The Annals of Probability
1994, Vol. 22, No. 4, 1715–1724

ON SOME ASYMPTOTIC PROPERTIES OF *U* STATISTICS AND ONE-SIDED ESTIMATES

BY ARUP BOSE AND RATAN DASGUPTA

Indian Statistical Institute

Let $\{X_i,\ 1 \le i \le n\}$ be independent and identically distributed random variables. For a symmetric function h of m arguments, with $\theta = Eh(X_1, \dots, X_m)$, we propose estimators θ_n of θ that have the property that $\theta_n \to \theta$ almost surely (a.s.) and $\theta_n \ge \theta$ a.s. for all large n. This extends the results of Gilat and Hill, who proved this result for $\theta = Eh(X_1)$. The proofs here are based on an almost sure representation that we establish for U statistics. As a consequence of this representation, we obtain the Marcinkiewicz–Zygmund strong law of large numbers for U statistics and for a special class of L statistics.

1. Introduction. Let $X, X_1, X_2, \dots$ be a sequence of independent and identically distributed (iid) observations from a distribution with finite mean μ. The usual estimate $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$ has the property that $\overline{X}_n \to \mu$ a.s. and it fluctuates around μ. However, in many practical situations it may be desirable to have an estimate μ_n of μ that is conservative in the sense that $\mu_n \to \mu$ a.s. and $\mu_n \ge \mu$ a.s. for all large n. We will then say μ_n converges to μ from above and write $\mu_n \to_+ \mu$ a.s. ($\mu_n \to_- \mu$ is defined in a similar manner). A candidate estimator for the convergence from above is one that puts more weight to the higher order statistics. Consider then the following estimator

$$\widehat{X}_n = \sum_{i=1}^{n} \left(\frac{1}{n} - \frac{n+1}{2n^{\alpha}} + \frac{i}{n^{\alpha}} \right) X_{(i)}, \tag{1.1}$$

where $X_{(1)} \le \cdots \le X_{(n)}$ is the order statistics of $X_1, \dots, X_n$ and $\alpha > 2$ is an appropriate constant. The following theorem was proved in Gilat and Hill (1992) (hereafter referred to as GH).

THEOREM 1.1 [Gilat and Hill (1992)]. *If* $E|X|^{1+\gamma} < \infty$ *for some* $\gamma > 0$, *then for any* α, $2 < \alpha < \min(2 + \gamma/(1+\gamma), 5/2)$, $\widehat{X}_n \to_+ \mu$ *a.s.*

The proof in GH is based on the following facts:

1. If $E|X|^{1+\gamma} < \infty$ for some $0 < \gamma < 1$, then $\overline{X}_n - \mu = o(n^{-\gamma/(1+\gamma)})$ a.s. This is known as the Marcinkiewicz–Zygmund strong law of large numbers; see, for example, Chow and Teicher (1978).

Received September 1993; revised November 1993.
AMS 1991 *subject classifications.* Primary 60F15; secondary 60G42, 62G05, 62G20, 62G30.
Key words and phrases. U statistics, L statistics, order statistics, Marcinkiewicz–Zygmund strong law, one-sided estimates.

2. If $E|X| < \infty$, then

$$\lim_{n\to\infty} \frac{2}{n^2} \sum_{i=1}^{n} iX_{(i)} = E\big(\max(X_1, X_2)\big) \quad \text{a.s.} \tag{1.2}$$

(This is Theorem 1.1 of GH with $k = 1$.)

Observe that $\widehat{X}_n$ may be expressed as

$$\widehat{X}_n = \overline{X}_n + \frac{1}{4n^{\alpha}} \sum_{i,j=1}^{n} |X_i - X_j| \tag{1.3}$$

and hence may also be identified as a U statistic, with varying kernel, of order 2.

We prove an almost sure representation theorem for U statistics (see Theorem 2.1), and use it to establish the Marcinkiewicz–Zygmund strong law of large numbers for U statistics (see Remark 2.1) and a special class of L statistics (see Remark 2.2). This may be used to give a different proof of Theorem 1.1. It may also be used to obtain the almost sure rate of convergence in Theorem 1.1 of GH (see Remark 2.2). In fact, representation (1.3) also suggests a way of extending the results of GH to the following situation.

Suppose we wish to estimate $\theta = Eh(X_1, \dots, X_m)$, where h is a symmetric function of its arguments. For instance, when θ is the population variance, we let $m = 2$ and $h(x_1, x_2) = (x_1 - x_2)^2/2$. We will use the notation $U_n(g)$ to denote the U statistics based on $X_1, \dots, X_n$ corresponding to the *symmetrized* version of the kernel function g. Consider the function

$$D(x_1, \dots, x_{2m}) = |h(x_1, \dots, x_m) - h(x_{m+1}, \dots, x_{2m})|.$$

Define the estimator

$$\theta_n = U_n(h) + a_n U_n(D), \tag{1.4}$$

where a_n is an appropriate sequence of positive constants converging to zero. We shall show that $\theta_n \to_+ \theta$ a.s (see Theorem 2.2).

It is interesting to note that the smaller the value of a_n, the lesser is the bias of the estimator. We will allow a wider choice of a_n than allowed by GH. In fact, as the proofs will show, our choice of a_n is rather tight.

We use a similar idea for the problem of quantile estimation. For a suitably constructed empirical distribution function G_n, we show that $G_n(x) \to_- F(x)$ a.s., where F is the cumulative distribution function of $h(X_1, \dots, X_m)$. The pth quantile of G_n is shown to be an upper estimate of the pth quantile of F (see Theorem 2.3).

In a subsequent paper, we will report finer asymptotic properties of our estimators.

2. Main results.

2.1. *An almost sure representation for U statistics.* If $U_n(h)$ is a U statistic with kernel h and $Eh^2 < \infty$, then it is known that

$$U_n = \widehat{U}_n + R_n, \tag{2.1}$$

where $\widehat{U}_n$ is the usual projection of U_n and $R_n = o(n^{-1}(\log n)^{\delta})$ a.s. for all $\delta > 1/2$; see, for example, Serfling [(1980), page 189]. We establish a similar representation under the weaker assumption $E|h|^{1+\gamma} < \infty$ for some $\gamma > 0$. This result is of independent interest. It will be used to obtain the Marcinkiewicz–Zygmund strong law of large numbers (to be abbreviated as MZSLLN) for U statistics and for a special class of L statistics. To state our result on U statistics, we will adopt the notations of Serfling (1980). Also, C will denote a generic positive constant throughout the paper.

THEOREM 2.1. *Suppose U_n is the U statistic based on the symmetric kernel h, where $E|h(X_1,\dots,X_m)|^{1+\gamma} < \infty$ for some $0 < \gamma \le 1$. Then*

$$U_n - \theta = \widehat{U}_n - \theta + R_{2n} + \cdots + R_{mn}, \tag{2.2}$$

where

$$R_{jn} = o\big(n^{-j\gamma/(1+\gamma)}(\log n)^{1/(1+\gamma)}(\log\log n)^{\delta}\big) \quad a.s. \tag{2.3}$$

for any $\delta > 1/(1+\gamma)$. Further, if for some $c > 1$, $\zeta_1 = \cdots = \zeta_{c-1} = 0$, then

$$\widehat{U}_n - \theta = R_{2n} = \cdots = R_{c-1,n} = 0 \quad a.s.$$

PROOF. Define as in Serfling [(1980), page 177]

$$\tilde{h}_1(x) = Eh(x, X_2,\dots,X_m) - \theta \tag{2.4}$$

and $\widehat{U}_n - \theta = (m/n)\sum_{i=1}^n \tilde{h}_1(X_i)$. Note that

$$U_n - \theta = \widehat{U}_n - \theta + R_n,$$

where

$$R_n = \sum_{j=2}^{m} \binom{m}{j}\binom{n}{j}^{-1} S_{jn} = \sum_{j=2}^{m} R_{jn}$$

is also a U statistic with kernel

$$H(x_1,\dots,x_m) = h(x_1,\dots,x_m) - \sum_{i=1}^{n} \tilde{h}_1(x_i) - \theta$$

and for each $j = 2, \dots, m$,

$$S_{jn} = \sum_{1 \le i_1 < \cdots < i_j \le n} g_j\big(X_{i_1}, \dots, X_{i_j}\big), \qquad n \ge j,$$

is a martingale and $Eg_j(x_1, \dots, x_{j-1}, X_j) = 0$. See Serfling [(1980), page 178] for the definition of the g_j's.

We first establish (2.3) for $j = 2$. Define

$$\lambda_n = n^{2\gamma/(1+\gamma)}(\log n)^{-1/(1+\gamma)}(\log\log n)^{-\delta}.$$

It is enough to show that for any $\varepsilon > 0$,

$$P\big(\lambda_n n^{-2}|S_{2n}| > \varepsilon \text{ i.o.}\big) = 0.$$

Since λ_n is nondecreasing for large n, it suffices to show that

$$\sum_{k=1}^{\infty} P(B_k) = \sum_{k=1}^{\infty} P\Big(\lambda_{2^{k+1}} \max_{2^k \le n \le 2^{k+1}} n^{-2}|S_{2n}| > \varepsilon\Big) < \infty.$$

For any $n \ge 2$,

$$E|S_{2n}|^{1+\gamma} = E\left|\sum_{1 \le i_1 < i_2 \le n} g_2\big(X_{i_1}, X_{i_2}\big)\right|^{1+\gamma} = E\left|\sum_{i_2=2}^{n}\sum_{i_1=1}^{i_2-1} g_2\big(X_{i_1}, X_{i_2}\big)\right|^{1+\gamma}$$
$$= E\left|\sum_{i_2=2}^{n} D_2(i_2)\right|^{1+\gamma},$$

where $D_2(i_2) = \sum_{i_1=1}^{i_2-1} g_2(X_{i_1}, X_{i_2})$, $2 \le i_2 \le n$, is a martingale difference sequence.

By Burkholder's inequality, the above expectation is bounded by

$$CE\left|\sum_{i_2=2}^{n} D_2^2(i_2)\right|^{(1+\gamma)/2} \le C\sum_{i_2=2}^{n} E|D_2(i_2)|^{1+\gamma}.$$

Now observe that for every fixed i_2, $g_2(X_{i_1}, X_{i_2})$, $1 \le i_1 \le i_2 - 1$, is a martingale difference. Thus, using the same argument again,

$$E|S_{2n}|^{1+\gamma} \le C\sum_{i_2=2}^{n}\sum_{i_1=1}^{i_2-1} E\big|g_2\big(X_{i_1}, X_{i_2}\big)\big|^{1+\gamma} \le Cn^2 \tag{2.5}$$

since $E|h|^{1+\gamma} < \infty$ implies $E|g_2|^{1+\gamma} < \infty$.

Using (2.5) and the maximal inequality for the martingale S_{2n},

$$\sum_{k=1}^{\infty} P(B_k) \le \sum_{k=1}^{\infty} P\Big(\lambda_{2^{k+1}} \sup_{2^k \le n \le 2^{k+1}} |S_{2n}| \ge \varepsilon 2^{2k}\Big)$$
$$\le C\sum_{k=1}^{\infty} (\lambda_{2^{k+1}})^{1+\gamma}\big(2^{2k}\varepsilon\big)^{-(1+\gamma)}\big(2^{k+1}\big)^2$$
$$\le C\sum_{k=1}^{\infty} k^{-1}(\log k)^{-\delta(1+\gamma)} < \infty,$$

since $\delta > 1/(1+\gamma)$. This completes the proof for $j = 2$.

A similar argument shows that

$$E|S_{jn}|^{1+\gamma} \le Cn^j.$$

Using the martingale property again, we have the required order for R_{jn}. The second part of the theorem is trivial. □

REMARK 2.1. Note that $(\widehat{U}_n - \theta)$ is a mean of iid random variables with zero mean and finite $(1+\gamma)$th absolute moment. Thus from Theorem 2.1 it follows that:

(i) If $0 < \gamma < 1$ and $\zeta_1 > 0$, then $U_n - \theta = o(n^{-\gamma/(1+\gamma)})$ a.s. This may be termed as the MZSLLN for U statistics.

(ii) If $0 < \gamma < 1$ and $\zeta_1 = \cdots = \zeta_{c-1} = 0$, $\zeta_c > 0$, for some $c > 1$, then

$$U_n - \theta = o\left(n^{-c\gamma/(1+\gamma)}(\log n)^{1/(1+\gamma)}(\log\log n)^{\delta}\right) \quad \text{a.s.}$$

for any $\delta > 1/(1+\gamma)$.

(iii) If $\gamma = 1$ and $\zeta_1 > 0$, then using the LIL for iid random variables, one may obtain

$$U_n - \theta = O\left(n^{-1/2}(\log\log n)^{1/2}\right) \quad \text{a.s.}$$

REMARK 2.2. GH have shown that if $E|X| < \infty$, then for any nonnegative integer k,

$$L_n(k) = \frac{k+1}{n^{k+1}}\sum_{i=1}^{n} i^k X_{(i)} \to E\left(\max(X_1,\dots,X_k)\right) = M_k \quad \text{a.s.},$$

where $X_{(1)} \le \cdots \le X_{(n)}$ are the order statistics of $X_1,\dots,X_n$. See also Helmers (1977) and van Zwet (1980).

Note that for each k, $L_n(k)$ is an L statistic and $L_n(0)$ is simply the sample mean. Given the MZSLLN for the sample mean, it is natural to ask if a similar result may be proved for $L_n(k)$.

(a) Remark 2.1 may be used to show that if $E|X|^{1+\gamma} < \infty$, for some $0 < \gamma < 1$, then

$$L_n(k) - M_k = o\left(n^{-\gamma/(1+\gamma)}\right) \quad \text{a.s.} \tag{2.6}$$

This may be proved as follows. First let $k = 1$. Consider the kernel $h(x_1,x_2) = \max(x_1,x_2)$ and the corresponding U statistic $U_n(h)$. From Remark 2.1 it follows that

$$U_n(h) - M_2 = o\left(n^{-\gamma/(1+\gamma)}\right) \quad \text{a.s.} \tag{2.7}$$

On the other hand,

$$\begin{aligned} U_n(h) &= \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \max(X_{(i)}, X_{(j)}) \\ &= \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} X_{(j)} \\ &= \binom{n}{2}^{-1} \sum_{1 \le j \le n} X_{(j)}(j-1) \\ &= \frac{n}{n-1} L_n(1) - \frac{2}{n-1} L_n(0) \end{aligned}$$

and hence (2.6) follows from the usual MZSLLN for sample mean and (2.7). For general k, it may be shown by algebraic calculations that

$$L_n(k) = U_n(h) + R_n,$$

where $h(x_1, \dots, x_k) = \max(x_1, \dots, x_k)$ and R_n involves linear combinations of $L_n(s)$, $1 \le s \le k-1$, with coefficients of smaller order. Hence by induction, the result follows for all k.

(b) when $\gamma = 0$, the above arguments can be used, along with the strong law of large numbers for U statistics (see Serfling (1980), Theorem A, page 190] to show that $L_n(k) - M_k = o(1)$ a.s. This is precisely Theorem 1.1 of GH.

2.2. *One sided convergence.* Recall the estimate

$$\theta = U_n(h) + a_n U_n(D).$$

Define $\Delta = E|h(X_1, \dots, X_m) - h(X_{m+1}, \dots, X_{2m})|$ and note that $\Delta > 0$ whenever F, the distribution of $h(X_1, \dots, X_m)$, is nondegenerate.

THEOREM 2.2. *Suppose* $E|h(X_1, \dots, X_m)|^{1+\gamma} < \infty$ *for some* $0 < \gamma \le 1$.

(i) *If* $\gamma < 1$, *then* $\theta_n \to_+ \theta$ a.s. *provided* $\liminf a_n n^{\gamma/(1+\gamma)} > 0$.
(ii) *If* $\gamma = 1$, *then* $\theta_n \to_+ \theta$ a.s. *provided* $\liminf a_n n^{1/2} (\log\log n)^{-1/2} = \infty$.

REMARK 2.3. Note that taking $m = 1$ and $h(x) = x$, we essentially get the estimator $\widehat{X}_n$ of GH with $a_n = n^{-(\alpha-2)}$, where α is as in Theorem 1.1. Theorem 2.2 is a stronger assertion than GH for any $0 < \gamma \le 1$.

PROOF. When F is degenerate, there is nothing to prove. When F is not degenerate, write

$$\theta_n - \theta = U_n(h) - \theta + a_n\big(U_n(D) - \Delta\big) + a_n \Delta.$$

From Remark 2.1, when $\gamma < 1$,

$$U_n(h) - \theta = o\big(n^{-\gamma/(1+\gamma)}\big) \quad \text{a.s.}$$

and by the strong law of large numbers for U statistics,

$$U_n(D) - \Delta = o(1) \quad \text{a.s.}$$

and hence (i) follows from the condition on the sequence a_n.

When $\gamma = 1$, using the LIL for U statistics,

$$|U_n(h) - \theta| = O\big(n^{-1/2}(\log\log n)^{1/2}\big) \quad \text{a.s.}$$

Thus again the result follows by using the given condition on a_n. □

REMARK 2.4. When $\zeta_1 = \cdots = \zeta_{c-1} = 0$ the range of values for a_n may be extended by using Remark 2.1.

REMARK 2.5. It may be noted that there are other estimators that will achieve positive convergence. In general, any estimator of the form

$$\theta_{1n} = U_n(h) + a_n E_n,$$

where E_n is such that $E_n - E = o(1)$ a.s. for some $E > 0$, will converge from above to θ. Some possible choices are $U_n(D^*)$ for a kernel D^* such that $E(D^*) > 0$, $\overline{X}_n^2$ when $E(X) \neq 0$, $U_n(|h|)$ and so forth. The choice $U_n(D)$ that we have used has the advantage that it has the same order of moments as $U_n(h)$, provides a U statistics representation for θ_n and also has the appeal that it is a dispersion index. Even though our later asymptotic results will be stated and proved for the estimator $\theta_n = U_n(h) + a_n U_n(D)$, it will be clear from the proofs that many of these results remain valid for θ_n with D replaced by any other suitable kernel D^*.

The optimality of $U_n(h)$ as an estimate of θ is well known. It will be an interesting problem to obtain guidelines for the choice of the perturbation that is added to it to obtain convergence from above. Our choice $U_n(D)$ may play a significant role in this respect.

2.3. *Estimation of quantiles.* Let F be the distribution of $h(X_1, \ldots, X_m)$, where h is a symmetric kernel, and let F_n be the empirical distribution function that puts equal mass at each W_i, the $N = \binom{n}{m}$ values of $h(X_{i_1}, \ldots, X_{i_m})$.

Define

$$G_n = F_n - a_n F_n(1 - F_n). \tag{2.8}$$

Observe that

$$G_n(x) = \frac{1}{N}\sum_{j=1}^{N} I(W_{(j)} \leq x)\left[1 - a_n\left(1 + \frac{1}{N} - \frac{2j}{N}\right)\right], \tag{2.9}$$

where $W_{(1)} \leq W_{(2)} \leq \cdots \leq W_{(N)}$ are the ordered values of the W_i's, $1 \leq i \leq N$. Note that G_n is an empirical distribution function that gives more weight to the higher order statistics and less to the lower ones. It may be noted that the estimator G_n resembles the estimator given in (12) of GH when $m = 1$.

In the following lemma we derive a probabilistic bound on the Kolmogorov–Smirov statistic D_n for the empirical distribution of a U statistic. This bound, which will be used in the proof of Theorem 2.3 on the upper convergence of quantiles, is slightly weaker than the best possible bound known for the empirical distribution of iid observations.

LEMMA 2.1. *Let F_n be the empirical distribution function of a U statistic with kernel $h(x_1, \dots, x_m)$. Define $D_n = \sup_x |F_n(x) - F(x)|$, where F is the distribution function of $h(x_1, \dots, x_m)$. Then*

$$P(D_n \geq t) \leq C_1 t^{-1} \exp\left(-C_2 n t^2\right) \tag{2.10}$$

for some constants C_1 and C_2.

PROOF. For any integer s, let $x_{s,k} = \phi(k/s)$, where $\phi(u) = \inf\{x\colon F(x) \geq u\}$ and define

$$\begin{aligned} D_{s,n} &= \max_{1 \leq k \leq s} \max\left\{|F_n(x_{s,k}) - F(x_{s,k})|, |F_n(x_{s,k}-) - F(x_{s,k}-)|\right\} \\ &= \max_{1 \leq k \leq s} \max\{\alpha_{n,s,k}, \beta_{n,s,k}\} \end{aligned}$$

as in Billingsley [(1991), page 276].

Then $D_n \leq D_{s,n} + 1/s$. Choosing $s = [2/t] + 1$, we have

$$P(D_n \geq t) \leq P\left(D_{s,n} \geq t - 1/s\right) \leq P\left(D_{s,n} \geq t/2\right). \tag{2.11}$$

For each s, the terms $\alpha_{n,s,k}$ and $\beta_{n,s,k}$ are U statistics with a kernel that is bounded by 1.

Using Theorem A of Serfling [(1980), page 201], for each t,

$$P(\alpha_{n,s,k} \geq t) \leq 2 \exp\left(-2\left[n/m\right]t^2\right). \tag{2.12}$$

Using Bonferroni's inequality and relations (2.11) and (2.12),

$$P(D_n \geq t) \leq 4\left(\left[2/t\right] + 1\right) \exp\left(-2\left[n/m\right]t^2\right).$$

This proves the lemma. □

For any p, let

$$\xi_p = \inf\{x\colon F(x) \geq p\}$$

be the pth quantile of F and let

$$\hat{\xi}_{p_n} = \inf\{x\colon G_n(x) \geq p\}$$

be the pth quantile of G_n. Also let

$$\psi_n(x) = x - a_n x(1-x), \qquad 0 \leq x \leq 1, \quad \text{and} \quad p_n = \psi_n^{-1}(p).$$

We prove the following theorem.

THEOREM 2.3. *Let* a_n *be such that* $\liminf a_n n^{1/2}(\log\log n)^{-1/2} = \infty$. *Then*:

(i) $\lim_n \sup_x |G_n(x) - F(x)| = 0$ *a.s.*
(ii) $G_n(x) \to_- F(x)$ *a.s. for every* x.
(iii) $\widehat{\xi}_{p_n} \to \xi_p$ *a.s. for all* $0 \le p \le 1$. *Further, if for all sufficiently large* n,

$$a_n p(1-p) > \left(3m/4\right)^{1/2} n^{-1/2}(\log n)^{1/2},$$

then $\widehat{\xi}_{p_n} \to_+ \xi_p$ *a.s. for all* $0 < p < 1$.
(iv) *Mean of* $G_n \to$ *mean of* F *a.s.*

PROOF. The first part (i) follows from the observation that $\|G_n - F_n\| \le a_n$ and the Glivenko–Cantelli theorem for U statistics, which says that $\|F_n - F\| \to 0$ a.s.

(ii) Note that

$$G_n - F = (1-a_n)(F_n - F) + a_n\left(F_n^2 - F^2\right) - a_n F(1-F).$$

For each fixed x the first term is $o(a_n)$ a.s. from the LIL for U statistics. The second term is obviously $o(a_n)$. Hence (ii) follows.

(iii) By using the Borel–Cantelli lemma, it easily follows from Lemma 2.1 that almost surely,

$$|D_n| \le C_0 n^{-1/2}(\log n)^{1/2}, \tag{2.13}$$

where C_0 may be chosen to be any number greater than $(3m/4)^{1/2}$; see the proof of Lemma 2.1. Note that $G_n = \psi_n(F_n)$ and for every n, $\psi_n(x)$ is strictly increasing in x.

Using Lemma (iii) of Serfling [(1980), page 3], first observe that

$$\begin{aligned} G_n^{-1}(p) = \widehat{\xi}_{p_n} \ge \xi_p \quad &\text{iff } F_n^{-1}\left(\psi_n^{-1}(p)\right) \ge F^{-1}(p) \\ &\text{iff } F\left(F_n^{-1}\left(\psi_n^{-1}(p)\right)\right) \ge p. \end{aligned} \tag{2.14}$$

Using the bound (2.13) on D_n and Lemma (ii) of Serfling [(1980), page 3], we get

$$\begin{aligned} F\left(F_n^{-1}\left(\psi_n^{-1}(p)\right)\right) &\ge F_n\left(F_n^{-1}\left(\psi_n^{-1}(p)\right)\right) - C_0 n^{-1/2}(\log n)^{1/2} \\ &\ge \psi_n^{-1}(p) - C_0 n^{-1/2}(\log n)^{1/2}. \end{aligned} \tag{2.15}$$

Solving the quadratic equation $\psi_n(x) = p$, one gets

$$\psi_n^{-1}(p) = p + a_n p(1-p) + O\left(a_n^2\right). \tag{2.16}$$

Using the condition on a_n, the result follows from (2.14)–(2.16).

(iv) Note that

$$E_{G_n}(Y) = U_n(h) - \frac{a_n}{2N^2} \sum_{j=1}^{N} \left(\frac{N+1}{2} - j \right) W_{(j)}.$$

Thus

$$\left| E_{G_n}(Y) - U_n(h) \right| \le \frac{a_n}{N^2} \sum_{j=1}^{N} \left| \frac{N+1}{2} - j \right| \left| W_{(j)} \right|$$
$$\le C \frac{a_n}{2N} \sum_{j=1}^{N} |W_{(j)}| = C a_n U_n(|h|).$$

Observe that $a_n \to 0$ and thus using the SLLN for U statistics, (iv) follows. □

Acknowledgment. We thank Arusharka Sen for some informative discussions. We also greatly appreciate the prompt and constructive comments of the referee.

REFERENCES

BILLINGSLEY, P. (1991). *Probability and Measure.* Wiley, New York.

CHOW, Y. S. and TEICHER, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales.* Springer, New York.

GILAT, D. and HILL, T. P. (1992). One-sided refinement of the strong law of large numbers and the Glivenko–Cantelli theorem. *Ann. Probab.* **20** 1213–1221.

HELMERS, R. (1977). A strong law of large numbers for linear combination of order statistics. Technical Report SW 50/77, Math. Centrum, Amsterdam.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

VAN ZWET, W. (1980). A strong law for linear functions of order statistics. *Ann. Probab.* **8** 986–990.

STATISTICS-MATHEMATICS UNIT
INDIAN STATISTICAL INSTITUTE
203 B.T. ROAD
CALCUTTA 700035
INDIA