

**THE CONVERGENCE OF U -STATISTICS
AND VON MISES FUNCTIONALS IN THE MEAN METRIC**

UDC 519.21

SH. SHARAKHMETOV

ABSTRACT. We prove global limit theorems for U -statistics and von Mises functionals. The rate of convergence in these theorems is also obtained.

Limit theorems for U -statistics are usually considered for the uniform metric (see [1]–[3]). This paper is devoted to limit theorems for the mean metric for which we study the convergence of distributions of U -statistics and von Mises functionals.

We prove global limit theorems and obtain the rate of convergence in these theorems for both degenerate and nondegenerate kernels.

The idea of the proofs of the results is based on the method of metric distances (see [4]). There are other approaches used in the theory of U -statistics, such as the martingale method, the method of characteristic functions, the method of orthogonal decompositions, and some others. It is worthwhile to mention that V. M. Zolotarev ([4, p. 404]) was the first to point out that the method of metric distances works in the case of nonlinear models, too.

The results of the paper are announced in [5].

1. RESULTS

Let x_1, \dots, x_n be independent identically distributed random variables assuming values in a measurable space (χ, A) . The U -statistic and von Mises functional are defined for x_1, \dots, x_n by

$$U_n = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \Phi(x_i, x_j) \quad \text{and} \quad V_n = \frac{1}{n^2} \sum_{i, j=1} \Phi(x_i, x_j),$$

respectively, where $\Phi: \chi^2 \rightarrow \mathbf{R}$ is a symmetric function of two variables.

Put

$$g(x_i) = \mathbf{E}(\Phi(x_i, x_j)/x_i) \quad \text{and} \quad y(x_i, x_j) = \Phi(x_i, x_j) - g(x_i) - g(x_j)$$

and note that $\mathbf{E} \Phi(x_i, x_j) = 0$ and $\sigma^2 = \mathbf{E} g^2(x_i)$. Let N denote a standard Gaussian random variable. Denote by χ_s the mean metric, that is

$$\chi_s(X, Y) = s \int |x|^{s-1} |\mathbf{P}(X < x) - \mathbf{P}(Y < x)| dx$$

for two random variables X and Y .

To make notation shorter, in what follows we write Φ_{ij} and g_i instead of $\Phi(x_i, x_j)$ and $g(x_i)$, respectively.

2000 *Mathematics Subject Classification.* Primary 60F05.

Theorem 1. Let $0 < \sigma < \infty$ and $\mathbf{E}|y_{12}|^{4/3} < \infty$. Then

$$\chi_1 \left(\frac{\sqrt{n}}{2\sigma} U_n, N \right) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Theorem 2. Let $0 < \sigma < \infty$ and $\mathbf{E}|y_{12}|^{4/3} < \infty$. Then

$$\chi_1 \left(\frac{\sqrt{n}}{2\sigma} V_n, N \right) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Theorem 3. Let $0 < \sigma < \infty$ and $\mathbf{E}|y_{12}|^{4/(3-\delta)} < \infty$ for some $0 < \delta \leq 1$. Then

$$\chi_1 \left(\frac{\sqrt{n}}{2\sigma} U_n, N \right) \leq n^{-\delta/2} \left(74 \frac{\mathbf{E}|g_1|^{2+\delta}}{\sigma^{2+\delta}} + 2^{1-\delta} \left(\frac{\mathbf{E}|y_{12}|^{4/(3-\delta)}}{\sigma} \right)^{(3-\delta)/4} \right).$$

Theorem 4. Let $0 < \sigma < \infty$ and $\mathbf{E}|y_{12}|^{4/(3-\delta)} < \infty$ for some $0 < \delta \leq 1$. Then

$$\chi_1 \left(\frac{\sqrt{n}}{2\sigma} V_n, N \right) \leq n^{-\delta/2} \left(74 \frac{\mathbf{E}|g_1|^{2+\delta}}{\sigma^{2+\delta}} + 2^{1-\delta} \left(\frac{\mathbf{E}|y_{12}|^{4/(3-\delta)}}{\sigma} \right)^{(3-\delta)/4} + \frac{\mathbf{E}|\Phi_{11}|}{2\sigma} \right).$$

Now we consider a degenerate kernel, that is, we deal with the case of $\sigma = 0$. Denote by S the linear operator $S: f \rightarrow \mathbf{E}(\Phi(x_1, x)f(x_1))$ and expand the kernel $\Phi(x, y)$ into a series with respect to its eigenvalues and eigenfunctions:

$$\Phi(x, y) = \sum_{k=1}^{\infty} \lambda_k e_k(x) e_k(y).$$

Theorem 5. Let $\sigma = 0$ and $\sum_{j=1}^{\infty} |\lambda_j| \mathbf{E}|e_j(x_j)|^3 < \infty$. Then

$$\chi_1(nU_n, U_\infty) \leq 74n^{-1/2} \sum_{j=1}^{\infty} |\lambda_j| \mathbf{E}|e_j(x_j)|^3,$$

where $U_\infty = \sum_{k=1}^{\infty} \lambda_k (N_k^2 - 1)$, and N_k is a standard Gaussian random variable, $k \geq 1$.

Theorem 6. Let $\sigma = 0$ and $\sum_{j=1}^{\infty} |\lambda_j| \mathbf{E}|e_j(x_j)|^3 < \infty$. Then

$$\chi_1(nV_n, V_\infty) \leq 74n^{-1/2} \sum_{j=1}^{\infty} |\lambda_j| \mathbf{E}|e_j(x_j)|^3,$$

where $V_\infty = \sum_{k=1}^{\infty} \lambda_k N_k^2$.

Remark. It will be seen from the proofs of the theorems that an analogous result holds for the metric χ_s , $s > 1$, too.

2. AUXILIARY RESULTS

To prove the theorems we need some auxiliary results.

Put $\tau_s(X, Y) = \mathbf{E}|X^{(s)} - Y^{(s)}|$, where $x^{(s)} = |x|^{s-1}x$ for $s \geq 1$, and

$$\widehat{\tau}_s(X, Y) = \inf \tau_s(X, Y),$$

where inf is taken over all joint distributions $P_{X,Y}$ whose marginal distributions $\mathbf{P}(X < x)$ and $\mathbf{P}(Y < y)$ are fixed.

Lemma 1 ([4, Theorem 1.3.1, p. 56]). *The metric $\chi_s(X, Y)$ is minimal for the metric $\tau_s(X, Y)$, that is, $\chi_s(X, Y) \leq \tau_s(X, Y)$ and $\widehat{\tau}_s(X, Y) = \chi_s(X, Y)$.*

Lemma 2. *Let ξ_1, \dots, ξ_n be independent identically distributed random variables such that $E\xi_1 = 0$, $E\xi_1^2 = \sigma^2$, and $E|\xi_1|^p < \infty$ for some $p \geq 2$. Then*

$$\chi_s\left(\frac{1}{\sqrt{n}\sigma} \sum_{i=1}^n \xi_i, N\right) \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where $1 \leq s \leq p$, and

$$\chi_s\left(\frac{1}{\sqrt{n}\sigma} \sum_{i=1}^n \xi_i, N\right) \leq cn^{-(p_0-2)/2} \frac{E|\xi_1|^{p_0}}{\sigma^{p_0}}$$

for $1 \leq s \leq p$, $p > 2$, and $p_0 = \min(3, p)$.

In Lemma 2, $c > 0$ is a constant depending only on s and p ; for example, one can take $c = 74$ in the case of $1 \leq s < p \leq 3$.

The first assertion of Lemma 2 follows from Theorem 2.1 in [6].

The second assertion of Lemma 2 follows from a nonuniform bound in the central limit theorem (see [7, Chapter V, Theorem 14] for the case of $p \geq 3$, and [8] for the case of $p \leq 3$).

Lemma 3 (see [2, Theorem 2.1.3 and relation (4.2.8)]). *Let $E\Phi_{12} = 0$, $\sigma = 0$, and $E|\Phi_{12}|^p < \infty$ for some $1 \leq p \leq 2$. Then $E|U_n|^p \leq 2^{2(2-p)}(c_n^2)^{1-p} E|y_{12}|^p$ and*

$$E\left|n^{-3/2} \sum_{i < j} \Phi_{ij}\right|^{4/3} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

3. PROOFS OF THE THEOREMS

We prove the results for the von Mises functional (Theorems 2, 4, and 6). The results for the U -statistic (Theorems 1, 3, and 5) can be proved similarly. The following representation is crucial for the proof of Theorems 1, 3, and 5:

$$V_n = \frac{n-1}{n}U_n + \frac{1}{n} \sum_{i=1}^n \Phi_{ii}.$$

Proof of Theorems 2 and 4. The triangle inequality for the metric χ_1 implies

$$(1) \quad \chi_1\left(\frac{\sqrt{n}}{2\sigma}V_n, N\right) \leq \chi_1\left(\frac{\sqrt{n}}{2\sigma}V_n, \frac{n-1}{\sigma n^{3/2}} \sum_{i=1}^n g_i\right) + \chi_1\left(\frac{n-1}{\sigma n^{3/2}} \sum_{i=1}^n g_i, N\right).$$

Using the relation between the metrics χ_1 and τ_1 (Lemma 1) and the Hoeffding representation (see [1]–[3])

$$V_n = \frac{1}{n^2} \sum_{i=1}^n \Phi_{ii} + \frac{2(n-1)}{n^2} \sum_{i=1}^n g_i + \frac{2}{n^2} \sum_{1 \leq i < j \leq n} y_{ij},$$

we get

$$\begin{aligned} \chi_1\left(\frac{\sqrt{n}}{2\sigma}V_n, \frac{n-1}{\sigma n^{3/2}} \sum_{i=1}^n g_i\right) &\leq \tau_1\left(\frac{\sqrt{n}}{2\sigma}V_n, \frac{n-1}{\sigma n^{3/2}} \sum_{i=1}^n g_i\right) \\ &= E\left|\frac{1}{2\sigma n^{3/2}} \sum_{i=1}^n \Phi_{ij} + \frac{1}{\sigma n^{3/2}} \sum_{i < j \leq n} y_{ij}\right| \\ &\leq \frac{E|\Phi_{11}|}{2\sqrt{n}\sigma} + \frac{1}{\sigma n^{3/2}} E\left|\sum_{1 \leq i < j \leq n} y_{ij}\right|. \end{aligned}$$

It follows from (1) that

$$(2) \quad \chi_1 \left(\frac{\sqrt{n}}{2\sigma} V_n, N \right) \leq \frac{\mathbf{E} |\Phi_{11}|}{2\sqrt{n}\sigma} + \frac{1}{\sigma n^{3/2}} \mathbf{E} \left| \sum_{1 \leq i < j \leq n} y_{ij} \right| + \chi_1 \left(\frac{n-1}{\sigma n^{3/2}} \sum_{i=1}^n g_i, N \right).$$

Lemmas 2 and 3 imply that

$$\chi_1 \left(\frac{n-1}{\sigma n^{3/2}} \sum_{i=1}^n g_i, N \right) \rightarrow 0 \quad \text{and} \quad \frac{1}{\sigma n^{3/2}} \mathbf{E} \left| \sum_{1 \leq i < j \leq n} y_{ij} \right| \rightarrow 0$$

as $n \rightarrow \infty$.

Theorem 2 is proved.

Theorem 4 also follows from relation (2). Indeed,

$$\chi_1 \left(\frac{n-1}{\sigma n^{3/2}} \sum_{i=1}^n g_i, N \right) \leq 74 \frac{\mathbf{E} |g_1|^{2+\delta}}{\sigma^{2+\delta}} n^{-\delta/2}$$

by Lemma 2. For $\mathbf{E} |g_1|^{2+\delta} < \infty$ we have

$$\frac{1}{n^{3/2}} \mathbf{E} \left| \sum_{i < j} y_{ij} \right| \leq \frac{1}{n^{3/2}} \left(\mathbf{E} \left| \sum_{i < j} y_{ij} \right|^{4/(3-\delta)} \right)^{(3-\delta)/4} \leq 2^{1-\delta} n^{-\delta/2} \left(\mathbf{E} |y_{12}|^{4/(3-\delta)} \right)^{(3-\delta)/4}$$

by Lemma 3.

Theorem 4 is proved. □

Proof of Theorem 6. Using the expansion of a kernel into a series with respect to a system of orthonormal functions $\{e_k\}$,

$$\phi(x, y) = \sum_{k=1}^{\infty} \lambda_k e_k(x) e_k(y),$$

we obtain

$$nVn = \frac{1}{n} \sum_{i,j=1}^n \Phi_{ij} = \frac{1}{n} \sum_{i,j=1}^n \sum_{k=1}^{\infty} \lambda_k e_k(x_i) e_k(x_j) = \sum_{k=1}^{\infty} \lambda_k \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n e_k(x_i) \right)^2 = \sum_{k=1}^{\infty} \lambda_k S_{nk}^2,$$

where $S_{nk} = n^{-1/2} \sum_{i=1}^n e_k(x_i)$ is a sum of independent identically distributed random variables such that $\mathbf{E} e_k(x_1) = 0$ and $\mathbf{E} e_k^2(x_1) = 1$.

Lemma 1 implies

$$\begin{aligned} \chi_1(n, V_n, V_{\infty}) &= \chi_1 \left(\sum_{k=1}^{\infty} \lambda_k S_{nk}^2, \sum_{k=1}^{\infty} \lambda_k N_k^2 \right) \leq \tau_1 \left(\sum_{k=1}^{\infty} \lambda_k S_{nk}^2, \sum_{k=1}^{\infty} \lambda_k N_k^2 \right) \\ &= \mathbf{E} \left| \sum_{k=1}^{\infty} \lambda_k (S_{nk}^2 - N_k^2) \right| \leq \sum_{k=1}^{\infty} |\lambda_k| \mathbf{E} |S_{nk}^2 - N_k^2| = \sum_{k=1}^{\infty} |\lambda_k| \tau_2(S_{nk}, N_k). \end{aligned}$$

We did not put any restriction on the joint distribution of s_{nk} and N_k , except for a condition posed on the marginal distributions of these random variables. Thus one can choose the marginal distributions such that the metric τ_2 for s_{nk} and N_k becomes equal to the minimal metric χ_2 . Therefore

$$\chi_1(n, V_n, V_{\infty}) \leq \sum_{k=1}^{\infty} |\lambda_k| \chi_2(s_{nk}, N_k).$$

Now we apply Lemma 2 to estimate $\chi_2(s_{nk}, N)$:

$$\chi_1(nV_n, V_\infty) \leq 74n^{-1/2} \sum_{k=1}^{\infty} |\lambda_k| \mathbb{E} |e_k(x_1)|^3.$$

Theorem 6 is proved. \square

4. DISCUSSION OF RESULTS AND SOME HISTORICAL REMARKS

We follow the monograph [4] when discussing the history of the topic.

The metric approach to the problems on the accuracy of approximations of distributions appeared in the theory of probability in the mid 1930s. In the course of time, the number of metrics involved in this approach was growing. Some of the metrics appeared in probability theory by analogy with the corresponding notions in the theory of functions (say, the uniform metric ρ as well as the total variation metric σ and mean metric χ_1); others are specific for probability theory.

The main idea of the method of metric distances is as follows: approaching given characteristics (say, the distributions) of certain random variables by those of other random variables and measuring the deviation between them, one must choose a metric that fits the mathematical model in a best possible way. For example, appropriate ideal metrics should be chosen in the case of linear models, while a compound metric depending on $|x - y|$ should be chosen in the case of a nonlinear model, provided the corresponding minimal metric can be expressed explicitly.

U -statistics and von Mises functionals are two of universal tools of the current probabilistic theory of summation. From the algebraic point of view, these two tools, being nonlinear objects, are more complicated than sums of random variables. On the other hand, the U -statistics and von Mises functionals contain an essential part of information on the dependence structure of random variables. A survey on the U -statistics and von Mises functionals, relevant references, and various examples can be found in [1]–[3].

Therefore it is quite natural to apply the method of metric distances to study the convergence of U -statistics and von Mises functionals to a limit law, as well as to choose the mean metric χ_1 as the measure of deviation.

Due to this setting, results, analogous to the central limit theorem (see Theorems 1–6 above), are proved just by applying two simple properties of the metric χ_1 . On the other hand it is well known that results of this type are nontrivial and their proof is rather complicated (most of those results have been obtained during the last 40 years, see [2]).

Theorems 1 and 2 are proved under optimal conditions posed on the kernel.

Estimates obtained in Theorems 3 and 4 cannot be improved as far as the order of decay with respect to n is concerned. Note that the absolute constants in Theorems 3 and 4 are much better than those in similar estimates obtained for the case of the uniform metric where the absolute constant is $12!$ (see [2]). To be honest, we note that the order of decay $O(n^{-1/2})$ is obtained in this paper under the conditions $\mathbb{E} |g_1|^3 < \infty$ and $\mathbb{E} |y_{12}|^2 < \infty$, while the same result for the uniform metric holds if $\mathbb{E} |g_1|^3 < \infty$ and $\mathbb{E} |y_{12}|^{5/3} < \infty$. Therefore estimates obtained in Theorems 3 and 4 seem to be nonoptimal as far as the condition posed on the moments of the kernel is concerned.

One also should mention that the conditions for estimates in the case of degenerate kernels (Theorems 5 and 6) are expressed in terms of moments of eigenvalues and eigenfunctions of kernels. It would be more natural to express these conditions in terms of kernels themselves. Perhaps, the correct rate of convergence in Theorems 5 and 6 is $o(n^{-1/2})$.

BIBLIOGRAPHY

1. L. J. Serfling, *Approximation theorems of mathematical statistics*, Wiley, New York, 1980. MR **82a**:62003
2. V. S. Korolyuk and Yu. V. Borovskikh, *Theory of U statistics*, “Naukova dumka”, Kiev, 1989; English transl., Kluwer, Dordrecht, 1993. MR **90k**:60044
3. P. K. Sen, *Sequential nonparametrics: invariance principles and statistical inference*, Wiley, New York, 1981. MR **83c**:62128
4. V. M. Zolotarev, *A modern theory of summation of independent random variables*, “Nauka”, Moscow, 1986. (Russian) MR **89a**:60003
5. Sh. Sharakhmetov, *Global limit theorems for U -statistic and von Mises functional*, Proceed. 5th Vilnius Conf., Vilnius, 1989. (Russian)
6. V. M. Kruglov, *Global limit theorems*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **VIII** (1976), 84–101. MR **55**:4324
7. V. V. Petrov, *Limit theorems for sums of independent random variables*, “Nauka”, Moscow, 1987; English transl., Oxford Univ. Press, Oxford, 1991. MR **88i**:60004
8. R. Michel, *On the constant in the nonuniform version of the Berry–Esseen theorem*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **55** (1981), 109–117. MR **82c**:60042
9. C. Gini, *Di una misura delle relazioni tra le graduatorie di due caratteri*, Ludovico Cecehini, Rome, 1914.

TASHKENT UNIVERSITY FOR ECONOMICS, UZBEKISTANSKII PROSPEKT 49, TASHKENT 700063, UZBEK-ISTAN

Received 1/SEP/1999

Translated by OLEG KLESOV