Sankhyā : The Indian Journal of Statistics
1979, Volume 40, Series B, Pts. 3 and 4, pp. 102-106.

IS THE MAXIMUM LIKELIHOOD ESTIMATE OF THE COMMON MEAN OF SEVERAL NORMAL POPULATIONS ADMISSIBLE?

By BIMAL KUMAR SINHA
Indian Statistical Institute

SUMMARY. Let $X_i \sim N(\mu, \sigma_i^2)$, $i = 1, \ldots, k$ be independent with $\sigma_1^2, \ldots, \sigma_k^2$ known and μ unknown. It is known that the maximum likelihood estimate (m.l.e.) $\hat{\mu}$, of μ, $\hat{\mu} = (\Sigma X_i/\sigma_i^2)/(\Sigma 1/\sigma_i^2)$, is admissible for μ in the class of all estimators under a fairly general type of loss which includes any positive power of the absolute error loss as a special case. In the case when σ_i^2's are unknown and are given independent $S_i^2 \sim \sigma_i^2 \; \chi^2_{m_i}/m_i$, $i = 1, \ldots, k$, also independent of the X_i's, it is shown that the m.l.e. μ^* of μ, $\mu^* = (\Sigma X_i/S_i^2)/(\Sigma 1/S_i^2)$, is inadmissible for μ even in an appropriate natural class of estimators of μ under a very general type of loss whenever it is known a priori that for some known i and j, $1 \leq i \neq j \leq k$, $\sigma_i^2 \leq \sigma_j^2$.

1. Introduction

For estimating the common unknown mean of several normal populations with known variances, the usual estimator is the weighted average of the sample means with weights proportional to the inverses of their population variances. Specifically, if $X_i \sim N(\mu, \sigma_i^2)$, $i = 1, \ldots, k$ be independent with σ_i^2's known, $-\infty < \mu < \infty$ unknown, the usual estimator of μ is

$$\hat{\mu} = \Big(\sum_1^k X_i/\sigma_i^2 \Big) \Big/ \Big(\sum_1^k 1/\sigma_i^2 \Big).$$

It is well known that $\hat{\mu}$ is the maximum likelihood estimate (m.l.e.) and also the unique best linear unbiased estimate of μ. On the other hand, since $\hat{\mu}$ is expressible as

$$\hat{\mu} = X_1 + \Big\{ \sum_2^k (X_i - X_1)/\sigma_i^2 \Big\} \Big/ \Big(\sum_1^k 1/\sigma \;\; \Big),$$

it follows that $\hat{\mu}$ is equivariant under the usual translation group of transformations. It is also known that $\hat{\mu}$ is the best equivariant estimator in the sense of having minimum mean square error. The estimator $\hat{\mu}$ enjoys yet another optimum property, namely, that of admissibility in the class of all estimators under a very general type of loss which includes any positive power of the absolute error loss as a special case (Blyth, 1951).

The problem discussed in this paper corresponds to the case when the variances σ_i^2's are unknown and along with the X_i's mentioned above, we are given k independent estimates $S_1^2, \ldots, S_k^2$ (also independent of the X_i's) of the k variances with the distributions $m_i S_i^2/\sigma_i^2 \sim \chi^2_{m_i}$, $i = 1, \ldots, k$. In this case a natural and reasonable estimator of μ is provided by

$$\mu^* = \left\{ \sum_1^k X_i/S_i^2 \right\} \Big/ \left(\sum_1^k 1/S_i^2 \right).$$

It is well known that μ^* is the m.l.e. of μ and also unbiased. Many authors have investigated properties of μ^* or variations of it. For $k = 2$, Zacks (1966) proposed a so-called 'testimator' for μ and investigated in details through numerical computations of the relevant variances conditions under which the testimator has more or less variance than μ^*. Graybill and Deal (1959) found out, for $k = 2$, conditions under which μ^* has less variance than both X_1 and X_2. A different set of necessary and sufficient conditions for arbitrary k has been given recently by Norwood and Hinkelmann (1977). However, unlike $\hat{\mu}$ as in the preceding paragraph, no exact optimum property of μ^* as regards to its admissibility or minimaxity is known so far. It seems difficult to establish such optimum properties of μ^* or their negation in the class of all estimators. It turns out however that μ^* is inadmissible for μ even in a natural subclass of estimators of μ for a very general type of loss whenever it is known a priori that for some *known* subscripts i and $j \neq i$, $\sigma_i^2 \leqslant \sigma_j^2$. The inadmissibility proof which is in the same spirit as in Stein (1964) and Sinha (1976) is given in Section 2.

2. Inadmissibility of the m.l.e. μ^*

In this section we prove that μ^* is inadmissible for μ for a very general type of loss under some prior information about the unknown variances. To be specific, assume that the loss function

$$L(\mu, d) = W(d-\mu) \qquad \ldots \quad (2.1)$$

satisfies

$$W(t) = W(-t),\ W(0) = 0,\ W(t) \text{ is strictly increasing in } |t|$$

and

$$\forall\ 0 < C < \infty;\ \int_0^\infty W(Ct)\,\phi(t)\,dt < \infty$$

where

$$\phi(t) = (\sqrt{2\pi})^{-1} \exp(-t^2/2). \qquad (2.2)$$

Assume further it is known a priori that for some *known* subscripts i and j, $1 \leqslant i \neq j \leqslant k$, $\sigma_i^2 \leqslant \sigma_j^2$. Without any loss of generality let us assume

$$\sigma_1^2 \leqslant \sigma_{i_0}^2 \text{ for some } known\ i_0 \geqslant 2. \qquad \ldots \quad (2.3)$$

The main result of the paper is the following.

Theorem 2.1 : *Under* (2.1) - (2.3), μ^* *is inadmissible for* μ *in* $\mathcal{C}$ *defined below.*

Proof: The proof is in the same spirit as in Stein (1964) and Sinha (1976). It is easy to verify that the underlying estimation problem remains invariant under the affine group G of transformations acting on the X_i's, S_i^2's, μ and σ_i^2's as

$$X_i \to aX_i + b,\ S_i^2 \to a^2 S_i^2, \qquad i = 1, \ldots, k$$

$$\mu \to a\mu + b,\ \sigma_i^2 \to a^2 \sigma_i^2, \qquad i = 1, \ldots, k \qquad \ldots \quad (2.4)$$

where $a \neq 0$, $-\infty < b < \infty$, and that any affine equivariant estimator ψ is of the form

$$\psi(X_1, \ldots, X_k; S_1^2, \ldots, S_k^2)$$

$$= X_1 + S_1 \phi\,(X_2 - X_1)/S_1, \ldots, (X_k - X_1)/S_1; S_2^2/S_1^2, \ldots, S_k^2/S_1^2). \quad \ldots \quad (2.5)$$

A subclass $\mathcal{C}$ of the class of affine equivariant estimators is obtained by choosing ϕ as

$$\phi((X_2 - X_1)/S_1, \ldots, (X_k - X_1)/S_1; S_2^2/S_1^2, \ldots, S_k^2/S_1^2)$$

$$= S_1^{-1} \sum_2^k (X_i - X_1)\, \phi_i\,(S_2^2/S_1^2, \ldots, S_k^2/S_1^2),$$

resulting in ψ as

$$\psi(X_1, \ldots, X_k; S_1^2, \ldots, S_k^2) = X_1 \left(1 - \sum_2^k \phi_i\right) + \sum_2^k X_i \phi_i \quad \ldots \quad (2.6)$$

where $\phi_i\,(S_2^2/S_1^2, \ldots, S_k^2/S_1^2)$ is abbreviated as ϕ_i, $i = 2, \ldots, k$. It is trivial to verify that $\mu^* \epsilon\, \mathcal{C}$ with

$$\phi_{i,\mu^*} = (S_1^2/S_i^2)\Big/\Big(\sum_1^k S_1^2/S_j^2\Big), \quad i \geqslant 2. \qquad \ldots \quad (2.7)$$

We now show that given any estimator $\psi\, \epsilon\, \mathcal{C}$, it is possible to improve over it in view of the condition (2.3) provided

$$\sum_{i \neq i_0} \phi_i \geqslant 0 \quad \text{and} \quad \phi_{i_0} > \frac{1}{2} \text{ hold with positive probability.} \qquad \ldots \quad (2.8)$$

Towards this end, let us first note that the risk of any $\psi\, \epsilon\, \mathcal{C}$ can be expressed as

$$R(\psi) = E\{W(\psi-\mu)\}$$

$$= E\left[E\left\{W\left(\left[\sigma_1^2\Big(1-\sum_2^k \phi_i\Big)^2+\sum_2^k \sigma_i^2\phi_i^2\right]^{\frac{1}{2}}\cdot|X|\right)\right\}\right] \qquad \ldots \quad (2.9)$$

where the inner expectation is w.r.t. $X \sim N(0, 1)$, keeping ϕ_i's fixed, and the outer expectation is w.r.t. the S_i^2's.

Define

$$\phi_{i_0}^* = \min\Big(\phi_{i_0}, \frac{1}{2}\Big)$$

and

$$\psi^* = X_1\Big(1-\phi_{i_0}^*-\sum_{i \neq i_0}\phi_i\Big)+\sum_{i \neq i_0} X_i\phi_i+X_{i_0}\phi_{i_0}^*. \qquad \ldots \quad (2.10)$$

We show that ψ^* which trivially $\epsilon\, \mathcal{C}$ is an improvement over ψ. It is enough to show in view of (2.9) that

$$\sigma_1^2\Big(1-\sum_2^k \phi_i\Big)^2+\sum_2^k \sigma_i^2\phi_i^2 \geqslant \sigma_1^2\Big(1-\sum_{i\neq i_0}\phi_i-\phi_{i_0}^*\Big)^2+\sum_{i\neq i_0}\sigma_i^2\phi_i^2\sigma+{}_{i_0}^2\phi_{i_0}^{*2}$$

$$\ldots \quad (2.11)$$

with probability 1, and that strict inequality in (2.11) holds with positive probability. But (2.11) after some simplification is equivalent to

$$\sigma_{i_0}^2\Big(\phi_{i_0}^2-\phi_{i_0}^{*2}\Big) \geqslant \sigma_1^2\Big(\phi_{i_0}-\phi_{i_0}^*\Big)\Big(2-\sum_2^k\phi_i-\sum_{i\neq i_0}\phi_i-\phi_{i_0}^*\Big)$$

$$\Longleftrightarrow \Big(\phi_{i_0}-\phi_{i_0}^*\Big)\Big[\Big(\sigma_{i_0}^2+\sigma_1^2\Big)\Big(\phi_{i_0}+\phi_{i_0}^*\Big)-2\sigma_1^2\Big(1-\sum_{i\neq i_0}\phi_i\Big)\Big] \geqslant 0.$$

$$\ldots \quad (2.12)$$

That (2.12) holds with probability one follows from the fact that over the set $\{\phi_{i_0} = \phi^*_{i_0}\}$, (2.12) holds trivially and over the complementary set $\left\{\phi_{i_0} > \phi^*_{i_0} = \frac{1}{2}\right\}$ also, (2.12) holds because

$$\phi_{i_0} + \phi^*_{i_0} > 1 \geqslant 2\sigma_1^2/(\sigma_1^2 + \sigma_{i_0}^2),$$

by (2.3). Also, the inequality in (2.12) and hence in (2.11) is strict with positive probability by (2.8).

Finally, since the condition (2.8) is satisfied with $\phi_{i_0} = \phi_{i_0, \mu^*}$, it follows that μ^* is inadmissible for μ and this completes the proof of the theorem.

Q.E.D.

Remark : The estimator ψ^* can be viewed as a sort of 'testimator' which uses ψ or a variant of it as an estimator of μ depending on whether $\phi_{i_0} \leqslant \frac{1}{2}$ or $> \frac{1}{2}$. If the apriori information (2.3) is strengthened as giving $\sigma_1^2 \leqslant \sigma_l^2$, $l = i_1, \ldots, i_r$, $1 \leqslant r \leqslant k-1$, $I = (i_1, \ldots, i_r) \subseteq (2, \ldots, k)$, then defining

$$\psi^* = X_1 (1 - \sum_{i \notin I} \phi_i - \sum_{i \in I} \phi^*_i) + \sum_{i \notin I} X_i \phi_i + \sum_{i \in I} X_i \phi^*_i$$

with $\phi^*_l = \min\left(\phi_l, \frac{1}{2}\right) l = i_1, \ldots, i_r$, it can be shown using the same arguments as given above that ψ^* improves over ψ provided at least one of the ϕ_l's $(l = i_1, \ldots, i_r)$ is greater than $\frac{1}{2}$ with positive probability.

References

Blyth, C. R. (1951) : On minimax statistical decision procedures and their admissibility. *Ann. Math. Statist.*, **22**, 22-42.

Graybill, F. A. and Deal, R. B. (1959) : Combining unbiased estimators. *Biometrics*, **15**, 543-550.

Norwood, T. E. and Hinkelmann, K. (1977) : Estimating the common mean of several normal populations. *Ann. Statist.*, **5**, 1047-1050.

Sinha, B. K. (1976) : On improved estimators of the generalized variance. *Jour. Mult. Analysis*, **6**, 617-625.

Stein, C. (1964) : Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.*, **16**, 155-160.

Zacks, S. (1966) : Unbiased estimation of the common mean. *Jour. Amer. Statist. Assoc.* **61**, 467-476.

Paper received : January, 1978.
Revised : January, 1979.