INADMISSIBILITY OF THE USUAL ESTIMATOR FOR THE MEAN OF A MULTIVARIATE NORMAL DISTRIBUTION

CHARLES STEIN
STANFORD UNIVERSITY

1. Introduction

If one observes the real random variables $X_1, \cdots, X_n$ independently normally distributed with unknown means $\xi_1, \cdots, \xi_n$ and variance 1, it is customary to estimate ξ_i by X_i. If the loss is the sum of squares of the errors, this estimator is admissible for $n \leqq 2$, but inadmissible for $n \geqq 3$. Since the usual estimator is best among those which transform correctly under translation, any admissible estimator for $n \geqq 3$ involves an arbitrary choice. While the results of this paper are not in a form suitable for immediate practical application, the possible improvement over the usual estimator seems to be large enough to be of practical importance if n is large.

Let X be a random n-vector whose expected value is the completely unknown vector ξ and whose components are independently normally distributed with variance 1. We consider the problem of estimating ξ with the loss function L given by

$$L(\xi, d) = (\xi - d)^2 = \Sigma(\xi_i - d_i)^2 \tag{1}$$

where d is the vector of estimates. In section 2 we give a short proof of the inadmissibility of the usual estimator

$$d = \xi_o(X) = X , \tag{2}$$

for $n \geqq 3$. For $n = 2$, the admissibility of ξ_o is proved in section 4. For $n = 1$ the admissibility of ξ_o is well known (see, for example, [1], [2], [3]) and also follows from the result for $n = 2$. Of course, all of the results concerning this problem apply with obvious modifications if the assumption that the components of X are independently distributed with variance 1 is replaced by the condition that the covariance matrix Σ of X is known and nonsingular and the loss function (1) is replaced by

$$L(\xi, d) = (\xi - d)'\Sigma^{-1}(\xi - d). \tag{3}$$

We shall give immediately below a heuristic argument indicating that the usual estimator ξ_o may be poor if n is large. With some additional precision, this could be made to yield a discussion of the infinite dimensional case or a proof that for sufficiently large n the usual estimator is inadmissible. We choose an arbitrary point in the sample space independent of the outcome of the experiment and call it the origin. Of course, in the way we have expressed the problem this choice has already been made, but in a correct coordinate-free presentation, it would appear as an arbitrary choice of one point in an affine space. Now

$$X^2 = (X - \xi)^2 + \xi^2 + 2\sqrt{\xi^2}Z \tag{4}$$