tfrac{kf(\Theta_p)(Z_k - \Theta_p)}{\sigma_p\sqrt{n}}, k = 1, ..., n,$$
$$W_n(t) = W_n(\tfrac{k}{n}) + (nt - k)\left(W_n(\tfrac{k+1}{n}) - W_n(\tfrac{k}{n})\right),$$
$$t \in [\tfrac{k}{n}, \tfrac{k+1}{n}],\ k = 0, ..., n-1.$$

Теорема 3 (функциональная центральная предельная теорема для выборочных квантилей) *Если условия* теоремы 2 *выполнены, то при $n \to +\infty$ последовательность случайных процессов $\{W_n(t), t \in [0,1]\}$ слабо сходится к стандартному броуновскому движению в пространстве $C[0,1]$.*

Теорема 4. *Пусть условия* теоремы 2 *выполнены и последовательность $\{N_r\}$ случайных величин, принимающих значения из множества натуральных чисел, такова, что при $r \to +\infty$*

$$r^{-1} N_r \to^P \lambda, \tag{6}$$

где λ – некоторая случайная величина, заданная на том же вероятностном пространстве, что и исходные случайные величины X_j. Тогда при $r \to +\infty$ последовательность распределений процессов

$\{W_{N_r}(t), t \in [0,1]\}$ *на* $C[0,1]$ *слабо сходится к стандартному броуновскому движению.*

В пространстве $D[0,1]$ с равномерной топологией рассмотрим последовательность случайных процессов:

$$H_n(t) = \frac{([nt]+1)f(\Theta_p)(Z_{[nt]+1} - \Theta_p)}{\sigma_p\sqrt{2n\ln\ln n}}, \ t \in [0,1], \ n \geq 3,$$

и шар Штрассена

$$K = \{x(t), t \in [0,1] : x(t) = \int_0^t h(z)dz, \int_0^1 (h(z))^2 dz \leq 1\}.$$

Теорема 5 (функциональный закон повторного логарифма для выборочных квантилей) *Множество предельных точек* $\{H_n, \ n \geq 3\}$ *в* $D[0,1]$ *почти наверное совпадает с* K.

2. Вспомогательные утверждения

Лемма 1([17], обобщение (22.15) [1]) *Пусть* $\{X_j\}_{j\in\mathbf{N}}$ *– строго стационарная последовательность ассоциированных случайных величин, заданных на одном вероятностном пространстве* $(\Omega, \mathcal{F}, P)$, $F(x)$ *– их общая функция распределения, являющаяся непрерывной. Пусть также* $0 \leq X_j \leq 1$, $j \in \mathbf{N}$, $\sigma_p^2 > 0$, *и существует положительное* ν *такое, что справедливо неравенство:*

$$\sum_{n=1}^{\infty} n^{\frac{13}{2}+\nu} cov(X_1, X_n) < +\infty.$$

Тогда

$$\mathbf{E}|Y_n(t) - Y_n(s)|^4 \leq C(n^{-\frac{1}{2}-\frac{\nu}{3}} + |t-s|^{\frac{6}{5}})$$

для всех s, t *из отрезка* $[0,1]$.

Лемма 2 ([1]) *Пусть для последовательности случайных величин* $\{X_j\}_{j\in\mathbf{N}}$ *выполнены все условия* леммы 1. *Тогда для всех* $s \in (0,1)$ *выполнено*

$$\sup_{s\leq t\leq s+mq} |Y_n(t) - Y_n(s)| \leq 3 \max_{1\leq i\leq m} |Y_n(s+iq) - Y_n(s)| + qn^{\frac{1}{2}},$$

где $m \in \mathbf{N}, s + mq < 1$.

В [1] соответствующее неравенство доказывается для последовательностей с ϕ-перемешиванием. Это доказательство дословно переносится на случай ассоциированной последовательности.

Лемма 3 (следствие 1a [10]) *Пусть для некоторого $d \in \mathbf{N}$ заданы случайные величины $\{\xi_n, n \in \mathbf{Z}_+^d\}$, определенные на одном вероятностном пространстве и имеющие конечный момент порядка $\gamma \geq 1$. Пусть также существуют числа $\alpha > 1$ и $\{u_n \geq 0, n \in \mathbf{Z}_+^d\}$ такие, что для любого прямоугольника R из $\mathbf{Z}_+^d$*

$$\begin{aligned} R &= R(b_1, ..., b_d; m_1, ..., m_d) \\ &= \{n = (n_1, ..., n_d) \in \mathbf{Z}_+^d : b_j < n_j \leq b_j + m_j, \ \forall j, j = 1, ..., d\} \end{aligned}$$

выполнено

$$\mathbf{E}\left|\sum_{n_1=b_1+1}^{b_1+m_1} \cdots \sum_{n_d=b_d+1}^{b_d+m_d} \xi_{n_1,\dots,n_d}\right|^{\gamma} \leq \left(\sum_{n\in R} u_n\right)^{\alpha}.$$

Тогда для любого R из $\mathbf{Z}_+^d$ выполнено

$$\mathbf{E}\left(\max\nolimits_{1\leq p_1\leq m_1} \dots \max\nolimits_{1\leq p_d\leq m_d} \left|\sum\nolimits_{n_1=b_1+1}^{b_1+p_1} \cdots \sum\nolimits_{n_d=b_d+1}^{b_d+p_d} \xi_{n_1,\dots,n_d}\right|\right)^{\gamma}$$

$$\leq \left(\tfrac{5}{2}\right)^d \left(1 - 2^{(1-\alpha)/\gamma}\right)^{-d\gamma} \left(\sum\nolimits_{n\in R} u_n\right)^{\alpha}.$$

Лемма 4 (частный случай леммы 1[7], обобщающей лемму 3.1 [6]) *Пусть X, Y – ассоциированные случайные величины, непрерывные функции $h : \mathbf{R} \to \mathbf{R}, g : \mathbf{R} \to \mathbf{R}$ имеют на всей числовой прямой ограниченные частные производные справа и слева, причем эти производные совпадают везде, кроме, может быть, конечного числа точек. Тогда справедливо неравенство:*

$$|cov(h(X), g(Y))| \leq M_1 M_2 cov(X, Y),$$

где

$$M_1 = \max\left\{\sup_{x\in R}\left|\frac{\partial h^+}{\partial x}\right|, \sup_{x\in R}\left|\frac{\partial h^-}{\partial x}\right|\right\},$$

$$M_2 = \max\left\{\sup_{x\in R}\left|\frac{\partial g^+}{\partial x}\right|, \sup_{x\in R}\left|\frac{\partial g^-}{\partial x}\right|\right\}.$$

Лемма 5 (замечание 3 [3], обобщение оценки [5]) *Пусть $\{\eta_j\}_{j\in\mathbf{N}}$ ассоциированная последовательность случайных величин, $\mathbf{E}\eta_j = 0$, $j \in \mathbf{N}$ и существуют числа $r > 2$, $\beta > 0$, $\nu \geq 0$ такие, что выполнены следующие условия:*

$$\sup_{k\in\mathbf{N}} \mathbf{E}|\eta_k|^{r+\beta} < +\infty,$$

$$u(n) = \sup_{k \in \mathbf{N}} \sum_{m:|m-k| \geq n} cov(\eta_k, \eta_m) = O(n^{-\nu}).$$

Тогда

$$\sup_{m \in \mathbf{N} \cup 0} \mathbf{E} \left| \sum_{k=m+1}^{m+n} \eta_k \right|^r = O\left(n^{\gamma(r,\beta,\nu)}\right),$$

где

$$\gamma(r,\beta,\nu) = \begin{cases} r - \beta(1+\nu)(r+\beta-2)^{-1}, \ 0 \leq \nu < \nu_0, \\ \frac{r}{2}, \ \nu \geq \nu_0, \end{cases}$$

а величина $\nu_0 = \frac{(r+\beta)(r-2)}{2\beta}$.

3. Доказательства

3.1. Доказательство теоремы 1

Для каждого натурального n определим события

$$A_n = \{\omega : \sup_{t \in J_n} |(F_n(t) - F(t)) - (F_n(\Theta_p) - p)| \geq 4n^{-\frac{5}{8}} \ln n\}$$
$$= \{\omega : \sup_{t \in J_n} |Y_n(t) - Y_n(\Theta_p)| \geq 4n^{-\frac{1}{8}} \ln n\}.$$

Введем числовую последовательность $\{n_k\}_{k \in \mathbf{N}}$, $n_k = \exp k$.
Зафиксируем $k \in \mathbf{N}$ и положим $q_k = n_{k+1}^{-\frac{5}{8}}$, $m_k = [n_{k+1}^{\frac{5}{24}}] + 1$.
Тогда, в силу леммы 2 и выбора q_k, m_k справедлива следующая цепочка неравенств:

$$P(\bigcup_{n_k \leq n < n_{k+1}} A_n) \leq P(\max_{n_k \leq n < n_{k+1}} \sup_{t \in J_n} |Y_n(t) - Y_n(\Theta_p)| \geq 4n_{k+1}^{-\frac{1}{8}} \ln n_k)$$

$$\leq P(\max_{n_k \leq n < n_{k+1}} \sup_{\Theta_p < t \leq \Theta_p + n^{-\frac{5}{12}}} |Y_n(t) - Y_n(\Theta_p)| \geq 4n_{k+1}^{-\frac{1}{8}} \ln n_k)$$

$$+P(\max_{n_k \leq n < n_{k+1}} \sup_{\Theta_p - n^{-\frac{5}{12}} \leq t < \Theta_p} |Y_n(t) - Y_n(\Theta_p)| \geq 4n_{k+1}^{-\frac{1}{8}} \ln n_k)$$

$$\leq P(\max_{n_k \leq n < n_{k+1}} \max_{1 \leq l \leq m_k} |Y_n(\Theta_p + lq_k) - Y_n(\Theta_p)| \geq n_{k+1}^{-\frac{1}{8}} \ln n_k)$$

$$+P(\max_{n_k \leq n < n_{k+1}} \max_{1 \leq l \leq m_k} |Y_n(\Theta_p - lq_k) - Y_n(\Theta_p)| \geq n_{k+1}^{-\frac{1}{8}} \ln n_k). \quad (7)$$

Определим для $0 \le s < t \le 1$, $i \in \mathbf{N}$, $j = 1, ..., m_k$ случайные величины

$$\xi_i(s,t) = (\mathbf{I}\{X_i \le t\} - F(t)) - (\mathbf{I}\{X_i \le s\} - F(s));$$
$$\xi_{i,j} = \xi_i(\Theta_p + (j-1)q_k, \Theta_p + jq_k),$$
$$\eta_{i,j} = \xi_i(\Theta_p - jq_k, \Theta_p - (j-1)q_k),$$

и для натуральных $j > m_k$ определим $\xi_{i,j} = 0$, $\eta_{i,j} = 0$. Тогда для всех $i \in \mathbf{N}$, $l = 1, ..., m_k$ справедливы следующие соотношения:

$$|Y_n(\Theta_p + lq_k) - Y_n(\Theta_p)| = \frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n}\xi_i(\Theta_p, \Theta_p + lq_k)\right| = \frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n}\sum_{j=1}^{l}\xi_{i,j}\right|,$$

$$|Y_n(\Theta_p - lq_k) - Y_n(\Theta_p)| = \frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n}\xi_i(\Theta_p - jq_k, \Theta_p)\right| = \frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n}\sum_{j=1}^{l}\eta_{i,j}\right|.$$

Продолжая цепочку неравенств (7), получим

$$P(\bigcup_{n_k \le n < n_{k+1}} A_n) \le P(\max_{n_k \le n < n_{k+1}} \max_{1 \le l \le m_k} |\sum_{j=1}^{n}\sum_{i=1}^{l}\xi_{i,j}| \ge n_{k+1}^{\frac{3}{8}} \ln n_k)$$

$$+P(\max_{n_k \le n < n_{k+1}} \max_{1 \le l \le m_k} |\sum_{j=1}^{n}\sum_{i=1}^{l}\eta_{i,j}| \ge n_{k+1}^{\frac{3}{8}} \ln n_k).$$

Обратимся к обозначениям леммы 3. Рассмотрим прямоугольник $R = (i_1, j_1; i_2, j_2)$, $n_k \le i_2 - i_1 < n_{k+1}$. Если $j_1 \ge m_k$, тогда

$$\mathbf{E}(\sum_{i=i_1+1}^{i_2}\sum_{j=j_1+1}^{j_2}\xi_{i,j})^4 = 0.$$

Если $j_2 \le m_k$, тогда в силу строгой стационарности последовательности $\{X_i\}$ и леммы 1 (с $\nu = \frac{1}{2}$) справедливо

$$\mathbf{E}(\sum_{i=i_1+1}^{i_2}\sum_{j=j_1+1}^{j_2}\xi_{i,j})^4 = \mathbf{E}(\sum_{i=1}^{i_2-i_1}\sum_{j=j_1+1}^{j_2}\xi_{i,j})^4 = \mathbf{E}\left(\sum_{i=1}^{i_2-i_1}\sum_{j=1}^{j_2}\xi_{i,j} - \sum_{i=1}^{i_2-i_1}\sum_{j=1}^{j_1}\xi_{i,j}\right)^4$$

$$= (i_2 - i_1)^2\mathbf{E}|Y_{i_2-i_1}(\Theta_p + j_2q_k) - Y_{i_2-i_1}(\Theta_p + j_1q_k)|^4$$

$$\le C(i_2 - i_1)^2((i_2 - i_1)^{-\frac{2}{3}} + (j_2 - j_1)^{\frac{6}{5}}n_{k+1}^{-\frac{3}{4}}) \le 2C(i_2 - i_1)^{\frac{4}{3}}(j_2 - j_1)^{\frac{6}{5}}$$

$$= \left((2C)^{\frac{5}{6}}(i_2 - i_1)^{\frac{10}{9}}(j_2 - j_1)\right)^{\frac{6}{5}} \le (\sum_{i,j \in R} u_{i,j})^{\frac{6}{5}},$$

с $u_{i,j} = (2C)^{\frac{5}{6}} i^{\frac{1}{9}}$. Наконец, для прямоугольников R с $j_1 \le m_k$, $j_2 > m_k$ путем таких же рассуждений получаем

$$\mathbf{E}(\sum_{i=i_1+1}^{i_2} \sum_{j=j_1+1}^{j_2} \xi_{i,j})^4 = \mathbf{E}(\sum_{i=i_1+1}^{i_2} \sum_{j=j_1+1}^{m_k} \xi_{i,j})^4 \le \left(\sum_{i=i_1}^{i_2} \sum_{j=j_1}^{m_k} u_{i,j} \right)^{\frac{6}{5}} \le \left(\sum_{i,j \in R} u_{i,j} \right)$$

Аналогично оцениваются $\mathbf{E}(\sum_{i,j\in R} \eta_{i,j})^4$. Таким образом, из леммы 3 будет следовать:

$$P(\bigcup_{n_k \le n < n_{k+1}} A_n) \le \frac{\mathbf{E}(\max\limits_{n_k \le n < n_{k+1}} \max\limits_{1 \le l \le m_k} |\sum_{j=1}^{n} \sum_{i=1}^{l} \xi_{i,j}|)^4}{(\ln n_k)^4 n_{k+1}^{\frac{3}{2}}}$$

$$+ \frac{\mathbf{E}(\max\limits_{n_k \le n < n_{k+1}} \max\limits_{1 \le l \le m_k} |\sum_{j=1}^{n} \sum_{i=1}^{l} \eta_{i,j}|)^4}{(\ln n_k)^4 n_{k+1}^{\frac{3}{2}}} \le C_1 \frac{(\sum_{i=1}^{n_{k+1}} \sum_{j=1}^{m_k} u_{i,j})^{\frac{6}{5}}}{(\ln n_k)^4 n_{k+1}^{\frac{3}{2}}}$$

$$\le C_2 \ln^{-4} n_k = C_2 k^{-4}$$

при этом константы C_1, C_2 не зависят от k. Отсюда вытекает, что

$$\sum_{k=1}^{\infty} P(\bigcup_{n_k \le n < n_{k+1}} A_n) \le C_2 \sum_{k=1}^{\infty} k^{-4} < +\infty.$$

Таким образом, по лемме Бореля-Кантелли, для всех достаточно больших натуральных n с вероятностью 1 выполнено (3). Теорема 1 доказана.

3.2. Доказательство теоремы 2

Сначала докажем (4). Для этого определим события B_n, $n \in \mathbf{N}$:

$$B_n = \left\{ Z_n \le \Theta_p - n^{-\frac{1}{2}} \ln n \right\}$$

$$= \left\{ \text{не менее } r \text{ из } X_i,\ i = 1, ..., n, \text{ не превосходят } \Theta_p - n^{-\frac{1}{2}} \ln n \right\}$$

$$= \left\{ \textstyle\sum_{j=1}^{n} \mathbf{I}\{X_j \le \Theta_p - n^{-\frac{1}{2}} \ln n\} \ge r \right\}$$

$$= \left\{ \frac{1}{n} \sum_{j=1}^{n} \left(\mathbf{I}\{X_j \le \Theta_p - n^{-\frac{1}{2}} \ln n\} - F(\Theta_p - n^{-\frac{1}{2}} \ln n) \right) \ge \right.$$

$$\left. \ge \frac{r}{n} - F(\Theta_p - n^{-\frac{1}{2}} \ln n) \right\}.$$

Зафиксируем n и рассмотрим случайные величины

$$\eta_j = \mathbf{I}\{X_j \le \Theta_p - n^{-\frac{1}{2}} \ln n\} - F(\Theta_p - n^{-\frac{1}{2}} \ln n).$$

Заметим, что $\{-\eta_j\}$ - набор ассоциированных случайных величин. Убедимся в том, что для $\{-\eta_j\}$ выполнены условия леммы 5. Действительно, $\mathbf{E}(-\eta_j) = 0, \forall j \in \mathbf{N}$,

$$\sup_{j \in \mathbf{N}} \mathbf{E}|-\eta_j|^{r+\beta} < +\infty, \ \forall r, \beta > 0,$$

поскольку $|-\eta_j| \leq 2$.

Проверим теперь, что ряд $\sum\limits_{k=1}^{\infty} cov(\eta_1, \eta_k)$ мажорируется сходящимся рядом $C_3 \sum_{m=1}^{\infty} cov^{\frac{1}{3}}(X_1, X_m)$. Определим для каждого $k \in \mathbf{N}$ функцию $h_k(x)$ следующим образом:

$$h_k(x) = \begin{cases} 1, \ x \leq \Theta_p - n^{-\frac{1}{2}} \ln n, \\ 0, \ x > \Theta_p - n^{-\frac{1}{2}} \ln n + a_k, \\ 1 - \frac{x - \Theta_p + n^{-\frac{1}{2}} \ln n}{a_k}, \ \Theta_p - n^{-\frac{1}{2}} \ln n < x \leq \Theta_p - n^{-\frac{1}{2}} \ln n + a_k, \end{cases}$$

последовательность положительных чисел $\{a_k\}$ зададим позднее. Тогда из леммы 4 для любого k и всех достаточно больших n с учетом условия на плотность $f(x)$ получим неравенство

$$\begin{aligned} &|cov(\mathbf{I}\{X_1 \leq \Theta_p - n^{-\frac{1}{2}} \ln n\}, \mathbf{I}\{X_k \leq \Theta_p - n^{-\frac{1}{2}} \ln n\})| \\ &\leq |cov(h_k(X_1), h_k(X_k))| + 3Ma_k \leq a_k^{-2} cov(X_1, X_n) + 3Ma_k. \end{aligned} \tag{8}$$

Если $cov(X_1, X_k) = 0$, то $cov(\eta_1, \eta_k) = 0$, поскольку (8) выполнено для любого $a_k > 0$. Если $cov(X_1, X_k) \neq 0$, то выберем $a_k = cov^{\frac{1}{3}}(X_1, X_k)$ и в силу (8) получим для всех номеров k:

$$|cov(\eta_1, \eta_k)| \leq C_3 cov^{\frac{1}{3}}(X_1, X_k).$$

Таким образом, в силу неравенства Гельдера и условия (2) имеем для любого $m \in \mathbf{N}$:

$$\sum_{k=1}^{m} cov(\eta_1, \eta_k) \leq C_3 \sum_{k=1}^{m} k^{\frac{7}{3}} k^{-\frac{7}{3}} cov^{\frac{1}{3}}(X_1, X_k)$$

$$\leq C_3 [\sum_{k=1}^{m} k^7 cov(X_1, X_k)]^{\frac{1}{3}} [\sum_{k=1}^{m} k^{-\frac{7}{2}}]^{\frac{2}{3}} \leq C_4.$$

Следовательно, ряд $\sum_{k=1}^{\infty} cov(\eta_1, \eta_k)$ сходится. Аналогично, для остатка ряда получаем

$$\begin{aligned} &\sum_{k=m}^{\infty} cov(\eta_1, \eta_k) \leq C_3 \sum_{k=m}^{\infty} cov^{\frac{1}{3}}(X_1, X_k) \\ &\leq C_3 [\sum_{k=m}^{\infty} k^7 cov(X_1, X_k)]^{\frac{1}{3}} [\sum_{k=m}^{\infty} k^{-\frac{7}{2}}]^{\frac{2}{3}} \leq (C_5 m^{-\frac{5}{2}})^{\frac{2}{3}} \leq C_6 m^{-\frac{3}{2}}. \end{aligned} \tag{9}$$

Выберем $r = \frac{5}{2}, \beta = \frac{1}{2}$. Тогда (9) влечет выполнение условий леммы 5 с $\nu_0 = \frac{3}{2}$. Итак, можно утверждать, что

$$\sup_{m \geq 0} \mathbf{E} \left| \sum_{k=m+1}^{m+n} \eta_k \right|^{\frac{5}{2}} \leq C_7 n^{\frac{5}{4}}. \quad (10)$$

Заметим далее, что

$$\frac{r}{n} - F(\Theta_p - n^{-\frac{1}{2}} \ln n) = F(\Theta_p) - F(\Theta_p - n^{-\frac{1}{2}} \ln n) = f(\Theta_p) n^{-\frac{1}{2}} \ln n (1 + o(1)) \quad (11)$$

Покажем теперь, что для достаточно больших n почти наверное

$$Z_n > \Theta_p - n^{-\frac{1}{2}} \ln n. \quad (12)$$

Заметим, что в силу (11) события B_n можно переписать

$$B_n = \{Z_n \leq \Theta_p - n^{-\frac{1}{2}} \ln n\} = \left\{ \sum_{i=1}^{n} \eta_i \geq \tau_n f(\Theta_p) \sqrt{n} \ln n \right\},$$

здесь числовая последовательность $\tau_n \to 1$ при $n \to \infty$. Обозначим $\tau = \inf_{n > 1}\{\tau_n\}$, при этом в силу (11) и условия на плотность $f(x)$ выполнено неравенство $\tau > 0$.

Заметим, что (10) представляет собой условие леммы 3 для последовательности $\{\eta_i,\ i \in \mathbf{N}\}$ при $d = 1$, $\gamma = \frac{5}{2}$, $\alpha = \frac{5}{4}$, $u_i = (C_7)^{\frac{4}{5}}$, $i \in \mathbf{N}$. Пусть числовая последовательность $\{n_k\}$ такая же, как в доказательстве теоремы 1. Тогда из леммы 3 и (11) следует справедливость цепочки неравенств

$$P\Big(\bigcup_{n_k \leq n < n_{k+1}} B_n\Big) \leq P\Big(\max_{n_k \leq n < n_{k+1}} \sum_{i=1}^{n} \eta_i \geq \tau f(\Theta_p) \sqrt{n}_k \ln n_k\Big)$$

$$\leq C_8 n_{k+1}^{\frac{5}{4}} (f(\Theta_p) \tau \ln n_k)^{-\frac{5}{2}} n_k^{-\frac{5}{4}} \leq C_9 k^{-\frac{5}{2}},$$

при этом константа C_9 не зависит от k. Отсюда

$$\sum_{k=1}^{\infty} P\Big(\bigcup_{n_k \leq n < n_{k+1}} B_n\Big) \leq C_9 \sum_{k=1}^{\infty} k^{-\frac{5}{2}} < +\infty,$$

и из леммы Бореля-Кантелли получаем (12).

Аналогично доказывается, что для достаточно больших n почти наверное выполнено

$$Z_n \leq \Theta_p + n^{-\frac{1}{2}} \ln n. \quad (13)$$

Таким образом, из (12) и (13) получаем для всех достаточно больших n с вероятностью 1:

$$|Z_n - \Theta_p| < n^{-\frac{1}{2}} \ln n. \qquad (14)$$

Замечая, что

$$F_n(Z_n) = \frac{r}{n} = p + O(n^{-1}),$$

и используя условия на производную плотности распределения в некоторой окрестности точки Θ_p и (14), получаем для достаточно больших n почти наверное:

$$\begin{aligned} &|F_n(Z_n) - F(Z_n) + f(\Theta_p)(Z_n - \Theta_p)| \\ &= |p + O(n^{-1}) - p - f(\Theta_p)(Z_n - \Theta_p) - \tfrac{1}{2}f\prime(\Delta_n)(Z_n - \Theta_p)^2 \\ &+ f(\Theta_p)(Z_n - \Theta_p)| \leq C_{10} n^{-1} \ln^2 n, \end{aligned} \qquad (15)$$

здесь Δ_n - случайные величины такие, что $\Theta_p < \Delta_n < Z_n$ или $\Theta_p > \Delta_n > Z_n$.
Из (15) и теоремы 1 следует, что почти наверное выполнено

$$\begin{aligned} &|n^{\frac{1}{2}} f(\Theta_p)(Z_n - \Theta_p) + n^{\frac{1}{2}}(F_n(\Theta_p) - p)| \leq |n^{\frac{1}{2}} f(\Theta_p)(Z_n - \Theta_p) \\ &- n^{\frac{1}{2}}(F_n(Z_n) - F(Z_n))| + |n^{\frac{1}{2}}(F_n(\Theta_p) - p) - n^{\frac{1}{2}}(F_n(Z_n) - F(Z_n))| \\ &\leq C n^{-\frac{1}{8}} \ln n \end{aligned} \qquad (16)$$

для достаточно больших n, поскольку $Z_n \in J_n$ для достаточно больших n почти наверное. Это доказывает (4).

Для доказательства (5) заметим, что из (16) вытекает:

$$\left| n^{\frac{1}{2}} f(\Theta_p)(Z_n - \Theta_p) - n^{\frac{1}{2}}(p - F_n(\Theta_p)) \right| \leq C n^{-\frac{1}{8}} \ln n \qquad (17)$$

почти наверное для достаточно больших n, а случайные величины $p - F_n(\Theta_p)$ для любого натурального n представляют собой нормированные суммы ассоциированных центрированных строго стационарных случайных величин $p - \mathbf{I}\{X_i \leq \Theta_p\}$, для которых выполнены все условия центральной предельной теоремы ([11]). Сходимость ряда их ковариаций $\sum_{n=1}^{\infty} cov(\mathbf{I}\{X_1 \leq \Theta_p\}, \mathbf{I}\{X_n \leq \Theta_p\})$ доказывается так же, как доказывалась сходимость ряда $\sum_{k=1}^{\infty} cov(\eta_1, \eta_k)$ при проверке условий теоремы 1. Значит, при $n \to \infty$

$$\frac{n^{\frac{1}{2}}(p - F_n(\Theta_p))}{\sigma_p} \to^d N(0, 1).$$

Из (17) следует, что последовательность распределений случайных ве
личин $n^{\frac{1}{2}}f(\Theta_p)(Z_n-\Theta_p)$ имеет тот же слабый предел, то есть выполнен

$$\frac{n^{\frac{1}{2}}f(\Theta_p)}{\sigma_p}(Z_n-\Theta_p)\to^d N(0,1)$$

при $n\to\infty$. Это оканчивает доказательство теоремы 2.

3.3. Доказательство теоремы 3

Рассмотрим еще одну последовательность случайных процессо
$\{W_n^*(t), t\in[0,1], n\in N\}$, определенных следующим образом:

$$W_n^*(0)=0, W_n^*(\tfrac{k}{n})=\tfrac{k(p-F_k(\Theta_p))}{\sigma_p\sqrt{n}}, k=1,...,n$$

$$W_n^*(t)=W_n^*(\tfrac{k}{n})+(nt-k)\left(W_n^*(\tfrac{k+1}{n})-W_n^*(\tfrac{k}{n})\right),$$

$$t\in[\tfrac{k}{n},\tfrac{k+1}{n}],\ k=0,...n-1.$$

Для случайных процессов $\{W_n^*(t)\}$ выполнены условия функци
ональной предельной теоремы [12], поскольку случайные величин
$p-\mathbf{I}\{X_i\le\Theta_p\}$ ассоциированные, центрированные, образуют строг
стационарную последовательность, и ряд их ковариаций сходится. Зна
чит, при $n\to\infty$ имеет место слабая сходимость вероятностных мер
пространстве $C[0,1]$:

$$W_n^*(t)\to^d W, \qquad (18$$

где W - стандартное броуновское движение. Докажем, что при $n\to$ о

$$\rho(W_n,W_n^*)\to^P 0. \qquad (19$$

Для этого выберем числовую последовательность $\{k_n\}, k_n\in N$ так
чтобы $k_n\to+\infty$, $k_nn^{-\frac{1}{2}}\to 0$ при $n\to\infty$. Имеем

$$\rho(W_n,W_n^*)\le \sup_{0\le t\le\frac{k_n}{n}}|W_n(t)|$$

$$+\sup_{0\le t\le\frac{k_n}{n}}|W_n^*(t)|+\sup_{\frac{k_n}{n}\le t\le 1}|W_n(t)-W_n^*(t)|.$$

При этом почти наверное

$$\sup_{0\le t\le\frac{k_n}{n}}|W_n(t)|=\max_{1\le k\le k_n}\left|\tfrac{kf(\Theta_p)(Z_k-\Theta_p)}{\sigma_p\sqrt{n}}\right|$$

$$\le\tfrac{k_nf(\Theta_p)}{\sigma_p\sqrt{n}}\max_{1\le k\le k_n}|Z_k-\Theta_p|\le\tfrac{2k_nf(\Theta_p)}{\sigma_p\sqrt{n}}\to 0,\ n\to\infty.$$

Далее, из (18) следует, что $\{W_n^*(t)\}$ является плотным семейством ме
на $C[0,1]$ и, следовательно, при $n\to\infty$

$$\sup_{0\le t\le\frac{k_n}{n}}|W_n^*(t)|\to^P 0.$$

Наконец , из задания процессов $\{W_n^*(t)\}$ и $\{W_n(t)\}$ и (4) получаем оценку:

$$\sup_{\frac{k_n}{n}\le t\le 1} |W_n(t) - W_n^*(t)| =$$

$$= \max_{k_n\le k\le 1}\left|\frac{kf(\Theta_p)(Z_k-\Theta_p)}{\sigma_p\sqrt{n}} + \frac{k(F_k(\Theta_p)-p)}{\sigma_p\sqrt{n}}\right| \le C_{12}n^{-\frac{1}{8}}\ln n$$

для достаточно больших n почти наверное. Тогда (19) выполнено, и слабые пределы $\{W_n^*(t)\}$ и $\{W_n(t)\}$ совпадают. Это оканчивает доказательство теоремы 3.

3.4. Доказательство теоремы 4

Отметим, что утверждение (18) означает справедливость принципа инвариантности для последовательности ассоциированных случайных величин $\{p - \mathbf{I}\{X_i \le \Theta\}, i \in N\}$. При его доказательстве ([12]) было установлено, что $\forall \epsilon, \eta > 0$ существуют $\delta > 0$ и n_0 такие, что для всех $n > n_0$ выполнено:

$$P\{\omega_\delta(W_n^*) > \epsilon\} < \eta,$$

где ω_δ - модуль непрерывности семейства функций $\{W_n^*(t,\omega) : \omega \in \Omega\}$. Тогда из (19) следует, что при $n > n_0$ имеет место неравенство:

$$P\{\omega_\delta(W_n) > 3\epsilon\} < 2\eta. \qquad (20)$$

Из (5) и теоремы 2 [7] следует асимптотическая нормальность выборочных квантилей, когда объем выборки является случайной величиной:

$$\frac{\sqrt{N_r}f(\Theta_p)}{\sigma_p}(Z_{N_r} - \Theta_p) \to^d N(0,1) \qquad (21)$$

при $r \to \infty$.

С учетом (20) и (21) утверждение теоремы получается, как в доказательстве теоремы 20.3 [1] (аналогичный результат для последовательностей с ϕ-перемешиванием.)

3.5. Доказательство теоремы 5

Рассмотрим еще одну последовательность случайных процессов:

$$H_n^*(t) = \frac{([nt]+1)(p - F_{[nt]+1}(\Theta_p))}{\sigma_p\sqrt{2n\ln\ln n}}, t \in [0,1],\ n \ge 3.$$

Для этого стационарной последовательности случайных величин $p - \mathbf{I}\{X_i \le \Theta\}$ выполнены условия функционального закона повторного логарифма [2], поскольку они являются ассоциированными, центрированными, $\mathbf{E}|p - \mathbf{I}\{X_i \leq \Theta\}|^3 < \infty$ и для этой последовательности коэффициент Кокса-Гримметта в силу (9) удовлетворяет соотношению

$$u(n) \leq C_{13} n^{-\frac{3}{2}}.$$

Значит, множество предельных точек семейства $\{H_n^*,\ n \geq 3\}$ в $D[0,1]$ с равномерной метрикой почти наверное есть множество K.
Покажем, что при $n \to \infty$ почти наверное выполнено:

$$\rho(H_n, H_n^*) \to 0. \qquad (22)$$

Действительно, из (17) следует, что существует $C_{14} < +\infty$ такое, что с вероятностью 1 для всех $k \in \mathbf{N}$

$$\left|\sqrt{k} f(\Theta_p)(Z_k - \Theta_p) - \sqrt{k}(p - F_k(\Theta_p))\right| \leq C_{14}.$$

Тогда выполнено почти наверное

$$\rho(H_n, H_n^*) = \sup_{0 \leq t \leq 1} |H_n(t) - H_n^*(t)|$$
$$\leq \frac{1}{\sqrt{2\sigma_p^2 \ln\ln n}} \sup_{0 \leq t \leq 1} \Big|\sqrt{[nt]+1} f(\Theta_p)(Z_{[nt]+1} - \Theta_p) -$$
$$-\sqrt{[nt]+1}(p - F_{[nt]+1}(\Theta_p))\Big|$$
$$\leq \frac{1}{\sqrt{2\sigma_p^2 \ln\ln n}} \sup_{k \in \mathbf{N}} \left|\sqrt{k} f(\Theta_p)(Z_k - \Theta_p) - \sqrt{k}(p - F_k(\Theta_p))\right|$$
$$\leq C_{15} (\ln\ln n)^{-\frac{1}{2}},$$

что доказывает (22). Таким образом, семейства $\{H_n,\ n \geq 3\}$ и $\{H_n^*,\ n \geq 3\}$ в равномерной топологии на $D[0,1]$ почти наверное имеют одинаковое множество предельных точек. Теорема доказана.

Литература

1. Биллингсли П. Сходимость вероятностных мер: Пер. с англ. М.: Наука, 1977. 352 с.

2. Булинский А.В., Функциональный закон повторного логарифма для ассоциированных случайных полей//*Фундаментальная и прикладная математика. 1995. Т.1. №3. С.623–639.*

3. Булинский А.В. Неравенства для моментов сумм ассоциированных мультииндексированных величин//*Теория вероятностей и ее применения. 1993. Т.38. Вып.2. С. 417–425.*

4. Bahadur R.R. A note on quantiles in the large samples//*Ann. Math. Statist. 1966. V.37. P. 577–580.*

5. Birkel T. Moment bounds for associated sequences//*Ann.Probab. 1988. V.16. №3. P. 1184–1193*

6. Birkel T. On the convergence rate in the central limit theorem for associated processes//*Ann. Probab. 1988. V.16. №4. P.1685–1698.*

7. Bulinski A.V. On the convergence rates in the CLT for positively and negatively dependent random fields//*In: Probability Theory and Mathematical statistics. Edited by I.A.Ibragimov and A.Yu.Zaitsev. Gordon and Breach Publishers. 1996. P.3–15.*

8. Dutta K., Sen P.K. On the Bahadur prepresentation of sample quantiles in some stationary multivariate autoregressive processes//*J. Multivariate Analysis. 1971. V.1. P.186–198.*

9. Mogyorodi J. Limit distributions for sequences of random variables with random indices//*4-th Prague Conference Infor.The.Statis.Dec.Fns Random Proc. 1965. P. 462–470.*

10. Moricz F. A general moment inequality for the maximum of the rectangular partial sums of multiple series//*Acta Math.Hung. 1983. V.41. P.337–346.*

11. Newman C.M. Normal fluctuations and the FKG-inequalities. *Commun.Math.Phys. 1980. V.74. P. 119–128.*

12. Newman C.M.,Wright A.L. An invariance principle for certain dependent sequences//*Ann. Probab. 1981. V.9. P. 671–675.*

13. Newman C.M. Asymptotic independence and limit theorems for positively and negatively dependent random variables//*In: Inequalities in Statistics and Probability. Ed. By Y.L.Tong. Harvard: IMS. 1984. P. 127–140.*

14. Sen P.K. Asymptotic normality of sample quantiles for m-dependent processes//*Ann. Math. Statist. 1968. V.39. P. 1724–1730.*

15. Sen R.K. On the Bahadur representation of sample quantiles for sequences of ϕ-mixing random variables//*J. Multivariate Anal. 1972. V.2. P. 77–95.*

16. Yoshihara K.-i. The Bahadur representation of sample quantiles for sequences of strongly mixing random variables//*Statistics and Probability Letters. 1995. V.24. P. 299–304.*

17. Yu H. A Glivenko-Cantelli lemma and weak convergence for empirical processes of associated seduences//*Probab. Theory Relat. Fields. 1993. V.95. P. 357–370.*

Summary

Ekisheva S.V. The Bahadur representation of sample quantile for associatedstochastic sequence

We obtain the Bahadur representation for the empirical distribution function in the case of associated strictly stationary random sequence. The result is used to prove the asymptotic normality of sample quantile, the functional central theorem and the functional law of the iterated logarithm for sample quantile.

Сыктывкарский университет *Поступила 05.05.1998*