*The Annals of Mathematical Statistics*
1972, Vol. 43, No. 1, 31-41

# WEAK CONVERGENCE OF *U*-STATISTICS AND VON MISES' DIFFERENTIABLE STATISTICAL FUNCTIONS

BY R. G. MILLER, JR.[1] AND PRANAB KUMAR SEN[2]

*Stanford University and University of North Carolina*

For partial cumulative sums of independent and identically distributed random variables (i.i.d.r.v.) with a finite (positive) variance, weak convergence to Brownian motion processes has been established by Donsker (1951, 1952). The result is extended here to differentiable statistical functions of von Mises (1947) and *U*-statistics of Hoeffding (1948). Along with the extension to generalized *U*-statistics, a few applications are briefly sketched.

**1. Introduction.** Let $\omega = \{X_1, X_2, \cdots\}$ be a sequence of i.i.d.r.v. with each $X_i$ having a distribution function (df) $F(x)$, $x \in R^p$, the $p(\geqq 1)$-dimensional Euclidean space. Let $g(X_1, \cdots, X_m)$, symmetric in its arguments, be a Borel-measurable *kernel* of *degree* $m(\geqq 1)$, and consider the *regular functional*

$$\theta(F) = \int_{R^{pm}} \cdots \int g(x_1, \cdots, x_m)\, dF(x_1) \cdots dF(x_m) \tag{1.1}$$

defined on $\mathscr{F} = \{F: |\theta(F)| < \infty\}$. For a sample $(X_1, \cdots, X_n)$, consider the *empirical df*

$$F_n(x, \omega) = n^{-1} \textstyle\sum_{i=1}^n c(x - X_i)\,, \qquad x \in R^p\,, \tag{1.2}$$

where $c(u)$ is 1 if all its $p$ arguments are nonnegative, and otherwise $c(u)$ is equal to 0. Then, the corresponding functional of $F_n$, viz.,

$$\theta(F_n, \omega) = \int_{R^{pm}} \cdots \int g(x_1, \cdots, x_m)\, dF_n(x_1, \omega) \cdots dF_n(x_m, \omega) \tag{1.3}$$

is termed a *differentiable statistical function* by von Mises (1947). Though $F_n$ unbiasedly estimates $F$, $\theta(F_n, \omega)$ is not necessarily an unbiased estimator of $\theta(F)$. The unbiased estimator (*U-statistic*) as considered by Hoeffding (1948) is defined by

$$U_n(\omega) = \tbinom{n}{m}^{-1} \textstyle\sum_{C_{n,m}} g(X_{i_1}, \cdots, X_{i_m})\,; \qquad C_{n,m} = \{1 \leqq i_1 < \cdots < i_m \leqq n\}\,, \tag{1.4}$$

where $n$ is $\geqq m$.

When $\theta(F)$ is stationary of order zero and $g \in L^2$, it has been shown by Hoeffding (1948) that (i) $n^{\frac{1}{2}}[U_n(\omega) - \theta(F)]$ is asymptotically normally distributed with zero mean and a finite (positive) variance, (ii) under an additional condition

Received May 6, 1971.

[1] Research supported in part by Public Health Service, Grant USPHS-5T1-25-13.

[2] Research supported in part by the Army Research Office, Durham, Grant DA-ARO-D-31-124-70-G6.

on $g$, $n^{\frac{1}{2}}[U_n(\omega) - \theta(F_n, \omega)] \to_P 0$ as $n \to \infty$, and hence, (iii) $n^{\frac{1}{2}}[\theta(F_n, \omega) - \theta(F)]$ is also aymptotically normally distributed. The earlier proof of (iii) by von Mises (1947) is more elaborate and complicated.

In particular, when $m = 1$, $\theta(F_n, \omega) = U_n(\omega) = n^{-1} \sum_{i=1}^n g(X_i) = \bar{g}_n(\omega)$, say. If we assume that $0 < \sigma^2 = E\{g(X_1) - \theta(F)\}^2 < \infty$, and let, for every $t \in I = [0, 1]$,

$$(1.5) \qquad Y_n(t, \omega) = \{[nt]\bar{g}_{[nt]}(\omega) + (nt - [nt])g(X_{[nt]+1}) - nt\theta(F)\}/\sigma n^{\frac{1}{2}} ,$$

where $[s]$ denotes the largest integer contained in $s\ (\geqq 0)$, then by the Donsker theorem [cf. Billingsley (1968, page 68)], as $n \to \infty$,

$$(1.6) \qquad Y_n(\omega) = [Y_n(t, \omega),\ t \in I] \to_{\mathscr{D}} W = [W_t,\ t \in I] ,$$

where $W_t$, $t \geqq 0$, is a standard Brownian motion, so that

$$(1.7) \qquad EW_t = 0 \quad \text{and} \quad E[W_s W_t] = s , \qquad 0 \leqq s \leqq t < \infty ,$$

and $\to_{\mathscr{D}}$ stands for convergence in distribution.

For $m \geqq 1$, proceeding as in Hoeffding (1948, pages 298–299) and Sproule (1969), it can be easily shown that the joint moments of $k[U_k(\omega) - \theta(F)]/n^{\frac{1}{2}}$ and $q[U_q(\omega) - \theta(F)]/n^{\frac{1}{2}}$, where $k = [ns]$, $q = [nt]$, $0 < s \leqq t \leqq 1$, have asymptotically (as $n \to \infty$) the structure (1.7) (apart from a multiplicative factor). Also, the weak convergence of $\{n_i[U_{n_i}(\omega) - \theta(F)]/n^{\frac{1}{2}},\ n_i = [nt_i],\ i = 1, \cdots, k\ (\geqq 1)\}$ (where $0 < t_1 < \cdots < t_k \leqq 1$) to a multinormal distribution follows trivially by the same projection technique as in Hoeffding (1948). Similar results hold for $\{\theta(F_n, \omega)\}$. This leads us intuitively to consider suitable processes for $\{U_n(\omega)\}$ or $\{\theta(F_n, \omega)\}$, and to study weak convergence results similar to (1.6). With this goal in mind, we define for every $h$: $0 \leqq h \leqq m$,

$$(1.8) \qquad g_h(x_1, \cdots, x_h) = Eg(x_1, \cdots, x_h, X_{h+1}, \cdots, X_m) : \qquad g_0 = \theta(F) ,$$

$$(1.9) \qquad \zeta_h(F) = Eg_h^2(X_1, \cdots, X_h) - \theta^2(F) ; \qquad \zeta_0(F) = 0 .$$

Then, our basic *assumptions* are

(I) $0 < \zeta_1(F) < \infty$, i.e., $\theta(F)$ is stationary of order zero, and

(IIa) $\zeta_m(F) < \infty$, or

(IIb) $\zeta^*(F) = \max_{1 \leqq i_1 \leqq \cdots \leqq i_m \leqq m} Eg^2(X_{i_1}, \cdots, X_{i_m}) < \infty$, i.e., $g \in L^2$ uniformly in $1 \leqq i_1 \leqq \cdots \leqq i_m \leqq m$.

Note that (IIb) $\Rightarrow$ (IIa) but not conversely. The same assumptions underlie the asymptotic normality results of Hoeffding (1948) and von Mises (1947).

Consider the space $C[0, 1]$ of all continuous functions on $I$, and associate with it the uniform topology.

$$(1.10) \qquad \rho(X(\cdot), Y(\cdot)) = \sup_{t \in I} |X(t) - Y(t)| ,$$

where both $X$ and $Y$ belong to $C[0, 1]$. For every $n\ (\geqq 1)$, we consider the

two processes $Y_n(\omega) = [Y_n(t, \omega), t \in I]$ and $Y_n^*(\omega) = [Y_n^*(t, \omega), t \in I]$, defined below. Let $Y_n(0, \omega) = 0$,

$$Y_n(k/n, \omega) = k[\theta(F_k, \omega) - \theta(F)]/m[n\zeta_1(F)]^{\frac{1}{2}}, \quad k = 1, \cdots, n, \tag{1.11}$$

and for $t \in [(k-1)/n, k/n]$, $k = 1, \cdots, n$, by linear interpolation,

$$\begin{aligned} Y_n(t, \omega) = {} & Y_n((k-1)/n, \omega) \\ & + n(t - (k-1)/n)[Y_n(k/n, \omega) - Y_n((k-1)/n, \omega)]. \end{aligned} \tag{1.12}$$

Similarly, we let $Y_n^*(t, \omega) = 0$ for $0 \leqq t \leqq m - 1$,

$$Y_n^*(k/n, \omega) = k[U_k(\omega) - \theta(F)]/m[n\zeta_1(F)]^{\frac{1}{2}}, \quad k = m, \cdots, n, \tag{1.13}$$

and for $t \in [(k-1)/n, k/n]$, $k = m, \cdots, n$, by linear interpolation,

$$\begin{aligned} Y_n^*(t, \omega) = {} & Y_n^*((k-1)/n, \omega) \\ & + n(t - (k-1)/n)[Y_n^*(k/n, \omega) - Y_n^*((k-1)/n, \omega)]. \end{aligned} \tag{1.14}$$

Thus, $Y_n(\omega)$ and $Y_n^*(\omega)$ both have continuous sample paths and they belong to $C[0, 1]$. Then, our main theorem of the paper is the following.

THEOREM 1. *Under* (I) *and* (IIa), $Y_n^*(\omega)$ *converges weakly in the uniform topology on* $C[0, 1]$ *to a standard Brownian motion* $W$. *The same result holds for* $Y_n(\omega)$ *provided* (IIa) *is replaced by* (IIb); *under the latter condition,*

$$\rho(Y_n(\omega), Y_n^*(\omega)) \to_P 0 \quad \textit{as} \quad n \to \infty. \tag{1.15}$$

In Section 2, some results on $\theta(F_n, \omega)$ and $U_n(\omega)$, having importance of their own, are established in a sequence of lemmas. These lemmas are then utilized in the proof of the theorem in Section 3. Lemmas 2.1—2.3 constitute the basis for the proof of the convergence of differentiable statistical functions, and Lemmas 2.4—2.5 the basis for the convergence of $U$-statistics. The asymptotic equivalence (1.15) of the differentiable statistical functions and $U$-statistics will follow from these two separate proofs, but a direct enumeration proof is also given in Lemma 2.6. Pedagogically, this means the proof for the convergence of $U$-statistics can be obtained from the proof for differentiable statistical functions with the aid of Lemma 2.6, or vice versa. A direct extension of Theorem 1 to two-sample $U$-statistics and von Mises' statistics is considered in Section 4. A few applications are also briefly sketched in Section 5. Throughout the paper, unless otherwise stated, we consider $m \geqq 2$.

In a related paper Loynes (1970) studies weak convergence of reverse martingales and cites $U$-statistics as an example. His results are complementary to ours since his process is constructed from the tail sequence $\{U_k, k \geqq n\}$ whereas our process involves the variables $\{U_k, m \leqq k \leqq n\}$.

**2. Some results on $\theta(F_n, \omega)$ and $U_n(\omega)$.** For every $h$ $(1 \leqq h \leqq m)$, define

$$V_{n,h}(\omega) = \int \cdots \int_{R^{ph}} g_h(x_1, \cdots, x_h) \prod_{j=1}^{h} d[F_n(x_j, \omega) - F(x_j)], \tag{2.1}$$

so that

$$nV_{n,1}(\omega) = \sum_{i=1}^{n} [g_1(X_i) - \theta(F)] = S_n(\omega)\,, \qquad \text{say.} \tag{2.2}$$

Now, writing $dF_n(x_j, \omega) = dF(x_j) + d[F_n(x_j, \omega) - F(x_j)]$, $1 \leqq j \leqq m$, we obtain from (1.1), (1.3), (2.1) and some simplifications that

$$[\theta(F_n, \omega) - \theta(F) - m\mathrm{V}_{n,1}(\omega)] = \sum_{h=2}^{m} \tbinom{m}{h} V_{n,h}(\omega)\,, \qquad n \geqq 1\,. \tag{2.3}$$

Let $\mathscr{B}_n$ be the $\sigma$-field generated by $\{X_i, i \leqq n\}$, and let

$$V^*_{n,2}(\omega) = n^2 V_{n,2}(\omega)\,, \qquad \theta_2{}^*(F) = \int g_2(x, x)\, dF(x)\,. \tag{2.4}$$

LEMMA 2.1. *$\{V^*_{n,2}(\omega) + n[\theta(F) - \theta_2{}^*(F)], \mathscr{B}_n\}$ forms a martingale sequence.*

PROOF. By (1.2), (2.1) and (2.4),

$$\begin{aligned} V^*_{n,2}(\omega) &= V^*_{n-1,2}(\omega) \\ &\quad + 2\sum_{i=1}^{n-1} \int_{R^{2p}} \cdots \int g_2(x_1, x_2) d[c(x_1 - X_i) - F(x_1)] d[c(x_2 - X_n) - F(x_2)] \\ &\quad + \int_{R^{2p}} \cdots \int g_2(x_1, x_2)\, d[c(x_1 - X_n) - F(x_1)]\, d[c(x_2 - X_n) - F(x_2)]\,. \end{aligned} \tag{2.5}$$

Also, for all $i \leqq n - 1$,

$$E\{d[c(x_1 - X_i) - F(x_1)]\, d[c(x_2 - X_n) - F(x_2)] \mid \mathscr{B}_{n-1}\} = 0\,, \tag{2.6}$$

$$\begin{aligned} E\{d[c(x_1 - X_n) - F(x_1)]\, d[c(x_2 - X_n) - F(x_2)] \mid \mathscr{B}_{n-1}\} \\ = (\delta_{x_1 x_2})\, dF(x_1) - dF(x_1)\, dF(x_2)\,, \end{aligned} \tag{2.7}$$

where $\delta_{x_1 x_2}$ is 1 or 0 according as $x_1 = x_2$ or not. Hence,

$$E\{V^*_{n,2}(\omega) \mid \mathscr{B}_{n-1}\} = V^*_{n-1,2}(\omega) + \theta_2{}^*(F) - \theta(F)\,, \qquad n \geqq 2\,, \tag{2.8}$$

and the lemma follows. □

Now, for all $r_j \geqq 1$, $j = 1, \cdots, l\ (\geqq 1)$, $\sum_{j=1}^{l} r_j = 2s$, $s \geqq 1$,

$$\begin{aligned} &|E(\{d[c(x_1 - X_1) - F(x_1)]\}^{r_1} \cdots \{d[c(x_l - X_l) - F(x_l)]\}^{r_l})| \\ &\qquad = 0\,, \qquad \text{if at least one of } r_1, \cdots, r_l = 1\,, \\ &\qquad \leqq dF(x_1) \cdots dF(x_l)\,, \qquad \text{otherwise}\,; \end{aligned} \tag{2.9}$$

and hence, the maximum $l$ for which (2.9) is different from zero is $s$, where $r_1 = \cdots = r_s = 2$.

LEMMA 2.2. *Under* (IIb), *for every* $\varepsilon > 0$,

$$\lim_{n\to\infty} P\{\omega : \max_{1\leqq k\leqq n} k|V_{k,2}(\omega)| > \varepsilon n^{\frac{1}{2}}\} = 0\,. \tag{2.10}$$

PROOF. For every $\varepsilon > 0$, there exists an $n_0(\varepsilon)$, such that

$$|\theta(F) - \theta_2{}^*(F)| < \tfrac{1}{2}\varepsilon n^{\frac{1}{2}} \quad \text{for all} \quad n \geqq n_0(\varepsilon)\,. \tag{2.11}$$

Hence, for $n \geqq n_0(\varepsilon)$, by definition in (2.4),

$$\begin{aligned} &P\{\omega : \max_{1\leqq k\leqq n} k|V_{k,2}(\omega)| > \varepsilon n^{\frac{1}{2}}\} \\ &\qquad \leqq P\{\omega : \max_{1\leqq k\leqq n} k^{-1}|V^*_{k,2}(\omega) + k[\theta(F) - \theta^*(F)]| > \tfrac{1}{2}\varepsilon n^{\frac{1}{2}}\}\,. \end{aligned} \tag{2.12}$$

Now, under (IIb), by (2.1) and (2.9),

$$(2.13)\quad E\{V_{k,2}^*(\omega) + k[\theta(F) - \theta_2^*(F)]\}^2 \leqq 2\{E[V_{k,2}^*(\omega)]^2 + k^2[\theta(F) - \theta_2^*(F)]^2\} \leqq C_1 k^2 \zeta^*(F)\,,$$

where $C_1$ $(<\infty)$ does not depend on $k$.

Hence, by Theorem 1 of Chow (1960) (i.e., the semi-martingale extension of the Hájek-Rényi inequality), Lemma 2.1, (2.12) and (2.13), it follows that for $n \geqq n_0(\varepsilon)$,

$$(2.14)\quad \begin{aligned} &P\{\omega : \max_{1\leqq k\leqq n} k|V_{k,2}(\omega)| > \varepsilon n^{\frac{1}{2}}\} \\ &\qquad \leqq (4C_1\zeta^*(F)/n\varepsilon^2)(1 + \textstyle\sum_{j=1}^{n-1}\{j^{-2} - (j+1)^{-2}\}j^2) \\ &\qquad \leqq (8C_1\zeta^*(F)[\log n]/n\varepsilon^2) \to 0 \quad \text{as} \quad n \to \infty\,. \quad \square \end{aligned}$$

LEMMA 2.3. *If $m \geqq 3$ and* (IIb) *holds, then for every $\varepsilon > 0$,*

$$(2.15)\quad \lim_{n\to\infty} P\{\omega : \max_{1\leqq k\leqq n} k|\textstyle\sum_{h=3}^m \binom{m}{h} V_{k,h}(\omega)| > \varepsilon n^{\frac{1}{2}}\} = 0\,.$$

PROOF. By (IIb), (2.1) and (2.9), for all $k \geqq 1$,

$$(2.16)\quad E\{[\textstyle\sum_{h=3}^m \binom{m}{h} V_{k,h}(\omega)]^2\} \leqq C_2 k^{-3}\zeta^*(F)\,, \qquad C_2 < \infty\,.$$

Hence, by the Bonferroni and Markov inequalities,

$$(2.17)\quad \begin{aligned} &P\{\omega : \max_{1\leqq k\leqq n} k|\textstyle\sum_{h=3}^m \binom{m}{h} V_{k,h}(\omega)| > \varepsilon n^{\frac{1}{2}}\} \\ &\qquad \leqq \textstyle\sum_{k=1}^n P\{\omega : k|\sum_{h=3}^m \binom{m}{h} V_{k,h}(\omega)| > \varepsilon n^{\frac{1}{2}}\} \\ &\qquad \leqq \textstyle\sum_{k=1}^n \{C_2\zeta^*(F)/n\varepsilon^2\}k^{-1} \leqq C_2(\log n)\zeta^*(F)/n\varepsilon^2\,, \end{aligned}$$

which converges to 0 as $n \to \infty$. $\square$

By (1.4), for all $n \geqq m$,

$$(2.18)\quad \begin{aligned} U_n(\omega) &= n^{-[m]} \textstyle\sum_{P_{n,m}} \int_{R^{pm}} \cdots \int g(x_1, \cdots, x_m) \prod_{j=1}^m d[c(x_j - X_{i_j})] \\ &= \theta(F) + \textstyle\sum_{h=1}^m \binom{m}{h} U_{n,h}(\omega)\,, \end{aligned}$$

where $U_{n,1}(\omega) = V_{n,1}(\omega)$ is given by (2.2), and for $2 \leqq h \leqq m$,

$$(2.19)\quad U_{n,h}(\omega) = n^{-[h]} \textstyle\sum_{P_{n,h}} \int_{R^{ph}} \cdots \int g_h(x_1, \cdots, x_h) \prod_{j=1}^h d[c(x_j - X_{i_j}) - F(x_j)]\,,$$

where $P_{n,h} = \{1 \leqq i_1 \neq \cdots \neq i_h \leqq n\}$ and $n^{-[k]} = \{n \cdots (n-k+1)\}^{-1}$. Let $\mathscr{C}_n$ be the $\sigma$-field generated by the unordered collection $\{X_1, \cdots, X_n\}$ and by $X_{n+1}$, $X_{n+2}, \cdots$, so that $\mathscr{C}_n$ is $\downarrow$ in $n \geqq 1$.

LEMMA 2.4. *For every $h(1 \leqq h \leqq m)$, $\{U_{n,h}(\omega), \mathscr{C}_n, n \geqq h\}$ forms a reverse martingale sequence.*

PROOF. For a $U$-statistic, the reverse martingale property is established by Berk (1966). By (1.8) and (2.19), $U_{n,h}(\omega)$ can be expressed as

$$(2.20)\quad \binom{n}{h}^{-1} \textstyle\sum_{C_{n,h}} g_h^*(X_{i_1}, \cdots, X_{i_h})\,,$$

where $g_h^*(X_1, \cdots, X_h) = g_h(X_1, \cdots, X_h) - \sum_{j=1}^h g_{h-1}(X_1, \cdots, X_{j-1}, X_{j+1}, \cdots, X_h) + \sum_{1 \leq j < k \leq h}^h g_{h-2}(X_1, \cdots, X_{j-1}, X_{j+1}, \cdots, X_{k-1}, X_{k+1}, \cdots, X_h) - \cdots + (-1)^h \theta(F)$. Therefore the lemma directly follows from (2.20) and Berk (1966). □

LEMMA 2.5. *Under* (IIa), *for every* $\varepsilon > 0$,

$$\lim_{n \to \infty} P\{\omega : \max_{m \leq k \leq n} k \left| \sum_{h=2}^m \binom{m}{h} U_{k,h}(\omega) \right| > n^{\frac{1}{2}} \varepsilon\} = 0 . \tag{2.21}$$

PROOF. By Lemma 2.4, $\{U_k^*(\omega) = \sum_{h=2}^m \binom{m}{h} U_{k,h}(\omega), \mathscr{C}_k; k \geqq m\}$ forms a reverse martingale sequence. Hence, reversing the order of the indexing set $\{m, \cdots, n\}$, and thereby, converting the above into a (forward) martingale, we obtain on using Theorem 1 of Chow (1960) (i.e., the semi-martingale extension of the Hájek-Rényi inequality) that for every $n > m$,

$$\begin{aligned} P\{\omega : \max_{m \leq k \leq n} k |U_k^*(\omega)| > \varepsilon n^{\frac{1}{2}}\} \\ \leqq (n\varepsilon^2)^{-1}\{m^2 E[U_m^*(\omega)]^2 + \sum_{k=m+1}^n (2k+1) E[U_k^*(\omega)]^2\} . \end{aligned} \tag{2.22}$$

Now, by (2.18), for every $n \geqq m$,

$$\begin{aligned} E[U_n^*(\omega)]^2 &= \binom{n}{m}^{-1} \sum_{h=1}^m \binom{m}{h}\binom{n-m}{m-h} \zeta_h(F) - m^2 n^{-1} \zeta_1(F) \\ &\leqq C n^{-2} \zeta_m(F) , \quad \text{where } C\ (< \infty) \text{ does not depend on } n. \end{aligned} \tag{2.23}$$

Hence, the right-hand side of (2.22) is bounded above by

$$\begin{aligned} C\zeta_m(F)\{1 + 3 \sum_{k=m+1}^n k^{-1}\}/(n\varepsilon^2) \\ \leqq 3C\zeta_m(F)\{\log n\}/(n\varepsilon^2) \to 0 \quad \text{as} \quad n \to \infty . \ \square \end{aligned} \tag{2.24}$$

Finally, we prove the following lemma on the asymptotic equivalence of $\theta(F_n, \omega)$ and $U_n(\omega)$.

LEMMA 2.6. *Under* (IIb), *for every* $\varepsilon > 0$,

$$P\{\omega : \max_{m \leq k \leq n} k |\theta(F_k, \omega) - U_k(\omega)| > \varepsilon n^{\frac{1}{2}}\} \to 0 \quad \textit{as} \quad n \to \infty . \tag{2.25}$$

PROOF. By (1.3) and (1.4), we have for every $n \geqq m$,

$$\theta(F_n, \omega) = n^{-m} n^{[m]} U_n(\omega) + n^{-m} \sum_{G_{n,m}} g(X_{i_1}, \cdots, X_{i_m}) , \tag{2.26}$$

where $G_{n,m}$ is the collection of $n^m - n^{[m]}\ (= \binom{m}{2} n^{m-1} + O(n^{m-2}))$ sets of $m$ indices out of $n$ with not all the $m$ indices distinct. Then, for $k \geqq m$,

$$\begin{aligned} k|\theta(F_k, \omega) - U_k(\omega)| \leqq k(1 - k^{-m} k^{[m]}) |U_k(\omega)| \\ + k^{-(m-1)} \left| \sum_{G_{k,m}} g(X_{i_1}, \cdots, X_{i_m}) \right| . \end{aligned} \tag{2.27}$$

Since $k(1 - k^{-m} k^{[m]}) = \binom{m}{2} + O(k^{-1})$, and $\{U_k(\omega), \mathscr{C}_k, k \geqq m\}$ forms a reverse martingale, by the extended Kolmogorov inequality, $P\{\max_{m \leq k \leq n} |U_k(\omega)| > \varepsilon n^{\frac{1}{2}}\} \to 0$ as $n \to \infty$, and hence, to prove (2.25), it suffices to show that

$$P\{\max_{m \leq k \leq n} [k^{-(m-1)} \left| \sum_{G_{k,m}} g(X_{i_1}, \cdots, X_{i_m}) \right|] > \varepsilon n^{\frac{1}{2}}\} \to 0 \quad \text{as} \quad n \to \infty . \tag{2.28}$$

Let $g^*(X_{i_1}, \cdots, X_{i_m}) = g(X_{i_1}, \cdots, X_{i_m}) - Eg(X_{i_1}, \cdots, X_{i_m})$. Since, by (IIb), $k^{-(m-1)}|\sum_{G_{k,m}} Eg(X_{i_1}, \cdots, X_{i_m})| \leqq [\zeta^*(F)]^{\frac{1}{2}} k^{-(m-1)}\{k^m - k^{[m]}\} = \binom{m}{2}[\zeta^*(F)]^{\frac{1}{2}}\{1 + O(k^{-1})\}$, for every $\varepsilon > 0$, it can be made smaller than $\varepsilon n^{\frac{1}{2}}$ when $n$ is taken adequately large. So, we complete the proof by showing that

$$P\{\omega : \max_{m \leqq k \leqq n} |k^{-(m-1)} \textstyle\sum_{G_{k,m}} g^*(X_{i_1}, \cdots, X_{i_m})| > \varepsilon n^{\frac{1}{2}}\} \to 0 \quad \text{as} \quad n \to \infty . \tag{2.29}$$

Now, in the expression $E\{\sum_{G_{k,m}} g^*(X_{i_1}, \cdots, X_{i_m})\}^2 = \sum_{G_{k,m}} \sum_{G_{k,m}} E\{g^*(X_{i_1}, \cdots, X_{i_m}) g^*(X_{j_1}, \cdots, X_{j_m})\}$, there are $\binom{m}{2}^2 k^{2m-2} + O(k^{2m-3})$ terms of the type $E\{g^*(X_{i_1}, X_{i_1}, X_{i_2}, \cdots, X_{i_{m-1}}) g^*(X_{j_1}, X_{j_1}, X_{j_2}, \cdots, X_{j_{m-1}})\}$ where $\{i_1, \cdots, i_{m-1}\} \cap \{j_1, \cdots, j_{m-1}\} = \emptyset$, and the repeated variable $X_{i_1}$ $(X_{j_1})$ can appear in any of the $\binom{m}{2}$ pairs of arguments. By independence, the contribution of these terms is equal to zero, while the contribution of the remaining terms is $O(k^{2m-3})$. Therefore, $E\{k^{-(m-1)} \sum_{G_{k,m}} g^*(X_{i_1}, \cdots, X_{i_m})\}^2 \leqq C\zeta^*(F)/k$, where $C < \infty$. Hence, by Boole's and Chebyshev's inequality, the left-hand side of (2.29) is bounded by

$$\begin{aligned} &\textstyle\sum_{k=m}^n P\{\omega : |k^{-(m-1)} \sum_{G_{k,m}} g^*(X_{i_1}, \cdots, X_{i_m})| > \varepsilon n^{\frac{1}{2}}\} \\ &\qquad \leqq (n\varepsilon^2)^{-1} \textstyle\sum_{k=m}^n C\zeta^*(F)/k \\ &\qquad \leqq [C\zeta^*(F)][(\log n)/n\varepsilon^2] \to 0 \quad \text{as} \quad n \to \infty . \quad \square \end{aligned} \tag{2.30}$$

**3. The proof of Theorem 1.** For $t \in I$, define $Y_n^0(\omega) = [Y_n^0(t, \omega), t \in I]$, by

$$Y_n^0(t, \omega) = \{S_{[nt]}(\omega) + (nt - [nt])[g_1(X_{[nt]+1}) - \theta(F)]\}/\{n\zeta_1(F)\}^{\frac{1}{2}} , \tag{3.1}$$

where $S_k(\omega)$, $k \geqq 1$, are defined by (2.2). Then by the Donsker theorem [cf. Billingsley (1968, page 68)],

$$Y_n^0(\omega) \to_{\mathscr{D}} W , \qquad \text{as} \quad n \to \infty . \tag{3.2}$$

We complete the proof of our theorem by showing that

$$\rho(Y_n(\omega), Y_n^0(\omega)) \to_P 0 , \qquad \rho(Y_n^*(\omega), Y_n^0(\omega)) \to_P 0 , \qquad \text{as} \quad n \to \infty . \tag{3.3}$$

Now, by (1.11), (1.12), (2.2), (2.3) and (3.1), under (IIb),

$$\begin{aligned} \rho(Y_n(\omega), Y_n^0(\omega)) &= \sup_{t \in I} |Y_n(t, \omega) - Y_n^0(t, \omega)| \\ &= \{\max_{1 \leqq k \leqq n} k|\theta(F_k, \omega) - \theta(F) - mV_{k,1}(\omega)|\}/\{m[n\zeta_1(F)]^{\frac{1}{2}}\} \\ &\leqq \{\max_{1 \leqq k \leqq n} k\tbinom{m}{2}|V_{k,2}(\omega)| \\ &\qquad + \max_{1 \leqq k \leqq n} k|\textstyle\sum_{h=3}^m \tbinom{m}{h} V_{k,h}(\omega)|\}/\{m[n\zeta_1(F)]^{\frac{1}{2}}\} \\ &\to_P 0 \quad \text{as} \quad n \to \infty , \\ &\qquad \text{by Lemma 2.2 and Lemma 2.3.} \end{aligned} \tag{3.4}$$

Hence, $Y_n(\omega) \to_{\mathscr{D}} W$. Also (3.2) along with $\lim_{n\to\infty} m/n = 0$ implies that

$$\sup_{0 \leqq t \leqq (m-1)/n} |Y_n^0(t, \omega)| \to_P 0 . \tag{3.5}$$

Further, by (1.13), (1.14), (2.18), (3.1) and (3.5),

$$\begin{aligned}
&\rho(Y_n^*(\omega), Y_n^0(\omega)) \\
&\quad = \sup_{t\in I} |Y_n^*(t, \omega) - Y_n^0(t, \omega)| \\
&\quad \leqq \sup_{0\leqq t\leqq (m-1)/n} |Y_n^0(t, \omega)| + \sup_{m/n\leqq t\leqq 1} |Y_n^*(t, \omega) - Y_n^0(t, \omega)| \\
&\quad = \sup_{0\leqq t\leqq (m-1)/n} |Y_n^0(t, \omega)| \\
&\qquad + \{\max_{m\leqq k\leqq n} k|U_k(\omega) - \theta(F) - mU_{k,1}(\omega)|\}/\{m[n\zeta_1(F)]^{\frac{1}{2}}\}\,,
\end{aligned} \tag{3.6}$$

where by Lemma 2.5, the second term on the right-hand side of (3.6) $\to_P 0$ as $n \to \infty$. Hence by (3.5) and (3.6), $\rho(Y_n^*(\omega), Y_n^0(\omega)) \to_P 0$ as $n \to \infty$, i.e., $Y_n^*(\omega) \to_{\mathscr{D}} W$.

Finally, $\rho(Y_n^*(\omega), Y_n(\omega)) \to_P 0$ follows from Lemma 2.6, or directly from (3.3)—(3.6) along with the triangle inequality

$$\rho(Y_n^*(\omega), Y_n(\omega)) \leqq \rho(Y_n^*(\omega), Y_n^0(\omega)) + \rho(Y_n^0(\omega), Y_n(\omega))\,.$$

**4. Generalizations.** For brevity and simplicity of presentation, we only consider the case of generalized $U$-statistics; for von Mises' functionals, the results follow on parallel lines. Let $\omega = \{X_1, X_2, \cdots; Y_1, Y_2, \cdots\}$ be a sequence of rv where each $X_i$ has a df $F(x)$ and each $Y_j$ has a df $G(y)$. For a regular functional $\theta(F, G)$ of degree $(m_1, m_2)$ $(m_1 \geqq 1, m_2 \geqq 1)$, viz.,

$$\theta(F, G) = \int \cdots \int g(x_1, \cdots, x_{m_1}; y_1, \cdots, y_{m_2}) \textstyle\prod_{i=1}^{m_1} dF(x_i) \prod_{j=1}^{m_2} dG(y_j)\,, \tag{4.1}$$

where the (real-valued) kernel $g(\cdot; \cdot)$ is symmetric in the $m_1$ arguments in the first $m_1$ places and also symmetric in the last $m_2$ arguments, the generalized $U$-statistic [cf. Lehmann (1951)] is defined, for $n_1 \geqq m_1$, $n_2 \geqq m_2$, by

$$U_{n_1,n_2}(\omega) = \tbinom{n_1}{m_1}^{-1}\tbinom{n_2}{m_2}^{-1} \textstyle\sum_{C_{n_1,m_1}} \sum_{C_{n_2,m_2}} g(X_{s_1}, \cdots, X_{s_{m_1}}; Y_{r_1}, \cdots, Y_{r_{m_2}})\,, \tag{4.2}$$

where $C_{n_1,m_1} = \{1 \leqq s_1 < \cdots < s_{m_1} \leqq n_1\}$ and $C_{n_2,m_2} = \{1 \leqq r_1 < \cdots < r_{m_2} \leqq n_2\}$.

The sample sizes $(n_1, n_2)$ will be defined by $n_1 + n_2 = n$, $n_1 = n\lambda_n$, $n_2 = n(1 - \lambda_n)$, and it is assumed that (i) $n\lambda_n$ and $n(1 - \lambda_n)$ are non-decreasing functions of $n$, and (ii) $\lambda_n$ tends to a constant $\lambda (0 < \lambda < 1)$ as $n \to \infty$. Let then

$$g_{10}(x_1) = Eg(x_1, X_2, \cdots, X_{m_1}; Y_1, \cdots, Y_{m_2})\,, \tag{4.3}$$

$$g_{01}(y_1) = Eg(X_1, \cdots, X_{m_1}; y_1, Y_2, \cdots, Y_{m_2})\,, \tag{4.4}$$

$$\zeta_{10}(F, G) = Eg_{10}^2(X_1) - \theta^2(F, G)\,, \qquad \zeta_{01}(F, G) = Eg_{01}^2(Y_1) - \theta^2(F, G)\,; \tag{4.5}$$

$$\zeta(F, G) = m_1^2\lambda^{-1}\zeta_{10}(F, G) + m_2^2(1 - \lambda)^{-1}\zeta_{01}(F, G)\,. \tag{4.6}$$

Define a process $[Z_n(t, \omega), t \in I]$ on $C[0, 1]$ by

$$Z_n(t, \omega) = 0 \quad \text{if} \quad [nt]\lambda_{[nt]} < m_1 \quad \text{or} \quad [nt](1 - \lambda_{[nt]}) < m_2\,, \tag{4.7}$$

$$Z_n(k/n, \omega) = \frac{k[U_{k\lambda_k, k(1-\lambda_k)}(\omega) - \theta(F, G)]}{[n\zeta(F, G)]^{\frac{1}{2}}}\,, \tag{4.8}$$

if $k\lambda_k \geqq m_1$, $k(1-\lambda_k) \geqq m_2$, and by linear interpolation, let

$$Z_n(t, \omega) = Z_n(k/n, \omega) + n(t - k/n)[Z_n((k+1)/n, \omega) - Z_n(k/n, \omega)] , \tag{4.9}$$

for $k/n < t \leqq (k+1)/n$, $k = 0, 1, \cdots, n-1$. Then we have the following.

THEOREM 4.1. *If $g \in L^2$ and $\zeta(F, G) > 0$, as $n \to \infty$, $Z_n(\omega) = [Z_n(t, \omega), t \in I]$ converges weakly in the uniform topology on $C[0, 1]$ to a standard Brownian motion.*

The proof is omitted because it is a straightforward duplication of the proof of Theorem 1. The theorem directly extends to regular functionals $\theta(F_1, \cdots, F_c)$ of $c$ ($\geqq 2$) independent distributions.

**5. Applications.** (i) *Asymptotic normality for random sample sizes.* For every (index) variable $r$ ($\geqq 1$), let $N_r$ be an integer-valued (nonnegative) random variable and $n_r$ be a real (positive) number, such that

$$\lim_{r\to\infty} n_r = \infty \quad \text{and} \quad \lim_{r\to\infty} (N_r/n_r) = 1 , \quad \text{in probability.} \tag{5.1}$$

This implies that for every $\delta > 0$,

$$\lim_{r\to\infty} P\{|n_r^{-1}N_r - 1| > \delta\} = 0 . \tag{5.2}$$

Now, by the tightness property [cf. Billingsley (1968, page 54)] of $W_t$, $t \geqq 0$, for every $\varepsilon > 0$ and $\eta > 0$, there exists a $\delta > 0$, such that

$$P\{\sup_{|t-s|<\delta} |W_t - W_s| > \varepsilon \,|\, t, s \in I\} < \eta . \tag{5.3}$$

By (5.1) and a well-known theorem of Cramér (1946, page 254) $N_r^{\frac{1}{2}}[\theta(F_{N_r}, \omega) - \theta(F)]$ and $n_r^{\frac{1}{2}}[\theta(F_{N_r}, \omega) - \theta(F)]$ both have the same limiting distribution, if they have one at all. Since $n_r^{\frac{1}{2}}[\theta(F_{n_r}, \omega) - \theta(F)]$ is asymptotically normally distributed [cf. Hoeffding (1948)] with zero mean and variance $m^2\zeta_1(F)$, we need only show that as $r \to \infty$

$$n_r^{\frac{1}{2}}|\theta(F_{N_r}, \omega) - \theta(F_{n_r}, \omega)| \to_P 0 \tag{5.4}$$

to establish that $N_r^{\frac{1}{2}}[\theta(F_{N_r}, \omega) - \theta(F)]$ is asymptotically normally distributed with zero mean and variance $m^2\zeta_1(F)$. But (5.4) follows readily from (1.11), (1.12), (5.1), (5.3), and Theorem 1, so the desired asymptotic normality follows.

A similar result for $U_{N_r}(\omega)$ has been obtained by Sproule (1969) by an indirect method involving some elaborate analysis. Our proof follows from Theorem 1, (5.2) and (5.3).

(ii) *Signed-rank statistic.* Consider the kernel

$$\phi(x_1, x_2) = \operatorname{sgn}(x_1 + x_2)[2 - \operatorname{sgn}^2(x_1 - x_2)] , \tag{5.5}$$

and assume that the $X_i$ have a continuous df $F(x)$, $x \in R^1$. Then

$$\theta(F) = \int_{-\infty}^{\infty} [1 - 2F(-x)]\, dF(x) \quad (=0 \text{ when } F \text{ is symmetric about } 0). \tag{5.6}$$

The corresponding $\theta(F_n, \omega)$ is given by

$$n\theta(F_n, \omega) = 2n^{-1} \textstyle\sum_{i=1}^{n} R_{ni} \operatorname{sgn}(X_i) = 2W_n^0, \qquad \text{say,} \tag{5.7}$$

where $R_{ni}$ = Rank of $|X_i|$ among $|X_1|, \cdots, |X_n|$, $\operatorname{sgn} u = 1, 0$ or $-1$ according as $u$ is $>$, $=$ or $< 0$, and $W_n^0$ is the classical Wilcoxon signed rank statistic. Also, here

$$\zeta_1(F) = \textstyle\int_{-\infty}^{\infty} [1 - 2F(-x)]^2 \, dF(x) - \theta^2(F), \tag{5.8}$$

which equals $\frac{1}{3}$ when $F$ is symmetric about 0. Thus, if we let

$$\begin{aligned} Y_n(t, \omega) &\\ &= \{(nt - [nt])W^0_{[nt]+1} + ([nt] + 1 - nt)W^0_{[nt]} - \tfrac{1}{2}nt\theta(F)\}/[n\zeta_1(F)]^{\frac{1}{2}}, \end{aligned} \tag{5.9}$$

$0 \leqq t \leqq 1$, we have from Theorem 1 that if $\zeta_1(F) > 0$,

$$Y_n(\omega) \to_{\mathscr{D}} W, \tag{5.10}$$

and thus the sequence $\{W_k^0 - \frac{1}{2}k\theta(F); k \geqq 1\}$ though not linear in the basic random variables is attracted by the Brownian motion processes. For example, by virtue of Theorem 1 and well-known results on Brownian motion processes [viz., Billingsley (1968, page 79)], we obtain that for all $\lambda > 0$,

$$\lim_{n\to\infty} P\{\sup_{t\in I} Y_n(t, \omega) > \lambda\} = P\{\sup_{t\in I} W_t > \lambda\} = (2/\pi)^{\frac{1}{2}} \textstyle\int_\lambda^\infty e^{-\frac{1}{2}t^2} \, dt, \tag{5.11}$$

$$\begin{aligned} \lim_{n\to\infty} P\{\omega : \sup_{t\in I} |Y_n(t, \omega)| > \lambda\} &\\ &= P\{\sup_{t\in I} |W_t| > \lambda\} \\ &= 1 - \textstyle\sum_{k=-\infty}^{\infty} (-1)^k[\Phi((2k+1)\lambda) - \Phi((2k-1)\lambda)], \end{aligned} \tag{5.12}$$

where $\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp(-\frac{1}{2}t^2) \, dt$, $-\infty < x < \infty$. There is evidence (see Miller (1971)) that the convergence to the limit in (5.11) and (5.12) is very rapid so the limiting probabilities are good approximations for $n$ as small as 10.

Weak convergence results studied in this paper will aid in the sequential analysis of $U$-statistics or von Mises' statistics, particularly, in the study of the OC functions where Brownian movement approximations greatly simplify the asymptotic results.

## REFERENCES


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

[2] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[3] Chow, Y. S. (1960). A martingale inequality and the law of large numbers. *Proc. Amer. Math. Soc.* **11** 107–111.

[4] Cramér, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

[5] Donsker, M. (1951). An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6** 1–12.

[6] Donsker, M. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **23** 277–281.

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

[8] LEHMANN, E. L. (1951). Consistency and unbiasedness of certain non-parametric tests. *Ann. Math. Statist.* **22** 165–179.

[9] LOYNES, R. M. (1970). An invariance principle for reversed martingales. *Proc. Amer. Math. Soc.* **25** 56–64.

[10] MILLER, R. G., JR. (1970). Sequential rank tests—one sample case. *Proc. 6th Berkeley Symp. Math. Statist. Prob.* (L. LeCam, *et al.*, eds.). Univ. of California Press.

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

[12] SPROULE, R. N. (1969). A sequential fixed-width confidence interval for the mean of a $U$-statistic. Doctoral Dissertation, Univ. of North Carolina, Chapel Hill.