Math-Net.Ru

Общероссийский математический портал

D. Ferger, Optimal bounds for the Prokhorov distance of the Miller–Sen process and Brownian motion, *Теория вероятн. и ее примен.*, 1997, том 42, выпуск 1, 202–208

DOI: <https://doi.org/10.4213/tvp1723>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.82.47.9

26 апреля 2023 г., 13:00:53

© 1997 г.

FERGER D.*

OPTIMAL BOUNDS FOR THE PROKHOROV DISTANCE OF THE MILLER-SEN PROCESS AND BROWNIAN MOTION

Мы рассматриваем U -процессы, введенные Миллером и Сенем. Установлена точная скорость сходимости к броуновскому движению в метрике Прохорова.

Ключевые слова и фразы: невырожденные U -процессы, метрика Прохорова, разложение Хёфдинга, процесс частичных сумм.

0. Introduction. For each $n \in \mathbb{N}$ let $X_{1n}, \dots, X_{nn}$ be independent random elements defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in a measurable space $(\mathcal{X}, \mathcal{B})$ and with common distribution ν_n . If $h: \mathcal{X}^m \rightarrow \mathbb{R}$, $m \in \mathbb{N}$, is a symmetric mapping, then the statistic

$$U_{kn} = \binom{k}{m}^{-1} \sum_{1 \leq i_1 < \dots < i_m \leq k} h(X_{i_1 n}, \dots, X_{i_m n}), \quad m \leq k \leq n,$$

is commonly known as a U -statistic of degree m with kernel h pertaining to the sample $X_{1n}, \dots, X_{kn}$.

Many of the classical results of probability theory dealing with sums of independent and identically distributed random variables have generalizations to U -statistics: the Central Limit Theorem (Hoeffding [8]), the Strong Law of Large Numbers (Hoeffding [9], Berk [2]), the Law of the Iterated Logarithm (Serfling [14]), the Berry-Esséen Theorem (Grams and Serfling [6]) and Donsker's Invariance Principle (Miller and Sen [12]).

Note that the Berry-Esséen theorem gives the best possible rate of convergence in the Central Limit Theorem. Since the Miller-Sen [12] result is a *functional* Central Limit Theorem one can be interested in a functional counterpart of the Berry-Esséen Theorem.

To be precise we consider the stochastic process Y_n introduced by Miller and Sen [12] as an extension of the partial sum process: Y_n is the continuous random polygon starting at the origin $(0, 0)$ and with vertices at the points $(k/n, Y_{kn})$, $k \in \{m, m+1, \dots, n\}$, where

$$Y_{kn} = \frac{k(U_{kn} - \theta_n)}{m\sqrt{n}\zeta_n}, \quad k \in \{m, m+1, \dots, n\},$$

with $\theta_n = \mathbb{E}h(X_{1n}, \dots, X_{mn})$ and ζ_n is some variance to be specified below. Thus Y_n is a random element in the space $C = C[0, 1]$ of all continuous real-valued functions on the unit interval endowed with the Borel- σ -field $\mathcal{B}(C)$ generated by the topology of uniform convergence. Let $\mu_n = \mathbb{P} \circ Y_n^{-1}$ denote the distribution of Y_n and let W be the *Wiener measure*. It is well known that the weak topology on the set $\mathcal{M}(C)$ of all probability measures on $\mathcal{B}(C)$ is metrizable by the *Prokhorov metric* ρ , which for two probability measures μ_1 and μ_2 on $\mathcal{B}(C)$ is defined by

$$\rho(\mu_1, \mu_2) = \inf \{ \varepsilon > 0: \mu_1(B) \leq \mu_2(B^\varepsilon) + \varepsilon \quad \forall B \in \mathcal{B}(C) \}.$$

Here B^ε is the ε -neighborhood of the set B , i.e.,

$$B^\varepsilon = \{ f \in C: d(f, B) < \varepsilon \},$$

---

*Mathematisches Institut, Justus-Liebig-Universität Gießen, Arndtstr. 2, D-35392 Gießen, Germany.

where $d(f, B)$ denotes the distance between f and B under the sup-norm d . Now, suppose — for a moment — that our triangular array $(X_{in})_{1 \leq i \leq n, n \in \mathbb{N}}$ arises in the usual way from a single sequence $(X_i)_{i \in \mathbb{N}}$ of independent and identically distributed random elements, i.e., $X_{in} = X_i$ for $1 \leq i \leq n, n \in \mathbb{N}$. In this framework the Miller-Sen [12] invariance principle can be formulated as

$$\lim_{n \rightarrow \infty} \rho(\mu_n, W) = 0. \quad (0.1)$$

We will derive exact rates of convergence in (0.1). Statements of that kind can be regarded as functional counterparts of the Berry-Esséen theorem.

Herewith, we do not only fill a gap in the extension of classical probability theory to U -statistics. In addition bounds of Berry-Esséen type for real-valued function of Y_n and the Brownian motion B can be deduced. To see this we quote a lemma due to Borovkov [3].

Lemma [3]. *Let $\Gamma, \Gamma_1, \Gamma_2, \dots$ be random elements in C which are defined on the same probability space. If $T: C \rightarrow \mathbb{R}$ is Lipschitz-continuous, i.e., there exists a constant $K_1 < \infty$ such that*

$$|T(f) - T(g)| \leq K_1 d(f, g) \quad \forall f, g \in C,$$

and if the distribution function of $T(\Gamma)$ is also Lipschitz-continuous, i.e., there exists a constant $K_2 < \infty$ such that

$$\left| \mathbb{P}\{T(\Gamma) \leq x\} - \mathbb{P}\{T(\Gamma) \leq y\} \right| \leq K_2 |x - y| \quad \forall x, y \in \mathbb{R},$$

then for all $n \in \mathbb{N}$

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}\{T(\Gamma_n) \leq x\} - \mathbb{P}\{T(\Gamma) \leq x\} \right| \leq (K_1 K_2 + 1) \rho(\mathbb{P} \circ \Gamma_n^{-1}, \mathbb{P} \circ \Gamma^{-1}).$$

So under the additional smoothness condition of Lipschitz continuity the Borovkov lemma is a sharpening of the Continuous Mapping Theorem.

In the rest of this section we introduce some notation. Let $h: \mathcal{X}^m \rightarrow \mathbb{R}$ be a symmetric mapping (kernel) satisfying $\mathbb{E}|h(X_{1n}, \dots, X_{mn})| < \infty$ for all $n \geq m$. We define the associated functions

$$h_{cn}(x_1, \dots, x_c) = \mathbb{E}[h(x_1, \dots, x_c, X_{c+1,n}, \dots, X_{mn})]$$

for each $c = 1, \dots, m-1$, and set $h_{mn} = h$.

Put $\theta_n = \mathbb{E}h(X_{1n}, \dots, X_{mn})$ and define recursively $g_{1n}(x_1) = h_{1n}(x_1) - \theta_n$ and

$$\begin{aligned} g_{cn} &= h_{cn}(x_1, \dots, x_c) - \theta_n - \sum_{i=1}^c g_{1n}(x_i) - \sum_{1 \leq i_1 < i_2 \leq c} g_{2n}(x_{i_1}, x_{i_2}) \\ &\quad - \dots - \sum_{1 \leq i_1 < \dots < i_{c-1} \leq c} g_{c-1,n}(x_{i_1}, \dots, x_{i_{c-1}}), \quad 2 \leq c \leq m. \end{aligned}$$

For $1 \leq c \leq m$ we set

$$S_{ck}^{(n)} = \sum_{1 \leq i_1 < \dots < i_c \leq k} g_{cn}(X_{i_1n}, \dots, X_{i_cn}), \quad \zeta_n = \text{Var} g_{1n}(X_{1n}),$$

and finally put $m_{n,r} \equiv m_n = \mathbb{E}|g_{1n}(X_{1n})|^r$.

1. The main results. In the following theorems we obtain the same exact rates of convergence as in the functional limit theorems for C -valued partial sum processes based on rowwise independent random variables (see Borovkov [3], Utev [16], Sakhanenko [13])

and Häusler [7]) or based on a sequence of i.i.d. random variables (see Komlós, Major, Tusnády [11]).

Theorem 1.1. Suppose that for some $r > 2$

$$M = \sup_{m \leq n \in \mathbb{N}} \mathbb{E} |h(X_{1n}, \dots, X_{mn})|^r < \infty, \quad (1.1)$$

$$\zeta_n > 0 \quad \text{for all } n \geq m \quad (1.2)$$

and

$$m_n^{-1} = O(n(\log n)^{-1}). \quad (1.3)$$

Then

$$\rho(\mu_n, W) \leq C_r \zeta_n^{-r/(2(r+1))} m_n^{1/(r+1)} n^{-(r-2)/(2(r+1))}, \quad (1.4)$$

where C_r is a positive constant.

Moreover, for each $r > 2$ there exists a kernel h and a triangular array (X_{in}) such that (1.1)–(1.3) hold and for some positive constant D_r

$$\rho(\mu_n, W) \geq D_r \zeta_n^{-r/(2(r+1))} m_n^{1/(r+1)} n^{-(r-2)/(2(r+1))} \quad \text{for eventually all } n \geq m. \quad (1.5)$$

This means that the rate in (1.4) is best possible.

If each row of the array comes from the same distribution, i.e., $\nu_n = \nu$ for all $n \in \mathbb{N}$, then condition (1.3) readily follows from (1.2). Note that the rate in (1.4) is now of the order $O(n^{-(r-2)/2(r+1)})$.

Next, let $(X_i)_{i \in \mathbb{N}}$ be a sequence of independent and identically distributed random elements. Let the triangular array (X_{in}) be defined by

$$X_{in} = X_i \quad \text{for all } 1 \leq i \leq n \text{ for all } n \in \mathbb{N}. \quad (1.6)$$

In this setup we can prove a sharper result.

Theorem 1.2. Let (X_{in}) be the triangular array (1.6) and suppose

$$\zeta_1 > 0 \quad (1.7)$$

and for some $r > 2$

$$\mathbb{E} |h(X_1, \dots, X_m)|^r < \infty. \quad (1.8)$$

Then

$$\rho(\mu_n, W) = o(n^{-(r-2)/(2(r+1))}). \quad (1.9)$$

Furthermore, for each $r > 2$ and for each sequence $(\omega_n) \subseteq \mathbb{R}$ tending to infinity, there exists a kernel h and an i.i.d. sequence $(X_i)_{i \in \mathbb{N}}$ such that (1.7) and (1.8) hold and

$$\limsup_{n \rightarrow \infty} \omega_n n^{(r-2)/(2(r+1))} \rho(\mu_n, W) = \infty. \quad (1.10)$$

In other words, the rate in (1.9) is optimal.

R e m a r k. Under the assumptions of Theorem 1.2 Horváth [10] proved the following strong invariance principle: There exists a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq 1} |Y_n(t) - B(t)| = o(\sqrt{\log \log n}) \right\} = 1, \quad \text{if } r = 2,$$

and

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq 1} |Y_n(t) - B(t)| = o(n^{-(r-2)/(2r)}) \right\} = 1, \quad \text{if } r > 2. \quad (1.11)$$

Note that the rates of convergence in (1.9) and (1.11) differ. Especially (1.11) does not imply (1.9) and vice versa.

2. Proofs. Recall $S_{1k}^{(n)} = \sum_{i=1}^k g_{1n}(X_{in})$, $0 \leq k \leq n$, being the first term of the Hoeffding-decomposition of U_{kn} (confer [14, p. 178]). The induced partial sum process S_n is defined by

$$S_n(t) = \frac{1}{\sqrt{n\zeta_n}} \sum_{i=1}^k g_{1n}(X_{in}),$$

if $t = k/n$, $0 \leq k \leq n$, and by linear interpolation on the subintervals $[(k-1)/n, k/n]$, $1 \leq k \leq n$. The following lemma deals with the closeness of S_n and Y_n . The distribution of S_n is denoted by τ_n : $\tau_n = P \circ S_n^{-1}$. For the rest of this section C is a generic constant that may depend on r , m and M .

Lemma 2.1. *Let $r \geq 2$ and suppose that (1.1) holds. Then*

$$P \left\{ \sup_{0 \leq t \leq 1} |Y_n(t) - S_n(t)| > \varepsilon \right\} \leq C \varepsilon^{-r} \zeta_n^{-r/2} n^{-r/2} \log n \quad (2.1)$$

for all $\varepsilon > 0$ and for all $n \geq m$. Moreover, we have that

$$\rho(\mu_n, \tau_n) = O \left(\zeta_n^{-r/(2(r+1))} n^{-r/(2(r+1))} (\log n)^{1/(r+1)} \right). \quad (2.2)$$

Proof. We have

$$\begin{aligned} P \left\{ \sup_{0 \leq t \leq 1} |Y_n(t) - S_n(t)| > \varepsilon \right\} &\leq P \left\{ \sup_{0 \leq t \leq m/n} |Y_n(t) - S_n(t)| > \frac{1}{2} \varepsilon \right\} \\ &+ P \left\{ \sup_{m/n \leq t \leq 1} |Y_n(t) - S_n(t)| > \frac{1}{2} \varepsilon \right\} = p_n + q_n. \end{aligned}$$

The first probability p_n is less than or equal to

$$\begin{aligned} &P \left\{ \sup_{0 \leq t \leq m/n} |Y_n(t)| > \frac{1}{4} \varepsilon \right\} + P \left\{ \sup_{0 \leq t \leq m/n} |S_n(t)| > \frac{1}{4} \varepsilon \right\} \\ &= P \left\{ \frac{|U_{mn} - \theta_n|}{\sqrt{n\zeta_n}} > \frac{1}{4} \varepsilon \right\} + P \left\{ \max_{1 \leq k \leq m} \frac{1}{\sqrt{n\zeta_n}} \left| \sum_{i=1}^k g_{1n}(X_{in}) \right| > \frac{1}{4} \varepsilon \right\} = p_{n1} + p_{n2}. \end{aligned}$$

By Markov's inequality and by (1.1) we obtain

$$p_{n1} = P \left\{ |h(X_{1n}, \dots, X_{mn}) - \theta_n| > \frac{1}{4} \varepsilon \sqrt{n\zeta_n} \right\} \leq C \varepsilon^{-r} n^{-r/2} \zeta_n^{-r/2}. \quad (2.3)$$

Using Markov's and Doob's inequalities we obtain

$$p_{n2} \leq C \varepsilon^{-r} n^{-r/2} \zeta_n^{-r/2} m_{n,r} \leq C \varepsilon^{-r} n^{-r/2} \zeta_n^{-r/2}, \quad \text{by (1.1).} \quad (2.4)$$

Combining (2.3) and (2.4) one finds that for all $n \geq 1$

$$p_n \leq C \varepsilon^{-r} \zeta_n^{-r/2} n^{-r/2}. \quad (2.5)$$

As to the second probability q_n note that by the Hoeffding decomposition

$$\begin{aligned} \sup_{m/n \leq t \leq 1} |Y_n(t) - S_n(t)| &= \frac{1}{m\sqrt{n\zeta_n}} \max_{m \leq k \leq n} \left| k \sum_{c=2}^m \binom{m}{c} \binom{k}{c}^{-1} S_{ck}^{(n)} \right| \\ &\leq \frac{1}{m\sqrt{n\zeta_n}} \sum_{c=2}^m \binom{m}{c} \max_{m \leq k \leq n} k \binom{k}{c}^{-1} |S_{ck}^{(n)}|, \end{aligned}$$

whence it follows that

$$q_n \leq \sum_{c=2}^m \mathbf{P} \left\{ \max_{c \leq k \leq n} k \binom{k}{c}^{-1} |S_{ck}^{(n)}| > \frac{1}{2} \binom{m}{c}^{-1} \varepsilon \sqrt{n \zeta_n} \right\}. \quad (2.6)$$

Fix $2 \leq c \leq m$. As part of the Hoeffding decomposition one also has that the sequence $(S_{ck}^{(n)}, \mathcal{F}_k^{(n)})_{c \leq k \leq n}$ with $\mathcal{F}_k^{(n)} = \sigma(X_{1n}, \dots, X_{kn})$ is a martingale. Therefore by Chow's inequality [4] for each $\gamma > 0$:

$$\begin{aligned} \mathbf{P} \left\{ \max_{c \leq k \leq n} k \binom{k}{c}^{-1} |S_{ck}^{(n)}| > \gamma \right\} &\leq \mathbf{P} \left\{ \max_{c \leq k \leq n} (k-c+1)^{-r(c-1)} |S_{ck}^{(n)}|^r > (c!)^{-r} \gamma^r \right\} \\ &\leq (c!)^r \gamma^{-r} \left\{ \sum_{k=c}^{n-1} \left[(k-c+1)^{-r(c-1)} - (k-c+2)^{-r(c-1)} \right] \mathbf{E} |S_{ck}^{(n)}|^r \right. \\ &\quad \left. + (n-c+1)^{-r(c-1)} \mathbf{E} |S_{cn}^{(n)}|^r \right\}. \end{aligned} \quad (2.7)$$

By Lemma 2.1 of Ferger [5] there exists a positive constant C_1 depending only on r and c such that

$$\mathbf{E} |S_{ck}^{(n)}|^r \leq C_1 \left[\int |g_{c,n}|^r d\nu_n^c \right] k^{rc/2}$$

for all $n \in \mathbf{N}$ and $c \leq k \leq n$, where ν_n^c denotes the c -fold product of the measure ν_n . A routine induction argument involving Jensen's inequality and the C_r -inequality (confer, e.g., [15, p. 843]) shows that for each $2 \leq c \leq m$ there exists a positive constant $A_{c,r}$ depending only on c and r such that

$$\int |g_{c,n}|^r d\nu_n^c \leq A_{c,r} \int |h|^r \nu_n^m \quad \text{for all } n \geq m.$$

Consequently, by (1.1), one gets

$$\mathbf{E} |S_{ck}^{(n)}|^r \leq B_{c,r} k^{rc/2} \quad \forall n \geq m \quad \forall 2 \leq c \leq m \quad \forall c \leq k \leq m,$$

where $B_{c,r}$ are positive constants depending only on r, M and c . Thus by (2.7) we obtain

$$\begin{aligned} \mathbf{P} \left\{ \max_{c \leq k \leq n} k \binom{k}{c}^{-1} |S_{ck}^{(n)}| > \gamma \right\} \\ \leq (c!)^r \gamma^{-r} \left(\sum_{k=c}^{n-1} \left[(k-c+1)^{-r(c-1)} - (k-c+2)^{-r(c-1)} \right] B_{c,r} k^{rc/2} \right. \\ \left. + (n-c+1)^{-r(c-1)} B_{c,r} n^{rc/2} \right). \end{aligned} \quad (2.8)$$

After some routine calculations upon using the Mean Value Theorem one sees that the third factor (...) on the right-hand side of (2.8) is less than or equal to $C \log n$ for all $n \geq m$. Put

$$\gamma = \frac{1}{2} \binom{m}{c}^{-1} \varepsilon \sqrt{n \zeta_n}.$$

Then (2.6)–(2.8) lead to

$$q_n \leq C \varepsilon^{-r} \zeta_n^{-r/2} n^{-r/2} \log n \quad \text{for all } n \geq m.$$

Now, recall (2.5) to see that (2.1) is true.

For the proof of (2.2) note that for all $\varepsilon > 0$

$$\begin{aligned}\rho(\mu_n, \tau_n) &\leq \varepsilon + \mathbf{P} \left\{ \sup_{0 \leq t \leq 1} |Y_n(t) - S_n(t)| > \varepsilon \right\} \\ &\leq \varepsilon + C \varepsilon^{-r} \zeta_n^{-r/2} n^{-r/2} \log n \quad \text{for all } n \geq m,\end{aligned}$$

by (2.1). For $\varepsilon = (n\zeta_n)^{-r/(2(r+1))}(\log n)^{1/(r+1)}$ we obtain (2.2). Lemma 2.1 is proved.
In view of the triangle inequality

$$\rho(\mu_n, W) \leq \rho(\mu_n, \tau_n) + \rho(\tau_n, W) \quad (2.9)$$

relation (2.2) is the key for the proof of our theorems as it reduces the problem to the well-known partial sum process S_n based on an array of rowwise independent random variables.

P r o o f o f T h e o r e m 1.1. Sakhanenko [13] showed that for all $r > 2$

$$\rho(\tau_n, W) \leq C \zeta_n^{-r/(2(r+1))} m_n^{1/(r+1)} n^{-(r-2)/(2(r+1))} \quad \forall n \geq m.$$

Thus (1.3), (2.2) and (2.9) imply (1.4).

For the proof of (1.5) choose $\mathcal{X} = \mathbf{R}$ endowed with the usual Borel- σ -algebra and $h: \mathcal{X}^m \rightarrow \mathbf{R}$ defined by

$$h(x_1, \dots, x_m) = \sum_{i=1}^m x_i. \quad (2.10)$$

Let $\xi_{1n}, \dots, \xi_{nn}$ be the independent and identically distributed random variables as introduced by Arak [1], i.e., ξ_{kn} takes on the values $-\alpha n^{-1/2}$, $\alpha n^{-1/2}$ and δ with probabilities

$$\frac{1}{2} \left(1 - \frac{1}{n} \delta + \frac{1}{\alpha} n^{-1/2} \delta^2 \right), \quad \frac{1}{2} \left(1 - \frac{1}{n} \delta - \frac{1}{\alpha} n^{-1/2} \delta^2 \right) \quad \text{and} \quad \frac{1}{n} \delta, \quad \text{respectively,}$$

where

$$\delta = n^{-(r-2)/(2(r+2))} \quad \text{and} \quad \alpha = \left\{ (1 - \delta^3) \left(1 - \frac{\delta}{n} \right)^{-1} \right\}^{1/2}.$$

Put

$$X_{in} = n^{1/2} \xi_{in}, \quad 1 \leq i \leq n, \quad n \in \mathbf{N}.$$

Then $\mathbf{E}X_{in} = 0$, $\text{Var}X_{in} = 1$, g_{1n} is the identity map, $\zeta_n = 1$, $m_n = \alpha^r (1 - \delta/n) + 1$ and $M \leq 2^{m(r-1)} < \infty$. We see that the kernel h and the array (X_{in}) meet the requirements (1.1)–(1.3). Arak [1] proved that

$$\rho(\tau_n, W) \geq \frac{1}{3} n^{-(r-2)/(2(r+1))} \quad \text{for eventually all } n \in \mathbf{N}.$$

By (2.2) we get

$$\rho(\mu_n, \tau_n) = O(n^{-r/(2(r+1))} \log n).$$

Since

$$m_n \leq 2 \quad \text{and} \quad \zeta_n = 1 \quad \text{for all } n \in \mathbf{N},$$

(1.5) is a consequence of the triangle inequality

$$\rho(\mu_n, W) \geq \rho(\tau_n, W) - \rho(\mu_n, \tau_n). \quad (2.11)$$

Theorem 1.1 is proved.

P r o o f o f T h e o r e m 1.2. In [11] it is proved that

$$\rho(\tau_n, W) = o \left(n^{-(r-2)/(2(r+1))} \right),$$

whence (1.9) follows from (2.2).

As to the optimality guaranteeing relation (1.10) we take $\mathcal{X} = \mathbb{R}$ and h as defined in (2.10). Now, Komlós, Major and Tusnady [11] also pointed out that given any real sequence (ω_n) tending to infinity there exists a sequence $(X_i)_{i \in \mathbb{N}}$ of independent and identically distributed random variables with $\mathbb{E}X_1 = 0$, $\text{Var}X_1 = 1$ and $\mathbb{E}|X_1|^r < \infty$ such that

$$\limsup \omega_n n^{(r-2)/(2(r+1))} \rho(\tau_n^*, W) = \infty,$$

where τ_n^* denotes the distribution of the partial sum process pertaining to the sequence $(X_i)_{i \in \mathbb{N}}$. But for our special h we have that $\tau_n = \tau_n^*$ and (1.10) is a consequence of (2.2), (2.11) and (2.12).

REFERENCES

1. Arak T. V. On an estimate of A. A. Borovkov. — *Theorie Probab. Appl.*, 1975, v. 20, p. 372–373.
2. Berk R. H. Limiting behaviour of posterior distributions when the model is incorrect. — *Ann. Math. Statist.*, 1966, v. 37, p. 51–58.
3. Borovkov A. A. On the rate of convergence for the invariance principle. — *Theorie Probab. Appl.*, 1973, v. 18, p. 207–225.
4. Chow Y. S. A martingale inequality and the law of large numbers. — *Proc. Amer. Math. Soc.*, 1960, v. 11, № 1, p. 107–111.
5. Ferger D. Moment inequalities for U -statistics with degeneracy of higher order. — *Sankyā, Ser. A*, 1996, v. 58, p. 142–148.
6. Grams W. F., and Serfling R. J. Convergence rates for U -statistics and related statistics. — *Ann. Statist.*, 1973, v. 1, № 1, p. 153–160.
7. Häusler E. An exact rate of convergence in the functional central limit theorem for special Martingale Difference Arrays. — *Z. Wahrscheinlichkeitstheor. verw. Geb.*, 1984, v. 65, p. 523–534.
8. Hoeffding W. A class of statistics with asymptotically normal distribution. — *Ann. Math. Statist.*, 1948, v. 19, p. 293–325.
9. Hoeffding W. The strong law of large numbers for U -statistics. — *Institute of Statistics Mimeo Series*, 1961, № 302, University of North Carolina.
10. Horváth L. Strong laws for randomly indexed U -statistics. — *Proc. Cambridge Philos. Soc.*, 1985, v. 98, № 3, p. 559–567.
11. Komlós J., Major P., Tusnady G. Weak convergence and embedding. — In: *Colloquia Mathematica Societatis János Bolyai*, 11. Limit Theorems of Probability Theory (Keszthely, 1974), Ed. by P. Révész. Amsterdam–London: North-Holland, p. 149–165.
12. Miller R. G., Sen P. K. Weak convergence of U -statistics and von Mises differentiable statistical functions. — *Ann. Math. Statist.*, 1972, v. 43, p. 31–41.
13. Сазаненко А. И. Оценки в принципе инвариантности. — *Труды Ин-та математики СО АН СССР*, 1985, т. 5, с. 27–44.
14. Serfling R. J. *Approximation Theorems of Mathematical Statistics*. New York: Wiley, 1980.
15. Shorack G. R., Wellner J. A. *Empirical processes with applications to statistics*. New York: Wiley, 1986.
16. Утев С. А. Замечание о скорости сходимости в принципе инвариантности. — *Сиб. матем. журн.*, 1981, т. 22, с. 206–209.

Поступила в редакцию
23.III.1993