AD-A069 565 CLEMSON UNIV S C DEPT OF MATHEMATICAL SCIENCES F/G 12/1
ESTIMATION OF THE MEAN OF A NORMAL DISTRIBUTION WITH SINGULAR C--ETC(U)
NOV 78 K ALAM N00014-75-C-0451
UNCLASSIFIED N103 NL

| OF |

AD
A069565

END
DATE
FILMED
7-79
DDC

12

6 ESTIMATION OF THE MEAN OF A NORMAL DISTRIBUTION WITH SINGULAR COVARIANCE MATRIX

By

10 KHURSHEED ALAM

CLEMSON UNIVERSITY

14 N103, TR-295

REPORT N103

9 TECHNICAL REPORT #295

11 Nov[redacted]978

12 12 p.

Research Supported in part by

THE OFFICE OF NAVAL RESEARCH

Task NR 042-271 Contract N00014-75-C-0451

15

407 183

mt

ESTIMATION OF THE MEAN OF A NORMAL DISTRIBUTION WITH SINGULAR COVARIANCE MATRIX

Khursheed Alam*

Clemson University

ABSTRACT

The problem of estimating the mean of a p-variate normal distribution has been of considerable interest to the statisticians since the pioneering work of Stein who showed that the maximum likelihood estimator (MLE) is inadmissible with respect to a quadratic loss function when $p \geq 3$. Certain families of estimators have been shown in the literature to dominate the MLE. In this paper we consider the case in which the covariance matrix of the normal distribution is singular. An application of the given result arises in a problem of estimating the mean of a multinomial distribution.

Key Words: Multivariate Normal Distribution; Quadratic Loss; Minimax; Admissible; Multinomial Distribution.

AMS Classification: 62F10

*The author's work was supported by the Office of Naval Research under Contract N00014-75-C-0451.

1. Introduction. Let X be a K-component ($K \geq 3$) random vector distributed according to a multivariate normal distribution $N(\mu,\Sigma)$ with mean μ and covariance Σ. For estimating μ let the loss function be given by

$$L(\delta;\mu,\Sigma) = (\delta-\mu)'A(\delta-\mu) \tag{1.1}$$

where $\delta = \delta(X)$ denotes any estimator of μ, and A is a given symmetric positive (semi-positive) definite matrix. The maximum likelihood estimator (MLE) is the vector X and is known to be minimax. On the other hand, for $A = I$ and $\Sigma = \sigma^2 I$, where I denotes the identity matrix, Stein (1955) showed that the MLE is inadmissible with respect to the given loss function. Since the pioneering work of Stein, the given problem has been examined by various authors. They have considered certain families of estimators which are shown to dominate the MLE. They are therefore minimax. The papers of Alam (1977) and Efron and Morris (1973, 1976) may be cited for reference. A list of other papers may be seen in the bibliography given in the two papers.

In the papers cited above, the minimax estimators are given for the case in which the covariance Σ is a non-singular matrix. In this paper we consider the case in which Σ is singular. Two cases may be considered: (i) Σ is known and (ii) Σ is unknown, but an estimate S/m is given, where S is distributed independent of X according to a Wishart distribution $W(S;\Sigma,m)$. Let $\phi: [0,\infty) \to [0,1]$. If Σ is non-singular, consider a class of estimators, given by

(1.2) $$\delta(X) = \phi(X'\Sigma^{-1}X)X \quad \text{for case (i), and}$$

(1.3) $$\eta(X) = \phi(X'S^{-1}X)X \quad \text{for case (ii).}$$

The author has shown (Alam (1977), Theorems 2.1 and 2.2*) that δ and η dominate the MLE for a certain class of functions ϕ. In the following section we shall extend the given result to the case in which Σ is singular. For this case, the estimators are given by substituting Σ^- for Σ^{-1} in (1.2) and S^- for S^{-1} in (1.3), where Σ^- and S^- are generalized inverse of Σ and S, respectively, satisfying the relations $\Sigma\Sigma^-\Sigma = \Sigma$ and $SS^-S = S$.

An application of the given result arises in the problem of estimating the mean of a multinomial distribution $M(n,\underset{\sim}{p})$, where $\underset{\sim}{p} = (p_1,\ldots,p_k)'$ denotes the cell probabilites and n represents the total of cell frequencies. To see this, let $X \overset{d}{\sim} M(n,p)$ where $\overset{d}{\sim}$ means "distributed as". The vector X/n is the maximum likelihood estimator of $\underset{\sim}{p}$, its covariance is a singular matrix and it is asymptotically normally distributed for large n.

2. Main results. First let Σ be known. Consider the estimator δ, given by (1.2), with the substitution of Σ^- for Σ^{-1}, where Σ^- is a generalized inverse of Σ. Let $X = QY$, where $Y \overset{d}{\sim} N(\nu,I)$, $QQ' = \Sigma$ and $\mu = Q\nu$. Since $\Sigma\Sigma^-\Sigma = \Sigma$ or $QQ'\Sigma^-QQ' = QQ'$, we have $Q'\Sigma^-QQ' = Q'$. Therefore

(2.1) $$Q'\Sigma^-QQ'\Sigma^-Q = Q'\Sigma^-Q.$$

That is, $Q'\Sigma^{-}Q$ is an idempotent matrix. Also

$$\text{Rank } Q'\Sigma^{-}Q = \text{Rank } \Sigma = \ell, \text{ say.} \tag{2.2}$$

Since $\Sigma A\Sigma\Sigma^{-}\Sigma = \Sigma\Sigma^{-}\Sigma A\Sigma$, or

$$Q(Q'A\Sigma\Sigma^{-}Q - Q'\Sigma^{-}\Sigma AQ)Q' = 0$$

we have

$$Q'AQQ'\Sigma^{-}Q = Q'\Sigma^{-}QQ'AQ. \tag{2.3}$$

That is, the matrices $Q'AQ$ and $Q'\Sigma^{-}Q$ commute. Therefore, there exists an orthogonal matrix P, say, which diagonalizes simultaneously the two matrices. Since $Q'\Sigma^{-}Q$ is idempotent, we have

$$P'Q\Sigma^{-}QP = I_{\ell}$$

$$P'Q'AQP = D$$

where I_{ℓ} denotes a K×K diagonal matrix with ℓ elements on the diagonal, each equal to 1 and the remaining elements equal to zero, and D denotes a diagonal matrix of rank r, where

$$r = \text{Rank } Q'AQ \leq \text{Rank } Q = \text{Rank } QQ' = \ell.$$

The risk, that is, the expected loss of δ is given by

$$\begin{aligned} R_{\delta} &= E\ (\delta(X)-\mu)'A(\delta(X)-\mu) \\ &= E\ (\phi(Y'Q'\Sigma^{-}QY)Y-\nu)'Q'AQ(\phi(Y'Q'\Sigma^{-}QY)Y-\nu) \\ &= E\ (\phi(Z'I_{\ell}Z)Z-\gamma)'D(\phi(Z'I_{\ell}Z)Z-\gamma) \end{aligned} \tag{2.4}$$

where $\gamma = P'\nu$ and $Z = P'Y \overset{d}{\sim} N(\gamma, I)$. Without loss of generality, we can assume that the first ℓ diagonal elements of I_ℓ are each equal to 1 and that $d_i = 0$ for $i > \ell$ where d_i denotes the ith diagonal element of D. Let $W = (Z_1, \ldots, Z_\ell)'$, $\gamma^* = (\gamma_1, \ldots, \gamma_\ell)'$ and let D^* be obtained from D by removing the last $K-\ell$ rows and columns. Then

$$R_\delta = E\,(\phi(W'W)W-\gamma^*)'D^*(\phi(W'W)W-\gamma^*). \tag{2.5}$$

Let $\alpha_1, \ldots, \alpha_k$ denote the characteristic roots of $A\Sigma$ or equivalently of $Q'AQ$. Let $\psi(A\Sigma)$ = trace $A\Sigma$ and $\alpha_0 = \frac{\psi(A\Sigma)}{\ell} / \max(\alpha_1, \ldots, \alpha_k) \leq 1$. Observe that the risk of the MLE is equal to $\psi(A\Sigma)$. Comparing (2.5) with (2.6) of [1] we obtain the following generalization of Theorem 2.1 of [1].

<u>Theorem 2.1</u>. Let $\ell \geq 3$. Then $R_\delta - \psi(A\Sigma) \leq 0$ if (i) $x^{t+1}(1-\phi(x))$ is nondecreasing in x, and (ii) $0 \leq x(1-\phi(x)) \leq 2\alpha_0\ell-4t-4$ for some value of $t \geq 0$.

When Σ is unknown we consider the estimator η, given by (1.3) with the substitution S^- for S^{-1}, where S^- denotes a generalized inverse of S. We obtain as above the following generalization of Theorem 2.1* of [1].

<u>Theorem 2.1*</u>. Let $\ell \geq 3$. Then $R_\eta - \psi(A\Sigma) \leq 0$ if (i) $x^{t+1}(1-\phi(x))$ is nondecreasing in x, and (ii) $0 \leq (m-n+2t+3)x(1-\phi(x)) \leq 2\alpha_0\ell-4t-4$ for some value of $t \geq 0$.

3. Application. Consider the multinomial distribution $M(n,\underset{\sim}{p})$ with k cells, and let $X \overset{d}{\sim} M(n,\underset{\sim}{p})$. The covariance matrix of X is given by $\Sigma = (\sigma_{ij})$, where $\sigma_{ij} = -np_i p_j$ $(i \neq j)$ and $\sigma_{ii} = np_i(1-p_i)$. Clearly, Σ is singular. For estimating $n\underset{\sim}{p}$ let the loss be given by (1.1). Two cases are considered: (a) $A = n^{-1}I$ and (b) $A = n^{-1}C$, where C is a diagonal matrix whose ith diagonal element is equal to p_i^{-1}. In case (a) the loss is proportional to the sum of squared errors. Since $n^{-1}C$ is a generalized inverse of Σ, the loss in case (b) is proportional to Mahalanobis distance function. The loss due to the MLE in case (b) leads to Pearson's Chi-square statistic, used for the goodness of fit test. Since X is asymptotically normally distributed for large values of n, Theorem 2.1 might be used to improve upon the maximum likelihood estimator X. For the application of the theorem it should be noted that $\ell = \text{rank } \Sigma = k-1$ and that the characteristic equation of $n^{-1}\Sigma$ is given by

$$(3.1) \qquad (1 - \textstyle\sum_{i=1}^{k} \frac{p_i^2}{p_i - \lambda}) \, \Pi_{i=1}^{k} (p_i - \lambda) = 0.$$

Let λ_0 denote the largest characteristic root of $n^{-1}\Sigma$. From (3.1) we have that $p_{[k-1]} \leq \lambda_0 \leq p_{[k]}$, where $p_{[i]}$ denotes the ith smallest value amongst $p_1, \ldots, p_k$. Therefore, in the case (a)

$$\alpha_0 = (1 - \textstyle\sum_{i=1}^{k} p_i^2)/\lambda_0(k-1).$$

In the case (b), the largest characteristic root of $A\Sigma$ is equal to 1, and $\alpha_0 = 1$.

Consider the case (b). Let $\phi(x) = 1 - \frac{k-3}{x}$. Then

$$\delta(x) = \left(1 - \frac{k-3}{X'\Sigma^- X}\right)X. \tag{3.2}$$

Since the diagonal matrix with the ith diagonal element equal to $(np_i)^{-1}$ is a generalized inverse of Σ, we substitute T^{-1} for Σ^- in (3.2) where T denotes a diagonal matrix whose ith diagonal element is equal to n_i, the maximum likelihood estimate of np_i. Then the risk of

$$\begin{aligned} \delta(x) &= \left(1 - \frac{k-3}{X'TX}\right)X \\ &= \left(1 - \frac{k-3}{n}\right)X \end{aligned} \tag{3.3}$$

with respect to the loss (1.1) with $A = n^{-1}C$ is given by

$$\begin{aligned} R_\delta &= (k-1) - \frac{(k+1)(k-3)}{n} + \frac{(k-1)(k-3)^2}{n^2} \\ &< (k-1) = R_X \end{aligned} \tag{3.4}$$

for $n > (k-1)(k+3)/(k+1)$. Therefore, the MLE is inadmissible, being dominated by δ. The estimate can be further improved by letting

$$\delta(x) = \left(1 - \frac{\nu(k-3)}{X'\Sigma^- X}\right)X$$

and minimizing R_δ for ν. The minimizing value of ν is given by

$$\nu_0 = \frac{k-3}{k-1}\left(1 + \frac{k-1}{n}\right)^{-1} .$$

Remark: The choice of $\phi(x) = 1 - \frac{k-2}{x}$ in (1.2) for estimating the mean of a multivariate normal distribution with non-singular covariance matrix was originally proposed by James & Stein (1961). It should be noted, however, that the relation (3.4) establishing the inadmissibility of the MLE is not based on the asymptotic property of the multinomial distribution.

REFERENCES

[1] Alam, K. (1977). Minimax estimators of a multivariate normal mean. Scand. J. Statist. (4), 125-130.

[2] Efron, B. and Morris, C. (1973). Stein's estimation rule and its competitors - an Empirical Bayes rule. Jour. Amer. Statist. Assoc. (68), 117-130.

[3] ______________________ (1976). Families of minimax estimators of the mean of a multivariate normal distribution. Ann. Statist. (4), 11-21.

[4] James, W. and Stein, C. (1961). Estimation with quadratic loss. Proc. Fourth Berkeley Symp. Math. Statist. Prob. (1), 361-379.

[5] 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.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
N103

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
Estimation of the mean of a normal distribution with singular covariance matrix.

5. TYPE OF REPORT & PERIOD COVERED

6. PERFORMING ORG. REPORT NUMBER
Technical Report #295

7. AUTHOR(s)
Khursheed Alam

8. CONTRACT OR GRANT NUMBER(s)
N00014-75-C-0451

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Clemson University
Dept. of Mathematical Sciences
Clemson, South Carolina 29631

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
NR 042-271

11. CONTROLLING OFFICE NAME AND ADDRESS
Office of Naval Research
Code 436
Arlington, Va. 22217

12. REPORT DATE
November 10, 1978

13. NUMBER OF PAGES
8

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Multivariate Normal, Multinomial, Minimax, Admissible.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper deals with the problem of estimating the mean of a multivariate normal distribution with a singular covariance matrix. A class of estimators is given which dominate the maximum likelihood estimator, under a quadratic loss.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)