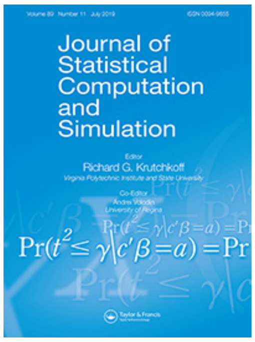

Journal of Statistical Computation and Simulation

ISSN: 0094-9655 (Print) 1563-5163 (Online) Journal homepage: https://www.tandfonline.com/loi/gscs20

Estimation of parameters in the tree order restriction by a randomized decision

R. Momeni, J. Etminan & M. Khanjari Sadegh

To cite this article: R. Momeni, J. Etminan & M. Khanjari Sadegh (2019) Estimation of parameters in the tree order restriction by a randomized decision, Journal of Statistical Computation and Simulation, 89:11, 1986-2005, DOI: 10.1080/00949655.2019.1606223

To link to this article: https://doi.org/10.1080/00949655.2019.1606223

Published online: 18 Apr 2019.

JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION
2019, VOL. 89, NO. 11, 1986–2005
https://doi.org/10.1080/00949655.2019.1606223

Taylor & Francis
Taylor & Francis Group

Estimation of parameters in the tree order restriction by a randomized decision

R. Momeni, J. Etminan and M. Khanjari Sadegh

Department of Statistics, University of Birjand, Birjand, Iran

ABSTRACT

For estimating the smallest location parameter in the location family of distributions which are constrained by the tree ordering $\theta_0 \le \theta_i$ for $1 \le i \le k$, the restricted maximum likelihood estimator diverges to $-\infty$ as $k \to \infty$ and therefore fails to dominate the corresponding unrestricted estimator in terms of the bias and hence the mean squared error (MSE). In this article, we propose a new procedure for the estimation of the location parameters based on a randomized decision. The proposed randomized estimator of θ_0 is improved via the smooth approach to construct the better estimator which remains bounded and decreases the growth rate of its bias and MSE. We show in the case of normal distributions that the MSE of the proposed estimator of θ_0 is less than that of the corresponding unrestricted estimator. By using a simulation study, the performance of the improved estimators is compared with that of the other restricted estimators in terms of three criteria (bias, MSE and coverage probability). The results show that the proposed estimator of θ_0 is substantially better than that of the alternative estimators. Unlike the other procedures, the proposed method for estimating $\theta_i, i = 1, \ldots, k$ performs well.

ARTICLE HISTORY
Received 7 October 2018
Accepted 8 April 2019

KEYWORDS
Bias; isotonic regression estimator (IRE); mean squared error (MSE); tree order restriction; unrestricted maximum likelihood estimator (UMLE)

1. Introduction

Consider the problem of estimation in $k+1$ univariate independent populations with location parameters $\theta_i, i = 0, 1, \ldots, k$ which are constrained by the tree order restriction, $\theta_0 \le \theta_i$, for $1 \le i \le k$. The tree order restriction plays an essential role in statistical inference, taking into account of this order to improve the efficiency of the estimators. It arises in situations where one wishes to compare several treatments θ_i with a control or standard θ_0, using prior information that all of the treatment parameters are at least as large as the control parameter. However, in some situations, this assumption is reasonable and therefore one would like to incorporate it into the statistical methods. Its application can be found in clinical trials and there are many examples in the ordered treatment parameters [1].

There exist considerable research works in literatures on the estimation of parameters under order restrictions in which the isotonic regression technique is used. Suppose

CONTACT J. Etminan jetminan@birjand.ac.ir

$\hat{\boldsymbol{\theta}}^{URE} = (\hat{\theta}_0^{URE}, \ldots, \hat{\theta}_k^{URE})$ is an unrestricted estimator (URE) for $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_k)$ which does not necessarily satisfy the set of constraint C on the parameters and $\boldsymbol{w} = (w_0, \ldots, w_k)$ is a vector of given positive weights.

Definition 1.1: The isotonic regression estimator $\hat{\boldsymbol{\theta}}^{IRE} = (\hat{\theta}_0^{IRE}, \ldots, \hat{\theta}_k^{IRE})$ of the vector $\hat{\boldsymbol{\theta}}^{URE}$ with weight $\boldsymbol{w}$ is the value of $\boldsymbol{\theta}$ which minimizes the weighted sum of squares:

$$\sum_{i=0}^{k} w_i(\hat{\theta}_i^{URE} - \theta_i)^2, \tag{1}$$

subject to the order restriction i.e. $\boldsymbol{\theta} \in C$ [1–3].

If $\hat{\boldsymbol{\theta}}^{URE}$ has an elliptically symmetric unimodal distribution, then the isotonic regression estimator is a restricted maximum likelihood estimator (RMLE) of $\boldsymbol{\theta}$ [2]. It is well known that the isotonic regression estimator $\hat{\boldsymbol{\theta}}^{IRE}$ is the orthogonal projection of $\hat{\boldsymbol{\theta}}^{URE}$ onto the convex cone C with respect to the given weight $\boldsymbol{w}$, geometrically.

Since the control group parameter θ_0 is the smallest parameter in the tree order restriction, it is first estimated. By using the nice form of the max–min formula in Barlow et al. [2], the isotonic regression estimator of θ_0 can be expressed explicitly for the tree order constraint C as follows:

$$\hat{\theta}_0^{IRE} = \min_S \left\{ \frac{\sum_{j\in S} w_j \hat{\theta}_j^{URE}}{\sum_{j\in S} w_j} \right\}, \tag{2}$$

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

In a special case of $k+1$ normal populations with the tree ordering on means θ_i, $i = 0, 1, \ldots, k$ and with commonly known variance σ^2, Lee [4] showed that if $0 \le \theta_i \le c$ for some fixed c, the RMLE of θ_0 diverges to $-\infty$ as $k \to \infty$ and demonstrated that when the weight of the control sample mean w_0 increases, then the absolute bias and MSE of the $\hat{\theta}_0^{IRE}$ will decrease. Various authors via the different procedures tried to decrease the bias and MSE of the $\hat{\theta}_0^{IRE}$ by increasing the weight w_0.

Hwang and Peddada [5] proved that the coverage probability of the confidence interval with fixed length centred at the $\hat{\theta}_0^{IRE}$ tends to zero, when k becomes very large. In the related result, Lee [6] proved that the isotonic regression for each i in the simple order has a smaller mean square error than the corresponding unrestricted estimator; therefore, Hwang and Peddada [5] transformed the tree order restriction to the simple order $\theta_0 \le \theta_1 \le \cdots \le \theta_k$, and then estimated parameters which perform well when populations are independent. But, it may perform poorly for some correlated patterns of populations. Hence, Cohen and Sackrowitz [7] introduced an estimating procedure in the normal populations with equal sample sizes based on the pairwise comparisons of each treatment with the control group in which the weight w_0 is increased and their estimator dominated the RMLE of θ_0 in the high dimensions k. Their estimator was motivated by the fact that each $M_i = \min(\hat{\theta}_0^{URE}, \hat{\theta}_i^{URE})$ has lower MSE than $\hat{\theta}_0^{URE}$ itself. Chaudhuri and Perlman [8] proposed an alternative estimator with this property based on the pairwise RMLE of $(\hat{\theta}_0^{URE}, \hat{\theta}_i^{URE})$

in which the bias of pairwise RMLE is less than that of pairwise minima. Betcher and Peddada [9] modified the Chaudhuri and Perlman's estimator in the analysis of covariance models with the various structures of covariance matrices. In fact, they decomposed the tree order problem into the collection of 2-dimensional simple orders by using the graph theoretic ideas and derived the estimators in the normal model with the common known variance.

Accordingly, in the tree order restriction, we found that by increasing k the isotonic regression of the smallest parameter θ_0, tends to the lower bound of the parameter space and hence underestimation will occur. Therefore, if the true value of the parameter be within an open interval of the natural parameter space, then the $\hat{\theta}_0^{IRE}$ away from the true value and hence the bias increases with the different growth rate in the various models. Also, according to the dependency between the control and treatment estimators, the overestimation can happen for treatment estimators. From this viewpoint, under the tree order restriction the isotonic regression estimator of the control group parameter $\hat{\theta}_0^{IRE}$ does not converge to the corresponding parameter when k increases. Hence, the magnitude of the bias and MSE of IRE (RMLE) will increase as $k \to \infty$, while those of the corresponding unrestricted estimator (URE) do not vary with k [4].

We note that this phenomenon occurs due to the increase of the probability $P(\hat{\theta}_{(1)}^{URE} \leq \hat{\theta}_0^{URE})$ as $k \to \infty$, where $\hat{\theta}_{(1)}^{URE} = \min_{i\geq 1} \hat{\theta}_i^{URE}$. Hence, based on (2) the magnitude of the bias and therefore the MSE of $\hat{\theta}_0^{IRE}$ are stochastically increasing in k. A natural way is that to prevent from the increasing of this inappropriate chance. So, we reverse the scenario of this inappropriate event by using a random decision and then improve the proposed estimators by conditioning expectation based on the decision-theoretic approach. Furthermore under some conditions, we demonstrate that the improved estimators perform better than the alternative estimators due to the random device.

The remainder of this paper is organized as follows. In Section 2, the asymptotic bias and MSE of the IRE (RMLE) are studied, and alternative estimators are introduced. In Section 3, according to the basic estimators URE and IRE, and based on a random decision, we construct a new estimator. On this basis, we give feasible probability to the unrestricted sample estimator to decrease the divergency and then by the use of a probabilistic combination between two underlying estimators URE and IRE the proposed randomized estimator is improved. On the other hand, in some situations of the sampling configurations the proposed estimator and alternatives may not have the isotonic property. In order to obtain such estimators that are the isotonic smooth estimators, we use a technique for arising of the isotonic regression. So, in Section 4, we transform the proposed estimator to an isotonic regression estimator through an iteration algorithm. In this section, we show how to adapt the random device to the case when the smooth estimators are not the isotonic regression estimators. In Section 5, by using a simulation study we compare the performance of the proposed estimator with the performance of alternative procedures in terms of the bias, MSE and coverage probability which are the most popular statistical criteria. According to the simulation results, the bias and MSE of our estimators remain bounded or grow at a much slower rate than those of the other alternatives. Also, the coverage probability of the proposed estimator is higher than those of the URE, IRE and other alternative estimators. Concluding remarks are given in Section 6.

2. The asymptotic bias and MSE of the isotonic regression estimator

Let X_{ij} for $j = 1, \dots, n_i$ and $i = 0, 1, \dots, k$ are independent random samples from $k+1$ populations that belong to a location family $\{P_\theta(x); \theta \in \Theta \subseteq \mathbb{R}\}$. Here $\boldsymbol{\theta} = (\theta_0, \dots, \theta_k)$ is the vector of location parameters in which the support of the underlying distributions or the distribution of the unrestricted estimators is either unbounded or at least is lower unbounded, as in the case of the normal distribution.

We can mathematically formulate the expression (2) in a simpler form as follows,

$$\hat{\theta}_0^{IRE} = \begin{cases} \hat{\theta}_0^{URE} & \text{if } r^{(0)} = 1 \\ \min_{i<r^{(0)}} \left\{ \dfrac{w_0 \hat{\theta}_0^{URE} + \sum_{j=1}^{i} w_{(j)} \hat{\theta}_{(j)}^{URE}}{w_0 + \sum_{j=1}^{i} w_{(j)}} \right\} & \text{if } r^{(0)} > 1, \end{cases} \tag{3}$$

where $r^{(0)}$ is the rank of $\hat{\theta}_0^{URE}$ among $\hat{\theta}_0^{URE}, \dots, \hat{\theta}_k^{URE}$ and $\hat{\theta}_{(r)}^{URE}$ is the rth order statistic of the unrestricted sample estimators $\hat{\theta}_1^{URE}, \dots, \hat{\theta}_k^{URE}$ for $r = 1, \dots, k$. The isotonic regression estimators of the treatment parameters are as follows:

$$\hat{\theta}_i^{IRE} = \max\left(\hat{\theta}_0^{IRE}, \hat{\theta}_i^{URE}\right), \quad i = 1, \dots, k. \tag{4}$$

The computation of the IRE depends heavily on the specific form of the observable sampling order.

Example 2.1: Consider a balanced design ($w_i = w$) and $k=3$ treatments. If the sample order is of the form $\hat{\theta}_{(1)}^{URE} \le \hat{\theta}_{(2)}^{URE} \le \hat{\theta}_{(3)}^{URE} \le \hat{\theta}_0^{URE}$ (i.e. the worst case), then based on the Equation (3) the IRE of θ_0 can be obtained as follows:

$$\hat{\theta}_0^{IRE} = \begin{cases} \dfrac{\hat{\theta}_0^{URE} + \hat{\theta}_{(1)}^{URE}}{2} & \text{if } \hat{\theta}_0^{URE} \le 2\hat{\theta}_{(2)}^{URE} - \hat{\theta}_{(1)}^{URE} \\ \dfrac{\hat{\theta}_0^{URE} + \hat{\theta}_{(1)}^{URE} + \hat{\theta}_{(2)}^{URE} + \hat{\theta}_{(3)}^{URE}}{4} & \text{if } \hat{\theta}_0^{URE} > 3\hat{\theta}_{(3)}^{URE} - (\hat{\theta}_{(1)}^{URE} + \hat{\theta}_{(2)}^{URE}) \\ \dfrac{\hat{\theta}_0^{URE} + \hat{\theta}_{(1)}^{URE} + \hat{\theta}_{(2)}^{URE}}{3} & \text{otherwise.} \end{cases} \tag{5}$$

This procedure is a corrected arithmetic means of the UREs according to the given order restriction.

On the other hand, by (3) and (4) we have $\hat{\theta}_0^{IRE} \le \hat{\theta}_0^{URE}$ and $\hat{\theta}_i^{IRE} \ge \hat{\theta}_i^{URE}$, respectively. Thus, when $\hat{\theta}_i^{URE}$, $i = 0, \dots, k$ are unbiased estimators, the bias of $\hat{\theta}_0^{IRE}$ is strictly negative and the bias of $\hat{\theta}_i^{IRE}$ for $i \ge 1$, is strictly positive, i.e.

$$b(\hat{\theta}_0^{IRE}) = E_{\boldsymbol{\theta}}(\hat{\theta}_0^{IRE}) - \theta_0 < 0, \quad b(\hat{\theta}_i^{IRE}) = E_{\boldsymbol{\theta}}(\hat{\theta}_i^{IRE}) - \theta_i > 0, \quad i \ge 1,$$

and from these inequalities, the magnitude of the bias of $\hat{\theta}_0^{IRE}$ is greatest when $\theta_0 = \theta_1 = \dots = \theta_k$, which is known as the least favourable case.

Corollary 2.2: *In the tree order model, when $\hat{\theta}_i^{URE}, i = 0, \ldots, k$ are unbiased estimators, the squared bias and MSE of the $\hat{\theta}_0^{IRE}$ diverge to ∞ as $k \to \infty$.*

Proof: By Theorem 1.3.6 of Robertson et al. [1] we have:

$$\sum_{i=0}^{k} w_i \hat{\theta}_i^{IRE} = \sum_{i=0}^{k} w_i \hat{\theta}_i^{URE},$$

i.e. the weighted sum is preserved. So,

$$w_0 E_{\boldsymbol{\theta}}(\hat{\theta}_0^{IRE} - \theta_0) = -\sum_{i=1}^{k} w_i E_{\boldsymbol{\theta}}(\hat{\theta}_i^{IRE} - \theta_i),$$

when $w_0 = w_i = 1$ and $\theta_i = \theta$ for $i = 1, \ldots, k$, by symmetry we have,

$$E_{\boldsymbol{\theta}}(\hat{\theta}_0^{IRE} - \theta_0) = -kE_{\boldsymbol{\theta}}(\hat{\theta}_1^{IRE} - \theta),$$

also, from (3) it is easy to see that,

$$\min(\hat{\theta}_0^{URE}, \hat{\theta}_{(1)}^{URE}) \leq \hat{\theta}_0^{IRE} \leq \frac{\hat{\theta}_0^{URE} + \hat{\theta}_{(1)}^{URE}}{2}.$$

Now, suppose $\hat{\theta}_1^{URE}, \ldots, \hat{\theta}_k^{URE}$ are independent and identically distributed (i.i.d) with common distribution function $F(.)$ and $l = sup\{x; F(x) = 0\}$, since $\hat{\theta}_{(1)}^{URE} \to l$ almost surely, as $k \to \infty$,

$$l \leq \hat{\theta}_0^{IRE} \leq \frac{\hat{\theta}_0^{URE} + l}{2},$$

therefore,

$$l - \theta_0 \leq E_{\boldsymbol{\theta}}(\hat{\theta}_0^{IRE}) - \theta_0 \leq \frac{l - \theta_0}{2},$$

especially, if $l = -\infty$ when $k \to \infty$ we have,

$$b(\hat{\theta}_0^{IRE}) \to -\infty, \quad MSE_{\boldsymbol{\theta}}(\hat{\theta}_0^{IRE}) = E_{\boldsymbol{\theta}}(\hat{\theta}_0^{IRE} - \theta_0)^2 \to \infty.$$ ■

For instance, in the tree order normal model, $\hat{\theta}_i^{URE} \overset{iid}{\sim} N(\theta_i, 1)$ with $w_i = 1$ for $i = 0, 1, \ldots, k$, Chaudhuri and Perlman [8], showed that

$$\left| E_{\boldsymbol{\theta}}(\hat{\theta}_0^{IRE}) - \theta_0 \right| = O(\sqrt{2 \log k}) \quad \text{and} \quad E_{\boldsymbol{\theta}}(\hat{\theta}_0^{IRE} - \theta_0)^2 = O(2 \log k).$$

So, both the squared bias and MSE of the smallest parameter in the tree ordering tend to infinity at a logarithmic rate in the least favourable case. Lee [4] established that if the weight w_0 corresponding to $\hat{\theta}_0^{URE}$ is chosen to be sufficiently large, then the MSE of the isotonic regression estimator $\hat{\theta}_0^{IRE}$ is strictly smaller than that of the corresponding unrestricted estimator.

Theorem 2.3 (Lee [4]): *In the normal model* $\hat{\theta}_i^{URE} \sim^{iid} N(\theta_i, \sigma^2)$, *let the means* $\theta_0, \ldots, \theta_k$, *the sample sizes* $n_0, \ldots, n_k$ *and the positive weights* $w_1, \ldots, w_k$ *be fixed. There exists a positive real W such that if* $w_0 \geq W$, *then:*

$$E(\hat{\theta}_0^{IRE} - \theta_0)^2 < E(\hat{\theta}_0^{URE} - \theta_0)^2. \tag{6}$$

Various estimation methods have been proposed by authors, tried in a generic manner to increase the weight w_0. Although Lee [6] showed that

$$E(\hat{\theta}_i^{IRE} - \theta_i)^2 < E(\hat{\theta}_i^{URE} - \theta_i)^2, \tag{7}$$

holds for all $i = 0, 1, \ldots, k$, if the tree order cone is replaced by a simple order cone, for example, $\theta_0 \leq \theta_1 \leq \cdots \leq \theta_k$. So, Hwang and Peddada [5] considered Lee's result in (7) and they chose only one of the simple orderings rather than considering all possible $k!$ simple orderings between $\theta_1, \ldots, \theta_k$ arbitrarily (e.g. $\theta_0 \leq \theta_1 \leq \cdots \leq \theta_k$). Under this consideration, they introduced the following estimators:

$$\hat{\theta}_0^{HP} = \min_{t \geq 0} \left\{ \frac{\sum_{j=0}^{t} w_j \hat{\theta}_j^{URE}}{\sum_{j=0}^{t} w_j} \right\}; \quad \hat{\theta}_i^{HP} = \max\left\{\hat{\theta}_i^{URE}, \hat{\theta}_0^{HP}\right\} \quad \text{for } i \geq 1.$$

Their estimators depend on the chosen simple order between treatment groups. Also, Tan and Peddada [10] observed that for nondiagonal covariance matrix, these estimators may be inconsistent, especially for two populations i.e. $k = 1$, treatment group. So, Cohen and Sackrowitz (CS) [7] introduced an estimator for equal sample sizes based on the pairwise minima as follows:

$$\hat{\theta}_0^{CS} = \frac{\sum_{i=0}^{k} \min(\hat{\theta}_0^{URE}, \hat{\theta}_i^{URE})}{k+1} \quad \hat{\theta}_i^{CS} = \hat{\theta}_0^{CS} + (\hat{\theta}_i^{URE} - \hat{\theta}_0^{URE})^+, \quad i = 1, \ldots, k. \tag{8}$$

The Cohen and Sackrowitz's estimator was motivated by the fact that each $\min(\hat{\theta}_0^{URE}, \hat{\theta}_i^{URE})$ has lower MSE than $\hat{\theta}_0^{URE}$ itself. On the other hand, Chaudhuri and Perlman (CP) [8] proposed an alternative estimator with this property based on the $\hat{\theta}_0^{IRE}(\hat{\theta}_0^{URE}, \hat{\theta}_i^{URE})$ instead of $\min(\hat{\theta}_0^{URE}, \hat{\theta}_i^{URE})$ in (8),where

$$\hat{\theta}_0^{IRE}\left(\hat{\theta}_0^{URE}, \hat{\theta}_i^{URE}\right) = \hat{\theta}_0^{URE} I_{\left\{\hat{\theta}_0^{URE} \leq \hat{\theta}_i^{URE}\right\}} + \left(\frac{\hat{\theta}_0^{URE} + \hat{\theta}_i^{URE}}{2}\right) I_{\left\{\hat{\theta}_0^{URE} > \hat{\theta}_i^{URE}\right\}}. \tag{9}$$

In fact, the MSE of pairwise RMLE i.e. $\hat{\theta}_0^{IRE}(\hat{\theta}_0^{URE}, \hat{\theta}_i^{URE})$ in the least favourable case is substantially smaller than that of pairwise minima i.e. $\min(\hat{\theta}_0^{URE}, \hat{\theta}_i^{URE})$ [11]. So, the CPE dominates the CSE in terms of the biases and MSEs. Also, Betcher and Peddada [9] modified the CP estimator for the analysis of covariance model with the different covariance structures in the normal model i.e. $\hat{\boldsymbol{\theta}}^{UMLE} \sim N_{k+1}(\boldsymbol{\theta}, \sigma^2 \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = [\sigma_{ij}]$ is known.

They estimated the components of $\boldsymbol{\theta}$ as follows:

$$\hat{\theta}_0^{BP} = \frac{\sum_{i=1}^k \sigma_{ii}^{-1}\hat{\theta}_0^{(0,i)RMLE}}{\sum_{i=1}^k \sigma_{ii}^{-1}}, \quad \hat{\theta}_i^{BP} = \max\left\{\hat{\theta}_0^{BP}, \hat{\theta}_i^{RMLE}\right\}, \quad \text{for } i \geq 1, \tag{10}$$

where,

$$\left(\hat{\theta}_0^{(0,i)RMLE}, \hat{\theta}_i^{RMLE}\right) = \begin{cases} \left(\hat{\theta}_0^{UMLE}, \hat{\theta}_i^{UMLE}\right) & \text{if } \hat{\theta}_0^{UMLE} \leq \hat{\theta}_i^{UMLE}, \\ \left(\alpha_i\hat{\theta}_0^{UMLE} + (1-\alpha_i)\hat{\theta}_i^{UMLE}\right)(1,1) & \text{otherwise,} \end{cases}$$

and $\alpha_i = (\sigma_{ii} - \sigma_{0i})/(\sigma_{00} + \sigma_{ii} - 2\sigma_{0i})$, for $i = 1, \ldots, k$.

In the next section, based on the two basic estimators i.e. URE and IRE, we propose an alternative estimator which dominates the unrestricted estimator (URE) and do not have the drawbacks of the isotonic regression estimator (IRE).

3. The proposed procedures (RE, SE)

In order to estimate the control group parameter θ_0 through a different argument, we introduce a probability via the random device and show that the dominancy of the improved estimators over the corresponding UMLEs. By (3), $\hat{\theta}_0^{IRE}$ is decreasing in k and diverges to $-\infty$. From this viewpoint, by increasing k, the IRE (RMLE) of the control group parameter which is the smallest parameter in the tree order restriction, performs poorly. In fact, this natural phenomenon occurs due to,

$$\lim_{k\to\infty} P\left(\hat{\theta}_{(1)}^{URE} \leq \hat{\theta}_0^{URE}\right) = 1. \tag{11}$$

We remove this divergency from the $\hat{\theta}_0^{IRE}$ through a probability which is allocated to the basic unrestricted estimator $\hat{\theta}_0^{URE}$. In fact, our idea implies Lee's [4] result which demonstrated that the MSE of $\hat{\theta}_0^{IRE}$ can be reduced by increasing the weight w_0. According to the Theorem 2.3, we increase the control group weight w_0 by using the allocated probability to the unrestricted sample estimator $\hat{\theta}_0^{URE}$.

We noted that $P(\hat{\theta}_{(1)}^{URE} \leq \hat{\theta}_0^{URE})$ will be increased as k increases, i.e. the chance of occurring the inappropriate event $\left\{\hat{\theta}_i^{URE} \leq \hat{\theta}_0^{URE}\right\}$ for at least one value of $i \geq 1$, converges to 1 as $k \to \infty$. In fact, the pressure of the tree order constraint directly stands on the control group parameter to satisfy the tree ordering. A certain way for preventing this pressure is to choose $\hat{\theta}_0^{URE}$ and $\hat{\theta}_0^{IRE}$ by a random mechanism. For such decision problems, a natural way to do this is simply to choose $\hat{\theta}_0^{URE}$ and $\hat{\theta}_0^{IRE}$ with probabilities p_k and $1 - p_k$, respectively [12]. Hence, we reverse the scenario of this inappropriate chance by the allocated probability to the unrestricted basic estimator $\hat{\theta}_0^{URE}$. It is obvious that the probability p_k must be dependent (i.e. increasing) on the number of treatments k. In order to improve upon the IRE, we first consider the estimator of the form,

$$\hat{\theta}_0^{RE} = \begin{cases} \hat{\theta}_0^{URE} & \text{with probability } p_k, \\ \hat{\theta}_0^{IRE} & \text{with probability } q_k = 1 - p_k, \end{cases} \tag{12}$$

where $\hat{\theta}_0^{URE}$ is the usual unrestricted estimator and $\hat{\theta}_0^{IRE}$ is the isotonic regression estimator (IRE) which is given in (3).

We seek the value of p_k to make the bias and hence the MSE of the proposed estimator of θ_0 as small as possible. One choice of p_k is the value that minimizes the MSE of $\hat{\theta}_0^{RE}$. Thus, the MSE of the proposed estimator can be reduced by a suitable choice of p_k. But this MSE depends on the unknown location parameters.

In this mechanism, there exist several ways for the selection of probability, p_k. An intuitive basic idea to determine p_k in (12) is raised from order restricted methodology which is employed at the hypothesis testing framework. In order to test the null hypothesis $H_0 : \theta_0 = \theta_1 = \cdots = \theta_k$ against the tree order restricted alternative $H_1 : \theta_0 \leq \theta_i;\ i \geq 1$, where the inequality is strict for at least one value of i, it might be reasonable to consider a test statistic based on the nature of alternative hypothesis that can be interpreted as the number of indices i such that $\hat{\theta}_0^{URE} \leq \hat{\theta}_i^{URE}$. Hence, this quantity can be defined as the weighted sum of indicator functions:

$$N(\hat{\boldsymbol{\theta}}^{URE}) = \sum_{i=0}^{k} w_i I_{\left\{\hat{\theta}_0^{URE} \leq \hat{\theta}_i^{URE}\right\}},$$

where the coefficient vector $\boldsymbol{w} = (w_0, w_1, \ldots, w_k)$ is the same vector of weights in the isotonic regression (Eq.1) and depends upon the precisions of the unrestricted sample estimators. So, in our stochastic mechanism the following proportion,

$$p_k = \frac{\sum_{i=0}^{k} w_i I_{\left\{\hat{\theta}_0^{URE} \leq \hat{\theta}_i^{URE}\right\}}}{\sum_{i=0}^{k} w_i} = \frac{w_0 + \sum_{i=r^{(0)}}^{k} w_{(i)}}{\sum_{i=0}^{k} w_i}, \tag{13}$$

is the probability of the selection of $\hat{\theta}_0^{URE}$ which is a function of the test statistic in the tree order problem.

Example 3.1: The randomized estimator in (8) with 2+1 populations constrained by the tree order parameters $\theta_0 \leq \theta_i, i = 1, 2$, and $w_i = w, i = 0, 1, 2$, derived as follows:

$$\hat{\theta}_0^{RE} = \begin{cases} \hat{\theta}_0^{URE} & \text{if } r^{(0)} = 1, \\ \begin{cases} \hat{\theta}_0^{URE} & \text{w.p. } \frac{2}{3}, \\ \hat{\theta}_0^{IRE} & \text{w.p. } \frac{1}{3}, \end{cases} & \text{if } r^{(0)} = 2, \\ \begin{cases} \hat{\theta}_0^{URE} & \text{w.p. } \frac{1}{3}, \\ \hat{\theta}_0^{IRE} & \text{w.p. } \frac{2}{3}, \end{cases} & \text{if } r^{(0)} = 3, \end{cases} \tag{14}$$

where $\hat{\theta}_0^{IRE}$ is obtained analogous to (5) for $k = 2$.

Thus, for general k when $r^{(0)}$ is the rank of $\hat{\theta}_0^{URE}$, it can be shown that the extension of the randomized estimator for equal weights (i.e. equal sample sizes) is as follows:

$$\hat{\theta}_0^{RE} = \begin{cases} \hat{\theta}_0^{URE} & \text{w.p. } p_k = \dfrac{k - r^{(0)} + 2}{k+1}, \\ \hat{\theta}_0^{IRE} & \text{w.p. } 1 - p_k = \dfrac{r^{(0)} - 1}{k+1}. \end{cases} \tag{15}$$

In the following, for comparing the MSEs, we first derive the MSE of $\hat{\theta}_0^{RE}$ when $\hat{\theta}_0^{URE}, \hat{\theta}_1^{URE}, \ldots, \hat{\theta}_k^{URE}$ are i.i.d with distribution F_θ where $\theta_0 = \theta_1 = \cdots = \theta_k = \theta$, i.e. the least favourable case,

$$\begin{aligned}
MSE_\theta(\hat{\theta}_0^{RE}) &= \sum_{r=1}^{k+1} E\left((\hat{\theta}_0^{RE} - \theta)^2 | r^{(0)} = r\right) \frac{1}{k+1} \\
&= \frac{1}{k+1} \sum_{r=1}^{k+1} \left\{ \frac{k-r+2}{k+1} E\left((\hat{\theta}_0^{URE} - \theta)^2 | r^{(0)} = r\right) \right. \\
&\quad \left. + \frac{r-1}{k+1} E\left((\hat{\theta}_0^{IRE} - \theta)^2 | r^{(0)} = r\right) \right\} \\
&= \frac{1}{k+1} \sum_{r=1}^{k+1} E\left((\hat{\theta}_0^{URE} - \theta)^2 | r^{(0)} = r\right) \\
&\quad + \frac{1}{k+1} \sum_{r=2}^{k+1} \frac{r-1}{k+1} E\left((\hat{\theta}_0^{IRE} - \theta)^2 - (\hat{\theta}_0^{URE} - \theta)^2 | r^{(0)} = r\right) \\
&= \mathrm{Var}(\hat{\theta}_0^{URE}) + \frac{1}{(k+1)^2} \sum_{r=2}^{k+1} (r-1) E\left((\hat{\theta}_0^{IRE} - \theta)^2 - (\hat{\theta}_0^{URE} - \theta)^2 | r^{(0)} = r\right).
\end{aligned}$$

Proposition 3.2: *Let $\hat{\mu}_i^{URE} \sim N(\mu_i, \sigma^2)$, and $w_i = w, i = 0, 1$, where $\mu_0 \le \mu_1$. Suppose that $\hat{\mu}_0^{URE}$ and $\hat{\mu}_1^{URE}$ are mutually independent. Then, the MSE of the proposed estimator $\hat{\mu}_0^{RE}$ is smaller than that of $\hat{\mu}_0^{URE}$ when $\mu_0 = \mu_1 = \mu$.*

Proof:

$$\begin{aligned}
MSE(\hat{\mu}_0^{RE}) &= \frac{1}{2} \mathrm{E}\left[(\hat{\mu}_0^{URE} - \mu)^2 | \hat{\mu}_0^{URE} \le \hat{\mu}_1^{URE}\right] + \frac{1}{4} \mathrm{E}\left[(\hat{\mu}_0^{URE} - \mu)^2 | \hat{\mu}_0^{URE} > \hat{\mu}_1^{URE}\right] \\
&\quad + \frac{1}{4} \mathrm{E}\left[\left(\frac{\hat{\mu}_0^{URE} + \hat{\mu}_1^{URE}}{2} - \mu\right)^2 \middle| \hat{\mu}_0^{URE} > \hat{\mu}_1^{URE}\right] \\
&= \mathrm{Var}(\hat{\mu}_0^{URE}) - \frac{1}{4} \mathrm{E}\left[(\hat{\mu}_0^{URE} - \mu)^2 | \hat{\mu}_0^{URE} > \hat{\mu}_1^{URE}\right] + \frac{1}{4} \mathrm{Var}\left(\frac{\hat{\mu}_0^{URE} + \hat{\mu}_1^{URE}}{2}\right) \\
&= \sigma^2 - \frac{1}{4} \int_{-\infty}^{\infty} 2(x-\mu)^2 \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \left[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right] \mathrm{d}x + \frac{\sigma^2}{8}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{9}{8}\sigma^2 - \frac{1}{4}\sigma^2 \int_{-\infty}^{\infty} 2z^2\phi(z)[1-\Phi(z)]\,\mathrm{d}z \\
&= \frac{9}{8}\sigma^2 - \frac{1}{4}\sigma^2 \\
&= \frac{7}{8}\sigma^2 < \sigma^2 = MSE(\hat{\mu}_0^{URE}),
\end{aligned}
$$

whereas,

$$
\begin{aligned}
&\mathrm{E}\left[\left(\frac{\hat{\mu}_0^{URE}+\hat{\mu}_1^{URE}}{2}-\mu\right)^2 \middle| \hat{\mu}_0^{URE} > \hat{\mu}_1^{URE}\right] \\
&\quad = \mathrm{E}\left[\left(\frac{\hat{\mu}_0^{URE}+\hat{\mu}_1^{URE}}{2}-\mu\right)^2 \middle| \hat{\mu}_0^{URE} < \hat{\mu}_1^{URE}\right] \\
&\quad = \mathrm{E}\left[\left(\frac{\hat{\mu}_0^{URE}+\hat{\mu}_1^{URE}}{2}-\mu\right)^2\right]
\end{aligned}
$$

and $\int_{-\infty}^{\infty} 2z^2\phi(z)[1-\Phi(z)]\,\mathrm{d}z$ is the second moment of the skewed-normal distribution with skewness parameter $\lambda = 1$ [13]. ■

Because of the complexity of the above calculations even for moderate k, in Section 5, we can rely on the Monte Carlo simulation study for comparison of the procedures.

Another generalization of the proposed procedure is based on the conditional expectation over the actions space. According to the randomized estimator $\hat{\theta}_0^{RE}$, we can use its expectation based on the Theorem 3.3, in which the obtained smoothed estimator has a smaller MSE than that of the randomized estimator. The results contained in this section are basically driven by the following theorem.

Theorem 3.3 ([12]): *Assume that A is a convex subset of R^k, and that for each $\theta \in \Theta$ the loss function $L(\theta, a)$ is a convex function of $a \in A$. Let δ^* be a randomized decision rule in D^* for which $E_{\delta^*(x,.)}(|a|) < \infty$ for all $x \in \mathcal{X}$. Then for the smoothed rule $\delta(x) = E_{\delta^*(x,.)}(a)$ we have:*

$$
L(\theta, \delta(x)) \leq L(\theta, \delta^*(x, .)) \quad \textit{for all } x \textit{ and } \theta.
$$

In particular, for any given loss function the risk of any estimator is reduced by taking its conditional expectation given a sufficient statistic [14].

Therefore, based on the Theorem 3.3 in order to reduce the MSE, a smoothed estimator is provided as,

$$
\hat{\theta}_0^{SE} = p_k\hat{\theta}_0^{URE} + (1-p_k)\hat{\theta}_0^{IRE}; \quad \hat{\theta}_i^{SE} = \max\left\{\hat{\theta}_0^{SE}, \hat{\theta}_i^{URE}\right\}, \quad i = 1, \dots, k, \tag{16}
$$

which, for a given $\hat{\boldsymbol{\theta}}^{URE}$, is the expected value of $\hat{\theta}_0^{RE}$ in the sense of the Theorem 3.3 and p_k is defined in (10). In constructing this estimator, by the use of the probability p_k as the smoother factor, we lead to the linear combination of the URE and IRE in which it tends to shrink the $\hat{\theta}_0^{IRE}$ toward $\hat{\theta}_0^{URE}$. Hence, it permits any smooth transition between the two extreme values of these estimators. Note that $\hat{\theta}_0^{IRE} \leq \hat{\theta}_0^{SE} \leq \hat{\theta}_0^{URE}$ and therefore via this combination the two basic estimators (IRE and URE) will approach each other.

According to the $\hat{\theta}_0^{SE}$ which is a smoothed estimator between two basic estimators, we adjusted the restricted estimator of θ_0. Therefore, we found that the IRE (RMLE) neither is undesirable nor fails in terms of the MSE for the tree order constraint, but it requires some modification on the IRE and URE of θ_0.

Example 3.4: Based on (13), the randomized estimator given in (14) can be written as follows:

$$\hat{\theta}_0^{SE} = \begin{cases} \hat{\theta}_0^{URE} & \text{if } r^{(0)} = 1, \\ \frac{2}{3}\hat{\theta}_0^{URE} + \frac{1}{3}\hat{\theta}_0^{IRE} & \text{if } r^{(0)} = 2, \\ \frac{1}{3}\hat{\theta}_0^{URE} + \frac{2}{3}\hat{\theta}_0^{IRE} & \text{if } r^{(0)} = 3, \end{cases}$$

in which the $\hat{\theta}_0^{IRE}$ is again obtained analogous to (5).

In a general case, when there are $k+1$ populations with equal weights ($w_i = w$), by using the randomized estimator in (12), we smooth this estimator as follows:

$$\hat{\theta}_0^{SE} = \left(\frac{k - r^{(0)} + 2}{k+1}\right)\hat{\theta}_0^{URE} + \left(\frac{r^{(0)} - 1}{k+1}\right)\hat{\theta}_0^{IRE}. \tag{17}$$

From the smoothed estimator in (14), we derive the MSE of $\hat{\theta}_0^{SE}$ in the following:

$$\begin{aligned} MSE\left(\hat{\theta}_0^{SE}\right) &= \sum_{r=1}^{k+1} E\Big((\hat{\theta}_0^{SE} - \theta_0)^2 \mid r^{(0)} = r\Big)\frac{1}{k+1} \\ &= \frac{1}{k+1}\sum_{r=1}^{k+1} E\left(\left(\frac{(k-r+2)\hat{\theta}_0^{URE} + (r-1)\hat{\theta}_0^{IRE}}{k+1} - \theta_0\right)^2 \middle| \, r^{(0)} = r\right) \\ &= \frac{1}{(k+1)^3}\sum_{r=1}^{k+1} E\Big(\big((k-r+2)(\hat{\theta}_0^{URE} - \theta_0) \\ &\quad +(r-1)(\hat{\theta}_0^{IRE} - \theta_0)\big)^2 \mid r^{(0)} = r\Big). \end{aligned}$$

It should be noted that even for moderate k, no analytic expressions for the biases and variances of $\hat{\theta}_0^{RE}$, $\hat{\theta}_0^{SE}$ and $\hat{\theta}_0^{IRE}$ are available. Thus, the mean squared error of these estimators cannot be compared analytically. Furthermore, this case remains unresolved. These questions are explored in a simulation study presented in Section 5.

In the sequel, we show under some specific conditions that the proposed smooth estimator and Chaudhuri and Perlman's (CPE) [8] estimator are the same. For this purpose, when $w_i = w, i = 0, 1, \ldots, k$, then CPE can be rewritten as follows:

$$
\begin{aligned}
\hat{\theta}_0^{CP} &= \frac{1}{k+1}\left(\hat{\theta}_0^{URE} + \sum_{i=1}^{k} \min\left(\hat{\theta}_0^{URE}, \frac{\hat{\theta}_0^{URE} + \hat{\theta}_i^{URE}}{2}\right)\right) \\
&= \frac{1}{k+1}\left(\hat{\theta}_0^{URE} + \sum_{i=1}^{r^{(0)}-1} \frac{\hat{\theta}_0^{URE} + \hat{\theta}_{(i)}^{URE}}{2} + \sum_{i=r^{(0)}+1}^{k+1} \hat{\theta}_0^{URE}\right) \\
&= \frac{1}{2(k+1)}\left(\left(2(k - r^{(0)} + 2) + (r^{(0)} - 1)\right)\hat{\theta}_0^{URE} + \hat{\theta}_{(1)}^{URE} + \cdots + \hat{\theta}_{(r^{(0)}-1)}^{URE}\right) \\
&= \frac{1}{2(k+1)}\left(\left(2k + 3 - r^{(0)}\right)\hat{\theta}_0^{URE} + \hat{\theta}_{(1)}^{URE} + \cdots + \hat{\theta}_{(r^{(0)}-1)}^{URE}\right) \\
&= \left(\frac{2k + 3 - r^{(0)}}{2(k+1)}\right)\hat{\theta}_0^{URE} + \frac{1}{2(k+1)}\left(\hat{\theta}_{(1)}^{URE} + \cdots + \hat{\theta}_{(r^{(0)}-1)}^{URE}\right),
\end{aligned}
\tag{18}
$$

now for some $i < r^{(0)}$ as in (3), we have,

$$
\hat{\theta}_0^{IRE} = \frac{\hat{\theta}_0^{URE} + \sum_{j=1}^{i} \hat{\theta}_{(j)}^{URE}}{i+1}. \tag{19}
$$

Hence, by the mathematical induction we can obtain the smooth estimator as follows:

$$
\begin{aligned}
\hat{\theta}_0^{SE} &= p_k\hat{\theta}_0^{URE} + (1 - p_k)\hat{\theta}_0^{IRE} \\
&= \left(\frac{ip_k + 1}{i+1}\right)\hat{\theta}_0^{URE} + (1 - p_k)\left(\frac{\hat{\theta}_{(1)}^{URE} + \cdots + \hat{\theta}_{(i)}^{URE}}{i+1}\right).
\end{aligned}
\tag{20}
$$

Now by comparison of the Equations (18) and (20), we see that if $i = r^{(0)} - 1$ and,

$$
p_k = \frac{(2k - r^{(0)} + 3)i - r^{(0)} + 1}{2i(k+1)}, \tag{21}
$$

then $\hat{\theta}_0^{SE} = \hat{\theta}_0^{CP}$.

4. Isotonicity of the proposed estimator

As mentioned in Section 3, $\hat{\theta}_0^{IRE} \le \hat{\theta}_0^{SE} \le \hat{\theta}_0^{URE}$. For estimating the control group parameter, $\hat{\theta}_0^{SE}$ may not be the isotonic regression. Suppose that the subscript i is the highest index of the treatment groups which is involved in making of $\hat{\theta}_0^{IRE}$ as in (3). From the Equation (16) we have,

$$
\begin{aligned}
\hat{\theta}_0^{SE} &= p_k\hat{\theta}_0^{URE} + (1 - p_k)\hat{\theta}_0^{IRE} \\
&= p_k\hat{\theta}_0^{URE} + (1 - p_k)\frac{w_0\hat{\theta}_0^{URE} + \sum_{j=1}^{i} w_{(j)}\hat{\theta}_{(j)}^{URE}}{w_0 + \sum_{j=1}^{i} w_{(j)}} \\
&= \frac{\left(w_0 + p_k\sum_{j=1}^{i} w_{(j)}\right)\hat{\theta}_0^{URE} + (1 - p_k)\sum_{j=1}^{i} w_{(j)}\hat{\theta}_{(j)}^{URE}}{w_0 + \sum_{j=1}^{i} w_{(j)}},
\end{aligned}
\tag{22}
$$

now two cases may occur:

(i) If $\hat{\theta}_0^{SE} \leq \hat{\theta}_{(i+1)}^{URE}$, then the isotonicity of the smooth proposed estimator, $\hat{\theta}_0^{SE}$, with new weights $w_0 + p_k \sum_{j=1}^{i} w_{(j)}, (1-p_k)w_{(1)}, \ldots, (1-p_k)w_{(i)}, w_{(i+1)}, \ldots, w_{(k)}$, would be maintained. We call this estimator the isotonic smoothed estimator (ISE) and will denote by $\hat{\theta}_0^{ISE}$.

(ii) If there exists a fixed $j = i+1, \ldots, r^{(0)} - 1$, such that $\hat{\theta}_{(j)}^{URE} \leq \hat{\theta}_0^{SE}$, then with above new weights the minimization in (3) does not hold and therefore $\hat{\theta}_0^{SE}$ is not an isotonic regression. But as follows, one can make an estimator for θ_0 based on $\hat{\theta}_0^{SE}$ having the isotonic regression property,

$$\begin{aligned}
\hat{\theta}_0^{IRSE} &= \min_{i+1 \leq l < r^{(0)}} \left\{ \frac{\left(w_0 + p_k \sum_{j=1}^{i} w_{(j)}\right) \hat{\theta}_0^{URE} + (1-p_k)\sum_{j=1}^{i} w_{(j)} \hat{\theta}_{(j)}^{URE} + \sum_{j=i+1}^{l} w_{(j)} \hat{\theta}_{(j)}^{URE}}{w_0 + p_k \sum_{j=1}^{i} w_{(j)} + (1-p_k)\sum_{j=1}^{i} w_{(j)} + \sum_{j=i+1}^{l} w_{(j)}} \right\} \\
&= \min_{i+1 \leq l < r^{(0)}} \left\{ \frac{\left(w_0 + \sum_{j=1}^{i} w_{(j)}\right) \hat{\theta}_0^{SE} + \sum_{j=i+1}^{l} w_{(j)} \hat{\theta}_{(j)}^{URE}}{w_0 + \sum_{j=1}^{i} w_{(j)} + \sum_{j=i+1}^{l} w_{(j)}} \right\}, \qquad (23)
\end{aligned}$$

This motivates us to derive smooth estimators that have the isotonic regression property as in (3). In order to achieve an estimator having this property, in the sequel we propose an iteration algorithm starting with $\hat{\theta}_0^{SE^{(0)}} = \hat{\theta}_0^{SE}$ as given in (22).

Algorithm:
Set $t = 1$,

Step 1. If $\hat{\theta}_0^{SE^{(t-1)}} \leq \hat{\theta}_{(i+1)}^{URE}$, where i is the highest index of treatment groups that is involved in the construction of $\hat{\theta}_0^{SE^{(t-1)}}$, return $\hat{\theta}_0^{ISE} = \hat{\theta}_0^{SE^{(t-1)}}$ and then stop. Otherwise, continue steps 2 to 5.

Step 2. In view of Equation (23), the isotonic regression estimator which is constructed based on $\hat{\theta}_0^{SE^{(t-1)}}$ is given by,

$$\hat{\theta}_0^{IRSE^{(t)}} = \min_{i+1 \leq l < r^{(0)}} \left\{ \frac{(w_0 + \sum_{j=1}^{i} w_{(j)}) \hat{\theta}_0^{SE^{(t-1)}} + \sum_{j=i+1}^{l} w_{(j)} \hat{\theta}_{(j)}^{URE}}{w_0 + \sum_{j=1}^{i} w_{(j)} + \sum_{j=i+1}^{l} w_{(j)}} \right\}.$$

Step 3. As the stochastic mechanism in (12), we may again make a randomized estimator of the control group as follows,

$$\hat{\theta}_0^{RE^{(t)}} = \begin{cases} \hat{\theta}_0^{SE^{(t-1)}} & \text{w.p. } p_{k-i}, \\ \hat{\theta}_0^{IRSE^{(t)}} & \text{w.p. } 1 - p_{k-i}, \end{cases}$$

where p_{k-i} is a newly allocated probability that can be determined as like as p_k, but with $k-i$ treatment groups.

Step 4. Now in view of (16), the smoothed estimator is given by,

$$\hat{\theta}_0^{SE^{(t)}} = p_{k-i}\hat{\theta}_0^{SE^{(t-1)}} + (1 - p_{k-i})\hat{\theta}_0^{IRSE^{(t)}},$$

Step 5. Set $t = t+1$, and go to step 1.

At the end of iteration t, the obtained estimator is more efficient and has lower MSE, since the weight of control group is increased and this agrees with Lee's [4] result in the Theorem 2.3. This algorithm is continued until the smoothed proposed estimator will be an isotonic regression estimator. For the CSE and CPE, it is clear that the isotonicity will not hold in some situations. It can be argued that the estimators which were constructed by this way satisfy in the tree order constraint.

Since expressions for the risk functions of the smooth estimators may not be expressed in simple closed forms, we use in the next section the Monte Carlo simulation to compare the behaviour of these estimators.

5. Simulation study

We conducted a large simulation study to compare the performance of the proposed estimators with other relevant procedures. Besides assesing the performance of unrestricted estimators (URE), we also assess the performance of four alternative estimators which are the isotonic regression, IRE (RMLE), Chaudhuri and Perlman [8] estimator (CPE), randomized estimator (RE) and smoothed estimator (SE). We generated independent observations with the small sample sizes $n_i = 1$ from $k + 1$ normal distributions i.e. $X_{ij} \sim N(\mu_i, \sigma^2)$, here $\boldsymbol{\theta} = \boldsymbol{\mu}$ and the values of mean parameters $(0, \mu, \dots, \mu)$ for $\mu = 0, 0.5, 1, 1.5, 2$ and standard deviations $(1, \sigma, ., .\sigma)$ for $\sigma = 0.5, 1, 1.5, 2$, which are constrained by the tree order restriction on the mean parameters $\mu_i \geq \mu_0 = 0; i = 1, \dots, k$. The values of biases, MSEs and coverage probabilities (CPs) are obtained for different patterns of $k = 4,9,14$ based on 100,000 repetitions and only a subset of results are summarized in Table 1.

We noted in the normal model that the unrestricted estimator (URE) and isotonic regression estimator (IRE) are equivalent to the unrestricted maximum likelihood estimator (UMLE) and restricted maximum likelihood estimator (RMLE), respectively. For convenience, we assume that the treatment means are homogeneous, so the average of treatment means is considered as the representative of treatment parameters. For any parameter, the coverage probability of a fixed width simultaneous confidence interval for $\mu_i, i = 1, \dots, k$ centred at the $\hat{\mu}_i$ is defined as $P(|\hat{\mu}_1 - \mu_1| < cs_1, \dots, |\hat{\mu}_k - \mu_k| < cs_k)$, where s_i is the standard deviation of the UMLE of μ_i. The constant c is chosen such that $P(|\bar{X}_1 - \mu_1| < cs_1, \dots, |\bar{X}_k - \mu_k| < cs_k) = 0.95$. The individual version of this criterion for the nodal parameter μ_0 is defined accordingly. This is the peakedness criterion of Birnbaum [15] for comparing estimators. Therefore, the estimator that has higher coverage probability is preferred, because it has a larger concentration of distribution around the true parameter.

As mentioned previously, the restricted estimators are biased. Table 1 shows that the absolute biases of the RE and SE for the control mean parameter are smaller than those of the RMLE, and compete very well with the CPE. Also, the biases of RE and SE for treatment

Table 1. Biases, MSEs and CPs of the UMLE, RMLE, CPE, RE and SE for $\mu_0 = 0$ and μ in the parentheses.

			Bias					MSE					CP				
k	μ	σ	UMLE	RMLE	CPE	RE	SE	UMLE	RMLE	CPE	RE	SE	UMLE	RMLE	CPE	RE	SE
4	0	1	−0.003	−0.648	−0.283	−0.394	−0.393	0.9998	0.822	0.679	0.784	0.637	0.949	0.970	0.974	0.968	0.974
			(0.0004)	(0.167)	(0.372)	(0.294)	(0.276)	(0.9998)	(0.721)	(0.722)	(0.759)	(0.660)	(0.950)	(0.974)	(0.974)	(0.973)	(0.975)
		2	0.002	−0.550	−0.175	−0.292	−0.293	0.998	0.952	0.855	0.897	0.790	0.949	0.957	0.964	0.960	0.969
			(0.008)	(0.561)	(0.780)	(0.697)	(0.689)	(4.017)	(2.393)	(2.392)	(2.406)	(2.334)	(0.949)	(0.974)	(0.974)	(0.974)	(0.974)
	1	1	0.002	−0.267	−0.098	−0.128	−0.128	1.002	0.672	0.803	0.795	0.747	0.950	0.975	0.972	0.968	0.974
			(−0.0004)	(0.067)	(0.143)	(0.124)	(0.119)	(1.0001)	(0.845)	(0.780)	(0.804)	(0.779)	(0.949)	(0.971)	(0.972)	(0.972)	(0.973)
		2	0.002	−0.316	−0.094	−0.138	−0.138	1.0003	0.826	0.891	0.877	0.831	0.950	0.967	0.961	0.962	0.967
			(0.007)	(0.325)	(0.432)	(0.402)	(0.400)	(4.003)	(2.729)	(2.576)	(2.619)	(2.595)	(0.950)	(0.974)	(0.974)	(0.974)	(0.974)
	2	1	0.0004	−0.078	−0.025	−0.029	−0.029	0.999	0.824	0.926	0.921	0.911	0.949	0.970	0.960	0.960	0.963
			(0.001)	(0.021)	(0.040)	(0.038)	(0.037)	(1.002)	(0.944)	(0.910)	(0.915)	(0.912)	(0.951)	(0.964)	(0.967)	(0.967)	(0.967)
		2	0.0002	−0.160	−0.045	−0.057	−0.057	0.999	0.842	0.933	0.923	0.906	0.950	0.966	0.958	0.958	0.961
			(0.003)	(0.163)	(0.209)	(0.201)	(0.200)	(4.00)	(3.171)	(3.036)	(3.060)	(3.053)	(0.950)	(0.974)	(0.974)	(0.974)	(0.974)
9	0	1	−0.0024	−0.916	−0.284	−0.574	−0.577	0.995	1.11	0.660	0.861	0.651	0.950	0.965	0.975	0.971	0.975
			(−0.0006)	(0.101)	(0.382)	(0.234)	(0.199)	(1.002)	(0.768)	(0.718)	(0.750)	(0.654)	(0.950)	(0.974)	(0.974)	(0.974)	(0.975)
		2	0.004	−0.937	−0.175	−0.510	−0.510	0.993	1.357	0.842	1.016	0.771	0.951	0.927	0.966	0.955	0.970
			(−0.002)	(0.416)	(0.784)	(0.603)	(0.583)	(3.999)	(2.470)	(2.358)	(2.399)	(2.274)	(0.950)	(0.975)	(0.975)	(0.975)	(0.975)
	1	1	−0.0007	−0.417	−0.100	−0.194	−0.192	0.997	0.633	0.794	0.730	0.654	0.950	0.974	0.972	0.970	0.975
			(0.001)	(0.047)	(0.150)	(0.111)	(0.100)	(0.998)	(0.870)	(0.773)	(0.806)	(0.783)	(0.950)	(0.972)	(0.973)	(0.973)	(0.973)
		2	0.001	−0.573	−0.095	−0.239	−0.238	0.997	0.895	0.883	0.847	0.733	0.950	0.962	0.963	0.964	0.973
			(−0.001)	(0.254)	(0.434)	(0.368)	(0.359)	(3.999)	(2.834)	(2.548)	(2.642)	(2.602)	(0.950)	(0.975)	(0.975)	(0.975)	(0.975)
	2	1	0.001	−0.135	−0.024	−0.043	−0.042	0.997	0.733	0.923	0.886	0.870	0.950	0.975	0.962	0.965	0.972
			(−0.001)	(0.014)	(0.040)	(0.034)	(0.032)	(0.998)	(0.948)	(0.903)	(0.912)	(0.908)	(0.951)	(0.966)	(0.969)	(0.969)	(0.969)
		2	0.004	−0.305	−0.042	−0.090	−0.091	0.993	0.772	0.924	0.786	0.834	0.951	0.971	0.959	0.962	0.969
			(−0.002)	(0.135)	(0.209)	(0.191)	(0.187)	(4.003)	(3.234)	(3.019)	(3.068)	(3.060)	(0.950)	(0.974)	(0.974)	(0.974)	(0.974)
14	0	1	0.001	−1.06	−0.281	−0.678	−0.677	0.995	1.362	0.656	0.972	0.711	0.950	0.958	0.974	0.970	0.974
			(0.0003)	(0.077)	(0.387)	(0.208)	(0.167)	(1.003)	(0.803)	(0.717)	(0.761)	(0.670)	(0.949)	(0.974)	(0.974)	(0.974)	(0.974)
		2	0.001	−1.193	−0.177	−0.656	−0.654	1.003	1.822	0.848	1.214	0.843	0.950	0.886	0.965	0.941	0.968
			(−0.0009)	(0.340)	(0.788)	(0.556)	(0.524)	(4.0005)	(2.593)	(2.364)	(2.455)	(2.297)	(0.950)	(0.975)	(0.975)	(0.975)	(0.975)
	1	1	0.002	−0.513	−0.979	−0.229	−0.229	1.002	0.652	0.794	0.717	0.611	0.949	0.973	0.972	0.970	0.974
			(−0.002)	(0.035)	(0.149)	(0.104)	(0.088)	(0.999)	(0.887)	(0.770)	(0.810)	(0.788)	(0.950)	(0.973)	(0.974)	(0.974)	(0.974)
		2	−0.006	−0.759	−0.102	−0.314	−0.314	1.004	1.057	0.890	0.876	0.697	0.949	0.952	0.962	0.963	0.973
			(0.002)	(0.217)	(0.436)	(0.353)	(0.338)	(3.995)	(2.934)	(2.551)	(2.680)	(2.630)	(0.951)	(0.975)	(0.975)	(0.975)	(0.975)
	2	1	0.003	−0.176	−0.022	−0.048	−0.048	1.000	0.685	0.926	0.875	0.849	0.949	0.975	0.961	0.964	0.973
			(−0.0007)	(0.012)	(0.041)	(0.034)	(0.031)	(1.002)	(0.957)	(0.904)	(0.915)	(0.912)	(0.949)	(0.967)	(0.970)	(0.970)	(0.970)
		2	−0.001	−0.423	−0.047	−0.122	−0.122	0.994	0.777	0.925	0.861	0.796	0.950	0.969	0.958	0.963	0.971
			(−0.001)	(0.120)	(0.210)	(0.187)	(0.182)	(4.008)	(3.303)	(3.023)	(3.087)	(3.078)	(0.951)	(0.975)	(0.975)	(0.975)	(0.975)

Table 2. Square root of the MSE ratio of the estimators relative to the UMLE, and coverage probabilities (CP) for the control group and treatment groups for different values of k on the set $\{(\mu,\sigma) : \mu = 0, 0.5, 1, 1.5, 2; \sigma = 0.5, 1, 1.5, 2\}$.

		Control group			Treatment groups		
	k	RMLE	CPE	SE	RMLE	CPE	SE
$\left(\frac{MSE(.)}{MSE(UMLE)}\right)^{1/2}$	4	0.881	0.906	0.883	0.894	0.883	0.867
	9	0.909	0.903	0.852	0.907	0.881	0.867
	14	0.951	0.902	0.844	0.917	0.880	0.871
	24	1.034	0.902	0.846	0.932	0.880	0.878
CP	4	0.970	0.967	0.970	0.970	0.969	0.971
	9	0.967	0.968	0.973	0.970	0.970	0.971
	14	0.960	0.968	0.973	0.971	0.971	0.972
	24	0.946	0.968	0.974	0.971	0.971	0.972

mean parameters (in parentheses) are smaller than those of CPE. According to Table 1, the growth rate of the biases of the RE and SE is about half of the RMLE.

From Table 1, it is remarkable that the MSE values of the smooth estimator (SE) are smaller than those of CPE and RMLE, especially in the least favourable case $\boldsymbol{\mu} = (0, 0, \ldots, 0)$. Also, the biases and hence the MSEs of the proposed estimators for treatment parameters perform better than the alternatives. In several cases, these values were substantially smaller than all its competitors. It is seen that the MSEs of the proposed estimators (RE, SE) appear to stabilize or decrease with k, especially for the least favourable case. In all cases for different patterns, the coverage probability of the proposed methods can be substantially higher than that of the other alternatives (UMLE, RMLE, CPE). Among these estimators for any pattern of parameters, SE has the most coverage probabilities and always exceeded the nominal level 0.95. By increasing k, unlike the RMLE and CPE, the coverage probabilities of the proposed procedures (RE, SE) do not decrease and therefore, the parameters are coveraged with the high probabilities by the proposed estimators.

Table 2 gives a summary of the relative performance of the estimators in terms of the averaged CPs and $\sqrt{MSE(.)/MSE(UMLE)}$ values. We extract from this table that the proposed estimator SE has comparatively good performance relative to the alternatives. Therefore, according to the simulation results, we found that the quantity p_k in (10) is a suitable choice to reduce the squared error of proposed estimators significantly, whereas the MSE of the UMLE is set to be equal to σ^2.

Since the smooth estimator and CPE may not be the isotonic regression, it is reasonable to derive the isotonic smooth estimator (ISE), as it is done in Section 4. Hence in the sequel, we also compare the performance of the isotonic smooth estimator with the other estimators IRE (RMLE) and CPE. Here, like Chaudhuri and Perlman [8], we present some results for $\sigma^2 = 1$ and small sample sizes $n_i = 1$. For k=1,2,3,5,10,20,100 normal populations with two configurations of $\boldsymbol{\mu} = (\mu_0, \mu_1, \ldots, \mu_k)$ are considered: $(0, \ldots, 0)$ that is the least favourable case and the interior case $(0, 1, \ldots, 1)$. The presented results are based on 100,000 repetitions, for each k.

For evaluation of the proposed algorithm in Section 4, we compute the percentage of times that the smooth estimator exceeded the natural treatment estimator which is included in the computation of the smooth estimator in each pattern. Based on the proportions in Table 3, it indicates that the proposed algorithm for obtaining the isotonic smooth

Table 3. The ratio of times that the smooth estimator is not the isotonic regression estimator.

k	2	3	5	10	20	50	100
$\boldsymbol{\mu} = (0, 0, \ldots, 0)$	0.06	0.14	0.30	0.59	0.81	0.92	0.97
$\boldsymbol{\mu} = (0, 1, \ldots, 1)$	0.02	0.05	0.14	0.35	0.56	0.75	0.84

Table 4. Bias and MSE of the estimators of μ_0 for $\boldsymbol{\mu} = (0, 0, \ldots, 0)$.

	Bias				MSE			
k	RMLE	CPE	SE	ISE	RMLE	CPE	SE	ISE
1	−0.279	− 0.138	−0.138	−0.138	0.747	0.811	0.811	0.811
2	−0.444	−0.186	−0.248	−0.249	0.730	0.754	0.704	0.703
3	−0.557	−0.209	−0.327	−0.329	0.762	0.727	0.654	0.653
5	−0.718	−0.236	−0.441	−0.445	0.879	0.703	0.628	0.628
10	−0.953	−0.258	−0.603	−0.610	1.17	0.682	0.669	0.673
20	−1.18	−0.266	−0.755	−0.766	1.60	0.662	0.778	0.792
100	−1.85	−0.283	−0.763	−0.769	3.177	0.655	0.687	0.690

Table 5. Bias and MSE of the estimators of μ_0 for $\boldsymbol{\mu} = (0, 1, \ldots, 1)$.

	Bias				MSE			
k	RMLE	CPE	SE	ISE	RMLE	CPE	SE	ISE
1	−0.100	− 0.050	−0.050	−0.050	0.832	0.898	0.898	0.898
2	−0.171	−0.068	−0.085	−0.085	0.746	0.860	0.826	0.826
3	−0.223	−0.074	−0.108	−0.108	0.701	0.844	0.781	0.780
5	−0.304	−0.083	−0.145	−0.147	0.653	0.826	0.719	0.717
10	−0.439	−0.087	−0.199	−0.202	0.634	0.811	0.642	0.638
20	−0.598	−0.097	−0.262	−0.269	0.700	0.807	0.591	0.587
100	−1.119	−0.117	−0.268	−0.273	1.187	0.797	0.545	0.539

estimators is very essential. As k increases, these proportions are increased, but the growth rate is greater for the least favourable case $(0, \ldots, 0)$.

The bias and MSE values of the control group estimators for two considered configurations are presented in Tables 4 and 5, respectively. These values for the treatment group estimators are listed in Tables 6 and 7 for two patterns, respectively. From Table 4, it is clear that the proposed estimators (SE, ISE) perform better than the RMLE and CPE in terms of the MSEs. The MSE values for proposed estimators seem to drop much more rapidly, which reflect a rapid decline in the least favourable case. We note that for certain values of the parameters on the boundary points, the RMLE never performed better than the proposed estimators. This is similar to our conclusion made above (cf. Table 1). In Table 5, for small values of k, it is seen that the RMLE has the smallest MSE when $\boldsymbol{\mu}$ lies in the interior of the parameter space, but it reverses for larger values of k. Although the biases of two proposed estimators for the control group are slightly larger than those of the CPE, the MSE values of the proposed estimators appear to stabilize or decrease with k. For instance, in $k = 100$, these values for SE and ISE are 0.545 and 0.539, while for the RMLE and CPE are 1.187 and 0.797, respectively.

For the evaluation of treatment estimators by various methods for simplicity, we assume all treatment means μ_i, $i \geq 1$ to be equal. So, for comparing the treatment estimators we compare one treatment mean μ_1. In Table 6, the proposed estimators (SE, ISE) for treatment mean μ_1 provide the smallest MSEs, with relative improvement by increasing k. As

Table 6. Bias and MSE of the estimators of μ_1 for $\boldsymbol{\mu} = (0, 0, \ldots, 0)$.

	Bias				MSE			
k	RMLE	CPE	SE	ISE	RMLE	CPE	SE	ISE
1	0.283	0.424	0.424	0.424	0.743	0.807	0.807	0.807
2	0.217	0.398	0.346	0.345	0.717	0.761	0.716	0.716
3	0.185	0.393	0.306	0.304	0.713	0.740	0.676	0.675
5	0.143	0.384	0.253	0.251	0.728	0.718	0.650	0.650
10	0.096	0.378	0.193	0.190	0.778	0.702	0.658	0.660
20	0.061	0.375	0.146	0.143	0.830	0.694	0.688	0.692
100	0.052	0.368	0.133	0.130	0.849	0.683	0.680	0.682

Table 7. Bias and MSE of the estimators of μ_1 for $\boldsymbol{\mu} = (0, 1, \ldots, 1)$.

	Bias				MSE			
k	RMLE	CPE	SE	ISE	RMLE	CPE	SE	ISE
1	0.100	0.149	0.149	0.149	0.834	0.801	0.801	0.801
2	0.083	0.147	0.134	0.133	0.833	0.785	0.785	0.785
3	0.072	0.146	0.123	0.123	0.833	0.775	0.775	0.775
5	0.062	0.149	0.114	0.113	0.850	0.775	0.780	0.780
10	0.039	0.145	0.093	0.091	0.872	0.765	0.781	0.782
20	0.033	0.151	0.086	0.084	0.903	0.769	0.797	0.798
100	0.030	0.160	0.075	0.074	0.923	0.781	0.785	0.779

seen in this table, the CPE is more biased than the proposed estimators. Also, in Table 7 the proposed estimators compete very well with the competitors in terms of the biases and MSEs.

In most cases, both proposed estimators (SE, ISE) have substantially lower MSE than the RMLE and CPE and compete very well with these estimators in terms of the biases. From simulation results in Tables 4–7, it is seen that the proposed algorithm for derivation of the ISE is successful. Thus we conclude in practical settings that the overall performance of the ISE is likely to be superior with respect to any alternative estimators over a large portion of the parametric space.

6. Concluding remarks

In this article, motivated by the Lee's [4] counterexample, we presented the modified estimation of the control group parameter in the tree order restriction as a problem of estimating a single target parameter θ_0 in the presence of an increasing number of treatment parameters $\theta_1, \ldots, \theta_k$. To find an improvement over the URE (UMLE) and IRE (RMLE), we first consider the random device based on the allocated probability. The improved estimators (RE, SE) allocate a greater weight to the unrestricted sample estimator as compared with the alternative procedures. In particular, the appearance of the probability in the proposed estimators is essential for optimality. The simulation results show that the proposed estimators (RE, SE) are substantially better than the IRE (RMLE) and standard ones of URE (UMLE). Also, they compete very well with the other estimators, CPE and CSE.

In general, it can be argued based on the minimization operator as in (3), that the CSE, CPE and SE may not be satisfied in the isotonic property. It is also essential to investigate the proposed approach in terms of the important property such as isotonicity of the estimators. Thus, we extended the proposed decision approach by using an iteration algorithm to solve the general problem (the isotonic regression problem) described in Section 1. This property is shared by the isotonic smooth procedure. We used the iteration algorithm to obtain isotonic smooth estimators that improve upon the other alternatives which does not have the isotonic properties such as CPE and CSE.

Although the performance of the proposed estimators (SE, ISE) is impressive when $\boldsymbol{\theta}$ lies in the boundary point $(0, \ldots, 0)$, the gains are much more for the interior case $(0, 1, \ldots, 1)$. On the other hand, the performance of the proposed estimators for treatment groups is substantially better than the alternatives. This suggests the desirability of the allocated probability to the basic unrestricted control estimator. So, it is seen that the proposed random device is successful.

The isotonic smooth estimator based upon the iteration procedure seems to be more desirable in this study because of the stability of its biases and MSEs. Unlike the CPE and CSE, this estimator has the isotonic property. It appears that the CSE and CPE procedures lose the isotonic regression property. The proposed method in this paper can be generalized to the other distributions in which the support of the underlying distributions or the distribution of the unrestricted estimators are at least unbounded exactly on the lower bound, and admit the divergency of the minimum location parameter to the lower bound of the natural parameter space.

Acknowledgements

We would like to thank the reviewer for carefully reading this article and giving several exact and constructive comments and valuable suggestions for the better presentation of the article.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Robertson T, Wright FT, Dykstra RL. Order restricted statistical inference. New York: John Wiley; 1988.
[2] Barlow RE, Bartholomew DJ, Bremner JM, Brunk HD. Statistical inference under order restrictions: the theory and application of isotonic regression. New York: Wiley; 1972.
[3] Silvapulle MJ, Sen PK. Constrained statistical inference. Hoboken: John Wiley; 2005.
[4] Lee CC. The quadratic loss of order restricted estimations for several treatment means and a control mean. Ann Stat. 1988;16:751–758.
[5] Hwang JTG, Peddada SD. Confidence interval estimation subject to order restrictions. Ann Stat. 1994;22:67–93.
[6] Lee CC. The quadratic loss of isotonic regression under normality. Ann Stat. 1981;11:467–477.
[7] Cohen A, Sackrowitz HB. Inference for the model of several treatments and a control. J Stat Plan Inference. 2002;107:89–101.
[8] Chaudhuri S, Perlman MD. On the bias and mean-square error of order-restricted maximum likelihood estimators. J Stat Plan Inference. 2005;130:229–250.
[9] Betcher J, Peddada SD. Statistical inference under order restrictions in analysis of covariance using a modified restricted maximum likelihood estimator. Sankhyā Indian J Stat. 2009;71,B(Part I):79–96.
[10] Tan X, Peddada SD. Asymptotic distribution of some estimators for parameters subject to order restrictions. Stat Appl. 2000;2:7–25.

[11] Chaudhuri S, Drton M. On the bias and mean-square error of the sample minimum and the maximum likelihood estimator for two ordered normal means. Seattle: Department of statistics, University of Washington; 2003 (Technical report No. 432).
[12] Berger JO. Statistical decision theory and Bayesian analysis. 2nd ed. New York: Springer; 1985.
[13] Azzalini A. A class of distributions which includes the normal ones. Scand J Stat. 1985;12:171–178.
[14] Lehmann EL, Casella G. Theory of point estimation. 2nd ed. Springer-Verlag New York, Inc; 1998.
[15] Birnbaum ZW. On random variables with comparable peakedness. Ann Math Stat. 1948;19:76–81.