The Annals of Statistics
1978, Vol. 6, No. 3, 610–618

ESTIMATES OF LOCATION: A LARGE DEVIATION COMPARISON[1]

By Gerald L. Sievers

Western Michigan University

This paper considers the estimation of a location parameter θ in a one-sample problem. The asymptotic performance of a sequence of estimates $\{T_n\}$ is measured by the exponential rate of convergence to 0 of

$$\max\{P_\theta(T_n < \theta - a), P_\theta(T_n > \theta + a)\}, \quad \text{say} \quad e(a).$$

This measure of asymptotic performance is equivalent to one considered by Bahadur (1967). The optimal value of $e(a)$ is given for translation invariant estimates. Some computational methods are reviewed for determining $e(a)$ for a general class of estimates which includes M-estimates, rank estimates and Hodges-Lehmann estimates. Finally, some numerical work is presented on the asymptotic efficiencies of some standard estimates of location for normal, logistic and double exponential models.

1. Introduction. Consider the problem of estimating a location (shift) parameter in a one-sample context. Let $X_1, X_2, \cdots$ denote a sequence of independent random variables having a common absolutely continuous distribution with cdf $F(x-\theta)$ and density $f(x-\theta)$. For each $n = 1, 2, \cdots$ let $T_n = T_n(X_1, \cdots, X_n)$ denote an estimate of θ. One measure of the performance of T_n is $P_\theta(|T_n - \theta| > a)$, the probability of the error of estimation exceeding a fixed number $a > 0$. A similar measure is given by the inaccuracy function

$$A_n(a, \theta, T_n) = \max\{P_\theta(T_n < \theta - a), P_\theta(T_n > \theta + a)\}.$$

This measure has been considered by Huber (1968, 1972) in the development of M-estimates.

If the sequence $\{T_n\}$ is consistent for θ, the inaccuracy tends to 0 as $n \to \infty$. Thus, to measure the asymptotic performance of $\{T_n\}$, consider

$$\lim_{n\to\infty}(-1/n)\log A_n(a, \theta, T_n) = e, \quad \text{say},$$

if this limit exists. The number e will be referred to as the inaccuracy rate of $\{T_n\}$. The inaccuracy rate e is identical to $e^* = \lim_{n\to\infty}(-1/n)\log P_\theta(|T_n - \theta|) > a)$, an asymptotic measure proposed by Bahadur (1967, 1971).

Huber (1968) has shown that an M-estimate is the translation invariant estimate which minimizes the inaccuracy for any finite n. Thus a sequence of M-estimates attains the largest possible inaccuracy rate in the class of sequences of translation invariant estimates. This work is explored in Section 2 and an upper bound on

Received November 1975; revised July 1977.

[1] This research was carried out during the author's sabbatical leave with the Department of Statistics, University of California, Berkeley.

AMS 1970 *subject classifications.* Primary 62G35; Secondary 60F10, 62G20.

Key words and phrases. Location parameter, large deviations, asymptotic efficiency, M-estimates.

e for translation invariant estimates is given. Bahadur (1960, 1967) implicitly and Fu (1975) explicitly have given an upper bound for e^* (or e) for consistent sequences of estimates as

$$e^* \leqq \inf\{\int f(x-\theta')\log(f(x-\theta')/f(x-\theta))\,dx : |\theta-\theta'| > a\}.$$

The Bahadur bound, involving a larger class of estimates, cannot be less than the bound for translation invariant estimates. In fact, the Bahadur bound may be too large in the sense that for some distributions F, it is not attained by any reasonable estimate. For instance, the Bahadur bound is attained by the sample mean in the normal case but does not appear to be attained by any familiar estimate in the logistic or double exponential case.

In this paper, the problem of evaluating and comparing inaccuracy rates for sequences of translation invariant estimates is considered. The optimal inaccuracy rate is given in Section 2 and a general approach for M-estimates and related estimates is discussed in Section 3. Some useful tools are to be found in Chernoff's (1952) paper. Finally, Section 4 contains some numerical work comparing the inaccuracy rates of the sample mean, the sample median, the maximum likelihood estimate and two M-estimates for the cases of normal, logistic and double exponential distributions.

2. The optimal inaccuracy rate. In this section it is shown that under quite general conditions, the optimal inaccuracy rate for translation invariant estimates is attained by a sequence of M-estimates.

Fix $a > 0$ and let $F_0(x) = F(x+a)$ and $F_1(x) = F(x-a)$. Let the ratio of the densities be $c(x) = f_1(x)/f_0(x)$. For testing the hypotheses $H_0: F_0$ vs. $H_1: F_1$, consider the test function

$$\phi_n(x_1, \cdots, x_n) = 1\,, \quad \gamma_n \quad \text{or} \quad 0 \qquad \text{if} \quad \prod_{i=1}^n c(x_i) > k_n\,, \quad = k_n\,, \quad \text{or} \quad < k_n\,,$$

respectively, where k_n and γ_n are chosen so that $E_{F_0}(\phi_n) = E_{F_1}(1-\phi_n) = \alpha_n$, say. Note that $\alpha_n = \inf_{\psi_n} \max\{E_{F_0}(\psi_n), E_{F_1}(1-\psi_n)\}$, where $\psi_n(x_1, \cdots, x_n)$ denotes an arbitrary test function.

Assuming that $c(x)$ is nondecreasing, an M-estimate of θ can be defined as follows. Let $l_n(\theta) = \prod_{i=1}^n c(x_i - \theta)$ and define $\theta_n' = \inf\{\theta : l_n(\theta) \leqq k_n\}$, $\theta_n'' = \sup\{\theta : l_n(\theta) \geqq k_n\}$, and

$$\begin{aligned} \hat{\theta}_n &= \theta_n' \qquad \text{with probability} \quad 1-\gamma_n \\ &= \theta_n'' \qquad \text{with probability} \quad \gamma_n\,. \end{aligned}$$

It follows that θ_n', θ_n'' and the estimate $\hat{\theta}_n$ are translation invariant.

As in Huber (1968), if $c(x)$ is nondecreasing it will follow that $\hat{\theta}_n$ attains the minimum inaccuracy within the class of translation invariant estimates of θ; that is, $A_n(a, \theta, \hat{\theta}_n) = \alpha_n$ and $A_n(a, \theta, T_n) \geqq \alpha_n$ for any translation invariant T_n. The specific inaccuracy rate is given in the following theorem.

THEOREM 2.1. *If $c(x)$ is nondecreasing, then*

(i) $\lim_{n\to\infty} (-1/n) \log A_n(a, \theta, \hat{\theta}_n) = -\log M$, *where* $M = \inf_{0<t<1} m(t)$ *and* $m(t) = \int f_1^t f_0^{1-t}\, dx$, *and*

(ii) $\limsup_{n\to\infty} (-1/n) \log A_n(a, \theta, T_n) \leqq -\log M$ *for any sequence of translation invariant estimates* $\{T_n\}$.

PROOF. Let $\psi_n(x_1, \cdots, x_n)$ denote an arbitrary test function for testing H_0: F_0 vs. H_1: F_1. Then Chernoff (1952) has shown that $(-1/n) \log [\inf_{\psi_n} \{E_{F_0}(\psi_n) + E_{F_1}(1 - \psi_n)\}] \to -\log M$ as $n \to \infty$. Then part (i) follows since $A_n(a, \theta, \hat{\theta}_n) = \alpha_n$ and $\alpha_n \leqq \inf_{\psi_n} \{E_{F_0}(\psi_n) + E_{F_1}(1 - \psi_n)\} \leqq 2\alpha_n$. Part (ii) is immediate.

REMARK 2.1. If $m(t) = m(1 - t)$, $0 < t < 1$, it follows from the convexity of $m(t)$ that $M = m(\frac{1}{2})$. This property holds for instance if the density $f(x)$ is symmetric about 0.

3. Inaccuracy rates for a general class of estimates. Let a sequence of functions $\{H_n(x_1, \cdots, x_n)\}$ be given satisfying the property: $h_n(\theta) \equiv H_n(x_1 - \theta, \cdots, x_n - \theta)$ is nonincreasing in θ. Let $\{k_n\}$ and $\{\gamma_n\}$ be sequences of constants, $0 \leqq \gamma_n \leqq 1$. For each n, define an estimate $\theta_n^* = \theta_n^*(X_1, \cdots, X_n)$ by $\theta_n' = \inf\{\theta : h_n(\theta) \leqq k_n\}$, $\theta_n'' = \sup\{\theta : h_n(\theta) \geqq k_n\}$, and

$$\begin{aligned} \theta_n^* &= \theta_n' \qquad \text{with probability} \quad 1 - \gamma_n \\ &= \theta_n'' \qquad \text{with probability} \quad \gamma_n . \end{aligned}$$

Note that θ_n', θ_n'' and θ_n^* are translation invariant and have continuous cdf's. Various M-estimates follow this prescription using $H_n(x_1, \cdots, x_n) = \sum_{i=1}^n \psi(x_i)$ for some nondecreasing function ψ. Special cases include the optimal estimate $\hat{\theta}_n$ of Section 2 and the maximum likelihood estimate when the density f is log concave ($\psi = -f'/f$ and $k_n \equiv 0$). Rank estimates of θ are also obtainable by choosing H_n to be a suitable function of the ranks of $|x_i|$.

The computation of the inaccuracy rate for $\{\theta_n^*\}$ is a nontrivial problem in general. One very basic approach is to note that if $h_n(\theta)$ is nonincreasing in θ, then

$$P_{F_0}(h_n(0) > k_n) \leqq P_{F_0}(\theta_n^* \geqq 0) = P_{F_0}(\theta_n^* > 0) \leqq P_{F_0}(h_n(0) \geqq k_n)$$

and

$$P_{F_1}(h_n(0) < k_n) \leqq P_{F_1}(\theta_n^* \leqq 0) = P_{F_1}(\theta_n^* < 0) \leqq P_{F_1}(h_n(0) \leqq k_n) .$$

The following theorem is then immediate.

THEOREM 3.1. *If* $h_n(\theta)$ *is nonincreasing in* θ *and*

$$\begin{aligned} \lim_{n\to\infty} (-1/n) \log P_{F_0}(h_n(0) > k_n) &= \lim_{n\to\infty} (-1/n) \log P_{F_0}(h_n(0) \geqq k_n) \\ &= e^+ , \quad \textit{say}, \qquad \textit{and} \end{aligned} \tag{3.1}$$

$$\begin{aligned} \lim_{n\to\infty} (-1/n) \log P_{F_1}(h_n(0) < k_n) &= \lim_{n\to\infty} (-1/n) \log P_{F_1}(h_n(0) \leqq k_n) \\ &= e^- , \quad \textit{say}, \end{aligned} \tag{3.2}$$

then the inaccuracy rate of $\{\theta_n^*\}$ *is* $\min\{e^+, e^-\}$.

The equality of the limits in (3.1) and (3.2) may be easy to verify in special cases, but such limits do not necessarily agree in general without some qualification. Various techniques in large deviations theory may be useful in the computation of the limits in (3.1) and (3.2).

REMARK 3.1. For the special case $h_n(0) = \sum_{i=1}^n \psi(x_i)$, for some nondecreasing function ψ, Chernoff's theorem (see Bahadur (1971), page 7) gives expressions for e^+ and e^-.

REMARK 3.2. The estimate θ_n^* was defined as a random choice of either θ_n' or θ_n'' and this randomization is frowned upon in practical circles. Instead, the estimate can be defined differently, for instance as $(\theta_n' + \theta_n'')/2$, without affecting the inaccuracy rate. To be specific, if $\bar{\theta}_n$ is any translation invariant estimate with $\theta_n' \leqq \bar{\theta}_n \leqq \theta_n''$, and if the hypothesis of Theorem 3.1 holds, then $\{\bar{\theta}_n\}$ has the same inaccuracy rate as $\{\theta_n^*\}$.

4. Applications. In this section a comparison of inaccuracy rates is made for selected estimates in the context of normal, logistic and double exponential distributions. If $\{T_n^{(i)}\}$ has inaccuracy rate $e_i(a)$ for $i = 1, 2$, then the asymptotic relative efficiency of $\{T_n^{(1)}\}$ to $\{T_n^{(2)}\}$ is defined to be $e_{1,2}(a) = e_1(a)/e_2(a)$. This efficiency can be directly interpreted as the ratio of respective (large) sample sizes required for the two estimates to have identical inaccuracies.

In what follows the underlying distribution will be indicated by a second subscript: N for normal, L for logistic and D for double exponential. Using Theorem 2.1 and Remark 2.1, the optimal inaccuracy rates are

$$e_{0,N}(a) = a^2/2\ ,$$

$$e_{0,L}(a) = a + \log((1 - \exp(-2a))/2a)\ ,$$

and

$$e_{0,D}(a) = a - \log(1 + a)\ .$$

Consider the sample mean $\bar{X}_n = \sum_{i=1}^n X_i/n$. Since the optimal estimate $\hat{\theta}_n$ of Section 2 is $\bar{X}_n$ in the normal case,

$$e_{\bar{X},N}(a) = a^2/2\ .$$

For the other distributions, the symmetry of f and Chernoff's theorem can be used to yield $e_{\bar{X}}(a) = -\log[\inf_{t\geqq 0} E_F \exp(t(X - a)))]$. Then using the well known moment-generating functions

$$e_{\bar{X},L}(a) = -\log[\inf_{t\geqq 0} t\pi \exp(-at)/\sin(t\pi)]$$

and

$$e_{\bar{X},D}(a) = ((1 + a^2)^{\frac{1}{2}} - 1) + \log[2((1 + a^2)^{\frac{1}{2}} - 1)/a^2]\ .$$

Consider next the sample median $M_n = \text{median}\{X_1, \cdots, X_n\}$. Bahadur (1971) page 25, shows that

$$\begin{aligned} e^* &= \lim_{n\to\infty} (-1/n) \log P_\theta(|M_n - \theta| > a) \\ &= -(\tfrac{1}{2}) \log(4p(1 - p)) \end{aligned}$$

where $p = P_F(X_1 > a)$. Since the inaccuracy rate equals e^*, it follows that

$$e_{M,N}(a) = -(\tfrac{1}{2}) \log [4\Phi(a)(1 - \Phi(a))]$$

where Φ is the standard normal cdf,

$$e_{M,L}(a) = (a/2) + \log [(1 + \exp(-a))/2]$$

and

$$e_{M,D}(a) = (-\tfrac{1}{2}) \log [\exp(-a)(2 - \exp(-a))] .$$

The inaccuracy rates for maximum likelihood (ML) estimates have already been given for the normal ($\bar{X}_n$) and the double exponential distribution (M_n) cases. For the logistic case, the ML estimate is the solution θ to the equation $\sum_{i=1}^n (1/(1 + \exp(-x_i + \theta))) = n/2$. Using Theorem 3.1, Remark 3.1 and the symmetry of f, it follows that

$$e_{ML,L}(a) = -\log [\inf_{t \geqq 0} \exp(-t/2) \textstyle\int_{-\infty}^{\infty} \exp(t/(1 + \exp(-x + a)))f(x)\, dx] .$$

As a final estimate, consider the M-estimate of Huber defined as the solution θ to the equation $\sum_{i=1}^n \psi(x_i - \theta) = 0$ with $\psi(x) = \max\{-k, \min\{k, x\}\}$ for some $k > 0$. This estimate is considered "in between" $\bar{X}_n$ and M_n in its sensitivity to extreme observations and has properties similar to a trimmed mean. The optimal estimate of Section 2 reduces to this estimate in the double exponential case if $k = a$ is used. Using Theorem 3.1, Remark 3.1 and the symmetry of f, the inaccuracy rate of this estimate is

$$e_H(a) = -\log [\inf_{t \geqq 0} l(t)] ,$$

where $l(t) = E_F(\exp(t\psi(X - a)))$. In particular

$$\begin{aligned} l_N(t) &= \exp(-kt)\Phi(a - k) + \exp(kt)(1 - \Phi(a + k)) \\ &\quad + \exp(-at + t^2/2)(\Phi(a + k - t) - \Phi(a - k - t)) , \\ l_L(t) &= \exp(-kt)/(1 + \exp(-a + k)) + \exp(-a - k - kt)/(1 + \exp(-a - k)) \\ &\quad + \exp(-at) \textstyle\int_{a-k}^{a+k} \exp(tx - x)/(1 + \exp(-x))^2\, dx \end{aligned}$$

and

$$\begin{aligned} l_D(t) &= \exp(-kt) + [t \exp(-a)/2(t - 1)][\exp((t - 1)k) - \exp(-(t - 1)k)] \\ &\qquad\qquad \text{if} \quad k \leqq a \\ &= (t/2(t + 1)) \exp(a - (t + 1)k) + (t/2(t - 1)) \exp(-a + (t - 1)k) \\ &\quad - \exp(-at)/(t^2 - 1) \qquad \text{if} \quad a < k . \end{aligned}$$

Tables 4.1—4.3 give selected values of the above inaccuracy rates and the corresponding efficiencies relative to the optimal rates. Where necessary, the integrals and infimums were computed numerically on a PDP-10 at the Computer Center of Western Michigan University. Figures 4.1—4.3 provide graphs of the efficiencies relative to the optimal rates.

The flatness of the efficiency curves for the normal and logistic cases was unexpected. This indicates that the local behavior of the estimates adequately reflects the overall behavior. Unfortunately, this does not carry over to the double exponential case where the sample median is locally efficient but its efficiency drops off quite fast. The efficiency of the sample mean is adequate for logistic distributions but too low for double exponential distributions. In this latter case, the sample mean becomes more efficient than the sample median for $a > 1.5$. The efficiency of the Huber estimate appears only mildly sensitive to the choice of k, especially for the normal and logistic cases. For the estimates and distributions considered here, the best compromise appears to be the Huber estimate with k chosen between $.5\sigma$ and 1.5σ.

TABLE 4.1
Inaccuracy rates and efficiencies (relative to the optimal rate) for the standard normal case

a	$e_0(a)$	$e_M(a)$	$e_H(a)$ ($k = 1$)	$e_H(a)$ ($k = 2$)
.01	$.(0)_4 5$	$.(0)_4 318$ (.637)	$.(0)_4 452$ (.904)	$.(0)_4 495$ (.990)
.10	.005	.00318 (.636)	.00452 (.903)	.00495 (.990)
.50	.125	.07928 (.634)	.11268 (.901)	.12369 (.990)
1.00	.50	.31374 (.627)	.44805 (.896)	.49443 (.989)
1.50	1.125	.69440 (.617)	.99765 (.887)	1.11110 (.988)
2.00	2.00	1.20995 (.605)	1.74615 (.873)	1.97121 (.986)
2.50	3.125	1.85079 (.592)	2.67117 (.855)	3.06979 (.982)
3.00	4.50	2.61139 (.580)	3.74633 (.833)	4.39670 (.977)

TABLE 4.2
Inaccuracy rates and efficiencies (relative to the optimal rate) for the logistic case

a	$e_0(a)$	$e_M(a)$	$e_{\bar{X}}(a)$	$e_{ML}(a)$	$e_H(a)$ ($k = 1$)	$e_H(a)$ ($k = 3$)
.01	$.(0)_4 167$	$.(0)_4 125$ (.750)	$.(0)_4 152$ (.912)	$.(0)_4 167$ (1.00)	$.(0)_4 156$ (.939)	$.(0)_4 163$ (.979)
.10	.00167	.00125 (.750)	.00152 (.912)	.00167 (1.00)	.00156 (.939)	.00163 (.979)
.50	.04132	.03093 (.749)	.03771 (.913)	.04132 (1.00)	.03869 (.936)	.04053 (.981)
1.00	.16144	.12011 (.744)	.14763 (.914)	.16137 (1.00)	.14981 (.928)	.15894 (.985)
1.50	.35032	.25827 (.737)	.32143 (.918)	.34962 (.998)	.32046 (.915)	.34663 (.989)
2.00	.59522	.43378 (.729)	.54842 (.921)	.59191 (.994)	.53432 (.898)	.59151 (.994)
2.50	.88380	.63574 (.719)	.81814 (.926)	.87347 (.988)	.77603 (.878)	.87997 (.996)
3.00	1.20576	.85544 (.709)	1.12163 (.930)	1.18096 (.979)	1.03369 (.857)	1.19832 (.994)

TABLE 4.3
Inaccuracy rates and efficiencies (relative to the optimal rate) for the double exponential case

a	$e_0(a)$	$e_M(a)$	$e_{\bar{X}}(a)$	$e_H(a)$ ($k = 1$)	$e_H(a)$ ($k = 2.5$)
.01	$.(0)_4497$	$.(0)_4495$ (.997)	$.(0)_4250$ (.503)	$.(0)_4378$ (.761)	$.(0)_4296$ (.595)
.10	.00469	.00455 (.970)	.00250 (.532)	.00377 (.804)	.00295 (.630)
.50	.09453	.08410 (.890)	.06069 (.642)	.08918 (.943)	.07196 (.761)
1.00	.30685	.25506 (.831)	.22599 (.737)	.30685 (1.00)	.26902 (.877)
1.50	.58371	.46257 (.793)	.46531 (.797)	.56804 (.973)	.55341 (.948)
2.00	.90139	.68846 (.764)	.75486 (.837)	.84073 (.933)	.88987 (.987)
2.50	1.24724	.92438 (.741)	1.07940 (.865)	1.11656 (1.895)	1.24724 (1.00)
3.00	1.61371	1.16603 (.723)	1.42936 (.886)	1.39204 (.863)	1.59823 (.990)

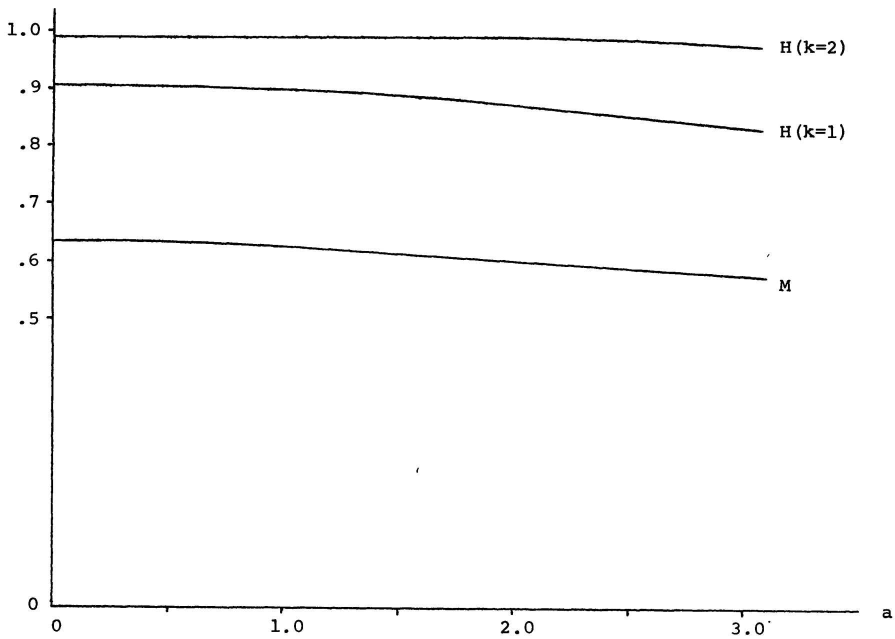

FIG. 4.1. Efficiencies relative to the optimal rate for the standard normal distribution

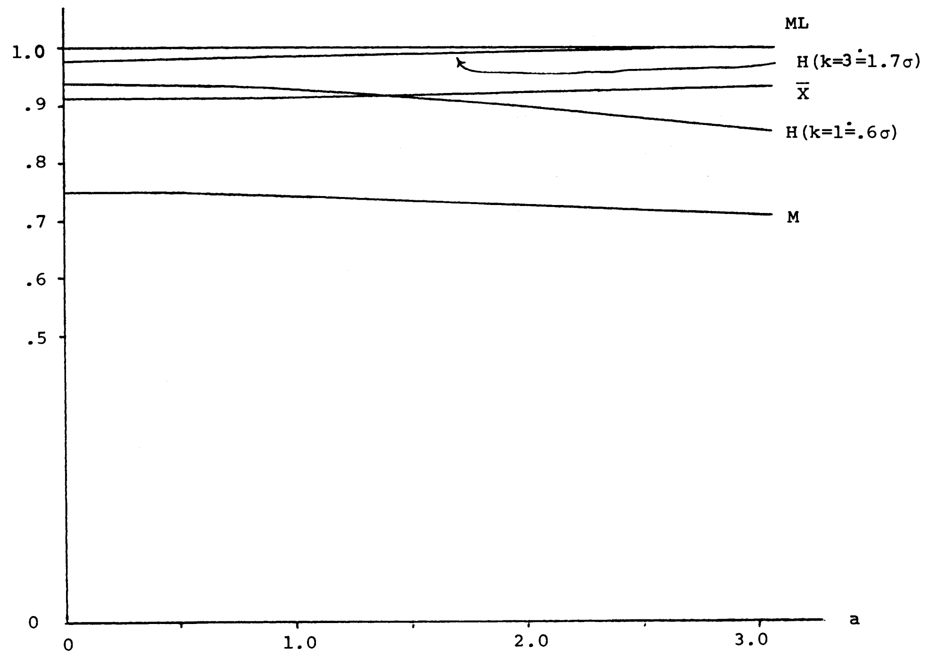

FIG. 4.2. Efficiencies relative to the optimal rate for the logistic distribution

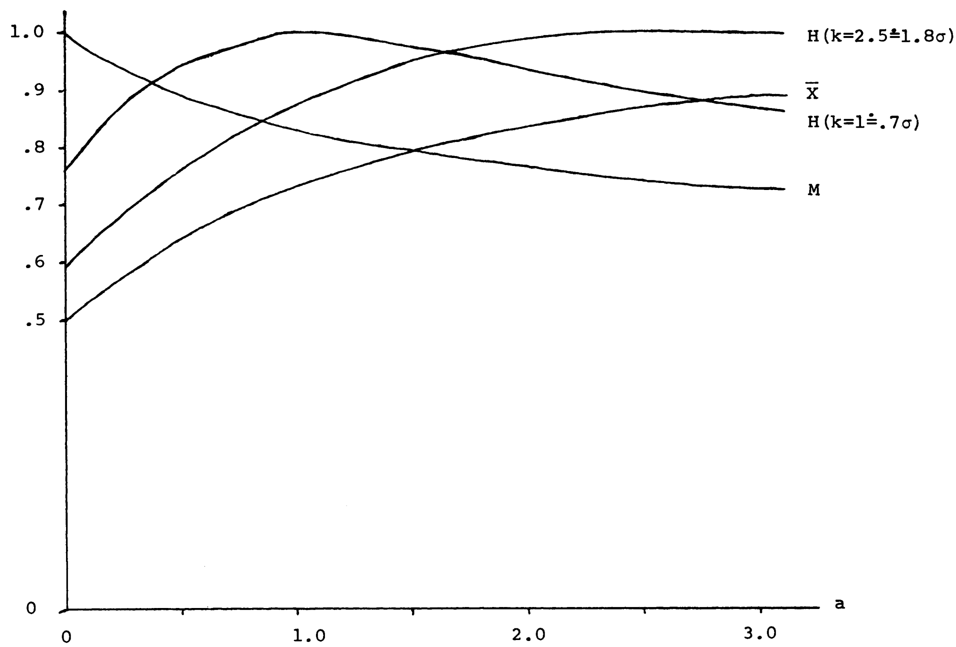

FIG. 4.3. Efficiencies relative to the optimal rate for the double exponential distribution

REFERENCES

[1] BAHADUR, R. R. (1960). Asymptotic efficiency of tests and estimates. *Sankhyā* **22** 229–252.

[2] BAHADUR, R. R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38** 303–324.

[3] BAHADUR, R. R. (1971). *Some Limit Theorems in Statistics.* SIAM, Philadelphia.

[4] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[5] FU, J. C. (1975). The rate of convergence of consistent point estimators. *Ann. Statist.* **3** 234–240.

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

[7] HUBER, P. J. (1968). Robust confidence limits. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **10** 269–278.

[8] HUBER, P. J. (1972). Robust statistics: a review. *Ann. Math. Statist.* **43** 1041–1067.

DEPARTMENT OF MATHEMATICS
WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MICHIGAN 49008