The Annals of Statistics
1975, Vol. 3, No. 1, 209–218

MINIMAX ESTIMATORS OF THE MEAN OF A MULTIVARIATE NORMAL DISTRIBUTION[1]

BY M. E. BOCK

University of Illinois

An extension of Strawderman's results yields minimax admissible estimates for the mean of a p-variate normal distribution where the known nonsingular covariance matrix is not necessarily the identity and $p > 2$. Minimax estimators for the case where the covariance matrix is unknown are also given.

1. Introduction and summary. Let X have the p-variate normal distribution with unknown mean vector θ and known nonsingular covariance matrix D. Define the risk of an estimator $\hat{\theta}(X)$ of θ to be

$$R(\hat{\theta}, \theta) = E_\theta[(\hat{\theta}(X) - \theta)'(\hat{\theta}(X) - \theta)] .$$

Let $p > 2$ and $g(X) = X$. Charles Stein [10] has shown that the usual minimax estimator g is inadmissible. For $D = I_p$, Strawderman [11] exhibited a class of admissible minimax estimators which were proper Bayes and spherically symmetric. Strawderman [12] also proved that there were no proper Bayes spherically symmetric minimax estimators for $p < 5$ if $D = I_p$. This paper extends these results to the case where the covariance matrix D is not necessarily the identity matrix. Certain forms of minimax estimators are also considered for the case where the covariance matrix is unknown.

In Section 2, a spherically symmetric estimator $\hat{\theta}$ is defined to be of the form $\hat{\theta}(X) = h(X'D^{-1}X)X$ where h is a real-valued function. Let d_L be the largest eigenvalue of D. It is shown that if $\operatorname{tr} D \leqq 2d_L$, no spherically symmetric minimax estimator which is essentially different from g exists, as noted independently by Brown [6]. For $\operatorname{tr} D > 2d_L$ a class of spherically symmetric minimax estimators is given and the class coincides with one given by Baranchik [1] for $D = I_p$. A subset of estimators in the class is exhibited which are proper Bayes, and thus, admissible. For $D = I_p$, these are the estimators given by Strawderman [11]. For $\operatorname{tr} D/d_L \leqq p/2 + 2$, it is shown that no proper Bayes spherically symmetric minimax estimators exist.

In Section 3, X has a p-variate normal distribution with unknown mean vector θ and covariance matrix $\sigma^2 D$. Let d_L be the largest eigenvalue of the known nonsingular matrix D. Assume $p > 2$ and σ^2 is an unknown positive constant. A random variable S is given such that (S/σ^2) has a chi-square (n)

Received March 1973; revised February 1974.

[1] This research is based on the author's Ph. D. dissertation submitted at the University of Illinois at Urbana-Champaign, October, 1973, financed by the Dept. of Agric. Econ.

Key words and phrases. Minimax, spherically symmetric, point estimation, admissible.

distribution independent of X. The risk is

$$R_1(\hat{\theta}; \theta, \sigma^2) = E_{\theta,\sigma^2}[(\hat{\theta}(X, S) - \theta)'(\hat{\theta}(X, S) - \theta)/\sigma^2] .$$

A class of minimax estimators of the form $h(X'D^{-1}X/S)X$ is given where h is a real-valued function. Unless $h(\cdot) = 1$ a.e., estimators of this form are minimax only if $\operatorname{tr} D > 2d_L$. In another formulation X is assumed to be p-variate normal with unknown mean vector θ and unknown nonsingular covariance matrix D. A random matrix $\mathscr{S}$ with Wishart (D, m, p) distribution independent of X is given. The risk for an estimator $\hat{\theta}(X, \mathscr{S})$ of θ is

$$R_2(\hat{\theta}; \theta, D) = E_{\theta,D}[(\hat{\theta}(X, \mathscr{S}) - \theta)'(\hat{\theta}(X, \mathscr{S}) - \theta)/\operatorname{tr} D]$$

and $p > 2$. The usual estimator $g(X, \mathscr{S}) = X$ is minimax, but it is essentially the only minimax estimator of the form $h(X'\mathscr{S}^{-1}X)X$ where h is a real-valued function.

2. Known covariance matrix. Assume X has a p-variate normal distribution with unknown mean vector θ and known nonsingular covariance matrix D. Let $\delta(X) = h(X'D^{-1}X)X$, where h is a real-valued function. For the risk function $R(\delta, \theta) = E_\theta(\delta(X) - \theta)'(\delta(X) - \theta)$ the following expression is obtained via Corollaries 1 and 2, Appendix:

$$R(\delta, \theta) = \operatorname{tr} D Eh^2(\chi^2_{(p+2,\theta'D^{-1}\theta)}) + \theta'\theta[Eh^2(\chi^2_{(p+4,\theta'D^{-1}\theta)}) - 2Eh(\chi^2_{(p+2,\theta'D^{-1}\theta)}) + 1] \tag{1}$$

where $\chi^2_{(j,\lambda)}$ denotes a chi-square (j) random variable with noncentrality parameter λ.

Let d_L be the largest characteristic root of D. Lemmas 1 and 2 are used to prove Theorem 1 which says that no minimax spherically symmetric estimator essentially different from $g(X) = X$ exists if $\operatorname{tr} D \leqq 2d_L$.

LEMMA 1. *Assume $p \geqq 2$. Unless $h(\cdot) = 1$ a.e., there exists $\lambda_\delta \geqq 0$ such that*

$$0 < 2(Eh^2(\chi^2_{(p+2,\lambda_\delta)}) - 1) + \lambda_\delta[Eh^2(\chi^2_{(p+4,\lambda_\delta)}) - 2Eh(\chi^2_{(p+2,\lambda_\delta)}) + 1] .$$

PROOF. Let X be given as above. For $p = 2$ and $D = I_p$, let $\delta_0(X) = h(X'X + \chi^2_{(n)})X$, where $\chi^2_{(n)}$ is a random variable with chi-square (n) distribution independent of X (let $\chi^2_{(n)} \equiv 0$ if $n = 0$). In view of (1), risk

$$R(\delta_0, \theta) = 2Eh^2(\chi^2_{(4+n,\theta'\theta)}) + \theta'\theta[Eh^2(\chi^2_{(6+n,\theta'\theta)}) - 2Eh(\chi^2_{(4+n,\theta'\theta)}) + 1] .$$

For $p = 2$ and $D = I_p$, the estimator $g(X) = X$ is minimax admissible with constant risk $R(g, \theta) = 2$. Farrel [8] has shown that an admissible estimator is essentially unique. Thus, unless $h(\cdot) = 1$ a.e., there exists a number $\lambda_{(h,n)} \geqq 0$ such that

$$0 < R(\delta_0, \theta) - R(g, \theta) = 2(Eh^2(\chi^2_{(4+n,\lambda_{(h,n)})}) - 1) + \lambda_{(h,n)}[Eh^2(\chi^2_{(6+n,\lambda_{(h,n)})}) - 2Eh(\chi^2_{(4+n,\lambda_{(h,n)})}) + 1]$$

for all 2×1 vectors θ such that $\theta'\theta = \lambda_{(h,n)}$.

Let $n = p - 2$ where $p \geqq 2$ and $\lambda_\delta = \lambda_{(h,p-2)}$ so that there exists $\lambda_\delta \geqq 0$ and

$$0 < 2(Eh^2(\chi^2_{(p+2,\lambda_\delta)}) - 1) + \lambda_\delta[Eh^2(\chi^2_{(p+4,\lambda_\delta)}) - 2Eh(\chi^2_{(p+2,\lambda_\delta)}) + 1]\,. \qquad \square$$

LEMMA 2. *Unless* $h(\cdot) = 1$ a.e., *if* $\delta(X) = h(X'D^{-1}X)X$ *is minimax and* $p > 2$, *then* $Eh^2(\chi^2_{(p+2,\lambda_\delta)}) < 1$ *where* λ_δ *is given in Lemma* 1.

PROOF. An estimator $\hat{\theta}$ is minimax if $R(\hat{\theta}, \theta) \leqq \operatorname{tr} D$ for all $p \times 1$ vectors θ. Let d_s be the smallest characteristic root of D. Unless $h(\cdot) = 1$ a.e., choose a $p \times 1$ vector θ_0 such that $\theta_0'\theta_0 d_s^{-1} = \theta_0' D^{-1}\theta_0 = \lambda_\delta$ where λ_δ is given in Lemma 1. Equation (1) and the minimaxity of δ imply that

$$0 \geqq \operatorname{tr} D(Eh^2(\chi^2_{(p+2,\lambda_\delta)}) - 1) + d_s\lambda_\delta[Eh^2(\chi^2_{(p+4,\lambda_\delta)}) - 2Eh(\chi^2_{(p+2,\lambda_\delta)}) + 1]\,.$$

Assume $Eh^2(\chi^2_{(p+2,\lambda_\delta)}) \geqq 1$. Then $\operatorname{tr} D/d_s \geqq p > 2$ and the above inequalities imply

$$0 \geqq d_s[2(Eh^2(\chi^2_{(p+2,\lambda_\delta)}) - 1) + \lambda_\delta[Eh^2(\chi^2_{(p+4,\lambda_\delta)}) - 2Eh(\chi^2_{(p+2,\lambda_\delta)}) + 1]]\,.$$

But the right-hand side of the above inequality is positive by Lemma 1. Thus, $Eh^2(\chi^2_{(p+2,\lambda_\delta)}) < 1$ unless $h(\cdot) = 1$ a.e. $\square$

THEOREM 1. *Let X have p-variate normal distribution with unknown mean θ and known nonsingular covariance matrix D. Assume $p > 2$ and d_L is the largest characteristic root of D. If* $\operatorname{tr} D \leqq 2d_L$, *then no estimator of the form* $\delta(X) = h(X'D^{-1}X)X$ *is a minimax estimator for θ under the quadratic loss* $(\hat{\theta} - \theta)'(\hat{\theta} - \theta)$ *where h is a real-valued function unless* $h(\cdot) = 1$ a.e.

PROOF. Assume $\delta(X) = h(X'D^{-1}X)X$ is minimax with $p > 2$, and $\operatorname{tr} D \leqq 2d_L$. Unless $h(\cdot) = 1$ a.e., choose a $p \times 1$ vector θ_0 such that $\theta_0'D^{-1}\theta_0 = \theta_0'\theta_0 d_L^{-1} = \lambda_\delta$ where λ_δ is given in Lemma 1. Then equation (1) and the minimaxity of δ imply that

$$0 \geqq \operatorname{tr} D(Eh^2(\chi^2_{(p+2,\lambda_\delta)}) - 1) + d_L\lambda_\delta[Eh^2(\chi^2_{(p+4,\lambda_\delta)}) - 2Eh(\chi^2_{(p+2,\lambda_\delta)}) + 1]\,.$$

Since $\operatorname{tr} D \leqq 2d_L$ and $Eh^2(\chi^2_{(p+2,\lambda_\delta)}) < 1$ (by Lemma 2), by Lemma 1 the right-hand side of the above inequality is positive, a contradiction. Thus, δ is not minimax unless $h(\cdot) = 1$ a.e. $\square$

The result given in Theorm 1 was given independently by Brown [6].

The following class of minimax spherically symmetric estimators is a generalization of a class given by Baranchik [1] for $D = I_p$.

THEOREM 2. *Let* $\operatorname{tr} D \geqq 2d_L$ *and* $p > 2$ *and* $r\colon [0, \infty) \to [0, 1]$. $\hat{\theta}(X) = (1 - cr(X'D^{-1}X)(X'D^{-1}X)^{-1})X$ *is a minimax estimator for θ if* $0 \leqq c \leqq 2((\operatorname{tr} D)\, d_L^{-1} - 2)$ *and r is monotone non-decreasing.*

PROOF. It suffices to show $R(\hat{\theta}, \theta) \leqq \operatorname{tr} D$ for all θ. By Corollaries 1 and 2, Appendix, and setting $r^*(a) = r(a)/a$

$$\begin{aligned} R(\hat{\theta}, \theta) - \operatorname{tr} D = c^2\{&(\operatorname{tr} D)E[(r^*(\chi^2_{(p+2+2K)}))^2] + \theta'\theta E[(r^*(\chi^2_{p+4+2K}))^2]\} \\ &+ 2c\theta'\theta E[r^*(\chi^2_{p+2+2K})] - 2c\{(\operatorname{tr} D)E[r^*(\chi^2_{p+2+2K})] \\ &+ \theta'\theta E[r^*(\chi^2_{p+4+2K})]\} \end{aligned}$$

where K is a Poisson $(\theta' D^{-1}\theta/2)$ random variable. Furthermore, by Lemmas 3 and 4, Appendix, and setting $\alpha(\theta) = \theta'\theta/\theta' D^{-1}\theta$,

$$\begin{aligned} R(\hat{\theta}, \theta) &- \operatorname{tr} D \\ &= c(\operatorname{tr} D)E[r(\chi^2_{p-2+2K})(p+2K)^{-1}(p-2+2K)^{-1}\{(c \cdot r(\chi^2_{p-2+2K}) \\ &\quad - 2\chi^2_{p-2+2K})(1+\alpha(\theta)2K(\operatorname{tr} D)^{-1}) + 2K(\operatorname{tr} D)^{-1}\alpha(\theta)2(p+2K)\}] \\ &\leqq c(\operatorname{tr} D)E[r(\chi^2_{p-2+2K})(p+2K)^{-1}(p-2+2K)^{-1}\{2(1+\alpha(\theta)2K(\operatorname{tr} D)^{-1}) \\ &\quad \times (p-2+2K-\chi^2_{p-2+2K}) + (c-2(p-2)) + \alpha(\theta)2K(\operatorname{tr} D)^{-1} \\ &\quad \times (c - 2(\operatorname{tr} D(\alpha(\theta))^{-1} - 2))\}] \end{aligned}$$

(since $r(\cdot) \leqq 1$). Since $c \leqq 2((\operatorname{tr} D)d_L^{-1} - 2) \leqq 2((\alpha(\theta))^{-1} \operatorname{tr} D - 2)$ and $c \leqq 2(p-2)$ (because $\operatorname{tr} D \leqq pd_L$),

$$\begin{aligned} R(\hat{\theta}, \theta) - \operatorname{tr} D &\leqq (\operatorname{tr} D)cE[r(\chi^2_{p-2+2K})(p+2K)^{-1}(p-2+2K)^{-1} \\ &\quad \times 2(1+\alpha(\theta)2K(\operatorname{tr} D)^{-1})(p-2+2K-\chi^2_{p-2+2K})] \\ &\leqq 0 \end{aligned}$$

(by Lemma 5, Appendix). □

For $p > 2$ and $\hat{\theta}_1(X) = (1 - c(X'D^{-1}X)^{-1})X$, Theorem 2 implies that $\hat{\theta}_1$ is minimax if $0 \leqq c \leqq 2((\operatorname{tr} D)d_L^{-1} - 2)$. If $D = I_p$, then $\hat{\theta}_1$ is the estimator given by James and Stein [9], which dominates the usual one, g. Theorem 3 shows that the bound on c given is precise.

THEOREM 3. *Let* $\hat{\theta}_1(X) = (1 - c(X'D^{-1}X)^{-1})X$ *and* $\operatorname{tr} D \geqq 2d_L$; *then* $\hat{\theta}_1$ *is not minimax if* $c > 2((\operatorname{tr} D)d_L^{-1} - 2)$ *and* $p > 2$.

PROOF. Assume $\theta' D^{-1} > 0$ is given and choose θ so that $\theta'\theta/\theta' D^{-1}\theta = d_L$. Then as in the proof of Theorem 2 with $r(\cdot) \equiv 1$,

$$\begin{aligned} R(\hat{\theta}_1, \theta) - \operatorname{tr} D &= E[c(\operatorname{tr} D)(p+2K)^{-1}(p-2+2K)^{-1}\{(c-2(p-2)) \\ &\quad + d_L(\operatorname{tr} D)^{-1}2K(c - 2(\operatorname{tr} Dd_L^{-1} - 2))\}] \\ &= E[c(\operatorname{tr} D)(p+2K)^{-1}(p-2+2K)^{-1}(p+2+2K)^{-1} \\ &\quad \times \{(c-2(p-2))(p+2+2K) \\ &\quad + d_L(\operatorname{tr} D)^{-1}(p-2+2K)\theta' D^{-1}\theta(c - 2(\operatorname{tr} Dd_L^{-1} - 2))\}], \end{aligned}$$

(by Lemma 3, Appendix). This is clearly positive if $\theta' D^{-1}\theta > 0$ and $c \geqq 2(p-2)$ even if $\operatorname{tr} D < 2d_L$. Assume $c < 2(p-2)$. Then

$$\begin{aligned} R(\hat{\theta}_1, \theta) - \operatorname{tr} D &= E[c(\operatorname{tr} D)(p+2K)^{-1}(p-2+2K)^{-1}(p+2+2K)^{-1} \\ &\quad \times \{-4(2(p-2)-c) \\ &\quad + d_L(\operatorname{tr} D)^{-1}(p-2+2K)\theta' D^{-1}\theta(c - 2(\operatorname{tr} Dd_L^{-1} - 2)) \\ &\quad - (2(p-2)-c)\operatorname{tr} D(\theta' D^{-1}\theta)^{-1}d_L^{-1})\}] \\ &> E[c(\operatorname{tr} D)(p+2K)^{-1}(p-2+2K)^{-1} \\ &\quad \times (p+2+2K)^{-1}\{-4(2(p-2)-c) \\ &\quad + d_L(\operatorname{tr} D)^{-1}(p-2+2K)\tfrac{7}{8}\theta' D^{-1}\theta(c - 2(\operatorname{tr} Dd_L^{-1} - 2))\}] \end{aligned}$$

if $\theta' D^{-1}\theta > 8(2(p-2)-c)(\text{tr } D)(c-2(\text{tr } Dd_L^{-1}-2))^{-1}d_L^{-1}$. Thus $R(\theta_1, \theta) - (\text{tr } D) > 0$ if $\theta' D^{-1}\theta > 8(2(p-2)-c)(\text{tr } D)(c-2(\text{tr } Dd_L^{-1}-2))^{-1}d_L^{-1}$, since $(p-2+2K) \geqq 1$. □

For $(\text{tr } D/d_L) > (p/2)+2$, the following estimator δ_1 is an example of a proper Bayes (and, thus, admissible) minimax spherically symmetric estimator. It is a generalization of the estimator given by Strawderman [9] for $D = I_p$. Let the conditional distribution of θ given λ be p-variate normal with zero mean and covariance matrix $D(1-\lambda)\lambda^{-1}$ where the unconditional density of λ is given by $\lambda^{-a}(1-a)$ for $0 < \lambda \leqq 1$. Let a be chosen such that $a < 1$ and such that $(\text{tr } D/d_L) \geqq p/2 + 3 - a$. The proper Bayes estimator with respect to this prior is

$$\delta_1(X) = \left[1 - \left(\frac{p+2-2a}{X'D^{-1}X} - \frac{2\exp[-\frac{1}{2}X'D^{-1}X]}{(X'D^{-1}X)\{\int_0^1 \lambda^{((p/2)-a)}\exp[-\lambda X'D^{-1}X/2]\,d\lambda\}}\right)\right]X.$$

It follows from Theorem 2 that δ_1 is minimax, setting $r(y) = 1 - [(p/2+1-a)\int_0^1 \lambda^{(p/2)-a}\exp[(1-\lambda)(y/2)\,dy]^{-1}$ for $y \geqq 0$ and $c = p+2-2a$. Theorem 4 demonstrates that the restriction on $\text{tr } D/d_L$ is necessary.

THEOREM 4. *No spherically symmetric estimator is proper Bayes minimax if* $\text{tr } D/d_L \leqq (p/2)+2$.

PROOF. Let $\delta(X) = h(X'D^{-1}X)X$ where h is a real-valued function and define $\omega(\cdot) = 1 - h(\cdot)$. If δ is minimax

$$\begin{aligned} 0 \geqq R(\delta,\theta) - \text{tr } D = {} & E[\omega^2(X'D^{-1}X)X'X] - 2\,\text{tr } DE[\omega(\chi^2_{(p+2,\theta'D^{-1}\theta)})] \\ & - 2\theta'\theta E[\omega(\chi^2_{(p+4,\theta'D^{-1}\theta)})] + 2\theta'\theta E[\omega(\chi^2_{(p+2,\theta'D^{-1}\theta)})] \end{aligned}$$

(by Corollaries 1 and 2, Appendix). Thus by Jensen's Inequality and Corollary 1, Appendix,

$$\begin{aligned} 0 \geqq {} & (E[\omega(\chi^2_{(p+2,\theta'D^{-1}\theta)})])^2\theta'\theta - 2\,\text{tr } DE[\omega(\chi^2_{(p+2,\theta'D^{-1}\theta)})] \\ & - 2\theta'\theta E[\omega(\chi^2_{(p+4,\theta'D^{-1}\theta)})] + 2\theta'\theta E[\omega(\chi^2_{(p+2,\theta'D^{-1}\theta)})]. \end{aligned}$$

For a given value of $\theta'D^{-1}\theta$ we may choose θ such that $\theta'\theta = d_L\theta'D^{-1}\theta$ and the above inequality becomes

$$\begin{aligned} 0 \geqq {} & d_L(E[\omega(\chi^2_{(p+2,\theta'D^{-1}\theta)})])^2\theta'D^{-1}\theta - 2\,\text{tr } DE[\omega(\chi^2_{(p+2,\theta'D^{-1}\theta)})] \\ & - 2d_L\theta'D^{-1}\theta E[\omega(\chi^2_{(p+4,\theta'D^{-1}\theta)})] + 2d_L\theta'D^{-1}\theta E[\omega(\chi^2_{(p+2,\theta'D^{-1}\theta)})]. \end{aligned}$$

Define $\psi(\lambda) = \lambda E[\omega(\chi^2_{(p+2,\lambda)})]$. Then

$$\frac{d}{d\lambda}(\psi(\lambda)) = E[\omega(\chi^2_{(p+2,\lambda)})] + \frac{\lambda}{2}\{E[\omega(\chi^2_{(p+4,\lambda)})] - E[\omega(\chi^2_{(p+2,\lambda)})]\}$$

so that the above inequality implies

$$\frac{d}{d\lambda}(\psi(\lambda)) \geqq \frac{\psi(\lambda)}{4\lambda}(\psi(\lambda) - 2(\text{tr } D/d_L - 2)).$$

Replacing "p" by "$\text{tr } D/d_L$" in the proof given by Strawderman [12], it may be shown that

$$0 \leqq E[\omega(\chi^2_{(p+2,\theta' D^{-1}\theta)})] \leqq 2(\operatorname{tr} D/d_L - 2)(\theta' D^{-1}\theta)^{-1}.$$

Furthermore, using the above inequality, the proof for Theorem 2 of Strawderman [11] implies that no estimator of the form $h(X'D^{-1}X)X$ is proper Bayes minimax if $\operatorname{tr} D/d_L \leqq p/2 + 2$. More detail is given in Bock [5]. □

The results of Brown [7] imply that the estimator $\hat{\theta}$ of Theorem 2 is admissible if and only if $\hat{\theta}$ is generalized Bayes and $\lim_{t\to\infty} cr(t) \geqq (p-2)$. Thus the estimator δ_1 is an admissible generalized Bayes spherically symmetric minimax estimator if the unconditional prior "density" for λ is λ^{-a} for $0 < \lambda \leqq 1$ and $a \leqq 2$. So for $\operatorname{tr} D/d_L \geqq p/2 + 1$, there exist admissible spherically symmetric minimax estimators.

Note that for other forms of loss functions such as the ones considered by Basar and Mintz [3], one may find proper Bayes estimators which are minimax because they are least favorable. No least favorable distribution for θ exists here.

3. Unknown covariance matrix. Assume X has a p-variate normal distribution with mean vector θ and covariance matrix $\sigma^2 D$ where σ^2 is an unknown positive constant, D is a known nonsingular matrix and $p > 2$. Let S be an independent random variable such that (S/σ^2) has a chi-square (n) distribution. (Regression is an example of this.) Redefine the risk for an estimator $\hat{\theta}$ of θ to be

$$R_1(\hat{\theta}; \theta, \sigma^2) = E_{(\theta,\sigma^2)}[(\hat{\theta} - \theta)'(\hat{\theta} - \theta)/\sigma^2].$$

Let $g(X) = X$ and note that g is minimax with constant risk, $\operatorname{tr} D$. The estimators given in Theorem 5 dominate g or have the same risk function.

THEOREM 5. *Assume* $r: [0, \infty] \to [0, 1]$ *is monotone non-decreasing. Let* $0 \leqq c < 2(\operatorname{tr} D/d_L - 2)(n+2)^{-1}$ *and assume* $\operatorname{tr} D > 2d_L$. *Then* $\hat{\theta}$ *is minimax where* $\hat{\theta}(X) = (1 - r(X'D^{-1}X/S)(X'D^{-1}X/S)^{-1})X$.

PROOF. It snffices to show $R_1(\hat{\theta}; \theta, \sigma^2) \leqq \operatorname{tr} D$ for all (θ, σ^2). Setting $\alpha(\theta) = \theta'\theta/\theta' D^{-1}\theta$ and letting K be a Poisson $(\theta' D^{-1}\theta/2\sigma^2)$ random variable, as in Theorem 2,

$$\begin{aligned} R_1(\hat{\theta}; \theta, \sigma^2) - \operatorname{tr} D \\ &\leqq c \operatorname{tr} DE[r(\chi^2_{p-2+2K}/\chi_n^2)\chi_n^2(p+2K)^{-1}(p-2+2K)^{-1} \\ &\quad \times \{2(1 + \alpha(\theta)2K(\operatorname{tr} D)^{-1})(p - 2 + 2K - \chi^2_{p-2+2K}) \\ &\quad + (c\chi_n^2 - 2(p-2)) + \alpha(\theta)2K(\operatorname{tr} D)^{-1}(c\chi_n^2 - 2(\operatorname{tr} D(\alpha(\theta))^{-1} - 2))\}] \\ &= c(n+2) \operatorname{tr} DE\Big[r(\chi^2_{p-2+2K}/\chi^2_{n+2})(p+2K)^{-1}(p-2+2K)^{-1} \\ &\quad \times \Big\{2(1 + \alpha(\theta)2K(\operatorname{tr} D)^{-1})\Big((p - 2 + 2K - \chi^2_{p-2+2K}) \\ &\quad + \frac{c}{2}(\chi^2_{n+2} - (n+2))\Big) + c(n+2) - 2(p-2) \\ &\quad + \alpha(\theta)2K(\operatorname{tr} D)^{-1}(c(n+2) - 2(\operatorname{tr} D(\alpha(\theta))^{-1} - 2))\Big\}\Big] \end{aligned}$$

by Lemma 4, Appendix. Applying Lemma 5, Appendix, to the above

$$R_1(\hat{\theta}; \theta, \sigma^2) - \operatorname{tr} D \leqq c(n+2)\operatorname{tr} DE[r(\chi^2_{p-2+2K}/\chi^2_{n+2})(p+2K)^{-1}(p-2+2K)^{-1} \times \{c(n+2) - 2(p-2) + \alpha(\theta)(\operatorname{tr} D)^{-1}2K(c(n+2) - 2(\operatorname{tr} D(\alpha(\theta))^{-1} - 2))\}] .$$

The above expression is $\leqq 0$ if $0 < c \leqq 2(\operatorname{tr} D/d_L - 2)(n+2)^{-1}$. □

For $D = I_p$, the theorem is given by Baranchik [2].

The following theorem shows that the assumption that $\operatorname{tr} D/d_L$ be greater than 2 is necessary for the minimaxity in Theorem 5 unless $h(\cdot) = 1$ a.e.

THEOREM 6. *If $p > 2$ and $\operatorname{tr} D \leqq 2d_L$ and h is a real-valued function, no estimator of the form $\delta(X, S) = h(X'D^{-1}X/S)X$ is minimax for θ under the quadratic loss function $(\hat{\theta} - \theta)'(\hat{\theta} - \theta)/\sigma^2$ unless $h(\cdot) = 1$ a.e.*

PROOF. Assume $\operatorname{tr} D \leqq 2d_L$. Using the proof of Theorem 1, it may be shown that for $\sigma^2 = 1$ there is a value of θ for which $R_1(\delta; \theta, 1) > \operatorname{tr} D$ unless $h(\cdot) = 1$ a.e. □

Theorem 5 and a proof similar to that of Theorem 3 gives the following theorem.

THEOREM 7. *For $p > 2$, let $\hat{\theta}_1(X, S) = (1 - cS(X'D^{-1}X)^{-1})X$ and $\operatorname{tr} D \geqq 2d_L$ and $c > 0$; then $\hat{\theta}_1$ is minimax if and only if $c \leqq 2(\operatorname{tr} D/d_L - 2)(n+2)^{-1}$.*

$\hat{\theta}_1$ is the estimator given by James and Stein [9] if $D = I_p$. Alternative forms of estimators have been given by Bhattacharya [4].

As an aside, consider the case where X has p-variate normal distribution with unknown mean θ and unknown covariance matrix D. Let $\mathscr{S}$ be a random matrix having independent Wishart distribution with m degrees of freedom and $E\mathscr{S} = mD$ where $m > p - 1$. Define the risk of an estimator $\hat{\theta}(X, \mathscr{S})$ of θ to be

$$R_2(\hat{\theta}; \theta, D) = E_{\theta,D}[(\hat{\theta}(X, \mathscr{S}) - \theta)'(\hat{\theta}(X, \mathscr{S}) - \theta)/\operatorname{tr} D] .$$

Then $g(X, \mathscr{S}) = X$ is minimax with constant risk, 1, but estimators of the form $\hat{\theta}(X, \mathscr{S}) = h(X'\mathscr{S}^{-1}X)X$ where h is real-valued are not minimax unless $h(\cdot) = 1$ a.e. This may be seen by noting that $X'\mathscr{S}^{-1}X$ is distributed as $X'D^{-1}X/S$ where S is independent of X and has χ^2_{m-p+1} distribution, according to Wijsman [13]. As in the proof of Theorem 6 (with $n = m - p + 1$, $\sigma^2 = 1$) for D such that $\operatorname{tr} D \leqq 2d_L$, there is a value of θ for which $R_2(\hat{\theta}; \theta, D) > 1$. Thus the estimator g is essentially the only minimax estimator of the form $h(X'\mathscr{S}^{-1}X)X$.

Acknowledgment. The author wishes to thank George Judge for comments and discussions. Thanks are due also to Larry Brown for helpful suggestions.

APPENDIX

Other authors seem to be aware of these corollaries but we are unaware of proofs of results as general as those given here.[2]

[2] Assume throughout that $\chi^2_{(j,\lambda)}$ has a chi-square (j) distribution with noncentrality parameter λ.

THEOREM A. *Let Y have p-variate normal distribution with mean η and identity covariance matrix. Let $h\colon [0, \infty) \to (-\infty, +\infty)$. Then for $\eta' = [\eta_1, \cdots, n_p]$ and $Y' = [Y_1, \cdots, Y_p]$, $E[h(Y'Y)Y_i] = \eta_i Eh(\chi^2_{(p+2,\eta'\eta)})$.*

PROOF. The Y_i's are independent. Therefore,

$$
\begin{aligned}
&E[h(Y'Y)Y_i]\\
&\quad = E\{E[h(Y_i^2 + \textstyle\sum_{j\neq i} Y_j^2)Y_i \mid \sum_{j\neq i} Y_j^2]\}\\
&\quad = E\left[e^{-\eta^2_i/2}\left(\int_{-\infty}^{+\infty} h(x^2 + \textstyle\sum_{j\neq i} Y_j^2)\,\dfrac{xe^{-x^2/2}e^{x\eta_i}}{(2\pi)^{\frac{1}{2}}}\,dx\right)\right]\\
&\quad = E\left[\frac{e^{-\eta_i^2/2}}{(2\pi)^{\frac{1}{2}}}\left\{\int_0^\infty h(y + \textstyle\sum_{j\neq i} Y_j^2)e^{-y/2}(e^{\eta_i y^{\frac{1}{2}}} - e^{-\eta_i y^{\frac{1}{2}}})\,\dfrac{dy}{2}\right\}\right]\\
&\quad = E\left[\frac{e^{-\eta_i^2/2}}{2(2\pi)^{\frac{1}{2}}}\left\{\int_0^\infty h(y + \textstyle\sum_{j\neq i} Y_j^2)e^{-y/2}\left(\sum_{k=0}^\infty \dfrac{2(\eta_i y^{\frac{1}{2}})^{2k+1}}{(2k+1)!}\right)dy\right\}\right]\\
&\quad = \eta_i E\left[\int_0^\infty h(y + \textstyle\sum_{j\neq i} Y_j^2)e^{-\eta_i^2/2}\left(\sum_{k=0}^\infty \dfrac{\left(\frac{\eta_i^2}{2}\right)^k}{k!}\,\dfrac{y^{[(2k+3)/2]-1}e^{-y/2}}{\Gamma\left(\frac{2k+3}{2}\right)2^{(2k+3)/2}}\right)dy\right]
\end{aligned}
$$

(because $\Gamma(2k) = \Gamma(k)\Gamma(k+\frac{1}{2})2^{2k-1}/\pi^{\frac{1}{2}}$). Thus

$$
\begin{aligned}
E[h(Y'Y)Y_i] &= \eta_i E[h(\chi^2_{(3,\eta_i^2)} + \textstyle\sum_{j\neq i} Y_j^2)] = \eta_i E[h(\chi^2_{(3,\eta_i^2)} + \chi^2_{(p-1,\sum_{j\neq i}\eta_j^2)})]\\
&= \eta_i E[h(\chi^2_{(p+2,\eta'\eta)})]\,. \qquad \square
\end{aligned}
$$

COROLLARY 1. *Let X have p-variate normal distribution with mean θ and nonsingular covariance matrix D. Let $h\colon [0, \infty) \to (-\infty, +\infty)$. Then*

$$E[h(X'D^{-1}X)X] = \theta Eh(\chi^2_{(p+2,\theta'D^{-1}\theta)})\,.$$

THEOREM B. *Given the hypotheses of Theorem A, we have*

$$E[h(Y'Y)Y_i^2] = E[h(\chi^2_{(p+2,\eta'\eta)})] + \eta_i^2 E[h(\chi^2_{(p+4,\eta'\eta)})]\,.$$

PROOF. Note that Y_i^2's are independent. Therefore,

$$
\begin{aligned}
&E[h(Y'Y)Y_i^2]\\
&\quad = E\{E[h(Y_i^2 + \textstyle\sum_{j\neq i} Y_i^2)Y_i^2 \mid \sum_{j\neq i} Y_j^2]\}\\
&\quad = E\left\{e^{-\eta^2_i/2}\sum_{k=0}^\infty \dfrac{(\eta_i^2/2)^k}{k!}\,E[h(\chi^2_{1+2k} + \textstyle\sum_{j\neq i} Y_j^2)\chi^2_{1+2k} \mid \sum_{j\neq i} Y_j^2]\right\}\\
&\quad = E\left\{e^{-\eta_i^2/2}\sum_{k=0}^\infty \dfrac{(\eta_i^2/2)^k}{k!}\,(1+2k)E[h(\chi^2_{(3+2k)} + \textstyle\sum_{j\neq i} Y_j^2) \mid \sum_{j\neq i} Y_j^2]\right\}\\
&\quad = E[h(\chi^2_{(3,\eta_i^2)} + \textstyle\sum_{j\neq i} Y_j^2)] + \left\{e^{-\eta_i^2/2}\sum_{k=0}^\infty \dfrac{(\eta_i^2/2)^k}{k!}\,(2k)Eh(\chi^2_{3+2k} + \textstyle\sum_{j\neq i} Y_j^2)\right\}\\
&\quad = E[h(\chi^2_{(3,\eta_i^2)} + \textstyle\sum_{j\neq i} Y_j^2)]\\
&\qquad + \left\{e^{-\eta_i^2/2}\eta_i^2\sum_{k=1}^\infty \dfrac{(\eta_i^2/2)^{k-1}}{(k-1)!}\,E[h(\chi^2_{5+2(k-1)} + \textstyle\sum_{j\neq i} Y_j^2)]\right\}
\end{aligned}
$$

$$= E[h(\chi^2_{(p+2,\eta'\eta)})] + (\eta_i^2)E[h(\chi^2_{5,\eta_i^2} + \textstyle\sum_{j\neq i} Y_j^2)]$$

(because $\sum_{j\neq i} Y_j^2 \sim \chi^2_{(p-1,\Sigma_{j\neq i}\eta_j^2)}$ and because $\sum_{j\neq i} Y_j^2$ and Y_i^2 are independent)

$$= E[h(\chi^2_{(p+2,\eta'\eta)})] + (\eta_i^2)E[h(\chi^2_{(p+4,\eta'\eta)})]\,, \qquad i = 1, \cdots, p\,. \ \square$$

COROLLARY 2. *Let W be a $p \times p$ positive definite matrix and assume the hypotheses of Corollary* 1. *Then*

$$E[h(X'D^{-1}X)X'WX] = \operatorname{tr}(WD)E[h(\chi^2_{(p+2,\theta'D^{-1}\theta)})] + \theta'W\theta E[h(\chi^2_{(p+4,\theta'D^{-1}\theta)})]\,.$$

LEMMA 3. *Let ϕ be a real-valued measurable function defined on the integers. Let $K \sim$ Poisson $(\lambda/2)$. Then if both sides exist,*

$$\lambda E[\phi(K)] = E[2K\phi(K-1)]\,.$$

LEMMA 4. *Let $h\colon [0, \infty) \to (-\infty, +\infty)$. Then if both sides exist,*

$$E[h(\chi^2_{(m)})] = E\left[\frac{mh(\chi^2_{(m+2)})}{\chi^2_{(m+2)}}\right].$$

LEMMA 5. *Let $s\colon [0, \infty) \to (0, \infty)$ and $t\colon [0, \infty) \to [0, \infty)$ be monotone non-decreasing and monotone non-increasing functions, respectively. Let W be a nonnegative random variable. Assume $E(W)$, $E(s(W))$, $E(Ws(W))$, $E(t(W))$, $E(Wt(W))$ exist and are finite. Then*

$$E[s(W)(E(W) - W)] \leqq 0 \leqq E[t(W)(E(W) - W)]\,.$$

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

[3] BASAR, T. and MINTZ, M. (1973). On a minimax estimate for the mean of a normal random vector under a generalized quadratic loss function. *Ann. Statist.* **1** 127–134.

[4] BHATTACHARYA, P. K. (1966). Estimating the mean of a multivariate normal population with general quadratic loss function. *Ann. Math. Statist.* **37** 1818–1825.

[5] BOCK, M. E. (1974). Certain minimax estimators of the mean of a multivariate normal distribution. Ph. D. Thesis, Univ. of Illinois.

[6] BROWN, L. D. (1973). Estimation with incompletely specified loss functions. Submitted to *J. Amer. Statist. Assoc.*

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

[8] FARRELL, R. (1964). Estimators of a location parameter in the absolutely continuous case. *Ann. Math. Statist.* **35** 949–999.

[9] JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 361–379.

[10] STEIN, C. (1955). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** 197–206.

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

[12] STRAWDERMAN, W. (1972). On the existence of proper Bayes minimax estimators of the mean of a multivariate normal distribution. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 51–55.

[13] WIJSMAN, R. (1957). Random orthogonal transformations and their use in some classical distribution problems in multivariate analysis. *Ann. Math. Statist.* **28** 415–423.

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
548 MATH. SCI.
WEST LAFAYETTE, INDIANA 47907