The Annals of Statistics
1974, Vol. 2, No. 6, 1323–1326

A COMPARISON OF EFFICIENT LOCATION ESTIMATORS[1]

BY FRIEDRICH-WILHELM SCHOLZ

University of California, Berkeley

For a given sufficiently regular distribution F two efficient location estimators are given. One is a linear combination of order statistics, called $L(F)$, and the other is an estimator derived from a rank test, called $R(F)$. The asymptotic variance of both estimators is then compared for various underlying distributions H and it is shown that the asymptotic variance of $R(F)$ is never larger than the one of $L(F)$.

0. Introduction and summary. The problem of estimating the center of a symmetric distribution has been treated quite extensively in statistical literature. Various classes of estimators have found special interest and we will deal with two of them: linear combinations or order statistics, and estimates that are derived from rank tests. Under suitable regularity conditions most of these estimators are asymptotically normal about the center of symmetry, with asymptotic variances depending on the underlying distribution. We therefore have a simple criterion, the asymptotic variances, for comparing the performance of two different estimators. Relative to this criterion it may happen that the performance of one estimator is never worse than the performance of the other. Chernoff and Savage (1958) give one example of such a comparison, although they consider this problem in terms of testing. For the problem of testing for shift in two samples, they showed that the asymptotic relative efficiency of the normal scores rank test relative to the t-test never falls below one. Using results of Hodges and Lehmann (1963), one can rephrase this result as follows for the estimation problem: The estimator derived from the normal scores one-sample rank test has an asymptotic variance which is always less or equal to the asymptotic variance of the sample mean. The asymptotic variances are equal if and only if the underlying distribution is normal, in which case both estimators are efficient.

This result suggests the following approach in finding other examples of this kind. We will compare only estimators which are efficient for a given fixed underlying distribution F. In Section 2 two such estimators are constructed. They are denoted by $L(F)$ and $R(F)$ and represent respectively a linear combination of order statistics and an estimator derived from a rank test. Section 3 presents a theorem which roughly states the following: If F is sufficiently regular

Received December 1972; revised January 1974.

[1] This research was partially supported by the National Science Foundation Grants GP-8690 and GP-31091X, and represents part of the author's doctoral dissertation, submitted at the Univ. of California, Berkeley.

AMS 1970 *subject classifications*. Primary 62G05, 62G20; Secondary 62G30.

Key words and phrases. Estimation, linear combination of order statistics, estimators derived from rank tests, efficiency.

(strongly unimodal, etc.), then the asymptotic variance of $R(F)$ is always less or equal to the asymptotic variance of $L(F)$ for all underlying distributions H satisfying certain regularity conditions.

1. Definitions. Let $X_1, \cdots, X_n$ be independent identically distributed random variables with distributions $H_\mu(x) = H(x - \mu)$. It is assumed that $H(x)$ has a density which is symmetric around zero. Otherwise H is unknown. We want to estimate the unknown location parameter μ. Since we deal with translation invariant estimators, we shall henceforth assume $\mu = 0$. For a more detailed account of the following two classes of estimators we refer to Jaeckel (1971).

One popular class of estimators is the class of linear combinations of order statistics, which we shall call L-estimators. We denote the L-estimator corresponding to the weight function g (Jaeckel denotes it by h) by $L_n(g)$. Under general regularity conditions on g and H we have

$$n^{\frac{1}{2}}L_n(g) \rightarrow N(O, \sigma_g^2(H))$$

where $\sigma_g^2(H) = \int_0^1 s^2(u)\,du$ with $s(u) = \int g(H(x))(I_{[u \leq H(x)]} - H(x))\,dx$ (cf. Chernoff, Gastwirth and Johns (1967), Bickel (1967) and Stigler (1969)). Denote the class of distributions H for which this asymptotic normality holds by $CL(g)$.

Another class of estimators was introduced by Hodges and Lehmann (1963). These estimators, called R-estimators, are derived from one-sample rank tests. The R-estimator corresponding to the score function J is denoted by $R_n(J)$. Under regularity conditions on J and H we have

$$n^{\frac{1}{2}}R_n(J) \rightarrow N(O, \sigma_J^2(H))$$

where $\sigma_J^2(H) = (\int_0^1 J^2(u)\,du) \cdot (\int (d/dx)\{J(H(x))\}\,dH(x))^{-2}$ (cf. Hodges and Lehmann (1963) and Puri and Sen (1971)). Denote the class of distributions H for which this asymptotic normality holds by $CR(J)$.

2. Efficient estimators for F. For $H_\mu(x) = F((x - \mu)/\sigma)$, where F is completely known and belongs to the class $\mathscr{F}$ defined below, we construct L- and R-estimators for μ which are efficient at F. Since these two estimators, which are denoted by $L(F)$ and $R(F)$ respectively, are location and scale invariant, we may restrict ourselves to $\mu = 0$ and $\sigma = 1$.

The class $\mathscr{F}$ consists of all distributions F satisfying the following regularity conditions:

(i) $F(x)$ has a density f with the properties $f(x) = f(-x)$ for $x \in \mathbb{R}$ and $f(x) > 0$ for $x \in \mathbb{R}$;

(ii) f is continuously differentiable on $\mathbb{R}$;

(iii) $I(f) = \int \psi_f^2(x)f(x)\,dx < \infty$, where $\psi_f(x) = -f'(x)/f(x)$;

(iv) ψ_f has a continuous derivative except at a finite number of points.

DEFINITION OF $L(F)$. For $F \in \mathscr{F}$ set $g(t) = [I(f)]^{-1} \cdot \psi_f'(F^{-1}(t))$ and define $L(F) = L_n(g)$. If $F \in CL(g)$, then $n^{\frac{1}{2}}L(F)$ is asymptotically normal with mean zero and variance $[I(f)]^{-1}$, i.e. $L(F)$ is efficient.

DEFINITION OF $R(F)$. For $F \in \mathscr{F}$ set $J(t) = \psi(F^{-1}(t))$ and define $R(F) = R_n(J)$. If $F \in CR(J)$, then $n^{\frac{1}{2}}R(F)$ is asymptotically normal with mean zero and variance $[I(f)]^{-1}$, i.e. $R(F)$ is efficient.

3. Comparison of $R(F)$ and $L(F)$. For a given $F \in \mathscr{F}$ we will now compare the asymptotic variances of $R(F)$ and $L(F)$ when the sample arises from a symmetric distribution H different from F.

The following additional assumptions on F will be imposed:

A1: $\psi_f(x) = -f'(x)/f(x)$ is non-decreasing,

A2: $F \in CR(J) \cap CL(g)$ where $J(t) = \psi_f(F^{-1}(t))$ and $g(t) = [I(f)]^{-1} \cdot \psi_f'(F^{-1}(t))$.

The underlying distribution H will be subject to the following restrictions:

B1: H has a density $h(x)$ and is symmetric around zero and $h(x) > 0$ on $\{x \mid 0 < H(x) < 1\}$;

B2: $H \in CR(J) \cap CL(g)$, where J and g are the same as in A2.

For underlying distributions H satisfying these regularity conditions, the asymptotic variances of $n^{\frac{1}{2}} \cdot R(F)$ and $n^{\frac{1}{2}}L(F)$ are given by:

$$\sigma_H^2(R(F)) = I(f)(\int [\psi_f'(F^{-1}(H(x)))/f(F^{-1}(H(x)))]h^2(x)\,dx)^{-2}$$

and $\sigma_H^2(L(F)) = \int_0^1 s^2(u)\,du$ with $s(u) = \int g(H(x))(I_{[u \leqq H(x)]} - H(x))\,dx$ where g is the same as in condition A2.

In the special case where F is equal to the normal distribution Φ, it was shown by Chernoff and Savage (1958) that $\sigma_H^2(R(\Phi)) \leqq \sigma_H^2(L(\Phi))$ for any underlying distribution H subject to above restrictions. Here equality holds if and only if $H(x) = \Phi(ax)$ for some $a > 0$. Gastwirth and Wolff (1968) gave a simpler proof for the same result, and we will use their method to prove the following theorem.

THEOREM. *Let $F \in \mathscr{F}$ satisfy conditions* A1 *and* A2. *Then*

$$\sigma_H^2(R(F)) \leqq \sigma_H^2(L(F)) \tag{1}$$

for all H satisfying B1 *and* B2. *If further $\psi_f' > 0$ where it is defined, then equality in* (1) *holds if and only if $H(x) = F(ax)$ for some $a > 0$.*

PROOF. Let $\psi = \psi_f$ and $g(u) = \psi'(F^{-1}(u)) \cdot I^{-1}(f)$. Since $g \geqq 0$ and $\int_0^1 g(u)\,du = 1$ it follows by Jensen's inequality:

$$\begin{aligned} M(H) &= \left[\int \frac{I(f)h(x)}{f(F^{-1}(H(x)))} \cdot g(H(x))\,dH(x)\right]^{-1} \\ &\leqq I(f)^{-1} \int g(H(x))f(F^{-1}(H(x)))\,dx \\ &= I(f)^{-1} \int_0^x g(H(y))\,dy \cdot f(F^{-1}(H(x)))|_{-\infty}^{+\infty} \\ &\qquad - I(f)^{-1} \int s(H(x))\psi(F^{-1}(H(x)))\,dH(x) \\ &= K(H)\,. \end{aligned}$$

Here we use that $s(H(x)) = -\int_0^x g(H(y))\,dy$. This follows easily from the identity:

$$s(u) = -\int_0^{H^{-1}(u)} g(H(x))\,dx + \int_0^\infty g(H(x))(1 - H(x))\,dx - \int_{-\infty}^0 g(H(x))H(x)\,dx\,.$$

Since $\int |s(F(x))|\, dF(x) < \infty$ one has for suitably chosen sequences u_n: $s(u_n) \times f(F^{-1}(u_n)) \to 0$ as $u_n \to 1$ or $u_n \to 0$. The Cauchy–Schwarz inequality then yields:

$$\sigma_H^2(R(F)) = I(f)M^2(H) \leqq I(f)K^2(H) \leqq I(f)^{-1} \int s^2(H(x))\, dH(x) \cdot \int \psi^2(x)\, dF(x)$$
$$= \int_0^1 s^2(u)\, du = \sigma_H^2(L(F)) .$$

The last assertion of the theorem follows easily by examining the inequalities. This concludes the proof.

REMARKS. The last assertion of the theorem is not necessarily true without the condition $\psi' > 0$, as can be seen by the following example. Let $\psi_k(x) = x$ for $|x| \leqq k$ and $\psi_k(x) = k \cdot \text{sign}(x)$ otherwise ($k > 0$). This ψ_k-function plays an important role in Huber's paper (1964). It corresponds to a distribution F_k which is normal in the middle but has double exponential tails. Since $\psi_k'(x) = 0$ for $|x| > k$, it is seen that the asymptotic variance $\sigma_H^2(R(F_k))$ is not changed if one modifies $H(x)$ for those x for which $H(|x|) > F_k(k)$. A similar remark applies to $\sigma_H^2(L(F_k))$. Here $L(F_k)$ takes on the form of a trimmed mean.

Thus $\sigma_H^2(R(F_k)) = \sigma_H^2(L(F_k))$ for all H which satisfy B1, B2 and $H(x) = F_k(x)$ for $|x| \leqq k$. So far it is not known to the author whether condition A1 is necessary for the assertion of the theorem to hold.

Acknowledgments. I would like to express my deep gratitude to Professor Erich L. Lehmann for his advice and guidance throughout the course of this research. I wish also to thank Professor Olkin for drawing my attention to the paper by Eaton and Olkin (1972) which gives solution methods to related extremal problems.

REFERENCES

BICKEL, P. J. (1965). Some contributions to the theory of order statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 575–591. Univ. of California Press.

CHERNOFF, H., GASTWIRTH, J. L. and JOHNS, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Ann. Math. Statist.* **38** 52–72.

CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29** 972–994.

EATON, M. L. and OLKIN, I. (1972). Applications of the Cauchy–Schwarz inequality to some extremal problems. *Inequalities-III*. Academic Press, New York.

GASTWIRTH, J. L. and WOLFF, S. (1968). An elementary method for obtaining lower bounds on the asymptotic power of rank tests. *Ann. Math. Statist.* **39** 2128–2130.

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

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

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

PURI, M. L. and SEN, P. K. (1971). *Nonparametric Methods in Multivariate Analysis*. Wiley, New York.

STIGLER, S. M. (1969). Linear functions of order statistics. *Ann. Math. Statist.* **40** 770–788.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195