Austrian Journal of Statistics
February 2017, Volume 46, 3–17.
http://www.ajs.or.at/
doi:10.17713/ajs.v46i2.446

Estimation of Order Restricted Normal Means when the Variances Are Unknown and Unequal

Najmeh Pedram
Persian Gulf University

Abouzar Bazyari
Persian Gulf University

Abstract

In the present paper, two normal distributions with parameters μ_i and σ_i^2 where there is an order restriction on the means when the variances are unknown and unequal are considered. Under the squared error loss function, a necessary and sufficient condition for the plug-in estimators to improve upon the unrestricted maximum likelihood estimators uniformly is given. Also under the modified Pitman nearness criterion; a class of estimators is considered that reduce to the estimators of a common mean when the unbiased estimators violate the order restriction. It is shown that the most critical case for uniform improvement with regard to the unbiased estimators is the one when two means are equal. To illustrate the results, two numerical examples are presented.

Keywords: maximum likelihood estimator, order restriction, Pitman nearness, squared error loss function.

1. Introduction

Let X_{ij} be the j th observation of the i th population and be mutually independently distributed as $N(\mu_i, \sigma_i^2)$, $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, n_i$, where the order restriction on the unknown parameters μ_i, $i = 1, 2, \cdots, k$ is defined as

$$\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k. \tag{1}$$

We consider the following squared error loss function of the estimators of μ_i, $i = 1, 2, \cdots, k$,

$$L(\mu_i, \hat{\mu}_i) = (\hat{\mu}_i - \mu_i)^2. \tag{2}$$

Then the risk is given by

$$R(\mu_i, \hat{\mu}_i) = E[L(\mu_i, \hat{\mu}_i)]. \tag{3}$$

The estimator $\hat{\mu}_i^*$ uniformly improves upon the estimator $\hat{\mu}_i^{**}$, $i = 1, 2, \cdots, k$, under the squared error loss function (2) if and only if

$$R(\mu_i, \hat{\mu}_i^*) \leq R(\mu_i, \hat{\mu}_i^{**}),$$

for all $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$. Note that $\bar{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i$ is the unrestricted maximum likelihood estimator of μ_i and is distributed as $N(\mu_i, \sigma_i^2/n_i)$.

Later, many authors, including Brown and Cohen (1974), Khatri and Shah (1974) and Bhattacharya *et al.* (1980) have given a class of improved estimators of the form

$$\hat{\mu}(\gamma) = \gamma \bar{X}_1 + (1-\gamma)\bar{X}_2,$$

where γ is a function of s_1^2 and s_2^2 .
Under the order restriction (1), the maximum likelihood estimator of μ_i is given by

$$\underset{t\geq i}{min}\,\underset{s\leq i}{max} \frac{\sum_{j=s}^{t} n_j \bar{X}_j/\sigma_j^2}{\sum_{j=s}^{t} n_j/\sigma_j^2}. \tag{4}$$

A possible alternative criterion to evaluate the goodness of estimators, mean squared error (MSE), is Pitmam nearness.
For comparing two estimators $T_i, (i = 1, 2)$ of a single parameter θ, Pitman (1937) proposed the following criterion: T_1 is said to be closer (better) than T_2 if

$$PN_\theta(T_1, T_2) = P\{|T_2 - \theta| > |T_1 - \theta|\} > \frac{1}{2}, \tag{5}$$

for all θ. The probability $PN_\theta(T_1, T_2)$ in (5) is usually called the Pitman nearness of T_1 relative to T_2.
Lee (1981) showed that the estimator (4) uniformly improves upon $\bar{X}_i$. Rao (1980) discussed the similarities and differences of MSE and PMN. Kelly (1989) strengthened Lee (1981)'s result and showed that (4) universally dominates $\bar{X}_i$.
Nayak (1990) defined modified Pitman nearness of an estimator T_1 of θ relative to the other estimator T_2 by

$$MPN_\theta(T_1, T_2) = P\{|T_1 - \theta| < |T_2 - \theta| | T_1 \neq T_2\}. \tag{6}$$

If $MPN_\theta(T_1, T_2) \geq 1/2$ for any parameter value, then T_1 is said to be closer to θ than T_2. Gupta and Singh (1992) have applied modified Pitman nearness to the estimation of ordered means of two normal population with common variance and have shown that MLE is closer than the unbiased estimator.
Hwang and Peddada (1994) showed that under arbitary order restriction on μ_i's, (4) universally dominates $\bar{X}_i$ to estimate μ_i if μ_i is a node and proposed estimation procedure also for nonnodal means. (μ_i is said to be a node if, for any j, it is known that either $\mu_j \leq \mu_i$ or $\mu_i \leq \mu_j$).

In this paper, we consider the estimation of two normal means when they are subject to the order restriction

$$\mu_1 \leq \mu_2, \tag{7}$$

and $\sigma_i^2, i = 1, 2$ are unknown and possibly unequal. If σ_i^2's are known, from (7) the restricted maximum likelihood estimators of μ_i's are given by

$$\hat{\mu}_1^* = min\left(\bar{X}_1, \frac{\frac{n_1}{\sigma_1^2}\bar{X}_1 + \frac{n_2}{\sigma_2^2}\bar{X}_2}{\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}}\right), \tag{8}$$

and

$$\hat{\mu}_2^* = max\left(\bar{X}_2, \frac{\frac{n_1}{\sigma_1^2}\bar{X}_1 + \frac{n_2}{\sigma_2^2}\bar{X}_2}{\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}}\right). \tag{9}$$

But, if we suppose that σ_i^2's are unknown, so we estimate σ_i^2 by $s_i^2 = \sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2/(n_i - 1)$ and replace σ_i^2 with s_i^2 in (8) and (9) and obtain the plug-in estimators as follows

$$\hat{\mu}_1 = min\left(\bar{X}_1, \frac{\frac{n_1}{s_1^2}\bar{X}_1 + \frac{n_2}{s_2^2}\bar{X}_2}{\frac{n_1}{s_1^2} + \frac{n_2}{s_2^2}}\right), \tag{10}$$

and

$$\hat{\mu}_2 = max\left(\bar{X}_2, \frac{\frac{n_1}{s_1^2}\bar{X}_1 + \frac{n_2}{s_2^2}\bar{X}_2}{\frac{n_1}{s_1^2} + \frac{n_2}{s_2^2}}\right). \quad (11)$$

? proposed another type of plug-in estimators $\tilde{\mu}_i'$ obtained by replacing s_i^2 with $\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2/n_i$ given in (10) and (11) and proposed results when $\mu_1 = \mu_2$. Chang and Shinozaki (2012) have considered a class of estimators of μ_i, $i = 1, 2$ of the form

$$\hat{\mu}_1(\gamma) = min\{\bar{X}_1, \gamma\bar{X}_1 + (1-\gamma)\bar{X}_2\}, \quad (12)$$

and

$$\hat{\mu}_2(\gamma) = max\{\bar{X}_2, \gamma\bar{X}_1 + (1-\gamma)\bar{X}_2\}. \quad (13)$$

Bazyari (2015) considered 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.
To illustrate the usefulness of order restriction we have taken the following examples.
Example 1. An experiment was conducted to evaluate the effect of exercise on the age at which a child starts to walk. Let Y denote the age (in months) at which a child starts to walk, the data on Y are given in Tabel 1. (The original experiment consisted of another treatment, however, here we consider only two treatments for simplicity.)

Table 1: The age at which a child first walks.

Treatment (i)			Age (in months)				n_i	$\bar{y}_i$	μ_i
1	9.00	9.50	9.75	10.00	13.00	9.50	6	10.125	μ_1
2	11.00	10.00	10.00	11.75	10.50	15.00	6	11.375	μ_2

The first treatment group received a special walking exercise for 12 minutes per day beginning at age 1 week and lasting 7 weeks. The second group received daily exercises but not the special walking exercises. For treatment i (i=1, 2), let μ_i be the mean age (in months) at which a child starts to walk. However, suppose that the researcher was prepared to assume that the walking exercises would not have negative effect of increasing the mean age at which a child starts to walk, and it was desired that this additional information be incorporated to improve on the statistical analysis. In this case, we have that $\mu_1 \leq \mu_2$.
Example 2. An experiment was done to evaluate the discrimination of men from women. Four psychological test scores, pictorial absurdities, paper form board, tool recognition and vocabulary were given to two different groups of 32 men and 32 women. The data on men and women are for 32 applicants for a professional position requiring 10 or more years of successful schooling (the completion of second-year high school in Ontario, up to a University degree). The 4 tests were each scored according to the number of questions answered successfully. The mean vectors of the two samples are

$$\bar{\boldsymbol{X}}_{\boldsymbol{1}} = (15.7, 15.91, 27.19, 22.75)', \quad \bar{\boldsymbol{X}}_{\boldsymbol{2}} = (12.34, 13.91, 16.66, 21.94)'.$$

Let $\boldsymbol{\mu_i} = (\mu_{i1}, \mu_{i2}, \mu_{i3}, \mu_{i4})'$ for $i = 1, 2$, denotes the mean variable for i^{th} group, where μ_{ij}, $j = 1, 2, 3, 4$, denotes the j^{th} element of mean vector $\boldsymbol{\mu_i}$. Suppose that the researcher is prepared to assume that the elements of mean vectors of two populations are subject to the order restriction

$$\mu_{21} < \mu_{11}, \quad \mu_{22} < \mu_{12}, \quad \mu_{23} < \mu_{13}, \quad \mu_{24} < \mu_{14}.$$

The rest of this paper is organized as follows. In section 2, we show that the plug-in estimator μ_i uniformly improves upon $\bar{X}_i$ if and only if for all σ_i^2's the risk difference $\bar{X}_i$ and $\hat{\mu}_i$ is nonnegative when $\mu_1 = \mu_2$. In section 3, with respect to modified Pitman nearness, we show that the estimator $\hat{\mu}_i(\gamma)$ improves upon $\bar{X}_i$ uniformly imporoves upon the $\bar{X}_i$ if and only if $MPN_{\mu_i}(\hat{\mu}_i(\gamma), \bar{X}_i) \geq \frac{1}{2}$ when $\mu_1 = \mu_2$, which is the most critical case for uniform improvement. Further, it is shown that $\hat{\mu}_i(\gamma)$ improves upon $\bar{X}_i$ if and only if $\hat{\mu}(\gamma)$ improves upon $\bar{X}_i$ for the same γ in estimating a common mean. To illustrate the results two numerical examples are presented in section 4. Concluding remarks are given in section 5.

2. Uniformly improved estimator of each of two ordered normal means

We show that the most critical case for $\hat{\mu}_i$ to improve upon $\bar{X}_i$ if and only uniformly is the one when $\mu_1 = \mu_2$.

Theorem 2.1. *The plug-in estimator $\hat{\mu}_1$ uniformly improves upon the unrestricted maximum likelihood estimator $\bar{X}_1$ if and only if for all σ_i^2's the risk of $\hat{\mu}_1$ is not larger than that of $\bar{X}_1$ when $\mu_1 = \mu_2$.*

Proof. Putting $\gamma = \frac{(\frac{n_1}{s_1^2})}{(\frac{n_1}{s_1^2} + \frac{n_2}{s_2^2})}$, $\hat{\mu}_1$ is expressed as

$$\hat{\mu}_1 = min(\bar{X}_1, \gamma\bar{X}_1 + (1-\gamma)\bar{X}_2), \tag{14}$$

and we calculate the risk difference of $\bar{X}_1$ and $\hat{\mu}_1$ as

$$\begin{aligned} &R(\mu_1, \bar{X}_1) - R(\mu_1, \hat{\mu}_1) \\ &= E[(\bar{X}_1 - \mu_1)^2 - \{\gamma(\bar{X}_1 - \mu_1) + (1-\gamma)(\bar{X}_2 - \mu_1)\}^2] I_{\bar{X}_1 \geq \bar{X}_2}, \end{aligned} \tag{15}$$

where I_d denotes the indicator function of the set satisfying the condition d. Making the transformations

$$Z_1 = \bar{X}_1 - \mu_1 \quad , \quad Z_2 = \bar{X}_2 - \mu_1, \tag{16}$$

Z_1 and Z_1 are mutually independently distributed as $N(0, \tau_1^2)$ and $N(\mu, \tau_2^2)$, respectively, where $\mu = \mu_2 - \mu_1 \geq 0, \tau_1^2 = \sigma_1^2/n_1$ and $\tau_2^2 = \sigma_2^2/n_2$. Noting that Z_1, Z_2 and γ are mutually independent, we have from (16)

$$\begin{aligned} R(\mu_1, \bar{X}_1) - R(\mu_1, \hat{\mu}_1) &= E[Z_1^2 - \{\gamma Z_1 + (1-\gamma)Z_2\}^2] I_{Z_1 \geq Z_2} \\ &= 2E[\gamma(1-\gamma)]E[(Z_1 - Z_2)Z_1 I_{Z_1 \geq Z_2}] \\ &+ E[(1-\gamma)^2]E[(Z_1^2 - Z_2^2)I_{Z_1 \geq Z_2}]. \end{aligned} \tag{17}$$

Making the further transformations

$$Y_1 = Z_1 - Z_2 \quad , \quad Y_2 = Z_1 + (\frac{\tau_1^2}{\tau_2^2})Z_2, \tag{18}$$

note that Y_1 and Y_2 are mutually independently distributed as $N(-\mu, \tau_1^2 + \tau_2^2)$ and $N\left((\frac{\tau_1^2}{\tau_2^2})\mu, \tau_1^2 + (\frac{\tau_1^4}{\tau_2^2})\right)$, respectively, and

$$Z_1 = \frac{Y_1(\frac{\tau_1^2}{\tau_2^2}) + Y_2}{1 + \frac{\tau_1^2}{\tau_2^2}} \quad , \quad Z_2 = \frac{Y_2 - Y_1}{1 + (\frac{\tau_1^2}{\tau_2^2})}.$$

Then, we have

$$
\begin{aligned}
&E[(Z_1 - Z_2)Z_1 I_{Z_1 \geq Z_2}] \\
&= E\left[\frac{\tau_2^2 Y_1 \left(Y_1(\frac{\tau_1^2}{\tau_2^2}) + Y_2\right)}{\tau_1^2 + \tau_2^2} I_{Y_1 \geq 0}\right] \\
&= E\left[\frac{\tau_2^2 \left(Y_1^2(\frac{\tau_1^2}{\tau_2^2}) + Y_1 E[E[Y_2]]\right)}{\tau_1^2 + \tau_2^2} I_{Y_1 \geq 0}\right] \\
&= E\left[\frac{\tau_1^2 (Y_1^2 + \mu Y_1)}{\tau_1^2 + \tau_2^2} I_{Y_1 \geq 0}\right] \\
&\geq \frac{\tau_1^2}{\tau_1^2 + \tau_2^2} E[Y_1^2 I_{Y_1 \geq 0}],
\end{aligned} \tag{19}
$$

and

$$
\begin{aligned}
&E[(Z_1^2 - Z_2^2) I_{Z_1 \geq Z_2}] \\
&= E\left[\left(\frac{\tau_2^2}{\tau_1^2 + \tau_2^2}\right)^2 \left[Y_1^2 \left(\frac{(\tau_1^2)^2 - (\tau_2^2)^2}{(\tau_2^2)^2}\right) + 2Y_1 Y_2 \left(\frac{\tau_1^2 + \tau_2^2}{\tau_2^2}\right)\right] I_{Y_1 \geq 0}\right] \\
&= E\left[\left(\frac{\tau_2^2}{\tau_1^2 + \tau_2^2}\right)^2 \left[Y_1^2 \left(\frac{(\tau_1^2)^2 - (\tau_2^2)^2}{(\tau_2^2)^2}\right) + 2Y_1 E[E[Y_2]] \left(\frac{\tau_1^2 + \tau_2^2}{\tau_2^2}\right)\right] I_{Y_1 \geq 0}\right] \\
&= E\left[\frac{(\tau_1^2 - \tau_2^2) Y_1^2 + 2\tau_1^2 \mu Y_1}{\tau_1^2 + \tau_2^2} I_{Y_1 \geq 0}\right] \\
&\geq \frac{\tau_1^2 - \tau_2^2}{\tau_1^2 + \tau_2^2} E[Y_1^2 I_{Y_1 \geq 0}],
\end{aligned} \tag{20}
$$

with equalities for $\mu = 0$ and strict inequalities for $\mu > 0$. Thus we have from (17), (19) and (20)

$$
\begin{aligned}
&R(\mu_1, \bar{X}_1) - R(\mu_1, \hat{\mu}_1) \\
&\geq \frac{E[Y_1^2 I_{Y_1 \geq 0}]}{\tau_1^2 + \tau_2^2} \{2\tau_1^2 E[\gamma(1-\gamma)] + (\tau_1^2 - \tau_2^2) E[(1-\gamma)^2]\} \\
&= \frac{E[Y_1^2 I_{Y_1 \geq 0}]}{\tau_1^2 + \tau_2^2} [\tau_1^2 - \{\tau_1^2 E[\gamma^2] + \tau_2^2 E[(1-\gamma)^2]\}] \\
&= \frac{E[Y_1^2 I_{Y_1 \geq 0}]}{E_{\mu_1 = \mu_2}[Y_1^2 I_{Y_1 \geq 0}]} \{R_{\mu_1 = \mu_2}(\mu_1, \bar{X}_1) - R_{\mu_1 = \mu_2}(\mu_1, \hat{\mu}_1)\},
\end{aligned} \tag{21}
$$

with equality for $\mu = 0$ and strict inequality for $\mu > 0$. Thus, we have shown that $\hat{\mu}_1$ unifrormly improves upon $\bar{X}_1$ if and only if for all σ_i^2's the risk difference is not positive when $\mu_1 = \mu_2$, which is the most critical case for uniform improvement. This completes the proof. □

Regarding the improved estimation of μ_2, we have a similar result as follows.

Corollary 2.2. *The plug-in estimator $\hat{\mu}_2$ uniformly improves upon the unrestricted maximum likelihood estimator $\bar{X}_2$ if and only if for all σ_i^2's the risk of $\hat{\mu}_2$ is not larger than that of $\bar{X}_2$ when $\mu_1 = \mu_2$.*

Proof. Since $\mu_1 \leq \mu_2$ can be written as $-\mu_2 \leq -\mu_1$, the result follows directly from theorem (2.1). □

3. Pitman dominates of new plug-in estimators

In this section, we consider estimators of μ_i of the form (12) and (13) and compare them with unbiased estimator $\bar{X}_i$. We first show that for the case when γ is a function of s_1^2 and s_2^2 the most critical case for $\hat{\mu}_i(\gamma)$ to be closer to μ_i than $\bar{X}_i$ is the one when $\mu_1 = \mu_2$. Further, it is shown that $\hat{\mu}_i(\gamma)$ improves upon $\bar{X}_i$ if and only if $\hat{\mu}(\gamma)$ dominates $\bar{X}_i$ in the estimation problem of a common mean.

Theorem 3.1. *Suppose that $0 \leq \gamma \leq 1$ is a function of s_1^2 and s_2^2. Then*

a)$MPN_{\mu_i}(\hat{\mu}_i(\gamma), \bar{X}_i) \geq \frac{1}{2}$ for all $\mu_1 \leq \mu_2$ and for all σ_1^2 and σ_2^2 if and only if for all σ_1^2 and σ_2^2, $MPN_{\mu_i}(\hat{\mu}_i(\gamma), \bar{X}_i) \geq \frac{1}{2}$ when $\mu_1 = \mu_2$.

b)$MPN_{\mu_i}(\hat{\mu}_i(\gamma), \bar{X}_i) \geq \frac{1}{2}$ for all $\mu_1 \leq \mu_2$ and for all σ_1^2 and σ_2^2 if and only if for all σ_1^2 and σ_2^2, $PN_{\mu}(\hat{\mu}_i(\gamma), \bar{X}_i) \geq 1/2$ to estimate μ when $\mu_1 = \mu_2 = \mu$.

Proof. We need only to give a proof for the case of μ_1.
a) Since $\hat{\mu}_1(\gamma) \neq \bar{X}_1$ if and only if $\bar{X}_2 < \bar{X}_1$ and $\gamma < 1$, we have

$$\begin{aligned}
&MPN_{\mu_1}(\hat{\mu}_1(\gamma), \bar{X}_1) \\
&= P\{|\hat{\mu}_1(\gamma) - \mu_1| < |\bar{X}_1 - \mu_1| | \hat{\mu}_1(\gamma) \neq \bar{X}_1\} \\
&= P\{|\gamma\bar{X}_1 + (1-\gamma)\bar{X}_2 - \mu_1| < |\bar{X}_1 - \mu_1| | \bar{X}_2 < \bar{X}_1, \gamma < 1\} \\
&= P\{-(\gamma\bar{X}_1 + (1-\gamma)\bar{X}_2 - \mu_1) < (\bar{X}_1 - \mu_1) | \bar{X}_2 < \bar{X}_1, \gamma < 1\} \\
&= P\{\bar{X}_1 - \mu_1 + \gamma\bar{X}_1 + \bar{X}_2 - \gamma\bar{X}_2 - \mu_2 > 0 | \bar{X}_2 < \bar{X}_1, \gamma < 1\} \\
&= P\{\bar{X}_1 - \mu_1 + \gamma\bar{X}_1 - \gamma\mu_1 + \bar{X}_2 - \mu_1 - \gamma\bar{X}_2 + \gamma\mu_1 > 0 | \bar{X}_2 - \mu_1 < \bar{X}_1 - \mu_1, \gamma < 1\} \\
&= P\{(1+\gamma)Z_1 + (1-\gamma)Z_2 > 0 | Z_2 < Z_1, \gamma < 1\}, \qquad (22)
\end{aligned}$$

where $Z_1 = \bar{X}_1 - \mu_1$ and $Z_2 = \bar{X}_2 - \mu_1$ are distributed as $N(0, \tau_1^2)$ and $N(\mu, \tau_2^2)$ respectively, $\mu = \mu_1 - \mu_2$ and $\tau_i^2 = \sigma_i^2/n_i$. Now, we consider the conditional probability

$$P\{0 < (1+\gamma)Z_1 + (1-\gamma)Z_2 | Z_2 < Z_1, s_1^2, s_2^2\} \equiv f(\mu),$$

as a function of μ. We need only to show that $f(0) \leq f(\mu)$. Putting $d = (1+\gamma)/(1-\gamma))$, we define the sets

$$\begin{aligned}
&A = \{(z_1, z_2) | z_2 \leq z_1, -dz_1 \leq z_2\}, \quad B = \{(z_1, z_2) | z_2 \leq z_1, -dz_1 > z_2\}, \\
&A_1 = \{(z_1, z_2) | z_2 \leq z_1, z_2 \geq 0\}, \quad and \quad A_2 = \{(z_1, z_2) | -dz_1 \leq z_2, z_2 < 0\}.
\end{aligned}$$

Since A_1 and A_2 are disjoint and $A = A_1 \cup A_2$, we have

$$\begin{aligned}
f(\mu) - f(0) &= \frac{P_\mu(A)}{P_\mu(A) + P_\mu(B)} - \frac{P_0(A)}{P_0(A) + P_0(B)} \\
&= \frac{\{P_\mu(A_1)P_0(B) - P_0(A_1)P_\mu(B)\} + \{P_\mu(A_2)P_0(B) - P_0(A_2)P_\mu(B)\}}{\{P_\mu(A) + P_\mu(B)\} \times \{P_0(A) + P_0(B)\}}.
\end{aligned}$$

We first show that $\{P_\mu(A_1)P_0(B) - P_0(A_1)P_\mu(B)\} > 0$ for $\mu > 0$. For that purpose, we note that

$$\begin{aligned}
P_\mu(B) &= \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\tau_2^2}} exp\{-\frac{(z_2-\mu)^2}{2\tau_2^2}\} \int_{z_2}^{-z_2/d} \frac{1}{\tau_1}\phi(z_1/\tau_1) dz_1 dz_2 \\
&< exp\{-\frac{\mu^2}{2\tau_2^2}\} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\tau_2^2}} exp\{-\frac{z_2^2}{2\tau_2^2}\} \int_{z_2}^{-z_2/d} \frac{1}{\tau_1}\phi(z_1/\tau_1) dz_1 dz_2 \\
&= exp\{-\frac{\mu^2}{2\tau_2^2}\} P_0(B). \qquad (23)
\end{aligned}$$

Similarly, we have

$$P_\mu(A_1) = \int_0^\infty \frac{1}{\sqrt{2\pi\tau_2^2}} exp\{-\frac{(z_2-\mu)^2}{2\tau_2^2}\} \int_{z_2}^\infty \frac{1}{\tau_1}\phi(z_1/\tau_1)dz_1dz_2$$
$$> exp\{-\frac{\mu^2}{2\tau_2^2}\}P_0(A_1). \tag{24}$$

From (23) and (24), we see that $\{P_\mu(A_1)P_0(B) - P_0(A_1)P_\mu(B)\} > 0$.
Next, we show that $\{P_\mu(A_2)P_0(B) - P_0(A_2)P_\mu(B)\} > 0$ for $\mu > 0$. We express $P_\mu(A_2)$ as

$$P_\mu(A_2) = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi\tau_2^2}} exp\{-\frac{(z_2-\mu)^2}{2\tau_2^2}\} \int_{-z_2/d}^\infty \frac{1}{\tau_1}\phi(z_1/\tau_1)dz_1dz_2$$
$$= P_\mu\{Z_2 < 0\}E_\mu[g(Z_2)|Z_2 < 0],$$

where $g(z_2) = \int_{-z_2/d}^\infty \phi(z_1/\tau_1)/\tau_1 dz_1$. Since $g(z_2)$ is an increasing function and the conditional distribution of $Z_2 < 0$ is stochastically smallest when $\mu = 0$, we have for $\mu > 0$

$$P_\mu(A_2) > P_\mu\{Z_2 < 0\}E_0[g(Z_2)|Z_2 < 0] = P_0\{A_2\}P_\mu\{Z_2 < 0\}/P_0\{Z_2\}. \tag{25}$$

Similarly, since $h(z_2) = \int_{z_2}^{-z_2/d} \phi(z_1/\tau_1)/\tau_1 dz_1$ is a decreasing function, we have

$$P_\mu(B) = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi\tau_2^2}} exp\{-\frac{(z_2-\mu)^2}{2\tau_2^2}\} \int_{z_2}^{-z_2/d} \frac{1}{\tau_1}\phi(z_1/\tau_1)dz_1dz_2$$
$$< P_\mu\{Z_2 < 0\}E_\mu[h(Z_2)|Z_2 < 0]$$
$$= P_0(B)P_\mu\{Z_2 < 0\}/P_0\{Z_2 < 0\}. \tag{26}$$

From (25) and (26), we have $\{P_\mu(A_2)P_0(B) - P_0(A_2)P_\mu(B)\} > 0$ and we have shown that $f(\mu) > f(0)$ for $\mu > 0$.

b)In the estimation problem of a common mean, as is stated in Kubokawa (1989) and according to the formula (26), $\hat{\mu}(\gamma)$ is closer to μ than $\bar{X}_1$ if and only if

$$P\{(1-\gamma)(U_2-U_1)^2 + 2U_1(U_2-U_1) \le 0\} \ge \frac{1}{2}, \tag{27}$$

where $U_i = \bar{X}_i - \mu$, $i = 1, 2$. Since

$$(1-\gamma)(U_2-U_1)^2 + 2U_1(U_2-U_1) = (U_2-U_1)\{(1+\gamma)U_1 + (1-\gamma)U_2\}, \tag{28}$$

the left-hand side of (27) is expressed as

$$P\{(1-\gamma)(U_2-U_1)^2 + 2U_1(U_2-U_1) \le 0\}$$
$$= P\{U_2 \ge U_1\}P\{(1+\gamma)U_1 + (1-\gamma)U_2 < 0|U_2 \ge U_1\}$$
$$+ P\{U_2 < U_1\}P\{(1+\gamma)U_1 + (1-\gamma)U_2 > 0|U_2 < U_1\}.$$

We notice that

$$P\{(1+\gamma)U_1 + (1-\gamma)U_2 < 0|U_2 \ge U_1\}$$
$$= P\{(1+\gamma)U_1 + (1-\gamma)U_2 > 0|U_2 < U_1\}.$$

Since U_1 and U_2 are symmetrically distributed about the origin, thus

$$P\{U_2 \ge U_1\} = P\{U_2 < U_1\} = \frac{1}{2}. \tag{29}$$

We see that the left-hand side of (26) is equal to

$$P\{(1+\gamma)U_1 + (1-\gamma)U_2 > 0|U_2 < U_1\},$$

which is $MPN_{\mu_1}(\hat{\mu}_1(\gamma), \bar{X}_1)$ given by (22) for the case $\mu_1 = \mu_2$. Therefore, we see from (a) that $MPN_{\mu_1}(\hat{\mu}_1(\gamma), \bar{X}_1)) \ge \frac{1}{2}$ for all $\mu_1 \le \mu_2$ and for all σ_i^2, $i = 1, 2$ if and only if $PN_\mu(\hat{\mu}(\gamma), \bar{X}_i) \ge \frac{1}{2}$ for all μ and for all σ_i^2, $i = 1, 2$. We complete the proof. □

Remark 3.2. In the estimation problem of a common mean, Kubokawa (1989) has given a sufficient condition on sample sizes n_1 and n_2 for $\hat{\mu}(\gamma)$ to be closer to μ than $\bar{X}_i$ for some specified class of γ .

Remark 3.3. We should mention about the general case when γ is a function of $s_i^2, i = 1, 2$ and $(\bar{X}_1 - \bar{X}_2)^2$. We first consider the case when we estimate μ_1 and suppose that $\hat{\mu_1}(\gamma_0)$ is closer to μ_1 than $\bar{X}_1$, where γ_0 is a function of s_i^2 and possibly $(\bar{X}_1 - \bar{X}_2)^2$. For any γ satisfying $\gamma_0 \leq \gamma < 1$ if $\gamma_0 < 1$, $\hat{\mu}_1(\gamma)$ is closer to μ_1 than $\bar{X}_1$. This is seen from (22), since (22) is true even when γ depends on $(\bar{X}_1 - \bar{X}_2)^2$ and (22) is an increasing function of γ.

4. Examples

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

Example 3. Consider two univariate normal distributions, when they are subject to the order restriction $\mu_1 \leq \mu_2$. Six different cases are considered here. We simulate the values of random samples $X_{11}, X_{12}, \cdots, X_{1n_1}$, from the univariate distributions $N(\mu_{1r}, s_{1r})$ with means μ_{1r}, $r = a, b, c$, and known variances s_{1r} respectively. Also the values of random samples $X_{21}, X_{22}, \cdots, X_{2n_2}$, from the univariate normal distributions $N(\mu_{2r}, s_{2r})$ with means μ_{2r}, $r = a, b, c$, and known variances s_{2r}, respectively. In each simulation, the process of computation is repeated 10000 times to get an estimate of sample means $\bar{X}_1$ and $\bar{X}_2$, isotonic estimators of means, i.e. $\hat{\mu}_1$ and $\hat{\mu}_1$ by (12) and (13), and the risk difference $RD_{\bar{X}_1,\hat{\mu}_1} = R(\mu_1, \bar{X}_1) - R(\mu_1, \hat{\mu}_1)$ and $RD_{\bar{X}_2,\hat{\mu}_2} = R(\mu_2, \bar{X}_2) - R(\mu_2, \hat{\mu}_2)$. For differente values of sample sizes and $r = a, b, c$ the results are given in Table 2. From the Table 2, it is completely clear that $\mu_{1a} \leq \mu_{2a}$, $\mu_{1b} \leq \mu_{2b}$ and $\mu_{1c} \leq \mu_{2c}$ and in case 2 (r=b)$[n_1 = 10, n_2 = 15, \mu_1 = 4\mu_2 = 4, s_1 = 2, s_2 = 3]$ and in case 1 (r=a) $[n_1 = 20, n_2 = 25, \mu_1 = 4\mu_2 = 4, s_1 = 5, s_2 = 6]$, the isotonic regression $\hat{\mu}_1$ uniformly has the smaller risk than the unrestricted maximum likelihood estimator, $\bar{X}_1$ and the isotonic regression $\hat{\mu}_2$ uniformly has the smaller risk than the unrestricted maximum likelihood estimator, $\bar{X}_2$, respectively. But in other cases the isotonic regression estimator $\hat{\mu}_1$ uniformly has not the smaller risk than the unrestricted maximum likelihood estimator, $\bar{X}_1$ and the isotonic regression estimator $\hat{\mu}_2$ uniformly has not the smaller risk than the unrestricted maximum likelihood estimator, $\bar{X}_2$, since $RD_{\bar{X}_1,\hat{\mu}_1} < 0$ and $RD_{\bar{X}_2,\hat{\mu}_2} < 0$, respectively. Figure 1 shows the risk deifference $RD_{\bar{X}_1,\hat{\mu}_1} = R(\mu_1, \bar{X}_1) - R(\mu_1, \hat{\mu}_1)$ as a function of $\mu_1 = \mu_2$, where $\mu = \mu_{2r} - \mu_{1r}$, for different values of r. Also, figure 2 shows the risk deifference $RD_{\bar{X}_2,\hat{\mu}_2} = R(\mu_2, \bar{X}_2) - R(\mu_2, \hat{\mu}_2)$ as a function of $\mu_1 = \mu_2$, where $\mu = \mu_{2r} - \mu_{1r}$, for different values of r.

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

	Sample sizes	$N(\mu_{1r}, s_{1r})$	$N(\mu_{2r}, s_{2r})$	$RD_{\bar{X}_1,\hat{\mu}_1}$	$RD_{\bar{X}_2,\hat{\mu}_2}$
$case1(r = a)$	$n_1 = 10$	$\mu_{1a} = 3$	$\mu_{2a} = 4$	1.179	-0.235
	$n_2 = 15$	$s_{1a} = 4$	$s_{2a} = 5$		
$case2(r = b)$	$n_1 = 10$	$\mu_{1b} = 4$	$\mu_{2b} = 4$	0.011	0.127
	$n_2 = 15$	$s_{1b} = 2$	$s_{2b} = 3$		
$case3(r = c)$	$n_1 = 10$	$\mu_{1c} = 3$	$\mu_{2c} = 3$	-0.139	-0.110
	$n_2 = 10$	$s_{1c} = 5$	$s_{2c} = 6$		
$case1(r = a)$	$n_1 = 20$	$\mu_{1a} = 4$	$\mu_{2a} = 4$	0.048	0.013
	$n_2 = 25$	$s_{1a} = 5$	$s_{2a} = 6$		
$case2(r = b)$	$n_1 = 20$	$\mu_{1b} = 5$	$\mu_{2b} = 5$	-0.019	-0.067
	$n_2 = 25$	$s_{1b} = 4$	$s_{2b} = 6$		
$case3(r = c)$	$n_1 = 20$	$\mu_{1c} = 7$	$\mu_{2c} = 7$	-0.037	-0.068
	$n_2 = 20$	$s_{1c} = 6$	$s_{2c} = 7$		

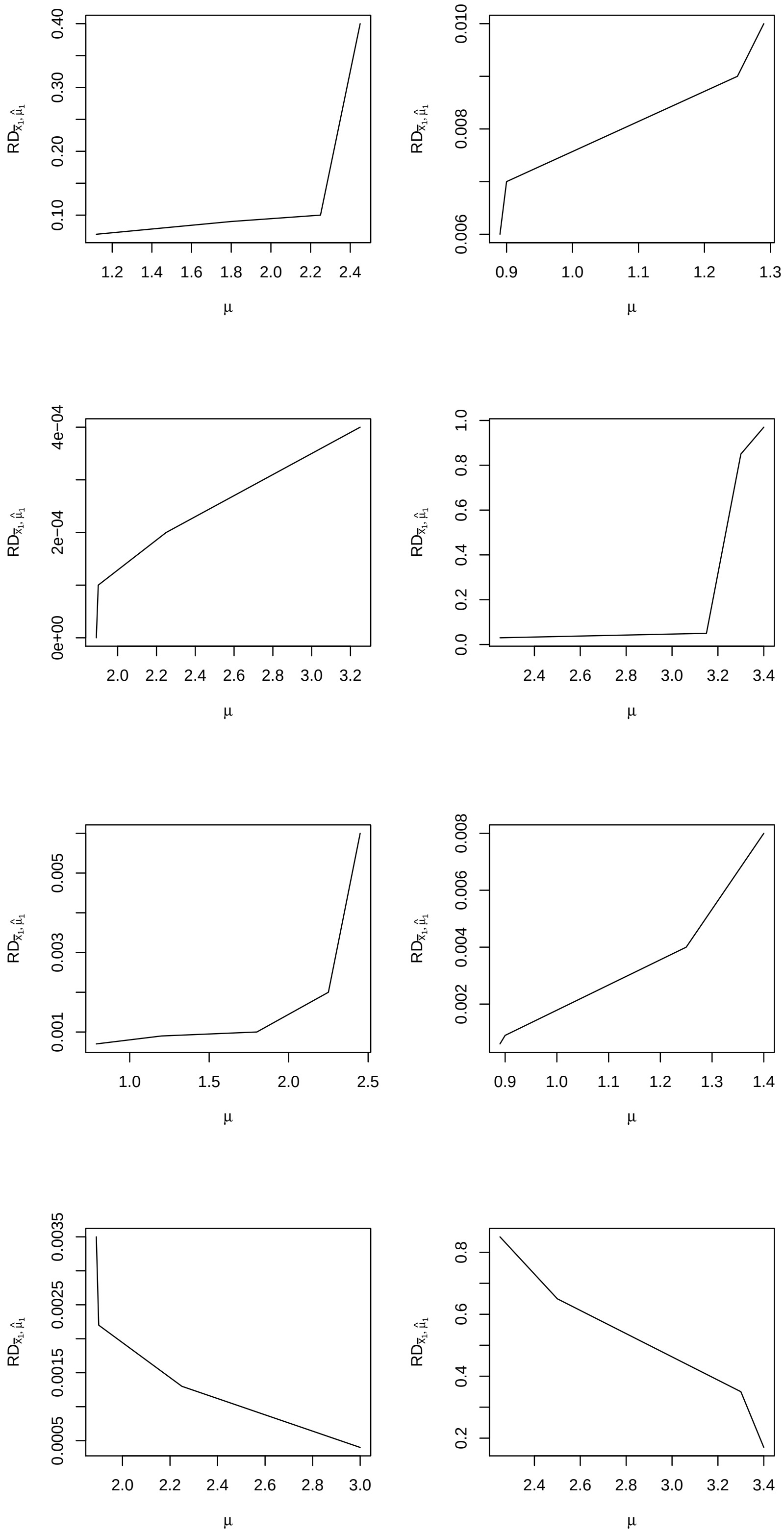

Figure 1: Risk difference $RD_{\bar{X}_1,\hat{\mu}_1} = R(\mu_1, \bar{X}_1) - R(\mu_1, \hat{\mu}_1)$.

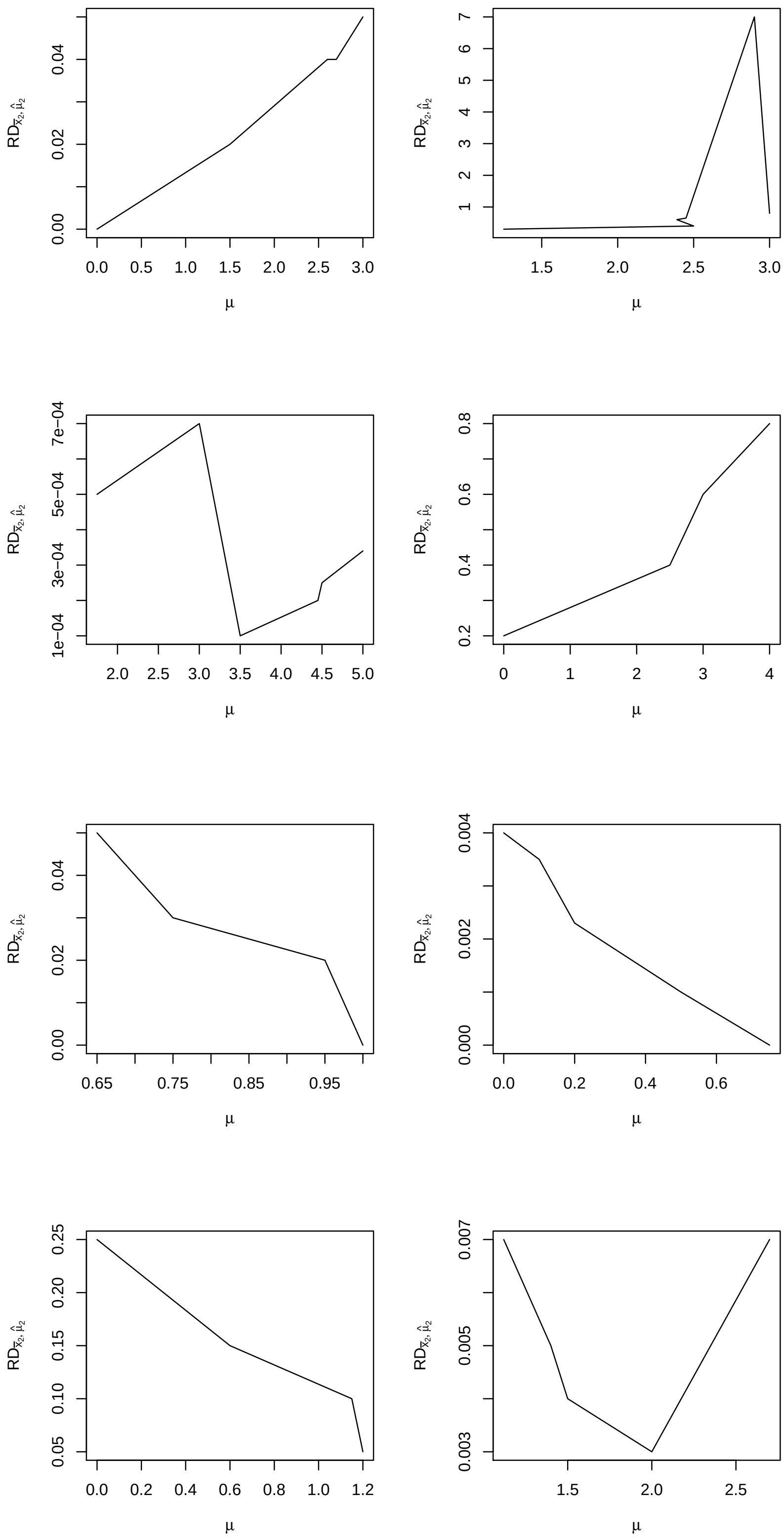

Figure 2: Risk difference $RD_{\bar{X}_2,\hat{\mu}_2} = R(\mu_2, \bar{X}_2) - R(\mu_2, \hat{\mu}_2)$.

Example 4. Consider two univariate normal distributions, when they are subject to the order restriction $\mu_1 \leq \mu_2$. Six different cases are considered here. We simulate the values of random samples $X_{11}, X_{12}, \cdots, X_{1n_1}$, from the univariate distributions $N(\mu_{1r}, s_{1r})$ with means μ_{1r}, $r = a, b, c$, and known variances s_{1r} respectively. Also the values of random samples $X_{21}, X_{22}, \cdots, X_{2n_2}$, from the univariate normal distributions $N(\mu_{2r}, s_{2r})$ with means μ_{2r}, $r = a, b, c$, and known variances s_{2r}, respectively. In each simulation, the process of computation is repeated 10000 times to get an estimate of sample means $\bar{X}_1$ and $\bar{X}_2$, isotonic estimators of means, i.e. $\hat{\mu}_1$ and $\hat{\mu}_1$ by (12) and (13), and $MPN_{\mu_i}(\hat{\mu}_i(\gamma), \bar{X}_i) \geq \frac{1}{2}$. For differente values of sample sizes and $r = a, b, c$ the results are given in Table 3. From the Table 3, it is completely clear that $\mu_{1a} \leq \mu_{2a}$, $\mu_{1b} \leq \mu_{2b}$ and $\mu_{1c} \leq \mu_{2c}$, and modified Pitman nearness of $(\bar{X}_1, \mu_1)$ is greater than $\frac{1}{2}$ for all of cases. Also, the modified Pitman nearness of $(\bar{X}_2, \mu_2)$ is greater than $\frac{1}{2}$ in cases 1,2,3 and 5. But in cases 4 and 6, the modified Pitman nearness of $(\bar{X}_2, \mu_2)$ is not greater than $\frac{1}{2}$. Figure 3 shows $MPN_{\mu_1} = MPN_{\mu_1}(\hat{\mu}_1(\gamma), \bar{X}_1)$ as a function of $\mu_1 = \mu_2$, where $\mu = \mu_{2r} - \mu_{1r}$, for different values of r. Also, figure 4 shows $MPN_{\mu_2} = MPN_{\mu_2}(\hat{\mu}_2(\gamma), \bar{X}_2)$ as a function of $\mu_1 = \mu_2$, where $\mu = \mu_{2r} - \mu_{1r}$, for different values of r.

Table 3: Simulation from two univariate normal distributions: the values of $MPN_{\mu_1} = MPN_{\mu_1}(\hat{\mu}_1(\gamma), \bar{X}_1)$ and $MPN_{\mu_2} = MPN_{\mu_2}(\hat{\mu}_2(\gamma), \bar{X}_2)$.

	Sample sizes	$N(\mu_{1r}, s_{1r})$	$N(\mu_{2r}, s_{2r})$	MPN_{μ_1}	MPN_{μ_2}
$case1(r = a)$	$n_1 = 15$	$\mu_{1a} = 6$	$\mu_{2a} = 7$	0.208	0.568
	$n_2 = 15$	$s_{1a} = 4$	$s_{2a} = 5$		
$case2(r = b)$	$n_1 = 10$	$\mu_{1b} = 4$	$\mu_{2b} = 5$	0.252	0.580
	$n_2 = 20$	$s_{1b} = 5$	$s_{2b} = 7$		
$case3(r = c)$	$n_1 = 15$	$\mu_{1c} = 3$	$\mu_{2c} = 3$	0.303	0.618
	$n_2 = 20$	$s_{1c} = 5$	$s_{2c} = 7$		
$case1(r = a)$	$n_1 = 20$	$\mu_{1a} = 9$	$\mu_{2a} = 9$	0.428	0.379
	$n_2 = 25$	$s_{1a} = 7$	$s_{2a} = 4$		
$case2(r = b)$	$n_1 = 20$	$\mu_{1b} = 6$	$\mu_{2b} = 7$	0.188	0.549
	$n_2 = 20$	$s_{1b} = 4$	$s_{2b} = 5$		
$case3(r = c)$	$n_1 = 15$	$\mu_{1c} = 5$	$\mu_{2c} = 7$	0.074	0.372
	$n_2 = 20$	$s_{1c} = 3$	$s_{2c} = 6$		

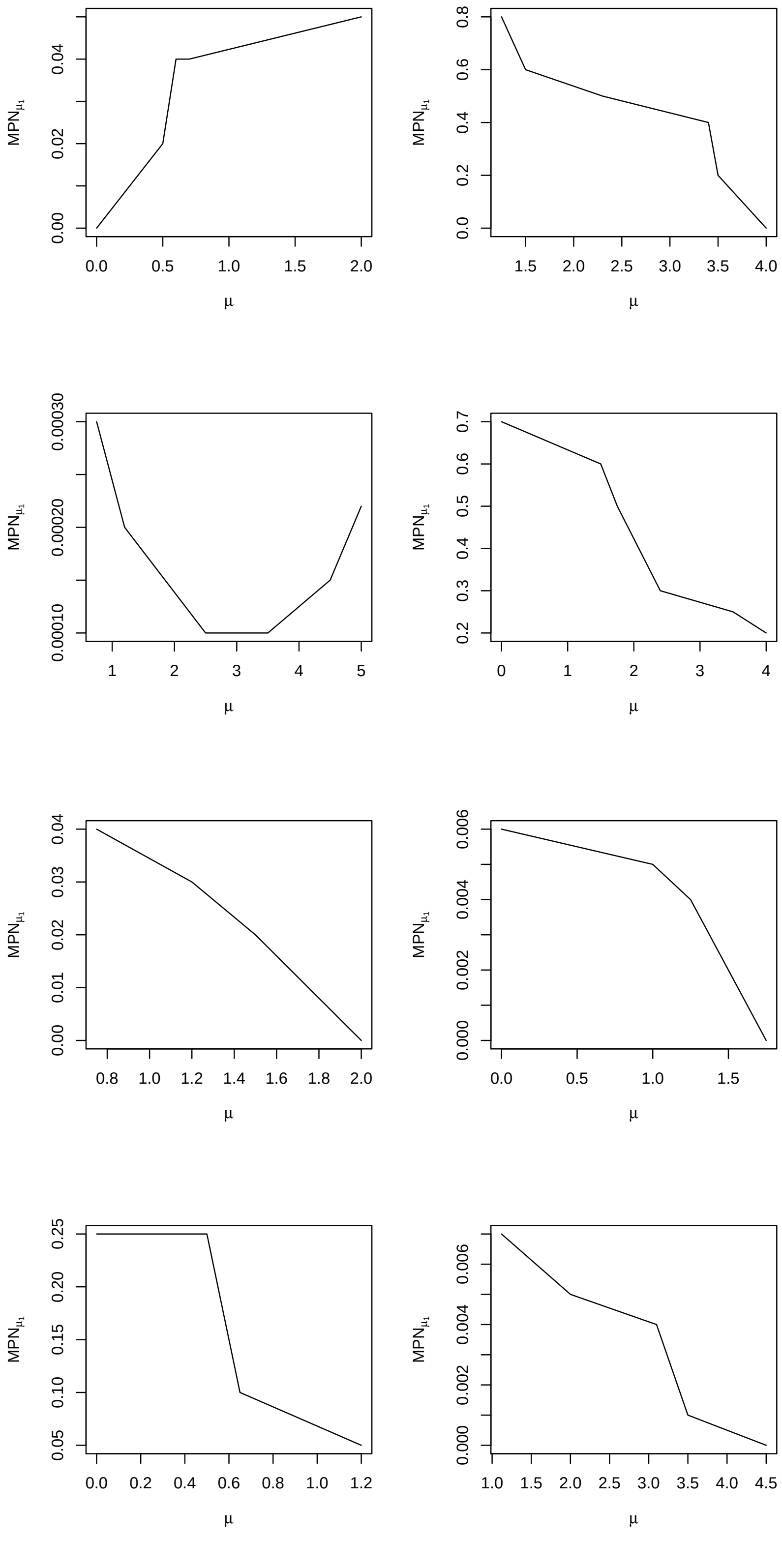

Figure 3: $MPN_{\mu_1} = MPN_{\mu_1}(\hat{\mu}_1(\gamma), \bar{X}_1)$.

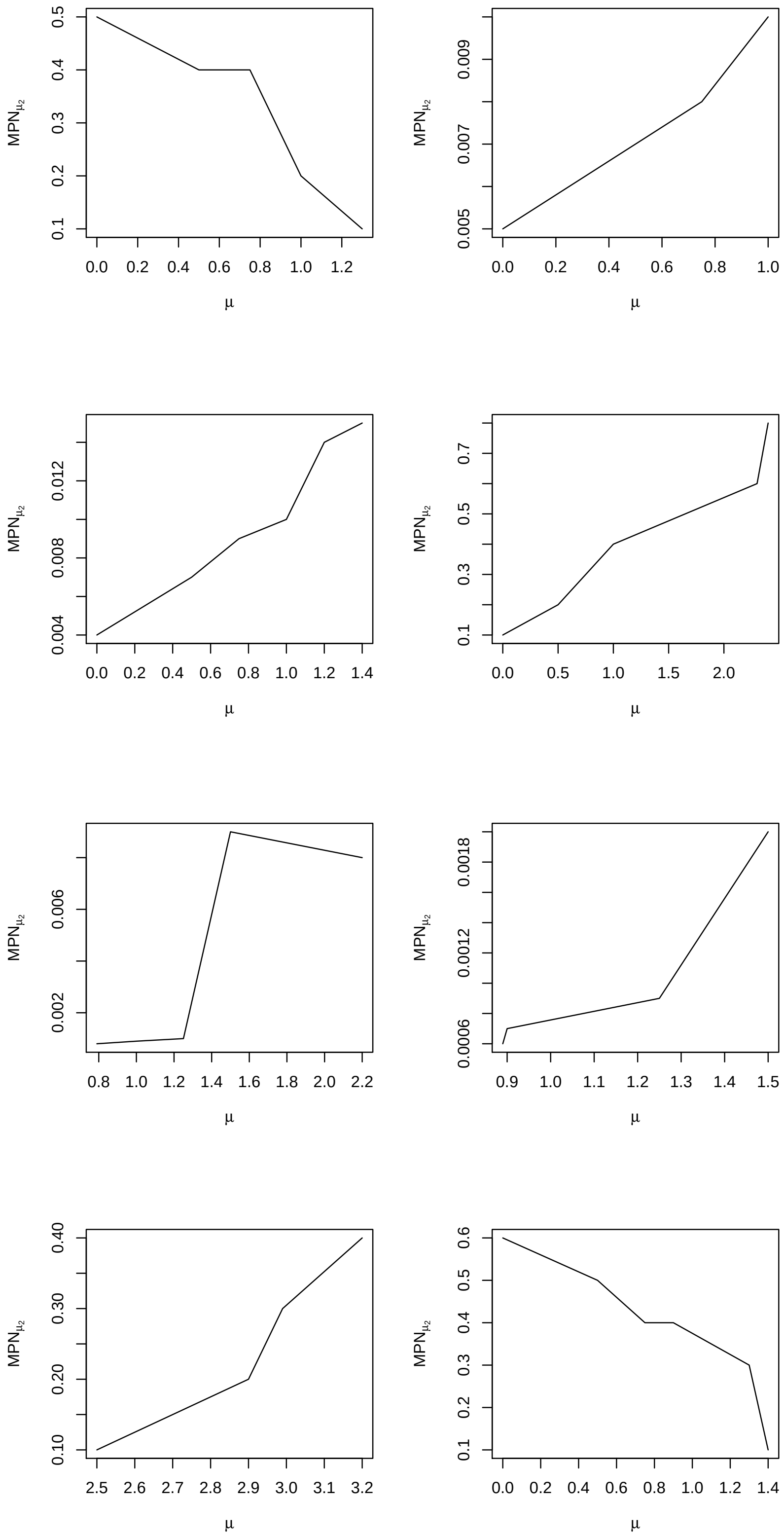

Figure 4: $MPN_{\mu_2} = MPN_{\mu_2}(\hat{\mu}_2(\gamma), \bar{X}_2)$.

5. Conclusion

In this paper, we have deal with the problem of estimating two ordered normal means under the squared error loss function when the variances are unknown and unequal. We showed that the plug-in estimator $\hat{\mu}_1$ uniformly improves upon the unrestricted maximum likelihood estimator $\bar{X}_1$ if and only if for all σ_i^2, the risk of $\hat{\mu}_1$ is not larger than that of $\bar{X}_1$ when $\mu_1 = \mu_2$, and showed that the plug-in estimator $\hat{\mu}_2$ uniformly improves upon the unrestricted maximum likelihood estimator $\bar{X}_2$ if and only if for all σ_i^2, the risk of $\hat{\mu}_2$ is not larger than that of $\bar{X}_2$ when $\mu_1 = \mu_2$. Also, under modified Pitman nearness criterion when the order restriction on variances is not present, it is shown that the most critical case for $\hat{\mu}_i(\gamma)$ to improve upon $\bar{X}_i$ is the one when $\mu_1 = \mu_2$ and that the problem of improving upon $\bar{X}_i$ reduces to the one of a common mean. Also, two numerical examples presented to illustrate the results. In example 1, the data simulated from different bivariate normal distributions. We showed that, in two cases, the isotonic regression estimators uniformly have the smaller risk than the unrestricted maximum likelihood estimator since the risk differences are positive and in the other cases, the isotonic regression estimators uniformly have the smaller risk than the unrestricted maximum likelihood estimator since the risk differences are negative. In example 2, the data simulated from different bivariate normal distributions. We showed that the modified Pitman nearness of $(\bar{X}_1, \mu_1)$ is greater than $\frac{1}{2}$ for all of cases. But, the modified Pitman nearness of $(\bar{X}_2, \mu_2)$ is greater than $\frac{1}{2}$ for some cases.

Acknowledgement

The authors would like to thank the editor and the anonymous referees for their helpful and constructive comments.

References

Bazyari A (2015). “On the Properties of Estimates of Monotonic Mean Vectors for Multivariate Normal Distributions.” *Journal of Statistical Theory and Applications*, **14**(1), 89–106.

Bhattacharya CG, *et al.* (1980). “Estimation of a Common Mean and Recovery of Interblock Information.” *The Annals of Statistics*, **8**(1), 205–211.

Brown L, Cohen A (1974). “Point and Confidence Estimation of a Common Mean and Recovery of Interblock Information.” *The Annals of Statistics*, pp. 963–976.

Chang YT, Shinozaki N (2012). “Estimation of Ordered Means of Two Normal Distributions with Ordered Variances.” *Journal of Mathematics and System Science*, **2**(1).

Gupta RD, Singh H (1992). “Pitman Nearness Comparisons of Estimates of Two Ordered Normal Means.” *Australian Journal of Statistics*, **34**(3), 407–414.

Hwang JG, Peddada SD (1994). “Confidence Interval Estimation Subject to Order Restrictions.” *The Annals of Statistics*, pp. 67–93.

Kelly RE (1989). “Stochastic Reduction of Loss in Estimating Normal Means by Isotonic Regression.” *The Annals of Statistics*, pp. 937–940.

Khatri C, Shah K (1974). “Estimation of Location Parameters from Two Linear Models under Normality.” *Communications in Statistics-Theory and Methods*, **3**(7), 647–663.

Kubokawa T (1989). “Closer Estimators of a Common Mean in the Sense of Pitman.” *Annals of the Institute of Statistical Mathematics*, **41**(3), 477–484.

Lee CIC (1981). "The Quadratic Loss of Isotonic Regression under Normality." *The Annals of Statistics*, pp. 686–688.

Nayak TK (1990). "Estimation of Location and Scale Parameters using Generalized Pitman Nearness Criterion." *Journal of Statistical Planning and Inference*, **24**(2), 259–268.

Pitman EJ (1937). "The "closest" Estimates of Statistical Parameters." **33**(02), 212–222.

Rao CR (1980). *Some Comments on the Minimum Mean Square Error as a Criterion of Estimation.*

Affiliation:

Najmeh Pedram and Abouzar Bazyari
Department of Statistics, Persian Gulf University, Bushehr, Iran
E-mail: ab_bazyari@yahoo.com

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 46 *Submitted:* 2016-03-28
February 2017 *Accepted:* 2016-08-01