The Annals of Statistics
1977, Vol. 5, No. 4, 646–657

UPPER BOUNDS ON ASYMPTOTIC VARIANCES OF M-ESTIMATORS OF LOCATION[1]

BY JOHN R. COLLINS

University of Calgary

If $X_1, \cdots, X_n$ is a random sample from $F(x-\theta)$, where F is an unknown member of a specified class $\mathscr{F}$ of approximately normal symmetric distributions, then an M-estimator of the unknown location parameter θ is obtained by solving the equation $\sum_{i=1}^n \psi(X_i - \hat{\theta}_n) = 0$ for $\hat{\theta}_n$. A suitable measure of the robustness of the M-estimator is $\sup\{V(\psi, F): F \in \mathscr{F}\}$, where $V(\psi, F) = \int \psi^2 \, dF / (\int \psi' \, dF)^2$ is (under regularity conditions) the asymptotic variance of $n^{\frac{1}{2}}(\hat{\theta}_n - \theta)$. A necessary and sufficient condition for F_0 in $\mathscr{F}$ to maximize $V(\psi, F)$ is obtained, and the result is specialized to evaluate $\sup\{V(\psi, F): F \in \mathscr{F}\}$ when the model for $\mathscr{F}$ is the gross errors model or the Kolmogorov model.

1. Introduction and summary. Let $X_1, \cdots, X_n$ be i.i.d. random variables with distribution function $F((x-\theta)/\sigma)$, where θ is an unknown location parameter to be estimated and σ is a (known or unknown) scale parameter. Following Huber [5], F is unknown, but is assumed to lie in a specified class of distributions $\mathscr{F}$ which is convex and vaguely compact. Assume further that the members of $\mathscr{F}$ are symmetric ($F(-x) = 1 - F(x-0)$ for all $x \geqq 0$) and that $\mathscr{F}$ contains the standard normal distribution $\Phi(x) = \int_{-\infty}^x \varphi(t)\,dt$, where $\varphi(x) = (2\pi)^{-\frac{1}{2}} \exp(-x^2/2)$. Two important specifications of $\mathscr{F}$ are the gross errors model,

$$\mathscr{F}_{1,\varepsilon} = \{F : F = (1-\varepsilon)\Phi + \varepsilon\, G \text{ for some symmetric } G\}, \tag{1.1}$$

and the Kolmogorov model,

$$\mathscr{F}_{2,\varepsilon} = \{F : F \text{ is symmetric and } \sup_x |F(x) - \Phi(x)| \leqq \varepsilon\}, \tag{1.2}$$

where in each model ε is a known number in $(0, 1)$.

For the case of σ known, M-estimators of θ (Huber [5]) are obtained by solving equations of the form

$$\sum_{i=1}^n \psi\left(\frac{X_i - \hat{\theta}_n}{\sigma}\right) = 0$$

for $\hat{\theta}_n$. Each ψ to be considered is assumed to lie in the class Ψ of continuous piecewise-smooth real-valued functions satisfying (i) $\psi(x) = -\psi(-x)$ for all x; (ii) $\psi(x) \geqq 0$ for all $x \geqq 0$; but $\psi \not\equiv 0$; and (iii) $\sup_x \max\{|\psi'(x-0), |\psi'(x+0)|\} < \infty$. Of particular interest are subclasses Ψ_c, defined for each $c > 0$ by: $\psi \in \Psi_c$ if $\psi \in \Psi$ and $\psi(x) > 0$ when $0 < x < c$, $\psi(x) = 0$ when $x \geqq c$.

Received August 1975; revised July 1976.

[1] Research supported by a grant from Purdue Research Foundation.

AMS 1970 *subject classifications*. Primary 62G05; Secondary 62G20, 62G35.

Key words and phrases. M-estimator, location parameter, asymptotic variance, robustness.

Under suitable regularity conditions on ψ and F([2], [3], [5] and [6]), $\hat{\theta}_n$ is a consistent estimator of θ, and $n^{\frac{1}{2}}(\hat{\theta}_n - \theta)$ is asymptotically normal with mean 0 and variance $\sigma^2 V(\psi, F)$, where

$$V(\psi, F) = \frac{\int \psi^2 \, dF}{(\int \psi' \, dF)^2}.$$

The problem considered in this paper is the following: given ψ, find $\sup\{V(\psi, F) : F \in \mathscr{F}\}$. One can regard the supremum as a measure of the robustness of the M-estimator based on ψ.

The problem of finding the ψ that minimizes $\sup\{V(\psi, F) : F \in \mathscr{F}\}$ was solved by Huber [5]. The minimax ψ has the form $\psi_0 = -f_0'/f_0$, where f_0 is the (necessarily absolutely continuous) density of the F_0 in $\mathscr{F}$ which minimizes the Fisher information. For nonminimax ψ, it will be seen that the least favorable F for ψ (i.e., the F in $\mathscr{F}$ which maximizes $V(\psi, F)$) typically does not have finite Fisher information, since it puts positive mass at least favorable points. We remark that the formally least favorable F may not satisfy the required regularity conditions. However, in typical cases the subset $\mathscr{F}'$ on which the regularity conditions hold is dense in $\mathscr{F}$, so that one obtains the correct value of $\sup\{V(\psi, F) : F \in \mathscr{F}'\}$.

One may question why one would consider any ψ other than the one minimizing $\sup\{V(\psi, F) : F \in \mathscr{F}\}$. The reason is that $\sup\{V(\psi, F) : F \in \mathscr{F}\}$ is just one of several reasonable numerical measures of robustness by which one can compare estimators. For some competing measures, see the table on page 392 of Hampel [4].

Section 2 contains some preliminary examples of finding $\sup\{V(\psi, F) : F \in \mathscr{F}\}$. Section 3 presents the main result: necessary and sufficient conditions for F in $\mathscr{F}$ to maximize $V(\psi, F)$. Sections 4 and 5 specialize the result to the gross errors model and the Kolmogorov model, respectively. The generalization of the results to the case of unknown σ is discussed in the concluding remarks in Section 6.

2. Preliminary examples. The following notation is used throughout: for $0 < \varepsilon < 1$ and $0 \leqq y \leqq \infty$, define $F_{(y),\varepsilon}$ in $\mathscr{F}_{1,\varepsilon}$ by

$$F_{(y),\varepsilon} = (1-\varepsilon)\Phi + \varepsilon G_{(y)}, \tag{2.1}$$

where

$$\begin{aligned} G_{(y)}(x) &= 0 & x &< -y \\ &= \tfrac{1}{2} & -y &\leqq x < y \\ &= 1 & x &\geqq y. \end{aligned}$$

EXAMPLE 2.1. *Minimax solution.* The ψ in Ψ_c which minimizes $\sup\{V(\psi, F) : F \in \mathscr{F}_{1,\varepsilon}\}$ has the form

$$\begin{aligned} \psi_0(x) &= x & 0 &\leqq |x| \leqq a \\ &= k \tanh\left[\tfrac{1}{2}k(c - |x|)\right] \operatorname{sgn}(x) & a &\leqq |x| \leqq c \\ &= 0 & |x| &\geqq c, \end{aligned}$$

with a and k determined by ε, provided that ε is less than a breakdown point ε_0 depending on c. This minimax property is proved in [2] by showing that $\psi_0(x) = -f_0'(x)/f_0(x)$, where f_0 is the density of a $F_0 \in \mathscr{F}_{1,\varepsilon}$ which maximizes $V(\psi_0, F)$. We note that F_0 is not the only F in $\mathscr{F}_{1,\varepsilon}$ which maximizes $V(\psi_0, F)$. To see this, let $F = (1-\varepsilon)\Phi + \varepsilon G^*$, where G^* is any symmetric distribution satisfying $G^*\{(a, c)\} = \frac{1}{2}$. Then, noting that $k^2 + 2\psi_0' \equiv \psi_0^2$ on the set (a, c), we have

$$V(\psi_0, F) = \frac{(1-\varepsilon)\int_0^c \psi_0^2\, d\Phi + \frac{1}{2}\varepsilon k^2 + 2\varepsilon \int_a^c \psi_0'\, dG^*}{2[(1-\varepsilon)\int_0^c \psi_0'\, d\Phi + \varepsilon \int_a^c \psi_0'\, dG^*]^2}\,.$$

Thus $\sup\{V(\psi_0, F) : F \in \mathscr{F}_{1,\varepsilon}\}$ is attained at all F of the above form satisfying

$$\int_a^c \psi_0'\, d\left[\frac{F_0 - (1-\varepsilon)\Phi}{\varepsilon}\right] = \int_a^c \psi_0'\, dG^*\,.$$

Note that the class of F's for which the equality holds is convex, and that in particular, there is a number $y \in (a, c)$ such that $F_{(y),\varepsilon}$ attains $\sup\{V(\psi_0, F): F \in \mathscr{F}_{1,\varepsilon}\}$.

EXAMPLE 2.2. *Hampel's piecewise linear ψ.* For $0 < a \leqq b < c$, define $\psi_{abc} \in \Psi_c$ by

$$\begin{aligned} \psi_{abc}(x) &= x & 0 \leqq |x| \leqq a \\ &= a \operatorname{sgn}(x) & a \leqq |x| \leqq b \\ &= \frac{c - |x|}{c - b} a \operatorname{sgn}(x) & b \leqq |x| \leqq c \\ &= 0 & |x| \geqq c\,. \end{aligned}$$

TABLE 1
Parameters of optimal "Hampels"

c	ε	optimal values of a	optimal values of b	$\sup\{V(\psi_{abc}, F): F \in \mathscr{F}_{1,\varepsilon}\}$
2.0	.001	1.7078	1.8558	1.4560
	.01	1.3022	1.6244	1.7774
	.05	0.8448	1.3014	2.9224
	.10	0.5967	1.1043	4.9149
	.20	0.3192	0.8031	15.2974
4.0	.001	2.5164	3.2060	1.0205
	.01	1.8253	2.6913	1.1123
	.05	1.2558	2.2331	1.4316
	.10	0.9797	1.9917	1.8645
	.20	0.6751	1.6931	3.1191
8.0	.001	2.6093	3.7677	1.0109
	.01	1.9120	3.2045	1.0751
	.05	1.3520	2.7443	1.3019
	.10	1.0845	2.5122	1.5929
	.20	0.7925	2.2353	2.3405

Let

$$A^* = \sup\{\textstyle\int_0^\infty \psi_{abc}^2\, dF : F \in \mathscr{F}_{1,\varepsilon}\} = (1-\varepsilon)\int_0^\infty \psi^2\, d\Phi + (\varepsilon a^2/2)$$

and

$$B^* = \inf\{\textstyle\int_0^\infty \psi'_{abc}\, dF : F \in \mathscr{F}_{1,\varepsilon}\} = (1-\varepsilon)\int_0^\infty \psi'_{abc}\, d\Phi - [\varepsilon a/2(c-b)]\,.$$

Assume that the parameters a, b, c and ε are such that $B^* > 0$, so that $\sup\{V(\psi, F) : F \in \mathscr{F}_{1,\varepsilon}\} \leqq A^*/[2(B^*)^2]$. To see that the supremum is equal to $A^*/[2(B^*)^2]$, note that $V(\psi, F_{(y_n),\varepsilon}) \to A^*/[2(B^*)^2]$ for any sequence $\{y_n\}$ for which $y_n \downarrow b$ as $n \to \infty$.

Given c and ε, one can obtain the optimal ψ of the form ψ_{abc} by finding the values of a and b that minimize $A^*/[2(B^*)^2]$. Table 1 presents optimal values of a and b and the corresponding minimum values of $\sup\{V(\psi_{abc}, F) : F \in \mathscr{F}_{1,\varepsilon}\}$.

EXAMPLE 2.3. *Monotone* ψ. Suppose that $\psi \in \Psi$ is monotone nondecreasing with ψ' monotone nonincreasing on $[0, \infty)$. Examples are: (i) $\psi(x) = x$; (ii) $\psi(x) = \Phi(x) - \frac{1}{2}$; (iii) Huber's estimator $\psi_k(x) = \{\min\{|x|, k\} \operatorname{sgn}(x)$, defined for $k > 0$. For such ψ, one sees immediately that $\sup\{V(\psi, F) : F \in \mathscr{F}_{1,\varepsilon}\} = V(\psi, F_{(\infty),\varepsilon})$. The same supremum is obtained if $\mathscr{F}_{1,\varepsilon}$ is replaced by a subclass $\mathscr{F}'_{1,\varepsilon}$ of distributions which are proper (i.e., $F\{(-\infty, \infty)\} = 1$) and which satisfy the required regularity conditions.

REMARK. The cases for which the values of $\sup\{V(\psi, F) : F \in \mathscr{F}\}$ are not always obvious occur when ψ is not monotone. For this reason the theory in the next section is developed for the class Ψ_c. The value of c is taken to be finite, but the results can easily be extended to the case $c = \infty$ to cover cases of nonmonotone ψ supported by the real line.

3. The general result. Let $\mathscr{F}$ be a class of df's satisfying the properties listed in Section 1. Let c be a fixed number in $(0, \infty)$, and assume that $\mathscr{F}$ satisfies:

$$F(c-0) - F(-c) > 0 \qquad \text{for all} \quad F \in \mathscr{F}\,. \tag{3.1}$$

The condition, imposed to eliminate 0/0 as a formal expression for $V(\psi, F)$, is satisfied in the practical cases where $\mathscr{F}$ is a "close neighborhood" of Φ.

Define Ψ_c' to be the subset of ψ's in Ψ_c for which the only possible discontinuities of ψ' are at $-c$ and c. The results, given here for ψ in Ψ_c', are easily modified to cover other ψ's in Ψ_c of interest (see Section 6).

For $\psi \in \Psi_c'$ and $F \in \mathscr{F}$, define

$$A_c(\psi, F) = \textstyle\int_0^c \psi^2\, dF\,,$$

and

$$B_c(\psi, F) = \textstyle\int_0^c \psi'\, dF\,,$$

with the convention that ψ' is defined at c by $\psi'(c-0)$ and that dF really means dF^*, where $F^*(0) = \frac{1}{2}$ and $F^*(x) = F(x)$ for $x > 0$. Define

$$V_c(\psi, F) = A_c(\psi, F)/[2B_c^2(\psi, F)]\,,$$

and note that $V_c(\psi, F)$ coincides with $V(\psi, F)$ unless $\psi'(c-0) \neq 0$ and $F(c) \neq F(c-0)$, so that $\sup\{V_c(\psi, F) : F \in \mathscr{F}\} = \sup\{V(\psi, F) : F \in \mathscr{F}\}$. We prefer to work with $V_c(\psi, F)$ since it is a continuous functional, while $V(\psi, F)$ need not be.

LEMMA 3.1. *Let* $\psi \in \Psi_c'$. *Then*:

(i) *There is a* F_0 *in* $\mathscr{F}$ *which attains* $\sup\{V_c(\psi, F) : F \in \mathscr{F}\}$,
(ii) $\sup\{V_c(\psi, F) : F \in \mathscr{F}\} < \infty$ *if and only if*

$$\inf\{B_c(\psi, F) : F \in \mathscr{F}\} > 0 . \tag{3.2}$$

PROOF. Suppose (3.2) holds. Then $\sup\{V_c(\psi, F) : F \in \mathscr{F}\} < \infty$, since $\sup\{A_c(\psi, F) : F \in \mathscr{F}\} \leqq (\frac{1}{2}) \sup\{\psi^2(x) : x \in [0, c]\} < \infty$. Also, if (3.2) holds, then $V_c(\psi, F)$ is a continuous function on the compact set $\mathscr{F}$, and (i) follows.

Suppose (3.2) fails. To complete the proof of (i) and (ii), it suffices to show that $B_c(\psi, F) = 0$ for some F in $\mathscr{F}$. Let $F^* \in \mathscr{F}$ satisfy $B_c(\psi, F^*) \leqq 0$. Since $\mathscr{F}$ is convex and contains Φ, $F_t = (1-t)\Phi + tF^*$ is in $\mathscr{F}$ for all $t \in [0, 1]$. Since $B_c(\psi, \Phi) = \int_0^c x\psi(x)\varphi(x)\,dx > 0$ and $B_c(\psi, F)$ is continuous, there is a $t_0 \in [0, 1)$ such that $B_c(\psi, F_{t_0}) = 0$. □

THEOREM 3.1. *Suppose* ψ *is in* Ψ_c' *and satisfies* (3.2). *Then* F_0 *maximizes* $V_c(\psi, F)$ *in* $\mathscr{F}$ *if and only if*

$$\int_0^c [2A_c(\psi, F_0)\psi'(x) - B_c(\psi, F_0)\psi^2(x)]\, d(F - F_0) \geqq 0 \tag{3.3}$$

for all $F \in \mathscr{F}$; *or equivalently, if and only if*

$$\int_0^c [2A_c(\psi, F_0)\psi'(x) - B_c(\psi, F_0)\psi^2(x)]\, dF \tag{3.4}$$

is minimized in $\mathscr{F}$ *by* F_0.

PROOF. Since $A_c(\psi, F)$ and $B_c(\psi, F)$ are linear functions of F and $A_c(\psi, F) > 0$ for all F in $\mathscr{F}$, it follows by Lemma 6 of Huber [5] that $1/V_c(\psi, F) = 2[B_c(\psi, F)]^2/A_c(\psi, F)$ is a convex function of F. Since $\mathscr{F}$ and $1/V_c(\psi, F)$ are convex, F_0 minimizes $1/V_c(\psi, F)$ iff

$$\frac{d}{dt} \frac{1}{V_c(\psi, F_t)}\bigg|_{t=0} \geqq 0 \tag{3.5}$$

for every $F \in \mathscr{F}$, where $F_t = (1-t)F_0 + tF$. A straightforward calculation shows that (3.5) is equivalent to

$$2\left(\int_0^c \psi^2\, dF_0\right)\left(-\int_0^c \psi'\, dF_0 + \int_0^c \psi'\, dF\right) - \left(\int_0^c \psi'\, dF_0\right)\left(-\int_0^c \psi^2\, dF_0 + \int_0^c \psi^2\, dF\right) \geqq 0 ,$$

which is (3.3). □

REMARK 3.1. Theorem 3.1 does not provide an explicit solution for the least favorable F_0 in terms of the given ψ. In typical applications, one guesses the form of F_0 (e.g., $F_{(x),\varepsilon}$ for some x), calculates the least favorable distribution of this form, and then checks condition (3.3).

REMARK 3.2. F_0 need not be unique. All that one can say in general about the subset of $\mathscr{F}$ on which $1/V_c(\psi, F)$ attains its minimum is that it is necessarily convex.

4. Application to the gross errors model. Let $0 < c < \infty$, $0 < \varepsilon < 1$, and $\mathscr{F} = \mathscr{F}_{1,\varepsilon}$, defined by (1.1). Then condition (3.1) is satisfied, so that Lemma 3.1 and Theorem 3.1 are true when $\mathscr{F} = \mathscr{F}_{1,\varepsilon}$. Condition (3.2) becomes

$$(4.1) \qquad (1-\varepsilon)\textstyle\int_0^c \psi'(x)\varphi(x)\,dx + \varepsilon \inf\{\psi'(x) : x \in [0, c)\} > 0 .$$

Note that since $B_c(\psi, \Phi) > 0$ and $\inf \psi'(x) < 0$, there is a "breakdown point" ε_0 in (0, 1) at which the left-hand side of (4.1) equals 0. In terms of ε_0, (4.1) can be written as $\varepsilon < \varepsilon_0$.

Theorem 3.1 can be rewritten as:

THEOREM 4.1. *Suppose $\psi \in \Psi_c'$ and (4.1) holds. Then $F_0 = (1-\varepsilon)\Phi + \varepsilon G_0$ maximizes $V_c(\psi, F)$ in $\mathscr{F}_{1,\varepsilon}$ if and only if*

$$G_0\{S\} = 1 ,$$

where S is the set of points in $[-c, c]$ at which $2A_c(\psi, F_0)\psi'(x) - B_c(\psi, F_0)\psi^2(x)$ attains its minimum.

PROOF. Since $\psi \in \Psi_c'$, ψ^2 attains a positive maximum at some $x_0 \in (0, c)$ where $\psi'(x_0) = 0$. Thus $\min\{2A_c(\psi, F_0)\psi' - B_c(\psi, F_0)\psi^2\} < 0$, so that the set S on which the minimum is attained satisfies $S \subset [-c, c]$.

Let $F = (1-\varepsilon)\Phi + \varepsilon G$ be any other member of $\mathscr{F}_{1,\varepsilon}$. Since $F - F_0 = \varepsilon(G - G_0)$, condition (3.3) becomes

$$\textstyle\int_0^c [2A_c(\psi, F_0)\psi'(x) - B_c(\psi, F_0)\psi^2(x)]\,d(G - G_0) \geqq 0 .$$

This holds for all $F \in \mathscr{F}_{1,\varepsilon}$ if and only if $G_0\{S\} = 1$. □

REMARK 4.1. An immediate consequence of Theorem 4.1 is that a necessary condition for the least favorable G_0 to have a density is that $2\psi' - k\psi^2$ be constant on the support of G_0, where $k = B_c(\psi, F_0)/A_c(\psi, F_0)$. This implies that on the support of G_0, ψ must have one of three special forms: $\psi(x) = a\tan[\frac{1}{2}ka(x-b)]$, $\psi(x) = a$, or $\psi(x) = a\tanh[\frac{1}{2}ka(b-x)]$. Note that all three forms appear in solutions to minimax problems in [2] and [5].

The simplest possible form for a least favorable F in $\mathscr{F}_{1,\varepsilon}$ is $F_{(y),\varepsilon}$, defined by (2.1).

COROLLARY 4.1. *Under the conditions of Theorem 4.1, $F_{(x_0),\varepsilon}$ maximizes $V_c(\psi, F)$ in $\mathscr{F}_{1,\varepsilon}$ if and only if x_0 is a number in $[0, c]$ which minimizes*

$$(4.2) \qquad 2A_c(\psi, F_{(x_0),\varepsilon})\psi'(x) - B_c(\psi, F_{(x_0),\varepsilon})\psi^2(x) .$$

PROOF. Immediate from Theorem 4.1. □

REMARK 4.2. To apply Corollary 4.1, one computes the x_0 in $[0, c]$ which maximizes

$$V_c(\psi, F_{(x),\varepsilon}) = \frac{(1-\varepsilon)\int_0^c \psi^2(y)\varphi(y)\,dy + \frac{1}{2}\varepsilon\psi^2(x)}{2[(1-\varepsilon)\int_0^c \psi'(y)\varphi(y)\,dy + \frac{1}{2}\varepsilon\psi'(x)]^2} ,$$

and then checks to see whether x_0 minimizes (4.2). The next theorem shows that this is always the case when $\psi \in \Psi_c'$ satisfies the following further conditions:

(4.3) $\quad \psi$ is twice differentiable and satisfies $\psi'' \leqq 0$ on $[0, c]$;

and

(4.4) $\quad \psi\psi'/\psi''$ is monotone nondecreasing on $[0, c]$.

The condition (4.3), implying that ψ is concave on $[0, c]$, is a natural requirement; however, (4.4) is a rather special condition.

THEOREM 4.2. *Suppose that* $\psi \in \Psi_c'$ *satisfies* (4.1), (4.3), *and* (4.4). *Let* $x_0 \in [0, c]$ *maximize* $V_c(\psi, F_{(x),\varepsilon})$. *Then* $F_{(x_0),\varepsilon}$ *maximizes* $V_c(\psi, F)$ *over the set of all* F *in* $\mathscr{F}_{1,\varepsilon}$.

PROOF. Note that for $x \in [0, c]$,

$$\frac{d}{dx}\left[\frac{1}{V_c(\psi, F_{(x),\varepsilon})}\right] = \frac{2B_c(\psi, F_{(x),\varepsilon})\varepsilon}{[A_c(\psi, F_{(x),\varepsilon})]^2}[A_c(\psi, F_{(x),\varepsilon})\psi''(x) - B_c(\psi, F_{(x),\varepsilon})\psi(x)\psi'(x)],$$

so that by (4.1), $d/dx[1/V_c(\psi, F_{(x),\varepsilon})]$ has the same sign as

$$A_c(\psi, F_{(x),\varepsilon})\psi''(x) - B_c(\psi, F_{(x),\varepsilon})\psi(x)\psi'(x).$$

The case $x_0 = 0$ is impossible since $\psi(0) = 0$ and $\psi'(0) \geqq 0$. In the case $x_0 \in (0, c)$, we have

$$0 = \frac{d}{dx}\left[\frac{1}{V_c(\psi, F_{(x),\varepsilon})}\right]\Bigg|_{x=x_0} = A_c(\psi, F_{(x_0),\varepsilon})\psi''(x_0) - B_c(\psi, F_{(x_0),\varepsilon})\psi(x_0)\psi'(x_0)$$
$$= \frac{d}{dx}[2A_c(\psi, F_{(x_0),\varepsilon})\psi'(x) - B_c(\psi, F_{(x_0),\varepsilon})\psi^2(x)]\Big|_{x=x_0}.$$

Conditions (4.3) and (4.4) imply that

$$A_c(\psi, F_{(x_0),\varepsilon})\psi''(x) - B_c(\psi, F_{(x_0),\varepsilon})\psi(x)\psi'(x) \begin{cases} \leqq 0 & 0 \leqq x \leqq x_0 \\ \geqq 0 & x_0 \leqq x \leqq c, \end{cases}$$

so that x_0 minimizes (4.2) and the conclusion follows from Corollary 4.1.

Similarly, in the case $x_0 = c$, one finds that $A_c(\psi, F_{(c),\varepsilon})\psi''(x) - B_c(\psi, F_{(c),\varepsilon})\psi(x)\psi'(x) \leqq 0$ for all $x \in [0, c]$, so that $x = c$ minimizes (4.2). □

EXAMPLE. Consider the following ψ proposed by Andrews [1]:

$$\psi_s(x) = \begin{cases} \sin(\pi x/c) & |x| \leqq c \\ 0 & |x| > c. \end{cases}$$

Since ψ_s satisfies (4.3) and (4.4), Theorem 4.2 applies and the result can be described as follows:

Define ε_1 and ε_2 by

$$\varepsilon_1/(1-\varepsilon_1) = \max\{0, 2[(c/\pi)B_c(\psi_s, \Phi) - A_c(\psi_s, \Phi)]\}$$

and

$$\varepsilon_2/(1-\varepsilon_2) = 2(c/\pi)B_c(\psi_s, \Phi) .$$

If $0 < \varepsilon \leqq \varepsilon_1$, then $F_{(x_0),\varepsilon}$ maximizes $V_c(\psi_s, F)$ in $\mathscr{F}_{1,\varepsilon}$, where

$$x_0 = \frac{c}{\pi}\cos^{-1}\left[\frac{2\pi(1-\varepsilon)A_c(\psi_s, \Phi) + \varepsilon\pi}{2(1-\varepsilon)cB_c(\psi_s, \Phi)}\right].$$

If $\varepsilon_1 < \varepsilon < \varepsilon_2$, then $F_{(c),\varepsilon}$ maximizes $V_c(\psi_s, F)$ in $\mathscr{F}_{1,\varepsilon}$. If $\varepsilon_2 \leqq \varepsilon < 1$, then $\sup\{V_c(\psi_s, F): F \in \mathscr{F}_{1,\varepsilon}\} = \infty$. For various values of c and ε, Table 2 presents the value of $\sup\{V_c(\psi_s, F): F \in \mathscr{F}_{1,\varepsilon}\}$ and the corresponding value of x_0 for which $F_{(x_0),\varepsilon}$ attains the supremum. Also tabulated are the values of ε_1 and ε_2 corresponding to each c.

TABLE 2
*Least favorable ε-contamination and the corresponding asymptotic variances when $\psi(x) = \sin(\pi x/c)$ for $|x| \leqq c$, $\psi(x) = 0$ otherwise**

c	ε_1	ε_2	ε = .001	.01	.05	.10	.20
2.0	0	.2454	1.8620 (2.0)	1.9890 (2.0)	2.7700 (2.0)	4.7392 (2.0)	43.196 (2.0)
3.0	.1201	.3673	1.2070 (2.3340)	1.2584 (2.3576)	1.5337 (2.4801)	2.0432 (2.7163)	4.6344 (3.0)
4.0	.2755	.4235	1.0692 (2.6429)	1.1134 (2.6610)	1.3401 (2.7476)	1.7182 (2.8737)	3.1514 (3.2301)
6.0	.3979	.4658	1.0187 (3.4692)	1.0749 (3.4897)	1.3565 (3.5867)	1.8023 (3.7228)	3.2593 (4.0641)
8.0	.4423	.4807	1.0127 (4.3690)	1.0931 (4.3944)	1.4930 (4.5132)	2.1165 (4.6786)	4.0740 (5.0852)
10.0	.4630	.4877	1.0139 (5.3033)	1.1271 (5.3339)	1.6879 (5.4773)	2.5565 (5.6762)	5.2370 (6.1611)

* The top entry for each case is $\sup\{V(\psi, F): F \in \mathscr{F}_{1,\varepsilon}\}$ and the bottom entry (in parenthéses) is the value of x_0 for which $V_c(\psi, F_{(x_0),\varepsilon})$ attains the supremum.

For various values of c and ε, Table 3 compares $\sup\{V_c(\psi, F): F \in \mathscr{F}_{1,\varepsilon}\}$ for the following ψ's: (i) the minimax ψ_0 (Example 2.1); (ii) the optimal Hampel ψ_{abc} (Example 2.2); and (iii) ψ_s. Notice that (1) there is not much loss of efficiency in using the optimal Hampel instead of ψ_0, and (2) in the range $c = 2$ to 4 and $\varepsilon = .05$ to .10, the maximum asymptotic variance for ψ_s is smaller than that of the optimal Hampel.

To conclude this section, a discussion of regularity conditions follows. Let $\mathscr{F}'_{1,\varepsilon}$ be the class of distributions of the form $F = (1-\varepsilon)\Phi + \varepsilon G$, where G is continuous and symmetric. For $\psi \in \Psi_c$, define the estimator as the Newton's method solution of $\sum \psi[(X_i - \theta)/\sigma] = 0$ with the sample median as starting value. Denote by C the condition that the estimator defined by ψ is consistent and asymptotically normally distributed with asymptotic variance $V(\psi, F)$. A

TABLE 3

Comparison of $\sup\{V(\psi, F) : F \in \mathscr{F}_{1,\varepsilon}\}$ *for three* ψ*'s which vanish off the set* $[-c, c]$

c	ψ	ε = 0	.001	.01	.05	.10	.20
2	ψ_0	1.3540	1.4503	1.7273	2.6401	4.1291	11.2118
	ψ_{abc}	1.3540	1.4560	1.7774	2.9224	4.9149	15.2974
	ψ_s	1.8467	1.8620	1.9890	2.7700	4.7392	43.1958
3	ψ_0	1.0302	1.0622	1.1663	1.5026	1.9633	3.3772
	ψ_{abc}	1.0302	1.0677	1.2033	1.6568	2.3024	4.4148
	ψ_s	1.2015	1.2070	1.2584	1.5337	2.0432	4.6344
4	ψ_0	1.0011	1.0151	1.0817	1.3143	1.6212	2.4632
	ψ_{abc}	1.0011	1.0205	1.1123	1.4316	1.8645	3.1191
	ψ_s	1.0645	1.0692	1.1134	1.3401	1.7182	3.1514
6	ψ_0	1.0000+	1.0096	1.0656	1.2600	1.5038	2.1145
	ψ_{abc}	1.0000+	1.0121	1.0820	1.3303	1.6547	2.5174
	ψ_s	1.0126	1.0187	1.0749	1.3565	1.8023	3.2593
8	ψ_0	1.0000+	1.0096	1.0653	1.2564	1.4914	2.0584
	ψ_{abc}	1.0000+	1.0110	1.0751	1.3019	1.5929	2.3405
	ψ_s	1.0040	1.0127	1.0931	1.4930	2.1165	4.0740
10	ψ_0	1.0000+	1.0096	1.0652	1.2561	1.4900	2.0479
	ψ_{abc}	1.0000+	1.0105	1.0722	1.2890	1.5643	2.2584
	ψ_s	1.0016	1.0139	1.1271	1.6879	2.5565	5.2375

proof is given in [2] that C holds in the special case where $G = \Phi$ and ψ is smooth, and the proof is extended in [3] to show that for any $\psi \in \Psi_c$, C holds uniformly for all F in $\mathscr{F}'_{1,\varepsilon}$ if ε is sufficiently small. Since $\mathscr{F}'_{1,\varepsilon}$ is dense in $\mathscr{F}_{1,\varepsilon}$, it follows that $\sup\{V(\psi, F) : F \in \mathscr{F}'_{1,\varepsilon}\} = \sup\{V_c(\psi, F) : F \in \mathscr{F}_{1,\varepsilon}\}$.

5. Application to the Kolmogorov model. Let $0 < c < \infty$, $0 < \varepsilon < 1$, and $\mathscr{F} = \mathscr{F}_{2,\varepsilon}$, defined by (1.2). We remark that, for sufficiently small ε, the ψ in Ψ_c that minimizes $\sup\{V(\psi, F) : F \in \mathscr{F}_{2,\varepsilon}\}$ has the form

$$\begin{aligned} \psi(x) &= k \tan(\tfrac{1}{2}kx) && 0 \leqq |x| \leqq a \\ &= x && a \leqq |x| \leqq b \\ &= m \tanh[\tfrac{1}{2}m(c - |x|)] \operatorname{sgn}(x) && b \leqq |x| \leqq c \\ &= 0 && |x| \geqq c\,. \end{aligned}$$

The proof, obtained by combining portions of proofs appearing in [2] and [5], is omitted.

Assume throughout this section that ε and c satisfy

$$\varepsilon < \tfrac{1}{2}[\Phi(c) - \tfrac{1}{2}]\,. \tag{5.1}$$

Then $\mathscr{F}_{2,\varepsilon}$ satisfies condition (3.1), so that Lemma 3.1 and Theorem 3.1 apply. As in Section 4, Theorem 3.1 will be specialized to ψ in Ψ_c' satisfying (4.3) and (4.4), in which case $\sup\{V(\psi, F) : F \in \mathscr{F}_{2,\varepsilon}\}$ can be determined explicitly.

Define $x_1 = \Phi^{-1}(\frac{1}{2} + \varepsilon)$ and $x_2 = \min\{\Phi^{-1}(1 - \varepsilon), c\}$, and note that $0 < x_1 < x_2 \leqq c$, by (5.1). Adopting the convention that each F in $\mathscr{F}_{2,\varepsilon}$ is normalized by setting $F(0) = \frac{1}{2}$, it is clear that a symmetric df F is in $\mathscr{F}_{2,\varepsilon}$ if and only if $F_*(x) \leqq F(x) \leqq F^*(x)$ for all $x \geqq 0$, where

$$\begin{aligned} F_*(x) &= \tfrac{1}{2} & 0 \leqq x < x_1 \\ &= \Phi(x) - \varepsilon & x \geqq x_1 \end{aligned}$$

and

$$\begin{aligned} F^*(x) &= \tfrac{1}{2} & x = 0 \\ &= \Phi(x) + \varepsilon & 0 < x < x_2 \\ &= 1 & x \geqq x_2 \,. \end{aligned}$$

For $y \geqq 0$, define $F_{\{y\}} \in \mathscr{F}_{2,\varepsilon}$ by

$$\begin{aligned} F_{\{y\}}(x) &= F_*(x) & 0 \leqq x < y \\ &= F^*(x) & x \geqq y \,. \end{aligned}$$

LEMMA 5.1. *Let $x_0 \in [0, c]$, and let g be a continuously differentiable function on $[0, c]$ that satisfies $g(c) \leqq 0$ and*

$$\begin{aligned} g'(x) &\leqq 0 & x \in [0, x_0] \\ &\geqq 0 & x \in [x_0, c] \,. \end{aligned}$$

Then $\int_0^c g\, dF$ is minimized in $\mathscr{F}_{2,\varepsilon}$ by $F_{\{x_0\}}$.

PROOF. Let $F \in \mathscr{F}_{2,\varepsilon}$. Then

$$\begin{aligned} &\int_0^c g\, d(F - F_{\{x_0\}}) \\ &\quad = g(c)[F(c) - F^*(c)] - g(0)[F(0) - F_*(0)] - \int_0^c [F(x) - F_{\{x_0\}}(x)]g'(x)\, dx \\ &\quad \geqq -\int_0^{x_0} [F(x) - F_*(x)]g'(x)\, dx - \int_{x_0}^c [F(x) - F^*(x)]g'(x)\, dx \\ &\quad \geqq 0 \,. \end{aligned}$$

□

Suppose that $\psi \in \Psi_c'$ satisfies (4.3). Then since $\psi'(c - 0) \leqq 0$, we can apply Lemma 5.1 with $g = \psi'$ and $x_0 = c$ to obtain that $F_{\{c\}}$ minimizes $\int_0^c \psi'\, dF$ in $\mathscr{F}_{2,\varepsilon}$. Condition (3.2) of Lemma 3.1 can be written as

$$\int_{x_1}^c \psi'(x)\varphi(x)\, dx + \psi'(c - 0)\min\{2\varepsilon, 1 - \Phi(c) + \varepsilon\} > 0 \,. \tag{5.2}$$

The analogue of Theorem 4.2 for the class $\mathscr{F}_{2,\varepsilon}$ is:

THEOREM 5.1. *Suppose that $\psi \in \Psi_c'$ satisfies* (5.2), (4.3), *and* (4.4). *If x_0 maximizes $V_c(\psi, F_{\{x\}})$ over all x in $[0, c]$, then $F_{\{x_0\}}$ maximizes $V_c(\psi, F)$ over all F in $\mathscr{F}_{2,\varepsilon}$.*

PROOF. On the set $[0, c]$ we have

$$\frac{d}{dx}\left[\frac{1}{V_c(\psi, F_{\{x\}})}\right] = \frac{4B_c(\psi, F_{\{x\}})H(x)}{[A_c(\psi, F_{\{x\}})]^2}[A_c(\psi, F_{\{x\}})\psi''(x) - B_c(\psi, F_{\{x\}})\psi(x)\psi'(x)] \,,$$

where

$$\begin{aligned} H(x) &= \Phi(x) - \tfrac{1}{2} + \varepsilon & x \in [0, x_1] \\ &= 2\varepsilon & x \in [x_1, x_2] \\ &= 1 - \Phi(x) + \varepsilon & x \in [x_2, c] \,. \end{aligned}$$

The same argument used in the proof of Theorem 4.2 yields

$$\begin{aligned} g'(x) &\leqq 0 \quad 0 \leqq x \leqq x_0 \\ &\geqq 0 \quad x_0 \leqq x \leqq c\,, \end{aligned}$$

where $g(x) = 2A_c(\psi, F_{\{x_0\}})\psi'(x) - B_c(\psi, F_{\{x_0\}})\psi^2(x)$. Note that $g(c) \leqq 0$.

By Lemma 5.1, $F_{\{x_0\}}$ minimizes

$$\int_0^c [2A_c(\psi, F_{\{x_0\}})\psi' - B_c(\psi, F_{\{x_0\}})\psi^2]\, dF$$

in $\mathscr{F}_{2,\varepsilon}$. The conclusion follows by Theorem 3.1. □

6. Concluding remarks. Let $0 < b < c < \infty$ and define Ψ_{bc} to be the class of ψ's in Ψ_c for which $\psi'(x-0) + \psi'(x+0) \geqq 0$ for all $x \in [0, b)$, and $\psi'(x)$ exists and is < 0 for all $x \in (b, c)$. Note that ψ_0 and ψ_{abc} (and, one can argue, all ψ's in Ψ_c of practical interest) have this form. If $\psi \in \Psi_{bc}$, then to find $\sup\{V(\psi, F) : F \in \mathscr{F}_{1,\varepsilon}\}$ one need only consider $F = (1-\varepsilon)\Phi + \varepsilon G$ for the class of symmetric G satisfying $G\{[b, c]\} = \frac{1}{2}$. For such F, one can replace $B(\psi, F)$ by a continuous functional in the obvious way to obtain theorems for the class Ψ_{bc} analogous to those obtained for Ψ_c'.

Another generalization is the following: let Ψ_∞ denote the class of ψ's in Ψ for which ψ' is continuous everywhere and both $\lim_{x\to\infty}\psi(x)$ and $\lim_{x\to\infty}\psi'(x)$ exist and are finite. Then, since ψ and ψ' are continuous on the compact set $[-\infty, \infty]$, Lemma 3.1 and Theorem 3.1 hold for the class Ψ_∞.

When the scale parameter σ in the model is unknown, one proposal for defining the M-estimator of θ is to solve

$$\sum_{i=1}^n \psi\left(\frac{X_i - \hat{\theta}_n}{\hat{\sigma}_n}\right) = 0$$

for $\hat{\theta}_n$, where $\hat{\sigma}_n = \hat{\sigma}_n(X_1, \cdots, X_n)$ is an estimator of σ that is unbiased when $F = \Phi$. Under regularity conditions, $\hat{\theta}_n$ is consistent and $n^{\frac{1}{2}}(\theta_n - \theta)$ is asymptotically normal with mean 0 and a variance which we choose to write in the form $\sigma^2\, \beta(F)V(\psi, F)$. If $\mathscr{F}$ is the gross errors or Kolmogorov model with ε "small," then $\beta(F)$ is "close to" 1 for all F in $\mathscr{F}$ (see Section 4 of [2] for a specific example of this). In such cases $\sup\{V(\psi, F) : F \in \mathscr{F}\}$ is a reasonably good measure of robustness in the scale unknown case.

Note that all results stated for neighborhoods of Φ essentially go through if Φ is replaced by any symmetric distribution H for which $\int \psi'\, dH > 0$.

Acknowledgment. The author wishes to thank the referees for their helpful comments.

REFERENCES

[1] ANDREWS, D. F., BICKEL, P. J., HAMPEL, F. R., HUBER, P. J., ROGERS, W. H. and TUKEY, J. W. (1972). *Robust Estimates of Location.* Princeton Univ. Press.

[2] COLLINS, J. R. (1976). Robust estimation of a location parameter in the presence of asymmetry. *Ann. Statist.* **4** 68–85.

[3] COLLINS, J. R. (1976). On the consistency of M-estimators. Purdue Univ. Dept. of Statistics Mimeograph Series No. 450.

[4] HAMPEL, F. R. (1974). The influence curve and its role in robust estimation. *J. Amer. Statist. Assoc.* **69** 383–393.

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

[6] HUBER, P. J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 221–233, Univ. of California Press.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF CALGARY
2920 24 AVENUE N. W.
CALGARY, ALBERTA, CANADA T2N 1N4