Ann. Inst. Statist. Math.
37 (1985), Part A, 109-129

ASYMPTOTIC DISTRIBUTION THEORY OF STATISTICAL FUNCTIONALS : THE COMPACT DERIVATIVE APPROACH FOR ROBUST ESTIMATORS

W. ESTY, R. GILLETTE, M. HAMILTON AND D. TAYLOR

(Received Aug. 30, 1983; revised Apr. 5, 1984)

Summary

Derivatives of statistical functionals have been used to derive the asymptotic distributions of L-, M- and R-estimators. This approach is often heuristic because the types of derivatives chosen have serious limitations. The Gâteaux derivative is too weak and the Fréchet derivative is too strong. In between lies the compact derivative. This paper obtains strong results in a rigorous manner using the compact derivative on $C_0(R)$. This choice of space allows results for a broader class of functionals than previous choices, and the fact that $\{\sqrt{n}(\tilde{F}_n-F)\}$ is often tight provides the compact set required. A major result is the derivation of the compact derivative of the inverse c.d.f. when the range space is endowed with the uniform norm. It has applications to the asymptotic theory of L-, M- and R-estimators. We illustrate the power of this result by applications to L-estimators in settings including the one sample problem, data grouped by quantiles, and censored survival time data.

1. Introduction

L-estimators, M-estimators, R-estimators, and minimum distance estimators can be expressed as functionals of estimates of the cumulative distribution function. The asymptotic distribution theory of such estimators, therefore, may be studied with the aid of functional analysis. Serfling [26] and Huber [17] both provide good expositions of the statistical functional approach to estimation. Derivatives of functionals (in various senses) and Taylor-like expansions have been useful tools of robust statistics. Unfortunately, their value has often been heuristic because of the limitations of the type of derivative used or of the space on which it was defined. Serfling (p. 216) notes that the Gâteaux de-

Key words and phrases: Robust estimation, statistical functional, compact derivative, grouped data, survival data.

rivative Taylor-like expansion is "Utilized rather *informally*, merely as a guiding concept", and (p. 220) that the Fréchet derivative is "somewhat too narrow for the purposes of statistical applications". Huber (p. 37) says "Unfortunately, the concept of Fréchet differentiability appears to be too strong: in too many cases the Fréchet derivative does not exist, and even when it does, the fact is difficult to establish". On the other hand, he notes (p. 40) "Mere Gâteaux differentiability does not suffice to establish asymptotic normality.... The most promising intermediate approach seems to be the one by Reeds [23], which is based on the notion of compact differentiability". However, the space on which Reeds defined the compact derivative forced restrictions on the functionals to which it applied (see remarks after Theorem 2).

The goal of this paper is to demonstrate the power and generality of the compact derivative along the space $C_0(R)$ for purposes of rigorously deriving the asymptotic distributions of robust estimators in realistic, applicable settings. A major result is the derivation of the compact derivative of the inverse c.d.f. when the range space is endowed with the uniform norm (Theorem 1). Consequences include rigorous derivations of the asymptotic distributions of L-, M- and R-estimators. We use robust L-estimators to illustrate the power of this approach. M- and R-estimators can be studied similarly and will be the subject of another paper. The hypotheses of this approach are simple and easily verified. If the true c.d.f. has a continuous positive derivative over a specified interval, the conclusions hold for a wider class of estimators than previous functional approaches (Theorems 2, 3, 4 and 5). Particular applications to one sample problems (Theorems 6 and 7), data grouped by quantiles (Theorem 8), and censored survival time data (Theorem 9) are given.

In the next section we explain why the compact derivative on $C_0(R)$ is the appropriate derivative on the appropriate space, and we give the background and notation of the statistical functional approach to L-estimation. Section 3 contains the major theoretical results, uninterrupted by proofs, which are in Section 4. The featured application to L-estimators is in Section 5, which contains results that are more general than any previously available. The final section contains a discussion of potential generalizations and applications.

2. Motivation and background

The statistical functional approach to estimation is important because parameters of distributions can be expressed as functionals of the c.d.f. For instance, the median is given by $T(F)=F^{-1}(1/2)$ and the mean of a nonnegative random variable by $T(F)=\int_0^\infty 1-F(x)dx$. Continuity of T at F (with the "right" norm on the space of differences of c.d.f.s) is the definition of robustness (Huber [17], p. 10). Differentiability of T at F is intended to permit an expansion of $\sqrt{n}(T(\tilde{F}_n) - T(F))$, where $\tilde{F}_n$ is some estimate of F, in terms of a linear functional of $\sqrt{n}(\tilde{F}_n - F)$ and a remainder term converging to 0 in probability. Then $\sqrt{n}(T(\tilde{F}_n) - T(F))$ will converge to the corresponding linear functional of the limit of $\sqrt{n}(\tilde{F}_n - F)$, which is often tractable. For instance, in the one sample problem when $\tilde{F}_n$ is the empirical distribution function the limit of $\sqrt{n}(\tilde{F}_n - F)$ is a Brownian bridge and $T(\tilde{F}_n)$ is asymptotically normal.

There are several types of derivatives which may be tried. The Gâteaux derivative, which yields the "influence curve", and parallels the directional derivative of functions on R^2, is too weak to complete rigorous proofs since $(\tilde{F}_n - F)$ is not in any particular direction. Another popular candidate, the Fréchet derivative, is strong enough to prove results, but too strong to obtain the most general results because there are important functionals for which it fails to exist, for instance, the median. On the other hand, Boos [8], restricting attention to the one sample problem and a certain class of L-estimators, used the Fréchet derivative to prove useful asymptotic distribution theorems.

In between the Gâteaux and Fréchet derivatives lies the compact derivative (Averbukh and Smolynov [2], [3]). It is strong enough to prove very general results and if $\{\sqrt{n}(\tilde{F}_n - F)\}$ is tight (which it is in many cases including the motivating one-sample data problem) the tightness provides the compact set of functions required.

Compact differentiation is most easily done when the arguments lie in a space of continuous functions. Thus, for continuous F, we let our estimate of F, $\tilde{F}_n$, also be a continuous c.d.f. so that $(\tilde{F}_n - F)$ is in $C_0(R)$, the space of continuous functions vanishing at infinity. Since the usual empirical distribution function in the one-sample data problem is discontinuous, this choice of space has been overlooked. But, whenever $\hat{F}_n$ is discontinuous, a continuous version, $\tilde{F}_n$, can easily be created for which the theory holds. Then a proof that $\sqrt{n}(T(\hat{F}_n) - T(\tilde{F}_n)) \xrightarrow{p} 0$ shows that $T(\hat{F}_n)$ has the same asymptotic theory as $T(\tilde{F}_n)$. This may often be proven directly. For the L-estimator case we state a result (Theorem 3) that allows the conversion of results for continuous estimates of F to results for discontinuous estimates of F. The choice of the space $C_0(R)$ greatly simplifies matters compared to the use of the space D of discontinuous functions that are right continuous and have right hand limits, the space in which $(\hat{F}_n - F)$ lies. Previous investigators have tended to focus on D.

The value of the compact derivative to asymptotic statistical theory was apparently first noticed by Reeds [23]. In his dissertation he thoroughly investigated the functional approach in the one-sample setting using the compact derivative on the space D. Reeds focused particularly on M-estimators. Fernholz [13], following Reeds, derived the asymptotic normal distribution for R-estimators in the one-sample problem. We also use the compact derivative, but we use a smaller domain for T than do Reeds and Fernholtz and thereby overcome the main difficulties in choosing a topology.

As Huber [17] points out, asymptotic results for L-estimators have not been completely established. There is no blanket statistical functional asymptotic distribution theorem covering all of the interesting cases simultaneously. For example, the results in Reeds [23], Boos [8], Mason [21], Reid [24] and Huber [17] cannot provide an asymptotic distribution for the L-estimator that is calculated by taking a weighted average of the median and the trimmed mean, both of which are L-estimators. Serfling [26] combines two approaches to get the result, but the variance is not expressed in closed form and is cumbersome to apply and his result only applies to the one-sample case. As indicated by Jaeckel [18] and Dodge and Lindstrom [10], weighted averages of L-estimators may be quite useful.

This article presents a general asymptotic distribution theorem for L-estimators with simple and clearly stated hypotheses and a rigorous mathematical proof. The theorem may be applied in a variety of estimation settings, three of which are treated in Section 5.

L-estimators will now be defined as certain functionals on the set of continuous c.d.f.'s. It can often be verified that the corresponding functionals of discontinuous c.d.f.'s, such as the empirical distribution function, have equivalent asymptotic results.

Let $\mathcal{M}$ be the collection of finite, regular (signed) Borel measures with compact support. Let $\mathcal{F}$ be the set of continuous c.d.f.'s. Consider the functional T_M; T_M: $\mathcal{F} \to \mathcal{R}$, defined by

$$T_M(F) = \int F^{-1}(t) M(dt) \,, \qquad F \in \mathcal{F},\ M \in \mathcal{M} \,, \tag{2.1}$$

where $F^{-1}(t) = \inf\{x \mid F(x) \geqq t\}$, for all $t \in (0, 1)$. Because F^{-1} is monotone, T_M is well-defined.

DEFINITION 1. Let $\hat{F} \in \mathcal{F}$ be a data-based estimate of the true underlying c.d.f. $F \in \mathcal{F}$. Any estimate $T_M(\hat{F})$ of form (2.1) is said to be an ***L*-estimate** of the parameter $T_M(F)$. Moreover, any functional T_M of form (2.1) is said to be an ***L*-estimator**.

By varying the choice of M, one can use this definition to generate all classical L-estimators (for specific examples, see Section 5 and Andrews et al. [1] and Huber [17]).

Although much research on robust estimation has been done with the assumption of a symmetric F, we make no symmetry requirements. Bickel and Lehmann [4], [5] demonstrate that the assumption of symmetric F is unnecessarily restrictive.

We now proceed with the definition of compact differentiability and the statements of theorems asserting the compact differentiability of the inverse functional and of L-estimators.

3. Asymptotic distribution theory

Under mild regularity conditions we show that the inverse map $F \to F^{-1}$ is compactly differentiable (Theorem 1). As the featured application we show that L-estimators are compactly differentiable (Theorem 2). This leads to easily applied asymptotic results for a very broad class of L-estimators in Section 5.

We now define compact differentiability. Let X and Z be Banach spaces and let Y be a closed linear subspace of X.

DEFINITION 2. Let $x \in X$. A map T: $x+Y \to Z$ is said to be **compactly differentiable** at x along Y if and only if there exists a continuous linear map D_x^T: $Y \to Z$ such that

$$\lim_{t \to 0} \sup_{y \in \mathcal{K}} \|[T(x+ty)-T(x)-D_x^T(ty)]/t\|=0 \tag{3.1}$$

for each set $\mathcal{K} \subseteqq Y$ for which the closure of $\mathcal{K}$ is compact. If $Y=X$, then we just say that T is compactly differentiable at x.

We are almost ready to state one of the major results of this paper. But first we need to present some more notation. Let a, b be real numbers with $0<a \leqq b<1$. Let $DL[a, b]$ denote the space of all bounded left continuous real valued functions defined on $[a, b]$ which have right hand limits ($DL[a, b]$ is the left continuous version of D). The space $DL[a, b]$ is a Banach space under the uniform norm $\| \ \|_\infty$. Now let $C_b(R)$ denote the bounded real valued continuous functions defined on the reals and let $C_0(R)$ denote those functions in $C_b(R)$ that vanish at infinity. The space $C_b(R)$ is also a Banach space under the uniform norm $\| \ \|_\infty$ and $C_0(R)$ is a closed subspace of $C_b(R)$. For a discussion of these spaces see Simmons [27]. Next let $G \in C_b(R)$ such that $\lim_{s \to -\infty} G(s) =0$ and $\lim_{s \to \infty} G(s)=1$. For each q, $a \leqq q \leqq b$, let

$$G^{-1}(q)=\inf \{s \mid G(s) \geqq q\} .$$

It is easy to establish that $G^{-1} \in DL[a, b]$, G^{-1} is always monotone increasing on $[a, b]$, and $G(G^{-1}(q))=q$ for all q in $[a, b]$.

We will now state a regularity condition and the first major theorem. Henceforth let a and b be fixed real numbers such that $0<a\leqq b<1$.

CONDITION A. Let F be a c.d.f. with a continuous positive derivative $F'(s)$ at all points s in a closed interval $[c, d]$, where $c<F^{-1}(a)\leqq F^{-1}(b)<d$.

THEOREM 1. *Let F be a c.d.f. that satisfies Condition A, and let T: $F+C_0(R)\rightarrow DL[a, b]$ be the map defined by*

$$T(F+H)=(F+H)^{-1}$$

for each $H\in C_0(R)$. Then T is compactly differentiable at F along $C_0(R)$ with compact derivative $D_F^T(H)=-H(F^{-1}(\cdot))/F'(F^{-1}(\cdot))$, $H\in C_0(R)$ or equivalently, for each subset $\mathcal{K}\subseteqq C_0(R)$ with compact closure it follows that

$$\lim_{t\rightarrow 0}\sup_{(H,q)\in\mathcal{K}\times[a,b]}\left|\frac{(F+tH)^{-1}(q)-F^{-1}(q)}{t}+\frac{H(F^{-1}(q))}{F'(F^{-1}(q))}\right|=0\ . \tag{3.2}$$

A variation of Theorem 1 already exists. Reeds [23] (unpublished) showed that when F is the uniform distribution on $[0, 1]$ and $DL[a, b]$ is endowed with an L^p-norm instead of the uniform norm, then T is compactly differentiable at F. Reeds' result for the uniform c.d.f. cannot be extended to a general c.d.f. unless further restrictions on F are imposed. Clearly, our case implies the case studied by Reeds. The fact that we are able to use the uniform norm instead of an L^p-norm is of fundamental importance.

CONDITION B. Let M be a finite, regular (signed) Borel measure with support contained in the interval $[a, b]$.

There are fundamental obstacles to allowing F to be discontinuous when dealing with the broad class of measures in Condition B, since the derivative may fail to exist (see the remark following Theorem 2). For discontinuous F, the results of Boos [8], Serfling [26], Mason [21], and Huber [17] remain the best available.

The next theorem gives the compact derivative for L-estimators.

THEOREM 2. *Let F and M satisfy Conditions A and B, respectively. Let T: $F+C_0(R)\rightarrow R$ be the map defined by*

$$T(F+H)=\int(F+H)^{-1}(q)M(dq)\ ,\qquad H\in C_0(R)\ .$$

Then T is compactly differentiable at F along $C_0(R)$ with compact derivative

$$D_F^T(H) = -\int_{[a,b]} [H(F^{-1}(q))/F'(F^{-1}(q))]M(dq) , \qquad H \in C_0(R) .$$

Theorem 2 follows immediately from Theorem 1 by passing limits through the integral sign, which is permissible because (3.2) yields uniform convergence and M is a finite signed measure.

Remark. The usual approach to deriving D_F^T allows H to be a discontinuous function. To accomplish this, one could attempt to replace $C_0(R)$ in Theorem 1 by a suitable Banach space D_0 of discontinuous functions (see the discussion of the space D in Billingsley [6]). This tactic cannot succeed without restricting M since, as simple examples show, if H has discontinuities in common with discontinuities of M, even the directional derivative may fail to exist, that is,

$$\lim_{t\to 0} \frac{1}{t}[T_M(F+tH) - T_M(F)]$$

may fail to exist. For instance, the median is not compactly differentiable along such a space D_0. The situation is different if we restrict M to be absolutely continuous (see Theorem 6.3.1 of Reeds [23]).

The asymptotic theory for L-estimators with discontinuous $\hat{F}$ concerns the expression

$$\sqrt{n}(T_M(\hat{F}) - T_M(F)) = \sqrt{n}(T_M(\hat{F}) - T_M(\tilde{F})) + \sqrt{n}(T_M(\tilde{F}) - T_M(F)) ,$$

where $\tilde{F}$ is a continuous version of $\hat{F}$. The theory in this paper applies to the second term and will therefore apply to discontinuous $\hat{F}$ if the first term converges in probability to 0. That this is so for many statistical functionals may be proven directly. For the L-estimator case it follows easily from the conclusion of the next theorem and the definition, (2.1), of T_M, if M is a finite signed measure. Moreover, under some lenient regularity conditions, the next theorem may be applied to M-estimators and R-estimators as well (Taylor [28]). Its application to L-estimators is illustrated in Section 5.

Before stating the next theorem, we need to present the following condition and additional notation.

CONDITION C. Let $(\Omega, \mathcal{A}, P)$ be a probability space. Let F be a c.d.f. that satisfies Condition A. For each positive integer n and $\omega \in \Omega$ let $\tilde{F}_n(\omega)$ be continuous and let $\hat{F}_n(\omega)$ be a c.d.f. that is a step function with a finite number of steps. The following assumptions about

$\tilde{F}_n$, $\hat{F}_n$ and F are made:

1) $(\tilde{F}_n - F)$: $\Omega \to C_0(R)$ is measurable with respect to P;

2) $c_n(\tilde{F}_n - F)$ is tight in $C_0(R)$, where $\{c_n\}$ is a sequence of positive numbers such that $c_n \to \infty$,

3) $\|c_n(\tilde{F}_n - F_n)\|_\infty \xrightarrow{P} 0$.

Usually c_n will be $\sqrt{n}$.

THEOREM 3. *Suppose* $\{\tilde{F}_n\}$, $\{\hat{F}_n\}$, *and* F *satisfy Condition C. Then* $\sup_{a \le q \le b} c_n |\tilde{F}_n^{-1}(q) - \hat{F}_n^{-1}(q)| \xrightarrow{P} 0$.

The proof will be given in the next section.

THEOREM 4. *Let* T *be as in* (2.1) *and let* F *and* M *satisfy Conditions A and B respectively. Suppose that the sequence of random elements* $\{\sqrt{n}(\tilde{F}_n - F)\}_{n=1}^\infty$ *is tight. Then for each* $\omega \in \Omega$

$$\sqrt{n}[T(\tilde{F}_n(\omega)) - T(F)] = -\int_{[a,b]} \frac{\sqrt{n}[\tilde{F}_n(F^{-1}(s), \omega) - s]}{F'(F^{-1}(s))} M(ds) + r_n(\omega) , \tag{3.3}$$

where $r_n = o_p(1)$ *as* $n \to \infty$.

The proof is given in Section 4.

THEOREM 5. *Let* T *be as in* (2.1) *and let* F *and* M *satisfy Conditions A and B respectively. Suppose that the sequence of random elements* $\{\sqrt{n}(\tilde{F}_n - F)\}_{n=1}^\infty$ *is tight and that the finite dimensional distributions of* $\sqrt{n}(\tilde{F}_n - F)$ *converge weakly to those of some Gaussian process* G *with* $\mathrm{E}[G(s)] = 0$ *and a continuous covariance kernel* $\mathrm{Cov}[G(s), G(t)]$ *on* $[c, d] \times [c, d]$. *Then*

$$\sqrt{n}[T(\tilde{F}_n) - T(F)] \xrightarrow{\mathcal{D}} N(0, \sigma_{T,F}^2) , \tag{3.4}$$

where

$$\sigma_{T,F}^2 = \iint_{[a,b]\times[a,b]} \frac{\mathrm{Cov}[G(F^{-1}(s)), G(F^{-1}(t))]}{F'(F^{-1}(s))F'(F^{-1}(t))} M(ds)M(dt) . \tag{3.5}$$

The proof is given in Section 4.

In many statistical settings the limiting Gaussian process G is already known, at least for the more popular c.d.f. estimators. This fact makes Theorem 5 particularly easy to apply. Examples are given in Section 5.

4. Proofs

The following technical lemma is used in proving Theorem 1.

LEMMA 1. *Let $\mathcal{K}$ be a subset of $C_0(R)$ having compact closure. Then, under Condition A,*

$$\lim_{t\to 0} \sup \{|(F+tH)^{-1}(q)-F^{-1}(q)| \,|\, (H, q) \in \mathcal{K}\times[a, b]\}=0 . \tag{4.1}$$

PROOF. By the Arzelà-Ascoli Theorem (Billingsley [6], Appendix 1), $\mathcal{K}$ is uniformly bounded. Set $K=\sup_{H\in\mathcal{K}} \|H\|_\infty$ and let $\varepsilon>0$. Choose $\eta_1>0$ so that $\eta_1 K<\min\{a(a-F(c)), b(F(d)-b)\}$. From now on let $t\in[-\eta_1, \eta_1]$. Then for $x\leqq c$, $y\geqq d$, $H\in\mathcal{K}$, we get

$$\begin{aligned} F(x)+t\cdot H(x) &\leqq F(x)+|t|\cdot\|H\|_\infty<F(x)+\eta_1 K \\ &<F(c)+a(a-F(c))=(1-a)F(c)+a^2 \\ &<(1-a)a+a^2=a<b=(1-b)b+b^2 \\ &<(1-b)F(d)+b^2=F(d)-b(F(d)-b) \\ &\leqq F(d)-\eta_1 K \\ &\leqq F(y)+tH(y) . \end{aligned}$$

Thus $c\leqq(F+tH)^{-1}(q)\leqq d$ for all $H\in\mathcal{K}$ and $q\in[a, b]$. Define $\mu=\inf\{|F'(s)| \,|\, s\in[c, d]\}$. For the moment assume $(F+tH)^{-1}(q)\neq F^{-1}(q)$. Then

$$\begin{aligned} &|(F+tH)^{-1}(q)-F^{-1}(q)| \\ &=|F((F+tH)^{-1}(q))-F(F^{-1}(q))|\left|\frac{(F+tH)^{-1}(q)-F^{-1}(q)}{F((F+tH)^{-1}(q))-F(F^{-1}(q))}\right| \\ &=|F((F+tH)^{-1}(q))-q|/|F'(\xi_{H,q,t})| , \end{aligned}$$

by the mean value theorem, where $\xi_{H,q,t}$ lies between $(F+tH)^{-1}(q)$ and $F^{-1}(q)$ and therefore lies in $[c, d]$. Thus

$$|(F+tH)^{-1}(q)-F^{-1}(q)|\leqq\mu^{-1}|F((F+tH)^{-1}(q))-q| . \tag{4.2}$$

If $(F+tH)^{-1}(q)=F^{-1}(q)$, then clearly (4.2) still holds. Because $q=(F+tH)((F+tH)^{-1}(q))=F((F+tH)^{-1}(q))+tH((F+tH)^{-1}(q))$, it follows that

$$F((F+tH)^{-1}(q))-q=-tH((F+tH)^{-1}(q)) . \tag{4.3}$$

By combining (4.2) and (4.3), we get

$$|(F+tH)^{-1}(q)-F^{-1}(q)|\leqq\mu^{-1}|t|\cdot\|H\|_\infty<|t|K/\mu . \tag{4.4}$$

Finally, choose $\eta_2<\eta_1$ so that $\eta_2 K/\mu<\varepsilon$. For $|t|\leqq\eta_2$, $\sup\{|(F+tH)^{-1}(q)-F^{-1}(q)| \,|\, (H, q)\in\mathcal{K}\times[a, b]\}<\varepsilon$; and therefore, (4.1) holds.

PROOF OF THEOREM 1. Since for any function g and sets U and V, $\sup_{(u,v)\in U\times V}|g(u, v)|=\sup_{u\in U}\sup_{v\in V}|g(u, v)|$, it will suffice to show (3.2) holds. Let $\mathcal{K}$ be a subset of $C_0(R)$ having compact closure. By the Arzelà-Ascoli Theorem, $\mathcal{K}$ is uniformly equicontinuous.

By Condition A, $c<F^{-1}(q)<d$, for $a\leqq q\leqq b$. Therefore, for fixed $\varepsilon>0$, there exists a $\delta>0$, $\delta<\min\{F^{-1}(a)-c, d-F^{-1}(b)\}$, so that if $|s-F^{-1}(q)|<\delta$, then

$$|H(s)-H(F^{-1}(q))|<\varepsilon\mu/2\,, \tag{4.5}$$

for every $H\in\mathcal{K}$ and $q\in[a, b]$. Here, as in the preceding proof, $\mu=\inf\{|F'(s)|\,|\,s\in[c, d]\}$.

Now choose δ_1, $0<\delta_1<\delta$, so that

$$|F'(s)-F'(F^{-1}(q))|<\varepsilon\mu^2/(2K)\,, \tag{4.6}$$

for every $q\in[a, b]$, whenever $|s-F^{-1}(q)|<\delta_1$. This can be done because F' is continuous, and hence uniformly continuous, on $[c, d]$. Now, by virtue of Lemma 1, there is an $\eta>0$ such that $|(F+tH)^{-1}(q)-F^{-1}(q)|<\delta_1$ for all $(H, q)\in\mathcal{K}\times[a, b]$ and $|t|\leqq\eta$. We now wish to show

$$\sup_{(H,q)\in\mathcal{K}\times[a,b]}\left|\frac{(F+tH)^{-1}(q)-F^{-1}(q)}{t}+\frac{H(F^{-1}(q))}{F'(F^{-1}(q))}\right|\leqq\varepsilon\,, \tag{4.7}$$

whenever $|t|\leqq\eta$.

First, assume $H(F^{-1}(q))=0$, for some $H\in\mathcal{K}$, and assume $(F+tH)^{-1}(q)\neq F^{-1}(q)$. Note that the argument between the absolute value signs of (4.7) is exactly zero if $(F+tH)^{-1}(q)=F^{-1}(q)$. Now, using (4.3), (4.5), and the mean value theorem, we get

$$\begin{aligned}
&|(F+tH)^{-1}(q)-F^{-1}(q)|/|t|\\
&=\left|\frac{(F+tH)^{-1}(q)-F^{-1}(q)}{F((F+tH)^{-1}(q))-F(F^{-1}(q))}\right|\frac{|t||H((F+tH)^{-1}(q))|}{|t|}\\
&=|H((F+tH)^{-1}(q))-H(F^{-1}(q))|/|F'(\xi_{H,q,t})|\\
&\leqq\mu^{-1}\varepsilon\mu/2=\varepsilon/2\,,
\end{aligned}$$

where $\xi_{H,q,t}$ lies between $(F+tH)^{-1}(q)$ and $F^{-1}(q)$.

Second, assume $H(F^{-1}(q))\neq0$. By virtue of (4.3), $(F+tH)^{-1}(q)\neq F^{-1}(q)$ as long as $t\neq0$. Then using $|t|<\eta$, (4.3), (4.5), (4.6), and the mean value theorem, we have

$$\begin{aligned}
&\left|\frac{(F+tH)^{-1}(q)-F^{-1}(q)}{t}+\frac{H(F^{-1}(q))}{F'(F^{-1}(q))}\right|\\
&=\left|\frac{-1}{t}\cdot\frac{(F+tH)^{-1}(q)-F^{-1}(q)}{F((F+tH)^{-1}(q))-q}\cdot tH((F+tH)^{-1}(q))+\frac{H(F^{-1}(q))}{F'(F^{-1}(q))}\right|\\
&=\left|\frac{H((F+tH)^{-1}(q))}{F'(\xi_{H,q,t})}-\frac{H(F^{-1}(q))}{F'(F^{-1}(q))}\right|
\end{aligned}$$

$$\leq \left| \frac{H((F+tH)^{-1}(q))-H(F^{-1}(q))}{F'(F^{-1}(q))} \right|$$
$$+|H((F+tH)^{-1}(q))|\cdot \left| \frac{F'(\xi_{H,q,t})-F'(F^{-1}(q))}{F'(\xi_{H,q,t})F'(F^{-1}(q))} \right|$$
$$\leq (\varepsilon/2)\mu\mu^{-1}+(K/\mu^2)(\varepsilon/2)(\mu^2/K)=\varepsilon ,$$

where $\xi_{H,q,t}$ lies between $(F+tH)^{-1}(q)$ and $F^{-1}(q)$.

Putting these two cases together, we have proved (4.7) and consequently proved

$$(4.8)\quad \lim_{t\to 0}\ \sup_{(H,q)\in \mathcal{K}\times[a,b]} \left| \frac{(F+tH)^{-1}(q)-F^{-1}(q)}{t}+\frac{H(F^{-1}(q))}{F'(F^{-1}(q))} \right| = 0 .$$

This proves Theorem 1.

The following technical lemma is used in proving Theorem 3. We shall assume $\hat{F}_n$ and $\tilde{F}_n$ are defined as before and for fixed $\omega\in\Omega$, $\hat{F}_n(\omega)$ jumps at $x_1, x_2, \cdots, x_m$.

LEMMA 2. *Let* $d_n(\omega)=\|\tilde{F}_n(\omega)-\hat{F}_n(\omega)\|_\infty$. *Suppose* $\sup_{1\leq i\leq m-1}|\hat{F}_n(x_{i+1},\omega)-\hat{F}_n(x_i,\omega)|<2d_n(\omega)<\min\{a/4,(1-b)/4\}$. *Then for all* q *in* $[a,b]$

$$(4.9)\quad |\tilde{F}_n^{-1}(q,\omega)-\hat{F}_n^{-1}(q,\omega)|\leq|\tilde{F}_n^{-1}(q-3d_n(\omega),\omega)-\tilde{F}_n^{-1}(q,\omega)|$$
$$+|\tilde{F}_n^{-1}(q+3d_n(\omega),\omega)-\tilde{F}_n^{-1}(q,\omega)| .$$

PROOF. Let $a\leq q\leq b$. Let i be the positive integer for which $q_i\leq q<q_{i+1}$, where $q_k=\hat{F}_n(x_k,\omega)$, $k=1,2,\cdots,m$. Note $0<q-3d_n(\omega)$ and $q+3d_n(\omega)<1$. Now set $x=\tilde{F}_n^{-1}(q-3d_n(\omega),\omega)$. We claim $x\leq x_i$. Suppose not, that is, suppose $x_i<x$. Then

$$q_i-d_n(\omega)=\hat{F}_n(x_i,\omega)-d_n(\omega)\leq\tilde{F}_n(x_i,\omega)<\tilde{F}_n(x)=q-3d_n(\omega) ,$$

$q_i\leq q-2d_n(\omega)<q_{i+1}-2d_n(\omega)\leq q_{i+1}-(q_{i+1}-q_i)=q_i$, a contradiction. Thus $\tilde{F}_n^{-1}(q-3d_n(\omega),\omega)=x\leq x_i=\hat{F}_n^{-1}(q_i,\omega)\leq\hat{F}_n^{-1}(q,\omega)$, so $\tilde{F}_n^{-1}(q-3d_n(\omega),\omega)\leq \hat{F}_n^{-1}(q,\omega)$. By using a similar argument we can also show $\hat{F}_n^{-1}(q,\omega)\leq \tilde{F}_n^{-1}(q+3d_n(\omega),\omega)$. These two inequalities clearly imply our assertion, hence our proof is complete.

PROOF OF THEOREM 3. Let δ and ε be positive real numbers. Since $\{c_n(\tilde{F}_n-F)\}$ is tight in $C_0(R)$, there is a set $\mathcal{K}\subseteq C_0(R)$ with compact closure so that

$$\mathrm{P}\{\omega\in\Omega\,|\,c_n(\tilde{F}_n-F)(\omega)\notin\mathcal{K}\}<\delta/2$$

for $n=1,2,\cdots$. Set $\Omega_{\mathcal{K}}=\{\omega\in\Omega\,|\,c_n(\tilde{F}_n-F)(\omega)\in\mathcal{K}\}$. By the Arzelà-Ascoli Theorem (Billingsley [6], Appendix 1), $\mathcal{K}$ is uniformly bounded and uniformly equicontinuous. Now let a', b' be real numbers such that $0<a'<a\leqq b<b'<1$ and $c<F^{-1}(a')<F^{-1}(a)\leqq F^{-1}(b)<F^{-1}(b')<d$. Next for each $H\in\mathcal{K}$, define the continuous function G_H on $[a', b']$ by the formula $G_H(q)=H(F^{-1}(q))/F'(F^{-1}(q))$. By virtue of the above fact it is easy to deduce that $\{G_H|H\in\mathcal{K}\}$ is a uniformly bounded and uniformly equicontinuous family of functions defined on the interval $[a', b']$. Hence there is an $\eta>0$ so that

$$\left|\frac{H(F^{-1}(q))}{F'(F^{-1}(q))}-\frac{H(F^{-1}(q'))}{F'(F^{-1}(q'))}\right|<\varepsilon/6\ ,$$

whenever $|q-q'|<\eta$, $q, q'\in[a', b']$. Let $M=\sup_{a'\leqq q\leqq b'}|F'(F^{-1}(q))|^{-1}$ and set $\alpha=\min\{\eta/3, (a-a')/4, (b'-b)/4, \varepsilon/18M\}$. Since $c_n(\tilde{F}_n-\hat{F}_n)\xrightarrow{\text{P}}0$, there exists a positive integer N so that for $n\geqq N$

$$\text{P}\{\omega\in\Omega\,|\,\|c_n(\tilde{F}_n(\omega)-\hat{F}_n(\omega))\|_\infty\geqq\alpha\}<\delta/2\ .$$

Set $\Omega_\alpha=\{\omega\in\Omega\,|\,\|c_n(\tilde{F}_n(\omega)-\hat{F}_n(\omega)\|_\infty<\alpha\}$. For the moment assume that we have already shown

$$\Omega_\alpha\cap\Omega\subseteqq\{\omega\in\Omega\,|\,\|c_n(\tilde{F}_n^{-1}(\omega)-\hat{F}_n^{-1}(\omega))\|_{[a,b]}<\varepsilon\}\ , \tag{4.10}$$

where $\|G\|_{[a,b]}=\sup_{a\leqq x\leqq b}|G(x)|$. Thus

$$\begin{aligned}\text{P}\{\omega\in\Omega\,|\,&\|c_n(\tilde{F}_n^{-1}(\omega)-\hat{F}_n^{-1}(\omega))\|_{[a,b]}\geqq\varepsilon\}\\ &\leqq\text{P}(\Omega_\alpha^C)+\text{P}(\Omega_{\mathcal{K}}^C)<\delta/2+\delta/2=\delta\ , \qquad n\geqq N\ ,\end{aligned}$$

so (4.10) implies $\|c_n(\tilde{F}_n^{-1}-\hat{F}_n^{-1})\|_{[a,b]}\xrightarrow{\text{P}}0$, which completes the proof. Thus we need to establish (4.10).

Now let $\omega\in\Omega_\alpha\cap\Omega_{\mathcal{K}}$ and set $d_n(\omega)=\|\tilde{F}_n(\omega)-\hat{F}_n(\omega)\|_\infty$. Because $\omega\in\Omega_\alpha$, $d_n(\omega)<\alpha/c_n$, which implies for suitably large n that $a'<q-3d_n(\omega)<q+3d_n(\omega)<b'$, for all q, $a\leqq q\leqq b$ and that $d_n(\omega)<\min\{a/4, (1-b)/4\}$. It is easy to see that $|\hat{F}_n(x_{i+1}, \omega)-\hat{F}_n(x_i, \omega)|\leqq 2d_n(\omega)$ for all i and therefore, by Lemma 2,

$$\begin{aligned}|\hat{F}_n^{-1}(q, \omega)-\tilde{F}_n^{-1}(q, \omega)|\leqq&|\tilde{F}_n^{-1}(q-3d_n(\omega), \omega)-\tilde{F}_n^{-1}(q, \omega)|\\ &+|\tilde{F}_n^{-1}(q+3d_n(\omega), \omega)-\tilde{F}_n^{-1}(q, \omega)|\end{aligned} \tag{4.11}$$

for all q, $a\leqq q\leqq b$. Now for q in $[a, b]$

$$\begin{aligned}&c_n|\tilde{F}_n^{-1}(q-3d_n(\omega), \omega)-\tilde{F}_n^{-1}(q, \omega)|\\ &\quad\leqq\sup_{H\in\mathcal{K}}\left|\frac{(F+(1/c_n)H)^{-1}(q-3d_n(\omega))-F^{-1}(q-3d_n(\omega))}{(1/c_n)}\right.\\ &\qquad\left.+\frac{H(F^{-1}(q-3d_n(\omega))}{F'(F^{-1}(q-3d_n(\omega))}\right|\\ &\qquad+\sup_{H\in\mathcal{K}}\left|\frac{(F+(1/c_n)H)^{-1}(q)-F^{-1}(q)}{(1/c_n)}+\frac{H(F^{-1}(q))}{F'(F^{-1}(q))}\right|\end{aligned}$$

$$+\sup_{H\in\mathcal{K}}\left|\frac{H(F^{-1}(q-3d_n(\omega))}{F'(F^{-1}(q-3d_n(\omega))}-\frac{H(F^{-1}(q))}{F'(F^{-1}(q))}\right|$$
$$+c_n|F^{-1}(q-3d_n(\omega))-F^{-1}(q)|\,.$$

Remember $\omega\in\Omega_\alpha\cap\Omega_{\mathcal{K}}$. So $|q-3d_n(\omega)-q|=3d_n(\omega)<\min\{\eta_1, \varepsilon/6c_nM\}$. By virtue of equicontinuity of $\{G_H\}$ and the mean value theorem

$$c_n|\tilde{F}_n^{-1}(q-3d_n(\omega), \omega)-\tilde{F}_n^{-1}(q, \omega)|$$
$$\leqq 2\sup_{(H, q')\in\mathcal{K}\times[a', b']}\left|\frac{(F+(1/c_n)H)^{-1}(q')-F^{-1}(q')}{(1/c_n)}+\frac{H(F^{-1}(q'))}{F'(F^{-1}(q'))}\right|$$
$$+\frac{1}{6}\varepsilon+\frac{1}{6}\varepsilon\,.$$

Now by virtue of Theorem 1 with a', b' playing the role of a, b in Theorem 1, there is an integer $N_1>N$ so that for $n>N_1$ we have

$$\sup_{(H,q')\in\mathcal{K}\times[a', b']}\left|\frac{(F+(1/c_n)H)^{-1}(q')-F^{-1}(q')}{(1/c_n)}+\frac{H(F^{-1}(q'))}{F'(F^{-1}(q'))}\right|<\varepsilon/12\,.$$

Thus $c_n|\tilde{F}_n^{-1}(q-2d_n(\omega), \omega)-\tilde{F}_n^{-1}(q, \omega)|<\varepsilon/2$. Similarly, we can show $c_n|\tilde{F}_n^{-1}(q+2d_n(\omega), \omega)-\tilde{F}_n^{-1}(q, \omega)|<\varepsilon/2$. It follows from (4.11) that $c_n\|\hat{F}_n^{-1}(\omega)-\tilde{F}_n^{-1}(\omega)\|_{[a,b]}<\varepsilon$, that is, (4.10) holds and our proof is complete.

PROOF OF THEOREM 4. Choose any $\varepsilon, \delta>0$. By (3.3)

$$r_n(\omega)=\sqrt{n}[T(\tilde{F}_n(\omega))-T(F)]+\int_{[a,b]}\frac{\sqrt{n}[\tilde{F}_n(F^{-1}(s), \omega)-s]}{F'(F^{-1}(s))}M(ds)\,.$$

By hypothesis $\{\sqrt{n}(\tilde{F}_n-F)\}$ is tight. Therefore, there exists a compact set $\mathcal{K}\subseteqq C_0(R)$ such that

$$\mathrm{P}\{\omega\in\Omega\,|\,\sqrt{n}(\tilde{F}_n(\omega)-F)\notin\mathcal{K}\}<\delta \qquad \text{for } n=1, 2, 3,\cdots\,.$$

By Theorem 1, there is a positive integer N such that, for $n>N$ and $\omega\in\{\omega'\in\Omega\,|\,\sqrt{n}(\tilde{F}_n(\omega')-F)\notin\mathcal{K}\}$, $|r_n(\omega)|<\varepsilon$. Thus, for $n\geqq N$,

$$\{\omega'\in\Omega\,|\,\sqrt{n}(\tilde{F}_n(\omega')-F)\in\mathcal{K}\}\subseteqq\{\omega'\in\Omega\,|\,|r_n(\omega)|<\varepsilon\}\,,$$

which implies that $\mathrm{P}\{\omega'\in\Omega\,|\,|r_n(\omega')|\geqq\varepsilon\}\leqq\mathrm{P}\{\omega'\in\Omega\,|\,\sqrt{n}(\tilde{F}_n(\omega')-F)\notin\mathcal{K}\}<\delta$ i.e., $r_n\xrightarrow{\mathrm{P}}0$ as $n\to\infty$. This proves Theorem 4.

Note how the compact set $\mathcal{K}$ required by the definition of "compact derivative" arises directly from the tightness of $\{\sqrt{n}(\tilde{F}_n-F)\}$. This is precisely the reason that the compact derivative is the appropriate definition of D_F^T to use as a means of finding the asymptotic distribution T.

LEMMA 3. *Let Q be a regular Borel measure on the real line with compact support contained in the interval $[c, d]$ (see Condition A). Let $L: C_0(R) \to R$ be the continuous linear functional given by $L(f) = \int f dQ$ for $f \in C_0(R)$. Let G be a Gaussian random element in $C_0(R)$ such that* $\mathrm{E}[G(s)] = 0$ *for each $s \in [c, d]$ and the covariance kernel* $\mathrm{Cov}[G(s), G(t)]$ *is continuous on $[c, d] \times [c, d]$. Then the random variable $L(G)$ is distributed as $N(0, \sigma^2_{L,G})$, where $\sigma^2_{L,G} = \int\int \mathrm{Cov}[G(s), G(t)] Q(ds) Q(dt)$.*

PROOF OF LEMMA 3. Although we do not know of an appropriate citation, Lemma 3 seems to be well-known. The proof is straightforward; only a sketch is given here. Define a sequence $\{L_n(f)\}$ such that $L_n(f)$ is a Riemann-Stieltjes approximating sum for $\int f dQ$ and such that $L_n(f) \to L(f)$ as $n \to \infty$. Then, $L_n(f)$ is normally distributed, and it is easy to show that the characteristic function of $L_n(f)$ goes to a characteristic function of the appropriate normal random variable as $n \to \infty$.

PROOF OF THEOREM 5. For each Borel set E in $[F^{-1}(a), F^{-1}(b)]$, define the Borel measure $MF(E) = M(F(E))$. By Billingsley ([7], Theorem 16.12) and Theorem 4,

$$\sqrt{n}[T(\tilde{F}_n) - T(F)] = -\int_{[F^{-1}(a), F^{-1}(b)]} \frac{\sqrt{n}[\tilde{F}_n(s) - F(s)]}{F'(s)} MF(ds) + r_n , \tag{4.12}$$

where $r_n = o_p(1)$. The hypotheses of tightness and convergence of the finite dimensional distributions of $\sqrt{n}(\tilde{F}_n - F)$ imply that $\sqrt{n}(\tilde{F}_n - F) \xrightarrow{\mathscr{D}} G$. Therefore, by Corollary 1, (4.12), and Billingsley [6], Section 5,

$$\sqrt{n}(T(\tilde{F}_n) - T(F)) \xrightarrow{\mathscr{D}} \int_{[F^{-1}(a), F^{-1}(b)]} \frac{-G(s)}{F'(s)} MF(ds) .$$

Define the measure Q by

$$Q(E) = \int_E -I_{[F^{-1}(a), F^{-1}(b)]}(s) \cdot \frac{1}{F'(s)} MF(ds) .$$

Then

$$\int_{[F^{-1}(a), F^{-1}(b)]} \frac{-G(s)}{F'(s)} MF(ds)$$
$$= \int_R G(s)(-I_{[F^{-1}(a), F^{-1}(b)]}(s))(F'(s))^{-1} MF(ds) = \int_R G(s) Q(ds) ,$$

which, by Lemma 3, is distributed as $N(0, \sigma^2_{L,G})$, where

$$\sigma^2_{L,G} = \int\int \mathrm{Cov}[G(s), G(t)] Q(ds) Q(dt) .$$

Finally, by Billingsley ([7], Theorem 16.12), we see that $\sigma^2_{L,G}=\sigma^2_{T,F}$.

5. Applications of Theorem 1

In this section, the symbol T denotes the functional representation of an L-estimator (see Definition 1) where the subscript M is suppressed. The proofs of the results of this section are found in Esty et al. [12].

APPLICATION 1: *One-sample data. Let* $\{X_i\}_{i=1}^n$ be i.i.d. with common distribution function F, which is continuous and strictly increasing on the interval $\{x\,|\,0<F(x)<1\}$. For sample size n, let $X_{n,1}<X_{n,2}<\cdots<X_{n,n}$ be the ordered $X_1,\cdots,X_n$. Let $\hat{F}_n(x)$ be the empirical distribution function given by $\hat{F}_n(x)=0$, if $x<X_{n,1}$, $\hat{F}_n(x)=k/n$, if $X_{n,k}\leqq x<X_{n,k+1}$ and $\hat{F}_n(x)=1$, if $x\geqq X_{n,n}$. Choose any sequence of positive numbers $\{d_i\}_{i=1}^\infty$ such that $\lim_{i\to\infty} d_i=0$. Define $X_{n,0}=X_{n,1}-d_n$. Let $\tilde{F}_n(x)$ be the continuous, piece-wise linear version of $\hat{F}_n(x)$ formed by linearly interpolating between $(X_{n,k}, \hat{F}_n(X_{n,k}))$ and $(X_{n,k+1}, \hat{F}_n(X_{n,k+1}))$, for $k=0,\cdots,n-1$. The c.d.f. estimator $\tilde{F}_n$ will now be used in Theorem 5. The $\hat{F}_n-F$ is not in $C_0(R)$ and therefore cannot be used directly in applying Theorem 5. The following two lemmas are apparently known, but, because we cannot find published proofs, they are given here for completeness.

LEMMA 4. *Let $Y_n(x)$ be the random element in $C_0(R)$ defined by $Y_n(x)=\sqrt{n}(\tilde{F}_n(x)-F(x))$, $-\infty<x<\infty$. Then $\{Y_n\}_{n=1}^\infty$ is a tight sequence of random elements.*

LEMMA 5. *$Y_n\xrightarrow{\mathcal{D}}G^0$, where G^0 is the random element in $C_0(R)$ such that G^0 is a (tied-down) Gaussian process with* $\mathrm{E}[G^0(t)]=0$, *for every $t\in R$, and* $\mathrm{Cov}[G^0(s), G^0(t)]=F(s)[1-F(t)]$, *for every $s<t$.*

This lemma is an extension of Theorem 13.1 in Billingsley [6].

THEOREM 6. *Under Conditions A and B.*

$$\sqrt{n}[T(\tilde{F}_n)-T(F)]\xrightarrow{\mathcal{D}}N(0,\sigma^2_{T,F}),$$

where

$$\sigma^2_{T,F}=\int_a^b\int_a^b[(s\wedge t-st]/[F'(F^{-1}(s))F'(F^{-1}(t))]M(ds)M(dt)\,. \tag{5.1}$$

Here $s\wedge t=\min\{s,t\}$.

The hypotheses of Theorem 3 are satisfied and Theorem 6 will hold for both the continuous and discontinuous versions, $\tilde{F}_n$ and $\hat{F}_n$.

Theorem 6 agrees with all previous asymptotic results for L-estimators for the one-sample problem (Reeds [23], Chapt. 6, Boos [8], Serfling [26], Mason [21], Huber [17]). It is more general, however, because it holds for a wider class of L-estimators.

As a specific case that is covered by Theorem 6 consider a weighted average of the median and the $100\cdot\alpha\%$ trimmed mean. In other words, let $T(\lambda)$ be associated with $M(s)=\lambda M_1(s)+(1-\lambda)M_2(s)$, $0<\lambda<1$, where $M_1(s)=0$, if $0\leqq s<1/2$, $M_1(s)=1$, if $1/2\leqq s\leqq 1$, and where $M_2(s)=0$, if $0\leqq s<\alpha$, $M_2(s)=(s-\alpha)/(1-2\alpha)$, if $s\in[\alpha, 1-\alpha]$, and $M_2(s)=1$, if $1-\alpha<s\leqq 1$ all for fixed $\alpha\in(0, 1/2]$. By substituting the above M into (5.1) and then simplifying the resulting expression, the asymptotic normal variance of $\sqrt{n}\,T(\lambda)$ is

$$\begin{aligned}\sigma^2_{T(\lambda),F}=&(\lambda/4)[F'(F^{-1}(1/2))]^{-2}+[\lambda(1-\lambda)/F'(F^{-1}(1/2))]\\ &\times\left[\alpha(F^{-1}(1-\alpha)-F^{-1}(\alpha))+\int_\alpha^{1/2}(F^{-1}(1-y)-F^{-1}(y))dy\right]\\ &+[2(1-\lambda)^2/(1-2\alpha)^2]\left[\int_\alpha^{1-\alpha}[(1-t)/F'(F^{-1}(t))]\right.\\ &\left.\times\int_\alpha^t[s/F'(F^{-1}(s))]dsdt\right].\end{aligned}$$

A useful, by-product of this calculation is the asymptotic covariance between the sample median and the sample $100\cdot\alpha\%$ trimmed mean,

$$\left[\alpha(F^{-1}(1-\alpha)-F^{-1}(\alpha))+\int_\alpha^{1/2}(F^{-1}(1-y)-F^{-1}(y))dy\right]\Big/[2nF'(F^{-1}(1/2))].$$

This form is much more convenient than that of Serfling ([26], p. 280). This method could be used to find the covariance of, for instance, the trimmed mean and the interquantile range.

Unfortunately, the asymptotic normal variance (5.1) depends on the true, unknown F. In order to form large sample confidence intervals for $T(F)$, a consistent estimator of $\sigma^2_{T,F}$ is required. The following statements describe one such estimator of $\sigma^2_{T,F}$.

DEFINITION 3. Choose a sequence of real numbers $\{h_n\}_{n=1}^\infty$ such that $h_n\to 0$ and $\sqrt{n}h_n\to\infty$. Define

$$\tilde{f}_n(x)=[\tilde{F}_n(x+h_n)-\tilde{F}_n(x-h_n)]/2h_n\,,\qquad x\in(-\infty,\infty)\,. \tag{5.2}$$

Then an estimator of $\sigma^2_{T,F}$ is

$$\tilde{\sigma}^2_n=\int_a^b\int_a^b(s\wedge t-st)/[\tilde{f}_n(\tilde{F}_n^{-1}(s))\tilde{f}_n(\tilde{F}_n^{-1}(t))]M(ds)M(dt)\,. \tag{5.3}$$

LEMMA 6. *Under Condition A,*

$$\sup_{a \leq s \leq b} \{|\tilde{f}_n(\tilde{F}_n^{-1}(s)) - F'(F^{-1}(s))|\} \xrightarrow{P} 0 .$$

THEOREM 7. *Under Conditions A and B,*

$$\sqrt{n}(T(\tilde{F}_n) - T(F))/\tilde{\sigma}_n \xrightarrow{\mathcal{D}} N(0, 1) .$$

As in Theorem 6, $T(\tilde{F}_n)$ may be replaced by $T(\hat{F}_n)$.

Notice that Theorem 7 allows one to calculate an approximate confidence interval for the unknown parameter $T(F)$.

APPLICATION 2: *Data grouped by quantiles.* Use the same notation as for the preceding one-sample application. In some situations, the statistician does not have access to the individual values X_{ni}, $i=1,\cdots,n$. For example, to protect the confidentiality of respondents, the interviewer may provide only the sample deciles or sample percentiles when releasing the survey data for analysis. It will now be demonstrated that the L-estimate can easily be calculated from the quantile information and that the associated large sample theory follows directly from Theorem 6.

The following notation is convenient. For sample size n, choose the $k(n)$ quantities ξ_{nj}, $j=1,\cdots,k(n)$, where $k(n)<n$ and $0=\xi_{n0}<\xi_{n1}<\cdots<\xi_{n,k(n)}<\xi_{n,k(n)+1}=1$. Define $e_n=\min\{\xi_{nj}-\xi_{n,j-1},\ j=1,\cdots,k(n)+1\}$. Let $e_n>1/n$ and let $Y_{n1}<Y_{n2}<\cdots<Y_{n,k(n)}$ be the associated set of sample quantiles defined by $Y_{nj}=\max\{X_{n,i}\,|\,\bar{F}_n(X_{n,i})\leqq\xi_{ni},\ i=1,\cdots,n\}$, $j=1,\cdots,k(n)$. Thus, the sample deciles correspond to $k(n)=9$, $\xi_{nj}=j/10$ the sample percentiles correspond to $k(n)=99$, $\xi_{nj}=j/100$, etc. The estimate of F is $\bar{F}_n$, a function formed by linearly interpolating between the points (Y_{nj}, ξ_{nj}) and $(Y_{n,j+1}, \xi_{n,j+1})$, for $j=0, 1,\cdots,k(n)$. Here, in a manner similar to that used to form $\tilde{F}_n$, define $Y_{n0}=Y_{n1}-d_n$ and $Y_{n,k(n)+1}=Y_{n,k(n)}+d_n$, where $\{d_n\}$ is a sequence of positive values with $d_n\to 0$.

It is important that the ξ_{ni}'s be far enough apart that the Y_{ni}'s are distinct (remember $e_n>1/n$) and also that the ξ_{ni}'s be close enough together that $\{\bar{F}_n\}$ is consistent for F. Let $u_n=\max\{\xi_{nj}-\xi_{n,j-1},\ j=1,\cdots,k(n)+1\}$.

THEOREM 8. *Suppose Conditions A and B hold. If* $u_n=o(n^{-1/2})$, *then* $\sqrt{n}(T(\bar{F}_n)-T(F))\xrightarrow{\mathcal{D}} N(0, \sigma^2_{T,F})$, *where* $\sigma^2_{T,F}$ *is given by* (5.1).

It seems clear that a result similar to Theorem 7 could be found for this quantile data case. Apparently, Theorem 8 is new. We know of no other work done of L-estimators for data grouped by quantiles.

APPLICATION 3: *Censored survival time data.* Let the X_i's be as defined in the one-sample data example. Let $F(0)=0$. Let $\{V_n\}_{n=1}^{\infty}$ be i.i.d. with common distribution F_V, which is continuous and strictly increasing on the interval $\{x|0<F_V(x)<1\}$. Let $\{V_i\}$ be independent of $\{X_i\}$. Observe the sequence of random variables $W_i=X_i\wedge V_i$. Then the distribution function of W, F_W, is given by $F_W(x)=1-(1-F(x))\cdot(1-F_V(x))$. The variable W_i is the type of measurement that one observes in right-censored medical follow-up studies or in product life testing, when the censoring is random. (Here V is the random censor time and X is the life time.) Such experiments are often conducted to estimate the survival distribution function F or to estimate some parameter related to that distribution. See Kalbfleish and Prentice [19] for a complete discussion of survival time data.

Let $\delta_i=I_{(-\infty,V_i]}(X_i)$, where I is the indicator function. That is, $\delta_i=0$, if the ith observation is censored, and $\delta_i=1$, if the ith observation is not censored. Because the experimenter knows whether or not each observation is censored, the data actually consists of the pairs (W_i, δ_i), $i=1,\cdots,n$. Let $W_{n,1}<\cdots<W_{n,n}$ be the ordered $W_1,\cdots,W_n$ and let $\delta_{n,i}$ be the censoring indicator variable that goes with $W_{n,i}$. Define $m=\sum_{i=1}^{n}\delta_{i,n}$; m is the number of uncensored, observed survival times, $m\leqq n$.

The Kaplan-Meier product-limit estimator of F (Kaplan and Meier [20]) is given by

$$\begin{aligned}\hat{F}_n(x)&=0\,, && \text{if } x<W_{n,1}\,,\\ &=1-\prod_{i=1}^{j}[(n-i)/(n-i+1)]^{\delta_i}\,, && W_{n,j}\leqq x<W_{n,j+1},\ j=1,\cdots,n-1\,,\\ &=1-\prod_{i=1}^{n}[(n-i)/(n-i+1)]^{\delta_i}\,, && W_{n,n}\leqq x\,.\end{aligned}$$

Notice that $\hat{F}_n$ is a step function that jumps only at the m observed survival times.

We shall use some of the large sample theory for $\hat{F}_n$ as derived by Breslow and Crowley [9]. Let $F^*(x)=\int_0^x(1-F_V(v))dF(v)$ be the "sub-distribution function" of censored observations.

LEMMA 7. *Let $\tau<\infty$ satisfy $F(\tau)<1$. Then the random element $\sqrt{n}(\hat{F}_n-F)\xrightarrow{\mathcal{D}}G^*$, over the interval $(0,\tau)$, where G^* is the Gaussian process with* $\mathrm{E}[G^*(t)]=0$, *for $t\in(0,\tau)$ and*

$$\mathrm{Cov}\,[G^*(s),G^*(t)]=(1-F(s))(1-F(t))\int_0^s(1-F_W(v))^{-2}dF^*(v)\,,$$

$0<s\leqq t<\tau$.

PROOF. See Breslow and Crowley [9].

In order to use Theorem 5, we need an estimator that is close to $\hat{F}_n$ and is in $C_0(R)$. Consider the estimator $\tilde{F}_n$ created by linear interpolation between $(W_{n,i}, \hat{F}_n(W_{n,i}))$ and $(W_{n,i+1}, \hat{F}_n(W_{n,i+1}))$, $i=0, 1\cdots, n-1$, where $W_{n,0}=\hat{F}_n(W_{n,0})=0$. Let $\tilde{F}_n(x)=\hat{F}_n(W_{n,n})$, if $x \geqq W_{n,n}$. Then $\tilde{F}_n \in C[0, \tau]$, the set of continuous functions on $[0, \tau]$.

LEMMA 8. *If* $F(\tau)<1$, *then*

$$\sqrt{n} \sup_{0\leq x\leq \tau} \{|\hat{F}_n(x)-\tilde{F}_n(x)|\} \xrightarrow{P} 0 .$$

Thus the hypotheses of Theorem 3 are satisfied and Theorem 9 will hold for both the continuous and discontinuous versions, $\tilde{F}_n$ and $\hat{F}_n$.

THEOREM 9. *Suppose Conditions A and B hold. If* τ *satisfies* $b< F(\tau)<1$, *then* $\sqrt{n}(T(\tilde{F}_n)-T(F)) \xrightarrow{\mathcal{D}} N(0, \sigma^2_{T,F})$, *where*

$$\sigma^2_{T,F}=\iint_{[a,b]\times[a,b]} \frac{(1-s)(1-t)R(s\wedge t)}{F'(F^{-1}(s))F'(F^{-1}(t))} M(ds)M(dt) , \tag{5.4}$$

where

$$R(u)=\int_0^{F^{-1}(u)} (1-F_W(z))^{-2}dF^*(z)=\int_0^u [1-F_V(F^{-1}(z))]^{-1}(1-z)^{-2}dz .$$

This result agrees with Reid [24] ((5.4) is identical to Reid's (3.6)), who used the Frèchet definition of D^T_F and required that M be absolutely continuous.

The asymptotic Gaussian process for the Kaplan-Meier estimator as derived by Meier [22] could be used in a manner analogous to Lemma 6 to produce a result similar to Corollary 5 if the censoring times are fixed rather than random.

6. Discussion and extensions

The results of this article are easily extended to the simultaneous estimation case, where the range of T is in R^k. For example, one can use the preceding theory to determine the asymptotic bivariate normal distribution of the interquantile range and the trimmed mean.

It should be a straight forward matter to allow the weighting function M to depend on n as do, for example, Groeneboom et al. ([15], Theorem 6.2). It would be more useful, but more difficult mathematically, if M were data dependent so that the associated L-estimator is adaptive. (See Hogg [16]).

Also, it may be possible to let M depend upon F. Cases where trimming is at a fixed distance from $T(F)$ (rather than, say, the extreme 10%) can be shown, by a change of variable argument, to correspond to such M's.

It also appears feasible to generalize the theorems of this article to cover the case where $\tilde{F}_n$ lies in the product space $C^k(R)$ (see Whitt [30]).

One of the authors (Taylor [28]) has used Theorem 1 to derive asymptotic distributions for a very broad class of estimators which includes L, M, R-estimators.

Acknowledgements

The work of M. Hamilton was supported in part by NIEHS Research Career Development Award No. 1–K04–ES0006. The work of D. Taylor was supported in part by a Faculty Research-Creativity Grant from the Office for Research, MSU.

MONTANA STATE UNIVERSITY

REFERENCES

[1] Andrews, D. F., Bickel, P. J. et al. (1972). *Robust Estimates of Location*, Princeton University Press.
[2] Averbukh, V. I. and Smolyanov, O. G. (1967). The theory of differentiation in linear topological spaces, *Russian Math. Surveys*, **22**, 201-258.
[3] Averbukh, V. I. and Smolyanov, O. G. (1968). The various definitions of the derivative in linear topological spaces, *Russian Math. Surveys*, **23**, 67-113.
[4] Bickel, P. J. and Lehmann, E. L. (1975). Descriptive statistics for nonparametric models II. Location, *Ann. Statist.*, **3**, 1045-1069.
[5] Bickel, P. J. and Lehmann, E. L. (1979). Descriptive statistics for nonparametric models IV. Spread, *Contributions to Statistics* (ed. J. Jureckova), D. Reidel Publ. Co., Boston.
[6] Billingsley, Patrick (1968). *Convergence of Probability Measures*, John Wiley and Sons, New York.
[7] Billingsley, Patrick (1979). *Probability and Measure*, John Wiley and Sons, New York.
[8] Boos, Dennis D. (1979). A differential for L-statistics., *Ann. Statist.*, **7**, 955-959.
[9] Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under censorship, *Ann. Statist.*, **2**, 437-453.
[10] Dodge, Yadolah and Lindstrom, F. Tom (1981). An alternative to least squares estimation when dealing with contaminated data, *Technical Report*, No. 79, Oregon State University, Department of Statistics.
[11] Ershov, A. A. (1978). Stable methods of estimating parameters, Survey *Automation and Remote Control*, **39**, 1152-1181.
[12] Esty, W., Gillette, R, Hamilton, M. and Taylor, D. (1983). Asymptotic distribution theory for statistical functionals: the compact derivative approach for robust estimators, *MSU Technical Report*.
[13] Fernholz, Luisa T. (1979). On compact differentiability and asymptotic properties of robust estimators of location, Ph.D. Thesis, Rutgers University.
[14] Gihman I. I. and Skorohod, A. V. (1974). *The Theory of Stochastic Processes*, Vol. I. Springer-Verlag, New York.

[15] Groeneboom, J., Oosterhoff, J. and Ruymgaart, F. H. (1979). Large deviation theorems for empirical probability measures, *Ann. Prob.*, **7**, 553-586.

[16] Hogg, Robert V. (1981). On adaptive robust inferences, In *Statistics and Related Topics* (eds. M. Csorgo, D. A. Dawson, J. N. K. Rao and A. K. Md. E. Saleh), North Holland Publishing Co., New York.

[17] Huber, Peter J. (1981). *Robust Statistics*, John Wiley and Sons, New York.

[18] Jaeckel, Louis A. (1971). Some flexible estimates of location, *Ann. Math. Statist.*, **42**, 1540-1552.

[19] Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, New York.

[20] Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations, *J. Amer. Statist. Ass.*, **53**, 457-481.

[21] Mason, David M. (1981). Asymptotic normality of linear combinations of order statistics with a smooth score function, *Ann. Statist.*, **9**, 899-908.

[22] Meier, P. (1975). Estimation of a distribution function from incomplete observations, In *Perspectives in Probability and Statistics*, 67-68, (ed. J. Gani), Applied Probability Trust, Sheffield.

[23] Reeds, James A. (1976). On the definition of von Mises functionals, Ph.D. Thesis, Harvard University.

[24] Reid, Nancy (1981). Influence functions for censored data, *Ann. Statist.*, **9**, 78-92.

[25] Royden, H. L. (1968). *Real Analysis*, 2nd edition, MacMillan Publ. Co., New York.

[26] Serfling, Robert J. (1980). *Approximation Theorems of Mathematical Statistics*, John Wiley and Sons, New York.

[27] Simmons, George F. (1963). *Intro. to Topology and Modern Analysis*, McGraw-Hill, New York.

[28] Taylor, Donald C. (1985). Asymptotic distribution theory for general statistical functionals, *Ann Inst. Statist. Math.*, **37**, 131-138.

[29] Von Mises, Richard (1947). On the asymptotic distribution of differentiable statistical functions, *Ann. Math. Statist.*, **18**, 309-348.

[30] Whitt, Ward (1970). Weak convergence of probability measures on the function space $C[0, \infty)$, *Ann. Math. Statist.*, **41**, 939-944.