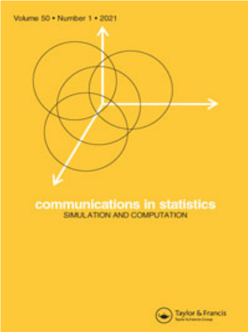

Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Simulation and Computation

ISSN: 0361-0918 (Print) 1532-4141 (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

Estimation of normal means in the tree order model by the weighting methods

Reza Momeni, Javad Etminan & M. Khanjari Sadegh

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
2021, VOL. 50, NO. 1, 282–294
https://doi.org/10.1080/03610918.2018.1554115

Estimation of normal means in the tree order model by the weighting methods

Reza Momeni, Javad Etminan, and M. Khanjari Sadegh

Department of Statistics, University of Birjand, Birjand, Iran

ABSTRACT

Consider $k+1$ independent normal populations with the tree order restriction on the mean parameters. For the tree order model, the restricted estimator of control group parameter is dominated by the unrestricted estimator when the number of treatment groups is large. We discuss two techniques for reducing of mean squared error via to the two weighting methods which are dissimilarity and conditional Bayesian criteria. Based on the bias and mean squared error criteria, the performance of the proposed estimators is compared with the alternative estimators in order to search for a better estimator. Although the superior estimator that uniformly dominates the others does not exist in general, but the proposed estimators dominate the corresponding unrestricted estimator and compete very well with the other alternative estimators introduced by the authors.

ARTICLE HISTORY
Received 26 May 2018
Accepted 11 November 2018

KEYWORDS
Bias; isotonic regression estimator; mean squared error; restricted maximum likelihood estimator; tree order constraint

MATHEMATICS SUBJECT CLASSIFICATION (2010):
62F30

1. Introduction

In many practical situations, statistical inference under order restrictions on the parameters is quite important. Estimation of parameters of the various types subject to order restrictions has received substantial interest during the past several decades. Isotonic regression inference concerns to the order restricted inference and includes situations in which a set of parameters is assumed, a priori to satisfy certain order restrictions. In the most common case, where the data are arranged in ordered groups, the mean value of a random variable is assumed to change monotonically with the complete ordering of the groups or the partial ordering of them. Related works to the statistical inference under order restrictions are reviewed by Barlow et al. (1972) and Robertson, Wright, and Dykstra (1988). Many authors have considered inference of restricted parameters of independent populations and proposed improvements on the usual unbiased estimator, especially for the normal distribution. Recently, one excellent context which serves as major references for this subject is Silvapulle and Sen (2005).

When some additional information regarding to order of the parameters θ_i is available, then for estimating the components of the parameter vector $\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_k)$ the isotonic regression technique can be used. The main advantage of considering order restrictions is that some information can be reflected in the model. For instance, if θ_0 is the yield average of a crop with no fertilizer added and θ_i is the yield average of the

CONTACT Javad Etminan jetminan@birjand.ac.ir Department of Statistics, University of Birjand, Birjand, Iran.

crop when the ith brand of fertilizer is added, then it is reasonable to expect $\theta_0 \leq \theta_i$ for all $i = 1, ..., k$ although one may have no information regarding the relative performance of the various brands $i = 1, ..., k$. Bartholomew (1961) called this order restriction as the tree order constraint. This restriction arises naturally when considering the problem of comparing several treatments to a control where treatments are as effective as the control and is applied in many applications e.g. biomedical and toxicology studies. Therefore, it is reasonable to take into account the order restrictions in making inferences about the group means. We are interested in utilizing the prior knowledge (i.e. tree order) to search for a better estimator with smaller mean squared error than the usual sample mean estimator, component-wise.

The classical approach for the restricted maximum likelihood estimator (RMLE) is the maximization of the likelihood function on the restricted parameter space. When the components of the parameter vector $\boldsymbol{\theta}$ are estimated simultaneously, the RMLE is known to perform better than the unrestricted maximum likelihood estimator (UMLE) (Brunk 1965). Although its aggregate mean squared error is usually less than that of the UMLE, but magnitude of the bias and hence mean squared error of a single component of the RMLE may increase without bound as the dimension k increases. When underlying distributions are independently normally distributed, the RMLE dominates the unrestricted one, under simple order $\theta_0 \leq \theta_1 \leq ... \leq \theta_k$ (Lee 1981; Kelly 1989). But, Lee (1988) recognized a shortcoming of the restricted maximum likelihood estimation for the tree order model. Under the tree order restriction, Lee (1988) and Fernandez, Rueda, and Salvador (1999) established that the RMLE of the control group θ_0 fails to dominate the unrestricted estimator in terms of the mean squared error, especially when k is large.

There are various suggestions for solving this drawback of the restricted maximum likelihood estimator in the tree order restriction. Lee (1988) established that by increasing of the weight of the control sample mean $\bar{X}_0$ the mean squared error of the control group estimator decreases. Hwang and Peddada (1994) studied alternative estimators to the RMLEs in order restricted elliptically models and introduced restricted estimators for some order restriction. They showed that their estimators have higher coverage probability than the unrestricted estimators. Cohen and Sackrowitz (2002), and Betcher and Peddada (2009) found some estimators in the normal models with equal known variances that dominate the corresponding unrestricted estimator as k increases.

In the rest of this paper, in Sec. 2, the restricted estimator in the tree order restriction with its shortcoming and two alternative estimators by authors are represented. Focus of this article is to construct the estimation of the control group parameter in the tree order model. So, in Sec. 3, we introduced two methods for the estimation of the control group parameter in the tree order model. In this section, we transform the tree order problem onto the several simple order restrictions and then introduced the weighted estimators via to the dissimilarity criterion and Bayesian approach. In the first method, we construct the weights by the Euclidean distance between the true order of parameters and the sample order which is dissimilarity criterion among two mentioned orders. In the second method, the posterior probabilities as the weights of any permutation of the simple orders are constructed by using of the Bayes rule. In each method, the resulting quantity is the weighted average of the simple order estimators with control group parameter θ_0 being the smallest parameter as required by the tree order constraint $\theta_0 \leq \theta_i$ for all $i = 1, ..., k$. These methods have higher accuracy in the sense of the least squares error. As in first approach the weights are proportional to the inverse of the sum of squared error in the selected simple orders. In another procedure, the weights are constructed based on the posterior probability of any simple order of parameters, given sample orders. On the other hand, for the construction of the simple orders from tree order, we are not guessing the many unknown inequalities between the various treatment means, when k is very large. We will consider a subset of all simple orders and compute RMLE in any selected simple order, then the weighted estimators by using of these methods are constructed. In Sec. 4, by using a simulation study the performance of the proposed estimators is compared with the UMLE, RMLE and two other estimators which are introduced in the literature. The simulation results show that the proposed estimators dominate the alternative estimators in terms of the mean squared error (MSE) and appear to stabilize in biases, when k is large. Some concluding remarks are given in Sec. 5.

2. Restricted maximum likelihood estimator

Suppose one has independent random samples from $k+1$ normal populations with means $\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_k)$ and a common variance σ^2. For the ith sample, let n_i denote the sample size and let $X_{ij}; i = 0, 1, ..., k; j = 1, ..., n_i$ be mutually independent variables in which $\bar{X}_i = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i}$ is the sample mean in ith sample, $i = 0, 1, ..., k$ which is an unbiased estimator. It is well known that the vector $\bar{\mathbf{X}} = (\bar{X}_0, \bar{X}_1, ..., \bar{X}_k)$ is the unrestricted maximum likelihood estimator (UMLE) of the population mean vector $\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_k)$.

In viewpoint of the geometrically, under the tree order restriction the parameter space for $\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_k)$ forms a symmetric polyhedral cone C in R^{k+1} defined as follows:

$$C = \left\{\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_k) \in R^{k+1} | \theta_0 \leq \theta_i; i = 1, ..., k\right\}. \quad (1)$$

The convex cone C has a linear subspace with spine:

$$L = \left\{\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_k) \in R^{k+1} | \theta_0 = \theta_1 = ... = \theta_k\right\}, \quad (2)$$

that is a one-dimensional line and is well known as the least favorable case at the homogeneity of means.

The classical approach for the restricted maximum likelihood estimator (RMLE) is the maximization of the likelihood function on the restricted parameter space which is obtained by applying the isotonic regression transformation to the unrestricted maximum likelihood estimator(UMLE) $\bar{\mathbf{X}} = (\bar{X}_0, \bar{X}_1, ..., \bar{X}_k)$, which is a continuous function of $\bar{\mathbf{X}}$, i.e. $\hat{\theta}_i^{RMLE} = f_i^{\theta \in C}(\bar{\mathbf{X}})$ (Robertson, Wright, and Dykstra 1988; Silvapulle and Sen 2005).

Definition 2.1. *For a given order restriction on the parameters $\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_k)$, the isotonic regression estimator of $\bar{\mathbf{X}} = (\bar{X}_0, \bar{X}_1, ..., \bar{X}_k)$ with given positive weights $\mathbf{w} = (w_0, ..., w_k)$ is the value of $\boldsymbol{\theta}$ that minimizes the weighted sum of squares:*

$$\sum_{i=0}^{k} (\bar{X}_i - \theta_i)^2 w_i \tag{3}$$

subject to the order restriction C on the parameters that is $\boldsymbol{\theta} \in C$, (Here C is the tree order cone).

For the special case, in the normal populations when $w_i = (\sigma_i^2)^{-1}$, the isotonic regression estimator is the RMLE subject to the tree order restriction C. The RMLE is obtained by the min-max formula in the explicit form that is (Barlow et al. 1972):

$$\hat{\theta}_0^{RMLE} = \min_{S \subseteq K} \frac{\sum_{j\in S} w_j \bar{X}_j}{\sum_{j \in S} w_j} \tag{4}$$

where the minimization is taken over all subsets S of $K = \{0, 1, ..., k\}$ containing element 0.

We describe the restricted maximum likelihood estimator $\hat{\theta}_0^{RMLE}$ as follows. With maintaining all the known inequalities $\theta_0 \le \theta_i$, for $i = 1, 2, ..., k$, we construct all $k!$ simple order restrictions between $\theta_0, \theta_1, ..., \theta_k$ by considering all $k!$ orderings between $\theta_1, ..., \theta_k$. Under each simple order restriction, we obtain the corresponding RMLE. Then $\hat{\theta}_0^{RMLE}$ is the minimum among all such RMLEs. Furthermore, the isotonic regression estimators for the treatment parameters are:

$$\hat{\theta}_i^{RMLE} = \max\left\{\hat{\theta}_0^{RMLE}, \bar{X}_i\right\} \quad for \quad i = 1, ..., k. \tag{5}$$

It is clear that $\hat{\theta}_0^{RMLE} \le \bar{X}_0$ and $\hat{\theta}_i^{RMLE} \ge \bar{X}_i$; for $i \ge 1$. Since these inequalities are strict with positive probabilities, hence the isotonic regression estimators are always biased. So, their efficiency is measured by the mean squared error (MSE) (Chaudhuri and Perlman 2005). On the other hand, Brunk (1965) showed that the total mean squared error of the RMLE is strictly smaller than that of the UMLE, i.e.

$$\sum_{i=0}^{k} E\left(\hat{\theta}_i^{RMLE} - \theta_i\right)^2 w_i < \sum_{i=0}^{k} E(\bar{X}_i - \theta_i)^2 w_i. \tag{6}$$

The estimate $\hat{\theta}^{RMLE} = (\hat{\theta}_0^{RMLE}, \hat{\theta}_1^{RMLE}, ..., \hat{\theta}_k^{RMLE})$ is the vector in the convex cone C which is the closest to $\bar{\mathbf{X}} = (\bar{X}_0, \bar{X}_1, ..., \bar{X}_k)$ in the sense that it minimizes:

$$\sum_{i=0}^{k} (\bar{X}_i - \theta_i)^2 w_i \tag{7}$$

among all $\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_k) \in C$. For positive weights $\mathbf{w} = (w_0, w_1, ..., w_k)$ the vector $\hat{\theta}^{RMLE}$ denote the projection of vector $\bar{\mathbf{X}} = (\bar{X}_0, \bar{X}_1, ..., \bar{X}_k)$ into the convex cone C with Euclidean distance between $\bar{\mathbf{X}}$ and $\hat{\theta}^{RMLE}$ that is:

$$d = \sqrt{\sum_{i=0}^{k} w_i \left(\hat{\theta}_i^{RMLE} - \bar{X}_i\right)^2} \tag{8}$$

The smaller value of d is equivalent to the higher accuracy of the restricted estimators. So, the choice of d^{-1} where d is dissimilarity criterion, as the suitable weight of the proposed method can be substantial increases the precision of the new estimator in the next section. Because of Eq. (4), $\hat{\theta}_0^{RMLE}$ is increasing in w_0 and the magnitude of the bias of $\hat{\theta}_0^{RMLE}$ is decreasing in w_0. But, Lee (1988) in Theorem 2.1 shows that the MSE of $\hat{\theta}_0^{RMLE}$ is unbounded, when k is very large. We present this theorem in the following.

Theorem 2.2. Lee (1988): *If the means $\theta_0, \theta_1, ..., \theta_k$ and the sample sizes $n_0, n_1, ..., n_k$ are bounded in the tree order normal models, then for sufficiently large k,*

$$E(\bar{X}_0 - \theta_0)^2 < E\left(\hat{\theta}_0^{RMLE} - \theta_0\right)^2. \tag{9}$$

Ever since Lee (1981) showed that the simple order estimators dominate the corresponding UMLEs if tree order restriction is replaced by the simple order. The reverse inequality (9) under the tree ordering is the first counterexample in the literature. Nevertheless, to overcome of this difficulty in (9), Lee (1988) demonstrated that by increasing of the control group weight the MSE reduction can be achieved. Hwang and Peddada (1994) established a similar phenomenon in terms of the coverage probability of any fixed width confidence interval of θ_0 centered at $\hat{\theta}_0^{RMLE}$. So, the RMLE of the control group mean under the tree order restrictions may perform poorly, in terms of the MSE and coverage probability. Therefore, Hwang and Peddada (1994) introduced a new estimation procedure for estimating of parameters in the elliptically unimodal distributions (e.g. normal and t-student distributions) subject to the tree order restriction. They chose one of the orderings rather than considering all possible $k!$ orderings between $\theta_1, ..., \theta_k$ arbitrarily, e.g. $\theta_0 \leq \theta_1 \leq ... \leq \theta_k$. They then introduced the following estimators:

$$\hat{\theta}_0^{HP} = \min_{0 \leq t} \left\{ \frac{\sum_{j=0}^{t} w_j \bar{X}_j}{\sum_{j=0}^{t} w_j} \right\}; \qquad \hat{\theta}_i^{HP} = \max\left\{\hat{\theta}_0^{HP}, \bar{X}_i\right\}; \quad i = 1, ..., k. \tag{10}$$

But their estimators depend on the chosen simple order between treatment groups and hence is not uniquely defined. Also, Tan and Peddada (2000) observed that for non-diagonal covariance matrix these estimators may be inconsistent, especially for $k = 2$. For this reason, Dunbar, Conaway, and Peddada (2001) proposed a modification to the Hwang and Peddadas procedure. As in the case of the RMLE, Dunbar, Conaway, and Peddada (2001) considered all possible $k!$ orderings of the parameters $\theta_1, \theta_2, ..., \theta_k$ with each ordering resulting in a simple order restriction between $\theta_0, \theta_1, ..., \theta_k$. They then estimated θ_0 by taking the mean of all such RMLEs of θ_0. Dunbar, Conaway, and Peddada (2001) also proved that their estimator has a smaller risk than the UMLE in terms of all convex monotone functions of the absolute loss function. This result holds true even if it is taken a subset of $k!$ orderings of the treatment parameters $\theta_1, \theta_2, ..., \theta_k$ rather than all $k!$ orderings. For this reason Cohen and Sackrowitz (2002) stated that the RMLE is undesirable and proposed an alternative estimator whose bias remains bounded. So, independently of the other works, the Cohen and Sackrowitz (2002) proposed a method for restrictive parameters in the tree order model that is:

$$\hat{\theta}_0^{CS} = \frac{\sum_{i=0}^{k} n_i \min(\bar{X}_0, \bar{X}_i)}{\sum_{i=0}^{k} n_i}, \tag{11}$$

$$\hat{\theta}_i^{CS} = \hat{\theta}_0^{CS} + (\bar{X}_i - \bar{X}_0)^+; \quad i = 1, 2, ..., k, \tag{12}$$

where $a^+ = \max(0, a)$. Like the RMLE, the magnitude of the bias of Cohen and Sackrowitz estimator (CSE) is again greatest under the equality of population parameters, which is the least favorable case for the bias of these estimators i.e. $\boldsymbol{\theta} \in L$.

The CS estimator $\hat{\theta}_0^{CS}$ was motivated by the fact that each $M_i = \min(\bar{X}_0, \bar{X}_i)$ has lower MSE for θ_0 than $\bar{X}_0$ itself (Chaudhuri and Drton (2003)). An alternative pair-wise estimator with these properties is the RMLE $\hat{\theta}_0^{RMLE}(\bar{X}_0, \bar{X}_i)$ based on the paired variable $(\bar{X}_0, \bar{X}_i)$ for $i = 1, ..., k$. Since each $\hat{\theta}_0^{RMLE}(\bar{X}_0, \bar{X}_i)$ also dominates the UMLE $\bar{X}_0$ as an estimator of θ_0. Hence, Betcher and Peddada (2009) by using of this principle, introduced a modification of the RMLE for the tree order restriction as follows:

$$\hat{\theta}_0^{BP} = \frac{\sum_{j=0}^{k} a_j \hat{\theta}_0^{RMLE}(\bar{X}_0, \bar{X}_i)}{\sum_{j=0}^{k} a_j}; \qquad \hat{\theta}_i^{BP} = \max\left\{\hat{\theta}_0^{BP}, \hat{\theta}_0^{RMLE}\right\}, \quad for \quad i \geq 1. \tag{13}$$

where a_j is the inverse of the variance of UMLE $\bar{X}_j$ and $\hat{\theta}_0^{RMLE}(\bar{X}_0, \bar{X}_i)$ is the RMLE of two linked parameters (θ_0, θ_j), that is:

$$\hat{\theta}_0^{RMLE}(\bar{X}_0, \bar{X}_i) = \begin{cases} \bar{X}_0 & if \quad \bar{X}_0 < \bar{X}_j \\ \dfrac{a_0 \bar{X}_0 + a_j \bar{X}_j}{a_0 + a_j} & if \quad \bar{X}_0 \geq \bar{X}_j \end{cases}$$

Based on the Theorem 6 in Chaudhuri and Drton (2003), the bias of $\hat{\theta}_0^{RMLE}(\bar{X}_0, \bar{X}_i)$ is substantially smaller than that of $M_i = \min(\bar{X}_0, \bar{X}_i)$ when $\boldsymbol{\theta} \in L$. Thus the pairwise RMLEs are preferable to the pairwise minima in the least favorable case that is $\boldsymbol{\theta} \in L$ (Chaudhuri and Drton (2003)). Therefore, the Betcher and Peddada estimator (BPE) dominates the CSE in terms of the bias. The BPE only for the case of equal known normal variances is constructed. Hence in the general case, the RMLE for the target parameter can be adjusted to reduce its bias, and hence decrease the MSE.

3. The proposed procedures

In this section we discuss two methods of calculating estimators for tree order parameters which are suitable weighting methods by using of the dissimilarity criterion and conditional Bayesian principal.

The control group parameter θ_0 in the tree order restriction is said the nodal parameter, because the inequality between θ_0 and every treatment parameter θ_i is known a priori. We first estimate the only nodal parameter θ_0 in the tree order constrained which is the smallest location parameter in the normal mean parameters. It will be convenient to consider a slightly more general setup in the previous section.

By Eq. (4), $\hat{\theta}_0^{\text{RMLE}}$ equals to $\bar{X}_0$ if the treatment means $\bar{X}_i, i = 1, ..., k$ are at least as large as $\bar{X}_0$, else computed $\frac{w_0 \bar{X}_0 + w_{(1)} \bar{X}_{(1)}}{w_0 + w_{(1)}}$ where $\bar{X}_{(1)} = \min\{\bar{X}_i; i \geq 1\}$. Now, if the computed value is smaller than the other treatment groups, then $\hat{\theta}_0^{RMLE}$ equals to the

$\frac{w_0\bar{X}_0+w_{(1)}\bar{X}_{(1)}}{w_0+w_{(1)}}$, otherwise we then compute the new weighted average of the $\bar{X}_0$ and two first order statistics of treatment means and so on.

For instance, suppose $k=4$ and the order of sample means is $\bar{X}_{(1)} \leq \bar{X}_{(2)} \leq \bar{X}_{(3)} \leq \bar{X}_0 \leq \bar{X}_{(4)}$. For notational simplicity, assume that $w_i = 1$. Here, in order to get the RMLE $\hat{\theta}_0^{RMLE}$, we have some reasonable facts of this estimator that are:

1. Smallest order statistics $\bar{X}_{(1)}$ is involved in the computation of $\hat{\theta}_0^{RMLE}$, certainly.
2. The highest order statistics $\bar{X}_{(4)}$ is not considered in the computation of $\hat{\theta}_0^{RMLE}$.
3. The first considerable value for $\hat{\theta}_0^{RMLE}$ equals to $\frac{\bar{X}_0+\bar{X}_{(1)}}{2}$ and the last feasible value is $\frac{\bar{X}_0+\bar{X}_{(1)}+\bar{X}_{(2)}+\bar{X}_{(3)}}{4}$, that is the worst case of the sample order.

Since $\bar{X}_0$ is an unbiased estimator of θ_0, it follows that $\hat{\theta}_0^{RMLE}$ tends to underestimate θ_0, so it has a negatively biased. By increasing of k (number of treatment groups) the magnitude of the bias of $\hat{\theta}_0^{RMLE}$ is unbounded.

Hence, rather than using Hwang and Peddada estimator (HPE) $\hat{\theta}_0^{HP}$ from one arbitrary choice of the inequalities between $\theta_1, \ldots, \theta_k$, we consider all $k!$ simple orderings and compute the weighted average of the resulting RMLEs of θ_0 under all $k!$ simple order restrictions. Rather than considering all possible simple orderings between $\theta_1, \ldots, \theta_k$, we select any subset of the simple orderings arbitrarily, when k be very large i.e. $k \geq 10$. Although the number of treatments k is usually small for practical reasons, frequently $k \in \{2, 3, 4\}$ (Bretz and Hothorn (2003)). More complex designs, for example, those including an additional treatment groups are possible, but will not be considered in this paper. So, we do not require to take a subset of the simple orders from permutations of the treatment means $\theta_1, \ldots, \theta_k$.

3.1. Weighted method by dissimilarity criterion

In this procedure, we choose the dissimilarity criterion between the observations and corresponding order as the weighting of the preliminary simple estimators. Dissimilarity is the relevant notion, which defines a positive real value function of two variables or estimators. For instance, a zero value of this criterion, means that two objects are similar and a large value implies a high dissimilarity. We construct the distance between the observing of the sample order and the true order restriction between the parameters as the dissimilarity criterion. Whatever, distance between sample order and the corresponding true order among the parameters be higher, the affected of the simple order estimator by these observations is lower. Hence, we derive the weighted estimator of the control group parameter based on the dissimilarity criterion in the following:

$$\hat{\theta}_0^{WA1} = \frac{\sum_{j=1}^{k!} \frac{\hat{\theta}_0^{(j)}}{d_j}}{\sum_{j=1}^{k!} \left(\frac{1}{d_j}\right)}, \tag{14}$$

where $\hat{\theta}_0^{(j)}$ is the control group estimator in the jth $(j = 1, \ldots, k!)$ selected simple order. For example, if the jth simple order be $\theta_0 \leq \theta_1 \leq \ldots \leq \theta_k$ we have:

$$\hat{\theta}_i^{(j)} = \min_{t \geq i} \max_{s \leq i} \frac{\sum_{l=s}^{t} w_l \bar{X}_l}{\sum_{l=s}^{t} w_l}; \quad for \quad i = 0, 1, ..., k, \tag{15}$$

and for $j = 1, ..., k!$ simple orders. The weights $\mathbf{d} = (d_1, d_2, ..., d_{k!})$ derived based on the Euclidean distance in each $k!$ simple order which is defined in (8). So, for jth simple order the weight d_j is given by:

$$d_j = \sqrt{\sum_{i=0}^{k} w_i \left(\hat{\theta}_i^{(j)} - \bar{X}_i\right)^2}; \quad for \quad j = 1, 2, ..., k!, \tag{16}$$

where $\mathbf{w} = (w_0, w_1, ..., w_k)$ is same as the given positive weights in the definition of the isotonic regression estimator in (3). The weight d_j in (16) means that whatever the distance between the RMLE and corresponding UMLE is higher in certain simple order, the effect of the RMLE extracted from corresponding simple order is lower on the evaluation of the proposed estimator $\hat{\theta}_0^{WA1}$.

Using (14), we then provide general strategies to estimate the non-nodal treatment parameters that are:

$$\hat{\theta}_i^{WA1} = \max\left\{\hat{\theta}_0^{WA1}, \bar{X}_i\right\}; \quad for \quad i = 1, ..., k. \tag{17}$$

3.2. Weighted method by Bayesian approach

In this subsection, an alternative method of involving a prior distribution, $\pi(\boldsymbol{\theta})$, in view point of the Bayesian principle is presented. To estimate the parameters in the tree order cone C, we apply a conditional approach for estimating population means satisfying $k!$ simple order restriction extracted from the tree order constraint.

In this conditional approach, the tree order restriction is considered to be $k!$ permutation of simple orders whose these variations can be described by a prior distribution. The prior distribution is updated with the use of Bayes rule, and hence the posterior distribution is constructed.

The tree order constraint on the k treatment means with a control can be expressed as the $k!$ of simple orderings among of the treatment means, $\theta_1, ..., \theta_k$. Thus, we choose the uniform non-informative prior distribution for $\theta_1, \theta_2, ..., \theta_k$ that is:

$$\pi(\theta_1, \theta_2, ..., \theta_k) = 1$$

where all $k!$ permutations of the simple orders of $\theta_1, ..., \theta_k$ are equally likely with proportion of $1/k!$. By using of the Bayes rule, the joint posterior distribution of $\theta_0, \theta_1, ..., \theta_k$ is as follows:

$$\pi(\theta_0, \theta_1, ..., \theta_k | \bar{\mathbf{x}}) \propto N\left(\bar{x}_0, \sigma_{\bar{x}_0}^2\right). \prod_{i=1}^{k} TN_{\theta_0^-}\left(\bar{x}_i, \sigma_{\bar{x}_i}^2\right) \tag{18}$$

where $TN_{\theta_0^-}(\bar{x}_i, \sigma_{\bar{x}_i}^2)$ is the truncated normal distribution with left truncation point θ_0^-. Therefore, these posterior probabilities are proportional to $N(\bar{x}_i, \sigma_{\bar{x}_i}^2)$ subject to the tree order restriction. By considering all $k!$ simple orders between $\theta_1, ..., \theta_k$, we estimate the posterior probabilities for each simple order.

Note that the posterior distribution is a conditional distribution upon the sample observations. The posterior probabilities of the order of populations are now used to make weightings about $k!$ simple orders and make evidence about order of parameters $\theta_1, \ldots, \theta_k$. Thus, by maintaining all the known inequalities $\theta_0 \leq \theta_i$, for $1 \leq i \leq k$, if the posterior probability of an arbitrary simple order i.e. $\pi(\theta_0 \leq \theta_1 \leq \ldots \leq \theta_k|\bar{\mathbf{x}})$ be large, then it is more likely to be the true order of population than another simple order and therefore the corresponding weight will be large. Thus, the weighted average is constructed based on the posterior probabilities as the second proposed weighting method.

Moreover, it would be a more plausible occurrence if corresponding posterior probability is large. Therefore, by using of these posterior probabilities as the weights, we propose the weighted estimator for the control group θ_0 in the following:

$$\hat{\theta}_0^{WA2} = \sum_{j=1}^{k!} \pi\left(\theta_1^{(j)} \leq \theta_2^{(j)} \leq \ldots \leq \theta_k^{(j)}|\bar{\mathbf{x}}\right)\hat{\theta}_0^{(j)} \tag{19}$$

where $\hat{\theta}_0^{(j)}$ is the RMLE of the smallest parameter (i.e. control group) in jth simple order, Eq. (15), and $\pi(\theta_1^{(j)} \leq \theta_2^{(j)} \leq \ldots, \leq \theta_k^{(j)}|\bar{\mathbf{x}})$ is the posterior probability of the jth simple order, where j is the permutation of the treatment groups $\{1, 2, \ldots, k\}$. Whatever, the posterior probability in this estimator is higher, we allocate higher weight to the preliminary estimator of the corresponding simple order. Hence, this simple order is affected than the others in the estimating procedure.

3.3. Dominance of two weighted estimators over UMLE

The two weighted proposed estimators of $\theta_0\left(\hat{\theta}_0^{WA1} \; and \; \hat{\theta}_0^{WA2}\right)$ dominate the corresponding UMLE $\bar{X}_0$ in the tree order restriction. We demonstrate this domination in Theorem 3.1 for one estimator e.g. $\hat{\theta}_0^{WA}$. Without loss of generality, we may assume that $w_j = 1$, i.e. the worse case.

Theorem 3.1. *Suppose that $\hat{\theta}_0^{(j)}$ is the RMLE of the smallest parameter in the jth simple order constructed between $\theta_1, \ldots, \theta_k$, in which θ_0 satisfied in the tree order restriction $\theta_0 \leq \theta_i; i = 1, \ldots, k$. If $\hat{\theta}_0^{WA} = \frac{\sum_{j=1}^{k!} \hat{\theta}_0^{(j)}}{k!}$ be a weighted estimator, then for all non-decreasing convex function ϕ we have:*

$$E\left\{\phi\left(\left|\hat{\theta}_0^{WA} - \hat{\theta}_0\right|\right)\right\} \leq E\{\phi(|\bar{X}_0 - \theta_0|)\} \tag{20}$$

Proof. Since ϕ is a non-decreasing function, by triangle inequality we have:

$$E\left\{\phi\left(\hat{\theta}_0^{WA} - \hat{\theta}_0|\right)\right\} = E\left\{\phi\left(\left|\frac{1}{k!}\sum_{j=0}^{k!}\hat{\theta}_0^{(j)} - \theta_0\right|\right)\right\} \leq E\left\{\phi\left(\frac{1}{k!}\sum_{j=0}^{k!}\left|\hat{\theta}_0^{(j)} - \theta_0\right|\right)\right\}$$

Since ϕ is a convex function, so by Jensen's inequality

$$E\left\{\phi\left(\frac{1}{k!}\sum_{j=0}^{k!}\left|\hat{\theta}_0^{(j)} - \theta_0\right|\right)\right\} \leq \frac{1}{k!}\sum_{j=0}^{k!} E\left\{\phi\left(\left|\hat{\theta}_0^{(j)} - \theta_0\right|\right)\right\} \tag{21}$$

Table 1. MSE and Bias values for RMLE, CSE, BPE, $\hat{\theta}_0^{WA1}$ and $\hat{\theta}_0^{WA2}$ estimators, least favorable case $\boldsymbol{\theta} = (0, 0, ..., 0)$.

	MSE					Bias				
k	RMLE	CSE	BPE	WA1	WA2	RMLE	CSE	BPE	WA1	WA2
1	0.733	0.733	0.733	0.733	0.733	−0.280	−0.280	−0.280	−0.280	−0.280
2	0.745	0.705	0.674	0.703	0.727	−0.447	−0.378	−0.331	−0.385	−0.419
3	0.786	0.680	0.610	0.676	0.702	−0.571	−0.436	−0.352	−0.446	−0.481
4	0.820	0.653	0.553	0.642	0.673	−0.649	−0.456	−0.347	−0.465	−0.504
5	0.875	0.642	0.514	0.626	0.661	−0.715	−0.469	−0.334	−0.477	−0.517
6	0.923	0.627	0.489	0.616	0.647	−0.769	−0.475	−0.336	−0.491	−0.528
7	1.004	0.642	0.482	0.627	0.658	−0.831	−0.497	−0.342	−0.511	−0.547
8	1.044	0.619	0.442	0.605	0.635	−0.866	−0.489	−0.325	−0.506	−0.542
9	1.103	0.617	0.418	0.603	0.632	−0.910	−0.503	−0.326	−0.520	−0.554
10	1.157	0.626	0.415	0.611	0.637	−0.947	−0.516	−0.332	−0.533	−0.564

From Kelly (1989) for each simple order j we know that,

$$E\left\{\phi\left(\left|\hat{\theta}_0^{(j)} - \theta_0\right|\right)\right\} \leq E\{\phi(|\bar{X}_0 - \theta_0|)\} \tag{22}$$

with a strict inequality for at least one value of the parameter. Now, if two estimators $\hat{\theta}_0^{(1)}$ and $\hat{\theta}_0^{(2)}$ dominate $\bar{X}_0$ for estimating θ_0 in terms of all non-decreasing convex functions of the absolute loss function, then the weighted average of estimators, $\frac{w_1\hat{\theta}_0^{(1)} + w_2\hat{\theta}_0^{(2)}}{w_1 + w_2}$, will dominate $\bar{X}_0$ (Dunbar, Conaway, and Peddada 2001). Therefore, by using of inequalities (21) and (22) we deduce that

$$E\left\{\phi\left(\left|\frac{1}{k!}\sum_{j=0}^{k!}\hat{\theta}_0^{(j)} - \theta_0\right|\right)\right\} \leq E\{\phi(|\bar{X}_0 - \theta_0|)\}$$

Thus, for any convex non-decreasing function ϕ we have

$$E\left\{\phi\left(\left|\hat{\theta}_0^{WA} - \hat{\theta}_0\right|\right)\right\} \leq E\{\phi(|\bar{X}_0 - \theta_0|)\}$$

Also, if k be very large, each obtained simple order estimator from a subset of these simple orders (i.e. $k' \leq k!$) dominates the corresponding sample mean $\bar{X}_0$ (Dunbar, Conaway, and Peddada 2001).

Hence, the two weighted estimators $\hat{\theta}_0^{WA1}$ and $\hat{\theta}_0^{WA2}$ analogous to Theorem 3.1, improve upon the corresponding unrestricted sample mean estimator $\bar{X}_0$ in terms of all nondecreasing convex functions of the absolute loss. This is a very strong property for an estimator to have and includes the MSE criterion. All results about the domination of our proposed estimators are represented in the next section.

4. Simulation study

We conducted a simulation study to evaluate the performance of the two proposed methods with the 1) UMLE 2) RMLE 3) CSE 4) BPE in the tree order normal means. Because of the BPE dominates the HPE (Betcher and Peddada 2009), so the HPE exclude from this comparison and just dominators are presented. We generated

Table 2. MSE and Bias values for RMLE, CSE, BPE, $\hat{\theta}_0^{WA1}$ and $\hat{\theta}_0^{WA2}$ estimators, interior case $\boldsymbol{\theta} = (0, 1, ..., 1)$.

	MSE					Bias				
k	RMLE	CSE	BPE	WA1	WA2	RMLE	CSE	BPE	WA1	WA2
1	0.821	0.821	0.821	0.821	0.821	−0.102	−0.102	−0.102	−0.102	−0.102
2	0.733	0.757	0.779	0.754	0.742	−0.163	−0.125	−0.082	−0.125	−0.147
3	0.694	0.736	0.779	0.732	0.720	−0.223	−0.151	−0.066	−0.148	−0.168
4	0.664	0.708	0.757	0.704	0.690	−0.268	−0.162	−0.043	−0.154	−0.178
5	0.651	0.706	0.772	0.702	0.688	−0.307	−0.171	−0.022	−0.161	−0.185
6	0.639	0.697	0.770	0.691	0.676	−0.332	−0.165	−0.007	−0.155	−0.178
7	0.617	0.677	0.745	0.670	0.655	−0.355	−0.161	−0.033	−0.150	−0.173
8	0.645	0.694	0.761	0.685	0.672	−0.407	−0.190	−0.019	−0.179	−0.201
9	0.623	0.673	0.745	0.665	0.650	−0.413	−0.172	−0.055	−0.162	−0.183
10	0.643	0.685	0.751	0.676	0.663	−0.450	−0.193	−0.047	−0.183	−0.203

observations from k normal distributions with mean parameters $\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_k) \in C$ and equal population variances $\sigma^2 = 1$ across groups. We consider two different configurations of the normal means that is the least favorable case $\boldsymbol{\theta} = (0, 0, ..., 0)$ i.e. in this configuration the magnitude of the biases is greatest and interior case $\boldsymbol{\theta} = (0, 1, ..., 1)$ with the small sample sizes of $n_i = 1, i = 0, 1, ..., k$ subject per groups. All simulation results are based on 50000 simulation runs by increasing of the number of populations $k = 1$ up to 10. The MSE and bias values for estimators in the least favorable and interior case are represented in the Tables 1 and 2, respectively. For computation of the posterior probabilities as the weights in the second method i.e. $\hat{\theta}_0^{WA2}$, we consider all $k!$ permutation of simple orders and then generated $\theta_0 \sim N(x_0, \sigma^2)$, then for $i = 1, 2, ..., k$, we generate $\theta_i \sim TN_{\theta_0^-}(x_i, \sigma^2)$ i.e. the truncated normal distribution with left truncation point θ_0^-. The proportion of each permutation of simple orders from these generated data to the total number of them is the posterior probability as the weight in the second proposed estimator $\hat{\theta}_0^{WA2}$.

In Table 1, as k increases, the RMLE performs poorly in terms of the MSE and bias. In several cases, the MSE of the RMLE is substantially larger than all its competitors. Instead, the proposed estimators of the control group θ_0 enjoy lower MSE than the RMLE, CSE and BPE. As k increases, the MSE values for two proposed estimators appear to stabilize or decrease. Also, the bias values of the proposed estimator are stabilized for higher values of k and these values are lower than that of the RMLE. In the least favorable case, the growth rate of the bias of BPE and latter WA1 estimator are smaller than those of the rest estimators. It might be expected that from the discussion in preceding section, the bias of the BPE is less than that of the CSE. This can be explained by the fact that of the bias of each pair-wise RMLE is less than that of the bias of each pair-wise minima. By comparison of the MSEs and biases, it is seen that the two weighting methods for estimating the control parameter θ_0 are desirable alternatives to the RMLE, and compete very well with the CSE and BPE, especially in terms of the MSE criterion.

By viewing of Table 2, for the interior case of parameter space $\boldsymbol{\theta} = (0, 1, ..., 1)$, the MSE values of the proposed estimators for the control group mean are lower than that of the competitor estimators. When the vector of population means lies in the interior case, the proposed estimators and RMLE are preferable for small values of k, while our estimators offer improvements for larger values of k. In all cases of Table 2, the proposed estimators improve upon the competitor estimators CSE and BPE in terms of the MSEs and biases.

We note that the RMLE never performed better than the two proposed estimators. In most cases, we observe that the proposed estimators achieve a substantial reduction in the MSE compared to the alternative estimators, especially for the weighted estimator when the dissimilarity weights are used i.e. $\hat{\theta}_0^{WA1}$. On the other hand, in all various patterns the MSE values of the proposed estimators are always lower than those of the unbiased UMLE which are equal to their variances and in this simulation they are set to be equal 1. Furthermore, the biases of the proposed estimators and CSE perform similarly for the least favorable case. Relative to the RMLE, the weighted proposed estimators perform well for all patterns of simulation. In most cases considered in our simulation study, the gain in the MSE of $\hat{\theta}_0^{WA1}$ is substantial. Also, the first proposed estimator $\hat{\theta}_0^{WA1}$ based on the dissimilarity weights improves upon $\hat{\theta}_0^{WA2}$ which is based on the posterior probability weights, but this superiority is negligible. Since proposed estimators dominate the BPE in terms of the MSE, these estimators dominate the HPE based on the transition property. Thus we conclude in overall that performances of the proposed estimators are superior in comparison of the relevant estimators.

5. Concluding remarks

In this article, we considered an experimental situation in which one wishes to compare several treatments with a control when it is believed a priori that all of the treatments are as effective as the control. Therefore, we had the problem of estimation of parameters in the tree order restriction under normality models with common variance, in the presence of an increasing number of treatment parameters $\theta_1, \ldots, \theta_k$. The research of Lee (1988) demonstrated that the RMLE of θ_0 by increasing of the number of treatment populations k, fails disastrously in terms of the MSE and solved this problem by increasing of the weight of the control group w_0. Because of the computational difficulties inherent in the RMLEs as expressed in Sec. 2, and also counterexample represented by Lee (1988), other estimators have been proposed by authors, recently. In some conditions, these estimators are more efficient than the corresponding RMLE and for some patterns are inconsistent (Chung and Shinozaki 2012). Therefore, depending on the configuration of the population means, their estimators can outperform the RMLE in certain regions of the tree order space.

In the present paper, we proposed two modified estimators in the tree order constraint and compared control group estimators with the corresponding UMLE, RMLE and other estimators that are listed in the literature. We showed that under squared error loss function, the proposed estimators of the control group parameter θ_0 have a smaller MSE than the UMLE and RMLE, especially when the number of populations k is large. Thus, two proposed estimators do not have the drawbacks of the RMLE.

On the other hand, the proposed estimators via to the weighting methods compete very well with the CSE and BPE procedures. If the true parameters are in the interior parameter space then the proposed estimators perform substantially better than the CSE and BPE in almost every situation. Of course, for least favorable case these dominations are true, in terms of the MSE. Also, unlike the RMLE of θ_0 under the tree order restriction, these estimators do not fail for large k in terms of the bias and MSE.

Acknowledgments

The authors thank the referee for carefully reading this manuscript and giving the useful comments that improved the presentation of the manuscript.

References

Barlow, R. E., D. J. Bartholomew, J. M. Bremner, and H. D. Brunk. 1972. *Statistical inference under order restrictions: the theory and application of isotonic regression*. New York: Wiley.

Bartholomew, D. J. 1961. A test of homogeneity of means under restricted alternatives. *Journal of the Royal Statistical Society, B* 23:239–81.

Betcher, J., and S. D. Peddada. 2009. Statistical inference under order restrictions in analysis of covariance using a modified restricted maximum likelihood estimator. *The Indian Journal of Statistics* 71-B (Part 1):79–96.

Bretz, F., and L. A. Hothorn. 2003. Statistical analysis of monotone or non-monotone dose-response data from in vitro toxicological assays. *ATLA* 31 (supplement 1):81–96.

Brunk, H. D. 1965. Conditional expectation given a σ-lattice and applications. *The Annals of Mathematical Statistics* 36 (5):1339–50. doi:10.1214/aoms/1177699895.

Chung, Y. T., and O. Y. Shinozaki. 2012. Improved estimators for common mean and ordered means of two normal distributions with ordered variances. *Journal of Statistical Planning and Inference* 142:2619–28. doi:10.1016/j.jspi.2012.03.006.

Cohen, A., and H. B. Sackrowitz. 2002. Inference for the model of several treatments and a control. *Journal of Statistical Planning and Inference* 107 (1–2):89–101. doi:10.1016/S0378-3758(02)00245-8.

Chaudhuri, S., and M. Drton. 2003. *On the bias and mean-square error of the sample minimum and the maximum likelihood estimator for two ordered normal means. Technical report no. 432*. Seattle: Department of Statistics. University of Washington.

Chaudhuri, S., and M. D. Perlman. 2005. On the bias and mean-square error of order-restricted maximum likelihood estimators. *Journal of Statistical Planning and Inference* 130 (1–2):229–50. doi:10.1016/j.jspi.2003.09.040.

Dunbar, S., M. Conaway, and S. D. Peddada. 2001. On improved estimation of parameters subject to order restrictions. *Statistics and Applications* 3:121–8.

Fernandez, M. A., C. Rueda, and B. Salvador. 1999. The loss of efficiency estimating linear functions under restrictions. *Scandinavian Journal of Statistics* 26:127–41.

Hwang, J. T. G., and S. D. Peddada. 1994. Confidence interval estimation subject to order restrictions. *The Annals of Statistics* 22 (1):67–93. doi:10.1214/aos/1176325358.

Kelly, R. 1989. Stochastic reduction of loss in estimating normal means by isotonic regression. *The Annals of Statistics* 17 (2):937–40. doi:10.1214/aos/1176347153.

Lee, C. I. C. 1981. The quadratic loss of isotonic regression under normality. *The Annals of Statistics* 9 (3):686–8. doi:10.1214/aos/1176345475.

Lee, C. I. C. 1988. Quadratic loss of order restricted estimators for treatment means with a control. *The Annals of Statistics* 16 (2):751–8. doi:10.1214/aos/1176350833.

Robertson, T., F. T. Wright, and R. L. Dykstra. 1988. *Order restricted statistical inference*. New York: Wiley.

Silvapulle, M. J., and P. K. Sen. 2005. *Constrained statistical inference: inequality order and shape restrictions*. Hoboken: John Wiley. New York, NY.

Tan, X., and S. D. Peddada. 2000. Asymptotic distribution of some estimators for parameters subject to order restrictions. *Statistics and Applications* 2:7–25.