The Annals of Statistics
1985, Vol. 13, No. 3, 1097–1113

ESTIMATION OF A SYMMETRIC DISTRIBUTION[1]

BY SHAW-HWA LO

Rutgers University

Suppose that F_0 is a population which is symmetric about zero, so that $F(\cdot) = F_0(\cdot - \theta)$ is symmetric about θ. We consider the problem of estimating F_0 (shape parameter), both θ and F_0, and F based on a random sample from F. First, some asymptotically minimax bounds are obtained. Then, some estimates are constructed which are asymptotically minimax-efficient (the risks of which achieve the minimax bounds uniformly). Furthermore, it is pointed out that one can estimate F_0, the shape of F, as well without knowing the location parameter θ as with knowing it. After a slight modification, Stone's (1975) estimator is proved to be asymptotically minimax-efficient in the Hellinger neighborhood.

1. Introduction and main results. In their fundamental work, Dvoretzky, Kiefer and Wolfowitz (1956) proved that the empirical distribution function (e.d.f.) is asymptotically minimax for estimating a distribution function (d.f.) belonging to the collection of all continuous d.f.'s. Twenty years later, Kiefer and Wolfowitz (1976) reopened this study and proved that the e.d.f. remains asymptotically minimax even if the collection of distributions is reduced to all concave (or convex) d.f.'s. Recently, Millar (1979) suggested a general problem: Given a collection C of distribution functions, when is the e.d.f. asymptotically minimax for estimation in this class?

In this paper, Millar gave sufficient conditions and some necessary conditions on the class C for making the e.d.f. asymptotically minimax. The collection of distribution functions with increasing failure rate (IFR), and the collection of distribution functions with decreasing failure rate (DFR) all satisfy the sufficient conditions. In the same paper, Millar also considered the estimation of distribution functions among the collection of symmetric distribution functions with a known center θ. He showed that the e.d.f. fails to be asymptotic minimax in this class; instead, the symmetrized e.d.f. around θ serves as an asymptotic minimax estimator.

Now some questions arise: (I) If we assume the symmetry around an unknown center θ, is it still possible to find a minimax estimator for the d.f. and if so what is it?

Suppose that F_0 is a distribution function which is symmetric about 0, so that $F(\cdot) = F_0(\cdot - \theta)$ is symmetric about θ. Consider the problem of estimating θ, F_0,

Received June 1982; revised March 1985.
[1] This research was supported by National Science Foundation Grants MCS-75-10376 and MCS-80-02698.
AMS 1980 *subject classifications*. Primary 62E20, 62G20; secondary 62G30.
Key words and phrases. Gaussian experiments, distribution functions, asymptotically minimax estimators, location parameters.

both θ and F_0, or F based on a random sample from F in each of the following cases:

A. F_0 known, θ unknown,
B. F_0 unknown, θ known ($= 0$ without loss of generality),
C. F_0 unknown, θ unknown,
D. F completely arbitrary.

Case A is the classical one in which the maximum likelihood estimator $\hat{\theta}_n$ is asymptotically optimal (efficient) in many ways. To estimate F in this case, one can simply replace θ by $\hat{\theta}_n$ and this turns out to be an optimal (efficient) estimator of F, with the limiting process Yf, where Y is a $N(0, 1/I(F_0))$ random variable, f is the density of F, and $I(F_0)$ is the usual Fisher information. In Case B, Millar (1979) has shown that e.d.f. symmetrized about zero is an asymptotically minimax estimator of $F_0(=F)$ (with limit process $W_s^0(F_0) = \frac{1}{2}(W^0(F_0) - W^0(1 - F_0))$, where $W^0(F_0)$ is a Brownian bridge process composed with F_0). For Case C, C. Stein has a paper in the *Proceedings of the Third Berkeley Symposium* (1956). In this famous paper, Stein asked the question, "When can one estimate θ without knowing the shape as well asymptotically as one does knowing the shape?" He gave a simple necessary condition which indicated that it is possible to estimate the location parameter θ adaptively. Complete definitive results were later obtained by Beran (1974, 1978), Sacks (1975), and Stone (1975). Bickel (1982) considered more general conditions for adaptation. He came up with some sufficient conditions of constructing estimators adaptively. A counterpart of the question is the following: (II) "Can one estimate the shape without knowing the center θ as well asymptotically as one does knowing it?" Case D has no structure (F is no longer assumed to be symmetric); in this case the e.d.f. is well known (see Dvoretzky, Kiefer and Wolfowitz, 1956; or Millar, 1979) to be an asymptotic minimax estimator of F. The basic aim in the article is to give complete answers to questions (I) and (II).

To answer (I) and (II), we first find asymptotic minimax bounds for estimating F (F_0 and θ jointly) and F_0 in Hellinger shrinking neighborhoods (Theorems 1 and 2). We then give the constructions of estimators $\hat{F}_n$, $\hat{F}_{0n}$ of F and F_0, respectively. We show that $\hat{F}_n$, $\hat{F}_{0n}$ achieve the asymptotic minimax bounds "uniformly" in the neighborhoods (Theorem 4). This is a stronger optimality property than asymptotic efficiency (note that an efficient estimator of F (or F_0) is not necessarily an asymptotically minimax-efficient estimator in this case).

The constructions of $\hat{F}_n$ involve the estimation of unknown center θ. We prove that, after a slight modification, Stone's estimator $\hat{\theta}_n$ (1975) will be asymptotically minimax-efficient (part (ii) of Theorem 3) in the sense that the mean square errors of $\hat{\theta}_n$ are always smaller or equal to $1/I(F)$ uniformly in the Hellinger neighborhood of F as n tends to infinity. Before explaining the meaning and implications of our findings with some details, we state the results as four theorems.

We observe random variables $X_1, X_2, \cdots, X_n$ and assume that they are i.i.d.

from a symmetric d.f. $F(x) = F_0(x - \theta)$ with density $f(x) = f_0(x - \theta)$ on the real line. Let $I(F) = I(F_0) = \int (f_0'/f_0)^2 f_0 \, dx$ denote the usual Fisher information for the location parameter. For any $c > 0$, consider the following Hellinger ball:

$$B_n(F; c) = \left\{ F_n(x) \colon F_n(x) \text{ is a d.f. symmetric about some } \theta_n \right.$$
$$\left. \text{and } h(F_n, F) = \left[\int_{-\infty}^{\infty} (\sqrt{dF_n} - \sqrt{dF})^2 \right]^{1/2} < c/\sqrt{n} \right\}.$$

Let $c(-\infty, \infty)$ denote the collection of all continuous functions on R. Consider the loss function $\ell \colon c(-\infty, \infty) \to R^+$ which is subconvex (i.e., $\ell^{-1}[0, \alpha]$ is closed, convex, and symmetric set in $c(-\infty, \infty)$ for every $\alpha \geq 0$), such as $\ell(y) = \| y \|_\infty = \sup_t | y(t) |$ or $\ell(y) = \int | y(t) |^2 \, dt$. For discussions of these loss functions, the readers are referred to Millar (1979).

The following theorem was first given by Lo (1981); one can derive similar results from a recent paper by Begun et al. (1983, Theorem 4.2, and Remarks 4.4 and 4.5).

THEOREM 1. *Let, F, $B_n(F; c)$ and ℓ be as described above, and let $I(F_0) < \infty$. Then one has the following inequality on the estimation of F:*

$$\lim_{c\to\infty} \lim_{n\to\infty} \inf_b \sup_{F_n \in B_n(F;c)} \int_{R^n} \int_{c(-\infty,\infty)} \ell[n^{1/2}(y - F_n)] b(x, dy) F_n^n(dx) \geq E\, \ell[W_s^0(F) + Yf], \tag{1.1}$$

where F_n^n denote the product of n copies of F_n, the infimum is taken over the class of all generalized procedures (see Millar, 1979, page 235), and $W_s^0(F)$, Yf are two independent processes distributed as described in cases A *and* B, *respectively.*

It is easy to see that $F_n \in B_n(F; c)$ if and only if $F_n(x - \theta) \in B_n(F_0; c)$. Let $\mathscr{F}_n(F_0; c)$ denote the collection of all $B_n(F; c)$ such that F is a translation of F_0 by some value. Let F_{0n} denote the shape of F_n; i.e., F_{0n} is a d.f. symmetric about zero and $F_{0n}(x) = F_n(x + \theta_n)$, where θ_n is the center of F_n.

THEOREM 2. *Suppose the assumptions in Theorem* 1 *hold. Then we have*

$$\begin{aligned} &\lim_{c\to\infty} \lim_{n\to\infty} \inf_b \sup_{F_n \in \mathscr{F}_n(F_0;c)} \int_{R^n} \int_{c(-\infty,\infty)} \ell[n^{1/2}(y - F_{0n})] b(x, dy) F_n^n(dx) \\ &\geq \lim_{c\to\infty} \lim_{n\to\infty} \inf_b \sup_{F_n \in B_n(F;c)} \int_{R^n} \int_{c(-\infty,\infty)} \ell[n^{1/2}(y - F_{0n})] b(x, dy) F_n^n(dx) \\ &\geq E\, \ell[W_s^0(F_0)]. \end{aligned} \tag{1.2}$$

For estimating F_n (or F_{0n}), we first estimate θ_n (= center of F_n).

The following theorem tells us that one can estimate θ_n uniformly well (asymptotically minimax-efficient) in the Hellinger ball $B_n(F; c)$.

THEOREM 3. (i) *There exists a nonrandomized and location invariant estimator* $\bar{\theta}_n$ *such that* $\bar{\theta}_n(-X_1, \cdots, -X_n) = -\bar{\theta}_n(X_1, \cdots, X_n)$, *and* $n^{1/2}(\bar{\theta}_n - \theta_n) = O_p(1)$, *uniformly in* $B_n(F; c)$.

(ii) *There exists a nonrandomized estimator* $\hat{\theta}_n$ *satisfying* (i) *above. Furthermore, there exists a positive random sequence* $\{\delta^2(\hat{\theta}_n; F_n)\}$ *under* $\{F_n\}$ *such that for any sequence* $\{F_n\}_{n=1}^{\infty}$, F_n *from* $\{B_n(F; c)\}$, $\mathscr{L}[n^{1/2}(\hat{\theta}_n - \theta_n)/\delta(\hat{\theta}_n; F_n) \mid F_n] \to N(0, 1)$ *as* $n \to \infty$, *and* $\lim_{n\to\infty}\sup_{F_n \in B_n(F;c)} \delta^2(\hat{\theta}_n; F_n) \le 1/I(F_0)$.

Let $\tilde{F}_n(x)$ denote the usual e.d.f. from a d.f. F_n. Let $\hat{F}_n(x; \theta)$ stand for the e.d.f. symmetrized about θ; i.e.,

$$\hat{F}_n(x; \theta) = \tfrac{1}{2}\{\tilde{F}_n(x) + 1 - \tilde{F}_n(2\theta - x)\}. \tag{1.3}$$

It is clear that $\hat{F}_n(x + \theta; \theta)$ is a symmetric d.f., symmetric about zero.

A loss function is of Kolmogorov type if $\ell(x) = g(\| x \|_\infty)$, where $\| \; \|_\infty$ denotes the sup norm and g is a continuous nondecreasing function defined on $[0, \infty)$.

THEOREM 4. (i) *If* $\hat{\theta}_n$ *satisfies part* (ii) *of Theorem* 3 *and the loss function* ℓ *is bounded and of Kolmogorov type, then*

$$\lim_{c\to\infty}\lim_{n\to\infty}\sup_{F_n\in B_n(F;c)} \int_{R^n} \ell[n^{1/2}(\hat{F}_n(\cdot\,; \hat{\theta}_n) - F_n)]F_n^n(dx) = E\ell[W_s^0(F) + Yf],$$

and hence $\hat{F}_n(\cdot\,; \hat{\theta}_n)$ *is a locally asymptotically minimax-efficient estimator.*

(ii) *If* $\bar{\theta}_n$ *satisfies part* (i) *of Theorem* 3, *then with any bounded continuous loss function* ℓ, $\hat{F}_n(\cdot + \bar{\theta}_n; \bar{\theta}_n)$ *achieves the lower bound in Theorem* 2; *i.e.,*

$$\begin{aligned}
&\lim_{c\to\infty}\lim_{n\to\infty}\sup_{F_n\in\mathscr{F}_n(F_0;c)} \int_{R^n} \ell[n^{1/2}(\hat{F}_n(\cdot + \bar{\theta}_n; \bar{\theta}_n) - F_{0n})]F_n^n(dx)\\
&\quad= \lim_{c\to\infty}\lim_{n\to\infty}\sup_{F_n\in B_n(F;c)} \int_{R^n} \ell[n^{1/2}(\hat{F}_n(\cdot + \bar{\theta}_n; \bar{\theta}_n) - F_{0n})]F_n^n(dx)\\
&\quad= E\ell[W_s^0(F_0)],
\end{aligned}$$

and hence $\hat{F}_n(\cdot + \bar{\theta}_n; \bar{\theta}_n)$ *is an asymptotically minimax-efficient estimator.*

Part (i) of Theorem 4 gives an affirmative answer to question (I). Part (ii) of Theorem 4 not only answers question (II) affirmatively, it also shows that our estimators behave uniformly well (minimax) on $\mathscr{F}_n(F_0; c)$ for large n.

REMARK 1. To see how much we gain from the knowledge of symmetry, note from Theorem 1 that

$$\operatorname{Var}[W_s^0(F)(x) + Yf(x)] = \begin{cases} \dfrac{F(x)[1 - 2F(x)]}{2} + \dfrac{f^2(x)}{I(F_0)} & \text{if } F(x) \le \dfrac{1}{2} \\[2ex] \dfrac{[1 - F(x)][2F(x) - 1]}{2} + \dfrac{f^2(x)}{I(F_0)} & \text{if } F(x) > \dfrac{1}{2}. \end{cases} \tag{1.4}$$

On the other hand,

$$\operatorname{Var}[W^0(F)(x)] = F(x)(1-F(x)). \tag{1.5}$$

(Note that $W^0(F)$ is the limiting process if we use the e.d.f. as our estimator.) Comparing (1.4) and (1.5), we find that the gain is

$$\frac{F(x)}{2} - \frac{f^2(x)}{I(F_0)} \qquad \text{if} \quad F(x) \le \frac{1}{2} \tag{1.6}$$

and

$$\frac{1-F(x)}{2} - \frac{f^2(x)}{I(F_0)} \quad \text{if} \quad F(x) > \frac{1}{2}. \tag{1.7}$$

We claim that both terms (1.6) and (1.7) are nonnegative. To see this, we may assume without loss of generality that f is symmetric about zero.

For any x, define

$$\rho_f(t) = \psi(t) f^{1/2}(t)/f(x),$$

where

$$\psi(t) = \begin{cases} -1 & \text{if} \quad t \le -|x| \\ 0 & \text{if} \quad -|x| < t < |x| \\ 1 & \text{if} \quad t \ge |x|. \end{cases}$$

Let

$$\beta = (-f'/f^{1/2}) - 2(\rho_f/\|\rho_f\|^2).$$

It is easy to check that $\beta \perp \rho_f$ (that is, $\langle \beta, \rho_f \rangle = 0$). So

$$I(F_0) = \|\beta\|^2 + (4/\|\rho_f\|^2)$$

and therefore

$$I(F_0) \ge 4/\|\rho_f\|^2 = 2f^2(x)/F(x).$$

The first part (1.6) of the claim thus follows. With the same arguments one can easily show that (1.7) is also nonnegative.

REMARK 2. Schuster (1975) considered the similar problem and proposed a similar estimator. He showed that the proposed estimator is asymptotically better than the e.d.f. for the cases of normal, double exponential and Cauchy. From Remark 1, one always gains by using this symmetrized estimator as long as $\hat{\theta}_n$ satisfies part (ii) of Theorem 3.

REMARK 3. Our estimators in Theorem 4 may be considered as a continuous version (piecewise linear) of original estimators since our loss functions are defined on $c(-\infty, \infty)$. The difference between the continuous version and the discrete version is $\sqrt{n}$ negligible ($=O(1/n)$).

REMARK 4. By replacing F by $\hat{F}$, one can estimate the real functional $T(F)$ by $T(\hat{F})$ optimally in the sense of Theorem 1 and Theorem 2 if the functional is sufficiently smooth (Hellinger differentiable). An example of such a functional is a quantile of F.

2. Proofs of theorems.

The proof of Theorem 1 was first given in Lo (1981). One can also derive similar results by using Theorem 4.2 (and Remarks 4.4 and 4.5) in a recent paper by Begun, *et al.*, (1983).

We will only provide a sketch of the proof here; for details, see Lo (1981).

PROOF OF THEOREM 1. Let

$$H(F) = \left\{h(x);\ \int_{-\infty}^{\infty} h^2(x)F(dx) < \infty,\ \int_{-\infty}^{\infty} h(x)F(dx) = 0\right\},$$

$$H_s(F) = \{h(x);\ h \in H(F) \text{ and } h(x) = h(2\theta - x)\},$$

$$H_1(F) = \{h(x);\ h(x) = \alpha(f'(x)/f(x)) \text{ for some } \alpha \in R\}.$$

Clearly, $H_s(F) \oplus H_1(F) \subset H(F)$, where "$\oplus$" denotes the direct sum of two orthogonal spaces. We parameterize the distributions near F by the space $H_s(F) \oplus H_1(F)$. By expanding the log likelihood ratio around F in the $c/\sqrt{n}$-Hellinger neighborhood, it is found that the whole experiment under F^n can be approximated by a Standard Gaussian experiment indexed by $H_s(F) \oplus H_1(F)$ as n tends to infinity.

Define a map from $H_s(F) \oplus H_1(F)$ to a Banach space (with sup norm) $B(F)$ as follows:

$$\tau(h)(t) = \int_{-\infty}^{t} h(x)F(dx),$$

where B is the collection of all continuous real functions $\gamma(t)$ defined on $[0, 1]$ satisfying $\gamma(0) = \gamma(1) = 0$, and $B(F) = \{\gamma \circ F;\ \gamma \in B\}$. It can be shown that $(\tau, H_s(F) \oplus H_1(F), B(F))$ forms an abstract Wiener space (see Kuo, 1975).

Let P_0^* denote the Standard Gaussian cylinder measure defined on the Hilbert space $H_s(F) \oplus H_1(F)$. The probability measure derived on $B(F)$ under the mapping τ is the same as the distribution of $W_s^0(F) + Yf$, where $W_s^0(F)$ and Y are two independent processes defined in the Introduction.

Now, using some standard results in asymptotic minimaxity theory for Gaussian experiments (see Propositions 2.1 and 3.1 in Millar, 1979; or Le Cam, 1982) and following the proof of Millar's Proposition 5.1, one establishes the desired results. □

PROOF OF THEOREM 2. The first inequality of (1.2) follows immediately due to the fact that $B_n(F;\ c) \subset \mathscr{F}_n(F_0;\ c)$. The second inequality follows from the same argument of Theorem 1, except here we replace $H_s(F) \oplus H_1(F)$ by $H_s(F)$. □

We can write our estimator $\hat{F}_n(\cdot;\hat{\theta}_n)$ of F_n as follows:

$$\begin{aligned}\hat{F}_n(\cdot;\hat{\theta}_n) &= (1/2n)[\textstyle\sum_{i=1}^n I_{(-\infty,\cdot]}(X_i) + \sum_{i=1}^n I_{(2\hat{\theta}_n-\cdot,\infty)}(X_i)] \\ &= \tfrac{1}{2}[\tilde{F}_n(\cdot) + 1 - \tilde{F}_n(2\hat{\theta}_n - \cdot)],\end{aligned} \tag{2.1}$$

where $X_1, X_2, \cdots, X_n$ are i.i.d. from F_n (with center θ_n), and I_A denotes the indicator function of A.

We can write

$$\begin{aligned}&\hat{F}_n(\cdot;\hat{\theta}_n) - F_n(\cdot) \\ &\quad= [\hat{F}_n(\cdot;\hat{\theta}_n) - F_n(\cdot - (\hat{\theta}_n - \theta_n))] + [F_n(\cdot - (\hat{\theta}_n - \theta_n)) - F_n(\cdot)] \\ &\quad= T_n(\cdot;\hat{\theta}_n) + S_n(\cdot;\hat{\theta}_n) \quad \text{(say)}.\end{aligned} \tag{2.2}$$

PROOF OF THEOREM 3. The constructions of these estimators are similar to those of Huber's M-estimators and Stone's estimators (see Stone (1975)). The details of the construction and the proof are deferred and given in the Appendix.

LEMMA 1. *If $\bar{\theta}_n$ satisfies part* (i) *of Theorem* 3, *then*

$$n^{1/2}\sup_x|S_n(x;\bar{\theta}_n) - [F(x - (\bar{\theta}_n - \theta_n)) - F(x)]| \to_p 0, \tag{2.3}$$

uniformly in $B_n(F;c)$.

Note that θ_n, θ are the centers of F_n and F, respectively.

PROOF. Since $S_n(x;\theta_n)$ can be written as $\int_x^{x-(\bar{\theta}_n-\theta_n)} dF_n$, (2.3) equals

$$\begin{aligned}&n^{1/2}\sup_x\left|\int_x^{x-(\bar{\theta}_n-\theta_n)} (dF_n - dF)\right| \\ &\quad\le n^{1/2}\sup_x\left|\left\{\int_{-\infty}^{\infty}(\sqrt{dF_n} - \sqrt{dF})^2\right\}^{1/2}\right. \\ &\qquad\left.\cdot\left\{\int_{-\infty}^{\infty} I_{[x,x-(\bar{\theta}_n-\theta_n)]\cup[x-(\bar{\theta}_n-\theta_n),x]}(\cdot)(\sqrt{dF_n} + \sqrt{dF})^2\right\}^{1/2}\right| \\ &\quad\le c\,\sup_x\left|\left\{\int_{x-|\bar{\theta}_n-\theta_n|}^{x+|\bar{\theta}_n-\theta_n|}(\sqrt{dF_n} + \sqrt{dF})^2\right\}^{1/2}\right|.\end{aligned}$$

The last expression above tends to zero in probability, uniformly in $B_n(F;c)$, in view of the fact that $\bar{\theta}_n - \theta_n = O_p(n^{-1/2})$ uniformly in $B_n(F;c)$,

$$\int_{x-|\bar{\theta}_n-\theta_n|}^{x+|\bar{\theta}_n-\theta_n|} dF = F(x + |\bar{\theta}_n - \theta_n|) - F(x - |\bar{\theta}_n - \theta_n|) \to_p 0$$

uniformly in $B_n(F;c)$ and $x \in R$, and

$$\left|\int_{x-|\bar{\theta}_n-\theta_n|}^{x+|\bar{\theta}_n-\theta_n|}(dF_n - dF)\right| \le \frac{2c}{\sqrt{n}}. \quad \square$$

LEMMA 2. *Let $\hat{\theta}_n$ be as described in part* (ii) *of Theorem* 3. *Then*

(2.4) $$n^{1/2}S_n(x;\,\hat{\theta}_n)/\delta(\hat{\theta}_n;\,F_n) \to_{\mathscr{L}} N(0,\,f^2(x)),\ \textit{uniformly in } B_n(F;\,c) \textit{ and in } x.$$

PROOF. From Lemma 1, $n^{1/2}S_n(\cdot;\,\theta_n) = n^{1/2}[F(\cdot - (\hat{\theta}_n - \theta_n)) - F(\cdot)] + o_p(1)$, uniformly in $B_n(F;\,c)$. The lemma follows from the fact that

$$n^{1/2}[F(x - (\hat{\theta}_n - \hat{\theta}_n)) - F(x)] = n^{1/2}(\hat{\theta}_n - \theta_n)f(x - \Delta(\hat{\theta}_n - \theta_n))$$

$$\text{for some} \quad 0 \le \Delta \le 1,$$

part (ii) of Theorem 3, and uniform continuity of $f(x)$ (since $I(F_0) < \infty$). □

LEMMA 3. *If $\bar{\theta}_n$ satisfies part* (i) *of Theorem* 3, *then*

(2.5) $$\sup_x |\, n^{1/2}T_n(x;\,\bar{\theta}_n) - n^{1/2}[\hat{F}_n(x;\,\theta_n) - F_n(x)]\,| \to_p 0$$

uniformly in $B_n(F;\,c)$ and therefore

(2.6) $$n^{1/2}T_n(x;\,\theta_n) \to_{\mathscr{L}} W_s^0(F)$$

uniformly in $B_n(F;\,c)$.

PROOF. Clearly,

(2.7) $$\begin{aligned} &n^{1/2}T_n(x;\,\bar{\theta}_n) - n^{1/2}[\hat{F}_n(x;\,\theta_n) - F_n(x)] \\ &= -n^{1/2}[F_n(x - (\bar{\theta}_n - \theta_n)) - F_n(x)] - \tfrac{1}{2}n^{1/2}[\tilde{F}_n(2\theta_n - x) - \tilde{F}_n(2\theta_n - x)]. \end{aligned}$$

By Lemma 1, up to $o_p(1)$, the first term on the right-hand side of (2.7) equals

(2.8) $$-n^{1/2}[F(x - (\bar{\theta}_n - \theta_n)) - F(x)] = n^{1/2}(\bar{\theta}_n - \theta_n)f(x) + o_p(1)$$

uniformly in $F_n \in B_n(F;\,c)$ and $x \in R$. This equality (2.8) follows from the same arguments as in Lemma 2.

The second term on the right-hand side of (2.7) can be further decomposed as

(2.9) $$\begin{aligned} &-\tfrac{1}{2}n^{1/2}[\tilde{F}_n(2\bar{\theta}_n - x) - F_n(2\bar{\theta}_n - x)] \\ &- \tfrac{1}{2}n^{1/2}[F_n(2\bar{\theta}_n - x) - F_n(2\theta_n - x)] \\ &-\tfrac{1}{2}n^{1/2}[F_n(2\theta_n - x) - \tilde{F}_n(2\theta_n - x)]. \end{aligned}$$

Since $|\,\bar{\theta}_n - \theta_n| = O_p(n^{-1/2})$ uniformly in $B_n(F;\,c)$, the first term and the third term of (2.9) cancel each other asymptotically (uniformly in $x \in R$). (See Theorem 2.11 of Stute (1982) or Shorack and Wellner (1982).)

Thus (2.9) equals (up to $o_p(1)$)

(2.10) $$\begin{aligned} &-\tfrac{1}{2}n^{1/2}[F_n(2\bar{\theta}_n - x) - F_n(2\theta_n - x)] \\ &\qquad = -\tfrac{1}{2}n^{1/2}[1 - F_n(x - 2(\bar{\theta}_n - \theta_n)) - 1 + F_n(x)] \\ &\qquad = \tfrac{1}{2}n^{1/2}[F_n(x - 2(\bar{\theta}_n - \theta_n)) - F_n(x)] \\ &\qquad = -n^{1/2}(\bar{\theta}_n - \theta_n)f(x) + o_p(1) \qquad \text{(by Lemma 1).} \end{aligned}$$

The first part of Lemma 3 follows now from (2.8) and (2.10).

The second part of Lemma 3 is a consequence of the first part of this lemma and the fact that $n^{1/2}[\hat{F}_n(x;\theta_n) - F_n(x)] \to_{\mathscr{L}} W_s^0(F)$ uniformly in $B_n(F; c)$. □

LEMMA 4. *The weak limits $N(0, f^2(x))$ and $W_s^0(F)$ obtained in Lemmas 2 and 3 are independent provided that $\hat{\theta}_n$ is location invariant and $\hat{\theta}_n(-x_1, -x_2, \cdots, -x_n) = -\hat{\theta}_n(x_1, x_2, x_n)$.*

PROOF. In view of (2.3) and (2.5), we only have to check that the processes $n^{1/2}[F(x - (\hat{\theta}_n - \theta_n)) - F(x)]\delta(\hat{\theta}_n; F_n)^{-1}$ and $n^{1/2}[\hat{F}_n(x;\theta_n) - F_n(x)]$ are asymptotically uncorrelated (both processes are asymptotically Gaussian) for each x.

To see this, we look at

$$E = \int_{-\infty}^{\infty} \cdots\cdots \int_{-\infty}^{\infty} [\hat{\theta}_n(x_1, x_2, \cdots, x_n) - \theta_n]\hat{F}_n(x;\theta_n) \prod_{i=1}^{n} dF_n(x_i)$$

$$= \int_{\infty}^{-\infty} \cdots \int_{\infty}^{-\infty} [\hat{\theta}_n(2\theta_n - x_1, \cdots, 2\theta_n - x_n) - \theta_n][1 - \hat{F}_n(2\theta_n - x;\theta_n)]$$

$$\cdot \prod_{i=1}^{n} dF_n(2\theta_n - x_i)$$

$$= \int_{\infty}^{-\infty} \cdots \int_{\infty}^{-\infty} [2\theta_n - \hat{\theta}_n(x_1, \cdots, x_n) - \theta_n]\hat{F}_n(x;\theta_n) \prod_{i=1}^{n} d(-F_n(x_i)) = -E,$$

and conclude that $E = 0$. The lemma thus follows. □

LEMMA 5. *Let $W(t)$ be a mean-zero Gaussian process on $(-\infty, \infty)$, and let X be a mean-zero normal random variable independent of $W(t)$. For any continuous function $f(t)$ such that $f(t) \geq 0$, we have*

$$\sup_{0\leq\alpha\leq1} Eg\{\sup_{-\infty<t<\infty} |W(t) + \alpha f(t)X|\} \leq Eg\{\sup_{-\infty<t<\infty} |W(t) + f(t)X|\},$$

where g is a positive nondecreasing function defined on $[0, \infty)$.

PROOF. Since g is a positive monotonic function, we only have to show that for every $\alpha \in [0, 1]$, and $c > 0$,

$$P\{\sup_{-\infty<t<\infty} |W(t) + \alpha f(t)X| \leq c\} > P\{\sup_{-\infty<t<\infty} |W(t) + f(t)X| \leq c\}.$$

Consider a finite set $\{-\infty < t_1 < t_2 < \cdots < \infty\}$. By Anderson's lemma (1954),

$$P\{\sup_{t_i} |W(t_i) + \alpha f(t_i)X| \leq c\} \geq P\{\sup_{t_i} |W(t_i) + f(t_i)X| \leq c\}.$$

The above inequality is true for all finite sets $\{t_i\}_{i=1}^{k}$. Let $\{t_i\}_{i=1}^{\infty}$ be a dense subset of R'. By taking the limit, the above inequality still holds if we take sup over all $\{t_i\}_{i=1}^{\infty}$.

The lemma follows from the fact that $W(t)$ is continuous with probability 1. □

PROOF OF THEOREM 4, PART (i). Let $\hat{\theta}_n$ be the estimator described in part (ii) of Theorem 3. From (2.2), we can write

$$n^{1/2}[\hat{F}_n(\cdot;\hat{\theta}_n) - F_n(\cdot)] = n^{1/2}T_n(\cdot;\hat{\theta}_n) + n^{1/2}S_n(\cdot;\hat{\theta}_n).$$

By assumption, ℓ is of Kolmogorov type. There exists a bounded continuous nondecreasing function g such that

$$\begin{aligned} &E\ell[n^{1/2}T_n(\cdot;\hat\theta_n) + n^{1/2}S_n(\cdot;\hat\theta_n)] \\ (2.11)\quad &= Eg\{\| n^{1/2}T_n(\cdot;\hat\theta_n) + n^{1/2}S_n(\cdot;\hat\theta_n)\|_\infty\} \\ &= Eg\{\| n^{1/2}T_n(\cdot;\hat\theta_n) + \delta(\hat\theta_n; F_n)\cdot(n^{1/2}S_n(\cdot;\hat\theta_n)/\delta(\hat\theta_n;F_n))\|_\infty\}. \end{aligned}$$

It follows from Theorem 3 that, for any $c > 0$,

$$\lim_{n\to\infty}\sup_{F_n\in B_n(F;c)}\delta(\hat\theta_n; F_n) \le 1/I(F_0)^{1/2}.$$

Therefore, for any $\{F_n\}$ from $\{B_n(F; c)\}$,

$$\begin{aligned} (2.12)\quad &\lim_{n\to\infty} Eg\{\| n^{1/2}T_n(\cdot;\hat\theta_n) + n^{1/2}S_n(\cdot;\hat\theta_n)\|_\infty\} \\ &\le \lim_{n\to\infty}\sup_{0\le\alpha\le 1/I(F_0)^{1/2}} Eg\{\| n^{1/2}T_n(\cdot;\hat\theta_n) + \alpha\cdot(n^{1/2}S_n(\cdot;\hat\theta_n)/\delta(\hat\theta_n; F_n)\|_\infty\}. \end{aligned}$$

We claim that the right-hand side of (2.12) cannot exceed

$$(2.13)\qquad \sup_{0\le\alpha\le 1/(F_0)^{1/2}} Eg\{\| W_s^0(F)(\cdot) + \alpha f(\cdot)Z\|_\infty\},$$

where Z is a standard normal random variable.

If the claim is false, then there must exist an $\varepsilon > 0$, a subsequence $\{n_k\}$ and a sequence $\{\alpha_{n_k}\}$, $0 \le \alpha_{n_k} \le 1/I(F_0)^{1/2}$, such that

$$\begin{aligned} (2.14)\quad &\lim_{n_k\to\infty} Eg\{\| n_k^{1/2}T_{n_k}(\cdot;\hat\theta_{n_k}) + \alpha_{n_k}\cdot(n_k^{1/2}S_{n_k}(\cdot;\hat\theta_{n_k})/\delta(\hat\theta_{n_k}; F_{n_k}))\|_\infty\} \\ &\qquad \ge \sup_{0\le\alpha\le 1/I(F_0)^{1/2}} Eg\{\| W_s^0(F)(\cdot) + \alpha f(\cdot)Z\|_\infty\} + \varepsilon. \end{aligned}$$

Let α_0 be the cluster point of $\{\alpha_{n_k}\}$. Clearly, $0 \le \alpha_0 \le 1/I(F_0)^{1/2}$. From (2.14), we have

$$\begin{aligned} (2.15)\quad &\lim_{n_k\to\infty} Eg\{\| n_k^{1/2}T_{n_k}(\cdot;\hat\theta_{n_k}) + \alpha_0(n_k^{1/2}S_{n_k}(\cdot;\hat\theta_{n_k})/\delta(\theta_{n_k}; F_{n_k}))\|_\infty\} \\ &\qquad \ge \sup_{0\le\alpha\le 1/I(F_0)^{1/2}} Eg\{\| W_s^0(F)(\cdot) + \alpha f(\cdot)Z\|_\infty\} + \varepsilon. \end{aligned}$$

The left-hand side of (2.15) converges to

$$Eg\{\| W_s^0(F)(\cdot) + \alpha_0 f(\cdot)Z\|_\infty\}$$

by Lemmas 2, 3, and the fact that loss function is bounded and continuous. Comparing with the right-hand side of (2.15), we obtain a contradiction. The claim thus follows.

By Lemma 5, the value of (2.13) equals

$$(2.16)\qquad Eg\{\| W_s^0(F)(\cdot) + Yf\|_\infty\},\quad \text{where}\quad Y \sim N(0, 1/I(F_0)).$$

Part (i) of the theorem follows easily from (2.12), (2.13), (2.16) and Theorem 1.

PROOF OF PART (ii). Recall that, for $F_n \in B_n(F; c)$,

$$n^{1/2}T_n(\cdot;\bar\theta_n) = n^{1/2}[\hat F_n(\cdot;\bar\theta_n) - F_n(\cdot - (\bar\theta_n - \theta_n))]$$

and

$$n^{1/2}T_n(\cdot + \bar\theta_n;\bar\theta_n) = n^{1/2}[\hat F_n(\cdot + \bar\theta_n;\bar\theta_n) - F_n(\cdot + \theta_n)],$$

where $F_n(\cdot + \theta_n)$ (the shape of F_n) is symmetric about 0. By arguments similar to those in Lemma 3, we obtain

$$| n^{1/2}T_n(x + \bar{\theta}_n; \bar{\theta}_n) - n^{1/2}[\hat{F}_n(x + \theta_n; \theta_n) - F_n(x + \theta_n)] | \to 0$$

in probability, uniformly in $F_n \in B_n(F; c)$ and $x \in R$. Combining these with the fact that

$$n^{1/2}[\hat{F}_n(x + \theta_n; \theta_n) - F_n(x + \theta_n)] \to_{\mathscr{L}} W_s^0(F_0),$$

we have

$$n^{1/2}T_n(x + \bar{\theta}_n; \bar{\theta}_n) \to_{\mathscr{L}} W_s^0(F_0)$$

uniformly in $B_n(F; c)$.

The second equality of part (ii) follows easily by going through a subsequence $\{n_k\}$ argument. The first equality of part (ii) holds because $\bar{\theta}_n$ is location invariant (see Appendix) and

$$\lim_{n\to\infty} \sup_{F_n \in B_n(F;c)} \int_{R^n} \ell[n^{1/2}(\hat{F}_n(\cdot + \bar{\theta}_n; \bar{\theta}_n) - F_{0n}(\cdot))]F_n^n(dx)$$

$$= \lim_{n\to\infty} \sup_{F_n^* \in B_n(F^*;c)} \int_{R^n} \ell[n^{1/2}(\hat{F}_n(\cdot + \bar{\theta}_n; \bar{\theta}_n) - F_{0n}^*(\cdot))]F_n^{*n}(dx)$$

for all $y \in R$ such that $F^*(\cdot) = F(\cdot - y)$. □

APPENDIX

PROOF OF THEOREM 3. (i) *Constructions of* $\bar{\theta}_n$. We use the construction in Huber (1964). Let $\psi(t)$ be a bounded function. Assume ψ is antisymmetric about zero. We further assume that ψ' exists and is positive on the whole real line (an example is $\psi(x) = \tan^{-1}(x)$). Let θ_n be the solution of

(A1) $$\sum_{i=1}^n \psi(x_i - t) = 0.$$

It is easy to see that $\bar{\theta}_n(-x_1, \cdots, -x_n) = -\bar{\theta}_n(x_1, \cdots, x_n)$ and $\bar{\theta}_n(x_1 + \alpha, x_2 + \alpha, \cdots, x_n + \alpha) = \bar{\theta}_n(x_1, x_2, \cdots, x_n) + \alpha$ for any real α. Since ψ is continuous and antisymmetric about 0, one can define a continuous functional

(A2) $$\lambda(\xi; F_n) = \int_{-\infty}^{\infty} \psi(t - \xi)\, dF_n(t).$$

It is clear that θ_n is the only root of $\lambda(\xi; F_n) = 0$ if F_n is symmetric about θ_n.

We claim that if $\bar{\theta}_n(x_1, \cdots, x_n)$ *is defined as above, and* $x_1, x_2, \cdots, x_n$ *are simple random samples from a certain symmetric d.f.* F_n, *then* $\bar{\theta}_n - \theta_n \to 0$ *in* F_n*-probability uniformly in* $B_n(F; c)$.

To prove this, consider $n^{-1}\sum_{i=1}^n \psi(x_i - \theta_n - \varepsilon)$, for some $\varepsilon > 0$, where $x_i \sim F_n(x)$ for all $i = 1, 2, \cdots, n$, and F_n is symmetric about θ_n.

If ψ is bounded and continuous, and if F is symmetric about θ, then by the law of large numbers

(A3) $n^{-1}\sum_{i=1}^{n}\psi(x_i-\theta-\varepsilon)-\int_{-\infty}^{\infty}\psi(t-\theta-\varepsilon)F_n(dt)\to 0$ in F_n-probability

for any sequence $\{F_n\}$ from $B_n(F; c)$. Since

$$\int_{-\infty}^{\infty}\psi(t-\theta-\varepsilon)F_n(dt)\to\int_{-\infty}^{\infty}\psi(t-\theta-\varepsilon)F(dt),$$

we have

(A4) $n^{-1}\sum_{i=1}^{n}\psi(x_i-\theta-\varepsilon)\to\lambda(\theta+\varepsilon, F)<0$ in F_n-probability.

Similarly, we have

(A5) $n^{-1}\sum_{i=1}^{n}\psi(x_i-\theta+\varepsilon)\to\lambda(\theta-\varepsilon, F)>0$ in F_n-probability.

From (A3), (A4), (A5) and the monotonicity λ, we conclude that

$$p\{\bar{\theta}_n\,\varepsilon\,[\theta-\varepsilon, \theta+\varepsilon]\,|\,F_n\}\to 1$$

for any sequence $\{F_n\}$ in $B_n(F; c)$. Furthermore, it is easy to see $\theta_n-\theta\to 0$ as $n\to\infty$, this implies $\bar{\theta}_n-\theta_n=O_{F_n}(1)$ uniformly in $B_n(F; c)$. The claim thus follows.

By the mean value theorem, there exist Δ_i, $0\le\Delta_i\le 1$, $1\le i\le n$, such that

$$\text{(A6)}\quad \begin{aligned} O &= \textstyle\sum_{i=1}^{n}\psi(x_i-\bar{\theta}_n)\\ &= \textstyle\sum_{i=1}^{n}\{\psi(x_i-\theta_n)-(\bar{\theta}_n-\theta_n)\psi'(x_i-\theta_n+\Delta_i(\bar{\theta}_n-\theta_n))\}.\end{aligned}$$

Therefore, under F_n,

(A7) $n^{1/2}(\bar{\theta}_n-\theta_n)=n^{1/2}\sum_{i=1}^{n}\psi(x_i-\theta_n)/n^{-1}\sum_{i=1}^{n}\psi'(x_i-\theta_i+\Delta_i(\bar{\theta}_n-\theta_n))$.

The numerator of the right-hand side in (A7) tends to $N(0, \int_{-\infty}^{\infty}\psi^2(x-\theta)F(dx))$ (since $\theta_n\to\theta$) uniformly in $B_n(F; c)$ and the denominator of the right-hand side in (A7) tends to a positive value, $\int_{-\infty}^{\infty}\psi'(x-\theta)F(dx)$ (ψ' is positive by assumption) uniformly in $B_n(F; c)$ by the above claim. Therefore, $n^{1/2}(\bar{\theta}_n-\theta_n)=O_p(1)$ uniformly in $B_n(F; c)$; thus part (i) of Theorem 3 follows. □

(ii) *Construction of* $\hat{\theta}_n$. We use $\bar{\theta}_n$ as our preliminary estimator, the construction of $\hat{\theta}_n$ is the same as that given by Stone (1975). Roughly speaking, $\hat{\theta}_n$ is the one-step maximum likelihood estimator, using a Newton-Raphson approach. For details of the construction, readers are referred to Stone (1975). We will show that this modified Stone's estimator $\hat{\theta}_n$ satisfies part (ii) of Theorem 3. (Note that the preliminary estimator $\bar{\theta}_n$ given here is different from that given by Stone (1975). The main reason is that we need a preliminary estimator $\bar{\theta}_n$ which behaves uniformly well in the neighborhood of F_θ.)

For any $\alpha>0$, let $\{F_n\}$ be an arbitrary sequence of symmetric d.f.'s F_n from $\{B_n(F; \alpha)\}$. Where F is a symmetric d.f., let θ_n be the center of F_n, and let θ be the center of F with finite Fisher Information. To establish the propositions and lemmas below, we only have to show that they are true under an arbitrary sequence $\{F_n\}$ from $\{B_n(F; \alpha)\}$.

To avoid confusion, we will use similar notation to that of Stone (1975). We observe $x_1 + \theta_n, x_2 + \theta_n, \cdots, x_n + \theta_n$ of size n from F_n and wish to estimate the unknown center θ_n from this sample. We will use "F_n^*" to denote the d.f. which is symmetric about origin (i.e., we shift F_n by θ_n). Therefore, at the nth stage, $x_1, x_2, \cdots, x_n$ are i.i.d. from F_n^* (even we observe $\{x_i + \theta_n\}_{i=1}^n$). To avoid ambiguity, we specify underlying d.f. everywhere. For example,

$$f(x; r; F_n^*) = \int \phi(x-y) F_n^*(dy),$$

$$f'(x; r; F_n^*) = (\partial/\partial x) f(x; r; F_n^*), L(x; r; F_n)$$

$$= f'(x; r; F_n^*)/f(x; r; F_n^*), A(r; c; F_n^*)$$

$$= \int L^2(x; r; F_n^*) g(x/c) f(x; r; F_n^*)\, dx, f_n(x; r; F_n^*)$$

$$= \frac{1}{n} \textstyle\sum_{i=1}^n \phi(x - x_i; r; F_n^*) (\text{where } x_i \sim F_n^*), \hat{f}_n(x; r; F_n^*)$$

$$= \frac{1}{2}(f_n(x + \bar{\theta}_n - \theta_n; r; F_n^*) + f_n(-x + \bar{\theta}_n - \theta_n; r; F_n^*)),$$

$$\hat{f}_n'(x; r; F_n^*) = (\partial/\partial x)\hat{f}_n(x; r; F_n^*), \hat{L}_n(x; r; F_n^*)$$

$$= \hat{f}_n'(x; r; F_n^*)/\hat{f}_n(x; r; F_n^*), \quad \text{and} \quad \hat{A}_n(r; c; F_n^*)$$

$$= \int \hat{L}_n^2(x; r; F_n^*) g(x/c) \hat{f}_n(x; r; F_n^*)\, dx,$$

where $\phi(x)$ is standard normal density, $\phi(x;\, r) = r^{-1}\phi(x/r)$, and g is defined in Stone (1975).

It is straightforward to check that all the propositions in Sections 2 and 3 of Stone (1975) hold uniformly in $\{F_n\}$ belonging to the Hellinger ball.

LEMMA A1. *For any bounded real value function $g(x)$ with finite support, and for $r_n^{-1} = o(n^{1/2})$, we have*

$$\sup_{F_n \in B_n(F;\alpha)} \left| \int g(x) f(x;\, r_n;\, F_n)\, dx - \int g(x) f(x;\, r_n;\, F)\, dx \right| \to 0 \quad \text{as} \quad n \to \infty.$$

PROOF. For any fixed x,

$$|f(x;\, r_n;\, F_n) - f(x;\, r_n;\, F)| = \left| \int \frac{1}{r_n} \phi\left(\frac{x-y}{r_n}\right) [F_n(dy) - F(dy)] \right|$$

$$\le \left\{ \int_{-\infty}^{\infty} \frac{1}{r_n^2} \phi^2\left(\frac{x-y}{r_n}\right) (\sqrt{dF_n} + \sqrt{dF})^2 \right\}^{1/2} \cdot \frac{\alpha}{\sqrt{n}}$$

$$\le \text{constant} \cdot \alpha/r_n\sqrt{n} \to 0, \quad \text{independent of } x \quad \text{and} \quad F_n \in B_n(F;\, \alpha)$$

$$\cdot\ (\text{since } r_n^{-1} = o(n^{1/2})).$$

Therefore,

$$\left| \int g(x)[f(x;\, r_n;\, F_n) - f(x;\, r_n;\, F)]\, dx \right| \le 2MA(\alpha/r_n\sqrt{n}) \to 0,$$

where $\| g(x) \|_\infty \le M$ and the support of $g \subset [-A, A]$. □

The proof of the following lemma is similar to that of Theorem 4.1 in Stone (1975).

LEMMA A2. *Assume that* $r_n^{-1} = o(n^{1/2})$. *Then*

$$\lim_{c\to\infty;n\to\infty} \inf_{F_n \in B_n(F;\alpha)} A(r_n;\, c;\, F_n) \ge I(F).$$

PROOF. According to Huber (1964),

$$I(F_0) = \sup_\psi E_F(\psi'(x))^2/E_F\psi^2(x),$$

where the sup extends over all $\psi \in C_c^1$ such that $E_F\psi^2(x) > 0$. Choose $\varepsilon > 0$. We can find a $\psi \in C_c^1$ such that $\int \psi^2 F(dx) > 0$ and

$$\text{(A8)} \qquad \left(\int \psi'(x)F(dx)\right)^2 \ge (1-\varepsilon)I(F) \int \psi^2(x)F(dx).$$

Thus for r_n sufficiently small

$$\text{(A9)} \qquad \left(\int \psi'(x)f(x;\, r_n;\, F)\, dx\right)^2 \ge (1-2\varepsilon)I(F) \int \psi^2(x)f(x;\, r_n;\, F)\, dx.$$

Choose $c > 0$ such that $\psi(x) = 0$ for $|x| \ge c$. Then

$$\begin{aligned}\left(\int \psi'(x)f(x;\, r_n;\, F)\, dx\right)^2 &= \left(\int \psi(x)f'(x;\, r_n;\, F)\, dx\right)^2 \\ &\le \left(\int \psi^2(x)f(x;\, r_n;\, F)\, dx\right) \int_{-c}^{c} L^2(x;\, r_n;\, F)f(x;\, r_n;\, F)\, dx.\end{aligned}$$

Therefore, $\int_{-c}^{c} L^2(x;\, r_n;\, F)f(x;\, r_n;\, F)\, dx \ge (1-2\varepsilon)I(F)$. On the other hand, with r_n and ψ given above, one also has

$$\left(\int \psi'(x)f(x;\, r_n;\, F_n)\, dx\right)^2 \le \left(\int \psi^2(x)f(x;\, r_n;\, F_n)\, dx\right) \cdot \left(\int_{-c}^{c} L^2(x;\, r_n;\, F_n)f(x;\, r_n;\, F_n)\, dx\right).$$

Therefore,

$$\text{(A10)} \qquad \int_{-c}^{c} L^2(x; r_n; F_n) \ge \left(\int \psi'(x)f(x; r_n; F_n)\, dx\right)^2 \Big/ \int \psi^2(x)f(x; r_n; F_n)\, dx.$$

By Lemma A1,

$$\text{(A11)} \qquad \int \psi'(x) f(x; r_n; F_n)\, dx \to \int \psi'(x) f(x; r_n; F)\, dx$$

and

$$\text{(A12)} \qquad \int \psi^2(x) f(x; r_n; F_n)\, dx \to \int \psi^2(x) f(x; r_n; F)\, dx$$

uniformly in $B_n(F; \alpha)$.

Thus the result follows from (A9)–(A12). □

In his Section 4, Stone (1975) constructed a nonadaptive estimator. Based on our preliminary estimator $\bar{\theta}_n$, we also consider the similar estimator.

$$\tilde{\theta}_n = \bar{\theta}_n - \frac{1}{A(r_n; c_n; F_n^*)} \cdot \int L(x; r_n; F_n^*) g(x/c)(f_n(x + \bar{\theta}_n - \theta_n; r_n; F_n^*) - f(x; r_n; F_n^*))\, dx.$$

One can mimic the proof of Theorem 4.2 in Stone (1975) to obtain the following proposition.

PROPOSITION A1. *Suppose that $c_n/n^{1-\varepsilon} r_n^6 = O_p(1)$ for some $\varepsilon > 0$, and $\{F_n\}$ is an arbitrary sequence from $\{B_n(F; \alpha)\}$. Then*

$$\text{(A13)} \qquad \mathscr{L}(n^{1/2}(\tilde{\theta}_n - \theta_n)/(\sigma(r_n, c_n)/A(r_n; c_n; F_n^*)) \mid F_n) \to N(0, 1) \quad \text{as} \quad n \to \infty,$$

where $\sigma^2(r_n, c_n)$ is the variance (up to $o_p(n^{-1/2})$) of the random variable

$$\int L(x; r_n; F_n^*) g(x/c_n) \phi(x - X_i; r_n; F_n^*)\, dx.$$

Furthermore,

$$\text{(A14)} \qquad \limsup_{n\to\infty} \sigma^2(r_n, c_n)/A^2(r_n; c_n; F_n^*) \le \lim_{n\to\infty} 1/A(r_n; c_n; F_n^*) \le 1/I(F).$$

PROOF. The first part of the proposition follows from the same arguments as given in Stone (1975, pages 278–279). The second part follows from Lemma A2 and the fact that $\sigma^2(r_n; c_n; F_n^*) \le A(r_n; c_n; F_n^*) + o(n^{-1/2})$ (see Stone, 1975, page 279, from (4.7) to (4.11)). □

The next proposition is similar to Theorem 5.1 in Stone (1975).

PROPOSITION A2. *Suppose that $c_n/n^{1-\varepsilon} r_n^5 = O_p(1)$ for some $\varepsilon > 0$. Then $\hat{A}_n(r_n; c_n; F_n^*)/A(r_n; c_n; F_n^*) \to 1$ in $\{F_n\}$ probability as $n \to \infty$.*

PROOF. The proof of this proposition is exactly the same as that of Theorem 5.1 in Stone (1975), so we omit it. (Note that Proposition 5.1 in Stone, 1975, holds in our case, too.) □

Define

$$\hat{\theta}_n = \bar{\theta}_n - \frac{1}{\hat{A}_n(r_n;\, c_n;\, F_n^*)}$$

$$\cdot \int \hat{L}_n(x;\, r_n;\, F_n^*) g(x/c_n)(f_n(x + \bar{\theta}_n - \theta_n;\, r_n;\, F_n^*) - f(x;\, r_n;\, F_n^*))\, dx.$$

PROPOSITION A3. *Suppose that* $c_n/n^{1-\varepsilon}r_n^6 = O_p(1)$ *for some* $\varepsilon > 0$. *Then* $\mathscr{L}[n^{1/2}(\hat{\theta}_n - \tilde{\theta}_n) \mid F_n] \to 0$ *in* $\{F_n\}$ *probability. Furthermore,*

$$\lim_{n\to\infty}\sup_{F_n\in B_n(F;\alpha)}\sigma^2\,(r_n,\, c_n)/A^2\,(r_n;\, c_n;\, F_n^* \le 1/I(F) \quad \textit{for any} \quad \alpha > 0.$$

PROOF. The first part of the theorem follows from the same arguments as given for Theorem 5.2 in Stone (1975). The second part of the theorem follows from the first part of this theorem and Proposition A1, and the fact that $\{F_n\}$ is an arbitrary sequence from $\{B_n(F;\, \alpha)\}$. □

PROOF OF THEOREM 3, PART (ii). It is easy to see that $\hat{\theta}_n$ satisfies part (i) of the theorem. Choose $c_n/n^{1-\varepsilon}r_n^6 = O_p(1)$ for some $\varepsilon > 0$, and $\delta(\hat{\theta}_n;\, F_n) = \sigma(r_n,\, c_n)/A(r_n;\, c_n;\, F_n^*)$. Then the theorem follows from Propositions A1, A2 and A3. □

Acknowledgements. This work is based on part of my dissertation completed at University of California, Berkeley (1981). I am indebted to Professor LeCam's guidance and encouragement. I thank Professor R. Beran, late Professor J. Kiefer and Professor P. W. Millar for their helpful conversations. I also thank the referees who gave me many valuable comments to improve the original manuscript.

REFERENCES

ANDERSON, T. W. (1954). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Ann. Math. Statist.* **25** 616.

BEGUN, J., HALL, W. J., HUANG, W. M. and WELLNER, J. (1983). Information and asymptotic efficiency in parametric-nonparametric models. *Ann. Statist.* **2** 435–452.

BERAN, R. J. (1974). Asymptotically efficient adaptive rank estimates in location models. *Ann. Statist.* **2** 63–74.

BERAN, R. J. (1978). An efficient and robust adaptive estimator of location. *Ann. Statist.* **6** 292–313.

BICKEL, P. J. (1982). On adaptive estimation. *Ann. Statist.* **10** 647–671.

DVORETZKY, A., KIEFER, J. and WOLFOWITZ, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669.

HUBER, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–101.

KIEFER, J. and WOLFOWITZ, J. (1976). Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrsch. verw. Gebiete* **34** 73–85.

KUO, H.-H. (1975). *Gaussian Measures in Banach Spaces.* Springer, New York.

LE CAM, L. (1982). *Asymptotic Method in Statistical Decision Theory.* To be published.

LO, S.-H. (1981). Locally asymptotically minimax estimation for symmetric distribution functions and shift parameters. Dissertation, University of California, Berkeley.

MILLAR, P. W. (1979). Asymptotic minimax theorems for the sample distribution function. *Z. Wahrsch. verw. Gebiete* **48** 233–252.

SACKS, J. (1975). An asymptotically efficient sequence of estimators of a location parameter. *Ann. Statist.* **3** 285–298.

SCHUSTER, E. F. (1975). Estimating the distribution function of a symmetric distribution. *Biometrika* **62** 631–635.

SHORACK, G. R. and WELLNER, J. A. (1982). Limit theorem and inequalities for the uniform empirical process indexed by intervals. *Ann. Probab.* **10** 639–652.

STEIN, C. (1956). Efficient nonparametric testing and estimation. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 187–196. Univ. California Press.

STONE, C. J. (1975). Adaptive maximum likelihood estimators of a location parameter. *Ann. Statist.* **3** 267–284.

STUTE, W. (1982). The oscillation behavior of empirical processes. *Ann. Probab.* **10** 86–107.

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903