Contents

Vol. 19, No. 4, 2009

Distributed worldwide by Springer. *Siberian Advances in Mathematics* ISSN 1055-1344.

ISSN 1055-1344, Siberian Advances in Mathematics, 2009, Vol. 19, No. 4, pp. 225–232. © Allerton Press, Inc., 2009.

Strong Laws for Certain Types of U-statistics Based on Negatively Associated Random Variables

K. Budsaba[1], P. Chen[2], K. Panishkan[3], and A. Volodin[4]**

[1]Thammasat University Rangsit Center, Pathumthani, 12121 Thailand
[2]Jinan University , Guangzhou, 510630 China
[3]Silpakorn University, Nakhon Pathom, 73000 Thailand
[4]University of Western Australia, Crawley, WA 6009, Australia

Received May 4, 2009

Abstract—We establish the Marcinkiewicz-Zygmund-type strong laws of large numbers for certain class of multilinear U-statistics based on negatively associated random variables.

DOI: 10.3103/S1055134409040014

Key words: *negative association, the Marcinkiewicz-Zygmund-type strong laws of large numbers, multilinear U-statistic.*

1. INTRODUCTION

The basic theory of U-statistics was developed by W. Hoeffding (1948). Since then the role of U-statistics has gradually increased both in Theoretical Probability and Statistics. We refer to the monograph by Koroljuk and Borovskich (1994) for not only the basic definitions, notions, and inequalities (moment and maximal inequalities), but also for laws of large numbers, weak convergence, functional limit theorems, approximations in limit theorems, asymptotic expansions, large deviations, and laws of the iterated logarithm for U-statistics based on a sequence of dependent random variables, that are discussed in detail in this book.

In the present paper, we provide strong laws of large numbers for certain multilinear-type U-statistics based on negatively associated random variables.

Recall that a finite family of random variables $\{X_i, 1 \le i \le n\}$ is said to be *negatively associated* (sometimes abbreviated to NA) if, for any disjoint subsets A and B of $\{1, 2, \cdots, n\}$ and any real coordinate-wise nondecreasing functions f on $\mathbf{R}^A$ and g on $\mathbf{R}^B$,

$$\mathrm{Cov}\,(f(X_i, i \in A), g(X_j, j \in B)) \le 0$$

whenever the covariance exists.

An infinite family is negatively associated if every finite subfamily is negatively associated.

This definition was introduced by Alam and Saxena (1981) and Joag-Dev and Proschan (1983). Joag-Dev and Proschan (1983) pointed and proved that a number of well known multivariate distributions posses the NA property, such as

(a) the multinomial distributions,

(b) the convolutions of various polynomial distributions,

(c) the multivariate hypergeometric distributions,

(d) the Dirichlet distributions,

(e) the generalized Dirichlet multinomial distributions,

(f) the negatively correlated normal distributions,

*The text was submitted by the authors in English.

**E-mail: andrei@maths.uwa.edu.au

(g) the permutation distributions,

(h) the joint sample distributions in the simple random sampling without replacement,

(i) the joint distributions of ranks.

Because of their wide applications in multivariate statistical analysis and Reliability Theory, the notion of negatively associated random variables attracts more and more attention. We refer the reader to Joag-Dev and Proschan (1983) for fundamental properties, to Newman (1984) for the central limit theorem, to Matula (1992) for the three series theorem, to Barbour et al. (1992) for Poisson approximation, to Shao (2000) for the moment and exponential inequalities, and to Shao and Su (1999) for the law of the iterated logarithm.

Trough this paper, C always stands for a positive constant which differs from one place to another.

2. MAIN RESULTS

Now we state the main results. The proofs will be presented in the next Section.

Theorem 1. *Let $\{X, X_n, n \geq 1\}$ be a negatively associated sequence of identically distributed random variables with $E|X|^p < \infty$ for some $p \geq 1$ and $EX = 0$. Let $\{k_n, n \geq 1\}$ be a sequence of positive integers and $\{a_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ be an array of real numbers satisfying the conditions*

$$k_n \leq Mn \ \text{for all}\ n \geq 1,$$

$$\sup_{n\geq 1} \max_{1\leq i\leq k_n} |a_{ni}| < \infty,$$

$$\sum_{i=1}^{k_n} a_{ni}^2 = O(n^{\delta}) \ \text{for some}\ \ \delta < \min\{1, 2/p\}.$$

Then, for every integer $m \geq 1$,

$$n^{-m/p} \sum_{1\leq i_1<\cdots<i_m\leq k_n} \prod_{j=1}^{m} a_{ni_j} X_{i_j} \to 0 \ a.s.$$

The second theorem sharpens Theorem 1 in the case $p = 1$.

Theorem 2. *Let $\{X, X_n, n \geq 1\}$ be a sequence of identically distributed negatively associated mean zero random variables. Let $\{a_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ be an array of real numbers under the conditions $k_n \leq Mn$ for some integer $M \geq 1$ and $\sup_{n\geq 1} \max_{1\leq i\leq k_n} |a_{ni}| < \infty$. Then, for every integer $m \geq 1$,*

$$n^{-m} \sum_{1\leq i_1<\cdots<i_m\leq k_n} \prod_{j=1}^{m} a_{ni_j} X_{i_j} \to 0 \ a.s.$$

3. LEMMAS

Hoffmann-Jørgensen's inequality (Hoffmann-Jørgensen, 1974) has now become a standard technique in proving strong limit theorems in the case of independent summands. But in the case of negatively associated random variables, we do not know whether the inequality of Hoffmann-Jørgensen holds or not. Recently, Shao (2000) proved exponential inequalities which may play the same role as Hoffmann-Jørgensen's inequality in the case of independent summands.

Lemma 1 (Shao, 2000). *Let $p \geq 1$ and let $\{Y_n, n \geq 1\}$ be a sequence of negatively associated mean zero random variables, with $E|Y_i|^p < \infty$ for every $1 \leq i \leq n$. Then*

$$E \max_{1\leq k\leq n} |\sum_{i=1}^{k} Y_i|^p \leq 2^{3-p} \sum_{i=1}^{n} E|Y_i|^p \quad \text{if}\ 1 \leq p \leq 2$$

and

$$E \max_{1\le k\le n} |\sum_{i=1}^{k} Y_i|^p \le 2(15p/\log p)^p \left\{ \left(\sum_{i=1}^{n} EY_i^2 \right)^{p/2} + \sum_{i=1}^{n} E|Y_i|^p \right\} \quad \text{if } p > 2.$$

Lemma 2 (Shao, 2000). *Let $\{Y_k, 1 \le k \le n\}$ be a sequence of negatively associated random variables with zero means and finite second moments. Let $S_k = \sum_{i=1}^{k} Y_i$ and $B_n = \sum_{i=1}^{n} EY_i^2$. Then, for all $x > 0$ and $a > 0$,*

$$P(\max_{1\le k\le n} |S_k| \ge x) \le 2P\left(\max_{1\le k\le n} |Y_k| \ge a \right) + 4\exp\left(-\frac{x^2}{8B_n} \right) + 4\left(\frac{B_n}{4(xa + B_n)} \right)^{x/(12a)}.$$

We also need the following result.

Lemma 3 (Wang et al., 2001). *Let $\{x_n, n \ge 1\}$ be an array of real numbers. Then, for all $n \ge m \ge 1$,*

$$\sum_{1\le i_1<\cdots<i_m\le n} \prod_{j=1}^{m} x_{i_j} = \sum_{\sum_{j=1}^{m} l_j r_j = m} A^{(m)}(l_j, r_j : 1 \le j \le m) \prod_{j=1}^{m} \left(\sum_{i=1}^{n} x_i^{r_j} \right)^{l_j},$$

where $A^{(m)}(l_j, r_j : 1 \le j \le m)$ and $\sum_{j=1}^{m} l_j r_j = m$ do not depend on n and $\{x_n, n \ge 1\}$.

4. PROOFS

In this Section, we present proofs of the main results of the paper.

Proof of Theorem 1. By the same argument as that in the proof of Theorem 3.1 in Li et al. (1995), we assume that $k_n = n$. We can also assume that $a_{ni} \ge 0$ for all $1 \le i \le n$ and $n \ge 1$. Otherwise, we can reduce the problem to the case of positive weights $\{a_{ni}^+, 1 \le i \le n, n \ge 1\}$ and $\{a_{ni}^-, 1 \le i \le n, n \ge 1\}$, where $a_{ni}^+ = \max(a_{ni}, 0)$ and $a_{ni}^- = \max(-a_{ni}, 0)$ are the respective positive and negative parts of a number.

Case 1 ($m = 1$). This case generalizes Theorem 3.1 in Li et al. (1995) to the case of NA random variables.

First of all, we show that

$$n^{-1/p} \sum_{i=1}^{n} a_{ni} X_i \to 0 \text{ in probability}. \tag{1}$$

We start with the case $1 \le p \le 2$. By Chebyshev's inequality, Lemma 1, and Hölder's inequality, we obtain that, for all $\varepsilon > 0$,

$$\begin{aligned} P(|\sum_{i=1}^{n} a_{ni} X_i| \ge \varepsilon n^{1/p}) &\le Cn^{-1} E|\sum_{i=1}^{n} a_{ni} X_i|^p \\ &\le Cn^{-1} \sum_{i=1}^{n} |a_{ni}|^p E|X|^p \le Cn^{-1} (\sum_{i=1}^{n} a_{ni}^2)^{p/2} n^{1-p/2} \\ &\le Cn^{(\delta-1)p/2} \to 0. \end{aligned}$$

We now consider the case $p > 2$. By Chebyshev's inequality and Lemma 1, we obtain that, for all $\varepsilon > 0$,

$$\begin{aligned} P(|\sum_{i=1}^{n} a_{ni} X_i| \ge \varepsilon n^{1/p}) &\le Cn^{-2/p} E|\sum_{i=1}^{n} a_{ni} X_i|^2 \\ &\le Cn^{-2/p} \sum_{i=1}^{n} a_{ni}^2 EX^2 \le Cn^{-2/p+\delta} \to 0. \end{aligned}$$

Hence, (1) holds.

By Theorem 3.2.1 in Stout (1974), it suffices to prove that

$$n^{-1/p}\sum_{i=1}^{n} a_{ni}X_i^s \to 0 \text{ a.s.},$$

where $\{X_n^s, n \geq 1\}$ is a symmetrized version of the sequence $\{X_n, n \geq 1\}$.

By the fundamental property of negatively associated random variables, one can prove (see Joag-Dev and Proschan, 1983) that $\{X_n^s, n \geq 1\}$ is also a sequence of negatively associated random variables. Hence we need only to prove the result for symmetric $\{X_n, n \geq 1\}$.

For an arbitrary fixed $\varepsilon > 0$, define $X'_{ni} = -\varepsilon z b^{-1} n^{1/p} I(X_i < -\varepsilon z b^{-1} n^{1/p}) + X_i I(|X_i| \leq \varepsilon z b^{-1} n^{1/p}) + \varepsilon z b^{-1} n^{1/p} I(X_i > \varepsilon z b^{-1} n^{1/p})$ and $X''_{ni} = X_i - X'_{ni}$ for $1 \leq i \leq n$ and $n \geq 1$, where $b = \sup_{n,i} |a_{ni}|$ and $0 < z < (\min\{1, 2/p\} - \delta)/12$.

Note that the condition $E|X|^p < \infty$ is equivalent to $\sum_{n=1}^{\infty} P(|X_n| > \alpha n^{1/p}) < \infty$ for any $\alpha > 0$. Hence, by the Borel-Cantelli lemma, $\sum_{i=1}^{n} |X''_{ni}|$ is bounded with probability 1. From here it follows that

$$n^{-1/p}|\sum_{i=1}^{n} a_{ni}X''_{ni}| \leq \max_{1\leq i\leq n} |a_{ni}| n^{-1/p} \sum_{i=1}^{n} |X''_{ni}| \to 0 \text{ a.s.} \tag{2}$$

If the statement

$$\limsup_{n\to\infty} n^{-1/p}|\sum_{i=1}^{n} a_{ni}X'_{ni}| \leq \varepsilon \text{ a.s.} \tag{3}$$

is true then, by (2),

$$\limsup_{n\to\infty} n^{-1/p}|\sum_{i=1}^{n} a_{ni}X_i| \leq \varepsilon \text{ a.s.}$$

Since $\varepsilon > 0$ is arbitrary, we obtain the statement of the theorem.

Hence it suffices to prove (3). By the Borel-Cantelli lemma, it suffices to show that

$$\sum_{n=1}^{\infty} P(|\sum_{i=1}^{n} a_{ni}X'_{ni}| \geq \varepsilon n^{1/p}) < \infty.$$

Further,

$$\begin{aligned} B_n &= \sum_{i=1}^{n} a_{ni}^2 E(X'_{ni})^2 \\ &= C\{n^{2/p} P(|X| > \varepsilon z b^{-1} n^{1/p}) + EX^2 I(|X| \leq \varepsilon z b^{-1} n^{1/p})\} \sum_{k=1}^{n} a_{nk}^2 \end{aligned}$$

$$\leq \begin{cases} Cn^{2/p-1+\delta}, & \text{if } 1 \leq p \leq 2 \\ Cn^{\delta}, & \text{if } p > 2 \end{cases} \leq Cn^{p'},$$

where $p' = \max\{2/p - 1 + \delta, \delta\}$, and $2/p - p' > 0$ by the assumption.

Take $x = \varepsilon n^{1/p}$ and $a = \varepsilon y n^{1/p}$, where $z < y < \min\{1, 2/p\} - \delta)/12$. By Lemma 2, we have that

$$\begin{aligned}
&P(|\sum_{i=1}^{n} a_{ni}X'_{ni}| \geq \varepsilon n^{1/p}) \\
&\leq 2P(\max_{1\leq i\leq n} |a_{ni}X'_{ni}| \geq a) + 4\exp\left(-\frac{x^2}{8B_n}\right) + 4\left(\frac{B_n}{4(xa+B_n)}\right)^{x/(12a)} \\
&\leq 2P(\max_{1\leq i\leq n} |a_{ni}X'_{ni}| \geq \varepsilon y n^{1/p}) + 4\exp\left(-Cn^{2/p-p'}\right) + Cn^{-(2/p-p')/(12y)}.
\end{aligned}$$

Note that

$$\sum_{n=1}^{\infty} \exp\left(-Cn^{2/p-p'}\right) < \infty \text{ and } \sum_{n=1}^{\infty} n^{-(2/p-p')/(12y)} < \infty.$$

Hence it remains to show

$$\sum_{n=1}^{\infty} P(\max_{1\leq i\leq n} |a_{ni}X'_{ni}| \geq \varepsilon y n^{1/p}) < \infty.$$

Note that

$$\max_{1\leq i\leq n} |a_{ni}X'_{ni}| \leq \sup_{n,i} |a_{ni}| \varepsilon z n^{1/p} (\sup_{n,i} |a_{ni}|)^{-1} = \varepsilon z n^{1/p} < \varepsilon y n^{1/p}.$$

Hence, $P(\max_{1\leq i\leq n} |a_{ni}X'_{ni}| \geq \varepsilon y n^{1/p}) = 0$ and we obtain the desired statement. □

Case 2 ($m \geq 2$). By Lemma 3, it suffices to show that

$$n^{-r/p} \sum_{i=1}^{n} a_{ni}^r X_i^r \to 0 \text{ a.s.} \tag{4}$$

for $r = 1, \cdots, m$. Note the following inequality which holds for all real numbers $\{b_i, i \geq 1\}$ and $r \geq 2$:

$$\left(\sum_{i=1}^{\infty} |b_i|^r\right)^{2/r} \leq \sum_{i=1}^{\infty} b_i^2.$$

Hence it suffices to prove (4) in the case $r = 1$ and $r = 2$.

If $r = 1$ then relation (4) holds by Case 1.

If $r = 2$ and $1 \leq p < 2$ then $2/p > 1$ and $E|X^2|^{p/2} < \infty$. Hence, by the Marcinkiewicz-Zygmund strong law of large number (see Sawyer, 1966),

$$n^{-2/p} \sum_{i=1}^{n} X_i^2 \to 0 \text{ a.s.}$$

Hence, (4) holds.

If $r = 2$ and $p \geq 2$ then, in order to prove (4), it suffices to show that

$$n^{-2/p} \sum_{i=1}^{n} a_{ni}^2 X_i^2 I(X_i \geq 0) \to 0 \text{ a.s.}$$

and

$$n^{-2/p} \sum_{i=1}^{n} a_{ni}^2 X_i^2 I(X_i < 0) \to 0 \text{ a.s.}$$

By the definition of negatively associated random variables (see Joag-Dev and Proschan, 1983), $\{X_n^2 I(X_n \geq 0), n \geq 1\}$ and $\{-X_n^2 I(X_n < 0), n \geq 1\}$ are also sequences of negatively associated random variables. Notice that

$$\sum_{i=1}^{n} (a_{ni}^2)^2 \leq \sup_{n,i} a_{ni}^2 \sum_{i=1}^{n} a_{ni}^2 = O(n^\delta) \text{ as } n \to \infty,$$

and $\delta \leq \min\{1, 2/p\} \leq \min\{1, 4/p\}$. Using Case 1 with $p/2$, we obtain

$$n^{-2/p} \sum_{i=1}^{n} a_{ni}^2 (X_i^2 I(X_i \geq 0) - EX^2 I(X \geq 0)) \to 0 \text{ a.s.}$$

and

$$n^{-2/p} \sum_{i=1}^{n} a_{ni}^2 (X_i^2 I(X_i < 0) - EX^2 I(X < 0)) \to 0 \text{ a.s.}$$

Taking the relation

$$n^{-2/p} \sum_{i=1}^{n} a_{ni}^2 \to 0 \text{ as } n \to \infty$$

into account, we obtain (4). □

Proof of Theorem 2. As in the previous proof, we assume without loss generality that $a_{ni} \geq 0$ for all $1 \leq i \leq n$ and $n \geq 1$.

Case 1 ($m = 1$). This case generalizes the result of Choi and Sung (1987) to the case of NA random variables. Note that the result of Choi and Sung (1987) sharpens Theorem 3.1 of Li et al. in the case $p = 1$. Fix $\Delta > 0$ and define

$$\begin{aligned} X_n^{(1)} &= X_n I(X_n > \Delta), \;\; X_n^{(2)} = X_n I(X_n < -\Delta), \\ X_n^{(3)} &= -\Delta I(X_n < -\Delta) + X_n I(|X_n| \leq \Delta) + \Delta I(X_i > \Delta), \\ X_n^{(4)} &= \Delta I(X_n < -\Delta) - \Delta I(X_n > \Delta). \end{aligned}$$

Then $X_n = X_n^{(1)} + X_n^{(2)} + X_n^{(3)} + X_n^{(4)}$, and by the definition of negatively associated random variables (see, Joag-Dev and Proschan, 1983), $\{X_n^{(j)}, n \geq 1\}$ is also a sequence of negatively associated random variables for each $j = 1, 2, 3, 4$.

By Theorem 1 of Matula (1992), we have

$$n^{-1} \Big| \sum_{i=1}^{n} a_{ni} X_i^{(1)} \Big| \leq \sup_{n,i} |a_{ni}| n^{-1} \sum_{i=1}^{n} X_i^{(1)} \to \sup_{n,i} |a_{ni}| EXI(X > \Delta) \text{ a.s.}$$

and

$$n^{-1} \Big| \sum_{i=1}^{n} a_{ni} X_i^{(2)} \Big| \leq -\sup_{n,i} |a_{ni}| n^{-1} \sum_{i=1}^{n} X_i^{(2)} \to -\sup_{n,i} |a_{ni}| EXI(X < -\Delta) \text{ a.s.}$$

Choosing Δ large enough, we have $\sup_{n,i} |a_{ni}| EXI(X > \Delta)$ and $-\sup_{n,i} |a_{ni}| EXI(X < -\Delta)$ can be made as small as desired. Also by choosing Δ large enough, we can make $|E \sum_{i=1}^{n} a_{ni} X_i^{(3)}|/n$ and $|E \sum_{i=1}^{n} a_{ni} X_i^{(4)}|/n$ as small as desired since $EX = 0$. It remains to show that

$$n^{-1} \sum_{i=1}^{n} a_{ni} (X_i^{(3)} - EX_i^{(3)}) \to 0 \text{ a.s.}$$

and

$$n^{-1}\sum_{i=1}^{n} a_{ni}(X_i^{(4)} - EX_i^{(4)}) \to 0 \text{ a.s.}$$

These relations follow from the fact that

$$\sum_{n=1}^{\infty} P(|\sum_{i=1}^{n} a_{ni}(X_i^{(3)} - EX_i^{(3)})| \geq \varepsilon n) < \infty$$

and

$$\sum_{n=1}^{\infty} P(|\sum_{i=1}^{n} a_{ni}(X_i^{(4)} - EX_i^{(4)})| \geq \varepsilon n) < \infty$$

for all $\varepsilon > 0$. In fact, taking $t > 2$, by Lemma 1, we have

$$\begin{aligned}
&\sum_{n=1}^{\infty} P(|\sum_{i=1}^{n} a_{ni}(X_i^{(3)} - EX_i^{(3)})| \geq \varepsilon n) \\
\leq\ & C\sum_{i=1}^{\infty} n^{-t}\left((\sum_{i=1}^{n} a_{ni}^2 E(X_i^{(3)} - EX_i^{(3)})^2)^{t/2} + \sum_{i=1}^{n} |a_{ni}|^t E|X_i^{(3)} - EX_i^{(3)}|^t\right) \\
\leq\ & C\sum_{n=1}^{\infty} n^{-t/2} + C\sum_{n=1}^{\infty} n^{-t+1} < \infty.
\end{aligned}$$

By the same argument,

$$\sum_{n=1}^{\infty} P(|\sum_{i=1}^{n} a_{ni}(X_i^{(4)} - EX_i^{(4)})| \geq \varepsilon n) < \infty.\square$$

Case 2 ($m \geq 2$) repeats the proof of the corresponding part of Theorem 1 and it is omitted. □

ACKNOWLEDGMENTS

The authors are grateful to the referee for carefully reading the manuscript and for providing some comments which helped them to improve the presentation. The research was supported by the National Natural Science Foundation of China and by the National Science and Engineering Research Council of Canada.

REFERENCES

Alam K. and Saxena K.M.L., 1981. Positive dependence in multivariate distribution. *Comm. Statist. A* 10, 1183-1196.

Barbour A. D., Holst L., and Janson S., 1992. *Poisson approximation*. Oxford university Press.

Choi, B.D., Sung, S.H., 1987. Almost sure convergence theorems of weighted sums of random variables. *Stochastic Anal. Appl.* 5, 365-377.

Hoeffding, W., 1948. A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* 19, 293-325.

Hoffmann-Jørgensen, J., 1974. Sums of independent Banach space valued random variables. *Studia Math.* 52, 159-186.

Joag-Dev K. and Proschan F., 1983. Negative association of random variables with applications. *Ann. Statist.* 11, 286-295.

Koroljuk, V. S.and Borovskich, Yu., 1994. *Theory of U-statistics*. Mathematics and its Applications, 273. Kluwer Academic Publishers Group, Dordrecht.

Li, D.L., Rao, M.B., Jiang, T.F., Wang X.C., 1995. Complete convergence and almost sure convergence of weighted sums of random variables. *J. Theoret. Probab.* 8, 49-76.

Matula P., 1992. A note on the almost sure convergence of sums of negatively dependent random variables. *Statist. Probab. Lett.* 15, 209-213.

Newman C.M., 1984. Asymptotic independence and limit theorems for positively and negatively dependent random variables. In Y.L.Tong (ed.), Inequalities in Statistics and Probability, Institute of Mathematical Statistics, Hayward, CA, pp.127-140.

Sawyer,S.,1966. Maximal inequalities of weak type. *Ann. Math.* 84,157-174.

Shao, Q.M., 2000. A comparison theorem on inequalities between negatively associated and independent random variables. *J. Theoret. Probab.* 13, 343-356.

Shao, Q.M. and Su C., 1999. The law of the iterated logarithm for negatively associated random variables. *Stoch. Process. Appl.* 83, 139-148.

Stout, W.F., 1974. *Almost sure convergence.* Academic Press, New York.

Wang Y.B., Yan J.G., Cheng F.Y., and Cai X.Z., 2001. On the strong stability for Jamison type weighted product sums of pairwise NQD series with different distribution. *Chinese Ann. Math.* 22(A), 701-706 (in Chinese).