1991

V.S. Korolyuk, Yu.V. Borovskikh

U-statistics of increasing degrees with asymptotically Poisson distributions

Department of Operations Research, Statistics, and System Theory   Report BS-R9120 August

CWI, nationaal instituut voor onderzoek op het gebied van wiskunde en informatica

CWI is the research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a non-profit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands organization for scientific research (NWO).

U-statistics of Increasing Degrees with Asymptotically Poisson Distributions

V.S. Korolyuk
Institute of Mathematics Ukrainian Academy of Sciences, Kiev, USSR

Yu. V. Borovskikh
Transport Institute, Leningrad, USSR

This paper is devoted to the limit theorem for U-statistics with increasing degrees provided that the counting measure of sampling variables has asymptotically Poisson distribution.

1980 Mathematics Subject Classification: 62E20
Key Words & Phrases: U-statistic, random measure, weak convergence, stochastic Poisson-Ito integral.

Let $X_{n1}, X_{n2},\ldots$ be a sequence of independent identically distributed random variables taking values in a measurable space $(\mathcal{X}, \mathcal{Q})$ having a common distribution π_n on this space. Consider an U-statistic

$$U_n = \begin{bmatrix} n \\ m \end{bmatrix}^{-1} \cdot \sum_{1\leqslant i_1<\ldots<i_m\leqslant n} \Phi(X_{ni_1}, \ldots, X_{ni_m}) \tag{1}$$

where $\Phi:\mathcal{X}^m \to R$ is a symmetric function, m is integer number, $1 \leqslant m \leqslant n$, which is called the degree of the U-statistic. The asymptotic theory, as $n\to\infty$, m being fixed, for U-statistics is constructed in [1]. The case when m depends on n, i.e. $m=m(n)$, so that $m\to\infty$ as $n\to\infty$ is studied very little. Some results (see [2],[3],[4]) give a hint that in this situation the asymptotics is essentially different to that in the case of fixed m.

The aim of this paper is to study the asymptotics of the U-statistics (1) under the condition

$$mn^{-1} \to \beta,\ 0<\beta<1 \tag{2}$$

as $n\to\infty$.

The general theorem

Let λ stand for the Radon measure (i.e. $\lambda(A)<\infty$ for all $A\in\mathcal{Q}$) having no atoms on the measurable space $(\mathcal{X}, \mathcal{Q})$ and $P_\lambda(A)=P_\lambda(\omega, A)$ denotes the Poisson measure with intensity λ, i.e.

$$P[P_\lambda(A)=k]=[k!]^{-1}\lambda^k(A)e^{-\lambda(A)},\ k=0,1,2,\ldots$$

and also events $\{\omega: P_\lambda(A_1)=q_1\}, \ldots, \{\omega: P_\lambda(A_p)=q_p\}$ are independent for any positive integer p, any disjoint subsets $A_1,\ldots,A_p\in\mathcal{Q}$ and any non-negative numbers $q_1,\ldots,q_p$. Further, put $\lambda_n(A)=n\pi_n(A)$, $A\in\mathcal{Q}$ and define a sequence of point measures

$$Q_n(A)=\delta_{X_{n1}}(A)+\delta_{X_{n2}}(A)+\ldots+\delta_{X_{nn}}(A) \tag{3}$$

based on the random variables $X_{n1}, X_{n2},\ldots,X_{nn}$ for any $A\in\mathcal{Q}$, with $\delta_x(A)$ standing for the Dirac measure.

The method of analysis proposed is based under condition (2) on the canonical Hoeffding decomposition ([1], p. 21)

$$U_n-\Theta(\lambda_n)=\sum_{c=1}^{m} \begin{bmatrix} m \\ c \end{bmatrix} \begin{bmatrix} n \\ c \end{bmatrix}^{-1} S_{nc}(g_{mc}) \tag{4}$$

Report BS-R9120
ISSN 0924-0659
CWI
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

where

$$\Theta(\lambda_n)=\int\cdots\int\Phi(y_1,\ldots,y_m)\prod_{s=1}^{m}\lambda_n(dy_s),$$

$$S_{nc}(g_{mc})=\sum_{1\leqslant i_1<\ldots<i_c\leqslant n} g_{mc}(X_{ni_1},\ldots,X_{ni_c}),$$

$$g_{mc}(x_1,\ldots,x_c)=\int\cdots\int\Phi(\lambda_1,\ldots,\lambda_m)\prod_{s=1}^{c}[\delta_{x_s}(dy_s)-\lambda_n(dy_s)]\times\prod_{s=c+1}\lambda_n(dy_s)$$

THEOREM 1. *Suppose that the weak convergence of random measures*

$$Q_n \xrightarrow{d} P_\lambda \tag{5}$$

takes place. Let the limit

$$\sigma^{-1}(m,n)g_{mc}(x_1,\ldots,x_c)\to g_\infty(x_1,\ldots,x_c) \text{ as } n\to\infty$$

exist for some $\sigma(m,n)>0$, *the functions* $g_\infty(x_1,\ldots,x_c)$ *being continuous,* $|g_\infty|\leqslant M^c$ *(M is a constant), and there exists a compact in* $\mathcal{X}^c$ *outside of which* g_∞ *equal zero; and also*

$$|\sigma^{-1}(m,n)g_{mc}(x_1,\ldots,x_c)|\leqslant M^c,\ c=0,1,2,\ldots$$

If condition (2) takes place then the weak convergence

$$\sigma^{-1}(m,n)[U_n-\Theta] \xrightarrow{d} \sum_{c=1}^{\infty}\frac{\beta^c}{c!}W_c(\lambda) \tag{6}$$

where

$$W_c(\lambda)=\int\cdots\int g_\infty(x_1,\ldots,x_c)P_\lambda(dx_1)\cdots P_\lambda(dx_c)$$

denotes c-dimensional stochastic Poisson-Wiener-Ito integral [5].

To prove (6) the representation (4) is rewritten into the form

$$\sigma^{-1}(m,n)[U_n-\Theta]=\sum_{c=1}^{m}[\frac{m}{n}]^c S_{nc}(g_\infty)+r_{nm}+\rho_{nm} \tag{7}$$

where

$$\rho_{nm}=\sum_{c=1}^{m}\begin{bmatrix}m\\c\end{bmatrix}\begin{bmatrix}n\\c\end{bmatrix}^{-1}S_{nc}\left[\sigma^{-1}(m,n)g_{mc}-g_\infty\right],$$

$$r_{nm}=\sum_{c=1}^{m}\left[\frac{m}{n}\right]^c\delta(c)S_{nc}(g_\infty),$$

in addition $\delta(1)=0$,

$$\delta(c)=\prod_{j=1}^{c-1}[1-jm^{-1}][1-jn^{-1}]^{-1}-1,\ c=2,3,\ldots$$

The detailed analysis shows that in (7)

$$r_{nm}\xrightarrow{P}0,\quad \rho_{nm}\xrightarrow{P}0$$

as $n\to\infty$. It follows from [4] and [6] that in (7)

$$\sum_{c=1}^{m}[mn^{-1}]^c S_{nc}(g_\infty) \xrightarrow{d} \sum_{c=1}^{\infty}\frac{\beta^c}{c!}W_c(\lambda)$$

Let's illustrate the more detailed estimates of this proof by the following example.

THE PARTICULAR CASE

Let in (1)

$$\Phi(x_1, \ldots, x_m) = h(x_1)\ldots h(x_m),$$

where $h:\mathcal{X}\to R$ is a continuous function with compact support. Denote

$$\mu_n = \int_{\mathcal{X}} h(x)\lambda_n(dx), \quad \mu = \int_{\mathcal{X}} h(x)\lambda(dx)$$

and let $\mu_n \neq 0$, $\mu \neq 0$. The assumption (5) implies that $\mu_n \to \mu$ as $n\to\infty$.

THEOREM 2. *If condition (5) holds then the weak convergence*

$$\mu_n^{-m} U_n \xrightarrow{d} \exp\left\{\int_{\mathcal{X}} \ln[1-\beta+\frac{\beta}{\mu}h(x)]P_\lambda(dx)\right\} \tag{8}$$

or, on the other form,

$$\mu_n^{-m}[U_n - \mu_n^m] \xrightarrow{d} \sum_{c=1}^{\infty} [\frac{\beta}{\mu}]\frac{1}{c!}\int_{\mathcal{X}}\cdots\int_{\mathcal{X}} \prod_{j=1}^{c} [h(x_j)-\mu] \prod_{j=1}^{c} P_\lambda(dx_j) \tag{9}$$

PROOF. Under the above conditions

$$g_{mc}(x_1, \ldots, x_c) = \mu_n^{m-c}[h(x_1)-\mu_n]\cdots[h(x_c)-\mu_n],$$
$$\Theta = \mu_n^m$$

Then one can choose $\sigma(m, n) = \mu_n^m$ and hence

$$g_\infty(x_1, \ldots, x_c) = \mu^{-c}[h(x_1)-\mu]\cdots[h(x_c)-\mu]$$

Reduce then representation (4) to the form

$$\mu_n^{-m}U_n = \sum_{c=0}^{m} \begin{bmatrix} m \\ c \end{bmatrix} \begin{bmatrix} n \\ c \end{bmatrix}^{-1} \mu_n^{-c} \sum_{1\leq i_1<\ldots<i_c\leq n} [h(X_{ni_1})-\mu_n]\cdots[h(X_{ni_c})-\mu_n]$$

Whence we derive as $n\to\infty$

$$\mu_n^{-m}U_n = \sum_{c=0}^{n} [\frac{m}{n\mu_n}]^c \sum_{1\leq i_1<\ldots<i_c\leq n} [h(X_{ni_1})-\mu_n]\cdots[h(X_{ni_c})-\mu_n] + R_n(\omega)$$

where $R_n(\omega) \xrightarrow{P} 0$. Then making use of the generating function for a symmetric polynomial we have

$$\mu_n^{-m}U_n = \prod_{j=1}^{n}\left[1+\frac{m}{n\mu_n}[h(X_{nj})-\mu_n]\right] + R_n(\omega) \tag{10}$$

In (10)

$$\prod_{j=1}^{n}\left[1+\frac{m}{n\mu_n}[h(X_{nj})-\mu_n]\right] =$$
$$= \exp\left\{\sum_{j=1}^{n} \ln\left[1+\frac{m}{n\mu_n}[h(X_{nj})-\mu_n]\right]\right\} =$$
$$= \exp\left\{\int_{\mathcal{X}} \ln[1-\frac{m}{n}+\frac{m}{n\mu_n}h(x)]Q_n(dx)\right\}$$

Under the condition (5) the weak convergence

$$\int_{\mathcal{X}} \ln[1 - \frac{m}{n} + \frac{m}{n\mu_n} h(x)]\, Q_n(dx) \xrightarrow{d} \int_{\mathcal{X}} \ln[1 - \beta + \frac{\beta}{\mu} h(x)] P_\lambda(dx)$$

takes place. This relation taken together with (10) implies (8).

Note that

$$Eexp\left\{ is \int_{\mathcal{X}} \ln[1 - \beta + \frac{\beta}{\mu} h(x)] P_\lambda(dx) \right\} =$$

$$\exp\left\{ \int_{\mathcal{X}} [e^{is \ln[1-\beta+\frac{\beta}{\mu}h(x)]} - 1] \lambda(dx) \right\}$$

for all $s \in R$.

The second statement from (9) follows directly from (6).

THE WEAK CONVERGENCE OF VON MISES FUNCTIONAL.

Consider the von Mises functional

$$V_n = n^{-m} \sum_{i_1=1}^{n} \sum_{i_m=1}^{n} \Phi(X_{ni_1}, \ldots, X_{ni_m})$$

and denote

$$EV_n = \int \cdots \int V_n(x_1, \ldots, x_n) \lambda_n(dx_1) \cdots \lambda_n(dx_n),$$

where $V_n(x_{n1}, \ldots, x_{nn}) = V_n$. We suppose that $|EV_n| < \infty$. V_n satisfies decomposition similar to (4)

$$V_n - EV_n = \sum_{c=1}^{m} \begin{bmatrix} m \\ c \end{bmatrix} n^{-c} \sum_{i_1=1}^{n} \cdots \sum_{i_c=1}^{n} g_{mc}(X_{ni_1}, \ldots, X_{ni_c})$$

Further, by means of measure Q_n the expression in the right-hand side of the latter relation can be represented in the form

$$V_n - EV_n = \sum_{c=1}^{m} \begin{bmatrix} m \\ c \end{bmatrix} n^{-c} \int \cdots \int g_{mc}(x_1, \ldots, x_c) \prod_{j=1}^{c} Q_n(dx_j)$$

Thus, under the conditions of Theorem 1 one can state the weak convergence

$$\sigma^{-1}(m,n)[V_n - EV_n] \xrightarrow{d} \sum_{c=1}^{\infty} \frac{\beta^c}{c!} \int \cdots \int g_{\infty}(x_1, \ldots, x_c) \prod_{j=1}^{c} P_\lambda(dx_j)$$

as $n \to \infty$ holds, i.e. the weak limits for U_n and V_n coincide.

REFERENCES

1. V.S. KOROLYUK and YU. V. BOROVSKIKH, *Theory of U-statistics*, Kiev, Naukova Dumka Publishers, 1989 (in Russian).
2. V.S. KOROLYUK and YU.V. BOROVSKIKH, *U-statistics with increasing degrees*, Doklady AN SSSR, 314 (1990), No. 3, pp. 547-551 (in Russian).
3. A.J. VAN ES and R. HELMERS, *Elementary symmetric polynomials of increasing order*, Probab. Theory and Related Fields, 80 (1988), No. 1, pp. 21-35.
4. I. KUBO and MUSTAFILD, *Law of small numbers and limit theorems of symmetric statistics*, Fifth Int. Vilnius Conf. on Probab. Theory and Math. Stat., Vilnius, June 26-July 1, 1989. Abst. of Commun., 1989, vol. 1, pp. 292-293.
5. Y. ITO and I. KUBO, *Calculus on Gaussian and Poisson white noises*, Nagoya Math. J., 111 (1988), No.1, pp. 41-84.
6. V.S. KOROLYUK and YU. V. BOROVSKIKH, *Algebraic decomposition and asymptotics for permanent of random matrix*, Asymptotic and Applied Problems in the Theory of Random Evolutions, Kiev, Inst. Math. Ukrain. Acad. Sci., 1990, pp. 68-74 (in Russian).