Available online at www.sciencedirect.com
SCIENCE DIRECT®
Statistics & Probability Letters 69 (2004) 261–269
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Almost sure limit theorems for U-statistics

Hajo Holzmann[1], Susanne Koch*, Aleksey Min[2]

Institut für Mathematische Stochastik, Georg-August-Universität Göttingen, Maschmühlenweg 8-10, 37073 Göttingen, Germany

Received 4 March 2004; received in revised form 5 April 2004

Available online 6 July 2004

Abstract

We relax the moment conditions from a result in almost sure limit theory for U-statistics due to Berkes and Csaki [(Stochastic Process. Appl. 94 (2001) 105)]. We extend this result to the case of convergence to stable laws and also prove a functional version.

MSC: 60F05; 60F17

Keywords: Almost sure limit theorem; Functional limit theorem; Stable distributions; U-statistics

1. Introduction

U-statistics generalize the concept of the sample mean of independent identically distributed (i.i.d.) random variables. The statistical interest in U-statistics stems from the fact that they form a class of unbiased estimators of a certain parameter with minimal variances. We begin by introducing some notation and recalling the concept of U-statistics.

Let $X_1, X_2, \dots$ be i.i.d random variables with common distribution function $F(x)$. Let $m \geqslant 1$ and let $h : \mathbb{R}^m \to \mathbb{R}$ be a measurable function symmetric in its arguments. The U-statistic with kernel h is defined by

$$U_n(h) = \binom{n}{m}^{-1} \sum_{1 \leqslant i_1 < i_2 < \cdots < i_m \leqslant n} h(X_{i_1}, \dots, X_{i_m}), \quad n \geqslant m.$$

*Corresponding author. Tel.: +49-551-3913509; fax: +49-551-3913505.
E-mail address: skoch@math.uni-goettingen.de (S. Koch).
[1] Acknowledges the financial support of the graduate school *Gruppen und Geometrie*.
[2] Acknowledges the financial support of the graduate school *Strömungsinstabilitäten und Turbulenz*.

0167-7152/$ - see front matter
doi:10.1016/j.spl.2004.06.019

The kernel h is called degenerate with respect to $F(x)$ if for all $1 \leqslant j \leqslant m$,

$$\int_{\mathbb{R}} h(x_1, \ldots, x_m)\, \mathrm{d}F(x_j) \equiv 0, \quad \text{where } -\infty < x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m < \infty.$$

Let

$$\theta = Eh(X_1, \ldots, X_m)$$

and for $i = 0, \ldots, m$ let

$$\bar{h}_i(x_1, \ldots, x_i) = Eh(x_1, \ldots, x_i, X_{i+1}, \ldots, X_m),$$
$$h_i(x_1, \ldots, x_i) = \sum_{k=0}^{i} (-1)^{i-k} \sum_{(j_1, \ldots, j_k) \subset \{1, \ldots, i\}} \bar{h}_k(x_{j_1}, \ldots, x_{j_k}),$$
$$\zeta_i = \operatorname{Var}(h_i(X_1, \ldots, X_i)).$$

Note that $\bar{h}_0 = \theta$ and $\bar{h}_m(x_1, \ldots, x_m) = h(x_1, \ldots, x_m)$. Furthermore, the h_i are degenerate for $i = 1, \ldots, m$ (see Denker, 1985). If $\zeta_1 > 0$ the U-statistic $U_n(h)$ is called non-degenerate and degenerate otherwise. The smallest integer c for which $\zeta_c > 0$ is called the critical parameter of the U-statistic $U_n(h)$. Without loss of generality we will assume $\theta = 0$ throughout the rest of this article.

The theory of U-statistics started to develop intensively after Hoeffding's (1948) fundamental article. He showed asymptotic normality of non-degenerate U-statistics under the assumption

$$Eh^2(X_1, \ldots, X_m) < \infty, \tag{1}$$

using the following representation of U-statistics:

$$U_n(h) = mU_n(h_1) + \sum_{k=2}^{m} \binom{m}{k} U_n(h_k), \tag{2}$$

where $mU_n(h_1)$ is a sum of i.i.d. random variables and $U_n(h_2), \ldots, U_n(h_m)$ are U-statistics with degenerate kernels.

In the degenerate case $n^{c/2} U_n(h)$ weakly converges to a multiple Wiener integral whenever h is square integrable (see e.g. Denker, 1985). Here c is the critical parameter of $U_n(h)$.

In the late 1980s, Brosamler (1988) and Schatte (1988) independently proved a new type of limit theorem. This type of statement extends the classical central limit theorem in the i.i.d. case to a pathwise version and is therefore called an almost sure central limit theorem (ASCLT). In the 1990s, many studies were done to prove almost sure limit theorems (ASLT) in different situations, for example in the case of independent but not necessarily identically distributed random variables (see Berkes and Dehling, 1993). Excellent surveys on this topic may be found in Atlagh and Weber (2000) as well as in Berkes (1998). Recently Berkes and Csaki (2001) obtained a general result in almost sure limit theory. They used it to prove almost sure versions of several classical limit theorems. In particular they stated the following theorem for U-statistics.

Theorem 1.1. *Let c be the critical parameter of the U-statistic $U_n(h)$. Under assumption* (1),

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbb{1}_{\{k^{c/2} U_k(h) < x\}} = G(x) \ \text{a.s. for any } x \in C_G,$$

where G is the limit distribution of $n^{c/2} U_n(h)$ and C_G denotes the set of continuity points of G.

In the present note we relax the moment condition in Theorem 1.1 and extend the statement in two directions. First we will obtain an ASLT with stable limiting distribution for a non-degenerate U-statistic. Furthermore we extend Theorem 1.1 to a functional version.

2. Preliminaries

Let $(Y_n)_{n\geqslant 1}$ be a sequence of random elements taking values in a Polish space $(\mathscr{S}, d)$ and let G be a probability measure on the Borel σ-field in $\mathscr{S}$. We say that $(Y_n)_{n\geqslant 1}$ satisfies the ASLT with limiting distribution G if with probability 1,

$$(\log n)^{-1} \sum_{k=1}^{n} \delta_{Y_k}/k \Rightarrow G, \quad n \to \infty.$$

Here δ_{Y_k} is the Dirac measure at Y_k and "$\Rightarrow$" denotes weak convergence of measures. Throughout this note the following lemma will be of fundamental interest.

Lemma 2.1. *Let $(Y_n)_{n\geqslant 1}$ be a sequence of $\mathscr{S}$-valued random elements which satisfies the ASLT with some limiting distribution G. Assume that Z_n is another sequence of $\mathscr{S}$-valued random elements on the same probability space such that almost surely, $d(Y_n, Z_n) \to 0$. Then Z_n also satisfies the ASLT with limiting distribution G.*

Proof. By a well-known principle in almost sure limit theory (see e.g. Lacey and Philipp, 1990), $(Y_n)_{n\geqslant 1}$ satisfies the ASLT with limiting distribution G, if and only if

$$(\log n)^{-1} \sum_{k=1}^{n} \frac{1}{k} \Psi(Y_k(\omega)) \to \int_{\mathscr{S}} \Psi(x) \, \mathrm{d}G(x) \quad \text{a.s.}$$

for any bounded Lipschitz function Ψ. Using the Lipschitz property of Ψ and the assumption that $d(Y_n, Z_n) \to 0$ we conclude that

$$\frac{1}{\log n} \left| \sum_{k=1}^{n} \frac{1}{k} [\Psi(Y_k(\omega)) - \Psi(Z_k(\omega))] \right| \leqslant \frac{C}{\log n} \sum_{k=1}^{n} \frac{1}{k} d(Y_k(\omega), Z_k(\omega)) \to 0 \quad \text{a.s.},$$

where C is a Lipschitz constant for Ψ. This proves the lemma. □

In the sequel we will make use of the following lemma which is a consequence of a more general result due to Giné and Zinn (1992).

Lemma 2.2. *Let $h(x_1, \dots, x_l)$ be measurable and degenerate. Let $q \in \left(\frac{l}{2}, l\right)$. If*

$$E\,|\,h(X_1, \dots, X_l)|^{l/q} < \infty, \tag{3}$$

then with probability 1

$$n^{-q} \sum_{1 \leqslant i_1 < \cdots < i_l \leqslant n} h(X_{i_1}, \dots, X_{i_l}) \to 0.$$

3. Relaxing the moment assumption

In this section, we will relax the moment assumption of Theorem 1.1. For the weak convergence of $n^{c/2} U_n(h)$ (where c denotes the critical parameter of $U_n(h)$) Koroljuk and Borovskich (1994) weakened assumption (1) to

$$E\,|\,h_k(X_1, \dots, X_k)|^{2k/(2k-c)} < \infty, \quad k = c, \dots, m. \tag{4}$$

We are going to prove the validity of Theorem 1.1 under these assumptions:

Theorem 3.1. *Let c be the critical parameter of the U-statistic $U_n(h)$. If* (4) *is satisfied then the statement of Theorem* 1.1 *is true.*

Proof. First of all note that, if c is the critical parameter of $U_n(h)$, then the functions $h_i(x_1, \dots, x_i) = 0$ a.s. for all $i = 1, \dots, c-1$, and so from the Hoeffding decomposition,

$$n^{c/2} U_n(h) - n^{c/2} \binom{m}{c} U_n(h_c) = n^{c/2} \sum_{k=c+1}^{m} \binom{m}{k} U_n(h_k). \tag{5}$$

By Theorem 1.1, $n^{c/2} \binom{m}{c} U_n(h_c)$ satisfies the ASLT.

To complete the proof it suffices to show that the sum on the right-hand side of (5) converges to zero a.s. This follows from Lemma 2.2 by letting $l = k$ and $q = k - c/2$ for $k = c+1, \dots, m$. An application of Lemma 2.1 finishes the proof. $\square$

4. Convergence to stable distributions

As we shall see in this section, under some mild moment conditions weak convergence of a sequence of non-degenerate U-statistics to a stable limit distribution implies the validity of the corresponding ASLT. Let G_α denote a stable law with characteristic exponent α.

Theorem 4.1. *Let for some* $\alpha \in (1, 2]$,

$$\frac{n^{1-1/\alpha}}{mL(n)} U_n(h) - A_n \Rightarrow G_\alpha, \tag{6}$$

where $L(n)$ is a slowly varying function for which $\liminf_{n \to \infty} L(n) > 0$. *If*

$$E\,|\,h_k(X_1, \dots, X_k)|^{\alpha k/(\alpha(k-1)+1)} < \infty, \quad k = 2, \dots, m, \tag{7}$$

then

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbb{1}_{\{(k^{1-1/\alpha})/U_k(h)/(mL(k))-A_k<x\}} = G_\alpha(x) \quad a.s.$$

Remarks. (1) Assumption (6) is very common in almost sure limit theory, when one deduces an ASLT from the validity of the corresponding weak limit theorem (see e.g. Berkes and Dehling, 1993).

(2) One has weak convergence in (6) if the distribution function of $h_1(X_1)$ belongs to the domain of attraction of G_α and if the moment condition

$$E\,|\,h(X_1,\ldots,X_m)|^{2\alpha/(\alpha+1)} < \infty \tag{8}$$

holds (see Heinrich and Wolf, 1993).

(3) It is not difficult to show that if the kernel h satisfies (8), then (7) holds.

(4) Theorem 4.1 will be true for any slowly varying function, if $E\,|\,h_k(X_1,\ldots,X_k)|^{p_k} < \infty$ for some $p_k > \alpha k/(\alpha(k-1)+1)$, $k = 2,\ldots,m$.

Proof of Theorem 4.1. Let us start by showing that

$$\frac{n^{1-1/\alpha}}{mL(n)}(U_n(h) - mU_n(h_1)) \to 0 \quad \text{a.s.} \tag{9}$$

Indeed

$$\frac{n^{1-1/\alpha}}{mL(n)}(U_n(h) - mU_n(h_1)) = \sum_{k=2}^{m} \binom{m}{k} \frac{n^{1-1/\alpha}}{mL(n)} U_n(h_k)$$

and letting $q = k - 1 + 1/\alpha$ and $l = k$ we can apply Lemma 2.2. This proves (9). Making use of (6) and (9), we also conclude that

$$\frac{n^{1-1/\alpha}}{L(n)} U_n(h_1) - A_n \Rightarrow G_\alpha. \tag{10}$$

It is known that weak convergence of normalized sums of real-valued i.i.d. random variables to some stable law G_α implies the corresponding ASLT (see Ibragimov and Lifshits, 1999). Hence (10) implies

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbb{1}_{\{k^{1-1/\alpha}/U_k(h_1)/L(k)-A_k<x\}} = G_\alpha(x) \quad \text{a.s.}$$

An application of Lemma 2.1 completes the proof. □

5. Functional version of the ASLT

5.1. The non-degenerate case

In the non-degenerate case the functional version of Theorem 1.1 can be deduced directly from Theorem 2 of Lacey and Philipp (1990), which deals with sums of i.i.d. random variables.

Throughout this section we assume that (1) holds. Let $D[0,1]$ denote the space of cadlag functions on [0,1] and let W denote the Wiener measure on $D[0,1]$ (see Billingsley, 1999). We introduce the following $D[0,1]$-valued random functions

$$Y_n(t) = \begin{cases} (m\sqrt{n\zeta_1})^{-1}\lfloor nt\rfloor U_{\lfloor nt\rfloor}(h) & : \quad t \in [m/n, 1], \\ 0 & : \quad t \in [0, \frac{m}{n}[, \end{cases} \tag{11}$$

where $\lfloor \cdot \rfloor$ denotes the integer part of a real number.

Theorem 5.1. *Suppose that $Eh^2(X_1, \ldots, X_m) < \infty$ and $\zeta_1 > 0$, then the sequence of random functions defined in* (11) *satisfies the ASLT with limiting measure W.*

Proof. First we want to see that, under suitable normalization, the difference between $U_n(h)$ and its first term in the Hoeffding decomposition (2) tends to 0 a.s. Since

$$U_n(h) - mU_n(h_1) = \sum_{k=2}^{m} \binom{m}{k} U_n(h_k)$$

and all kernels h_c on the right-hand side are degenerate with degree $c \geqslant 2$, we can apply Lemma 2.2 with $q = c - 1/2$ and $l/q = 3/2$ to conclude that $n^{1/2}(U_n(h) - mU_n(h_1)) \to 0$ a.s. Then also

$$n^{-1/2} \max_{m \leqslant k \leqslant n} k(U_k(h) - mU_k(h_1)) \to 0 \quad \text{a.s.} \tag{12}$$

Let

$$Z_n(t) = \frac{1}{\sqrt{n\zeta_1}} \lfloor nt\rfloor U_{\lfloor nt\rfloor}(h_1), \quad t \in [0,1].$$

Note that $nU_n(h_1) = \sum_{i=1}^{n} \bar{h}(X_i)$ is a sum of i.i.d. random variables. From the ASCLT for i.i.d. random variables (see Lacey and Philipp, 1990, Theorem 2), the $Z_n(t)$ satisfy the ASLT with limiting measure W. From (12) we deduce that $\|Y_n - Z_n\|_\infty \to 0$, where $\|\cdot\|_\infty$ denotes the supremum norm on $D[0,1]$. Hence we can apply Lemma 2.1 to $Y_n(t)$ and $Z_n(t)$ to prove the theorem. $\square$

5.2. The degenerate case

In order to obtain a functional ASLT in the degenerate case, we need the following version of a general result by Berkes and Csaki (2001) for function-valued random elements.

Theorem 5.2. *Let X_n, $n \geqslant 1$, be independent random variables taking values in some measurable space $(E, \mathscr{B})$ and let $f_l : E^l \to D[0,1]$, $l \in \mathbb{N}$, be measurable mappings such that $f_l(X_1, \ldots, X_l) \Rightarrow G$ for some distribution G on $D[0,1]$. Assume that for each $1 \leqslant k < l$, where $l - k \geqslant m$ for some fixed m, there exists a mapping $f_{k,l} : E^{l-k} \to D[0,1]$ such that*

$$E(\|f_l(X_1, \ldots, X_l) - f_{k,l}(X_{k+1}, \ldots, X_l)\|_\infty) \leqslant C\left(\frac{k}{l}\right)^\alpha$$

for some C, $\alpha > 0$. *Then the sequence* $(f_l(X_1, \ldots, X_l))_{l \geqslant 1}$ *satisfies the ASLT with limiting distribution* G.

Proof. The proof proceeds along the lines of the proof in Berkes and Csaki (2001). $\square$

From now on we restrict ourselves to kernels of degree 2. In this case the distribution invariance principle for degenerate U-statistics was obtained by Neuhaus (1977) (for the general case see Denker et al., 1985). It states that for a degenerate kernel $h : \mathbb{R}^2 \to \mathbb{R}$ the sequence of $D[0,1]$-valued random elements,

$$\frac{2}{n} \sum_{1 \leqslant i < j \leqslant \lfloor nt \rfloor} h(X_i, X_j) \tag{13}$$

converges in distribution to

$$\sum_{k \geqslant 1} \lambda_k (w_k(t) - t)^2,$$

where the w_k are independent Brownian motions and the λ_k are the eigenvalues of the integral operator associated with h (see Neuhaus, 1977). The distribution of this limiting process will be denoted by G.

Theorem 5.3. *Assume that the kernel* $h : \mathbb{R}^2 \to \mathbb{R}$ *is degenerate and satisfies* (1). *Then the sequence defined in* (13) *satisfies the ASLT with limiting distribution* G.

Proof. We want to apply Theorem 5.2. Denote $\|h\|_2^2 = Eh^2(X_1, X_2)$ and let

$$f_l(x_1, \ldots, x_l)(t) = \frac{2}{l} \sum_{1 \leqslant i < j \leqslant \lfloor lt \rfloor} h(x_i, x_j),$$

$$f_{k,l}(x_{k+1}, \ldots, x_l)(t) = \frac{2}{l} \sum_{k+1 \leqslant i < j \leqslant \lfloor lt \rfloor} h(x_i, x_j),$$

where $l - k \geqslant 2$, $t \in [0,1]$ and the empty sum is 0. By Theorem 5.2, it will be enough to show that

$$E(\|f_l(x_1, \ldots, x_l) - f_{k,l}(x_{k+1}, \ldots, x_l)\|_\infty^2) \leqslant Ck/l. \tag{14}$$

Evidently,

$$\|f_l(x_1, \ldots, x_l) - f_{k,l}(x_{k+1}, \ldots, x_l)\|_\infty^2 \leqslant 4l^{-2} E \left[\max_{2+k \leqslant n \leqslant l} \left(\sum\nolimits_n' h(x_i, x_j) \right)^2 \right.$$
$$\left. + \max_{2 \leqslant n < 2+k} \left(\sum_{1 \leqslant i < j \leqslant n} h(x_i, x_j) \right)^2 \right],$$

where $\sum_n'$ denotes summation over all indices $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant n$ with $i < j$. Observe that due to the degeneracy of h,

$$\sum\nolimits_n' h(X_i, X_j), \quad 2 + k \leqslant n \leqslant l,$$

is a martingale with respect to the canonical filtration $\mathscr{F}_n = \sigma(X_1, \ldots, X_n)$. Using Doob's inequality and once more the degeneracy of h we estimate

$$\begin{aligned} E\left[\max_{2+k \leqslant n \leqslant l} \left(\sum\nolimits_n' h(X_i, X_j)\right)^2\right] &\leqslant 4E\left(\sum\nolimits_l' h(X_i, X_j)\right)^2 \\ &= 4\sum\nolimits_l' Eh^2(X_i, X_j) \\ &\leqslant 4kl\|h\|_2^2. \end{aligned} \tag{15}$$

Similarly,

$$E\left[\max_{2 \leqslant n < 2+k} \left(\sum_{1 \leqslant i < j \leqslant n} h(x_i, x_j)\right)^2\right] \leqslant 4\binom{k+2}{2}\|h\|_2^2. \tag{16}$$

From (15) and (16) it follows that,

$$\begin{aligned} E(\|f_l(x_1, \ldots, x_l) - f_{k,l}(x_{k+1}, \ldots, x_l)\|_\infty^2) &\leqslant 4l^{-2}4\|h\|_2^2\left(kl + \binom{k+2}{2}\right) \\ &\leqslant Ck/l. \end{aligned}$$

The theorem is proved. □

Acknowledgements

We are grateful to H. Dehling for drawing our attention to a helpful publication by E. Giné and J. Zinn.

References

Atlagh, M., Weber, M., 2000. Le théorème central limite presque sûr. Exposition. Math. 18, 97–126.

Berkes, I., 1998. Results and problems related to the pointwise central limit theorem. In: Szyszkowicz, B. (Ed.), Proceedings of the International Conference on Asymptotic Methods in Probability and Statistics. Elsevier, Amsterdam, pp. 59–96.

Berkes, I., Csaki, E., 2001. A universal result in almost sure limit theory. Stochastic Process. Appl. 94, 105–134.

Berkes, I., Dehling, H., 1993. Some limit theorems in log density. Ann. Probab. 21, 1640–1670.

Billingsley, P., 1999. Convergence of Probability Measures, 2nd Edition. Wiley, New York.

Brosamler, G.A., 1988. An almost everywhere central limit theorem. Math. Proc. Cambridge Philos. Soc. 104, 561–574.

Denker, M., 1985. Asymptotic Distribution Theory in Nonparametric Statistics. Vieweg & Sohn, Braunschweig/Wiesbaden.

Denker, M., Grillenberger, C., Keller, G., 1985. A note on invariance principles for v. Mises statistics. Metrika 32, 197–214.

Giné, E., Zinn, J., 1992. Marcinkiewicz type laws of large numbers and convergence of moments for U-statistics. In: Dudley, R.M., Hahn, M.G., Kuelbs, J. (Eds.), Probability in Banach Spaces, 8: Proceedings of the Eighth International Conference, Birkhäuser, pp. 273–291.

Heinrich, L., Wolf, W., 1993. On the convergence of U-statistics with stable limit distribution. J. Multivariate Anal. 44, 226–278.

Hoeffding, W., 1948. A class of statistics with asymptotically normal distribution. Ann. Math. Statist. 39, 325–346.
Ibragimov, I.A., Lifshits, M.A., 1999. On almost sure limit theorems. Theory Probab. Appl. 44, 254–272.
Koroljuk, V.S., Borovskich, Yu.V., 1994. Theory of *U*-Statistics. Kluwer Academic Publishers, Dordrecht, Boston, London.
Lacey, M., Philipp, W., 1990. A note on the almost sure central limit theorem. Statist. Probab. Lett. 9, 201–205.
Neuhaus, G., 1977. Functional limit theorems for *U*-statistics in the degenerate case. J. Multivariate Anal. 7, 424–439.
Schatte, P., 1988. On strong versions of the central limit theorem. Math. Nachr. 137, 249–256.