James-Stein estimators for the mean vector of a multivariate normal population based on independent samples from two normal populations with common covariance structure

Shahjahan Khan*
Department of Mathematics & Statistics
Sultan Qaboos University, Al-Khodh 123, Muscat, Oman
and
Zahirul Hoque†
Department of Mathematics & Computing
University of Southern Queensland, Toowoomba, Australia

Abstract

The paper considers shrinkage estimators of the mean vector of a multivariate normal population based on independent random samples from two multivariate normal populations with different mean vectors but common covariance structure. The shrinkage and the positive-rule shrinkage estimators are defined by using the preliminary test approach when uncertain prior information regarding the equality of the two population mean vectors is available. The properties and performances of the estimators are investigated. The performances of the estimators are compared based on the unbiasedness and quadratic risk criteria. The relative performances of the estimators are discussed under different conditions. The shrinkage estimator dominates the maximum likelihood estimator, and the positive-rule shrinkage estimator uniformly over performs the shrinkage estimator with respect to the quadratic risk.

Keyword and Phrases: Two-sample problem; uncertain prior information; preliminary test approach; multivariate normal, noncentral chi-square and F-distributions; incomplete beta ratio; bias and quadratic bias; quadratic risk; and admissibility.

AMS 1991 Subject Classification: Primary 62F30; and Secondary 62J05.

---

*On leave from Department of Mathematics & Computing, University of Southern Queensland, Toowoomba, Queensland, Australia

†On leave from Department of Statistics, Chittagong University, Bangladesh

1 Introduction

The seminal work of Stein (1956), and James and Stein (1961) has generated a flurry of studies by many researchers in search of improved estimators. It is well known that the popular maximum likelihood estimator (MLE) of the mean vector is based exclusively on the sample information. If the dimension of the population $p > 2$, it is well known that the James-Stein shrinkage estimator dominates the usual MLE of the mean vector under the squared error loss criterion. Procedures are available in the literature to include *uncertain prior information* which may be available in the form of a given value of the mean vector and can be expressed by a null hypothesis, in addition to the sample data for the estimation of the unknown mean vector. Bancroft (1944) introduced the idea of preliminary test estimator to remove the uncertainty in the suspected value of the mean vector through an appropriate statistical test. Saleh (1973) applied the idea for the multivariate normal case with unknown covariance matrix, while Bancroft and Han (1976) discussed the preliminary test estimator (PTE) of the mean vector of two multivariate normal populations with identity covariance matrix. Saleh and Sen (eg.1978, 1985) published a series of papers on the topic both in the parametric and non-parametric context. Ahmed and Saleh (1989) gave a comprehensive study of the two-sample problem with unknown but common covariance matrix, including the shrinkage estimator. However, their study did not cover the positive-rule shrinkage estimator for the two-sample problem. From the single population multivariate normal case, it is well known that the shrinkage estimator (SE) is unstable when the value of the test statistic is too close to zero, or even it can be negative. To address this problem, modified James-Stein estimator, namely the positive-rule shrinkage estimator (PRSE) is adopted. By the definition of this estimator, it can never be negative. More recent work in the area includes Sclove et al. (1972), Stein (1981), Maata and Casella (1990), Kubokawa (1991), Chang et al. (1993), Chang (1995), and Khan and Saleh (1997, 1998).

Unlike the original James-Stein estimator which was obtained for a single population multivariate normal model, in this paper we consider two multivariate normal populations of the same dimension, $p > 2$. It is assumed that the two populations have a common but unknown covariance matrix and that the mean vectors are not equal. Based on the two independent samples from the two populations we wish to estimate the mean vector of either population when it is *apriori* suspected that the two mean vectors are equal, but not sure. First we discuss an appropriate test procedure to test the hypothesis of equality of the two population mean vectors based on the two independent samples. Such a test removes the uncertainty in the null hypothesis. Then using the preliminary test approach we define the usual James-Stein shrinkage estimator (SE) of the mean vector, for the two-sample case, as a function of the sample data as well as the uncertain prior information. The shrinkage estimator defined in this fashion becomes a function of the test statistic. To overcome the shortcoming of the SE and achieve further improvement in terms of better statistical properties, we define the PRSE. The bias, quadratic bias and quadratic risk functions as well as the mean square error matrices of the two shrinkage estimators are obtained. The relative performance of the two estimators are investigated based on the above three criteria with a view of selecting the better one.

As in the case of one population multivariate normal problem, both the shrinkage and positive-rule shrinkage estimators of the mean vector dominate the commonly used maximum likelihood estimator (MLE) for the two population multivariate normal problem. The study reveals the fact that the PRSE uniformly over performs the SE under the quadratic risk criterion. This domination of the PRSE over the SE is for all p and for each Δ, a measure of the difference between the true population mean vectors of the two populations (see equations 3.5 and 4.1). But the dominance of the PRSE over the SE is not uniform under the quadratic bias criterion. If the prior information regarding the value of the mean vector is not too far from its actual value, that is Δ is not too large, then the PRSE dominates the SE with respect to the quadratic bias criterion. However, for very large values of Δ the opposite is true, although the difference appears to be very small (see Figure 1). In real life users are more interested in the minimization of the risk than the bias, in that sense the PRSE is obviously a better choice than the SE, which in turn dominates the MLE. A similar picture of the dominance of the PRSE over the SE emerges from the analysis of the mean square error matrices.

The findings of this paper reaffirm the fact regarding the dominance of positive rule-shrinkage estimator over the shrinkage estimator based on the one sample problem with respect to the quadratic risk criterion when $p > 2$. Moreover, the PRSE is free from the potential problems of instability and negative sign of the SE.

Specifications of the two-sample multivariate normal problem is given in the next section. Section 3 deals with definition of different estimators of the mean vector and the test statistic to test the uncertainty in the prior information. The bias of the two shrinkage estimators are provided in section 4. The derivation and analysis of the quadratic risk functions are provided in section 5. Some concluding remarks are given in the final section.

2 The Two-Sample Problem

Consider a p-dimensional ($p > 2$) multivariate normal population with unknown mean vector $\boldsymbol{\mu}_1$ and covariance matrix $\Sigma = \sigma^2 I_p$. Let $\boldsymbol{X}_{11}, \boldsymbol{X}_{12},..., \boldsymbol{X}_{1n_1}$ be a random sample of size n_1 from the above population. Similarly, let $\boldsymbol{X}_{21}, \boldsymbol{X}_{22}...\ \boldsymbol{X}_{2n_2}$ be another independent random sample of size n_2 from a second p-variate normal population with mean vector $\boldsymbol{\mu}_2$ and common covariance matrix Σ. The mean vectors as well as the covariance matrix is assumed to be unknown. However, it is suspected that the mean vectors of the two populations are equal, but not sure. This *uncertain prior information* regarding the equality of the two mean vectors can be expressed by the null hypothesis, H_o: $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, and the uncertainty in the H_o can be removed by testing it out at a pre-selected level of significance. We want to estimate the mean vector of the first population, $\boldsymbol{\mu}_1$ based on the above two random samples and an appropriate test statistic to remove the uncertainty in the H_o. Although the proposed shrinkage estimators involve preliminary test statistic, unlike the PTE the SE and the PRSE do not depend on the pre-selected level of significance.

The traditional maximum likelihood estimator (MLE) of $\boldsymbol{\mu}_1$ is based only on the sample information from the first population. Since additional information is available from the second sample, it could be used in an appropriate way to improve the quality of the estimator of $\boldsymbol{\mu}_1$. Moreover, the uncertain prior information provided by the H_o can also be incorporated for further improvement of the estimator of $\boldsymbol{\mu}_1$. Bancroft (1944) proposed the preliminary test estimator (PTE) to remove such uncertainty in the H_o. However, the PTE is an extreme choice between the unrestricted MLE and the restricted (by H_o) MLE of $\boldsymbol{\mu}_1$. Also, it depends on the choice of the level of significance. In this paper we use the information from both the samples as well as the H_o to define a Stein-type shrinkage estimator (SE) of $\boldsymbol{\mu}_1$ by adopting the preliminary test approach. Such an estimator is biased but dominates over the usual unbiased MLE when $p > 2$ (cf. Ahmed and Saleh 1989). However, as the value of the test statistic to test H_o: $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ against the alternative hypothesis H_a: $\boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, approaches to zero, the shrinkage estimator becomes unstable, and it can even change the sign of the MLE . To avoid such a potential problem we define a modified Stein-type estimator, namely, the positive-rule shrinkage estimator (PRSE) of $\boldsymbol{\mu}_1$. The two Stein-type shrinkage estimators are compared based on well known statistical criteria of unbiasedness, mean square error and risk under quadratic loss. Relative performances of the estimators are discussed under various conditions.

3 Estimators for $\boldsymbol{\mu}_1$

The unrestricted maximum likelihood estimator (MLE) of $\boldsymbol{\mu}_1$ is the sample mean based on the random sample from the first population. That is, the MLE of $\boldsymbol{\mu}_1$ is

$$\tilde{\boldsymbol{\mu}}_1 = \bar{\boldsymbol{X}}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1} \boldsymbol{X}_{1j} \tag{3.1}$$

when no other information is used. When the second sample is available and H_0 is true, the restricted MLE of $\boldsymbol{\mu}_1$ is given by

$$\hat{\boldsymbol{\mu}}_1 = \bar{\boldsymbol{X}} = \frac{1}{n_1+n_2}\sum_{i=1}^{2}\sum_{j=1}^{n_i} \boldsymbol{X}_{ij}, \tag{3.2}$$

the combined or pooled mean of the two samples. The restricted estimator performs better than the unrestricted MLE when the H_o is true. To test the null hypothesis, the likelihood ratio test can be based on the following test statistic

$$T^2 = \frac{n_1 n_2}{n_1+n_2}(\bar{\boldsymbol{X}}_2 - \bar{\boldsymbol{X}}_1)' S^{-1} (\bar{\boldsymbol{X}}_2 - \bar{\boldsymbol{X}}_1) \tag{3.3}$$

where $\bar{\boldsymbol{X}}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} \boldsymbol{X}_{ij}$ is the unrestricted MLE of $\boldsymbol{\mu}_i$ based on the i^{th} random sample for i = 1,2, and

$$S = \sum_{i=1}^{2}\sum_{j=1}^{n_i} (\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}}_i)'(\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}}_i),$$

the sum of the two sample sum of squares. The MLE of σ^2 is found to be $\tilde{\sigma}^2 = \frac{1}{(n_1+n_2)p}S$. Thus the unrestricted MLE of Σ, $\tilde{\Sigma} = \frac{1}{(n_1+n_2)p}SI_p$ is a biased estimator, while $\tilde{\Sigma}^* = \frac{1}{(n_1+n_2-2)p}SI_p$ is an unbiased estimator of Σ. Khan (1997) has shown this result in the context of the multivariate Student-t population. The above T^2 statistic follows a modified Hotelling's T^2 distribution (cf. Anderson, 1984, p.167) and it is well known that

$$T^2 \sim \frac{p}{m} F_{p,m}(\Delta_1), \tag{3.4}$$

that is, under the alternative hypothesis the above T^2 statistic follows a scaled non-central F-distribution with p and $m = (n_1 + n_2 - p - 1)$ degrees of freedom and noncentrality parameter

$$\Delta_1 = \boldsymbol{\delta}'\Sigma^{-1}\boldsymbol{\delta} \tag{3.5}$$

in which $\boldsymbol{\delta} = \boldsymbol{\mu_2} - \boldsymbol{\mu_1}$, the difference between the two population mean vectors. Under the null hypothesis $\boldsymbol{\delta} = \mathbf{0}$, and hence T^2 has a scaled central F-distribution. Khan (1998) used the above test statistic to define preliminary test estimator for the two-sample problem with diagonal covariance matrix. In this paper two James-Stein shrinkage estimators are pursued for the general covariance matrix problem. The shrinkage estimator (SE) of $\boldsymbol{\mu}_1$ is defined by using the MLE of $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, and the T^2 statistic as follows:

$$\hat{\boldsymbol{\mu}}_1^s = \tilde{\boldsymbol{\mu}}_1 + McT^{-2}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1) \tag{3.6}$$

where $M = \frac{n_2}{n_1+n_2}$ and $0 < c < \frac{2(p-2)}{(N-p+3)}$ is the shrinkage constant with $N = n_1+n_2-2$.

Ahmed and Saleh (1989) defined such a James-Stein estimator of $\boldsymbol{\mu}_1$ and discussed its dominance over the usual MLE of $\boldsymbol{\mu}_1$ when $p > 2$, (also see Anderson 1984, p.171) under the squared error loss function. Although the SE provides smoother transition, compared to the PTE, between $\tilde{\boldsymbol{\mu}}_1$ and $\hat{\boldsymbol{\mu}}_1$, and does not depend on the level of significance, it has its own shortcomings. As the value of T^2 approaches to zero, the SE becomes unstable. It can even be negative. This is a very serious setback for the SE and an obvious matter of concern for its users. To overcome this problem we define the following positive-rule shrinkage estimator (PRSE) by modifying the shrinkage estimator. Thus for the two-sample problem the PRSE of $\boldsymbol{\mu}_1$ is defined as

$$\hat{\boldsymbol{\mu}}_1^{s+} = \hat{\boldsymbol{\mu}}_1 + (1 - cT^{-2})I(T^2 > c)(\tilde{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_1) \tag{3.7}$$

where $I(T^2 > c)$ is an indicator function that assumes only two values, 0, or 1, depending on the value of the argument, and $\hat{\boldsymbol{\mu}}_1$ is as defined in (3.2). In the forthcoming sections of the paper, we investigate the performances of the above two James-Stein type estimators as well as the MLE based on various criteria of good estimators.

The following representations of the SE and the PRSE are useful in computing the bias, mean squared error and risk under quadratic loss:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_1^s - \boldsymbol{\mu}_1 &= (\tilde{\boldsymbol{\mu}}_1 - \boldsymbol{\mu}_1) + cMT^{-2}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1) & (3.8)\\ \hat{\boldsymbol{\mu}}_1^{s+} - \boldsymbol{\mu}_1 &= (\hat{\boldsymbol{\mu}}_1^s - \boldsymbol{\mu}_1) + M(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 \le c) & \\ &\quad -cMT^{-2}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 \le c). & (3.9) \end{aligned}$$

By definition both the preliminary test estimator and the Stein-type shrinkage estimator of $\boldsymbol{\mu}_1$ are convex combination of the unrestricted MLE, $\tilde{\boldsymbol{\mu}}_1$ and the restricted MLE, $\hat{\boldsymbol{\mu}}_1$. When the observed value of the test statistic is insignificant then both estimators become the same as the unrestricted MLE of $\boldsymbol{\mu}_1$. However, for significant (large) value of the test statistic, they are different. Moreover, when the value of the test statistic is less than the shrinkage constant c, the shrinkage factor $(1 - cT^{-2})$ becomes negative. Likewise, the positive-rule shrinkage estimator of $\boldsymbol{\mu}_1$ is a convex combination of $\tilde{\boldsymbol{\mu}}_1$ and $\hat{\boldsymbol{\mu}}_1$, but it can never change the sign of $\tilde{\boldsymbol{\mu}}_1$, the unrestricted MLE of $\boldsymbol{\mu}_1$, and hence the nomenclature.

4 The Bias of the SE and PRSE

It is well known that the MLE of $\boldsymbol{\mu}_1$, $\tilde{\boldsymbol{\mu}}_1$ is an unbiased estimator of $\boldsymbol{\mu}_1$. However, both the SE, $\hat{\boldsymbol{\mu}}_1^s$ and the PRSE, $\hat{\boldsymbol{\mu}}_1^{s+}$ are biased estimators of $\boldsymbol{\mu}_1$ under the alternative hypothesis. The amount of bias of the two shrinkage estimators are given by the following theorems.

4.1 Theorem

For the two-sample multivariate normal problem described in section 2, the bias of the SE of $\boldsymbol{\mu}_1$ is given by

$$B(\hat{\boldsymbol{\mu}}_1^s; \boldsymbol{\mu}_1) = cmM\boldsymbol{\delta}E[\chi_{p+2}^{-2}(\Delta)] \tag{4.1}$$

where $c = \frac{p-2}{m+2}$, the optimal shrinkage constant that maximizes the quadratic risk of the SE, $\Delta = n_1 M\boldsymbol{\delta}'\Sigma^{-1}\boldsymbol{\delta}$ and $\chi_{p+2}^{-2}(\Delta)$ is an inverted noncentral chi-square variable with $(p+2)$ d.f. and noncentrality parameter Δ. Note

$$E[\chi_{p+2}^{-2}(\Delta)] = \sum_{r=0}^{\infty} \frac{1}{p+2r}\, p(r) \text{ where } R \sim \text{ Poisson } \left(\frac{\Delta}{2}\right). \tag{4.2}$$

Proof: From the presentation of $(\hat{\boldsymbol{\mu}}_1^s - \boldsymbol{\mu}_1)$ in (3.8), the bias of the SE is given by

$$B(\hat{\boldsymbol{\mu}}_1^s; \boldsymbol{\mu}_1) = cME[T^{-2}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)]. \tag{4.3}$$

To compute the above expectation, and other forthcoming expressions, consider the following transformation

$$Y = \sqrt{n_1 M}\Sigma^{-\frac{1}{2}}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1). \tag{4.4}$$

Since both $\tilde{\boldsymbol{\mu}}_1$ and $\tilde{\boldsymbol{\mu}}_2$ are independently distributed as multivariate normal vectors with mean vectors $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, and covariance matrices $\frac{\Sigma}{n_1}$ and $\frac{\Sigma}{n_2}$ respectively, the statistic

$$(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1) \sim N_p\left(\boldsymbol{\delta},\ \frac{\Sigma}{n_1 M}\right) \text{ and } Y \sim N_p\left(\sqrt{n_1 M}\Sigma^{-\frac{1}{2}}\boldsymbol{\delta},\ I_p\right). \tag{4.5}$$

Then $Y'Y = n_1 M(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1) \sim \chi_p^2(\Delta)$, a noncentral chi-square variable with noncentrality parameter Δ and p degrees of freedom. Moreover, the test statistic T^2 can be expressed as

$$T^2 = \frac{Y'Y}{\chi_m^2} = \frac{p}{m}F_{p,m}(\Delta). \tag{4.6}$$

Therefore, we get

$$E[T^{-2}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)] = m\boldsymbol{\delta}E[\chi_{p+2}^{-2}(\Delta)] \tag{4.7}$$

by applying Theorem 1 from Appendix B2 of Judge and Bock (1978, p.321-324). Hence the proof.

4.2 Theorem

For the two-sample multivariate normal problem described in Section 2, the bias of the PRSE of $\boldsymbol{\mu}_1$ is given by

$$
\begin{aligned}
B(\hat{\boldsymbol{\mu}}_1^{s+};\boldsymbol{\mu}_1) = B(\hat{\boldsymbol{\mu}}_1^{s};\boldsymbol{\mu}_1) + M\boldsymbol{\delta}G_{p+2,m}(h_2;\Delta) \\
-cmM\boldsymbol{\delta}E\left[\chi_{p+2}^{-2}(\Delta)I\left(\frac{\chi_{p+2}^2(\Delta)}{\chi_{m+2}^2}\le c\right)\right]
\end{aligned}
\tag{4.8}
$$

where $h_2 = \frac{mc}{p+2}$, $B(\hat{\boldsymbol{\mu}}_1^{s};\boldsymbol{\mu}_1)$ is the expression of bias for the SE specified in equation (4.1); $G_{p+2,m}(h_2;\Delta)$ is the distribution function of the noncentral $F_{p+2,m}(\Delta)$ variable evaluated at h_2; $I\left(\frac{\chi_{p+2}^2(\Delta)}{\chi_{m+2}^2}\le c\right)$ is an indicator function; and

$$
G_{p+i,m+i}(h_i;\Delta) = \sum_{r=0}^{\infty}\frac{e^{-\frac{\Delta}{2}}(\frac{\Delta}{2})^r}{r!}IB_{h_i^*}\left(\frac{p+i+2r}{2},\frac{m+i}{2}\right)
\tag{4.9}
$$

in which $IB_{h_i^*}(a,b)$ is the complement of the incomplete beta function ratio with arguments a and b and evaluated at $h_i^* = \frac{(p+i)}{(p+i)+(p+i+2r)c}$ for $i=-2,0,2,4$.

Proof: From the representation of $(\boldsymbol{\mu}_1^{s+}-\boldsymbol{\mu}_1)$ in (3.9) the bias of the PRSE can be written as

$$
\begin{aligned}
B(\hat{\boldsymbol{\mu}}_1^{s+};\boldsymbol{\mu}_1) &= B(\hat{\boldsymbol{\mu}}_1^{s};\boldsymbol{\mu}_1) + ME[(\tilde{\boldsymbol{\mu}}_2-\tilde{\boldsymbol{\mu}}_1)I(T^2\le c)] \\
&\quad -cME[T^{-2}(\tilde{\boldsymbol{\mu}}_2-\tilde{\boldsymbol{\mu}}_1)I(T^2\le c)].
\end{aligned}
\tag{4.10}
$$

Now applying the transformation in (4.4) we can write

$$
\begin{aligned}
E[(\tilde{\boldsymbol{\mu}}_2-\tilde{\boldsymbol{\mu}}_1)I(T^2\le c)] &= E\left[\frac{\Sigma^{\frac{1}{2}}}{\sqrt{n_1 M}}Y\ I\left(\frac{Y'Y}{\chi_m^2}\le c\right)\right] \\
&= \boldsymbol{\delta}E\left[I\left(\frac{\chi_{p+2}^2}{\chi_m^2}(\Delta)\le c\right)\right] \\
&= \boldsymbol{\delta}G_{p+2,m}(h_2,\Delta)
\end{aligned}
\tag{4.11}
$$

where the result from Appendix B2 of Judge and Bock (1978) has been applied. Note the shrinkage constant is optimal in the sense of minimizing the quadratic risk of the estimator (cf. Ahmad and Saleh, 1989 for instance).

Similarly, we get

$$
E[T^{-2}(\tilde{\boldsymbol{\mu}}_2-\tilde{\boldsymbol{\mu}}_1)I(T^2\le c)] = m\boldsymbol{\delta}E\left[\chi_{p+2}^{-2}(\Delta)I\left(\frac{\chi_{p+2}^2}{\chi_{m+2}^2}(\Delta)\le c\right)\right].
\tag{4.12}
$$

Combining the results in (4.11) and (4.12) and plugging into the expression in (4.10) the theorem is proved, after simplification and adjustments.

The graph in Figure 1, is obtained by setting $\Sigma = I_p$, $n_1 = 15$ and $n_2 = 15$. But such a graph can be obtained for any set of reasonable values of Σ, n_1 and n_2. Of course the shape of the quadratic bias function will change as the value of either Σ, or $(n_1,\ n_2)$ or p changes.

Analysis of Bias: Clearly, both the SE and PRSE are biased estimators of $\boldsymbol{\mu}_1$ and an increasing function of $\boldsymbol{\delta}$. However, under the null hypothesis $\boldsymbol{\delta} = \mathbf{0}$ and hence both the SE and the PRSE are unbiased estimators when the H_0 holds good. But when the H_0 is not true, the bias of the two estimators can be compared by using the difference of the amount of biases. Thus from the bias expressions of PRSE and SE we get the difference as follows:

$$D(\hat{\boldsymbol{\mu}}_1^{s+}; \hat{\boldsymbol{\mu}}_1^{s}) = \boldsymbol{\delta} M\{G_{p+2,m}(h_2; \Delta) - cmE\left[\chi_{p+2}^{-2}(\Delta) I\left(\frac{\chi_{p+2}^2}{\chi_{m+2}^2}(\Delta) \le c\right)\right]. \tag{4.13}$$

The above difference can be represented as

$$\begin{aligned}\boldsymbol{\delta} M \sum_{r=0}^{\infty} \Bigg\{ IB_{h_2^*}\left(\frac{p+2+2r}{2}; \frac{m}{2}\right) - \frac{m(p-2))}{(m+2)(p)} \\ \times IB_{h_2^*}\left(\frac{p+2r}{2}; \frac{m+2}{2}\right)\Bigg\} \times p(r)\end{aligned} \tag{4.14}$$

where $p(r)$ is the $p.m.f.$ of a Poisson variable with parameter $\frac{\Delta}{2}$.

Now, from the previous specifications, we have that $\frac{m(p-2))}{(m+2)(p)}$ is less than 1 for each value of r. Also, from the property of the incomplete beta function ratio we have, $IB_{h_2^*}(\frac{p+2+2r}{2}; \frac{m}{2}) \ge IB_{h_2^*}(\frac{p+2r}{2}; \frac{m+2}{2})$. Hence the value of the expression in (4.14) is always positive for each value of r. Therefore, the PRSE will always have at least as large amount of bias as the SE when the null hypothesis is not true. However, as shown in the Figure 1, the quadratic bias functions of the SE and PRSE are almost identical for a wide range of values of Δ starting from 0, and differ only slightly for very large values of Δ.

The amount of bias of the SE is a function of $E[\chi_{p+2}^{-2}(\Delta)]$ which in turn depends on the value of $\boldsymbol{\delta}$, the difference between the two population mean vectors. The function $E[\chi_{p+2}^{-2}(\Delta)]$ is a decreasing function of Δ. However, when $\boldsymbol{\delta} = \mathbf{0}$, the value of $\Delta = 0$, and the SE of $\boldsymbol{\mu}_1$ becomes an unbiased estimator. Since the expressions of bias of the estimators are vectors, straightforward comparison of vectors is not meaningful, if not impossible. However, to have a clear picture of the behaviour of the bias function of the estimators, we compute the quadratic bias functions of the SE and PRSE. The plot of the quadratic bias function against the values of Δ in Figure 1 shows that the quadratic bias of the SE as well as the PRSE is the lowest at $\Delta = 0$. Then it begins to grow up as Δ grows larger and larger, and then start declining after reaching a

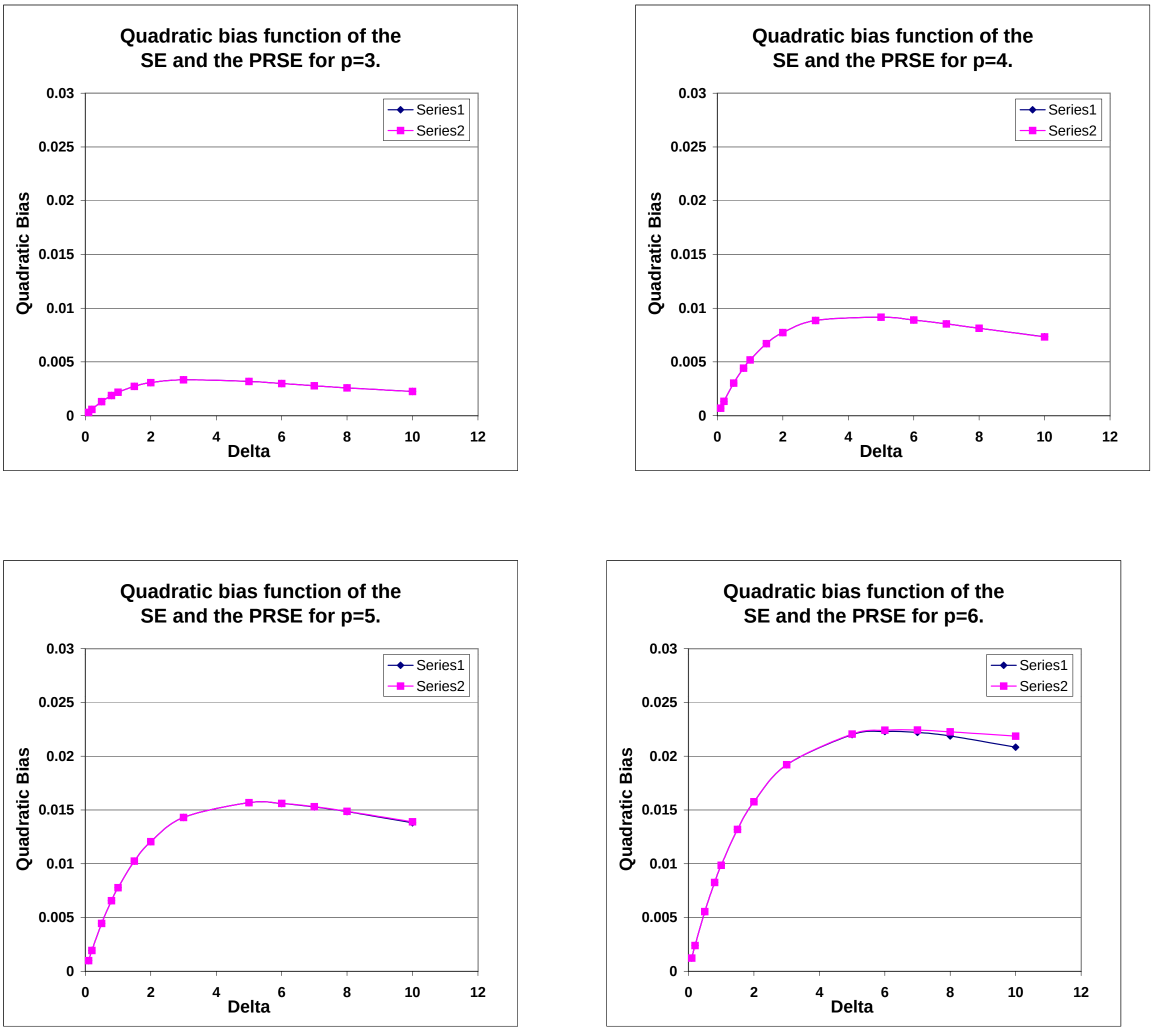

Figure 1: Quadratic bias of the SE and PRSE for p=3, 4, 5 and 6.

maximum for some medium value of Δ. From the graphs in Figure 1 it is evident that the quadratic bias function of both the shrinkage estimators have a very similar growth pattern for all values of p and Δ.

The bias of the PRSE is a function of $E[\chi^{-2}_{p+2}(\Delta) I(\frac{\chi^2_{p+2}}{\chi^2_{m+2}}(\Delta) \le c)]$ as well as $G_{p+2,m}(h_2; \Delta)$ in addition to that of the SE. Both of the above functions are again dependent on the value of Δ. Hence, like the SE, the PRSE is an unbiased estimator when $\boldsymbol{\delta} = \mathbf{0}$. Moreover, the quadratic bias of the PRSE increases as Δ moves away from 0, much like that of the SE.

The values of the quadratic bias for both the shrinkage estimators depend on the dimension of the population (p). As the value of p increases the quadratic bias of the estimators increases for every value of Δ, except for 0. For a wide range of initial values of Δ, starting from $\Delta = 0$, the quadratic bias function for both the estimators are almost identical. But as the value of p increases the value of the quadratic bias functions of both the SE and PRSE also increase. Thus, based on the criterion of quadratic bias, the SE and PRSE perform about the same for every p, and for a wide range of values of Δ. The difference between the quadratic bias functions of the estimators reduces as the size of either one or both samples increases.

5 The Risk Functions

In this section, we compute the expressions of the quadratic risks for the estimators. The quadratic risk of an estimator $\hat{\boldsymbol{\theta}}$, based on a random sample of size n, in estimating the mean vector $\boldsymbol{\theta}$ of a multivariate normal population with covariance matrix Ω is given by the expected value of the quadratic loss function as follows:

$$R(\hat{\boldsymbol{\theta}};\boldsymbol{\theta}) = E[L(\hat{\boldsymbol{\theta}},\boldsymbol{\theta})] \tag{5.1}$$

where

$$L(\hat{\boldsymbol{\theta}},\boldsymbol{\theta}) = n(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})'\Omega^{-1}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}). \tag{5.2}$$

The quadratic risk of the MLE of $\boldsymbol{\mu}_1$ is known to be

$$R(\tilde{\boldsymbol{\mu}}_1;\boldsymbol{\mu}_1) = p, \tag{5.3}$$

a constant, since $n_1(\tilde{\boldsymbol{\mu}}_1-\boldsymbol{\mu}_1)'\sum^{-1}(\tilde{\boldsymbol{\mu}}_1-\boldsymbol{\mu}_1)$ follows a central chi-square distribution with mean p.

5.1 Theorem

For the two-sample multivariate normal problem described in Section 2, the quadratic risk of the SE of $\boldsymbol{\mu}_1$ is given by

$$R(\hat{\boldsymbol{\mu}}_1^s;\boldsymbol{\mu}_1) = p - cmM\left\{2(p-2)-c(m+2)\right\}E[\chi_p^{-2}(\Delta)] \tag{5.4}$$

where $E[\chi_p^{-2}(\Delta)] = \sum_{r=0}^{\infty}\frac{p(r)}{(p+2r-2)}$ *with* $R \sim$ *Poisson* $(\frac{\Delta}{2})$.

Proof: The quadratic risk of the SE of $\boldsymbol{\mu}_1$ can be written as

$$\begin{aligned} R(\hat{\boldsymbol{\mu}}_1^s;\boldsymbol{\mu}_1) = p \quad &+ \quad n_1c^2M^2E[T^{-4}(\tilde{\boldsymbol{\mu}}_1-\boldsymbol{\mu}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2-\tilde{\boldsymbol{\mu}}_1)] \\ &+ \quad 2n_1cME[T^{-2}(\tilde{\boldsymbol{\mu}}_1-\boldsymbol{\mu}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2-\tilde{\boldsymbol{\mu}}_1)] \end{aligned} \tag{5.5}$$

where p is the quadratic risk of the unrestricted MLE of $\boldsymbol{\mu}_1$. Now using the transformation in equation (4.4) the second term in (5.5) can be expressed as

$$c^2Mm(m+2)E[\chi_p^{-2}(\Delta)]. \tag{5.6}$$

Noting that conditional on $(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)$, the expected value of $(\tilde{\boldsymbol{\mu}}_1 - \boldsymbol{\mu}_1)$ is

$$-M\{(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1) - (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)\}, \tag{5.7}$$

the last term on the right hand side of (5.5) becomes

$$-2cM^2 n_1 \left\{ E[T^{-2}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)] - (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)' E[T^{-2}\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)] \right\}. \tag{5.8}$$

Then, since

$$\begin{aligned} E[T^{-2}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)] &= \frac{m}{n_1 M} \text{ and} \\ (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)' E[T^{-2}\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)] &= \frac{m}{n_1 M} \Delta E[\chi_{p+2}^{-2}(\Delta)], \end{aligned}$$

the expression in (5.8) becomes

$$-2cMm + 2cMm\Delta E[\chi_{p+2}^{-2}(\Delta)]. \tag{5.9}$$

Finally, putting the results of (5.6) and (5.9) into (5.5) and applying the relation

$$\Delta E[\chi_{p+2}^{-2}(\Delta)] = 1 - (p-2)E[\chi_p^{-2}(\Delta)], \tag{5.10}$$

the right hand side of (5.5), after simplification, becomes

$$p - cMm\{2(p-2) - c(m+2)\}E[\chi_p^{-2}(\Delta)]. \tag{5.11}$$

Hence the theorem. The optimal value of the shrinkage constant is obtained by maximizing the above quadratic risk function with respect to c.

5.2 Theorem

For the two-sample multivariate normal problem described in Section 2, the quadratic risk of the PRSE of $\boldsymbol{\mu}_1$ is given by

$$\begin{aligned} R(\hat{\boldsymbol{\mu}}_1^{s+}; \boldsymbol{\mu}_1) \;\; = \;\; & R(\hat{\boldsymbol{\mu}}_1^{s}; \boldsymbol{\mu}_1) + M\Big\{ 2cmG_{p,\,m}(h_0;\Delta) - pG_{p+2,m}(h_2;\Delta) \\ & -c^2 m(m+2) E\left[\chi_p^{-2}(\Delta) I\left(\frac{\chi_p^2(\Delta)}{\chi_{m+4}^2} \le c \right)\right] \Big\} \\ & + M\Delta \Big\{ 2G_{p+2,m}(q_2;\Delta) - G_{p+4,m}(q_4;\Delta) \\ & -2cmE\left[\chi_{p+2}^{-2}(\Delta) I\left(\frac{\chi_{p+2}^2(\Delta)}{\chi_{m+2}^2} \le c \right)\right] \Big\} \end{aligned} \tag{5.12}$$

where $h_i = \frac{mc}{p+i}$; $R(\hat{\boldsymbol{\mu}}_1^{s}; \boldsymbol{\mu}_1)$ is the quadratic risk of the SE of $\boldsymbol{\mu}_1$ as given in (5.4); and

$$G_{p+i,m+i}(h_i;\Delta) = \sum_{r=0}^{\infty} \frac{e^{-\frac{\Delta}{2}}(\frac{\Delta}{2})^r}{r!} IB_{h_i^*}\left(\frac{p+i+2r}{2}, \frac{m+i}{2} \right) \tag{5.13}$$

in which $IB_{h_i^*}(a,b)$ is the complement of an incomplete beta function ratio with arguments a and b, and evaluated at h_i with $h_i^* = \frac{(p+i)}{(p+h)+(p+i+2r)c}$ for $i = -2, 0, 2, 4$.

Proof: By definition, the quadratic risk of the PRSE of $\boldsymbol{\mu}_1$ is given by

$$R(\hat{\boldsymbol{\mu}}_1^{s+}; \boldsymbol{\mu}_1) = E[n_1(\hat{\boldsymbol{\mu}}_1^{s+} - \boldsymbol{\mu}_1)'\Sigma^{-1}(\hat{\boldsymbol{\mu}}_1^{s+} - \boldsymbol{\mu}_1)]. \tag{5.14}$$

Using the presentation of $(\hat{\boldsymbol{\mu}}_1^{s+}; \boldsymbol{\mu}_1)$ as given in (3.9), expanding the resulting quadratic forms and simplifying the terms, the above risk function can be written as

$$\begin{aligned} R(\hat{\boldsymbol{\mu}}_1^{s+}; \boldsymbol{\mu}_1) &= E[n_1(\tilde{\boldsymbol{\mu}}_1 - \boldsymbol{\mu}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_1 - \boldsymbol{\mu}_1)] \qquad (5.15)\\ &+ c^2M^2n_1E[T^{-4}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)]\\ &+ M^2n_1E[(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 \le c)]\\ &+ c^2M^2n_1E[T^{-4}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 \le c)]\\ &+ 2cMn_1E[(\tilde{\boldsymbol{\mu}}_1 - \boldsymbol{\mu}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)T^{-2}]\\ &+ 2Mn_1E[(\tilde{\boldsymbol{\mu}}_1 - \boldsymbol{\mu}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 \le c)]\\ &- 2cMn_1E[T^{-2}(\tilde{\boldsymbol{\mu}}_1 - \boldsymbol{\mu}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 \le c)]\\ &+ 2cM^2n_1E[T^{-2}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 \le c)]\\ &- 2c^2M^2n_1E[T^{-4}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 \le c)]\\ &- 2cM^2n_1E[T^{-2}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 \le c)] \end{aligned}$$

where $I^2(A) = I(A)$ for the indicator function, $I(A)$ of the set A, has been used.

Clearly, the first term in (5.15) is p, the quadratic risk of the unrestricted MLE of $\boldsymbol{\mu}_1$. Also, the terms 8 and 10 cancel each other. For the evaluation of the remaining terms in (5.15), we use the transformation in (4.4), and apply the results from the Appendix B2 of Judge and Bock (1978), as before. Let t_i denote the i^{th} term in (5.15) for $i = 1, 2, ..., 10$. Then we have $t_1 = p$; and

$$\begin{aligned} t_2 &= c^2M^2n_1, E\left[\frac{\chi_4^2}{Y'Y}\frac{1}{n_1M}\right] = c^2Mm(m+2)E[\chi_p^{-2}(\Delta)];\\ t_3 &= M^2n_1E\left[\frac{Y'Y}{n_1M}\,I\left(\frac{Y'Y}{\chi_m^2} \le c\right)\right]\\ &= MpG_{p+2,m}(h_2;\Delta) + M\Delta G_{p+4,m}(h_4;\Delta);\\ t_4 &= c^2M^2n_1, E\left[\frac{\chi_m^4}{Y'Y}\frac{1}{n_1M}\,I\left(\frac{Y'Y}{\chi_m^2} \le c\right)\right]\\ &= c^2Mm(m+2)E\left[\chi_p^{-2}(\Delta)I\left(\frac{\chi_p^2(\Delta)}{\chi_{m+4}^2} \le c\right)\right];\\ t_5 &= -2cM^2n_1E\left[\frac{\chi_m^2}{n_1M}\right] + 2cM\Delta E\left[\frac{\chi_m^2}{Y'Y}\right]\\ &= -2cMm + 2cMm\Delta E[\chi_{p+2}^{-2}(\Delta)]; \end{aligned}$$

where the result on the conditional expectation of $(\tilde{\boldsymbol{\mu}}_1 - \boldsymbol{\mu}_1)$, given $(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)$, has been applied from (5.7);

$$\begin{aligned} t_6 &= -2M^2 n_1 E\left[\frac{Y'Y}{n_1 M}\, I\left(\frac{Y'Y}{\chi^2_m} \le c\right)\right] + 2Mn_1 \boldsymbol{\delta}'\Sigma^{-1} E\left[\frac{\Sigma^{\frac{1}{2}}}{\sqrt{Mn_1}} Y I\left(\frac{Y'Y}{\chi^2_m} \le c\right)\right] \\ &= -2MpG_{p+2,m}(h_2;\Delta) - 2M\Delta G_{p+4,m}(h_4;\Delta) + 2M\Delta G_{p+2,m}(h_2;\Delta), \end{aligned}$$

in which the previous result on the conditional expectation has been used and the expression has been simplified;

$$\begin{aligned} t_7 &= 2cM^2 n_1 E\left[\frac{\chi^2_m}{Mn_1}\, I\left(\frac{Y'Y}{\chi^2_m} \le c\right)\right] - 2cMn_1 E\left[\frac{\chi^2_m}{Y'Y}\boldsymbol{\delta}'\frac{\Sigma^{\frac{1}{2}}}{\sqrt{Mn_1}} Y\, I\left(\frac{Y'Y}{\chi^2_m} \le c\right)\right] \\ &= 2cMmG_{p,m}(h_0;\Delta) - 2cMm\Delta E\left[\chi^{-2}_{p+2}(\Delta) I\left(\frac{\chi^2_{p+2}(\Delta)}{\chi^2_{m+2}} \le c\right)\right] \end{aligned}$$

when simplified after using the earlier conditional expectation and results from Appendix B2 of Judge and Bock (1978); and

$$\begin{aligned} t_9 &= -2c^2 M^2 n_1 E\left[\frac{\chi^4_m}{Y'Y}\frac{1}{Mn_1} I\left(\frac{Y'Y}{\chi^2_m} \le c\right)\right] \\ &= -2c^2 Mm(m+2)E\left[\chi_p^{-2}(\Delta)\; I\left(\frac{\chi^2_p(\Delta)}{\chi^2_{m+4}} \le c\right)\right]. \end{aligned}$$

Now collecting $t_1 - t_{10}$ in (5.15) and simplifying, the quadratic risk of the PRSE of $\boldsymbol{\mu}_1$ becomes, on regrouping of the terms and simplification,

$$\begin{aligned} R(\hat{\boldsymbol{\mu}}_1^{s+};\boldsymbol{\mu}_1) &= \left\{p + c^2 Mm(m+2)E[\chi_p^{-2}(\Delta)] - 2cMm + 2cm\Delta E[\chi^{-2}_{p+2}(\Delta)]\right\} \\ &+ M\Delta\Big\{2G_{p+2,m}(h_2;\Delta) - G_{p+4,m}(h_4;\Delta) \\ &\qquad -2cmE\left[\chi^{-2}_{p+2}(\Delta) I\left(\frac{\chi_{p+2}(\Delta)}{\chi^2_{m+2}} \le c\right)\right]\Big\} \\ &- M\Big\{pG_{p+2,m}(q_2;\Delta) - 2cmG_{p,m}(q_0;\Delta) \\ &\qquad +c^2 m(m+2)E\left[\chi_p^{-2}(\Delta) I\left(\frac{\chi^2_p(\Delta)}{\chi_{m+4}} \le c\right)\right]\Big\} \end{aligned} \tag{5.16}$$

where the terms inside the first curly braces is $R(\hat{\boldsymbol{\mu}}_1^s;\boldsymbol{\mu}_1)$, the quadratic risk of the SE of $\boldsymbol{\mu}_1$. Hence the proof.

Figure 2, displays the quadratic risk functions of the SE, PRSE and MLE for different values of p and Δ when $\Sigma = I_p$, $n_1 = 10$ and $n_2 = 15$.

Analysis of Risk: It is well known that the quadratic risk of the MLE is fixed at a constant value, p for the one sample problem. The risk of both the SE and the PRSE

Figure 2: The graph of the quadratic risk of the MLE, SE and PRSE for $p = 4, 6, 8, 10$.

is always smaller than p, regardless of the value of p (> 2) and Δ. Thus the MLE is not admissible in the estimation of the mean vector of the two-sample multivariate normal problem, a result well known for the one sample multivariate problem with $p > 2$. However, in this section our objective is to compare the risks of the SE and the PRSE. The quadratic risk function of both shrinkage estimators depends on Δ. When the null hypothesis is true, then $\Delta = 0$, and the difference between the risks of the SE and PRSE is the largest. This is true for all values of p. However, for fixed sample sizes, this difference decreases as the value of p increases.

The scenario around the behaviour of the two risk curves are not much different as Δ departs from 0 and the value of p changes. The risks of both the estimators are considerably lower than that of the MLE in the neighborhood of $\Delta = 0$. The quadratic risk curves of both the SE and PRSE grow larger and larger as the value of Δ increases, and approaches to that of the MLE (which is fixed at p) from below but never crosses it.

For a smaller value of p (say $p = 4$, in Figure 2), the risk curve of the PRSE remains at a lower level than that of the SE before the later curve moves towards the former from below. Then as Δ grows larger the quadratic risk of the PRSE approaches that of the SE. The risk curves of the two estimators merges to a single curve for very large values of Δ. From Figure 2, it is clear that the dominance of the PRSE over the SE becomes less significant as the dimension of the multivariate normal population grows higher. Thus the PRSE uniformly dominates the SE under the quadratic risk criterion, and does more so on the higher dimension. The difference between the quadratic risk functions of the estimators decreases as the sample size of one or both samples increases.

5.3 Theorem

For the two-sample multivariate normal problem described in Section 2, the relationship between the quadratic risks of the SE and PRSE of $\boldsymbol{\mu}_1$ *is given by*

$$R(\hat{\boldsymbol{\mu}}_1^{s+}; \boldsymbol{\mu}_1) \leq R(\hat{\boldsymbol{\mu}}_1^{s}; \boldsymbol{\mu}_1) \tag{5.17}$$

for all p *and for all* Δ.

Proof: First note that the relation in (3.9) can be written as

$$(\hat{\boldsymbol{\mu}}_1^{s+} - \boldsymbol{\mu}_1) = (\hat{\boldsymbol{\mu}}_1^{s} - \boldsymbol{\mu}_1) - M(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 < c)(cT^{-2} - 1). \tag{5.18}$$

Then from the definition, the difference between the quadratic risks of the PRSE and the SE of $\boldsymbol{\mu}_1$ is given by

$$\begin{aligned} D(\hat{\boldsymbol{\mu}}_1^{s+}; \hat{\boldsymbol{\mu}}_1^{s}) = & -M^2 E[n_1(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 < c)(cT^{-2} - 1)^2] \\ & -2M^2 E[n_1(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)'\Sigma^{-1}(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)I(T^2 < c)(cT^{-2} - 1)] \end{aligned} \tag{5.19}$$

where the conditional mean of $(\tilde{\boldsymbol{\mu}}_1 - \boldsymbol{\mu}_1)$, given $(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1)$, which is,

$$-M\{(\tilde{\boldsymbol{\mu}}_2 - \tilde{\boldsymbol{\mu}}_1) - (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)\} \tag{5.20}$$

has been used in the calculation. Since the definition of the PRSE requires $I(T^2 < c)$, we have $cT^{-2} > 1$, and hence $(cT^{-2} - 1) > 0$ in the last factor of both terms on the right hand side of the above equation (5.19). Because every other factor on the right hand side of the above equation are positive, and hence both terms are negative, the value of $D(\hat{\boldsymbol{\mu}}_1^{s+}; \hat{\boldsymbol{\mu}}_1^{s})$ is always negative. Thus the quadratic risk of the PRSE is always less than or equal to that of the SE. Hence the proof.

6 Concluding Remarks

It has been revealed by the foregoing analysis that the SE and the PRSE have almost identical quadratic bias for all values of $\Delta > 0$ and for all moderate values of p.

So, if the criterion of comparison is to minimize the quadratic bias of the estimator, the SE and PRSE are not much different if p is not too large. However, in real life this particular criterion is not so popular. In fact, people are more interested in the minimization of the risk than the bias of the estimators. It has been observed in Figure 2 that there is uniform domination of the PRSE over the SE, for all values of Δ and for all p. Moreover, the dominance of the PRSE over the SE is larger near $\Delta = 0$ and it is the largest at $\Delta = 0$ for all values of p.

From the analysis of the quadratic risk functions of the SE and PRSE it is evident that the risk of the PRSE is always less than or equal to that of the SE. Thus PRSE uniformly over performs the SE for all p (> 2) and for all values of Δ. Hence, if the criteria of comparison is to minimize the quadratic risk function, the obvious best choice is the PRSE.

In this paper we have considered the populations with equal covariance structure. The case of unequal covariance matrix remains to be an open problem.

ACKNOWLEDGEMENTS

The first author thankfully acknowledges the thoughtful advice and valuable suggestion of Professor A.K.Md.E. Saleh, Carleton University, Canada that significantly improved the derivation of results of the paper. The excellent research facilities provided by the Sultan Qaboos University, Muscat, Oman is also acknowledged.

REFERENCES

1. Ahmed, S.E. and Saleh, A.K.Md.E. (1989). Pooling multivariate data. Jou. Statist. Comput. Simul., Vol. 31, 149-167.

2. Anderson, T.W. (1984). Introduction to Multivariate Analysis, 2nd ed., Wiley, New York.

3. Bancroft, T.A. (1944). On biases in estimation due to the use of the preliminary tests of significance. Ann. Math. Statist., Vol. 15, 190-204.

4. Bancroft, T.A. and Hon, C.P. (1976). On pooling mean in multivariate normal distributions. Essays in Probability and Statistics, Ikeda et al. ed., Tokyo, 353-366.

5. Chang, C.H.; Lin, J.J. and Pal, N. (1993). Improvements over the James-Stein estimator: a risk analysis. Jou. Statist. Comput. Simulation, **48**, 117-126.

6. Chang, Y.T. (1995). Two-Stage James-Stein estimators of the mean based on prior knowledge. Comm in Statist.: Theory & Method, **24**, 2211-2227.

7. Han, C.P. an Bancroft, T.A. (1968). On pooling means when variance is unknown. Jou. Amer. Statist. Assoc., Vol. 63, 1333- 1342.

8. James, W. and Stein, C. (1961). Estimation with quadratic loss. Proceedings of the Fourth Berkeley Symposium on Math. Statist. and Probability, University of California Press, Berkeley, Vol. 1, 361-379.

9. Judge, G.G. and Bock, M.E. (1978). The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics. North- Holland, New York.

10. Khan, S. (1997). Likelihood based inference on the mean vectors of two multivariate Student-t populations. Far East Journal of Theoretical Statistics. Vol. 1, 77-91.

11. Khan, S. (1998). On the estimation of mean vector of Student-t population with uncertain prior information. Pakistan Journal of Statistics. **14**, 161-175.

12. Khan, S. and Saleh, A.K.Md.E. (1997). Shrinkage pre-test estimator of the intercept parameter for a regression model with multivariate Student-t errors. Biometrical Journal. Vol. 39, 131-147.

13. Khan, S. and Saleh, A.K.Md.E. (1998). Comparison of estimators of the mean based on p-samples from multivariate Student-t population. Communications in Statistics: Theory & Methods. Vol. 27(1), 193-210.

14. Maatta, J.M. and Casella, G. (1990). Developments in decision-theoretic variance estimation, Statistical Sciences, Vol. 5, 90-101.

15. Saleh, A.K.Md.E. (1973). Pretest estimator for the means of multivariate normal distribution. Carleton Mathematical Series, No. 94, Carleton University, Canada.

16. Saleh, A.E.Md.E. and Sen, P.K. (1985). Shrinkage least squares estimation in a general multivariate linear model. Proceedings of the Fifth Pannonian Symposium on Mathematical Statistics, 307-325.

17. Saleh, A.K.Md.E. and Sen, P.K. (1978). Nonparametric estimation of location parameter after a preliminary test on regression. Annals of Statistics, Vol 6, 154-168.

18. Sclove, S.L., Morris, C. and Radhakrishnan, R. (1972). Non-optimality of preliminary-test estimators for the mean of a multivariate normal distribution. Ann. Math. Statist. Vol. 43, 1481-1490.

19. Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, Proceedings of the Third Berkeley Symposium on Math. Statist. and Probability, University of California Press, Berkeley, Vol. 1, 197-206.

20. Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. Ann. Statist. Vol. 9, 1135-1151.