The Annals of Mathematical Statistics
1970, Vol. 41, No. 2, 642–645

NOTES

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

BY A. J. BARANCHIK

Columbia University

0. Introduction and summary. A family of estimators, each of which dominates the "usual" one, is given for the problem of simultaneously estimating means of three or more independent normal random variables which have a common unknown variance. Charles Stein [4] established the existence of such estimators (for the case of a known variance) and later, with James [3], exhibited some, both for the case of unknown common variances considered here and for other cases as well. Alam and Thompson [1] have also obtained estimators which dominate the usual one. The class of estimators given in this paper contains those of James and Stein and also those of Alam and Thompson.

1. A family of minimax estimators for the mean of a multivariate normal distribution. Given a p-dimensional ($p \geqq 3$) normal random vector X with unknown mean vector θ and covariance matrix of the form $\sigma^2 I$, and, independent of X, a statistic S which is distributed as σ^2 times a χ^2 random variable on n degrees of freedom, the problem is to estimate θ when the loss function is

$$L(\hat{\theta};\theta,\sigma^2) = (\hat{\theta}-\theta)'(\hat{\theta}-\theta)/\sigma^2. \tag{1.1}$$

Setting $F = X'X/S$, we will establish the following minimax theorem.

THEOREM. *Relative to the loss function* (1.1) *an estimator of the form*

$$\varphi(X,S) = (1-r(F)/F)X \tag{1.2}$$

is minimax if

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

PROOF. James and Stein ([3], page 366) obtained this result for $r(\cdot)$ any constant satisfying (ii). Since the "usual" estimator, X, is minimax it will suffice to show that

$$E\,||\varphi(X,S)-\theta||^2 - E\,||X-\theta||^2 \tag{1.3}$$

is not positive for all parameter values (θ, σ^2). Here we use the notational convention that, for a vector u, $||u||^2 = u'u$. Setting $g(F) = 1-r(F)/F$, (1.3) becomes

$$E[X'Xg^2(F)] - 2\theta' E[g(F)X] + ||\theta||^2 - p\sigma^2. \tag{1.4}$$

Received September 14, 1966; revised October 7, 1969.
[1] Research supported by National Science Foundation Grants NSF-GP-3694 and 9640 at Columbia University, Department of Mathematical Statistics.

Computing conditionally, given $S = s$, we obtain the conditional expectations (1.5)–(1.10):

$$(1.5)\quad E[X'Xg^2(X'X/s)] = e^{-||\theta||^2/2\sigma^2} \sum_{k=0}^{\infty} \frac{(||\theta||^2/2\sigma^2)^k}{k!} E\left[\sigma^2 \chi_{p+2k}^2 g^2\left(\frac{\sigma^2 \chi_{p+2k}^2}{s}\right)\right],$$

where χ_{p+2k}^2 is a chi-squared random variable with $p+2k$ degrees of freedom. To compute

$$(1.6)\qquad \theta' E[g(X'X/s)X],$$

we make an orthogonal transformation, mapping X into a random variable Y and θ into $(||\theta||, 0, \cdots, 0)'$. This does not affect the values of σ^2 and s. Then (1.6) is equal to

$$(1.7)\qquad ||\theta||\, E[g(Y'Y/s)Y_1],$$

where Y_1 is the first component of Y. Writing out (1.7) in terms of the distribution of Y it becomes

$$\frac{\sigma^2 ||\theta||\, e^{-||\theta||^2/2\sigma^2}}{(2\pi\sigma^2)^{\frac{1}{2}p}} \frac{d}{d||\theta||}\left[\int \cdots \int g(\textstyle\sum y_i^2/s)\, e^{-(\Sigma y_i^2 - 2||\theta|| y_i)/2\sigma^2} \displaystyle\prod_{i=1}^{p} dy_i\right],$$

or

$$(1.8)\qquad ||\theta||\, \sigma^2 e^{-||\theta||^2/2\sigma^2} \frac{d}{d||\theta||} e^{||\theta||^2/2\sigma^2} E\left[g\left(\frac{\sigma^2 \chi_{p+2K}^2}{s}\right)\right],$$

where K is a Poisson random variable with mean $||\theta||^2/2\sigma^2$. Thus (1.7) equals

$$(1.9)\qquad 2\sigma^2 \sum_{k=0}^{\infty} e^{-||\theta||^2/2\sigma^2} \left(\frac{||\theta||^2}{2\sigma^2}\right)^k kE\left[g\left(\frac{\sigma^2 \chi_{p+2k}^2}{s}\right)\Big/ k!\right].$$

Combining (1.5) and (1.9), and noting that $E[2K] = ||\theta||^2/\sigma^2$, (1.4) (conditional on $S = s$) becomes

$$(1.10)\quad \sigma^2 e^{-||\theta||^2/2\sigma^2} \sum_{k=0}^{\infty} \frac{(||\theta||^2/2\sigma^2)^k}{k!} \cdot \left\{E\left[\chi_{p+2k}^2 g^2\left(\frac{\sigma^2 \chi_{p+2k}^2}{s}\right)\right] - 4kE\left[g\left(\frac{\sigma^2 \chi_{p+2k}^2}{s}\right)\right] - p + 2k\right\}.$$

Averaging (1.10) over S and writing $S = \sigma^2 \chi_n^2$, we see that our theorem will be proved if we show that

$$(1.11)\qquad E[\chi_{p+2k}^2 g^2(\chi_{p+2k}^2/\chi_n^2) - 4kg(\chi_{p+2k}^2/\chi_n^2) - p + 2k]$$

is not positive for each value of $k = 0, 1, \cdots$. In the computations which follow we write $U = \chi_{p+2k}^2/\chi_n^2$ and will use the notation

$$(1.12)\qquad r(U) = (1 - g(U))U$$

and the fact that

$$(1.13)\qquad g(U) \geqq 1 - 2\frac{p-2}{n+2}U^{-1}.$$

It follows from (1.12) and the fact that $E\chi^2_{p+2k} = p+2k$ that (1.11) equals

$$E[-2r(U)\chi_n{}^2 + r(U)(1-g(U))\chi_n{}^2 + 4kr(U)/U],$$

which is

$$(1.14)\qquad E[r(U)\chi_n{}^2(-1-g(U)+4k/\chi^2_{p+2k})].$$

Using (1.13) we see that (1.14) is bounded above by

$$(1.15)\qquad E[r(U)Z] = E[E[r(\chi^2_{p+2k}/\chi_n{}^2)Z \mid \chi_n{}^2]],$$

where

$$Z = \chi_n{}^2\left[-2 + \left(4k + 2\frac{p-2}{n+2}\chi_n{}^2\right)\Big/\chi^2_{p+2k}\right].$$

Fixing $\chi_n{}^2$, we define the constant a by

$$(1.16)\qquad -2 + \left(4k + 2\frac{p-2}{n+2}\chi_n{}^2\right)\Big/a = 0.$$

From condition (i), we have the inequality

$$\begin{aligned} E[r(\chi^2_{p+2k}/\chi_n{}^2)Z \mid \chi_n{}^2] &\leqq r(a/\chi_n{}^2)E[Z \mid \chi_n{}^2; \chi^2_{p+2k} \leqq a]P[\chi^2_{p+2k} \leqq a] \\ &\quad + r(a/\chi_n{}^2)E[Z \mid \chi_n{}^2; \chi^2_{p+2k} > a]P[\chi^2_{p+2k} > a] \\ &= r(a/\chi_n{}^2)E[Z \mid \chi_n{}^2] \\ &= r(a/\chi_n{}^2)\chi_n{}^2\left[-2 + \left(4k + 2\frac{p-2}{n+2}\chi_n{}^2\right)\Big/(p-2+2k)\right]. \end{aligned}$$

Multiplying through by $(p-2+2k)/2(p-2)$ and using (1.15) and (1.16), we see that (1.2) will be minimax if

$$(1.17)\qquad E\left[r\left(\frac{2k}{\chi_n{}^2} + \frac{p-2}{n+2}\right)\chi_n{}^2[-1+\chi_n{}^2/(n+2)]\right]$$

is less than or equal to 0. But, by condition (i), (1.17) is bounded above by

$$\begin{aligned} & r\left(\frac{2k+p-2}{n+2}\right)E\{\chi_n{}^2[-1+\chi_n{}^2/(n+2)] \mid \chi_n{}^2 < n+2\}P[\chi_n{}^2 < n+2] \\ &\qquad + r\left(\frac{2k+p-2}{n+2}\right)E\{\chi_n{}^2[-1+\chi_n{}^2/(n+2)] \mid \chi_n{}^2 \geqq n+2\}P[\chi_n{}^2 \geqq n+2] \\ &= r\left(\frac{2k+p-2}{n+2}\right)E\{\chi_n{}^2[-1+\chi_n{}^2/(n+2)]\} = 0, \end{aligned}$$

which completes the proof.

2. Some examples. The theorem of Section 1 will now be used to obtain the estimators of James and Stein [3] and of Alam and Thompson [1] as well as some others.

EXAMPLE 1. Setting r equal to a constant c we obtain the estimators of [3], for $0 \leqq c \leqq 2(p-2)/(n+2)$. These estimators may be improved upon [see [2]] by replacing $(1-c/F)$ by $\max(0, 1-c/F)$. It is worth noting that the "improved" estimators also satisfy the conditions of the theorem (here we take $r(F)$ equal to c, if $c < F$, and equal to F, otherwise).

EXAMPLE 2. Setting $r(F) = c/(1+cF^{-1})$, we have, for $0 \leqq c \leqq (p-2)/(n+2)$,

$$\left(\frac{X'X}{X'X+cS}\right)X,$$

the estimators given in [1]. It is easy to see that this $r(F)$ satisfies the theorem, and, hence, the estimators all dominate X and are minimax.

We conclude with an example which is not as intuitively pleasing as those given above but which is, nevertheless, minimax.

EXAMPLE 3. Define $r(F)$ to be $c(0 \leqq c \leqq (p-2)/(n+2))$ if $F > c$ and 0 otherwise. This satisfies the conditions of the theorem and gives the estimator

$$\begin{aligned} \varphi(X) &= (1-c/F)X, \qquad \text{if} \quad F > c, \\ &= X, \qquad \text{if} \quad F \leqq c. \end{aligned}$$

REFERENCES

[1] ALAM, K. and THOMPSON, J. R. (1964). Estimation of the mean of a multivariate normal distribution. Indiana Univ. Technical Report.

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

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

[4] 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. California Press.