İSTATİSTİK: JOURNAL OF THE TURKISH STATISTICAL ASSOCIATION
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 1300-4077 | 00 | 0000 | 0001 | 000

İSTATİSTİK

SIMULTANEOUS ESTIMATION OF NORMAL MEANS UNDER SIMPLE ORDER

Rahul Bhattacharya* and Soumyadeep Das
University of Calcutta, India

Abstract : Simultaneous estimation of several normal means is considered under simple order restriction. Maximum likelihood estimators of normal means are derived under such order restriction and compared empirically with the existing estimators using Pitman closeness as well as mean square type precision criteria.

Key words: Maximum likelihood estimator, Order restriction, Pitman nearness.

1. Introduction

Estimation of order restricted normal means for several populations with unknown as well as unequal variances arises in many inferential problems. Various estimators of normal means have been proposed and investigated over the decades including the works of Gupta and Singh (1992), Hwang and Peddada (1994), Chang and Shinozaki (2012) and Bazyari (2015), among others. Most of them have given improved estimators by simply comparing them with the sample mean, that is, unrestricted maximum likelihood estimators of the normal means. For two normal populations, the works of Brown and Cohen (1974), Khatri and Shah (1974) and Bhattacharya et al. (1980) provided a class of improved estimators, which is a weighted combination of the sample means with data dependent weights. The latest addition in the list is the estimators of Pedram and Bazyari (2017). Starting from the maximum likelihood estimators for known variances, they obtained their plug-in estimators by replacing the variances with usual unbiased estimators. Considering two independent normal populations, they compared the performance of their plug-in estimator with that of the sample mean in terms of both mean square errors and modified Pitman closeness criterion (Nayak, 1990). Although simultaneous estimation of normal means is considered under an order restriction, performance of their estimator is compared only component-wise. However, none of the component-wise estimators were found to be uniformly "better" than the usual estimator for the unrestricted configuration. Order restrictions among the unknown parameters connects the independent parameters and hence component-wise performance evaluation ignores the connection and may lead to ambiguity. That is, some individual estimator may have larger mean square error than that of the usual estimator and others less.

For a neat description, consider independent observations $X_{ij},\ j = 1,2,...,n_i$, from the i th population, $i = 1,2,...,k$, where X_{ij} is distributed as normal with mean μ_i and variance $\sigma_i^2,\ j = 1,2,...,n_i,\ i = 1,2,...,k$. It is further known that the components of $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_k)^T$ satisfy the simple order, that is,

$$\mu_1 \le \mu_2 \le \cdots \le \mu_k. \tag{1.1}$$

Simple order restriction among multiple normal means often arises in practical situations. Several real life situations, including those mentioned in Ma and Liu (2013) and Pedram and Bazyari (2017) reported the existence of natural orderings. However, we start with an example

* Corresponding author. E-mail address: bhattya@yahoo.com (R. Bhattacharya)

provided in Shi (2004) based on the marks of students in an entrance examination.

Example: Consider 7 courses, namely, language (full marks 120), Mathematics (full marks 120), Politics (full marks 100), Physics (full marks 100), Chemistry (full marks 100), English (full marks 100) and Biology (full marks 70) in the Chinese notational college entrance examination, 1992. Table 1 gives the average total scores and the corresponding variances in five districts of Jillin province, with 100 students from each district.

TABLE 1. Summary measures for the selected students of Jillin.

	District 1	District 2	District 3	District 4	District 5
Sample mean	388.270	384.610	398.000	395.170	418.010
Sample Variance	4013.917	5354.438	4269.380	3582.821	4928.749

The conditions of education in these five districts are different. However, past experience established that the condition of education in the $(i+1)$ th district is better than that for the i th district for each $i = 1, 2, 3, 4$. If μ_i denotes the mean score for the i th district, observations from the past experience can be expressed by the simple order restriction $\mu_1 \le \mu_2 \le \mu_3 \le \mu_4 \le \mu_5$. Description of the real example compels the statistician to take into account the possibility of order restrictions among the unknown mean parameters. On the other hand, restrictions among the parameters bridge the independent populations and hence any inference procedure should be developed for the whole set of parameters to use the connected feature of independent populations. Consequently, in the current work, we derive estimators of the normal means under simple order (Robertson et al., 1988, Silvapulle and Sen, 2006) following the celebrated maximum likelihood method of estimation and explore the performance empirically. Starting with the formulation of the problem, we derive the maximum likelihood estimators of mean under simple order for three independent normal populations. The exact expression of the estimator together with related large sample properties are also derived in section 2. Proposed estimator is compared from the perspective of simultaneous estimation with those of Pedram and Bazyari (2017) and the usual estimator for no restriction in terms of both Corrected Pitman Nearness (Keating et al., 1993) and a relevant mean square type precision measure. All these can be found in section 3 with relevant numerical computations. Section 4 provides different types of estimates for the real example of Shi (2004). Finally, in section 5, we investigate some related issues and the scope of possible extension.

2. Maximum likelihood estimators under simple order For the purpose of estimation, we consider the method of maximum likelihood under order restrictions. Naturally, maximizing the log likelihood function (apart from an additive constant)

$$\ell(\mu_1, \mu_2, ..., \mu_k, \sigma_1^2, \sigma_2^2, ..., \sigma_k^2) = -\sum_{i=1}^{k} \left\{ n_i log\sigma_i + \frac{\sum\limits_{j=1}^{n_i}(X_{ij} - \mu_i)^2}{2\sigma_i^2} \right\}$$

subject to the order restriction (1.1), the maximum likelihood (ML) estimator $\widehat{\boldsymbol{\mu}}_M$ of $\boldsymbol{\mu}$ can be obtained. However, Pedram and Bazyari (2017) considered the ML estimator ,

$$\underset{t \ge i}{min}\ \underset{s \le i}{max} \frac{\sum_{j=s}^{t} \sum_{r=1}^{n_j} \frac{X_{rj}}{\sigma_j^2}}{\sum_{j=s}^{t} \frac{n_j}{\sigma_j^2}}, \ i = 1, 2, ..., k.$$

of $\mu_i$under (1.1) assuming $(\sigma_1^2, \sigma_2^2, ..., \sigma_k^2)$ known and then suggested to replace σ_j^2 by the respective unbiased estimators to get the corresponding plug-in estimators. For $k = 2$, they expressed their plug-in estimator in the reduced form

$$\widehat{\mu}_1 = min\left(\bar{X}_1, \frac{\sum_{i=1}^{2} \frac{n_i \bar{X}_i}{s_i'^2}}{\sum_{i=1}^{2} \frac{n_i}{s_i'^2}}\right)$$

and

$$\widehat{\mu}_2 = max\left(\bar{X}_2, \frac{\sum_{i=1}^{2} \frac{n_i \bar{X}_i}{s_i'^2}}{\sum_{i=1}^{2} \frac{n_i}{s_i'^2}}\right)$$

where $\bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$, $s_i'^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2 = \frac{n_i}{n_i - 1}s_i^2$, and compared $\widehat{\mu}_i$ with the unrestricted estimator $\bar{X}_i$ for each i, in terms of both mean square error (MSE) and the modified Pitman nearness criterion of Nayak (1990). However, the estimator $\widehat{\mu}_i$ is derived on an ad-hoc basis and hence does not coincide with the corresponding maximum likelihood (ML) estimator under order restriction. Moreover, the performance measures are compared separately for each of $\widehat{\mu}_1$ and $\widehat{\mu}_2$ and hence fails to capture any insight of the simultaneous estimation problem.

On the contrary, we consider ML estimation of $\boldsymbol{\mu}$ under (1.1) in the current work. Such an estimator $\widehat{\boldsymbol{\mu}}_M = (\widehat{\mu}_{M1}, \widehat{\mu}_{M2}, ..., \widehat{\mu}_{Mk})^T$ is not easy to obtain and requires rigorous application of non-linear optimization techniques (Bazarra et al., 1993).

However, we provide the estimator in the general form. Naturally, $\widehat{\boldsymbol{\mu}}_M$ is the solution of the optimization problem:

$$\begin{aligned} &\text{Minimize } -\ell(\mu_1, \mu_2, ..., \mu_k, \sigma_1^2, \sigma_2^2, ..., \sigma_k^2) \\ &\text{subject to } \mu_1 \le \mu_2 \le \cdots \le \mu_k. \end{aligned}$$

Since, the restriction $\mu_1 \le \mu_2 \le \cdots \le \mu_k$ is equivalent to the inequalities $\mu_i - \mu_{i+1} \le 0,\ i = 1, 2, ..., k-1$, we have the following Lagrangian,

$$F(\mu_1, \mu_2, ..., \mu_k, \sigma_1^2, \sigma_2^2, ..., \sigma_k^2) = -\ell(\mu_1, \mu_2, ..., \mu_k, \sigma_1^2, \sigma_2^2, ..., \sigma_k^2) + \sum_{i=1}^{k-1} l_i(\mu_i - \mu_{i+1}) \tag{2.1}$$

with the non-negative multipliers $l_i,\ i = 1, 2, ..., k-1$. Then $\widehat{\boldsymbol{\mu}}_M$ is the solution of the equations

$$\frac{\partial F}{\partial \mu_1} = -\frac{n_1(\bar{X}_1 - \mu_1)}{\sigma_1^2} + l_1 = 0. \tag{2.2}$$

$$\frac{\partial F}{\partial \mu_k} = -\frac{n_k(\bar{X}_k - \mu_k)}{\sigma_k^2} - l_{k-1} = 0. \tag{2.3}$$

$$\frac{\partial F}{\partial \mu_i} = -\frac{n_i(\bar{X}_i - \mu_i)}{\sigma_i^2} + l_i - l_{i-1} = 0,\ i = 2, 3, ..., k-1. \tag{2.4}$$

$$\frac{\partial F}{\partial \sigma_i} = \frac{1}{\sigma_i} - \frac{s_i^2 + (\bar{X}_i - \mu)^2}{\sigma_i^3} = 0,\ i = 1, 2, ..., k. \tag{2.5}$$

$$l_i(\mu_i - \mu_{i+1}) = 0, \text{ for } i = 1, 2, ..., k-1. \tag{2.6}$$

with $l_i \geq 0$, for $i = 1, 2, ..., k-1$.

Now define two sets $M = \{1, 2, ..., k-1\}$ and $M' = \{i_1, i_2, ..., i_s\} \subseteq M$ with $1 \leq i_1 < i_2 < \cdots i_s \leq k-1$ such that $l_i > 0$ for $i \in M'$.

If $s = 0$, $l_i = 0$ for every $i \in M$ and hence $\mu_i \leq \mu_{i+1}$ for $i = 1, 2, ..., k-1$, which gives $\widehat{\boldsymbol{\mu}}_M = (\bar{X}_1, \bar{X}_2, ..., \bar{X}_k)^T$, provided the condition $\bar{X}_1 \leq \bar{X}_2 \leq \cdots \leq \bar{X}_k$ is satisfied.

Next assume that $s > 0$, that is, M' is non-empty. Then, for every $i \in M'$, $\mu_i = \mu_{i+1}$ and we get from equations (2.2) to (2.5) that, $\widehat{\boldsymbol{\mu}}_M$ satisfies $l_i = \sum_{j=1}^{i} \frac{n_j(\bar{X}_j - \widehat{\mu}_{Mj})}{s_j^2 + (\bar{X}_j - \widehat{\mu}_{Mj})^2} > 0$. It should be noted that, depending on the natures of $i_1, i_2, ..., i_s$, the number of equal adjacent $\mu_i's$ varies. For example, if $i_{r+1} = i_r + 1$, $\mu_{i_r} = \mu_{i_r+1} = \mu_{i_r+2}$ but if $i_{r+1} > i_r + 1$, $\mu_{i_r} = \mu_{i_r+1}$ and $\mu_{i_{r+1}} = \mu_{i_{r+1}+1}$, separately. Estimation of such common adjacent parameters are discussed in Lemma 1. Moreover, for $i \in M - M'$, $\mu_i \leq \mu_{i+1}$ and we get $\widehat{\mu}_{Mi} = \bar{X}_i$ provided $\widehat{\mu}_{Mi} \leq \widehat{\mu}_{M(i+1)}$.

Lemma 1 : *If* $\widetilde{\Omega} = \{\boldsymbol{\theta} \;:\; \mu_l = \mu_{l+1} = ... = \mu_{l+r} = \mu,\; 1 \leq l < l + r - 1 \leq k\}$, *with* $\boldsymbol{\theta} = (\mu_1, \mu_2, ..., \mu_k, \sigma_1^2, \sigma_2^2, ..., \sigma_k^2)^T$, *then the ML estimator* $\widetilde{\mu}_M(l, l+1, ..., l+r)$ *of* μ *satisfies the* $(2r+1)$ *degree polynomial equation*

$$\widetilde{h}_{2r+1}(z) = \sum_{i \in A} n_i(z - \bar{X}_i) \prod_{\substack{j \neq i \\ j \in A}} \left\{s_j^2 + (z - \bar{X}_j)^2\right\} = 0 \tag{2.7}$$

with $A = \{l, l+1, ..., l+r\}$.

Proof : For $\boldsymbol{\theta} \in \widetilde{\Omega}$, the log likelihood function $\ell(\mu_1, \mu_2, ..., \mu_k, \sigma_1^2, \sigma_2^2, ..., \sigma_k^2)$ can be expressed as

$$-\sum_{i \in A} \{n_i log\sigma_i + \frac{n_i[s_i^2 + (\bar{X}_i - \mu)^2]}{2\sigma_i^2}\} - \sum_{i \in A^c} \{n_i log\sigma_i + \frac{n_i[s_i^2 + (\bar{X}_i - \mu_i)^2]}{2\sigma_i^2}\},$$

where $A^c = \{1, 2, ..., k\} - A$. Then the likelihood equations give $\mu = \dfrac{\sum_{i \in A} \frac{n_i \bar{X}_i}{\widehat{\sigma}_i^2}}{\sum_{i \in A} \frac{n_i}{\widehat{\sigma}_i^2}}$ and $\sigma_i^2 = s_i^2 + (\bar{X}_i - \mu)^2$, $i \in A$. Replacing σ_i^2 in the expression of μ, we get $\sum_{i \in A} \frac{n_i(\mu - \bar{X}_i)}{s_i^2 + (\mu - \bar{X}_i)^2} = 0$. Rationalizing the above equation, we get the $(2r+1)$ degree polynomial equation $\widetilde{h}_{2r+1}(z) = 0$, where μ is replaced by the dummy variable z.

The following Lemma traces the asymptotic nature of the roots of the polynomial equation (2.7).

Lemma 2 : *Suppose for each* $i = 1, 2, ..., k$, $n_i \to \infty$*in such a way that* $\frac{n_i}{n} \to \phi_i \in [0, 1]$*where* $n = \sum_{j=1}^{k} n_j$*. Then for any* $\boldsymbol{\theta} \in \widetilde{\Omega}$ *and large* n*, the equation (2.7) has exactly one real root converging almost surely to*μ *provided* $\phi_i > 0$ *for at least one* $i \in A$.

Proof : Since, $\bar{X}_i \to \mu$ and $s_i^2 \to \sigma_i^2$ almost surely for $i \in A$, $\frac{1}{n}\widetilde{h}_{2r+1}(z)$ in (2.7) converges almost surely to the $(2r+1)$ degree polynomial

$$h_{2r+1}(z) = (z - \mu) \sum_{i \in A} \phi_i \prod_{\substack{j \neq i \\ j \in A}} \{\sigma_j^2 + (z - \mu)^2\}. \tag{2.8}$$

As $h_{2r+1}(z)$ is a $(2r+1)$ degree polynomial in z with real coefficients, the equation $h_{2r+1}(z)=0$ has at least one real root. However, the expression of $h_{2r+1}(z)$ suggests that μ is always a root of the equation $h_{2r+1}(z)=0$. Then the other roots of $h_{2r+1}(z)=0$ are the roots of the equation $g_{2r}(z)=\sum_{i\in A}\phi_i\prod_{\substack{j\neq i\\ j\in A}}\{\sigma_j^2+(z-\mu)^2\}=0$. Naturally, $g_{2r}(z)$ is a strictly positive quantity for every z as long as $\phi_i>0$ for some $i\in A$. Since $g_{2r}(z)$ is a strictly positive polynomial of degree $2r$, the roots of $g_{2r}=0$ are complex conjugates. Thus the only real root of $h_{2r+1}(z)=0$ is μ. Hence, for large n, the equation (2.7) has exactly one real root, which converges almost surely to μ.

However, in small samples, the equation (2.7) may have multiple real roots and in such a situation, the actual ML estimator of the common parameter is to be decided by examining the likelihood function.

3. Evaluating the performance of the ML estimator

3.1. Performance measures

Performance of a real valued estimator is, in general, evaluated through mean square error (MSE) or equivalently by the expected value of the squared distance between the estimator and the concerned parameter. However, if the estimator is vector valued, sum of the component-wise MSE's could be a reasonable measure. Thus we define relative precision measure between two k component estimators $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ of the vector parameter $\boldsymbol{\theta}$,

$$RP_{\boldsymbol{\theta}}(\boldsymbol{T}_1/\boldsymbol{T}_2)=\frac{E||\boldsymbol{T}_2-\boldsymbol{\theta}||^2}{E||\boldsymbol{T}_1-\boldsymbol{\theta}||^2},$$

where $||\boldsymbol{T}_i-\boldsymbol{\theta}||$ denotes the Euclidean distance between $\boldsymbol{T}_i$ and $\boldsymbol{\theta},\ i=1,2$ and a value higher than unity indicated that $\boldsymbol{T}_1$ is better than $\boldsymbol{T}_2$. However, one can use other distance metrics instead of the Euclidean distance to derive the relative precision.

Naturally the relative measure of precision depends on the existence of higher order moments and hence, as an alternative, consider the corrected Pitman Nearness (CPN) criterion of Keating et al. (1993). For the estimators $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ of the vector parameter $\boldsymbol{\theta}$, CPN is defined by

$$CPN_{\boldsymbol{\theta}}(\boldsymbol{T}_1/\boldsymbol{T}_2)=P_{\boldsymbol{\theta}}\left\{||\boldsymbol{T}_1-\boldsymbol{\theta}||<||\boldsymbol{T}_2-\boldsymbol{\theta}||\right\}+\frac{1}{2}P_{\boldsymbol{\theta}}\left\{||\boldsymbol{T}_1-\boldsymbol{\theta}||=||\boldsymbol{T}_2-\boldsymbol{\theta}||\right\}.$$

Then $\boldsymbol{T}_1$ is closer (i.e. better) to $\boldsymbol{\theta}$ than $\boldsymbol{T}_2$ if $CPN_{\boldsymbol{\theta}}(\boldsymbol{T}_1/\boldsymbol{T}_2)\geq\frac{1}{2}$ for every $\boldsymbol{\theta}$. It is easy to observe that $CPN_{\boldsymbol{\theta}}(\boldsymbol{T}_1/\boldsymbol{T}_2)>$ or $<\frac{1}{2}$ as $P_{\boldsymbol{\theta}}\{||\boldsymbol{T}_1-\boldsymbol{\theta}||<||\boldsymbol{T}_2-\boldsymbol{\theta}||\big|\boldsymbol{T}_1\neq\boldsymbol{T}_2\}>$ or $<\frac{1}{2}$. The conditional probability $P_{\boldsymbol{\theta}}\{||\boldsymbol{T}_1-\boldsymbol{\theta}||<||\boldsymbol{T}_2-\boldsymbol{\theta}||\big|\boldsymbol{T}_1\neq\boldsymbol{T}_2\}$ is defined by Nayak (1990) as the Modified Pitman Nearness measure. Since, both the formulations are equivalent, we continue with CPN for our purpose.

3.2. Numerical computations

Performance evaluation requires specification of the competitors in addition to the performance measures. Consequently, we consider the estimator of Pedram and Bazyari (2017) (indicated by $\widehat{\boldsymbol{\mu}}_P$) and $\widehat{\boldsymbol{\mu}}_R=(\bar{X}_1,\bar{X}_2,\cdots,\bar{X}_k)^T$, the minimum variance unbiased estimator (MVUE) as well as the maximum likelihood estimator of $\boldsymbol{\mu}$ in the unrestricted situation. For the purpose of computation, we consider several combinations of $(\mu_1,\mu_2,\cdots,\mu_k)$ satisfying (1.1). For each such combination, performance measure are computed for fixed sets of $(\sigma_1,\sigma_2,\cdots,\sigma_k)$ and $(n_1,n_2,\cdots,n_k)$. An extensive simulation study is conducted with 10,000 iterations to compute the performance measures for $k=3,4$. Although, performance measures are computed for a large number of configurations, we provide only few representatives in Tables 2 and 3. Moreover, few figures are included to understand the comparison graphically. Figure 1 compares $\widehat{\boldsymbol{\mu}}_M$, $\widehat{\boldsymbol{\mu}}_P$ and $\widehat{\boldsymbol{\mu}}_R$ in terms of both relative precision (RP) and corrected Pitman nearness (CPN) for $(n_1,n_2,n_3)=(25,30,45)$ and $(\sigma_1,\sigma_2,\sigma_3)=(2,4,3)$. For the computation, we fix $\mu_1=2.4,\mu_2=2.5$ and vary μ_3 over the range [2.5,4]. For each of the RP and CPN, we also add the lines of equivalence (indicated by E), that is, horizontal lines at 0.50 and 1.0, respectively for CPN and RP. Figure 2 give the same comparison for different combinations of (n_1,n_2,n_3) and $(\sigma_1,\sigma_2,\sigma_3)$ considering the configurations $\mu_1=0,\mu_2=0.1,\mu_3\geq 0.1$.

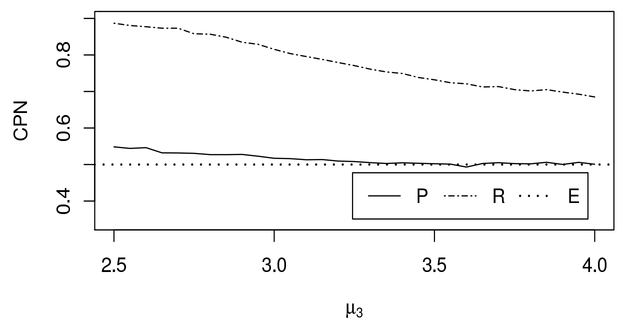
TABLE 2. Comparison between $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_P$ and $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_R$ for $k=3$.

(μ_1,μ_2,μ_3)	$(\sigma_1,\sigma_2,\sigma_3)$	(n_1,n_2,n_3)	$CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$	$CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$	$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$	$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$
(3.00,3.50,4.00)	(1,3,2)	(15,20,25)	0.557	0.751	1.002	1.646
(2.40,2.50,2.60)	(2,4,3)	(25,30,45)	0.544	0.879	1.003	2.176
(0.00,0.50,1.00)	(3,2,1)	(20,20,20)	0.562	0.703	1.001	1.347
(0.00,0.50,0.75)	(3,2,1)	(20,30,40)	0.482	0.747	1.001	1.451
(0.00,0.10,0.20)	(3,2,1)	(15,30,45)	0.513	0.792	0.998	1.821
(0.10,0.20,0.30)	(2,3,4)	(25,25,25)	0.484	0.820	0.998	1.688

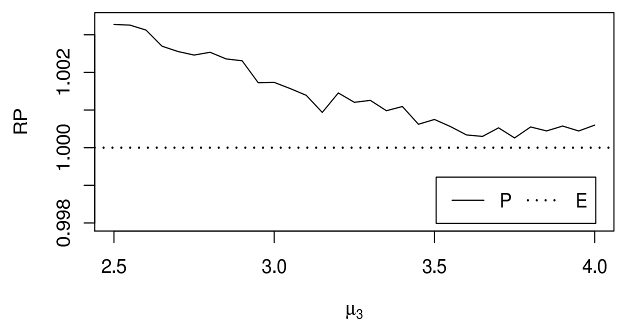
Performance measures like CPN and RP are also computed for $k=4$ and are provided in Table 3 and Figure 4 for different combinations of $(\mu_1,\mu_2,\mu_3,\mu_4)$, $(\sigma_1,\sigma_2,\sigma_3,\sigma_4)$ and (n_1,n_2,n_3,n_4). To be specific, for Figure 3, μ_1,μ_2 and μ_4 are kept fixed at 0, 0.5 and 1.25, respectively, with $(\sigma_1,\sigma_2,\sigma_3,\sigma_4)=(6,5,4,3)$ and $(n_1,n_2,n_2,n_4)=(50,50,50,50)$ and varying μ_3 between μ_2 and μ_4, CPN and RP are calculated. Figure 4 provides a similar comparison with fixed $\mu_1=0,\mu_2=0.1,\mu_4=0.3,(\sigma_1,\sigma_2,\sigma_3,\sigma_4)=(0.2,0.1,0.4,0.3)$ and $(n_1,n_2,n_3,n_4)=(15,20,25,30)$.

TABLE 3. Comparison between $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_P$ and $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_R$ for $k=4$.

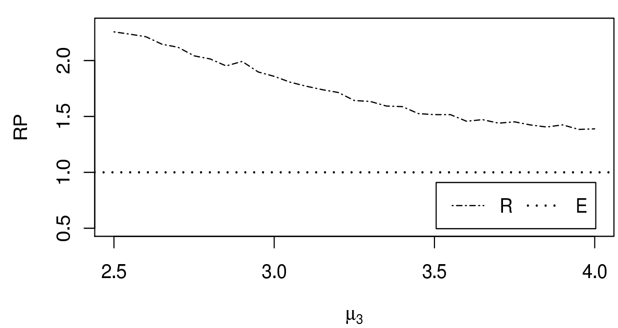
$(\mu_1,\mu_2,\mu_3,\mu_4)$	$(\sigma_1,\sigma_2,\sigma_3,\sigma_4)$	(n_1,n_2,n_3,n_4)	$CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$	$CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$	$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$	$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$
(0.00,0.50,0.75,1.25)	(2.0,6.0,4.0,8.0)	(15,20,25,30)	0.525	0.887	1.005	2.799
(2.00,2.10,2.20,2.30)	(3.0,2.0,1.0,4.0)	(20,20,25,30)	0.505	0.934	1.001	3.356
(0.00,0.50,1.00,1.25)	(6.0,5.0,4.0,3.0)	(50,50,50,50)	0.515	0.802	1.002	1.615
(2.00,2.10,2.20,2.30)	(4.0,3.0,2.0,1.0)	(20,20,25,30)	0.495	0.931	1.007	2.457
(0.00,0.10,0.30,0.60)	(1.0,2.0,3.0,4.0)	(15,20,25,30)	0.526	0.902	0.999	2.276
(0.00,0.50,1.00,1.50)	(1.0,2.0,3.0,4.0)	(20,30,40,50)	0.375	0.515	0.997	1.176
(0.00,0.10,0.20,0.30)	(0.2,0.1,0.4,0.3)	(15,20,25,30)	0.159	0.200	0.970	0.670



(a)$CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$ and $CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$

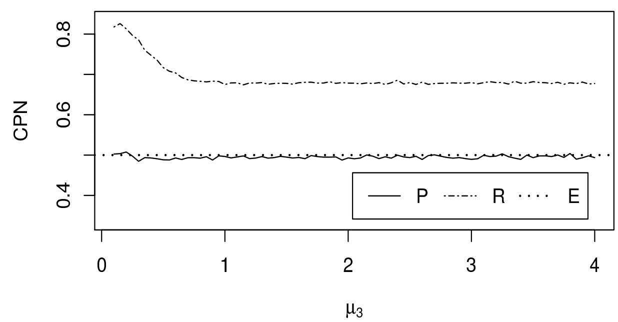


(b)$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$

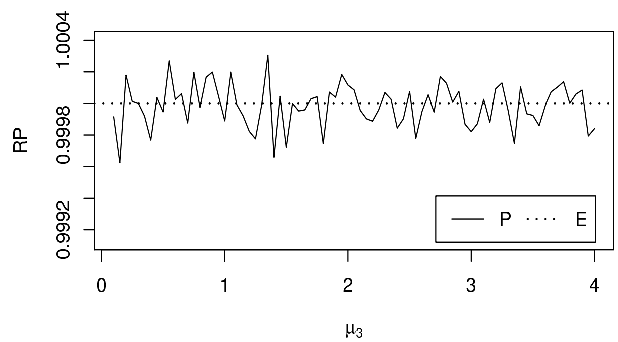


(c)$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$

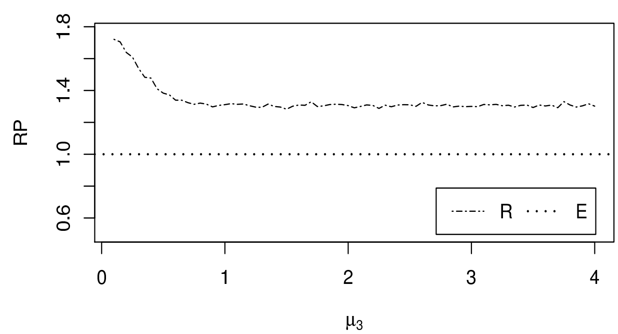
FIGURE 1. Comparisons between $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_P$ (indicated by P) and $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_R$ (indicated by R) for $(n_1, n_2, n_3) = (25, 30, 45)$ and $(\sigma_1, \sigma_2, \sigma_3) = (2, 4, 3)$ with $\mu_1 = 2.4, \mu_2 = 2.5$ and $\mu_3 \in [2.5, 4]$.



(a)$CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$ and $CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$

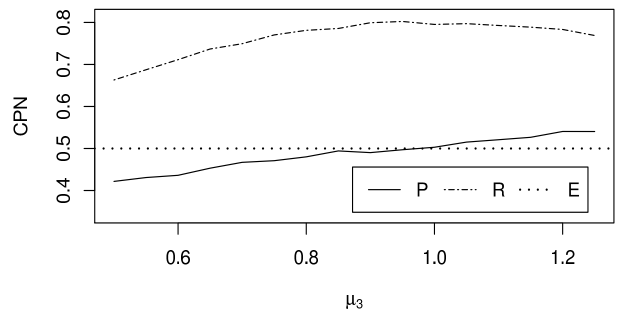


(b)$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$

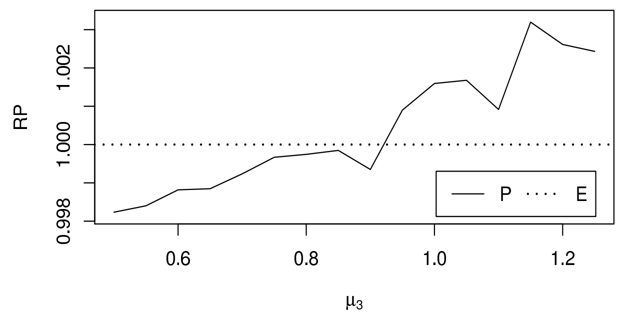


(c)$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$

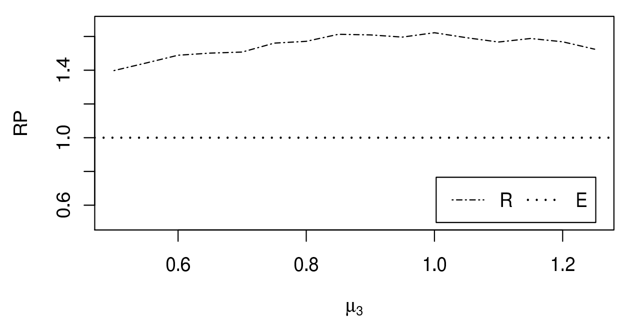
FIGURE 2. Comparisons between $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_P$ (indicated by P) and $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_R$ (indicated by R) for $(n_1,n_2,n_3)=(30,35,40)$ and $(\sigma_1,\sigma_2,\sigma_3)=(2,1.5,1)$ with $\mu_1=0,\mu_2=0.1$ and $\mu_3\in[0.1,4]$.



(a)$CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$ and $CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$

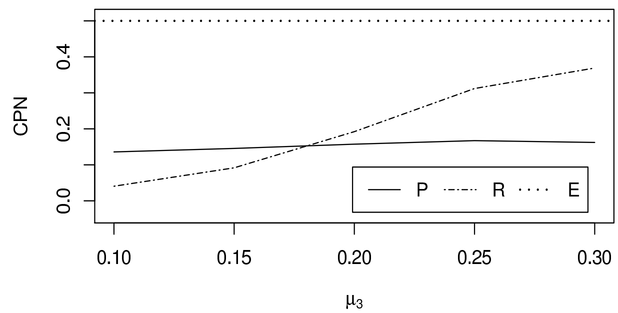


(b)$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$

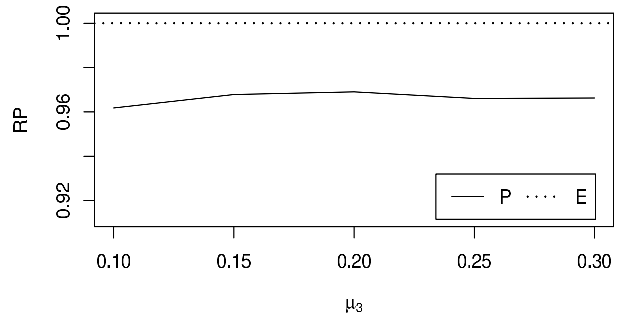


(c)$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$

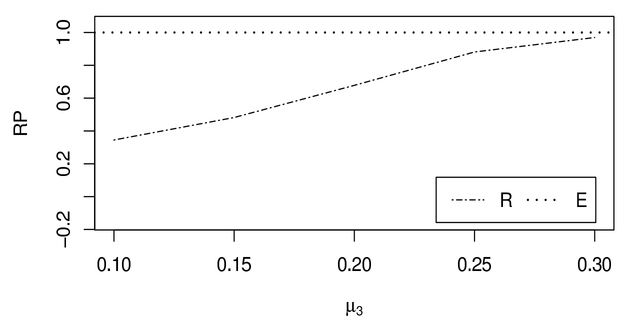
FIGURE 3. Comparisons between $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_P$ (indicated by P) and $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_R$ (indicated by R) for $(n_1, n_2, n_3, n_4) = (50, 50, 50, 50)$ and $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = (6, 5, 4, 3)$ with $\mu_1 = 0, \mu_2 = 0.5, \mu_4 = 1.25$ and $\mu_3 \in [0.5, 1.25]$.



(a)$CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$ and $CPN_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$



(b)$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_P)$



(c)$RP_{\boldsymbol{\mu}}(\widehat{\boldsymbol{\mu}}_M/\widehat{\boldsymbol{\mu}}_R)$

FIGURE 4. Comparisons between $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_P$ (indicated by P) and $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_R$ (indicated by R) for $(n_1,n_2,n_3,n_4)=(15,20,25,30)$ and $(\sigma_1,\sigma_2,\sigma_3,\sigma_4)=(0.2,0.1,0.4,0.3)$ with $\mu_1=0,\mu_2=0.1,\mu_4=0.35$ and $\mu_3\in[0.1,0.3]$.

Remarks: A close inspection of Tables 2 and 3 and Figures 1, 2, 3 and 4 reveal that none of the estimators $\widehat{\boldsymbol{\mu}}_P$ and $\widehat{\boldsymbol{\mu}}_M$ is uniformly better than $\widehat{\boldsymbol{\mu}}_R$. Naturally, there are situations, where $\widehat{\boldsymbol{\mu}}_M$ outperforms $\widehat{\boldsymbol{\mu}}_P$ and $\widehat{\boldsymbol{\mu}}_R$ and vice versa. It is interesting to observe that $\widehat{\boldsymbol{\mu}}_R$, the MVUE is not uniformly better than $\widehat{\boldsymbol{\mu}}_M$. This is not unusual as the derivation of MVUE does not take into account any restriction among the parameters. Again, $\widehat{\boldsymbol{\mu}}_P$ is derived by plugging in the usual unbiased estimators of variance. But under order restriction, variance estimators vary depending on the ordering of the estimated mean. Although empirical evidence does not reveal uniform superiority of $\widehat{\boldsymbol{\mu}}_M$, it takes into account both order restrictions and variance estimators and hence are more sensible than $\widehat{\boldsymbol{\mu}}_P$ and $\widehat{\boldsymbol{\mu}}_R$.

4. Illustration with real data

In order to investigate the applicability of the proposed estimator, we consider the already introduced (see section 1) example of Shi(2014) and calculate the relevant estimates. In particular, we calculate $\widehat{\boldsymbol{\mu}}_M, \widehat{\boldsymbol{\mu}}_P$ and $\widehat{\boldsymbol{\mu}}_R$ together with the respective variance estimates $\widehat{\boldsymbol{\sigma}}^2_M, \widehat{\boldsymbol{\sigma}}^2_P$ and $\widehat{\boldsymbol{\sigma}}^2_R$, for the above mentioned data and provide them in Table 4. The figures of Table 4 show that $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_P$ are more or less the same but $\widehat{\boldsymbol{\mu}}_R$ is slightly different from these. This is quite expected as $\widehat{\boldsymbol{\mu}}_R$ does not take into account any order restriction. However, the sample sizes (i.e. 100) are quite large and hence, we get similar figures for $\widehat{\boldsymbol{\mu}}_M$ and $\widehat{\boldsymbol{\mu}}_P$.

TABLE 4. Estimates for the examination data od Shi (2004).

Estimates	District 1	District 2	District 3	District 4	District 5
$\widehat{\boldsymbol{\mu}}_M$	386.7020	386.7020	396.4612	396.4612	418.0100
$\widehat{\boldsymbol{\sigma}}^2_M$	4016.376	5358.815	4271.748	3584.488	4928.749
$\widehat{\boldsymbol{\mu}}_P$	386.7019	386.7019	396.4613	396.4613	418.0100
$\widehat{\boldsymbol{\sigma}}^2_P$	4016.376	5358.814	4271.748	3584.488	4928.749
$\widehat{\boldsymbol{\mu}}_R$	388.2700	384.6100	398.0000	395.1700	418.0100
$\widehat{\boldsymbol{\sigma}}^2_R$	4013.917	5354.438	4269.380	3582.821	4928.749

5. Concluding remarks

Simultaneous estimation of normal means under order restriction for several independent univariate populations is considered both theoretically and numerically. However, there are several types of restrictions apart from the simple order (for example, simple tree order). Naturally, the analytic derivation of exact maximum likelihood estimator under such orderings is not straightforward. Moreover, involvement of multiple and/or multivariate populations make the situation even more complicated. All these issues, therefore, provide the scope of further development.

Acknowledgements

The authors are grateful to the anonymous referees and the associate editor for providing their helpful comments, which have led to significant improvement over an earlier version.

References

[1] Bazaraa, MS., Sherali, HD. and Shetty, CM. (1993). *Nonlinear Programming: Theory and Algorithms*, Wiley.

[2] 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.

[3] Bhattacharya, CG. et al. (1980). Estimation of a common mean and recovery of interblock information. *The Annals of Statistics*, 8(1), 205-211.

[4] Brown, L. and Cohen, A. (1974). Point and confidence estimation of a common mean and recovery of intrablock information. *The Annals of Statistics*, 2, 963-976.

[5] Chang, YT., Shinozaki, N. (2012). Estimation of ordered means of two normal distributions with ordered variances. *Journal of Mathematics and System Science*, 2(1).

[6] Gupta, RD., Singh, H. (1992). Pitman nearness comparisons of estimators of two ordered normal means. *Austrian Journal of Statistics*, 34(3), 407-414.

[7] Hwang, JG., Peddada, SD. (1994).Confidence interval estimation subject to order restrictions. *The Annals of Statistics*, 67-93.

[8] Keating, JP., Mason, RL. and Sen, PK. (1993). Pitman's Measure of Closeness, A Comparison of Statistical Estimators. *Society for Industrial and Applied Mathematics*.

[9] 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.

[10] Nayak, TK. (1990). Estimation of location and scale parameters using generalized Pitman nearness criterion. *Journal of Statistical Planning and Inference*, 24(2), 259-268

[11] Pedram N, Bazyari, A. (2017). Estimation of order restricted normal means when the variances are unknown and unequal. *Austrian Journal of Statistics*, 14, 3-17.

[12] Robertson, T., Wright, FT. and Dykstra, RL. (1988).*Order Restricted Statistical Inference*. John Wiley, New York.

[13] Shi, N Z (1994). Maximum likelihood estimation of means and variances of normal populations under simultaneous order restrictions. *Journal Of Multivariate Analysis*, 50, 282-293.

[14] Silvapulle, MJ., Sen, PK. (2005). *Constrained Statistical Inference : Inequality, Order and Shape Restrictions.* John Wiley, New York.