INVARIANCE PRINCIPLES FOR U-STATISTICS AND VON MISES' FUNCTIONALS IN THE NON-I.D. CASE

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1404

June 1982

INVARIANCE PRINCIPLES FOR U-STATISTICS AND VON MISES' FUNCTIONALS IN THE NON-I.D. CASE *

By PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

For independent but non necessarily identically distributed random vectors, weak as well as strong invariance principles for U-statistics and von Mises' differentiable statistical functions are established under appropriate regularity conditions. Martingale characterizations of some decompositions of these statistics play a fundamental role in this context.

1. Introduction.

Let $\{X_i, i \geq 1\}$ be a sequence of independent random vectors (r.v.) with distribution functions (d.f.) $\{F_i, i \geq 1\}$, all belonging to a common class of d.f.'s. Let $g(x_1, \dots, x_m)$ be a Borel-measurable function *(kernel)* of *degree* m (≥ 1), and, without any loss of generality, we may assume that g(.) is a symmetric function of its m arguments (vectors). For $n \geq m$, a statistic of the form

$$(1.1) \qquad U_n = \binom{n}{m}^{-1} \Sigma_{C_{n,m}} g(X_{i_1}, \dots, X_{i_m}) \quad (\text{where } C_{n,m} = \{1 \leq i_1 < \dots < i_m \leq n\})$$

is termed a *U-statistic*. A closely related one, termed the *von Mises' differentiable statistical function*, is the following

$$(1.2) \qquad V_n = n^{-m} \Sigma_{i_1=1}^{n} \cdots \Sigma_{i_m=1}^{n} g(X_{i_1}, \dots, X_{i_m}).$$

When all the F_i are the same, i.e., the X_i are identically distributed (i.d.) r.v., central limit theorems for these statistics have been elaborately studied by von

---

AMS Subject Classification : 60F17, 62E20

Key Words & Phrases: functional central limit theorem, martingale(-difference), Skorokhod-Strassen embedding of Wiener process, tightness.

* Work partially supported by the National Heart, Lung and Blood Institute, Contract NIH-NHLBI-71-2243-L from the National Institutes of Health.

Mises (1947) and Hoeffding (1948); asymptotic normality results in the non-i.d. case were also considered by Hoeffding (1948). Functional central limit theorems for these statistics have been studied by Loynes (1970) and Miller and Sen (1972), among others, and, based on the Skorokhod-Strassen embedding of Wiener processes, Sen (1974) has obtained a strong invariance principle for these statistics ; all these results are confined to the i.d. case where a special decomposition of U_n, due to Hoeffding (1961), and the basic reverse martingale property of U-statistics play a vital role.

In the non-i.d. case, neither the reverse matingale property of U_n nor the martingale structure underlying the Hoeffding (1961) decomposition may hold, and the simple proofs of the functional central limit theorems in Miller and Sen (1972) or Sen (1974) may not workout well. Nothing particularly has been done, so far, on such invariance principles in the non-i.d. case (though the case of stationary dependent sequence has received considerable attention during the past few years). The basic purpose of this study is to incorporate a sequential decomposition of U_n (viz., Sen (1960)) along with the Hoeffding (1961) decomposition which enable us to use some martingale-difference representations for these statistics, and this provide an easy approach to the study of the desired invariance principles, under no extra regularity conditions.

Along with the preliminary notions, the basic results on U-statistics are stated in Section 2 while their derivations are presented in Section 3. Section 4 deals with the parallel results for von Mises' functions. For simplicity of presentations, throughout the paper, the specific case of m=2 has been condidered. The case of m=1 is trivial, while, for $m \geq 3$, a very similar but more lengthy treatment holds.

2. Preliminary notions and basic results for U-statistics.

For every $i,j \geq 1$, let

$$\theta_{ij} = \int \ldots \int g(x_1,x_2)dF_i(x_1)dF_j(x_2). \tag{2.1}$$

Then, by (1.1) and (2.1),

(2.2) $$\theta_{(n)} = EU_n = \binom{n}{2}^{-1}\Sigma_{1\le i<j\le n}\,\theta_{ij}\ ,\ \text{for every } n \ge 2.$$

Our first goal is to study weak invariance principles relating to the (partial) sequence $\{n^{-\frac{1}{2}}k(U_k-\theta_{(k)});2\le k\le n\}$ and the tail-sequence $\{n^{\frac{1}{2}}(U_k - \theta_{(k)}),k \ge n\}$, which would extend the results of Miller and Sen(1972) and Loynes(1970) to the non-i.d. case. Also, we like to extend the results in Sen (1974) to the non-i.d. case.

As in Hoeffding (1948), we let for every $i(\neq j) \ge 1$,

(2.3) $$\psi_{1(i)j}(x_i) = Eg(x_i,X_j) - \theta_{ij}\ ,\ \psi_{2(ij)}(x_i,x_j) = g(x_i,x_j)- \theta_{ij}\ ;$$

(2.4) $$\zeta_{c(i_1\ldots i_c)i_{c+1}\ldots i_2;j_{c+1}\ldots j_2} = E\psi_{c(i_1\ldots i_c)i_{c+1}\ldots i_2}(X_{i_1},\ldots,X_{i_c}).\ \psi_{c(i_1\ldots i_c)j_{c+1}\ldots j_2}(X_{i_1},\ldots,X_{i_c})$$

for all possible combinations of distinct i_r, j_r and c=1,2. For a fixed c(=1,2) and $n(\ge m)$, the average over the possible ζ_c's with $1 \le i_r,j_r \le n$ and distinct i_r,j_r, is denoted by $\zeta_{c,n}$. Then, from Hoeffding (1948), we obtain that

(2.5) $$E\{U_n - \theta_{(n)}\}^2 = \binom{n}{2}^{-1}\Sigma_{c=1}^{2}\binom{c}{2}\binom{n-2}{2-c}\zeta_{c,n}\ ,\ n\ge 2.$$

The asymptotic normality of $\{U_n - \theta_{(n)}\}/\{n^{-1}\zeta_{1,n}\}^{\frac{1}{2}}$, under some extra (mild) regularity conditions, has been established by Hoeffding (1948). In particular, the following conditions relate to the asymptotic normality result. Let

(2.6) $$\bar{\psi}_{1(i),n}(x_i) = (n-1)^{-1}\Sigma_{j=1(j\neq i)}^{n}\,\psi_{1(i)j}(x_i),\ \text{for } i=1,\ldots,n.$$

Assume that there exists a number A, such that for every $n\ge 2$,

(2.7) $$\sup_{i,j\ge 1}\int\ldots\int g^2(x_1,x_2)dF_i(x)dF_j(x) \le A < \infty .$$

Further,

(2.8) $$\sup_{n}\max_{1\le i\le n} E|\bar{\psi}_{1(i),n}(X_i)|^{2+\delta} < \infty ,$$

and

(2.9) $$\lim_{n\to\infty}\left\{\Sigma_{i=1}^{n}E|\bar{\psi}_{1(i),n}(X_i)|^{2+\delta}/\{\Sigma_{i=1}^{n}E\bar{\psi}^{2}_{1(i),n}(X_i)\}^{1+\delta/2}\right\} = 0,$$

where δ $(0 < \delta\le 1)$ is some positive number; Hoeffding(1948) took $\delta = 1$. In particular, (2.9) insures that $n\zeta_{1,n}$ goes to ∞, as $n \to \infty$, though possibly in an arbitrary manner. In the i.d. case, $\bar{\psi}_{1(i),n}(x) \equiv \bar{\psi}_1(x)$ does not depend on (i,n) and $\zeta_{1,n} = \zeta_1$ is also independent of n. Hence, we may ,instead of (2.8)-(2.9), appeal to the classical Lindeberg condition, and this will only require that $\zeta_1 > 0$. Note that by (2.5), (2.6) and (2.9),

$$\begin{aligned} n^2E\{U_n - \theta_{(n)}\}^2 &= \{4n\ \zeta_{1,n}\}\{1 + o(1)\} \\ &= 4\{\Sigma_{i=1}^{n}\ E\bar{\psi}^2_{1(i),n}(X_i)\}\{\ 1 + o(1)\ \}. \end{aligned} \tag{2.10}$$

The ingenuity of the Hoeffding (1948) approach lies in the (quadratic mean) approximation of $n[U_n - \theta_{(n)}]$ by $2\ \Sigma_{i=1}^{n}\ \bar{\psi}_{1(i),n}(X_i)$ and incorporating the central limit theorem for the triangular array $\{Y_{ni} = \bar{\psi}_{1(i),n}(X_i),\ i=1,\dots,n;\ n \geq 1\}$. For our desired invariance principles, we need to establish some maximal inequalities relating to $\{\ k[U_k - \theta_{(k)}]-2\ \Sigma_{i=1}^{k}Y_{ki}\ ; k \geq 2\}$ and some invariance principles relating to the triangular array $\{\ S_{nk} = \Sigma_{i=1}^{k}Y_{ni},\ k \leq n;\ n \geq 2\}$. Towards this, we have the following.

<u>Theorem 2.1</u>. *(I) If* $\lim_{n\to\infty}\{(\log n)/(n\zeta_{1,n})\} = 0$, *then, for every* $\varepsilon > 0$, *under (2.7),*

$$P\{\max_{k \leq n} |k(U_k - \theta_{(k)}) - 2S_{kk}\ | > \varepsilon(n\ \zeta_{1,n})^{\frac{1}{2}}\} \to 0, \text{ as } n \to \infty\ , \tag{2.11}$$

while, if $\lim_{n\to\infty}\{\ n^{-1}\ \Sigma_{k=1}^{n}\ (k\zeta_{1,k})^{-1}\} = 0$, *then, as* $n \to \infty$,

$$P\{\max_{k \leq n} \zeta_{1,k}^{-\frac{1}{2}}\ |k(U_k - \theta_{(k)})\ -2S_{kk}\ | > \varepsilon\ n^{\frac{1}{2}}\ \} \to 0. \tag{2.12}$$

(II) If $\lim_{n\to\infty}\{(n\zeta_{1,n})^{-1}\} = 0$, *then, for every* $\varepsilon > 0$, *as* $n \to \infty$,

$$P\{\ \sup_{k \geq n}\ |(U_k - \theta_{(k)}) - 2k^{-1}S_{kk}| > \varepsilon(n^{-1}\zeta_{1,n})^{\frac{1}{2}}\ \} \to 0, \tag{2.13}$$

while, if $\lim_{n\to\infty}\{n\ \Sigma_{k\geq n}(k^3\zeta_{1,k})^{-1}\} = 0$, *then, as* $n \to \infty$,

$$P\{\ \sup_{k \geq n}\ \zeta_{1,k}^{-\frac{1}{2}}|\ (U_k - \theta_{(k)}) - 2k^{-1}S_{kk}\ | > \varepsilon\ n^{\frac{1}{2}}\ \} \to 0. \tag{2.14}$$

(III) If $\Sigma_{k\geq 2}\ (k^2\zeta_{k,1})^{-1} < \infty$, *then, for every* $\varepsilon > 0$, *as* $n \to \infty$,

$$P\{\ \sup_{k \geq n}\ |\{k(U_k - \theta_{(k)})\ -2S_{kk}\}|/(k\zeta_{1,k})^{\frac{1}{2}} > \varepsilon\ \} \to 0. \tag{2.15}$$

The proof of the theorem is considered in Section 3. The maximal inequalities in (2.11) - (2.15) along with proper construction of stochastic processes lead us to the desired results. In the i.d. case, $\{S_{kk};\ k \geq 1\}$ (or $\{k^{-1}S_{kk};\ k \geq 1\}$)

form a martingale (or reverse martingale) sequence, and hence, the forward (or backward) invariance principle for U-statistics, considered by Miller and Sen (1972) and Loynes (1970), follows directly by an appeal to the functional central limit theorem for martingales (or reverse martingales). In the non-i.d. case, neither the forward nor the reverse martingale property holds, and hence, a more elaborate treatment is necessary. Note that in order to establish weak invariance principles, one needs to study the convergence of finite dimensional distributions (f.d.d.) and the tightness of appropriate stochastic processes constructed from these statistics. By virtue of Theorem 2.1, it suffices to construct such processes from the S_{kk} or $k^{-1}S_{kk}$. In this respect, if we consider arbitrary $c(\geq 1)$ and $0<t_1 <\ldots<t_c(\leq 1)$, then, for any non-null $\underset{\sim}{\lambda} =(\lambda_1,\ldots,\lambda_c)$, on letting $n_j = [nt_j]$ (or $[n/t_j]$), $j=1,\ldots,c$, $Z_n = \Sigma_{j=1}^{c} \lambda_j S_{n_j n_j}$ (or $\Sigma_{j=1}^{c}\lambda_j n_j^{-1} S_{n_j n_j}$), we may, by virtue of the definitions of the S_{kk} and Y_{kj} , express Z_n as $(Z_{n1}+\ldots+Z_{nn})$ (or $Z_{n_1 1}+\ldots + Z_{n_1 n_1}$) where the Z_{ni} form a triangular array of row-wise independent r.v.'s. Thus, the Liapounoff-type condition in (2.8)-(2.9) may again be called on to verify the asymptotic normality of Z_n . Thus, the convergence of f.d.d.'s may be established under conditions similar to (2.8)-(2.9). However, these finite dimensional distributions, though asymptotically multinormal, may not conform to that of a Wiener process, particularly when $\zeta_{1,n}$ does not converge to a positive constant. Thus, in the non-i.d. case, for the invariance principles to be studied, the question of having a Wiener process in the picture remains of some good interest. Secondly, even if, such a Wiener process does not come to the picture, the question of weak convergence to some appropriate Gaussian function needs to be addressed. In both the cases, the basic task is to establish the tightness property of the stochastic processes under consideration, and, then to examine the covariance structure to decide whether the Gaussian function is of Wiener structure or not. This will be studied here.

Let us introduce the following stochastic processes $Z_n^{(j)} = \{Z_n^{(j)}(t);0\leq t\leq 1\}$,

$j=1,\ldots,4$, all defined on the space $D[0,1]$, where

$$(2.16)\qquad Z_n^{(1)}(t) = \begin{cases} 0, & 0\le t<2/n, \\ [nt](U_{[nt]} - \theta_{([nt])})/2\{n\zeta_{1,n}\}^{\frac{1}{2}}, & t\ge 2/n\,; \end{cases}$$

$$(2.17)\qquad Z_n^{(2)}(t) = \{\zeta_{1,n}/\zeta_{1,[nt]}\}^{\frac{1}{2}}\, Z_n^{(1)}(t),\ 2/n\le t\le 1, \text{ and } 0, \text{ otherwise};$$

$$(2.18)\qquad Z_n^{(3)}(t) = n^{\frac{1}{2}}\{U_{n(t)} - \theta_{(n(t))}\}/2\zeta_{1,n}^{\frac{1}{2}}\ ;\ n(t) = \min\{\, k: n/k\le t\},\ 0\le t\le 1;$$

$$(2.19)\qquad Z_n^{(4)}(t) = \{\zeta_{1,n}/\zeta_{1,n(t)}\}^{\frac{1}{2}}\, Z_n^{(3)}(t)\ ,\ 0\le t\le 1.$$

Our goal is to study weak conveegence of these processes to appropriate Gaussian functions. Corresponding to the d.f. $\{F_i, i\ge 1\}$ and the kernel $g(x,y)$, we define

$$(2.20)\qquad \bar{F}_{(n)} = n^{-1}\Sigma_{i=1}^{n} F_i\ ,\ \text{for } n\ge 1,$$

$$(2.21)\qquad \bar{g}_{1(n)}(x) = \int g(x,y)\,d\bar{F}_{(n)}(y)\ ,$$

$$(2.22)\qquad \zeta_1(\bar{F}_{(n)}) = \int \bar{g}^2_{1(n)}(x)\,d\bar{F}_{(n)}(x) - \left(\iint g(x,y)\,d\bar{F}_{(n)}(x)\,d\bar{F}_{(n)}(y)\right)^2$$

and

$$(2.23)\qquad \Delta_n^2 = n^{-1}\Sigma_{i=1}^{n}\left(\int \bar{g}_{1(n)}(x)\,d[F_i(x)-\bar{F}_{(n)}(x)]\right)^2 .$$

Then, following Hoeffding (1948) and Sen(1969), we have

$$(2.24)\qquad \zeta_{1,n} = \zeta_1(\bar{F}_{(n)}) - \Delta_n^2 + O(n^{-1});\quad \Delta_n^2 \le \zeta_1(\bar{F}_{(n)}).$$

As such, if $\bar{F}_{(n)}$ converges to some $\bar{F}$, $\zeta_1(F)$ is a continuous functional of F in some neighbourhood of $\bar{F}$ and Δ_n^2 converges to some $\Delta^2 < \zeta_1(\bar{F})$, then, $\zeta_{1,n} \to \zeta^* = \zeta_1(\bar{F}) - \Delta^2 > 0$, as $n\to\infty$. There may be other situations where $\zeta_{1,n}$ may converge to some positive limit ζ^* without requiring that $\bar{F}_{(n)}$ converges to a limit $\bar{F}$.

Theorem 2.2. *If (2.7)-(2.9) hold and $\zeta_{1,n} \to \zeta^* > 0$, then, for each $j(=1,\ldots,4)$, $Z_n^{(j)}$ converges in law to $Z = \{Z(t); 0\le t\le 1\}$, a standard Wiener process on* $[0,1]$. *The result holds if in (2.16)-(2.19), $\theta_{(k)}$ is replaced by $\theta(\bar{F}_{(k)}) = \iint g(x,y)\,d\bar{F}_{(k)}(x).d\bar{F}_{(k)}(y)$, for all possible k.*

In the other cases where $\zeta_{1,n}$ may not converge to a positive limit, though the f.d.d.'s of the $Z_n^{(j)}$ may converge to some multinormal distributions, the covariance function may not conform to that of a Brownian motion. Nevertheless, we may like to study the tightness property of these processes. In the sequel, it will be assumed that $n\zeta_{1,n}$ is ↑ in n. Also, let $h_n(k/n) = \zeta_{1,n}/\zeta_{1,k}$, $k \leq n$. Note that $h_n(k/n) \geq k/n$, $\forall\, k \leq n$. We may also need the following :

$$\max\{ h_n(k/n) : k \leq n\} = O(1), \tag{2.25}$$

$$|h_n(q/n) - h_n(k/n)| \leq K|q/n - k/n|^{\alpha} , \tag{2.26}$$

for some $\alpha > 1/r$, for every n,k and q, such that $q/n > k/n \geq \delta > 0$, where K may depend on δ but not on n.

Theorem 2.3. *If* $\underline{\lim}\ \zeta_{1,n} > 0$, *then, under (2.7)-(2.9), the tightness of* $Z_n^{(1)}$ *and* $Z_n^{(3)}$ *are in order, while, if in addition, (2.25) and (2.26) hold, then* $Z_n^{(2)}$ *and* $Z_n^{(4)}$ *are also tight.*

When $\underline{\lim}\ \zeta_{1,n}$ may not be positive, we may still be able to establish the tightness property of the $Z_n^{(j)}$, provided we impose more stringent conditions on the kernels. For every $n \geq q > k \geq 2$, we define

$$Z_{n;q,k} = n^{-\frac{1}{2}}\{\Sigma_{i=1}^{k}[\bar{\Psi}_{1(i),q}(X_i) - \bar{\Psi}_{1(i),k}(X_i)] + \Sigma_{i=k+1}^{q}\ \bar{\Psi}_{1(i),q}(X_i)\} . \tag{2.27}$$

Theorem 2.4. *If for some* $r > 2$, *for every n,q and k, we have*

$$E|Z_{n;q,k}|^r / \{EZ^2_{n;q,k}\}^{r/2} \leq K < \infty , \tag{2.28}$$

then, the tightness property of the $Z_n^{(j)}$ *holds under the hypothesis of Theorem 2.3, without the condition that* $\underline{\lim}\ \zeta_{1,n} > 0$.

Finally, we consider the strong invariance principle for U-statistics. In Sen (1974), the basic tool of martingales and reverse martingales were incorporated in the derivation of the main result. In the non-i.d. case, these martingale or reverse martingale property may not hold and hence that method may not be applicable. By virtue of (2.15), the desired strong invariance principle would follow if we can establish the Skorokhod-Strassen embedding for the sequence $\{S_{kk}; k \geq n \geq 2\}$. However, for a double sequence of independent (row-wise) r.v., we may not have such a powerful result to utilize in the current context. Keeping in mind (2.6) and (2.21), we conceive of a sequence $\{\psi_k^*(X_k) ; k \geq 1\}$ of random variables and assume that there exists a positive $c > 1$, such that

$$\xi_n = \Sigma_{i=1}^{n} E[\bar{\Psi}_{1(i),n}(X_i) - \psi_i^*(X_i)]^2 = O((\log n)^{-c}) \tag{2.29}$$

and that $s_n^{*2} = \Sigma_{k=1}^n E[\psi_k^*(X_k)]^2 \to \infty$, as $n \to \infty$. More specifically, we assume that

$$(2.30) \qquad \lim_{n\to\infty} n^{-1} s_n^{*2} = \zeta^* > 0 .$$

Typically (2.29) and (2.30) hold when $\overline{F}_{(n)} \to \overline{F}$ as $n \to \infty$, so that $\psi_k^*(x) = \int g(x,y)d\overline{F}(y) - \int\int g(x,y)dF_k(x)d\overline{F}(y)$, $k \geq 1$, where the $\psi_k^*(X_k)$ need not be i.i.d. (as in the multisample situation where the F_i belong to a finite class $\{G_1,\ldots,G_m\}$ of d.f.'s.). Let $Z = \{Z(t): 0 \leq t < \infty\}$ be a random process on $[0,\infty)$, where $Z(t) = Z(k)$ for $k \leq t < k+1$ and

$$(2.31) \qquad \begin{aligned} Z(k) &= 0 , \quad k < 2, \\ &= k(U_k - \theta(\overline{F}_{(k)}))/2(\zeta^*)^{\frac{1}{2}} , \quad k \geq 2. \end{aligned}$$

Finally, let $W = \{W(t): 0 \leq t < \infty\}$ be a standard Wiener process on $[0,\infty)$.

Theorem 2.5. *Under (2.29) and the hypothesis of Theorem 2.2,*

$$(2.32) \qquad Z(t) = W(t) + o(t^{\frac{1}{2}}) \text{ almost surely, as } t \to \infty .$$

The result can be extended to the case where s_n^* may not go to ∞ at the rate of $n^{\frac{1}{2}}$, but at a slower rate. In that case, we need stronger order of convergence in (2.29). We will not elaborate it further.

3. Proofs of the theorems.

We start with the following (sequential) decomposition of U_n [c.f. Sen(1960)]. For m=2, by (1.1), (2.1), (2.2) and (2.3),

$$(3.1) \qquad \binom{n}{2}[U_n - \theta_{(n)}] = \Sigma_{i=2}^n \{\Sigma_{j=1}^{i-1} [g(X_i,X_j) - \theta_{ij}]\} = \Sigma_{i=2}^n \Sigma_{j=1}^{i-1} \psi_{2(ij)}(X_i,X_j) .$$
$$= \Sigma_{i=2}^n U_i^* , \text{ say.}$$

Further, we write, for every $i \geq 2$,

$$(3.2) \qquad U_i^* = \Sigma_{j=1}^{i-1} \psi_{1(i)j}(X_i) + \Sigma_{j=1}^{i-1} \psi_{1(j)i}(X_j) + \Sigma_{j=1}^{i-1}\{\psi_{2(ij)}(X_i,X_j) - \psi_{1(i)j}(X_i) - \psi_{1(j)i}(X_j)\}$$
$$= U_{i1}^* + U_{i2}^* + U_i^{**} , \text{ say.}$$

Let $\mathcal{B}_k$ be the sigma-field generated by $(X_1,\ldots,X_k)$, for $k \geq 1$. Then, note that

$$(3.3) \qquad E(U_i^{**} \mid \mathcal{B}_{i-1}) = 0 \text{ and } E(U_{i1}^* \mid \mathcal{B}_{i-1}) = 0 , \ \forall\, i \geq 2 ;$$

however, this martingale (difference) property does not hold for the U_{i2}^*. Also, note that for every $n \geq 2$, (3.4) $(n-1)^{-1}\Sigma_{i=2}^{n}\{U_{i1}^{*}+U_{i2}^{*}\} = \Sigma_{i=1}^{n} \bar{\Psi}_{1(i),n}(X_i) = \Sigma_{k=1}^{n} Y_{nk} = S_{nn}$,

where the Y_{nk} and S_{nn} are defined after (2.10). Hence,for every $n \geq 2$,

(3.5) $n[U_n - \theta_{(n)}] - 2S_{nn} = 2(n-1)^{-1} \Sigma_{i=2}^{n} U_i^{**} = R_n$, say,

and, conventionally, we let $R_1 = 0$. Note that by (3.3) and (3.5),

(3.6) $E[|R_n| \mid \mathcal{B}_{n-1}] \geq |E(R_n \mid \mathcal{B}_{n-1})| = \frac{n-2}{n-1} |R_{n-1}|$, $\forall\, n \geq 2$.

Therefore, by the Birnbaum-Marshall (1961) inequality, for every sequence $\{a_k\}$ of positive numbers, $t > 0$ and $n_1, n_2 : n_2 > n_1 \geq 2$,

(3.7) $P\{\max_{n_1 \leq k \leq n_2} a_k|R_k| \geq t\} \leq t^{-2} \Sigma_{k=n_1}^{n_2} (b_k^2 - k^{-2}(k-1)^2 b_{k+1}^2) ER_k^2$,

where

(3.8) $b_k = \max\{ a_k, k^{-1}(k-1)a_{k+1},\ldots,(n_2-1)^{-1}(k-1)a_{n_2}\}$, for $k=n_1,\ldots,n_2$.

Note that by (3.20, (3.3) and (2.7),for every $n \geq 2$,

(3.9) $ER_n^2 = 4(n-1)^{-2}\Sigma_{k=2}^{n}E\{(U_k^{**})^2\} \leq 8A(n-1)^{-2}\Sigma_{i=2}^{n}(i-1) = 4An/(n-1) \leq 8A$.

Thus, if we let $n_1=2$, $n_2=n$, $a_k=1, 2\leq k\leq n$ (so that $b_k=1$, $2\leq k\leq n$) and $t= \varepsilon(n\zeta_{1,n})^{\frac{1}{2}}$, then, (2.11) follows from (3.7) and (3.9) when $(n\zeta_{1,n})^{-1}\log n \to 0$ as $n \to \infty$. For (2.12), we take $t= \varepsilon n^{\frac{1}{2}}$ and $a_k = \zeta_{1,k}^{-\frac{1}{2}}$ and the proof follows on parallel lines. For the proof of (2.13), we take $n_1= n$, $n_2 \to \infty$, $t = \varepsilon(n^{-1}\zeta_{1,n})^{\frac{1}{2}}$ and $a_k = k^{-1}$, for $k \geq n$. Note that $(n\zeta_{1,n}^{-1}) \Sigma_{k\geq n} k^{-3}ER_k^2 = O(n\zeta_{1,n}^{-1} \Sigma_{k\geq n} k^{-3}) = O((n\zeta_{1,n})^{-1})$, and hence, (3.7) insures(2.13) under the assumed condition. The proofs of (2.14) and (2.15) are very similar, and hence, omitted. This completes the proof of Theirem 2.1.

In the proof of Theorem 2.2, we may note first that by (2.1), (2.2), (2.7) and (2.20),

(3.10) $|\theta_{(k)} - \theta(\bar{F}_{(k)})| \leq (k-1)^{-1}A^{\frac{1}{2}}$, for every $k \geq 2$,

so that $\max\{ k|\theta_{(k)}- \theta(\bar{F}_{(k)})|/\sqrt{n} : 2\leq k\leq n\}$ and $\max\{ n^{\frac{1}{2}} |\theta_{(k)}- \theta(\bar{F}_{(k)})|:k\geq n\}$ both converge to 0 as $n \to \infty$. Hence, we may use $\theta_{(k)}$ and $\theta(\bar{F}_{(k)})$ interchangeably . Since the convergence of the f.d.d.'s follow along the line of Hoeffding (1948, Theorem 8.1), we shall only prove the tightness property of the $Z_n^{(j)}$, $j=1,\ldots,4$. For this purpose, we define for each $k \geq 2$,

$$(3.11)\qquad W_{k1} = \Sigma_{i=2}^{k}(i-1)^{-1}\Sigma_{j=1}^{i}\ \psi_{1(i)j}(X_i) = \Sigma_{i=2}^{k}\ \overline{\Psi}_{1(i-1)}(X_i)\ ,\ \text{say}\ ,$$

$$(3.12)\qquad W_{k2} = \Sigma_{i=2}^{k}(i-1)^{-1}\Sigma_{j=1}^{i}\ \psi_{1(j)i}(X_j) = \Sigma_{i=2}^{k}\ \overline{\Psi}_{1(i-1)i}\ ,\ \text{say.}$$

Then, by (3.1), (3.2), (3.11) and (3.12), for every $q \geq k \geq 2$, we have

$$(q-1)^{-1}\Sigma_{i=2}^{q}\ U_{ij}^{*} - (k-1)^{-1}\Sigma_{i=2}^{k}\ U_{ij}^{*}$$

$$(3.13)\qquad = (W_{qj} - W_{kj}) - (q-1)^{-1}\Sigma_{i=2}^{q}W_{ij} + (k-1)^{-1}\Sigma_{i=2}^{k}W_{ij}\ ,\ \text{for } j=1,2.$$

By (3.4) and (3.13), for every $q \geq k \geq 2$ and n,

$$(3.14)\qquad n^{-\frac{1}{2}}|(S_{qq} - S_{kk})| \leq \Sigma_{j=1}^{2}n^{-\frac{1}{2}}\{|W_{qj} - W_{kj}| + |\tfrac{1}{q-1}\Sigma_{i=2}^{q}\ W_{ij} - \tfrac{1}{k-1}\Sigma_{i=2}^{k}W_{ij}|\}.$$

Now, under (2.8), for some $\delta > 0$,

$$(3.15)\ n^{-(1+\delta/2)}E|\ W_{q1} - W_{k1}|^{2+\delta} = n^{-(1+\delta/2)}E|\Sigma_{k+1}^{q}\ \overline{\Psi}_{1(i-1)}(X_i)\ |^{2+\delta}$$

$$\leq\ c_{2+\delta}\ (q-k)^{1+\delta/2}\ n^{-(1+\delta/2)}\ \{(q-k)^{-1}\ \Sigma_{k+1}^{q}E\ |\overline{\Psi}_{1(i-1)}(X_i)|^{2+\delta}\ \}$$

$$\leq\ C\ [(q-k)/n]^{1+\delta/2}\ ,\ \text{where } c_{2+\delta} \text{ and } C \text{ are finite positive constants.}$$

Similarly,

$$(3.16)\qquad \begin{aligned} &n^{-(1+\delta/2)}E|\ W_{q2} - W_{k2}\ |^{2+\delta} = n^{-(1+\delta/2)}\ E|\Sigma_{i=k+1}^{q}\ \overline{\Psi}_{1(i-1)i}\ |^{2+\delta} \\ &\leq n^{-(1+\delta/2)}(\Sigma_{i=k+1}^{q}\ i^{1+\delta/2}E|\overline{\Psi}_{1(i-1)i}|^{2+\delta}\ )(\Sigma_{i=k+1}^{q}i^{-(2+\delta)/2(1+\delta)})^{1+\delta}\end{aligned}$$

where, under (2.8), $E|\overline{\Psi}_{1(i-1)i}|^{2+\delta} = O(i^{-1-\delta/2})$, so that by some standard steps, the right hand side of (3.16) can be bounded by

$$(3.17)\qquad C[(q-k)/n]^{1+\delta/2}\ ,\ \text{for every } k \leq q \leq n;\ C < \infty\ .$$

Also, note that

$$(3.18)\qquad (q-1)^{-1}\Sigma_{i=2}^{q}\ W_{ij} - (k-1)^{-1}\Sigma_{i=2}^{k}\ W_{ij}$$

$$= [(q-k)/(q-1)]\{(q-k)^{-1}\Sigma_{i=k+1}^{q}W_{ij} - (k-1)^{-1}\Sigma_{i=2}^{k}\ W_{ij}\},\ \text{for } j=1,2.$$

Thus, under (2.8), a moment-bound similar to (3.15) and (3.17) applies to (3.18) as well. Using (3.4) and (3.13) through (3.18), we conclude that under (2.8), for every $n \geq q \geq k \geq 2$, we have

$$(3.19)\qquad E|n^{-\frac{1}{2}}\ (S_{qq} - S_{kk}\ )|^{2+\delta} \leq K[(q-k)/n]^{1+\delta/2}\ ,\ \delta > 0,$$

where K is a positive constant, independent of n. By virtue of Theorem 12.3 of Billingsley (1968) and our Theorem 2.1, (3.19) insures the tightness of $Z_n^{(1)}$ when

$\zeta_{1,n} \to \zeta^* > 0$. The tightness of $Z_n^{(2)}$ follows from (2.17) and the tightness of $Z_n^{(1)}$ where we use the fact that for every $t \in (0,1]$, $\zeta_{1,n}/\zeta_{1,[nt]} \to 1$, as $n \to \infty$. For the process $Z_n^{(3)}$ in (2.18), we again use Theorem 2.1 and note that

$$(3.20) \quad n^{\frac{1}{2}}(k^{-1}S_{kk} - q^{-1}S_{qq}) = n^{\frac{1}{2}}[(k^{-1}-q^{-1})S_{kk} - q^{-1}(S_{qq}-S_{kk})].$$

As such, proceeding as in (3.14) through (3.19), we obtain that for every $n \leq k < q < \infty$, under (2.8), for some $\delta > 0$,

$$(3.21) \quad E|n^{\frac{1}{2}}(k^{-1}S_{kk} - q^{-1}S_{qq})|^{2+\delta} \leq K[(n/k-n/q)]^{1+\delta/2},$$

where $K(< \infty)$ does not depend on n. Thus, Theorem 12.3 of Billingsley (1968) may again be recalled to verify the tightness of $Z_n^{(3)}$. The tightness of $Z_n^{(4)}$ follows from (2.19), the tightness of $Z_n^{(3)}$ and the fact that $\zeta_{1,k} \to \zeta^* > 0$, as $k \to \infty$. This completes the proof of Theorem 2.2.

To prove Theorem 2.3, we note that Theorem 2.1, (3.19) and the assumption that $\underline{\lim}\, \zeta_{1,n} > 0$ insure the tightness of $Z_n^{(1)}$, while (3.21) and the other conditions insure the tightness of $Z_n^{(3)}$. Further, (2.25)-(2.26) and the tightness of $Z_n^{(1)}$ (or $Z_n^{(3)}$) insure the tightness of $Z_n^{(2)}$ (or $Z_n^{(4)}$). Hence, the details are omitted.

For Theorem 2.4, we may note that under (2.28), for some $r > 2$,

$$(3.22) \quad E|Z_n^{(1)}(t) - Z_n^{(1)}(s)|^r \leq K|t-s|^{r/2},$$

for every $0 \leq s \leq t \leq 1$, where $K\ (< \infty)$ does not depend on n. A similar moment-bound holds for the $Z_n^{(3)}$ also. Hence, Theorem 12.3 of Billingsley (1968) and (3.22) insure the tightness of $Z_n^{(1)}$ and $Z_n^{(3)}$. The case of $Z_n^{(2)}$ and $Z_n^{(4)}$ follows by using (2.25)-(2.26) along with the tightness of $Z_n^{(1)}$ and $Z_n^{(3)}$.

Finally, we consider the proof of Theorem 2.5. Note that by (2.15) and Theorem 2.2, for every positive ε, as $n \to \infty$,

$$(3.23) \quad P\{\sup_{k \geq n} (k\zeta_{1,k})^{-\frac{1}{2}}|k(U_k - \theta(\overline{F}_{(k)})) - 2S_{kk}| > \varepsilon\} \to 0.$$

Let us define now

$$(3.24) \quad S_k^* = \Sigma_{i=1}^k \psi_i^*(X_i), \quad k \geq 1.$$

Note that the S_k^* involve independent (but, possibly non-i.d.) summands, and, by

(2.29), (3.24) and the definition of the S_{kk} , we have

$$(3.25)\quad \begin{aligned} &P\{\sup_{k \geq n} k^{-\frac{1}{2}}\zeta_{1,k}^{-\frac{1}{2}} |S_{kk} - S_k^*| > \varepsilon \} \\ &\leq \Sigma_{k\geq n} P\{(k\zeta_{1,k})^{-\frac{1}{2}}|S_{kk} - S_k^*| > \varepsilon \} \\ &\leq \Sigma_{k\geq n} (k\zeta_{1,k})^{-1} E(S_{kk} - S_k^*)^2 \\ &= \Sigma_{k\geq n} (k\zeta_{1,k})^{-1} \xi_k = \Sigma_{k\geq n} \zeta_{1,k}^{-1} [O(k^{-1}(\log k)^c)] , \end{aligned}$$

where $c > 1$ and the $\zeta_{1,k}$ converge to a positive limit ζ^* . Hence, the right hand side of (3.25) converges to 0 as $n \to \infty$. Thus, from (3.23) and (3.25), we obtain that as $n \to \infty$,

$$(3.26)\quad (n\zeta_{1,n})^{-\frac{1}{2}} | n[U_n - \theta(\overline{F}_{(n)})] -2S_n^* | \to 0 \text{ , almost surely.}$$

On the other hand, for the sequence $\{ S_k^* ; k \geq 1\}$, we directly appeal to the Skorokhod-Strassen embedding of Wiener process [c.f. Strassen (1967)] and conclude that under (2.8),

$$(3.27)\quad S_n^*/\sqrt{\zeta^*} = W(n) + o(n^{\frac{1}{2}}) \text{ almost surely , as } n \to \infty .$$

Thus, (2.31), (3.26) and (3.27) imply (2.32). Q.E.D.

4. Invariance principles for the von Mises' functional. Recall that by (1.1) and (1.2);

$$(4.1)\quad n(V_n - U_n) = n^{-1}\Sigma_{i=1}^n g(X_i,X_i) - U_n , \forall n \geq 2.$$

In addition to (2.7), we now assume that

$$(4.2)\quad \sup_{i\geq 1} \int g^2(x,x)dF_i(x) < B < \infty .$$

Let us now write

$$(4.3)\quad \theta_i^o = Eg(X_i,X_i) \text{ and } \zeta_i^o = Var\{g(X_i,X_i)\} , i \geq 1 ;$$

$$(4.4)\quad \overline{\theta}_n^o = n^{-1}\Sigma_{i=1}^n\theta_i^o \text{ and } \theta_{(n)}^* = \overline{\theta}_n^o - \theta_{(n)} , n \geq 2.$$

Then, under (4.2), by the Kolmogorov law of large numbers, as $n \to \infty$,

$$(4.5)\quad n^{-1}\Sigma_{i=1}^n g(X_i,X_i) - \overline{\theta}_n^o \to 0 \text{ , almost surely .}$$

Further, as in (3.21), for some $\delta >0$,

$$(4.6)\quad E|k^{-1}S_{kk}|^{2+\delta} = O(k^{-1-\delta/2}), \text{ for every } k \geq 2.$$

Hence, by (2.15), (4.6) and the Borel-Cantelli Lemma, under (2.8),

$$(4.7)\quad U_n - \theta_{(n)} \to 0 \text{ almost surely, as } n \to \infty .$$

Therefore, by (4.1), (4.5) and (4.7), we obtain that as $n \to \infty$,

$$(4.8) \quad n(U_n - V_n) - \overline{\theta}_n^o + \theta_{(n)} \to 0, \text{ almost surely.}$$

Consequently, on letting $\theta^*_{(n)} = n^{-1}[(n-1)\theta_{(n)} + \overline{\theta}_n^o]$ and noting that $n(V_n - \theta^*_{(n)}) = n(U_n - \theta_{(n)}) + \{n(V_n - U_n) - \overline{\theta}_n^o + \theta_{(n)}\}$, we conclude that as $n \to \infty$,

$$(4.9) \quad n(V_n - \theta^*_{(n)}) - n(U_n - \theta_{(n)}) \to 0, \text{ almost surely.}$$

Thus, if in (2.18),(2.19) and (2.31), we replace the $k(U_k - \theta_{(k)})$ by the $k(V_k - \theta^*_{(k)})$ and denote the resulting processes by $Z_n^{o(3)}$, $Z_n^{o(4)}$ and Z^o, respectively, then the invariance principles in Theorems 2.2, 2.3, 2.4 and 2.5 hold for these processes as well. Similarly, under (4.2),

$$(4.10) \quad \max\{ n^{-\frac{1}{2}} |k^{-1}\Sigma_{i=1}^k g(X_i,X_i) - \theta^*_{(k)}| : k \le n \} \xrightarrow{p} 0, \text{ as } n \to \infty,$$

while, by the Theorems in Section 2,

$$(4.11) \quad \max\{ n^{-\frac{1}{2}} |U_k - \theta_{(k)}| : k \le n \} \xrightarrow{p} 0, \text{ as } n \to \infty.$$

Consequently, under (4.2) and (2.7)-(2.9), as $n \to \infty$,

$$(4.12) \quad \max\{ n^{-\frac{1}{2}} | k(V_k - \theta^*_{(k)}) - k(U_k - \theta_{(k)})| : k \le n \} \xrightarrow{p} 0.$$

Thus, if in (2.16)-(2.17), we replace the $U_k - \theta_{(k)}$ by $V_k - \theta^*_{(k)}$, and denote the resulting processes by $Z_n^{o(1)}$ and $Z_n^{o(2)}$, respectively, then, under (4.2) and the hypotheses of Theorems 2.2, 2.3 and 2.4, the invariance principle holds for $Z_n^{o(1)}$ and $Z_n^{o(2)}$ also. This leads us to the following.

Theorem 4.1. *Under the additional assumption (4.2), the invariance principles in Section 2 hold for the von Mises' functionals as well.*

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* New York: Wiley.

BIRNBAUM, Z.W. and MARSHALL, A.W. (1961). Some multivariate Chebyshev inequalities with extensions to continuous parameter processes. *Ann.Math.Statist. 32*, 687-703.

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

HOEFFDING, W. (1961). The strong law of large numbers for U-statistics. *Inst. Statist. Univ. North Carolina, Mimeo. Report. No. 302.*

LOYNES, R.M. (1970). An invariance principle for reverse martingales. *Proc.Amer. Math. Soc. 25*, 56-64.

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.

SEN, P.K. (1960). On some convergence properties of U-statistics. *Calcutta Statist.Assoc.Bull. 10*, 1-19.

SEN, P.K. (1969). On a robustness property of a class of nonparametric tests based on U-statistics. *Calcutta Statist.Assoc.Bull.18*, 51-60.

SEN, P.K. (1974). Almost sure behaviour of U-statistics and von Mises' differentiable statistical functions. *Ann.Statist. 2*, 387-395.

SEN, P.K. (1981). *Sequential Nonparametrics: Invariance Principles and Statistical Inference*. New York : Wiley.

STRASSEN, V. (1967). Almost sure behaviour of sums of independent random variables and martingales. *Proc. Fifth Berkeley Symp. Math.Statist.Probability. 2*, 315-343 (ed: L.LeCam et. al, Univ. Calif. Press).

VON MISES, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann.Math.Statist. 18*, 309-348.