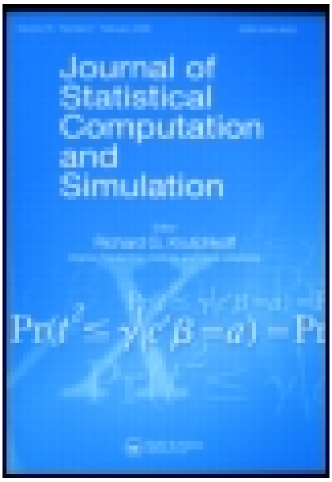

Journal of Statistical Computation and Simulation
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gscs20

Simultaneous confidence intervals for pairwise multiple comparisons in a two-way unbalanced design with unequal variances

Guoyi Zhang[a]
[a] Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131-0001, USA
Published online: 03 Jul 2014.

Journal of Statistical Computation and Simulation, 2014
http://dx.doi.org/10.1080/00949655.2014.935735

Simultaneous confidence intervals for pairwise multiple comparisons in a two-way unbalanced design with unequal variances

Guoyi Zhang*

Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131-0001, USA

(*Received 17 April 2014; accepted 13 June 2014*)

In this research, we propose simultaneous confidence intervals for all pairwise multiple comparisons in a two-way unbalanced design with unequal variances, using a parametric bootstrap approach. Simulation results show that Type 1 error of the multiple comparison test is close to the nominal level even for small samples. They also show that the proposed method outperforms Tukey–Kramer procedure when variances are heteroscedastic and group sizes are unequal.

Keywords: ANOVA; parametric bootstrap; multiple comparison; simulations; unequal variance

1. Introduction

Consider the ANOVA problem of *ab* normal populations with unequal population variances σ_{ij}^2 and unequal group sizes n_{ij}. The two-way ANOVA model is as follows:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}, \tag{1}$$

where $\epsilon_{ijk} \overset{\text{iid}}{\sim} N(0, \sigma_{ij}^2),\ i = 1, 2, \ldots, a,\ j = 1, 2, \ldots, b, k = 1, 2, \ldots, n_{ij}$. α_is, β_js and γ_{ij}s are subject to the constraints (first introduced by Scheffé [1] to solve the model identifiability problem)

$$\sum_{i=1}^{a} w_i\alpha_i = 0, \quad \sum_{j=1}^{b} v_j\beta_j = 0, \quad \sum_{i=1}^{a} w_i\gamma_{ij} = 0, \quad \sum_{j=1}^{b} v_j\gamma_{ij} = 0, \tag{2}$$

where $w_1, \ldots, w_a$ and $v_1, \ldots, v_b$ are nonnegative weights. The multiple comparison procedure (MCP) applies when the family of interest is the set of all pairwise comparisons of factor-level means. Pairwise comparisons of the factor A level means $\mu_{i\cdot}$ can be described as

$$H_0 : \mu_{i\cdot} - \mu_{i'\cdot} = 0 \quad \text{vs.} \quad H_\alpha : \text{at least one of } \mu_{i\cdot} - \mu_{i'\cdot} \neq 0. \tag{3}$$

Pairwise comparisons of the factor B level means $\mu_{\cdot j}$ can be described as

$$H_0 : \mu_{\cdot j} - \mu_{\cdot j'} = 0 \quad \text{v.s.} \quad H_\alpha : \text{at least one of} \quad \mu_{\cdot j} - \mu_{\cdot j'} \neq 0. \tag{4}$$

*Email: gzhang123@gmail.com

Scheffé's [1] method, the Bonferroni inequality-based method, and Tukey [2] method are widely used for MCP among the group means for balanced design with constant variance. Hochberg and Tamhane [3, p.81] showed that the Tukey procedure is optimal in the sense of that it gives the shortest confidence intervals given the joint confidence level at least $1-\alpha$. In practice, unequal treatment sample sizes are often encountered. Several approximation procedures have been developed to deal with such imbalance. The most commonly used approximation procedure is the Tukey–Kramer procedure.[2,4] Hayter [5] proved that the Tukey–Kramer approximation is conservative. Dunnett,[6] Stoline,[7] and Spurrier [8] showed that Tukey–Kramer approximation yields narrower confidence intervals than other approximation procedures unless the treatment sample sizes are severely imbalanced. Cheung and Chan [9] extended Tukey's procedure to two-way unbalanced designs. However, evaluating a $b-1$ (or $a-1$) dimensional integral in their methods brings computational difficulties.

Research on MCP under the assumption of heteroscedasticity is limited. Kaiser and Bowden [10] discussed simultaneous confidence intervals for all linear contrasts in an one-way ANOVA with unequal variances. Recently, Krishnamoorthy et al. [11] proposed a parametric bootstrap (PB) test for equality of factor means in one-way ANOVA. Xu et al. [12] showed that PB test performs better than the generalized F-test for two-way ANOVA and also performs very satisfactorily even for small samples. Inspired by Krishnamoorthy et al.,[11] Zhang [13] proposed PB MCP for one-way ANOVA. To our knowledge, there is no practical MCP available for two-way unbalanced design with unequal variances. Our research intends to fill this gap.

This paper is organized as follows. In Section 2, we briefly review Tukey–Kramer's methods. In Section 3, we propose PB algorithm of multiple comparisons for two-way ANOVA. In Section 4, we present simulation studies and compared our proposed methods to the Tukey–Kramer procedure. Section 5 gives conclusions.

2. Tukey–Kramer MCP

For unbalanced data, the most commonly used procedure is the Tukey–Kramer procedure.[2,4] The Tukey–Kramer procedure utilizes the studentized range denoted by

$$q(r,v)=\frac{w}{s}, \tag{5}$$

where degrees of freedom of the range distribution are $r=a$ or b and $v=N-ab$, with $N=\sum_i\sum_j n_{ij}$; s is the square root of an estimate s^2 of the constant variance σ^2, and w is the range for the set of observations. Tukey–Kramer procedure gives the confidence interval estimates of $\mu_{i\cdot}-\mu_{i'\cdot}$ as

$$\bar{y}_{i\cdot\cdot}-\bar{y}_{i'\cdot\cdot}\pm q_{a,N-ab}(\alpha)\sqrt{\mathrm{MSE}}\sqrt{\frac{1}{2}\left(\frac{1}{n_{i\cdot}}+\frac{1}{n_{i'\cdot}}\right)},$$

where $n_{i\cdot}=\sum_{j=1}^{b}n_{ij}$, $n_{i'\cdot}=\sum_{j=1}^{b}n_{i'j}$, $s_{ij}^2=(1/(n_{ij}-1))\sum_{k=1}^{n_{ij}}(y_{ijk}-\bar{y}_{ij\cdot})^2$.

$$\mathrm{MSE}=\frac{\mathrm{SSE}}{N-ab}=\frac{1}{N-ab}\sum_a\sum_b(n_{ij}-1)s_{ij}^2$$

and $q_{a,N-ab}(\alpha)$ is the upper αth quantile of the studentized range distribution with $a, N-ab$ degrees of freedom. Similarly, Tukey's multiple comparison confidence limits for all pairwise comparisons $\mu_{\cdot j} - \mu_{\cdot j'}$ with family confidence coefficient of at least $1 - \alpha$ are

$$\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot j'\cdot} \pm q_{b,N-ab}(\alpha)\sqrt{\text{MSE}}\sqrt{\frac{1}{2}\left(\frac{1}{n_{\cdot j}} + \frac{1}{n_{\cdot j'}}\right)},$$

where $n_{\cdot j} = \sum_{i=1}^{a} n_{ij}$ and $n_{\cdot j'} = \sum_{i=1}^{a} n_{ij'}$, and $q_{b,N-ab}(\alpha)$ is the upper αth quantile of the studentized range distribution with $b, N - ab$ degrees of freedom.

3. The PB method for multiple comparisons in a two-way unbalanced design with unequal variances

Krishnamoorthy et al. [11] first introduced PB method for testing equality of factor means (overall test) in one-way ANOVA. The PB method has been shown to be the best among nine current existing methods for overall test under the assumption of heteroscedastic variances.[14] Inspired by Krishnamoorthy et al.,[11] Zhang [13] first introduced PB method to MCP in one-way ANOVA, and showed that PB method works very well. In this section, we extend our work on PB method from one-way design to two-way unbalanced design with unequal variances. We will first discuss MCP for factor A level means, then discuss MCP for factor B level means.

3.1. *MCP for factor A level means*

In this section, we propose the MCP for factor A level means $\mu_{i\cdot\cdot}$. First we discuss the point estimator and variance estimator of $\mu_{i\cdot\cdot}$. Next we derive the PB pivotal variables. Last, we propose the test statistic and computational algorithm.

Under heteroscedastic variances and unequal sizes, the estimate of factor-level means would be a weighted average of the corresponding cell means. We propose the estimator of factor A level means $u_{i\cdot}$ as follows:

$$\begin{aligned}
\bar{Y}_{i\cdot\cdot} &= \frac{\sum_j v_j \bar{Y}_{ij}}{\sum_j v_j} \\
&= \frac{\sum_j v_j(\mu + \alpha_i + \beta_j + \gamma_{ij} + \bar{\epsilon}_{ij})}{\sum_j v_j} \\
&= \mu + \alpha_i + \frac{\sum_j v_j \beta_j}{\sum_j v_j} + \frac{\sum_j v_j \gamma_{ij}}{\sum_j v_j} + \frac{\sum_j v_j \bar{\epsilon}_{ij}}{\sum_j v_j} \\
&= \mu + \alpha_i + \frac{\sum_j v_j \bar{\epsilon}_{ij}}{\sum_j v_j},
\end{aligned}$$

where v_js are the nonnegative weights described in Equation (2), $\bar{Y}_{ij} = \sum_{k=1}^{n_{ij}} Y_{ijk}/n_{ij}$ are cell means, and $\bar{\epsilon}_{ij} \sim N(0, \sigma_{ij}^2/n_{ij})$. The following two sets of weights v_i are suggested

$$v_1 = v_2 = \cdots = v_b = \frac{1}{b} \tag{6}$$

and

$$v_1 = \frac{n_{\cdot 1}}{N},\ v_2 = \frac{n_{\cdot 2}}{N},\ \ldots,\ v_b = \frac{n_{\cdot b}}{N}. \tag{7}$$

In general, σ_{ij}^2s are unknown, and are replaced by $S_{ij}^2 = \sum_{k=1}^{n_{ij}} (Y_{ijk} - \bar{Y}_{ij})^2/(n_{ij} - 1)$. Let s_{ij}^2 be the observed value of S_{ij}^2, and $\bar{y}_{ij}$ be the observed value of $\bar{Y}_{ij}$. It is easy to show that variance of

$\bar{Y}_{i\cdot\cdot}$ is

$$v(\bar{Y}_{i\cdot\cdot}) = \frac{1}{(\sum_j v_j)^2}\left(\sum_j \frac{\sigma_{ij}^2}{n_{ij}} v_j^2\right) \tag{8}$$

and estimate of the variance is

$$\hat{v}(\bar{Y}_{i\cdot\cdot}) = \frac{1}{(\sum_j v_j)^2}\left(\sum_j \frac{s_{ij}^2}{n_{ij}} v_j^2\right). \tag{9}$$

The PB pivot variable can be developed as follows. For a given $(\bar{y}_{11}, \bar{y}_{12}, \ldots, \bar{y}_{ab}; s_{11}^2, s_{12}^2, \ldots, s_{ab}^2)$, $\bar{Y}_{Bij} \sim N(0, s_{ij}^2/n_{ij})$, $S_{Bij}^2 \sim s_{ij}^2 \chi_{n_{ij}-1}^2/(n_{ij}-1)$. Hence $\bar{Y}_{Bij} \overset{d}{=} Z_i(s_{ij}/\sqrt{n_{ij}})$ and $S_{B_{ij}}^2 \overset{d}{=} s_{ij}^2 \chi_{n_{ij}-1}^2/(n_{ij}-1)$, where $\overset{d}{=}$ means the same distribution. Let $\bar{Y}_{Bi\cdot\cdot}$ be the PB estimator of $\mu_{i\cdot}$, we have the following variance estimate:

$$v(\bar{Y}_{Bi\cdot\cdot}) = \frac{1}{(\sum_j v_j)^2}\left(\sum_j \frac{s_{ij}^2}{n_{ij}} v_j^2\right). \tag{10}$$

The test statistic $q_{ii'}^A$ can be formed as the following equation:

$$q_{ii'}^A = \frac{|\bar{Y}_{Bi\cdot\cdot} - \bar{Y}_{Bi'\cdot\cdot}|}{\sqrt{v(\bar{Y}_{Bi\cdot\cdot}) + v(\bar{Y}_{Bi'\cdot\cdot})}}, \tag{11}$$

which has the same distribution as

$$q_{ii'}^A \overset{d}{=} \frac{\left|\frac{\sum_j v_j Z_i (s_{ij}/\sqrt{n_{ij}})}{\sum_j v_j} - \frac{\sum_j v_j Z_{i'} (s_{i'j}/\sqrt{n_{i'j}})}{\sum_j v_j}\right|}{\sqrt{(1/(\sum_j v_j)^2)\sum_j (s_{ij}^2 \chi_{n_{ij}-1}^2/n_{ij}(n_{ij}-1))v_j^2 + (1/(\sum_j v_j)^2)\sum (s_{i'j}^2 \chi_{n_{i'j}-1}^2/n_{i'j}(n_{i'j}-1))v_j^2}}, \tag{12}$$

for $i < i', i = 1, \ldots, a-1, i' = i+1, \ldots, a$. For a given $(n_{11}, n_{12} \ldots, n_{ab})$, $(\bar{y}_{11}, \ldots, \bar{y}_{ab})$ and $(s_{11}^2, \ldots, s_{ab}^2)$, let

$$q_{ii'}^{A0} = \frac{|\bar{y}_{i\cdot\cdot} - \bar{y}_{i'\cdot\cdot}|}{\sqrt{v(\bar{Y}_{Bi\cdot\cdot}) + v(\bar{Y}_{Bi'\cdot\cdot})}} \quad \text{for } i = 1, \ldots, a-1,\ i' = i+1, \ldots, a. \tag{13}$$

Given a significance level α, the multiple comparison confidence limits for simultaneous comparisons $\mu_{i\cdot} - \mu_{i'\cdot}$ with family confidence coefficient at least $1-\alpha$ are

$$\bar{y}_{i\cdot\cdot} - \bar{y}_{i'\cdot\cdot} \pm q_\alpha^A \sqrt{v(\bar{Y}_{Bi\cdot\cdot}) + v(\bar{Y}_{Bi'\cdot\cdot})}, \tag{14}$$

where q_α^A can be estimated using the PB method given in Algorithm 1.

ALGORITHM 1 *For a given* $(n_{11}, n_{12} \ldots, n_{ab})$, $(\bar{y}_{11}, \ldots, \bar{y}_{ab})$ *and* $(s_{11}^2, \ldots, s_{ab}^2)$*:*
For $l = 1, \ldots, L$
Generate $Z_i \sim N(0,1)$ *and* $\chi_{n_{ii'}-1}^2, i = 1, \ldots, a$
Compute $q_{ii'}^A$ *using* (12) *for* $i = 1, \ldots, a-1,\ i' = i+1, \ldots, a$
Find $q_l = max(q_{ii'})$
(end loop)
q_α^A *is the* $1-\alpha$ *percentile of the simulated distribution of* q.

3.2. *MCP for factor B level means*

PB MCP for factor B level means can be derived similarly. We propose the following estimators for factor B level means $u_{\cdot j}, j = 1, 2, \ldots, b$,

$$\bar{Y}_{\cdot j\cdot} = \mu + \beta_j + \frac{\sum_i w_i \bar{\epsilon}_{ij}}{\sum_i w_i}, \tag{15}$$

where w_i's are the nonnegative weights described in Equation (2). The following two sets of weights w_i are suggested

$$w_1 = w_2 = \cdots = w_a = \frac{1}{a}, \tag{16}$$

$$w_1 = \frac{n_{1\cdot}}{N},\ w_2 = \frac{n_{2\cdot}}{N},\ \ldots,\ w_a = \frac{n_{a\cdot}}{N}. \tag{17}$$

Let $\bar{Y}_{B\cdot j\cdot}$ be the PB estimator of factor B level mean $\mu_{\cdot j}$. The variance of $\bar{Y}_{B\cdot j\cdot}$ can be found as follows:

$$v(\bar{Y}_{B\cdot j\cdot}) = \frac{1}{(\sum_i w_i)^2}\left(\sum_i \frac{s_{ij}^2}{n_{ij}} w_i^2\right). \tag{18}$$

The test statistic

$$q_{jj'}^B = \frac{|\bar{Y}_{B\cdot j\cdot} - \bar{Y}_{B\cdot j'\cdot}|}{\sqrt{v(\bar{Y}_{B\cdot j\cdot}) + v(\bar{Y}_{B\cdot j'\cdot})}} \tag{19}$$

has the same distribution as

$$q_{jj'}^B \overset{d}{=} \frac{\left|\frac{\sum_i w_i Z_j (s_{ij}/\sqrt{n_{ij}})}{\sum_i w_i} - \frac{\sum_i w_i Z_{j'} (s_{ij'}/\sqrt{n_{ij'}})}{\sum_i w_i}\right|}{\sqrt{(1/(\sum_i w_i)^2) \sum_i (s_{ij}^2 \chi_{n_{ij}-1}^2 / n_{ij}(n_{ij}-1)) w_i^2 + (1/(\sum_i w_i)^2) \sum (s_{ij'}^2 \chi_{n_{ij'}-1}^2 / n_{ij'}(n_{ij'}-1)) v_{j'}^2}} \tag{20}$$

for $j < j',\ j = 1, \ldots, b-1,\ j' = j+1, \ldots, b$. For a given $(n_{11}, n_{12} \ldots, n_{ab})$, $(\bar{y}_{11}, \ldots, \bar{y}_{ab})$ and $(s_{11}^2, \ldots, s_{ab}^2)$, let

$$q_{jj'}^{B0} = \frac{|\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot j'\cdot}|}{\sqrt{v(\bar{Y}_{B\cdot j\cdot}) + v(\bar{Y}_{B\cdot j'\cdot})}} \quad \text{for } j = 1, \ldots, b-1,\ j' = j+1, \ldots, b. \tag{21}$$

Given a significance level α, the multiple comparison confidence limits for simultaneous comparisons $\mu_{\cdot j} - \mu_{\cdot j'}$ with family confidence coefficient at least $1 - \alpha$ are

$$\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot j'\cdot} \pm q_\alpha^B \sqrt{v(\bar{Y}_{B\cdot j\cdot}) + v(\bar{Y}_{B\cdot j'\cdot})}, \tag{22}$$

where q_α^B can be estimated using the PB method given in Algorithm 2.

ALGORITHM 2 *For a given* $(n_{11}, n_{12} \ldots, n_{ab})$, $(\bar{y}_{11}, \ldots, \bar{y}_{ab})$ *and* $(s_{11}^2, \ldots, s_{ab}^2)$*:*
For $l = 1, \ldots, L$
Generate $Z_i \sim N(0, 1)$ *and* $\chi_{n_{ij'}-1}^2, j = 1, \ldots, b$
Compute $q_{jj'}^B$ *using* (20) *for* $j = 1, \ldots, b-1, j' = j+1, \ldots, b$
Find $q_l = max(q_{jj'})$
(end loop)
q_α^B *is the* $1 - \alpha$ *percentile of the simulated distribution of* q.

4. Simulations

In this section, we use simulation to study the proposed MCP of two-way ANOVA under the assumption of heteroscedastic variances and unequal sizes. We compare our method with Tukey–Kramer procedure. The simulation settings follow from Krishnamoorthy et al. [11] and Xu et al.[12]

The tests we consider are location-scale invariant. Consider $Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$, $i = 1, \ldots, a, j = 1, \ldots, b$ and $\epsilon_{ijk} \sim N(0, \sigma_{ij})$. Without loss of generality, we take $\mu_{ij} = 0$ in simulation studies. The sample statistics $\bar{y}_{ij}$ and s_{ij}^2 are generated independently as $\bar{y}_{ij} \sim N(0, \sigma_{ij}^2/n_{ij}), s_{ij}^2 \sim \sigma_{ij}^2 \chi_{n_{ij}-1}^2/(n_{ij} - 1)$.

The simulation study was performed with factors: (1) number of factor levels: $a = 2$ and $b = 3$; (2) population standard deviation $\boldsymbol{\sigma}_i = (\sigma_{11}, \ldots, \sigma_{23}) : \boldsymbol{\sigma}_1^2 = (1, 1, 1, 1, 1, 1), \boldsymbol{\sigma}_2^2 = (0.1, 0.1, 0.1, 0.5, 0.5, 0.5), \boldsymbol{\sigma}_3^2 = (1, 1, 1, 0.5, 0.5, 0.5), \boldsymbol{\sigma}_4^2 = (0.1, 0.2, 0.3, 0.4, 0.5, 1.0)$,

Table 1. Simulation results of multiple comparisons with $\alpha = 0.05$. Numbers in the table are estimated Type 1 errors. Tukey–Kramer methods (TKA and TKB) work well for balanced design with unequal variances, but are very conservative for unbalanced design with unequal variances in general. The proposed PB methods (MCPA1, MCPA2, MCPB1, MCPB2) work well for all the settings.

		$\alpha = 0.05$					
	$\boldsymbol{\sigma}_i^2$	MCPA1	MCPA2	TKA	MCPB1	MCPB2	TKB
$\mathbf{n}_1$	$\boldsymbol{\sigma}_1^2$	0.0375	0.0515	0.0506	0.0495	0.0460	0.0490
	$\boldsymbol{\sigma}_2^2$	0.0465	0.0530	0.0535	0.0515	0.0565	0.0571
	$\boldsymbol{\sigma}_3^2$	0.0400	0.0460	0.0503	0.0380	0.0420	0.0497
	$\boldsymbol{\sigma}_4^2$	0.0470	0.0395	0.05395	0.0450	0.0510	0.0576
	$\boldsymbol{\sigma}_5^2$	0.0515	0.0415	0.0527	0.0515	0.0415	0.0517
	$\boldsymbol{\sigma}_6^2$	0.0520	0.064	0.0674	0.0490	0.0505	0.0836
$\mathbf{n}_2$	$\boldsymbol{\sigma}_1^2$	0.0395	0.0520	0.0500	0.0440	0.0585	0.0490
	$\boldsymbol{\sigma}_2^2$	0.0505	0.0480	0.0503	0.0430	0.0480	0.0516
	$\boldsymbol{\sigma}_3^2$	0.0415	0.0455	0.0527	0.0510	0.0465	0.0523
	$\boldsymbol{\sigma}_4^2$	0.0520	0.0495	0.0512	0.0475	0.0380	0.0574
	$\boldsymbol{\sigma}_5^2$	0.0585	0.0550	0.0494	0.0470	0.0500	0.0495
	$\boldsymbol{\sigma}_6^2$	0.0500	0.0550	0.0614	0.051	0.0460	0.0746
$\mathbf{n}_3$	$\boldsymbol{\sigma}_1^2$	0.0480	0.0520	0.0498	0.0445	0.0390	0.0497
	$\boldsymbol{\sigma}_2^2$	0.0470	0.0395	0.0218	0.0430	0.0500	0.0502
	$\boldsymbol{\sigma}_3^2$	0.0620	0.0545	0.0731	0.0465	0.0415	0.0547
	$\boldsymbol{\sigma}_4^2$	0.0440	0.0405	0.0270	0.0450	0.0435	0.0462
	$\boldsymbol{\sigma}_5^2$	0.0430	0.0435	0.0428	0.0460	0.0375	0.0482
	$\boldsymbol{\sigma}_6^2$	0.0545	0.0540	0.0305	0.0490	0.0540	0.0573
$\mathbf{n}_4$	$\boldsymbol{\sigma}_1^2$	0.0530	0.0485	0.0501	0.0520	0.0495	0.0500
	$\boldsymbol{\sigma}_2^2$	0.0525	0.0485	0.0085	0.0505	0.0570	0.0473
	$\boldsymbol{\sigma}_3^2$	0.0555	0.0465	0.0973	0.0525	0.0500	0.0516
	$\boldsymbol{\sigma}_4^2$	0.0595	0.0545	0.0133	0.0460	0.0490	0.0351
	$\boldsymbol{\sigma}_5^2$	0.0405	0.0555	0.0364	0.0625	0.0495	0.0416
	$\boldsymbol{\sigma}_6^2$	0.0485	0.0515	0.0087	0.0415	0.0480	0.0310

Table 2. Simulation results of multiple comparisons with $\alpha = 0.1$. Numbers in the table are estimated Type 1 errors. Tukey–Kramer methods (TKA and TKB) work well for balanced design with unequal variances, but in general very conservative for unbalanced design with unequal variances. The proposed PB methods (MCPA1, MCPA2, MCPB1, MCPB2) work well for all the settings.

		$\alpha = 0.1$					
	σ_i^2	MCPA1	MCPA2	TKA	MCPB1	MCPB2	TKB
$\mathbf{n}_1$	σ_1^2	0.0850	0.1015	0.1012	0.0940	0.0825	0.0982
	σ_2^2	0.0895	0.0995	0.1076	0.0865	0.0945	0.1110
	σ_3^2	0.0890	0.0885	0.1019	0.0885	0.0800	0.1039
	σ_4^2	0.0965	0.1010	0.1055	0.0880	0.1000	0.1108
	σ_5^2	0.0915	0.0990	0.1018	0.0925	0.0915	0.1032
	σ_6^2	0.1110	0.1110	0.1265	0.099	0.0970	0.1356
$\mathbf{n}_2$	σ_1^2	0.0945	0.0940	0.1011	0.1050	0.1010	0.1006
	σ_2^2	0.0920	0.1095	0.1038	0.0995	0.0985	0.1071
	σ_3^2	0.1020	0.1020	0.1014	0.0955	0.1000	0.1028
	σ_4^2	0.1020	0.0925	0.1020	0.1010	0.0885	0.1008
	σ_5^2	0.0895	0.1030	0.1007	0.1000	0.1040	0.1032
	σ_6^2	0.0990	0.1025	0.1098	0.0850	0.1075	0.1199
$\mathbf{n}_3$	σ_1^2	0.0930	0.1035	0.0985	0.1015	0.0905	0.1017
	σ_2^2	0.0940	0.0860	0.0517	0.0835	0.0815	0.09865
	σ_3^2	0.0945	0.0915	0.1380	0.0985	0.0835	0.1069
	σ_4^2	0.0870	0.1125	0.0630	0.0890	0.0835	0.0911
	σ_5^2	0.0930	0.1035	0.0821	0.092	0.0965	0.0989
	σ_6^2	0.0910	0.1060	0.0652	0.1065	0.0980	0.0980
$\mathbf{n}_4$	σ_1^2	0.0980	0.1045	0.1009	0.1025	0.0965	0.0997
	σ_2^2	0.0890	0.0940	0.0243	0.0980	0.1045	0.0960
	σ_3^2	0.1030	0.0890	0.1605	0.1005	0.1050	0.1060
	σ_4^2	0.1005	0.1075	0.0388	0.1120	0.0835	0.0710
	σ_5^2	0.0895	0.0905	0.0761	0.0940	0.0960	0.0895
	σ_6^2	0.1170	0.1015	0.0231	0.0940	0.0965	0.0535

$\boldsymbol{\sigma}_5^2 = (0.3, 0.9, 0.4, 0.7, 0.5, 1), \boldsymbol{\sigma}_6^2 = (0.01, 0.1, 0.1, 0.1, 0.1, 1)$; (3) Significance level α: .05 and .1; (4) group sizes $\mathbf{n}_i = (n_{11}, \ldots, n_{23}) : \mathbf{n}_1 = (5, 5, 5, 5, 5, 5), \mathbf{n}_2 = (10, 10, 10, 10, 10, 10), \mathbf{n}_3 = (3, 3, 4, 5, 6, 6), \mathbf{n}_4 = (4, 6, 8, 12, 16, 20)$; (5) weight variable w_i and v_j: two sets for w_i and two sets for v_j. For a given sample size and parameter configuration, we generated 2500 observed vectors $(\bar{y}_{11}, \ldots, \bar{y}_{ab}, s_{11}^2, \ldots, s_{ab}^2)$ and used 5000 runs to estimate Type 1 error (p-value). The following is used to derive p-value of simultaneous tests (3): (a) calculate $q_m^0 = \max(q_{ii'}^{A0})$ using Equation (13), use Algorithm 1 to find q_α^A, the $1 - \alpha$ percentile of the simulated distribution of q; (b) repeat step (a) for 2500 times, p-value is the proportion of the 2500 simulations when $q_m^0 > q_\alpha^A$. p-value for simultaneous tests (4) can be derived similarly.

Tables 1 and 2 give the results of multiple comparisons of the proposed methods and Tukey–Kramer procedure for $\alpha = 0.05$ and $\alpha = 0.1$ respectively. In both tables, 'MCPA1' means MCP for factor A levels using weights from Equation (6); 'MCPA2' means MCP for factor B levels with weights from Equation (7); 'MCPB1' means MCP for factor B levels with weights from Equation (16); 'MCPB2' means MCP for factor B levels with weights from Equation (17). 'TKA' and 'TKB' are the Tukey–Kramer MCP for factor A and B levels, respectively. Numbers in table are simulated estimates of Type 1 errors. We consider four different sizes, $\mathbf{n}_1 = (5,5,5,5,5,5)$, $\mathbf{n}_2 = (10,10,10,10,10,10)$, $\mathbf{n}_3 = (3,3,4,5,6,6)$, $\mathbf{n}_4 = (4,6,8,12,16,20)$. $\boldsymbol{\sigma}_i^2$ is a vector of unequal variances, we consider $\boldsymbol{\sigma}_1^2 = (1,1,1,1,1,1)$, $\boldsymbol{\sigma}_2^2 = (0.1,0.1,0.1,0.5,0.5,0.5)$, $\boldsymbol{\sigma}_3^2 = (1,1,1,0.5,0.5,0.5)$, $\boldsymbol{\sigma}_4^2 = (0.1,0.2,0.3,0.4,0.5,1.0)$, $\boldsymbol{\sigma}_5^2 = (0.3,0.9,0.4,0.7,0.5,1)$, and $\boldsymbol{\sigma}_6^2 = (0.01,0.1,0.1,0.1,0.1,1)$.

From Tables 1 and 2, we can see that the estimates of Type 1 errors of the proposed MCP are close to the nominal levels under all the settings. With balanced design $\mathbf{n}_1$ and $\mathbf{n}_2$, all the simulated p-values of 'TKA' and 'TKB' are close to nominal levels even with unequal variances. However, for unbalanced design $\mathbf{n}_3$ and $\mathbf{n}_4$, Tukey–Kramer MCP only provide valid inference for equal variance case $\boldsymbol{\sigma}_1^2$; when variance are unequal ($\boldsymbol{\sigma}_2^2, \boldsymbol{\sigma}_3^2, \boldsymbol{\sigma}_4^2, \boldsymbol{\sigma}_5^2$ and $\boldsymbol{\sigma}_6^2$), Tukey–Kramer tests are very conservative in general. For example, $\alpha = 0.05$, $\mathbf{n}_3 = (3,3,4,5,6,6)$ and $\boldsymbol{\sigma}_2^2 = (0.1,0.1,0.1,0.5,0.5,0.5)$, simulated p-value of 'TKA' is 0.0218 compared with 0.0470 for 'MCPA1'. With $\alpha = 0.05$, $\mathbf{n}_4 = (4,6,8,12,16,20)$ and $\boldsymbol{\sigma}_2^2 = (0.1,0.1,0.1,0.5,0.5,0.5)$, simulated p-value of 'TKA' is only 0.0085 compared with 0.0525 for 'MCPA1'. The advantages of our proposed method are obvious.

5. Conclusions

MCP applies when the family of interest is the set of all pairwise comparisons of factor-level means. In a two-way unbalanced design with unequal variances, facts such as degree of variance and sample size heterogeneity, the shape of the population etc. can all affect the rates of Type I error and power characteristics. In this research, we proposed an MCP for a two-way unbalanced design with unequal variances based on PB approach. The proposed MCP is easy to use and have computational advantage. Simulation studies show that Type 1 errors of MCP are close to the nominal level for all the settings. They also show that the proposed method outperform Tukey–Kramer comparison procedure when variances are heteroscedastic and group sizes are unequal. Future research will consider MCP for multi-way unbalanced design with unequal variances.

Acknowledgements

The author thanks the referee for helpful comments and constructive suggestions to improve the manuscript.

References

[1] Scheffé H. The analysis of variance. New York: Wiley; 1959.
[2] Tukey JW, 1953. The problem of multiple comparisons. Unpublished manuscript, Princeton University.
[3] Hochberg Y, Tamhane AC. Multiple comparison procedures. New York: Wiley; 1987.
[4] Kramer CY. Extension of multiple range test to group means with unequal numbers of replications. Biometrics 1956;12:307–310.
[5] Hayter AJ. Pairwise comparisons of generally correlated means. J Am Stat Assoc. 1989;84:208–213.
[6] Dunnett CW. Pairwise multiple comparisons in the homogeneous variances, unequal sample size case. J Am Stat Assoc. 1980;75:789–795.
[7] Stoline MR. The status of multiple comparisons: simultaneous estimation of all pairwise comparisons in one-way ANOVA designs. Am Statist. 1981;35:134–141.
[8] Spurrier JD. An improved gt2 method for simultaneous confidence intervals on pairwise differences. Technometrics 1981;23:189–192.
[9] Cheung SH, Chan WS. Simultaneous confidence intervals for pairwise multiple comparisons in a two-way unbalanced design. Biometrics 1996;52:463–472.
[10] Kaiser L, Bowden D. Simultaneous confidence intervals for all linear contrasts of means with heterogenous variances. Commun Stat Theory Methods 1983;12:73–88.

[11] Krishnamoorthy K, Lu F, Mathew T. A parametric bootstrap approach for ANOVA with unequal variances: fixed and random models. Comput Stat Data Anal. 2007;51:5731–5742.
[12] Xu L, Yang F, Abula A, Qin S. A parametric bootstrap approach for two-way ANOVA in presence of possible interactions with unequal variances. J Multivariate Anal. 2013;115:172–180.
[13] Zhang G. A parametric bootstrap approach for one-way ANOVA under unequal variances with unbalanced data. Commun Stat Simul Comput. Forthcoming 2014. Available from http://www.tandfonline.com/doi/abs/10.1080/03610918.2013.794288?journalCode=lssp20#preview.
[14] Yiğit E, Gokpinar F. A simulation study on tests for one-way ANOVA under the unequal variance assumption. Commun Fac Sci Univ Ankara, Ser A1, Math Stat. 2010;59(2):15–34.