ON SOME ROBUST ESTIMATES OF LOCATION[1]

By Peter J. Bickel

University of California, Berkeley

1. Summary. During the past 15 years various approaches have been proposed to deal with the lack of robustness of the sample mean as an estimate of the population mean when the distribution sampled is contaminated by gross errors, i.e., has heavier tails than the normal distribution. First, Tukey and the Statistical Research Group at Princeton in [9] suggested and investigated the properties of "trimmed" and "Winsorized" means. More recently, Hodges and Lehmann [6], proposed estimates related to the well-known robust Wilcoxon and normal scores tests, among others. Finally Huber in [7] considered essentially the class of maximum likelihood estimates and found those members of this class which minimize the maximum variance over various classes of contaminated distributions. For a review of work in these directions in testing as well as estimation the reader is referred to Elashoff [3].

In Theorems 3.1 and 3.2 we state the main results of the asymptotic theory of the Winsorized and trimmed means and outline the proof. An alternative method of trimming and Winsorizing (not equivalent to that of Tukey) which encompasses the efficient estimates proposed by Huber and which generalizes to higher dimensions is discussed in Section 2.

The fourth section (Theorem 4.1) gives the minimum efficiency with respect to the families of all symmetric and symmetric unimodal distributions, of the Winsorized and trimmed means with respect to the mean. The lower bounds found for the trimmed means (for small trimming proportions) in the unimodal case compare well with that found by Hodges and Lehmann in [5] for the median of averages of pairs, the Hodges-Lehmann estimate. However, the Winsorized mean (for unimodal distributions) has minimum efficiency $\frac{1}{3}$ with respect to the mean whatever be the trimming proportion used. For all distributions, the minimum efficiency is 0.

Also in the fourth section (Theorem 4.2) we compare the trimmed mean to the H-L estimate and find that while the latter can be infinitely more efficient than the former, the H-L estimate, for small trimming proportions, $\alpha = .05$, is at least 90 per cent (approximately) as efficient. This would suggest that unless the computations involved are prohibitive, the H-L estimate is to be preferred in any situation where the degree of contamination and type of distribution is not known with great precision. The same remarks apply to the Winsorized mean with only somewhat less force since the lower bounds involved are .74 for all symmetric distributions and .79 for symmetric unimodal distributions.

Finally we compare the principal estimate proposed by Huber in [7] (Proposal

Received 30 September 1964; revised 6 January 1965.

[1] Prepared with the partial support of the National Science Foundation, Grant GP-2593.

2) to the mean and the Hodges-Lehmann estimate, both for all symmetric densities and for the symmetric unimodal family. Results similar to those already mentioned in connection with the trimmed mean are obtained in Theorems 5.1 and 5.2.

2. Some definitions. We shall assume, henceforth, that X_i, for $1 \leqq i \leqq n$ are a sample from a population with a continuous strictly increasing distribution $F(x - \theta)$ where θ is unknown and F is symmetric about 0, i.e., $F(x) = 1 - F(-x)$. We denote by $x(\alpha)$ the α quantile of F, i.e., the solution of the equation $F(x) = \alpha$. We shall also assume that F is absolutely continuous with respect to Lebesgue measure and possesses a density f continuous and strictly positive on its convex support $C = \{x: 0 < F(x) < 1\}$. Finally we denote by $W_1 < W_2 < \cdots < W_n$ the order statistics of the sample $X_1, \cdots, X_n$.

Following Tukey we first define the α-trimmed mean of the sample by

$$\bar{X}_\alpha = \{n - 2[\alpha n]\}^{-1} \sum_{i=[\alpha n]+1}^{n-[\alpha n]} W_i, \tag{2.1}$$

where $[\alpha n]$ is the greatest integer in αn for $0 \leqq \alpha < \frac{1}{2}$.

Similarly we define the α-Winsorized mean by

$$X_\alpha^* = n^{-1}[[\alpha n]W_{[\alpha n]} + \sum_{i=[\alpha n]+1}^{n-[\alpha n]} W_i + [\alpha n]W_{n-[\alpha n]+1}]. \tag{2.2}$$

The mean will be referred to as usual by $\bar{X}$ and the Hodges-Lehmann estimate, $\text{med}_{i \leqq j} (X_i + X_j)/2$ by M.

We will also consider the estimate given in Proposal 2 of Huber, which is obtained as the unique solution T of the system of equations

$$\sum_{i=1}^n \psi(k, (x_i - T)/s) = 0, \qquad \sum_{i=1}^n \psi^2(k, (x_i - T)/s) = 0,$$

where $\psi(\lambda, t) = \rho_1'(t)$. We shall denote this estimate by $H(k)$.

As an alternative to the method of Tukey which seems capable of extension to higher dimensions we define the λ trimmed and Winsorized means about $\hat{\theta}$, as follows. Let $\hat{\theta}$ be a consistent, asymptotically normal, estimate of θ.

Define the λ metrically trimmed mean about $\hat{\theta}$ by

$$\bar{X}_\lambda(\hat{\theta}) = N^{-1} \sum_{i=1}^n X_i I(|X_i - \hat{\theta}| \leqq \lambda), \tag{2.3}$$

where $N = \sum_{i=1}^n I(|X_i - \hat{\theta}| \leqq \lambda)$ and $I(A)$ is the indicator of the event A.

Similarly, the λ-Winsorized mean about $\hat{\theta}$ may be defined by

$$X_\lambda^*(\hat{\theta}) = n^{-1} [(\hat{\theta} - \lambda)N_1(\hat{\theta} - \lambda) + \sum_{i=1}^n X_i I(|X_i - \hat{\theta}| \leqq \lambda) + (\hat{\theta} + \lambda)N_2(\hat{\theta} + \lambda)] \tag{2.4}$$

where $N_1(\hat{\theta} - \lambda) = \sum_{i=1}^n I(X_i \leqq \hat{\theta} - \lambda)$, $N_2(\hat{\theta} + \lambda) = \sum_{i=1}^n I(X_i \geqq \hat{\theta} + \lambda)$.

If trimming and Winsorizing metrically about a point are considered as operations, clearly two estimates of particular interest are those points which remain invariant under each of these operations. Winsorizing leads to the estimate proposed by Huber in [6] (scale known), the point which minimizes $\sum_{i=1}^n \rho_1(X_i - \theta)$, where $\rho_1(t) = \frac{1}{2}t^2$ for $|t| \leqq \lambda$, $\rho_1(t) = \lambda |t| - \frac{1}{2}\lambda^2$ for $|t| \geqq \lambda$.

Trimming leads to the estimate characterized by Huber as minimizing $\sum_{i=1}^{n} \rho_2(X_i - \theta)$ where $\rho_2(t) = \frac{1}{2}t^2$ for $|t| \leqq \lambda$, $\rho(t) = \frac{1}{2}\lambda^2$ for $|t| \geqq \lambda$.

The first of these estimates has been shown by Huber to possess a minimax property for the contaminated normal model. The asymptotic normality of the second has not yet been established. Asymptotic normality of metrically trimmed and Winsorized means seem in general to hold only under regularity conditions. We hope to investigate this class of estimates in a subsequent paper.

3. Asymptotic theory. We now state three theorems which give the necessary formulae for the asymptotic variance of $\bar{X}_\alpha$, X_α^*, and $H(k)$ which we shall need in Sections 4 and 5.

THEOREM 3.1. *Under the conditions of Section* 2.

(a) $$\mathcal{L}(n^{\frac{1}{2}}(\bar{X}_\alpha - \theta)) \to N(0, \sigma_1^2(\alpha)) \quad as \quad n \to \infty,$$

where $\sigma_1^2(\alpha) = (1 - 2\alpha)^{-2}[\int_{x(\alpha)}^{x(1-\alpha)} t^2\, dF(t) + 2\alpha x_\alpha^2]$.

(b) *If moreover* $E(X_1^2) < \infty$ *then also*

$$\text{Var}\,(n^{\frac{1}{2}}(\bar{X}_\alpha - \theta)) \to \sigma_1^2(\alpha).$$

THEOREM 3.2. *Under the above conditions*

(a) $\mathcal{L}(n^{\frac{1}{2}}(X_\alpha^* - \theta)) \to N(0, \sigma_2^2(\alpha))$ *as* $n \to \infty$, *where*

$$\sigma_2^2(\alpha) = \int_{x(\alpha)}^{x(1-\alpha)} t^2\, dF(t) + 2\alpha[x(1 - \alpha) + \alpha/f(x_\alpha)]^2.$$

(b) *If* $E(X_1^2) < \infty$ *we may also conclude that*

$$\text{Var}\,(n^{\frac{1}{2}}(X_\alpha^* - \theta)) \to \sigma_2^2(\alpha).$$

Theorem 3.1(a) was stated in Tukey and Harris [4]. Theorem 3.2(a) appears at least implicitly in [9]. Both parts of both theorems follow readily from Bickel [1]. For convenience we sketch here a different method of proof for 3.1 and 3.2(a) first used by Sethuraman in [8] in another connection.

Let $R_n = n^{-\frac{1}{2}} \sum_{k=[\alpha n]+1}^{n-[\alpha n]} [W_k - E(W_k \mid W_{[\alpha n]}, W_{n-[\alpha n]+1})]$, $S_n = n^{\frac{1}{2}}(W_{[\alpha n]} - x(\alpha))$, $T_n = n^{\frac{1}{2}}(W_{n-[\alpha n]+1} - x(1 - \alpha))$. It is well known that conditional on W_r, W_s, $r < s$, $W_{r+1}, \cdots, W_{s-1}$ are distributed as the order statistics of a sample of $s - r - 1$ from a population with density $f(x)[F(W_s) - F(W_r)]^{-1}$ for $W_r \leqq x \leqq W_s$ and 0 otherwise. It follows from the above, Theorem 1 of [8] and the normal convergence of quantiles that (R_n, S_n, T_n) are jointly normal with mean 0 and covariance matrix $\|d_{ij}\|$ where $d_{ij} = 0$ for $i \neq 1$, $d_{23} = \alpha^2/f^2(x(\alpha))$, $d_{11} = \int_{x(\alpha)}^{x(1-\alpha)} t^2 f(t)\, dt$, $d_{22} = d_{33} = \alpha(1 - \alpha)/f^2(x(\alpha))$. Upon remarking that,

$$n^{\frac{1}{2}}(\bar{X}_\alpha - \theta)$$
$$= (n/n - 2m)R_n + n^{\frac{1}{2}}\{\int_{W_{[n\alpha]}}^{W_{n-[n\alpha]+1}} tf(t)\, dt/[F(W_{n-[n\alpha]+1}) - F(W_{[n\alpha]})]\}$$

and employing Taylor's theorem, Theorem 3.1(a) follows. Theorem 3.2(a) may be proved similarly.

We also state without proof the following theorem due to Huber [7].

THEOREM 3.3 (Huber). *Under the above conditions,*

$$\mathcal{L}(n^{\frac{1}{2}}(H(k) - \theta)) \to N(0, \sigma_3^2(k)) \quad as \quad n \to \infty,$$

where

$$\sigma_3^2(k) = (\int_{-q}^{q} dF(t))^{-2}(\int_{-q}^{q} t^2 \, dF(t) + 2q^2 \int_q^\infty dF(t))$$

and q satisfies

$$q^2\beta(k)/k^2 = \int_{-q}^{q} t^2 \, dF(t) + 2q^2 \int_q^\infty dF(t) \tag{3.1}$$

and

$$\beta(k) = \int_{-k}^{k} t^2 \, d\Phi(t) + 2k^2 \int_k^\infty d\Phi(t)$$

where Φ is the standard normal distribution.

It is interesting to note that for each fixed F, $H(k)$ (a metrically Winsorized mean) has an asymptotic variance which is the same as that of an α trimmed mean (in the sense of Tukey).

4. Comparison of the Tukey estimates to $\bar{X}$ and M. Let $\mathfrak{F}$ be the family of all symmetric distributions possessing the regularity conditions of Section 2. Let $\mathcal{G}$ be the family of all symmetric unimodal distributions which possess the above regularity conditions. Define $e_1(\alpha)$ to be the efficiency, in terms of the ratio of asymptotic variances, of $\bar{X}_\alpha$ to $\bar{X}$ and similarly $e_2(\alpha)$ the efficiency of X_α^* to $\bar{X}$, where dependence upon F is understood. Then we have,

$$\inf_{F \varepsilon \mathfrak{F}} e_1(\alpha) = (1 - 2\alpha)^2, \tag{4.1}$$

$$\inf_{F \varepsilon \mathfrak{F}} e_2(\alpha) = 0. \tag{4.2}$$

Equation (4.2) follows immediately by choosing distributions with arbitrarily small $x(\alpha)$.

Equation (4.1) is also immediate since

$$\int_{-\infty}^{\infty} t^2 \, dF(t) = \int_{x(\alpha)}^{x(1-\alpha)} t^2 \, dF(t) + 2 \int_{x(1-\alpha)}^{\infty} t^2 \, dF(\mathrm{t})$$
$$\geqq \int_{x(\alpha)}^{x(1-\alpha)} t^2 \, dF(t) + 2\alpha x_{1-\alpha}$$

and the lower bound may clearly be approached by distributions concentrating their mass outside $(x(\alpha), x(1 - \alpha))$ more and more closely to the endpoints.

We can also establish

THEOREM 4.1.

$$\inf_{F \varepsilon \mathcal{G}} e_1(\alpha) = (1 + 4\alpha)^{-1}, \tag{4.3}$$

$$\inf_{F \varepsilon \mathcal{G}} e_2(\alpha) = \tfrac{1}{3}. \tag{4.4}$$

PROOF. We prove (4.4) first. Let (1) $x(\alpha) = -\lambda$, (2) $f(x(\alpha)) = k$, (3) $\int_{-\lambda}^{\lambda} x^2 f(x) \, dx = c$. We wish first to minimize

$$e_2(\alpha) = \int_{-\infty}^{\infty} x^2 f(x) \, dx / \{c + 2\alpha[\lambda + \alpha/k]^2\}$$

for all $F \varepsilon \mathcal{G}$ which also satisfy (1), (2) and (3).

This is equivalent to minimizing $\int_\lambda^\infty x^2 f(x) \, dx$ subject to the above side conditions and upon applying the method of undetermined multipliers we find that $\int_\lambda^\infty (x^2 - \delta^2)f(x)\,dx$ is minimized by

$$f(x) = k, \qquad \lambda \leqq x \leqq \alpha/k + \lambda,$$
$$= 0, \qquad x \geqq \alpha/k + \lambda,$$

for $\delta = \alpha/k + \lambda$, yielding as a minimum

$$(4.5)\quad e_2(c, k, \lambda, \alpha) = 1 - [2\alpha^2\lambda/k + \tfrac{4}{3}(\alpha^3/k^2)] \cdot (c + 2\alpha\lambda^2 + 4\alpha^2\lambda/k + 2\alpha^3/k^2)^{-1}.$$

Hence e_2 is minimized when c is. The minimum of $\int_{-\lambda}^{\lambda} x^2 f(x)\,dx$ subject to f unimodal, $f(\lambda) = f(-\lambda) = k$, f symmetric, and $\int_{-\lambda}^{\lambda} f(x)\,dx = 1 - 2\alpha$ may clearly be approached and, if $k = (1 - 2\alpha)/2\lambda$, is achieved for $f(x) \equiv k$ for all $-\lambda \leqq x \leqq \lambda$. The minimum of $e_2(c, k, \lambda, \alpha) = e_2(k, \lambda, \alpha)$ is given by

$$(4.6)\quad e_2(k, \lambda, \alpha) = 1 - 3\alpha^2\{[(\lambda k + \alpha) - \tfrac{1}{3}\alpha]/[(\lambda k + \alpha)^3 + (2\alpha^3 + 3\alpha^2\lambda k)]\}.$$

Letting $t = \lambda k + \alpha$, we obtain

$$(4.7)\quad e_2(k, \lambda, \alpha) = e_2(t, \alpha) = 1 - 3\alpha^2(t - \tfrac{1}{3}\alpha)(t^3 + 3\alpha^2 t - \alpha^3)^{-1}.$$

By the unimodal property of f the range of t is clearly the open interval $(\alpha, \frac{1}{2})$. It is easy to see (by differentiation) that $(t - \frac{1}{3}\alpha)(t^3 + 3\alpha^2 t - \alpha^3)^{-1}$ approaches its maximum as $t \to \alpha$ and since appropriate sequences of distributions may readily be constructed we find that $\inf e_2(t, \alpha) = \frac{1}{3}$.

We now prove (4.3). As before, fix $x(\alpha) = -\lambda$, $f(\lambda) = k$, $\int_{-\lambda}^{\lambda} x^2 f(x)\,dx = c$. We wish to minimize

$$(1 - 2\alpha)^2\left(\int_{-\infty}^{\infty} x^2 f(x)\,dx\right)\left(\int_{-\lambda}^{\lambda} x^2 f(x)\,dx + 2\alpha x_\alpha^2\right)^{-1} = e_1(\alpha).$$

Under the full set of side conditions we first minimize $\int_\lambda^\infty x^2 f(x)\,dx$ which as before yields a minimum of

$$(4.8)\quad e_1(\alpha, c, k, \lambda) = (1 - 2\alpha)^2[1 + (2\alpha^2\lambda/k + \tfrac{2}{3}\alpha^3/k^2)(c + 2\alpha\lambda^2)^{-1}].$$

The problem is now to maximize c subject to the given side conditions. We readily find that the maximum value of c which is approachable but in general not attainable is $(1 - 2\alpha)\lambda^2/3$ from the following

Lemma 4.1. *$\int_{-a}^{a} x^2 h(x)\,dx$ is maximized among all symmetric unimodal probability densities by $h(x) \equiv \frac{1}{2}a$, $|x| \leqq a$.*

Proof. Suppose h^* is any other such density. Without loss of generality take h^* continuous. Then there exists a t, where $0 \leqq t < a$, such that $h^*(x) < \frac{1}{2}a$ for $a > x > t$ and $h^*(x) \geqq \frac{1}{2}a$ for $0 \leqq x \leqq t$. Then

$$\int_0^a x^2 h^*(x)\,dx - (2a)^{-1}\int_0^a x^2\,dx = \int_0^t x^2[h^*(x) - (2a)^{-1}]\,dx - \int_t^a x^2[(2a)^{-1}h^*(x)]\,dx < t^2\{\int_0^t [h^*(x) - (2a)^{-1}]\,dx - \int_t^a [(2a)^{-1} - h^*(x)]\,dx\} = 0. \qquad \text{QED.}$$

It follows that the infimum of $e_1(\alpha, c, k, \lambda)$ equals

(4.9) $$e_1(\alpha, k, \lambda) = (1-2\alpha)^2[(1 + (2\alpha^2\lambda/k + \tfrac{4}{3}\alpha^3/k^2)((1-2\alpha)\lambda^2/3 + 2\alpha\lambda^2)^{-1}].$$

It remains only to maximize k. The maximum is clearly $k = (1 - 2\alpha)/2\lambda$. So in this case the least favorable distribution always exists and is any uniform distribution on a symmetric interval. We readily conclude that the minimum value is given by (4.3).

For $\alpha = .05$, a rather usual trimming proportion, we obtain an infimum value of .833 which is very close to that of the estimate M with respect to $\bar{X}$.

In the next theorem we consider the behavior of the α trimmed mean, the more successful of the two estimates proposed by Tukey with respect to M, the estimate considered by Hodges and Lehmann.

We have

THEOREM 4.2. *Let $e_3(\alpha)$ be the efficiency of M with respect to the trimmed mean $\bar{X}_\alpha$. Then*

(4.10) $$\inf_{F\varepsilon\mathfrak{F}} e_3(\alpha) = (27/2000)[(1 - 2\alpha)c(\alpha) + 10\alpha](1 - 2\alpha)^2(3c^2 - 10c + 15)^2$$

where $c(\alpha) = 1 + \tfrac{4}{3}[[3(\alpha^2 + \alpha)]^{\frac{1}{2}} - 3\alpha](1 - 2\alpha)^{-1}$. *Also the supremum of the efficiency is equal to*

(4.11) $$\sup_{F\varepsilon\mathfrak{F}} e_3(\alpha) = \infty.$$

PROOF. Let us recall that

(4.12) $$e_3(\alpha) = 12(1 - 2\alpha)^{-2}\left(\int_{-\infty}^{\infty} f^2(x)\,dx\right)^2\left[\int_{x(\alpha)}^{x(1-\alpha)} t^2 f(t)\,dt + 2\alpha x_\alpha^2\right].$$

Equation (4.11) clearly follows by letting $x(\alpha)$ tend to $-\infty$, keeping all other terms fixed. Intuitively this corresponds to taking distributions with heavier and heavier tails, i.e., guessing wrong about the trimming proportion. On the other hand,

(4.13) $$\inf_{F\varepsilon\mathfrak{F}} e_3(\alpha) = \inf_{F\varepsilon\mathfrak{F}} 12(1 - 2\alpha)^{-2}\left(\int_{x(\alpha)}^{x(1-\alpha)} f^2(x)\,dx\right)^2 \cdot\left[\int_{x(\alpha)}^{x(1-\alpha)} x^2 f(x)\,dx + 2\alpha x_\alpha^2\right]$$

since clearly one can make $\int_{x(1-\alpha)}^{\infty} f^2(x)\,dx$ arbitrarily small subject to $\int_{x(1-\alpha)}^{\infty} f(x)\,dx = \alpha$. Let $h(x) = [1/(1 - 2\alpha)]f(x)$ for $x(\alpha) \leqq x \leqq x(1 - \alpha)$, 0 otherwise, and denote the quantity whose infimum is taken on the right of (4.13) by $e_4(\alpha)$. Then

(4.14) $$e_4(\alpha) = 12(1 - 2\alpha)^2\left(\int_{x(\alpha)}^{x(1-\alpha)} h^2(x)\,dx\right)^2 \cdot\left[(1 - 2\alpha)\int_{x(\alpha)}^{x(1-\alpha)} x^2 h(x)\,dx + 2\alpha x_\alpha^2\right].$$

Let $x(\alpha) = -\lambda$, $\int_{-\lambda}^{\lambda} x^2 h(x)\,dx = c$. Minimizing $\int_{-\lambda}^{\lambda} h^2(x)\,dx$ subject to $\int_{-\lambda}^{\lambda} x^2 h(x)\,dx = c$ and $\int_{-\lambda}^{\lambda} h(x)\,dx = 1$ we find by using undetermined multipliers in a fashion similar to that used by Hodges and Lehmann in [4] that the minimum is achieved by

$$(4.15)\qquad h(x) = b_1(x^2 - a_1^2), \quad x^2 \leqq \lambda^2, \qquad = 0, \quad \text{otherwise,}$$

if $5c/3 \leqq \lambda^2 \leqq 3c$.

$$(4.16)\qquad h(x) = b_2(a_2^2 - x^2), \quad x^2 \leqq \lambda^2, \qquad = 0, \quad \text{otherwise,}$$

if $3c \leqq \lambda^2 \leqq 5c$.

$$(4.17)\qquad h(x) = b_3(a_3^2 - x^2), \quad x^2 \leqq 5c, \qquad = 0, \quad \text{otherwise,}$$

for $\lambda^2 \geqq 5c$, where,

$$(4.18)\qquad a_1^2 = a_2^2 = .2\lambda^2(3\lambda^2 - 5c)|\lambda^2 - 3c|^{-1},$$

$$(4.19)\qquad b_1 = b_2 = 1.875\lambda^{-5}\,|\lambda^2 - 3c|,$$

$$(4.20)\qquad a_3^2 = 5c,$$

$$(4.21)\qquad b_3 = .75(5c)^{-(3/2)}.$$

Solution of the variational problem for the range $c \leqq \lambda^2 \leqq \frac{5}{3}c$, for densities, is as we shall see unnecessary. We remark only that formally the solution given by (4.15) though no longer a density continues to minimize $\int h^2(t)\,dt$ subject to $\int h(t)\,dt = 1$ and $\int t^2h(t)\,dt = c$.

Substituting (4.18), (4.19) in (4.15), (4.16) we obtain after some computations, $\inf_{F\varepsilon\mathfrak{F}} e_4(\alpha)$ equal to

$$(4.22)\qquad e_4(\alpha, \lambda, c) = (108/64)\lambda^{-10}(1 - 2\alpha)^2(3\lambda^4 - 10c\lambda^2 + 15c^2)^2 \cdot ((1 - 2\alpha)c + 2\alpha\lambda^2).$$

Since the infimum is clearly independent of the choice of $\lambda \neq 0$ we choose $\lambda = 5^{\frac{1}{2}}$ (leading to the restriction $c \geqq 1$) and obtain

$$(4.23)\qquad e_4(\alpha, c) = (27/2000)((1 - 2\alpha)c + 10\alpha)(3c^2 - 10c + 15)^2(1 - 2\alpha)^2.$$

For the range $c \geqq 1$ it is easy to show that the minimum of the above expression is reached for $c = c(\alpha)$ given in the statement of the theorem. Now, for $c \geqq 3$, this expression though not giving the true infimum gives a lower bound to the efficiency and (from the preceding remark) since the minimizing $c(\alpha)$ is always <3, (4.10) will be established if we can dispense with the case $c < 1$. But, evaluating the corresponding lower bound as a function of c, we find that it assumes its minimum for $c = 1$. The conclusion of the theorem now follows.

Table 1 gives values of $\inf_{F\varepsilon\mathfrak{F}} e_3(\alpha)$ for various common values of α. Since they are all very high, we conclude that the Hodges-Lehmann estimate is preferable to the trimmed mean unless very precise knowledge of the required α is available.

Let $e_5(\alpha)$ denote the efficiency of M with respect to the α Winsorized mean. Then, it is easy to see that,

$$\inf_{F\varepsilon\mathfrak{F}} e_5(\alpha) = \inf_{F\varepsilon\mathfrak{F}} e_3(\alpha)(1 - 2\alpha)^2 \tag{4.24}$$

since it is clear that in general one can approach the least favorable densities of Theorem 4.2 so that all the quantities involved in $e_5(\alpha)$ remain fixed but $f(x(\alpha))$ tends to ∞. This is however clearly not possible if we restrict ourselves to $\mathcal{G}$. Computation of the exact lower bound in this case seems to be especially tedious. However, we can approximate the bound as follows. For $f \varepsilon \mathcal{G}$, $f(x(\alpha)) \leqq (1 - 2\alpha)/2x(1 - \alpha)$. From this remark and the arguments of Theorem 4.2 it

TABLE 1

$\alpha =$	.01	.02	.03	.04	.05	.06	.07	.08	.10
$\inf_{F\varepsilon\mathfrak{F}} e_3(\alpha)$	.89	.90	.91	.91	.91	.91	.90	.89	.865

TABLE 2

$\alpha =$	.01	.02	.03	.04	.05	.06	.07	.08	.10
$\inf_{F\varepsilon\mathfrak{F}} e_5(\alpha)$	.85	.83	.80	.77	.74	.70	.665	.63	.55
Lower bound of $e_5(\alpha)$ for $F \varepsilon \mathcal{G}$	.85	.84	.82	.80	.79	.77	.75	.73	.70

follows that

$$\inf_{F\varepsilon\mathcal{G}} e_5(\alpha) \geqq \inf_{c\geqq 1} (27/2000)(1 - 2\alpha)^2 \cdot[(1 - 2\alpha)^3 c + 10\alpha](3c^2 - 10c + 15)^2. \tag{4.25}$$

Values of the bound as well as the infimum for all distributions are given in Table 2 for selected values of α. Computation of $e(M, X_\alpha^*)$ for various distributions at $\alpha = .05$, indicate that the bound is within 2 per cent of the actual bound. Its accuracy of course improves for increasing α. For $\alpha = .01$ the bound would seem to underestimate the true infimum considerably, but since the value given by the estimate is already quite high (.85) this would seem to be satisfactory.

5. Comparison of H to $\bar{X}$ and M. Let $\tau = 2k^2/\beta(k)$. Since τ is a strictly monotone increasing function of k, we can and shall use it as an equivalent parametrization of the lower bounds we shall consider. We have

THEOREM 5.1. *Let $e_6(\tau)$ denote the efficiency of Huber's (Proposal 2) estimate with respect to the mean. Then,*

$$(5.1)\qquad \inf_{F\in\mathfrak{F}} e_6(\tau) = (1 - 2/\tau)^2,$$

$$(5.2)\qquad \inf_{F\in\mathcal{G}} e_6(\tau) = \tau/6, \qquad 2 \leqq \tau \leqq 3.6,$$
$$= \tfrac{1}{27}(\tau - 2u)^2(3 - 2u)^3[\tau(1 - u)]^{-2}, \qquad \tau \geqq 3.6,$$

where $u(\tau) = \frac{1}{12}[(10 + \tau) - (\tau^2 + 20\tau - 44)^{\frac{1}{2}}]$.

PROOF. Fix q, $\int_q^\infty dF(t) = \alpha$, $\int_{-q}^q t^2\, dF(t) = c$.

Relation (3.1) may now be rephrased as

$$(5.3)\qquad c = [2(1 - \alpha\tau)/\tau]q^2.$$

From (5.3) it is clear that $\alpha \leqq 1/\tau$, and hence by the same argument used to establish (4.1), it follows that $\inf_{F\in\mathfrak{F}} e_6(\tau) \geqq (1 - 2\alpha)^2 \geqq (1 - 2/\tau)^2$ and that the last bound may be approached by distributions concentrating a mass of $1 - 2/\tau + 2\epsilon$ near the origin and of $1/\tau - \epsilon$ on the outside of the points $\pm q$ in such a way as to satisfy (5.3), and letting $\epsilon \to 0$, and the outside mass converge to point mass at $\pm q$.

To obtain (5.2) we proceed as in the proof of Theorem 4.1 by fixing $f(q) = v$ ensuring always that (5.3) is satisfied.

We readily obtain that the infimum over all unimodal densities satisfying the given side conditions is given by

$$(5.4)\qquad e_6(\alpha, v, q, c, \tau) = (1 - 2\alpha)^2\{1 + (2\alpha^2 q/v + \tfrac{2}{3}\alpha^3/v^2)(c + 2\alpha q^2)^{-1}\}$$

as in (4.8).

Upon substituting from (5.3) and letting $q = 1$ with no loss in generality we obtain

$$(5.5)\qquad e_6(\alpha, v, t) = (1 - 2\alpha)^2[1 + (\alpha^2\tau/3)(3/v + \alpha/v^2)].$$

We now wish to maximize v subject to fixed α and c given by (5.3). We require

LEMMA 5.1. *Let* $\mathfrak{F}(c)$ *be the family of all symmetric unimodal densities on* $[-1, 1]$ *such that*

$$\int_{-1}^1 t^2 f(t)\, dt = c, \qquad (c \leqq \tfrac{1}{3}).$$

Then, $\sup_{\lambda\in\mathfrak{F}(c)} f(1) = \frac{3}{2}c$.

PROOF. Clearly unimodality and $f(1) > \frac{3}{2}c$ imply $\int t^2 f(t)\, dt > c$, while then the densities

$$f(t) = \tfrac{3}{2}c - \epsilon, \qquad a \leqq |t| \leqq 1,$$
$$= b, \qquad |t| \leqq a,$$

where $a = \{[(1 - 3c)^2 + 16\epsilon]^2 - (1 - 3c)\}(4\epsilon)^{-1}$ and $b = (1 - 3c)/2a + \frac{3}{2}c$, are members of $\mathfrak{F}(c)$ and approximate the supremum.

From the lemma it follows that the lower bound of $e_6(\alpha, v, \tau)$, equals, if $u = \alpha\tau$,

$$e_6(u, \tau) = \tfrac{1}{27}(\tau - 2u)^2(3 - 2u)^3[\tau(1 - 2u)]^{-2}.$$

The theorem now follows upon minimizing with respect to u and employing the

TABLE 3

$k =$	1.10	1.20	1.30	1.40	1.50	1.60	1.70	1.80	2.00
$\inf_{F\varepsilon\mathfrak{F}} e_6(k)$	.27	.31	.35	.39	.43	.46	.50	.53	.60
$\inf_{F\varepsilon\mathfrak{G}} e_6(k)$	.65	.72	.75	.80	.81	.85	.87	.88	.94

restrictions $(6 - \tau)/4 \leqq u \leqq 1$, of which the first follows from $c \leqq q^2(1 - 2\alpha)/3$.

Table 3 gives values of $\inf_{F\varepsilon\mathfrak{F}} e_6$ and $\inf_{F\varepsilon\mathfrak{G}} e_6(\tau)$ for various values of k.

Finally we have the comparison of M and $H(k)$ in

THEOREM 5.2. *Let $e_7(\tau)$ denote the efficiency of M with respect to $H(k)$. Then,*

$$\sup_{F\varepsilon\mathfrak{G}} e_7(\tau) = \infty, \tag{5.6}$$

$$\inf_{F\varepsilon\mathfrak{F}} e_7(\tau) = \tfrac{27}{8}\tau^{-5}(1 - 2\alpha)^{-2} \cdot [32\alpha^2\tau^2 + (8\tau^2 - 80\tau)\alpha + (3\tau^2 - 20\tau - 60)]^2 \tag{5.7}$$

for $2 \leqq \tau \leqq 7.415$, $\alpha(\tau) = .5 - .685(1 - 2/\tau)$,

$$\begin{aligned} \inf_{F\varepsilon\mathfrak{F}} e_7(\tau) &= \tfrac{27}{8}(\tau^{-5}(3\tau^2 - 20\tau + 60)^2), &&\textit{for } 7.45 \leqq \tau \leqq 10, \\ &= .864, &&\textit{for } \tau \geqq 10. \end{aligned} \tag{5.8}$$

PROOF. (5.6) follows readily upon taking $q = 1$ and defining a family of densities by

$$\begin{aligned} f_\epsilon(t) &= \epsilon^{-3/4} && \text{on } (-\epsilon, \epsilon), \\ &= a(\epsilon, \alpha) && \text{on } (-1, 1) - (-\epsilon, \epsilon) \\ &= b(\epsilon, \alpha), && \text{on } (-1, 1)^c, \end{aligned}$$

where $(6 - \tau)/4 < \alpha < 1/\tau$ and $a(\epsilon, \alpha)$, $b(\epsilon, \alpha)$ are chosen so as to satisfy (5.3) with $\int_q^\infty dF(t) = \alpha$ and $b \leqq a$. This is clearly possible for ϵ sufficiently small since $\int_0^\epsilon t^2 f_\epsilon(t)\, dt$, and $\int_0^\epsilon f_\epsilon(t)\, dt$ both converge to 0 as $\epsilon \to 0$.

Then clearly as $\epsilon \to 0$, $\int f^2(t)\, dt \to \infty$, a eventually becomes less than $\epsilon^{-3/4}$, c, α are fixed and (5.6) follows.

Equations (5.7) and (5.8) are obtained in the same fashion as Theorem 4.2. Fixing α, c, $q^2 = 5$ we find as in (4.17) to (4.20) that, if $\tilde{c} = c/(1 - 2\alpha)$, the lower bound is given by,

$$\begin{aligned} e_7(\alpha, \tau, c) &= (27/2000)((1 - 2\alpha)\tilde{c} + 10\alpha) \\ &\quad \cdot (3\tilde{c}^2 - 10\tilde{c} + 15)^2(1 - 2\alpha)^2 \quad \text{for } 1 \leqq \tilde{c} \leqq 3. \\ &= .864(1 - 2\alpha)^3(1 + 10\alpha/c) \quad \text{for } \tilde{c} \leqq 1. \end{aligned} \tag{5.9}$$

Upon substituting (5.3) we obtain the infimum given by

$$\begin{aligned} e_7(\alpha, \tau) &= \tfrac{27}{8}\tau^5(1 - 2\alpha)^2 \cdot \{32\alpha^2\tau^2 + [8\tau^2 - 80\tau]\alpha + [3\tau^2 - 20\tau + 60]\}^2 \\ &\qquad \text{for } (10 - 3\tau)/4\tau \leqq \alpha \leqq (10 - \tau)/8\tau \\ &= .864(1 - 2\alpha)^3(1 - \alpha\tau)^{-1} \qquad \text{for } \alpha \geqq (10 - \tau)/8\tau. \end{aligned} \tag{5.10}$$

TABLE 4

$k =$	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	2.00
$\inf_{F \epsilon \mathfrak{F}} e_7(k)$	.65	.85	.87	.90	.91	.91	.90	.88	.86

TABLE 5*

	Normal		Rectangular		Laplace		Cauchy	
$e_1(\alpha)$	.995	.970	.960	.833	1.06	1.21		
$e_2(\alpha)$	.997	.980	.986	.971	1.01	1.05		
$e_3(\alpha)$	.960	.985	1.040	1.200	1.41	1.24	6.72	2.67
$e_5(\alpha)$	.958	.974	1.010	1.030	1.48	1.425	8.35	5.88

* The first column under each distribution is for $\alpha = .01$, the second for $\alpha = .05$.

TABLE 6*

	Normal		Rectangular		Laplace		Cauchy	
$e_6(k)$	.965	.990	.963	1	1.31	1.18		
$e_7(k)$	.990	.964	1.040	1	1.14	1.29	1.35	2.06

* The first column under each distribution is for $k = 1.5$, the second for $k = 2.0$.

As before, we remark that the right side of (5.10) still provides a lower bound, though not the infimum if $\alpha \leqq (10 - 3\tau)/4\tau$, $\tilde{c} \geqq 3$.

Upon minimizing the expressions given in (5.10) we find that the first is minimized for α given in (5.7) which lies in the given range for $\tau \leqq 7.415$ and by $\alpha(\tau) = 0$ for $7.415 \leqq \tau \leqq 10$. The second is minimized by $\alpha(\tau) = (6 - \tau)/4\tau$ for $\tau \leqq 6$ and by $\alpha(\tau) = 0$ for $\tau \geqq 6$. Upon comparing the expressions so gotten we obtain the given result.

Table 4 gives the values of $\inf_{F \epsilon \mathfrak{F}} e_7(\tau)$ for various values of k.

Finally, we give Tables 5 and 6 which give actual values of e_1, e_2, e_3 and e_5 for selected values of α and e_6, e_7 for selected values of k and particular classical underlying distributions.

The conclusion of this investigation would seem to be that whereas all the proposed "nonparametric" estimates of location behave satisfactorily when compared to the mean, with the possible exception of the Winsorized mean, the Hodges-Lehmann estimate M would seem to be the "safest" among them.

I should like to thank Professor E. L. Lehmann for many valuable comments. I am also indebted to Professors R. Elashoff and R. Purves for helpful discussions.

REFERENCES

[1] Bickel, P. J. (1964). Some contributions to the theory of order statistics. (Unpublished).
[2] Cramér, H. (1956). *Mathematical Methods of Statistics*. Princeton Univ. Press.

[3] ELASHOFF, R. M. (1964). Robust estimation. Unpublished monograph.

[4] HARRIS, T. E. and TUKEY, J. W. (1949). Sampling from contaminated distributions (3). *Statist. Research Group.* Princeton Univ. (Unpublished).

[5] HODGES, J. L. JR. and LEHMANN, E. L. (1956). Efficiency of some nonparametric competitors of the t-test. *Ann. Math. Statist.* **27** 324–335.

[6] HODGES, J. L. JR. and LEHMANN, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.

[7] HUBER, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–102.

[8] SETHURAMAN, J. (1961). Some limit distributions connected with fractile graphical analysis. *Sankhyā Ser. A* **23** 79–90.

[9] TUKEY, J. W. (1949). Memorandum Reports 31–34. *Statist. Research Group,* Princeton Univ. (unpublished).