Hybrid Serial Gatekeeping Procedures for All-Pairwise Comparisons in Multi-Sample Models

Taka-aki Shiraishi*

In q multi-sample models, we consider multiple comparison tests for all-pairwise differences of means. Let $\mathcal{H}^{(p)}$ be the family of null hypothesis among $k^{(p)}$ means for $p = 1, \ldots, q$. The family $\mathcal{H}^{(1)} \succ \ldots \succ \mathcal{H}^{(q)}$ has the order of priority. This paper describes procedures for performing multiple comparison tests at level α based on serial gatekeeping methods. In the p-th stage, a test procedure under unrestricted means or a test procedure under order restricted means is used. The power of the proposed tests is much superior to the serial gatekeeping methods based on Bonferroni tests which are proposed by Maurer et al. (1995).

Keywords: Multiple comparisons, Closed testing procedures, Parametric tests

1 Introduction

The homoscedastic k sample model is expressed as

$$X_{ij} = \mu_i + e_{ij} \quad (j = 1, \ldots, n_i;\ i = 1, \ldots, k),$$

where e_{ij}'s are independent and identically distributed normal with mean 0 and variance σ^2 unknown. Then Tukey (1953) and Kramer (1953) proposed single-step procedures as multiple comparison tests of level α for all-pairwise comparisons of $\{$the null hypothesis $H_{(i,i')} : \mu_i = \mu_{i'}$ vs. the alternative $H^A_{(i,i')} : \mu_i \neq \mu_{i'} \big| (i,i') \in \mathcal{U}_k\}$, where $\mathcal{U}_k := \{(i,i') \mid 1 \leq i < i' \leq k\}$.

As multi-step procedures, Ryan (1960), Einot and Gabriel (1975), and Welsch (1977) gave the REGW (Ryan–Einot–Gabriel–Welsch) methods. The REGW test procedures are included in the SPSS system. Shiraishi (2011) proposed closed testing procedures which are more powerful than the REGW tests. Shiraishi and Sugiura (2018) showed that the closed testing procedures are more powerful than the Tukey-Kramer method. When the simple order restrictions

$$\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k \tag{1.1}$$

is satisfied, Hayter (1990) proposed single-step simultaneous tests for $\{$the null hypothesis $H_{(i,i')}$ vs. the alternative $H^{OA}_{(i,i')} : \mu_i < \mu_{i'} \big| (i,i') \in \mathcal{U}_k\}$ under the equal sample sizes $n_1 = \ldots = n_k$. Shiraishi (2014) proposed closed testing procedures and showed that (i) the proposed multi-step procedures are more powerful than the single-step procedure of Hayter (1990), and (ii) confidence regions induced by the multi-step procedures are equivalent to simultaneous confidence intervals. Under unequal sample sizes, Shiraishi and Matsuda (2016) proposed closed testing procedures based on Bartholomew's tests.

Gatekeeping procedures became to be used in recent years as a convenient way to handle relationships between multiple hierarchical objectives. To solve questions concerning different objectives, null hypotheses are divided into q ordered families, $\mathcal{F}^{(1)} \succ \ldots \succ \mathcal{F}^{(q)}$. Westfall and Krishen (2001) proposed the serial gatekeeping procedures in which the hypotheses in $\mathcal{F}^{(p+1)}$ are tested if and only if all hypotheses in $\mathcal{F}^{(p)}$ are rejected ($1 \leq p \leq q-1$). The individual test procedures are based on Bonferroni tests.

In this paper, we consider q multi-sample models. For the p-th multi-sample model such that $1 \leq p \leq q$, $(X^{(p)}_{i1}, \ldots, X^{(p)}_{in^{(p)}_i})$ is a random sample of size $n^{(p)}_i$ from the i-th normal population with mean $\mu^{(p)}_i$ and variance $\sigma^2_{(p)}$ ($i = 1, \ldots, k^{(p)}$), that is, $P(X^{(p)}_{ij} \leq x) = \Phi((x - \mu^{(p)}_i)/\sigma_{(p)})$, where $\Phi(x)$ is a standard normal distribution function. Futhermore, for fixed p, $X^{(p)}_{ij}$'s are independent.

*Faculty of Sciences and Engineering, Nanzan University, E-Mail: marble@nanzan-u.ac.jp

We need not assume that $\boldsymbol{X}^{(1)},\dots,\boldsymbol{X}^{(q)}$ are independent, where $\boldsymbol{X}^{(p)} := \left(X_{11}^{(p)},\dots,X_{k^{(p)}n_{k^{(p)}}^{(p)}}^{(p)}\right)$.
Let

$$\mathcal{H}^{(p)} := \left\{ H_{ii'}^{(p)} : \ \mu_i^{(p)} = \mu_{i'}^{(p)} \,\middle|\, (i,i') \in \mathcal{U}_{k^{(p)}} \right\} \tag{1.2}$$

be the family of null hypothesis among $k^{(p)}$ means for $p = 1,\dots,q$, where

$$\mathcal{U}_{k^{(p)}} := \{(i,i') \mid 1 \le i < i' \le k^{(p)}\}. \tag{1.3}$$

The family $\mathcal{H}^{(1)},\dots,\mathcal{H}^{(q)}$ has the order of priority.

$$\mathcal{H}^{(1)} \succ \dots \succ \mathcal{H}^{(q)}. \tag{1.4}$$

This paper describes procedures for performing multiple comparison tests at level α based on serial gatekeeping methods. In the p-th stage, a test procedure under unrestricted means or a test procedure under order restricted means is used. The proposed hybrid procedures are the parametric methods assuming the normal distribution. The methods of Tukey (1953), Kramer (1953), Hayter (1990), Shiraishi (2011, 2014), and Shiraishi and Matsuda (2016) are used in the hybrid gatekeeping procedures. The power of the proposed tests is much superior to the serial gatekeeping methods based on Bonferroni tests which are usually used.

2 Multiple comparison test procedures under unrestricted means in the p-th multi-sample model

The unbiased estimators for $\mu_i^{(p)}$, overall mean $\nu^{(p)} = \sum_{i=1}^{k^{(p)}} n_i^{(p)}\mu_i^{(p)}/n^{(p)}$, and $\sigma_{(p)}^2$, respectively, are given by $\hat{\mu}_i^{(p)} = \bar{X}_{i\cdot}^{(p)}$, $\hat{\nu}^{(p)} = \bar{X}_{\cdot\cdot}^{(p)}$, and

$$V_E^{(p)} = \frac{1}{m^{(p)}} \sum_{i=1}^{k^{(p)}} \sum_{j=1}^{n_i^{(p)}} (X_{ij}^{(p)} - \bar{X}_{i\cdot}^{(p)})^2 \tag{2.1}$$

where $\bar{X}_{i\cdot}^{(p)} := (1/n_i^{(p)}) \sum_{j=1}^{n_i^{(p)}} X_{ij}^{(p)}$, $\bar{X}_{\cdot\cdot}^{(p)} := (1/n^{(p)}) \sum_{i=1}^{k^{(p)}} \sum_{j=1}^{n_i^{(p)}} X_{ij}^{(p)}$,

$$m^{(p)} := n^{(p)} - k^{(p)}, \text{ and } n^{(p)} := \sum_{i=1}^{k^{(p)}} n_i^{(p)}. \tag{2.2}$$

The ratio $F_t^{(p)} := \sum_{i=1}^{k^{(p)}} n_i^{(p)}(\bar{X}_{i\cdot}^{(p)} - \bar{X}_{\cdot\cdot}^{(p)})^2/\{(k^{(p)}-1)V_E^{(p)}\}$ is used to test the null hypothesis of no treatment effects,

$$H_0^{(p)} : \ \mu_1^{(p)} = \dots = \mu_{k^{(p)}}^{(p)}, \tag{2.3}$$

as follows. We reject $H_0^{(p)}$ at level α if $F_t^{(p)} > F_{m^{(p)}}^{k^{(p)}-1}(\alpha)$, where $F_{m^{(p)}}^{k^{(p)}-1}(\alpha)$ denotes the upper $100\alpha\%$ point of F-distribution with degrees of freedom $(k^{(p)}-1, m^{(p)})$.

For specified i,i' such that $1 \le i < i' \le k^{(p)}$, if we are interested in testing the null hypothesis $H_{(i,i')}^{(p)} : \mu_i^{(p)} = \mu_{i'}^{(p)}$ vs. the alternative $H_{(i,i')}^{(p)A} : \mu_i^{(p)} \ne \mu_{i'}^{(p)}$, we can use the two-sided two-sample t-test. In this section, we consider test procedures for all-pairwise comparisons of

$$\left\{\text{the null hypothesis } H_{(i,i')}^{(p)} \text{ vs. the alternative } H_{(i,i')}^{(p)A} \,\middle|\, (i,i') \in \mathcal{U}_{k^{(p)}}\right\}, \tag{2.4}$$

where $\mathcal{U}_{k^{(p)}}$ is defined by (1.3).

Tukey (1953) and Kramer (1956) proposed single-step procedures as multiple comparison tests of level α. We introduce two distribution functions of $TA(t|k^{(p)},m^{(p)})$ and $TA^*(t|k^{(p)},m^{(p)},\boldsymbol{\lambda}_n^{(p)})$.

$$TA(t|k^{(p)},m^{(p)}) := k^{(p)} \int_0^\infty \Big[\int_{-\infty}^\infty \{\Phi(x) - \Phi(x - \sqrt{2}\cdot ts)\}^{k^{(p)}-1} d\Phi(x)\Big] g(s|m^{(p)})ds, \tag{2.5}$$

$$TA^*(t|k^{(p)},m^{(p)},\boldsymbol{\lambda}_n^{(p)}) := \sum_{j=1}^{k^{(p)}} \int_0^\infty \left[\int_{-\infty}^\infty \prod_{\substack{i=1\\ i\neq j}}^{k^{(p)}} \left\{ \Phi\left(\sqrt{\frac{\lambda_{ni}^{(p)}}{\lambda_{nj}^{(p)}}} \cdot x \right) - \Phi\left(\sqrt{\frac{\lambda_{ni}^{(p)}}{\lambda_{nj}^{(p)}}} \cdot x - \sqrt{\frac{\lambda_{ni}^{(p)}+\lambda_{nj}^{(p)}}{\lambda_{nj}^{(p)}}} \cdot ts \right) \right\} d\Phi(x) \right] g(s|m^{(p)})ds.$$

where

$$\boldsymbol{\lambda}_n^{(p)} := (\lambda_{n1}^{(p)},\ldots,\lambda_{nk^{(p)}}^{(p)}),\ \lambda_{ni}^{(p)} := n_i^{(p)}/n^{(p)} \quad (i=1,\ldots,k^{(p)}), \tag{2.6}$$

$$g(s|m^{(p)}) := \frac{(m^{(p)})^{m^{(p)}/2}}{\Gamma(m^{(p)}/2)2^{(m^{(p)}/2-1)}} s^{m^{(p)}-1}\exp(-m^{(p)}s^2/2), \tag{2.7}$$

and $m^{(p)}$ is defined in (2.2).

$TA(t/\sqrt{2}|p)$ is referred to as studentized range distribution. We put

$$T_{i'i}^{(p)} := \frac{\bar{X}_{i'\cdot}^{(p)} - \bar{X}_{i\cdot}^{(p)}}{\sqrt{V_E^{(p)}\left(\frac{1}{n_i^{(p)}}+\frac{1}{n_{i'}^{(p)}}\right)}} \quad ((i,i')\in\mathcal{U}_{k^{(p)}}). \tag{2.8}$$

Then we get, for $t>0$,

$$TA(t|k^{(p)},m^{(p)}) \le P_{(p)0}\left(\max_{(i,i')\in\mathcal{U}_{k^{(p)}}} |T_{i'i}^{(p)}| \le t \right) \le TA^*(t|k^{(p)},m^{(p)},\boldsymbol{\lambda}_n^{(p)}) \tag{2.9}$$

holds, where $P_{(p)0}(\cdot)$ stands for probability measure under the null hypothesis $H_0^{(p)}$. When $n_1^{(p)} = \ldots = n_{k^{(p)}}^{(p)}$ is satisfied, both of the inequalities of (2.9) become an equality.

The left hand side of the inequality (2.9) is derived from main theorem of Hayter (1984). The right hand side of the inequality (2.9) is given by Shiraishi (2006). For a given α such that $0<\alpha<1$, we put

$$ta(k^{(p)},m^{(p)};\alpha) := \text{a solution of } t \text{ satisfying the equation } TA(t|k^{(p)},m^{(p)}) = 1-\alpha. \tag{2.10}$$

[2.1] Single-step tests based on t-statistics

The Tukey-Kramer simultaneous test of level α for the null hypotheses $\{H_{(i,i')}^{(p)} \mid (i,i')\in\mathcal{U}_{k^{(p)}}\}$ consists in rejecting $H_{(i,i')}^{(p)}$ for $(i,i')\in\mathcal{U}_{k^{(p)}}$ such that $|T_{i'i}^{(p)}| > ta(k^{(p)},m^{(p)};\alpha)$. From the left inequality of (2.9), we find that the Tukey-Kramer simultaneous test is conservative. Under the condition of $\max_{1\le i\le k^{(p)}} n_i^{(p)}/\min_{1\le i\le k^{(p)}} n_i^{(p)} \le 2$, Shiraishi (2006) found that the values of $TA^*(t|k^{(p)},m^{(p)},\boldsymbol{\lambda}_n^{(p)}) - TA(t|k^{(p)},m^{(p)})$ is nearly equal to 0 for various values of t from numerical integration. Therefore the conservativeness of the Tukey-Kramer method is small.

The closure of $\mathcal{H}^{(p)}$ is given by

$$\overline{\mathcal{H}^{(p)}} = \left\{ \bigwedge_{\boldsymbol{v}\in V} H_{\boldsymbol{v}}^{(p)} \mid \emptyset \subsetneq V \subset \mathcal{U}_{k^{(p)}} \right\},$$

where $\bigwedge$ denotes the conjunction symbol (Refer to Enderton (2001)). Then, we get

$$\bigwedge_{\boldsymbol{v}\in V} H_{\boldsymbol{v}}^{(p)} : \text{for any } (i,i')\in V,\ \mu_i^{(p)} = \mu_{i'}^{(p)} \text{ holds.} \tag{2.11}$$

For an integer $J^{(p)}$ and disjoint sets $I_1^{(p)},\ldots,I_{J^{(p)}}^{(p)} \subset \{1,\ldots,k^{(p)}\}$, we define the null hypothesis $H^{(p)}(I_1^{(p)},\ldots,I_{J^{(p)}}^{(p)})$ by

$$H^{(p)}(I_1^{(p)},\ldots,I_{J^{(p)}}^{(p)}) : \text{for any integer } j \text{ such that } 1\le j\le J^{(p)} \text{ and for any } i,i'\in I_j^{(p)},\ \mu_i^{(p)} = \mu_{i'}^{(p)} \text{ holds.} \tag{2.12}$$

From (2.11) and (2.12), for any nonempty $V \subset \mathcal{U}_{k^{(p)}}$, there exist an integer $J^{(p)}$ and disjoint sets $I_1^{(p)}, \ldots, I_{J^{(p)}}^{(p)}$ such that

$$\bigwedge_{\boldsymbol{v} \in V} H_{\boldsymbol{v}}^{(p)} = H^{(p)}(I_1^{(p)}, \ldots, I_{J^{(p)}}^{(p)}) \tag{2.13}$$

and $\#(I_j^{(p)}) \geq 2$ $(j = 1, \ldots, J^{(p)})$, where $\#(A)$ stands for the cardinal number of set A. For $H^{(p)}(I_1^{(p)}, \ldots, I_{J^{(p)}}^{(p)})$ of (2.13), we set

$$M^{(p)} := M^{(p)}(I_1^{(p)}, \ldots, I_{J^{(p)}}^{(p)}) = \sum_{j=1}^{J^{(p)}} \ell_j^{(p)}, \ \ell_j^{(p)} := \#(I_j^{(p)}). \tag{2.14}$$

Let us put

$$T^{(p)}(I_j^{(p)}) := \max_{i<i',\ i,i' \in I_j^{(p)}} \left| T_{i'i}^{(p)} \right| \quad (j = 1, \ldots, J^{(p)}).$$

Then, we propose the stepwise procedure [2.2].

[2.2] Stepwise procedure based on t-statistics

For $\ell^{(p)} = \ell_1^{(p)}, \ldots, \ell_{J^{(p)}}^{(p)}$, we define $\alpha(M^{(p)}, \ell^{(p)})$ by

$$\alpha(M^{(p)}, \ell^{(p)}) := 1 - (1 - \alpha)^{\ell^{(p)}/M^{(p)}}. \tag{2.15}$$

Corresponding to (2.5), we put

$$TA(t|\ell^{(p)}, m^{(p)}) := \ell^{(p)} \int_0^{\infty} \Big[\int_{-\infty}^{\infty} \{\Phi(x) - \Phi(x - \sqrt{2} \cdot ts)\}^{\ell^{(p)}-1} d\Phi(x) \Big] g(s|m^{(p)}) ds. \tag{2.16}$$

By obeying the notation $ta(k^{(p)}, m^{(p)}; \alpha)$, we get

$$TA(ta(\ell^{(p)}, m^{(p)}; \alpha(M^{(p)}, \ell^{(p)}))|\ell^{(p)}, m^{(p)}) = 1 - \alpha(M^{(p)}, \ell^{(p)}), \tag{2.17}$$

that is, $ta(\ell^{(p)}, m^{(p)}; \alpha(M^{(p)}, \ell^{(p)}))$ is an upper $100\alpha(M^{(p)}, \ell^{(p)})\%$ point of the distribution $TA(t|\ell^{(p)}, m^{(p)})$.

(a) $J^{(p)} \geq 2$

Whenever $ta\left(\ell_j^{(p)}, m^{(p)}; \alpha(M^{(p)}, \ell_j^{(p)})\right) < T^{(p)}(I_j^{(p)})$ holds for an integer j such that $1 \leq j \leq J^{(p)}$, we reject the hypothesis $\bigwedge_{\boldsymbol{v} \in V} H_{\boldsymbol{v}}^{(p)}$.

(b) $J^{(p)} = 1$ $(M^{(p)} = \ell_1^{(p)})$

Whenever $ta\left(M^{(p)}, m^{(p)}; \alpha\right) < T^{(p)}(I_1^{(p)})$ holds, we reject the hypothesis $\bigwedge_{\boldsymbol{v} \in V} H_{\boldsymbol{v}}^{(p)}$.

By using the methods of (a) and (b), when $\bigwedge_{\boldsymbol{v} \in V} H_{\boldsymbol{v}}^{(p)}$ is rejected for any V such that $(i, i') \in V \subset \mathcal{U}_{k^{(p)}}$, the null hypothesis $H_{(i,i')}^{(p)}$ is rejected as a multiple comparison test.

As a closed testing procedure under assuming normality for $k^{(p)}$-sample model, the REGW (Ryan–Einot–Gabriel–Welsch) method is utilized. The REGW method is also stated in Hsu (1996). In order to introduce the REGW method, we define the hypothesis $H^{(p)}(I^{(p)})$ by

$$H^{(p)}(I^{(p)}) : \ \mu_i^{(p)} = \mu_{i'}^{(p)} \text{ for } i, i' \in I^{(p)}$$

and we put $\imath^{(p)} = \#(I^{(p)})$, where $I^{(p)}$ $(I^{(p)} \subset \{1, \ldots, k^{(p)}\})$ and $\#(I^{(p)}) \geq 2$. Suppose $k^{(p)} \geq 4$. We define $\alpha^*(\imath^{(p)})$ by

$$\alpha^*(\imath^{(p)}) = \begin{cases} 1 - (1 - \alpha)^{\imath^{(p)}/k^{(p)}} & (2 \leq \imath^{(p)} \leq k^{(p)} - 2) \\ \alpha & (\imath = k^{(p)} - 1, k^{(p)}). \end{cases} \tag{2.18}$$

[2.3] REGW method
If $ta\left(\imath^{(p)}, m^{(p)}; \alpha^*(\imath^{(p)})\right) < T^{(p)}(I^{(p)})$ for any $I^{(p)}$ such that $i, i' \in I^{(p)}$, $H^{(p)}_{(i,i')}$ is rejected.

Suppose $\ell^{(p)}_j = \imath^{(p)} = \ell^{(p)}$. Then, since

$$1-(1-\alpha)^{\ell^{(p)}/M^{(p)}} \geq 1-(1-\alpha)^{\ell^{(p)}/k^{(p)}},$$

in testing the null hypothesis $\bigwedge_{\boldsymbol{v}\in V} H^{(p)}_{\boldsymbol{v}}$, the rejection region for the closed testing procedure [2.2] includes the one for the closed testing procedure [2.3]. Therefore, the closed testing procedure [2.2] is more powerful than the closed testing procedure [2.3].

[2.4] Stepwise procedure based on F-statistics
Let us put

$$S^{(p)}(I^{(p)}_j) := \sum_{i\in I^{(p)}_j} n^{(p)}_i \left(\bar{X}^{(p)}_{i\cdot} - \bar{X}^{(p)}_{I^{(p)}_j}\right)^2 / \{(\ell^{(p)}_j - 1)V^{(p)}_E\} \ (j = 1, \ldots, J^{(p)}), \tag{2.19}$$

where $I^{(p)}_j$ is defined in (2.13), $\ell^{(p)}_j$ is defined in (2.14), and $\bar{X}^{(p)}_{I^{(p)}_j} := \sum_{i\in I^{(p)}_j} n^{(p)}_i \bar{X}^{(p)}_{i\cdot} / \sum_{i\in I^{(p)}_j} n^{(p)}_i$. In the procedure [2.2], replace $ta\left(\ell^{(p)}_j, m^{(p)}; \alpha(M^{(p)}, \ell^{(p)}_j)\right) < T^{(p)}(I^{(p)}_j)$ and $ta\left(M^{(p)}, m^{(p)}; \alpha\right) < T^{(p)}(I^{(p)}_1)$ with $F^{\ell^{(p)}_j - 1}_{m^{(p)}}\left(\alpha(M^{(p)}, \ell^{(p)}_j)\right) < S^{(p)}(I^{(p)}_j)$ and $F^{M^{(p)}-1}_{m^{(p)}}(\alpha) < S(I^{(p)}_1)$, respectively. Then, this procedure also becomes a closed test.

3 Multiple comparison test procedures under order restricted means in the p-th multi-sample model

When the simple order restrictions

$$\mu^{(p)}_1 \leq \mu^{(p)}_2 \leq \ldots \leq \mu^{(p)}_{k^{(p)}} \tag{3.1}$$

is satisfied, we consider the null hypothesis $H^{(p)}_0$ vs. the alternative $H^{(p)A} : \mu^{(p)}_1 \leq \mu^{(p)}_2 \leq \ldots \leq \mu^{(p)}_{k^{(p)}}$ with at least one strict inequality, which is equivalent to $H^{(p)}_0 : \mu^{(p)}_1 = \mu^{(p)}_{k^{(p)}}$ vs. $H^{(p)A} : \mu^{(p)}_1 < \mu^{(p)}_{k^{(p)}}$, where $H^{(p)}_0$ is defined by (2.3). We define $\{\hat{\mu}^{(p)o}_i | \ i = 1, \ldots, k^{(p)}\}$ by $\{u_i | \ i = 1, \ldots, k^{(p)}\}$ which minimize $\sum_{i=1}^{k^{(p)}} \lambda^{(p)}_{ni} \left(u_i - \bar{X}^{(p)}_{i\cdot}\right)^2$ under simple order restrictions $u_1 \leq u_2 \leq \ldots \leq u_{k^{(p)}}$, that is.,

$$\sum_{i=1}^{k^{(p)}} \lambda^{(p)}_{ni} \left(\hat{\mu}^{(p)o}_i - \bar{X}^{(p)}_{i\cdot}\right)^2 = \min_{u_1\leq\ldots\leq u_{k^{(p)}}} \sum_{i=1}^{k^{(p)}} \lambda^{(p)}_{ni} \left(u_i - \bar{X}^{(p)}_{i\cdot}\right)^2.$$

$\hat{\mu}^{(p)o}_1, \ldots, \hat{\mu}^{(p)o}_{k^{(p)}}$ are computed by using the pool-adjacent-violators algorithm stated in Robertson et al. (1988). Accordingly, we find

$$\tilde{\mu}^{(p)o}_i = \max_{1\leq a\leq i} \min_{i\leq b\leq k^{(p)}} \frac{\sum_{j=a}^{b} \lambda^{(p)}_{nj} \bar{X}^{(p)}_{j\cdot}}{\sum_{j=a}^{b} \lambda^{(p)}_{nj}} = \max_{1\leq a\leq i} \min_{i\leq b\leq k^{(p)}} \frac{\sum_{j=a}^{b} n^{(p)}_j \bar{X}^{(p)}_{j\cdot}}{\sum_{j=a}^{b} n^{(p)}_j}. \tag{3.2}$$

We put

$$\bar{X}^{2(p)}_{k^{(p)}} := \frac{1}{\sigma^2_{(p)}} \sum_{i=1}^{k^{(p)}} n^{(p)}_i \left(\hat{\mu}^{(p)o}_i - \sum_{j=1}^{k^{(p)}} \lambda^{(p)}_{nj} \bar{X}^{(p)}_{j\cdot}\right)^2. \tag{3.3}$$

We define $\tilde{\nu}_1^{(p)o}, \ldots, \tilde{\nu}_{k^{(p)}}^{(p)o}$ by

$$\sum_{i=1}^{k^{(p)}} \lambda_{ni}^{(p)} \left(\tilde{\nu}_i^{(p)o} - Y_i^{(p)}\right)^2 = \min_{u_1 \leq \ldots \leq u_{k^{(p)}}} \sum_{i=1}^{k^{(p)}} \lambda_{ni}^{(p)} \left(u_i - Y_i^{(p)}\right)^2,$$

where $Y_1^{(p)}, \ldots, Y_{k^{(p)}}^{(p)}$ are independent and $Y_i^{(p)} \sim N(0, 1/\lambda_{ni}^{(p)})$ $(i = 1, \ldots, k^{(p)})$. Let $P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)})$ be the probability that $\tilde{\nu}_1^{(p)o}, \ldots, \tilde{\nu}_{k^{(p)}}^{(p)o}$ takes exactly L distinct values, where $\boldsymbol{\lambda}_n^{(p)}$ is defined by (2.6). Then, for positive constant c, $P(L, k^{(p)}; c\boldsymbol{\lambda}_n^{(p)}) = P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)})$ holds. Futhermore, from Theorem 2.3.1 of Robertson et al. (1988), we get

$$P_{(p)0}(\bar{\chi}_{k^{(p)}}^{2(p)} \geq t) = \sum_{L=2}^{k^{(p)}} P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)}) P\left(\chi_{L-1}^2 \geq t\right) \quad (t > 0), \tag{3.4}$$

where χ_{L-1}^2 is a chi-square variable with $L-1$ degrees of freedom. The recurrence formula of computing $P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)})$ is written in Robertson et al. (1988). The fundamental algorithm of $P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)})$ based on sinc integral is stated in section 7.4 of Shiraishi et al. (2019).

Since $P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)})$ depends on L and $k^{(p)}$ for

$$\lambda_{n1}^{(p)} = \ldots = \lambda_{nk^{(p)}}^{(p)} = 1/k^{(p)}, \tag{3.5}$$

we simply write $P(L, k^{(p)})$ instead of $P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)})$. Barlow et al. (1972) offers the following recurrence formula.

$$\begin{aligned} &P(1, k^{(p)}) = \frac{1}{k^{(p)}}, \\ &P(L, k^{(p)}) = \frac{1}{k^{(p)}} \left\{ (k^{(p)} - 1) P(L, k^{(p)} - 1) + P(L-1, k^{(p)} - 1) \right\}, \qquad (2 \leq L \leq k^{(p)} - 1) \\ &P(k^{(p)}, k^{(p)}) = \frac{1}{k^{(p)}!}. \end{aligned}$$

In $\bar{\chi}_{k^{(p)}}^{2(p)}$ defined by (3.3), replace $\sigma_{(p)}^2$ with the estimator $V_E^{(p)}$. Then the subsequent statistic is denoted by

$$\bar{B}_{(p)}^2 := \frac{\sum_{i=1}^{k^{(p)}} n_i^{(p)} (\tilde{\mu}_i^{(p)o} - \bar{X}_{\cdot\cdot}^{(p)})^2}{V_E^{(p)}} = \frac{\bar{\chi}_{k^{(p)}}^{2(p)}}{V_E^{(p)}/\sigma_{(p)}^2}. \tag{3.6}$$

Since $\bar{\chi}_{k^{(p)}}^{2(p)}$ and $V_E^{(p)}$ are independent, from (3.4), (3.6) and the relationship with $m^{(p)} V_E^{(p)}/\sigma_{(p)}^2 \sim \chi_{m^{(p)}}^2$, we find, for $t > 0$,

$$P_{(p)0}(\bar{B}_{(p)}^2 \geq t) = \sum_{L=2}^{k^{(p)}} P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)}) P\left((L-1) F_{m^{(p)}}^{L-1} \geq t\right), \tag{3.7}$$

where $F_{m^{(p)}}^{L-1}$ denotes the random variable having the F-distribution with $L-1$ and $m^{(p)}$ degrees of freedom. For a given α such that $0 < \alpha < 1$, we give the following equation of t.

$$\sum_{L=2}^{k^{(p)}} P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)}) P\left((L-1) F_{m^{(p)}}^{L-1} \geq t\right) = \alpha$$

We denote a solution of this equation by $\bar{b}^2(k^{(p)}, m^{(p)}, \boldsymbol{\lambda}_n^{(p)}; \alpha)$. Thus, from (3.7), as a test of level α for the null hypothesis $H_0^{(p)}$ vs. the alternative $H^{(p)A}$, we can propose to reject $H_0^{(p)}$ when the value of $\bar{B}_{(p)}^2$ is greater than $\bar{b}^2(k^{(p)}, m^{(p)}, \boldsymbol{\lambda}_n^{(p)}; \alpha)$.

For a given α such that $0 < \alpha < 1$, we give the following equation of t.

$$\sum_{L=2}^{k^{(p)}} P(L, k^{(p)}; \boldsymbol{\lambda}_n^{(p)}) P\left(\chi_{L-1}^2 \geq t\right) = \alpha$$

We denote a solution of this equation by $\bar{c}^2(k^{(p)}, \boldsymbol{\lambda}_n^{(p)}; \alpha)$. Hence, from (3.4), we can reject $H_0^{(p)}$ when the value of $\bar{\chi}^{2(p)}_{k^{(p)}}$ is greater than $\bar{c}^2(k^{(p)}, \boldsymbol{\lambda}_n^{(p)}; \alpha)$.

For specified i, i' such that $(i, i') \in \mathcal{U}_{k^{(p)}}$, if we are interested in testing of

$$\text{the null hypothesis } H_{(i,i')}^{(p)} : \mu_i^{(p)} = \mu_{i'}^{(p)} \text{ vs. the alternative } H_{(i,i')}^{(p)OA} : \mu_i^{(p)} < \mu_{i'}^{(p)}, \tag{3.8}$$

we can use the one-sided two-sample t-test. We consider test procedures for all-pairwise comparisons of $\left\{\text{the null hypothesis } H_{(i,i')}^{(p)} \text{ vs. the alternative } H_{(i,i')}^{(p)OA} \middle| (i,i') \in \mathcal{U}_{k^{(p)}}\right\}$, where $\mathcal{U}_{k^{(p)}}$ is defined by (1.3). Under the equality of sample sizes $n_1^{(p)} = \ldots = n_{k^{(p)}}^{(p)}$, Hayter (1990) proposed single-step simultaneous tests for $\left\{\text{the null hypothesis } H_{(i,i')}^{(p)} \text{ vs. the alternative } H_{(i,i')}^{(p)OA} \middle| (i,i') \in \mathcal{U}_{k^{(p)}}\right\}$. Shiraishi (2014) proposed closed testing procedures. It is shown that (i) the proposed multi-step procedures are more powerful than the single-step procedure of Hayter (1990), and (ii) confidence regions induced by the multi-step procedures are equivalent to simultaneous confidence intervals.

We add the condition (C1) of equal sample sizes.

$$(C1) \qquad n_1^{(p)} = n_2^{(p)} = \ldots = n_{k^{(p)}}^{(p)} \ (p = 1, \ldots, q).$$

Then $T_{i'i}^{(p)}$ of (2.8) is given by

$$T_{i'i}^{(p)} = \frac{\sqrt{n_1^{(p)}}(\bar{X}_{i'\cdot}^{(p)} - \bar{X}_{i\cdot}^{(p)})}{\sqrt{2V_E^{(p)}}}. \tag{3.9}$$

We put

$$D(t|k^{(p)}) := P\left(\max_{1 \le i < i' \le k^{(p)}} \frac{Z_{i'} - Z_i}{\sqrt{2}} \le t\right), \tag{3.10}$$

where $Z_i \sim N(0,1)$ $(i = 1, \ldots, k^{(p)})$ and $Z_1, \ldots, Z_{k^{(p)}}$ are independent. Shiraishi (2014) gives

$$\lim_{n^{(p)} \to \infty} P_{(p)0}\left(\max_{1 \le i < i' \le k^{(p)}} T_{i'i}^{(p)} \le t\right) = D(t|k^{(p)}). \tag{3.11}$$

Let $U_E^{(p)}$ be a random variable distributed to χ^2-distribution with $m^{(p)}$ degrees of freedom that is independent of $Z_1, \ldots, Z_{k^{(p)}}$. Then we define $TD(t)$ by

$$\begin{aligned} TD(t|k^{(p)}, m^{(p)}) &:= P_{(p)0}\left(\max_{1 \le i < i' \le k^{(p)}} T_{i'i}^{(p)} \le t\right) \\ &= P\left(\max_{1 \le i < i' \le k^{(p)}} \frac{Z_{i'} - Z_i}{\sqrt{2U_E^{(p)}/m^{(p)}}} \le t\right) \\ &= \int_0^\infty D(ts|k^{(p)}) g(s|m^{(p)}) ds, \end{aligned} \tag{3.12}$$

where $g(s|m^{(p)})$ is defined by (2.7). From Shiraishi et al. (2019), we get the following recurrence formula.

$$H_1(t,x) := P\left(\frac{Z_1 - x}{\sqrt{2}} \le t\right) = \Phi(\sqrt{2} \cdot t + x), \tag{3.13}$$

$$\begin{aligned} H_r(t,x) := &\int_{-\infty}^x H_{r-1}(t,y)\varphi(y) dy \\ &+ H_{r-1}(t,x)\{\Phi(\sqrt{2} \cdot t + x) - \Phi(x)\} \quad (2 \le r \le k^{(p)} - 1), \end{aligned} \tag{3.14}$$

$$D(t|k^{(p)}) = \int_{-\infty}^\infty H_{k-1}(t,x)\varphi(x) dx. \tag{3.15}$$

Futhermore from (3.12) and (3.15), we get

$$TD(t|k^{(p)}, m^{(p)}) = \int_0^\infty \left\{\int_{-\infty}^\infty H_{k-1}^{(p)}(ts,x)\varphi(x) dx\right\} g(s|m^{(p)}) ds. \tag{3.16}$$

For a given α such that $0 < \alpha < 1$, we put

$$td(k^{(p)}, m^{(p)}; \alpha) := \text{a solution of } t \text{ satisfying the equation } TD(t|k^{(p)}, m^{(p)}) = 1 - \alpha. \quad (3.17)$$

By using (3.12), we can derive single step procedures proposed by Hayter (1990).

[3.1] Single-step tests based on one-sided t-test statistics

The simultaneous test of level α for the {null hypothesis $H^{(p)}_{(i,i')}$ vs. alternative hypothesis $H^{(p)OA}_{(i,i')}$: $\mu^{(p)}_i < \mu^{(p)}_{i'}$ | $(i,i') \in \mathcal{U}_{k^{(p)}}$} consist in rejecting $H^{(p)}_{(i,i')}$ for $(i,i') \in \mathcal{U}_{k^{(p)}}$ such that $T_{i'i} > td(k^{(p)}, m^{(p)}; \alpha)$.

Next we introduce closed testing procedures. The closure of $\mathcal{H}^{(p)}$ under the order restrictions (3.1) is given by

$$\overline{\mathcal{H}}^{(p)o} = \left\{ \bigwedge_{\boldsymbol{v} \in V} H^{(p)}_{\boldsymbol{v}} \;\middle|\; \emptyset \subsetneq V \subset \mathcal{U}_{k^{(p)}} \right\} = \left\{ \bigwedge_{\boldsymbol{v} \in V^+} H^{(p)}_{\boldsymbol{v}} \;\middle|\; \emptyset \subsetneq V \subset \mathcal{U}_{k^{(p)}} \right\}. \quad (3.18)$$

where

$$V^+ := \{(i, i+1) | \text{ For } (i_0, i'_0) \in V, \ i_0 \le i < i+1 \le i'_0\}. \quad (3.19)$$

Then, we get

$$\bigwedge_{\boldsymbol{v} \in V} H^{(p)}_{\boldsymbol{v}} = \bigwedge_{\boldsymbol{v} \in V^+} H^{(p)}_{\boldsymbol{v}} : \text{for any } (i,i') \in V, \ \mu^{(p)}_i = \mu^{(p)}_{i'} \text{ holds.} \quad (3.20)$$

Let $I^{(p)o}_1, \ldots, I^{(p)o}_{J^{(p)}}$ be disjoint sets satisfying the following property (C2).

(**C2**) There exist integers $\ell^{(p)}_1, \ldots, \ell^{(p)}_{J^{(p)}} \ge 2$ and integers $0 \le s^{(p)}_1 < \ldots < s^{(p)}_{J^{(p)}} < k^{(p)}$ such that

$$I^{(p)o}_j = \{s^{(p)}_j + 1, s^{(p)}_j + 2, \ldots . s^{(p)}_j + \ell^{(p)}_j\} \ (j = 1, \ldots, J^{(p)}) \quad (3.21)$$

$$\text{and } s^{(p)}_j + \ell^{(p)}_j \le s^{(p)}_{j+1} \ (j = 1, \ldots, J^{(p)} - 1).$$

We define the null hypothesis $H^{(p)o}(I^{(p)o}_1, \ldots, I^{(p)o}_{J^{(p)}})$ by

$$H^{(p)o}(I^{(p)o}_1, \ldots, I^{(p)o}_{J^{(p)}}) : \text{for any } j \text{ such that } 1 \le j \le J^{(p)} \text{ and for any } i, i' \in I^{(p)o}_j, \ \mu^{(p)}_i = \mu^{(p)}_{i'} \text{ holds.} \quad (3.22)$$

The elements of $I^{(p)o}_j$ are consecutive integers and $\ell^{(p)}_j = \#(I^{(p)o}_j) \ge 2$. From (3.22), for any nonempty $V \subset \mathcal{U}_{k^{(p)}}$, there exist an integer $J^{(p)}$ and some subsets $I^{(p)o}_1, \ldots, I^{(p)o}_{J^{(p)}} \subset \{1, \ldots, k^{(p)}\}$ satisfying (C2) such that

$$\bigwedge_{\boldsymbol{v} \in V} H^{(p)}_{\boldsymbol{v}} = \bigwedge_{\boldsymbol{v} \in V^+} H^{(p)}_{\boldsymbol{v}} = H^{(p)o}(I^{(p)o}_1, \ldots, I^{(p)o}_{J^{(p)}}). \quad (3.23)$$

Futhermore $H^{(p)o}(I^{(p)o}_1, \ldots, I^{(p)o}_{J^{(p)}})$ is expressed as

$$H^{(p)o}(I^{(p)o}_1, \ldots, I^{(p)o}_{J^{(p)}}) : \ \mu^{(p)}_{s^{(p)}_j + 1} = \mu^{(p)}_{s^{(p)}_j + 2} = \ldots = \mu^{(p)}_{s^{(p)}_j + \ell^{(p)}_j} \ (j = 1, \ldots, J^{(p)}). \quad (3.24)$$

Let us put

$$T^{(p)o}(I^{(p)o}_j) = \max_{s^{(p)}_j + 1 \le i < i' \le s^{(p)}_j + \ell^{(p)}_j} T^{(p)}_{i'i} \quad (j = 1, \ldots, J^{(p)}),$$

where $I^{(p)o}_j$ is defined in (C2) and $T^{(p)}_{i'i}$ is defined by (3.9).

Corresponding to (3.10), (3.12) and (3.17), for $\ell^{(p)}$ such that $2 \le \ell^{(p)} \le k^{(p)}$, we put

$$D(t|\ell^{(p)}) := P\left(\max_{1 \le i < i' \le \ell^{(p)}} \frac{Z_{i'} - Z_i}{\sqrt{2}} \le t \right), \quad (3.25)$$

$$TD(t|\ell^{(p)}, m^{(p)}) := P\left(\max_{1 \le i < i' \le \ell^{(p)}} \frac{Z_{i'} - Z_i}{\sqrt{2U^{(p)}_E / m^{(p)}}} \le t \right)$$

and

$$td(\ell^{(p)}, m^{(p)}; \alpha) := \text{a solution of } t \text{ satisfying the equation}$$
$$TD(t|\ell^{(p)}, m^{(p)}) = 1 - \alpha, \tag{3.26}$$

where Z_i and $U_E^{(p)}$ are random variables used in (3.12).

Corresponding to (3.12), we have

$$TD(t|\ell^{(p)}, m^{(p)}) = \int_0^\infty D(ts|\ell^{(p)}) g(s|m^{(p)}) ds.$$

Then, we propose the stepwise procedure [3.2].

[3.2] Stepwise procedure based on one-sided t-test statistics

For $H^{(p)o}(I_1^{(p)o}, \ldots, I_{J^{(p)}}^{(p)o})$ of (3.23), we set

$$M^{(p)} = M^{(p)}(I_1^{(p)o}, \ldots, I_{J^{(p)}}^{(p)o}) = \sum_{j=1}^{J^{(p)}} \ell_j^{(p)}. \tag{3.27}$$

For $\ell^{(p)} = \ell_1^{(p)}, \ldots, \ell_J^{(p)}$, we define $\alpha(M^{(p)}, \ell^{(p)})$ by (2.15). By obeying the notation $td(\ell^{(p)}, m^{(p)}; \alpha)$, we get

$$TD(td(\ell^{(p)}, m^{(p)}; \alpha(M^{(p)}, \ell^{(p)}))|\ell^{(p)}, m^{(p)}) = 1 - \alpha(M^{(p)}, \ell^{(p)}), \tag{3.28}$$

that is, $td(\ell^{(p)}, m^{(p)}; \alpha(M^{(p)}, \ell^{(p)}))$ is an upper $100\alpha(M^{(p)}, \ell^{(p)})\%$ point of the distribution $TD(t|\ell^{(p)}, m^{(p)})$.

(a) $J^{(p)} \geq 2$

Whenever $td\left(\ell_j^{(p)}, m^{(p)}; \alpha(M^{(p)}, \ell_j^{(p)})\right) < T^{(p)o}(I_j^{(p)o})$ holds for an integer j such that $1 \leq j \leq J^{(p)}$, we reject the hypothesis $\bigwedge_{\boldsymbol{v} \in V^+} H_{\boldsymbol{v}}^{(p)}$.

(b) $J^{(p)} = 1$ $(M^{(p)} = \ell_1^{(p)})$

Whenever $td\left(M^{(p)}, m^{(p)}; \alpha\right) < T^{(p)o}(I_1^{(p)o})$, we reject the hypothesis $\bigwedge_{\boldsymbol{v} \in V^+} H_{\boldsymbol{v}}^{(p)}$.

By using the methods of (a) and (b), when $\bigwedge_{\boldsymbol{v} \in V^+} H_{\boldsymbol{v}}^{(p)}$ is rejected for any V such that $(i, i') \in V \subset \mathcal{U}_{k^{(p)}}$, the null hypothesis $H_{(i,i')}^{(p)}$ is rejected as a multiple comparison test, where V^+ is defined by (3.19).

We do not suppose the condition (C1) of equal sample sizes from now on. The discussions of (3.18)-(3.24) do not depend on the condition (C1).

For $I_j^{(p)o}$ of (3.21) and $j = 1, \ldots, J^{(p)}$, we define $\tilde{\mu}_{s_j^{(p)}+1}^{(p)o}(I_j^{(p)o}), \ldots, \tilde{\mu}_{s_j^{(p)}+\ell_j^{(p)}}^{(p)o}(I_j^{(p)o})$ by $u_{s_j^{(p)}+1}, \ldots, u_{s_j^{(p)}+\ell_j^{(p)}}$ which minimize $\sum_{i \in I_j^{(p)o}} \lambda_{ni}^{(p)} \left(u_i - \bar{X}_{i\cdot}^{(p)}\right)^2$ under simple order restrictions $u_{s_j^{(p)}+1} \leq u_{s_j^{(p)}+2} \leq \ldots \leq u_{s_j^{(p)}+\ell_j^{(p)}}$, i.e.,

$$\sum_{i \in I_j^{(p)o}} \lambda_{ni}^{(p)} \left(\tilde{\mu}_i^{(p)o}(I_j^{(p)o}) - \bar{X}_{i\cdot}^{(p)}\right)^2 = \min_{u_{s_j^{(p)}+1} \leq \ldots \leq u_{s_j^{(p)}+\ell_j^{(p)}}} \sum_{i \in I_j^{(p)o}} \lambda_{ni}^{(p)} \left(u_i - \bar{X}_{i\cdot}^{(p)}\right)^2.$$

Corresponding to (3.2), we get

$$\tilde{\mu}_{s_j^{(p)}+r}^{(p)o}(I_j^{(p)o}) = \max_{s_j^{(p)}+1 \leq a \leq s_j^{(p)}+r} \ \min_{s_j^{(p)}+r \leq b \leq s_j^{(p)}+\ell_j^{(p)}} \frac{\sum_{i=a}^b n_i^{(p)} \bar{X}_{i\cdot}^{(p)}}{\sum_{i=a}^b n_i^{(p)}} \quad (r = 1, \ldots, \ell_j^{(p)}).$$

We put

$$\bar{B}_{(p)}^2(I_j^{(p)o}) := \frac{\sum_{i \in I_j^{(p)o}} n_i^{(p)} \left(\tilde{\mu}_i^{(p)o}(I_j^{(p)o}) - \bar{X}_{\cdot\cdot}^{(p)}(I_j^{(p)o})\right)^2}{V_E^{(p)}}, \tag{3.29}$$

where

$$\bar{X}_{\cdot\cdot}^{(p)}(I_j^{(p)o}) := \frac{\sum_{i \in I_j^{(p)o}} \sum_{t=1}^{n_i^{(p)}} X_{it}^{(p)}}{\sum_{i \in I_j^{(p)o}} n_i^{(p)}}.$$

Let $P(L, \ell_j^{(p)}; \boldsymbol{\lambda}_n^{(p)}(I_j^{(p)o}))$ be the probability that $\tilde{\mu}_{s_j^{(p)}+1}^{(p)o}(I_j^{(p)o}), \ldots, \tilde{\mu}_{s_j^{(p)}+\ell_j^{(p)}}^{(p)o}(I_j^{(p)o})$ takes exactly L distinct values under $H_0^{(p)}$, where
$\boldsymbol{\lambda}_n^{(p)}(I_j^{(p)o}) := (n_{s_j^{(p)}+1}^{(p)}/n^{(p)}, n_{s_j^{(p)}+2}^{(p)}/n^{(p)}, \ldots, n_{s_j^{(p)}+\ell_j^{(p)}}^{(p)}/n^{(p)})$. Then, from (3.7), for $t > 0$, under $H^{(p)o}(I_1^{(p)o}, \ldots, I_{J^{(p)}}^{(p)o})$ of (3.24), we get

$$\begin{aligned} P(\bar{B}_{(p)}^2(I_j^{(p)o}) \geq t) &= P_{(p)0}(\bar{B}_{(p)}^2(I_j^{(p)o}) \geq t) \\ &= \sum_{L=2}^{\ell_j^{(p)}} P(L, \ell_j^{(p)}; \boldsymbol{\lambda}_n^{(p)}(I_j^{(p)o})) P\left((L-1)F_{m^{(p)}}^{L-1} \geq t\right). \end{aligned} \tag{3.30}$$

For a given α such that $0 < \alpha < 0.5$, we put

$$\begin{aligned} &\bar{b}^2(\ell_j^{(p)}, \boldsymbol{\lambda}_n^{(p)}(I_j^{(p)o}), m^{(p)}; \alpha) := \text{a solution of } t \text{ satisfying the equation} \\ &\qquad P_{(p)0}(\bar{B}_{(p)}^2(I_j^{(p)o}) \geq t) = \alpha. \end{aligned} \tag{3.31}$$

We put

$$\bar{\chi}_{(p)}^2(I_j^{(p)o}) := \frac{\sum_{i \in I_j^{(p)o}} n_i^{(p)} \left(\tilde{\mu}_i^{(p)o}(I_j^{(p)o}) - \bar{X}_{\cdot\cdot}^{(p)}(I_j^{(p)o})\right)^2}{\sigma_{(p)}^2}.$$

In order to discuss the asymptotic theory, we add the condition (C3).

$(C3)$ $$\lim_{n^{(p)} \to \infty} (n_i^{(p)}/n^{(p)}) = \lambda_i^{(p)} > 0 \ \ (1 \leq i \leq k^{(p)},\ 1 \leq p \leq q)$$

We define $\breve{\mu}_1^{(p)o}, \ldots, \breve{\mu}_{\ell_j^{(p)}}^{(p)o}$ by

$$\sum_{i=1}^{\ell_j^{(p)}} \lambda_{s_j^{(p)}+i}^{(p)} \left(\breve{\mu}_i^{(p)o} - Z_i\right)^2 = \min_{u_1 \leq \cdots \leq u_{\ell_j^{(p)}}} \sum_{i=1}^{\ell_j^{(p)}} \lambda_{s_j^{(p)}+i}^{(p)} (u_i - Z_i)^2$$

where $Z_1, \ldots, Z_{\ell_j^{(p)}}$ are independent and $Z_i \sim N(0, 1/\lambda_{s_j^{(p)}+i}^{(p)})$ $(i = 1, \ldots, \ell_j^{(p)})$. Let $P(L, \ell_j^{(p)}; \boldsymbol{\lambda}^{(p)}(I_j^{(p)o}))$ be the probability that $\breve{\mu}_1^{(p)o}, \ldots, \breve{\mu}_{\ell_j^{(p)}}^{(p)o}$ takes exactly L distinct values, where
$\boldsymbol{\lambda}^{(p)}(I_j^{(p)o}) := (\lambda_{s_j^{(p)}+1}^{(p)}, \ldots, \lambda_{s_j^{(p)}+\ell_j^{(p)}}^{(p)})$. Then, for $t > 0$, under the condition (C3), we get

$$\lim_{n^{(p)} \to \infty} P_{(p)0}\left(\bar{\chi}_{(p)}^2(I_j^{(p)o}) \geq t\right) = \sum_{L=2}^{\ell_j^{(p)}} P(L, \ell_j^{(p)}; \boldsymbol{\lambda}^{(p)}(I_j^{(p)o})) P\left(\chi_{L-1}^2 \geq t\right). \tag{3.32}$$

Futhermore, for $t > 0$, under the condition (C3),

$$\lim_{n^{(p)} \to \infty} P_{(p)0}(\bar{B}_{(p)}^2(I_j^{(p)o}) \geq t) = \lim_{n^{(p)} \to \infty} P_{(p)0}\left(\bar{\chi}_{(p)}^2(I_j^{(p)o}) \geq t\right) \tag{3.33}$$

holds. For a given α such that $0 < \alpha < 0.5$, we put

$$\begin{aligned} &\bar{c}^2\left(\ell_j^{(p)}, \boldsymbol{\lambda}^{(p)}(I_j^{(p)o}); \alpha\right) := \text{a solution of } t \text{ satisfying the equation} \\ &\qquad \lim_{n^{(p)} \to \infty} P_{(p)0}\left(\bar{\chi}_{(p)}^2(I_j^{(p)o}) \geq t\right) = \alpha. \end{aligned} \tag{3.34}$$

Under (C3), we have

$$\lim_{n^{(p)}\to\infty} \bar{b}^2(\ell_j^{(p)}, \boldsymbol{\lambda}_n^{(p)}(I_j^{(p)o}), m^{(p)}; \alpha) = \bar{c}^2\left(\ell_j^{(p)}, \boldsymbol{\lambda}^{(p)}(I_j^{(p)o}); \alpha\right).$$

Then, we propose the stepwise procedure [3.3].

[3.3] Stepwise procedure based on $\bar{B}^2_{(p)}$ statistics

For $H^{(p)o}(I_1^{(p)o}, \ldots, I_{J^{(p)}}^{(p)o})$ of (3.23) and for $\ell^{(p)} = \ell_1^{(p)}, \ldots, \ell_J^{(p)}$, we define $M^{(p)}$ and $\alpha(M^{(p)}, \ell^{(p)})$ by (3.27) and (2.15) respectively.

(a) $J^{(p)} \geq 2$

Whenever $\bar{b}^2\left(\ell_j^{(p)}, \boldsymbol{\lambda}_n^{(p)}(I_j^{(p)o}), m^{(p)}; \alpha(M^{(p)}, \ell_j^{(p)})\right) < \bar{B}^2_{(p)}(I_j^{(p)o})$ holds for an integer j such that $1 \leq j \leq J$, we reject the hypothesis $\bigwedge_{\boldsymbol{v}\in V^+} H_{\boldsymbol{v}}^{(p)}$.

(b) $J^{(p)} = 1$ $(M^{(p)} = \ell_1^{(p)})$

Whenever $\bar{b}^2\left(\ell_1^{(p)}, \boldsymbol{\lambda}_n^{(p)}(I_1^{(p)o}), m^{(p)}; \alpha\right) < \bar{B}^2_{(p)}(I_1^{(p)o})$, we reject the hypothesis $\bigwedge_{\boldsymbol{v}\in V^+} H_{\boldsymbol{v}}^{(p)}$.

By using the methods of (a) and (b), when $\bigwedge_{\boldsymbol{v}\in V^+} H_{\boldsymbol{v}}^{(p)}$ is rejected for any V such that $(i,i') \in V \subset \mathcal{U}_{k^{(p)}}$, the null hypothesis $H_{(i,i')}^{(p)}$ is rejected as a multiple comparison test, where V^+ is defined by (3.19).

4 Serial gatekeeping procedures

Suppose that the families $\mathcal{H}^{(1)}, \ldots, \mathcal{H}^{(q)}$ of null hypotheses has the order (1.4) of priority. Futhermore suppose that, for some p, simple order restrictions $\mu_1^{(p)} \leq \mu_2^{(p)} \leq \ldots \leq \mu_{k^{(p)}}^{(p)}$ hold. Let us put the set

$$O_q := \{p \mid \mu_1^{(p)} \leq \mu_2^{(p)} \leq \ldots \leq \mu_{k^{(p)}}^{(p)} \text{ is satisfied and } 1 \leq p \leq q\}. \tag{4.1}$$

Then we propose multiple test procedures for all-pairwise comparisons of

$$\left\{\text{the null hypothesis } H_{(i,i')}^{(p)} \text{ vs. the alternative } H_{(i,i')}^{(p)A} \text{ or } H_{(i,i')}^{(p)OA} \,\middle|\, (i,i') \in \mathcal{U}_{k^{(p)}},\ 1 \leq p \leq q\right\}, \tag{4.2}$$

where we choose $H_{(i,i')}^{(p)A}$ as the alternative hypothesis for $p \in O_q^c \cap \{1, \ldots, q\}$ and choose $H_{(i,i')}^{(p)OA}$ for $p \in O_q$. In sections 2.1 and 2.2, we state multiple tests among $\mu_1^{(p)}, \ldots, \mu_{k^{(p)}}^{(p)}$ for fixed p. In this section, we discuss multiple tests among $\mu_1^{(p)}, \ldots, \mu_{k^{(p)}}^{(p)}$ for all p's. Since the serial gatekeeping procedure is a closed testing procedure, we introduce closed testing procedures for $\bigcup_{p=1}^q \mathcal{H}^{(p)}$.

The closure of $\bigcup_{p=1}^q \mathcal{H}^{(p)}$ is given by

$$\overline{\bigcup_{p=1}^q \mathcal{H}^{(p)}} \equiv \left\{ \bigwedge_{g=1}^h \left(\bigwedge_{\boldsymbol{v}\in V^{(p_g)}} H_{\boldsymbol{v}}^{(p_g)} \right) \,\middle|\, \begin{array}{l} \text{there exist integer } h \text{ and integers } p_1, \ldots, p_h \\ \text{such that } 1 \leq h \leq q,\ 1 \leq p_1 < \ldots < p_h \leq q, \text{ and} \\ \emptyset \subsetneq V^{(p_g)} \subset \mathcal{U}_{k^{(p_g)}}\ (1 \leq g \leq h) \text{ hold} \end{array} \right\}$$

Then, we get

$$\bigwedge_{g=1}^h \left(\bigwedge_{\boldsymbol{v}\in V^{(p_g)}} H_{\boldsymbol{v}}^{(p_g)} \right) : \text{for any } g \text{ such that } 1 \leq g \leq h \text{ and for any } (i,i') \in V^{(p_g)},$$

$$\mu_i^{(p_g)} = \mu_{i'}^{(p_g)} \text{ holds.}$$

Then, from (2.13) and (3.23), for any nonempty $V^{(p_g)} \subset \mathcal{U}_{k^{(p_g)}}$, there exist an integer $J^{(p_g)}$ and disjoint sets $I_1^{(p_g)*}, \ldots, I_{J^{(p_g)}}^{(p_g)*}$ such that

$$\bigwedge_{\boldsymbol{v} \in V^{(p_g)}} H_{\boldsymbol{v}}^{(p_g)} = H^{(p_g)*}(I_1^{(p_g)*}, \ldots, I_{J^{(p_g)}}^{(p_g)*}) \tag{4.3}$$

and $\#(I_j^{(p_g)*}) \geq 2$ $(j = 1, \ldots, J^{(p_g)})$,
where $H^{(p_g)*}(I_1^{(p_g)*}, \ldots, I_{J^{(p_g)}}^{(p_g)*})$ stands for

$$H^{(p_g)*}(I_1^{(p_g)*}, \ldots, I_{J^{(p_g)}}^{(p_g)*}) := \begin{cases} H^{(p_g)}(I_1^{(p_g)}, \ldots, I_{J^{(p_g)}}^{(p_g)}) & (p_g \in O_q^c \cap \{1, \ldots, q\}) \\ H^{(p_g)o}(I_1^{(p_g)o}, \ldots, I_{J^{(p_g)}}^{(p_g)o}) & (p_g \in O_q), \end{cases}$$

and, for $j = 1, \ldots, J_{p_g}$, $I_j^{(p_g)*}$ stands for

$$I_j^{(p_g)*} := \begin{cases} I_j^{(p_g)} & (p_g \in O_q^c \cap \{1, \ldots, q\}) \\ I_j^{(p_g)o} & (p_g \in O_q). \end{cases}$$

Hence we get

$$\bigwedge_{g=1}^{h} \left(\bigwedge_{\boldsymbol{v} \in V^{(p_g)}} H_{\boldsymbol{v}}^{(p_g)} \right) = \bigwedge_{g=1}^{h} H^{(p_g)*}(I_1^{(p_g)*}, \ldots, I_{J^{(p_g)}}^{(p_g)*}), \tag{4.4}$$

where $1 \leq p_1 < \ldots < p_h \leq q$.

[4.1] Hybrid serial gatekeeping procedure

For integer p such $1 \leq p \leq q$, in ascending order, perform multiple comparison test of level α based on stepwise procedure [2.4] or [3.3], where we choose [2.4] for $p \in O_q^c \cap \{1, \ldots, q\}$ and choose [3.3] for $p \in O_q$. Then we reject null hypotheses in $\bigcup_{p=1}^{q} \mathcal{H}^{(p)}$ obeying the following (b1)-(b3).

(b1) When there is a null hypothesis in $\mathcal{H}^{(1)}$ that is not rejected by stepwise procedure [2.4] or [3.3], only the null hypothesis rejected in $\mathcal{H}^{(1)}$ is rejected as a multiple comparison test.

(b2) When there exists an integer q_0 satisfying $q_0 < q$ that, for any p such that $1 \leq p \leq q_0$, all the null hypotheses in $\mathcal{H}^{(p)}$ are rejected by stepwise procedure [2.4] or [3.3] and there is a null hypothesis in $\mathcal{H}^{(q_0+1)}$ that is not rejected, all the null hypotheses in $\bigcup_{p=1}^{q_0} \mathcal{H}^{(p)}$ are rejected as a multiple comparison test and only the null hypothesis rejected in $\mathcal{H}^{(q_0+1)}$ is rejected.

(b3) When, for any p satisfying $1 \leq p \leq q$, all the null hypotheses in $\mathcal{H}^{(p)}$ are rejected by stepwise procedure [2.4] or [3.3], all the null hypotheses in $\bigcup_{p=1}^{q} \mathcal{H}^{(p)}$ are rejected as a multiple comparison test.

Theorem 2.1 The test procedure [4.1] is a multiple comparison test of level α.
proof. It is enough to show that [4.1] is a closed testing procedure of level α.

The test of level α for $H^{(p_1)*}(I_1^{(p_1)*}, \ldots, I_{J^{(p_1)}}^{(p_1)*})$ is executed. Futhermore, when $h \geq 2$ in (4.4), the test of level 0 for $H^{(p_g)*}(I_1^{(p_g)*}, \ldots, I_{J^{(p_g)}}^{(p_g)*})$ is executed for any g such that $2 \leq g \leq h$. As these results, we are able to decide whether or not the hypothesis $\bigwedge_{g=1}^{h} \left(\bigwedge_{\boldsymbol{v} \in V^{(p_g)}} H_{\boldsymbol{v}}^{(p_g)} \right)$ is rejected. Let us refer to this method as A-procedure.

(1). The case of (b1)

Suppose that $H_{(i_1,i_1')}^{(1)}$ is rejected by using the procedure [4.1]. Then, from A-procedure, $p_1 = 1$ holds and there exists j such that all the null hypotheses (4.4) satisfying $i_1,\ i_1' \in I_j^{(1)*}$ and $1 \leq j \leq J^{(1)}$ are rejected. Suppose that $H_{(i_2,i_2')}^{(1)}$ is not rejected by using the procedure [4.1]. Then, from A-procedure, $p_1 = 1$ holds and there exists j' such that some null hypothesis (4.3) satisfying $i_2,\ i_2' \in I_{j'}^{(1)*}$ and $1 \leq j' \leq J^{(1)}$ is not rejected. Let $H^{(1)*}(I_{01}^{(1)*}, \ldots, I_{0J^{(1)}}^{(1)*})$

be the null hypothesis that is not rejected. Then, from A-procedure, $H^{(1)}_{(i_2,i'_2)}$ is not rejected. From A-procedure, $H^{(1)*}(I^{(1)*}_{01}, \ldots, I^{(1)*}_{0J^{(1)}}) \bigwedge \left(\bigwedge_{g=2}^{h} H^{(p_g)*}(I^{(p_g)*}_{1}, \ldots, I^{(p_g)*}_{J^{(p_g)}})\right)$ is not rejected. Therefor all the null hypotheses in $\bigcup_{p=2}^{q} \mathcal{H}^{(p)}$ are not rejected.

(2). The case of (b2)

Since any null hypothesis (4.4) satisfying $1 \le p_1 \le q_0$ is rejected, for any p such that $1 \le p \le q_0$, all the null hypotheses in $\mathcal{H}^{(p)}$ are rejected. Suppose that $H^{(q_0+1)}_{(i_1,i'_1)}$ in $\mathcal{H}^{(q_0+1)}$ is rejected by using the procedure [4.1]. Then, from A-procedure, $p_1 = q_0 + 1$ holds and there exists j such that all the null hypotheses (4.4) satisfying $i_1,\ i'_1 \in I^{(q_0+1)*}_{j}$ and $1 \le j \le J^{(q_0)}$ are rejected. Suppose that $H^{(q_0+1)}_{(i_2,i'_2)}$ in $\mathcal{H}^{(q_0+1)}$ is not rejected by using the procedure [4.1]. Then, from A-procedure, $p_1 = q_0+1$ holds and there exists j' such that some null hypothesis (4.3) satisfying $i_2,\ i'_2 \in I^{(q_0+1)*}_{j'}$ and $1 \le j' \le J^{(q_0+1)}$ is not rejected. Let $H^{(q_0+1)*}(I^{(q_0+1)*}_{01}, \ldots, I^{(q_0+1)*}_{0J^{(q_0+1)}})$ be the null hypothesis that is not rejected. Then, from A-procedure, $H^{(q_0+1)}_{(i_2,i'_2)}$ is not rejected. From A-procedure, $H^{(q_0+1)*}(I^{(q_0+1)*}_{01}, \ldots, I^{(q_0+1)*}_{0J^{(q_0+1)}}) \bigwedge \left(\bigwedge_{g=2}^{h} H^{(p_g)*}(I^{(p_g)*}_{1}, \ldots, I^{(p_g)*}_{J^{(p_g)}})\right)$ $(q_0 + 1 < p_2)$ is not rejected. Therefor all the null hypotheses in $\bigcup_{p=q_0+2}^{q} \mathcal{H}^{(p)}$ are not rejected.

(3) The case of (b3)

From A-procedure, it is self-evident that all the null hypotheses in $\left\{H^{(p)}_{(i,i')} \mid (i,i') \in \mathcal{U}_{k^{(p)}},\ 1 \le p \le q\right\}$ are rejected.

From (1)-(3), the closed testing procedure of level α for (4.2) based on A-procedure is equivalent to the procedure [4.1]. □

Even if we replace [2.4] with [2.1], [2.2] or [2.3] in the hybrid serial gatekeeping procedure [4.1], Theorem 2.1 still holds. Futhermore even if we replace [3.3] with [3.1] or [3.2] in [4.1], Theorem 2.1 still holds under the condition (C1) of equal sample sizes.

5 Application to multivariate multi-sample models

Let $\{\boldsymbol{X}_{ij} = (X^{(1)}_{ij}, \ldots, X^{(q)}_{ij})' \mid j = 1, \ldots, n_i,\ i = 1, \ldots, k\}$ be a set of independent vector-values random variables. Furhtermore suppose that the mean vector and variance-covariance matrix of $\boldsymbol{X}_{ij}$ are given by

$$E(\boldsymbol{X}_{ij}) = \boldsymbol{\mu}_i = (\mu^{(1)}_i, \ldots, \mu^{(q)}_i)' \ \text{ and } \ V(\boldsymbol{X}_{ij}) = \Sigma = (\sigma_{pp'})_{p,p'=1,\ldots,q},$$

respectively, and that the cumulative distribution of $\boldsymbol{X}_{ij}$ is denoted by $\boldsymbol{F}(\boldsymbol{x} - \boldsymbol{\mu}_i)$. Then this model becomes a q variate k sample model. Futhermore the model is limited to

$$(C4) \qquad k^{(p)} = k\ (p = 1, \ldots, q) \text{ and } n^{(p)}_i = n_i\ (i = 1, \ldots, k,\ p = 1, \ldots, q)$$

in the model of (1). $\sigma_{pp} = \sigma^2_{(p)}$ holds for $p = 1, \ldots, q$. The notations of (4), (6) and (7) are simplified to

$$\mathcal{U}_k = \{(i,i') \mid 1 \le i < i' \le k\},\ V^{(p)}_E = \frac{1}{m} \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X^{(p)}_{ij} - \bar{X}^{(p)}_{i\cdot})^2$$

$$m = n - k,\ \text{and } n = \sum_{i=1}^{k} n_i.$$

In the q variate k sample model, gatekeeping procedures [4.1] gives multiple comparison tests of level α for all-pairwise comparisons of

$$\Big\{\text{the null hypothesis } H^{(p)}_{(i,i')} \text{ vs. the alternative } H^{(p)A}_{(i,i')} \text{ or } H^{(p)OA}_{(i,i')} \Big|\ (i,i') \in \mathcal{U}_k,\ 1 \le p \le q\Big\}$$

under the limits of (C4).

National cancer mortality data is posted on Cancer Registry and Statistics. Cancer Information Service, National Cancer Center, Japan (Vital Statistics of Japan). In the data, we use mortality rate by age group (to population 100,000), site, sex, year of death and we show the results for the application of the proposed procedures. We divide four groups of 15 to 19 years old, 20 to 24 years old, 25 to 29 years old, and 30 to 34 years old in male. Table 1 is mortality rate data of all parts consisting of observations from 2014 to 2019 in the four groups. Futhermore Table 2 is mortality rate data of leukemia. When Table 1 and Table 2 are combined, it becomes two-dimensional 4 samples, that is, $q = 2$ and $k = 4$. We find $n_i^{(p)} = 6$ $(p = 1, 2; i = 1, \dots, 4)$. We wish to confirm differences among mortality rates of all parts at first. Next we confirm differences among mortality rates of leukemia, that is, $\mathcal{H}^{(1)} \succ \mathcal{H}^{(2)}$. From Table 1, estimators of $\mu_i^{(1)}$ $(i = 1, 2, 3, 4)$

$$\bar{X}_{1\cdot}^{(1)} = 2.57 < \bar{X}_{2\cdot}^{(1)} = 3.35 < \bar{X}_{3\cdot}^{(1)} = 4.57 < \bar{X}_{4\cdot}^{(1)} = 7.17.$$

Futhermore we get values of statistics

$$T_{21}^{(1)} = 2.970,\ T_{31}^{(1)} = 7.582,\ T_{41}^{(1)} = 17.438,$$
$$T_{32}^{(1)} = 4.612,\ T_{42}^{(1)} = 14.469,\ T_{43}^{(1)} = 9.856.$$

From Table 16 of Shiraishi and Sugiura (2018), we have $td(4, 20; 0.05) = 2.508$. Since $T_{i'i}^{(1)} > td(4, 20; 0.05)$ holds for any $(i, i') \in \mathcal{U}_4$, by using multiple comparison tests of level 0.05 for all-pairwise comparisons of $\{$the null hypothesis $H_{(i,i')}^{(1)}$ vs. the alternative $H_{(i,i')}^{(1)OA} \mid (i, i') \in \mathcal{U}_4\}$, all the null hypotheses in $\{H_{(i,i')}^{(1)} \mid (i, i') \in \mathcal{U}_4\}$ are rejected. From Table 2, we get the values of statistics

$$T_{21}^{(2)} = 2.569,\ T_{31}^{(2)} = 2.202,\ T_{41}^{(2)} = 3.853,$$
$$T_{32}^{(2)} = -0.367,\ T_{42}^{(2)} = 1.284,\ T_{43}^{(2)} = 1.651.$$

We give the values of $ta\,(\ell, m\alpha(M, \ell))$ for the stepwise procedure [2.2] with $\alpha = 0.05$ and $m = 20$. By using the stepwise procedure [2.2] of level 0.05 for $\{$the null hypothesis $H_{(i,i')}^{(2)}$ vs. the alternative $H_{(i,i')}^{(2)A} \mid (i, i') \in \mathcal{U}_4\}$, the null hypotheses $H_{(1,2)}^{(2)}$ and $H_{(1,4)}^{(2)}$ are rejected. The other null hypotheses are retained. Hence we consider the hybrid serial gatekeeping procedure replaced [2.4] and [3.3] with [2.2] and [3.1] respectively in the hybrid serial gatekeeping procedure [4.1]. By using this hybrid serial gatekeeping procedure of level 0.05 for $\{$the null hypothesis $H_{(i,i')}^{(1)}$ vs. the alternative $H_{(i,i')}^{(1)OA} \mid (i, i') \in \mathcal{U}_4\} \bigcup \{$the null hypothesis $H_{(i,i')}^{(2)}$ vs. the alternative $H_{(i,i')}^{(2)A} \mid (i, i') \in \mathcal{U}_4\}$, we find

$$\mu_1^{(1)} < \mu_2^{(1)} < \mu_3^{(1)} < \mu_4^{(1)},\ \ \mu_1^{(2)} \neq \mu_2^{(2)},\ \mu_1^{(2)} \neq \mu_4^{(2)}.$$

6 Discussion

When the families of null hypotheses $\mathcal{F}_p = \{H_{pj} \mid j = 1, \dots, m_p\}$ $(p = 1, \dots, q)$ has the order of priority, $\mathcal{F}_1 \succ \dots \succ \mathcal{F}_q$, Maurer et al. (1995) proposed a multiple comparison test using a closed test procedure called the serial gatekeeping method. The serial gatekeeping procedures are based on Bonferroni tests and the test procedure of Holm (1979). In the serial gatekeeping method, the tests are performed in the order of the null hypothesis family $\mathcal{F}_1, \dots, \mathcal{F}_q$. If a null hypothesis in the $\mathcal{F}_p$ $(1 \leq p < q)$ is not rejected, test procedures for the subsequent null hypothesis family $\mathcal{F}_{p+1}, \dots, \mathcal{F}_q$ are not performed. As a closed test procedure that covers this shortcoming, Dmitrienko et al. (2003) proposed a multiple comparison test called the parallel gatekeeping procedure. In the parallel gatekeeping procedure, Bonferroni's method is used. Since the parallel gatekeeping procedure is not simple, it is difficult to propose the parallel gatekeeping procedure based on [2.1]-[3.3] as a multiple comparison test of level α. It is simple to use the hybrid serial gatekeeping procedure replaced [2.4] and [3.3] with [2.1] and [3.1] respectively in [4.1]. Under

Table 1: Motality rate data of all parts in male

age group	2014	2015	2016	2017	2018	2019
15 − 19	3.2	2.8	2.6	2.3	2.6	1.9
20 − 24	3.1	3.7	3.1	3.7	3.3	3.2
25 − 29	4.5	4.8	5.0	4.4	4.3	4.4
30 − 34	8.3	7.1	7.3	7.4	6.6	6.3

Table 2: Motality rate data of leukemia in male

age group	2014	2015	2016	2017	2018	2019
15 − 19	0.8	0.8	0.8	0.5	0.7	0.4
20 − 24	0.9	1.0	0.6	1.1	1.0	0.8
25 − 29	1.0	0.9	0.9	0.9	0.6	0.9
30 − 34	1.2	1.1	1.1	0.9	0.9	0.9

simple order restrictions of (3.1), Shiraishi et al. (2019) investigate the all-pairs power proposed by Ramsey (1978). As the results, the order of the power is following:

$$[3.3] \geq [3.2] > [2.4] \geq [2.2] > [2.3] > [3.1] > [2.1]. \tag{6.1}$$

In the all-pairs power of specified alternatives, [3.3] is a little superior to [3.2] and [2.4] is a little superior to [2.2]. Shiraishi (2014) proves that the rejection region of stepwise procedure [3.2] includes that of single step procedure [3.1].

As a distribution-free method, we the hybrid gatekeeping procedure of [3.7] based on rank statistics. By using the asymptotic theory of Hájeck et al. (1999), under contigious local alternatives, we derive asymptotic power of [3.1]-[3.6]. Then corresponding to (6.1), under simple order restrictions of (3.1), the order of the asymptotic all-pairs power is following:

$$[3.6] \geq [3.5] > [3.3] \geq [3.2] > [3.4] > [3.1].$$

We suppose the reverse order restrictions

$$\mu_1^{(p)} \geq \mu_2^{(p)} \geq \ldots \geq \mu_{k^{(p)}}^{(p)} \tag{6.2}$$

Then we put $Y_{ij}^{(p)} = -X_{ij}^{(p)}$ $(j = 1, \ldots n_i^{(p)};\ i = 1, \ldots, k^{(p)})$. $(Y_{i1}^{(p)}, \ldots, Y_{in_i^{(p)}}^{(p)})$ is a random sample of size $n_i^{(p)}$ from the i-th normal population with unknown mean $\mu_i'^{(p)} = -\mu_i^{(p)}$ $(i = 1, \ldots, k^{(p)})$ and unknown variance $\sigma_{(p)}^2$. (6.2) is equivalent to the simple order restrictions of $\mu_i'^{(p)}$'s: $\mu_1'^{(p)} \leq \mu_2'^{(p)} \leq \ldots \leq \mu_{k^{(p)}}'^{(p)}$. By replacing $X_{ij}^{(p)}$ with $Y_{ij}^{(p)}$ in all statistics of sections 2 and 3, we can discuss the multiple comparison procedures under the restrictions (6.2).

Table 3: Critical values $ta\,(\ell, m\alpha(M, \ell))$ for the stepwise procedure [2.2] with $\alpha = 0.05$ and $m = 20$

$M \setminus \ell$	2	3	4
4	2.417	$\diamond$	2.799
3	$\diamond$	2.530	
2	2.086		

Aacknowledgements

This research was supported in part by a Grant-in-Aid for Co-operative Research (C) 18K11204 from the Japanese Ministry of Education. This research was also supported in part by Nanzan University Pache Research Subsidy I-A-2 for 2021 academic year.

References

[1] Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). *Statistical Inference under Order Restrictions.* Wiley.

[2] Cancer Registry and Statistics. Cancer Information Service, National Cancer Center, Japan (Vital Statistics of Japan), https://ganjoho.jp/reg_stat/statistics/stat/summary.html. (2020).

[3] Dmitrienko, A., Offen, W. and Westfall, P. H. (2003). Gatekeeping strategies for clinical trials that do not require all primary effects to be significant. *Statistics in Medicine*, **22**, 2387–2400.

[4] Einot, I. and Gabriel, K. R. (1975). A study of the powers of several methods of multiple comparisons. *Journal of the American Statistical Association*, **70**, 574–583.

[5] Enderton, H. B. (2001). *A Mathematical Introduction to Logic: Second Edition.* Academic Press.

[6] Hayter, A. J. (1984). A proof of the conjecture that the Tukey-Kramer multiple comparisons procedure is conservative. *Annals of Statistics,* **12**, 61–75.

[7] Hayter, A. J. (1990). A one-sided Studentized range test for testing against a simple ordered alternative. *Journal of the American Statistical Association*, **85**, 778–785.

[8] Kramer, C. Y. (1956). Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics*, **8**, 75–86.

[9] Hsu, J. C. (1996). *Multiple Comparisons-Theory and Methods.* Chapman & Hall.

[10] Maurer, W., Hothorn, L. and Lehmacher, W. (1995). Multiple comparisons in drug clinical trials and preclinical assays: a priori ordered hypotheses. *In Biometrie in der ChemischPharmazeutischen Industrie*, **6**, 3-18.

[11] Ramsey, P. H. (1978). Power differences between pairwise multiple comparisons. *J. Amer. Statist. Assoc.*, **73**, 479–485.

[12] Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference.* Wiley.

[13] Ryan, T. H. (1960). Significance tests for multiple comparison of proportions, variances, and other statistics. *Psychological Bulletin*, **57**, 318–328.

[14] Shiraishi, T. (2006). The upper bound for the distribution of Tukey-Kramer's statistic. *Bulletin of the Computational Statistics*, **19**, 77-87.

[15] Shiraishi, T. (2011). Closed testing procedures for pairwise comparisons in multi-sample models. *Biometric Society of Japan* **32**, 33–47, (in Japanese).

[16] Shiraishi, T. (2014). Closed testing procedures in multi-sample models under a simple ordered restriction. *Journal of Japan Statistical Society, Japanese Issue*, **43**, 215–245.

[17] Shiraishi, T. and Matsuda, S. (2016). Closed testing procedures based on Bartholomew ' s tests for the differences among mean responses under a simple ordered restriction. *Journal of Japan Statistical Society, Japanese Issue,* **45**, 247–271.

[18] Shiraishi, T. and Sugiura, H. (2015). The Upper $100\alpha^{\star}$th percentiles of the distributions used in multiple comparison procedures under a simple order restriction. *Journal Japan Statistical Society, Japanese Issue*, **44**, 271–314.

[19] Shiraishi, T. and Sugiura, H. (2018). *Theory of Multiple Comparison Procedures and its Computation.* Kyoritsu-Shuppan Co., Ltd. (in Japanese).

[20] Shiraishi, T., Sugiura, H. and Matsuda, S. (2019). *Pairwise Multiple Comparisons: Theory and Computation.* SpringerBriefs in Statistics, Springer.

[21] Tukey, J. W. (1953). *The Problem of Multiple Comparisons. The Collected Works of John W. Tukey (1994),* **Vol. VIII** *Multiple Comparisons*, Chapman & Hall.

[22] Welsch, R. E. (1977). Stepwise Multiple Comparison Procedures. *Journal of the American Statistical Association*, **72**, 566–575.

[23] Westfall, P. H. and Krishen, A. (2001). Optimally weighted, fixed sequence and gatekeeper multiple testing procedures. *Journal of Statistical Planning and Inference*, **99**, 25–41.