The Annals of Statistics
1982, Vol. 10, No. 2, 643–645

L- AND *R*-ESTIMATION AND THE MINIMAX PROPERTY

BY JEROME SACKS[1] AND DONALD YLVISAKER[2]

Northwestern University and University of California, Los Angeles

Let $\{X_i\}$ be a sample from $F(x-\theta)$ where F is in a class $\mathscr{F}$ of symmetric distributions on the line and θ is the location parameter to be estimated. Huber has shown that maximum likelihood estimation has a minimax property over a convex $\mathscr{F}$. Here a simple convex $\mathscr{F}$ is given for which neither L- nor R-estimation has the minimax property. In particular, this example shows that a recent assertion concerning L-estimation is not true.

0. Introduction. Let $\mathscr{F}$ be a collection of symmetric distributions on the line and consider the location parameter problem based on a sample from $F(x-\theta)$, where $F \in \mathscr{F}$ and θ denotes the location parameter to be estimated. This set-up is basic in the development of a theory of robustness by Huber (1964), and it is known that when $\mathscr{F}$ is convex there is a maximum likelihood estimate of θ which minimizes the maximum (over $\mathscr{F}$) asymptotic variance. There are natural competitors to M-estimates in these settings, namely L-estimates (linear functions of order statistics) and R-estimates (estimates based on ranks). In the case that $\mathscr{F}$ is the class of ε-contaminated normals, i.e. $F = (1-\varepsilon)\Phi + \varepsilon H$ with H symmetric, Jaeckel (1971) showed that appropriate L- and R-estimates also enjoy the minimax property mentioned above. On the other hand, there are $\mathscr{F}$'s given in Sacks and Ylvisaker (1972), in particular the ε-normal family $\mathscr{F} = \{F \mid \sup_x |F(x) - \Phi(x)| \le \varepsilon, F \text{ symmetric}\}$ for ε large enough, for which no L-estimate has the minimax property. In this note, we give a simple convex class $\mathscr{F}$ for which there is neither an L- nor an R-estimate that is asymptotically minimax. While it is not surprising that there should be such an $\mathscr{F}$—see the discussion of this property in Huber (1981) page 97, for instance—we are not aware that any has been constructed for R-estimation, let alone for L- and R-estimation simultaneously.

It has been claimed in Gribkova and Egorov (1978) that L-estimation has the minimax property for a convex $\mathscr{F}$ provided each density $f = F'$ in the class satisfies $\lim_{|x|\to\infty} f(x) = 0$. Our example shows this assertion to be false. Moreover, it will be clear that even monotonicity of each $f(|x|)$ is not sufficient to bring about the property. There remains the interesting question of finding general conditions on $\mathscr{F}$ which would guarantee asymptotically minimax L- or R-estimates.

1. The Example. Begin with the density f_0 on R^1 defined by

$$f_0(x) = \begin{cases} \dfrac{3}{5}\dfrac{e^x}{(1+e^x)^2}, & |x| \le \ln 2, \\[2ex] \dfrac{2}{15} e^{-(1/3)(|x| - \ln 2)}, & |x| > \ln 2, \end{cases}$$

and take $F_0' = f_0$. From Huber (1964) one can identify F_0 as the least informative distribution in the contaminated logistic model $\mathscr{F}_0 = \{F \mid F = (3/5)L + (2/5)H, H \text{ symmetric}\}$, where $L(x) = e^x/(1+e^x)$, $x \in R^1$. (The contamination constant $\varepsilon = 2/5$, leading to the

Received October 1981; revised December 1981.

[1] This author's work was supported in part by NSF Grant MCS-80-01789.

[2] This author's work was supported in part by NSF Grant MCS-80-02732.

AMS 1980 *subject classification.* Primary 62G35; secondary 62G20.

Key words and phrases. Location parameter estimation, robust estimation, minimax property.

breakpoints at $\pm\ln 2$ in Huber's analysis, has been chosen for computational simplicity.) In any event, we rely only on the fact that f_0 is a valid density.

For the location estimation problem under F_0, there are three efficient estimates to mention: an M-estimate with a skew-symmetric ψ function given by

$$\psi_0(x) = -\frac{f_0'(x)}{f_0(x)} = \begin{cases} \dfrac{e^x - 1}{e^x + 1}, & 0 \le x \le \ln 2 \\ \dfrac{1}{3}, & \ln 2 \le x < \infty, \end{cases}$$

an R-estimate determined through the skew-symmetric weight function

$$J_0(t) = -\frac{f_0'(F_0^{-1}(t))}{f_0(F_0^{-1}(t))} = \begin{cases} \dfrac{5}{3}(2t-1), & \dfrac{1}{2} \le t \le \dfrac{3}{5}, \\ \dfrac{1}{3}, & \dfrac{3}{5} \le t \le 1, \end{cases}$$

on [0, 1], and an L-estimate with a symmetric weight function on [0, 1] given by

$$w_0(t) = -\frac{(\ln f_0)''(F_0^{-1}(t))}{I(F_0)} = \begin{cases} \dfrac{30}{13}(5t-1)(4-5t), & \dfrac{1}{2} \le t \le \dfrac{3}{5}, \\ 0, & \dfrac{3}{5} \le t \le 1. \end{cases}$$

Here $I(F_0) = \int f_0'^2/f_0 = 13/135$ is the information number of F_0. If sampling is done from a suitably smooth and symmetric F, the asymptotic variances associated with the three estimates above are denoted by $V_M^0(F)$, $V_R^0(F)$ and $V_L^0(F)$. In particular,

$$V_R^0(F) = \int_0^1 J_0^2(t)\,dt \Big/ \left\{\frac{5}{3}\int_{F^{-1}(2/5)}^{F^{-1}(3/5)} 2f^2(x)\,dx\right\}^2 \tag{1.1}$$

and one finds that $\int_0^1 J_0^2(t)\,dt = I(F_0) = \{V_R^0(F_0)\}^{-1}$.

Let $\Delta > I^{-1}(F_0)$ be given. We will produce a symmetric G satisfying

$$I(F_0) \le I(\lambda G + (1-\lambda)F_0), \qquad 0 \le \lambda \le 1, \tag{1.2}$$

so that F_0 is least informative in $\mathscr{F} = \{F \mid F = \lambda G + (1-\lambda)F_0, 0 \le \lambda \le 1\}$. Moreover it will be the case that $\max_{\mathscr{F}} V_R^0(F) > I^{-1}(F_0)$ and $\max_{\mathscr{F}} V_L^0(F) > \Delta$. This is the desired conclusion. By way of contrast, one has $\max_{\mathscr{F}} V_M^0(F) = V_M^0(F_0) = I^{-1}(F_0)$.

To put (1.2) in a more useful way, set

$$u_0(x) = I(f_0) + 4\frac{(f_0^{1/2})''}{f_0^{1/2}} = \begin{cases} \dfrac{148}{135} - \dfrac{8e^x}{(1+e^x)^2}, & |x| \le \ln 2, \\ \dfrac{28}{135}, & |x| > \ln 2. \end{cases}$$

It is noted in Huber (1981) page 82 that (1.2) is equivalent to $\int u_0 g \le 0$, or to

$$\int_{-\ln 2}^{\ln 2} \left\{\frac{2e^x}{(1+e^x)^2} - \frac{2}{9}\right\} g(x)\,dx \ge \frac{7}{135}. \tag{1.3}$$

Clearly (1.3) involves only the central portion of g. Let us take

$$g(x) = c\left\{\frac{2e^x}{(1+e^x)^2} - \frac{2}{9}\right\}, \qquad |x| \le \ln 2,$$

in order to have equality in (1.3). Some calculation will give $c = 21/\{10(1 + 4\ln 2)\}$, and then

$$\text{(1.4)} \qquad \begin{aligned} &\text{(i)} \int_{-\ln 2}^{\ln 2} g(x)\,dx = \frac{7(3 - 2\ln 2)}{15(1 + 4\ln 2)} < .2 = \int_{-\ln 2}^{\ln 2} f_0(x)\,dx, \\ &\text{(ii)} \frac{5}{3}\int_{-\ln 2}^{\ln 2} 2g^2(x)\,dx = \frac{49}{135(1 + 4\ln 2)} \le \frac{13}{135} = \frac{5}{3}\int_{-\ln 2}^{\ln 2} 2f_0^2(x)\,dx. \end{aligned}$$

To complete the construction of $\mathscr{F}$ we make an appropriate extension of g outside $[-\ln 2, \ln 2]$. This is to be done so that

$$\text{(1.5)} \qquad \begin{aligned} &\text{(i)}\ G^{-1}\left(\frac{3}{5}\right) > \Delta', \qquad \Delta' \text{ suitably large,} \\ &\text{(ii)} \frac{5}{3}\int_{G^{-1}(2/5)}^{G^{-1}(3/5)} 2g^2(x)\,dx < \frac{13}{135}, \end{aligned}$$

a task which is possible because of (1.4). From (1.1) and (1.5) it follows that $\max_{\mathscr{F}} V_R^0(F) \ge V_R^2(G) > I_0^{-1}(F_0)$. Furthermore, it is easy to argue, from the form of $V_L^0(G)$ for instance (Jaeckel, 1971), that a sufficiently large value of $G^{-1}(3/5)$ results in $V_L^0(G) > \Delta$, since w_0 assigns positive weight to $[2/5, 3/5]$.

Finally, observe that g is still defined only on $[G^{-1}(2/5), G^{-1}(3/5)]$ and we have already obtained the desired properties of asymptotic variance over $\mathscr{F}$. Thus g can be taken to be a symmetric density, positive and monotone on $x \ge 0$ for example. Any $f \in \mathscr{F}$ would then share these properties, being a convex combination of f_0 and g. Moreover, by construction, $I(F)$ is finite on $\mathscr{F}$ and is uniquely minimized at F_0.

REFERENCES

GRIBKOVA, N. V. and EGOROV, V. A. (1978). Robust estimations of the location parameter which are linear combinations of order statistics. *Vestnik Leningrad Univ. Mat. Meh. Astronom* vyp. **3** 24–27.

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

HUBER, PETER J. (1981). *Robust Statistics.* Wiley, New York.

JAECKEL, LOUIS A. (1971). Robust estimations of location: symmetry and asymmetric contamination. *Ann. Math. Statist.* **42** 1020–1034.

SACKS, JEROME and YLVISAKER, DONALD (1972). A note on Huber's robust estimation of a location parameter. *Ann. Math. Statist.* **43** 1068–1075.

DEPARTMENT OF MATHEMATICS
NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS 60201

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, LOS ANGELES
LOS ANGELES, CALIFORNIA 90024