The Annals of Statistics
1976, Vol. 4, No. 1, 223–226

ADMISSIBLE MINIMAX ESTIMATION OF A MULTIVARIATE NORMAL MEAN WITH ARBITRARY QUADRATIC LOSS

BY JAMES O. BERGER

Purdue University

The problem of estimating the mean of a p-variate ($p \geqq 3$) normal distribution is considered. It is assumed that the covariance matrix Σ is known and that the loss function is quadratic. A class of minimax estimators is given, out of which admissible minimax estimators are developed.

1. Introduction. Let $X = (X_1, \cdots, X_p)^t$ be an observation from a p-variate normal population with mean vector $\theta = (\theta_1, \cdots, \theta_p)^t$ and known positive definite covariance matrix Σ. Assume that the loss incurred in estimating θ by $\delta(X) = (\delta_1(X), \cdots, \delta_p(X))^t$ is the quadratic loss $(\delta - \theta)^t Q(\delta - \theta)$. Assume also that Q is positive definite and that $p \geqq 3$.

The special situation $Q = \Sigma^{-1}$ has been considered by several authors (James and Stein (1960), Baranchik (1970), Strawderman (1971) among them), and wide classes of minimax and admissible estimators of θ have been found. Results for arbitrary Q and Σ, however, are incomplete. Bhattacharya (1966) and Bock (1975) found some particular minimax estimators for the general situation. In this paper, a different and simpler class of minimax estimators is given, out of which admissible minimax estimators are developed.

2. A class of minimax estimators. Define $||x||^2 = x^t\Sigma^{-1}Q^{-1}\Sigma^{-1}x$, and let I denote the $p \times p$ identity matrix. Estimators of the form

$$\delta(X) = (I - r(||X||^2)Q^{-1}\Sigma^{-1}/||X||^2)X \tag{1}$$

will be considered, where r is a measurable function from $R^1 \rightarrow R^1$.

THEOREM 1. *The estimator* δ, *given by* (1), *is minimax if*

(i) $0 \leqq r(\cdot) \leqq 2(p-2)$, *and*
(ii) $r(\cdot)$ *is nondecreasing*.

PROOF. This theorem has since been extended to a more general theorem, dealing with a wide class of densities, in Berger (1975). The above theorem is a special case of Theorem 3 in that paper. □

It has been brought to the author's attention that the above result was independently discovered by Malcolm Hudson (1974).

3. Admissible minimax estimators. The obvious question which arises is how

Received July 1974; revised May 1975.
AMS 1970 *subject classifications.* Primary 62C15; Secondary 62F10, 62H99.
Key words and phrases. Admissible, minimax, normal mean, generalized Bayes, quadratic loss.

should the function r, in (1), be chosen? In this section, choices of r which give rise to admissible minimax estimators are developed.

Let α denote the smallest characteristic root of the matrix ΣQ. The following choices of r will be considered:

$$r_c(t) = \frac{t \int_0^\alpha \lambda^{(p/2-c+1)} \exp\{-\lambda t/2\}\, d\lambda}{\int_0^\alpha \lambda^{(p/2-c)} \exp\{-\lambda t/2\}\, d\lambda}, \qquad c < 1 + p/2 . \tag{2}$$

A simple integration by parts in the numerator of the above expression gives

$$r_c(t) = (p - 2c + 2) - \frac{2\alpha^{(p/2-c+1)} \exp\{-\alpha t/2\}}{\int_0^\alpha \lambda^{(p/2-c)} \exp\{-\lambda t/2\}\, d\lambda} . \tag{3}$$

Note that when $Q = \Sigma^{-1}$ (and hence $\alpha = 1$), r_c gives rise to the admissible minimax estimator found in Strawderman (1971).

THEOREM 2. *Assume that δ is given by* (1), *with $r = r_c$.*

(a) *If $3 - p/2 \leqq c < 1 + p/2$, then δ is minimax.*
(b) *If $3 - p/2 \leqq c < 2$, then δ is admissible.*
(c) *If $3 - p/2 \leqq c < 1$, then δ is proper Bayes.*

PROOF. To prove (a), it is only necessary to verify conditions (i) and (ii) of Theorem 1. From (2), it is clear that $r_c(\cdot) > 0$. Using (3) and the assumption that $c \geqq 3 - p/2$, it is also clear that $r_c(\cdot) < 2(p - 2)$. Hence condition (i) of Theorem 1 is satisfied. From (3) and the fact that $\exp\{(\alpha - \lambda)t/2\}$ is nondecreasing in t for $0 \leqq \lambda \leqq \alpha$, it follows that $r_c(t)$ is nondecreasing in t. Condition (ii) of Theorem 1 is thus satisfied and the conclusion follows.

To prove (b) and (c), δ must first be shown to be a generalized Bayes estimator. The notation will be considerably simplified by considering only the case $Q = I$ and $\Sigma = A$, where A is a $p \times p$ diagonal matrix with diagonal elements $a_i > 0$. Since Q and Σ are positive definite, it is easy to check that the problem can always be transformed into this diagonal case. Note that $||X||^2 = \sum X_i^2/a_i^2$ and that $\alpha = \min\{a_i\}$. For notational convenience, define $b_i(\lambda) = a_i(a_i - \lambda)/\lambda$.

For $c < 1 + p/2$, consider the generalized prior density

$$g_c(\theta) = \int_0^\alpha \left[\prod_{i=1}^p b_i(\lambda)^{-\frac{1}{2}}\right] \exp\{-\tfrac{1}{2} \textstyle\sum_{i=1}^p \theta_i^2/b_i(\lambda)\}\lambda^{-c}\, d\lambda . \tag{4}$$

It is easy to check that $g_c(\cdot)$ is a bounded function for the given choice of c. (Clearly $b_i(\lambda)$ behaves like a_i^2/λ near $\lambda = 0$.) Note also that g_c has finite mass if $c < 1$.

The generalized Bayes estimator of θ, with respect to g_c, is given componentwise by

$$\delta_i^c(X) = \frac{\int \theta_i \exp\{-\frac{1}{2} \sum_{i=1}^p (X_i - \theta_i)^2/a_i\} g_c(\theta)\, d\theta}{\int \exp\{-\frac{1}{2} \sum_{i=1}^p (X_i - \theta_i)^2/a_i\} g_c(\theta)\, d\theta} . \tag{5}$$

Consider first the numerator of the above expression. Using the definition of $g_c(\theta)$, interchanging orders of integration (g_c is a bounded function), and completing squares gives

$$(6)\qquad \int \theta_i \exp\{-\tfrac{1}{2}\textstyle\sum_{i=1}^{p}(X_i-\theta_i)^2/a_i\}g_c(\theta)\,d\theta = \int_0^\alpha \int_{R^p}\theta_i \exp\{\varphi\}[\prod_{j=1}^{p} b_j(\lambda)^{-\frac{1}{2}}]\lambda^{-c}\,d\theta\,d\lambda\,,$$

where

$$\varphi = -\tfrac{1}{2}\textstyle\sum_{j=1}^{p}([a_j^{-1}+b_j(\lambda)^{-1}][\theta_j - X_j/\{1+a_jb_j(\lambda)^{-1}\}]^2 + X_j^2/\{a_j+b_j(\lambda)\})\,.$$

Integrating out over θ, on the right hand side of (6), gives that the numerator of (5) equals

$$\int_0^\alpha [X_i/\{1+a_ib_i(\lambda)^{-1}\}][\prod_{j=1}^{p}\{1+a_j^{-1}b_j(\lambda)\}^{-\frac{1}{2}}] \exp\{-\tfrac{1}{2}\textstyle\sum_{j=1}^{p}[X_j^2/\{a_j+b_j(\lambda)\}]\}\lambda^{-c}\,d\lambda\,.$$

Using the identities

$$a_j + b_j(\lambda) = a_j + a_j(a_j-\lambda)/\lambda = a_j^2/\lambda\,,$$
$$\{1+a_jb_j(\lambda)^{-1}\}^{-1} = 1-\lambda/a_j\,,$$
$$\{b_j(\lambda)a_j^{-1}+1\}^{-\frac{1}{2}} = (\lambda/a_j)^{\frac{1}{2}}\,,$$

it is thus clear that the numerator of (5) equals

$$(7)\qquad \int_0^\alpha (1-\lambda/a_i)X_i\exp\{-\lambda\|X\|^2/2\}\lambda^{(p/2-c)}[\prod_{j=1}^{p}a_j^{-\frac{1}{2}}]\,d\lambda\,.$$

It can similarly be shown that

$$(8)\qquad \int \exp\{-\tfrac{1}{2}\textstyle\sum_{i=1}^{p}(X_i-\theta_i^2)/a_i\}g_c(\theta)\,d\theta = \int_0^\alpha \exp\{-\lambda\|X\|^2/2\}\lambda^{(p/2-c)}[\prod_{j=1}^{p}a_j^{-\frac{1}{2}}]\,d\lambda\,.$$

Combining (5), (7), (8), and (2) gives

$$\delta_i^c(X) = \frac{\int_0^\alpha(1-\lambda/a_i)X_i\exp\{-\lambda\|X\|^2/2\}\lambda^{(p/2-c)}\,d\lambda}{\int_0^\alpha\exp\{-\lambda\|X\|^2/2\}\lambda^{(p/2-c)}\,d\lambda} = [1-r_c(\|X\|^2)/(a_i\|X\|^2)]X_i\,.$$

Thus $\delta=\delta^c$ is indeed the generalized Bayes estimator with respect to g_c. Part (c) of the theorem follows immediately from this and the observation that g_c has finite mass if $c<1$.

To prove part (b) of the theorem, a result from Brown (1971) will be used. Note first that (8) and a change of variables give

$$\begin{aligned} f^*(X) &= \int_{R^p}[\prod_{i=1}^{p}(2\pi a_i)^{-\frac{1}{2}}]\exp\{-\tfrac{1}{2}\textstyle\sum_{i=1}^{p}(X_i-\theta_i)^2/a_i\}g_c(\theta)\,d\theta \qquad \text{(definition)}\\ &= (2\pi)^{-p/2}[\prod_{i=1}^{p}a_i^{-1}]\int_0^\alpha\exp\{-\lambda\|X\|^2/2\}\lambda^{(p/2-c)}\,d\lambda\\ &= (2\pi)^{-p/2}[\prod_{i=1}^{p}a_i^{-1}]\|X\|^{-(p-2c+2)}\int_0^{\alpha\|X\|^2}\exp\{-\lambda/2\}\lambda^{(p/2-c)}\,d\lambda\\ &\leqq K|X|^{(2c-p-2)}\,. \end{aligned}$$

(Here $|X|$ denotes the usual Euclidean norm of X.) Using Corollary 4.3.4 and Theorem 5.1.1 (B) of Brown (1971), together with the assumption that $c<2$, it can be concluded that δ is admissible. □

Note that for the estimator of Theorem 2 to be minimax, it is necessary to have $p\geqq 3$. Clearly admissible, minimax estimators of the given form do exist for $p\geqq 3$. Proper Bayes versions, however, exist only if $p\geqq 5$.

REFERENCES

[1] BARANCHIK, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. *Ann. Math. Statist.* **41** 642–645.

[2] BERGER, J. (1975). Minimax estimation of location vectors for a wide class of densities. *Ann. Statist.* **3** 1318–1328.

[3] BHATTACHARYA, P. K. (1966). Estimating the mean of a multivariate normal population with general quadratic loss function. *Ann. Math. Statist.* **37** 1819–1824.

[4] BOCK, M. E. (1975). Minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.* **3** 209–218.

[5] BROWN, L. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Statist.* **42** 855–903.

[6] HUDSON, M. (1974). Empirical Bayes estimation. Technical Report # 58, Stanford Univ.

[7] JAMES, W. and STEIN, C. (1960). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 361–379. Univ. of California Press.

[8] STRAWDERMAN, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Statist.* **42** 385–388.

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
MATHEMATICAL SCIENCES BUILDING
WEST LAFAYETTE, INDIANA 47907