On the conservative multivariate Tukey-Kramer type procedures for multiple comparisons among mean vectors

Takashi Seo[a] *, Takahiro Nishiyama[b]

[a] Department of Mathematical Information Science , Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo, 162-8601 Japan

[b] Graduate School of Science, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo, 162-8601 Japan

Abstract

In this paper, conservative simultaneous confidence intervals for multiple comparisons among mean vectors in multivariate normal distributions are considered. Some properties of the multivariate Tukey-Kramer type procedures for pairwise comparisons and for comparisons with a control are presented. The upper bounds for the conservativeness of the procedures are discussed. Finally, numerical results by Monte Carlo simulations are given.

Key Words: Comparisons with a control; Conservativeness; Coverage probability; Multivariate Tukey-Kramer Procedure; Monte Carlo simulation; Pairwise comparisons

1. Introduction

Consider the simultaneous confidence intervals for multiple comparisons among mean vectors from the multivariate normal populations. Let $\boldsymbol{\mu}_i$ be the mean vector from ith population. Let $\boldsymbol{M} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k]$ be the unknown $k \times p$ matrix and $\widehat{\boldsymbol{M}} = [\widehat{\boldsymbol{\mu}}_1, \ldots, \widehat{\boldsymbol{\mu}}_k]$ be the estimator of $\boldsymbol{M}$ such that $\mathrm{vec}(\boldsymbol{X})$ is distributed as $N_{kp}(\mathbf{0}, \boldsymbol{V} \otimes \boldsymbol{\Sigma})$,

*Corresponding author. Tel.: +81-3-5228-8199; fax: +81-3-3260-4293.
E-mail address: seo@rs.kagu.tus.ac.jp.

where $\boldsymbol{X} = \widehat{\boldsymbol{M}} - \boldsymbol{M}$, $\boldsymbol{V}$ is a known $k \times k$ positive definite matrix and $\boldsymbol{\Sigma}$ is an unknown $p \times p$ positive definite matrix, and vec$(\cdot)$ denotes the column vector formed by stacking the columns of the matrix under each other. Also, let $\boldsymbol{S}$ be an unbiased estimator of $\boldsymbol{\Sigma}$ such that $\nu\boldsymbol{S}$ is independent of $\widehat{\boldsymbol{M}}$ and is distributed as a Wishart distribution $\mathrm{W}_p(\boldsymbol{\Sigma}, \nu)$. Then the simultaneous confidence intervals for multiple comparisons among mean vectors are given by

$$\boldsymbol{a}'\boldsymbol{M}\boldsymbol{b} \in \left[\ \boldsymbol{a}'\widehat{\boldsymbol{M}}\boldsymbol{b} \pm t(\boldsymbol{b}'\boldsymbol{V}\boldsymbol{b})^{1/2}(\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a})^{1/2}\right], \quad \forall \boldsymbol{a} \in \mathbb{R}^p,\ \forall \boldsymbol{b} \in \mathbb{B}, \tag{1}$$

where $\mathbb{R}^p$ is the set of any nonzero real p-dimensional vectors and $\mathbb{B}$ is a subset that consists of r vectors in the k-dimensional space. We note that the value $t(> 0)$ is the upper α percentile of the $T^2_{\max}$-type statistic,

$$T^2_{\max} = \max_{\boldsymbol{b} \in \mathbb{B}} \left\{ \frac{(\boldsymbol{X}\boldsymbol{b})'\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{b}}{\boldsymbol{b}'\boldsymbol{V}\boldsymbol{b}} \right\}, \tag{2}$$

which the coverage probability for (1) is $1 - \alpha$.

In many experimental situations, pairwise comparisons and comparisons with a control are standard for multiple comparisons. In the case of pairwise comparisons, we note that

$$\mathbb{B} = \mathbb{C} \equiv \{\boldsymbol{c} \in \mathbb{R}^k : \boldsymbol{c} = \boldsymbol{e}_i - \boldsymbol{e}_j,\ 1 \le i < j \le k\},$$

where $\boldsymbol{e}_i$ is the ith unit vector of the k-dimensional space. We can also express (1) as

$$\boldsymbol{a}'(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \in \left[\ \boldsymbol{a}'(\widehat{\boldsymbol{\mu}}_i - \widehat{\boldsymbol{\mu}}_j) \pm t_{\max\cdot p}\ (d_{ij}\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a})^{1/2}\right], \forall \boldsymbol{a} \in \mathbb{R}^p,\ 1 \le i < j \le k,$$

where $t^2_{\max\cdot p}$ is the upper α percentile of $T^2_{\max\cdot p}$ statistic,

$$T^2_{\max\cdot p} = \max_{i<j}\{(\boldsymbol{x}_i - \boldsymbol{x}_j)'(d_{ij}\boldsymbol{S})^{-1}(\boldsymbol{x}_i - \boldsymbol{x}_j)\},$$

and $d_{ij} = v_{ii} - 2v_{ij} + v_{jj}$.

In the case of pairwise comparisons with $\boldsymbol{V} = \boldsymbol{I}$, the $T^2_{\max}$ type statistic is reduced as the same as half of the multivariate studentized range statistic $R^2_{\max}$ (see, e.g., Seo and Siotani(1992)). Seo, Mano and Fujikoshi(1994) proposed the multivariate Tukey-Kramer procedure which is a simple procedure by replacing with the upper percentile of the multivariate Studentized range statistic as an approximation to the one of $T^2_{\max}$ type statistic for any positive definite matrix $\boldsymbol{V}$. The multivariate generalized Tukey conjecture is known as the statement that the multivariate Tukey-Kramer procedure yields the conservative simultaneous confidence intervals for all pairwise comparisons among mean vectors. This conjecture has been affirmatively proved in the case of three correlated mean vectors by Seo, Mano and Fujikoshi(1994). Relating to this conjecture, Seo(1996) discussed how the approximate simultaneous confidence level by the multivariate Tukey-Kramer procedure is close to $1-\alpha$. The related discussion for the univariate case is referred to Somerville(1993).

Some approximations for the upper percentiles of $T^2_{\max}$ statistic are discussed, for example, by Siotani(1959a, 1959b, 1960), Seo and Siotani(1992), Seo(1995), Krishnaiah(1979) and Siotani, Hayakawa and Fujikoshi(1985).

In the case of comparisons with a control, we have

$$\mathbb{B} = \mathbb{D} \equiv \{\boldsymbol{d} \in \mathbb{R}^k : \boldsymbol{d} = \boldsymbol{e}_i - \boldsymbol{e}_k,\ i = 1, \ldots, k-1\},$$

where k-th population is the control. Then we can write (1) as

$$\boldsymbol{a}'(\boldsymbol{\mu}_i - \boldsymbol{\mu}_k) \in [\ \boldsymbol{a}'(\widehat{\boldsymbol{\mu}}_i - \widehat{\boldsymbol{\mu}}_k) \pm t_{\max\cdot c}\ (d_{ik}\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a})^{1/2}], \forall \boldsymbol{a} \in \mathbb{R}^p,\ i = 1, \ldots, k-1.$$

where $t^2_{\max\cdot c}$ is the upper α percentile of $T^2_{\max\cdot c}$ statistic,

$$T^2_{\max\cdot c} = \max_{i=1,\ldots,k-1}\{(\boldsymbol{x}_i - \boldsymbol{x}_k)'(d_{ik}\boldsymbol{S})^{-1}(\boldsymbol{x}_i - \boldsymbol{x}_k)\}.$$

Seo(1995) proposed the conservative procedure for the case of comparisons with a control, which is similar to the multivariate Tukey-Kramer procedure. In this paper, we discuss the bound for these conservative procedures for pairwise comparisons and comparisons with a control.

The organization of the paper is as follows. In Section 2, the conservativeness of the multivariate Tukey-Kramer procedure is discussed. In Section 3, a conservative procedure for comparisons with a control by Seo(1995) is discussed and its upper bound for the conservativeness is presented. Finally, we also give some numerical results by Monte Carlo simulations.

2. The multivariate Tukey-Kramer procedure

The simultaneous confidence intervals for all pairwise comparisons by the multivariate Tukey-Kramer procedure are given by

$$\boldsymbol{a}'(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \in \left[\ \boldsymbol{a}'(\widehat{\boldsymbol{\mu}}_i - \widehat{\boldsymbol{\mu}}_j) \pm t_p\sqrt{d_{ij}\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a}}\right], \forall \boldsymbol{a} \in \mathbb{R}^p,\ 1 \le i < j \le k, \qquad (3)$$

where ${t_p}^2$ is the upper α percentile of $T^2_{\max\cdot p}$ statistic with $\boldsymbol{V} = \boldsymbol{I}$, that is, $t_p^2 = q^2/2$ and $q^2 \equiv q^2_{p,k,\nu}(\alpha)$ is the upper α percentile of the p-variate Studentized range statistic with parameters k and ν. By a reduction of relating to the coverage probability of (3), Seo, Mano and Fujikoshi(1994) proved that the coverage probability in the case $k = 3$ is equal or greater than $1 - \alpha$ for any positive definite matrix $\boldsymbol{V}$. Using the same reduction, Seo(1996) discussed the bound of conservative simultaneous confidence levels.

Consider the probability

$$Q(t, \boldsymbol{V}, \mathbb{B}) = \Pr\{\ (\boldsymbol{X}\boldsymbol{b})'(\nu\boldsymbol{S})^{-1}(\boldsymbol{X}\boldsymbol{b}) \le t(\boldsymbol{b}'\boldsymbol{V}\boldsymbol{b}),\ \forall \boldsymbol{b} \in \mathbb{B}\ \}, \qquad (4)$$

where t is any fixed constant. Without loss of generality, we may assume $\boldsymbol{\Sigma} = \boldsymbol{I}_p$ when we consider the probability (4).

When $t = t_p^*(\equiv t_p^2/\nu)$ and $\mathbb{B} = \mathbb{C}$, the coverage probability (4) is the same as the coverage probability of (3). The conservativeness of the simultaneous confidence intervals (3) means that $Q(t_p^*, \boldsymbol{V}, \mathbb{C}) \ge Q(t_p^*, \boldsymbol{I}, \mathbb{C}) = 1 - \alpha$. The inequality is known as the multivariate generalized Tukey conjecture. Then we have the following theorem for the case $k = 3$ by using same line of the proof of Theorem 3.2 in Seo, Mano and Fujikoshi(1994).

Theorem 1. *Let $Q(t, \boldsymbol{V}, \mathbb{B})$ be the coverage probability (4) with a known matrix $\boldsymbol{V}$ for the case $k = 3$. Then, for any positive definite matrix $\boldsymbol{V}$, it holds that*

$$1 - \alpha = Q(t_p^*, \boldsymbol{I}, \mathbb{C}) \leq Q(t_p^*, \boldsymbol{V}, \mathbb{C}) < Q(t_p^*, \boldsymbol{V}_0, \mathbb{C}),$$

where $t_p^* = t_p^2/\nu$, $\mathbb{C} = \{\boldsymbol{c} \in \mathbb{R}^k : \boldsymbol{c} = \boldsymbol{e}_i - \boldsymbol{e}_j,\ 1 \leq i < j \leq k\}$ and $\boldsymbol{V}_0$ satisfies with one of $\sqrt{d_{ij}} = \sqrt{d_{il}} + \sqrt{d_{jl}}$, $i \neq j \neq l$.

This is an extended result of Seo(1996). For the case $k = 3$, we note that there is no exist a positive definite matrix such that $\sqrt{d_{12}} = \sqrt{d_{13}} + \sqrt{d_{23}}$ or $\sqrt{d_{13}} = \sqrt{d_{12}} + \sqrt{d_{23}}$ or $\sqrt{d_{23}} = \sqrt{d_{12}} + \sqrt{d_{13}}$. That is, V_0 is a positive semi-definite matrix.

For a special case that $\boldsymbol{V}$ is a diagonal matrix, we consider the statistic $T^2_{\max \cdot p}$ with the case $\boldsymbol{V} = \mathrm{diag}(n_1^{-1}, \ldots, n_k^{-1})$. Then we have

$$T^2_{\max \cdot p} = \max_{i<j} \left[\frac{(\boldsymbol{x}_i - \boldsymbol{x}_j)' \boldsymbol{S}^{-1} (\boldsymbol{x}_i - \boldsymbol{x}_j)}{n_i^{-1} + n_j^{-1}} \right].$$

Without loss of generality, we can assume $n_i \leq n_j$ and put $a_{ij}^2 = n_i/n_j$. Then we have

$$T^2_{\max \cdot p} = \max_{i<j} \left[\frac{(\boldsymbol{u}_i - a_{ij}\boldsymbol{u}_j)' \boldsymbol{S}^{-1} (\boldsymbol{u}_i - a_{ij}\boldsymbol{u}_j)}{1 + a_{ij}^2} \right],$$

where $\boldsymbol{u}_i = \sqrt{n_i}(\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i) \sim N_p(\boldsymbol{0}, \boldsymbol{I})$ and $\nu \boldsymbol{S} \sim W_p(\boldsymbol{I}, \nu)$. Then, when $a_{ij} \to 0$, we have

$$T^2_{\max \cdot p} = \widetilde{T}^2_{\max} = \max_{i=1,\ldots,k-1} \{\boldsymbol{u}_i' \boldsymbol{S}^{-1} \boldsymbol{u}_i\}.$$

Also, when $k = 3$, the distribution of $\widetilde{T}^2_{\max}$ statistic is the same as that of $T^2_{\max \cdot p}$ statistic with $\sqrt{d_{12}} = \sqrt{d_{13}} + \sqrt{d_{23}}$ or $\sqrt{d_{13}} = \sqrt{d_{12}} + \sqrt{d_{23}}$ or $\sqrt{d_{23}} = \sqrt{d_{12}} + \sqrt{d_{13}}$, that is, one of $\sqrt{d_{ij}} = \sqrt{d_{il}} + \sqrt{d_{jl}}$, $i \neq j \neq l$. From Theorem 1, it follows that $Q(t_p^*, \boldsymbol{V}, \mathbb{C}) < \Pr\{\widetilde{T}^2_{\max} < t_p^{*2}\}$ for any diagonal matrix $\boldsymbol{V}$. Therefore, we have the following corollary.

Corollary 2. *Let $Q(t, \boldsymbol{V}, \mathbb{B})$ be the coverage probability (4) with a diagonal matrix $\boldsymbol{V}$ for the case $k = 3$. Then, for any diagonal matrix $\boldsymbol{V}$, it holds that*

$$1 - \alpha = Q(t_p^*, \boldsymbol{I}, \mathbb{C}) \leq Q(t_p^*, \boldsymbol{V}, \mathbb{C}) < \Pr\{\widetilde{T}_{\max}^2 < t_p^2\},$$

where $t_p^* = t_p^2/\nu$, $\mathbb{C} = \{\boldsymbol{c} \in \mathbb{R}^k : \boldsymbol{c} = \boldsymbol{e}_i - \boldsymbol{e}_j,\ 1 \leq i < j \leq k\}$,

$$\widetilde{T}_{\max}^2 = \max_{i=1,\dots,k-1} \{\boldsymbol{u}_i' \boldsymbol{S}^{-1} \boldsymbol{u}_i\},$$

and $\boldsymbol{u}_i$, $i = 1, \dots, k-1$ are independent identically distributed as $N_p(\boldsymbol{0}, \boldsymbol{I})$ and $\nu \boldsymbol{S}$ is distributed as $W_p(\boldsymbol{I}, \nu)$.

3. A conservative procedure for comparisons with a control

In this section, a conservative procedure for comparisons with a control by Seo(1995) is discussed. The simultaneous confidence intervals for all comparisons with a control are given by

$$\boldsymbol{a}'(\boldsymbol{\mu}_i - \boldsymbol{\mu}_k) \in \left[\ \boldsymbol{a}'(\widehat{\boldsymbol{\mu}}_i - \widehat{\boldsymbol{\mu}}_k) \pm t_{\max\cdot c}\sqrt{d_{ik}\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a}}\right], \forall \boldsymbol{a} \in \mathbb{R}^p,\ i = 1, \dots, k-1, \quad (5)$$

where $t_{\max\cdot c}^2$ is the upper α percentile of $T_{\max\cdot c}^2$ statistic. Seo(1995) conjectured conservative simultaneous confidence intervals given by

$$\boldsymbol{a}'(\boldsymbol{\mu}_i - \boldsymbol{\mu}_k) \in \left[\ \boldsymbol{a}'(\widehat{\boldsymbol{\mu}}_i - \widehat{\boldsymbol{\mu}}_k) \pm t_c\sqrt{d_{ik}\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a}}\right], \forall \boldsymbol{a} \in \mathbb{R}^p,\ i = 1, \dots, k-1$$

with $t_c = t_c(\alpha; p, k, \nu, \boldsymbol{V}_1)$ and $\boldsymbol{V}_1$ satisfies with $d_{ij} = d_{ik} + d_{jk}$, $1 \leq i < j \leq k-1$. This conjecture has been affirmatively proved in the case of $k = 3$ by Seo(1995).

The coverage probability (4) with $t = t_c^* = t_c^2/\nu$ and $\mathbb{B} = \mathbb{D}$ is the same as the coverage probability of (5). Then we have the following theorem for the case $k = 3$.

Theorem 3. *Let $Q(t, \boldsymbol{V}, \mathbb{B})$ be the coverage probability (4) with a known matrix $\boldsymbol{V}$ for the case $k = 3$. Then, for any positive definite matrix $\boldsymbol{V}$, it holds that*

$$1 - \alpha = Q(t_c^*, \boldsymbol{V}_1, \mathbb{D}) \leq Q(t_c^*, \boldsymbol{V}, \mathbb{D}) < Q(t_c^*, \boldsymbol{V}_2, \mathbb{D}),$$

where $t_c^* = t_c^2(\alpha; p, k, \nu, \boldsymbol{V}_1)/\nu$, $\mathbb{D} = \{\boldsymbol{d} \in \mathbb{R}^k : \boldsymbol{d} = \boldsymbol{e}_i - \boldsymbol{e}_k,\ i = 1, \ldots, k-1\}$ *and* $\boldsymbol{V}_1$ *satisfies with* $d_{12} = d_{13} + d_{23}$ *and* $\boldsymbol{V}_2$ *satisfies with* $\sqrt{d_{12}} = |\sqrt{d_{13}} - \sqrt{d_{23}}|$.

In the case of multiple comparisons with a control and $k = 3$, there is no exist a positive definite matrix such that $\sqrt{d_{12}} = |\sqrt{d_{13}} - \sqrt{d_{23}}|$. That is, V_2 is a positive semi-definite matrix.

From Theorem 3, it is noted that the multivariate Tukey-Kramer type procedure for multiple comparisons with a control yields the conservativeness for any positive definite matrix $\boldsymbol{V}$ when the case $k = 3$. For the case $k \geq 4$, we can conjecture that the following simultaneous confidence intervals are conservative.

$$\boldsymbol{a}'\boldsymbol{M}\boldsymbol{d} \in \left[\boldsymbol{a}'\widehat{\boldsymbol{M}}\boldsymbol{d} \pm t_c\,(\boldsymbol{d}'\boldsymbol{V}\boldsymbol{d})^{\frac{1}{2}}(\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a})^{\frac{1}{2}}\ \right], \quad \forall \boldsymbol{a} \in \mathbb{R}^p, \forall \boldsymbol{d} \in \mathbb{D},$$

where $t_c = t_c(\alpha; p, k, \nu, \boldsymbol{V}_1)$ and $\boldsymbol{V}_1$ satisfies with the conditions $d_{ij} = d_{ik} + d_{jk}$, $1 \leq i < j \leq k-1$. That is, it may be expected that the procedure give the conservative and good approximate simultaneous confidence intervals. Further, we have the following corollary.

Corollary 4. *Let* $Q(t, \boldsymbol{V}, \mathbb{B})$ *be the coverage probability (4) with a diagonal matrix* $\boldsymbol{V}$ *for the case* $k = 3$. *Then, for any diagonal matrix* $\boldsymbol{V}$, *it holds that*

$$1 - \alpha = \Pr\{\widetilde{T}^2_{\max} < \widetilde{t}^2_c\} \leq Q(\widetilde{t}^*_c, \boldsymbol{V}, \mathbb{D}) < \Pr\{T^2_k < \widetilde{t}^2_c\},$$

where $\mathbb{D} = \{\boldsymbol{d} \in \mathbb{R}^k : \boldsymbol{d} = \boldsymbol{e}_i - \boldsymbol{e}_k,\ i = 1, \ldots, k-1\}$, $\widetilde{t}^{*2}_c = \widetilde{t}^2_c/\nu$, $\widetilde{t}^2_c$ *is the upper* α *percentile of* $\widetilde{T}^2_{\max}$ *statistic defined in Corollary 2, and* T^2_k *statistic is* $np/(n-p+1)F_{p,n-p+1}$ *statistic with* p *and* $n-p+1$ *degrees of freedoms.*

4. Numerical examinations

This section gives some numerical results of the coverage probability for $T^2_{\max\cdot p}$ and $T^2_{\max\cdot c}$ statistics and the upper percentiles of their statistics by Monte Carlo simulation. The Monte Carlo simulations are made from 10 replications of 1000000 simulations for each of parameters based on normal random vectors based on $N_{kp}(\mathbf{0},\ \boldsymbol{V}\otimes \boldsymbol{I}_{\mathbf{p}})$. The sample covariance matrix $\boldsymbol{S}$ is computed on the basis of random vectors from $N_p(\mathbf{0},\ \boldsymbol{I}_{\mathbf{p}})$. Also, we note that $\boldsymbol{S}$ is formed independently in each time with ν degrees of freedom. The average of 10 replications based on 1000000 simulations is used as the simulated value of the statistic.

Table 1 gives the upper percentiles $t_p(\boldsymbol{V}) \equiv t_p(\alpha; p, k, \nu, \boldsymbol{V})$ of $T_{\max \cdot p}$ $(= \sqrt{T^2_{\max \cdot p}})$ and the upper bounds of the coverage probability for the following parameters: $\alpha = 0.1, 0.5, 0.01$, $p = 1, 2, 5$, $k = 3$, $\nu = 20, 40, 60$, and $\boldsymbol{V} = \boldsymbol{I}, \boldsymbol{V}_0$, that is,

$$\boldsymbol{V}_0 = \begin{bmatrix} 4 & 0 & 2 \\ 0 & 4 & 2 \\ 2 & 2 & 2 \end{bmatrix}.$$

Here we note that $\boldsymbol{V}_0$ is a positive semi-definite matrix such that $\sqrt{d_{12}} = \sqrt{d_{13}} + \sqrt{d_{23}}$.

A distribution of vec$(\boldsymbol{X})$ is said to be singular or degenerate when the rank of $\boldsymbol{V}$ is less than k, say $r(< k)$. In this case, the total probability mass concentrates on a linear set of exactly rp dimensions with probability one(see, Cramér (1946)).

To have a natural definition of a singular multivariate normal distribution, consider the following factorization of $\boldsymbol{V}$,

$$\boldsymbol{V} = \boldsymbol{H} \begin{bmatrix} \boldsymbol{D}_\lambda & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{H}', \quad \boldsymbol{H} = (\boldsymbol{H}_1, \boldsymbol{H}_2),$$

where $\boldsymbol{H} : k \times k$, $\boldsymbol{H}_1 : k \times r$ being column orthogonal, i.e. $\boldsymbol{H}_1'\boldsymbol{H}_1 = \boldsymbol{I}_r$, and $\boldsymbol{D}_\lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_r)$ with $\lambda_i > 0$ for all $i = 1, \ldots, r$. Putting $\boldsymbol{B} = \boldsymbol{H}_1\boldsymbol{D}_{\sqrt{\lambda}}$ and $\boldsymbol{D}_{\sqrt{\lambda}} = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_r})$, we have $\boldsymbol{V} = \boldsymbol{B}\boldsymbol{B}'$. Note that $\boldsymbol{B}$ is not unique because $\boldsymbol{B}$ can be replaced by $\boldsymbol{B}\boldsymbol{L}$, where $\boldsymbol{L}$ is any $r \times r$ orthogonal matrix. Using one of such matrices, $\boldsymbol{B}$, and $\mathbf{z} \sim N_r(\mathbf{0}, \boldsymbol{I}_{\mathbf{r}})$, the characteristic function of $\mathbf{y} = \boldsymbol{B}\mathbf{z} + \boldsymbol{\mu}$ can be written as

$$\begin{aligned} \mathrm{E}\left[\exp\{i\boldsymbol{t}'\mathbf{y}\}\right] &= \exp\{i\boldsymbol{t}'\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}'(\boldsymbol{B}\boldsymbol{B}')\boldsymbol{t}\} \\ &= \exp\{i\boldsymbol{t}'\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}'\boldsymbol{V}\boldsymbol{t}\} \end{aligned}$$

which is the same form as characteristic function of p variates singular normal distribution with covariance matrix $\boldsymbol{V}$ whose rank is $r < k$. For the case of $\boldsymbol{V}_0$ in Table 1, we have $r = 2$, $\boldsymbol{D}_\lambda$ and $\boldsymbol{B}$ are given by

$$\boldsymbol{D}_\lambda = \begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix}, \ \boldsymbol{B} = \begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix},$$

respectively. Therefore, in this case, the three dimensional random numbers are produced from the ones generated by bivariate normal distribution.

It can be seen from some simulation results in Table 1 that the upper percentiles with $\boldsymbol{V} = \boldsymbol{I}$ are always largest values and those with $\boldsymbol{V} = \boldsymbol{V}_0$ are the lower limit values in any positive definite matrix $\boldsymbol{V}$ for each parameter.

It is noted from Table 1 that the upper bounds for the conservativeness of pairwise multiple comparisons can be obtained. For example, when $p = 2$, $\alpha = 0.1$ and $\nu = 20$, we note that $0.9 \leq Q(t_p^*, \boldsymbol{V}, \mathbb{C}) < 0.958$ for any positive definite $\boldsymbol{V}$.

Table 2 gives the upper percentiles $t_c(\boldsymbol{V}) \equiv t_c(\alpha; p, k, \nu, \boldsymbol{V})$ of $T_{\max\cdot c}$ $(= \sqrt{T^2_{\max\cdot c}})$ and the upper bounds of the coverage probability for the following parameters: $\alpha = 0.1, 0.5, 0.01$, $p = 1, 2, 5$, $k = 3$, $\nu = 20, 40, 60$, and $\boldsymbol{V} = \boldsymbol{I}, \boldsymbol{V_{11}}, \boldsymbol{V_{12}}, \boldsymbol{V_{21}}, \boldsymbol{V_{22}}$, that is,

$$\begin{aligned} \boldsymbol{V}_{11} &= \begin{bmatrix} 1 & 0 & 0.5 \\ 0 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}, \ \boldsymbol{V}_{21} = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 4 \end{bmatrix}, \\ \boldsymbol{V}_{12} &= \begin{bmatrix} 2 & 0.5 & 1 \\ 0.5 & 2 & 1.5 \\ 1 & 1.5 & 2 \end{bmatrix}, \ \boldsymbol{V}_{22} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 4 \end{bmatrix}. \end{aligned}$$

Here we note that $\boldsymbol{V}_{11}, \boldsymbol{V}_{12}$ are positive definite matrices such that $d_{12} = d_{13} + d_{23}$ and $\boldsymbol{V}_{21}, \boldsymbol{V}_{22}$ are positive semi-definite matrices such that $\sqrt{d_{12}} = |\sqrt{d_{13}} - \sqrt{d_{23}}|$. On computing $t_p(\boldsymbol{V}_{21})$ for the case of $\boldsymbol{V}_{21}$, we have $r = 2$, $\boldsymbol{D}_\lambda$ and $\boldsymbol{B}$ are given by

$$\boldsymbol{D}_\lambda = \begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix}, \ \boldsymbol{B} = \begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 0 \\ \sqrt{2} & \sqrt{2} \end{bmatrix},$$

respectively.

It can be seen from some simulation results in Table 2 that the upper percentiles with $\boldsymbol{V} = \boldsymbol{V}_{11}, \boldsymbol{V}_{12}$ are always largest values and those with $\boldsymbol{V} = \boldsymbol{V}_{21}, \boldsymbol{V}_{22}$ are the lower limit values in any positive definite matrix $\boldsymbol{V}$ for each parameter.

In Table 2, the upper bounds for the conservativeness of multiple comparisons with a control are presented. For example, when $\alpha = 0.1$ and $\nu = 20$, we note that $0.9 \leq Q(t_c^*, \boldsymbol{V}, \mathbb{D}) < 0.948$ for any positive definite $\boldsymbol{V}$. It may be noted from simulation study that the coverage probabilities do not depend on p. Also, it may be noted that the coverage probabilities are large as ν is large.

Acknowledgements

The authors wish to thank Mr. Jun Kikuchi and Ms. Mariko Tomita for the Monte Carlo simulation in numerical examinations. The research of the first author was supported in part by Grant-in-Aid for Encouragement of Young Scientists under Contract Number 15700238.

References

Cramér, H., 1946. Mathematical Methods of Statistics. Prinston University Press. Prinston, New Jersey, pp. 297.

Krishnaiah, P. R., 1979. Some developments on simultaneous test procedures. Developments in Statistics, Vol. 2 (Krishnaiah, P. R. ed.) North-Holland, Amsterdam 157–201.

Seo, T., 1995. Simultaneous confidence procedures for multiple comparisons of mean vectors in multivariate normal populations. Hiroshima Math. J. 25, 387–422.

Seo, T., 1996. A note on the conservative multivariate Tukey-Kramer multiple comparison procedure. Amer. J. Math. Manage. Sci. 16, 251–266.

Seo, T., Mano, S. and Fujikoshi, Y., 1994. A generalized Tukey conjecture for multiple comparisons among mean vectors. J. Amer. Statist. Assoc. 89, 676–679.

Siotani, M. 1959a. On the extreme value of the generalized distances of the individual points in the multivariate normal sample. Ann. Inst. Statist. Math. 10, 183–208.

Siotani, M. 1959b. On the range in the multivariate case. Proc. Inst. Statist. Math. 6, 155–165.

Siotani, M. 1960. Note on multivariate confidence bounds. Ann. Inst. Statist. Math. 11, 167–182.

Seo, T. and Siotani, M., 1992. The multivariate Studentized range and its upper percentiles. J. Japan Statist. Soc. 22, 123–137.

Siotani, M., Hayakawa, T. and Fujikoshi, Y., 1985. Modern Multivariate Analysis : A Graduate Course and Handbook. American Sciences Press, Ohio.

Somerville, P. N., 1993. On the conservatism of the Tukey-Kramer multiple comparison procedure. Statist. Prob. Lett. 16, 343–345.

Table 1: Simulation results for pairwise comparison

p	ν	α	$\boldsymbol{V}=\boldsymbol{I}$	$\boldsymbol{V}=\boldsymbol{V}_0$	Q
1	20	0.01	3.280	2.843	0.996
		0.05	2.530	2.085	0.980
		0.1	2.177	1.725	0.958
	40	0.01	3.089	2.706	0.996
		0.05	2.434	2.021	0.980
		0.1	2.113	1.684	0.959
	60	0.01	3.027	2.660	0.996
		0.05	2.403	2.000	0.981
		0.1	2.092	1.670	0.959
2	20	0.01	4.006	3.533	0.996
		0.05	3.200	2.723	0.980
		0.1	2.826	2.342	0.958
	40	0.01	3.655	3.265	0.996
		0.05	2.999	2.578	0.980
		0.1	2.678	2.239	0.959
	60	0.01	3.552	3.182	0.996
		0.05	2.936	2.532	0.981
		0.1	2.631	2.206	0.960
5	20	0.01	5.897	5.269	0.996
		0.05	4.835	4.224	0.980
		0.1	4.357	3.746	0.958
	40	0.01	4.893	4.457	0.996
		0.05	4.172	3.709	0.981
		0.1	3.826	3.344	0.960
	60	0.01	4.633	4.245	0.996
		0.05	3.994	3.570	0.981
		0.1	3.680	3.234	0.961

Table 2: Simulation results for multiple comparisons with a control

p	ν	α	$\boldsymbol{V}=\boldsymbol{I}$	$\boldsymbol{V}=\boldsymbol{V}_{11}$	$\boldsymbol{V}=\boldsymbol{V}_{12}$	$\boldsymbol{V}=\boldsymbol{V}_{21}$	$\boldsymbol{V}=\boldsymbol{V}_{22}$	Q
1	20	0.01	3.125	3.145	3.146	2.845	2.845	0.995
		0.05	2.377	2.410	2.409	2.086	2.086	0.974
		0.1	2.026	2.064	2.064	1.724	1.724	0.948
	40	0.01	2.951	2.969	2.967	2.705	2.705	0.995
		0.05	2.292	2.321	2.321	2.021	2.021	0.975
		0.1	1.970	2.005	2.006	1.684	1.684	0.948
	60	0.01	2.898	2.913	2.913	2.659	2.659	0.996
		0.05	2.266	2.292	2.291	2.000	2.000	0.976
		0.1	1.951	1.986	1.986	1.671	1.671	0.951
2	20	0.01	3.837	3.858	3.858	3.533	3.533	0.995
		0.05	3.035	3.068	3.068	2.723	2.723	0.974
		0.1	2.661	2.701	2.701	2.342	2.342	0.948
	40	0.01	3.516	3.531	3.531	3.264	3.264	0.995
		0.05	2.853	2.877	2.878	2.577	2.577	0.974
		0.1	2.529	2.561	2.561	2.239	2.239	0.948
	60	0.01	3.421	3.435	3.436	3.184	3.184	0.996
		0.05	2.798	2.820	2.820	2.532	2.532	0.978
		0.1	2.488	2.518	2.518	2.207	2.207	0.953
5	20	0.01	5.672	5.702	5.701	5.270	5.270	0.995
		0.05	4.621	4.660	4.661	4.223	4.223	0.974
		0.1	4.147	4.193	4.193	3.745	3.745	0.948
	40	0.01	4.735	4.751	4.751	4.457	4.457	0.995
		0.05	4.011	4.037	4.037	3.710	3.710	0.975
		0.1	3.661	3.694	3.694	3.345	3.345	0.948
	60	0.01	4.496	4.505	4.506	4.244	4.244	0.997
		0.05	3.847	3.669	3.869	3.571	3.571	0.983
		0.1	3.528	3.556	3.557	3.235	3.235	0.962