Statistica Sinica 1(1991), 185-202

SECOND-ORDER DIFFERENTIABILITY AND JACKKNIFE

Jun Shao

University of Ottawa

Abstract: We establish some asymptotic results for the statistic $T(F_n)$ and the corresponding jackknife estimators, where F_n is the empirical distribution and T is a second-order differentiable functional in some sense. In particular, second-order asymptotic representations of $T(F_n)$ and the jackknife estimator are obtained and the jackknife variance and bias estimators are shown to be consistent.

Key words and phrases: Asymptotic representation, bias, consistency, statistical functional, variance.

1. Introduction

Let $X_1, \ldots, X_n$ be independent and identically distributed (i.i.d.) random samples from an unknown population distribution F and let F_n be the corresponding empirical distribution, i.e., $F_n(x) = n^{-1} \sum_i I_{X_i}(x)$, where $I_y(x) = 1$ if $y \leq x$ and $= 0$ otherwise. In many situations the parameter of interest is $T(F)$, where T is a functional defined on a set $\mathbf{N}$ (to be specified in Section 2). A nonparametric estimator of $T(F)$ is $T(F_n)$. Under weak conditions,

$$n^{1/2}[T(F_n) - T(F)] \to N(0, \sigma^2) \quad \text{in distribution,}$$

where σ^2 is usually unknown and σ^2/n is called the asymptotic variance of $T(F_n)$. Let B_n denote the bias or the asymptotic bias of $T(F_n)$ as an estimator of $T(F)$ (see Section 3). To assess the performance of $T(F_n)$ or make other statistical inferences, we need estimators of σ^2/n and B_n. An essential requirement of variance and bias estimators is their weak consistency. That is, if ν_n and b_n are estimators of σ^2/n and B_n, respectively, then we require

$$n\nu_n - \sigma^2 \to_p 0 \quad \text{and} \quad n(b_n - B_n) \to_p 0,$$

where $\to_p$ denotes convergence in probability.

The jackknife, originally introduced by Quenouille (1956) for bias reduction, is a convenient nonparametric method for variance and bias estimation and for improving the estimator $T(F_n)$ in some situations. For each i, let $F_n^{(i)}$ be the empirical distribution corresponding to the samples $X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n$. The jackknife estimators of σ^2/n and B_n are respectively

$$\nu_n^J = \frac{n-1}{n}\sum_{i=1}^n [T(F_n^{(i)}) - \overline{T}]^2, \quad \overline{T} = \frac{1}{n}\sum_{i=1}^n T(F_n^{(i)}),$$

and

$$b_n^J = (n-1)[\overline{T} - T(F_n)].$$

The jackknife estimator of $T(F)$, designed to reduce the bias in estimating $T(F)$, is

$$T^J(F_n) = T(F_n) - b_n^J.$$

The asymptotic validity of the jackknife (e.g., the weak consistency of ν_n^J and b_n^J and performance of $T^J(F_n)$) was established in some special situations. Beran (1984) and Parr (1985) used a differentiable functional approach to establish a general result. However, the differentiability assumption on T they assumed is too strong and does not hold for some commonly used statistics. The main purpose of this note is to establish properties of the jackknife estimators under a weak differentiability assumption on T. After introducing several versions of (second-order) differentiability of T in Section 2, we study asymptotic properties of $T(F_n)$ in Section 3 and prove the weak consistency of the jackknife estimators ν_n^J and b_n^J in Sections 4 and 5. In Section 6, asymptotic properties of the jackknife estimator $T^J(F_n)$ are established, which show that $T^J(F_n)$ reduces the bias in some sense. The last section contains some examples.

2. Second-Order Differentiability

Let $\mathbf{F} = \{$all distributions on the real line $\mathbf{R}\}$ and $\|\cdot\|$ be a norm defined on the space $\mathbf{F}^\Delta = \{c(G-K) : c \in \mathbf{R}, G, K \in \mathbf{F}, \|G-K\| < \infty\}$. Some useful norms are the sup-norm $\|G-K\|_\infty = \sup_x |G(x) - K(x)|$ and the L_p-norm $(p \geq 1)$ $\|G-K\|_{L_p} = [\int |G(x) - K(x)|^p dx]^{1/p}$. Let $\mathbf{N} = \{G \in \mathbf{F} : \|G-K\| < \infty$ for any $K \in \mathbf{F}\}$. Suppose that $\mathbf{V}$ is a normed vector space generated by $\|\cdot\|$ and $\mathbf{N} \subset \mathbf{V}$. We assume $F \in \mathbf{N}$ and the functional T is defined on $\mathbf{N}$.

Definition 2.1. A functional f defined on $\mathbf{N}$ is *second-order Fréchet differentiable at $H \in \mathbf{N}$ with respect to* (w.r.t.) *a norm* $\|\cdot\|$ if there is a real-valued function $\phi_f(x,y,H)$ on $\mathbf{R}^2$ such that $\phi_f(x,y,H) = \phi_f(y,x,H)$, $\iint \phi_f(x,y,H)dH(x)dH(y) = 0$ and

$$\lim_{\|G-H\|\to 0, G\in\mathbf{N}} \frac{f(G) - f(H) - \iint \phi_f(x,y,H)dG(x)dG(y)}{\|G-H\|^2} = 0.$$

The differentiability of f according to Definition 2.1 is called second-order since $q(G) = \iint \phi_f(x,y,H)dG(x)dG(y)$ is a quadratic functional. f is second-order differentiable at H simply means that $f(G) - f(H)$ can be approximated by a quadratic functional for G near H (w.r.t. the topology generated by $\|\ \|$) and that the rate of the approximation is $o(\|G - H\|^2)$. If $f(G) - f(H)$ can be approximated by a linear functional $l(G)$ for G near H, i.e.,

$$\lim_{\|G-H\|\to 0, G\in \mathbf{N}} \frac{f(G) - f(H) - l(G)}{\|G - H\|} = 0,$$

then f is first-order Fréchet differentiable at H, which is a well understood concept in the literature. See Serfling (1980) for a detailed discussion of the first-order differentiability of a functional. For the study of the jackknife bias estimator b_n^J, however, assuming the first-order differentiability of T is not enough, since the bias of $T(F_n)$ is directly related to the second-order differential of T (see Theorems 3.1 and 3.2). Also, the first-order (Fréchet or compact) differentiability of T does not ensure the consistency of the jackknife variance estimator ν_n^J. Parr (1985) proved the consistency of ν_n^J under a first-order strong Fréchet differentiability assumption on T, which is much stronger than the first-order Fréchet differentiability (in the ordinary sense) and is not comparable with the second-order Fréchet differentiability in Definition 2.1.

Beran (1984) proved the weak consistency of ν_n^J under the conditions that T is second-order Fréchet differentiable at F w.r.t. the sup-norm $\|\cdot\|_\infty$ with $\phi_T(x,y,F)$ satisfying

$$\iint \phi_T^2(x,y,F)dF(x)dF(y) < \infty \quad \text{and} \quad \int |\phi_T(x,x,F)|dF(x) < \infty \tag{2.1}$$

and that

$$\sup_{G\in \mathbf{N}} \left| \frac{T(G) - T(F) - \iint \phi_T(x,y,F)dG(x)dG(y)}{\|G - F\|_\infty^2} \right| < \infty. \tag{2.2}$$

Condition (2.2), however, is unnecessarily restrictive. Under the assumption that T is second-order Fréchet differentiable at F w.r.t. $\|\cdot\|_\infty$ with $\phi_T(x,y,F)$ satisfying (2.1), the weak consistency of ν_n^J was proved in Shao and Wu (1989). In Section 4, we will prove the weak consistency of ν_n^J when T is differentiable in a different and/or weaker sense.

Definition 2.2. Let $\mathbf{S}$ be a class of subsets of $\mathbf{V}$. A functional f defined on $\mathbf{N}$ is *second-order* $\mathbf{S}$*-differentiable at* $H \in \mathbf{N}$ w.r.t. $\|\cdot\|$ if there is a real-valued function

$\phi_f(x,y,H)$ such that $\phi_f(x,y,H) = \phi_f(y,x,H)$, $\iint \phi_f(x,y,H)dH(x)dH(y) = 0$ and for any $\Lambda \in \mathbf{S}$,

$$\lim_{t\to 0} \frac{f(H+tK) - f(H) - Q(tK)}{t^2} = 0 \tag{2.3}$$

uniformly for all $K \in \Lambda$ and $H + tK \in \mathbf{N}$, where $Q(tK) = \iint \phi_f(x,y,H)\, d[H+tK](x)d[H+tK](y)$.

If $\mathbf{S}$ is chosen to be $\mathbf{B}$={bounded subsets of $\mathbf{V}$}, Proposition 2.1 shows that $\mathbf{B}$-differentiability (Definition 2.2) is equivalent to Fréchet differentiability (Definition 2.1). If $\mathbf{S}$ is chosen to be $\mathbf{C}$={compact subsets of $\mathbf{V}$}, $\mathbf{C}$-differentiability is referred to as compact differentiability. Therefore, compact differentiability is weaker than Fréchet differentiability.

Proposition 2.1. *A functional f on* $\mathbf{N}$ *is second-order* $\mathbf{B}$*-differentiable at H* w.r.t. $\|\cdot\|$ *if and only if f is second-order Fréchet differentiable at H* w.r.t. $\|\cdot\|$.

Proof. (i) $\mathbf{B}$-differentiability implies Fréchet differentiability. Suppose that $G \in \mathbf{N}$ and $\|G - H\| \to 0$. Let $\Lambda = \{K/\|K\| : K \in \mathbf{V}\}$. Then $\Lambda \in \mathbf{B}$ and $(G-H)/\|G-H\| \in \Lambda$. The result follows by replacing t and K in (2.3) by $\|G-H\|$ and $(G-H)/\|G-H\|$, respectively.
(ii) Fréchet differentiability implies $\mathbf{B}$-differentiability. For any $\Lambda \in \mathbf{B}$, there is a constant c such that $\|K\| \le c$ for $K \in \Lambda$. For all $K \in \Lambda$ with $G = H + tK \in \mathbf{N}$, as $t \to 0$,

$$\frac{f(H+tK) - f(H) - Q(tK)}{t^2} \le c^2 \frac{f(G) - f(H) - Q(G-H)}{\|G-H\|^2} \to 0.$$

When $\|\cdot\|$ is the sup-norm and F is continuous, it is convenient to consider the functional

$$\tau(G) = T(G \circ F), \quad G \in \mathbf{N}_\tau, \tag{2.4}$$

where $\mathbf{N}_\tau$={ all distributions on $[0,1]$ }. Note that $\mathbf{N}$=$\mathbf{F}$ and $K \circ F \in \mathbf{N}$ if $K \in \mathbf{N}_\tau$. Hence $T(K \circ F)$ is well defined. Let $\mathbf{D}$ be the space of right continuous real-valued functions on $[0,1]$ which have left hand limits. Definitions 2.1 and 2.2 are applicable to τ with $\mathbf{N}$ and $\mathbf{V}$ replaced by $\mathbf{N}_\tau$ and $\mathbf{D}$ (equipped with $\|\cdot\|_\infty$), respectively. Let U be the uniform distribution function on $[0,1]$. Then $\tau(U) = T(F)$. The result in Proposition 2.2 indicates a relation between the differentiability of T at F and the differentiability of τ at U.

Proposition 2.2. *If T is second-order Fréchet differentiable at F* w.r.t. $\|\cdot\|_\infty$ *and F is continuous, then τ is second-order Fréchet and compactly differentiable at U (uniform distribution on* $[0,1]$ *)* w.r.t. $\|\cdot\|_\infty$.

Proof. From Definition 2.1, there is a function $\phi_T(x,y,F)$ such that for $G \in \mathbf{N}$,

$$R_T(G,F)/\|G-F\|_\infty^2 \to 0 \tag{2.5}$$

as $\|G-F\|_\infty \to 0$, where $R_T(G,F) = T(G)-T(F) - \iint \phi_T(x,y,F)dG(x)dG(y)$. Define $\phi_\tau(s,t,U) = \phi_T(F^{-1}(s), F^{-1}(t), F)$, where $F^{-1}(t) = \inf\{x : F(x) \geq t\}$, and

$$R_\tau(K,U) = \tau(K) - \tau(U) - \iint \phi_\tau(s,t,U)dK(s)dK(t), \quad K \in \mathbf{N}_\tau. \tag{2.6}$$

If $K \in \mathbf{N}_\tau$, then $K \circ F \in \mathbf{N}$ and $R_\tau(K,U) = R_T(K \circ F, F)$. Then the result follows from (2.5) and $\|K \circ F - F\|_\infty = \|K \circ F - U \circ F\|_\infty \leq \|K-U\|_\infty$.

3. A Second-Order Representation for $T(F_n)$

In this section we establish a second-order representation (3.3)-(3.4) for $T(F_n)$.

Theorem 3.1. *Let τ be defined in (2.4). Suppose that F is continuous and that τ is second-order compactly differentiable at U (uniform distribution on $[0,1]$) w.r.t. $\|\cdot\|_\infty$ with $\phi_\tau(s,t,U)$ satisfying*

$$\iint \phi_\tau^2(s,t,U)dsdt < \infty \quad \textit{and} \quad \int |\phi_\tau(t,t,U)|dt < \infty. \tag{3.1}$$

(i) *Let $\sigma^2 = 4\int[\int \phi_\tau(s,t,U)dt]^2 ds$. Then*

$$n^{1/2}[T(F_n) - T(F)] \to N(0,\sigma^2) \quad \textit{in distribution.} \tag{3.2}$$

(ii) *There are* i.i.d. *random variables ξ_i satisfying $E_F\xi_i = 0$ and $E_F\xi_i^2 = \sigma^2$ such that*

$$T(F_n) = T(F) + \frac{1}{n}\sum_{i=1}^n \xi_i + \Gamma_n \tag{3.3}$$

with

$$n\Gamma_n \to \sum_{k=1}^\infty \lambda_k Y_k \quad \textit{in distribution,} \tag{3.4}$$

where λ_k are real numbers and Y_k are independent χ_1^2 random variables.
(iii) *Let $\gamma = \int \phi_\tau(t,t,U)dt$. Then*

$$\gamma = \sum_{k=1}^\infty \lambda_k = E\Big(\sum_{k=1}^\infty \lambda_k Y_k\Big). \tag{3.5}$$

If, in addition, (2.2) *holds with* T *and* F *replaced by* τ *and* U*, respectively, then*

$$nE_F[T(F_n) - T(F)] \to \gamma. \tag{3.6}$$

Proof. Let $Z_i = F(X_i)$, $i = 1, \ldots, n$, and U_n be the empirical distribution corresponding to $Z_1, \ldots, Z_n$. Since F is continuous, Z_i are i.i.d. from the uniform distribution U, $F_n = U_n \circ F$ and $T(F_n) - T(F) = \tau(U_n) - \tau(U)$. From the second-order compact differentiability of τ,

$$\begin{aligned} T(F_n) - T(F) &= \iint \phi_\tau(s,t,U)dU_n(s)dU_n(t) + R_\tau(U_n,U) \\ &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\phi_\tau(Z_i,Z_j,U) + R_\tau(U_n,U) \end{aligned}$$

and therefore,

$$T(F_n) - T(F) = \frac{1}{n}\sum_{i=1}^{n}\xi_i + \frac{1}{n^2}\sum_{1\le i,j\le n} h(Z_i,Z_j,U) + R_\tau(U_n,U), \tag{3.7}$$

where

$$\xi_i = 2\int \phi_\tau(Z_i,t,U)dt, \quad h(s,t,U) = \phi_\tau(s,t,U) - \int[\phi_\tau(s,u,U) + \phi_\tau(u,t,U)]du$$

and R_τ is given in (2.6). As a direct consequence of result (4.4) in Section 4,

$$R_\tau(U_n,U) = o_p(n^{-1}). \tag{3.8}$$

Note that $E_F\xi_i = 0$ and $E_F\xi_i^2 = \sigma^2$. Then (3.3)-(3.4) follows by applying Theorem 6.4.1B in Serfling (1980) to $h(s,t,U)$, and (3.2) follows from (3.3)-(3.4) and the central limit theorem. From the remark on p.227 of Serfling (1980), (3.5) holds. Taking expectation on both sides of (3.7), we obtain that

$$E_F[T(F_n) - T(F)] = n^{-1}\gamma + E_F[R_\tau(U_n,U)].$$

Hence (3.6) follows from

$$E_F[R_\tau(U_n,U)] = o(n^{-1}).$$

From (2.2), $n|R_\tau(U_n,U)| \le n\|U_n - U\|_\infty^2$ for all n. From Dvoretzky, Kiefer and Wolfowitz's inequality $P_F\{n^{1/2}\|U_n - U\|_\infty \ge t\} \le c_1 e^{-t^2}$, we have

$$E[nR_\tau(U_n,U)]^2 \le E_F[n^2\|U_n - U\|_\infty^4] = \int_0^\infty P_F\{n^2\|U_n - U\|_\infty^4 \ge t\}dt \le c_2,$$

where c_1 and c_2 are positive constants. This shows that $\{nR_\tau(U_n, U) : i = 1, 2, \dots\}$ is uniformly integrable (Serfling (1980), p.13). Hence $E_F[R_\tau(U_n, U)] = o(n^{-1})$ by (3.8).

The next theorem proves a similar result for second-order Fréchet differentiable T w.r.t. $\|\cdot\|_\infty$ or the r-norm defined by

$$\|\cdot\|_r = \|\cdot\|_\infty + \|\cdot\|_{L_r}, \quad r = 1, 2. \tag{3.9}$$

The advantage of using the r-norm is that more functionals are differentiable w.r.t. r-norm (see Example 7.2), since from $\|\cdot\|_r \geq \max\{\|\cdot\|_{L_r}, \|\cdot\|_\infty\}$, it is apparent that differentiability of T w.r.t. $\|\cdot\|_{L_r}$ or $\|\cdot\|_\infty$ implies differentiability of T w.r.t. $\|\cdot\|_r$.

Theorem 3.2. (i) *Suppose that T is second-order Fréchet differentiable at F w.r.t. the sup-norm $\|\cdot\|_\infty$ with $\phi_T(x, y, F)$ satisfying* (2.1). *Then* (3.2)-(3.5) *hold with $\xi_i = 2\int \phi_T(X_i, y, F)dF(y)$, $\sigma^2 = E_F\xi_i^2$ and $\gamma = \int \phi_T(x, x, F)dF(x)$.*
(ii) *Results* (3.2)-(3.5) *also hold if $\|\cdot\|_\infty$ in* (i) *is replaced by the r-norm $\|\cdot\|_r$ defined in* (3.9) *and $\int\{F(x)[1 - F(x)]\}^{r/2}dx < \infty (r = 1$ or $2)$.*
(iii) *Assume the conditions in either* (i) *or* (ii). *Then* (3.6) *holds under condition* (2.2) *or condition* (2.2) *with $\|\cdot\|_\infty$ replaced by $\|\cdot\|_r$.*

Remark. For $r = 1$, $\int\{F(x)[1 - F(x)]\}^{r/2}dx < \infty$ is almost the same as $E_F X_1^2 < \infty$ (see Serfling (1980), p.276) and is implied by $E_F|X_1|^{2+t} < \infty$ with $t > 0$. For $r = 2$, $\int F(x)[1 - F(x)]dx < \infty$ is equivalent to $E_F|X_1| < \infty$.

Proof. (i) Note that if F is continuous, then the results follow directly from Proposition 2.2 and Theorem 3.1. For general F, let

$$R_T(F_n, F) = T(F_n) - T(F) - \frac{1}{n^2} \sum_{1 \leq i \leq j \leq n} \phi_T(X_i, X_j, F).$$

Following the proof of Theorem 3.1, we obtain the result if we can show

$$R_T(F_n, F) = o_p(n^{-1}). \tag{3.10}$$

From the second-order Fréchet differentiability of T, for any $\epsilon > 0$, there is a $\delta_\epsilon > 0$ such that

$$P_F\{n|R_T(F_n, F)| \geq \epsilon_0\} \leq P_F\{n\|F_n - F\|_\infty^2 \geq \epsilon_0/\epsilon\} + P_F\{\|F_n - F\|_\infty \geq \delta_\epsilon\},$$

where ϵ_0 is arbitrary. Then (3.10) follows from $\|F_n - F\|_\infty = O_p(n^{-1/2})$.

(ii) From the proof of (i), results (3.2)-(3.5) follow from (3.10), which is implied by

$$\|F_n - F\|_r = O_p(n^{-1/2}), \quad r = 1,2. \tag{3.11}$$

Note that $\|F_n - F\|_\infty = O_p(n^{-1/2})$,

$$E_F\|F_n - F\|_{L_1} \le \int \{E_F[F_n(x) - F(x)]^2\}^{1/2} dx \le n^{-1/2} \int \{F(x)[1 - F(x)]\}^{1/2} dx$$

and

$$E_F\|F_n - F\|_{L_2}^2 = n^{-1} \int F(x)[1 - F(x)]dx.$$

Hence (3.11) holds under the given conditions.

(iii) If (2.2) holds with a norm $\|\cdot\|$, then (3.6) follows from (3.10) and the uniform integrability of $\{n\|F_n - F\|^2\}$. Using Dvoretzky, Kiefer and Wolfowitz's inequality, $E_F[n^2\|F_n - F\|_\infty^4]$ is bounded. Hence the results follow from the uniform integrability of $\{n\|F_n - F\|_{L_r}^2\}$, $r = 1,2$. Let $\eta_i(x) = I_{X_i}(x) - F(x)$. Then

$$\begin{aligned} n\|F_n - F\|_{L_1}^2 &= n\Big[\int \Big|\frac{1}{n}\sum_{i=1}^n \eta_i(x)\Big|dx\Big]^2 = \frac{1}{n}\iint \Big|\sum_{1\le i,j\le n} \eta_i(x)\eta_j(y)\Big|dxdy \\ &\le \frac{1}{n}\sum_{i=1}^n \Big[\int |\eta_i(x)|dx\Big]^2 + \frac{1}{n}\iint \Big|2\sum_{1\le i<j\le n} \eta_i(x)\eta_j(y)\Big|dxdy. \end{aligned} \tag{3.12}$$

From $E_F\big[\int|\eta_1(x)|dx\big]^2 \le \big\{\int\{E_F[\eta_i^2(x)]\}^{1/2}dx\big\}^2 = \big\{\int\{F(x)[1-F(x)]\}^{1/2}dx\big\}^2 < \infty$, the first term in (3.12) is uniformly integrable. From

$$\begin{aligned} &E_F\Big\{\frac{1}{n}\iint \Big|2\sum_{1\le i<j\le n} \eta_i(x)\eta_j(y)\Big|dxdy\Big\}^2 \\ &\le \Big\{\frac{1}{n}\iint \Big[E_F\Big|2\sum_{1\le i<j\le n} \eta_i(x)\eta_j(y)\Big|^2\Big]^{1/2} dxdy\Big\}^2 \\ &\le \Big\{\iint [E_F\eta_1^2(x)E_F\eta_2^2(y)]^{1/2} dxdy\Big\}^2 \\ &= \Big\{\int \{F(x)[1 - F(x)]\}^{1/2} dx\Big\}^4, \end{aligned}$$

the second term in (3.12) is also uniformly integrable. Hence $\{n\|F_n - F\|_{L_1}^2\}$ is uniformly integrable. Similarly,

$$n\|F_n - F\|_{L_2}^2 = \frac{1}{n}\sum_{i=1}^n \int \eta_i^2(x)dx + \frac{2}{n}\int \sum_{1\le i<j\le n} \eta_i(x)\eta_j(x)dx.$$

Then $\{n\|F_n - F\|_{L_2}^2\}$ is uniformly integrable since $E_F \int \eta_1^2(x)dx = \int F(x)[1 - F(x)]dx$ and

$$E_F\left[\frac{2}{n}\int \sum_{1\le i<j\le n} \eta_i(x)\eta_j(x)dx\right]^2 = \frac{n-1}{n}\iint E_F[\eta_1(x)\eta_2(y)]^2 dxdy$$
$$\le \left\{\int F(x)[1-F(x)]dx\right\}^2.$$

In view of (3.3)-(3.5), $B_n = \gamma/n$ can be considered as the asymptotic bias of $T(F_n)$. Result (3.6) may hold under weaker conditions than those in Theorems 3.1 and 3.2.

4. Consistency of ν_n^J

Let ν_n^J be the jackknife variance estimator defined in Section 1.

Theorem 4.1. *Suppose that F is continuous and τ in* (2.4) *is second-order compactly differentiable at U* w.r.t. $\|\cdot\|_\infty$ *with $\phi_\tau(s,t,U)$ satisfying* (3.1). *Let σ^2 be given in* (3.2). *Then*

$$n\nu_n^J \to_p \sigma^2. \tag{4.1}$$

Proof. We use the same notation as in the proof of Theorem 3.1. For each i, let $U_n^{(i)}$ be the empirical distribution corresponding to $Z_1,\dots,Z_{i-1},Z_{i+1},\dots,Z_n$. Then

$$T(F_n^{(i)}) - T(F) = \tau(U_n^{(i)}) - \tau(U) = w_{n,i} + r_{n,i},$$

where $w_{n,i} = n^{-2}\sum_{1\le k,j\le n,\, k,j\ne i}\phi_\tau(Z_k,Z_j,U)$ and $r_{n,i} = R_\tau(U_n^{(i)},U)$ (R_τ is given in (2.6)). Let $\bar{w}$ and $\bar{r}$ be the averages of $w_{n,i}$ and $r_{n,i}$, respectively. Then

$$n\nu_n^J = (n-1)\sum_{i=1}^n (w_{n,i}-\bar{w})^2 + (n-1)\sum_{i=1}^n (r_{n,i}-\bar{r})^2 + 2(n-1)\sum_{i=1}^n (w_{n,i}-\bar{w})r_{n,i}. \tag{4.2}$$

The first term on the right side of (4.2) $\to_p \sigma^2$, since it is the jackknife variance estimator of a V-statistic (see Sen (1977)). By Cauchy-Schwarz inequality, (4.1) follows from

$$(n-1)\sum_{i=1}^n r_{n,i}^2 \to_p 0, \tag{4.3}$$

which is implied by

$$n^2 \max_{i\le n}[R_\tau(U_n^{(i)},U)]^2 \to_p 0. \tag{4.4}$$

Let U_n^* be the continuous version of U_n, i.e., U_n^* is the distribution function corresponding to a uniform distribution of mass $(n+1)^{-1}$ in each of the $n+1$ intervals $[Z_{(i-1)}, Z_{(i)}]$, where $Z_{(i)}$ are ordered Z_i's, $Z_{(0)} = 0$ and $Z_{(n+1)} = 1$. From Donsker's theorem and Prohorov's theorem, for any $\epsilon > 0$, there exists a compact $\Lambda \subset \mathbf{D}$ such that for $n > 1$,

$$P\{n^{1/2}(U_n^* - U) \in \Lambda\} > 1 - \epsilon.$$

Note that $\|U_n^* - U_n\|_\infty \le n^{-1}$ and for each i, $\|U_n^{(i)} - U_n\|_\infty \le n^{-1}$. Then $\|U_n^{(i)} - U_n^*\|_\infty \le 2n^{-1}$ for all $i \le n$ and for $n > 1$,

$$P\{\rho[n^{1/2}(U_n^{(i)} - U), \Lambda] \le 2n^{-1/2},\ i = 1, \dots, n\} > 1 - \epsilon,$$

where $\rho[G, \Lambda] = \inf\{\|K - G\|_\infty : K \in \Lambda\}$. We now show that for sufficiently large n, $\rho[n^{1/2}(U_n^{(i)} - U), \Lambda] \le 2n^{-1/2}$, $i = 1, \dots, n$, implies $n|R_\tau(U_n^{(i)}, U)| < \epsilon$, $i = 1, \dots, n$. Suppose not. Then there is a sequence $\{n_m\}$ such that

$$n_m|R_\tau(U_{n_m}^{(i_m)}, U)| \ge \epsilon \tag{4.5}$$

and $\rho[G_m, \Lambda] \le 2n_m^{-1/2}$ for all m, where $G_m = n_m^{1/2}(U_{n_m}^{(i_m)} - U)$. For each m, there is a $K_m \in \Lambda$ such that $\|G_m - K_m\|_\infty \le 2n_m^{-1/2}$. Since Λ is compact, there is a $K \in \Lambda$ and a subsequence $\{m_j\}$ such that $\|K_{m_j} - K\|_\infty \to 0$ as $j \to \infty$. Then $\|G_{m_j} - K\|_\infty \to 0$ and the set $\Lambda_1 = \{K, G_{m_j}, j = 1, 2, \dots\}$ is compact. Since τ is second-order compactly differentiable at U, $\lim_{j\to\infty} n_{m_j}|R_\tau(U + n_{m_j}^{-1/2} G_{m_j}, U)| = 0$. This contradicts (4.5). Therefore

$$P\{n|R_\tau(U_n^{(i)}, U)| < \epsilon, i = 1, \dots, n\} > 1 - \epsilon$$

for sufficiently large n. This completes the proof since ϵ is arbitrary.

It was shown in Shao and Wu (1989, Theorem 5) that (4.1) holds if T is second-order Fréchet differentiable at F w.r.t. $\|\cdot\|_\infty$. Compared with the result there, our Theorem 4.1 requires less condition on T (see Proposition 2.2) but assumes the continuity of F. The following result shows the consistency of ν_n^J when T is second-order Fréchet differentiable at F w.r.t. the r-norm.

Theorem 4.2. *Suppose that T is second-order Fréchet differentiable at F* w.r.t. $\|\cdot\|_r$ *defined in* (3.9) ($r = 1$ *or* 2) *with* $\phi_T(x, y, F)$ *satisfying* (2.1) *and that* $\int\{F(x)[1 - F(x)]\}^{r/2} < \infty$. *Then* (4.1) *holds with* σ^2 *given in Theorem* 3.2.

Proof. Let $w_{n,i} = n^{-2}\sum_{1\le k,j\le n,\, k,j\ne i}\phi_T(X_k,X_j,F)$ and $r_{n,i} = T(F_n^{(i)}) - T(F) - w_{n,i}$. Then (4.2) holds and it remains to show (4.3). Let $\epsilon_0 > 0$ be given and

$$\zeta_n = (n-1)\sum_{i=1}^n \|F_n^{(i)} - F\|_\infty^4 \quad \text{and} \quad t_n(r) = (n-1)\sum_{i=1}^n \|F_n^{(i)} - F\|_{L_r}^4, r = 1,2.$$

From the second-order Fréchet differentiability of T, for any $\epsilon > 0$, there is a $\delta_\epsilon > 0$ such that

$$P_F\Big\{(n-1)\sum_{i=1}^n r_{n,i}^2 > \epsilon_0\Big\} \le P_F\{\zeta_n + t_n(r) > \epsilon_0/8\epsilon\} + P_F\{\max_{i\le n}\|F_n^{(i)} - F\|_r > \delta_\epsilon\}.$$

Since $(\max_{i\le n}\|F_n^{(i)} - F\|_r)^4 \le \sum_{i=1}^n \|F_n^{(i)} - F\|_r^4$, (4.3) follows from $\zeta_n + t_n(r) = O_p(1)$. From $\zeta_n \le 8 + 8n^2\|F_n - F\|_\infty^4 = O_p(1)$, it remains to show that $t_n(r) = O_p(1)$, $r = 1,2$. Note that

$$t_n(1) \le 8(n-1)\sum_{i=1}^n\Big[\int |F_n^{(i)}(x) - F_n(x)|dx\Big]^4 + 8n^2\Big[\int |F_n(x) - F(x)|dx\Big]^4$$

$$\le \frac{64}{(n-1)^3}\sum_{i=1}^n\Big[\int |I_{X_i}(x) - F(x)|dx\Big]^4 + \Big[\frac{64n}{(n-1)^3} + 8n^2\Big]\Big[\int |F_n(x) - F(x)|dx\Big]^4.$$

From the proof of Theorem 3.2(ii), $\big[\int |F_n(x) - F(x)|dx\big]^4 = O_p(n^{-2})$. Since

$$E_F\Big[\int |I_{X_1}(x) - F(x)|dx\Big]^2 \le \Big\{\int \{F(x)[1 - F(x)]\}^{1/2}dx\Big\}^2 < \infty,$$

$n^{-3}\sum_{i=1}^n\big[\int |I_{X_i}(x) - F(x)|dx\big]^4 \to_p 0$ by Marcinkiewicz law of large numbers. Hence $t_n(1) = O_p(1)$. A similar argument shows that

$$t_n(2) \le \frac{64}{(n-1)^3}\sum_{i=1}^n\Big\{\int [I_{X_i}(x) - F(x)]^2dx\Big\}^2 + O_p(1).$$

Since $E_F\int[I_{X_1}(x) - F(x)]^2dx = \int F(x)[1 - F(x)]dx < \infty$, $t_n(2) = O_p(1)$ follows from Marcinkiewicz law of large numbers.

5. Consistency of b_n^J

Let b_n^J be the jackknife bias estimator defined in Section 1. To show the consistency of b_n^J, we need a differentiability assumption which is stronger than that in Definition 2.1.

Definition 5.1. A functional f on $\mathbf{N}$ is *uniformly second-order Fréchet differentiable at* $H \in \mathbf{N}$ w.r.t. $\|\cdot\|$ if there is a real-valued function $\phi_f(x,y,K)$ defined on $\mathbf{R}^2 \times \{K \in \mathbf{N} : \|K - H\| < \delta\}$ for a fixed $\delta > 0$ such that $\phi_f(x,y,K) = \phi_f(y,x,K)$, $\iint \phi_f(x,y,K)dK(x)dK(y) = 0$ and

$$\lim_{\|G-H\|+\|K-H\|\to 0,\, G,K\in \mathbf{N}} \frac{f(G) - f(K) - \iint \phi_f(x,y,K)dG(x)dG(y)}{\|G-K\|^2} = 0.$$

It can be seen from Example 7.1 that this uniform second-order Fréchet differentiability is much weaker than the strong second-order Fréchet differentiability defined in Parr (1985).

Lemma 5.1. *Let $g(x,K)$ be a real-valued function defined on $\mathbf{R}\times\{K \in \mathbf{N} : \|K - F\| < \delta\}$ for a fixed $\delta > 0$, where the norm $\|\cdot\|$ satisfies $\|F_n - F\| \to_p 0$. Suppose that $E_F|g(X_1,F)| < \infty$ and for any $c > 0$, $\sup_{|x|\le c} |g(x,F_n) - g(x,F)| \to_p 0$. Suppose also that there is a function $h(x) \ge 0$ such that $E_F h(X_1) < \infty$ and $\lim_{n\to\infty} P_F\{|g(x,F_n)| \le h(x), x \in \mathbf{R}\} = 1$. Then*

$$\zeta_n = \int |g(x,F_n) - g(x,F)|dF_n(x) \to_p 0.$$

Proof. Let $S_n = \{\|F_n - F\| < \delta$ and $|g(x,F_n)| \le h(x)$ for all $x\}$ and S_n^c be its complement. Let $\epsilon > 0$ be given. Then there is a $c > 0$ such that

$$a = \int_{|x|>c} [h(x) + |g(x,F)|]dF(x) < \epsilon/3.$$

Let $a_n = \int_{|x|>c}[h(x) + |g(x,F)|]dF_n(x)$ and $b_n = \sup_{|x|\le c} |g(x,F_n) - g(x,F)|$. Then

$$P_F\{\zeta_n \ge \epsilon\} \le P_F\{a_n \ge \epsilon/2\} + P_F\{b_n \ge \epsilon/2\} + P_F(S_n^c).$$

Under the given conditions, $P_F(S_n^c) \to 0$, $b_n \to_p 0$ and $a_n \to_p a$. Hence $P_F\{\zeta_n \ge \epsilon\} \to 0$.

Theorem 5.1. (i) *Suppose that T is uniformly second-order Fréchet differentiable at F* w.r.t. $\|\cdot\|_\infty$ *with $\phi_T(x, y, K)$ satisfying* (2.1) *and $\phi_T(x, x, K)$ satisfying the conditions in Lemma* 5.1. *Let γ be given in Theorem* 3.2. *Then*

$$nb_n^J \to_p \gamma. \tag{5.1}$$

(ii) *Result* (5.1) *also holds if $\|\cdot\|_\infty$ in* (i) *is replaced by $\|\cdot\|_r$ and $E_F|X_1|^{2/r} < \infty$* ($r = 1$ *or* 2).
(iii) *Assume the conditions in either* (i) *or* (ii). *If $\phi_T(x, x, K)$ is defined on* $\mathbf{R} \times \mathbf{N}$ *and satisfies*

$$\Phi_n = \frac{1}{n}\sum_{i=1}^{n} \phi_T(X_i, X_i, F_n) \text{ is uniformly integrable} \tag{5.2}$$

(*which is implied by* $\sup_n E_F|\Phi_n|^t < \infty$ *for* $t > 1$) *and T satisfies*

$$\sup_{G,H\in\mathbf{N}} \left| \frac{T(G) - T(H) - \iint \phi_T(x, y, H)dG(x)dG(y)}{\|G - H\|^2} \right| < \infty, \tag{5.3}$$

where $\|\cdot\|$ is either $\|\cdot\|_\infty$ or $\|\cdot\|_r$, then

$$nE_F b_n^J \to \gamma. \tag{5.4}$$

Proof. (i) From the uniform second-order Fréchet differentiability of T at F, for any $0 < \epsilon < \epsilon_0$, there is a $\delta_\epsilon > 0$ such that

$$P_F\{|s_n| > \epsilon_0\} \le P_F\Big\{n\sum_{i=1}^{n} \|F_n^{(i)} - F_n\|_\infty^2 \ge \epsilon_0/\epsilon\Big\} + P_F\Big\{\max_{0\le i\le n} \|F_n^{(i)} - F\|_\infty > \delta_\epsilon\Big\}, \tag{5.5}$$

where $F_n^{(0)} = F_n$ and

$$s_n = (n-1)\sum_{i=1}^{n} \Big[T(F_n^{(i)}) - T(F_n) - \iint \phi_T(x, y, F_n)dF_n^{(i)}(x)dF_n^{(i)}(y)\Big]. \tag{5.6}$$

Since $\|F_n^{(i)} - F_n\|_\infty \le n^{-1}$, $s_n \to_p 0$. A straightforward calculation shows that

$$\begin{aligned}\frac{n-1}{n}\sum_{i=1}^{n}&\Big[\frac{1}{(n-1)^2}\sum_{1\le k,j\le n, k,j\ne i} \phi_T(X_k, X_j, F_n)\Big] \\ &= \frac{1}{n}\sum_{i=1}^{n} \phi_T(X_i, X_i, F_n) + \frac{2(n-2)}{n(n-1)}\sum_{1\le i<j\le n} \phi_T(X_i, X_j, F_n).\end{aligned}$$

Since $\sum_{1\leq i,j\leq n}\phi_T(X_i,X_j,F_n) = n^2\iint\phi_T(x,y,F_n)dF_n(x)dF_n(y) = 0$, we have

$$nb_n^J = \frac{1}{n-1}\sum_{i=1}^n \phi_T(X_i,X_i,F_n) + o_p(1). \tag{5.7}$$

Hence (5.1) holds since by the law of large numbers and Lemma 5.1 with $g(x,K) = \phi_T(x,x,K)$,

$$\frac{1}{n}\sum_{i=1}^n \phi_T(X_i,X_i,F_n) \to_p \gamma. \tag{5.8}$$

(ii) Let s_n be given in (5.6). By the second-order differentiability of T, (5.5) holds with $\|\cdot\|_\infty$ replaced by $\|\cdot\|_r$. From the proof of (i), (5.1) holds if

$$(n-1)\sum_{i=1}^n \|F_n^{(i)} - F_n\|_{L_r}^2 = O_p(1), \quad r = 1,2. \tag{5.9}$$

From $F_n^{(i)} - F_n = (n-1)^{-1}(I_{X_i} - F_n)$, the left side of (5.9) is equal to

$$\frac{1}{n}\sum_{i=1}^n \|I_{X_i} - F_n\|_{L_r}^2 \leq \frac{4}{n-1}\sum_{i=1}^n \|I_{X_i} - F\|_{L_r}^2, \quad r = 1,2.$$

From the inequality

$$\int |I_y(x) - F(x)|dx \leq |y| + E_F|X_1|, \tag{5.10}$$

$E_F\|I_{X_1} - F\|_{L_1}^2 \leq E_F[|X_1| + E_F|X_1|]^2$. Also, $E_F\|I_{X_1} - F\|_{L_2}^2 = \int F(x)[1 - F(x)]dx \leq E_F|X_1|$. Hence (5.9) holds under the given conditions.

(iii) By conditions (5.2) and (5.8), $E_F\big[\frac{1}{n}\sum_{i=1}^n \phi_T(X_i,X_i,F_n)\big] \to \gamma$. Hence (5.4) holds if $E_F s_n \to 0$, which follows from the uniform integrability of $\{s_n\}$. If (5.3) holds with $\|\cdot\|_\infty$, then (5.4) holds since $|s_n| \leq (n-1)\sum_{i=1}^n \|F_n^{(i)} - F_n\|_\infty^2 \leq 1$. Suppose that (5.3) holds with $\|\cdot\|_r$. Then

$$|s_n| \leq 2 + 2(n-1)\sum_{i=1}^n \|F_n^{(i)} - F_n\|_{L_r}^2 \leq 2 + \frac{8}{n-1}\sum_{i=1}^n \|I_{X_i} - F\|_{L_r}^2.$$

Note that $\|I_{X_i} - F\|_{L_r}^2$ are i.i.d. random variables with

$$E_F\|I_{X_1} - F\|_{L_r}^2 = E_F\Big[\int |I_{X_1}(x) - F(x)|^r dx\Big]^{2/r} \leq E_F(|X_1| + E_F|X_1|)^{2/r}$$

(by inequality (5.10)). Hence if $E_F|X_1|^{2/r} < \infty$, $\{(n-1)^{-1}\sum_{i=1}^n \|I_{X_i} - F\|_{L_r}^2\}$ is uniformly integrable, which implies the uniform integrability of $\{s_n\}$. This completes the proof.

6. Asymptotic Properties of $T^J(F_n)$

The jackknife estimator $T^J(F_n) = T(F_n) - b_n^J$ is designed to remove the leading term γ/n (see Theorems 3.1 and 3.2) in the bias of $T(F_n)$. The following result rigorously justifies this "bias reduction" property of the jackknife method.

Theorem 6.1. (i) *Suppose that T is uniformly second-order Fréchet differentiable at F* w.r.t. $\|\cdot\|_\infty$ *with $\phi_T(x,y,K)$ satisfying* (2.1) *and $\phi_T(x,x,K)$ satisfying the conditions in Lemma* 5.1. *Let $\xi_i, \sigma^2, \lambda_k$ and Y_k be the same as in Theorem* 3.2. *Then*

$$n^{1/2}[T^J(F_n) - T(F)] \to N(0,\sigma^2) \quad \textit{in distribution,} \tag{6.1}$$

and

$$T^J(F_n) = T(F) + \frac{1}{n}\sum_{i=1}^n \xi_i + \Gamma_n^J \tag{6.2}$$

with

$$n\Gamma_n^J \to \sum_{k=1}^\infty \lambda_k(Y_k - 1) \quad \textit{in distribution.} \tag{6.3}$$

(ii) *Results* (6.1)-(6.3) *hold if* $\|\cdot\|_\infty$ *in* (i) *is replaced by* $\|\cdot\|_r$ *and* $E_F|X_1|^{2/r} < \infty$ $(r = 1$ *or* $2)$.

(iii) *Assume the conditions in either* (i) *or* (ii) *hold. If* (3.6), (5.2) *and* (5.3) *hold* (*with* $\|\cdot\| =$ *either* $\|\cdot\|_\infty$ *or* $\|\cdot\|_r$), *then*

$$nE_F[T^J(F_n) - T(F)] \to 0. \tag{6.4}$$

Proof. Let $h(x,y,F) = \phi_T(x,y,F) - \int[\phi_T(x,u,F) + \phi_T(u,y,F)]du$. From (3.10), (5.6)-(5.8),

$$T^J(F_n) = T(F) + \frac{1}{n}\sum_{i=1}^n \xi_i + \frac{2}{n(n-1)}\sum_{1\le i<j\le n} h(X_i,X_j,F) + o_p(n^{-1}).$$

Hence results (6.1)-(6.3) follows from Theorem 5.5.2 in Serfling (1980) under the given conditions. Result (6.4) follows from (3.6) and (5.4) under the given conditions.

From (3.1) and (6.1), the jackknife estimator $T^J(F_n)$ has the same first order asymptotic distribution as $T(F_n)$. The second-order asymptotic representations of $T(F_n)$ and $T^J(F_n)$, however, are different: Γ_n^J in (6.3) converges in distribution to a random variable with mean zero whereas Γ_n in (3.3) converges in distribution to a random variable with mean γ (which is not necessarily zero). Hence the jackknife estimator $T^J(F_n)$ is less biased. Formal results such as (6.4) may be established under weaker conditions than those in Theorem 6.1(iii). Note that both $T(F_n)$ and $T^J(F_n)$ have the same asymptotic variance ($= \sigma^2/n$ if $\sigma^2 > 0$ and $= \sum_k \lambda_k^2/n^2$ if $\sigma^2 = 0$). Hence the jackknife estimator $T^J(F_n)$ improves $T(F_n)$ in the sense that it reduces the bias but does not increase the asymptotic variance.

7. Examples

Example 7.1. Functions of mean. Suppose that $\mathbf{F}$ consists of all the distributions with finite mean and $T(G) = g(\mu_G)$, where $\mu_G = \int x dG(x)$ and g is a real-valued function which is second-order differentiable in a neighborhood of μ_F. Let

$$\phi_T(x,y,K) = g'(\mu_K)[(x+y)/2 - \mu_K] + g''(\mu_K)[xy - \mu_K(x+y) + \mu_K^2]/2. \quad (7.1)$$

Let $\bar{X}$ be the average of X_i. Then $\phi_T(x,x,F_n) = g'(\bar{X})(x-\bar{X}) + g''(\bar{X})(x-\bar{X})^2/2$. If $E_F X_1^2 < \infty$ and g'' is continuous at μ_F, it can be easily shown that (2.1) holds and $\phi(x,x,F_n)$ satisfies the conditions in Lemma 5.1. If there is a constant $c > 0$ such that $|g''(x)| \le c$, then

$$\left|\frac{1}{n}\sum_{i=1}^n \phi_T(X_i,X_i,F_n)\right| = \frac{1}{2n}|g''(\bar{X})|\sum_{i=1}^n (X_i - \bar{X})^2 \le \frac{c}{2n}\sum_{i=1}^n X_i^2$$

and (5.2) holds since $\{\frac{1}{n}\sum_{i=1}^n X_i^2\}$ is uniformly integrable under $E_F X_1^2 < \infty$. The following result shows that T is uniformly second-order Fréchet differentiable at F w.r.t. the L_1-norm. Note that T is not necessarily strongly Fréchet differentiable at F (Definition 2 in Parr (1985)).

Proposition 7.1. *If $g''(t)$ is defined on $[\mu_F - a, \mu_F + a]$, $a > 0$, and is continuous at μ_F, then T is uniformly second-order Fréchet differentiable at F* w.r.t. $\|\cdot\|_{L_1}$ *with $\phi_T(x,y,K)$ given in* (7.1). *If, in addition, g'' is defined on $\mathbf{R}$ and is bounded, then* (5.3) *holds with $\|\cdot\|_{L_1}$.*

Proof. For $G, K \in \mathbf{N}$ with $\|G-K\|_{L_1} < a$, we have $|\mu_G - \mu_K| \le \|G-K\|_{L_1} < a$ and

$$T(G) - T(K) = g'(\mu_K)(\mu_G - \mu_K) + 2^{-1}g''(\zeta_{G,K})(\mu_G - \mu_K)^2,$$

where $\zeta_{G,K}$ is on the line segment between μ_G and μ_K. Since $\iint xydG(x)dG(y) = \mu_G^2$,

$$|R_T(G,K)| = \left|T(G) - T(K) - \iint \phi_T(x,y,K)dG(x)dG(y)\right|$$
$$= 2^{-1}|g''(\zeta_{G,K}) - g''(\mu_K)|(\mu_G - \mu_K)^2 \le 2^{-1}|g''(\zeta_{G,K}) - g''(\mu_K)|\,\|G - K\|_{L_1}^2.$$

Hence the first assertion follows from the continuity of g'' at μ_F and (5.3) follows from the boundedness of g''.

Example 7.2. L-statistics. Let $J(t)$ be a real-valued differentiable function on $[0,1]$ and $T(G) = \int xJ[G(x)]dG(x)$. $T(F_n)$ is called an L-statistic (see Serfling (1980), Chapter 8). Define $a(x,K) = \int [K(u) - I_x(u)]J[K(u)]du$, $b(x,y,K) = -\int [K(u) - I_x(u)][K(u) - I_y(u)]J'[K(u)]du$ and

$$\phi_T(x,y,K) = [a(x,K) + a(y,K) + b(x,y,K)]. \tag{7.2}$$

If $E_F X_1^2 < \infty$, it can be shown that (2.1), (5.2) and the conditions in Lemma 5.1 hold (with $h(x) = |x| + E_F|X_1| + 1$). The following result shows the differentiability of T. Note that T may not be second-order differentiable w.r.t. $\|\cdot\|_\infty$ unless J is trimmed (see Proposition 7.2(ii)).

Proposition 7.2. (i) *If J' is continuous on* $[0, 1]$, *then* (5.3) *holds with* $\|\cdot\|_2$ *and T is uniformly second-order Fréchet differentiable at F* w.r.t. $\|\cdot\|_2$ *with* $\phi_T(x,y,K)$ *given in* (7.2).
(ii) *Suppose that $J(t) = 0$ for $t \in [0, \alpha) \cup (\beta, 1]$, $0 < \alpha < \beta < 1$, and that J' is continuous on $[\alpha, \beta]$. Then T is uniformly second-order Fréchet differentiable at* F w.r.t. $\|\cdot\|_\infty$ *with $\phi_T(x,y,K)$ given in* (7.2). *Furthermore, for each $K \in$* N, *there is a constant c_K such that $|\phi_T(x,y,K)| \le c_K$.*

Proof. (i) Let $R_T(G,K) = T(G) - T(K) - \iint \phi_T(x,y,K)dG(x)dG(y)$. From Serfling (1980) (p.289),

$$R_T(G,K) = \int W[G(x),K(x)][G(x) - K(x)]^2dx,$$

where $W(s,t) = \left[\int_t^s J(u)du - J(t)(s-t) - 2^{-1}J'(t)(s-t)^2\right]/(s-t)^2$ if $s \ne t$ and $W(s,t) = 0$ if $s = t$. From the continuity of J', $\sup_x |W[G(x),K(x)]| \le \|J'\|_\infty < \infty$ and $\sup_x |W[G(x),K(x)]| \to 0$ as $\|G - K\|_\infty \to 0$. Hence the results in (i) hold.
(ii) Let c and d be two real numbers such that $F(c) > \beta$ and $F(d) < \alpha$, and $\delta = \min\{\alpha - F(d), F(c) - \beta\}$. Note that if both $G(x)$ and $K(x)$ are in $[0, \alpha) \cup (\beta, 1]$,

then $W[G(x), K(x)] = 0$. Suppose that G and K satisfy $\|G - F\|_\infty < \delta$ and $\|K - F\|_\infty < \delta$. Then if $x \le d$, both $G(x)$ and $K(x) \le \alpha$ and if $x \ge d$, both $G(x)$ and $K(x) \ge \beta$. Hence

$$|R_T(G,K)| = \left| \int_d^c W[G(x), K(x)][G(x) - K(x)]^2 dx \right| \le \|G - K\|_\infty^2 \int_d^c |W[G(x), K(x)]| dx.$$

Then the first assertion follows from $\int_d^c |W[G(x), K(x)]| dx \to 0$ as $\|G - K\|_\infty \to 0$. Similarly, for each K, there are constants $c(K)$ and $d(K)$ such that

$$a(x, K) = \int_{d(K)}^{c(K)} [K(u) - I_x(u)] J[K(u)] du$$

and

$$b(x, y, K) = -\int_{d(K)}^{c(K)} [K(u) - I_x(u)][K(u) - I_y(u)] J'[K(u)] du.$$

Thus the second assertion holds since J and J' are bounded.

References

Beran, R. (1984). Jackknife approximations to bootstrap estimates. *Ann. Statist.* **12**, 101-118.

Parr, W. C. (1985). Jackknifing differentiable statistical functionals. *J. Roy. Statist. Soc. Ser. B* **47**, 56-66.

Quenouille, M. H. (1956). Notes on bias in estimation. *Biometrika* **43**, 353-360.

Sen, P. K. (1977). Some invariance principles relating to jackknifing and their role in sequential analysis. *Ann. Statist.* **5**, 316-329.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley, New York.

Shao, J. and Wu, C. F. J. (1989). A general theory for jackknife variance estimation. *Ann. Statist.* **17**, 1176-1197.

Department of Mathematics, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5.

(Received April 1989; accepted June 1990)