Reprinted from JOURNAL OF MULTIVARIATE ANALYSIS
All Rights Reserved by Academic Press, New York and London
Vol. 39, No. 1, October 1991
Printed in Belgium

On Hadamard Differentiability of Extended Statistical Functional

JIAN-JIAN REN

University of Nebraska-Lincoln

AND

PRANAB KUMAR SEN

University of North Carolina at Chapel Hill

Communicated by C. R. Rao

It has been shown (Reeds, 1976, Ph.D. dissertation, Harvard University) that the remainder term of a form of the Taylor expansion, involving Hadamard derivative, of the statistical functional is asymptotically negligible. This result is extended to a more general form with respect to weighted empirical processes in order to establish some (uniform) linear functional approximations, which is usually needed for drawing statistical conclusions (in a large sample). © 1991 Academic Press, Inc.

1. INTRODUCTION

In nonparametric models, a parameter θ $(= T(F))$ is regarded as a functional $T(\cdot)$ on a space $\mathscr{F}$ of distribution functions (d.f.) F. Thus, the same functional of the sample d.f. F_n (i.e., $T(F_n)$) is regarded as a natural estimator of θ. Using a form of the Taylor expansion involving the derivatives of the functional, Von Mises [8] expressed $T(F_n)$ as

$$T(F_n) = T(F) + T'_F(F_n - F) + \text{Rem}(F_n - F; T(\cdot)), \tag{1.1}$$

where T'_F is the derivative of the functional at F and $\text{Rem}(F_n - F; T(\cdot))$ is the remainder term in this first-order expansion. Note that $F_n(x) = (1/n) \sum_{i=1}^n I(X_i \leqslant x)$ is based on n independent and identically dis-

Received February 8, 1990; revised October 18, 1990.
AMS 1980 subject classifications: 62F35, 62G99.
Key words and phrases: extended statistical functional, Hadamard differentiability, robust (M-)estimation, statistical functional, uniform asymptotic linearity, weighted empirical processes.

0047-259X/91 $3.00

tributed random variables (i.i.d.r.v.) $X_1, ..., X_n$, each having the d.f. F, and that T'_F is a linear functional. Hence, $T'_F(F_n - F)$ is an average of n i.i.d.r.v.'s. For drawing statistical conslusions (in a large sample), T'_F plays the basic role, and in this context, it remains to show that $\text{Rem}(F_n - F; T(\cdot))$ is asymptotically negligible to the desired extent. Approptiate differentiability conditions are usually incorporated towards this verification.

We observe that a statistical functional induces a functional on the space $D[0, 1]$ (of right continuous functions having left-hand limits) in the following way:

$$\tau(G) = T(G \circ F), \qquad G \in D[0, 1]. \tag{1.2}$$

Thus, (1.1) can be written equivalently as

$$\tau(U_n) = \tau(U) + \tau'_U(U_n - U) + \text{Rem}(U_n - U; \tau(\cdot)), \tag{1.3}$$

where U_n is the empirical d.f. of the $F(X_i)$, $1 \leqslant i \leqslant n$, and U is the classical uniform d.f. on $[0, 1]$ (i.e., $U(t) = t$, $0 \leqslant t \leqslant 1$). Since the expansion in (1.1), written in (1.3), is based on some kind of differentiation, it is quite natural to inquire about the right form of such a differentiation to suit the desired purpose. The current literature is based on an extensive use of the Fréchet derivatives which are generally too stringent. Less restrictive concepts involve the Gâteaux and Hadamard (or compact) derivatives (viz., Kallianpur [4], Reeds [6], and Fernholz [2], among others). Using the Hadamard differentiability (along with some other regularity conditions), Reeds [6] has shown that

$$\sqrt{n}\, \text{Rem}(U_n - U; \tau(\cdot)) \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty \tag{1.4}$$

so that noting that $\tau'_U(U_n - U) = (1/n) \sum_{i=1}^{n} IC(X_i; F, T)$, where $IC(x; F, T)$ is the influence function of T at F, and assuming that $\sigma^2 = \text{Var}_F\{IC(X_i; F, T)\} < \infty$, one obtains that

$$\sqrt{n}\,(T(F_n) - \theta) \overset{P}{\sim} \sqrt{n}\,\tau'_U(U_n - U) \xrightarrow{\mathscr{D}} N(0, \sigma^2). \tag{1.5}$$

In the context of the law of iterate logarithm or some almost sure (a.s.) representation for TF_n), one may require a stronger mode of convergence in (1.4), and this, in turn, may require a more stringent differentiability condition. However, in a majority of statistical applications, Hadamard differentiability suffices, and we shall explore this concept in the context of extended statistical functions arising in robust (M-)estimation in simple linear models.

Our main results in Section 3 extend the result (1.4) to a more general form with respect to the weighted empirical process:

$\{S_n^*(\cdot, u); u \in R, n \geqslant 1\}$, where, for a sequence $\{c_{ni}\}$ with $\sum_{i=1}^n c_{ni}^2 = 1$, $S_n^*(\cdot, u) = \sum_{i=1}^n c_{ni} I(Y_i \leqslant F^{-1}(\cdot) + c_{ni} u)$. Our results deal with the remainder term uniformly over a variable u in a compact set so that they, as an application, may be applied to uniform asymptotic linearity of some statistics, viz., the M-estimators in linear regression models.

Consider the simple linear model:

$$X_i = \mathbf{b}^{\mathrm{T}} \mathbf{c}_i + e_i, \qquad i \geqslant 1, \tag{1.6}$$

where the $\mathbf{c}_i$ are known p-vectors of regression constants, $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)^{\mathrm{T}}$ is the vector of unknown (regression) parameters, $p \geqslant 1$, and e_i are i.i.d.r.v.'s with d.f. $F\ (\in \mathscr{F})$. Based on a suitable score function $\psi: R \to R$, an M-estimator $\hat{\boldsymbol{\beta}}_n$ of $\boldsymbol{\beta}$ is defined as a solution (with respect to $\boldsymbol{\theta}$) of the equations

$$\sum_{i=1}^n \mathbf{c}_i \psi(X_i - \boldsymbol{\theta}^{\mathrm{T}} \mathbf{c}_i) \equiv \mathbf{0}, \tag{1.7}$$

where "$\equiv \mathbf{0}$" accommodates the possibility of left-hand side being closest to $\mathbf{0}$ when equality in (1.7) is unattainable (such a case may arise when ψ is not continuous everywhere). Setting $Y_i = X_i - \boldsymbol{\beta}^{\mathrm{T}} \mathbf{c}_i$ (i.i.d.r.v.'s with d.f. F), we shall see that the empirical function

$$\mathbf{S}_n^*(t, \mathbf{u}) = \sum_{i=1}^n \mathbf{c}_{ni} I(Y_i \leqslant F^{-1}(t) + \mathbf{c}_{ni}^{\mathrm{T}} \mathbf{u}), \qquad t \in [0, 1], \mathbf{u} \in R^p, \tag{1.8}$$

arises typically in the study of the asymptotic properties of $\hat{\boldsymbol{\beta}}_n$, where the $\mathbf{c}_{ni}$ are suitably normalized version of the $\mathbf{c}_i$. For example, we may set of every $n \leqslant p$, $\mathbf{C}_n = \sum_{i=1}^n \mathbf{c}_i \mathbf{c}_i^{\mathrm{T}} = (c_{nij})_{1 \leqslant i, j \leqslant p}$; $\mathbf{C}_n^0 = \mathrm{Diag}(c_{n11}^{1/2}, ..., c_{npp}^{1/2})$; $\mathbf{c}_{ni} = (\mathbf{C}_n^0)^{-1} \mathbf{c}_i = (c_{ni1}, ..., c_{nip})^{\mathrm{T}}$, $1 \leqslant i \leqslant n$. Thus letting $\mathbf{u} = \mathbf{C}_n^0(\boldsymbol{\theta} - \boldsymbol{\beta})$ and

$$\mathbf{M}_n(\mathbf{u}) = \sum_{i=1}^n \mathbf{c}_{ni} \psi(Y_i - \mathbf{c}_{ni}^{\mathrm{T}} \mathbf{u}), \tag{1.9}$$

we see that (1.7) is equivalent to

$$\mathbf{M}_n(\mathbf{u}) \equiv \mathbf{0} \qquad \text{(with respect to } \mathbf{u}\text{).} \tag{1.10}$$

The solution of the implicit (set of) equations is greatly facilitated by the following type of (Jurečková-)uniform asymptotic linearity for M-processes: for every finite real number $K > 0$, as $n \to \infty$,

$$\sup\{\|\mathbf{M}_n(\mathbf{u}) - \mathbf{M}_n(\mathbf{0}) + \mathbf{Q}_n \mathbf{u} \gamma\|; \|\mathbf{u}\| \leqslant K\} \xrightarrow{P} 0, \tag{1.11}$$

where $\|\cdot\|$ stands for the Euclidean norm, $\gamma = \int \psi' \, dF > 0$ and

$\mathbf{Q}_n = \sum_{i=1}^n \mathbf{c}_{ni}\mathbf{c}_{ni}^{\mathrm{T}} = (\mathbf{C}_n^0)^{-1}\mathbf{C}_n(\mathbf{C}_n^0)^{-1}$. Under various conditions on the $\{\mathbf{c}_{ni}\}$, the score function ψ and the d.f. F, (1.11) has been established (Jurečková [3]), and this provides an easy access to the study of the asymptotic properties of the M-estimator $\hat{\boldsymbol{\beta}}_n$. We consider a different approach here. For simplicity of presentation, we consider the case of $p=1$, i.e., the c_{ni} are real numbers. Since for each $u \in R$, $M_n(u)$ is a linear functional of $S_n^*(t, u)$, viz.,

$$M_n(t, u) = \int \psi(F^{-1}(t))\, dS_n^*(t, u),$$

$M_n(u)$ could be the Hadamard derivative of a certain functional τ. Thus, using the results of this paper, for a proper functional τ, we have, for any $K > 0$, as $n \to \infty$

$$\sup_{|u| \leqslant K} \left| \sum_{i=1}^n c_{ni}\, c_{ni} \left(\tau\left(\frac{S_n^*(\cdot, u)}{\sum_{i=1}^n c_{ni}}\right) - \tau\left(\frac{S_n^*(\cdot, 0))}{\sum_{i=1}^n c_{ni}}\right)\right) \right.$$
$$\left. - [M_n(u) - M_n(0)] \right| \xrightarrow{P} 0. \tag{1.12}$$

Therefore, (1.11) follows from showing

$$\sup_{|u| \leqslant K} \left| \sum_{i=1}^n c_{ni} \left(\tau\left(\frac{S_n^*(\cdot, u)}{\sum_{i=1}^n c_{ni}}\right) - \tau\left(\frac{S_n^*(\cdot, 0)}{\sum_{i=1}^n c_{ni}}\right)\right) + u\gamma \right| \xrightarrow{P} 0.$$

Some notations along with basic assumptions are presented in Section 2. In the same section, the notion of statistical functional and the concept of Hadamard differentiability are also introduced. The main results along with part of their derivations are considered in Section 3. The proof of Theorem 3.1 is given separately in Section 4.

2. Preliminary Notions

Consider the $D[0, 1]$ space (of right continuous real valued functions with left-hand limits) endowed with the Skorohod-J_1 (denoted by $\|\cdot\|_s$) topology. The space $C[0, 1]$ of real valued functions, endowed with the uniform (denoted by $\|\cdot\|$) topology, is a subspace of $D(0, 1]$. For every $u \in R$, denote by

$$S_n^*(t, u) = \sum_{i=1}^n c_{ni} I(Y_i \leqslant F^{-1}(t) + c_{ni}u), \qquad t \in [0, 1] \tag{2.1}$$

where Y_i are i.i.d. random variables with d.f. and c_{ni} are all given real numbers. It is easy to see that, for every $u(\in R)$, $S_n^*(\cdot, u)$ is an element of $D[0, 1]$. The population counterpart of the $I(Y_i \leqslant F^{-1}(t) + c_{ni}u)$ are the $F(F^{-1}(t) + c_{ni}u)$, and this leads us to consider the following:

$$S_n(t, u) = \sum_{i=1}^{n} c_{ni} F(F^{-1}(t) + c_{ni}u), \qquad t \in [0, 1], u \in R. \tag{2.2}$$

We also write

$$c_{ni} = c_{ni}^+ - c_{ni}^-; \qquad c_{ni}^+ = \max\{0, c_{ni}\}; \qquad c_{ni}^- = -\min\{0, c_{ni}\}; \tag{2.3}$$

$$S_n^{*+}(t, u) = \sum_{i=1}^{n} c_{ni}^+ I(Y_i \leqslant F^{-1}(t) + c_{ni}^+ u), \qquad t \in [0, 1], u \in R \tag{2.4}$$

$$S_n^{*-}(t, u) = \sum_{i=1}^{n} c_{ni}^+ c_{ni}^- I(Y_i \leqslant F^{-1}(t) + c_{ni}^- u), \qquad t \in [0, 1], u \in R \tag{2.5}$$

so that $S_n^*(t, u) = S_n^{*+}(t, u) - S_n^{*-}(t, u)$. Then let

$$W_n(t, u) = S_n^*(t, (2u-1)K) - S_n(t, (2u-1)K), \tag{2.6}$$

$$W_n^0(t, u) = S_n(t, (2u-1)K) - t \sum_{i=1}^{n} c_{ni}, \tag{2.7}$$

where K is a positive real number and $(t, u) \in [0, 1]^2$. Also, let f be a function defined on $[0, 1]^2$ and let us denote by

$$\omega_f(\delta) = \sup\{|f(t, u) - f(s, v)|; |t - s| \leqslant \delta, |u - v| \leqslant \delta\}. \tag{2.8}$$

Some assumptions, which may be required for our main results, are given below:

(A1) $c_{ni} \geqslant 0$, $i = 1, 2, ..., n$;

(A2) $\sum_{i=1}^n c_{ni}^2 = 1$, $\overline{\lim}_{n \to \infty} \max_{1 \leqslant i \leqslant n} c_{ni}^2 = 0$;

(B) F is absolutely continuous and has a positive and continuous derivative F' with limits at $\pm\infty$.

In order to prove our results in Section 3, some basic concepts about statistical functional and Hadamard differentiability are needed.

Definition. Let $X_1, ..., X_n$ be a sample from a population with d.f. F and let $T_n = T_n(X_1, ..., X_n)$ be a statistics. If T_n can be written as a functional T of the empirical d.f. F_n corresponding to the sample $X_1, ..., X_n$, i.e., $T_n = T(F_n)$, where T does not depend on n, then T will be called a *statistical functional*. The domain of the definition of T is assumed to contain the empirical d.f.'s F_n for all $n \geqslant 1$, as well as the population d.f. F, and the range of T will be the set of real numbers.

As we saw earlier, any statistical functional T induces a functional τ on $D[0, 1]$ by the relation given in (1.2). In Sections 3 and 4, we will always assume that functional τ is induced by a statistical functional T.

Let V and W be the topological vector spaces and $L(V, W)$ be the set of continuous linear transformation from V to W. Let $\mathscr{A}$ be an open set of V.

DEFINITION. A functional $T: \mathscr{A} \to W$ is *Hadamard differentiable* (*or compact differentiable*) at $F \in \mathscr{A}$ if there exists $T'_F \in L(V, W)$ such that for any compact set Γ of V,

$$\lim_{t \to 0} \frac{T(F+tH) - T(F) - T'_F(tH)}{t} = 0 \tag{2.9}$$

uniformly for any $H \in \Gamma$. The linear function T'_F is called the *Hadamard derivative* of T at F.

For our current study, we actually consider an extended statistical functional, i.e., the domain of the definition of T is assumed to contain $S_n^*(F(\cdot), u)/\sum_{i=1}^n c_{ni}$ for all $n \geqslant 1$ and $u \in R$, as well as the population d.f. F, and the Hadamard differentiability of the extended statistical functional is just the same as the definition above treating $S_n^*(F(\cdot), u)$ as an element of $D[0, 1]$ for a fixed $u \in R$.

For convenience sake, in (2.9), we usually denote

$$\operatorname{Rem}(tH) = T(F+tH) - T(F) - T'_F(tH),$$

then, correspondingly in Sections 3 and 4, we always use the notation

$$\operatorname{Rem}(tH) = \tau(U+tH) - \tau(U) - \tau'_U(tH),$$

where H is an arbitrary element of $D[0, 1]$.

3. MAIN RESULTS

THEOREM 3.1. *Suppose* $\tau: D[0, 1] \to R$ *is a functional and is Hadamard differentiable at* U. *Assume* (A1), (A2), *and* (B). *Then, for any* $K > 0$,

$$\sup_{|u| \leqslant K} \left| \sum_{i=1}^n c_{ni} \operatorname{Rem}\left(\frac{S_n^*(\cdot, u)}{\sum_{i=1}^n c_{ni}} - U(\cdot) \right) \right| \xrightarrow{P} 0, \qquad as \quad n \to \infty. \tag{3.1}$$

Therefore, we have, as $n \to \infty$

$$\sup_{|u| \leqslant K} \left| \sum_{i=1}^{n} c_{ni} \left\{ \tau\left(\frac{S_n^*(\cdot, u)}{\sum_{i=1}^{n} c_{ni}} \right) - \tau(U(\cdot)) \right\} - \tau'_U \left(S_n^*(\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{ni} \right) \right| \xrightarrow{P} 0. \tag{3.2}$$

Remark. We notice that (1.4) is just the special case of (3.1) for $c_{ni} = 1/n$, $1 \leqslant i \leqslant n$.

The proof of Theorem 3.1 will be given in Section 4. When (A1) is not satisfied, we have the following theorem.

THEOREM 3.2. *Suppose* $\tau: D[0, 1] \to R$ *is a functional and is Hadamard differentiable at U. Assume* (A2) *and* (B). *Then, for any* $K > 0$, *as* $n \to \infty$,

$$\sup_{|u| \leqslant K} \left| \sum_{i=1}^{n} c_{ni}^{+} \tau\left(\frac{S_n^{*+}(\cdot, u)}{\sum_{i=1}^{n} c_{ni}^{+}} \right) - \sum_{i=1}^{n} c_{ni}^{-} \tau\left(\frac{S_n^{*-}(\cdot, u)}{\sum_{i=1}^{n} c_{ni}^{-}} \right) - \tau(U) \sum_{i=1}^{n} c_{ni} - \tau'_U \left(S_n^*(\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{ni} \right) \right| \xrightarrow{P} 0. \tag{3.3}$$

Proof. Denote $\bar{c}_{ni}^{+} = c_{ni}^{+}/d_n^{+}$ and $\bar{c}_{ni}^{-} = \bar{c}_{ni}^{-}/d_n^{-}$ for $(d_n^{+})^2 = \sum_{i=1}^{n} (c_{ni}^{+})^2$ and $(d_n^{-})^2 = \sum_{i=1}^{n} (c_{ni}^{-})^2$. Since $\sum_{i=1}^{n} (\bar{c}_{ni}^{+})^2 = 1$ and $\sum_{i=1}^{n} (\bar{c}_{ni}^{-})^2 = 1$, therefore, for $\bar{c}_{ni}^{+}$ and $\bar{c}_{ni}^{-}$, (A1) and (A2) are satisfied, and $0 < d_n^{+}, d_n^{-} < 1$.

We observe, for $|u| \leqslant K$,

$$\frac{S_n^{*+}(t, u)}{\sum_{i=1}^{n} c_{ni}^{+}} = \frac{\sum_{i=1}^{n} \bar{c}_{ni}^{+} I(Y_i \leqslant F^{-1}(t) + c_{ni}^{+} u)}{\sum_{i=1}^{n} \bar{c}_{ni}^{+}} = \frac{\sum_{i=1}^{n} \bar{c}_{ni}^{+} I(Y_i \leqslant F^{-1}(t) + \bar{c}_{ni}^{+} u_1)}{\sum_{i=1}^{n} \bar{c}_{ni}^{+}} \qquad |u_1| \leqslant K.$$

Therefore, by Theorem 3.1, we have, for any $K > 0$, as $n \to \infty$,

$$\sup_{|u| \leqslant K} \left| \sum_{i=1}^{n} \bar{c}_{ni}^{+} \left\{ \tau\left(\frac{S_n^{*+}(\cdot, u)}{\sum_{i=1}^{n} c_{ni}^{+}} \right) - \tau(U(\cdot)) \right\} - \tau'_U \left(\frac{S_n^{*+}(\cdot, u)}{d_n^{+}} - U(\cdot) \sum_{i=1}^{n} \bar{c}_{ni}^{+} \right) \right| \xrightarrow{P} 0.$$

Since $0 < d_n^{+} < 1$, we have, as $n \to \infty$,

$$\sup_{|u| \leqslant K} \left| \sum_{i=1}^{n} c_{ni}^{+} \left\{ \tau\left(\frac{S_n^{*+}(\cdot, u)}{\sum_{i=1}^{n} c_{ni}^{+}} \right) - \tau(U(\cdot)) \right\} - \tau'_U \left(S_n^{*+}(\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{ni}^{+} \right) \right| \xrightarrow{P} 0. \tag{3.4}$$

Similarly, we can show that, as $n \to \infty$,

$$\sup_{|u| \leqslant K} \left| \sum_{i=2}^{n} c_{ni}^{-} \left\{ \tau\left(\frac{S_n^{*-}(\cdot, u)}{\sum_{i=1}^{n} c_{ni}^{-}} \right) - \tau(U(\cdot)) \right\} - \tau_U' \left(S_n^{*-}(\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{ni}^{-} \right) \right| \xrightarrow{P} 0. \tag{3.5}$$

Therefore, (3.3) follows from (3.4) and (3.5). ∎

Remark. As an application to M-estimators of regression, our Theorem 3.2 allows regression coefficients $\{\mathbf{c}_i\}$ to be positive or negative in order to establish (1.11) by the method given in Section 1.

4. Proof of Theorem 3.1

In this section, (A1), (A2), and (B) are assumed. First, we notice that (A1) and (A2) imply that

$$\sum_{i=1}^{n} c_{ni} \to \infty, \qquad \text{as} \quad n \to \infty. \tag{4.1}$$

We also notice that (B) implies that F' is bounded and uniformly continuous.

Due to the uniformity over u in result (3.1), the proof will be achieved through the bivariate version of $S_n^*(\cdot, \cdot)$. Unfortunately, $S_n^*(t, (2u-1)K) \notin D[0, 1]^2$ (Neuhaus [5]); therefore any existing results do not directly involve $S_n^*(\cdot, \cdot)$. We will deal with this problem through the bivariate smoothed version of S_n^* in our proof.

In order to see the difference between $S_n^*(\cdot, \cdot)$ and its smoothed version, we first will show that, with probability one, the biggest jump of $S_n^*(t, u)$ is no greater than $2 \max_{1 \leqslant i \leqslant n} c_{ni}$.

For any $0 < K < \infty$, consider

$$S_n^*(t, u) = \sum_{i=1}^{n} c_{ni} I(Y_i \leqslant F^{-1}(t) + c_{ni} u), \qquad t \in [0, 1], |u| \leqslant K.$$

For each i,

$$I(Y_i \leqslant F^{-1}(t) + c_{ni} u) = \begin{cases} 1 & \text{if} \quad F(Y_i - c_{ni} u) \leqslant t \\ 0 & \text{otherwise,} \end{cases}$$

the curve l_{ni}: $t = F(Y_i - c_{ni} u)$ is nonincreasing and continuous in u. Hence, for each $n \geqslant 1$ and $Y_1, ..., Y_n$, $[0, 1] \times [-K, K]$ is divided into finite pieces

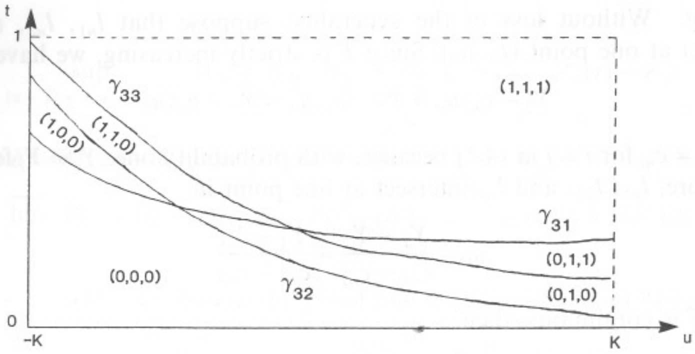

FIGURE 1

by smooth curves l_{ni}, $1 \leqslant i \leqslant n$, shown (for $n=3$) in Fig. 1 and the value of $S_n^*(t, u)$ is a constant in each different pice, or region:

region(0, 0, 0): $S_n^*(t, u) = 0$; region(0, 1, 0): $S_n^*(t, u) = c_{n2}$;

region(0, 1, 1): $S_n^*(t, u) = c_{n2} + c_{n3}$; region(1, 0, 0): $S_n^*(t, u) = c_{n1}$;

region(1, 1, 0): $S_n^*(t, u) = c_{n1} + c_{n2}$; region(1, 1, 1): $S_n^*(t, u) = c_{n1} + c_{n2} + c_{n3}$.

If l_{31}, l_{32}, l_{33} intersect at one point shown as Fig. 2, the biggest jump of $S_3^*(t, u)$ is $c_{31} + c_{32} + c_{33}$. However, Lemma 4.1 shows that, with probability one, no more than two curves will intersect at one point.

LEMMA 4.1. *For each n, no more than two l's intersect at one point in region* $[0, 1] \times [-K, K]$ *with probability one.*

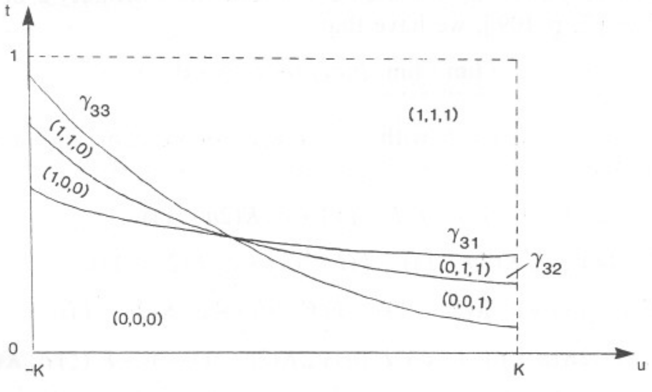

FIGURE 2

Proof. Without loss of the generality, suppose that l_{n1}, l_{n2}, and l_{n3} intersect at one point (t_0, u_0). Since F is strictly increasing, we have

$$Y_1 - c_{n1} u_0 = Y_2 - c_{n2} u_0 = Y_3 - c_{n3} u_0 \tag{4.2}$$

and $c_{ni} \neq c_{nj}$ for $i \neq j$ in (4.2) because, with probability one, $Y_i \neq Y_j$ for $i \neq j$. Therefore, l_{n1}, l_{n2}, and l_{n3} intersect at one point iff

$$u_0 = \frac{Y_1 - Y_2}{c_{n1} - c_{n2}} = \frac{Y_1 - Y_3}{c_{n1} - c_{n3}}.$$

Since F is continuous, then

$$P\left\{\frac{Y_1 - Y_2}{c_{n1} - c_{n2}} = \frac{Y_1 - Y_3}{c_{n1} - c_{n3}}\right\} = P\left\{Y_2 = \frac{c_{n2} - c_{n3}}{c_{n1} - c_{n3}}\, Y_1 + \frac{c_{n1} - c_{n2}}{c_{n1} - c_{n3}}\, Y_3\right\} = 0.$$

For each n, the probability that more than two l's intersect at one point only depends on $\{c_{n1}, c_{n2}, ..., c_{nn}\}$. Hence, with probability one, no more than two l's intersect at one point for each $n \geqslant 1$. ∎

Lemma 4.1 implies that, with probability one, no more than two l's intersect at one point along $\{l_{n1}, ..., l_{nn}; n \geqslant 1\}$. Therefore, with probability one, the largest jump of $S_n^*(t, u)$ is no larger than $2 \max_{1 \leqslant i \leqslant n} c_{ni}$. Let $\bar{S}_n^*(t, u)$ (obtained by smoothing $S_n^*(t, u)$ through the regions shown in Fig. 1.) be a bivariate continuous version of $S_n^*(t, u)$; then $\bar{S}_n^*(t, (2u-1)K)$ is an element of $C[0, 1]^2$ and

$$\|\bar{S}_n^*(\cdot, \cdot) - S_n^*(\cdot, \cdot)\| \leqslant 2 \max_{1 \leqslant i \leqslant n} c_{ni}, \qquad \text{a.s.} \tag{4.3}$$

LEMMA 4.2. *For any* $\varepsilon > 0$, $\lim_{\delta \to 0} \overline{\lim}_{n \to \infty} P(\omega_{W_n}(\delta) \geqslant \varepsilon) = 0$.

Proof. Consider a weighted empirical process $Z_n(t) = \sum_{i=1}^n c_{ni}[I(F(Y_i) \leqslant t) - t]$, for $0 \leqslant t \leqslant 1$. From the Corollary 2 of Shorack and Wellner [7, p. 109], we have that

$$\lim_{\delta \to 0} \overline{\lim_{n \to \infty}} P(\omega_{Z_n}(\delta) \geqslant \varepsilon) = 0. \tag{4.4}$$

Since, for $0 \leqslant x, y, u, v \leqslant 1$ with $|x - y| \leqslant \delta$, $|u - v| \leqslant \delta$, and a constant $M \geqslant 1$, by (B),

$$\begin{aligned}
&|F(F^{-1}(x) + c_{ni}K(2u-1)) - F(F^{-1}(y) + c_{ni}K(2n-1))| \\
&\quad \leqslant |F(F^{-1}(x) + c_{ni}K(2u-1)) - F(F^{-1}(x) + c_{ni}K(2v-1))| \\
&\qquad + |F(F^{-1}(x) + c_{ni}K(2v-1)) - F(F^{-1}(y) + c_{ni}K(2v-1))| \\
&\quad = |F'(\xi)\, c_{ni} 2K(u-v)| + |x + F'(\eta)\, c_{ni}K(2v-1) - y - F'(\zeta)\, c_{ni}K(2v-1)| \\
&\quad \leqslant M\delta + \delta + KM \max_{1 \leqslant i \leqslant n} |c_{ni}|,
\end{aligned}$$

then

$$\sup_{|x-y|\leqslant\delta, |u-v|\leqslant\delta} |W_n(x,u)-W_n(y,v)| \leqslant \sup_{|t-s|\leqslant\delta'} |Z_n(t)-Z_n(s)|,$$

where $\delta' = M(2\delta + K\max_{1\leqslant i\leqslant n}|c_{ni}|)$. By (A2), we have

$$\overline{\lim_{n\to\infty}}\, P(\omega_{W_n}(\delta)\geqslant\varepsilon) \leqslant \overline{\lim_{n\to\infty}}\, P(\omega_{Z_n}(\delta')\geqslant\varepsilon) \leqslant \overline{\lim_{n\to\infty}}\, P(\omega_{Z_n}(\delta'')\geqslant\varepsilon), \quad (4.5)$$

where $\delta'' = 3M\delta$. Therefore, the proof follows from (4.4) and (4.5). ∎

PROPOSITION 4.3. *Let* $T_n(t,u) = \bar{S}_n^*(t,(2u-1)K) - t\sum_{i=1}^n c_{ni}$, $(t,u)\in [0,1]^2$, *and let* $\{P_n; n\geqslant 1\}$ *be the sequence of probability measures corresponding to* T_n, $n\geqslant 1$. *Then,* $\{P_n\}$ *is relatively compact.*

Proof. Note that $T_n((t,u)\in C[0,1]^2$ and $S_n^*(0,-K)\equiv 0$, for $n\geqslant 1$. By (4.3), we have

$$|T_n(0,0)| = |\bar{S}_n^*(0,-K)| \leqslant 2\max_{1\leqslant i\leqslant n} c_{ni}, \qquad \text{a.s.}$$

Hence, $P_n^0 = P\pi_0^{-1}$ converges in distribution. By virtue of Neuhaus [5] (discussion on pp. 1290–1291), it suffices to show that any $\varepsilon > 0$,

$$\lim_{\delta\to 0}\overline{\lim_{n\to\infty}}\, P(\omega_{T_n}(\delta)\geqslant\varepsilon) = 0. \quad (4.6)$$

Since for any (t,u), $(s,v)\in[0,1]^2$,

$$\begin{aligned}
&|T_n(t,u)-T_n(s,v)| \\
&\quad = \left|[\bar{S}_n^*(t,(2u-1)K) - \bar{S}_n^*(s,(2v-1)K)] - (t-s)\sum_{i=1}^n c_{ni}\right| \\
&\quad \leqslant |[\bar{S}_n^*(t,(2u-1)K) - \bar{S}_n^*(s,2v-1)K)] \\
&\qquad - [S_n^*(t,(2u-1)K) - S_n^*(s,(2v-1)K)]| \\
&\qquad + |W_n(t,u) - W_n(s,v)| + |W_n^0(t,u) - W_n^0(s,v)|
\end{aligned}$$

by (4.3), we have

$$\omega_{T_n}(\delta) \leqslant 4\max_{1\leqslant i\leqslant n} c_{ni} + \omega_{W_n}(\delta) + \omega_{W_n^0}(\delta), \qquad \text{a.s.}$$

Hence, by (A2) and Lemma 4.2, (4.6) follows from showing

$$\omega_{W_n^0}(\delta)\to 0, \qquad \text{as}\quad n\to\infty,\ \delta\to 0. \quad (4.7)$$

Let $G(t, u) = (2u - 1)\, KF'(F^{-1}(t))$, then, by $\sum_{i=1}^{n} c_{ni}^2 = 1$,

$$\begin{aligned}
&\sup_{(t,u) \in [0,1]^2} |W_n^0(t, u) - G_n(t, u)| \\
&= \sup_{t \in [0,1], |u| \leqslant K} \left| \sum_{i=1}^{n} c_{ni} [F(F^{-1}(t) + v_{ni} u) - t] - uF'(F^{-1}(t)) \right| \\
&= \sup_{t \in [0,1], |u| \leqslant K} \left| u \sum_{i=1}^{n} c_{ni}^2 (F'(\xi_{ni}) - F'(F^{-1}(t))) \right| \\
&\leqslant K \sum_{i=1}^{n} c_{ni}^2 \sup_{t \in [0,1]} |F'(\xi_{ni}) - F'(F^{-1}(t))|,
\end{aligned}$$

where ξ_{ni} is between $F^{-1}(t)$ and $F^{-1}(t) + c_{ni} u$. By (A2) and the uniform continuity of F', we have

$$\sup_{(t,u) \in [0,1]^2} |W_n^0(t, u) - G_n(t, u)| \to 0, \qquad \text{as } n \to \infty. \tag{4.8}$$

Note that $G \in C[0, 1]^2$ and $\omega_{W_n^0}(\delta) \leqslant 2 \| W_n^0 - G \| + \omega_G(\delta)$, then (4.7) follows from (4.8) and the uniform continuity of G on $[0, 1]^2$. ∎

Let Γ be a set in $D[0, 1]$ and $H \in D[0, 1]$; define

$$\operatorname{dist}(H, \Gamma) = \inf_{G \in \Gamma} \|H - G\|. \tag{4.9}$$

LEMMA 4.4. *Let Q: $D[0, 1] \times R \to R$ and suppose that for any compact set Γ in $D[0, 1]$,*

$$\lim_{t \to 0} Q(H, t) = 0 \tag{4.10}$$

uniformly for $H \in \Gamma$. Let $\varepsilon > 0$ and let α_n, β_n be sequences of real numbers such that $\alpha_n \to 0$, $\beta_n \to 0$, as $n \to \infty$. Then, for any compact set Γ in $D[0, 1]$, there exists a positive integer N such that, if $\operatorname{dist}(H, \Gamma) \leqslant \alpha_n$, then

$$|Q(H, \beta_n)| < \varepsilon, \qquad \text{for } n \geqslant N.$$

Proof. Suppose not. Then, for a real number $\varepsilon > 0$, there exists a compact set Γ in $D[0, 1)$ and sequence $\{H_k\} \subset D[0, 1]$ with $\operatorname{dist}(H_k, \Gamma) \leqslant \alpha_{n_k}$ such that

$$|Q(H_k, \beta_{n_k})| \geqslant \varepsilon. \tag{4.11}$$

Since $\operatorname{dist}(H_k, \Gamma) \leqslant \alpha_{n_k}$, we can choose $H_k^* \in \Gamma$ such that $\|H_k - H_k^*\| \leqslant \alpha_{n_k}$. Since $\{H_k^*\} \subset \Gamma$ and Γ is a compact set, $\{H_k^*\}$ has an accumulation point $H^* \in \Gamma$. Therefore, we can choose a subsequence of

$\{H_k^*\}$ also denoted by $\{H_k^*\}$ such that $H_k^* \to H^*$, as $k \to \infty$. Since $\alpha_{n_k} \to 0$, we also have $H_k \to H^*$, as $k \to \infty$, and the set $\Gamma_1 = \{H_k; k \geqslant 1\} \cup \{H^*\}$ is compact. By (4.10), we have $Q(H_k, t) \to 0$, as $t \to 0$, uniformly for $H_k \in \Gamma_1$. This contradicts (4.11). ∎

Proof of Theorem 3.1. By Proposition 4.3, $\{P_n\}$ is relatively compact in $C[0, 1]^2$, where $P_n(A) = P(T_n \in A)$. Since $C[0, 1]^2$ is complete and separable, by Prohorov's theorem (Billingsley [1, Theorem 6.2]), $\{P_n\}$ is tight, i.e., for any $\varepsilon > 0$, there exists a compact set Γ in $C[0, 1]^2$ such that $PT_n \in \Gamma) > 1 - \varepsilon$, $n \geqslant 1$. By (4.3), we have

$$P\{T_n \in \Gamma, \|\bar{S}_n^*(\cdot, \cdot) - S_n^*(\cdot, \cdot)\| \leqslant 2 \max_{1 \leqslant i \leqslant n} c_{ni}\} \geqslant 1 - \varepsilon, \qquad \text{for} \quad n \geqslant 1. \tag{4.12}$$

Let $\Gamma_1 = \{T_n(\cdot, u); T_n \in \Gamma, u \in [0, 1]\}$, then Γ_1 is a compact set in $C[0, 1]$ and is also a compact set in $D[0, 1]$, because $C[0, 1]$ is a subspace of $D[0, 1]$. Since $T_n \in \Gamma$ implies $T_n(\cdot, u) \in \Gamma_1$ for any $u \in [0, 1]$, i.e.,

$$\left[\bar{S}_n^*(\cdot, (2u-1)K) - U(\cdot) \sum_{i=1}^{n} c_{ni}\right] \in \Gamma_1, \qquad \text{for any} \quad u \in [0, 1], \tag{4.13}$$

and, since $\|\bar{S}_n^*(\cdot, \cdot) - S_n^*(\cdot, \cdot)\| \leqslant 2 \max_{1 \leqslant i \leqslant n} c_{ni}$ implies

$$\begin{aligned} &\|\bar{S}_n^*(\cdot, (2u-1)K) - S_n^*(\cdot, (2u-1)K)\| \\ &\qquad \leqslant 2 \max_{1 \leqslant i \leqslant n} c_{ni}, \quad \text{for any } u \in [0, 1], \end{aligned} \tag{4.14}$$

then, by (4.12) and the fact (4.13) and (4.14) imply

$$\operatorname{dist}\left(\left[S_n^*(\cdot, (2u-1)K) - U(\cdot) \sum_{i=1}^{n} c_{ni}\right], \Gamma_1\right) \leqslant 2 \max_{1 \leqslant i \leqslant n} c_{ni}, \qquad \forall u \in [0, 1],$$

we have, for $n \geqslant 1$,

$$\begin{aligned} P\Bigg\{&\operatorname{dist}\left(\left[S_n^*(\cdot, (2u-1)K) - U(\cdot) \sum_{i=1}^{n} c_{ni}\right], \Gamma_1\right) \\ &\geqslant 2 \max_{1 \leqslant i \leqslant n} c_{ni}, \forall u \in [0, 1]\Bigg\} > 1 - \varepsilon. \end{aligned} \tag{4.15}$$

Since $\tau: D[0, 1] \to R$ is Hadamard differentiable at U, by the definition of Hadamard differentiability, (4.10) holds for $Q(H, t) = \operatorname{Rem}(tH)/t$. By Lemma 4.4, (4.1), and (A2), for the above compact set Γ_1, there exists a positive integer N such that, for $n \geqslant N$ if $\operatorname{dist}(H, \Gamma_1) \leqslant 2 \max_{1 \leqslant i \leqslant n} c_{ni}$, then

$$\left|\sum_{i=1}^{n} c_{nik} \operatorname{Rem}\left(\frac{H}{\sum_{i=1}^{n} c_{nik}}\right)\right| < \varepsilon.$$

Therefore, taking $H = [S_n^*(\cdot, (2u-1)K) - U(\cdot) \sum_{i=1}^n c_{ni}]$ for $n \geqslant N$ and $u \in [0, 1]$, we have that $\{\text{dist}([S_n^*(\cdot, (2u-1)K) - U(\cdot) \sum_{i=1}^n c_{ni}], \Gamma_1) \leqslant 2 \max_{1 \leqslant i \leqslant n} c_{ni}, \forall u \in [0, 1]\}$ implies

$$\left| \sum_{i=1}^n c_{ni} \operatorname{Rem} \left(\frac{S_n^*(\cdot, (2u-1)K)}{\sum_{i=1}^n c_{ni}} - U(\cdot) \right) \right| < \varepsilon, \quad \text{for } u \in [0, 1]. \quad (4.16)$$

Since (4.16) implies

$$\sup_{|u| \leqslant K} \left| \sum_{i=1}^n c_{ni} \operatorname{Rem} \left(\frac{S_n^*(\cdot, u)}{\sum_{i=1}^n c_{ni}} - U(\cdot) \right) \right| \leqslant \varepsilon,$$

by (4.15) we have, for $n \geqslant N$,

$$1 - \varepsilon < P \left\{ \sup_{|u| \leqslant K} \left| \sum_{i=1}^n c_{ni} \operatorname{Rem} \left(\frac{S_n^*(\cdot, u)}{\sum_{i=1}^n c_{ni}} - U(\cdot) \right) \right| \leqslant \varepsilon \right\}. \blacksquare$$

Remark. If in Theorem 3.1,

$$\sup_{|u| \leqslant K} \left| \sum_{i=1}^n c_{ni} \operatorname{Rem} \left(\frac{S_n^*(\cdot, u)}{\sum_{i=1}^n c_{ni}} - U(\cdot) \right) \right|$$

is not measurable, we replace $|u| \leqslant K$ by $u \in Q_K$, where $Q_K = \{\text{all rational numbers in } [-K, K]\}$; then, it is measurable. Our main results in Section 3 will be slightly different, but still good enough for the study of (1.11).

ACKNOWLEDGMENTS

The authors are grateful to the referees for their useful comments on the manuscripts.

REFERENCES

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

[2] FERNHOLZ, L. T. (1983). *Von Mises Calculus for Statistical Functional*. Lecture Notes in Statistics, Vol. 19. Springer-Verlag, New York.

[3] JUREČKOVÁ, J. (1984). *M-*, *L-*, and *R*-estimators. In *Handbook of Statistics*, Vol. 4: *Nonparametric Methods* (P. R. Krishnaiah and P. K. Sen, Eds.), pp. 463–485, North-Holland, Amsterdam.

[4] KALLIANPUR, G. (1963). Von Mises function and maximum likelihood estimation. *Sankhyā A* **23** 149–158.

[5] NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Amer. Math. Soc.* **42** 1285–1295.

[6] REEDS, J. A. (1976). *On the Definition of Von Mises Functionals*. Ph.D. dissertation. Harvard University.

[7] SHORACK, G., AND WELLNER, J. A. (1986). *Empirical Processes with Application to Statistics*. Wiley, New York.

[8] VON MISES, R. (1947). On the asymptotic distributions of differentiable statistical functions. *Amer. Math. Soc.* **18** 309–348.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium