The Annals of Statistics
1977, Vol. 5, No. 2, 405–407

NOTE ON CONDITIONS FOR WEAK CONVERGENCE OF VON MISES' DIFFERENTIABLE STATISTICAL FUNCTIONS

BY N. BÖNNER AND H.-P. KIRSCHNER

University of Freiburg

In establishing weak convergence of von Mises' differentiable statistical functions to a normal distribution usually square integrability conditions with respect to the underlying kernel function are assumed. It is shown that these conditions can be weakened by assuming integrability of the von Mises' functional itself. In addition it is pointed out that in nontrivial cases the conditions of square integrability of the kernel do not hold whereas weak convergence of the von Mises' functional can still be proved.

Let $\Phi(x_1, \cdots, x_m)$ be a symmetric kernel of degree m on $\mathbb{R}_{k\cdot m}$, $x_i \in \mathbb{R}_k$, and $\Theta(F_n) = \int_{\mathbb{R}_{k\cdot m}} \Phi(x_1, \cdots, x_m)\, dF_n(x_1) \cdots dF_n(x_m)$ be the corresponding von Mises' functional of the empirical df $F_n(x)$; let U_n be the U-statistic with respect to Φ and denote by $\Theta(F) = E_F \Phi(X_1, \cdots, X_m)$ the estimable parameter, where $X_1, \cdots, X_m$ are i.i.d. random variables, distributed according to the df F (for the terminology cf. [1] or [3]). In [1] and [2] regularity conditions are given for weak convergence of U_n and $\Theta(F_n)$, respectively, to a normal distribution. With rigor this result can be stated as follows.

THEOREM. *Let* $\{X_i\}$ *be a sequence of* i.i.d. *random variables with values in* $\mathbb{R}_k$, *distributed according to the* df F.

(a) *Suppose*

$$\text{Var}_F\, \Phi(X_1, \cdots, X_m) < \infty\,; \tag{1}$$

then, for Φ *being stationary of order zero,*

$$\mathfrak{L}(n^{\frac{1}{2}}(U_n - \Theta(F))) \to \mathscr{N}(0, \zeta_1 m^2)\,, \qquad as \quad n \to \infty\,. \tag{2}$$

(b) *Suppose*

$$\max\{E_F \Phi^2(X_{i_1}, \cdots, X_{i_m}),\ 1 \leqq i_1 \leqq \cdots \leqq i_m \leqq m\} < \infty\,. \tag{3}$$

Then it follows

$$n^{\frac{1}{2}}(U_n - \Theta(F_n)) \to_P 0\,, \qquad as \quad n \to \infty\,; \tag{4}$$

hence $n^{\frac{1}{2}}(U_n - \Theta(F))$ *and* $n^{\frac{1}{2}}(\Theta(F_n) - \Theta(F))$ *have the same limiting distribution if they have any, and, for* Φ *being stationary of order zero,*

$$\mathfrak{L}(n^{\frac{1}{2}}(\Theta(F_n) - \Theta(F))) \to \mathscr{N}(0, \zeta_1 m^2)\,, \qquad as \quad n \to \infty\,. \tag{5}$$

Received June 1975; revised October 1975.

AMS 1970 *subject classifications*. Primary 62E20; Secondary 62E15.

Key words and phrases. Von Mises' differentiable statistical functions, U-statistics, weak convergence.

Note that condition (1) is only needed in proving the weak convergence of the U-statistic, whereas (3) yields the asymptotic equivalence in probability—(4)—of U_n and $\Theta(F_n)$.

In the sequel it is shown that condition (3) is too restrictive in the sense that nontrivial cases in which (5) and in particular (4) hold are excluded. Since in addition in the literature there seems to be some confusion about the assumptions necessary in proving (5) (viz. [1], where it is assumed that the variance of $\Theta(F_n)$ exists), for the sake of completeness we prove

THEOREM 1. *Let $\{X_i\}$ be a sequence of* i.i.d. *random variables with values in $\mathbb{R}_k$, distributed according to the* df F. *Let Φ be a symmetric kernel of degree $m \geqq 1$ on $\mathbb{R}_{k\cdot m}$. Suppose*

$$\text{(3')} \qquad \max\{|E_F\Theta(F_k)|,\ 1 \leqq k \leqq m\} < \infty\ .$$

Then, for all $n \geqq 1$, $E_F\Theta(F_n) \in \mathbb{R}$, and $\Theta(F_n) - U_n = O(n^{-1})$ a.s. *as $n \to \infty$. Hence $n^{\frac{1}{2}}(U_n - \Theta(F))$ and $n^{\frac{1}{2}}(\Theta(F_n) - \Theta(F))$ have the same limiting distribution if they have any.*

PROOF. First, it readily follows from simple combinatorial arguments that

$$\text{(6)} \qquad \Theta(F_n) = n^{-m} \sum_{k=1}^{n\wedge m} \sum_{(1)} \sum_{(2)} \frac{m!}{\prod_{j=1}^{k} \nu_j!} \Phi(\underbrace{x_{i_1}, \cdots, x_{i_1}}_{\nu_1\text{-times}}, \cdots, \underbrace{x_{i_k}, \cdots, x_{i_k}}_{\nu_k\text{-times}})$$

where summation $\sum_{(1)}$ extends over all $1 \leqq i_1 < i_2 < \cdots < i_k \leqq n$, $1 \leqq k \leqq n \wedge m$, and $\sum_{(2)}$ extends over all $(\nu_1, \cdots, \nu_k)$, $\nu_j \geqq 1$, $\forall j \in \{1, \cdots, k\}$, $\sum_{j=1}^{k} \nu_j = m$, $m, n \geqq 1$. For $n \geqq m$, (6) turns out to be the well-known representation of $\Theta(F_n)$ as a linear combination of U-statistics, though no precise reference for (6) is known to the authors.

Define

$$\varphi_k(x_1, \cdots, x_k) = \sum_{(2)} \frac{m!}{\prod_{j=1}^{k} \nu_j!} \Phi(\underbrace{x_1, \cdots, x_1}_{\nu_1\text{-times}}, \cdots, \underbrace{x_k, \cdots, x_k}_{\nu_k\text{-times}})\ , \qquad k \leqq m\ ,$$

with summation $\sum_{(2)}$ as in formula (6). In a fashion similar to $\Theta(F)$ define $\Theta_k(F) = E_F \varphi_k(X_1, \cdots, X_k)$. We shall prove first by induction that, for all $k \in \{1, \cdots, m\}$, $\Theta_k(F) \in \mathbb{R}$. Obviously $\Theta(F_1) = \varphi_1(x_1)$ and therefore $\Theta_1(F) \in \mathbb{R}$. Suppose $\Theta_s(F) \in \mathbb{R}$, $\forall s \leqq s_0 < m$, then (6) yields

$$(s_0 + 1)^m \Theta(F_{s_0+1}) - \sum_{k=1}^{s_0} \sum_{(1)} \varphi_k(x_{i_1}, \cdots, x_{i_k}) = \varphi_{s_0+1}(x_1, \cdots, x_{s_0+1})\ .$$

Hence $\Theta_{s_0+1}(F) \in \mathbb{R}$.

Now $\Theta_k(F) \in \mathbb{R}$, $\forall k \leqq m$, together with (6) implies $E_F\Theta(F_n) \in \mathbb{R}$, $\forall n \geqq 1$. Since $\Theta_k(F)$ is estimable with respect to φ_k, by Theorem 3.2.6., page 60, in [3], we have $U_n^{(k)} = \binom{n}{k}^{-1} \sum_{(1)} \varphi_k(X_{i_1}, \cdots, X_{i_k}) \to_{\text{a.s.}} \Theta_k(F)$, as $n \to \infty$. Therefore the theorem follows immediately from

$$\text{(7)} \qquad \Theta(F_n) - U_n = \left(\frac{n^{[m]}}{n^m} - 1\right) U_n + \sum_{k=1}^{m-1} \binom{n}{k} n^{-m} U_n^{(k)}\ . \qquad \square$$

Note that the reverse martingale property of $\{U_n^{(k)}\}$, for each k, implies that $\max_{m \leqq i \leqq n} |U_i^{(k)}| = O_P(1)$ under (3′), $1 \leqq k \leqq m$, so that by (7), we conclude that under (3′), $n^{-\frac{1}{2}} \max_{m \leqq i \leqq n} i|\Theta(F_i) - U_i| \to_P 0$ as $n \to \infty$. Thus, for the results of [2], also (3′) suffices.

Also note that by the same arguments used at the beginning of the proof $\operatorname{Var}_F \Theta(F_k) < \infty$, $\forall\, k \leqq m$, implies $\operatorname{Var}_F \Theta(F_n) < \infty$, $\forall\, n \geqq 1$. Finally we note that, of course, (3) implies (3′) but the converse is not true.

We shall now give an example such that (3) is not fulfilled but the relations (4) and (5) hold.

Let F be a continuous df on $\mathbb{R}$ with $F(x) + F(-x) = 1$, for all $x \in \mathbb{R}$, such that $\int |x|\, dF(x) < \infty$ but $\int x^2\, dF(x) = \infty$. Let $\mathfrak{J}_A$ be the indicator function of the set $A \subset \mathbb{R}$. Define

$$\Phi(x_1, x_2, x_3) = \prod_{i=1}^{3} \{\mathfrak{J}_{(-\infty,0]}(x_i) - \tfrac{1}{2}\mathfrak{J}_{(0,\infty)}(x_i)\} \qquad \text{and}$$
$$\tilde{\Phi}(x_1, x_2, x_3) = (x_2 - x_3)\mathfrak{J}_{\{x_2\}}(x_1) + (x_2 - x_1)\mathfrak{J}_{\{x_2\}}(x_3) + (x_1 - x_2)\mathfrak{J}_{\{x_1\}}(x_3)\,,$$

for $x_i \in \mathbb{R}$, $i = 1, 2, 3$.

Now let $\{X_i\}$ be a sequence of i.i.d. random variables, distributed according to F. Then, for the symmetric kernel Φ, one easily computes $E_F\, \Phi(X_1, X_2, X_3) = 4^{-3}$, $\operatorname{Var}_F \Phi_1(X_1) = 9 \cdot 4^{-6}$, $\operatorname{Var}_F \Phi(X_1 X_2, X_3) < \infty$, and both condition (1) and (3) hold. Thus, for $\Theta(\Phi, F_n)$ based on Φ, it follows $\mathfrak{L}(n^{\frac{1}{2}}(\Theta(\Phi, F_n) - 4^{-3})) \to \mathscr{N}(0, 3^4 \cdot 4^{-6})$, as $n \to \infty$.

Now consider $\hat{\Phi} = \Phi + \tilde{\Phi}$. It readily follows from the definition and the continuity of F that $E_F\, \hat{\Phi}^2(X_1, X_2, X_2) = \infty$, hence (3) does not hold for the symmetric kernel $\hat{\Phi}$.

On the other hand, one easily computes that $U_n(\Phi) = U_n(\hat{\Phi})$ a.s. and $\Theta(\Phi, F_n) = \Theta(\hat{\Phi}, F_n)$ a.s., for all $n \geqq 1$, such that in particular $E_F|\Theta(\hat{\Phi}, F_k)| < \infty$, for $k = 1, 2, 3$, hence (3′). Finally one obtains $\mathfrak{L}(n^{\frac{1}{2}}(\Theta(\hat{\Phi}, F_n) - 4^{-3})) \to \mathscr{N}(0, 3^4 \cdot 4^{-6})$, as $n \to \infty$.

Summing up this example one can say that condition (3) is too sensitive against small variations on sets of $F \otimes F \otimes F$-measure zero.

REFERENCES

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

[2] 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.

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

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
DER ALBERT-LUDWIGS-UNIVERSITÄT
78 FREIBURG IM BREISGAU
HERMANN-HERDER-STRASSE 10
GERMANY