The Annals of Statistics
1988, Vol. 16, No. 2, 609–618

ROBUSTNESS OF ESTIMATORS FOR DIRECTIONAL DATA

BY DAIJIN KO AND PETER GUTTORP

Virginia Commonwealth University and University of Washington

Some standard robustness concepts, developed for linear data, are found wanting in the case of directional data. We introduce a standardized bias robustness allowing for uniform robustness considerations of statistics for bounded parameter spaces. Specifically, we verify the nonrobustness in this sense of the directional mean, the directional dispersion and the maximum likelihood estimator of the concentration parameter of the von Mises–Fisher distributions.

1. Introduction. Despite much recent research on the robustness of estimators of location and scale for linear data, the robustness of estimators for directional data has not been investigated comprehensively. Mardia (1975) and discussant mentioned some aspects of robustness. Collett (1980), Fisher, Lewis and Willcox (1981) and Kimber (1985) studied outliers in directional data, Lenth (1981) used the M-estimator to robustify the circular mean, Barnett and Lewis (1984) reviewed their work and Fisher (1985) studied spherical medians. Wehrly and Shine (1981) and Watson (1986) evaluated the robustness of the directional mean, which is the maximum likelihood estimator of the location of the von Mises–Fisher distribution, via an influence function introduced by Hampel (1968, 1974) and concluded that the estimator is robust since the influence function is bounded.

In this paper we study measures of robustness using the influence function and introduce the concept of standardized bias robustness. We derive the relationship between unstandardized bias robustness and standardized bias robustness and compute the influence functions of the directional mean and dispersion and the maximum likelihood estimator of the concentration parameter of the von Mises–Fisher distribution, connecting some earlier computations in the literature. We show that these statistics are not robust in the standardized bias robustness sense.

2. Robustness of estimators. The influence function of a functional T at the underlying probability distribution F is defined as [Hampel, Ronchetti, Rousseeuw and Stahel (1986)]

$$\mathrm{IF}(x; T, F) = \lim_{s\to 0} \frac{1}{s}\left[T\{(1-s)F + s\delta_x\} - T(F)\right],$$

where δ_x denotes the point mass at x. The gross error sensitivity of the estimator

Received November 1986; revised August 1987.

AMS 1980 *subject classification*. 62F35.

Key words and phrases. Directional data, influence function, robustness, standardized bias robustness, standardized gross error sensitivity.

T at F is

$$\gamma(T, F) = \sup \|\mathrm{IF}(x; T, F)\|,$$

where $\|\cdot\|$ denotes the Euclidean norm and the supremum is taken over all x, where $\mathrm{IF}(x; T, F)$ exists. It measures approximately the largest influence that a small amount of contamination of fixed size can have on the value of estimator. Hence γs may be regarded as an approximate upper bound to the asymptotic bias of the estimator, where s denotes the amount of contamination. It is desirable that $\gamma(T, F)$ be finite, in which case we say that T is bias-robust or B-robust at F [Rousseeuw (1981)].

Often the gross error sensitivity is bounded if the parameter space is bounded. It is not surprising because commonly the gross error sensitivity approximates maximum bias, which is bounded on a bounded parameter space. We therefore need to modify the concept of B-robustness when we deal with bounded parameter spaces. A modification of the influence curve on a bounded parameter space can be found in Lambert (1981). She defined influence functions for testing by applying Hampel's influence function to the logarithm of P-values.

In the case of location, it is natural to consider the maximum bias or the gross error sensitivity relative to a measure of dispersion of the underlying distribution. Measures of dispersion of a distribution on the real line are defined in Bickel and Lehmann (1976) and can easily be extended to spherical distributions. For linear distributions, it is often the case that $\sqrt{n}\,[T[F_n] - T(F)]$ converges to $N[0, S^2(F)]$, where $T[F_n]$ is a location estimator and $S(F)$ is a measure of dispersion of the distribution F. In this case, the gross error sensitivity relative to $S(F)$ measures the maximum bias relative to the asymptotic accuracy of $T[F_n]$.

Definition 1. Let T be a functional. The standardized influence function of T with respect to a functional S is defined by

$$\mathrm{SIF}(x; T, F, S) = \frac{1}{S(F)}\mathrm{IF}(x; T, F),$$

for F with $S(F) \neq 0$.

The SIF of T with respect to S measures the influence of x in units of the functional S. In particular, for a location functional T and a measure of dispersion S, SIF measures relative influence with respect to the dispersion. For a dispersion functional S, SIF of S with respect to S itself is the influence function of the log-transformed functional $\log S$ at F with $S(F) > 0$, which is analogous to Lambert's (1981) approach.

Definition 2. The standardized gross error sensitivity of T with respect to the functional S at a family $\mathscr{F}$ of distributions is defined by

$$\begin{aligned}\gamma^*(T, \mathscr{F}, S) &= \sup_{\mathscr{F}} \gamma(T, F)/S(F)\\ &= \sup_{\mathscr{F}} \sup_{x} \mathrm{SIF}(x; T, F, S).\end{aligned}$$

DEFINITION 3. T is called standardized bias robust at $\mathscr{F}$ with respect to S, or SB-robust at $\mathscr{F}$, if it has a finite standardized gross error sensitivity at $\mathscr{F}$.

$\gamma^*(T, \mathscr{F}, S)$ measures the maximum asymptotic bias in the unit of the functional S of the distribution. If $\mathscr{F}$ is a singleton set, SB-robustness at $\mathscr{F}$ coincides with B-robustness of Rousseeuw (1981) provided that the functional S is nonzero and finite. However, B-robustness at F for every F in $\mathscr{F}$ does not imply SB-robustness at $\mathscr{F}$. If T is SB-robust at $\mathscr{F}$ with respect to S, then T is SB-robust at $\mathscr{G}$ for any nonempty sub-family $\mathscr{G}$ of $\mathscr{F}$.

THEOREM 1. *Let T be a B-robust scale equivariant estimator at F_1. Then T is* SB-*robust at* $\{F_\sigma | \sigma > 0\}$ *with respect to S such that* $S(F_\sigma) = \sigma$, *where* $F_\sigma(x) = F_1(x/\sigma)$.

PROOF. Let X have distribution $(1 - s)F_1 + s\delta_z$. Then σX has distribution $(1 - s)F_\sigma + s\delta_{\sigma z}$ and $T\{(1 - s)F_\sigma + s\delta_{\sigma z}\} = \sigma T\{(1 - s)F_1 + s\delta_{\sigma z}\}$. Hence

$$\begin{aligned}\mathrm{IF}(z; T, F_\sigma) &= \lim_{s\to 0} \frac{1}{s}\left[T\{(1 - s)F_\sigma + s\delta_{\sigma z}\} - T(F_\sigma)\right]\\ &= \lim_{s\to 0} \frac{1}{s}\left[\sigma T\{(1 - s)F_1 + s\delta_{\sigma z}\} - \sigma T(F_1)\right]\\ &= \sigma \mathrm{IF}(\sigma z; T, F_1)\end{aligned}$$

and

$$\frac{\gamma(T, F_\sigma)}{S(F_\sigma)} = \gamma(T, F_1). \qquad \square$$

The theorem indicates that for the problem of estimating location of linear data, it is sufficient to consider B-robustness at a single distribution provided that we only consider scale-equivariant estimators. This seems to be a reason why Rousseeuw's definition has been used and that efforts have been concentrated on bounding the gross error sensitivity γ at a single distribution, such as the standard normal distribution, instead of the standardized gross error sensitivity at a family of distributions.

It is a different story if the sample space is not a Euclidean space. Then scale equivariance may not be a natural thing to require for location estimators. Many families of distributions are not even closed under scale transformation. Actually the term "scale" may not mean anything; for example, the family of von Mises–Fisher distributions on the unit sphere or circle has no natural scale.

3. Robustness of estimators for directional data.

3.1. *Location.* Dispersion for directional data should be interpreted a little differently from dispersion for linear data. On the unit sphere or unit circle there is no notion of scale transformation or scale equivariance, because the unit length is determined as the unit distance on the sample space. The most frequently used family of distributions for directional data, the family of von Mises–Fisher distributions, is not a location-scale family.

We can define a measure of dispersion of a distribution on Ω_q, the unit sphere in the q-dimensional Euclidean space $\mathbb{R}^q$ as a functional S satisfying the following conditions. Let X and Y be random unit vectors with unimodal distributions F and G with modal vector $T(X)$ and $T(Y)$, respectively. Then a real-valued functional S is called a dispersion on Ω_q if

(1) $S(F) \leq S(G)$ whenever $d(Y, T(Y))$ is stochastically larger than $d(X, T(X))$, where $d(\cdot, \cdot)$ is a metric on Ω_q;

(2) $$S(F) = S(G), \quad \text{if } Y = \Gamma X \text{ for an orthogonal matrix } \Gamma;$$

(3) $$S(\delta_c) = 0, \quad \text{if } c \text{ is a fixed point on } \Omega_q.$$

In particular, directional dispersion S defined by $S^2(F) = 1 - \|C(F)\|$, where $\|C(F)\|$ is the Euclidean norm of $C(F) = [\int x\,dF_1, \dots, \int x\,dF_q]^T$ and $F_1, \dots, F_q$ are marginal distributions on the coordinates, is a dispersion on Ω_q in this sense. By defining the directional mean $T(F)$ of F as $C(F)/\|C(F)\|$ for $\|C(F)\| \neq 0$, we have

$$S^2(F) = 1 - \|C(F)\| = \int [1 - X^T T(F)]\, dF = \frac{1}{2} \int d^2(X, T(F))\, dF,$$

where $d(\cdot, \cdot)$ is the Euclidean metric on $\mathbb{R}^q$. The sample directional mean and dispersion can be obtained by evaluating T at the empirical measure given by the data. The sample directional mean is the maximum likelihood estimator of the location parameter of the von Mises–Fisher family; see Wehrly and Shine (1981) and Watson (1986).

Theorem 2. *The influence function of the directional mean T at F with $C(F) \neq 0$ is*

$$\mathrm{IF}(x; T, F) = [x\|C(F)\|^2 - C(F)\{x^T C(F)\}]/\|C(F)\|^3,$$

for $x \in \Omega_k$. The norm of the influence function is

$$\|\mathrm{IF}(x; T, F)\| = \{1 - |x^T T(F)|^2\}^{1/2}/\|C(F)\|.$$

Proof. Since $C\{(1 - s)F + s\delta_x\} = C(F) + s\{x - C(F)\}$,

$$\begin{aligned}\mathrm{IF}(x; T, F) &= \lim_{s \to 0} [T\{(1 - s)F + s\delta_x\} - T(F)]/s \\ &= d/ds|_{s=0} T\{(1 - s)F + s\delta_x\} \\ &= d/ds|_{s=0} [C(F) + s\{x - C(F)\}]/\|C(F) + s\{x - C(F)\}\| \\ &= [x\|C(F)\|^2 - C(F)\{x^T C(F)\}]/\|C(F)\|^3.\end{aligned}$$

Using the fact that $x^Tx = 1$ and $C(F) = \|C(F)\|T(F)$, we have

$$\|\mathrm{IF}(x; T, F)\|^2 = \left[x^Tx\|C(F)\|^4 - 2x^TC(F)\{x^TC(F)\}\|C(F)\|^2 + C(F)^TC(F)\{x^TC(F)\}^2\right]/\|C(F)\|^6$$
$$= \left[1 - \{x^TT(F)\}^2\right]/\|C(F)\|^2. \qquad \square$$

In particular, when $q = 2$, $T(F) = (\cos_0\theta, \sin_0\theta)^T$, $x = (\cos\theta, \sin\theta)^T$ and $\rho = \|C(F)\|$, the norm of the influence function is $|\sin(\theta - \theta_0)|/\rho$. Let the directional dispersion be defined by $S(F) = (1 - \|C(F)\|)^{1/2}$. Then the norm of the standardized influence function of T at F with respect to the directional dispersion S is

$$\{1 - \langle x^TT(F)\rangle^2\}^{1/2}\|C(F)\|^{-1}[1 - \|C(F)\|]^{-1/2}$$
$$= \left[-\{d^2(x, T(F)) - 2\}^2/4 + 1\right]\|C(F)\|^{-1}[1 - \|C(F)\|]^{-1/2},$$

since $x^TT(F) = 1 - d^2(x, T(F))/2$ (see Figure 1). We note that the most

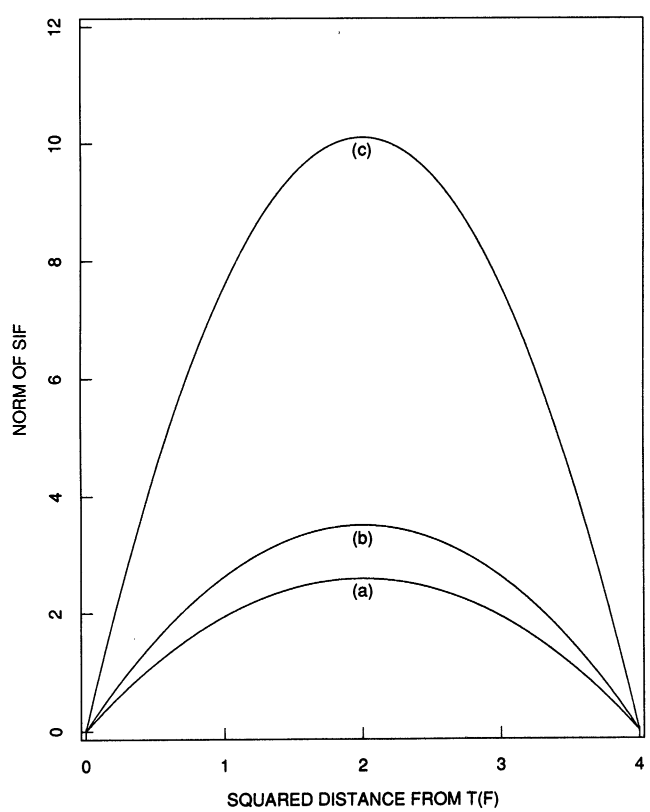

FIG. 1. *Norm of standardized influence function of directional mean (a) when* $\|C(F)\| = 2/3$; *(b) when* $\|C(F)\| = 9/10$; *and (c) when* $\|C(F)\| = 99/100$.

influential point on Ω_q is the point orthogonal to the directional mean; i.e., the point x with $x^T T(F) = 0$ or, equivalently, $d^2(x, T(F)) = 2$. The point farthest from the directional mean has little influence. We also note that the norm of the standardized influence function is bounded from below by $\|\mathrm{SIF}(x; T, F_0, S)\|$ with $\|C(F_0)\| = 2/3$. If $\mathscr{F}$ is a family of distributions such that $\|C(F)\| \neq 0$ for all $F \in \mathscr{F}$ and $\sup\|C(F)\| = 1$ or $\inf\|C(F)\| = 0$, where the supremum and the infimum are taken over $\mathscr{F}$, then the standardized gross error sensitivity $\sup\|C(F)\|^{-1}(1 - \|C(F)\|)^{-1/2}$ is infinite and T is not SB-robust at the family $\mathscr{F}$. In other words, the asymptotic bias of T of a small contamination could be very large compared to the dispersion.

On Ω_q, we can only expect to encounter outliers in samples when the main mass of the data is sufficiently concentrated about a particular point. In the case of a sample having a low concentration, that is, a small $\|C(F)\|$, the standardized gross error sensitivity would be very large. It would be, however, difficult to find a datum that is sufficiently separated from others to provide evidence of being an outlier. We, therefore, concentrate on the family $\mathscr{F}$ such that $\inf\|C(F)\| > 0$ and $\sup\|C(F)\| = 1$, where the supremum and the infimum are taken over $\mathscr{F}$.

For the directional mean, the maximum bias of T at F for s-contamination is defined by $b(T, F) = \sup d[T\{(1 - s)F + sH\}, T(F)]$, where the supremum is taken over the family of distributions on Ω_q such that $\|C\{(1 - s)F + sH\}\| \neq 0$ and d is a metric on Ω_q. For the Riemannian metric $d(x, y)$ on Ω_q (the length of the shortest arc between x and y), we can calculate the maximum bias as follows.

Theorem 3. *The maximum bias of T at F of s-contamination is*

$$b(T, F; s) = \begin{cases} \arcsin[\{s/(1 - s)\}/\|C(F)\|], & \text{if } (1 - s)\|C(F)\| \geq s, \\ \pi, & \text{otherwise}. \end{cases}$$

Proof.

$$\begin{aligned} &d[T\{(1 - s)F + sH\}, T(F)] \\ &\quad = d[\{(1 - s)C(F) + SC(H)\}/\|(1 - s)C(F) + sC(H)\|, C(F)/\|C(F)\|]. \end{aligned}$$

Since $\|C(F)\| \leq 1$ for $s \leq (1 - s)\|C(F)\|$ the distance given previously is maximized when $sC(H)$ is orthogonal to $(1 - s)C(F) + sC(H)$. Then the maximized distance is $\arcsin[\{s/(1 - s)\}/\|C(F)\|]$. □

If $C(F)$ is close to 1 and s is small, the maximum bias can be approximated by $s\|C(F)\|^{-1}$. We notice the same result of nonrobustness by observing the standardized maximum bias defined by $\sup b(T, F; s)/S(F)$, where supremum is taken over a family $\mathscr{F}$ of distributions. Using $S(F) = \{1 - \|C(F)\|\}^{1/2}$, for example, for $\mathscr{F}$ with $\sup\|C(F)\| = 1$, the standardized maximum bias at $\mathscr{F}$ is infinity, so the bias relative to the dispersion $\{1 - \|C(F)\|\}^{1/2}$ can be very large even for a small contamination. This result accords with intuition. For highly concentrated distributions on Ω_q, we are effectively dealing with a distribution on the $(q - 1)$-dimensional hyperplane, so we expect to see properties similar to those of sample mean of data in $\mathbb{R}^{q-1}$. For example, the von Mises distribution $M(\mu_0, \kappa)$ on Ω_2 with the directional mean vector $(\cos\mu_0, \sin\mu_0)$, whose density is given by $g(\theta; \mu_0, \kappa) = [2\pi I_0(\kappa)]^{-1}\exp[\kappa\cos(\theta - \mu_0)]$, where $I_0(\kappa)$ is the modified Bessel function of the first kind and order zero, can be approximated by the normal distribution $N(\mu_0, \kappa^{-1})$ on the real line for large κ [Mardia (1972)].

Wehrly and Shine (1981) derived the influence function for the circular median [Mardia (1972), page 28] at a symmetric unimodal circular distribution F with the modal angle μ_0. The influence function is given by

$$\mathrm{IF}(\theta; C\text{-Median}, F) = \tfrac{1}{2}\,\mathrm{sgn}[\theta - \mu_0]/\{f(\mu_0) - f(\mu_0 + \pi)\},$$

for $\mu_0 - \pi < \theta < \mu_0 + \pi$, where $\mathrm{sgn}(x) = 1$, 0 or -1 as $x > 0$, $x = 0$ or $x < 0$ and f is the density corresponding to F. Let $S(F) = K(F)^{-1/2}$, where $K(F) = A_2^{-1}(\|C(F)\|)$, A_2^{-1} is the inverse function of $A_2(\cdot) = I_1(\cdot)/I_0(\cdot)$ and $I_p(\cdot)$ is the modified Bessel function of the first kind and order p. S is a measure of dispersion corresponding to the scale parameter of the normal distribution that approximates a von Mises distribution with large concentration parameter. The standardized gross error sensitivity with respect to S at the von Mises distribution $M(\mu_0, \kappa)$ is $\pi\kappa^{1/2}I_0(\kappa)e^{-\kappa}(1 - e^{-2\kappa})^{-1}$. When κ is large $I_0(\kappa) = C\kappa^{-1/2}e^{\kappa}(1 + O(\kappa^{-1}))$, where $0.39 < C < 0.40$. Hence the standardized gross error sensitivity at $\mathscr{F} = \{M(\mu_0, \kappa) | \kappa \geq m > 0\}$ is bounded. This is an example of an SB-robust estimator. Lenth (1981) used the circular median as the starting point of the M-estimator. In contradistinction to the linear case, it is not known that the circular median is the most SB-robust statistics. Another robust estimator is the least median of squares estimator given by Rousseeuw (1984).

3.2. *Dispersion.*

THEOREM 4. *The influence function of the directional dispersion* $S = (1 - \|C(\cdot)\|)^{1/2}$ *is* $-\{x^T T(F) - \|C(F)\|\}(1 - \|C(F)\|)^{-1/2}/2$.

PROOF.

$$\begin{aligned}\mathrm{IF}(x; \|C(\cdot)\|, F) &= \lim_{s\to 0}\left[\|C\{(1-s)F + s\delta_x\}\| - \|C(F)\|\right]/s\\ &= d/ds|_{s=0}\|C(F) + s(x - C(F))\|\\ &= [x - C(F)]^T C(F)/\|C(F)\|\\ &= x^T T(F) - \|C(F)\|.\end{aligned}$$

The proof is completed by using the chain rule. □

The standardized influence function of the directional dispersion S with respect to S itself at $\{F\}$ is $-\{x^T T(F) - \|C(F)\|\}(1 - \|C(F)\|)^{-1}/2$, which is $4^{-1}(1 - \|C(F)\|)^{-1} d^2(x, T(F)) - 1/2$, a linear function of $d^2(x, T(F))$, where d is the Euclidean metric on $\mathbb{R}^q$ (see Figure 2). We note that the SIF is bounded from below by $d^2(x, T(F))/4 - 1/2$. The most influential point is opposite to the directional mean. The gross error sensitivity at F is

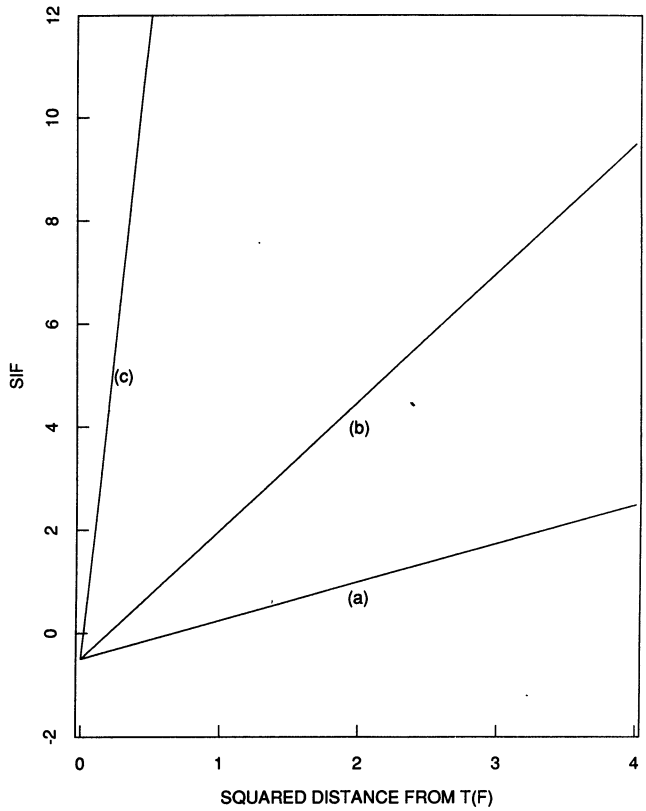

FIG. 2. *Standardized influence function of directional dispersion* (*a*) *when* $\|C(F)\| = 2/3$; (*b*) *when* $\|C(F)\| = 9/10$; *and* (*c*) *when* $\|C(F)\| = 99/100$.

$(1 + \|C(F)\|)(1 - \|C(F)\|)^{-1/2}/2$, which is bounded if $\|C(F)\| \neq 1$. Since the standardized gross error sensitivity of S with respect to itself at $\{F\}$ is $(1 + \|C(F)\|)(1 - \|C(F)\|)^{-1}/2$, the standardized gross error sensitivity at $\mathscr{F}$ with $\sup\|C(F)\| = 1$ is infinity. Furthermore, the standard gross error sensitivity at $\mathscr{F}$ with respect to the constant functional 1 is infinity. This strongly indicates that the directional dispersion is not robust when we deal with highly concentrated data on Ω_q.

3.3. *Concentration.* The von Mises–Fisher distribution with the mean directional vector μ and the concentration parameter κ on Ω_q has a density function of the form

$$f(x; \mu, \kappa) = (2\pi)^{-q/2}\left\{I_{(q/2)-1}(\kappa)\kappa^{-(q/2)+1}\right\}^{-1}\exp\{\kappa\mu^T x\},$$

for $x \in \Omega_q$, where $I_p(\cdot)$ is the modified Bessel function of the first kind and order P. The maximum likelihood estimator of the concentration parameter κ is given by $\hat{\kappa} = A_q^{-1}(\|C(F_n)\|)$, where F_n is the empirical measure of the data $x_1, \dots, x_n$

on Ω_q and A_q^{-1} is the inverse function of $A_q(\cdot) = I_{(q/2)}(\cdot)/I_{(q/2)-1}(\cdot)$. See Watson (1983) for details.

THEOREM 5. *The influence function of the functional K defined by $K(F) = A_q^{-1}(\|C(F)\|)$ at F is*

$$(x^T T(F) - \|C(F)\|)\{1 - \|C(F)\|^2 - (q-1)\|C(F)\|/K(F)\}^{-1}.$$

PROOF. By the chain rule,

$$\mathrm{IF}(x; K, F) = d/dx|_{x=\|C(F)\|} A_q^{-1}(x)\mathrm{IF}(x; \|C(\cdot)\|, F).$$

Since $d/dx A_q(x) = 1 - (A_q(x))^2 - (q-1)A_q(x)/x$ [see Watson (1983), page 193],

$$\begin{aligned} d/dx A_q^{-1}(x) &= \left\{1 - \left(A_q\left(A_q^{-1}(x)\right)\right)^2 - (q-1)A_q\left(A_q^{-1}(x)\right)\right)/A_q^{-1}(x)\right\}^{-1} \\ &= \left\{1 - x^2 - (q-1)x/A_q^{-1}(x)\right\}^{-1}. \end{aligned}$$

Using Theorem 4,

$$\mathrm{IF}(x; K, F) = (x^T T(F) - \|C(F)\|)\{1 - \|C(F)\|^2 - (q-1)\|C(F)\|/K(F)\}^{-1}.$$

□

The influence function of K is bounded at F with $0 < \|C(F)\| < 1$. However, the standardized gross error sensitivity with respect to S (or even with respect to a nonzero constant functional) at $\mathscr{F}$ with $\sup\|C(F)\| = 1$ is infinity, indicating nonrobustness of the estimator especially when the data are highly concentrated. This supports the claim of Fisher (1982), Kimber (1985) and Watson (1986) that this estimator is not robust.

Acknowledgments. The authors are most grateful to the Associate Editor and the referees for their thorough reading of and constructive suggestions on an earlier draft of this paper.

REFERENCES

BARNETT, V. and LEWIS, T. (1984). *Outliers in Statistical Data*, 2nd ed. Wiley, New York.

BICKEL, P. J. and LEHMANN, E. L. (1976). Descriptive statistics for nonparametric models. III. Dispersion. *Ann. Statist.* **4** 1139–1158.

COLLETT, D. (1980). Outliers in circular data. *Appl. Statist.* **29** 50–57.

FISHER, N. I. (1982). Robust estimation of the concentration parameter of Fisher's distribution on the sphere. *Appl. Statist.* **31** 152–154.

FISHER, N. I. (1985). Spherical medians. *J. Roy. Statist. Soc. Ser. B* **47** 342–348.

FISHER, N. I., LEWIS, T. and WILLCOX, M. E. (1981). Tests of discordancy for samples from Fisher's distribution on the sphere. *Appl. Statist.* **30** 230–237.

HAMPEL, F. R. (1968). Contributions to the theory of robust estimation. Ph.D. thesis, Univ. California, Berkeley.

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

HAMPEL, F. R., RONCHETTI, E. M., ROUSSEEUW, P. J. and STAHEL, W. A. (1986). *Robust Statistics*. Wiley, New York.

KIMBER, A. C. (1985). A note on the detection and accommodation of outliers relative to Fisher's distribution on the sphere. *Appl. Statist.* **34** 169–172.

LAMBERT, D. (1981). Influence functions for testing. *J. Amer. Statist. Assoc.* **76** 649–657.

LENTH, R. V. (1981). Robust measure of location for directional data. *Technometrics* **23** 77–81.

MARDIA, K. V. (1972). *Statistics of Directional Data*. Academic, New York.

MARDIA, K. V. (1975). Statistics of directional data (with discussion). *J. Roy. Statist. Soc. Ser. B* **37** 349–393.

ROUSSEEUW, P. J. (1981). A new infinitesimal approach to robust estimation. *Z. Wahrsch. verw. Gebiete* **56** 127–132.

ROUSSEEUW, P. J. (1984). Least median of squares regression. *J. Amer. Statist. Assoc.* **79** 871–880.

WATSON, G. S. (1983). *Statistics on Spheres*. Wiley, New York.

WATSON, G. S. (1986). Some estimation theory on the sphere. *Ann. Inst. Statist. Math.* **38** 263–275.

WEHRLY, T. E. and SHINE, E. P. (1981). Influence curves of estimators for directional data. *Biometrika* **68** 334–335.

DEPARTMENT OF BIOSTATISTICS
MEDICAL COLLEGE OF VIRGINIA
VIRGINIA COMMONWEALTH UNIVERSITY
RICHMOND, VIRGINIA 23298

DEPARTMENT OF STATISTICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195