The Annals of Probability
1977, Vol. 5, No. 2, 287-290

ALMOST SURE CONVERGENCE OF GENERALIZED U-STATISTICS[1]

BY PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

Almost sure convergence of generalized U-statistics and von Mises' differentiable statistical functions is studied with the help of the general $L \log L$ martingale convergence theorem.

1. Introduction. Let $\{X_{ij}, j \geqq 1\}$ be a sequence of independent and identically distributed random vectors (defined on a probability space $(\Omega, \mathscr{A}, P)$) with each X_{ij} having a distribution function (df) $F_i(x)$, $x \in R^p$, the $p\ (\geqq 1)$-dimensional Euclidean space, for $i = 1, \cdots, c\ (\geqq 2)$; all these c sequences are assumed to be mutually independent. Consider an *estimable parameter* (a functional of $\mathbf{F} = (F_1, \cdots, F_c)$ defined on an appropriate space of c-tuples of df's):

$$(1.1) \qquad \theta(\mathbf{F}) = \int \cdots \int \phi(x_{ij}, j = 1, \cdots, m_i, i = 1, \cdots, c) \prod_{i=1}^{c} \prod_{j=1}^{m_i} dF_i(x_{ij}),$$

where $m_1, \cdots, m_c$ are nonnegative integers, $\mathbf{m} = (m_1, \cdots, m_c)\ (\neq \mathbf{0})$ is the *degree* (vector) of $\theta(\mathbf{F})$ and the Borel measurable *kernel* ϕ can always be so chosen that it is symmetric in the elements within each of its c sets of arguments. The corresponding *generalized U-statistic* based on a sample of size $\mathbf{n} = (n_1, \cdots, n_c)$ is

$$(1.2) \qquad U(\mathbf{n}) = \prod_{i=1}^{c} \binom{n_i}{m_i}^{-1} \sum\nolimits_{(\mathbf{n})}^{*} \phi(X_{ij_\alpha}, \alpha = 1, \cdots, m_i, i = 1, \cdots, c),$$

where $\mathbf{n} \geqq \mathbf{m}$ (i.e., $n_i \geqq m_i$, $i = 1, \cdots, c$) and the summation $\sum_{(\mathbf{n})}^{*}$ extends over all possible $1 \leqq j_1 < \cdots < j_{m_i} \leqq n_i$, $i = 1, \cdots, c$. $U(\mathbf{n})$ is an unbiased and symmetric estimator of $\theta(F)$; we may refer to Puri and Sen (1971, Chapter 3) for various properties of $U(\mathbf{n})$.

In the context of weak convergence of generalized U-statistics, Sen (1974a) has established some Kolmogorov-type inequalities which for square integrable (with respect to $\mathbf{F}$) kernels insure that

$$(1.3) \qquad U(\mathbf{n}) \to \theta(\mathbf{F}) \qquad \text{almost surely (a.s.) as} \quad \mathbf{n} \to \infty .$$

Since $U(\mathbf{n})$ is symmetric in each of $\{X_{i1}, \cdots, X_{in_i}\}$, $i = 1, \cdots, c$, one might naturally inquire whether for (1.3), the square integrability condition may be relaxed. In fact, for $c = 1$, through the reverse martingale property of the classical U-statistics, Berk (1966) has elegantly shown that (1.3) holds whenever ϕ is integrable. For $c \geqq 2$, i.e., for a multiple array of random variables, in

Received September 22, 1975; revised May 3, 1976.

[1] Work supported partially by the Air Force Office of Scientific Research, A.F. S.C., U.S.A.F., Grant No. AFOSR 74-2736, and partially by a grant of the Deutsche Forschungsgemeinschaft during the visit of the author at the University of Freiburg, W. Germany.

AMS 1970 *subject classification.* Primary 60B10.

Key words and phrases. Almost sure convergence, generalized U-statistics, reverse martingales and submartingales, von Mises' differentiable statistical functions.

view of the lack of the complete ordering of the elements of the index set $\mathbb{N} = \{\mathbf{n}: \mathbf{n} \geqq \mathbf{m}\}$, it is difficult to provide a direct extension of Berk's approach. Nevertheless, by using some specific directional reverse martingale properties of $\{U(\mathbf{n})\}$, we are able to show that (1.3) holds under a condition slightly more stringent than Berk's. The main theorem of this note is the following.

THEOREM 1. *If* $E\{|\phi|(\log^+ |\phi|)^{c-1}\} < \infty$, *then* (1.3) *holds.*

The proof is outlined in Section 2. It may be noted that the condition of the theorem is sufficient but not necessary. In Section 3, we present some special cases where (1.3) holds under the ideal condition that ϕ is integrable. The case of von Mises' differentiable statistical functions is also treated briefly in Section 3. In the sequel, we let $p = 1$ and $c = 2$; the case of $c > 2$ follows by induction, while the case of $p > 1$ poses no additional problem.

2. The proof of Theorem 1. Let $\mathscr{F}_n^{(i)}$ be the σ-field generated by the ordered random variables $X_{i,1} \leqq \cdots \leqq X_{i,n}$ and $X_{in+1}, X_{in+2}, \cdots$, for $n \geqq 1$ and $i = 1, 2$. Note that for each i, $\mathscr{F}_n^{(i)}$ is nonincreasing in n and $\mathscr{F}_{n_1}^{(1)}$ and $\mathscr{F}_{n_2}^{(2)}$ are mutually independent for every $\mathbf{n} \geqq \mathbf{1}$. Let then $\mathscr{F}^*(\mathbf{n})$ be the σ-field generated by $\mathscr{F}_{n_1}^{(1)}$ and $\mathscr{F}_{n_2}^{(2)}$, for $\mathbf{n} \geqq \mathbf{1}$, so that $\mathscr{F}^*(\mathbf{n})$ is nonincreasing in each of its arguments. Note that by (1.2), for every $k \geqq 0$, $n_1' \geqq n_1 \geqq m_1$ and $n_2 \geqq m_2$,

$$
\begin{aligned}
&E\{U(n_1, n_2 + k) \mid \mathscr{F}^*(n_1', n_2)\} \\
&\quad = \binom{n_1}{m_1}^{-1}\binom{n_2+k}{m_2}^{-1} \textstyle\sum^*_{(n_1, n_2+k)} E\{\phi(X_{1j_1}, \cdots, X_{1j_{m1}}, \\
&\qquad\qquad X_{2j_1'}, \cdots, X_{2j'_{m_2}}) \mid \mathscr{F}^*(n_1', n_2)\} \\
&\quad = \binom{n_2+k}{m_2}^{-1} \textstyle\sum_{1 \leqq v_1 < \cdots < v_{m_2} \leqq n_2+k} E\{\phi(X_{11}, \cdots, X_{1m_1}, \\
&\qquad\qquad X_{2v_1}, \cdots, X_{2v_{m_2}}) \mid \mathscr{F}^*(n_1', n_2)\} \\
&\quad = U(n_1', n_2 + k) \quad \text{a.e.}
\end{aligned} \tag{2.1}
$$

For every $\mathbf{n} \geqq \mathbf{m}$, we define

$$\mathbf{Y}(\mathbf{n}) = (U(n_1, n_2 + k) - \theta(\mathbf{F}), k = 0, 1, \cdots) \tag{2.2}$$

and consider the sequence

$$\{\mathbf{Y}(\mathbf{n}), \mathscr{F}^*(\mathbf{n}); n_1 \geqq m_1\}, \qquad \text{defined for each} \quad n_2 \geqq m_2. \tag{2.3}$$

It follows from (2.1), (2.2) and (2.3) that

$$E\{\mathbf{Y}(\mathbf{n}) \mid \mathscr{F}^*(n_1', n_2)\} = \mathbf{Y}(n_1', n_2) \qquad \forall\, n_1' \geqq n_1 \geqq m_1 \quad \text{and} \quad n_2 \geqq m_2. \tag{2.4}$$

Thus, if we let

$$Z(\mathbf{n}) = \|\mathbf{Y}(\mathbf{n})\| = \sup_{k \geqq n_2} |U(n_1, k) - \theta(\mathbf{F})|, \quad \mathbf{n} \geqq \mathbf{m}, \tag{2.5}$$

then, by virtue of (2.4) and the convexity of the sup-norm, for every $n_2 \geqq m_2$,

$$\{Z(n, n_2), \mathscr{F}^*(n, n_2); n \geqq m_1\} \quad \text{is a nonnegative reverse submartingale.} \tag{2.6}$$

Consequently, by the two celebrated inequalities of Doob (1953, page 317), we have for every $\lambda > 0$,

$$\begin{aligned}
&\lambda P\{\sup_{\mathbf{n}} Z(\mathbf{n}) \geqq \lambda\} \\
&\quad \leqq \sup_{n_2} E\{\sup_{n_1' \geqq n_1} |U(n_1', n_2) - \theta(\mathbf{F})|\} \\
&\quad \leqq \sup_{n_2} [\{e/(e-1)\}\{1 + E(|U(\mathbf{n}) - \theta(\mathbf{F})| \log^+ |U(\mathbf{n}) - \theta(\mathbf{F})|)\}] \\
&\quad \leqq \{e/(e-1)\}\{1 + E(|\phi - \theta(\mathbf{F})| \log^+ |\phi - \theta(\mathbf{F})|)\}\,,
\end{aligned} \tag{2.7}$$

where the last inequality follows from the fact that $U(\mathbf{n}) - \theta(\mathbf{F}) = E\{\phi - \theta(\mathbf{F}) \mid \mathscr{F}^*(\mathbf{n})\}$, $(\mathbf{n} \geqq \mathbf{m})$ and the Jensen inequality. Having proved this, we can now virtually repeat the modified Cairoli arguments of Symthe (1973, pages 167–168) [who considered the a.s. convergence of $\bar{X}(\mathbf{n}) = (\sum_{\mathbf{i} \leqq \mathbf{n}} X_{\mathbf{i}})/||\mathbf{n}||$, where the $X_{\mathbf{i}}$ are independent and identically distributed random variables and $\mathbf{i} = (i_1, \cdots, i_r) \geqq \mathbf{1}$ for some positive integer r] and complete the proof of (1.3) on the same line. For brevity, the details are therefore omitted. □

3. Some general remarks. By virtue of (2.1), it follows that

$$E\{U(n\mathbf{1}) \mid \mathscr{F}^*(n'\mathbf{1})\} = U(n'\mathbf{1}) \qquad \text{for every} \quad n' \geqq n \geqq \max(m_1, m_2)\,, \tag{3.1}$$

and, as such, by noting that $EU(n\mathbf{1}) = \theta(\mathbf{F})$, we obtain by the reverse martingale convergence theorem that whenever $E\phi$ exists,

$$U(n\mathbf{1}) \to \theta(\mathbf{F}) \quad \text{a.s.,} \qquad \text{as} \quad n \to \infty\,. \tag{3.2}$$

The same argument holds if one considers the sequence $\{U([n\lambda], n - [n\lambda])$, $n \geqq n_0$, where $[s]$ denotes the largest integer $\leqq s$, $\lambda \in (0, 1)$ and $n_0 = \min\{n: [n\lambda] \geqq m_1\, n - [n\lambda] \geqq m_2\}$. Thus, for the diagonal case, for (1.3), the integrability of ϕ suffices.

For a multidimensional array of random variables, $\{X_{\mathbf{i}}, \mathbf{i} \geqq \mathbf{1}\}$, Smythe (1973) has studied the necessity of $E|X|(\log^+ |X|)^{r-1} < \infty$ for the strong law of large numbers to hold. Though, as indicated in Section 2, the proofs involve similar techniques, in our case, $E|\phi|(\log^+ |\phi|)^{c-1} < \infty$ is only a sufficient condition for (1.3) to hold. Towards this, we consider the following theorem where (1.3) holds under the ideal condition that ϕ is integrable.

THEOREM 2. *Suppose that*

$$\theta(\mathbf{F}) = \sum_{\mathbf{0} \leqq \mathbf{s} \leqq \mathbf{m}} \alpha(\mathbf{s}) \prod_{j=1}^{c} \theta_{\mathbf{s}}^{(j)}(F_j)\,, \tag{3.3}$$

where the $\alpha(s)$ are real (known) constants and the $\theta_{\mathbf{s}}^{(j)}(F_j)$ are functionals of only the jth df F_j, $j = 1, \cdots, c$. *Then,* (1.3) *holds whenever ϕ is integrable.*

PROOF. Because of the linearity in (3.3), it suffices to prove (1.3) for the particular case of $\theta_t(\mathbf{F}) = \prod_{j=1}^{c} \theta_{\mathbf{s}}^{(j)}(F_j)$. A simple summation shows that in this case

$$U(\mathbf{n}) = U_1(n_1) \cdots U_c(n_c)\,, \quad \forall\, \mathbf{n} \geqq \mathbf{m}\,, \tag{3.4}$$

where the $U_j(n_j)$ are the classical one-sample U-statistics, and hence, $U_j(n_j) \to \theta_{\mathbf{s}}^{(j)}(F_j)$ a.s. as $n_j \to \infty$, for every $j (= 1, \cdots, c)$, when ever ϕ is integrable. Hence, (1.3) holds whenever ϕ is integrable. □

Let us define the empirical df's by

(3.5) $$F_{n_i}^{(i)}(x) = n_i^{-1} \sum_{j=1}^{n_i} c(x - X_{ij}) , \quad n_i \geqq 1 , \qquad i = 1, \cdots, c ,$$

where for a p-vector u, $c(u) = 1$ iff all the p components of u are nonnegative, and is 0, otherwise. Let then $\mathbf{F}(\mathbf{n}) = (F_{n_1}^{(1)}, \cdots, F_{n_c}^{(c)})$. The von Mises' differentiable statistical function $\theta(\mathbf{F}(\mathbf{n}))$ can now be defined as

(3.6) $$\int \cdots \int \phi(x_{ij}, j = 1, \cdots, m_i, i = 1, \cdots, c) \prod_{i=1}^{c} \prod_{j=1}^{m_i} dF_{n_i}^{(i)}(x_{ij}) = n_1^{-m_1} \cdots n_c^{-m_c} \tilde{\sum}_{(\mathbf{n})} \phi(X_{ij_\alpha}, \alpha = 1, \cdots, m_i, i = 1, \cdots, c) ,$$

where the summation $\tilde{\sum}_{(\mathbf{n})}$ extends over all possible $j_\alpha = 1, \cdots, n_i, i = 1, \cdots, c$. $\theta(\mathbf{F}(\mathbf{n}))$ is an alternative estimator of $\theta(\mathbf{F})$, though it is not necessarily unbiased. From (1.2) and (3.6), we obtain by some standard steps that for $\mathbf{n} \geqq \mathbf{m}$,

(3.7) $$\theta(\mathbf{F}(\mathbf{n})) = \sum_{\mathbf{0} \leqq \mathbf{s} \leqq \mathbf{m}} h(\mathbf{n}, \mathbf{s}) U(\mathbf{n}; \mathbf{s}) ,$$

where $h(\mathbf{n}, \mathbf{0}) = \prod_{i=1}^{c} \{n_i^{-m_i} n_i^{[m_i]}\} = 1 + O(n_0^{-1})$, $n_0 = \min(n_1, \cdots, n_c)$ and for $\mathbf{s} \neq \mathbf{0}$, $h(\mathbf{n}, \mathbf{s}) = O(n^{-\mathbf{s}'\mathbf{1}})$, $U(\mathbf{n}, \mathbf{0}) = U(\mathbf{n})$, defined by (1.2) and for each $\mathbf{s}$, $U(\mathbf{n}, \mathbf{s})$ is a generalized U-statistic of degree $\leqq \mathbf{m}$ (with at least one strict inequality sign). Thus, if we assume that

(3.8) $$E\{|\phi|(\log^+ |\phi|)^{c-1}\} < \infty ,$$

where the m_i arguments in the ith set of ϕ are not necessarily all distinct, $i = 1, \cdots, c$, then by (3.7), Theorem 1 and (3.8), we conclude that for every $\varepsilon > 0$,

(3.9) $$n_0^{1-\varepsilon} |U(\mathbf{n}) - \theta(F(\mathbf{n}))| \to 0 \quad \text{a.s.}, \qquad \text{as} \quad n_0 \to \infty ,$$

so that $\theta(\mathbf{F}(\mathbf{n}))$ also converges a.s. to $\theta(\mathbf{F})$, as $\mathbf{n} \to \infty$.

Acknowledgment. The author is grateful to Professor Ron Pyke and the referee for their valuable comments on the manuscript.

REFERENCES

[1] BERK, R. H. (1966). Limiting behaviour of posterior distributions when the model is incorrect. *Ann. Math. Statist.* **37** 51-58.

[2] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[3] PURI, M. L. and SEN, P. K. (1971). *Nonparametric Methods in Multivariate Analysis.* Wiley, New York.

[4] SEN, P. K. (1972). A Hájek-Rényi type inequality for generalized U-statistics. *Calcutta Statist. Assoc. Bull.* **21** 171-179.

[5] SEN, P. K. (1974a). Weak convergence of generalized U-statistics. *Ann. Probability* **2** 90-102.

[6] SEN, P. K. (1974b). Almost sure behaviour of U-statistics and von Mises' differentiable statistical functions. *Ann. Statist.* **2** 387-396.

[7] SMYTHE, R. T. (1973). Strong laws of large numbers for r-dimensional arrays of random variables. *Ann. Probability* **1** 164-170.

DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27514