AISM (2006) 58: 739–756
DOI 10.1007/s10463-005-0025-5

Youhei Oono · Nobuo Shinozaki

Estimation of error variance in ANOVA model and order restricted scale parameters

Received: 5 August 2004 / Revised: 6 June 2005 / Published online: 30 August 2006

Abstract We consider the estimation of error variance in the analysis of experiments using two level orthogonal arrays. We address the estimator which is the minimum of all the estimators which we obtain by pooling some sums of squares for factorial effects. Under squared error loss, we discuss whether or not this estimator uniformly improves upon the best positive multiple of error sum of squares. We show that when we have two factorial effects, we obtain uniform improvement. However, we show that when we have more than two factorial effects, we cannot necessarily obtain uniform improvement. Further, the above results are applied to the problem of estimating the smallest scale parameter of chi-square distributions.

Keywords Two-level orthogonal arrays · Stein's estimator · Squared error loss · Uniform improvement · Simple tree order restriction · Isotonic regression estimator · Random effects model

1 Introduction

We consider the estimation of error variance σ^2 based on experiments using two-level orthogonal arrays. Let each of p factorial effects be assigned to one column and the error term to ν_0 columns. Let S_i be the sum of squares for the ith factorial effect and let S_0 be that for the error term. Assume that random errors are independently distributed as $N(0, \sigma^2)$. Then S_0 and $S_i,\ i = 1, \ldots, p$ are independently distributed as $\sigma^2\chi^2_{\nu_0}$ and $\sigma^2\chi^2_1(\lambda_i),\ i = 1, 2, \ldots, p$ respectively, where $\chi^2_{\nu_0}$ denotes

Y. Oono (✉)
School of Science for OPEN and Environmental Systems, Graduate School of Science and Technology, Keio University, 3-14-1, Hiyoshi, Kohoku, Yokohama 223-8522, Japan
E-mail: oono@ae.keio.ac.jp

N. Shinozaki
Department of Administration Engineering, Faculty of Science and Technology, Keio University, 3-14-1, Hiyoshi, Kohoku, Yokohama 223-8522, Japan

a central χ^2 distribution with ν_0 dfs, and $\chi_1^2(\lambda_i)$ a non-central χ^2 distribution with 1 df and noncentrality parameter λ_i. Note that $\lambda_i = 0$ implies that the ith factorial effect is inactive.

When we estimate σ^2 under the squared error loss

$$L(\sigma^2, \hat{\sigma^2}) = \left(\hat{\sigma^2} - \sigma^2\right)^2, \tag{1}$$

it is well-known that the estimator

$$\delta_0 = \frac{S_0}{(\nu_0 + 2)} \tag{2}$$

is the best among positive multiples of S_0. Stein (1964) showed that for the case $p = 1$, $\gamma_1 = \min\{S_0/(\nu_0 + 2), (S_0 + S_1)/(\nu_0 + 3)\}$ uniformly improves upon δ_0. Brown (1968) and Brewster and Zidek (1974) generalized Stein's result for the case $p = 1$. Further, Gelfand and Dey (1988) generalized Stein's result for a class of nested linear models and showed that for the case $p \geq 2$

$$\gamma_p = \min\left(\frac{S_0}{\nu_0 + 2}, \frac{S_0 + S_1}{\nu_0 + 3}, \frac{S_0 + S_1 + S_2}{\nu_0 + 4}, \ldots, \frac{S_0 + \sum_{i=1}^p S_i}{\nu_0 + p + 2}\right) \tag{3}$$

uniformly improves upon δ_0. See Oono and Shinozaki (2006) for a related result. One may also refer to Maatta and Casella (1990) for tracing the history of developments in decision-theoretic variance estimation, starting with Stein (1964)'s discovery.

In Gelfand and Dey (1988)'s estimator, the pooling order of $S_i,\ i = 1, \ldots, p$ must be determined in advance of observing data, that is in the order $S_1, S_2, \ldots, S_p$. However in ANOVA model, it is usual that the pooling order is not determined in advance. For instance, one may test whether each factorial effect is active or not and then pool sums of squares corresponding to all nonsignificant effects with error sum of squares and obtain an estimator of σ^2. Nagata (1989) showed by a Monte Carlo simulation study that for $p = 2$ one estimator of this type has a good performance as compared with the unbiased estimator when the significance level of the preliminary test is 0.50.

Here we address an estimator

$$\delta_p = \min\left(\frac{S_0}{\nu_0 + 2}, \frac{S_0 + S_{(1)}}{\nu_0 + 3}, \frac{S_0 + S_{(1)} + S_{(2)}}{\nu_0 + 4}, \ldots, \frac{S_0 + \sum_{i=1}^p S_{(i)}}{\nu_0 + p + 2}\right), \tag{4}$$

where $S_{(i)},\ i = 1, \ldots, p$, $(S_{(1)} \leq S_{(2)} \leq \ldots \leq S_{(p)})$ denote the order statistics of $S_i,\ i = 1, \ldots, p$. Unlike γ_p, in δ_p the pooling order of $S_i,\ i = 1, \ldots, p$ is not determined in advance. However, we should remark that δ_p is not precisely interpreted as a preliminary test estimator, since it is not decided stepwise whether we pool $S_{(i)}$ or not. Oono and Shinozaki (2004) have addressed an estimator of the form

$$\zeta_p = \frac{S_0}{\nu_0 + 2} - \frac{1}{\nu_0 + p + 2} \sum_{i=1}^p \left(\frac{S_0}{\nu_0 + 2} - S_i\right)^+, \tag{5}$$

where $x^+ = \max(x, 0)$, and have shown that for $p \le 2$, ζ_p uniformly improves upon δ_0 but that for $p \ge 3$, ζ_p uniformly improves upon δ_0 only when $p \le 6$ and ν_0 is small.

We should remark that George (1990) also mentioned δ_p as one generalization of Stein (1964)'s estimator for the case $p \ge 2$. Kubokawa et al. (1993) derived the asymptotic risk expansion for δ_2 and analytically demonstrated that δ_2 is asymptotically better than δ_0. However it has not been well established whether or not δ_p uniformly improves upon δ_0 for $p \ge 2$ so far.

In Sect. 2, we discuss whether or not δ_p uniformly improves upon δ_0. We show that for $p = 2$, δ_p uniformly improves upon δ_0. However we show partially through numerical evaluation that δ_p does not uniformly improve upon δ_0 for $p \ge 12$ when $\nu_0 = 1$, for $p \ge 5$ when $2 \le \nu_0 \le 3$, for $p \ge 4$ when $4 \le \nu_0 \le 12$ and for $p \ge 3$ when $13 \le \nu_0 \le 20$.

In Sect. 3, the results of Sect. 2 are applied to the estimation of the smallest scale parameter of χ^2 distributions. Several authors have studied the estimation of order restricted scale parameters of gamma distributions. See, for example, Kushary and Cohen (1989), Kaur and Singh (1991), Vijayasree and Singh (1993), Hwang and Peddada (1994), Iliopoulos and Kourouklis (2000), Chang and Shinozaki (2002) and Oono (2005). Some other related researches can be traced through the bibliography of Kourouklis (2001).

Let V_0 and V_i, $i = 1, \dots, p$ be independently distributed as $\sigma_0^2 \chi_{\nu_0}^2$ and $\sigma_i^2 \chi_1^2$, $i = 1, \dots, p$. Assume that it is known that σ_i^2's are subject to the simple tree order restriction

$$\sigma_0^2 \le \sigma_j^2, \quad j = 1, \dots, p. \tag{6}$$

The above setup arises naturally when considering the additive random effects model. See, for example, Sect. 3.5 of Lehmann and Casella (1998). For simplicity, let us consider the random effects two-way layout

$$X_{ijk} = \mu + A_i + B_j + \epsilon_{ijk}, \quad i = 1,\ 2, \quad j = 1,\ 2, \quad k = 1, \dots, n. \tag{7}$$

Assume that the unobservable random effects A_i, B_j and the error term ϵ_{ijk} are independently distributed as $N(0, \sigma_A^2)$, $N(0, \sigma_B^2)$ and $N(0, \sigma_\epsilon^2)$. Let $V_0 = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^n (X_{ijk} - \bar{X}_{i\cdot\cdot} - \bar{X}_{\cdot j\cdot} + \bar{X}_{\cdots})^2$, $V_1 = 2n \sum_{i=1}^2 (\bar{X}_{i\cdot\cdot} - \bar{X}_{\cdots})^2$ and $V_2 = 2n \sum_{j=1}^2 (\bar{X}_{\cdot j\cdot} - \bar{X}_{\cdots})^2$, where $\bar{X}_{i\cdot\cdot} = \sum_{j=1}^2 \sum_{k=1}^n X_{ijk}/(2n)$, $\bar{X}_{\cdot j\cdot} = \sum_{i=1}^2 \sum_{k=1}^n X_{ijk}/(2n)$ and $\bar{X}_{\cdots} = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^n X_{ijk}/(4n)$. Then V_0, V_1 and V_2 are independently distributed as $\sigma_0^2 \chi_{\nu_0}^2$, $\sigma_1^2 \chi_1^2$ and $\sigma_2^2 \chi_1^2$ with $\nu_0 = 4n - 3$, $\sigma_0^2 = \sigma_\epsilon^2$, $\sigma_1^2 = \sigma_\epsilon^2 + 2n\sigma_A^2$ and $\sigma_2^2 = \sigma_\epsilon^2 + 2n\sigma_B^2$, and we have the simple tree order restriction Eq. (6) with $p = 2$.

When we consider the estimation of σ_0^2 under squared error loss, $V_0/(\nu_0 + 2)$ is the best among positive multiples of V_0. However, since the information (Eq. 6) is available, a reasonable estimator of σ_0^2 may be the isotonic regression of $\{V_0/(\nu_0 + 2),\ V_1, \dots,\ V_p\}$ with weights $\{\nu_0 + 2,\ 1, \dots,\ 1\}$, that is

$$\hat{\sigma_0^2}^{\mathrm{ST}} = \min\left(\frac{V_0}{\nu_0 + 2}, \frac{V_0 + V_{(1)}}{\nu_0 + 3}, \frac{V_0 + V_{(1)} + V_{(2)}}{\nu_0 + 4}, \dots, \frac{V_0 + \sum_{i=1}^p V_{(i)}}{\nu_0 + p + 2} \right), \tag{8}$$

where $V_{(i)},\ i = 1, \ldots, p$ denote the order statistics of $V_i,\ i = 1, \ldots, p$. See Bartholomew et al. (1972) or Roberson et al. (1988) as for the construction of isotonic regression estimators. Here, we are interested in whether or not $\hat{\sigma}_0^{2\,\mathrm{ST}}$ uniformly improves upon $V_0/(\nu_0 + 2)$ under squared error loss. We show that $\hat{\sigma}_0^{2\,\mathrm{ST}}$ uniformly improves upon $V_0/(\nu_0 + 2)$ for $p = 2$ but that $\hat{\sigma}_0^{2\,\mathrm{ST}}$ does not uniformly improve upon $V_0/(\nu_0 + 2)$ for larger p.

2 Estimation of error variance in ANOVA model

In this Section, we discuss whether or not δ_p uniformly improves upon δ_0 under squared error loss. We discuss this problem for the case when $p = 2$ in Sect. 2.1 and for the case when $p \geq 3$ in Sect. 2.2 separately.

2.1 The case when $p = 2$

Here, we show that δ_2 uniformly improves upon δ_0. The following well-known Lemma, which can be obtained by integration by parts method, is very useful to evaluate the risk difference of δ_0 and δ_2. See Efron and Morris (1976) or Shinozaki (1995).

Lemma 2.1 *Let T be distributed as χ_n^2 and let $f(\cdot)$ be an absolutely continuous function. Then $E[Tf(T)] = nE[f(T)] + 2E[Tf^{'}(T)]$, provided that both expectations exist.*

Let $\mathcal{J}_i^2$ be the set of (S_0, S_1, S_2) such that $\delta_2 = (S_0 + S_i)/(\nu_0 + 3)$ and let $\mathcal{J}_{12}^2$ be the set of (S_0, S_1, S_2) such that $\delta_2 = (S_0 + S_1 + S_2)/(\nu_0 + 4)$. Further let J_i^2 (or J_{12}^2) be the indicator function of the set $\mathcal{J}_i^2$ (or $\mathcal{J}_{12}^2$). Then δ_2 can be written as

$$\delta_2 = \frac{S_0}{\nu_0 + 2} - g(U_1, U_2), \tag{9}$$

where

$$g(x_1, x_2) = \frac{1}{\nu_0 + 3}(x_1 J_1^2 + x_2 J_2^2) + \frac{1}{\nu_0 + 4}(x_1 + x_2) J_{12}^2 \tag{10}$$

and $U_i = S_0/(\nu_0 + 2) - S_i$. Now we evaluate the risk difference of δ_0 and δ_2. Without loss of generality we set $\sigma^2 = 1$. Let us denote the risk when we estimate σ^2 by $\hat{\sigma}^2$ as $R(\sigma^2, \hat{\sigma}^2) = E[L(\sigma^2, \hat{\sigma}^2)]$. Then from Eq. (9), we have the risk difference as

$$R(\sigma^2, \delta_0) - R(\sigma^2, \delta_2) = 2E\left[\left(\frac{S_0}{\nu_0 + 2} - 1\right) g(U_1, U_2)\right] - E\left[\{g(U_1, U_2)\}^2\right]. \tag{11}$$

To evaluate the first term on the right-hand side of Eq. (11), we apply Lemma 2.1 with $T = S_0$ and $f(T) = g(U_1, U_2)$, and we have

$$\begin{aligned} E[S_0 g(U_1, U_2)] &= \nu_0 E[g(U_1, U_2)] + 2E\left[S_0 \frac{g(1,1)}{\nu_0 + 2}\right] \\ &= (\nu_0 + 2)E[g(U_1, U_2)] + 2E\left[g(S_1, S_2)\right]. \end{aligned} \tag{12}$$

Thus we have

$$\begin{aligned} &2E\left[\left(\frac{S_0}{\nu_0 + 2} - 1\right) g(U_1, U_2)\right] \\ &= \frac{4}{\nu_0 + 2}\left\{\frac{1}{\nu_0 + 3}\left(E\left[S_1 J_1^2\right] + E\left[S_2 J_2^2\right]\right) + \frac{1}{\nu_0 + 4} E\left[(S_1 + S_2) J_{12}^2\right]\right\}. \end{aligned} \tag{13}$$

To evaluate the second term on the right-hand side of Eq. (11), we utilize the inequality

$$\begin{aligned} \{g(U_1, U_2)\}^2 \le{}& \frac{2(\nu_0 + 3)}{(\nu_0 + 2)(\nu_0 + 4)^2} \\ &\times \left\{\left(\frac{S_0}{\nu_0 + 2} - S_1\right) g_1(U_1, U_2) + \left(\frac{S_0}{\nu_0 + 2} - S_2\right) g_2(U_1, U_2)\right\}, \end{aligned} \tag{14}$$

where

$$g_1(x_1, x_2) = \frac{(\nu_0 + 2)(\nu_0 + 4)}{(\nu_0 + 3)^2} x_1 J_1^2 + \left(x_1 - \frac{x_2}{\nu_0 + 3}\right) J_{12}^2 \tag{15}$$

and

$$g_2(x_1, x_2) = \frac{(\nu_0 + 2)(\nu_0 + 4)}{(\nu_0 + 3)^2} x_2 J_2^2 + \left(x_2 - \frac{x_1}{\nu_0 + 3}\right) J_{12}^2. \tag{16}$$

The inequality (14) can be confirmed since one needs to add

$$\frac{\nu_0 + 2}{(\nu_0 + 3)^2(\nu_0 + 4)}(U_1^2 J_1^2 + U_2^2 J_2^2) + \frac{1}{(\nu_0 + 2)(\nu_0 + 4)}(U_1 - U_2)^2 J_{12}^2, \tag{17}$$

which is clearly nonnegative, to $\{g(U_1, U_2)\}^2$ to obtain the right-hand side of Eq. (14). Note that $g_1(U_1, U_2)$ and $g_2(U_1, U_2)$ are absolutely continuous functions of S_0, S_1 and S_2. To evaluate the expectation of Eq. (14), we introduce auxiliary random variables $K_i,\ i = 1,\ 2$ distributed independently as Poisson distribution with mean λ_i such that K_i is independent of S_0, and S_i given K_i is distributed as $\sigma^2 \chi^2_{1+2K_i}$. From Lemma 2.1, we evaluate the expectation of each term on the right-hand side of Eq. (14) as

$$E[S_0 g_i(U_1, U_2)] = (\nu_0 + 2)E[g_i(U_1, U_2)] + 2E[g_i(S_1, S_2)], \tag{18}$$

$$\begin{aligned} &E[S_1 g_1(U_1, U_2) \mid K_1, K_2] \\ &\quad = (1 + 2K_1)E\left[g_1(U_1, U_2) \mid K_1, K_2\right] - 2E[g_1(S_1, 0) \mid K_1, K_2] \end{aligned} \tag{19}$$

and

$$E[S_2 g_2(U_1, U_2) \mid K_1, K_2] \\ = (1 + 2K_2)E\left[g_2(U_1, U_2) \mid K_1, K_2\right] - 2E[g_2(0, S_2) \mid K_1, K_2]. \quad (20)$$

Using Eqs. (14), (18), (19) and (20) we have

$$E[\{g(U_1, U_2)\}^2] \le \frac{2(\nu_0 + 3)}{(\nu_0 + 2)(\nu_0 + 4)^2} E\left[\frac{2}{\nu_0 + 2}\{g_1(S_1, S_2) + g_2(S_1, S_2)\} \right. \\ \left. + 2\{g_1(S_1, 0) + g_2(0, S_2)\}\right] \quad (21)$$

$$= \frac{4}{(\nu_0 + 2)(\nu_0 + 4)} \left\{\left(E\left[S_1 J_1^2\right] + E\left[S_2 J_2^2\right]\right) \right. \\ \left. + E\left[(S_1 + S_2) J_{12}^2\right]\right\}.$$

Thus we see from Eqs. (11), (13) and (21) that

$$R(\sigma^2, \delta_0) - R(\sigma^2, \delta_2) \ge \frac{4}{(\nu_0 + 2)(\nu_0 + 3)(\nu_0 + 4)} \left(E\left[S_1 J_1^2\right] + E\left[S_2 J_2^2\right]\right), \quad (22)$$

which is clearly positive. Summarizing the above we have the following Theorem.

Theorem 2.1 *δ_2 uniformly improves upon δ_0 under squared error loss.*

2.2 The case when $p \ge 3$

We discuss whether or not δ_p uniformly improves upon δ_0 for the case $p \ge 3$. We should mention that Oono and Shinozaki (2004) have shown that the case when $\lambda_i = 0,\ i = 1, \ldots, p$ is the most critical one for ζ_p to improve upon δ_0 uniformly in the sense that ζ_p uniformly improves upon δ_0 if and only if $R(\sigma^2, \zeta_p) \le R(\sigma^2, \delta_0)$ when $\lambda_i = 0,\ i = 1, \ldots, p$. This case may also be the most critical one for δ_p to improve upon δ_0 uniformly, since in this case δ_p is stochastically smallest and may shrink δ_0 too much. Here we evaluate the risk difference of δ_0 and δ_p only for the case when $\lambda_i = 0,\ i = 1, \ldots, p$, and show that δ_p does not uniformly improve upon δ_0 for larger p. Let $\{i_1, \ldots, i_l\}$ be a subset of the set $\{1, \ldots, p\}$ and let $\mathcal{J}^p_{i_1 \ldots i_l}$ be the set of $(S_0, S_1, \ldots, S_p)$ such that $\delta_p = (S_0 + \sum_{j=1}^{l} S_{i_j})/(\nu_0 + l + 2)$. Further, let $J^p_{i_1 \cdots i_l}$ be the indicator function of the set $\mathcal{J}^p_{i_1 \cdots i_l}$. Then δ_p can be written as

$$\delta_p = \frac{S_0}{\nu_0 + 2} - \frac{1}{\nu_0 + p + 2} h(U_1, \ldots, U_p), \quad (23)$$

where

$$h(x_1, \ldots, x_p) = \sum_{l=1}^{p} \frac{\nu_0 + p + 2}{\nu_0 + l + 2} \sum_{\{i_1, \ldots, i_l\}} \left(\sum_{j=1}^{l} x_{i_j}\right) J^p_{i_1 \cdots i_l} \quad (24)$$

and $U_i = S_0/(\nu_0+2) - S_i$. We note that the summation $\sum_{\{i_1,\ldots,i_l\}}$ is taken over arbitrary subset $\{i_1,\ldots,i_l\}$ of the set $\{1,\ldots,p\}$. Without loss of generality we set $\sigma^2 = 1$. Let R_0 and E_0 denote the risk and the expectation both when $\lambda_i = 0,\ i = 1,\ldots,p$. Then we have from Eq. (23)

$$\begin{aligned} R_0(\sigma^2,\delta_0) - R_0(\sigma^2,\delta_p) &= \frac{2}{\nu_0+p+2} E_0\left[\left(\frac{S_0}{\nu_0+2} - 1\right) h(U_1,\ldots,U_p)\right] \\ &\quad - \frac{1}{(\nu_0+p+2)^2} E_0\left[\left\{h(U_1,\ldots,U_p)\right\}^2\right]. \end{aligned} \tag{25}$$

Similarly with Eq. (12), we have from Lemma 2.1

$$E_0[S_0 h(U_1,\ldots,U_p)] = (\nu_0+2) E_0[h(U_1,\ldots,U_p)] + 2E_0\left[h(S_1,\ldots,S_p)\right]. \tag{26}$$

Applying Eq. 26 to the first term on the right-hand side of Eq. (25), we have

$$\begin{aligned} R_0(\sigma^2,\delta_0) - R_0(\sigma^2,\delta_p) &= \frac{4}{(\nu_0+2)(\nu_0+p+2)} E_0\left[h(S_1,\ldots,S_p)\right] \\ &\quad - \frac{1}{(\nu_0+p+2)^2} E_0\left[\left\{h(U_1,\ldots,U_p)\right\}^2\right]. \end{aligned} \tag{27}$$

To evaluate the right-hand side of Eq. (27), we need the following Lemmas 2.3, 2.4 and 2.5. Lemma 2.2 is used to show Lemma 2.3. The proofs of these Lemmas are rather technical and we give them in Appendix A.

Lemma 2.2 *Let $\mathcal{L}^p_{i_1\cdots i_l}$ be the set of $(S_0,\ldots,S_p)$ such that $S_0/(\nu_0+2) \ge S_j$ if and only if $j \in \{i_1,\ldots,i_l\}$. If $(S_0,\ldots,S_p) \in \mathcal{J}^p_{i_1\cdots i_l}$, then $(S_0,\ldots,S_p) \in \mathcal{L}^p_{i_1\cdots i_h}$ for some $\{i_1,\ldots,i_h\} \supseteq \{i_1,\ldots,i_l\}$.*

Lemma 2.3 *Let $h(\cdot,\ldots,\cdot)$ be defined as in Eq. (24). Further, let $L^p_{i_1\cdots i_l}$ be the indicator function of the set $\mathcal{L}^p_{i_1\cdots i_l}$ and let*

$$h_1(x_1,\ldots,x_p) = \sum_{l=1}^{p} \frac{\nu_0+p+2}{\nu_0+l+2} \sum_{\{i_1,\ldots,i_l\}} \left(\sum_{j=1}^{l} x_{i_j}\right) L^p_{i_1\cdots i_l}. \tag{28}$$

Then (i) $h(S_1,\ldots,S_p) \le h_1(S_1,\ldots,S_p)$ and (ii) $h(U_1,\ldots,U_p) \ge h_1(U_1,\ldots,U_p)$.

Lemma 2.4 *Let*

$$h_2(x_1,\ldots,x_p) = \sum_{l=1}^{p-1} \frac{p-l}{\nu_0+l+2} \sum_{\{i_1,\ldots,i_l\}} \left(\sum_{j=1}^{l} x_{i_j}\right) L^p_{i_1\cdots i_l} \tag{29}$$

and

$$h_3(x_1,\ldots,x_p) = \sum_{l=1}^{p-1} \left\{\left(\frac{\nu_0+p+2}{\nu_0+l+2}\right)^2 - 1\right\} \sum_{\{i_1,\ldots,i_l\}} \left(\sum_{j=1}^{l} x_{i_j}\right)^2 L^p_{i_1\cdots i_l}. \tag{30}$$

Then for $p \geq 3$*,*

$$h_1(S_1, \dots, S_p) = \sum_{i=1}^{p} S_i I_{\frac{S_0}{\nu_0+2} > S_i} + h_2(S_1, \dots, S_p) \tag{31}$$

and

$$\{h_1(U_1, \dots, U_p)\}^2 = \left(\sum_{i=1}^{p} U_i^+ \right)^2 + h_3(U_1, \dots, U_p), \tag{32}$$

where I_C *is the indicator function of a set* C *and* $a^+ = \max(0, a)$.

Lemma 2.5 *For* $p \geq 3$*,*

$$E_0[h_2(S_1, \dots, S_p)] \leq \frac{p(p-1)}{\nu_0 + p + 1} \left\{ E_0[S_1 L_1^2] + \frac{p-2}{\nu_0 + 3} E_0[S_1 L_1^3] \right\} \tag{33}$$

and

$$\begin{aligned} &E_0[h_3(U_1, \dots, U_p)] \\ &\quad \geq p(p-1) \left\{ \left(\frac{\nu_0 + p + 2}{\nu_0 + p + 1} \right)^2 - 1 \right\} \left\{ E_0[U_1^2 L_1^2] + (p-2) E_0[U_1 U_2 L_{12}^3] \right\}. \end{aligned} \tag{34}$$

We have from Lemmas 2.3, 2.4 and 2.5,

$$\begin{aligned} E_0[h(S_1, \dots, S_p)] &\leq E_0[h_1(S_1, \dots, S_p)] \\ &= \sum_{i=1}^{p} E_0 \left[S_i I_{\frac{S_0}{\nu_0+2} > S_i} \right] + E_0[h_2(S_1, \dots, S_p)] \\ &\leq \sum_{i=1}^{p} E_0 \left[S_i I_{\frac{S_0}{\nu_0+2} > S_i} \right] + \frac{p(p-1)}{\nu_0 + p + 1} \\ &\qquad \times \left\{ E_0[S_1 L_1^2] + \frac{p-2}{\nu_0 + 3} E_0[S_1 L_1^3] \right\}. \end{aligned} \tag{35}$$

Similarly we have

$$\begin{aligned} E_0[\{h(U_1, \dots, U_p)\}^2] &\geq E_0 \left[\left(\sum_{i=1}^{p} U_i^+ \right)^2 \right] \\ &+ p(p-1) \left\{ \left(\frac{\nu_0 + p + 2}{\nu_0 + p + 1} \right)^2 - 1 \right\} \left\{ E_0[U_1^2 L_1^2] + (p-2) E_0[U_1 U_2 L_{12}^3] \right\}. \end{aligned} \tag{36}$$

As shown in Oono and Shinozaki (2004), we can easily confirm from Lemma 2.1 that

$$\begin{aligned} R_0(\sigma^2, \delta_0) - R_0(\sigma^2, \zeta_p) &= \frac{4}{(\nu_0+2)(\nu_0+p+2)} \sum_{i=1}^{p} E_0\left[S_i I_{\frac{S_0}{\nu_0+2} > S_i}\right] \\ &\quad - \frac{1}{(\nu_0+p+2)^2} E_0\left[\left(\sum_{i=1}^{p} U_i^+\right)^2\right] \\ &= \frac{2p(\nu_0+2p+1)}{(\nu_0+p+2)^2(\nu_0+2)} E_0\left[S_1 I_{\frac{S_0}{\nu_0+2} > S_1}\right] \\ &\quad - \frac{p(p-1)}{(\nu_0+p+2)^2} E_0[U_1^+ U_2^+]. \end{aligned} \tag{37}$$

Applying Eq. (35) and (36) to Eq. (27) and noting the first equality of Eq. (37), we evaluate the risk difference as

$$\begin{aligned} & R_0(\sigma^2, \delta_0) - R_0(\sigma^2, \delta_p) \\ & \le R_0(\sigma^2, \delta_0) - R_0(\sigma^2, \zeta_p) + \frac{4p(p-1)}{(\nu_0+2)(\nu_0+p+1)(\nu_0+p+2)} \\ & \times \left\{ E_0[S_1 L_1^2] + \frac{p-2}{\nu_0+3} E_0[S_1 L_1^3] \right\} - \frac{p(p-1)}{(\nu_0+p+2)^2} \\ & \times \left\{ \left(\frac{\nu_0+p+2}{\nu_0+p+1}\right)^2 - 1 \right\} \left\{ E_0[U_1^2 L_1^2] + (p-2) E_0[U_1 U_2 L_{12}^3] \right\}. \end{aligned} \tag{38}$$

If the right-hand side of Eq. (38) is negative, then δ_p does not uniformly improve upon δ_0. For $1 \le \nu_0 \le 20$, using Mathematica, we have numerically evaluated the values of $E_0[S_1 I_{\frac{S_0}{\nu_0+2} > S_1}]$, $E_0[U_1^+ U_2^+]$, $E_0[S_1 L_1^2]$, $E_0[U_1^2 L_1^2]$, $E_0[S_1 L_1^3]$ and $E_0[U_1 U_2 L_{12}^3]$ in Table 1. Based on Table 1 and the inequality (38) and noting the second equality of (37), we can numerically confirm the following for $2 \le \nu_0 \le 20$.

Result 2.1 δ_p does not uniformly improve upon δ_0 for $5 \le p \le 25$ when $2 \le \nu_0 \le 3$, for $4 \le p \le 25$ when $4 \le \nu_0 \le 12$, and for $3 \le p \le 25$ when $13 \le \nu_0 \le 20$.

When $\nu_0 = 1$, the numerical value of the right-hand side of Eq. (38) is positive for $p \ge 3$, and we can not determine whether or not δ_p uniformly improves upon δ_0 based on Table 1 and the inequality (38) unfortunately. Further, similar remark applies to the case when p is large. However, as formally stated in the following Proposition, δ_p does not uniformly improve upon δ_0 for large p. The proof is rather technical and we give it in Appendix B.

Proposition 2.1 *δ_p does not uniformly improve upon δ_0 for $p \ge 12$ when $\nu_0 = 1$, for $p \ge 10$ when $\nu_0 = 2$, for $p \ge 9$ when $3 \le \nu_0 \le 4$, and for $p \ge 24$ when $\nu_0 \ge 5$.*

Combining Result 2.1 and Proposition 2.1, we see that δ_p does not uniformly improve upon δ_0 for $p \ge 12$ when $\nu_0 = 1$, for $p \ge 5$ when $2 \le \nu_0 \le 3$, for $p \ge 4$ when $4 \le \nu_0 \le 12$ and for $p \ge 3$ when $13 \le \nu_0 \le 20$. We should mention that

Table 1 Numerical evaluation 1

ν_0	$E_0[S_1 I_{\frac{S_0}{\nu_0+2}>S_1}]$	$E_0[U_1^+ U_2^+]$	$E_0[S_1 L_1^2]$	$E_0[U_1^2 L_1^2]$	$E_0[S_1 L_1^3]$	$E_0[U_1 U_2 L_{12}^3]$
1	0.057669	0.107241	0.020089	0.033023	0.008656	0.020310
2	0.089443	0.151270	0.030652	0.052254	0.012576	0.031881
3	0.109551	0.174162	0.037112	0.064870	0.014695	0.039361
4	0.123417	0.187827	0.041453	0.073793	0.015978	0.044595
5	0.133555	0.196763	0.044563	0.080441	0.016818	0.048464
6	0.141289	0.202994	0.046898	0.085587	0.017401	0.051440
7	0.147384	0.207553	0.048714	0.089690	0.017824	0.053802
8	0.152310	0.211015	0.050165	0.093037	0.018142	0.055720
9	0.156375	0.213722	0.051351	0.095821	0.018388	0.057311
10	0.159785	0.215891	0.052339	0.098172	0.018583	0.058650
11	0.162688	0.217664	0.053174	0.100184	0.018740	0.059794
12	0.165188	0.219137	0.053888	0.101926	0.018869	0.060782
13	0.167364	0.220379	0.054507	0.103448	0.018977	0.061644
14	0.169275	0.221439	0.055048	0.104790	0.019068	0.062403
15	0.170966	0.222353	0.055244	0.105983	0.019145	0.063076
16	0.172474	0.223149	0.055948	0.107049	0.019212	0.063677
17	0.173827	0.223848	0.056326	0.108007	0.019270	0.064217
18	0.175048	0.224466	0.056666	0.108874	0.019321	0.064704
19	0.176154	0.225016	0.056974	0.109662	0.019365	0.065147
20	0.177162	0.225509	0.057254	0.110381	0.019405	0.065551

we may be able to confirm that δ_p does not uniformly improve upon δ_0 for $p \geq 3$ also when $\nu_0 > 20$ by numerically evaluating the value of the right-hand side of Eq. (38) and combining the result with Proposition 2.1.

We remark that it is implied by our Monte Carlo simulation study over ten million iterations for the case when $\lambda_i = 0,\ i = 1, \ldots, p$ that δ_3 does not uniformly improve upon δ_0 also when $1 \leq \nu_0 \leq 12$.

3 Estimation of the smallest scale parameter

Let V_0 and $V_i,\ i = 1, \ldots, p$ be independently distributed as $\sigma_0^2 \chi_{\nu_0}^2$ and $\sigma_i^2 \chi_1^2,\ i = 1, \ldots, p$ respectively. Assume that σ_i^2's are subject to the simple tree order restriction Eq. (6). Here we consider the estimation of the smallest scale parameter σ_0^2 and discuss whether or not the isotonic regression estimator $\hat{\sigma_0^2}^{\mathrm{ST}}$ as defined in Eq. (8) uniformly improves upon $V_0/(\nu_0 + 2)$ under squared error loss. We first show that for $p = 2$, $\hat{\sigma_0^2}^{\mathrm{ST}}$ uniformly improves upon $V_0/(\nu_0 + 2)$ by using Theorem 2.1 and the following well-known Lemma.

Lemma 3.1 *Let V_i be distributed as $\sigma_i^2 \chi_{\nu_i}^2$, where $\sigma_i^2 \geq \sigma_0^2$. Then there exists an auxiliary random variable W_i satisfying the following two conditions. (a) V_i given W_i is distributed as $\sigma_0^2 \chi_{\nu_i}^2(W_i)$. (b) W_i is distributed as $\tau_i^2/(2\sigma_0^2)\chi_{\nu_i}^2$, where $\tau_i^2 = \sigma_i^2 - \sigma_0^2$.*

Theorem 3.1 *For the case $p = 2$, $\hat{\sigma_0^2}^{\mathrm{ST}}$ uniformly improves upon $V_0/(\nu_0 + 2)$ under squared error loss.*

Proof From Lemma 3.1, we can imagine auxiliary independent random variables $W_i,\ i = 1,\ 2$ such that V_0 and $V_i,\ i = 1,\ 2$ given $W_i,\ i = 1,\ 2$ are independently distributed as $\sigma_0^2 \chi_{\nu_0}^2$ and $\sigma_0^2 \chi_1^2(W_i),\ i = 1,\ 2$ respectively. Given W_i's, by applying Theorem 2.1 with $S_i = V_i,\ i = 0,\ 1,\ 2$ and $\lambda_i = W_i,\ i = 1,\ 2$, we have

$$E[L(\sigma_0^2, \hat{\sigma_0^2}^{\mathrm{ST}})|W_1,\ W_2] < E[L(\sigma_0^2, V_0/(\nu_0 + 2))|W_1,\ W_2]. \tag{39}$$

Taking the expectation on both sides of Eq. (39) over W_i's, we see that $R(\sigma_0^2, \hat{\sigma_0^2}^{\mathrm{ST}}) < R(\sigma_0^2, V_0/(\nu_0 + 2))$, which completes the proof.

In the following, we discuss whether or not $\hat{\sigma_0^2}^{\mathrm{ST}}$ uniformly improves upon $V_0/(\nu_0 + 2)$ for $p \geq 3$. We remark that the case $\sigma_i^2 = \sigma_0^2,\ i = 1, \ldots, p$ may possibly be the most critical one for $\hat{\sigma_0^2}^{\mathrm{ST}}$ to improve upon $V_0/(\nu_0 + 2)$ since in this case $\hat{\sigma_0^2}^{\mathrm{ST}}$ is stochastically smallest and may shrink $V_0/(\nu_0 + 2)$ too much. Note that the risks of $\hat{\sigma_0^2}^{\mathrm{ST}}$ and $V_0/(\nu_0 + 2)$ when $\sigma_i^2 = \sigma_0^2,\ i = 1, \ldots, p$ are equal to $R_0(\sigma_0^2, \delta_p)$ and $R_0(\sigma_0^2, \delta_0)$. Thus we see from the results of Section 2.2 that $\hat{\sigma_0^2}^{\mathrm{ST}}$ does not uniformly improve upon $V_0/(\nu_0 + 2)$ for $p \geq 12$ when $\nu_0 = 1$, for $p \geq 5$ when $2 \leq \nu_0 \leq 3$, for $p \geq 4$ when $4 \leq \nu_0 \leq 12$ and for $p \geq 3$ when $13 \leq \nu_0 \leq 20$. We finally give the following two Remarks.

Remark 3.1 Our results indicate that the isotonic regression estimator $\hat{\sigma_0^2}^{\mathrm{ST}}$ of the smallest scale parameter under simple tree order restriction fails to improve upon the usual estimator $V_0/(\nu_0 + 2)$ for larger p. Not surprisingly, similar phenomenon is reported by Lee (1988) and Hwang and Peddada (1994) for the problem of estimating the smallest location parameter of p elliptically symmetric distributions under simple tree order restriction. They showed that for sufficiently large p, the isotonic regression estimator of the smallest location parameter tends to $-\infty$ and fails to improve upon the usual estimator.

Remark 3.2 Recently, (Cohen et al. 2000) have pointed out that while the isotonic regression estimator has desirable property for simple order model, it is prone to behavior which is somewhat unintuitive and unappealing to our sensibilities for many order restricted models including the simple tree order model. Actually, as stated in Remark 3.1, the isotonic regression estimator under simple tree order model fails to improve upon $V_0/(\nu_0 + 2)$ for larger p. This behavior may cause us to seek an alternative estimation procedure. Oono and Shinozaki (2006) have generalized the result of Hwang and Peddada (1994) and have given an estimator which not only has desirable property in the sense of Cohen et al. (2000) but also uniformly improves upon $V_0/(\nu_0 + 2)$.

A Appendix

Proof of Lemma 2.2 Let $\mathcal{M}$ be the set of $(S_0, \ldots, S_p)$ such that $S_0/(\nu_0 + 2) \geq S_{i_j}$ for $j = 1, \ldots, l$. Note that $\mathcal{M} = \bigcup_{\{i_1,\ldots,i_h\}} \mathcal{L}^p_{i_1 \cdots i_h}$, where $\bigcup$ is taken over all the sets $\{i_1, \ldots, i_h\}$ such that $\{i_1, \ldots, i_h\} \supseteq \{i_1, \ldots, i_l\}$. Then we need only to show that if $(S_0, \ldots, S_p) \in \mathcal{J}^p_{i_1 \cdots i_l}$ then $(S_0, \ldots, S_p) \in \mathcal{M}$. Equivalently, supposing that $(S_0, \ldots, S_p) \notin \mathcal{M}$, we show that $(S_0, \ldots, S_p) \notin \mathcal{J}^p_{i_1 \cdots i_l}$. From $(S_0, \ldots, S_p) \notin \mathcal{M}$, wee see that $S_0/(\nu_0+2) < S_{i_j}$ for at least one j, $j = 1, \ldots, l$. We first consider the case when $S_0/(\nu_0+2) \geq S_{i_j}$ for some j's, $j = 1, \ldots, l$. Without loss of generality, we assume that $S_0/(\nu_0+2) \geq S_{i_j}$ for $j = 1, \ldots, m$ $(< l)$ and that $S_0/(\nu_0+2) < S_{i_j}$ for $j = m+1, \ldots, l$. Let us denote $\xi_{i_1 \cdots i_l} = (S_0 + \sum_{j=1}^{l} S_{i_j})/(\nu_0 + l + 2)$. Then we can easily confirm that $\xi_{i_1 \cdots i_m} < \xi_{i_1 \cdots i_l}$, which implies $(S_0, \ldots, S_p) \notin \mathcal{J}^p_{i_1 \cdots i_l}$. In the following we consider the case when $S_0/(\nu_0+2) < S_{i_j}$ for all j, $j = 1, \ldots, l$. Then we can easily confirm that $S_0/(\nu_0+2) < \xi_{i_1 \cdots i_l}$, which implies $(S_0, \ldots, S_p) \notin \mathcal{J}^p_{i_1 \cdots i_l}$. This completes the proof. □

Proof of Lemma 2.3 We omit the proof of (i) since it can be discussed similarly with that of (ii). Without loss of generality we assume $(S_0, \ldots, S_p) \in \mathcal{J}^p_{i_1 \cdots i_k}$. We show that (ii) is true. We see from Lemma 2.2 that $(S_0, \ldots, S_p) \in \mathcal{L}^p_{i_1 \cdots i_h}$ for some $\{i_1, \ldots, i_h\} \supseteq \{i_1, \ldots, i_k\}$. Let us denote $\xi_{i_1 \cdots i_k} = (S_0 + \sum_{j=1}^{k} S_{i_j})/(\nu_0 + k + 2)$. Then we have

$$\begin{aligned} h(U_1, \ldots, U_p) &= \frac{\nu_0 + p + 2}{\nu_0 + k + 2} \sum_{j=1}^{k} U_{i_j} \\ &= (\nu_0 + p + 2)\left(\frac{S_0}{\nu_0 + 2} - \xi_{i_1 \cdots i_k}\right) \end{aligned} \tag{40}$$

and

$$\begin{aligned} h_1(U_1, \ldots, U_p) &= \frac{\nu_0 + p + 2}{\nu_0 + h + 2} \sum_{j=1}^{h} U_{i_j} \\ &= (\nu_0 + p + 2)\left(\frac{S_0}{\nu_0 + 2} - \xi_{i_1 \cdots i_h}\right). \end{aligned} \tag{41}$$

Since $(S_0, \ldots, S_p) \in \mathcal{J}^p_{i_1 \cdots i_k}$ implies $\xi_{i_1 \cdots i_k} \leq \xi_{i_1 \cdots i_h}$, we see from Eqs. (40) and (41) that (ii) is true. This completes the proof. □

Proof of Lemma 2.4 We omit the proof of Eq. (32) since it can be discussed similarly with that of Eq. (31). Since we have from Eqs. (28) and (29)

$$h_1(S_1, \ldots, S_p) - h_2(S_1, \ldots, S_p) = \sum_{l=1}^{p} \sum_{\{i_1, \ldots, i_l\}} \left(\sum_{j=1}^{l} S_{i_j}\right) L^p_{i_1 \cdots i_l},$$

we need only to show that

$$\sum_{l=1}^{p} \sum_{\{i_1, \ldots, i_l\}} \left(\sum_{j=1}^{l} S_{i_j}\right) L^p_{i_1 \cdots i_l} = \sum_{i=1}^{p} S_i I_{\frac{S_0}{\nu_0+2} > S_i}. \tag{42}$$

If $(S_0, \ldots, S_p) \in \mathcal{L}^p_{i_1 \cdots i_l}$ for some $\{i_1, \ldots, i_l\}$, then both sides of Eq. (42) are equal to $\sum_{j=1}^{l} S_{i_j}$. If $(S_0, \ldots, S_p) \notin \mathcal{L}^p_{i_1 \cdots i_l}$ for any $\{i_1, \ldots, i_l\}$, then both sides of Eq. (42) are equal to 0. This completes the proof. □

Proof for (33) in Lemma 2.5 Since $S_j,\ j = 1, \ldots, p$ are identically distributed as χ_1^2 when $\lambda_i = 0,\ i = 1, \ldots, p$, we have

$$E_0[S_{i_1} L^p_{i_1 \cdots i_l}] = E_0[S_1 L^p_{1 \ldots l}]. \tag{43}$$

Thus we have

$$\sum_{\{i_1, \ldots, i_l\}} \left\{ \sum_{j=1}^{l} E_0 \left[S_{i_j} L^p_{i_1 \cdots i_l} \right] \right\} = l \binom{p}{l} E_0 \left[S_1 L^p_{1 \ldots l} \right]. \tag{44}$$

We see from Eqs. (30) and (44) that the left-hand side of Eq. (33) is expressed as

$$\sum_{l=1}^{p-2} \frac{p - l}{\nu_0 + l + 2} l \binom{p}{l} E_0 \left[S_1 L^p_{1 \ldots l} \right] + \frac{p(p-1)}{\nu_0 + p + 1} E_0 \left[S_1 L^p_{1 \cdots p-1} \right]. \tag{45}$$

On the other hand, we have from Eq. (43)

$$\begin{aligned} E_0[S_1 L_1^2] &= E_0 \left[S_1 (L_1^3 + L_{12}^3) \right] \\ &= E_0 \left[S_1 (L_1^4 + 2L_{12}^4 + L_{123}^4) \right] \\ &= \cdots \\ &= E_0 \left[S_1 \sum_{l=1}^{p-1} \binom{p-2}{l-1} L^p_{1 \ldots l} \right] = \sum_{l=1}^{p-1} \binom{p-2}{l-1} E_0[S_1 L^p_{1 \ldots l}]. \end{aligned} \tag{46}$$

Similarly with Eq. (46), we have

$$E_0[S_1 L_1^3] = \sum_{l=1}^{p-2} \binom{p-3}{l-1} E_0[S_1 L^p_{1 \ldots l}], \tag{47}$$

where we define $\binom{0}{0} = 1$. We see from Eqs. (46) and (47) that the right-hand side of Eq. (33) is expressed as

$$\begin{aligned} &\frac{p(p-1)}{\nu_0 + p + 1} \sum_{l=1}^{p-2} \left\{ \binom{p-2}{l-1} + \frac{p-2}{\nu_0 + 3} \binom{p-3}{l-1} \right\} E_0[S_1 L^p_{1 \ldots l}] \\ &\quad + \frac{p(p-1)}{\nu_0 + p + 1} E_0[S_1 L^p_{1 \cdots p-1}] \\ &= \sum_{l=1}^{p-2} \left\{ \frac{(p-l)(\nu_0 + p - l + 2)}{(\nu_0 + 3)(\nu_0 + p + 1)} l \binom{p}{l} \right\} E_0[S_1 L^p_{1 \ldots l}] \\ &\quad + \frac{p(p-1)}{\nu_0 + p + 1} E_0[S_1 L^p_{1 \cdots p-1}], \end{aligned} \tag{48}$$

where we have the last equality by

$$\binom{p-2}{l-1} = \frac{l(p-l)}{p(p-1)} \binom{p}{l} \quad \text{and} \quad \binom{p-3}{l-1} = \frac{l(p-l)(p-l-1)}{p(p-1)(p-2)} \binom{p}{l}. \tag{49}$$

Thus from Eqs. (45) and (48), we need only to show that

$$\frac{(\nu_0+l+2)(\nu_0+p-l+2)}{(\nu_0+p+1)(\nu_0+3)} \geq 1, \tag{50}$$

for $l = 1, \ldots, p-2$, which can be easily verified. □

Proof for Eq. (34) in Lemma 2.5 Similarly with Eq. (43), we have

$$E_0[U_{i_1}^2 L_{i_1\cdots i_l}^p] = E_0[U_1^2 L_{1\cdots l}^p] \tag{51}$$

and

$$E_0[U_{i_1} U_{i_2} L_{i_1\cdots i_l}^p] = E_0[U_1 U_2 L_{1\cdots l}^p]. \tag{52}$$

Thus we have

$$\sum_{\{i_1,\ldots,i_l\}} E_0\left[\left(\sum_{j=1}^{l} U_{i_j}\right)^2 L_{i_1\cdots i_l}^p\right] = l\binom{p}{l}\left\{E_0\left[U_1^2 L_{1\cdots l}^p\right] + (l-1)E_0\left[U_1 U_2 L_{1\cdots l}^p\right]\right\}. \tag{53}$$

We see from Eqs. (30) and (53) that the left-hand side of Eq. (34) is expressed as

$$\sum_{l=1}^{p-1} Q(l)(p-l)l\binom{p}{l}\left\{E_0\left[U_1^2 L_{1\cdots l}^p\right] + (l-1)E_0\left[U_1 U_2 L_{1\cdots l}^p\right]\right\}, \tag{54}$$

where

$$Q(l) = \frac{1}{p-l}\left\{\left(\frac{\nu_0+p+2}{\nu_0+l+2}\right)^2 - 1\right\}. \tag{55}$$

On the other hand, similarly with Eqs. (46) and (47) we have

$$E_0[U_1^2 L_1^2] = \sum_{l=1}^{p-1}\binom{p-2}{l-1} E_0[U_1^2 L_{1\cdots l}^p] \tag{56}$$

and

$$E_0[U_1 U_2 L_{12}^3] = \sum_{l=2}^{p-1}\binom{p-3}{l-2} E_0[U_1 U_2 L_{1\cdots l}^p], \tag{57}$$

where we define $\binom{0}{0} = 1$. We see from Eqs. (56) and (57) that the right-hand side of Eq. (34) is expressed as

$$\begin{aligned} p(p-1)Q(p-1)&\left\{\sum_{l=1}^{p-1}\binom{p-2}{l-1}E_0[U_1^2 L_{1\cdots l}^p]\right. \\ &\left. +(p-2)\sum_{l=2}^{p-1}\binom{p-3}{l-2}E_0[U_1 U_2 L_{1\cdots l}^p]\right\} \\ = \sum_{l=1}^{p-1} Q(p-1)&(p-l)l\binom{p}{l}\left\{E_0\left[U_1^2 L_{1\cdots l}^p\right] + (l-1)E_0\left[U_1 U_2 L_{1\cdots l}^p\right]\right\}, \end{aligned} \tag{58}$$

where we have the last equality by Eq. (49) and

$$\binom{p-3}{l-2} = \frac{l(l-1)(p-l)}{p(p-1)(p-2)}\binom{p}{l}. \tag{59}$$

Thus from Eqs. (54) and (58), we need only to show that

$$Q(l) \geq Q(p-1), \tag{60}$$

for $l = 1, 2, \ldots, p-1$. We see that Eq. (60) is true since $Q(l)$ is a decreasing function of l, which can be easily verified. □

B Appendix

Proof of Proposition 2.1 Without loss of generality we set $\sigma^2 = 1$. We first note that the risk of δ_p when $\lambda_i = 0,\ i = 1, \ldots, p$ can be expressed as

$$R_0(\sigma^2, \delta_p) = \mathrm{Var}_0[\delta_p] + (E_0[\delta_p] - 1)^2, \tag{61}$$

where Var_0 is the variance when $\lambda_i = 0,\ i = 1, \ldots, p$. Based on Eq. (61), we give the condition on p such that

$$R_0(\sigma^2, \delta_p) > R_0(\sigma^2, \delta_0) = 2/(\nu_0 + 2), \tag{62}$$

which implies that δ_p does not uniformly improve upon δ_0. To evaluate the variance of δ_p, we note that δ_p can be written as

$$\delta_p = \delta_p^1 + \delta_p^2, \tag{63}$$

where $\delta_p^1 = S_0/(\nu_0 + p + 2)$ and $\delta_p^2 = \min\left\{\frac{pS_0}{(\nu_0+2)(\nu_0+p+2)}, \frac{(p-1)S_0+(\nu_0+p+2)S_{(1)}}{(\nu_0+3)(\nu_0+p+2)}, \ldots, \frac{\sum_{i=1}^{p} S_{(i)}}{\nu_0+p+2}\right\}$. Since δ_p^1 and δ_p^2 are both increasing in S_0, their covariance is non-negative and we see that

$$\begin{aligned} \mathrm{Var}_0[\delta_p] &= \mathrm{Var}_0[\delta_p^1] + \mathrm{Var}_0[\delta_p^2] + 2\mathrm{Cov}_0[\delta_p^1, \delta_p^2] > \mathrm{Var}_0[\delta_p^1] \\ &= \frac{2\nu_0}{(\nu_0 + p + 2)^2}, \end{aligned} \tag{64}$$

where Cov_0 is the covariance when $\lambda_i = 0,\ i = 1, 2, \ldots, p$.

To evaluate the bias of δ_p, we utilize the inequality

$$h(U_1, \ldots, U_p) \geq \sum_{i=1}^{p} U_i^+ + \frac{1}{\nu_0 + p + 1}\sum_{\{i,j\}} U_i^+ I_{\frac{S_0}{\nu_0+2} < S_j}, \tag{65}$$

whose proof is given later in this Appendix. Using Eq. (65) and taking the expectation of Eq. (23), we have

$$E_0[\delta_p] \leq \frac{\nu_0}{\nu_0 + 2} - \frac{p}{\nu_0 + p + 2}a_{\nu_0} - \frac{p(p-1)}{(\nu_0 + p + 1)(\nu_0 + p + 2)}b_{\nu_0}, \tag{66}$$

where $a_{\nu_0} = E_0[U_1^+]$ and $b_{\nu_0} = E_0[U_1 L_1^2]$. Since the right-hand side of Eq. (66) is clearly smaller than 1, we see from Eq. (66) that

$$(E_0[\delta_p] - 1)^2 \geq \left\{ \frac{2}{\nu_0 + 2} + \frac{p}{\nu_0 + p + 2} a_{\nu_0} + \frac{p(p-1)}{(\nu_0 + p + 1)(\nu_0 + p + 2)} b_{\nu_0} \right\}^2 . \tag{67}$$

Thus we see from Eqs. (61), (64) and (67) that if

$$\frac{2\nu_0}{(\nu_0 + p + 2)^2} + \left\{ \frac{2}{\nu_0 + 2} + \frac{p}{\nu_0 + p + 2} a_{\nu_0} + \frac{p(p-1)}{(\nu_0 + p + 1)(\nu_0 + p + 2)} b_{\nu_0} \right\}^2 \geq \frac{2}{\nu_0 + 2} \tag{68}$$

is true, then Eq. (62) is true. We give the condition for p to satisfy Eq. (68). We consider the two cases, $1 \leq \nu_0 \leq 4$ and $\nu_0 \geq 5$ separately.

Case 1 $1 \leq \nu_0 \leq 4$. Using Mathematica, we have numerically evaluated the values of a_{ν_0} and b_{ν_0} in Table 2. Based on Table 2, we can easily confirm that Eq. (68) is true for $p \geq 12$ when $\nu_0 = 1$, for $p \geq 10$ when $\nu_0 = 2$ and for $p \geq 9$ when $3 \leq \nu_0 \leq 4$.

Case 2 $\nu_0 \geq 5$. We should remark that we can figure out a necessary and sufficient condition for p to satisfy Eq. (68) by numerically evaluating the values of a_{ν_0} and b_{ν_0}. However, in this case, we analytically demonstrate that Eq. (68) is true for $p \geq 24$. Since $b_{\nu_0} > 0$, we can easily confirm that Eq. (68) is true if p satisfies

$$\left\{ a_{\nu_0}^2 (\nu_0 + 2)^2 + 4a_{\nu_0}(\nu_0 + 2) - 2\nu_0 \right\} p + 4(\nu_0 + 2) \left\{ a_{\nu_0}(\nu_0 + 2) - \nu_0 \right\} \geq 0. \tag{69}$$

Noting that $S_0 + S_1$ and U_1 are independently distributed, we evaluate a_{ν_0} as

$$\begin{aligned} a_{\nu_0} &= \frac{1}{\nu_0 + 2} E_0 \left[(S_0 + S_1) \{1 - (\nu_0 + 3)U_1\}^+ \right] \\ &= \frac{1}{\nu_0 + 2} E_0[S_0 + S_1] E_0 \left[\{1 - (\nu_0 + 3)U_1\}^+ \right] \\ &= \frac{\nu_0 + 1}{\nu_0 + 2} P_0 \left(U_1 < \frac{1}{\nu_0 + 3} \right) \left\{ 1 - (\nu_0 + 3) E_0 \left[U_1 \mid U_1 < \frac{1}{\nu_0 + 3} \right] \right\}, \end{aligned} \tag{70}$$

Table 2 Numerical evaluation 2

ν_0	a_{ν_0}	b_{ν_0}
1	0.145330	0.047103
2	0.223607	0.072337
3	0.272519	0.087986
4	0.305971	0.098613

where P_0 is the probability when $\lambda_i = 0,\ i = 1, \ldots, p$. Noting that U_1 is distributed as $Beta(1/2, \nu_0/2)$ when $\lambda_i = 0,\ i = 1, \ldots, p$, it can be shown that for $\nu_0 \geq 5$

$$E_0\left[U_1 \mid U_1 < \frac{1}{\nu_0 + 3}\right] \leq \frac{1}{3(\nu_0 + 3)} \quad \text{and} \quad P_0\left(U_1 < \frac{1}{\nu_0 + 3}\right) \geq \frac{11}{20}, \tag{71}$$

which is Lemma A2 in Oono and Shinozaki (2004). We have from Eqs. (70) and (71)

$$a_{\nu_0} \geq \frac{11}{30}\frac{\nu_0 + 1}{\nu_0 + 2}. \tag{72}$$

Since the left-hand side of Eq. (69) is increasing in a_{ν_0}, we see from Eq. (72) that Eq. (69) is true if

$$p \geq \frac{120(19\nu_0^2 + 27\nu_0 - 22)}{121\nu_0^2 - 238\nu_0 + 1441}. \tag{73}$$

Thus we need only to show that the right-hand side of Eq. (73) is smaller than 24, which can be easily verified. This completes the proof.

□

Proof for Eq. (65) in the proof of Proposition 2.1 Without loss of generality we assume $(S_0, \ldots, S_p) \in \mathcal{J}^p_{i_1\cdots i_l}$. Then we see from Lemma 2.2 that $(S_0, \ldots, S_p) \in \mathcal{L}^p_{i_1\cdots i_h}$ for some $\{i_1, \ldots, i_h\} \supseteq \{i_1, \ldots, i_l\}$. Let us denote $\xi_{i_1\cdots i_l} = (S_0 + \sum_{j=1}^{l} S_{i_j})/(\nu_0 + l + 2)$. Then we have the right-hand side of Eq. (65) as

$$\left(1 + \frac{p - h}{\nu_0 + p + 1}\right) \sum_{j=1}^{h} U_{i_j} = \left(1 + \frac{p - h}{\nu_0 + p + 1}\right)(\nu_0 + h + 2)\left(\frac{S_0}{\nu_0 + 2} - \xi_{i_1\cdots i_h}\right) \tag{74}$$

On the other hand we have from Eq. (40)

$$h(U_1, \ldots, U_p) = \left(1 + \frac{p - h}{\nu_0 + h + 2}\right)(\nu_0 + h + 2)\left(\frac{S_0}{\nu_0 + 2} - \xi_{i_1\cdots i_l}\right). \tag{75}$$

Since $(S_0, \ldots, S_p) \in \mathcal{J}^p_{i_1\cdots i_l}$ implies $\xi_{i_1\cdots i_l} \leq \xi_{i_1\cdots i_h}$, we see from Eqs. (74) and (75) that Eq. (65) is true. This completes the proof. □

Acknowledgements The authors are very grateful to two anonymous referees and an associate editor for valuable and insightful comments.

References

Barlow, R.E., Bartholomew, D.J., Bremner, J.M., Brunk, H.D. (1972). R.A. Bradley, D.G. Kendall, J.S. Hunter, G.S. Watson (Eds.), *Statistical inference under order restrictions: The theory and application of isotonic regression.* New York: Wiley

Brewster, J.F., Zidek, J.V. (1974). Improving on equivariant estimators. *The Annals of Statistics*, *2*, 21–38.

Brown, L. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *The Annals of Mathematical Statistics*, *39*, 29–48.

Chang, Y-T., Shinozaki, N. (2002). A comparison of restricted and unrestricted estimators in estimating linear functions of ordered scale parameters of two gamma distributions. *Annals of the Institute of Statistical Mathematics*, *54*, 848–860.

Cohen, A., Kemperman, J.H.B., Sackrowitz, H.B. (2000). Properties of likelihood inference for order restricted models. *Journal of Multivariate Analysis*, *72*, 50–77.

Efron, B., Morris, C. (1976). Families of minimax estimators of the mean of a multivariate normal distributions. *The Annals of Statistics*, *4*, 11–21.

Gelfand, A.E., Dey, D.K. (1988). Improved estimation of the disturbance variance in a linear regression model. *Journal of Econometrics*, *39*, 387–395.

George, E.I. (1990). Developments in decision-theoretic variance estimation:comment. *Statistical Science*, *5*, 107–109.

Hwang, J.T.G., Peddada, S.D. (1994). Confidence interval estimation subject to order restrictions. *The Annals of Statistics*, *22*, 67–93.

Iliopoulos, G., Kourouklis, S. (2000). Interval estimation for the ratio of scale parameters and for ordered scale parameters. *Statistics & Decisions*, *18*, 169–184.

Kaur, A., Singh, H. (1991). On the estimation of ordered means of two exponential populations. *Annals of the Institute of Statistical Mathematics*, *43*, 347–356.

Kourouklis, S. (2001). Estimating the smallest scale parameter: universal domination results. In: Ch. A. Charalambides, M. V. Koutras, N. Balakrishnan, (Eds.), *Probability and Statistical Models with Applications.* London/Boca Raton: Chapman & Hall/CRC

Kubokawa, T., Morita, K., Makita, S., Nagakura, K. (1993). Estimation of the variance and its applications. *Journal of Statistical Planning and Inference*, *35*, 319–333.

Kushary, D., Cohen, A. (1989). Estimating ordered location and scale parameters. *Statistics & Decisions*, *7*, 201–213.

Lee, C.I.C. (1988). Quadratic loss of order restricted estimators for treatment means with a control. *The Annals of Statistics*, *16*, 751–758.

Lehmann, E.L., Casella, G. (1998). G. Casella, S. Fienberg, I. Olkin (Eds.), *Theory of Point Estimation* New York: Springer

Maatta, J.M., Casella, G. (1990). Developments in decision-theoretic variance estimation. *Statistical Science*, *5*, 90–120.

Nagata, Y. (1989). Estimation with pooling procedures of error variance in ANOVA for orthogonal arrays. *Journal of the Japanese Society for Quality Control*, *19*, 12–19 (in Japanese).

Oono, Y. (2005). *On the improved estimation of error variance and order restricted normal means and variances*, PhD. dissertation, Keio University.

Oono, Y., Shinozaki, N. (2004). Estimation of error variance in the analysis of experiments using two-level orthogonal arrays. *Communications in Statistics – Theory and Methods*, *33*, 75–98.

Oono, Y., Shinozaki, N. (2006). On a class of improved estimators of variance and estimation under order restriction. *Journal of Statistical Planning and Inference, 136*, 2584–2605.

Robertson, T., Wright, F.T., Dykstra, R.L. (1988). In: V. Barnett, R.A. Bradley, J.S. Hunter, D.G. Kendall, R.G. Miller, A.F.M. Smith, S.M. Stigler, G.S. Watson, (Eds.), *Order restricted statistical inference*. New York: Wiley

Shinozaki, N. (1995). Some modifications of improved estimators of a normal variance. *Annals of the Institute of Statistical Mathematics*, *47*, 273–286.

Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Annals of the Institute of Statistical Mathematics*, *16*, 155–160.

Vijayasree, G., Singh, H. (1993). Mixed estimators of two ordered exponential means. *Journal of Statistical Planning and Inference*, *35*, 47–53.