Title	A note on estimating the mean vector of a multivariate normal distribution with general quadratic loss function
Sub Title	
Author	篠崎, 信雄(Shinozaki, Nobuo)
Publisher	慶応義塾大学工学部
Publication year	1974
Jtitle	Keio engineering reports Vol.27, No.7 (1974.) ,p.105- 112
JaLC DOI	
Abstract	We investigate the problem of estimating the mean vector of a multivariate normal random vector x with general quadratic loss function. Some estimators which are different from BHATTACHARYA's one and have simpler form are shown to have uniformly smaller risk than x. A generalization of BHATTACHARYA's result is also given.
Notes	
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50001004-00270007-0105

KEIO ENGINEERING REPORTS
VOL. 27 NO. 7

A NOTE ON ESTIMATING THE MEAN VECTOR OF A MULTIVARIATE NORMAL DISTRIBUTION WITH GENERAL QUADRATIC LOSS FUNCTION

By

NOBUO SHINOZAKI

FACULTY OF ENGINEERING
KEIO UNIVERSITY
YOKOHAMA 1974

KEIO ENGINEERING REPORTS
VOL. 27, NO. 7, pp. 105-112, 1974

A NOTE ON ESTIMATING THE MEAN VECTOR OF A MULTIVARIATE NORMAL DISTRIBUTION WITH GENERAL QUADRATIC LOSS FUNCTION

NOBUO SHINOZAKI

Dept. of Mathematics, Keio University, Yokohama, 223 Japan

(Received, Jul. 10, 1974)

ABSTRACT

We investigate the problem of estimating the mean vector of a multivariate normal random vector x with general quadratic loss function. Some estimators which are different from BHATTACHARYA's one and have simpler form are shown to have uniformly smaller risk than x. A generalization of BHATTACHARYA's result is also given.

1. Introduction

Let $\boldsymbol{x}$ be a p-dimensional random vector, $p \geqq 3$, having a normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\sigma^2 I$. For estimating $\boldsymbol{\mu}$ with loss function

$$L_1(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}}) = (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})'(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})$$

and known σ^2, STEIN (1955) has shown that $\boldsymbol{x}$ is inadmissible. An estimator with uniformly smaller risk than that of $\boldsymbol{x}$ has been given by JAMES and STEIN (1961) for the case where σ^2 is unknown and an independent estimator of σ^2 is available, which is distributed as $\sigma^2\chi^2_f$ (chi-square with f degrees of freedom). The estimator has been improved upon by BARANCHIK (1964). ALAM and THOMPSON (1964) have considered a family of estimators that dominate $\boldsymbol{x}$. BARANCHIK (1970) has shown that $\boldsymbol{x}$ is dominated by a general class of estimators, including those given by the above authors. The class has been extended further by STRAWDERMAN (1973).

BHATTACHARYA (1966) has given an estimator with uniformly smaller risk than that of $\boldsymbol{x}$ under the loss function

$$L_2'(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}}) = (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})' D (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}),$$

where D is a diagonal matrix. He also mentioned that the more general problem where the covariance matrix of $\boldsymbol{x}$ is $\sigma^2\Sigma$ instead of $\sigma^2 I$ and the diagonal matrix D involved in the loss function is replaced by M, Σ and M being arbitrary but known positive definite matrices, can be reduced to the above simpler problem.

In this paper we also consider the problem of estimating the mean vector of a multivariate normal distribution with general quadratic loss function. Some estimators which are different from BHATTACHARYA's one and have simpler form are shown to have uniformly smaller risk than $\boldsymbol{x}$. A generalization of BHATTACHARYA's result is also given. An application of the estimators gives an improvement of the least squares estimator of the parameter vector in a usual linear statistical model with normal errors. (See SCLOVE (1968).)

2. Main results

Given a p-dimensional ($p\geqq 3$) normal random vector $\boldsymbol{x}$ with unknown mean vector $\boldsymbol{\mu}$ and covariance matrix of the form $\sigma^2\Sigma$, where Σ is known and nonsingular, and, given a statistic S which is independent of $\boldsymbol{x}$, and is distributed as $\sigma^2\chi^2_f$, the problem we are going to study is to estimate $\boldsymbol{\mu}$ when the loss function is given by

$$L_1(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}})=(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})'(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}}),$$

or more generally, by

$$L_2(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}})=(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})'M(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}}),$$

where M is a given positive definite matrix.

In Theorems 1 and 2 we show that relative to each of the loss functions L_1 and L_2 a class of estimators of the form $c\boldsymbol{x}$ dominates $\boldsymbol{x}$ under some conditions, where c is a scalar function of $\boldsymbol{x}$ and S. In Theorems 3 and 4 we give a class of estimators of the form $A\boldsymbol{x}$ which always dominates $\boldsymbol{x}$, where A is a matrix depending on $\boldsymbol{x}$ and S. In doing these we need the following lemmas.

Lemma 1. Let $\boldsymbol{z}$ be a p-dimensional random vector having a normal distribution with mean vector $\boldsymbol{\xi}$ and covariance matrix I and put $v=\boldsymbol{z}'\boldsymbol{z}$. Let $r(\cdot)$ be a bounded measurable function. Let K be a random variable having a Poisson distribution with mean $\boldsymbol{\xi}'\boldsymbol{\xi}/2$. Then the following relations hold:

$$(1)\qquad E\left\{z_i\frac{r(v)}{v}\right\}=\frac{2\xi_i}{\boldsymbol{\xi}'\boldsymbol{\xi}}E\left[KE\left\{\frac{r(\chi^2_{p+2K})}{\chi^2_{p+2K}}\middle|K\right\}\right],\qquad \text{if } \boldsymbol{\xi}\neq\boldsymbol{0},$$

$$=0,\qquad \text{if } \boldsymbol{\xi}=\boldsymbol{0}.$$

$$(2)\qquad E\left\{z_i^2\frac{r(v)}{v}\right\}=E\left[\left\{\frac{1}{p+2K}+\frac{\xi_i^2}{\boldsymbol{\xi}'\boldsymbol{\xi}}\frac{2K}{p+2K}\right\}E\{r(\chi^2_{p+2K})|K\}\right],\qquad \text{if } \boldsymbol{\xi}\neq\boldsymbol{0},$$

$$=p^{-1}E\{r(\chi^2_p)\},\qquad \text{if } \boldsymbol{\xi}=\boldsymbol{0}.$$

$$(3)\qquad E\left\{z_i^2\frac{r(v)}{v^2}\right\}=E\left[\left\{\frac{1}{p+2K}+\frac{\xi_i^2}{\boldsymbol{\xi}'\boldsymbol{\xi}}\frac{2K}{p+2K}\right\}E\left\{\frac{r(\chi^2_{p+2K})}{\chi^2_{p+2K}}\middle|K\right\}\right],\qquad \text{if } \boldsymbol{\xi}\neq\boldsymbol{0},$$

$$=p^{-1}E\{r(\chi^2_p)/\chi^2_p\},\qquad \text{if } \boldsymbol{\xi}=\boldsymbol{0}.$$

These can be proved along the same line as in BARANCHIK (1973).

Lemma 2. Let $\boldsymbol{x}$ be a p-dimensional random vector having a normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\sigma^2\Sigma$ and put $v=\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}$. Let $r(\cdot)$ be a bounded measurable function. Let K be a random variable having a Poisson distribution with mean $\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}/(2\sigma^2)$. Then we have the following relations.

I (1) $$E\left\{\boldsymbol{x}'\boldsymbol{\mu}\frac{r(v)}{v}\right\}=\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}E\left[2KE\left\{\frac{r(\sigma^2\chi^2_{p+2K})}{\chi^2_{p+2K}}\middle|K\right\}\right], \quad \text{if } \boldsymbol{\mu}\neq\boldsymbol{0},$$
$$=0, \quad \text{if } \boldsymbol{\mu}=\boldsymbol{0}.$$

(2) $$E\left\{\boldsymbol{x}'\boldsymbol{x}\frac{r(v)}{v}\right\}=E\left[\left\{\frac{tr\Sigma}{p+2K}+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2K}{p+2K}\right\}E\{r(\sigma^2\chi^2_{p+2K})|K\}\right], \quad \text{if } \boldsymbol{\mu}\neq\boldsymbol{0},$$
$$=p^{-1}tr\Sigma E\{r(\sigma^2\chi^2_p)\}, \quad \text{if } \boldsymbol{\mu}=\boldsymbol{0}.$$

(3) $$E\left[\boldsymbol{x}'\boldsymbol{x}\frac{r(v)}{v^2}\right\}=\sigma^{-2}E\left[\left\{\frac{tr\Sigma}{p+2K}+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2K}{p+2K}\right\}E\left\{\frac{r(\sigma^2\chi^2_{p+2K})}{\chi^2_{p+2K}}\middle|K\right\}\right], \quad \text{if } \boldsymbol{\mu}\neq\boldsymbol{0},$$
$$=(p\sigma^2)^{-1}tr\Sigma E\{r(\sigma^2\chi^2_p)/\chi^2_p\}, \quad \text{if } \boldsymbol{\mu}=\boldsymbol{0}.$$

II (1) $$E\left\{\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}\frac{r(v)}{v}\right\}=E[2KE\{r(\sigma^2\chi^2_{p+2K})/\chi^2_{p+2K}\}|K], \quad \text{if } \boldsymbol{\mu}\neq\boldsymbol{0},$$
$$=0, \quad \text{if } \boldsymbol{\mu}=\boldsymbol{0}.$$

(2) $$E\left[\boldsymbol{x}'\Sigma^{-2}\boldsymbol{x}\frac{r(v)}{v}\right\}=\sigma^{-2}E\left[\left\{\frac{tr\Sigma^{-1}}{p+2K}+\frac{\boldsymbol{\mu}'\Sigma^{-2}\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2K}{p+2K}\right\}E\left\{\frac{r(\sigma^2\chi^2_{p+2K})}{\chi^2_{p+2K}}\middle|K\right\}\right], \quad \text{if } \boldsymbol{\mu}\neq\boldsymbol{0},$$
$$=(p\sigma^2)^{-1}tr\Sigma^{-1}E\{r(\sigma^2\chi^2_p)/\chi^2_p\}, \quad \text{if } \boldsymbol{\mu}=\boldsymbol{0}.$$

Proof. Let $\Sigma=PDP'$ be the spectral decomposition of Σ. Let d_i be the i-th diagonal element of the matrix D and define $D^{-1/2}$ as the diagonal matrix with i-th diagonal element $(d_i)^{-1/2}$. If we define $\boldsymbol{z}=D^{-1/2}P'\boldsymbol{x}$ and $\boldsymbol{\xi}=D^{-1/2}P'\boldsymbol{\mu}$, then $\boldsymbol{z}$ is normally distributed with mean vector $\boldsymbol{\xi}$ and covariance matrix σ^2I and $v=\boldsymbol{z}'\boldsymbol{z}$. Therefore the desired expected values can be expressed by linear combinations of those in Lemma 1 and some arithemetic leads to the desired results.

In the following we assume that $r(\cdot,\cdot)$ is a function satisfying the following conditions:

I (a) For each fixed y, $r(\cdot, y)$ is monotone non-decreasing.
(b) For each fixed x, $r(x, \cdot)$ is monotone non-increasing.

II (a) $0\leqq r(\cdot,\cdot)\leqq 2(p-2)/(f+2)$.
(b) $0<r(\cdot,\cdot)<2(p-2)/(f+2)$ in some rectangular region.

Theorem 1. If the maximum eigenvalue d_1 of Σ is less than $2^{-1}tr\Sigma$, then in connection with the loss function L_1 the estimator

$$\Phi_1(\boldsymbol{x}, S)=\left\{1-\frac{tr\Sigma-2d_1}{(p-2)d_1}\frac{r(F,S)}{F}\right\}\boldsymbol{x}$$

has uniformly smaller risk than $\boldsymbol{x}$, where $F=\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}/S$.

Proof. We adopt the method of the proof given by BARANCHIK (1970) and prove that

$$E[\{\Phi_1(\boldsymbol{x},S)-\boldsymbol{\mu}\}'\{\Phi_1(\boldsymbol{x},S)-\boldsymbol{\mu}\}]-E\{(\boldsymbol{x}-\boldsymbol{\mu})'(\boldsymbol{x}-\boldsymbol{\mu})\}$$

is negative for all parameter values $(\boldsymbol{\mu},\sigma^2)$. As a matter of fact we prove it for the case $\boldsymbol{\mu}\neq\boldsymbol{0}$ but the argument is valid for the case $\boldsymbol{\mu}=\boldsymbol{0}$ if we interpret $\boldsymbol{\mu}'\boldsymbol{\mu}/(\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu})$ to be 0. Let $a=(tr\Sigma-2d_1)/\{(p-2)d_1\}$ and $A=\{\Phi_1(\boldsymbol{x},S)-\boldsymbol{\mu}\}'\{\Phi_1(\boldsymbol{x},S)-\boldsymbol{\mu}\}-(\boldsymbol{x}-\boldsymbol{\mu})'(\boldsymbol{x}-\boldsymbol{\mu})$. Then

$$A=-2a\boldsymbol{x}'\boldsymbol{x}r(F,S)/F+2a\boldsymbol{\mu}'\boldsymbol{x}r(F,S)/F+a^2\boldsymbol{x}'\boldsymbol{x}r^2(F,S)/F^2.$$

From I of Lemma *2* the conditional expectation of A/a given $S=s$ is given by

$$(2.1)\qquad -2sE\left[\left\{\frac{tr\Sigma}{p+2K}+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2K}{p+2K}\right\}E\left\{r\left(\frac{\sigma^2\chi^2_{p+2K}}{s},s\right)\Big|K\right\}\right]$$
$$+2s\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}E\left[2KE\left\{\frac{r(\sigma^2\chi^2_{p+2K}/s,s)}{\chi^2_{p+2K}}\Big|K\right\}\right]$$
$$+\frac{as^2}{\sigma^2}E\left[\left\{\frac{tr\Sigma}{p+2K}+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2K}{p+2K}\right\}E\left\{\frac{r^2(\sigma^2\chi^2_{p+2K}/s,s)}{\chi^2_{p+2K}}\Big|K\right\}\right],$$

where K is a Poisson random variable with mean $\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}/(2\sigma^2)$. Averaging (2.1) over $S=\sigma^2\chi^2_f$, we see that it suffices to prove that

$$(2.2)\qquad E\left[\chi^2_f r\left(\frac{\chi^2_{p+2k}}{\chi^2_f},\sigma^2\chi^2_f\right)\left\{-2\left(\frac{tr\Sigma}{p+2k}+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2k}{p+2k}\right)+2\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2k}{\chi^2_{p+2k}}\right.\right.$$
$$\left.\left.+a\left(\frac{tr\Sigma}{p+2k}+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2k}{p+2k}\right)\frac{\chi^2_f}{\chi^2_{p+2k}}r\left(\frac{\chi^2_{p+2k}}{\chi^2_f},\sigma^2\chi^2_f\right)\right\}\right]$$

is negative for each value of $k=0,1,\cdots$. Using the fact that

$$a\left(\frac{tr\Sigma}{p+2k}+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2k}{p+2k}\right)\leqq\frac{tr\Sigma-2d_1}{p-2}$$

and the condition II on $r(\cdot,\cdot)$, we see that (2.2) is less than

$$2E\left\{\chi^2_f r\left(\frac{\chi^2_{p+2k}}{\chi^2_f},\sigma^2\chi^2_f\right)g(\chi^2_{p+2k},\chi^2_f)\right\},$$

where

$$g(\chi^2_{p+2k},\chi^2_f)=-\left(\frac{tr\Sigma}{p+2k}+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2k}{p+2k}\right)+\frac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\frac{2k}{\chi^2_{p+2k}}+\frac{tr\Sigma-2d_1}{f+2}\frac{\chi^2_f}{\chi^2_{p+2k}}.$$

Fixing χ^2_f, we define the constant b so that $g(b,\chi^2_f)=0$. By condition I (a) on $r(\cdot,\cdot)$ we have the inequality

$$E[\chi^2_f r(\chi^2_{p+2k}/\chi^2_f,\sigma^2\chi^2_f)g(\chi^2_{p+2k},\chi^2_f)|\chi^2_f]$$
$$\leqq\chi^2_f r(b/\chi^2_f,\sigma^2\chi^2_f)E\{g(\chi^2_{p+2k},\chi^2_f)|\chi^2_f\}.$$

Since $\mu'\mu/\mu'\Sigma^{-1}\mu \leqq d_1$, we have

$$(2.3)\qquad E\{g(\chi^2_{p+2k}, \chi^2_f)|\chi^2_f\} \leqq -\frac{tr\Sigma}{p+2k} + d_1\frac{4k}{(p+2k)(p+2k-2)} + \frac{tr\Sigma - 2d_1}{f+2}\,\frac{\chi^2_f}{p+2k-2}.$$

Let $h(\chi^2_f)$ be the right-hand side of (2.3) and define the constant c by $h(c)=0$. From the definition of b we have $b/\chi^2_f = c_1/\chi^2_f + c_2$, where c_1 and c_2 are some constants. Using (2.3) we see that our theorem will be proved if

$$(2.4)\qquad E\{\chi^2_f r(c_1/\chi^2_f + c_2, \sigma^2\chi^2_f)h(\chi^2_f)\}$$

is shown to be less than or equal to 0. By condition I on $r(\cdot,\cdot)$ (2.4) is bounded above by

$$r(c_1/c+c_2, \sigma^2 c)E\{\chi^2_f h(\chi^2_f)\}$$

$$= r(c_1/c+c_2, \sigma^2 c)f\left\{-\frac{tr\Sigma}{p+2k} + d_1\frac{4k}{(p+2k)(p+2k-2)} + \frac{tr\Sigma - 2d_1}{p+2k-2}\right\} \leqq 0,$$

which completes the proof.

Remark. If $\Sigma = I$, then we obtain STRAWDERMAN's result.

As a corollary of Theorem 1 we obtain the following

Theorem 2. If the maximum eigenvalue d_1 of $M\Sigma$ is less than $1/2trM\Sigma$, then in connection with the loss function L_2 the estimator

$$\Phi_2(\boldsymbol{x}, S) = \left\{1 - \frac{trM\Sigma - 2d_1}{(p-2)d_1}\,\frac{r(F,S)}{F}\right\}\boldsymbol{x}$$

has uniformly smaller risk than $\boldsymbol{x}$, where $F = \boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}/S$.

This will be obvious.

Remarks. 1. It can be shown that $[1-(trM\Sigma - 2d_1)r(F,S)/\{(p-2)d_1F\}]^+\boldsymbol{x}$ dominated $\Phi_2(\boldsymbol{x}, S)$ if they are not equivalent to each other, where $a^+ = \max(0, a)$.
2. We have assumed so for that Σ and M are nonsingular. However if the matrices are singular, we can reduce the case to the nonsingular one. If M is nonsingular, then the estimator

$$\boldsymbol{x} - \frac{trM\Sigma - 2d_1}{(s-2)d_1}\,\frac{r(F,S)}{F}\Pi_\Sigma\boldsymbol{x}$$

behaves in the same way as $\Phi_2(\boldsymbol{x}, S)$, where s is the rank of Σ, Π_Σ is the orthogonal projector onto the range space of Σ and $F = \boldsymbol{x}'\Sigma^{\dagger}\boldsymbol{x}/S$, where $\Sigma^{\dagger}$ denotes the Moore-Penrose inverse of Σ.

Theorem 3. In connection with the loss function L_1 the estimator

$$\Phi_3(\boldsymbol{x}, S) = \boldsymbol{x} - d_p\frac{r(F,S)}{F}\Sigma^{-1}\boldsymbol{x}$$

has uniformly smaller risk than $\boldsymbol{x}$, where $F = \boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}/S$ and d_p is the minimum eigenvalue of Σ.

We can prove this along the same line as in the proof of Theorem 1 using II of Lemma 2.

Remarks. 1. If $\Sigma=I$, then we obtain STRAWDERMAN'S result.

2. If we set $r(\cdot,\cdot)$ equal to a constant $b(p-2)/(f+2)$, where $0<b<2$, then the difference of the risk functions of $\boldsymbol{x}$ and $\Phi_3(\boldsymbol{x}, S)$ is

$$\frac{bf}{f+2}(p-2)^2 d_p \sigma^2\left[(2p-bd_p tr\Sigma^{-1})E\left\{\frac{1}{(p+2K)(p+2K-2)}\right\}\right.$$

$$\left.+\left(2-bd_p\frac{\boldsymbol{\mu}'\Sigma^{-2}\boldsymbol{\mu}}{\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}}\right)E\left\{\frac{2K}{(p+2K)(p+2K-2)}\right\}\right]$$

$$\geqq \frac{bf}{f+2}(p-2)^2 d_p \sigma^2 E\left(\frac{1}{p+2K-2}\right),$$

where K is a Poisson random variable with mean $\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}/(2\sigma^2)$.

As a corollary of Theorem 3 we obtain the following

Theorem 4. In connection with the loss function L_2 the estimator

$$\Phi_4(\boldsymbol{x}, S)=\boldsymbol{x}-d_p\frac{r(F,S)}{F}M^{-1}\Sigma^{-1}\boldsymbol{x}$$

has uniformly smaller risk than $\boldsymbol{x}$, where $F=\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}/S$ and d_p is the minimum eigenvalue of $M\Sigma$.

Remarks. 1. It can be shown that $M^{-1/2}\{I-d_p r(F,S)/F(M^{1/2}\Sigma M^{1/2})^{-1}\}^+M^{1/2}\boldsymbol{x}$ dominates $\Phi_4(\boldsymbol{x}, S)$ if they are equivalent to each other, where $M^{1/2}$ is the symmetric positive definite matrix such that $(M^{1/2})^2=M$, $M^{-1/2}=(M^{1/2})^{-1}$ and A^+ denotes the matrix UD^+U' for the symmetric matrix A with spectral decomposition UDU', where D^+ denotes the diagonal matrix with i-th diagonal element $d_i^+=\max(0, d_i)$ for the diagonal matrix D with i-th diagonal element d_i.

2. If Σ is singular but M is nonsingular, then the estimator

$$\boldsymbol{x}-d_p\frac{r(F,S)}{S}PM^{-1}\Sigma^{\dagger}\boldsymbol{x}$$

behaves in the same way as $\Phi_4(\boldsymbol{x}, S)$, where $F=\boldsymbol{x}'\Sigma^{\dagger}\boldsymbol{x}/S$, $\Sigma^{\dagger}$ is the Moore-Penrose inverse of Σ and P is the orthogonal projector onto the range space of Σ when the inner product of $\boldsymbol{a}$ and $\boldsymbol{b}$ is defined by $\boldsymbol{a}'M\boldsymbol{b}$.

3. An improvement of the least squares estimator in a linear statistical model with normal errors.

Consider the linear statistical model

$$\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{e},$$

where $\boldsymbol{y}$ is an $n\times 1$ vector of observations, X is a known $n\times p$ matrix of rank p ($p\geqq 3$), $\boldsymbol{\beta}$ is a $p\times 1$ vector of unknown parameters and $\boldsymbol{e}$ is an $n\times 1$ vector of errors having a normal distribution with mean vector $\boldsymbol{0}$ and covariance matrix $\sigma^2 I$. The least squares estimator $\hat{\boldsymbol{\beta}}=(X'X)^{-1}X'\boldsymbol{y}$ is distributed as p-variate normal distribution with mean vector $\boldsymbol{\beta}$ and covariance matrix $\sigma^2(X'X)^{-1}$, and the residual sum of squares $S=(\boldsymbol{y}-X\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-X\hat{\boldsymbol{\beta}})$ is distributed independently of $\hat{\boldsymbol{\beta}}$ as $\sigma^2\chi^2_{n-p}$.

Suppose we want to estimate $\boldsymbol{\beta}$ by $\boldsymbol{x}$ with loss function

$$(\boldsymbol{\beta}-\boldsymbol{x})'M(\boldsymbol{\beta}-\boldsymbol{x}),$$

where M is a given positive definite matrix. In this case from Theorem 2 we see that if the maximum eigenvalue d_1 of $M(X'X)^{-1}$ is less than $2^{-1}trM(X'X)^{-1}$, then the estimator

$$\Phi_2(\hat{\boldsymbol{\beta}}, S)=\left\{1-\frac{trM(X'X)^{-1}-2d_1}{(p-2)d_1}\,\frac{r(F,S)}{F}\right\}\hat{\boldsymbol{\beta}}$$

dominates $\hat{\boldsymbol{\beta}}$, where $F=\hat{\boldsymbol{\beta}}'X'X\hat{\boldsymbol{\beta}}/S$. And from Theorem 4 we see that the estimator

$$\Phi_4(\hat{\boldsymbol{\beta}}, S)=\hat{\boldsymbol{\beta}}-d_p\frac{r(F,S)}{F}M^{-1}X'X\hat{\boldsymbol{\beta}}$$

always dominates $\hat{\boldsymbol{\beta}}$, where d_p is the minimum eigenvalue of $M(X'X)^{-1}$. These estimators are different from the one given by BHATTACHARYA. Their form is simpler and it is not necessary to make the transformation of the least squares estimator, which was needed to obtain the estimator given by BHATTACHARYA.

4. A generalization of BHATTACHARYA'S result.

Here we give a generalization of BHATTACHARYA's result. Following BHATTACHARYA we may assume the following without loss of generality: Let $\boldsymbol{x}$ be a p-dimensional normal random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\sigma^2 I$, and, let S be a statistic which is distributed independently of $\boldsymbol{x}$ as $\sigma^2\chi^2_f$. Let the loss suffered from estimating $\boldsymbol{\mu}$ by $\hat{\boldsymbol{\mu}}$ be

$$L_2'(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}})=(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})'D(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}}),$$

where D is a diagonal matrix with i-th diagonal element d_i and suppose $d_1\geqq d_2\geqq\cdots\geqq d_p>0$.

Let $r_j(\cdot,\cdot)$ be a function satisfying the conditions given in Section 2 with replacement of p by j ($j\geqq 3$).

A generalization of BHATTACHARYA'S *result.* Let $\boldsymbol{x}'_{(i)}=(x_1,\cdots,x_i)$ and $F_i=\boldsymbol{x}'_{(i)}\boldsymbol{x}_{(i)}/S$, for $i=3,\cdots,p$. Let

$$f_j(\boldsymbol{x}, S)=r_j(F_j, S)/F_j, \qquad j=3,\cdots,p,$$

and

$$h_i(\boldsymbol{x}, S)=1-\sum_{j=3}^{p} a_j f_j(\boldsymbol{x}, S)/d_i, \qquad i=1,2,$$

$$=1-\sum_{j=i}^{p} a_j f_j(\boldsymbol{x}, S)/d_i, \qquad i=3, \cdots, p,$$

where $\{a_3, \cdots, a_p\}$ is any set of constants satisfying the following inequalities:

$$0 \leqq a_p \leqq d_p, \quad 0 \leqq a_{p-1} \leqq d_{p-1}-a_p, \cdots, \quad 0 \leqq a_i \leqq d_i-(a_{i+1}+\cdots+a_p), \cdots,$$

$$0 \leqq a_3 \leqq d_3-(a_4+\cdots+a_p), \text{ and, } \quad a_j>0, \qquad \text{for some } j.$$

Let H be the diagonal matrix with i-th diagonal element $h_i(\boldsymbol{x}, S)$. Then $H\boldsymbol{x}$ has uniformly smaller risk than $\boldsymbol{x}$.

The proof can be carried out in just the same way as the one given by BHATTACHARYA.

If we take $a_p=d_p, a_{p-1}=d_{p-1}-d_p, \cdots, a_i=d_i-d_{i-1}, \cdots$, and $a_3=d_3-d_4$, then BHATTACHARYA's result follows. If we take $a_p=d_p$, and $a_{p-1}=\cdots=a_3=0$, then $H\boldsymbol{x}=\boldsymbol{x}-d_p r_p(F_p, S)F_p^{-1}D^{-1}\boldsymbol{x}$, which is the estimator given in Theorem 4. Therefore the above generalization may be considered to include as special cases not only BHATTACHARYA's result but also Theorem 4.

Acknowledgement

The author would like to thank Prof. YASUTOSHI WASHIO, Keio University, and Dr. MASAAKI SIBUYA, IBM Japan, for their many helpful suggestions and encouragements.

REFERENCES

ALAM, K. and THOMPSON, J. R. (1964): An estimate of the mean of a multivariate normal distribution, Technical Report, Indiana Univ.

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

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

BARANCHIK, A. J. (1973): Inadmissibility of maximum likelihood estimators in some multiple regression problems with three or more independent variables, Ann. Statist. **1** 312-321.

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

JAMES, W. and STEIN, C. (1961): Estimation with quadratic loss, Proc. Fourth Berkeley Symp. Math. Statist. Prob. **1** 361-379. Univ. of California Press.

SCLOVE, S. L. (1968): Improved estimators for coefficients in linear regression, J. Amer. Statist. Assoc. **63** 596-606.

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. Univ. of California Press.

STRAWDERMAN W. E. (1973): Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variances, Ann. Statist. **1** 1189-1194.