The Annals of Probability
1974, Vol. 2, No. 1, 90-102

WEAK CONVERGENCE OF GENERALIZED U-STATISTICS[1]

BY PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

Wichura (1969) studied an invariance principle for partial sums of a multi-dimensional array of independent random variables. It is shown that a similar invariance principle holds for a broad class of generalized U-statistics for which the different terms in the partial sums are not independent. Weak convergence of generalized U-statistics for random sample sizes is also studied. The case of (generalized) von Mises' functional is treated briefly.

1. Introduction. Let $\{X_{ji}, i \geqq 1\}$, $j = 1, \cdots, c$, be $c\,(\geqq 2)$ independent sequence of independent random vectors (irv), where X_{ji} has a distribution function (df) $F_j(x)$, $x \in R^p$, the $p\,(\geqq 1)$-dimensional Euclidean space, for $j = 1, \cdots, c$. Let $g(X_{ji}, i = 1, \cdots, m_j, j = 1, \cdots, c)$ be a Borel-measurable *kernel* of *degree* $\mathbf{m} = (m_1, \cdots, m_c)$, where we assume (without any loss of generality) that g is symmetric in the $m_j\,(\geqq 1)$ arguments (vectors) of the jth set, for $j = 1, \cdots, c$. Let $m_0 = m_1 + \cdots + m_c$, $\mathbf{F} = (F_1, \cdots, F_c)$, and consider a functional of $\mathbf{F}$

$$\theta(\mathbf{F}) = \int_{R^{pm_0}} \cdots \int g(x_{11}, \cdots, x_{cm_c}) \prod_{j=1}^{c} \prod_{i=1}^{m_j} dF_j(x_{ji}) \tag{1.1}$$

defined on $\mathscr{F} = \{\mathbf{F}: |\theta(F)| < \infty\}$. $\theta(\mathbf{F})$ is called an *estimable parameter* or a *regular functional* of F over $\mathscr{F}$.

For a set of samples of sizes $\mathbf{n} = (n_1, \cdots, n_c)$ with $n_j \geqq m_j$, $1 \leqq j \leqq c$, the *generalized U-statistic* for $\theta(\mathbf{F})$ is defined by

$$U(\mathbf{n}) = \prod_{j=1}^{c} \binom{n_j}{m_j}^{-1} \sum_{(\mathbf{n})}^{*} g(X_{j\alpha}, \alpha = i_{j1}, \cdots, i_{jm_j}, 1 \leqq j \leqq c)\,, \tag{1.2}$$

where the summation $\sum_{(\mathbf{n})}^{*}$ extends over all $1 \leqq i_{j1} < \cdots < i_{jm_j} \leqq n_j$, $1 \leqq j \leqq c$. For various properties of $U(\mathbf{n})$, including its asymptotic normality, we may refer to Fraser (1957) and Puri and Sen (1971), among others. For the asymptotic normality, it is assumed that $\theta(\mathbf{F})$ is *stationary of order zero* [i.e., (2.3) holds] and essentially

$$\lim_{n\to\infty} n_j/n = \lambda_j : 0 < \lambda_j < 1\,, \qquad j = 1, \cdots, c\,, \tag{1.3}$$

where $n = n_1 + \cdots + n_c$. Weak convergence of a stochastic process derived

---

Received March 1, 1972; revised May 22, 1973.

[1] Work sponsored by the Aerospace Research Laboratories, Air Force Systems Command, U. S. Air Force Contract No. F33615-71-C-1927. Reproduction in whole or in part permitted for any purpose of the U. S. Government.

AMS 1970 *subject classifications.* 60B10, 60G99.

Key words and phrases. $D\{[0, 1]^c\}$ space, Generalized U-statistics, invariance principle, Gaussian processes, relative compactness, random indices, Von Mises' differentiable statistical functions, and weak convergence.

from the tail sequence of one-sample U-statistics to a Wiener process has been studied by Loynes (1970), while Miller and Sen (1972) consider a Donsker-type invariance principle for one-sample U-statistics. They show that a process derived from $\{U([k\lambda_1], \cdots, [k\lambda_c]), k \leqq n\}$ converges weakly to a one-dimensional Wiener process. The more general and natural case of a c-dimensional time-parameter where we use the entire set $\{U(\mathbf{k}): \mathbf{m} \leqq \mathbf{k} \leqq \mathbf{n})$ (or the entire tail set $\{U(\mathbf{k}): \mathbf{k} \geqq \mathbf{n}\}$) is considered here, and it is shown that weak convergence to appropriate multi-dimensional Gaussian processes hold under no extra regularity conditions; here $\mathbf{a} \leqq \mathbf{b}$ means that $a_i \leqq b_i$ for all $1 \leqq i \leqq c$. It may be noted that for $c \geqq 2$, the ordering of $\mathbf{n}$ is not defined, and as a result, the treatment of Loynes (1970) or of Miller and Sen (1972), resting on the reverse martingale property of U-statistics, does not work out. Also, as is usually the case with generalized U-statistics, $U(\mathbf{n})$ in (1.2) involves a set of summands which are not all stochastically independent. Thus, Theorem 2 (or its Corollary 1) of Wichura (1969) does not lead us to the desired result. The task is accomplished here by first extending Theorem 1 of Wichura (1969) to more general summands, and then using a decomposition of $U(\mathbf{n})$ which fits into this extension.

The preliminary notions and the basic theorems are considered in Section 2. The proofs of the theorems are presented in Section 3. The case of generalized von Mises' (1947) functionals is treated in Section 4. In the last section, the case of random sample sizes is also considered. The results are useful in the developing area of sequential inference based on generalized U-statistics.

2. Statement of the main results. For every $d_j: 0 \leqq d_j \leqq m_j, 1 \leqq j \leqq c$, let

$$(2.1)\qquad g_{d_1\cdots d_c}(x_{ji}, i = 1, \cdots, d_j, 1 \leqq j \leqq c) = E\{g(x_{j1}, \cdots, x_{jd_j}, X_{jd_j+1}, \cdots, X_{jm_j}, 1 \leqq j \leqq c)\},$$

so that $g_{00\cdots 0} = \theta(F)$ and $g_{m_1\cdots m_c}(\ \ ) = g(\ \ )$. Let then

$$(2.2)\qquad \zeta_{d_1\cdots d_c}(\mathbf{F}) = Eg^2_{d_1\cdots d_c}(X_{ji}, 1 \leqq i \leqq d_j, 1 \leqq j \leqq c) - \theta^2(\mathbf{F}),$$

so that $\zeta_{0\cdots 0}(\mathbf{F}) = 0$. We assume that

(i) $\theta(\mathbf{F})$ is stationary of order 0, i.e.,

$$(2.3)\qquad \sigma_j^2 = m_j^2 \zeta_{\delta_{j1}\cdots\delta_{jc}}(\mathbf{F}) > 0 \quad \text{for every} \quad 1 \leqq j \leqq c,$$

(where $\delta_{ab} = 1$ or 0 according as $a = b$ or not), and

(ii) g is square integrable, i.e.,

$$(2.4)\qquad \zeta_{m_1\cdots m_c}(\mathbf{F}) < \infty \quad \text{i.e.,} \qquad g \in L^2.$$

Later on, we shall see that (2.3) may be replaced by $\max_{1\leqq j\leqq c} \sigma_j^2 > 0$, and the necessary modifications are trivial. We know that for $n_j \geqq m_j$, $1 \leqq j \leqq c$, under (1.3), (2.3) and (2.4).

$$(2.5)\qquad r^2(\mathbf{n}) = V(U(\mathbf{n})) = \prod_{j=1}^c \binom{n_j}{m_j}^{-1} \sum_{d_1=0}^{m_1} \cdots \sum_{d_c=0}^{m_c} \prod_{j=1}^c \left\{\binom{m_j}{d_j}\binom{n_j-m_j}{m_j-d_j}\right\} \zeta_{d_1\cdots d_c}(\mathbf{F}) = \sum_{j=1}^c n_j^{-1}\sigma_j^2 + O(n^{-2}), \qquad n = n_1 + \cdots + n_c.$$

Let now $E_c = [0, 1]^c$ be the c-dimensional unit cube in R^c, $\mathbf{t} = (t_1, \cdots, t_c) \in E_c$, and let $[\mathbf{nt}] = ([n_1 t_1], \cdots, [n_c t_c])$ where $[s]$ denotes the largest integer $\leqq s$. Then, for every $\mathbf{n}$ ($\geqq \mathbf{m}$), we define a process $W(\mathbf{n}) = \{W(\mathbf{t}; \mathbf{n}) : \mathbf{t} \in E_c]$ by letting

$$(2.6) \qquad W(\mathbf{t}, \mathbf{n}) = \psi([\mathbf{nt}]; n)[U([\mathbf{nt}]) - \theta(\mathbf{F})], \quad \forall [\mathbf{nt}] \geqq \mathbf{m}, \quad \text{and} \quad 0, \text{ otherwise,}$$

where for $\mathbf{k} = (k_1, \cdots, k_c)$, $k_j > 0$, $j = 1, \cdots, c$,

$$(2.7) \qquad \psi(\mathbf{k}, n) = n^{-\frac{1}{2}}(\textstyle\sum_{j=1}^{c} \sigma_j \lambda_j^{\frac{1}{2}})(\sum_{j=1}^{c} \sigma_j \lambda_j^{\frac{1}{2}} k_j^{-1})^{-1},$$

so that $\psi(\mathbf{k}, n)$ is $n^{-\frac{1}{2}}$ times a harmonic mean of $k_1, \cdots, k_c$. Consider now c independent copies of a standard Brownian motion on $[0, 1]$ and denote these by $W_j = \{W_j(t) : 0 \leqq t \leqq 1\}$, $j = 1, \cdots, c$. Finally, the space $D_c = D\{[0, 1]^c\}$ of all real functions on E_c with no discontinuities of the second kind and associated (extended) Skorokhod J_1-topology are defined as in Neuhaus (1971). Then, we have the following.

THEOREM 2.1. *Under* (1.3), (2.3) *and* (2.4), $W(\mathbf{n})$ *converges in law in the extended Skorokhod J_1-topology on D_c to a Gaussian function* $W = \{W(\mathbf{t}) : \mathbf{t} \in E_c\}$, *where*

$$(2.8) \qquad \begin{aligned} W(\mathbf{t}) &= (\textstyle\sum_{j=1}^{c} \sigma_j \lambda_j^{\frac{1}{2}})(\sum_{j=1}^{c} \sigma_j \lambda_j^{-\frac{1}{2}} t_j^{-1})^{-1}[\sum_{j=1}^{c} \sigma_j \lambda_j^{-\frac{1}{2}} t_j^{-1} W_j(t_j)], \ \mathbf{t} > \mathbf{0}, \\ &= 0, \quad \textit{with probability} \quad 1, \textit{ if } \quad t_j = 0 \quad \textit{for some} \quad j : 1 \leqq j \leqq c. \end{aligned}$$

We define a process $W^*(\mathbf{n}) = \{W^*(\mathbf{t}; \mathbf{n}); \mathbf{t} \in E_c\}$ as follows. Considering the tail set $\{U(\mathbf{k}) : \mathbf{k} \geqq \mathbf{n}\}$, let

$$(2.9) \qquad W^*(\mathbf{t}; \mathbf{n}) = r^{-1}(\mathbf{n})[U([\mathbf{n}/\mathbf{t}]) - \theta(\mathbf{F})], \qquad \mathbf{t} \in E_c,$$

where $[\mathbf{n}/\mathbf{t}] = ([n_1/t_1], \cdots, [n_c/t_c])$. Let then

$$(2.10) \qquad \mathbf{w} = (w_1, \cdots, w_c)'; \ w_j = \sigma_j \lambda_j^{-\frac{1}{2}}(\textstyle\sum_{j=1}^{c} \sigma_j^2/\lambda_j)^{-\frac{1}{2}}, \qquad 1 \leqq j \leqq c;$$

$$(2.11) \qquad \mathbf{w}(\mathbf{t}) = (W_1(t_1), \cdots, W_c(t_c))', \qquad \mathbf{t} \in E_c,$$

where the $W_j(t)$ are defined earlier, and let

$$(2.12) \qquad W^* = \{W^*(\mathbf{t}) : \mathbf{t} \in E_c\}; \ W^*(\mathbf{t}) = \mathbf{w}'\mathbf{W}(\mathbf{t}), \qquad \mathbf{t} \in E_c.$$

THEOREM 2.2. *Under* (1.3), (2.3) *and* (2.4), *as* $n \to \infty$,

$$(2.13) \qquad W^*(\mathbf{n}) \to_{\mathscr{D}} W^*, \quad \textit{in the Skorokhod } J_1\textit{-topology on} \quad D_c.$$

Theorems 2.1 and 2.2 provide multi-sample extensions of the weak convergence results of Miller and Sen (1972) and Loynes (1970), respectively. Related results on von Mises' (1947) functionals are considered in Section 4.

3. Proofs of the theorems. For simplicity of the proofs, we explicitly consider the case of $c = 2$; an essentially same but more laborious proof holds for general c ($\geqq 2$). First, we consider certain basic lemmas needed in the subsequent steps of the proof.

Let $\mathscr{B}_n^{(j)}$ be the σ-field generated by $\{X_{j1}, \cdots, X_{jn}\}$ for $n \geqq 1$, $1 \leqq j \leqq c$; $\mathscr{B}_{n_1n_2}^{(12)}$ denote the product σ-field $\mathscr{B}_{n_1}^{(1)} \times \mathscr{B}_{n_2}^{(2)}$ for $n_1 \geqq 1$, $n_2 \geqq 1$. Further, let $S_{ii'} = S(X_{11}, \cdots, X_{1i}, X_{21}, \cdots, X_{2i'})$, for $i, i' \geqq 1$, be a sequence of random variables such that

$$E(S_{ij} \mid \mathscr{B}_{i'j}^{(12)}) = S_{i'j} \quad \text{a.e. (almost everywhere)} \tag{3.1}$$

for every $i \geqq i' \geqq 1$ and $j \geqq 1$, and for every $j \geqq j' \geqq 1$, $n_1 \geqq 1$,

$$E(\mathbf{S}_{n_1}^{(j)} \mid \mathscr{B}_{n_1j'}^{(12)}) = \mathbf{S}_{n_1}^{(j')} \quad \text{a.e.,} \tag{3.2}$$

where $\mathbf{S}_k^{(j)} = (S_{1j}, \cdots, S_{kj})'$ for $k \geqq 1$, $j \geqq 1$. Finally, assume that for every $i \geqq 1$, $j \geqq 1$,

$$E(S_{ij}) = 0\,, \qquad \sigma_{ij}^2 = E(S_{ij}^2) < \infty\,. \tag{3.3}$$

LEMMA 3.1. *Under* (3.1), (3.2) *and* (3.3), *for every* $n_1 \geqq 1$, $n_2 \geqq 1$,

$$E[(\max_{1\leqq i\leqq n_1} \max_{1\leqq j\leqq n_2} |S_{ij}|)^2] \leqq 16\sigma_{n_1n_2}^2\,. \tag{3.4}$$

Since Doob's ((1953) page 317) inequality holds for nonnegative submartingales, the proof of the lemma follows precisely on the same line as in the proof of Theorem 1 of Wichura (1969), and hence, the details are omitted. The extension of (3.4) for a general c ($\geqq 2$) is immediate; we need to replace 16 by 4^c and $\sigma_{n_1n_2}^2$ by $E[S^2(X_{11}, \cdots, X_{1n_1}, \cdots, X_{c1}, \cdots, X_{cn_c})]$.

Consider now a two-sample U-statistic $U_{n_1n_2}$ and denote by $r^2(n_1, n_2) = \operatorname{Var}(U_{n_1n_2})$; note that by (2.5), $r^2(n_1, n_2)$ is non-increasing in each of n_1 and n_2.

LEMMA 3.2. *For every* $N_j \geqq n_j \geqq m_j\ (> 1)$, $j = 1, 2$,

$$\begin{aligned} &E[(\max_{n_1\leqq k\leqq N_1} \max_{n_2\leqq q\leqq N_2} r^{-1}(n_1, n_2)|U_{kq} - U_{kN_2} - U_{N_1q} + U_{N_1N_2}|)^2] \\ &\quad \leqq 16r^{-2}(n_1, n_2)[r^2(n_1, n_2) - r^2(n_1, N_2) - r^2(N_1, n_2) + r^2(N_1, N_2)] \\ &\quad \leqq 16, \qquad \textit{uniformly in} \quad N_1(\geqq n_1) \quad \textit{and} \quad N_2(\geqq n_2)\,. \end{aligned} \tag{3.5}$$

PROOF. For every $r \geqq 1$, $s \geqq 1$, we let

$$h_{rs}^{(N_1N_2)} = U_{N_1-r+1\,N_2-s+1} - U_{N_1-r\,N_2-s+1} - U_{N_1-r+1\,N_2-s} + U_{N_1-r\,N_2-s}\,, \tag{3.6}$$

so that for every $1 \leqq i \leqq N_1 - n_1$, $1 \leqq j \leqq N_2 - n_2$,

$$\begin{aligned} S_{ij} = \sum_{r=1}^{i} \sum_{s=1}^{j} h_{rs}^{(N_1N_2)} &= U_{N_1-i+1\,N_2-j+1} - U_{N_1-i+1\,N_2} \\ &\quad - U_{N_1\,N_2-j+1} + U_{N_1N_2}\,. \end{aligned} \tag{3.7}$$

Let, now $\mathscr{C}_n^{(j)}$ be the σ-field generated by the unordered collection $\{X_{j1}, \cdots, X_{jn}\}$ and by $X_{jn+1}, X_{jn+2}, \cdots, j = 1, 2$, and $\mathscr{C}_{n_1n_2}^{(12)}$ be the product σ-field $\mathscr{C}_{n_1}^{(1)} \times \mathscr{C}_{n_2}^{(2)}$ for $n_1 \geqq m_1$, $n_2 \geqq m_2$. It follows by standard arguments that $E(U_{kq} \mid \mathscr{C}_{k'q'}) = U_{k'q'}$ a.e. for every $k \leqq k'$ and $q \leqq q'$. Consequently, it follows by some routine steps that for every $1 \leqq i \leqq N_1 - n_1$, $1 \leqq j \leqq N_2 - n_2$,

$$E(S_{ij} \mid \mathscr{C}_{i'j}^{*}) = S_{i'j} \quad \text{a.e.,} \qquad i' \leqq i\,, \tag{3.8}$$

$$E(\mathbf{S}_{N_1-n_1}^{(j)} \mid \mathscr{C}_{N_1-n_1,j'}^{*}) = \mathbf{S}_{N_1-n_1}^{(j')} \quad \text{a.e.,} \qquad j \geqq j'\,, \tag{3.9}$$

where $\mathscr{C}^*_{kq} = \mathscr{C}^{(12)}_{N_1-kN_2-q}$, $1 \leqq k \leqq N_1 - n_1$, $1 \leqq q \leqq N_2 - n_2$, and $\mathbf{S}_k^{(j)} = (S_{1j}, \cdots, S_{kj})'$, for $k \geqq 1$, $j \geqq 1$. Further, by (3.7), $E(S_{ij}) = 0$ for all $i \geqq 1$, $j \geqq 1$. Finally, $\mathscr{C}^*_{kq}$ is $\uparrow$ in k and q. Consequently, the same proof as in Lemma 3.1 holds, and (3.5) follows by noting that by (3.7),

$$(3.10)\qquad E(S^2_{N_1-n_1N_2-n_2}) = r^2(n_1, n_2) - r^2(n_1, N_2) - r^2(N_1, n_2 + r^2(N_1, N_2) \leqq r^2(n_1, n_2) \quad \text{for every} \quad N_1 \geqq n_1, N_2 \geqq n_2 ,$$

as $r^2(i, j)$ is $\downarrow$ in i and j. □

We further note that $\{U_{kN_2}, \mathscr{C}_{kN_2}; n_1 \leqq k \leqq N_1\}$ has the reverse martingale property, so that by reversing the index set, we obtain that

$$(3.11)\qquad E[(\max_{n_1 \leqq k \leqq N_1} r^{-1}(n_1, N_2)|U_{kN_2} - U_{N_1N_2}|)^2] \leqq 4r^{-2}(n_1, N_2)[r^2(n_1, N_2) - r^2(N_1, N_2)] \leqq 4 ,$$

uniformly in $N_1 \geqq n_1$, $N_2 \geqq n_2$. Similarly,

$$(3.12)\qquad E[(\max_{n_2 \leqq q \leqq N_2} r^{-1}(N_1, n_2)|U_{N_1q} - U_{N_1N_2}|)^2] \leqq 4r^{-2}(N_1, n_2)[r^2(N_1, n_2) - r^2(N_1, N_2)] \leqq 4 ,$$

uniformly in $N_1 \geqq n_1$, $N_2 \geqq n_2$. Finally,

$$(3.13)\qquad E[|U_{N_1N_2} - \theta(\mathbf{F})|^2] = r^2(N_1, N_2) \leqq r^2(n_1, n_2) ,$$

uniformly in $N_1 \geqq n_1$, $N_2 \geqq n_2$. Hence, by Lemma 3.2, (3.11), (3.12), (3.13), the c_r-inequality and the Chebychev inequality, we obtain that for every $\varepsilon > 0$, there exists a positive K_ε ($< \infty$), such that for every $N_1 \gneqq n_1$, $N_2 \geqq n_2$,

$$(3.14)\qquad P\{\max_{n_1 \leqq k \leqq N_1} \max_{n_2 \leqq q \leqq N_2} |U_{kq} - \theta(\mathbf{F})| > r(n_1, n_2)K_\varepsilon\} < \varepsilon ,$$

and hence,

$$(3.15)\qquad P\{\sup_{\mathbf{k} \geqq \mathbf{n}} |U_{\mathbf{k}} - \theta(\mathbf{F})| > r(\mathbf{n})K_\varepsilon\} < \varepsilon .$$

We now consider a typical decomposition for $U(\mathbf{n})$. Let us write $U^*_{00}(n) = \theta(\mathbf{F})$ and for $\mathbf{k} \geqq \mathbf{0}$,

$$(3.16)\qquad U_{\mathbf{k}}^*(\mathbf{n}) = \sum_{d_1=0}^{k_1} \sum_{d_2=0}^{k_2} (-1)^{d_1+d_2} \binom{k_1}{d_1}\binom{k_2}{d_2} U_{d_1d_2}(\mathbf{n}) ; \qquad \mathbf{0} \leqq \mathbf{k} \leqq \mathbf{m} ,$$

where for $0 \leqq d_1 \leqq m_1$ and $0 \leqq d_2 \leqq m_2$, we have

$$(3.17)\qquad U_{d_1d_2}(\mathbf{n}) = \prod_{j=1}^{2} \binom{n_j}{d_j}^{-1} \sum\nolimits^*_{(\mathbf{n})} g_{d_1d_2}(X_{j\alpha_{ji}}, 1 \leqq i \leqq d_j, 1 \leqq j \leqq 2) ,$$

where the summation $\sum^*_{(\mathbf{n})}$ extends over all $1 \leqq \alpha_{ji} < \cdots < \alpha_{jd_j} \leqq m_j$, $j = 1$, 2. Then, by (1.2), (3.16), (3.17) and a few routine steps, we obtain that

$$(3.18)\qquad U(\mathbf{n}) = \sum_{k_1=0}^{m_1} \sum_{k_2=0}^{m_2} \binom{m_1}{k_1}\binom{m_2}{k_2} U_{\mathbf{k}}^*(\mathbf{n}) ,$$

where each $U_{\mathbf{k}}^*(\mathbf{n})$ ($\mathbf{0} \leqq \mathbf{k} \leqq \mathbf{m}$) is a generalized U-statistic. A few readjustments lead us to

$$(3.19)\qquad U(\mathbf{n}) = \theta(\mathbf{F}) + U_1(n_1) + U_2(n_2) + \sum_{d_1=1}^{m_1} \sum_{d_2=1}^{m_2} \binom{m_1}{d_1}\binom{m_2}{d_2} U_{\mathbf{d}}^*(\mathbf{n}) ,$$

where

$$U_1(n_1) = \binom{n_1}{m_1}^{-1} \sum\nolimits_{(n_1)}^{*} [g_{m_1 0}(X_{1i_1}, \cdots, X_{1i_{m_1}}) - \theta(\mathbf{F})] , \tag{3.20}$$

$$U_2(n_2) = \binom{n_2}{m_2}^{-1} \sum\nolimits_{(n_2)}^{*} [g_{0 m_2}(X_{2i_1}, \cdots, X_{2i_{m_2}}) - \theta(\mathbf{F})] \tag{3.21}$$

and the summation $\sum_{(n_j)}^{*}$ extends over all $1 \leqq i_1 < \cdots < i_{m_j} \leqq n_j$, $j = 1, 2$.

We first consider the proof of Theorem 2.2 which is relatively simpler. We write [by (3.19)]

$$\begin{aligned} U_3^*(\mathbf{n}) &= [U(\mathbf{n}) - \theta(\mathbf{F})] - U_1(n_1) - U_2(n_2) \\ &= \sum_{d_1=1}^{m_1} \sum_{d_2=1}^{m_2} \binom{m_1}{d_1}\binom{m_2}{d_2} U_{\mathbf{d}}^*(\mathbf{n}) , \end{aligned} \tag{3.22}$$

where $U_3^*(\mathbf{n})$ is a generalized U-statistic for which

$$\begin{aligned} [r^*(\mathbf{n})]^2 &= E[U_3^*(\mathbf{n})]^2 = r^2(\mathbf{n}) - E[U_1(n_1)]^2 - E[U_2(n_2)]^2 \\ &= \binom{n_1}{m_1}^{-1}\binom{n_2}{m_2}^{-1} \sum_{d_1=1}^{m_1} \sum_{d_2=1}^{m_2} \binom{m_1}{d_1}\binom{m_2}{d_2}\binom{n_1-m_1}{m_1-d_1}\binom{n_2-m_2}{m_2-d_2}\zeta_{d_1 d_2}(\mathbf{F}) \\ &= n_1^{-1}n_2^{-1}m_1^2 m_2^2 \zeta_{11}(\mathbf{F}) + O(n^{-3}). \end{aligned} \tag{3.23}$$

Thus, from (2.5) and (3.23), we note that under (1.3),

$$r^*(\mathbf{n})/r(\mathbf{n}) = O(n^{-\frac{1}{2}}) , \tag{3.24}$$

and hence, using (3.15) for $\{U_3^*(\mathbf{k}), \mathbf{k} \geqq \mathbf{n})$ and (3.24), we conclude that for every $\varepsilon > 0$ and $\eta > 0$, there exists an $n_0[= n_0(\varepsilon, \eta)]$, such that for $n \geqq n_0$,

$$P\{\sup_{\mathbf{k} \geqq \mathbf{n}} |U_3^*(\mathbf{k})| > \eta \cdot r(\mathbf{n})\} < \varepsilon . \tag{3.25}$$

Therefore, under (1.3), (2.3) and (2.4), as $n \to \infty$,

$$\sup_{\mathbf{k} \geqq \mathbf{n}} r^{-1}(\mathbf{n}) |U(\mathbf{k}) - \theta(\mathbf{F}) - U_1(k_1) - U_2(k_2)| \to_P 0 . \tag{3.26}$$

Let us now define for each $j\,(= 1, 2)$,

$$W_j^*(t, n_j) = \sigma_j^{-1} n_j^{\frac{1}{2}} U_j([n_j/t]) , \quad 0 < t \leqq 1 , \quad \text{and} \quad 0 \quad \text{for} \quad t = 0 , \tag{3.27}$$

$$w_j(\mathbf{n}) = r^{-1}(\mathbf{n}) \sigma_j n_j^{-\frac{1}{2}} . \tag{3.28}$$

Then, from (2.9), (3.26), (3.27) and (3.28), we obtain that

$$\sup_{\mathbf{t} \in E_2} |W^*(\mathbf{t}; \mathbf{n}) - \sum_{j=1}^{2} w_j(\mathbf{n}) W_j^*(t_j, n_j)| \to_P 0 \qquad \text{as } n \to \infty . \tag{3.29}$$

Now, by (1.3), (2.5), (2.10) and (3.28), we obtain that

$$\lim_{n \to \infty} w_j(\mathbf{n}) = w_j \qquad \text{for} \quad j = 1, 2 . \tag{3.30}$$

Also, $W_j^*(n_j) = \{W_j^*(t_j n_j), 0 \leqq t \leqq 1\}, j = 1, 2$, are stochastically independent, and by the results of Loynes (1970), as $n \to \infty$,

$$W_j^*(n_j) \to_{\mathscr{D}} W_j = \{W_j(t) \colon 0 \leqq t \leqq 1\} , \qquad j = 1, 2 \tag{3.31}$$

in the Skorokhod J_1-topology on $D[0, 1]$, where W_1 and W_2 are independent copies of a standard Brownian motion. (2.13) follows from (3.27) through (3.31). □

To prove Theorem 2.1, we note on using (3.23) and some standard inequalities among the $\{\zeta_{\mathbf{d}}(\mathbf{F}): \mathbf{0} \leqq \mathbf{d} \leqq \mathbf{m}\}$ that $r^*(\mathbf{n}) \leqq C(\mathbf{F})(n_1 n_2)^{-\frac{1}{2}}$, $\forall\, \mathbf{n} \geqq \mathbf{m}$, where $C(\mathbf{F}) < \infty$ (whenever (2.4) holds) and it does not depend on $\mathbf{n}$. First, we show that under (1.3), (2.3) and (2.4) as $n \to \infty$,

$$\max_{\mathbf{m} \leq \mathbf{k} \leq \mathbf{n}} \psi(\mathbf{k}, n)|U_3^*(\mathbf{k})| \to_P 0 . \tag{3.32}$$

For this, we partition the set $A(\mathbf{n}) = \{\mathbf{k} : \mathbf{m} \leqq \mathbf{k} \leqq \mathbf{n}\}$ into three disjoint subsets: $A_1(\mathbf{n}) = \{\mathbf{k} \in A(\mathbf{n}) : \min_{1 \leq j \leq 2} k_j \leqq \varepsilon_1 n^{\frac{1}{2}}\}$, $A_2(\mathbf{n}) = \{\mathbf{k} \in A(\mathbf{n}) : \varepsilon_1 n^{\frac{1}{2}} < \min_{1 \leq j \leq 2} k_j \leqq [n^{\frac{4}{5}}]\}$, and $A_3(\mathbf{n}) = \{\mathbf{k} \in A(\mathbf{n}) : \min_{1 \leq j \leq 2} k_j > [n^{\frac{4}{5}}]\}$, where ε_1 (> 0) is an arbitrarily small number. The proof of (3.32) rests on the fact that $[\max_{\mathbf{k}} \psi(\mathbf{k}, n) \cdot \min_{\mathbf{k}} r(\mathbf{k})]$ for $\mathbf{k} \in A_i(\mathbf{n})$, $i = 1, 2, 3$, all converge to zero, as $n \to \infty$. The proof follows by using Lemma 3.2 and a few routine steps with are omitted for intended brevity.

On $D[0, 1]$, we now consider independent processes $W_j(n_j) = (W_j(t, n_j): 0 \leqq t \leqq 1\}$, $j = 1, 2$, where for $0 \leqq t \leqq 1$,

$$\begin{aligned} W_j(t, n_j) &= ([n_j t]/\sigma_j n_j^{\frac{1}{2}}) U_j([n_j t]) , && m_j \leqq [n_j t] \leqq n_j , \\ &= 0,\ t \leqq (m_j - 1)/n_j ; && j = 1, 2 . \end{aligned} \tag{3.33}$$

Also, let for every $\mathbf{t} > \mathbf{0}$.

$$\begin{aligned} w_j(\mathbf{t}, \mathbf{n}) &= \psi([\mathbf{nt}]; n)\{\sigma_j n_j^{\frac{1}{2}}/[n_j t_j]\} \\ &= n^{-\frac{1}{2}}(\textstyle\sum_{j=1}^2 \sigma_j \lambda_j^{\frac{1}{2}})(\sum_{j=1}^2 \sigma_j \lambda_j^{\frac{1}{2}}/[n_j t_j])^{-1}(\sigma_j n_j^{\frac{1}{2}}/[n_j t_j]) \\ &= (\textstyle\sum_{j=1}^2 \sigma_j \lambda_j^{\frac{1}{2}})(\sum_{j=1}^2 \sigma_j \lambda_j^{\frac{1}{2}}/[n_j t_j])^{-1}(\lambda_j^{\frac{1}{2}} \sigma_j/[n_j t_j])[1 + o(1)] . \end{aligned} \tag{3.34}$$

Then, for every $n_j t_j \geqq 1$, $1 \leqq j \leqq 2$, $w_j(\mathbf{t}, \mathbf{n})$ is positive and is bounded from above by $(\sum_{j=1}^2 \sigma_j \lambda_j^{\frac{1}{2}})(n_j/n\lambda_j)^{\frac{1}{2}} = (\sum_{j=1}^2 \sigma_j \lambda_j^{\frac{1}{2}})[1 + o(1)]$, $j = 1, 2$. Also, for every $\mathbf{t} > \mathbf{0}$,

$$\begin{aligned} &\lim_{n \to \infty} w_j(\mathbf{t}, \mathbf{n}) \\ &\quad = (\textstyle\sum_{j=1}^2 \sigma_j \lambda_j^{\frac{1}{2}})(\sum_{j=1}^2 \sigma_j \lambda_j^{\frac{1}{2}} t_j^{-1})^{-1}(\sigma_j \lambda_j^{-\frac{1}{2}} t_j^{-1}) , \qquad j = 1, 2 . \end{aligned} \tag{3.35}$$

Finally, it follows from Miller and Sen (1972) that $W_1(n_1)$ and $W_2(n_2)$ are stochastically independent, each converging in law to a standard Brownian motion. Consequently, for every $\varepsilon > 0$ and $\eta > 0$, there exist a $\delta > 0$ and an n_0, such that for $n \geqq n_0$,

$$P\{\sup_{0 \leq t \leq \delta} |W_j(t; n_j)| > \varepsilon\} < \tfrac{1}{2}\eta , \qquad j = 1, 2 , \tag{3.36}$$

$$\sup_{0 \leq t \leq 1} |W_j(t; n_j)| = O_p(1) , \qquad j = 1, 2 . \tag{3.37}$$

Thus, (2.8) follows from (2.6), (2.7), (3.22) and (3.32) through (3.37). □

4. Weak convergence of von Mises' functionals. We define the empirical dfs by

$$F_j(x, n_j) = n_j^{-1} \textstyle\sum_{i=1}^{n_j} c(x - X_{ji}) , \qquad x \in R^p,\ n_j \geqq 1, \text{ for } j = 1, \cdots, c , \tag{4.1}$$

where $c(u) = 1$ or 0 according as all the p components of u are nonnegative or at least one negative; we let $F(\cdot, \mathbf{n}) = (F_1(\cdot, n_1), \cdots F_c(\cdot, n_c))$. Then, the von Mises (1947) functional corresponding to (1.1) is

$$\theta(\mathbf{F}(\cdot, \mathbf{n})) = \int_{R^{pm_0}} \cdots \int g(x_{11}, \cdots, x_{cm_c}) \prod_{j=1}^{c} \prod_{i=1}^{m_j} dF_j(x_{ji}; n_j) . \tag{4.2}$$

Here, we define a process $\tilde{W}(\mathbf{n}) = \{\tilde{W}(\mathbf{t}; \mathbf{n}); \mathbf{t} \in E_c\}$ as in (2.6), where we replace $\{U(\mathbf{k}); \mathbf{k} \in A(\mathbf{n})\}$ by $\{\theta(\mathbf{F}(\cdot, \mathbf{k})); \mathbf{k} \in A(\mathbf{n})\}$. Similarly, on replacing $\{U(\mathbf{k}); \mathbf{k} \geqq \mathbf{n}\}$ by $\{\theta(\mathbf{F}(\cdot, \mathbf{k}); \mathbf{k} \geqq \mathbf{n}\}$ in (2.9), we define $\tilde{W}^*(\mathbf{n}) = \{\tilde{W}^*(\mathbf{t}; \mathbf{n}); \mathbf{t} \in E_c\}$. Finally we strengthen (2.4) to

$$\zeta^*(\mathbf{F}) = \max_{1 \leqq j \leqq c} \max_{1 \leqq \alpha_{j1} \leqq \cdots \leqq \alpha_{jm_j} \leqq m_j} E\{g^2(X_{1\alpha_{11}}, \cdots, X_{c\alpha_{cm_c}})\} < \infty . \tag{4.3}$$

THEOREM 4.1. *Under* (1.3), (2.3) *and* (4.3), *as* $n \to \infty$, $\tilde{W}(\mathbf{n})$ *and* $\tilde{W}^*(\mathbf{n})$ *converge in law in the extended Skorokhod* J_1*-topology on* D_c *to* W *and* W^*, *respectively, which are defined in* (2.8) *and* (2.12).

PROOF. Again, we consider the case of $c = 2$, and for $\mathbf{0} \leqq \mathbf{d} \leqq \mathbf{m}$, define

$$V_{\mathbf{d}}^*(\mathbf{n}) = \int_{R^{pd_0}} \cdots \int g_{\mathbf{d}}(x_{11}, \cdots, x_{1d_1}, x_{21}, \cdots, x_{2d_2}) \times \prod_{j=1}^{2} \prod_{i=1}^{d_j} d[F_j(x_{ji}, n_j) - F_j(x_{ji})] , \tag{4.4}$$

where $d_0 = d_1 + d_2$. Then, we may write

$$\begin{aligned} \theta(\mathbf{F}(\cdot, \mathbf{n})) &= \textstyle\sum_{\mathbf{d}=\mathbf{0}}^{\mathbf{m}} \binom{m_1}{d_1}\binom{m_2}{d_2} V_{\mathbf{d}}(\mathbf{n}) \\ &= \theta(\mathbf{F}) + V_1(n_1) + V_2(n_2) + \textstyle\sum_{\mathbf{d}=\mathbf{1}}^{\mathbf{m}} \binom{m_1}{d_1}\binom{m_2}{d_2} V_{\mathbf{d}}^*(\mathbf{n}) , \end{aligned} \tag{4.5}$$

where

$$V_1(n_1) = \int_{R^{pm_1}} \cdots \int g_{m_1 0}(x_1, \cdots, x_{m_1}) \prod_{i=1}^{m_1} dF_1(x_i, n_1) - \theta(\mathbf{F}) ; \tag{4.6}$$

$$V_2(n_2) = \int_{R^{pm_2}} \cdots \int g_{0 m_2}(x_1, \cdots, x_{m_2}) \prod_{i=1}^{m_2} dF_2(x_i, n_2) - \theta(\mathbf{F}) . \tag{4.7}$$

Now, proceeding as in the proof of Lemma 2.6 of Miller and Sen (1972), we have for every $k \geqq m_j$,

$$E[V_j(k) - U_j(k)]^2 \leqq C(\mathbf{F})k^{-3} , \qquad \text{for } j = 1, 2 , \tag{4.8}$$

where, under (4.3), $C(\mathbf{F}) < \infty$, and $U_j(n_j)$, $j = 1, 2$, are defined in (3.20) and (3.21). Since, $\psi(\mathbf{k}, n) \leqq n^{-\frac{1}{2}}\delta^{-1}k_j$ for $j = 1, 2$, by (4.8), for each $j\,(= 1, 2)$,

$$\begin{aligned} P(\max_{\mathbf{m} \leq \mathbf{k} \leq \mathbf{n}} \psi(\mathbf{k}, n)|U_j(k_j) - V_j(k_j)| > \varepsilon\} &\leqq \textstyle\sum_{k=m_j}^{n_j} n^{-1}\delta^{-2}k^2 E[U_j(k) - V_j(\mathbf{k})]^2 \varepsilon^{-2} \\ &\leqq C(\mathbf{F})n^{-1}\delta^{-2}\varepsilon^{-2} \textstyle\sum_{k=m_j}^{n_j} k^{-1} = C(\mathbf{F})(\delta\varepsilon)^{-2} \cdot O(n^{-1}\log n) , \end{aligned} \tag{4.9}$$

so that for every $\varepsilon > 0$, the right-hand side of (4.9) converges to 0 as $n \to \infty$. Similarly, on noting that $r^2(\mathbf{n})$, defined by (2.5) is bounded below by $\min_{1 \leqq j \leqq c}(\sigma_j^2/n_j)$, we obtain that as $n \to \infty$, under (1.3), for every $\varepsilon > 0$, $j = 1, 2$,

$$\begin{aligned} P\{\max_{\mathbf{k} \geq \mathbf{n}} r^{-1}(\mathbf{n})|U_j(k_j) - V_j(k_j)| > \varepsilon\} &\leqq \varepsilon^{-2}r^{-2}(\mathbf{n}) \textstyle\sum_{k=n_j}^{\infty} E\{[U_j(k) - V_j(k)]^2\} \\ &\leqq \varepsilon^{-2}r^{-2}(\mathbf{n})C(\mathbf{F}) \cdot \textstyle\sum_{k=n_j}^{\infty} k^{-3} \\ &= C(\mathbf{F})\varepsilon^{-2}[O(n^{-1})] \to 0 \qquad \text{as } n \to \infty . \end{aligned} \tag{4.10}$$

Consequently, to prove the theorem, all we need to show is this under (1.3), (2.3) and (4.3), for every $\mathbf{d} > \mathbf{1}$,

$$P\{\max_{\mathbf{m}\leq\mathbf{k}\leq\mathbf{n}} \psi(\mathbf{k}, n)|U_{\mathbf{d}}^*(\mathbf{k}) - V_{\mathbf{d}}^*(\mathbf{k})| > \varepsilon\} < \eta \,, \tag{4.11}$$

$$P\{\max_{\mathbf{k}\geq\mathbf{n}} r^{-1}(\mathbf{n})|U_{\mathbf{d}}^*(\mathbf{k}) - V_{\mathbf{d}}^*(\mathbf{k})| > \varepsilon\} < \eta \,, \tag{4.12}$$

where both ε (> 0) and η (> 0) are arbitrarily small, and n is chosen adequately large. Note that $V_{11}^*(\mathbf{k}) = U_{11}^*(\mathbf{k})$ for all $\mathbf{k} > \mathbf{m}$. Also, by extending the proof of Lemma 2.6 of Miller and Sen (1972), we have for every $\mathbf{d} \geqq \mathbf{1}$ [under (4.3)],

$$E\{[U_{\mathbf{d}}^*(\mathbf{k}) - V_{\mathbf{d}}^*(\mathbf{k})]^2\} \leqq C(\mathbf{F})k_1^{-d_1}k_2^{-d_2}\{k_1^{-2} + k_2^{-2}\} \,, \tag{4.13}$$

for $\mathbf{k} \geqq \mathbf{m}$. Consequently, if $d_1 \geqq 2$, $d_2 \geqq 2$, the proof for (4.11) and (4.12) follow trivially by using (4.13), the Chebychev and the Bonferroni inequalities. Thus, we need to consider the case where $\min_{1\leq j\leq 2} d_j = 1$, but $d_1 + d_2 \geqq 3$. We consider explicitly the case of $\mathbf{d} = (1, 2)$; the case of (2, 1) follows similarly, while for any other $\mathbf{d}$: $1 = \min(d_1, d_2) < \max(d_1\, d_2)(\geqq 2)$, a similar but more laborious proof holds.

By direct evaluation from (3.16) and (4.4), we have

$$V_{12}^*(\mathbf{k}) - U_{12}^*(\mathbf{k}) = \frac{1}{k_2} U_{12}^*(\mathbf{k}) + \frac{1}{k_2 - 1} U_{12}^{**}(\mathbf{k}) \,, \tag{4.14}$$

for every $\mathbf{k} \geqq \mathbf{m}$, where for $m_2 \geqq 2$,

$$\begin{aligned} U_{12}^{**}(\mathbf{k}) = \frac{1}{k_1 k_2} \textstyle\sum_{i_1=1}^{k_1} \sum_{i_2=1}^{k_2} \int_{R^{3p}} \cdots \int g_{12}(x_1, y_1, y_2) \\ \times\, d[c(x_1 - X_{1i_1}) - F_1(x_1)] \\ \times\, d[c(y_1 - X_{2i_2}) - F_2(y_1)]\, d[c(y_2 - X_{2i_2}) - F_2(y_2)] \,. \end{aligned} \tag{4.15}$$

Consequently, to prove (4.11) and (4.12) for $\mathbf{d} = (1, 2)$, it suffices to show that as $n \to \infty$,

$$\max_{\mathbf{m}\leq\mathbf{k}\leq\mathbf{n}} |U_{12}^*(\mathbf{k})| = O_p(1) \,, \qquad \max_{\mathbf{m}\leq\mathbf{k}\leq\mathbf{n}} |U_{12}^{**}(\mathbf{k})| = O_p(1) \,; \tag{4.16}$$

$$\sup_{\mathbf{k}\geq\mathbf{n}} |U_{12}^*(\mathbf{k})| = O_p(1) \,, \qquad \sup_{\mathbf{k}\geq\mathbf{n}} |U_{12}^{**}(\mathbf{k})| = O_p(1) \,; \tag{4.17}$$

note that $\psi(\mathbf{k}, n)/k_2 = O(n^{-\frac{1}{2}})$, $\forall\, \mathbf{k} \leq \mathbf{n}$ and $r^{-1}(\mathbf{n})k_2^{-1} \to 0$ as $n \to \infty$, for every $\mathbf{k} \geqq \mathbf{n}$. Since $U_{12}^*(\mathbf{k})$ is a generalized U-statistic, the proof of (4.16) and (4.17) for $\{U_{12}^*(\mathbf{k})\}$ follows directly from (3.15) and the fact that as in Section 3, $E[U_{12}^*(\mathbf{k})]^2 \leqq C(\mathbf{F})k_1^{-1}k_2^{-2}$ for every $\mathbf{k} \geqq \mathbf{m}$.

Now, we define the sequence of σ-fields $\{\mathscr{C}_{kq},\, k \geqq m_1,\, q \geqq m_2\}$ as in Section 3 [following (3.7)]. Then it follows from (4.15) that for every $k' \geqq k$ and $q' \geqq q$,

$$E[U_{12}^{**}(k, q) \,|\, \mathscr{C}_{k'q'}] = U_{12}^{**}(k', q') \quad \text{a.e.} \tag{4.18}$$

Consequently, following the same method of approach as in the proof of Lemma 3.2, we obtain from Lemma 3.1 that for every $\mathbf{N} \geqq \mathbf{n} \geqq \mathbf{m}$,

$$\begin{aligned} &E[(\max_{1\leq j\leq 2} \max_{n_j\leq k_j\leq N_j} |U_{12}^{**}(k_1, k_2)|)^2] \\ &\leqq 16E\{E[U_{12}^{**}(\mathbf{n})]^2 - E[U_{12}^{**}(n_1, N_2)]^2 - E[U_{12}^{**}(N_1 n_2)] + E[U_{12}^{**}(\mathbf{N})]^2\} \\ &\leqq 16E[U_{12}^{**}(\mathbf{n})]^2 \leqq 16C(\mathbf{F})n_1^{-1}n_2^{-2} \,. \end{aligned} \tag{4.19}$$

Thus, the proof of (4.16) and (4.17) for $\{U_{12}^{**}(\mathbf{k})\}$ follows immediately from (4.19) and the Chebychev inequality. □

5. Weak convergence for random indices. We now consider the case where in (1.3) we allow $(\lambda_1, \cdots, \lambda_c)' = \boldsymbol{\lambda}$ to be a stochastic vector with positive elements. More precisely, let $\{\mathbf{N}_n = (N_n^{(1)}, \cdots, N_n^{(c)})', n \geqq 1\}$ be a sequence of vectors with positive integer valued random variables, such that

$$n^{-1}\mathbf{N}_n \to_P \boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_c)' , \qquad \text{as } n \to \infty , \tag{5.1}$$

where λ_j, $j = 1, \cdots, c$ are positive random variables defined on the same probability space as of the original $\{X_{ji}, i \geqq 1\}$, $j = 1, \cdots, c$.

We define $W(\mathbf{N}_n)$ as in (2.6) when $\mathbf{N}_n \geqq \mathbf{1}$; otherwise, we let $W(\mathbf{N}_n) = 0$. Similarly, we define $W^*(\mathbf{N}_n)$ as in (2.9) when $\mathbf{N}_n \geqq \mathbf{1}$; otherwise, we let $W^*(\mathbf{N}_n) = 0$. Finally, we define $\tilde{W}(\mathbf{N}_n)$ and $\tilde{W}^*(\mathbf{N}_n)$ as in Section 4, with n being replaced by $\mathbf{N}_n$. Our basic problem is to study the weak convergence of $W(\mathbf{N}_n)$, $W^*(\mathbf{N}_n)$, $\tilde{W}(\mathbf{N}_n)$ and $\tilde{W}^*(\mathbf{N}_n)$, when $n \to \infty$ and (5.1) holds.

THEOREM 5.1. *Under* (2.3), (2.4) *and* (5.1), $W(\mathbf{N}_n)$ *and* $W^*(\mathbf{N}_n)$ *converge in law in the extended Skorokhod J_1-topology on D_c to W and W^*, respectively, defined in* (2.8) *and* (2.12), *while under* (2.3), (4.3) *and* (5.1), $\tilde{W}(\mathbf{N}_n)$ *and* $\tilde{W}^*(\mathbf{N}_n)$ *weakly converge to W and W^* respectively.*

PROOF. We shall only consider the case of $W(\mathbf{N}_n)$, as the other cases follow similarly. By Theorem 2.1 and Lemma 3.2, we can verify the conditions of Theorem 2 of Mogyorodi (1967) which insures that the finite dimensional distributions of $\{W(\mathbf{N}_n)\}$ converge to those of W. To establish the tightness of $\{W(\mathbf{N}_n)\}$, we define the *supremum metric* $\rho(x, y) = \sup_{\mathbf{t} \in E_c}|x(\mathbf{t}) - y(\mathbf{t})|$, and the *modulus of continuity* (for $\delta > 0$) $\omega_\delta(x) = \sup\{|x(\mathbf{t}) - x(\mathbf{s})| : \mathbf{t}, \mathbf{s} \in E_c, |\mathbf{t} - \mathbf{s}| < \delta\}$. Now, using our Lemma 3.2, (3.11) and (3.12) in place of Theorem 1 of Wichura (1969) and thereby extending his (2a) to our statistics, we obtain by the same technique as in his proof of Theorem 3 (on pages 686–687) that under (1.3), (2.3) and (2.4), for every $\varepsilon > 0$, as $n \to \infty$,

$$\lim_{\delta \downarrow 0} \limsup_n P\{\omega_\delta(W(\mathbf{n})) > \varepsilon\} = 0 . \tag{5.2}$$

Consider now a sequence of generalized U-statistics

$$\begin{aligned} U'([\mathbf{nt}]) &= \textstyle\prod_{j=1}^{c} \binom{n_j}{m_j}^{-1} \sum_{(\mathbf{nt})}^{**} g^*(X_{j\alpha_{ji}}, 1 \leqq i \leqq m_j, 1 \leqq j \leqq c) \\ & \qquad t_j \geqq n_j^{-1}k_n^{(j)} , \qquad\qquad 1 \leqq j \leqq c ; \\ &= 0 , \qquad \text{if} \quad t_j < n_j^{-1}k_n^{(j)} \quad \text{for some} \quad j\,(= 1, \cdots, c) , \end{aligned} \tag{5.3}$$

where $g^*(x_{11}, \cdots, x_{cm_c}) = g(x_{11}, \cdots, x_{cm_c}) - \theta(\mathbf{F})$, the summation $\sum_{(\mathbf{nt})}^{**}$ extends over all $k_n^{(j)} + 1 \leqq \alpha_{j1} < \cdots < \alpha_{jm_j} \leqq [n_j t_j]$, and $\lim_{n\to\infty} k_n^{(j)} = \infty$ but $\lim_{n\to\infty} n^{-\frac{1}{2}}k_n^{(j)} = 0$, $1 \leqq j \leqq c$. Then, on replacing $U([\mathbf{nt}]) - \theta(\mathbf{F})$ in (2.6) by $U'([\mathbf{nt}])$, $\mathbf{t} \in \mathbf{E}_c$, we define a parallel process $W'(\mathbf{n}) = \{W'(\mathbf{t}; \mathbf{n}) : \mathbf{t} \in \mathbf{E}_c\}$.

LEMMA 5.2. *Under* (1.3), (2.3) *and* (2.4), $\rho(W(\mathbf{n}), W'(\mathbf{n})) \to 0$ *almost surely as* $n \to \infty$.

PROOF. Here also, we let $c = 2$. For $\mathbf{0} \leqq \mathbf{l} \leqq \mathbf{m} \leqq \mathbf{k}_n$, $\mathbf{k} - \mathbf{k}_n \geqq \mathbf{m} - \mathbf{l}$, we define

$$U_{\mathbf{l}}(\mathbf{k}, \mathbf{k}_n) = \left\{\prod_{j=1}^{2} \binom{k_n^{(j)}}{l_j} \binom{k_j - k_n^{(j)}}{m_j - l_j}\right\}^{-1} \times \hat{\sum}_{(\mathbf{k})}^{*} g^*(X_{j\alpha_{ji}}, 1 \leqq i \leqq m_j, 1 \leqq j \leqq 2)\,, \tag{5.4}$$

where the summation $\hat{\sum}_{(\mathbf{k})}^{*}$ extends over all $1 \leqq \alpha_{jl} < \cdots < \alpha_{jl_j} \leqq k_n^{(j)} < \alpha_{ji_{l_j+1}} < \cdots \leqq \alpha_{ji_{m_j}} \leqq k_j$, $1 \leqq j \leqq 2$. Thus $U_{\mathbf{l}}(\mathbf{k}, \mathbf{k}_n)$ may be interpreted as a generalized U-statistic based on four samples of sizes $(k_n^{(1)}, k_1 - k_n^{(1)}, k_n^{(2)}, k_2 - k_n^{(2)})$. Consequently, as in our Lemma 3.2, (3.11), (3.12), (3.13), (3.14) and (3.15), it can be shown that for every $\varepsilon > 0$, there exist a positive K_ε ($< \infty$) and an $n_0(\varepsilon)$, such that for $n \geqq n_0(\varepsilon)$, under (1.3) and (2.4),

$$P(\max_{\mathbf{0}\leqq \mathbf{l} \leqq \mathbf{m}} \sup_{\mathbf{k} > \mathbf{k}_n} |U_{\mathbf{l}}(\mathbf{k}, \mathbf{k}_n)| > K_\varepsilon\} < \varepsilon\,. \tag{5.5}$$

Also, by (1.2), (5.3) and (5.4), we obtain that for $[\mathbf{nt}] \geqq \mathbf{k}_n$,

$$\begin{aligned} U([\mathbf{nt}]) - \theta(\mathbf{F}) \\ = \sum_{\mathbf{l}=\mathbf{0}}^{\mathbf{m}} \prod_{j=1}^{2} \left\{\binom{m_j}{l_j}\binom{k_j - k_n^{(j)}}{m_j - l_j}\binom{k_n^{(j)}}{l_j}\binom{k_j}{m_j}^{-1}\right\} U_{\mathbf{l}}([\mathbf{nt}], \mathbf{k}_n)\,, \end{aligned} \tag{5.6}$$

where we let $k_j = [n_j t_j]$, $j = 1, 2$, Consequently, by (5.5) and (5.6), as $n \to \infty$,

$$\begin{aligned} &\sup_n \max_{\mathbf{k}_n \leqq \mathbf{k} \leqq \mathbf{n}} \{n^{-\frac{1}{2}}[\min(k_1, k_2)]|U(\mathbf{k}) - U'(\mathbf{k})|\} \\ &\leqq \{\sup_n \max_{\mathbf{k}_n \leqq \mathbf{k} \leqq \mathbf{n}} \max_{\mathbf{0}\leqq \mathbf{l} \leqq \mathbf{m}} |U_{\mathbf{l}}(\mathbf{k}, \mathbf{k}_n)|\}\{O([n^{-1}(k_n^{(1)})^2], [n^{-1}(k_n^{(2)})^2])\} \\ &= O_p((n^{-\frac{1}{2}}k_n^{(1)})^2, (n^{-\frac{1}{2}}k_n^{(2)})^2) = o_p(1)\,. \end{aligned} \tag{5.7}$$

Thus, by (2.6), (2.7) (3.39) and (5.7), as $n \to \infty$,

$$\sup_{n^{-1}\mathbf{k}_n \leqq \mathbf{t} \leqq \mathbf{1}} |W(\mathbf{t}; \mathbf{n}) - W'(\mathbf{t}; \mathbf{n})| \to 0\,, \quad \text{almost surely (a.s.)} \tag{5.8}$$

On the other hand, for $\mathbf{t} \in E_c^{(n)} = E_c - \{\mathbf{t}: n^{-1}\mathbf{k}_n \leqq \mathbf{t} \leqq \mathbf{1}\}$, $W'(\mathbf{t}; \mathbf{n}) = 0$ [by (5.3)], so that

$$\begin{aligned} &\sup_{\mathbf{t} \in E_c^{(n)}} |W(\mathbf{t}; \mathbf{n}) - W'(\mathbf{t}; \mathbf{n})| \\ &= \sup_{\mathbf{t} \in E_c^{(n)}} |W(\mathbf{t}; \mathbf{n})| \\ &\leqq \{\max_{\mathbf{m} \leqq \mathbf{k} \leqq \mathbf{n}} |U(\mathbf{k}) - \theta(\mathbf{F})|\}\{\sup_{\mathbf{t} \in E_c^{(n)}} \psi(\{\mathbf{nt}\}, n)\} \\ &\leqq \{\sup_{\mathbf{k} \leqq \mathbf{m}} |U(\mathbf{k}) - \theta(\mathbf{F})|\}\{O([n^{-\frac{1}{2}}k_n^{(1)}]^2, [n^{-\frac{1}{2}}k_n^{(2)}]^2)\} \\ &= O(1) \cdot o(1) = o(1)\,, \quad \text{a.s.,} \qquad \text{as } n \to \infty\,. \end{aligned} \tag{5.9}$$

The lemma follows from (5.8) and (5.9).

LEMMA 5.3. *If* $A \in A$, *then for every* $\varepsilon > 0$ *and* $\eta > 0$, *there exists a* δ (> 0), *sufficiently small, such that* $\limsup_n P\{\omega_\delta(W'(\mathbf{n})) > \varepsilon \mid A\} < \eta$.

PROOF. By (5.2) and Lemma 5.2, for every $\varepsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_n P\{\omega_\delta(W'(\mathbf{n})) > \varepsilon\} < \eta\,.$$

Also, by definition in (5.3), $W'(\mathbf{n})$ depends only on the set $\{X_{ji}, k_n^{(j)} < i < n_j, j = 1, \cdots, c\}$. Hence, using Rényi's (1958) idea of mixing sequence of sets and proceeding as in Lemma 3 of Blum, Hanson and Rosenblatt (1963), the result follows. □

We return now to the proof of Theorem 5.1. By virtue of (5.1), for every $\delta' > 0$, $P\{|n^{-1}\mathbf{N}_n - \lambda| > \delta'\} \to 0$ as $n \to \infty$. As such, it can be shown by standard steps [as in Mogyorodi (1967)] that for every $\varepsilon > 0$, there exist an $\eta > 0$ and an integer n_0, such that $P\{\rho W(\mathbf{N}_n), W'(\mathbf{N}_n)) > \varepsilon\} < \eta$ for $n \geqq n_0$. Hence, to establish the tightness of $\{W(\mathbf{N}_n)\}$, it suffices to show that for every $\varepsilon > 0$,

$$\limsup_n P\{\omega_\delta(W'(\mathbf{N}_n)) > \varepsilon\} \to 0 \qquad \text{as } \delta \downarrow 0 . \tag{5.10}$$

We note that by (1.2), (2.6), (2.7), (5.2), (5.3), (5.6) and Lemma 5.2, for every $\varepsilon > 0$ and $\eta > 0$, there exist a $\delta > 0$ and an $n_0(\varepsilon, \eta)$, such that for $n \geqq n_0(\varepsilon, \eta)$,

$$P\{\max_{k:|\mathbf{k}-\mathbf{n}|<\delta n} \rho(W'(\mathbf{k}), W'(\mathbf{n})) > \varepsilon\} < \eta . \tag{5.11}$$

Hence, on using the inequality that $\omega_\delta(W'(\mathbf{N}_n)) \leqq 2\rho(W'(\mathbf{N}_n), W'([n\boldsymbol{\lambda}])) + \omega_\delta(W([n\boldsymbol{\lambda}]))$, it suffices to show that

$$\limsup_n P\{\omega_\delta(W'([n\boldsymbol{\lambda}])) > \varepsilon\} \to 0 \qquad \text{as } \delta \downarrow 0 , \tag{5.12}$$

$$P\{\rho(W'(\mathbf{N}_n), W'([n\boldsymbol{\lambda}])) > \varepsilon\} \to 0 \qquad \text{as } n \to \infty . \tag{5.13}$$

Consider the events $A(\mathbf{h}) = \{\boldsymbol{\lambda}: a_0(\eta) + h_j\delta' < \lambda_j \leqq a_0(\eta) + (h_j + 1)\delta', 1 \leqq j \leqq c\}$, $\mathbf{h} \geqq \mathbf{0}$, $a_0(\eta) > 0$, and let $\mathbf{a}(\mathbf{h}) = (a_0(\eta) + (h_j + \frac{1}{2})\delta', 1 \leqq j \leqq c)$, $\mathbf{h} \geqq \mathbf{0}$. Then, we have

$$P\{\omega_\delta(W'([n\boldsymbol{\lambda}])) > \varepsilon\} \leqq P\{\min_{1\leqq j\leqq c} \lambda_j \leqq a_0(\eta)\} + \textstyle\sum_{\mathbf{h}=\mathbf{0}}^{\infty} P\{A(\mathbf{h})\}P\{\omega_{\delta'}(W'([n\boldsymbol{\lambda}])) > \varepsilon \mid A(\mathbf{h})\} , \tag{5.14}$$

where we let $P(B \mid A) = 0$ when $P(A) = 0$. Since, when $A(\mathbf{h})$ holds, $|\boldsymbol{\lambda} - \mathbf{a}(\mathbf{h})| < \delta'$, it can be shown on using (5.11) and Lemma 5.3 that for every $\varepsilon > 0$ and $\eta > 0$, there exists a $\delta'(0 < \delta' < \frac{1}{2}a_0(\eta))$, and an $n_0(\varepsilon, \eta)$, such that for $n \geqq n_0(\varepsilon, \eta)$ and all $\mathbf{h} \geqq \mathbf{0}$, $P\{\omega_{\delta'}(W'([n\boldsymbol{\lambda}])) > \varepsilon \mid A(\mathbf{h})\} < \eta/2$, $\forall\, \mathbf{h} > \mathbf{0}$. Hence (5.12) holds. Finally, on writing

$$\begin{aligned} &P\{\rho(W'(\mathbf{N}_n), W'([n\boldsymbol{\lambda}])) > \varepsilon\} \\ &\quad \leqq P\{|n^{-1}\mathbf{N}_n - \boldsymbol{\lambda}| > \delta'\} + P\{\min_{1\leqq j\leqq c} \lambda_j \leqq a_0(\eta)\} + \textstyle\sum_{\mathbf{h}=\mathbf{0}}^{\infty} P\{A(\mathbf{h})\} \\ &\qquad P\{\rho(W'(\mathbf{N}_n), W'([n\boldsymbol{\lambda}])) > \varepsilon, |n^{-1}\mathbf{N}_n - \boldsymbol{\lambda}| \leqq \delta' \mid A(\mathbf{h})\} , \end{aligned} \tag{5.15}$$

the proof of (5.13) follows on parallel lines.

The theorem is useful in the context of sequential procedures based on generalized U-statistics; we may refer to Williams and Sen (1973) for such an application.

As an illustration of the uses of Theorems 2.1, 2.2, 4.1 and 5.1, we consider a simple case where $p = 1$, $c = 2$ and

$$\theta(F_1, F_2) = \textstyle\int_{-\infty}^{\infty} F_1(x)\, dF_2(x) = P\{X_{1i} \leqq X_{2j}\} , \tag{5.16}$$

and we assume that both F_1 and F_2 are continuous everywhere. Then

$\theta(F(\cdot\, \mathbf{n}) = U(\mathbf{n})$ is the Wilcoxon rank statistic

$$\text{(5.17)} \qquad \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} c(X_i, X_j)\,; \qquad c(u\ v) = 1\,, \quad u \leqq v$$
$$= 0\,, \quad u > u\,.$$

Here $m_1 = m_2 = 1$, and the summands $\{c(X_i, X_j), 1 \leqq i \leqq n, 1 \leqq j \leqq n_2\}$ are not all independent, so that Wichura's (1969) results do not hold. If the two distributions F_1 and F_2 are mutually overlapping, then (2.3) holds, while (2.4) holds for all F_1, F_2, as $c(u, v)$ is bounded. If $F_1 \equiv F_2$, $\theta(F_1, F_2) = \frac{1}{2}$, $\sigma_1^2 = \sigma_2^2 = \frac{1}{12}$, so that the results further simplify. Theorem 5.1 for the Wilcoxon statistic is useful for the problem of sequential testing and estimating $\theta(F_1, F_2)$.

Acknowledgments. The author is grateful to the associate editor and the referee for their useful comments on the paper.

REFERENCES

[1] Blum, J. R., Hanson, D. L., and Rosenblatt, J. I. (1963). On the central limit theorem for the sum of a random number of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiéte.* **1** 389–393.

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

[3] Fraser, D. A. S. (1957). *Nonparametric Methods in Statistics.* Wiley, New York.

[4] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

[5] Loynes, R. M. (1970). An invariance principle for reverse martingales. *Proc. Amer. Math. Soc.* **25** 56–64.

[6] Miller, R. G. Jr., and Sen, P. K. (1972). Weak convergence of U-statistics and von Mises' differentiable statistical functions. *Ann. Math. Statist.* **43** 31–41.

[7] Mises, R. von. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* **18** 309–348.

[8] Mogyorodi, J. (1967). Limit distributions for sequences of random variables with random indices. *Trans. Fourth Prague Confer. Infor. Th. Statist. Dec. Func. Rand. Proc.* 463–470.

[9] Neuhaus, G. (1971). On weak convergence of stochastic processes with multi-dimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295.

[10] Puri, M. L., and Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis.* Wiley, New York.

[11] Rényi, A. (1958). On mixing sequence of sets. *Acta. Math. Acad. Sci. Hungar.* **9** 215–228.

[12] Wichura, M. J. (1969). Inequalities with applications to the weak convergence of random processes with multi-dimensional time parameters. *Ann. Math. Statist.* **40** 681–687.

[13] Williams, G. W., and Sen, P. K. (1973). Asymptotically optimal sequential estimation of regular functionals of several distributions based on generalized U-statistics. *J. Multivar. Anal.* **3** in press.

Department of Biostatistics
School of Public Health
University of North Carolina
Chapel Hill, North Carolina 27514