Functional Jackknifing:
Rationality and General Asymptotics

by

P.K. Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1800

March 1980

FUNCTIONAL JACKKNIFING : RATIONALITY AND GENERAL ASYMPTOTICS *

BY PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

Though jackknifing serves the dual purpose of bias reduction and variance estimation, the pseudo variables, it generates, may not generally preserve robustness for general statistical functionals. These pseudo variables are incorporated in a differentiable functional detour of jackknifing, and along with its rationality, the related asymptotic theory is studied systematically. Second order asymptotic distributional representations for the classical jackknifed estimators are also considered.

1. Introduction. The *jackknife technique*, originally conceived for possible reduction of (smaller order) *bias* of an estimator, generates *pseudo variables* which also provide a (strongly) consistent estimator of the *asymptotic variance* of the *jackknifed estimator;* for some detailed studies of these properties of jackknifing, we may refer to Miller(1974) and Sen(1977), among others. To filter *robustness* under jackknifing, one should choose a robust initial estimator; otherwise, the pseudo variables, generated by jackknifing, may lead to a less robust jackknifed version. For example, let $X_1,\ldots,X_n$ be n independent and identically distributed random variables (i.i.d.r.v.) with a distribution function (d.f.) F having finite (but unknown) mean μ and variance σ^2. For normal F, the classical estimator of μ is the sample mean $\bar{X}_n = n^{-1}\Sigma_{i=1}^n X_i$; the corresponding jackknifed version is also $\bar{X}_n$, and the jackknifed variance estimator agrees with the sample variance $s_n^2 = (n-1)^{-1}\Sigma_{i=1}^n (X_i - \bar{X}_n)^2$. However, none of $\bar{X}_n$ and s_n^2 is robust against outliers or gross errors. In such a case (as well as for a general *estimable parameter* $\theta(F)$), one may choose a more robust initial estimator $T_n = T(X_1,\ldots,X_n)$ of θ [viz., the

AMS Subject Classifications : 60F05, 62E20, 62F12

Key Words and Phrases: Asymptotic normality; asymptotic variance; bias reduction; Hadamard derivative; jackknifed variance; perturbations; pseudo variables; robustness; second order representations; statistical functional; two-step jackknifing.

* Work supported by the Office of Naval Research, Contract : N00014-83-K-0387.

trimmed mean, median, R- and M-estimators etc.,) and construct the corresponding jackknifed version T_n^* (as well as the jackknifed variance estimator V_n^*) from the pseudo variables $T_{n,i}$ given by

(1.1) $T_{n,i} = nT_n - (n-1)T(X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n)$, $i=1,\ldots,n$;

(1.2) $T_n^* = n^{-1}\Sigma_{i=1}^n T_{n,i}$ and $V_n^* = (n-1)^{-1}\Sigma_{i=1}^n (T_{n,i} - T_n^*)^2$.

In this setup, we may take $T_n = T(F_n)$, a functional of the empirical d.f. F_n. Note that if G_n^* stands for the empirical d.f. of the $T_{n,i}$ in (1.1), then

(1.3) $T_n^* = \int x dG_n^*(x)$ and $V_n^* = n(n-1)^{-1}\{\int x^2 dG_n^*(x) - (\int x dG_n^*(x))^2\}$.

For the particular case of $\bar{X}_n$, $G_n^* = F_n$, so that $T_n^* = \bar{X}_n$ and $V_n^* = s_n^2$. However, for a general *statistical functional* $T(F)$ $(= \theta)$, F_n and G_n^* are generally not equivalent, and moreover, the $T_{n,i}$ in (1.1) are not independent, so that G_n^* may not possess all the properties of F_n . In such a case, the appropriateness of the linear functional in (1.3) may be justified on the basis of the inherent reverse martingale structure of the resampling scheme in jackknifing [viz., Sen(1977)], although, on the ground of robustness, other functionals may appear to be more appropriate. Indeed, Hinkley and Wang(1980) advocated the use of the *trimmed jackknifing* by prescribing the usual trimmed mean on the pseudo variables in (1.1). Thus, one may raise the issue in favor of a general functional :

(1.4) $T_n^{o*} = T^o(G_n^*)$, for suitable $T^o(.)$ on $D[0,1]$;

we term T_n^{o*} a *functional jackknifed estimator (FJE)*.

It can be shown that under fairly general regularity conditions [on $T(.)$ and F], there exists a d.f. G^* (which may depend on F), such that as $n \to \infty$,

(1.5) $||G_n^* - G^*|| = \sup_x |G_n^*(x) - G^*(x)| \to 0$, almost surely (a.s.).

Thus, a minimal requirement for the rationality of the FJE T_n^{o*} in (1.4) is that

(1.6) $T^o(G^*) = T(F)$, for all F belonging to a class $\mathcal{F}$.

We denote by $\mathcal{J}_T$ the class of all $T^o(.)$ for which (1.6) holds (for the given $T(.)$). Then, the consistency of T_n^{o*} (as an estimator of $\theta = T(F)$), in a weaker or stronger sense, can be established under (1.5) and (1.6) , whenever $T^o(.)$ is continuous in the metric in (1.5) (i.e., the uniform topology).

For the asymptotic normality (or weak invariance principles) for the FJE, it is clear that (1.5)-(1.6) may not suffice; we may find it convenient to incorporate the weak convergence of

(1.7) $n^{\frac{1}{2}}(G_n^* - G^*)$, to an appropriate Gaussian function W^*,

along with the usual *Hadamard differentiability* of $T^o(.)$ in a conventional proof. However, the pseudo variables in (1.1) are generally not independent, and, moreover, we have a triangular scheme, for which (1.7) may hold (under additional regularity conditions) but W^* may not have a covariance function reducible to that of a Brownian bridge. Thus, even if (1.7) holds, the asymptotic distribution (or the variance) of the FJE may not agree with that of T_n or the classical jackknifed estimator T_n^*. Generally, under adequate regularity conditions on T(.), for each i (=1,...,n), $T_{n,i}$ can be expressed as a smooth function of X_i, perturbed by another stochastic element α_{ni} which is generally $O_p(n^{-\frac{1}{2}})$. Thus, the usual weak convergence results on the emprical d.f. processes , subject to perturbations [viz., Rao and Sethuraman(1976)] may not be adaptable here. For the trimmed jackknifing, Hinkley and Wang (1980) considered an alternative approach, which may not work out in full generality for a general $T^o(.)$; indeed, (1.7) plays an important role in this context, and it needs to be exploited more systematically.

The main purpose of this study is to focus on the basic regularity conditions on T(.) and $T^o(.)$ and on their fruitful incorporation in the study of the asymptotic theory of FJE. This enables us to study the general robustness of T_n^{o*} and also the *second order asymptotic distributional representations (SOADR)* for the classical jackknifed estimators. For (functional) jackknifing, one usually needs the *second order Hadamard differentiability* of T(.) [and $T^o(.)$], and in Section 2, along with the preliminary notions, SOADR results for T_n^* are considered. The principal asymptotic results on FJE are presented in Section 3. In the light of these, convergence properties of V_n^* and some allied variance estimators are studied in Section 4. The concluding section is devoted to some general discussions and statistical implications of the results obtained in earlier sections.

2. FJE : Preliminary notions. Let $L(A,B)$ be the set of continuous linear transformations from a topological vector space A to another B, and let C be a class of compact subsets of A , such that every subset consisting of a single point belongs to C . Also, let A^{o} be an open subset of A . A function $T: A^{o} \to B$ is said to be *Hadamard (or compact) differentiable* at $F \in A$, if there exists a $T_F' \in L(A,B)$, such that for any $K \in C$,

(2.1) $$\lim_{t\to 0}\{ t^{-1}[T(F+tJ) - T(F) - T_F'(tJ)] \} = 0 ,$$

uniformly for $J \in K$; T_F' is called the *compact derivative of* T at F. We may refer to Fernholz(1983) for a detailed account of such differentiability conditions. In the context of jackknifing, usually, we need the second order compact-differentiability of T (at F), that is, we assume that for any $K \in C$,

(2.2) $$T(F + (G-F)) = T(F) + \int T_1(F;x)d[G(x) - F(x)] + \tfrac{1}{2}\int\int T_2(F;x,y)d[G(x)-F(x)]d[G(y)-F(y)] + R_2(F;G-F) ,$$

where

(2.3) $$|R_2(F;G-F)| = o(||G - F||^2) , \text{ uniformly in } G \in K .$$

The functions $T_1(.)$ and $T_2(.)$ are called the first- and second-order compact derivatives of $T(.)$, and we can always normalize them in such a way that

(2.4) $$\int T_1(F;x)dF(x) = 0 , \quad T_2(F;x,y) = T_2(F;y,x) ,$$

(2.5) $$\int T_2(F;x,y)dF(y) \equiv 0 \equiv \int T_2(F;x,y)dF(x) .$$

We assume that the d.f. F is defined on $R = (-\infty,\infty)$, and denote by

(2.6) $$F_n(x) = n^{-1}\Sigma_{i=1}^{n} I(X_i \le x) , \quad x \in R ;$$

(2.7) $$F_{n-1}^{(i)}(x) = (n-1)^{-1}\Sigma_{j=1(\neq i)}^{n} I(X_j \le x), \quad x \in R ; \; i =1,\dots,n.$$

Then, from (1.1) and (2.6)-(2.7), we obtain that

(2.8) $$T_{n,i} = T(F_n) + (n-1)\{ T(F_n) - T(F_{n-1}^{(i)}) \} , \text{ for } i =1,\dots,n.$$

Note that $F_n(x) = n^{-1}\Sigma_{i=1}^{n} F_{n-1}^{(i)}(x)$, $x \in R$, and further,

(2.9) $$\max_{1\le i\le n} \{ \sup_x | F_{n-1}^{(i)}(x) - F_n(x) | \} = n^{-1} , \text{ for every } n \ge 1.$$

Therefore, by (2.2), (2.3), (2.4), (2.5), (2.8) and (2.9), we have

$$(2.10)\quad T_{n,i} = T_n + (n-1)\{ \int T_1(F_n;x)d[F_n(x)-F_{n-1}^{(i)}(x)] - \tfrac{1}{2}\int\int T_2(F_n;x,y)d[F_n(x)-F_{n-1}^{(i)}(x)]d[F_n(y)-F_{n-1}^{(i)}(y)]\} + o(n^{-1})$$

$$= T_n + \int T_1(F_n;x)d[I(X_i \le x) - F_n(x)] - (2(n-1))^{-1}\int\int T_2(F_n;x,y)d[I(X_i\le x)-F_n(x)]d[I(X_i\le y)-F_n(y)] + o(n^{-1})$$

$$= T_n + T_1(F_n;X_i) - (2(n-1))^{-1}T_2(F_n;X_i,X_i) + o(n^{-1}),\ i=1,\dots,n.$$

As a result, by (1.2), (2.4) and (2.10), we obtain that

$$(2.11)\quad T_n^* = T_n - (2(n-1))^{-1}\int T_2(F_n;x,x)dF_n(x) + o(n^{-1}).$$

Thus, we have the second order representation of the classical jackknifed estimator:

$$(2.12)\quad R_n^* = n(T_n^* - T_n) = -n(2(n-1))^{-1}\int T_2(F_n;x,x)dF_n(x) + o(1)$$

$$= -(2(n-1))^{-1} \Sigma_{i=1}^n T_2(F_n;X_i,X_i) + o(1).$$

Thus, the key role is played by the second order von Mises' functional

$$(2.13)\quad \bar{T}_{2n}^* = n^{-1}\Sigma_{i=1}^n T_2(F_n;X_i,X_i) = \int T_2(F_n;x,x)dF_n(x),$$

and, we have

$$(2.14)\quad (n-1)(T_n^* - T_n) + \tfrac{1}{2}\bar{T}_{2n}^* \to 0 \text{ as } n\to\infty .$$

In fact, by (2.2) through (2.5) and the fact that $||F_n-F|| = O_p(n^{-1/2})$, we have

$$(2.15)\quad T_n = T(F) + \int T_1(F;x)d[F_n(x)-F(x)] + \tfrac{1}{2}\int\int T_2(F;x,y)d[F_n(x)-F(x)]d[F_n(y)-F(y)] + o_p(n^{-1})$$

$$= T(F) + n^{-1}\Sigma_{i=1}^n T_1(F,X_i) + (2n^2)^{-1}\Sigma_{i=1}^n T_2(F;X_i,X_i) + (2n^2)^{-1} \Sigma_{1\le i\ne j\le n} T_2(F;X_i,X_j) + o_p(n^{-1})$$

$$= T(F) + \bar{T}_{1n} + (2n)^{-1}\bar{T}_{2n} + ((n-1)/2n)U_n^{(2)} + o_p(n^{-1}) ,$$

where

$$(2.16)\quad \bar{T}_{1n} = n^{-1}\Sigma_{i=1}^n T_1(F;X_i) \text{ and } \bar{T}_{2n} = n^{-1}\Sigma_{i=1}^n T_2(F;X_i,X_i)$$

are averages over i.i.d.r.v.'s, while

$$(2.17)\quad U_n^{(2)} = \binom{n}{2}^{-1}\Sigma_{1\le i<j\le n} T_2(F;X_i,X_j)$$

is a Hoeffding (1948) U-statistic with mean zero [by (2.5)]and is stationary of order 1 (i.e., whenever $E_F T^2 < \infty$, $E(U_n^{(2)})^2 = O(n^{-2})$). Thus, from (2.14) through (2.17), we obtain that as $n\to\infty$, (2.18) $(n-1)\{ T_n^* - T(F) - \bar{T}_{1n}\} + \frac{1}{2}(\bar{T}_{2n}^* - \frac{n-1}{n} \bar{T}_{2n}) - (\frac{n-1}{n})^2 \frac{1}{2}nU_n^{(2)} \overset{P}{\to} 0$.

Though $EU_n^{(2)} = 0$, $\nu = E_F T_2(F;X_1,X_1)$ may not be equal to 0. However, note that

(2.19) $\bar{T}_{2n}^* - \bar{T}_{2n} = \int [T_2(F_n;x,x) - T_2(F;x,x)]dF_n(x)$,

so that if the functional in (2.13) is continuous in the metric $||F_n - F||$ and ν is finite, by using the fact that $||F_n - F|| \to 0$ a.s., as $n \to \infty$, we claim that

(2.20) $\bar{T}_{2n}^* - \bar{T}_{2n} \to 0$ a.s. and $\bar{T}_{2n}^* \to \nu$ a.s., as $n \to \infty$.

Consequently, on letting $R_n^{**} = (n-1)(T_n^* - T(F) - \bar{T}_{1n})$, we have

(2.21) $(n-1)\{ T_n^* - T_n\} + \nu/2 \to 0$ a.s., as $n \to \infty$;

(2.22) $R_n^{**} - \frac{1}{2}(n U_n^{(2)}) \overset{P}{\to} 0$, as $n \to \infty$.

At this stage, we note that whenever $0 < \sigma_1^2 = E_F T_1^2(F;X_1) < \infty$,

(2.23) $n^{\frac{1}{2}} \bar{T}_{1n} \xrightarrow{\mathfrak{D}} \mathcal{N}(0, \sigma_1^2)$,

while, under $E_F T_2^2(F;X_1,X_2) < \infty$, it follows from Gregory (1977) that there exist a set of (finite or infinite collection of) eigenvalues of $T_2(.)$ corresponding to orthonormal eigen-functions $\{\tau_k(.), k \geq 0\}$, such that

(2.24) $\int T_2(F;x,y)\tau_k(x)dF(x) = \lambda_k\tau_k(y)$ a.e. (F) , $\forall$ $k \geq 0$,

(2.25) $\int \tau_k(x)\tau_q(x)dF(x) = \delta_{kq}$, $\forall$ $k,q \geq 0$,

where δ_{kq} is the usual Kronecker delta. Note that the $\tau_k(.)$ and λ_k may as well depend on F. Then, as $n \to \infty$,

(2.26) $P_F\{ n\cdot U_n^{(2)} \leq x \} \to P\{ \Sigma_{k\geq 0} \lambda_k(Z_k^2 - 1) \leq x \}$, $x \in R$; $\lambda_0 = 0$,

where the Z_k are i.i.d.r.v.'s having the standard normal distribution.[An extension of this result for the joint asymptotic d.f. of $n^{\frac{1}{2}} \bar{T}_{1n}$ and $n\cdot U_n^{(2)}$ is due to Hall (1979).] Thus, we arrive at the following :

THEOREM 2.1. *Under the assumed regularity conditions, as* $n \to \infty$,

(2.27) $R_n^* + \nu/2 \to 0$, *a.s.* ,

(2.28) $2R_n^{**} \xrightarrow{\mathfrak{D}} \Sigma_{k\geq 0} \lambda_k(Z_k^2 - 1)$,

where the λ_k *and* Z_k *are defined as in (2.24)-(2.26).*

Now, (2.27) is an improvement over (3.20) of Sen(1977). It not only provides a sharper result on R_n^* , but also yields the even stronger result that for $\nu = 0$, $R_n^* \to 0$ a.s., as $n \to \infty$. Also, (2.28) provides the SOADR result for the classical jackknifed estimator under second order hadamard differentiability conditions. This result is different from the parallel SOADR results for M-estimators, studied recently by Jurečková(1985) and Jurečková and Sen(1986), among others. (2.27) also suggests that jackknifing (in the classical sense) essentially amounts to a second order bias adjustment (i.e., $n^{-1}\nu/2$) without making any other functional change in T_n. The positive effect of this feature is that if T_n is any robust estimator, T_n^* also remains so, while the negative effect is that the classical jackknifing while reducing the bias (to a smaller order) shares the same lack of robustness property with the initial estimator T_n when the later is not so robust. This negative feature generally calls for the choice of more robust initial estimators and/or the use of functional jackknifing to induce robustness.

Looking at (1.1) and (2.10), we gather that the FJE T_n^{o*} in (1.4) can generally be interpreted as a measure of location of the $T_{n,i}$. In this respect, for any d.f. G , defined on R, we define $G(x;a) = G(x-a)$, for $x,a \in R$. Then, a statistical functional $\tau(G)$ is termed *translation-equivariant* if for every $a \in R$ and $G \in \mathfrak{F}$,

(2.29) $$\tau(G(.;a)) = a + \tau(G(.;0)).$$

It is clear that the functional for T_n^* in (1.3) is translation-equivariant. Similarly, for the trimmed jackknifing, considered by Hinkley and Wang(1980), the corresponding functional is also translation-equivariant. We assume that $T^o(.)$ satisfies (2.29). For notational simplicity, we let

(2.30) $$T_{n,i}^* = T_{n,i} - T_n \text{ , for } i=1,\dots,n;$$

(2.31) $$G_n^{**}(x) = n^{-1} \Sigma_{i=1}^n I(T_{n,i}^* \leq x) \text{ ,} x \in R \text{ .}$$

Then, by (1.4) and (2.31), for translation-equivariant $T^o(.)$, we have

(2.32) $$T_n^{o*} = T^o(G_n^*) = T_n + T^o(G_n^{**}) \text{ .}$$

Looking at (2.10), we also define

(2.33) $$G_n^{o*}(x) = n^{-1} \Sigma_{i=1}^n I(T_1(F;X_i) \leq x) \text{ , } x \in R \text{ ;}$$

(2.34) $$G_F^{o*}(x) = P_F\{ T_1(F;X_1) \leq x \} \text{ , } x \in R \text{ .}$$

Note that by (2.4), $E_F T_1(F;X_1) = 0$. Thus, identifying G^* as the d.f. of $T(F) + T_1(F,X_1)$, we may replace (1.6) by the following :

(2.35) $T^{o}(.)$ is translation-equivariant with $T^{o}(G_F^{o*}) = 0$.

For the trimmed jackknifing, Hinkley and Wang (1980) assumed that $T_1(F;X_i)$ has a d.f. symmetric about 0, and this ensures (2.35), while, for the classical jackknifing, $E_F T_1(F;X_i) = 0$ ensures (2.35). Other conditions on $T(.)$ and $T^{o}(.)$ will be introduced as and when they are needed.

3. FJE: General asymptotics. Note that by virtue of (2.4)-(2.5), the Hadamard derivative of $T_1(F;x)$ is given by

(3.1) $T_{1,1}(F;x,y) = T_2(F;x,y) - T_1(F;y)$.

Thus, for every $i(=1,\ldots,n)$, we have

$$T_1(F_n;X_i) = T_1(F;X_i) + \int T_{1,1}(F;X_i,y)d[F_n(y)-F(y)] + o(||F_n - F||)$$
$$= T_1(F;X_i) + n^{-1}\Sigma_{j=1}^{n} T_2(F;X_i,X_j) - n^{-1}\Sigma_{j=1}^{n} T_1(F;X_j) + o_p(n^{-\frac{1}{2}})$$

(3.2) $$= T_1(F;X_i) - \bar{T}_{1n} + \check{T}_{2n}(X_i) + o_p(n^{-\frac{1}{2}}) ,$$

where $\bar{T}_{1n}$, defined by (2.16), is $O_p(n^{-\frac{1}{2}})$, while under $E_F T_2^2(.) < \infty$, by (2.4)-(2.5), $E\check{T}_{2n}^2(X_i) = O(n^{-1})$, so that $\check{T}_{2n}(X_i) = O_p(n^{-\frac{1}{2}})$. This shows that $T_1(F_n;X_i)$ can be expressed as $T_1(F;X_i)$ perturbed by stochastic factors $[O_p(n^{-\frac{1}{2}})]$ which may play a dominant role in the weak convergence of the related empirical processes. As such, we study first these weak convergence results and incorprate them subsequently in the study of the main results on the FJE. By (2.10), (2.15) and (3.2), we have

(3.3) $T_{n,i} = T(F) + T_1(F;X_i) + \check{T}_{2n}(X_i) + o_p(n^{-\frac{1}{2}})$, for every $i = 1,\ldots,n$.

Thus, if we define the empirical d.f. $\hat{G}_n^*$ by letting

(3.4) $\hat{G}_n^*(x) = n^{-1}\Sigma_{i=1}^{n} I(T_1(F;X_i) + \check{T}_{2n}(X_i) \leq x)$, $x \in R$,

then, noting that G_n^* is the empirical d.f. of the $T_{n,i}$, we obtain by using (3.3) and some standard steps that

(3.5) $n^{\frac{1}{2}} ||G_n^*(x) - \hat{G}_n^*(x - T(F))|| \overset{p}{\rightarrow} 0$, as $n \rightarrow \infty$.

As such, first, we consider the weak convergence of the empirical process related to $\hat{G}_n^*$. We define then a process $\omega_n = \{ \omega_n(t), t \in [0,1] \}$, by letting

(3.6) $\omega_n(t) = n^{-\frac{1}{2}}\Sigma_{j=1}^{n} T_2(F; F^{-1}(t), X_j)$, $t \in [0,1]$.

Then, by (2.4)-(2.5), $E_F\omega_n(t) = 0, \forall t \in [0,1]$, and under $E_F T_2^2 < \infty$, for every

$0 \le s \le t \le 1$,

(3.7) $E[\omega_n(t) - \omega_n(s)]^2 = \int_{-\infty}^{\infty} \{T_2(F;F^{-1}(t),y) - T_2(F;F^{-1}(s),y)\}^2 dF(y) \quad (< \infty)$.

We assume that there exist some finite and positive numbers K, γ and β, such that

(3.8) $E|\omega_n(t) - \omega_n(s)|^{\gamma} \le K|t-s|^{1+\beta}, \quad 0 \le s \le t \le 1.$

Note that whenever $T_2(.)$ satisfies a local Lipschitz condition, (3.7) ensures (3.8) with $\gamma = 2$ and $\beta = 1$. Also, we may note that (3.8) ensures the *tightness* of $\omega_n(.)$, while the convergence of the finite dimensional distributions (f.d.d.) of $\omega_n(.)$ to those of a Gaussian function $\omega = \{\omega(t), t \in [0,1]\}$ follows firectly by a routine use of the standard central limit theorem for i.i.d.r.v.'s. Thus, we conclude that

(3.9) $\sup\{ |\omega_n(t)| : t \in [0,1] \} = O_p(1)$,

and, for every $\varepsilon > 0$ and $\eta > 0$, there exist a positive $\delta_o (< 1)$ and an integer n_o, such that for every $n \ge n_o$ and $\delta : 0 < \delta \le \delta_o$,

(3.10) $P\{ \sup\{ |\omega_n(t) - \omega_n(s)| : 0 \le s \le t \le s+\delta \le 1\} > \varepsilon \} < \eta$.

Then, defining G_n^{o*} as in (2.33), we have the following.

Lemma 3.1. Whenever ω_n converges weakly to ω, i.e., (3.9) and (3.10) hold, as $n \to \infty$,

(3.11) $\sup_x \{ n^{1/2} | \hat{G}_n^*(x) - G_n^{o*}(x - n^{-1/2}\omega_n(F^{-1}(t_x))) | \} \xrightarrow{p} 0$,

where $x = T_1(F;F^{-1}(t_x))$, $x \in R$.

Proof. By (3.9), for every $\varepsilon > 0$, there exists a positive K $(= K_\varepsilon)$, such that

(3.12) $P\{ \sup_{0 \le t \le 1} |\omega_n(t)| > K \} < \varepsilon \quad , \forall\, n \ge n_o$.

Note that by definition [in (3.2)], $\check{T}_{2n}(X_i) = n^{-1/2}\omega_n(F(X_i))$, so that by (3.12), $\max\{ |\omega_n(F(X_i))| : 1 \le i \le n\} \le K$, with a probability $\ge 1 - \varepsilon$, $\forall\, n \ge n_o$. Consider a partition of R into $(-\infty, x-2n^{-1/2}K)$, $[x-2n^{-1/2}K, x+2n^{-1/2}K]$, $(x+2n^{-1/2}K, \infty)$, for a given x. For all i : $T_1(F;X_i) < x - 2n^{-1/2}K$, by (3.12), $T_1(F;X_i) + \check{T}_{2n}(X_i) = T_1(F;X_i) + n^{-1/2}\omega_n(F(X_i)) < x - n^{-1/2}K \le x - n^{-1/2}\omega_n(F^{-1}(t_x))$, with a probability $\ge 1 - \varepsilon$; similarly, for all i : $T_1(F;X_i) \ge x + 2n^{-1/2}K$, $T_1(F;X_i) + \check{T}_{2n}(X_i) > x + n^{-1/2}\omega_n(F^{-1}(t_x))$, with a probability $\ge 1 - \varepsilon$. Finally, by (3.10), for all i: $x - 2n^{-1/2}K \le T_1(F;X_i) \le x + 2n^{-1/2}K$, we have

(3.13) $T_1(F;X_i) + \check{T}_{2n}(X_i) = T_1(F;X_i) + n^{-1/2}\omega_n(F^{-1}(t_x)) + o_p(n^{-1/2})$, $\forall\, n \ge n_o$.

The rest of the proof follows by some standard arguments, and hence, is omitted.

We define G_F^{o*} as in (2.34) and assume that it has a continuous density function g_F^{o*} a.e. Then, on using (3.9), we obtain that as $n \to \infty$,

$$(3.14) \quad \sup_x |n^{\frac{1}{2}}\{G_F^{o*}(x) - G_F^{o*}(x-n^{-\frac{1}{2}}\omega_n(F^{-1}(t_x))) - n^{-\frac{1}{2}}g_F^{o*}(x)\omega_n(F^{-1}(t_x))\}| \overset{p}{\to} 0.$$

Thus, by (3.5), (3.9), (3.11), (3.14) and the fact that by the classical Kolmogorov-Smirnov result, $|| G_n^{o*} - G_F^{o*} || = O_p(n^{-\frac{1}{2}})$, we obtain that

$$(3.15) \quad n^{\frac{1}{2}}|| G_n^{*}(x) - G_F^{o*}(x - T(F)) || = O_p(1).$$

Similarly, defining G_n^{**} as in (2.31), we have

$$(3.16) \quad n^{\frac{1}{2}}|| G_n^{**}(x) - G_F^{o*}(x) || = O_p(1).$$

Thus, by (2.4), (2.32), (2.35), (3.2) and (3.16), we have

$$\begin{aligned}(3.17) \quad n^{\frac{1}{2}}(T_n^{o*} - T_n) &= n^{\frac{1}{2}} T^{o}(G_n^{**}) \\ &= n^{\frac{1}{2}} \int T_1^{o}(G_F^{o*};x)d[G_n^{**}(x) - G_F^{o*}(x)] + o_p(1) \\ &= n^{\frac{1}{2}} \int T_1^{o}(G_F^{o*};x)dG_n^{**}(x) + o_p(1) \\ &= n^{-\frac{1}{2}} \Sigma_{i=1}^{n} T_1^{o}(G_F^{o*}; T_{n,i}^{*}) + o_p(1) .\end{aligned}$$

Thus, if we define the first order remainder term R_n^{o*} by the left hand side of (3.17), we obtain that

$$(3.18) \quad R_n^{o*} = o_p(1) \quad \text{when} \quad n^{-\frac{1}{2}}\Sigma_{i=1}^{n}T_1^{o}(G_F^{o*}; T_{n,i}^{*}) = o_p(1) .$$

Note that (3.18) holds for the particular functional $T^{o}(G) = \int xdG$ (i.e., for the classical jackknifing), but not, in general, for the FJE (as may easily be verified with the case of the trimmed jackknifing, treated in Hinkley and Wang (1980)). Hence for FJE, T_n^{o*} and T_n (or T_n^{*}) are not asymptotically equivalent (upto the order $n^{-\frac{1}{2}}$), so that stronger results (such as in Theorem 2.1) may not be generally true for the FJE. Thus, for the FJE, there remains the issue of a representation for the first order remainder term R_n^{o*} (in terms of independent summands) as well as for the normalized form $n^{\frac{1}{2}}(T_n^{o*} - T(F))$. These will be studied here.

Note that if $T_2(.)$ is regular in the sense that

$$(3.19) \quad E_F\{ T_2(F_n; X_1,X_1) - T_2(F; X_1,X_1)\}^2 \to 0 \text{ , as } n \to \infty \quad ,$$

we have

$$(3.20)\quad n^{-1}\{\max_{1\le i\le n} | T_2(F_n;X_i,X_i) - T_2(F;X_i,X_i)|^2\}$$
$$\le n^{-1}\Sigma_{i=1}^{n} [T_2(F_n;X_i,X_i) - T_2(F;X_i,X_i)]^2$$
$$= \int [T_2(F_n;x,x) - T_2(F;x,x)]^2 dF_n(x) \overset{p}{\to} 0, \text{ as } n\to\infty.$$

Hence, from (2.10), (2.15)-(2.18), (2.30) and (3.2), we have

$$(3.21)\quad \max_{1\le i\le n} | T^*_{n,i} - T_1(F;X_i) - \bar{T}_{1n} - \check{T}_{2n}(X_i) | = o_p(n^{-\frac{1}{2}}).$$

Thus, by the translation-equivariance of $T^o(.)$ and (3.21), we obtain that

$$(3.22)\quad T_n^{o*} = T_n - \bar{T}_{1n} + T^o(\hat{G}_n^*) + o_p(n^{-\frac{1}{2}}).$$

Hence, parallel to (3.17), we have

$$(3.23)\quad R_n^{o*} = n^{\frac{1}{2}}(T_n^{o*} - T_n) = n^{\frac{1}{2}}[T^o(\hat{G}_n^*) - \bar{T}_{1n}] + o_p(1)$$
$$= n^{-\frac{1}{2}}\Sigma_{i=1}^{n} \{T_1^o(G_F^{o*}; T_1(F;X_i)+\check{T}_{2n}(X_i)) - T_1(F;X_i)\} + o_p(1).$$

Hence, if we assume that $T_1^o(.)$ admits the following expansion:

$$(3.24)\quad T_1^o(G_F^{o*};T_1(F;X_i)+n^{-\frac{1}{2}}t) = T_1^o(G_F^{o*};T_1(F;X_i)) + n^{-\frac{1}{2}}t\, T_{11}^o(G_F^{o*};T_1(F;X_i)) + o_p(n^{-\frac{1}{2}}),\ \forall\, |t| \le T < \infty,$$

and denote by

$$(3.25)\quad T_1^o(G_F^{o*};T_1(F;x)) - T_1(F;x) = \psi_F(x),\ x\in R,$$

then, from (3.23) through (3.35), we obtain that

$$(3.26)\quad R_n^{o*} = n^{-\frac{1}{2}}\Sigma_{i=1}^{n} \psi_F(X_i) + n^{-\frac{1}{2}}\Sigma_{i=1}^{n} \check{T}_{2n}(X_i)T_{11}^o(G_F^{o*};T_1(F;X_i)) + o_p(1).$$

Next, we note that

$$(3.27)\quad n^{-\frac{1}{2}}\Sigma_{i=1}^{n} \check{T}_{2n}(X_i)T_{11}^o(G_F^{o*};T_1(F;X_i))$$
$$= n^{-3/2}\Sigma_{i=1}^{n}\Sigma_{j=1}^{n} T_2(F;X_i,X_j)T_{11}^o(G_F^{o*}; T_1(F,X_i))$$
$$= n^{-3/2}\Sigma_{i=1}^{n} T_2(F;X_i,X_i)T_{11}^o(G_F^{o*};T_1(F;X_i)) + n^{-3/2}\Sigma_{1\le i\ne j\le n} T_2(F;X_i,X_j)T_{11}^o(G_F^{o*};T_1(F;X_i))$$
$$= \tilde{U}_{n(1)} + \tilde{U}_{n(2)},\ \text{say.}$$

Note that if we assume that

(3.28) $\upsilon(F) = E_F| T_2(F;X_i,X_i)T_{11}^o(G_F^{o*};T_1(F;X_i))|$ exists and is finite,

then, by the Chebyshev inequality, we obtain that

$$(3.29)\quad \tilde{U}_{n(1)} = O_p(n^{-\frac{1}{2}}).$$

Similarly, we may write

(3.30) $\tilde{U}_{n(2)} = n^{\frac{1}{2}}(1-n^{-1})\{\binom{n}{2}^{-1}\Sigma_{1\leq i<j\leq n}\ \phi_F(X_i,X_j)\}$,

where $\phi_F(x,y) = [\ T_2(F;x,y)T^o_{11}(G^{o*}_F;T_1(F;x)) + T_2(F;x,y)T^o_{11}(G^{o*}_F;T_1(F;y))]/2$ is a symmetric kernel of degree 2. Note that by (2.4)-(2.5), $E_F\phi_F(X_i,X_j) = 0$, but $E_F\phi_F(x,X_i)$ need not be equal to 0 (a.e.), so that $\tilde{U}_{n(2)}$ will generally have asymptotically normal distribution. Thus, if we assume that

(3.31) $E_F\phi^2_F(X_1,X_2) < \infty$,

and denote by

(3.32) $\phi_{F,1}(x) = E_F\phi_F(x,X_1)$, $x \in R$ and $\zeta_1(F) = E_F\phi^2_{F,1}(X_1)$,

then, by the classical results of Hoeffding (1948) [on U-statistics], we have

(3.33) $\tilde{U}_{n(2)} - 2n^{-\frac{1}{2}}\Sigma^n_{i=1}\ \phi_{F,1}(X_i) \overset{p}{\to} 0$;

(3.34) $2n^{-\frac{1}{2}}\Sigma^n_{i=1}\ \phi_{F,1}(X_i) \sim \mathcal{N}(0,\ 4\zeta_1(F))$.

We may recall,at this stage, that by (2.5)

(3.35) $2\phi_{F,1}(x) = \int T_2(F;x,y)T^o_{11}(G^{o*}_F;T_1(F;y))dF(y)$ a.e.,

so that from (3.26), (3.27), (3.29), (3.30), (3.33), (3.34) and (3.35), we arrive at the following .

THEOREM 3.1. *Under the regularity conditions assumed above,*

(3.36) $R^{o*}_n = n^{-\frac{1}{2}}\Sigma^n_{i=1}\{\ \psi_F(X_i) + 2\ \phi_{F,1}(X_i)\ \} + o_p(1)$,

where ψ_F *and* ϕ_1 *are defined by (3.25) and (3.35), respectively. Thus,*

(3.37) $R^{o*}_n \sim \mathcal{N}(0,\ \gamma^{o*2})$; $\gamma^{o*2} = E_F[\ \psi_F(X_1)+2\phi_{F,1}(X_1)]^2$.

Therefore, the FJE T^{o*}_n *and* T_n *are asymptotically (first order) equivalent when*

(3.38) $T^o_1(G^{o*}_F;T_1(F;x)) - T_1(F;x) + \int T_2(F;x,y)T^o_{11}(G^{o*}_F;T_1(F;y))dF(y) = 0$ (a.e. F).

It is easy to verify that (3.38) may not generally hold for the FJE , although for the classical jackknifed estimator, it holds trivially. Also, by (2.15) and (3.22), we have

(3.39) $n^{\frac{1}{2}}(\ T^{o*}_n - T(F)) = n^{\frac{1}{2}}\ T^o(\hat{G}^*_n\ ) + o_p(1)$

$= n^{-\frac{1}{2}}\ \Sigma^n_{i=1}\{\ T^o_1(G^{o*}_F;T_1(F;X_i)) + 2\ \phi_{F,1}(X_i)\} + o_p(1)$

$= n^{-\frac{1}{2}}\ \Sigma^n_{i=1}\{\ T^o_1(G^{o*}_F;T_1(F;X_i)) + \int T_2(F;X_i,y)T^o_{11}(G^{o*}_F;T_1(F;y))dF(y)\} + o_p(1).$

$= n^{-\frac{1}{2}}\ \Sigma^n_{i=1}\ T^{o*}_1(T,F;\ X_i) + o_p(1)$, say.

In this setup, $T_1^{o*}(T,F;x)$ may be identified as the *influence function* of the functional $T^o(.)$ at the point F ; it depends also on the functional T(.) through the $T_{n,i}$. We denote by

$$(3.40) \qquad \sigma_o^2 = \int \{T_1^{o*}(T,F;x)\}^2 dF(x) = E_F\{T_1^{o*}(T,F:X_1)\}^2.$$

THEOREM 3.2. *Under the assumed regularity conditions, for the FJE,*

$$(3.41) \qquad n^{\frac{1}{2}}(T_n^{o*} - T(F)) = n^{-1/2} \Sigma_{i=1}^n T_1^{o*}(T,F;X_i) + o_p(1)$$
$$\sim \mathcal{N}(0, \sigma_o^2).$$

It may be noted that for the classical jackknifed estimator T_n^* , the influence function reduces to $T_1(F;x)$, so that in (3.41), σ_o^2 reduces to σ_1^2 , defined in (2.23). However, for a general FJE, the two influence functions, $T_1^{o*}(T,F:x)$ and $T_1(F;x)$,are not the same, and this makes it possible to choose $T^o(.)$ skillfully, so that the corresponding influence function reflects robustness in a generally interpretable manner [viz., Huber (1981)]. This is perhaps the main point in favor of using a FJE instead of the classical jackknifing. However, we may note that whereas the asymptotic variance of the original and the classical jackknifed estimators are the same , for a general FJE, σ_o^2 in (3.40) may not be equal to σ_1^2 in (2.23). In fact, if T_n is an asymptotically optimal estimator of T(F), then σ_o^2 is $\geq \sigma_1^2$, so that while advocating more robustness, the FJE may lead to some loss of efficiency. On the other hand, if T_n is not an asymptotically efficient estimator of T(F), then it may be possible to choose an appropriate FJE (i.e, $T^o(.)$), such that we may have simultaneously more robust and more efficient estimator of T(F). However, for T(F), under fairly general regularity conditions, $T_n = T(F_n)$ is an asymptotically optimal (nonparametric) estimator, and hence, the last point generally does not merit serious considerations, although such a consideration may arise in a parametric setup. Since robustness in a local sense is mostly confined to parametric models, FJE may find a better case in the parametrics than in the nonparametrics. We shall make more comments on it in the concluding section.

4. FJE: Estimation of asymptotic variance. We may note that by (1.3), (3.3), (3.9) and the usual law of large numbers for the i.i.d. $T_1(F;X_i)$,

(4.1) $\quad V_n^* \to \sigma_1^2$, in probability, as $n \to \infty$.

In fact, if we assume that for each r (= 0,1,2),

(4.2) $\quad \int T_1^r(G;x)T_2^{2-r}(G;x,x)dG(x)$ is a continuous functional of G in some neighbourhood of F (with respect to the norm $||G - F||$), in the Hadamard-sense,

then, by using (2.10) along with the a.s. convergence of $||F_n - F||$ to 0 and the Khintchine strong law of large numbers on the $T_1(F;X_i)$ (and $T_2(F;X_i,X_i)$), we obtain that under (4.2),

(4.3) $\quad V_n^* \to \sigma_1^2$ a.s. , as $n \to \infty$.

Thus, in the classical jackknifing, V_n^* , based on the pseudo variables in (1.1), serves a useful role in the estimation of the asymptotic variance σ_1^2 .

In view of Theorems 3.1 and 3.2, we may note that for the FJE, in general, V_n^* may not consistently estimate σ_o^2 , the asymptotic variance of $n^{\frac{1}{2}}(T_n^{o*} - T(F))$. In some specific cases, such as the trimmed jackknifing treated in Hinkley and Wang (1980), it may be possible to introduce some sample statistics $h_{ni} = h_n(F_n;X_i)$, $i = 1,\ldots,n$, such that

(4.4) $\quad \max_{1\le i\le n} | h_{ni} - T_1^{o*}(T,F;X_i) | \to 0$, as $n \to \infty$,

(in probability or a.s.), where the $T_1^{o*}(T,F;X_i)$ are defined in (3.39), so that defining $V_n^{o*} = (n-1)^{-1} \Sigma_{i=1}^n (h_{ni} - \bar{h}_n)^2$; $\bar{h}_n = n^{-1}\Sigma_{i=1}^n h_{ni}$, we have

(4.5) $\quad V_n^{o*} \to \sigma_o^2$, as $n \to \infty$,

(in probability or a.s.). For a general FJE, a natural way to choose these h_{ni} is to employ the so called *two-step jackknifing*. Toward this, we notice that $T_n^{o*} = T^o(G_n^*)$, where G_n^* is the empirical d.f. of the $T_{n,i}$, defined in (1.1). Let $T_{n-1}^{(i)}$ (and $T_{n-2}^{(ij)}$) be the statistic T_n computed from a sample of size n-1 (and n-2) deleting X_i (and X_i,X_j) from the given sample of size n, for $i \neq j =1,\ldots,n$. For each i (=1,...,n) , let us then define

(4.6) $\quad T_{n,i:j} = (n-1)T_{n-1}^{(i)} - (n-2)T_{n-2}^{(ij)}$, for $j = 1,\ldots,n$ ($j\neq i$).

If we denote the empirical d.f.'s for the sample observations in the sample of sizes n-1 and n-2 (resulting from the deletion of X_i and (X_i,X_j) from the complete sample of size n) by $F_{n-1}^{(i)}$ and $F_{n-2}^{(ij)}$, respectively, then, using the same expansions as in earlier sections, we have the following:

$$(4.7)\quad \begin{aligned} T_{n,i:j} &= (n-1)T(F_{n-1}^{(i)}) - (n-2)T(F_{n-2}^{(ij)}) \\ &= T(F_{n-1}^{(i)}) - (n-2)[\,T(F_{n-2}^{(ij)}) - T(F_{n-1}^{(i)})\,] \\ &= T(F_n) + [\,T(F_{n-1}^{(i)}) - T(F_n)\,] - (n-2)[\,T(F_{n-2}^{(ij)}) - T(F_{n-1}^{(i)})\,] \\ &= T(F_n) + T_1(F_n;X_j) - (2(n-2))^{-1}[\,2T_2(F_n;X_i,X_j) + T_2(F_n;X_j,X_j)\,] \\ &\qquad + r_{n,i:j}\,, \end{aligned}$$

where

$$(4.8)\quad \max_{1\le i\ne j\le n}\{n|r_{n,i:j}|\} \to 0 \text{ a.s., as } n\to\infty\ .$$

As such, using (2.10) along with (4.7) and (4.8), we have

$$(4.9)\quad \max_{1\le i\ne j\le n}|\,T_{n,i:j} - T_{n,j} + n^{-1}T_2(F_n;X_i,X_j)| = o(n^{-1}) \text{ a.s., as } n\to\infty\ .$$

For the $T_{n,i:j}$, $j=1,\dots,n(\ne i)$, we denote the empirical d.f. by $G_{n-1}^{*(i)}$, for $i=1, \dots, n$, while, as in Section 1, G_n^* stands for the empirical d.f. of the $T_{n,i}$, $i=1,\dots,n$. Then, using (4.9) and a technique very similar to that in Lemma 3.1, it follows that

$$(4.10)\quad \max_{1\le i\le n}\ \sup_x\{\,n^{\frac{1}{2}}\,|\,G_{n-1}^{*(i)}(x) - G_n^*(x)\,|\} \overset{p}{\to} 0\ , \text{ as } n\to\infty\ .$$

At the second step of jackknifing, we identify that the FJE based on the $T_{n,i:j}$ (for $j=1,\dots,n$; $j\ne i$) is noting but $T^o(G_{n-1}^{*(i)})$, for $i=1,\dots,n$. Thus, the pseudo variables generated by these FJE are given by

$$(4.11)\quad nT^o(G_n^*) - (n-1)T^o(G_{n-1}^{*(i)}) = Q_{n,i}\ , \text{ say, for } i=1,\dots,n.$$

Using (4.9), (4.10) and (4.11), we obtain that for $T^o(.)$, satisfying (2.2),

$$(4.12)\quad \begin{aligned} Q_{n,i} &= T^o(G_n^*) - (n-1)[\,T^o(G_{n-1}^{*(i)}) - T^o(G_n^*)] \\ &= T^o(G_n^*) - (n-1)\int T_1^o(G_n^*;x)\,d[G_{n-1}^{*(i)}(x) - G_n^*(x)] + O(n\,||G_{n-1}^{*(i)} - G_n^*||^2\,) \\ &= T^o(G_n^*) - (n-1)\int T_1^o(G_n^*;x)\,dG_{n-1}^{*(i)}(x) + o_p(1) \\ &= T^o(G_n^*) - \Sigma_{j=1(j\ne i)}^n\ T_1^o(G_n^*;\ T_{n,i:j}\,) + o_p(1) \\ &= T_n^{o*} - \Sigma_{j=1(j\ne i)}^n\ \{\,T_1^o(G_n^*;T_{n,j}) - n^{-1}T_2(F_n;X_i,X_j)T_{11}^o(G_n^*;T_{n,j})\} + o_p(1) \end{aligned}$$

$$= T_n^{o*} - n\int T_1^o(G_n^*;x)dG_n^*(x) + T_1^o(G_n^*;T_{n,i}) + n^{-1}\Sigma_{j=1}^n T_2(F_n;X_i,X_j)T_{11}^o(G_n^*;T_{n,j}) - n^{-1}T_2(F_n;X_i,X_i)T_{11}^o(G_n^*;T_{n,i}) + o_p(1)$$

$$= T_n^{o*} + T_1^o(G_n^*;T_{n,i}) + n^{-1}\Sigma_{j=1}^n T_2(F_n;X_i,X_j)T_{11}^o(G_n^*;T_{n,j}) + o_p(1) ,$$

where in the penultimate step, we have made use of (3.20) and similar inequalities on the $T_{11}^o(G_n^*;T_{n,i})$. Thus, making use of (2.4)-(2.5) [on $T_1^o(.)$ and $T_2(.)$], we obtain from (4.12) that

(4.13) $\bar{Q}_n = n^{-1}\Sigma_{i=1}^n Q_{n,i} = T_n^{o*} + 0 + 0 + o_p(1) = T_n^{o*} + o_p(1)$,

so that as $n \to \infty$,

(4.14) $\max_{1\le i\le n} | \{Q_{n,i} - \bar{Q}_n\} - \{T_1^o(G_n^*;T_{n,i}) + n^{-1}\Sigma_{j=1}^n T_2(F_n;X_i,X_j)T_{11}^o(G_n^*;T_{n,j})\}| \overset{p}{\to} 0$.

Note that by definition,

(4.15) $n^{-1} \Sigma_{i=1}^n \{T_1^o(G_n^*; T_{n,i})\}^2 = \int\{T_1^o(G_n^*;x)\}^2 dG_n^*(x)$,

where, in Sectuon 3, we have shown that $||G_n^* - G_F^*|| \overset{p}{\to} 0$, as $n \to \infty$; G_F^* being the true d.f. of $T_1^o(G_F^{o*};T_1(F,X_1))$. Thus, if we assume that $T_1^o(.)$ is square integrable and the functional on the right hand side of (4.15) is continuous in the hadamard sense, then, the left hand side of (4.15) converges in probability to

(4.16) $\int \{T_1^o(G_F^{o*};x)\}^2 dG_F^*(x) = \int \{T_1^o(G_F^{o*};T(F;x))\}^2 dF(x)$.

A very similar treatment applies to the other two terms in the expansion of $n^{-1} \Sigma_{i=1}^n (Q_{n,i} - \bar{Q}_n)^2$, using only the leading terms in (4.14). Thus, if we define

(4.17) $V_n^{**} = (n-1)^{-1} \Sigma_{i=1}^n (Q_{n,i} - \bar{Q}_n)^2$,

we arrive at the following.

THEOREM 4.1. *If the functionals* T *and* T^o *are both second order hadamard differentiable and their Hadamard derivatives satisfy the continuity and integrability conditions discussed before, then defining* σ_o^2 *as in (3.40), we have*

(4.18) $V_n^{**} \to \sigma_o^2$, *in probability, as* $n \to \infty$.

In passing, we may remark that (4.10) may be improved to an a.s. convergence (under no extra regularity conditions), and hence, in (4.12), (4.13) and (4.15), we may also replace $o_p(1)$ by $o(1)$ a.s., as $n \to \infty$, so that (4.18) holds a.s., as $n \to \infty$. Note that the construction of V_n^{**} is based on the FJE at the first step and the classical jackknifing at the second stage. Thus, V_n^{**} may be regarded as a two-step jackknifed variance estimator. This two-step jackknifing avoids the arbitrariness in the choice of the h_{ni} in (4.4) and provides natural estimates of the $T_1^{o*}(T,F;X_i)$.

5. Some general discussions. In Section 3, we have mainly stressed on the asymptotic normality of the FJE and on the representation of the first order remainder term. It is possible to extend the asymptotic normality result to a weak invariance principle for the partial sequence $\{ n^{-\frac{1}{2}}k(T_k^{o*} - T(F)); k \leq n\}$, as $n \to \infty$. A key to this weak invariance principle is provided by the following well known result on the emprical d.f. F_n :

(5.1) $$\max_{1 \leq k \leq n} \sup_x \{n^{-\frac{1}{2}} k \mid F_k(x) - F(x) \mid \} = O_p(1) .$$

As such, in (3.5), Lemma 3.1 and elsewhere, we may replace the G_n^* and G_n^{o*} by G_k^* and G_k^{o*} , n by $n^{-\frac{1}{2}}k$ and take a maximum of the resulting quantities over $k \leq n$, and obtain the same convergence results. Thus, (3.39) may be extended to

(5.2) $$\max_{k \leq n} \{n^{-\frac{1}{2}} \mid k(T^{o*} - T(F)) - \Sigma_{i=1}^{k} T_1^{o*}(T,F;X_i) \mid \} = o_p(1), \text{ as } n \to \infty .$$

Since the weak invariance principle holds for the partial sums $\{n^{-\frac{1}{2}}\Sigma_{i=1}^{k} T_1^{o*}(T,F:X_i)$, $k \leq n\}$ (under the usual square integrability condition), by (5.2), the same result holds for $\{n^{-\frac{1}{2}}k(T_k^{o*} - T(F)), k \leq n\}$. This extends Theorems 3.3.1 and 3.3.2 of Sen (1981) to FJE .

We may recall that in Section 3 , we have stressed the utility of the FJE from the basic consideration of robustness. However, we may note that whereas in the classical jackknifing, under the usual second order Hadamard differentiability of $T(.)$, we are able to reduce the bias of T_n^* to $o(n^{-1})$, such a stronger result on T_n^{o*} may need extra regularity conditions on the $T(.)$ and $T^o(.)$. However, in the asymptotic case, as we have shown that the bias of the FJE is $o(n^{-\frac{1}{2}})$, this refinement may not be really of much importance. Although, the FJE may not thus serve the primary function of bias reduction to the extent the classical jackknifing does. Further, in the FJE, we may need extra manipulations (as in Section 4) to serve the dual role of variance estimation. Thus, in advocating for the use of a general FJE, we should take into account the increased robustness aspects of the FJE which need to be sufficient to counterbalance the deficiency in the other two aspects.

We have confined ourselves , so far, to univariate d.f,'s. There is no problem in considering multivariate d.f.'s (for the X_i); the concepts of compact-differentiability remains in tact for this general case too, and the manipulations can be made on parallel lines. Further, we can also consider a vector of statistical functionals and employ FJE in the same manner as in Sections 1 and 2. Because of the coordinate wise representation, the parallel results for the vector case follow on similar lines. Use of multi-sample statistical functionals for FJE also poses no problem. For example , for a two-sample statistics T_{n_1,n_2} , the pseudo variables in (1.1) should be taken as $n_1n_2T_{n_1,n_2} - (n_1-1)n_2T_{n_1-1,n_2} - n_1(n_2-1)T_{n_1,n_2-1} + (n_1-1)(n_2-1)T_{n_1-1,n_2-1}$, and the results in Sections 2, 3 and 4 can be extended in a natural fashion.

Finally, we may comment on the choice of the functional $T^o(.)$ on which to base the FJE. For the classical jackknifing, the reverse martingale structure of the conditional expectation of T_{n-1} given the non-increasing tail sigma-field $\mathcal{C}_n; n \geq 1$, viz., Sen(1977), provides the desired tools for the asymptotic study. For a general FJE, such a stronger justification has to be found out. Since robustness consideration dominates the choice of a FJE, it has to be done so judiciously that the possible increase in the asymptotic variance σ_o^2 (over σ_1^2) and other negative features with the bias reduction and variance estimation can be justified rationally. Note that (2.35) is not needed for the classical jackknifing, but sans (2.35), the FJE may be undesirable, so that in the robustness picture, the effect of possible departures from (2.35) needs to be studied too. For the trimmed jackknifing, Hinkley and Wang (1980) assumed the symmetry of the d.f. of the $T_1(F;X_i)$ which ensures (2.35). It appears that any departure from this symmetry may affect the trimmed jackknifing, while the classical jackknifing is not affected by this deviation. In parametric models, for local robustness aspects, of course, suitable $T^o(.)$ can be constructed to meet these requirements. However, judged from the nonparametric (global) robustness aspects, such a $T^{O}(.)$ may not work out well. Recently, Fernholz (1983) has worked out neatly the von Mises' Calculus for statistical functionals covering the parametric as well as nonparametric cases. This treatment allows us to have a broader choice of T^{O} satisfying the regularity conditions assumed in the earlier sections. In particular, choice of M-functionals or L-functionals may preserve the global robustness on a more general framework, and may be advocated on general considerations.

REFERENCES

[1] FERNHOLZ, L.T. (1983). *Von Mises Calculus for Statistical Functionals*. Lecture Notes in Statist. 19, Springer-Verlag, New York.

[2] GREGORY, G.G. (1977). Large sample theory for U-statistics and test of fit. *Ann. Statist. 5* ,110-123.

[3] HALL, P. (1979). On the invariance principle for U-statistics. *Stoch. Proc. Appl. 9*, 163-174.

[4] HINKLEY, D. and WANG, H.-L. (1980). A trimmed jackknife. *J. Roy. Statist. Soc. B 42*, 347-356.

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

[6] HUBER, P.J. (1981). *Robust Statistics*. Wiley, New York.

[7] JUREČKOVÁ, J. (1985). Representation of M-estimators with the second order asymptotic distribution. *Statist. Dec. 3*, 263-276.

[8] JUREČKOVÁ, J. and SEN, P.K. (1986). A second order asymptotic distributional representation of M-estimators with discontinuous score functions. *Ann. Probability* (under consideration of publication).

[9] MILLER, R.G. Jr. (1974). An unbalanced jackknife. *Ann. Statist. 2*, 880-891.

[10] PARR, W.C. (1983). Jackknifing differentiable statistical functionals. Tech. Report No. 194, Univ. Florida, Gainesville.

[11] RAO, J.S. and SETHURAMAN, J. (1975). Weak convergence of empirical distribution functions of random variables subject to perturbations and scale factors. *Ann. Statist. 3*, 299-313.

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

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