

## On the Properties of Estimates of Monotonic Mean Vectors for Multivariate Normal Distributions

Abouzar Bazyari

*Department of Statistics, Persian Gulf University,  
Bushehr, Iran  
ab\_bazyari@yahoo.com*

Received 13 February 2014

Accepted 16 November 2014

### Abstract

Problems concerning estimation of parameters and determination the statistic, when it is known a priori that some of these parameters are subject to certain order restrictions, are of considerable interest. In the present paper, we consider the estimators of the monotonic mean vectors for two dimensional normal distributions and compare those with the unrestricted maximum likelihood estimators under two different cases. One case is that covariance matrices are known, the other one is that covariance matrices are completely unknown and unequal. We show that when the covariance matrices are known, under the squared error loss function which is similar to the mahalanobis distance, the obtained multivariate isotonic regression estimators, motivated by estimators given in Robertson et al. (1988), which are the estimators given by Sasabuchi et al. (1983) and Sasabuchi et al. (1992), have the smaller risk than the unrestricted maximum likelihood estimators uniformly, but when the covariance matrices are unknown and unequal, the estimators have the smaller risk than the unrestricted maximum likelihood estimators only over some special sets which are defined on the covariance matrices. To illustrate the results two numerical examples are presented.

*Keywords:* Maximum likelihood estimator, Multivariate normal distribution, Monotonic mean vectors, Squared error loss function.

2000 Mathematics Subject Classification: primary 62F30, secondary 62H12

### 1. Introduction

Suppose that  $\mathbf{X}_{i1}, \mathbf{X}_{i2}, \dots, \mathbf{X}_{in_i}$  are random vectors from a  $p$ -dimensional normal distribution with unknown mean vector  $\mu_i$  and known nonsingular covariance matrix  $\Sigma_i, i = 1, 2, \dots, k$ . Consider the monotonic mean vectors  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  where  $\mu_v \leq \mu_i$  means that all the elements of  $\mu_i - \mu_v$  are nonnegative. In fact the vector set  $\mu_1, \mu_2, \dots, \mu_k$  is said to be isotonic with respect to  $\leq$  if  $v \leq i$ , then  $\mu_v \leq \mu_i$ . This kind of order restriction on the mean vectors representation is common, for instance, in selection and ranking problem for finding the largest element of several normal means (see Shimodaira, 2000). Also in the problems with testing hypothesis, Silvapulle and Sen (2005) discuss other examples from different areas, especially in medicine.

Table 1. Serum enzyme level in rats

Dosage	Enzyme	1	2	3	4	5	6	7	8	9	10
0	SDH	18	27	16	21	26	22	17	27	26	27
	SGPOT	101	103	90	98	101	92	123	105	92	88
	SGPT	65	67	52	58	64	60	66	63	68	56
1500	SDH	25	21	24	19	21	22	20	25	24	27
	SGPOT	113	99	102	144	109	135	100	95	89	98
	SGPT	65	63	70	73	67	66	58	53	58	65
5000	SDH	22	21	22	30	25	21	29	22	24	21
	SGPOT	88	95	104	92	103	96	100	122	102	107
	SGPT	54	56	71	59	61	57	61	59	63	61
15000	SDH	31	26	28	24	33	23	27	24	28	29
	SGPOT	104	123	105	98	167	111	130	93	99	99
	SGPT	57	61	54	56	45	49	57	51	51	48

As an application of order restrictions on the mean vectors for several populations, the problem we are considering comes from Dietz (1989).

**Example 1 (Dietz, 1989).** Vinylidene fluoride is suspected of causing liver damage. An experiment was carried out to evaluate its effects. Four groups of 10 male Fischer-344 rats received, by inhalation exposure, one of several dosages of vinylidene fluoride. Among the response variables measured on the rats were three serum enzymes: SDH, SGPOT, and SGPT. Increased levels of these serum levels are often associated with liver damage. It is of interest to test whether or not these enzyme levels are affected by vinylidene fluoride. The data are given in Table 1. (Serum enzyme levels are in international units/liter; dosage of vinylidene in parts/million.)

Let  $X_{ij} = (X_{ij1}, X_{ij2}, X_{ij3})'$  denote the observations on the three enzymes for  $j$ th subject  $j = 1, 2, \dots, 10$  in treatment  $i$  ( $i = 1, 2, 3, 4$ ). Let  $\mu_{ik}$  denote the mean response for  $i$ th treatment (i.e., dose) and  $k$ th variable and let  $\mu_i = (\mu_{i1}, \mu_{i2}, \mu_{i3})'$  for  $i = 1, 2, 3, 4$ . The researcher, especially if four populations reflect increasing levels of a treatment, may be interested in the alternative hypothesis  $\mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4$ , with at least one strict inequality. Now, one formulation of the null and alternative hypothesis is

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 \quad \text{and} \quad H_1 : \mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4,$$

(with at least one strict inequality in  $H_1$ .)

The basic works on the ordered means started by Bartholomew (1959), who considered problem of testing the homogeneity of  $k$  univariate normal means against an order restricted alternative hypothesis. The most well known and extensively studied approach to obtain the test statistic is the likelihood ratio method. He used the well-known method, Pool Adjacent Violators Algorithm to estimate the unknown ordered means. Much of the developments and theories on this subject are assembled in Barlow et al. (1972), Robertson et al. (1988) and in Silvapulle and Sen (2005). Sasabuchi et al. (1983) extended Bartholomew's (1959) problem to multivariate normal mean vectors with known covariance matrices. They derived the likelihood ratio test and proposed an iterative algorithm for computing the bivariate isotonic regression. Sasabuchi et al. (1992) generalized the iterative algorithm to multivariate isotonic regression. Fernando and Kulatunga (2007) proposed a

FORTTRAN program for the computation of multivariate isotonic regression and studied the convergence of the algorithm when the dimension is greater than or equal to five through Monte Carlo simulation. Hansohm and Hu (2012) gave a more general framework for multivariate isotonic regression and an algorithm based on Dykstras method is used to compute the multivariate isotonic regression. For more review and details on this kind of problems of testing specially in multivariate normal distributions one may refer to Bazyari (2012), Bazyari and Chinipardaz (2012), Bazyari and Chinipardaz (2013) and Bazyari and Pesarin (2013).

Suppose that  $X_{ij}$  be the  $j$ th  $p \times 1$  observation vector of the  $i$ th population and mutually independently distributed as  $N_p(\mu_i, \Sigma_i)$ ,  $i = 1, 2, \dots, k$ ,  $j = 1, 2, \dots, n_i$ , where the following monotonic order restriction on the unknown parameters  $\mu_i$ ,  $i = 1, 2, \dots, k$ ,  $j = 1, 2, \dots, n_i$ , is defined as

$$\mu_1 \leq \mu_2 \leq \dots \leq \mu_k, \tag{1.1}$$

(with at least one strict inequality), where  $\mu_v \leq \mu_i$  means that all the elements of the vector  $\mu_i = \mu_v$  are nonnegative.

For the  $i$  th population, let  $\mathbf{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{X}_{ij}$  and  $\mathbf{S}_i = \frac{1}{n_i-1} \sum_{j=1}^{n_i} (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)'$  be the sample mean vector and sample covariance matrix. We note that  $(n_i - 1)\mathbf{S}_i$  is distributed as Wishart distribution  $W_p(n_i - 1, \Sigma_i)$ .

The rest of this paper is organized as follows. In Section 2, the obtained results by some authors are provided and the squared error loss function of the estimators of mean vectors is given. The closed forms of the multivariate isotonic regression estimators for two dimensional normal distributions and also the main theorems are given in Section 3. In Section 4, the proofs of the theorems of Section 3 are derived. In Section 5, the application of the results is presented. Concluding remarks are given in Section 6. The complete source programs are written in SPLUS software.

## 2. Literature review

In the mentioned problem if  $p = 1$ , then it is clear that we have univariate case and replace  $\sigma_i^2$  instead of  $\Sigma_i$ . In this case, for  $k > 2$ , when the variances are known and  $\mu_1 = \mu_2 = \dots = \mu_k = \mu$ , then the unbiased estimator of  $\mu$  is given by

$$\hat{\mu} = \frac{\sum_{i=1}^k \binom{n_i}{\sigma_i^2} \bar{X}_i}{\sum_{i=1}^k \binom{n_i}{\sigma_i^2}}, \tag{2.1}$$

where  $\bar{X}_i = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i}$ ,  $i = 1, 2, \dots, k$ , is the mean sample of  $i$  th population. But if we suppose that variances are unknown, so a suitable estimator is Graybill and Deal (1959) estimator as follows

$$\hat{\mu}_{GD} = \frac{\sum_{i=1}^k \binom{n_i}{\hat{\sigma}_i^2} \bar{X}_i}{\sum_{i=1}^k \binom{n_i}{\hat{\sigma}_i^2}}. \tag{2.2}$$

They found out conditions under which  $\hat{\mu}_{GD}$  (which is also unbiased for  $\mu$ ) has smaller variance than  $\bar{X}_j$ . Now, suppose that the order restriction  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  is imposed among several univariate normal mean vectors. Robertson et al. (1988), showed that the restricted maximum likelihood

estimator of  $\mu_i$  is

$$\min_{s \geq i+1} \max_{r \leq i} \frac{\sum_{j=r}^{s-1} \left(\frac{n_j}{\sigma_j^2}\right) \bar{X}_j}{\sum_{j=r}^{s-1} \left(\frac{n_j}{\sigma_j^2}\right)}. \tag{2.3}$$

Lee (1981), studied the estimator given in (2.3) and showed that this estimator uniformly improves upon  $\bar{X}_j$ . Kelly (1989), considered Lee’s results and showed that (2.3) universally dominates  $\bar{X}_j$ . Shi and Jiang (1998), supposed that the data draw from  $j$  univariate normal populations with unknown variances and discussed some properties of the maximum likelihood estimates of means under the restriction and proposed an algorithm for obtaining the estimates.

The above results are obtained in univariate normal distribution. In this case, the common mean problem has received considerable attention in the other statistical literatures, for instant, Krishnamoorthy and Lu (2003), Lin and Lee (2005) and the references therein. In multivariate case i.e.  $p \geq 2$ , Chiou and Cohen (1985), showed that under  $\mu_1 = \mu_2 = \mu$ , the estimators  $\bar{\mathbf{X}}_1$  and  $\bar{\mathbf{X}}_2$  dominate the estimator

$$\hat{\mu}_{CC} = \left( \sum_{i=1}^2 n_i \mathbf{S}_i^{-1} \right)^{-1} \sum_{i=1}^2 n_i \mathbf{S}_i^{-1} \bar{\mathbf{X}}_i, \tag{2.4}$$

although Grabill and Deal (1959) got the opposite result in univariate two sample case. Loh (1991) derived the unbiased estimator for the common mean of two multivariate normal distributions with unknown covariance matrices with respect to the natural quadratic loss function. Also they presented a Monte Carlo swindle to evaluate its risk. Zhou and Mathew (1994) proposed several combined tests for testing the common mean vector, but the problem of multiple comparisons had not been discussed when the null hypothesis was rejected. For multivariate distributions, when the covariance matrices are unknown, Jordan and Krishnamoorthy (1995), provided a confidence region of  $\mu$  centered at a weighted Graybill and Deal estimator

$$\hat{\mu}_{JK} = \left( \sum_{i=1}^2 d_i n_i \mathbf{S}_i^{-1} \right)^{-1} \sum_{i=1}^2 d_i n_i \mathbf{S}_i^{-1} \bar{\mathbf{X}}_i, \tag{2.5}$$

to find the confidence intervals for the common mean of several normal populations, where  $d_i$ ’s are some positive constants defined by

$$d_i = \frac{(\text{Var}(T_i)^2)^{-1}}{\sum_{i=1}^k (\text{Var}(T_i)^2)^{-1}},$$

where  $T_i^2 = \sum_{i=1}^k (\bar{\mathbf{X}}_i - \mu_i)' \mathbf{S}_i^{-1} (\bar{\mathbf{X}}_i - \mu_i)$  and  $\text{Var}(T_i^2) = \frac{2p(n_i-1)^2(n_i-2)}{(n_i-p-2)^2(n_i-p-4)}$ .

Thus Graybill and Deals method can be treated as a special case of Jordan and Krishnamoorthy (1995)’s method with  $d_1 = d_2 = \dots = d_k$ . Lin et al. (2007), considered the confidence region for the common mean vector of several multivariate normal populations when the covariance matrices are unknown and unequal. A generalized confidence region is derived using the concepts of generalized method based on the generalized p-value.

We consider the following squared error loss function of the estimators of  $\mu_i$

$$L(\mu_i, \hat{\mu}_i, \mathbf{M}) = (\hat{\mu}_i - \mu_i)' \mathbf{M} (\hat{\mu}_i - \mu_i), \tag{2.6}$$

where  $\mathbf{M}$  is a positive definite matrix. Then the risk is given by

$$R(\mu_i, \hat{\mu}_i, \mathbf{M}) = E[L(\mu_i, \hat{\mu}_i, \mathbf{M})].$$

In fact, the given loss function in (2.6) is similar to the mahalanobis distance which the positive definite matrix  $\mathbf{M}$  is replaced instead of  $S^{-1}$ .

### 3. The main theorems

As we expressed, in this paper we consider two multivariate normal populations such that the monotonic order restriction given in (1.1) is imposed. Suppose that the covariance matrices are known. Then motivated by univariate restricted maximum likelihood estimators given by Robertson et al. (1988), we obtain the following isotonic estimators of  $\mu_1$  and  $\mu_2$  in the multivariate case

$$\hat{\mu}_1^* = \bar{\mathbf{X}}_1 (1 - I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}) + \hat{\mu}^* I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}, \tag{3.1}$$

and

$$\hat{\mu}_2^* = \bar{\mathbf{X}}_2 (1 - I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}) + \hat{\mu}^* I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}, \tag{3.2}$$

where  $I_d$  denotes the indicator function of set satisfying the condition  $d$  and

$$\hat{\mu}^* = \left( \left( \frac{\Sigma_1}{n_1} \right)^{-1} + \left( \frac{\Sigma_2}{n_2} \right)^{-1} \right)^{-1} \left( \left( \frac{\Sigma_1}{n_1} \right)^{-1} \bar{\mathbf{X}}_1 + \left( \frac{\Sigma_2}{n_2} \right)^{-1} \bar{\mathbf{X}}_2 \right). \tag{3.3}$$

Now, suppose that the covariance matrices are unknown and unequal, then motivated by the idea of Graybill and Deal (1959) and also Robertson et al. (1988) in univariate case, we estimate the unknown covariance matrices,  $\Sigma_i$  by  $S_i$ ,  $i = 1, 2$ . Therefore, we have

$$\hat{\mu}_1 = \bar{\mathbf{X}}_1 (1 - I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}) + \hat{\mu} I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}, \tag{3.4}$$

and

$$\hat{\mu}_2 = \bar{\mathbf{X}}_2 (1 - I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}) + \hat{\mu} I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}, \tag{3.5}$$

where

$$\hat{\mu} = \left( \left( \frac{S_1}{n_1} \right)^{-1} + \left( \frac{S_2}{n_2} \right)^{-1} \right)^{-1} \left( \left( \frac{S_1}{n_1} \right)^{-1} \bar{\mathbf{X}}_1 + \left( \frac{S_2}{n_2} \right)^{-1} \bar{\mathbf{X}}_2 \right), \tag{3.6}$$

where  $\mathbf{G} = \left( \frac{S_1}{n_1} + \frac{S_2}{n_2} \right)^{-1}$ .

We present the following theorems for the obtained estimators

**Theorem 1.** Under the squared error loss function given in (2.6), the isotonic regression estimator  $\hat{\mu}_1^*$  uniformly has the smaller risk than the unrestricted maximum likelihood estimator,  $\bar{\mathbf{X}}_1$ .

**Theorem 2.** Under the squared error loss function given in (2.6), the isotonic regression estimator  $\hat{\mu}_2^*$  uniformly has the smaller risk than the unrestricted maximum likelihood estimator,  $\bar{\mathbf{X}}_2$ .

These results establish for the estimators  $\hat{\mu}_1$  and  $\hat{\mu}_2$  but under the following special sets which are defined on the covariance matrices.

**Theorem 3.** Under the squared error loss function given in (2.6), the isotonic regression estimator  $\hat{\mu}_1$  uniformly has the smaller risk than the unrestricted maximum likelihood estimator  $\bar{\mathbf{X}}_1$  over the set  $\mathbf{A}_1 = \{\Sigma : n_2 \Sigma_1 \geq n_1 \Sigma_2\}$ .

**Theorem 4.** Under the squared error loss function given in (2.6), the isotonic regression estimator  $\hat{\mu}_2$  uniformly has the smaller risk than the unrestricted maximum likelihood estimator  $\bar{\mathbf{X}}_2$  over the set  $\mathbf{A}_2 = \{\Sigma : n_2 \Sigma_1 \leq n_1 \Sigma_2\}$ .

Note that, for  $p \times p$  matrices  $A$  and  $B$ ,  $A \geq B$ , means that  $A - B$  is a non-negative definite matrix.

#### 4. The proofs of the given theorems

##### Proof of theorem 1.

Put  $\mathbf{U} = \left( \left( \frac{\Sigma_1}{n_1} \right)^{-1} + \left( \frac{\Sigma_2}{n_2} \right)^{-1} \right)^{-1} \left( \frac{\Sigma_1}{n_1} \right)^{-1} = \frac{\Sigma_2}{n_2} \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1}$  and define the positive definite matrix

$$\mathbf{M} = \left( \frac{\Sigma_1}{n_1} \right)^{-1} \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^2 \left( \frac{\Sigma_1}{n_1} \right)^{-1}.$$

To proof the theorem 1, we must show that the following inequality is true

$$R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1^*, \mathbf{M}) > 0, \tag{4.1}$$

the left hand side of the inequality denotes the risk difference of  $\bar{\mathbf{X}}_1$  and  $\hat{\mu}_1^*$ .

Then the estimator  $\bar{\mathbf{X}}_1$  and  $\hat{\mu}_1^*$  given in (3.1) can be express as

$$\hat{\mu}_1^* = \bar{\mathbf{X}}_1 (1 - I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}) + (\mathbf{U} \bar{\mathbf{X}}_1 + (I - \mathbf{U}) \bar{\mathbf{X}}_2) I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}. \tag{4.2}$$

Therefore, the risk difference of  $\bar{\mathbf{X}}_1$  and  $\hat{\mu}_1^*$  is as

$$\begin{aligned} R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1^*, \mathbf{M}) &= E[(\bar{\mathbf{X}}_1 - \mu_1)' \mathbf{M} (\bar{\mathbf{X}}_1 - \mu_1)] \\ &\quad - (\hat{\mu}_1^* - \mu_1)' \mathbf{M} (\hat{\mu}_1^* - \mu_1) I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2} \\ &\quad - E[(\bar{\mathbf{X}}_1 - \mu_1)' \mathbf{M} (\bar{\mathbf{X}}_1 - \mu_1)] \\ &\quad - (\mathbf{U} \bar{\mathbf{X}}_1 + (I - \mathbf{U}) \bar{\mathbf{X}}_2 - \mu_1)' \mathbf{M} (\mathbf{U} \bar{\mathbf{X}}_1 + (I - \mathbf{U}) \bar{\mathbf{X}}_2 - \mu_1) I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}. \end{aligned} \tag{4.3}$$

Now, define the transformations

$$\mathbf{Z}_1 = \bar{\mathbf{X}}_1 - \mu_1 \quad \text{and} \quad \mathbf{Z}_2 = \bar{\mathbf{X}}_2 - \mu_1. \tag{4.4}$$

Then it is clearly that the vectors  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  are mutually independently distributed as  $N_p \left( 0, \frac{\Sigma_1}{n_1} \right)$  and  $N_p \left( \mu_2 - \mu_1, \frac{\Sigma_2}{n_2} \right)$  respectively, and

$$\mathbf{Z}_1 = \left( I + \frac{\Sigma_1}{n_1} \left( \frac{\Sigma_2}{n_2} \right)^{-1} \right)^{-1} \left( \frac{\Sigma_1}{n_1} \left( \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_1 + \mathbf{Y}_2 \right),$$

and

$$\mathbf{Z}_2 = \left( I + \frac{\Sigma_1}{n_1} \left( \frac{\Sigma_2}{n_2} \right)^{-1} \right)^{-1} (\mathbf{Y}_2 - \mathbf{Y}_1),$$

where  $\mathbf{Y}_1 = \mathbf{Z}_1 - \mathbf{Z}_2$  and  $\mathbf{Y}_2 = \mathbf{Z}_1 + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{Z}_2$ , then  $\mathbf{Y}_1$  and  $\mathbf{Y}_1$  are mutually independent and we have

$$\mathbf{Y}_1 \sim N_p \left( \mu_1 - \mu_2, \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right),$$

and

$$\mathbf{Y}_2 \sim N_p \left( \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} (\mu_2 - \mu_1), \frac{\Sigma_1}{n_1} + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1} \right).$$

Thus from (4.3), we have

$$\begin{aligned} R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1^*, \mathbf{M}) &= E[\mathbf{Z}'_1 \mathbf{M} \mathbf{Z}_1 - (\mathbf{U} \mathbf{Z}_1 + (\mathbf{I} - \mathbf{U}) \mathbf{Z}_2)' \\ &\quad \times \mathbf{M} (\mathbf{U} \mathbf{Z}_1 + (\mathbf{I} - \mathbf{U}) \mathbf{Z}_2)] I_{\mathbf{Z}_1 > \mathbf{Z}_2} \\ &= E[\mathbf{Z}'_1 \mathbf{U}' \mathbf{M} (\mathbf{I} - \mathbf{U}) (\mathbf{Z}_1 - \mathbf{Z}_2) \\ &\quad + (\mathbf{Z}_1 - \mathbf{Z}_2)' (\mathbf{I} - \mathbf{U})' \mathbf{M} \mathbf{U} \mathbf{Z}_1] I_{\mathbf{Z}_1 > \mathbf{Z}_2} \\ &\quad + E[\mathbf{Z}'_1 (\mathbf{I} - \mathbf{U})' \mathbf{M} (\mathbf{I} - \mathbf{U}) \mathbf{Z}_1 \\ &\quad - \mathbf{Z}'_2 (\mathbf{I} - \mathbf{U})' \mathbf{M} (\mathbf{I} - \mathbf{U}) \mathbf{Z}_2] I_{\mathbf{Z}_1 > \mathbf{Z}_2}, \end{aligned}$$

Now, since

$$\begin{aligned} &E[\mathbf{Z}'_1 \mathbf{U}' \mathbf{M} (\mathbf{I} - \mathbf{U}) (\mathbf{Z}_1 - \mathbf{Z}_2) + (\mathbf{Z}_1 - \mathbf{Z}_2)' (\mathbf{I} - \mathbf{U})' \mathbf{M} \mathbf{U} \mathbf{Z}_1] I_{\mathbf{Z}_1 > \mathbf{Z}_2} \\ &= E \left[ \left( \mathbf{Y}_2 \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{Y}_1 \right)' \left( \mathbf{I} + \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1} \right)^{-1} \mathbf{U}' \mathbf{M} (\mathbf{I} - \mathbf{U}) \mathbf{Y}_1 \right. \\ &\quad \left. + \mathbf{Y}_1 (\mathbf{I} - \mathbf{U})' \mathbf{M} \mathbf{U} \left( \mathbf{I} + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \right)^{-1} \left( \mathbf{Y}_2 + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{Y}_1 \right) \right] I_{\mathbf{Y}_1 > 0}. \end{aligned}$$

and

$$\begin{aligned} &E[\mathbf{Z}'_1 (\mathbf{I} - \mathbf{U})' \mathbf{M} (\mathbf{I} - \mathbf{U}) \mathbf{Z}_1 - \mathbf{Z}'_2 (\mathbf{I} - \mathbf{U})' \mathbf{M} (\mathbf{I} - \mathbf{U}) \mathbf{Z}_2] I_{\mathbf{Z}_1 > \mathbf{Z}_2} \\ &= E \left[ \left( \mathbf{Y}_2 + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{Y}_1 \right)' \left( \mathbf{I} + \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1} \right)^{-1} (\mathbf{I} - \mathbf{U})' \mathbf{M} (\mathbf{I} - \mathbf{U}) \left( \mathbf{I} + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \right)^{-1} \right. \\ &\quad \times \left( \mathbf{Y}_2 + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{Y}_1 \right)' - (\mathbf{Y}_2 - \mathbf{Y}_1)' \left( \mathbf{I} + \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1} \right)^{-1} (\mathbf{I} - \mathbf{U})' \mathbf{M} (\mathbf{I} - \mathbf{U}) \\ &\quad \left. \times \left( \mathbf{I} + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \right)^{-1} (\mathbf{Y}_2 - \mathbf{Y}_1) \right] I_{\mathbf{Y}_1 > 0}. \end{aligned}$$

Then

$$\begin{aligned}
 R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1^*, \mathbf{M}) &= E \left[ \mathbf{Y}_1 \left( \frac{\Sigma_1}{n_1} \right)^{-1} \frac{\Sigma_2}{n_2} \mathbf{Y}_1 + \mathbf{Y}_1' \frac{\Sigma_2}{n_2} \left( \frac{\Sigma_1}{n_1} \right)^{-1} \mathbf{Y}_2 \right] I_{\mathbf{Y}_1 > 0} \\
 &\quad + E \left[ \mathbf{Y}_1' \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \frac{\Sigma_2}{n_2} \mathbf{Y}_1 + \mathbf{Y}_1 \frac{\Sigma_2}{n_2} \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_1 \right. \\
 &\quad + \mathbf{Y}_1' \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \left( \frac{\Sigma_1}{n_1} \right)^2 \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_1 \\
 &\quad \left. - \mathbf{Y}_1' \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \left( \frac{\Sigma_2}{n_2} \right)^2 \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_1 \right] I_{\mathbf{Y}_1 > 0} \\
 &= E \left[ \mathbf{Y}_2' \left( \frac{\Sigma_1}{n_1} \right)^{-1} \frac{\Sigma_2}{n_2} \mathbf{Y}_1 + \mathbf{Y}_1' \frac{\Sigma_2}{n_2} \left( \frac{\Sigma_1}{n_1} \right)^{-1} \mathbf{Y}_2 \right] + E[\mathbf{Y}_1' \mathbf{Y}_1] I_{\mathbf{Y}_1 > 0}.
 \end{aligned}$$

Since the vectors  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$  are mutually independent, and with paying attention to the increasing order restriction defined on the unknown parameters  $\mu_1, \mu_2$  and  $\mu_2 - \mu_1 > 0$ , we have

$$E \left[ \mathbf{Y}_2' \left( \frac{\Sigma_1}{n_1} \right)^{-1} \frac{\Sigma_2}{n_2} \mathbf{Y}_1 + \mathbf{Y}_1' \left( \frac{\Sigma_2}{n_2} \right) \left( \frac{\Sigma_1}{n_1} \right)^{-1} \mathbf{Y}_2 \right] I_{\mathbf{Y}_1 > 0} = E \left[ (\mu_2 - \mu_1)' \mathbf{Y}_1 + \mathbf{Y}_1' (\mu_2 - \mu_1) \right] I_{\mathbf{Y}_1 > 0} > 0.$$

Therefore

$$R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1^*, \mathbf{M}) > 0.$$

**Proof of theorem 2.**

The proof of this theorem is similar to the proof of theorem 1. The difference point is that in this theorem we define  $\mathbf{M} = \left( \frac{\Sigma_2}{n_2} \right)^{-1} \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^2 \left( \frac{\Sigma_2}{n_2} \right)^{-1}$ . Then it is clear that the estimator  $\hat{\mu}_2^*$  given in (3.2) can be express as

$$\hat{\mu}_2^* = \bar{\mathbf{X}}_2 (1 - I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}) + (\mathbf{U} \bar{\mathbf{X}}_1 + (\mathbf{I} - \mathbf{U}) \bar{\mathbf{X}}_2) I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}.$$

The risk difference of  $\bar{\mathbf{X}}_2$  and  $\hat{\mu}_2^*$  is as

$$\begin{aligned}
 R(\mu_2, \bar{\mathbf{X}}_2, \mathbf{M}) - R(\mu_2, \hat{\mu}_2^*, \mathbf{M}) &= E[(\bar{\mathbf{X}}_2 - \mu_2)' \mathbf{M} (\bar{\mathbf{X}}_2 - \mu_2) \\
 &\quad - (\hat{\mu}_2^* - \mu_2)' \mathbf{M} (\hat{\mu}_2^* - \mu_2)] I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2} \\
 &= E[(\bar{\mathbf{X}}_2 - \mu_2)' \mathbf{M} (\bar{\mathbf{X}}_2 - \mu_2) \\
 &\quad - (\mathbf{U} \bar{\mathbf{X}}_2 + (\mathbf{I} - \mathbf{U}) \bar{\mathbf{X}}_2 - \mu_2)' \mathbf{M} (\mathbf{U} \bar{\mathbf{X}}_1 + (\mathbf{I} - \mathbf{U}) \bar{\mathbf{X}}_2 - \mu_2)] I_{\bar{\mathbf{X}}_1 > \bar{\mathbf{X}}_2}.
 \end{aligned} \tag{4.5}$$

Again consider the following transformations

$$\mathbf{Z}_1 = \bar{\mathbf{X}}_1 - \mu_2 \quad \text{and} \quad \mathbf{Z}_2 = \bar{\mathbf{X}}_2 - \mu_2.$$

Then it is clearly that the vectors  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  are mutually independently distributed as  $N_p \left( \mu_1 - \mu_2, \frac{\Sigma_1}{n_1} \right)$  and  $N_p \left( 0, \frac{\Sigma_2}{n_2} \right)$  respectively, and also define the vectors

$$\mathbf{Z}_1 = \left( \mathbf{I} + \frac{\Sigma_2}{n_2} \left( \frac{\Sigma_1}{n_1} \right)^{-1} \right)^{-1} (\mathbf{Y}_1 + \mathbf{Y}_2),$$



and

$$\mathbf{Z}_2 = \left( I + \frac{\Sigma_1}{n_1} \left( \frac{\Sigma_2}{n_2} \right)^{-1} \right)^{-1} \left( \frac{\Sigma_1}{n_1} \left( \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_2 - \mathbf{Y}_1 \right),$$

where  $\mathbf{Y}_1 = \mathbf{Z}_1 - \mathbf{Z}_2$  and  $\mathbf{Y}_2 = \mathbf{Z}_2 + \frac{\Sigma_2}{n_2} \left( \frac{\Sigma_1}{n_1} \right)^{-1} \mathbf{Z}_1$ , then  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$  are mutually independent and we have

$$\mathbf{Y}_2 \sim N_p \left( \frac{\Sigma_2}{n_2} \left( \frac{\Sigma_1}{n_1} \right)^{-1} (\mu_2 - \mu_1), \frac{\Sigma_2}{n_2} + \frac{\Sigma_2}{n_2} \left( \frac{\Sigma_1}{n_1} \right)^{-1} \frac{\Sigma_2}{n_2} \right).$$

From (4.5), we have

$$\begin{aligned} R(\mu_2, \bar{\mathbf{X}}_2, \mathbf{M}) - R(\mu_2, \hat{\mu}_2^*, \mathbf{M}) &= E[\mathbf{Z}_2'(I - \mathbf{U})\mathbf{M}\mathbf{U}(\mathbf{Z}_2 - \mathbf{Z}_1) \\ &\quad + (\mathbf{Z}_2 - \mathbf{Z}_1)'\mathbf{U}'\mathbf{M}(I - \mathbf{U})\mathbf{Z}_2]I_{\mathbf{Z}_1 > \mathbf{Z}_2} \\ &= E\left[-\mathbf{Y}_2' \left( \frac{\Sigma_2}{n_2} \right)^{-1} \frac{\Sigma_1}{n_1} \mathbf{Y}_1 - \mathbf{Y}_1' \frac{\Sigma_1}{n_1} \left( \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_2 \right. \\ &\quad + \mathbf{Y}_1' \frac{\Sigma_1}{n_1} \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_1 + \mathbf{Y}_1' \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \frac{\Sigma_1}{n_1} \mathbf{Y}_1 \\ &\quad - \mathbf{Y}_1' \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \left( \frac{\Sigma_2}{n_2} \right)^2 \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_1 \\ &\quad \left. + \mathbf{Y}_1' \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \left( \frac{\Sigma_1}{n_1} \right)^{-1} \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_1 \right] I_{\mathbf{Y}_1 > 0} \\ &= E \left[ -\mathbf{Y}_2' \left( \frac{\Sigma_2}{n_2} \right)^{-1} \frac{\Sigma_1}{n_1} \mathbf{Y}_1 - \mathbf{Y}_1' \frac{\Sigma_1}{n_1} \left( \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_2 \right] I_{\mathbf{Y}_1 > 0} + E[\mathbf{Y}_1' \mathbf{Y}_1] I_{\mathbf{Y}_1 > 0}. \end{aligned}$$

Since the vectors  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$  are mutually independent and  $\mu_2 - \mu_1 > 0$ , with paying attention to the distribution of  $\mathbf{Y}_2$ , we have

$$E \left[ -\mathbf{Y}_2' \left( \frac{\Sigma_2}{n_2} \right)^{-1} \frac{\Sigma_1}{n_1} \mathbf{Y}_1 - \mathbf{Y}_1' \left( \frac{\Sigma_1}{n_1} \right) \left( \frac{\Sigma_2}{n_2} \right)^{-1} \mathbf{Y}_2 \right] I_{\mathbf{Y}_1 > 0} = 2(\mu_2 - \mu_1)' E[\mathbf{Y}_1 I_{\mathbf{Y}_1 > 0}] > 0.$$

Therefore

$$R(\mu_2, \bar{\mathbf{X}}_2, \mathbf{M}) - R(\mu_2, \hat{\mu}_2^*, \mathbf{M}) > 0.$$

**Remark.** For two dimensional normal distributions the estimators given in (3.1) and (3.2) are the obtained estimators by the well-known method, Pool Adjacent Violators Algorithm proposed by Sasabuchi et al. (1983) and also Sasabuchi et al. (1992) for more than two dimensional normal distributions.

**Proof of theorem 3.**

Put  $\mathbf{U} = \left( \left( \frac{\Sigma_1}{n_1} \right)^{-1} + \left( \frac{\Sigma_2}{n_2} \right)^{-1} \right)^{-1} \left( \frac{\Sigma_1}{n_1} \right)^{-1}$  and  $\mathbf{M} = \left( \frac{\Sigma_1}{n_1} \right)^{-1} \left( \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2} \right)^2 \left( \frac{\Sigma_1}{n_1} \right)^{-1}$ . Then the estimator  $\hat{\mu}_1$

given in (3.4) can be express as

$$\hat{\mu}_1 = \bar{\mathbf{X}}_1(1 - I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}) + (\mathbf{U}\bar{\mathbf{X}}_1 + (I - \mathbf{U})\bar{\mathbf{X}}_2)I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}.$$

The risk difference of  $\bar{\mathbf{X}}_1$  and  $\hat{\mu}_1$  is as

$$\begin{aligned} R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1, \mathbf{M}) &= E[(\bar{\mathbf{X}}_1 - \mu_1)' \mathbf{M}(\bar{\mathbf{X}}_1 - \mu_1)] \\ &\quad - (\hat{\mu}_1 - \mu_1)' \mathbf{M}(\hat{\mu}_1 - \mu_1) I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2} \\ &= E[(\bar{\mathbf{X}}_1 - \mu_1)' \mathbf{M}(\bar{\mathbf{X}}_1 - \mu_1)] \\ &\quad - (\mathbf{U}\bar{\mathbf{X}}_1 + (I - \mathbf{U})\bar{\mathbf{X}}_2 - \mu_1)' \mathbf{M}(\mathbf{U}\bar{\mathbf{X}}_1 + (I - \mathbf{U})\bar{\mathbf{X}}_2 - \mu_1) I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}. \end{aligned} \tag{4.6}$$

Define the transformations  $\mathbf{Z}_1 = \bar{\mathbf{X}}_1 - \mu_1$  and  $\mathbf{Z}_2 = \bar{\mathbf{X}}_2 - \mu_1$ . It is clearly that the vectors  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  are mutually independently distributed as  $N_p\left(0, \frac{\Sigma_1}{n_1}\right)$  and  $N_p\left(\mu_2 - \mu_1, \frac{\Sigma_2}{n_2}\right)$  respectively, and

$$\mathbf{Z}_1 = \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \left(\frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{G}^{-1} \mathbf{Y}_1 + \mathbf{Y}_2\right),$$

and

$$\mathbf{Z}_2 = \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} (\mathbf{Y}_2 - \mathbf{G}^{-1} \mathbf{Y}_1),$$

where  $\mathbf{Y}_1 = \mathbf{G}(\mathbf{Z}_1 - \mathbf{Z}_2)$  and  $\mathbf{Y}_2 = \mathbf{Z}_1 + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{Z}_2$ , then  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$  are mutually independent and we have

$$\mathbf{Y}_1 \sim N_p\left(\mu_1 - \mu_2, \frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2}\right),$$

and

$$\mathbf{Y}_2 \sim N_p\left(\frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} (\mu_2 - \mu_1), \frac{\Sigma_1}{n_1} + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1}\right).$$

Thus from (4.6), we have

$$\begin{aligned} &R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1^*, \mathbf{M}) \\ &= E[\mathbf{Z}_1' \mathbf{M} \mathbf{Z}_1 - (\mathbf{U} \mathbf{Z}_1 + (I - \mathbf{U}) \mathbf{Z}_2)' \mathbf{M} (\mathbf{U} \mathbf{Z}_1 + (I - \mathbf{U}) \mathbf{Z}_2)] I_{\mathbf{G} \mathbf{Z}_1 > \mathbf{G} \mathbf{Z}_2} \\ &= E[\mathbf{Z}_1' \mathbf{U}' \mathbf{M} (I - \mathbf{U}) (\mathbf{Z}_1 - \mathbf{Z}_2) + (\mathbf{Z}_1 - \mathbf{Z}_2)' (I - \mathbf{U})' \mathbf{M} \mathbf{U} \mathbf{Z}_1] I_{\mathbf{G} \mathbf{Z}_1 > \mathbf{G} \mathbf{Z}_2} \\ &\quad + E[\mathbf{Z}_1' (I - \mathbf{U})' \mathbf{M} (I - \mathbf{U}) (\mathbf{Z}_1 + \mathbf{Z}_2' (I - \mathbf{U})' \mathbf{M} (I - \mathbf{U}) \mathbf{Z}_2)] I_{\mathbf{G} \mathbf{Z}_1 > \mathbf{G} \mathbf{Z}_2}. \end{aligned}$$

Now, since

$$\begin{aligned} &E[\mathbf{Z}_1' \mathbf{U}' \mathbf{M} (I - \mathbf{U}) (\mathbf{Z}_1 - \mathbf{Z}_2) + (\mathbf{Z}_1 - \mathbf{Z}_2)' (I - \mathbf{U})' \mathbf{M} \mathbf{U} \mathbf{Z}_1] I_{\mathbf{G} \mathbf{Z}_1 > \mathbf{G} \mathbf{Z}_2} \\ &= E\left[\left(\mathbf{Y}_2 + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{G}^{-1} \mathbf{Y}_1\right)' \left(I + \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1}\right)^{-1} \mathbf{U}' \mathbf{M} (I - \mathbf{U}) \mathbf{G}^{-1} \mathbf{Y}_1\right. \\ &\quad \left.+ \mathbf{Y}_1 \mathbf{G}^{-1} (I - \mathbf{U})' \mathbf{M} \mathbf{U} \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \left(\mathbf{Y}_2 + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{G}^{-1} \mathbf{Y}_1\right)\right] I_{\mathbf{Y}_1 > 0}, \end{aligned}$$

and

$$\begin{aligned}
 & E[\mathbf{Z}'_1(I-\mathbf{U})'\mathbf{M}(I-\mathbf{U})(\mathbf{Z}_1+\mathbf{Z}'_2(I-\mathbf{U})'\mathbf{M}(I-\mathbf{U})\mathbf{Z}_2)]I_{\mathbf{G}\mathbf{Z}_1>\mathbf{G}\mathbf{Z}_2} \\
 = & E\left[\left(\mathbf{Y}_2+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\mathbf{G}^{-1}\mathbf{Y}_1\right)'\left(I+\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\right)^{-1}(I-\mathbf{U}'\mathbf{M}(I-\mathbf{U})\left(I+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1}\right. \\
 & \times\left.\left(\mathbf{Y}_2+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\mathbf{G}^{-1}\mathbf{Y}_1\right)'-(\mathbf{Y}_2-\mathbf{G}^{-1}\mathbf{Y}_1)'\left(I+\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\right)^{-1}(I-\mathbf{U})'\mathbf{M}(I-\mathbf{U})\right. \\
 & \times\left.\left(I+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1}(\mathbf{Y}_2-\mathbf{G}^{-1}\mathbf{Y}_1)\right]I_{\mathbf{Y}_1>0}.
 \end{aligned}$$

Then

$$\begin{aligned}
 R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1^*, \mathbf{M}) = & E[\mathbf{Y}'_2\left(I+\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\right)^{-1}\mathbf{M}(I-\mathbf{U})\mathbf{G}^{-1}\mathbf{Y}_1 \\
 & +\mathbf{Y}'_1\mathbf{G}^{-1}(I-\mathbf{U})'\mathbf{M}\mathbf{Y}'_2\left(I+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1}\mathbf{Y}_2]I_{\mathbf{Y}_1>0} \\
 & +E[\mathbf{Y}'_1\mathbf{G}^{-1}\left(\frac{\Sigma_1}{n_1}+\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\mathbf{U}'\mathbf{M}(I-\mathbf{U})\mathbf{G}^{-1}\mathbf{Y}_1 \\
 & +\mathbf{Y}_1\mathbf{G}^{-1}(I-\mathbf{U})\mathbf{M}\mathbf{U}\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_1}{n_1}+\frac{\Sigma_2}{n_2}\right)^{-1}\mathbf{G}^{-1}\mathbf{Y}_1 \\
 & +\mathbf{Y}'_1\mathbf{G}^{-1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\left(I+\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\right)^{-1} \\
 & \times(I-\mathbf{U})'\mathbf{M}(I-\mathbf{U})\left(I+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1}\frac{1}{n_1} \\
 & \times\Sigma_1\left(\frac{\Sigma_2}{n_2}\right)^{-1}\mathbf{G}^{-1}\mathbf{Y}_1-\mathbf{Y}'_1\mathbf{G}^{-1}\left(I+\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\right)^{-1} \\
 & \times(I-\mathbf{U})'\mathbf{M}(I-\mathbf{U})\left(I+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1}\mathbf{G}^{-1}\mathbf{Y}_1]I_{\mathbf{Y}_1>0}.
 \end{aligned}$$

It is obvious that

$$E\left[\left(I+\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\right)^{-1}\mathbf{M}(I-\mathbf{U})\mathbf{G}^{-1}\right]=\left(\frac{\Sigma_1}{n_1}\right)^{-1}\left(\frac{\Sigma_2}{n_2}\right)\left(\frac{\Sigma_1}{n_1}\right)^{-1}E\left[\frac{\mathbf{S}_1}{n_1}\right]=\left(\frac{\Sigma_1}{n_1}\right)^{-1}\frac{\Sigma_2}{n_2}.$$

So we have

$$\begin{aligned}
 & E[\mathbf{Y}'_2\left(I+\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\right)^{-1}\mathbf{M}(I-\mathbf{U})\mathbf{G}^{-1}\mathbf{Y}_1 \\
 & +\mathbf{Y}'_1\mathbf{G}^{-1}(I-\mathbf{U})'\mathbf{M}\mathbf{Y}'_2\left(I+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1}\mathbf{Y}_2]I_{\mathbf{Y}_1>0} \\
 = & 2E\left(\mathbf{Y}'_2\left(\frac{\Sigma_1}{n_1}\right)^{-1}\frac{\Sigma_2}{n_2}\right)E(\mathbf{Y}_1I_{\mathbf{Y}_1>0})=2(\mu_2-\mu_1)'E(\mathbf{Y}_1I_{\mathbf{Y}_1>0})>0.
 \end{aligned}$$

Due to  $n_2\Sigma_1 \geq n_1\Sigma_2$ , we have

$$\begin{aligned} & [\mathbf{Y}'_1\mathbf{G}^{-1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\left(I+\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\right)^{-1} \\ & \times (I-\mathbf{U})'\mathbf{M}(I-\mathbf{U})\left(I+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1}\frac{1}{n_1} \\ & \times \Sigma_1\left(\frac{\Sigma_2}{n_2}\right)^{-1}\mathbf{G}^{-1}\mathbf{Y}_1-\mathbf{Y}'_1\mathbf{G}^{-1}\left(I+\left(\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\right)^{-1}(I-\mathbf{U})'\mathbf{M}(I-\mathbf{U}) \\ & \times \left(I+\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1}\mathbf{G}^{-1}\mathbf{Y}_1]I_{\mathbf{Y}_1>0} \geq 0. \end{aligned}$$

Then

$$E\left[\mathbf{Y}'_1\mathbf{G}^{-1}\left(\frac{\Sigma_1}{n_1}+\frac{\Sigma_2}{n_2}\right)^{-1}\frac{\Sigma_1}{n_1}\mathbf{U}'\mathbf{M}(I-\mathbf{U})\mathbf{G}^{-1}\mathbf{Y}_1+\mathbf{Y}_1\mathbf{G}^{-1}(I-\mathbf{U})\mathbf{M}\mathbf{U}\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_1}{n_1}+\frac{\Sigma_2}{n_2}\right)^{-1}\mathbf{G}^{-1}\mathbf{Y}_1\right] \geq 0.$$

From the obtained results, we have

$$R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1^*, \mathbf{M}) > 0.$$

**Proof of theorem 4.**

Put  $\mathbf{M} = \left(\frac{\Sigma_2}{n_2}\right)^{-1}\left(\frac{\Sigma_1}{n_1}+\frac{\Sigma_2}{n_2}\right)^{-1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}$ , then the estimator  $\hat{\mu}_2$  given in (3.5) can be express as

$$\hat{\mu}_2 = \bar{\mathbf{X}}_2(1 - I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}) + (\mathbf{U}\bar{\mathbf{X}}_1 + (I - \mathbf{U})\bar{\mathbf{X}}_2)I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}.$$

The risk difference of  $\bar{\mathbf{X}}_2$  and  $\hat{\mu}_2$  is as

$$\begin{aligned} R(\mu_2, \bar{\mathbf{X}}_2, \mathbf{M}) - R(\mu_2, \hat{\mu}_2, \mathbf{M}) &= E[(\bar{\mathbf{X}}_2 - \mu_2)'\mathbf{M}(\bar{\mathbf{X}}_2 - \mu_2) \\ & - (\hat{\mu}_2 - \mu_2)'\mathbf{M}(\hat{\mu}_2 - \mu_2)]I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2} \\ &= E[(\bar{\mathbf{X}}_2 - \mu_2)'\mathbf{M}(\bar{\mathbf{X}}_2 - \mu_2) \\ & - (\mathbf{U}\bar{\mathbf{X}}_2 + (I - \mathbf{U})\bar{\mathbf{X}}_2 - \mu_1)'\mathbf{M}(\mathbf{U}\bar{\mathbf{X}}_1 + (I - \mathbf{U})\bar{\mathbf{X}}_2 - \mu_1)]I_{\mathbf{G}\bar{\mathbf{X}}_1 > \mathbf{G}\bar{\mathbf{X}}_2}. \end{aligned} \tag{4.7}$$

Consider the transformations  $\mathbf{Z}_1 = \bar{\mathbf{X}}_1 - \mu_2$  and  $\mathbf{Z}_2 = \bar{\mathbf{X}}_2 - \mu_2$ . Then it is clearly that the vectors  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  are mutually independently distributed as  $N_p\left(\mu_1 - \mu_2, \frac{\Sigma_1}{n_1}\right)$  and  $N_p\left(0, \frac{\Sigma_2}{n_2}\right)$  respectively. Then

$$\mathbf{Z}_1 = \left(I + \frac{\Sigma_2}{n_2}\left(\frac{\Sigma_1}{n_1}\right)^{-1}\right)^{-1}(\mathbf{G}^{-1}\mathbf{Y}_1 + \mathbf{Y}_2),$$

and

$$\mathbf{Z}_2 = \left(I + \frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1}\left(\frac{\Sigma_1}{n_1}\left(\frac{\Sigma_2}{n_2}\right)^{-1}\mathbf{Y}_2 - \mathbf{G}^{-1}\mathbf{Y}_1\right),$$

where  $\mathbf{Y}_1 = \mathbf{G}(\mathbf{Z}_1 - \mathbf{Z}_2)$  and  $\mathbf{Y}_2 = \mathbf{Z}_2 + \frac{\Sigma_2}{n_2}\left(\frac{\Sigma_1}{n_1}\right)^{-1}\mathbf{Z}_1$ , then  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$  are mutually independent and we have

$$\mathbf{Y}_2 \sim N_p\left(\frac{\Sigma_2}{n_2}\left(\frac{\Sigma_1}{n_1}\right)^{-1}(\mu_1 - \mu_2), \frac{\Sigma_2}{n_2} + \frac{\Sigma_2}{n_2}\left(\frac{\Sigma_1}{n_1}\right)^{-1}\frac{\Sigma_2}{n_2}\right).$$

From (4.7), we have

$$\begin{aligned}
 R(\mu_2, \bar{\mathbf{X}}_2, \mathbf{M}) - R(\mu_2, \hat{\mu}_2^*, \mathbf{M}) &= E[\mathbf{Z}'_2(I - \mathbf{U})\mathbf{M}\mathbf{U}(\mathbf{Z}_2 - \mathbf{Z}_1) \\
 &\quad + \mathbf{Z}_2 - \mathbf{Z}_1)' \mathbf{U}' \mathbf{M} (I - \mathbf{U}) \mathbf{Z}_2] I_{\mathbf{G}\mathbf{Z}_1 > \mathbf{G}\mathbf{Z}_2} \\
 &\quad + E[-\mathbf{Y}'_2 \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1} \mathbf{Y}_1 - \mathbf{Y}'_1 \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{Y}_2] \\
 &\quad + \mathbf{Y}_1 \mathbf{G}^{-1} \left(I + \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1}\right)^{-1} \mathbf{U}' \mathbf{M} \mathbf{U} \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \mathbf{G}^{-1} \mathbf{Y}_1 \\
 &\quad - \mathbf{Y}_1 \mathbf{G}^{-1} \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \mathbf{U}' \mathbf{M} \mathbf{U} \\
 &\quad \times \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \mathbf{G}^{-1} \mathbf{Y}_1 + \mathbf{Y}_1 \mathbf{G}^{-1} \mathbf{U}' \mathbf{M} (I - \mathbf{U}) \\
 &\quad \times \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \mathbf{G}^{-1} \mathbf{Y}_1 \\
 &\quad + \mathbf{Y}_1 \mathbf{G}^{-1} \left(I + \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1}\right)^{-1} (I - \mathbf{U})' \mathbf{M} \mathbf{U} \mathbf{G}^{-1} \mathbf{Y}_1] I_{\mathbf{Y}_1 > 0}.
 \end{aligned}$$

It is obvious that

$$E\left[-\mathbf{Y}'_2 \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1} \mathbf{Y}_1 - \mathbf{Y}'_1 \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1} \mathbf{Y}_2\right] I_{\mathbf{Y}_1 > 0} > 0.$$

Due to  $n_2 \Sigma_1 \leq n_1 \Sigma_2$ , we have

$$\begin{aligned}
 &E[\mathbf{Y}'_1 \mathbf{G}^{-1} \left(I + \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1}\right)^{-1} \mathbf{U}' \mathbf{M} \mathbf{U} \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \mathbf{G}^{-1} \mathbf{Y}_1 \\
 &- \mathbf{Y}'_1 \mathbf{G}^{-1} \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \mathbf{U}' \mathbf{M} \mathbf{U} \\
 &\times \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \mathbf{G}^{-1} \mathbf{Y}_1] I_{\mathbf{Y}_1 > 0} \geq 0,
 \end{aligned}$$

and

$$\begin{aligned}
 &E[\mathbf{Y}_1 \mathbf{G}^{-1} \mathbf{U}' \mathbf{M} (I - \mathbf{U}) \left(I + \frac{\Sigma_1}{n_1} \left(\frac{\Sigma_2}{n_2}\right)^{-1}\right)^{-1} \mathbf{G}^{-1} \mathbf{Y}_1 \\
 &+ \mathbf{Y}_1 \mathbf{G}^{-1} \left(I + \left(\frac{\Sigma_2}{n_2}\right)^{-1} \frac{\Sigma_1}{n_1}\right)^{-1} (I - \mathbf{U})' \mathbf{M} \mathbf{U} \mathbf{G}^{-1} \mathbf{Y}_1] I_{\mathbf{Y}_1 > 0} \geq 0.
 \end{aligned}$$

Therefore

$$R(\mu_2, \bar{\mathbf{X}}_2, \mathbf{M}) - R(\mu_2, \hat{\mu}_2^*, \mathbf{M}) > 0.$$

This completes the proof.

### 5. Application of results

In this section, to illustrate the results the following numerical examples are presented.

**Example 2.** Consider two bivariate normal distributions, i.e. the case  $k = 2$ , when they are subject to the order restriction  $\mu_1 < \mu_2$ . Three different cases are considered here. We simulate the

Table 2. Simulation from two bivariate normal distributions: the values of risks difference  $\hat{\mu}_1^*$  and  $\hat{\mu}_2^*$

	Sample sizes	$N_2(\mu_{1r}, \Sigma_{1r})$	$N_2(\mu_{2r}, \Sigma_{2r})$	$RD_{\bar{X}_1, \hat{\mu}_1^*}$	$RD_{\bar{X}_2, \hat{\mu}_2^*}$
Case 1 ( $r = a$ )	$n_1 = 20$	$\mu_{1a} = (1, 2)$	$\mu_{2a} = (2, 2)$	1.043	2.560
	$n_2 = 20$	$\Sigma_{1a} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$	$\Sigma_{2a} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$		
Case 2 ( $r = b$ )	$n_1 = 20$	$\mu_{1b} = (3, 1)$	$\mu_{2b} = (5, 2)$	0.841	1.754
	$n_2 = 15$	$\Sigma_{1b} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\Sigma_{2b} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$		
Case 3 ( $r = c$ )	$n_1 = 20$	$\mu_{1c} = (2, 5)$	$\mu_{2c} = (4, 5)$	0.963	0.802
	$n_1 = 10$	$\Sigma_{1c} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$	$\Sigma_{2c} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$		

values of random vectors  $\mathbf{X}_{11}, \mathbf{X}_{12}, \dots, \mathbf{X}_{1n_1}$ , from the bivariate normal distributions  $N_2(\mu_{1r}, \Sigma_{1r})$  with mean vectors  $\mu_{1r}$ ,  $r = a, b, c$ , and known nonsingular covariance matrices  $\Sigma_{1r}$  respectively, as  $N_2\left(\mu_{1a} = (1, 2), \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\right)$ ,  $N_2\left(\mu_{1b} = (3, 1), \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$  and  $N_2\left(\mu_{1c} = (2, 5), \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\right)$ . Also the values of random vectors  $\mathbf{X}_{21}, \mathbf{X}_{22}, \dots, \mathbf{X}_{2n_2}$ , from the bivariate normal distributions  $N_2(\mu_{2r}, \Sigma_{2r})$  with mean vectors  $\mu_{2r}$ ,  $r = a, b, c$ , and known nonsingular covariance matrices  $\Sigma_{2r}$  respectively, as  $N_2\left(\mu_{2a} = (2, 2), \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$ ,  $N_2\left(\mu_{2b} = (5, 2), \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\right)$  and  $N_2\left(\mu_{2c} = (4, 5), \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\right)$ .

It is completely clear that  $\mu_{1a} < \mu_{2a}$ ,  $\mu_{1b} < \mu_{2b}$  and  $\mu_{1c} < \mu_{2c}$ . In each simulation, the process of computation is repeated 10000 times to get an estimate of sample mean vectors  $\bar{\mathbf{X}}_1$  and  $\bar{\mathbf{X}}_2$ , isotonic estimators of mean vectors, i.e.  $\hat{\mu}_1^*$  and  $\hat{\mu}_2^*$  by (3.1) and (3.2) respectively, and the risk difference  $RD_{\bar{X}_1, \hat{\mu}_1^*} = R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1^*, \mathbf{M})$  and  $RD_{\bar{X}_2, \hat{\mu}_2^*} = R(\mu_2, \bar{\mathbf{X}}_2, \mathbf{M}) - R(\mu_2, \hat{\mu}_2^*, \mathbf{M})$ , where  $\mathbf{M} = \left(\frac{\Sigma_1}{n_1}\right)^{-1} \left(\frac{\Sigma_1}{n_1} + \frac{\Sigma_2}{n_2}\right)^2 \left(\frac{\Sigma_1}{n_1}\right)^{-1}$ . For different values of sample sizes and  $r = a, b, c$  the results are given in Table 2. Also in this simulation, the values of  $\hat{\mu}_1$  and  $\hat{\mu}_2$  by (3.4) and (3.5) respectively, and the risk difference  $RD_{\bar{X}_1, \hat{\mu}_1} = R(\mu_1, \bar{\mathbf{X}}_1, \mathbf{M}) - R(\mu_1, \hat{\mu}_1, \mathbf{M})$  and  $RD_{\bar{X}_2, \hat{\mu}_2} = R(\mu_2, \bar{\mathbf{X}}_2, \mathbf{M}) - R(\mu_2, \hat{\mu}_2, \mathbf{M})$  are computed and the results are given in Table 3. From the table 2, it is clear that the isotonic regression estimator  $\hat{\mu}_1^*$  uniformly has the smaller risk than the unrestricted maximum likelihood estimator,  $\bar{\mathbf{X}}_1$ , in three cases  $RD_{\bar{X}_1, \hat{\mu}_1^*} > 0$ . Also the isotonic regression estimator  $\hat{\mu}_2^*$  uniformly has the smaller risk than the unrestricted maximum likelihood estimator,  $\bar{\mathbf{X}}_2$ , in three cases  $RD_{\bar{X}_2, \hat{\mu}_2^*} > 0$ . In table 3, case 2, estimator  $\hat{\mu}_1$  uniformly has not the smaller risk than the unrestricted maximum likelihood estimator  $\bar{\mathbf{X}}_1$ ,  $RD_{\bar{X}_1, \hat{\mu}_1} = -1.716$  since in this case  $n_2\Sigma_1 < n_1\Sigma_2$ . But in this case,  $\hat{\mu}_2$  uniformly has the smaller risk than the unrestricted maximum likelihood estimator  $\bar{\mathbf{X}}_2$ , since  $\bar{\mathbf{X}}_2$ ,  $RD_{\bar{X}_2, \hat{\mu}_2} = 3.915$  and  $n_2\Sigma_1 < n_1\Sigma_2$ . In table 3, case 3, estimator  $\hat{\mu}_1$  uniformly has not the smaller risk than the unrestricted maximum likelihood estimator  $\bar{\mathbf{X}}_1$ , but  $\hat{\mu}_2$  uniformly has the smaller risk than the unrestricted maximum likelihood estimator  $\bar{\mathbf{X}}_2$ .

All the results in the previous sections are establish in the univariate case. To show this verity the following numerical example is presented.

**Example 3.** We consider the estimation of linear functions  $c_1\mu_1 + c_2\mu_2$ , where  $c_1$  and  $c_2$  are arbitrary constants. We show that the plug-in estimator  $c_1\hat{\mu}_1 + c_2\hat{\mu}_2$  uniformly improves upon the unrestricted

Table 3. Simulation from two bivariate normal distributions: the values of risks difference  $\hat{\mu}_1$  and  $\hat{\mu}_2$

	Sample sizes	$N_2(\mu_{1r}, \Sigma_{1r})$	$N_2(\mu_{2r}, \Sigma_{2r})$	$RD_{\bar{X}_1, \hat{\mu}_1}$	$RD_{\bar{X}_2, \hat{\mu}_2}$
Case 1 ( $r = a$ )	$n_1 = 20$	$\mu_{1a} = (1, 2)$	$\mu_{2a} = (2, 2)$	3.162	2.560
	$n_2 = 20$	$\Sigma_{1a} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$	$\Sigma_{2a} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$		
Case 2 ( $r = b$ )	$n_1 = 20$	$\mu_{1b} = (3, 1)$	$\mu_{2b} = (5, 2)$	-1.716	3.915
	$n_2 = 15$	$\Sigma_{1b} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\Sigma_{2b} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$		
Case 3 ( $r = c$ )	$n_1 = 20$	$\mu_{1c} = (2, 5)$	$\mu_{2c} = (4, 5)$	-2.550	3.244
	$n_1 = 10$	$\Sigma_{1c} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$	$\Sigma_{2c} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$		

maximum likelihood estimator  $c_1\bar{X}_1 + c_2\bar{X}_2$  if and only if for all  $\sigma_i^2$ 's the risk difference

$$R(c_1\mu_1 + c_2\mu_2, c_1\bar{X}_1 + c_2\bar{X}_2) - R(c_1\mu_1 + c_2\mu_2, c_1\hat{\mu}_1 + c_2\hat{\mu}_2),$$

is nonnegative when  $\mu_1 = \mu_2$ , which is the most critical case for uniform improvement. In this case, the risk difference is nonnegative if and only if the risk of  $(c_1 + c_2)\hat{\mu}_{GD}$  is not larger than that of  $c_1\bar{X}_1 + c_2\bar{X}_2$  to estimate  $(c_1 + c_2)\mu_1$ . From this, we also show that when  $c_1c_2 > 0$ ,  $c_1\hat{\mu}_1 + c_2\hat{\mu}_2$  does not uniformly improve upon  $c_1\bar{X}_1 + c_2\bar{X}_2$  and give some sufficient conditions on  $c_1$ ,  $c_2$ ,  $n_1$  and  $n_2$  for  $c_1\hat{\mu}_1 + c_2\hat{\mu}_2$  to improve upon  $c_1\bar{X}_1 + c_2\bar{X}_2$  uniformly for the case when  $c_1c_2 < 0$ . However, we show that if  $\sigma_i^2$ 's are known, the restricted maximum likelihood estimator  $c_1\hat{\mu}_1^* + c_2\hat{\mu}_2^*$  always uniformly improves upon  $c_1\bar{X}_1 + c_2\bar{X}_2$ . These results again make a clear contrast between the two cases: variances known or unknown. In Fig. 1 we show the risk difference  $R(\mu_1, \bar{X}_1) - R(\mu_1, \hat{\mu}_1)$  as a function of  $\mu = \mu_1 - \mu_2$  with  $\sigma_i^2$ 's fixed for the case when  $n_1 = n_2 = 11$  (Fig. 1 (a)-(c)), which implies  $\hat{\mu}_1$  uniformly improves upon  $\bar{X}_1$  by Theorem 1, and also for the case when  $n_1 = n_2 = 10$  (Fig. 1 (d)-(h)), which implies  $\hat{\mu}_1$  does not uniformly improve upon  $\bar{X}_1$  by Theorem 1. If  $\frac{\sigma_1^2}{\sigma_2^2}$  is small (Fig. 1 (a), (b), (d) and (e)), the risk difference is monotone decreasing in  $\mu$ , and if  $\frac{\sigma_1^2}{\sigma_2^2}$  is large (Fig. 1 (c) and (f)(h)), the risk difference is decreasing in  $\mu$  after increasing monotonically. When  $n_1 = n_2 = 10$ ,  $\frac{\sigma_1^2}{\sigma_2^2}$  is sufficiently large (Fig. 1 (h)), the risk difference is negative for  $\mu = 0$ , since the most critical case for the GraybillDeal estimator  $\hat{\mu}_{GD}$  to have smaller risk than  $\bar{X}_1$  is the case when  $\frac{\sigma_1^2}{\sigma_2^2} \rightarrow \infty$  by Graybill and Deal (1959).

Fig (a)= $n_1 = n_2 = 11, \sigma_1^2 = 1, \sigma_2^2 = 0.1^2$ . Fig (b)= $n_1 = n_2 = 11, \sigma_1^2 = 1, \sigma_2^2 = 1$ . Fig(c)= $n_1 = n_2 = 11, \sigma_1^2 = 1, \sigma_2^2 = 10^2$ . Fig (d)= $n_1 = n_2 = 10, \sigma_1^2 = 1, \sigma_2^2 = 0.1^2$ . Fig (e)= $n_1 = n_2 = 10, \sigma_1^2 = 1, \sigma_2^2 = 1$ . Fig (f)= $n_1 = n_2 = 10, \sigma_1^2 = 1, \sigma_2^2 = 2^2$ . Fig (g)= $n_1 = n_2 = 10, \sigma_1^2 = 1, \sigma_2^2 = 5^2$ . Fig (h)= $n_1 = n_2 = 10, \sigma_1^2 = 1, \sigma_2^2 = 10^2$ .

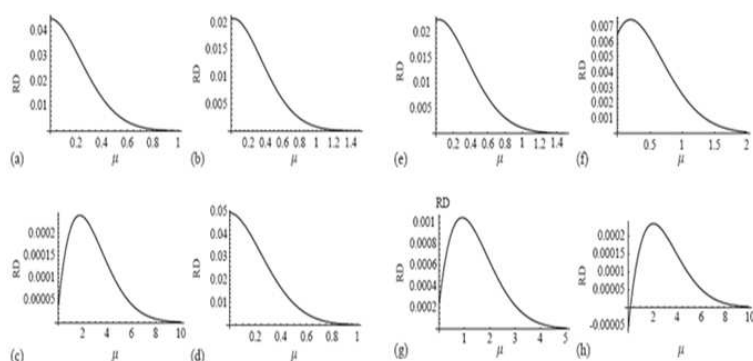


Fig. 1. The risk difference  $RD=R(\mu_1 - \bar{X}_1) - R(\mu_1 - \hat{\mu}_1)$

## 6. Concluding remarks

In this study, a comparison between the estimators of the monotonic mean vectors and unrestricted maximum likelihood estimators in two dimensional normal distributions was done under two different cases for covariance matrices with unequal sample sizes. Under the squared error loss function of the estimators of mean vectors when the covariance matrices were known, the obtained estimators of ordered mean vectors had the smaller risk than the estimators of unrestricted maximum likelihood uniformly, but when the covariance matrices were unknown and unequal, the estimators had the smaller risk than the estimators of unrestricted maximum likelihood only over some special sets which are defined on the covariance matrices. Also, two numerical examples presented to illustrate the results. In example 2, the data simulated from different bivariate normal distributions. We showed that the isotonic regression estimators uniformly have the smaller risk than the unrestricted maximum likelihood estimator since the risk differences are positive. In this example, also the isotonic regression estimators uniformly have the smaller risk than the unrestricted maximum likelihood estimator over the sets  $A_1$  and  $A_2$ . In example 3, the plug-in estimator  $c_1\hat{\mu}_1 + c_2\hat{\mu}_2$  uniformly improved upon the unrestricted maximum likelihood estimator  $c_1\bar{X}_1 + c_2\bar{X}_2$  if and only if for all  $\sigma_i^2$ 's the risk difference is nonnegative when  $\mu_1 = \mu_2$ , which was the most critical case for uniform improvement.

## Acknowledgements

The author is thankful to the Editor-in-Chief and two anonymous referees for their important and valuable comments and suggestions that led to considerable improvements of this paper. The author thank to Dr. H. Haghbin, Assistance Professor of Statistics, Department of Statistics, Persian Gulf University and also the family of Majid Shahidzadeh for encouragement, discussions, and financial support. I lovingly dedicate this paper to my wife Soheila Shahidzadeh who supported me each step of the way.



## References

- [1] R. E. Barlow, D. J. Bartholomew, J. M. Bremner, and H. D. Brunk. *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*. New York: John Wiley, 1972.
- [2] D. J. Bartholomew. A test of homogeneity for ordered alternatives, *Biometrika*, **46**, 36-48, 1959.
- [3] A. Bazyari. On the computation of some properties of testing homogeneity of multivariate normal mean vectors against an order restriction. *METRON, International Journal of Statistics*, vol. **LXX**, n. **1**, 71-88, 2012.
- [4] A. Bazyari and R. Chinipardaz. A test for order restriction of several multivariate normal mean vectors against all alternatives when the covariance matrices are unknown but common, *Journal of Statistical Theory and Applications*, **11(1)**, 23-45, 2012.
- [5] A. Bazyari and R. Chinipardaz. Upper Bound for p-Value of the Test of Multivariate Normal Ordered Mean Vectors Against all Alternatives, *Communications in Statistics, Theory and Methods*, **42**, 1748-1758, 2013.
- [6] A. Bazyari and F. Pesarin. Parametric and permutation testing for multivariate monotonic alternatives, *Statistics and Computing*, **23**, 639-652, 2013.
- [7] W. Chiou and A. Cohen. On estimating a common multivariate normal mean vector, *Annals of the Institute of Statistical Mathematics*, **37**, 499-506, 1985.
- [8] E. J. Dietz. Multivariate generalizations of Jonckheere's test for ordered alternatives, *Communications in Statistics, Theory and Methods*, **18**, 3763-3783, 1989.
- [9] W. T. P. S. Fernando, D. D. S. Kulatunga. On the computation and some applications of multivariate isotonic regression, *Computational Statistics and Data Analysis*, **52**, 702-712, 2007.
- [10] F. A. Graybill and R. B. Deal. Combining unbiased estimators, *Biometrics*, **15**, 543-550, 1959.
- [11] J. Hansohm and X. Hu. A convergent algorithm for a generalized multivariate isotonic regression problem. *Statistical Papers*, **53**, 107-115, 2012.
- [12] S. M. Jordan and K. Krishnamoorthy. Confidence intervals for the common mean of several normal populations, *Canadian Journal of Statistics*, **23**, 283-297, 1995.
- [13] R. E. Kelly. Stochastic reduction of loss in estimating normal means by isotonic regression, *Annals of Statistics*, **17**, 937-940, 1989.
- [14] K. Krishnamoorthy and Y. Lu. Inferences on the common mean of several normal populations based on the generalized variable method, *Biometrics*, **59**, 237-247, 2003.
- [15] C. I. C. Lee. The quadratic loss of isotonic regression under normality, *Annals of Statistics*, **9**, 686-88, 1981.
- [16] S. H. Lin and J. C. Lee. Generalized inferences on the common mean of several normal populations, *Journal of Statistical Planning and Inference*, **134**, 568-582, 2005.
- [17] S. H. J. C. Lee and R. S. Wang. Generalized inferences on the common mean vector of several multivariate normal populations, *Journal of Statistical Planning and Inference*, **137**, 2240-2249, 2007.
- [18] W. L. Loh. Estimating the common mean of two multivariate normal distributions, *Annals of Statistics*, **19**, 283-296, 1991.
- [19] T. Robertson, F. T. Wright and R. L. Dykstra. *Order Restricted Statistical Inference*. New York: John Wiley, 1988.
- [20] S. Sasabuchi, M. Inutsuka and D. D. S. Kulatunga. A multivariate version of isotonic regression, *Biometrik*, **70**, 465-472, 1983.
- [21] S. Sasabuchi, M. Inutsuka and D. D. S. Kulatunga. An algorithm for computing multivariate isotonic regression, *Hiroshima Mathematical Journal*, **22**, 551-560, 1992.
- [22] N. Z. Shi and H. Jiang. Maximum likelihood estimation of isotonic normal means with unknown variances, *Journal of Multivariate Analysis*, **64**, 183-195, 1998.
- [23] H. Shimodaira. Approximately Unbiased one-sided Tests of the Maximum of Normal Means Using Iterated Bootstrap Corrections. Tech Rep no 2000-07, Department of Statistics, Stanford University, Stanford, 2000.
- [24] M. J. Silvapulle and P. K. Sen. *Constrained Statistical Inference: Inequality, Order, and Shape Restrictions*, New York: John Wiley, 2005.

*Abouzar Bazyari*

- [25] L. Zhou and T. Mathew. Combining independent tests in multivariate linear models, *Journal of Multivariate Analysis*, **51**, 265-276, 1994.