ROBUSTNESS OF THE HODGES-LEHMANN ESTIMATES FOR SHIFT[1]

By Arnljot Høyland

University of Oslo and *University of California, Berkeley*

0. Introduction. Let $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ be $(m + n) = N$ independent observations from continuous distributions

$$P(X_i \leqq u) = F(u) \quad \text{and} \quad P(Y_j \leqq u) = F(u - \Delta), \qquad (i = 1, \cdots, m; j = 1, \cdots, n). \tag{0.1}$$

Two problems often studied in this setup are (i) either to test the hypothesis $\Delta = 0$ against $\Delta > 0 (\Delta \neq 0)$ or (ii) to estimate the shift Δ.

The classical approach to the testing problem is based on the statistic $t = (\bar{Y} - \bar{X})/\{(1/m + 1/n)[\sum_i (X_i - \bar{X})^2 + \sum_j (Y_j - \bar{Y})^2]/(m + n - 2)\}^{\frac{1}{2}}$ known to be approximately normally distributed under general assumptions on F. Similarly the classical estimate for Δ is

$$\hat{\Delta} = \bar{Y} - \bar{X}; \qquad \bar{Y} = \sum_j Y_j/n, \qquad \bar{X} = \sum_i X_i/m. \tag{0.2}$$

Both these methods are known to be vulnerable to gross errors. For the testing problem, rank tests, such as the Wilcoxon and the Normal Scores (Fisher-Yates) tests have been in use for several years and shown to be more robust against gross errors than the classical one. At the same time little efficiency (in the Pitman sense) is lost when after all no gross errors are present and normality assumptions hold.

Similar robust methods for the corresponding estimation problem have recently been proposed by J. L. Hodges, Jr. and E. L. Lehmann [5]. They study small-sample as well as large-sample properties of a large class of such estimates, derived from corresponding test statistics $h(X_1, \cdots, X_m, Y_1, \cdots, Y_n)$ for the hypothesis $\Delta = 0$ against $\Delta > 0$, satisfying the following two conditions:

(A) $h(x_1, \cdots, x_m, y_1 + a, \cdots, y_n + a)$ is a nondecreasing function of a for all x and y, and such that

(B) when $\Delta = 0$, the distribution of $h(X_1, \cdots, X_m, Y_1, \cdots, Y_n)$ is symmetric about a fixed point μ (independent of F),

(i) for all $F \varepsilon \mathfrak{F}_0$, or (ii) for all $F \varepsilon \mathfrak{F}_1$.

(Throughout this paper we shall use the same short notation that was used in [5]. Furthermore $P_{\Delta_0}(\cdot)$, $E_{\Delta_0}(\cdot)$, etc. will be used to indicate that the expression in question is being computed for the case $\Delta = \Delta_0$.)

Here as throughout this paper, $\mathfrak{F}_0$ denotes the class of all continuous distributions and $\mathfrak{F}_1$ the class of all continuous distributions symmetric about zero.

Received 6 February 1964; revised 10 August 1964.

[1] This paper was prepared with the partial support of the Office of Naval Research, Contract Nonr-222-43. This paper in whole or in part may be reproduced for any purpose of the United States Government.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

For suitable functions satisfying (A) and (B)

$$\Delta^{**} = (\Delta' + \Delta'')/2 \tag{0.3}$$

is proposed as an estimate for Δ, where

$$\Delta' = \sup\{\Delta: h(x, y - \Delta) > \mu\} \quad \text{and} \quad \Delta'' = \inf\{\Delta: h(x, y - \Delta) < \mu\}. \tag{0.4}$$

The test statistic of the Mann-Whitney version of the Wilcoxon two-sample test satisfies conditions (A) and (B) with $\mu = mn/2$. The corresponding estimate Δ^{**} is in [5] shown to be the median of the set of mn differences $Y_j - X_i$,

$$\Delta^{**} = \text{med}_{i \leqq j}(Y_j - X_i), \qquad (i = 1, \cdots, m; j = 1, \cdots, n). \tag{0.5}$$

In the same paper Hodges and Lehmann also consider the problem of estimating θ when $Z_1, \cdots, Z_N$ are N independent observations from a distribution $P(Z_i \leqq u) = F(u - \theta)$, $(i = 1, \cdots, N)$ where $F \varepsilon \mathfrak{F}_1$. In this case the estimate is based on a test statistic $h(Z_1, \cdots, Z_N)$ for the hypothesis $\theta = 0$ against $\theta > 0$ with the following properties:

(C) $h(z_1 + a, \cdots, z_N + a)$ is a nondecreasing function of a for each z.

(D) The distribution of h is, when $\theta = 0$, symmetric about a fixed point μ (independent of F) for all $F \varepsilon \mathfrak{F}_1$.

For suitable functions of this type

$$\theta^* = (\theta' + \theta'')/2 \tag{0.6}$$

is proposed as an estimate for θ, where

$$\theta' = \sup\{\theta: h(z - \theta) > \mu\}, \qquad \theta'' = \inf\{\theta: h(z - \theta) < \mu\}. \tag{0.7}$$

The Wilcoxon one-sample test statistic satisfies Conditions (C) and (D) with $\mu = N(N + 1)/4$ for all $F \varepsilon \mathfrak{F}_1$. The corresponding estimate θ^* is in [5] shown to be the median of the $\binom{N}{2} + N$ averages $(Z_i + Z_j)/2$,

$$\theta^* = \text{med}_{i \leqq j}\{(Z_i + Z_j)/2\}. \tag{0.8}$$

An alternative estimate of the shift Δ, based on statistics of the form (0.8) is suggested by Lehmann [8] for the case $F \varepsilon \mathfrak{F}_1$, namely

$$\begin{aligned} \Delta^* = \text{med}_{i \leqq j}\{(Y_i + Y_j)/2\} - \text{med}_{k \leqq l}\{(X_k + X_l)/2\}, \\ (i, j = 1, \cdots, n; k, l = 1, \cdots, m). \end{aligned} \tag{0.9}$$

The asymptotic behavior of the procedures Δ^* (0.9) and Δ^{**} (0.5) is studied in [8] and [5]. There, as throughout this paper, it is assumed that the corresponding pairs of sample sizes $m(N)$ and $n(N)$ tend to infinity in such a way that $m(N)/N \to \lambda$, where λ is a positive number smaller than 1. It is shown that the asymptotic efficiency of Δ^{**} (0.5) relative to $\hat{\Delta}$ (0.2), denoted ARE $(\Delta^{**}, \hat{\Delta})$, in the sense of reciprocal ratio of asymptotic variances, is the same as the ARE $(\Delta^*, \hat{\Delta})$ and equal to $12\sigma_x^2[\int f^2(x)\,dx]^2$, where f denotes the density corresponding to F. This efficiency has been shown [3] always to be $\geqq .864$ and equal to $3/\pi$ when f is normal.

Hence a natural question to ask is the following: Which one of the two estimates, Δ^* (0.9) or Δ^{**} (0.5) should be preferred in a given situation? The choice might be based on the robustness of these estimates against the deviations from the model (0.1) with $F \varepsilon \mathfrak{F}_1$ which are considered most likely to occur.

One assumption of (0.1) which might turn out not to hold, is the assumption that the distribution of the Y's is the same as the distribution of the X's, only shifted by Δ. Another such assumption is the assumption of symmetry, $F \varepsilon \mathfrak{F}_1$, used to justify Δ^*. After establishing a few general results in Section 1, we shall in this paper mainly be concerned with the study of the robustness of Δ^* (0.9) and Δ^{**} (0.5) against these deviations from the model.

1. General remarks.

1.1. *Some definitions and lemmas.* Since it is convenient for our purpose to have a uniquely defined median for all $F \varepsilon \mathfrak{F}_0$, we shall adopt the following definition.

DEFINITION. *The median θ of a distribution $F \varepsilon \mathfrak{F}_0$ is defined as $(\theta_1 + \theta_2)/2$ where $\theta_1 = \inf (\theta: F(\theta) = \frac{1}{2})$ and $\theta_2 = \sup (\theta: F(\theta) = \frac{1}{2})$.*

With this definition the following lemmas are easily seen to be true and are given here only for the purpose of convenient reference.

LEMMA 1.1. *Let θ be the median of $F \varepsilon \mathfrak{F}_0$. Assume F to be symmetric about ζ,* i.e., *$F(\zeta - x) + F(\zeta + x) = 1$. Then θ coincides with the point of symmetry ζ.*

LEMMA 1.2. *Let $X_1, \cdots, X_n$ be independent random variables with symmetric distributions $F_1, \cdots, F_n$ and medians $\theta_1, \cdots, \theta_n$ respectively. Then for any set of real constants $a_1, \cdots, a_n$, $Z = \sum_i a_i X_i$ is symmetrically distributed with median $\sum_i a_i \theta_i$.*

1.2. *Results from* [5] *that are true in more general models.* Let $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ be $m + n = N$ independent random variables with distributions

$$P(X_i \leqq u) = F(u - \zeta) \quad \text{and} \quad P(Y_j \leqq u) = G(u - \zeta - \Delta), \qquad (i = 1, \cdots, m; j = 1, \cdots, n). \tag{1.1}$$

Consider the following four models, where $\mathfrak{F}_{01}$ denotes the class of all continuous distributions with median zero, and $\mathfrak{F}_1$ the class of all continuous distributions symmetric about zero.

Model I: $F \varepsilon \mathfrak{F}_{01}$, $G \varepsilon \mathfrak{F}_{01}$.
Model II: $F \varepsilon \mathfrak{F}_{01}$, $G \varepsilon \mathfrak{F}_{01}$, $F = G$.
Model III: $F \varepsilon \mathfrak{F}_1$, $G \varepsilon \mathfrak{F}_1$.
Model IV: $F \varepsilon \mathfrak{F}_1$, $G \varepsilon \mathfrak{F}_1$, $F = G$.

In all four models F, G, ζ and Δ are assumed unknown, and the problem is to estimate Δ. *The "shift" Δ between two distributions is thus defined to be the difference between their medians.*

Let $h(X, Y)$ be a test statistic satisfying (A), and Δ^{**}, Δ' and Δ'' be defined as in (0.3) and (0.4) with μ denoting a constant independent of F and G, not necessarily a symmetry point for $h(X, Y)$ when $\Delta = 0$.

Similarly let $h(Z)$ be a test statistic satisfying (C), and θ^*, θ' and θ'' be defined as in (0.6) and (0.7) with μ denoting any constant independent of F, not necessarily a symmetry point for $h(Z)$ when $\theta = 0$. Then Δ^{**} and θ^* are easily shown to have the same regularity and invariance properties in Model I as shown for Model IV ([5], Sections 6 and 7). Furthermore, the following lemma and theorem can be shown to be true with only slight modifications of the proofs given in [5].

LEMMA 1.3. *The distribution of* $h(X, Y)$ *is symmetric about* μ *for* $\Delta = 0$, *if the following three conditions hold:*

$$h(x + a, y + a) = h(x, y), \qquad \text{for all real } a, \tag{1.2}$$

$$h(x, y) + h(-x, -y) = 2\mu, \qquad \text{(a.e. } P_0) \tag{1.3}$$

$$F \varepsilon \mathfrak{F}_1, \qquad G \varepsilon \mathfrak{F}_1 . \tag{1.4}$$

THEOREM 1.1 (Hodges and Lehmann). *Let the test function h satisfy* (1.2) *and* (1.3). *Then the distribution of* Δ^{**}, *defined by* (0.3) *and* (0.4) *with* μ *being the value from* (1.3), *is symmetric about* Δ *if* $F \varepsilon \mathfrak{F}_1$, $G \varepsilon \mathfrak{F}_1$.

Let us now consider the class of estimates Δ^{**} obtained from

$$h(X, Y) = \textstyle\sum_{j=1}^{n} E_\psi[V^{(s_j)}], \tag{1.5}$$

where $s_1, \cdots, s_n$ denote the ranks of $Y_1, \cdots, Y_n$ in the combined sample and where $V^{(1)}, \cdots, V^{(m+n)}$ denote an ordered sample from a distribution Ψ. Then Lemma 2(i) in [5] is still true when $F \varepsilon \mathfrak{F}_1$, $G \varepsilon \mathfrak{F}_1$, and we have the following result.

THEOREM 1.2. *Let* $h(X, Y)$ *be defined by* (1.5) *with* Ψ *symmetric. Then* Δ^{**} *defined by* (0.3) *and* (0.4) *is symmetrically distributed about* Δ *in Models* III *and* IV *and, typically, approximately median unbiased in Model* II.

We notice that Theorem 1.2 in particular applies when $h(X, Y)$ is the Wilcoxon two-sample test statistic and also when $h(X, Y)$ is the Normal Scores test statistic.

As indicated in the Introduction, estimates for shift may easily be constructed on the basis of estimates of the form (0.8) in the following way. Let $\theta_G{}^*$ and $\theta_F{}^*$ be (0.6)-estimates of the medians of G and F, both belonging to $\mathfrak{F}_1$, and define Δ^* by

$$\Delta^* = \theta_G{}^* - \theta_F{}^*. \tag{1.6}$$

Let $Z_1, \cdots, Z_N$ be independently and identically distributed with distribution $F \varepsilon \mathfrak{F}_{01}$ and $s_1, \cdots, s_n$ denote the ranks of the positive Z's among the absolute values $|Z_1|, \cdots, |Z_N|$. Consider the statistic

$$h(Z) = \textstyle\sum_{j=1}^{n} E_\psi[V^{(s_j)}], \tag{1.7}$$

where $V^{(1)}, \cdots, V^{(N)}$ denote the ordered absolute values of a sample of size N from a distribution Ψ. In [5] it is shown that h then satisfies (C) and (D) with $\mu = \frac{1}{2}NE_\psi|Z_1|$ for all $F \varepsilon \mathfrak{F}_1$ and furthermore that the corresponding θ^* (0.6) is symmetrically distributed about θ. By Lemma 1.2 Δ^* is then symmetrically distributed about Δ when $F \varepsilon \mathfrak{F}_1$ and $G \varepsilon \mathfrak{F}_1$, and we have the following result.

THEOREM 1.3. *Let $h(Z)$ be defined by* (1.5) *with* Ψ *symmetric. Then* Δ^* *defined by* (1.6) *is symmetrically distributed about* Δ *in Models* III *and* IV.

2. Asymptotic properties of the estimates Δ^{**} (0.5) and Δ^* (0.9).

2.1. *Introduction.* In the remaining part of the paper we shall restrict attention to the estimates (0.9), (0.5) and (0.2),

$$\Delta^* = \operatorname{med}_{i \leqq j}\{(Y_i + Y_j)/2\} - \operatorname{med}_{k \leqq l}\{(X_k + X_l)/2\}, \tag{2.1}$$

$$\Delta^{**} = \operatorname{med}_{i \leqq j}\{Y_j - X_i\}, \tag{2.2}$$

and

$$\hat{\Delta} = \bar{Y} - \bar{X} \tag{2.3}$$

and study their behavior in Models I to IV. A first natural question to ask is in which of these models Δ^*, Δ^{**} and $\hat{\Delta}$ are reasonable estimates of Δ. We shall say that an estimate is reasonable if its distribution or asymptotic distribution in some sense is centered on the corresponding parameter.

TABLE 1a

Estimate	Model I	Model II	Model III	Model IV
$\hat{\Delta}$	Difference between means	Δ	Δ	Δ
Δ^*	?	?	Δ	Δ
Δ^{**}	?	Δ	Δ	Δ

As for $\hat{\Delta}$, its expected value is $EY - EX$ when these exist. We shall assume they do. This difference coincides with Δ in Models II, III and IV, but may be very different from Δ in Model I. Utilizing the results of Theorems 1.2 and 1.3 we have so far the following table of the quantities (Table 1a) of which Δ^*, Δ^{**} and $\hat{\Delta}$ are consistent estimates.

Our first task will be to complete this table.

2.2. *Some definitions and lemmas.*

DEFINITION. *The pseudomedian of a distribution (occasionally we shall say pseudomedian of X instead of pseudomedian of the distribution of X) $F \varepsilon \mathfrak{F}_0$ is defined as the median of the distribution of $(X_1 + X_2)/2$ where X_1 and X_2 are independently and identically distributed according to F.*

By this definition and Lemma 1.2, the following results are immediate.

(i) *The median and the pseudomedian coincide when F is symmetric. Furthermore*

(ii) *if $F \varepsilon \mathfrak{F}_0$ and $G \varepsilon \mathfrak{F}_0$ have pseudomedians η_F and η_G respectively, then $G(x) = F(x - \Delta)$ implies $\eta_G = \eta_F + \Delta$.*

(ii) implies that the pseudomedian of $F \varepsilon \mathfrak{F}_0$ may be used as an ordinary location parameter.

LEMMA 2.1. *Let $K \varepsilon \mathfrak{F}_0$ and have pseudomedian zero. The corresponding density* *is k. Then*

(i) $\int K(-x)k(x)\,dx = \frac{1}{2}$

(ii) $\int k(-x)k(x)\,dx \leqq \int k^2(x)\,dx$, *with equality if and only if K is symmetric, and*

(iii) $\int K^2(-x)k(x)\,dx \geqq \frac{1}{3}$, *with equality if and only if K is symmetric.*

PROOF. (i) follows from the fact that the median of the convolution $K*K$, by definition is zero. The first part of (ii) follows by the Schwarz inequality. As for the second part, $\int k(x)k(-x)\,dx = \int k^2(x)\,dx$ implies $\int k(x)k(-x)\,dx = \int k^2(-x)\,dx$ and hence $\int [k(x) - k(-x)]^2\,dx = 0$, which in turn implies $k(x) = k(-x)$ a.e. The first part of (iii) follows by the following argument:

$$\int [K(x) - (1 - K(-x))]^2 k(x)\,dx \geqq 0. \tag{2.4}$$

(2.4) may also be written

$$\int K^2(-x)k(x)\,dx + 2\int K(x)K(-x)k(x)\,dx \geqq \tfrac{2}{3}. \tag{2.5}$$

If we integrate the last integral of (2.5) by parts, we get

$$2\int K(x)K(-x)k(x)\,dx = \int K^2(-x)k(x)\,dx \tag{2.6}$$

and the result follows. From (2.4) to (2.6) we see that $\int K^2(-x)k(x)\,dx = \frac{1}{3}$ implies

$$\int [K(x) - (1 - K(-x))]^2 k(x)\,dx = 0. \tag{2.7}$$

Hence the second part of (iii) will follow if we can prove that (2.7) implies $K(x) = 1 - K(-x)$. This may however be proved exactly as the second part of Lemma 4.1 in [7].

LEMMA 2.2. *Let X and Y be independent random variables with distributions F and G, $F \varepsilon \mathfrak{F}_0$, $G \varepsilon \mathfrak{F}_0$ and densities f and g respectively. Then*

(i) $\int F^2 g\,dx \geqq [\int Fg\,dx]^2$,

(ii) $\int F^2 g\,dx + \int G^2 f\,dx \geqq \frac{2}{3}$.

(iii) *In particular, if the median of* $(Y - X) = 0$, i.e., $\int Fg\,dx = \frac{1}{2}$, *then* $\lambda\int F^2 g\,dx + (1 - \lambda)\int G^2 f\,dx > \frac{1}{4}$, *where* $0 < \lambda < 1$.

PROOF. (i) follows by the Schwarz inequality. (ii) follows from

$$\int (F - G)^2 (f + g)\,dx \geqq 0$$

by evaluation and integration by parts. (iii) $\int Fg\,dx = \frac{1}{2}$ implies $\int F^2 g\,dx \geqq \frac{1}{4}$, $\int G^2 f\,dx \geqq \frac{1}{4}$ such that $\lambda\int F^2 g\,dx + (1 - \lambda)\int G^2 f(x)\,dx \geqq \frac{1}{4}$. A necessary condition for equality is $\int F^2 g\,dx = \int G^2 f\,dx = \frac{1}{4}$ which violates (ii). q.e.d.

As in [5] we shall study the asymptotic behavior of our procedures as $N \to \infty$. The test functions, estimates and constants corresponding to sample size N will be denoted h_N, θ_N^*, μ_N, etc. To make the passages to the limit under the integral sign permissible, we will from now on assume the distributions considered to satisfy the regularity conditions of Lemma 3a in [4]. Such a distribution (F) is thus assumed;

R_1 : to be continuous,

(2.8) R_2 : to be differentiable in each of the open intervals $(-\infty, a_1)$, (a_1, a_2), $\cdots$, (a_{s-1}, a_s), (a_s, ∞),

R_3 : and to have a bounded derivative in each of these intervals.

The following theorem from [5] is true under the more general assumptions $F \in \mathfrak{F}_{01}$, $G \in \mathfrak{F}_{01}$, with μ_N being any constant independent of F and G, since the proof given in [5] does not utilize the assumptions $F = G$, $F \in \mathfrak{F}_1$ and μ_N being a symmetry point of h_N when $\Delta = 0$.

THEOREM 2.1 (Hodges and Lehmann). *Let a, c_1, $\cdots$ be real constants and let $\Delta_N = -a/c_N$ or $\theta_N = -a/c_N$. Let Ψ be the continuous distribution function of a random variable with mean zero and unit variance, and suppose*

$$\lim_{N\to\infty} P_N\{c_N(h_N - \mu_N) \leqq u\} = \Psi[(u + aB)/A],$$

where P_N indicates that the probability is computed for the parameter values Δ_N or θ_N and where h_N stands for $h_N(X_1, \cdots, X_{m(N)}, Y_1, \cdots, Y_{n(N)})$ or *$h_N(Z_1, \cdots, Z_N)$. Then for any fixed Δ and θ, $\lim_{N\to\infty} P_\Delta\{c_N(\Delta_N^{**} - \Delta) \leqq a\} = \Psi(aB/A)$, or $\lim_{N\to\infty} P_\theta\{c_N(\theta_N^{*} - \theta) \leqq a\} = \Psi(aB/A)$.*

This theorem will be used repeatedly in the following to establish the asymptotic variances of our estimates.

2.3. *Model* I. To study the behavior of Δ^* in Model I we shall need the following theorem.

THEOREM 2.2. *Let $Z_1, \cdots, Z_N$ be N independent observations from a distribution $P(Z_i \leqq z) = K(z - \eta)$ where $K \in \mathfrak{F}_0$ and has pseudomedian 0. Let $h(Z)$ be defined by $\binom{N}{2}^{-1} W_1(Z)$ where $W_1(Z)$ is the one-sample Wilcoxon statistic, that is, $W_1(Z) =$ Number of pairs (i, j) with $1 \leqq i \leqq j \leqq N$ such that $z_i + z_j > 0$. Furthermore, let $\eta_N^*(Z)$ be defined exactly as θ^* in (0.6) and (0.7) with $\mu_N = \frac{1}{2} + 1/(N-1)$(or equivalently as in (0.8)). Then for every fixed η*

$$\lim_{N\to\infty} P_\eta\{N^{\frac{1}{2}}(\eta_N^* - \eta) \leqq u\} = \phi(uB/A),$$

where

$$A^2 = 4[\textstyle\int K^2(-z)k(z)\,dz - \frac{1}{4}] \tag{2.9}$$

and

$$B = 2\textstyle\int k(-z)k(z)\,dz \tag{2.10}$$

and where ϕ, as throughout this paper, denotes the distribution function of the standard normal distribution.

PROOF. The theorem will be proved by utilizing Theorem 2.1. Hence we need an expression for $\lim_{N\to\infty} P_N\{N^{\frac{1}{2}}(h_N - \mu_N) \leqq u\}$, where P_N indicates that the probability is computed for the value $\eta_N = -a/N^{\frac{1}{2}}$ of the pseudomedian. Define for all (i, j), $(i = 1, \cdots, N, j = 1, \cdots, N)$,

$$\varphi(Z_i, Z_j) = 1 \quad \text{when } Z_i + Z_j > 0,$$
$$= 0 \quad \text{otherwise.}$$

Then

$$h_N(Z) = \binom{N}{2}^{-1} \sum_{i<j}^{N} \varphi(Z_i, Z_j) + \binom{N}{2}^{-1} \sum_{i=1}^{N} \varphi(Z_i, Z_i) \tag{2.11}$$

and

$$N^{\frac{1}{2}}(h_N - \mu_N) = Q_N + R_N + S_N, \tag{2.12}$$

where

$$Q_N = N^{\frac{1}{2}} \binom{N}{2}^{-1} \sum_{i<j}^{N} [\varphi(Z_i, Z_j) - E_{\eta_N} \varphi(Z_i, Z_j)], \tag{2.13}$$

$$R_N = N^{\frac{1}{2}}[E_{\eta_N} \varphi(Z_i, Z_j) - \tfrac{1}{2}], \tag{2.14}$$

and

$$S_N = N^{\frac{1}{2}}[\binom{N}{2}^{-1} \sum_{i=1}^{N} \varphi(Z_i, Z_i) - 1/(N-1)]. \tag{2.15}$$

$\binom{N}{2}^{-1} \sum_{i<j}^{N} \varphi(Z_i, Z_j)$ is easily seen to be a generalized U-statistic [6].

$$E_{\eta_N} \varphi(Z_i, Z_j) = P_{\eta_N}[(Z_i + Z_j) > 0] = 1 - \int K(-z - \eta_N) k(z - \eta_N)\, dz \tag{2.16}$$

and

$$\begin{aligned} &\operatorname{cov}_{\eta_N}(\varphi(Z_1, Z_2), \varphi(Z_1, Z_3)) \\ &\quad = P_{\eta_N}[(Z_1 + Z_2) > 0, (Z_1 + Z_3) > 0] - P_{\eta_N}[(Z_1 + Z_2) > 0]^2, \end{aligned} \tag{2.17}$$

where Z_1, Z_2 and Z_3 are independent observations from $K(z - \eta_N)$. Hence

$$\begin{aligned} &\operatorname{cov}_{\eta_N} [\varphi(Z_1, Z_2), \varphi(Z_1, Z_3)] \\ &\quad = \int K^2(-z - \eta_N) k(z - \eta_N)\, dz - [\int K(-z - \eta_N) k(z - \eta_N)\, dz]^2. \end{aligned} \tag{2.18}$$

When $N \to \infty$, $\eta_N \to 0$ and the right side of (2.18) $\to A^2/4$, defined in (2.9). Furthermore, since the expression by Lemma 2.1 is $\geqq \frac{1}{12}$, Lehmann's extension of Hoeffding's theorem on generalized U-statistics applies [8]. Hence

$$\lim_{N\to\infty} P_N(Q_N \leqq u) = \phi(u/A). \tag{2.19}$$

As for R_N we find by using (2.16), Lemma 2.1(i) and introducing $\eta_N = -a/N^{\frac{1}{2}}$, that

$$R_N = -2a \int \frac{K(-z + 2a/N^{\frac{1}{2}}) - K(-z)}{2a/N^{\frac{1}{2}}} k(z)\, dz.$$

Hence $R_N \to -2a \int k(-z) k(z)\, dz = -aB$ as $N \to \infty$.

As for S_N, $|S_N| \leqq N^{\frac{1}{2}}\{\binom{N}{2}^{-1} \sum |\varphi(Z_i, Z_i)| + 1/(N-1)\} \leqq 3N^{\frac{1}{2}}/(N-1)$ and therefore $S_N \to 0$ as $N \to \infty$.

By repeated use of Slutsky's theorem (see, for example [1], p. 254),

$$\lim_{N\to\infty} \{P_N(Q_N + R_N + S_N) \leqq u\} = \phi((u + aB)/A). \qquad \text{q.e.d.}$$

Therefore, in Model I η^* is a consistent estimate for the pseudomedian, and Δ^* thereby a consistent estimate for the difference between the pseudomedians of the distributions of Y and X. This difference is in general different from Δ.

In the study of Δ^{**} in Model I we shall need the following theorem.

THEOREM 2.3. *Let $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ be $m + n = N$ independent observations from distributions $P(X_i \leqq u) = F(u - \zeta)$, $(i = 1, \cdots, m)$ and $P(Y_j \leqq u) = G(u - \eta)$, $(j = 1, \cdots, n)$, $F \varepsilon \mathfrak{F}_{01}$, $G \varepsilon \mathfrak{F}_{01}$. Furthermore, let ω denote the median of the distribution of $(Y - X)$ and $h(X, Y)$ be defined as $W_2(X, Y)/mn$ where $W_2(X, Y)$ is the Mann-Whitney test statistic. Finally let Δ_N^{**} be defined as in (0.3) and (0.4) with $\mu_N = \frac{1}{2}$, or equivalently as in (0.5).*

Then for every fixed ω, $\lim_{N\to\infty} P_\omega(N^{\frac{1}{2}}(\Delta_N^{**} - \omega) \leqq u) = \phi(uB/A)$, where

$$(2.20) \quad A^2 = \lambda^{-1}\int K^2(x)f(x)\,dx + (1 - \lambda)^{-1}\int F^2(x)k(x)\,dx - 1/[4\lambda(1 - \lambda)],$$

$$(2.21) \quad B = \int f(x)k(x)\,dx,$$

$\lambda = \lim_{N\to\infty} m(N)/N$, and $K(u)$ denote the distribution of $Y - \omega$ with corresponding density k.

PROOF. First we notice that by (1.2), ζ may without loss of generality be taken to be 0. Then there is a one-to-one correspondence between η and ω.

Define for all (i, j), $(i = 1, \cdots, m; j = 1, \cdots, n)$

$$\begin{aligned} \varphi(X_i, Y_j) &= 1 \qquad \text{when } Y_j > X_i \\ &= 0 \qquad \text{otherwise,} \end{aligned}$$

and write $N^{\frac{1}{2}}[h_N(X, Y) - \frac{1}{2}] = Q_N + R_N$, where

$$Q_N = N^{\frac{1}{2}}\{(mn)^{-1}\textstyle\sum_{ij} [\varphi(X_i, Y_j) - E_\omega\varphi(X_i, Y_j)]\}$$

and

$$R_N = N^{\frac{1}{2}}\{E_\omega\varphi(X_i, Y_j) - \tfrac{1}{2}\}.$$

$(mn)^{-1}\sum_{ij} \varphi(X_i, Y_j)$ is easily seen to be a generalized U-statistic, and its mean and variance may be determined the usual way.

Suppose the distribution G depends on N in such a way that $\omega_N = -a/N^{\frac{1}{2}}$. Then as $N \to \infty$, $\omega_N \to 0$ and var $Q_N \to A^2$. By Lemma 2.2(ii) $A^2 > 0$; once again Lehmann's extension of Hoeffding's theorem on generalized U-statistics applies, and we may conclude that $\lim_{N\to\infty} P_N(Q_N \leqq u) = \phi(u/A)$.

By the same approach as in the proof of Theorem 2.2, R_N is shown to tend to $-a\int f(y)k(y)\,dy$ as $N \to \infty$. The proof is then completed by Slutsky's theorem.

Theorem 2.3 implies that Δ^{**} in Model I is a consistent estimate of the median of the distribution of $(Y - X)$. This is in general different from Δ.

Hence in Model I, $\hat{\Delta}$, Δ^* and Δ^{**} estimate different aspects of the model, and the choice between them will have to be based on what "type of shift" one wants to estimate.

2.4. *Model* II. Since Model II is contained in Model I, Δ^* is a consistent es-

TABLE 1b

Estimate	Model I	Model II	Model III	Model IV
$\hat{\Delta}$	Difference between EY and EX	Δ	Δ	Δ
Δ^*	Difference between η_G and η_F	Δ	Δ	Δ
Δ^{**}	Median of $(Y - X)$	Δ	Δ	Δ

timate for the difference between the pseudomedians of the distributions of Y and of X. Since $F = G$, this difference coincides with Δ.

2.5. *Concluding remarks.* Table 1b summarizes the results of this section, giving the quantities of which Δ^*, Δ^{**} and $\hat{\Delta}$ are consistent estimates.

By the regularity properties proved in [5], the estimates all have continuous distributions. Since the asymptotic relative behavior of the three estimates in Model IV is known ([5] and [8]), we shall restrict our investigation to Models II and III.

3. Robustness against symmetry. Model II.

3.1. ARE $(\Delta^*, \hat{\Delta})$ *and* ARE (Δ^*, Δ^{**}) *in Model* II. Let η_F and η_G be the pseudomedians of X and Y respectively. Since in Model II, $F(x - \Delta) = G(x)$, we have by Section 2.2 that $\eta_G - \eta_F = \Delta$, and the model may be rewritten in the following way:

$$P(X \leqq u) = K(u - \eta), \qquad P(Y \leqq u) = K(u - \eta - \Delta) \tag{3.1}$$

where $K \in \mathfrak{F}_0$ and has pseudomedian zero but is not necessarily symmetric. The shift Δ is to be estimated.

THEOREM 3.1. *Let* $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ *be* $(m + n) = N$ *independent observations from the distributions* (3.1). *Furthermore, let*

$$\eta_G^* = \operatorname{med}_{i \leqq j}\{(Y_i + Y_j)/2\}, \qquad \eta_F^* = \operatorname{med}_{k \leqq l}\{(X_k + X_l)/2\} \tag{3.2}$$
$$(i, j = 1, \cdots, n; k, l = 1, \cdots, m)$$

and

$$\Delta_N^* = \eta_G^* - \eta_F^*. \tag{3.3}$$

Assume that B *given by* (2.10) *is finite. Then for every fixed* Δ,

$$\lim_{N \to \infty} P_\Delta\{N^{\frac{1}{2}}(\Delta_N^* - \Delta) \leqq u\} = \phi(\lambda^{\frac{1}{2}}(1 - \lambda)^{\frac{1}{2}}Bu/A),$$

where A^2 *and* B *are given by* (2.9) *and* (2.10).

PROOF. By Theorem 2.2, $m(N)^{\frac{1}{2}}(\eta_F^* - \eta)$ has a limiting normal distribution $(0, A^2/B^2)$, that is, with mean zero and variance A^2/B^2. Hence by Slutsky's theorem $N^{\frac{1}{2}}(\eta_F^* - \eta)$ has a limiting normal distribution $(0, A^2/\lambda B^2)$. Similarly $N^{\frac{1}{2}}(\eta_G^* - \eta - \Delta)$ has a limiting normal distribution $(0, A^2/(1 - \lambda)B^2)$. Since $A^2 \geqq \frac{1}{12}$ and B is finite, neither of the limiting distributions is singular, and the theorem follows.

THEOREM 3.2. *Let* $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ *be* $(m + n) = N$ *independent observations from the distributions* (3.1). *Let* $\hat{\Delta}$ *and* Δ^{**} *be defined by* (0.2) *and*

(0.5) and Δ^* *by* (0.9). *Then*

(i) ARE (Δ^*, Δ^{**})

$$= [\int k(x)k(-x)\,dx]^2/[\int k^2(x)\,dx]^2 \cdot 12[\int K^2(-x)k(x)\,dx - \tfrac{1}{4}],$$

(ii) ARE $(\Delta^*, \hat{\Delta}) = \sigma_x^2[\int k(x)k(-x)\,dx]^2/\{\int K^2(-x)k(x)\,dx - \tfrac{1}{4}\}$, *where* σ_x^2 *denotes the variance of* X,

(iii) ARE $(\Delta^*,\Delta^{**}) \leqq 1$ *with equality if and only if* K *is symmetric, and similarly*

(iv) ARE $(\Delta^*, \hat{\Delta}) \leqq$ ARE $(\Delta^{**}, \hat{\Delta})$ *with equality if and only if* K *is symmetric.*

Proof. In [5] is shown that the asymptotic variance of $N^{\frac{1}{2}}(\Delta_N^{**} - \Delta)$ is $1/[12\lambda(1 - \lambda)(\int k^2(x)\,dx)^2]$. This result and Theorem 3.1 imply (i). (ii) is immediate. (iii) and (iv) follow by the use of Lemma 2.1. Note that (i) and (iii) do not require that σ_x^2 be finite.

3.2. *Lower bound for* ARE $(\Delta^*, \hat{\Delta})$ *in Model* II. Since Hodges and Lehmann have shown [3] that $12\sigma_x^2[\int f^2(x)\,dx]^2 \geqq .864$, implying ARE $(\Delta^{**}, \hat{\Delta}) \geqq .864$ in Model II, a natural question is whether or not a similar lower bound exists for ARE $(\Delta^*, \hat{\Delta})$.

The following example shows, however, that there exist distributions $K(x)$ with pseudomedian 0 for which $\int k(x)k(-x)\,dx = 0$. Since the denominator of ARE $(\Delta^*, \hat{\Delta})$ is bounded away from zero, this implies that there exist nonsymmetric distributions for which ARE $(\Delta^*, \hat{\Delta}) = 0$. An example of such distributions is the one with density

$$(3.4)\qquad \begin{aligned} k(x) &= 1/2^{\frac{1}{2}}a && \text{for } -a < x < 0,\\ &= 0 && \text{for } 0 \leqq x \leqq a,\\ &= (2^{\frac{1}{2}} - 1)/2^{\frac{1}{2}}a && \text{for } a < x < 2a, \end{aligned}$$

where a is a real number.

It is easily verified that this distribution has pseudomedian 0, and finite variance, while $\int k(x)k(-x)\,dx = 0$.

3.3. *An example.* Assume Model II with

$$\begin{aligned} F(x) &= 1 - \exp\{-x/2\}, && x \geqq 0,\\ &= 0, && \text{otherwise}, \end{aligned}$$

and let us derive the ARE $(\Delta^*, \hat{\Delta})$ and ARE $(\Delta^{**}, \hat{\Delta})$ in this case. Using the property that F is the χ^2 distribution with two degrees of freedom, the pseudomedian η_0 is easily determined to be ≈ 1.678. The distribution of $X - \eta_0$ will have pseudomedian zero. Hence in this case $K(u) = 1 - \exp\{-(u + \eta_0)/2\}$, $u \geqq -\eta_0$. Numerical evaluation gives $\int K^2(-x)k(x)\,dx = .408$, $\int k(x)k(-x)\,dx = .157$ and ARE $(\Delta^*, \hat{\Delta}) = .62$.

For comparison we also compute the ARE $(\Delta^{**}, \hat{\Delta})$ and get ARE $(\Delta^{**}, \hat{\Delta}) = 3$. In this case of rather extreme asymmetry therefore we have an example where Δ^{**} in the limit is about "five times as efficient" as Δ^*.

4. Robustness against $F \neq G$. Model III.

4.1. ARE $(\Delta^*, \hat{\Delta})$ *in Model* III. For convenience we return to the notation used in Section 1.2.

THEOREM 4.1. *Let* $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ *be* $(m + n) = N$ *independent observations from the distributions* (1.1) *with* $F \varepsilon \mathfrak{F}_1$, $G \varepsilon \mathfrak{F}_1$. *Furthermore let* $\hat{\Delta}$ *be defined by* (0.2),

$$\begin{aligned} \zeta_2^* &= \text{med}_{i \leqq j} [(Y_i + Y_j)/2], \\ \zeta_1^* &= \text{med}_{k \leqq l} [(X_k + X_l)/2] \quad (i, j = 1, \cdots, n; k, l = 1, \cdots, m) \end{aligned} \tag{4.1}$$

and

$$\Delta_N^* = \zeta_2^* - \zeta_1^*. \tag{4.2}$$

Assume that $\int f^2(x)\,dx$ *and* $\int g^2(x)\,dx$ *are both finite.*

(i) *Then for every fixed* Δ, $\lim_{N \to \infty} P_\Delta\{N^{\frac{1}{2}}(\Delta_N^* - \Delta) \leqq u\} = \phi(u/D)$, *where*

$$D^2 = \{12\lambda[\int f^2(x)\,dx]^2\}^{-1} + \{12(1 - \lambda)[\int g^2(x)\,dx]^2\}^{-1}. \tag{4.3}$$

(ii) *Furthermore*

$$\text{ARE}\,(\Delta^*, \hat{\Delta}) = [\sigma_x^2/\lambda + \sigma_y^2/(1 - \lambda)]D^{-2}, \tag{4.4}$$

where σ_x^2 *and* σ_y^2 *as usual denote the variances of* X *and* Y *respectively, assumed to exist.*

(i) is proved exactly as Theorem 3.1. (ii) is immediate.

4.2. ARE $(\Delta^{**}, \hat{\Delta})$ *in Model* III.

THEOREM 4.2. *Let* $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ *be* $(m + n) = N$ *independent observations from the distributions* (1.1) *with* $F \varepsilon \mathfrak{F}_1$, $G \varepsilon \mathfrak{F}_1$ *and let* $\hat{\Delta}$ *and* Δ^{**} *be defined as in* (0.2) *and* (0.5)

(i) *Then for every fixed* Δ, $\lim_{N \to \infty} P_\Delta\{N^{\frac{1}{2}}(\Delta_N^{**} - \Delta) \leqq u\} = \phi(Bu/A)$, *where*

$$A^2 = \lambda^{-1}\int G^2(x)f(x)\,dx + (1 - \lambda)^{-1}\int F^2(x)g(x)\,dx - 1/4\lambda(1 - \lambda) \tag{4.5}$$

and

$$B = \int f(x)g(x)\,dx. \tag{4.6}$$

(ii) *Furthermore,*

$$\text{ARE}(\Delta^{**}, \hat{\Delta}) = \frac{[(1 - \lambda)\sigma_x^2 + \lambda\sigma_y^2][\int f(x)g(x)\,dx]^2}{(1 - \lambda)\int G^2(x)f(x)\,dx + \lambda\int F^2(x)g(x)\,dx - \frac{1}{4}}. \tag{4.7}$$

PROOF. Part (i) is really only a corollary of Theorem 2.3. Since F and G are symmetric, $\omega = \Delta$. Part (ii) is immediate.

The most interesting application of the results in Theorem 4.1 and Theorem 4.2 is to the nonparametric Behrens-Fisher situation where F and G represent the same distribution except for an unknown scale parameter. This case will be discussed in the next section, and we shall first apply the general results to a few examples.

For the computations we shall require repeated use of the following two lemmas which are given here for the purpose of convenient reference. The first one is due to Sheppard (1898) (see, for example [2]), the other also contains well-known results.

LEMMA 4.1. *Let (X_1, X_2) have a binormal distribution with means $(0, 0)$, variances $(1, 1)$ and correlation coefficient ρ. Then*

$$P(X_1 \leqq 0,\ X_2 \leqq 0) = 4^{-1}(1 + 2\pi^{-1} \text{ Arcsine } \rho). \tag{4.8}$$

LEMMA 4.2. *Let ϕ and φ denote the distribution function and the density of the standard normal distribution. Then for all positive real numbers a, b and c*

(i)

$$\begin{aligned}\int \phi(u/b)\phi(u/c)a^{-1}\varphi(u/a)\,du \\ = 4^{-1}\{1 + 2\pi^{-1} \text{ Arcsine } [a^2/((a^2 + b^2)(a^2 + c^2))^{\frac{1}{2}}]\}.\end{aligned} \tag{4.9}$$

(ii)

$$\int \varphi(u/a)\varphi(u/b)\,du = ab/(2\pi(a^2 + b^2))^{\frac{1}{2}}. \tag{4.10}$$

PROOF. (i) Let U_1, U_2 and U_3 be independent with normal distributions $(0, a^2)$, $(0, b^2)$ and $(0, c^2)$ respectively. Then $\int \phi(u/b)\phi(u/c)a^{-1}\varphi(u/a)\,du = P(U_2 \leqq U_1, U_3 \leqq U_1) = P(U_2 - U_1 \leqq 0, U_3 - U_1 \leqq 0) = P[(U_2 - U_1)\cdot(a^2 + b^2)^{-\frac{1}{2}} \leqq 0, (U_3 - U_1)(a^2 + c^2)^{-\frac{1}{2}} \leqq 0]$. The result now follows by the use of Lemma 4.1. (ii) is obtained by direct integration.

EXAMPLE 4.2.1. Suppose one wants to estimate the amount Δ by which the response of certain experimental units is increased by a specific treatment. Of the $m + n = N$ experimental units at hand n are selected at random for treatment and the remaining m serve as controls. The model used in this situation is often the one given by (1.1). Let us now specify further and consider the following gross error model in which independent normal observations are contaminated by a proportion ϵ of gross errors.

$$\begin{aligned} P(X_i \leqq u) &= F(u - \zeta) \\ &= (1 - \epsilon)\phi[(u - \zeta)/\sigma] + \epsilon\phi[(u - \zeta)/a\sigma], \qquad (i = 1, \cdots, m).\end{aligned} \tag{4.11}$$

$$\begin{aligned} P(Y_j \leqq u) &= G(u - \zeta - \Delta) \\ &= (1 - \epsilon)\phi[(u - \zeta - \Delta)/c\sigma] + \epsilon\phi[(u - \zeta - \Delta)/a\sigma], \ (j = 1, \cdots, n).\end{aligned} \tag{4.12}$$

ϵ will typically be less than 5 per cent and a a number between 2 and 4. ζ, σ, Δ and c are all assumed unknown.

Let us study the asymptotic behavior of the estimates $\hat{\Delta}$ (0.2), Δ^* (0.9) and Δ^{**} (0.5) in this situation, using Theorems 4.1 and 4.2.

Direct evaluation of (4.3) and (4.4) by the use of Lemma 4.2 gives

$$\begin{aligned}&\text{ARE}(\Delta^*, \hat{\Delta}) \\ &\quad = 3\pi^{-1} \frac{\{(1 - \lambda)[1 + \epsilon(a^2 - 1)] + \lambda[c^2 + \epsilon(a^2 - c^2)]\}\cdot A^2\cdot B^2(c)}{(1 - \lambda)B^2(c) + \lambda A^2},\end{aligned} \tag{4.13}$$

where

$$A = [(1 - \epsilon)^2 + \epsilon^2 a^{-1} + 2^{\frac{3}{2}}\epsilon(1 - \epsilon)\cdot(1 + a^2)^{-\frac{1}{2}}] \tag{4.14}$$

and

$$B(c) = [(1 - \epsilon)^2 c^{-1} + \epsilon^2 a^{-1} + 2^{\frac{3}{2}}\epsilon(1 - \epsilon)(c^2 + a^2)^{-\frac{1}{2}}]. \tag{4.15}$$

For the evaluation of ARE $(\Delta^{**}, \hat{\Delta})$ one needs:

$$\begin{aligned} \int F^2(x)g(x)\,dx = {} & 4^{-1}\{1 + (1 - \epsilon)^3 \cdot 2\pi^{-1} \text{ Arcsine } \alpha(c, 1) \\ & + 2\epsilon(1 - \epsilon)^2 \cdot 2\pi^{-1} \text{ Arcsine } (\alpha(c, 1) \cdot \alpha(c, a))^{\frac{1}{2}} \\ & + \epsilon^2(1 - \epsilon) \cdot 2\pi^{-1} \text{ Arcsine } \alpha(c, a) + (1 - \epsilon)^2 \cdot \epsilon \cdot 2\pi^{-1} \text{ Arcsine } \alpha(a, 1) \\ & + 2\epsilon^2(1 - \epsilon) \cdot 2\pi^{-1} \text{ Arcsine } (\alpha(a, 1) \cdot \alpha(1, 1))^{\frac{1}{2}} + \epsilon^3/3\}. \end{aligned} \tag{4.16}$$

$$\begin{aligned} \int G^2(x)f(x)\,dx = {} & 4^{-1}\{1 + (1 - \epsilon)^3 \cdot 2\pi^{-1} \text{ Arcsine } \alpha(1, c) \\ & + 2\epsilon(1 - \epsilon)^2 \cdot 2\pi^{-1} \text{ Arcsine } (\alpha(1, c) \cdot \alpha(1, a))^{\frac{1}{2}} \\ & + \epsilon^2(1 - \epsilon) \cdot 2\pi^{-1} \text{ Arcsine } \alpha(1, a) \\ & + (1 - \epsilon)^2 \cdot \epsilon \cdot 2\pi^{-1} \text{ Arcsine } \alpha(a, c) \\ & + 2\epsilon^2(1 - \epsilon) \cdot 2\pi^{-1} \text{ Arcsine } (\alpha(a, c) \cdot \alpha(1, 1))^{\frac{1}{2}} + \epsilon^3/3\}, \end{aligned} \tag{4.17}$$

where

$$\alpha(u, v) = u^2/(u^2 + v^2) \tag{4.18}$$

and

$$\begin{aligned} & \int f(x)g(x)\,dx \\ & \quad = (2\pi\sigma^2)^{-\frac{1}{2}}[(1 - \epsilon)^2/(1 + c^2)^{\frac{1}{2}} + \epsilon(1 - \epsilon)/(a^2 + c^2)^{\frac{1}{2}} \\ & \qquad + \epsilon(1 - \epsilon)/(a^2 + 1)^{\frac{1}{2}} + \epsilon^2/2^{\frac{1}{2}}a]. \end{aligned} \tag{4.19}$$

As a numerical illustration we have computed the ARE $(\Delta^*, \hat{\Delta})$ and the ARE $(\Delta^{**}, \hat{\Delta})$ for $\lambda = \frac{1}{2}$ (equal sample size), gross error of order 3σ, $c = 1, 2^{\frac{1}{2}}$ and 2, and $\epsilon = .01, .02, .03, .04$ and $.05$. The results are given in Table 2 and Figure 1. We know from [5] that ARE $(\Delta^{**}, \hat{\Delta})$ = ARE $(\Delta^*, \hat{\Delta})$ when $c = 1$.

TABLE 2

ARE$(\cdot, \hat{\Delta})$. Gross error model. $\lambda = \frac{1}{2}$

ϵ	$c = 1$	$c = 2^{\frac{1}{2}}$		$c = 2$	
	ARE$(\cdot, \hat{\Delta})$	ARE$(\Delta^*, \hat{\Delta})$	ARE$(\Delta^{**}, \hat{\Delta})$	ARE$(\Delta^*, \hat{\Delta})$	ARE$(\Delta^{**}, \hat{\Delta})$
.00	.955	.955	.935	.955	.886
.01	1.009	.985	.964	.969	.898
.02	1.060	1.013	.992	.982	.910
.03	1.108	1.040	1.018	.995	.921
.04	1.153	1.066	1.043	1.007	.933
.05	1.196	1.091	1.067	1.019	.943

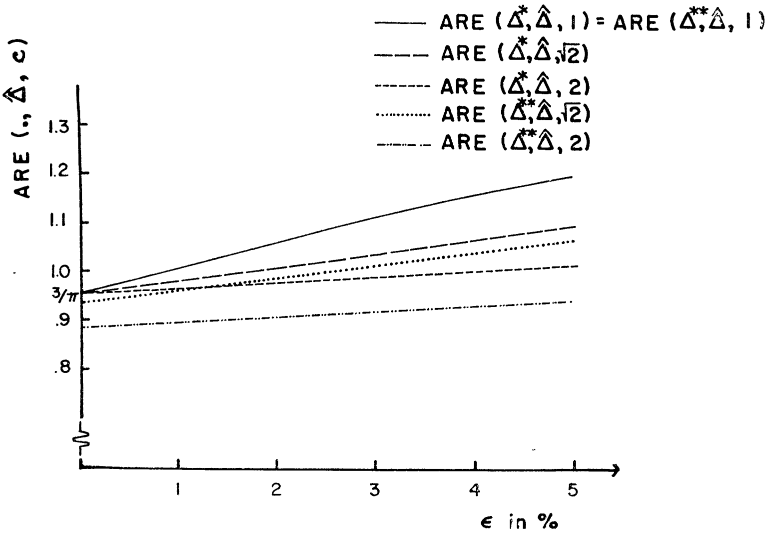


FIG. 1. Gross error model, $\lambda = \frac{1}{2}$, $a = 3$

The common values of these efficiencies for different values of ϵ are given in Table 2, column 2 under ARE $(\cdot, \hat{\Delta})$. In Figure 1 ARE $(\Delta^*, \hat{\Delta}, c_0)$ denotes the asymptotic efficiency of Δ^* relative to $\hat{\Delta}$ when $c = c_0$, c being the parameter occurring in (4.12). ARE $(\Delta^{**}, \hat{\Delta}, c_0)$ is defined similarly.

These numerical results seem to indicate that in this model the ARE $(\Delta^*, \hat{\Delta})$ as well as the ARE $(\Delta^{**}, \hat{\Delta})$ increase with increasing proportion (ϵ) of gross error (of course only up to a certain ϵ-value), and furthermore that Δ^* here is preferable to Δ^{**}.

In this example ARE $(\Delta^*, \hat{\Delta}) \geqq$ ARE $(\Delta^{**}, \hat{\Delta})$ for all calculated values, and one might be tempted to look for a general theorem such as Theorem 3.2 in Model III. The example treated in Section 4.3.1 shows that such a theorem is impossible when $\lambda \neq \frac{1}{2}$. The following example shows that even when $\lambda = \frac{1}{2}$, it is easy to find situations where ARE $(\Delta^{**}, \hat{\Delta}) >$ ARE $(\Delta^*, \hat{\Delta})$ in Model III.

EXAMPLE 4.2.2. Assume Model III with

$$F(x) = \phi(x) \quad \text{and} \quad G(x) = 2^{-1}\phi(x) + 2^{-1}\phi(x/10) \tag{4.20}$$

and take $\lambda = \frac{1}{2}$. By introducing (4.20) into (4.7) and (4.4) and evaluating the integrals by Lemmas 4.1 and 4.2, one obtains ARE $(\Delta^{**}, \hat{\Delta}) = 1.526$ and ARE $(\Delta^*, \hat{\Delta}) = 1.094$. This shows an example of Model III with $\lambda = \frac{1}{2}$, where ARE $(\Delta^{**}, \Delta^*) > 1$.

4.3. *The Behrens-Fisher situation in Model* III. We shall now consider the special case of Model III where F and G represent the same symmetric distribution, except for an unknown scale parameter $c(>0)$. The corresponding densities are denoted f and g, and we shall assume that the variances exist. Then

$$F(x) = G(cx), \qquad f(x) = cg(cx), \tag{4.21}$$

$$c^2\sigma_x^2 = \sigma_y^2. \tag{4.22}$$

Since we shall study the asymptotic ($N \to \infty$) efficiencies of Δ^{**} and Δ^* relative to $\hat{\Delta}$ as functions of λ and c, it will be convenient to use the notation ARE $(\cdot, \hat{\Delta}; \lambda, c)$.

When (4.21) and (4.22) are introduced into (4.4) and (4.7), the following theorem is immediate.

THEOREM 4.3. *Let $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ be $m + n$ independent observations from distributions* (1.1) *with $F \varepsilon \mathfrak{F}_1$ and $G(x) = F(x/c)$. Furthermore, let $\hat{\Delta}$, Δ^* and Δ^{**} be defined by* (0.2), (0.9) *and* (0.5). *Then*

(i)

$$\text{ARE}\,(\Delta^*, \hat{\Delta}; \lambda, c) = 12\sigma_x^2[\textstyle\int f^2(x)\,dx]^2 \tag{4.23}$$

and hence independent of λ as well as of c.

(ii)

$$\text{ARE}\,(\Delta^{**}, \hat{\Delta}; \lambda, c) = \frac{\sigma_x^2[(1-\lambda) + \lambda c^2]\cdot[\int f(x)f(cx)\,dx]^2}{(1-\lambda)\int F^2(x/c)f(x)\,dx + \lambda\int F^2(cx)f(x)\,dx - \frac{1}{4}} \tag{4.24}$$

and

(iii)

$$\text{ARE}\,(\Delta^{**}, \hat{\Delta}; \lambda, c) = \text{ARE}\,(\Delta^{**}, \hat{\Delta}; 1-\lambda, 1/c). \tag{4.25}$$

PROOF. (i) and (ii) follow from (4.4) and (4.7). (iii) is seen immediately to be true by writing (4.24) in an alternative way:

$$\text{ARE}\,(\Delta^{**}, \hat{\Delta}; \lambda, c) = \frac{\sigma_x^2\{(1-\lambda)[\int f(x)f(cx)\,dx]^2 + \lambda[\int f(x)f(x/c)\,dx]^2\}}{(1-\lambda)\int F^2(x/c)f(x)\,dx + \lambda\int F^2(cx)f(x)\,dx - \frac{1}{4}}. \tag{4.26}$$

THEOREM 4.4. *Assume Model* III *with $F(x) = G(cx)$, and let $\hat{\Delta}$ and Δ^{**} be defined by* (0.2) *and* (0.5). *In addition to the usual regularity conditions* (2.8) *we assume that $f(x)$ is continuous at $x = 0$. Then*

$$\lim_{c\to 0} \text{ARE}\,(\Delta^{**}, \hat{\Delta}; \lambda, c) = 4\sigma_x^2 f^2(0) = \lim_{c\to\infty} \text{ARE}\,(\Delta^{**}, \hat{\Delta}; \lambda, c) \tag{4.27}$$

(*which we recognize as the asymptotic efficiency of the median relative to the mean*).

PROOF. (4.24) may be rewritten as follows:

$$\text{ARE}\,(\Delta^{**}, \hat{\Delta}; \lambda, c) = \frac{\sigma_x^2[(1-\lambda) + \lambda c^2][\int f(x)f(cx)\,dx]^2}{\frac{1}{4} - 2(1-\lambda)\int_0^\infty F(-x/c)F(x/c)f(x)\,dx - 2\lambda\int_0^\infty F(cx)F(-cx)f(x)\,dx}. \tag{4.28}$$

Due to the regularity assumptions made, we are permitted to determine the limiting value of (4.28) as $c \to 0$, by letting $c \to 0$ under the integral signs. The second half follows from (4.25).

In addition to (4.27) we have the result from [5], also easily obtained from (4.24), that ARE $(\Delta^{**}, \hat{\Delta}; \lambda, 1) = 12\sigma_x^2[\int f^2(x)\,dx]^2$. Considering ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c)$ as a function of c for fixed λ, we hence know its value for $c = 1$ and its limiting values as $c \to 0$ and as $c \to \infty$. An interesting question is whether or not there exists a broad class of distributions for which

$$\text{(4.29)} \qquad \text{ARE } (\Delta^{**}, \hat{\Delta}; \lambda, c) \leqq \text{ARE } (\Delta^{*}, \hat{\Delta}; \lambda, c) \qquad \text{for all } \lambda \text{ and } c$$

or equivalently

$$\text{(4.29)}' \qquad \text{ARE } (\Delta^{**}, \hat{\Delta}; \lambda, c) \leqq \text{ARE } (\Delta^{**}, \hat{\Delta}; \lambda, 1) \qquad \text{for all } \lambda \text{ and } c$$

The examples considered in Sections 4.3.1 and 4.3.2 show that the rectangular distribution has this property, while the normal distribution requires the condition $\lambda = \frac{1}{2}$ (equal sample size). Let us therefore restrict ourselves to determining a class $\mathfrak{F}^*$ of distributions for which

$$\text{(4.30)} \qquad \text{ARE } (\Delta^{**}, \hat{\Delta}; \tfrac{1}{2}, c) \leqq \text{ARE } (\Delta^{*}, \hat{\Delta}; \tfrac{1}{2}, c) \qquad \text{for all } c.$$

This class $\mathfrak{F}^*$ contains at least the normal and the rectangular distribution.

THEOREM 4.5. *Assume Model* III *with* $F(x) = G(cx)$. *In addition to the regularity conditions of Theorem* 4.4, $f(x)$ *is assumed to satisfy regularity conditions of the form* R_2 *and* R_3 *in* (2.8). *Let* $\hat{\Delta}$, Δ^* *and* Δ^{**} *be defined by* (0.2), (0.9) *and* (0.5). *Then, if*

$$\text{(4.31)} \qquad f(x)\cdot f'(cx) - cf(cx)f'(x) \geqq 0 \qquad \text{for} \quad 0 \leqq x, \quad 0 < c \leqq 1,$$

ARE $(\Delta^{**}, \hat{\Delta}; \frac{1}{2}, c)$ *is nonincreasing from* $12\sigma_x^2[\int f^2(x)\,dx]^2$ *toward* $4\sigma_x^2 f^2(0)$ *as* c *decreases from* 1 *toward* 0, *and also when* c *increases from* 1 *toward* ∞.

PROOF. (4.26) may be written

$$\text{(4.32)} \qquad \begin{aligned} &\text{ARE } (\Delta^{**}, \hat{\Delta}; \lambda, c) \\ &= \frac{4\sigma_x^2\{(1-\lambda)[\int_0^\infty f(x)f(cx)\,dx]^2 + \lambda[\int_0^\infty f(x)f(x/c)\,dx]^2\}}{\frac{1}{4} - 2(1-\lambda)\int_0^\infty F(-x/c)F(x/c)f(x)\,dx - 2\lambda\int_0^\infty F(cx)F(-cx)f(x)\,dx}. \end{aligned}$$

Due to the regularity assumptions made, the integrals in (4.32) may for any $c > 0$ be differentiated with respect to c by differentiation under the integral sign. The derivative of the denominator is $-2\int_0^\infty xf(cx)f(x)\{(1-\lambda)[2F(x) - 1] - \lambda[2F(cx) - 1]\}\,dx$ which for $\lambda \leqq \frac{1}{2}$ is $\leqq -4\lambda\int_0^\infty xf(cx)f(x)\cdot[F(x) - F(cx)]\,dx$, and hence $\leqq 0$ for $0 < c \leqq 1$. The denominator of (4.32) is therefore nonincreasing in c for $0 < c \leqq 1$ and $\lambda \leqq \frac{1}{2}$, and hence in particular for $\lambda = \frac{1}{2}$.

The derivative of the numerator is

$$8\sigma_x^2\int_0^\infty f(x)f(cx)\,dx\cdot\int_0^\infty x[(1-\lambda)f(x)f'(cx) - \lambda cf(cx)f'(x)]\,dx,$$

which for $\lambda = \frac{1}{2}$ is equal to $4\sigma_x^2\int_0^\infty f(x)f(cx)\,dx\cdot\int_0^\infty x[f(x)f'(cx) - cf(cx)f'(x)]\,dx$. (4.31) implies that this derivative is $\geqq 0$ for $0 < c \leqq 1$. Hence we have shown that ARE $(\Delta^{**}, \hat{\Delta}; \frac{1}{2}, c)$ is nondecreasing in c for $0 < c \leqq 1$. From (4.24) follows that ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c)$ is continuous in c as $c \downarrow 0$. The rest follows from (4.25).

COROLLARY 4.1 *(4.31) is a sufficient condition for a distribution F satisfying the regularity conditions of Theorem 4.5, to belong to* $\mathfrak{F}^*$.

Let us finally illustrate the results of Theorems 4.3, 4.4 and 4.5 by considering a few examples, specifying F.

4.3.1. *Normal distributions.* Let

$$F(x) = \phi(x/\sigma_x), \qquad G(x) = \phi(x/c\sigma_x), \tag{4.33}$$

where c and σ_x are unknown. The normal distribution is easily seen to satisfy all regularity conditions of Theorem 4.5 and in particular (4.31), hence the theorem applies.

Since the integrals (4.26) in this case can easily be evaluated by Lemma 4.2, the results can be sharpened and made more specific.

THEOREM 4.6. *Assume Model* III *with F and G given by* (4.33). *Let $\hat{\Delta}$, Δ^* and Δ^{**} be defined by* (0.2), (0.9) *and* (0.5). *Then*

(i) ARE $(\Delta^*, \hat{\Delta}) = 3/\pi$ *for all c and λ,*

(ii) $$\text{ARE}\,(\Delta^{**}, \hat{\Delta}) = \frac{\lambda c^2/(1+c^2) + (1-\lambda)/(1+c^2)}{\lambda \text{ Arcsine}\,[c^2/(1+c^2)] + (1-\lambda) \text{ Arcsine}\,[1/(1+c^2)]},$$

(iii) ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) \leqq 1$ *for all λ and c,*

(iv) ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) =$ ARE $(\Delta^{**}, \hat{\Delta}; 1-\lambda, 1/c)$,

(v) ARE $(\Delta^{**}, \hat{\Delta}; \lambda, \frac{1}{2}) = 3/\pi$, *for all λ,*

(vi) $\lim_{c\to 0}$ ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) = \lim_{c\to\infty}$ ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) = 2/\pi$, *for all λ.*

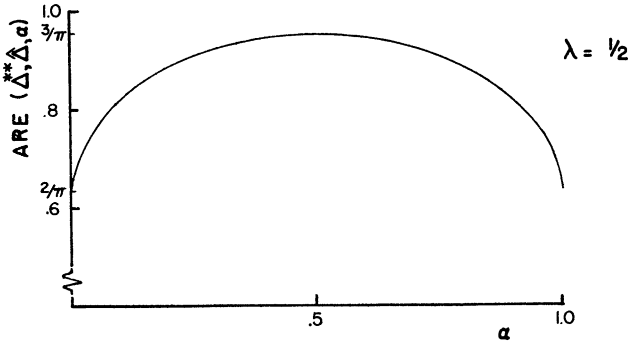
If in particular $\lambda = \frac{1}{2}$, then

TABLE 3

ARE$(\Delta^{**}, \hat{\Delta})$ as a Function of c for $\lambda = \frac{1}{2}, \frac{3}{4}$ and $\frac{9}{10}$

Normal Distributions

c^2	$\alpha = c^2/(1+c^2)$	ARE$(\Delta^{**}, \hat{\Delta})$ $\lambda = \frac{1}{2}$	$\lambda = \frac{3}{4}$	$\lambda = \frac{9}{10}$
0	0.000	.637	.637	.637
$\frac{1}{7}$	.125	.840	.866	.910
$\frac{1}{4}$	.200	.887	.914	.949
$\frac{1}{3}$	.250	.909	.933	.961
$\frac{1}{2}$	.333	.934	.952	.968
$\frac{2}{3}$	.400	.947	.958	.964
$\frac{4}{5}$	.455	.953	.958	.961
1	.500	.955	.955	.955
$\frac{4}{3}$	.571	.951	.942	.936
2	.667	.934	.923	.917
4	.800	.887	.872	.866
7	.875	.840	.828	.824
∞	1.000	.637	.637	.637



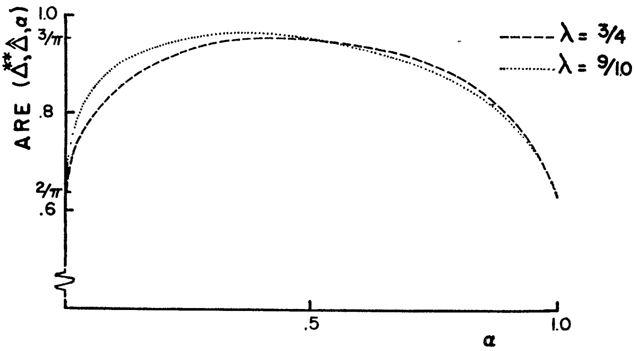


FIG. 2. (*upper*) Normal distributions, $\lambda = \frac{1}{2}$
FIG. 3. (*lower*) Normal distributions, $\lambda = \frac{3}{4}, \frac{9}{10}$

(vii) ARE $(\Delta^{**}, \hat{\Delta}; \frac{1}{2}, c)$ *decreases strictly from* $3/\pi$ *to* $2/\pi$ *when* c *decreases from* 1 *to* 0, *and also when* c *increases from* 1 *to* ∞.

(viii) ARE $(\Delta^{**}, \hat{\Delta}; \frac{1}{2}, c) \leqq$ ARE $(\Delta^{*}, \hat{\Delta}; \frac{1}{2}, c)$ *with equality if and only if* $c = 1$.

PROOF. (iii) follows since Arcsine $[c^2/(1 + c^2)] \geqq c^2/(1 + c^2)$. (vii) follows since $\partial/\partial c$ ARE $(\Delta^{**}, \hat{\Delta}; \frac{1}{2}, c) \gtreqless 0$ according as $c \lesseqgtr 1 (c > 0)$. The rest of the theorem is immediate.

As a numerical illustration, the ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c)$ has been computed as a function of c for $\lambda = \frac{1}{2}$, $\lambda = \frac{3}{4}$ and $\lambda = \frac{9}{10}$. The results are given in Table 3 and Figures 2 and 3. In connection with Figures 2, 3, 4 and 5 we consider the asymptotic efficiency of Δ^{**} relative to $\hat{\Delta}$ as a function of $\alpha = c^2/(1 + c^2)$ instead of c. This function is on the graphs denoted ARE $(\Delta^{**}, \hat{\Delta}, \alpha)$ for short. We notice from Figure 3 that there exist combinations of λ- and c-values such that ARE $(\Delta^{**}, \hat{\Delta}) >$ ARE $(\Delta^{*}, \hat{\Delta})$ and thereby see that Theorem 4.6(viii) needs condition $\lambda = \frac{1}{2}$ to hold.

4.3.2. *Rectangular distributions.* Let

$$F(x) = (\omega - \zeta + x)/2\omega; \qquad \zeta - \omega \leqq x \leqq \zeta + \omega, \tag{4.34}$$

(4.35) $G(x) = (c\omega - \zeta - \Delta + x)/2\omega c; \quad \zeta + \Delta - c\omega \leqq x \leqq \zeta + \Delta + c\omega,$

where $\omega > 0$, $c > 0$, ζ and Δ all are unknown constants. As in the previous example we are able to evaluate the integrals in (4.26). The results in this case can be summarized in the following theorem.

THEOREM 4.7. *Assume Model* III *with F and G given by* (4.34) *and* (4.35). *Let* $\hat{\Delta}$, Δ^* *and* Δ^{**} *be defined by* (0.2), (0.9) *and* (0.5). *Then*

(i) ARE $(\Delta^*, \hat{\Delta}) = 1$ for all c and λ,

$$\text{(ii)} \quad \text{ARE } (\Delta^{**}, \hat{\Delta}) = \frac{(1-\lambda) + \lambda c^2}{3(1-\lambda) - 2(1-\lambda)c + \lambda c^2}, \qquad \textit{for } 0 < c \leqq 1,$$

$$= \frac{(1-\lambda) + \lambda c^2}{3\lambda c^2 - 2\lambda c + 1 - \lambda}, \qquad \textit{for } 1 \leqq c,$$

(iii) ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) \leqq 1$ *for all* λ *and c and hence* ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) \leqq$ ARE $(\Delta^*, \hat{\Delta}; \lambda, c)$, *with equality if and only if* $c = 1$.

(iv) ARE $(\Delta^{**}, \hat{\Delta}; \lambda, 1) = 1$, *for all* λ,

(v) $\lim_{c \to 0}$ ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) = \lim_{c \to \infty}$ ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) = \frac{1}{3}$, *for all* λ,

(vi) ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c)$ *decreases strictly from* 1 to $\frac{1}{3}$ *when c decreases from* 1 to 0, *and also when c increases from* 1 *to* ∞.

The proof is immediate.

As a numerical illustration the ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c)$ has been computed as a function of c for $\lambda = \frac{1}{2}, \frac{2}{3}, \frac{3}{4}$ and $\frac{9}{10}$. The results are given in Table 4 and Figures 4 and 5.

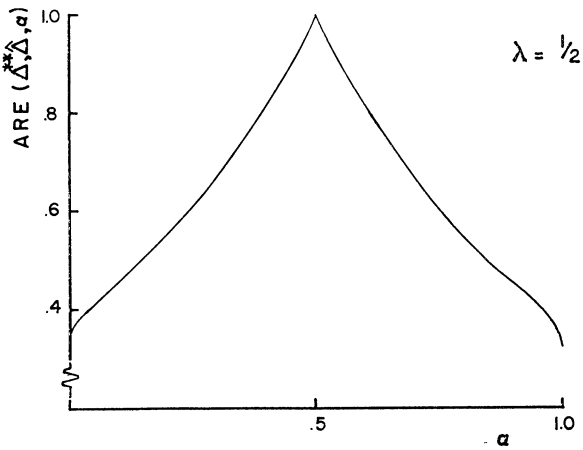
Before we proceed with an example of a symmetric distribution not satisfying (4.31), let us remark that the beta distributions $f(x) = k(a - x)^q(a + x)^q$; $-a \leqq x \leqq a$; for $q \geqq 1$, is easily seen to satisfy all required regularity conditions and (4.31), and hence belong to $\mathfrak{F}^*$.

TABLE 4

ARE$(\Delta^{**}, \hat{\Delta})$ *as a Function of c for* $\lambda = \frac{1}{2}, \frac{2}{3}, \frac{3}{4}$ *and* $\frac{9}{10}$.

Rectangular Distributions

c^2	$\alpha = c^2/(1 + c^2)$	ARE$(\Delta^{**}, \hat{\Delta})$			
		$\lambda = \frac{1}{2}$	$\lambda = \frac{2}{3}$	$\lambda = \frac{3}{4}$	$\lambda = \frac{9}{10}$
0.000	0.0	.333	.333	.333	.333
.111	.1	.455	.478	.500	.600
.250	.2	.556	.600	.636	.765
.429	.3	.674	.729	.768	.875
.667	.4	.820	.864	.891	.950
1.000	.5	1.000	1.000	1.000	1.000
1.500	.6	.820	.784	.769	.745
2.333	.7	.674	.637	.623	.603
4.000	.8	.556	.529	.520	.507
9.000	.9	.455	.442	.438	.432
∞	1.0	.333	.333	.333	.333



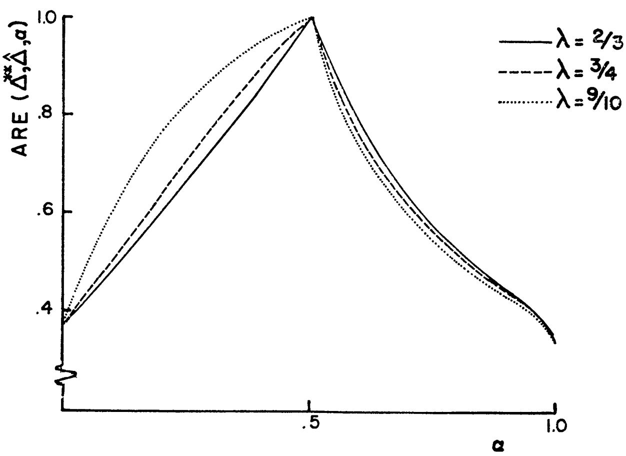


FIG. 4. (*upper*) Rectangular distributions, $\lambda = \frac{1}{2}$
FIG. 5. (*lower*) Rectangular distributions, $\lambda = \frac{2}{3}, \frac{3}{4}, \frac{9}{10}$

4.3.3. *Double exponential distributions.* Let

$$
\begin{aligned}
F(x) &= \tfrac{1}{2}e^{x/a}, & x &\leqq 0, \\
&= 1 - \tfrac{1}{2}e^{-x/a}, & x &> 0, \qquad (4.36)\\
G(x) &= F(x/c),
\end{aligned}
$$

where a and c are unknown positive constants. F satisfies the regularity conditions of Theorem 4.5, but not (4.31). As in the previous examples, the integrals in (4.26) may be evaluated, and we have the following results.

THEOREM 4.8. *Assume Model* III *with F and G defined by* (4.36). *Let $\hat{\Delta}$, Δ^* and Δ^{**} be defined by* (0.2), (0.9) *and* (0.5). *Then*

TABLE 5

ARE$(\Delta^{**}, \hat{\Delta})$ *as a Function of c for* $\lambda = \frac{1}{2}, \frac{3}{4}$ *and* $\frac{9}{10}$

Double Exponential Distributions

c	$\alpha = c/(1+c)$	ARE$(\Delta^{**}, \hat{\Delta})$ $\lambda = \frac{1}{2}$	$\lambda = \frac{3}{4}$	$\lambda = \frac{9}{10}$
0	0.0	2.000	2.000	2.000
$\frac{1}{9}$	.1	1.883	1.852	1.765
$\frac{2}{8}$	.2	1.749	1.668	1.525
$\frac{3}{7}$	.3	1.623	1.533	1.427
$\frac{4}{6}$	.4	1.533	1.479	1.436
1	.5	1.500	1.500	1.500
$\frac{6}{4}$	.6	1.533	1.571	1.589
$\frac{7}{3}$	.7	1.623	1.670	1.690
$\frac{8}{2}$	.8	1.749	1.782	1.794
9	.9	1.883	1.894	1.898
∞	1.0	2.000	2.000	2.000

(i) ARE $(\Delta^*, \hat{\Delta}) = \frac{3}{2}$ *for all* λ *and* c,

(ii) ARE$(\Delta^{**}, \hat{\Delta})$

$$= \frac{2\left[(1-\lambda)/(1+c)^2 + \lambda c^2/(1+c)^2\right]}{1-(1-\lambda)\left[2c/(1+c) - c/(2+c)\right] - \lambda\left[2/(1+c) - 1/(1+2c)\right]}$$

(iii) ARE $(\Delta^{**}, \hat{\Delta}; \lambda, 1) = \frac{3}{2}$ *for all* λ,

(iv) $\lim_{c\to 0}$ ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) = \lim_{c\to\infty}$ ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c) = 2$, for all λ. *In particular*, if $\lambda = \frac{1}{2}$, *then*

(v) ARE $(\Delta^{**}, \hat{\Delta}; \frac{1}{2}, c)$ *is strictly increasing from* $\frac{3}{2}$ *to* 2 *as* c *decreases from* 1 *to* 0 *and also when* c *increases from* 1 *to* ∞.

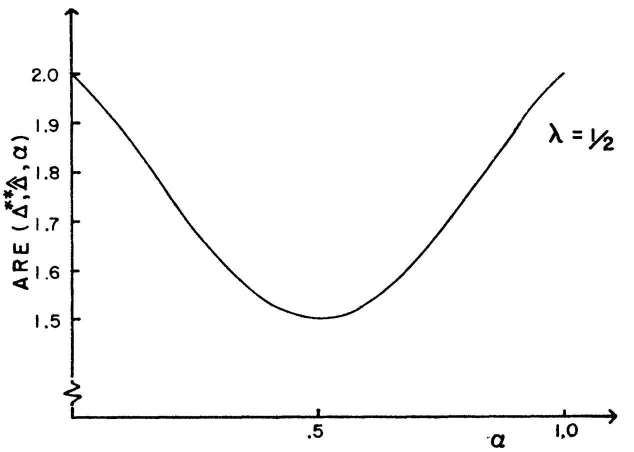
(vi) ARE $(\Delta^{**}, \hat{\Delta}; \frac{1}{2}, c) \geqq$ ARE $(\Delta^*, \hat{\Delta}; \frac{1}{2}, c)$ *with equality if and only if* $c = 1$.

As a numerical illustration the ARE $(\Delta^{**}, \hat{\Delta}; \lambda, c)$ has been computed as a function of c for $\lambda = \frac{1}{2}, \frac{3}{4}$ and $\frac{9}{10}$. The results are given in Table 5 and Figures 6 and 7.

In connection with Figures 6 and 7 we consider the asymptotic efficiency of Δ^{**} relative to $\hat{\Delta}$ as a function of $\alpha = c/(1+c)$ instead of c. This function is on the graphs denoted ARE $(\Delta^{**}, \hat{\Delta}, \alpha)$ for short.

The examples treated show that when we have a Behrens-Fisher situation in Model III and $\lambda = \frac{1}{2}$, then there exists a class of distributions, including the normal distribution and the beta distribution, for which Δ^* is preferable to Δ^{**}, but that there also exist distributions for which Δ^{**} is preferable to Δ^*. The examples further seem to indicate that heavy tails on the distributions make Δ^{**} favorable.

The choice between the two estimates Δ^* and Δ^{**} then will have to be made on the basis of the type of distribution one expects in each case.



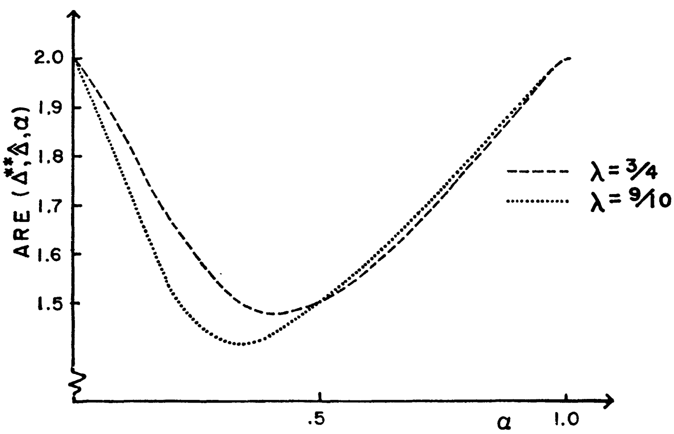


FIG. 6. (*upper*) Double exponential distributions, $\lambda = \frac{1}{2}$
FIG. 7. (*lower*) Double exponential distributions, $\lambda = \frac{3}{4}, \frac{9}{10}$

5. Acknowledgment. This work was done as part of a Ph.D. thesis at the University of California, Berkeley. I wish to express my sincere thanks to Professor E. L. Lehmann for suggesting the problem and for continued encouragement and guidance.

REFERENCES

[1] CRAMÉR, HARALD (1945). *Mathematical Methods of Statistics.* Almqvist and Wiksells, Uppsala.
[2] DAVID, F. N. (1953). A note on the evaluation of the multivariate normal integral. *Biometrika.* **40** 458–459.
[3] HODGES, J. L., JR. and LEHMANN, E. L. (1956). The efficiency of some nonparametric competitors of the *t*-test. *Ann. Math. Statist.* **27** 324–335.
[4] HODGES, J. L., JR. and LEHMANN, E. L. (1961). Comparison of the normal scores and Wilcoxon tests. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 307–317.

[5] HODGES, J. L., JR. and LEHMANN, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.
[6] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.
[7] LEHMANN, E. L. (1951). Consistency and unbiasedness of certain nonparametric tests. *Ann. Math. Statist.* **22** 165–179.
[8] LEHMANN, E. L. (1963). Robust estimation in analysis of variance. *Ann. Math. Statist.* **34** 957–966.