The Annals of Statistics
1980, Vol. 8, No. 2, 279–284

MINIMAX ESTIMATION OF LOCATION PARAMETERS FOR SPHERICALLY SYMMETRIC DISTRIBUTIONS WITH CONCAVE LOSS[1]

BY ANN COHEN BRANDWEIN AND WILLIAM E. STRAWDERMAN

The City University of New York and Rutgers University

For $p > 4$ and one observation X on a p-dimensional spherically symmetric distribution, minimax estimators of θ whose risks are smaller than the risk of X (the best invariant estimator) are found when the loss is a nondecreasing concave function of quadratic loss. For n observations $X_1, X_2, \cdots, X_n$, we have classes of minimax estimators which are better than the usual procedures, such as the best invariant estimator, $\bar{X}$, or a maximum likelihood estimator.

1. Introduction. Since Stein [9] and Brown [5] proved the inadmissibility of the best invariant estimator of the mean vector in 3 or more dimensions, many classes of minimax estimators which are better than the best invariant estimator have been found for quadratic loss and general quadratic loss. (See Brandwein and Strawderman [4] for a discussion of these results.) However, for other loss functions, little is known about improving on the best invariant estimator even for the multivariate normal distribution. (Berger [2] has some results for losses which are polynomials in the coordinates of $(\delta - \theta)$). Here we provide minimax estimators for the mean of a spherically symmetric distribution for the following loss function:

$$L(\delta, \theta) = f(\|\delta - \theta\|^2)$$

where f is a nondecreasing concave function. The risk of any estimator δ is $E_\theta L(\delta, \theta)$. An estimator δ_1 is better than another estimator δ_2, if it has a smaller $(\leq)$ risk for all θ (the risk of δ_1 must dominate the risk of δ_2) and strictly smaller for some θ.

Given X, one observation on a spherically symmetric distribution about θ, we find classes of minimax estimators with respect to any nondecreasing concave loss function of $\|\delta - \theta\|^2$. Specifically, when the loss is a nondecreasing concave function of quadratic loss, e.g. $\|\delta - \theta\|$, for values of "a" given in Theorem 2.1, $\delta_{a,r}(X) = (1 - ar(\|X\|^2)\|X\|^{-2})X$ is better (has smaller risk than X for all θ and strictly smaller for some θ) than X provided (i) $0 < r(\cdot) \leq 1$; (ii) $r(\|X\|^2)$ is nondecreasing; and (iii) $r(\|X\|^2)/\|X\|^2$ is nonincreasing.

Of particular interest is the problem of estimating the mean of a p-dimensional $(p \geq 4)$ multivariate normal distribution for these nonquadratic loss functions. In Section 3, we will find classes of minimax estimators for the normal mean when the loss is $f(\|\delta - \theta\|^2) = \|\delta - \theta\|^q$ for $0 < q < 2$.

Received February 1978; revised January 1979.

[1]Research supported by National Science Foundation Grants MCS77-01654 and MCS75-23343-A01.

AMS 1970 *subject classifications.* Primary 62C99; secondary 62F10, 62H99.

Key words and phrases. Minimax estimation, spherically symmetric, multivariate, location parameter.

It is important to recognize that the results given in Section 2 apply to the multiple observation case. When sampling n observations $X_1, X_2, \cdots, X_n$, the problem becomes an estimation problem for one observation, since any spherically symmetric, translation invariant estimator based on $X_1, X_2, \cdots, X_n$, also has a spherically symmetric distribution. This situation is discussed in Section 4.

2. Minimax estimators of location parameters with respect to concave functions of quadratic loss. Consider X a $p \times 1$ random vector having a spherically symmetric distribution about θ. It is well known [8] that the random vector X is the best invariant procedure and minimax with respect to the following loss:

$$L(\delta, \theta) = f(\|\delta - \theta\|^2) \tag{2.1}$$

where $f(\cdot)$ is a nondecreasing concave function and

$$\|\delta - \theta\|^2 = \Sigma_{i=1}^p(\delta_i - \theta_i)^2.$$

The problem of finding minimax estimators which are better than X with respect to the loss $L(\delta, \theta)$ given in (2.1) has been solved for the special case of quadratic loss $f(\cdot)$, the identity (see Brandwein [3]). However, for a general concave $f(\cdot)$, such results have not been given until now, as far as we know, even for the multivariate normal distribution. In this section, we find classes of minimax estimators with respect to loss (2.1), which are better than X for $p \geqslant 4$.

The estimators we look at are of the form

$$\delta_{a,r}(X) = \big(1 - \big(ar(\|X\|^2)/\|X\|^2\big)\big)X. \tag{2.2}$$

Using these estimators, Brandwein [3] proved the following theorem for quadratic loss (i.e., $f(u) \equiv u$):

THEOREM A. *If X has a p-dimensional ($p \geqslant 4$) spherically symmetric distribution about θ, then the risk of $\delta_{a,r}(X)$, given above by* (2.2), *is less than or equal to that of X for all θ with strict inequality for some θ provided*:

(1) $0 < r(\cdot) \leqslant 1$,

(2) $r(\|X\|^2)$ *is nondecreasing*,

(3) $r(\|X\|^2)/\|X\|^2$ *is nonincreasing*,

and

(4) $0 < a \leqslant (2(p-2)/p)/E_0(\|X\|^{-2})$

where E_0 denotes the expected value when $\theta = \mathbf{0}$.

To prove $\delta_{a,r}(X)$ is better than $\delta_{0,r}(X) = X$ under the conditions of this theorem, Brandwein shows that if,

$$\Delta_{a,r}(X) = \|X - \theta\|^2 - \|\delta_{a,r}(X) - \theta\|^2, \tag{2.3}$$

then the difference in risks,

$$R(\delta_{0,r}, \theta) - R(\delta_{a,r}, \theta) = E_\theta\Delta_{a,r}(X) \geqslant 0$$

for all θ and strictly positive for $\theta = \mathbf{0}$. Specifically, if G is any cdf for

$R = \|X - \theta\|$, and conditions (1)–(4) of Theorem A hold, then

$$E_\theta \Delta_{a,r}(X) = \int E_\theta(\Delta_{a,r}(X)|R)\, dG(R) \geq 0 \quad \text{for} \quad 0 \leq a \leq (2(p-2)/p)/E_G R^{-2};$$

(2.4) and

$$E_0 \Delta_{a,r}(X) = \int E_0(\Delta_{a,r}(X)|R)\, dG(R) > 0 \quad \text{for} \quad 0 < a \leq (2(p-2)/p)/E_G R^{-2} \quad \text{and} \quad 0 < r(\cdot) \leq 1$$

where E_G denotes the expected value when R has distribution function G. Since the conditional distribution of $X|(\|X - \theta\|^2 = R^2)$, (denoted by $X|R$), has a spherical uniform distribution on the surface of the sphere $(\|X - \theta\|^2 = R^2)$, $E_0(\|X\|^{-2}) = ER^{-2}$. This with (2.4) implies the theorem is true.

With these preliminary remarks given, we can now prove the following theorem.

THEOREM 2.1. *If X has a p-dimensional ($p \geq 4$) spherically symmetric distribution about θ then $\delta_{a,r}(X)$ given by* (2.2), *is better than X and is thus minimax with respect to the concave loss $f(\|\delta - \theta\|^2)$ given by* (2.1), *provided*:

(i) $0 \leq r(\cdot) \leq 1$,

(ii) $r(\|X\|^2)$ *is nondecreasing,*

(iii) $r(\|X\|^2)/\|X\|^2$ *is nonincreasing,*

(iv) $0 < E_G f'(R^2) < \infty$,

and

(v) $0 \leq a \leq (2(p-2)/p)/E_H R^{-2}$

where $H(R) = \int_0^R f'(S^2)\, dG(S) / \int_0^\infty f'(S^2)\, dG(S)$, *$G$ is the* cdf *of R and E_H denotes the expected value when R has* cdf *H. Moreover, if $0 < r(\cdot) \leq 1$ and $0 < a \leq (2(p-2)/p)/E_H R^{-2}$, $\delta_{a,r}$ is better than X.*

PROOF. We begin by showing that the difference in risks, between $\delta_{0,r}(X) = X$ and $\delta_{a,r}(X)$ is nonnegative.

Clearly,

$$\begin{aligned} R(\delta_{0,r}, \theta) - R(\delta_{a,r}, \theta) &= E_\theta f(\|X - \theta\|^2) - E_\theta f(\|\delta_{a,r}(X) - \theta\|^2) \\ &= E\big[E_\theta\big[f(\|X - \theta\|^2) \\ &\quad - f(\|X - \theta\|^2 - \Delta_{a,r}(X)) \big| \|X - \theta\|^2 = R^2\big]\big] \\ &= Ef(R^2) - E_\theta f(R^2 - \Delta_{a,r}(X)) \end{aligned} \tag{2.5}$$

where $\Delta_{a,r}(X)$ is given by (2.3).

Since $f(\cdot)$ is a nondecreasing concave function, for any points s and t, $f(s) \leq f(t) + f'(t)(s - t)$. Thus, $f(R^2 - \Delta_{a,r}(X)) \leq f(R^2) + f'(R^2)(-\Delta_{a,r}(X))$. This, together with (2.5) implies that

$$\begin{aligned} R(\delta_{0,r}, \theta) - R(\delta_{a,r}, \theta) &\geq E_\theta f'(R^2)\Delta_{a,r}(X) = E\big[f'(R^2)E_\theta(\Delta_{a,r}(X)|R)\big] \\ &= \int f'(R^2) E_\theta(\Delta_{a,r}(X)|R)\, dG(R) \end{aligned} \tag{2.6}$$

$$= \int f'(R^2)\, dG(R) \left[\frac{\int E_\theta(\Delta_{a,r}(X)|R) f'(R^2)\, dG(R)}{\int f'(R^2)\, dG(R)} \right]$$

$$= \int f'(R^2)\, dG(R) \int E_\theta(\Delta_{a,r}(X)|R)\, dH(R)$$

where G is the cdf of $R = \|X - \theta\|$ and $H(R) = \int_0^R f'(S^2)\, dG(S) / \int_0^\infty f'(S^2) dG(S)$. Moreover, since $f(\cdot)$ is nondecreasing, $f'(R^2) \geqslant 0$ and thus H is a cdf by (iv).

We see from (2.4) that

$$\int E_\theta(\Delta_{a,r}(X)|R)\, dH(R) \geqslant 0 \tag{2.7}$$

if $p \geqslant 4$ and $0 \leqslant a \leqslant (2(p-2)/p) E_H R^{-2}$. Since $\int f'(R^2)\, dG(R) > 0$, combining (2.7) we have $R(\delta_{0,r}, \theta) - R(\delta_{a,r}, \theta) \geqslant 0$ for $0 \leqslant a \leqslant (2(p - 2/p)/E_H R^{-2}$.

Moreover, since by (2.4) $\int E_0(\Delta_{a,r}(X)|R)\, dH(R) > 0$ for $0 < a \leqslant (2(p - 2)/p) E_H R^{-2}$ and $0 < r(\cdot) \leqslant 1$, $\delta_{a,r}(X)$ is better than X for a in this interval, provided $\int f'(R^2)\, dG(R) > 0$, which is so by (iv).

This completes the proof of the theorem.

Note that when $r(\cdot) \equiv 1$, our class of minimax estimators reduces to the Stein class of estimators $\delta_a(X) = (1 - (a/\|X\|^2))X$. In addition, as stated in [3] due to the work of Baranchik [1], it is clear that $\delta_a^+(X) = \max(0, (1 - a/\|X\|^2))X$ is also minimax with respect to quadratic loss. Thus, it is clear from the proof of Theorem 2.1, that for the same a's, δ_a^+ will also be minimax with respect to the loss function $f(\|\delta - \theta\|^2)$.

3. Examples.

3.1. *Multivariate normal distribution.* Of particular interest is the problem of estimating the mean θ of a multivariate normal distribution with covariance matrix the identity (MVN(θ, I)), for certain nondecreasing concave loss functions of $\|\delta - \theta\|^2$. As is well known, the multivariate normal distribution is the only spherically symmetric distribution with independent coordinates (see Kac [7]).

For quadratic loss, we know from the work of James and Stein [6] and Baranchik [1], that $\delta_{a,r}(X) = (1 - a(r(\|X\|^2)/\|X\|^2))X$ is minimax for $0 \leqslant a \leqslant 2(p-2)$ and better than X for $0 < a < 2(p-2)$ provided $r(\|X\|^2)$ is a nondecreasing function and $0 < r(\cdot) \leqslant 1$. We will now show that for $0 < a \leqslant 2(p-2)(1 - 3/p)$, and, in addition to the other assumptions on $r(\cdot)$, $r(\|X\|^2)/\|X\|^2$ is nonincreasing, $\delta_{a,r}(X)$ is better than X (and thus minimax) when the loss is $\|\delta - \theta\|$ and $p \geqslant 4$. This will be a simple application of Theorem 2.1.

If $f(\|\delta - \theta\|^2) = \|\delta - \theta\|$, $f(u) = u^{\frac{1}{2}}$ and it follows that $f'(R^2) = 1/(2R)$. Also, it is well known that when $X \sim$ MVN(θ, I) that $\|X - \theta\|^2 = R^2$ has a chi-square distribution with p degrees of freedom (χ_p^2). So, R has a density which we will denote by $g(R)$ and $E_H R^{-2}$ is given by

$$E_H R^{-2} = \int_0^\infty R^{-3} g(R)\, dR \Big/ \int_0^\infty R^{-1} g(R)\, dR$$

$$= E_G R^{-3} / E_G R^{-1}.$$

Since $R^2 \sim \chi_p^2$,

$$E_H R^{-2} = E_G R^{-3} / E_G R^{-1} = \Gamma\left(\frac{p-3}{2}\right) / 2\Gamma\left(\frac{p-1}{2}\right) = 1/(p-3).$$

Applying Theorem 2.1, $\delta_{a,r}(X)$ is better than X under conditions (i)–(iv) for $0 < a \leqslant 2(p-2)(1-3/p)$, when the loss is $\|\delta - \theta\|$.

Similarly, with respect to the class of loss functions,

$$L(\delta, \theta) = \|\delta - \theta\|^q \quad \text{for} \quad 0 \leqslant q \leqslant 2$$

we have $E_H R^{-2} = (p + q - 4)^{-1}$. Therefore, $\delta_{a,r}(X)$ is better than X when $0 < a \leqslant 2(p-2)(1-(4-q)/p)$.

We have therefore exhibited a class of minimax estimators which are better than X for the multivariate normal distribution tor losses other than quadratic loss.

3.2. *Uniform distribution on the sphere* $(\|X - \theta\|^2 \leqslant S^2)$. If X has a p-dimensional uniform distribution on $(\|X - \theta\|^2 \leqslant S^2)$ then, $R = \|X - \theta\|$ has a density, of the form

$$g(R) = pR^{p-1}/S^p \quad \text{for} \quad 0 \leqslant R \leqslant S$$
$$= 0 \quad \text{elsewhere.}$$

With respect to quadratic loss, Brandwein and Strawderman proved in [4] that $\delta_{a,r}(X)$ is better than X for $0 < a \leqslant (2(p-2)/(p+2))S^2$ if $p > 4$. When H is the cdf defined in Theorem 2.1 and the loss is $\|\delta - \theta\|$, $E_H R^{-2} = \int_0^S R^{p-4}\, dR \,/ \int_0^S R^{p-2}\, dR = ((p-1)/(p-3))S^2$, and so $\delta_{a,r}(X)$ is better than X for $0 < a \leqslant (2(p-2)/(p+2))S^2[1 - 6/(p(p-1))]$.

Similarly, for the general loss function $L(\delta, \theta) = \|\delta - \theta\|^q$ for $0 \leqslant q \leqslant 2$, we have $\delta_{a,r}(X)$ is better than X for $0 < a \leqslant (2(p-2)/(p+2))S^2[1 - (2(4-q)/(p(p+q-2)))]$.

4. Multiple observations. When sampling n observations, $X_1, X_2, \cdots, X_n$ on a spherically symmetric distribution about θ, the problem of estimating θ reduces to the problem of taking one observation. In [4], Brandwein shows that estimators based on n observations, which are spherically symmetric and translation invariant also have spherically symmetric distributions about θ. Moreover, Pitman's estimator, a maximum likelihood estimator, and $\bar{X}$ are all such estimators. Thus, we have classes of estimators which improve on the usual estimators based on n observations for the nonquadratic loss functions.

We remark that these results have potential applicability to improving on robust invariant estimators of multivariate location parameters (see also remarks in Brandwein [3]).

REFERENCES

[1] BARANCHIK, A. J. (1964). Multiple regression and estimation of the mean of a multivariate normal distribution. Stanford Univ., Technical Report No. 51.

[2] BERGER, JAMES O. (1978). Minimax estimation of a multivariate normal mean under polynomial loss. *J. Multivariate Anal.* **8** 173–180.

[3] BRANDWEIN, ANN C. (1979). Minimax estimation of location parameters for spherically symmetric distributions under general quadratic loss. *J. Multivariate Anal.* **9**.

[4] BRANDWEIN, ANN C. and STRAWDERMAN, W. E. (1978). Minimax estimation of location parameters for spherically symmetric unimodal distributions. *Ann. Statist.* **6** 377–416.

[5] BROWN, L. D. (1966). On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Statist.* **37** 1087–1136.

[6] JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Probability* **1** 361–379 Univ. California Press.

[7] KAC, M. (1945). A remark on independence of linear and quadratic forms involving independent Gaussian variables. *Ann. Math. Statist.* **16** 400–401.

[8] KIEFER, J. (1957). Invariance, minimax sequential estimation and continuous time processes. *Ann. Math. Statist.* **28** 537–601.

[9] STEIN, C. (1955). Inadmissibility of the usual estimator for the mean vector of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Probability* **1** 197–206 Univ. California Press.

STATISTICS DEPARTMENT
BARUCH COLLEGE
THE CITY UNIVERSITY OF NEW YORK
17 LEXINGTON AVE.
NEW YORK, N.Y. 10010

DEPARTMENT OF STATISTICS
HILL CENTER FOR MATH SCIENCE
RUTGERS UNIVERSITY
NEW BRUNSWICK, N.J. 08904