NOTE ON CLOSED TESTING PROCEDURE AND SEQUENTIALLY REJECTIVE STEP DOWN PROCEDURE

Imada, Tsunehisa
Department of Management, Tokai University

https://doi.org/10.5109/2232322

出版情報：Bulletin of informatics and cybernetics. 49, pp.35-51, 2017-12. Research Association of Statistical Sciences

NOTE ON CLOSED TESTING PROCEDURE AND SEQUENTIALLY REJECTIVE STEP DOWN PROCEDURE

by

Tsunehisa Imada

Reprinted from the Bulletin of Informatics and Cybernetics
Research Association of Statistical Sciences, Vol.49

FUKUOKA, JAPAN
2017

Bulletin of Informatics and Cybernetics, Vol. 49, 2017

NOTE ON CLOSED TESTING PROCEDURE AND SEQUENTIALLY REJECTIVE STEP DOWN PROCEDURE

By

Tsunehisa IMADA*

Abstract

Among various types of stepwise multiple comparison procedures for normal means we focus on the closed testing procedure and the sequentially rejective step down procedure and discuss the relation between them. First, we consider the multiple comparison with a control. Specifically, we indicate that the power of the sequentially rejective step down procedure is not higher than that of the closed testing procedure and two procedures are equivalent when we use same critical values for them. Next, we consider the all-pairwise multiple comparison. Ryan-Einot-Gabriel-Welsch's procedure using Tukey-Welsh's allocation of the significance level is the well known closed testing procedure. When we test an intersection of mutually disjoint plural hypotheses by it, we should test each hypothesis allocating an specified significance level to it. It is accompanied with computational complications when the number of populations is large. Here, we propose a method of testing the intersection of mutually disjoint plural hypotheses at a time in the closed testing procedure. Next, among several types of sequentially rejective step down procedures for the all-pairwise multiple comparison we focus on Holland-Copenhaver's procedure and indicate that the power of Holland-Copenhaver's procedure is not higher than that of the proposed closed testing procedure specifying the total number of populations. We give simulation results regarding the power of the test intended to compare the procedures.

Key Words and Phrases: All-pairwise multiple comparison, Holland-Copenhaver's procedure, Multiple comparison with a control, Power of the test, Ryan-Einot-Gabriel-Welsch's procedure.

1. Introduction

There are independent normal random variables X_1, X_2, …, X_K. Assume X_k is distributed according to normal $N(\mu_k, \sigma^2)$ for $k = 1, 2, \ldots, K$. Here, the common σ^2 is unknown. We consider the multiple comparison for $\mu_1, \mu_2, \ldots, \mu_K$ using a sample $X_{k1}, X_{k2}, \ldots, X_{kn_k}$ from $N(\mu_k, \sigma^2)$ for $k = 1, 2, \ldots, K$.

First, we consider the multiple comparison with a control intended to compare μ_1 with $\mu_2, \mu_3, \ldots, \mu_K$ simultaneously. Dunnett (1955) proposed the single step procedure. Dunnett and Tamhane (1991) proposed the sequentially rejective step down procedure intended to obtain higher power. On the other hand, it is possible to apply the closed testing procedure (cf. Marcus *et al.* (1976)) to the multiple comparison with a control

* Department of Management, Tokai University, 9–1–1 Toroku Higashi-ku Kumamoto, Japan 862-8652 Japan. tel +81–96–386–2731, timada@ktmail.tokai-u.jp

for normal means. We indicate that the power of the sequentially rejective step down procedure is not higher than that of the closed testing procedure and two procedures are equivalent when we use same critical values for them. We give simulation results regarding the power of the test intended to compare two stepwise procedures.

Next, we consider the all-pairwise multiple comparison for $\mu_1, \mu_2, \ldots, \mu_K$. Tukey (1953) proposed the single step procedure. Ryan-Einot-Gabriel-Welsch's procedure (cf. Ryan (1960), Einot and Gabriel (1975) and Welsch (1977)) is the closed testing procedure using Tukey-Welsh's allocation of the significance level for testing the intersection of mutually disjoint plural hypotheses. Here, we construct another type of closed testing procedure which enables us to test the intersection of mutually disjoint plural hypotheses at a time. On the other hand, Holm (1979) proposed a simple sequentially rejective step down procedure. Shaffer (1986) and Holland-Copenhaver (1987) modified Holm's procedure intended to obtain higher power. Here, focusing on Holland-Copenhaver's procedure, we indicate that the power of Holland-Copenhaver's procedure is not higher than that of our proposed closed testing procedure specifying the total number of populations. We give simulation results regarding the power of the test intended to compare two types of closed testing procedures and Holland-Copenhaver's procedure.

In Sections 2 we discuss the multiple comparison with a control. In Sections 3 we discuss the all-pairwise multiple comparison. In Section 4 we give concluding remarks.

2. Multiple comparison with a control

2.1. Single step procedure

First, we discuss the single step procedure proposed by Dunnett (1955). Intended to compare μ_1 and μ_k $(k > 1)$ we set up a null hypothesis and its alternative hypothesis as

$$H_{1k} : \mu_1 = \mu_k \quad \text{vs.} \quad H_{1k}^A : \mu_1 \neq \mu_k \tag{1}$$

or

$$H_{1k} : \mu_1 = \mu_k \quad \text{vs.} \quad H_{1k}^A : \mu_1 > \mu_k. \tag{2}$$

We consider the simultaneous test of $H_{12}, H_{13}, \ldots, H_{1K}$ based on the single step procedure. We focus on (1), because the following discussion is similar for (2). We use the statistic

$$S_{1k} = \sqrt{\frac{n_1 n_k}{n_1 + n_k}}(\bar{X}_1 - \bar{X}_k)s^{-1}$$

for testing H_{1k}. Here

$$\bar{X}_k = \frac{1}{n_k}\sum_{i=1}^{n_k} X_{ki} \quad (k = 1, 2, \ldots, K), \quad s = \sqrt{\frac{1}{N-K}\sum_{k=1}^{K}\sum_{i=1}^{n_k}(X_{ki} - \bar{X}_k)^2}$$

where $N = \sum_{k=1}^{K} n_k$. If $|S_{1k}| > c$ for a specified critical value c, H_{1k} is rejected. Otherwise, it is retained. We determine c so that

$$P(\max_{k>1} |S_{1k}| > c) = \alpha$$

for a specified significance level α when $H_{12}, H_{13}, \ldots, H_{1K}$ are true. Then

$$P(\max_{k>1} |S_{1k}| > c)$$

$$= 1 - P(|S_{12}| < c, |S_{13}| < c, \ldots, |S_{1K}| < c)$$

$$= 1 - \int_0^{\infty} \left[\int_{-\infty}^{\infty} \prod_{k=2}^{K} \left\{ \Phi\left(\frac{\sqrt{\lambda_{1k}} z + c s_0}{\sqrt{1-\lambda_{1k}}} \right) - \Phi\left(\frac{\sqrt{\lambda_{1k}} z - c s_0}{\sqrt{1-\lambda_{1k}}} \right) \right\} \phi(z) dz \right] g(s_0) ds_0. \quad (3)$$

Here

$$\lambda_{1k} = \frac{n_k}{n_1 + n_k} \quad (k = 2, 3, \ldots, K),$$

$\Phi(\cdot)$ is the cumulative distribution function of $N(0,1)$, $\phi(\cdot)$ is the probability density function of $N(0,1)$ and $g(s_0)$ is the probability density function of $s_0 = s/\sigma$ given by

$$g(s_0) = \frac{\psi^{\psi/2}}{2^{(\psi-2)/2}\Gamma[\psi/2]} s_0^{\psi-1} \exp\left[-\frac{\psi s_0^2}{2} \right]$$

where $\psi = N - K$.

2.2. Closed testing procedure

Let $I = \{2, 3, \ldots, K\}$. Let I_q be an arbitrary subset of I. $\sharp(I_q)$ denotes the number of elements of I_q. Defining the hypothesis H_{1I_q} as

$$H_{1I_q} : \mu_1 = \mu_i \quad \text{for all } i \in I_q,$$

we obtain

$$H_{1I_q} = \cap_{i \in I_q} H_{1i}.$$

Let F be the family consisting of all H_{1I_q}s. F is closed. Specifically, for two hypotheses chosen from F arbitrarily, their intersection is also included in F.

We construct the stepwise multiple comparison procedure for F applying the closed testing procedure. For testing each H_{1I_q} in F we use the statistic

$$S_{1I_q} = \max_{i \in I_q} |S_{1i}|.$$

Assuming H_{1I_q} is true, we determine c_{I_q} so that $P(S_{1I_q} > c_{I_q}) = \alpha$. Then

$$P(S_{1I_q} > c_{I_q}) = 1 - \int_0^{\infty} \left[\int_{-\infty}^{\infty} \prod_{i \in I_q} \left\{ \Phi\left(\frac{\sqrt{\lambda_{1i}} z + c_{I_q} s_0}{\sqrt{1-\lambda_{1i}}} \right) - \Phi\left(\frac{\sqrt{\lambda_{1i}} z - c_{I_q} s_0}{\sqrt{1-\lambda_{1i}}} \right) \right\} \phi(z) dz \right] g(s_0) ds_0.$$

We test the hypotheses in F hierarchically as follows.

Step 1.

Case 1. If $S_{1I} \leq c_I$, we retain all hypotheses in F and stop the test.

Case 2. If $S_{1I} > c_I$, we reject H_{1I} and go to the next step.

Step 2.

We test all H_{1I_q}s satisfying $\sharp(I_q) = K - 2$.

Case 1. If $S_{1I_q} \leq c_{I_q}$, we retain H_{1I_q} and all hypotheses induced by H_{1I_q}.

Case 2. If $S_{1I_q} > c_{I_q}$, we reject H_{1I_q}.

Step 3.

If all H_{1I_q}s satisfying $\sharp(I_q) = K - 3$ are retained at Step 2, we stop the test. Otherwise, we test all H_{1I_q}s satisfying $\sharp(I_q) = K - 3$ which are not retained at Step 2.

Case 1. If $S_{1I_q} \leq c_{I_q}$, we retain H_{1I_q} and all hypotheses induced by H_{1I_q}.

Case 2. If $S_{1I_q} > c_{I_q}$, we reject H_{1I_q}.

We repeat similar judgments till up to Step $K - 1$.

2.3. Sequentially rejective step down procedure

The sequentially rejective step down procedure consists of $K-1$ steps of tests. Assuming all $H_{12}, H_{13}, \ldots, H_{1K}$ are true, we determine c_m $(m = 1, 2, \ldots, K-1)$ as the minimum c satisfying

$$P(\max_{k=l_1,l_2,\ldots,l_m} |S_{1k}| > c) \leq \alpha$$

for $l_1, l_2, \ldots, l_m$ chosen from $2, 3, \ldots, K$ arbitrarily. Here

$$P(\max_{k=l_1,l_2,\ldots,l_m} |S_{1k}| > c)$$

$$= 1 - \int_0^\infty \left[\int_{-\infty}^\infty \prod_{j=1}^m \left\{ \Phi\left(\frac{\sqrt{\lambda_{1l_j}} z + cs_0}{\sqrt{1-\lambda_{1l_j}}} \right) - \Phi\left(\frac{\sqrt{\lambda_{1l_j}} z - cs_0}{\sqrt{1-\lambda_{1l_j}}} \right) \right\} \phi(z) dz \right] g(s_0) ds_0.$$

If $n_2 = n_3 = \cdots = n_K$, $P(\max_{k=l_1,l_2,\ldots,l_m} |S_{1k}| > c)$ does not depend on the choice of $l_1, l_2, \ldots, l_m$ from $2, 3, \ldots, K$. In this case c_m is determined by

$$P(\max_{k=2,3,\ldots,m+1} |S_{1k}| > c_m) = \alpha.$$

Arranging $|S_{12}|, |S_{13}|, \ldots, |S_{1K}|$ in order of a size of value, assume

$$|S_{(1)}| \leq |S_{(2)}| \leq \cdots \leq |S_{(K-1)}|.$$

$H_{(1)}, H_{(2)}, \ldots, H_{(K-1)}$ denote hypotheses corresponding to $S_{(1)}, S_{(2)}, \ldots, S_{(K-1)}$. Then, we test $H_{(1)}, H_{(2)}, \ldots, H_{(K-1)}$ sequentially as follows.

Step 1.

Case 1. If $|S_{(K-1)}| \leq c_{K-1}$, we retain $H_{(1)}, H_{(2)}, \ldots, H_{(K-1)}$ and stop the test.

Case 2. If $|S_{(K-1)}| > c_{K-1}$, we reject $H_{(K-1)}$ and go to the next step.

Step 2.

Case 1. If $|S_{(K-2)}| \leq c_{K-2}$, we retain $H_{(1)}, H_{(2)}, \ldots, H_{(K-2)}$ and stop the test.

Case 2. If $|S_{(K-2)}| > c_{K-2}$, we reject $H_{(K-2)}$ and go to the next step.

We repeat similar judgments till up to Step $K-1$.

For the critical values of the sequentially rejective step down procedure and those of the closed testing procedure, it is clear that

$$c_{K-1} = c_I, \quad c_k = \max_{\sharp(I_q)=k} c_{I_q} \quad \text{for } k = 1, 2, \ldots, K-2.$$

If $n_2 = n_3 = \cdots = n_K$ and $\sharp(I_q) = k$, $c_{I_q} = c_k$ which implies the critical values of two stepwise procedures are same. Although we may use $c_{\sharp(I_q)}$ instead of c_{I_q} for unbalanced sample sizes in the closed testing procedure, the procedure is more conservative. In the next Subsection we indicate that two procedures are equivalent when we use same critical values $c_1, c_2, \ldots, c_{K-1}$.

2.4. Relation between two stepwise procedures

In this Subsection we indicate that the power of the sequentially rejective step down procedure is not higher than that of the closed testing procedure. Furthermore, we indicate that two procedures are equivalent when we use same critical values.

First, we indicate that the power of the sequentially rejective step down procedure is not higher than that of the closed testing procedure. It is sufficient to indicate that H_{1k} rejected by the sequentially rejective step down procedure is also rejected by the closed testing procedure. If $|S_{1k}| > c_{K-1}$ and $k \in I_q$, $S_{1I_q} > c_{K-1} = c_I \geq c_{I_q}$. This implies H_{1k} is rejected by the closed testing procedure. Next, assume $c_{l+1} \geq |S_{1k}| > c_l$ for some l $(1 \leq l \leq K-2)$. Then, $S_{1I_q} > c_l \geq c_{I_q}$ for each I_q satisfying $k \in I_q$ and $\sharp(I_q) \leq l$. We indicate that $S_{1I_q} > c_{I_q}$ for each I_q satisfying $k \in I_q$ and $\sharp(I_q) > l$. Since H_{1k} is rejected by the sequentially rejective step down procedure, $j(l')$ $(2 \leq j(l') \leq K)$ exists satisfying $|S_{1j(l')}| > c_{l'}$ for each $l < l' \leq K-1$. Then $|S_{1I}| > c_I$. Assume $\sharp(I_q) = K-2$. If $j(K-1) \in I_q$, $S_{1I_q} > c_{K-1} > c_{I_q}$. If $j(K-1) \notin I_q$, $j(K-2) \in I_q$ and $S_{1I_q} > c_{K-2} \geq c_{I_q}$. Assume $\sharp(I_q) = K-3$. If $j(K-1) \in I_q$ or $j(K-2) \in I_q$, $S_{1I_q} > c_{K-2} > c_{I_q}$. If $j(K-1), j(K-2) \notin I_q$, $j(K-3) \in I_q$ and $S_{1I_q} > c_{K-3} \geq c_{I_q}$. Continuing similar steps, we can confirm that $S_{1I_q} > c_{I_q}$ for each I_q satisfying $k \in I_q$ and $\sharp(I_q) > l$. Therefore, since $S_{1I_q} > c_{I_q}$ for each I_q satisfying $k \in I_q$, H_{1k} is rejected by the closed testing procedure.

H_{1k} rejected by the closed testing procedure is occasionally retained by the sequentially rejective step down procedure. The example is given in the next section. However, when we use same critical values $c_1, c_2, \ldots, c_{K-1}$ for two procedures, it is possible to indicate that H_{1k} rejected by the closed testing procedure is also rejected by the sequentially rejective step down procedure. If $|S_{1k}| > c_{K-1}$, H_{1k} is rejected by the sequentially rejective step down procedure. Next, assume $c_{l+1} \geq |S_{1k}| > c_l$ for some l $(1 \leq l \leq K-2)$. Since $S_{1I} > c_{K-1}$, $j(K-1)$ $(2 \leq j(K-1) \leq K)$ exists satisfying $|S_{1j(K-1)}| > c_{K-1}$. Assume $k \in I_q$, $j(K-1) \notin I_q$ and $\sharp(I_q) = K-2$. Since $S_{1I_q} > c_{K-2}$, $j(K-2)$ $(2 \leq j(K-2) \leq K)$ exists satisfying $|S_{1j(K-2)}| > c_{K-2}$. Continuing similar steps, we obtain $j(l')$ $(l' = l+1, l+2, \ldots, K-1)$ satisfying $|S_{1j(l')}| > c_{l'}$. Therefore, H_{1k} is rejected by the sequentially rejective step down procedure. Specifically, the sequentially rejective step down procedure and the closed testing procedure are equivalent when we use same critical values for two procedures.

2.5. Simulation results

In Subsection 2.4 we indicated that the power of the sequentially rejective step down procedure is not higher than that of the closed testing procedure and two procedures are equivalent when we use same critical values for them. In this Subsection we compare two stepwise procedures in terms of simulation results regarding the power of the test for unbalanced sample sizes.

Let $K = 5$. We set up four types of $(n_1, n_2, n_3, n_4, n_5)$s as follows.

Sam.1. $(10, 20, 10, 20, 10)$
Sam.2. $(20, 10, 20, 10, 20)$
Sam.3. $(10, 30, 10, 30, 10)$
Sam.4. $(30, 10, 30, 10, 30)$

Let $\alpha = 0.05$. Table 1 gives critical values of the closed testing procedure. Table 2 gives those of the sequentially rejective step down procedure.

We give an example that a hypothesis rejected by the closed testing procedure is retained by the sequentially rejective step down procedure. For example, assume

$$|S_{12}| = 2.000, |S_{13}| = 2.255, |S_{14}| = 2.400, |S_{15}| = 2.500$$

in Sam.1. By the closed testing procedure $H_{12}, H_{13}, H_{14}, H_{15}$ are rejected. By the sequentially rejective step down procedure H_{14}, H_{15} are rejected and H_{12}, H_{13} are retained.

Table 1 : Critical values of the closed testing procedure

	Sam.1	Sam.2	Sam.3	Sam.4
$c_{\{2,3,4,5\}}$	2.480	2.516	2.452	2.507
$c_{\{2,3,4\}}$	2.382	2.416	2.350	2.409
$c_{\{2,3,5\}}$	2.395	2.408	2.375	2.399
$c_{\{2,4,5\}}$	2.382	2.416	2.350	2.409
$c_{\{3,4,5\}}$	2.395	2.408	2.375	2.399
$c_{\{2,3\}}$	2.252	2.264	2.233	2.257
$c_{\{2,4\}}$	2.235	2.270	2.203	2.263
$c_{\{2,5\}}$	2.252	2.264	2.233	2.257
$c_{\{3,4\}}$	2.252	2.264	2.233	2.257
$c_{\{3,5\}}$	2.260	2.254	2.250	2.245
$c_{\{4,5\}}$	2.252	2.264	2.233	2.257
$c_{\{2\}}$	1.998	1.993	1.989	1.986
$c_{\{3\}}$	1.998	1.993	1.989	1.986
$c_{\{4\}}$	1.998	1.993	1.989	1.986
$c_{\{5\}}$	1.998	1.993	1.989	1.986

Table 2 : Critical values of the sequentially rejective step down procedure

	Sam.1	Sam.2	Sam.3	Sam.4
c_4	2.480	2.516	2.452	2.507
c_3	2.395	2.416	2.375	2.409
c_2	2.260	2.270	2.250	2.263
c_1	1.998	1.993	1.989	1.986

Next, we consider the power of the test. Here, we focus on the all-pairs power defined by Ramsey (1978). We set up four types of $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$ as follows.

Case 1. $(0, \delta, \delta, \delta, \delta)$, Case 2. $(0, \delta, \delta, \delta, 0)$, Case 3. $(0, \delta, \delta, 0, 0)$, Case 4. $(0, \delta, 0, 0, 0)$.

Here $\delta = 1.0, 1.5$. Since the power depends on the unknown σ^2, let $\sigma^2 = 1$.

The power of the sequentially rejective step down procedure can be obtained by the formulation derived by Dunnett *et al.* (2001). Since it is difficult to formulate the power of the close testing procedure, the power is calculated by Monte Carlo simulation with 1,000,000 times of experiments. Table 3 gives the power for two procedures. Here CT and SD mean the closed testing procedure and the sequentially rejective step down procedure, respectively. Table shows that the difference of the power between two procedures is fairly small in each case.

Table 3 : Power comparison

			Sam.1	Sam.2	Sam.3	Sam.4
Case 1.	$\delta = 1.0$	CT	0.354	0.488	0.384	0.596
		SD	0.351	0.487	0.382	0.596
	$\delta = 1.5$	CT	0.821	0.935	0.838	0.965
		SD	0.820	0.935	0.837	0.965
Case 2.	$\delta = 1.0$	CT	0.354	0.407	0.403	0.491
		SD	0.351	0.404	0.398	0.489
	$\delta = 1.5$	CT	0.818	0.895	0.846	0.938
		SD	0.816	0.894	0.842	0.937
Case 3.	$\delta = 1.0$	CT	0.355	0.491	0.393	0.605
		SD	0.350	0.491	0.383	0.605
	$\delta = 1.5$	CT	0.801	0.919	0.827	0.954
		SD	0.799	0.918	0.820	0.954
Case 4.	$\delta = 1.0$	CT	0.544	0.530	0.614	0.594
		SD	0.544	0.529	0.614	0.594
	$\delta = 1.5$	CT	0.916	0.910	0.950	0.944
		SD	0.916	0.910	0.950	0.944

3. All-pairwise comparison

3.1. Single step procedure

First, we discuss the single step procedure proposed by Tukey (1953). Intended to compare μ_i and μ_j $(i < j)$ we set up a null hypothesis and its alternative hypothesis as

$$H_{ij} : \mu_i = \mu_j \quad \text{vs.} \quad H_{ij}^A : \mu_i \neq \mu_j .$$

We consider the simultaneous test of all H_{ij}s based on the single step procedure. We use the statistic

$$S_{ij} = \sqrt{\frac{n_i n_j}{n_i + n_j}} (\bar{X}_i - \bar{X}_j) s^{-1}$$

for testing H_{ij}. If $|S_{ij}| > c$ for a specified critical value c, H_{ij} is rejected. Otherwise, it is retained. We want to determine c so that

$$P(\max_{i<j} |S_{ij}| > c) = \alpha$$

for a specified significance level α when all H_{ij}s are true. If it is difficult, we want to determine c so that

$$P(\max_{i<j} |S_{ij}| > c) \leq \alpha .$$

If $n_1 = n_2 = \cdots = n_K$,

$$P(\max_{1 \leq i < j \leq K} |S_{ij}| > c) = 1 - K \int_0^\infty \left[\int_{-\infty}^\infty \{\Phi(z) - \Phi(z - \sqrt{2} c s_0)\}^{K-1} \phi(z) dz \right] g(s_0) ds_0 . \tag{4}$$

If sample sizes are unbalanced,

$$P(\max_{1 \leq i < j \leq K} |S_{ij}| > c) \leq 1 - K \int_0^\infty \left[\int_{-\infty}^\infty \{\Phi(z) - \Phi(z - \sqrt{2} c s_0)\}^{K-1} \phi(z) dz \right] g(s_0) ds_0 . \tag{5}$$

Although (5) had been called Tukey-Cramer's conjecture, it was proved by Hayter (1984). We determine c so that

$$K\int_0^\infty \left[\int_{-\infty}^\infty \{\Phi(z)-\Phi(z-\sqrt{2}cs_0)\}^{K-1}\phi(z)dz\right] g(s_0)ds_0 = 1-\alpha. \tag{6}$$

If $n_1 = n_2 = \cdots = n_K$, c satisfies the significance level α exactly. Otherwise, c is a conservative critical value for α. In both cases, assuming $X_1^*, X_2^*, \ldots, X_K^*$ are mutually independent and each of them is distributed according to $N(0, K\sigma^2/N)$ independently of s, let

$$S_{ij}^* = \sqrt{\frac{N}{2K}}(X_i^* - X_j^*)s^{-1}$$

for $1 \le i < j \le K$. Note

$$P(\max_{1\le i<j\le K}|S_{ij}^*| > c) = 1 - K\int_0^\infty \left[\int_{-\infty}^\infty \{\Phi(z)-\Phi(z-\sqrt{2}cs_0)\}^{K-1}\phi(z)dz\right] g(s_0)ds_0. \tag{7}$$

On the other hand, we can obtain a conservative critical value using Bonferroni's inequality

$$P(\max_{i<j}|S_{ij}| > c) < \sum_{i<j} P(|S_{ij}| > c). \tag{8}$$

Each S_{ij} is distributed according to t_ψ under H_{ij} where t_ψ is the t-distribution with ψ degrees of freedom. If we determine $c^{(1)}_{K(K-1)/2}$ so that

$$P(|t_\psi| > c^{(1)}_{K(K-1)/2}) = \frac{2\alpha}{K(K-1)},$$

we obtain

$$P(\max_{i<j}|S_{ij}| > c^{(1)}_{K(K-1)/2}) < \alpha$$

by (8). On the other hand, using Sidak's inequality and the inequality given by Hsu (see page 227, Corollary A.1.1, 1996). we obtain the inequality

$$P(\max_{i<j}|S_{ij}| > c) \le 1 - \prod_{i<j} P(|S_{ij}| \le c). \tag{9}$$

If we determine $c^{(2)}_{K(K-1)/2}$ so that

$$P(|t_\psi| > c^{(2)}_{K(K-1)/2}) = 1-(1-\alpha)^{\frac{2}{K(K-1)}},$$

we obtain

$$P(\max_{i<j}|S_{ij}| > c^{(2)}_{K(K-1)/2}) \le \alpha$$

by (9). Since

$$1-(1-\alpha)^{\frac{2}{K(K-1)}} > \frac{2\alpha}{K(K-1)},$$

we obtain

$$c^{(1)}_{K(K-1)/2} > c^{(2)}_{K(K-1)/2},$$

which means $c^{(2)}_{K(K-1)/2}$ is less conservative compared to $c^{(1)}_{K(K-1)/2}$. On the other hand, for c determined by (6) we obtain $c^{(2)}_{K(K-1)/2} \geq c$ by (7) and

$$P(\max_{i<j} |S^*_{ij}| > c^{(2)}_{K(K-1)/2}) \leq \alpha.$$

For $c > 0$ and arbitrary positive integer m we use the inequality

$$m \int_0^\infty \left[\int_{-\infty}^\infty \{\Phi(z) - \Phi(z - \sqrt{2}cs_0)\}^{m-1} \phi(z) dz \right] g(s_0) ds_0 \geq P(|t_\psi| \leq c)^{\frac{m(m-1)}{2}} \tag{10}$$

in hereafter discussions. Furthermore, we define $c^{(1)}_m$ and $c^{(2)}_m$ so that

$$P(|t_\psi| > c^{(1)}_m) = \frac{\alpha}{m},\ P(|t_\psi| > c^{(2)}_m) = 1 - (1-\alpha)^{\frac{1}{m}}.$$

Then

$$c^{(i)}_1 < c^{(i)}_2 < c^{(i)}_3 < \cdots\cdots \quad \text{for} \quad i = 1, 2.$$

3.2. Ryan-Einot-Gabriel-Welsch's procedure

Next, we discuss the closed testing procedure called Ryan-Einot-Gabriel-Welsch's procedure. For arbitrary subset $I^* = \{i_1, i_2, \ldots, i_k\}$ $(1 \leq i_1 < i_2 < \cdots < i_k \leq K)$ of $I = \{1, 2, \ldots, K\}$ we define

$$H_{I^*} : \mu_{i_1} = \mu_{i_2} = \cdots = \mu_{i_k}.$$

Letting F be the family consisting of all H_{ij}s and all kinds of intersections of plural H_{ij}s, F is closed. Each hypothesis in F is equal to single H_{I^*} or $H_{I_1} \cap H_{I_2} \cap \cdots \cap H_{I_q}$ where $I_1, I_2, \ldots, I_q$ are disjoint. We discuss Ryan-Einot-Gabriel-Welsch's procedure for F. When we test H_{I^*} where $I^* = \{i_1, i_2, \ldots, i_k\}$, we use the statistic

$$S_{I^*} = \max_{i,j \in I^*\ (i<j)} |S_{ij}|.$$

The critical value c for testing H_{I^*} is determined so that

$$\sharp(I^*) \int_0^\infty \left[\int_{-\infty}^\infty \{\Phi(z) - \Phi(z - \sqrt{2}cs_0)\}^{\sharp(I^*)-1} \phi(z) dz \right] g(s_0) ds_0 = 1 - \alpha.$$

Since c depends on $\sharp(I^*)$, let $c_{\sharp(I^*)}$ denote c determined by the above equation. If $S_{I^*} > c_{\sharp(I^*)}$, H_{I^*} is rejected. Otherwise, it is retained. Next, we discuss how to test $H_{I_1} \cap H_{I_2} \cap \cdots \cap H_{I_q}$. Letting $M_1 = \sharp(I_1) + \sharp(I_2) + \cdots + \sharp(I_q)$, allocate

$$1 - (1-\alpha)^{\sharp(I_i)/M_1}$$

to H_{I_i} for $(i = 1, 2, \ldots, q)$. This is called Tukey-Welsh's allocation of α. For $i = 1, 2, \ldots, q$ we determine $c_{\sharp(I_i), M_1}$ so that

$$\sharp(I_i) \int_0^\infty \left[\int_{-\infty}^\infty \{\Phi(z) - \Phi(z - \sqrt{2}c_{\sharp(I_i), M_1} s_0)\}^{\sharp(I_i)-1} \phi(z) dz \right] g(s_0) ds_0 = 1 - (1-\alpha)^{\sharp(I_i)/M_1}.$$

Specifically, intended to test $H_{I_1} \cap H_{I_2} \cap \cdots \cap H_{I_q}$ we set up the critical value $c_{\sharp(I_i),M_1}$ for testing H_{I_i} for $i = 1, 2, \ldots, q$. If $S_{I_i} > c_{\sharp(I_i),M_1}$ for at least one i, $H_{I_1} \cap H_{I_2} \cap \cdots \cap H_{I_q}$ is rejected. Otherwise, it is retained. It is indicated that the probability that $H_{I_1} \cap H_{I_2} \cap \cdots \cap H_{I_q}$ is rejected when it is true is not greater than α. We specified the way to test each hypothesis in F satisfying the specified significance level α. We test the hypotheses in F hierarchically. Specifically, if a hypothesis and all hypotheses deriving it are rejected, we reject the hypothesis. Otherwise we retain it.

3.3. Another approach for closed testing procedure

We discuss another approach for closed testing procedure. When we test $H_{I_1} \cap H_{I_2} \cap \cdots \cap H_{I_q}$ where $I_1, I_2, \ldots, I_q$ are disjoint by Ryan-Einot-Gabriel-Welsch's procedure, we test each of $H_{I_1}, H_{I_2}, \ldots, H_{I_q}$. Here, we discuss the closed testing procedure testing $H_{I_1} \cap H_{I_2} \cap \cdots \cap H_{I_q}$ at a time using the statistic $\max\{S_{I_1}, S_{I_2}, \ldots, S_{I_q}\}$. When the value of s_0 is given, $S_{I_1}, S_{I_2}, \ldots, S_{I_q}$ are independent and

$$P(\max\{S_{I_1}, S_{I_2}, \ldots, S_{I_q}\} > c) = 1 - \int_0^\infty \prod_{i=1}^q P(S_{I_i} \le c|s_0) g(s_0) ds_0 \tag{11}$$

for $c > 0$. Since

$$\int_0^\infty \prod_{i=1}^q P(S_{I_i} \le c|s_0) g(s_0) ds_0 \ge \prod_{i=1}^q \int_0^\infty P(S_{I_i} \le c|s_0) g(s_0) ds_0 = \prod_{i=1}^q P(S_{I_i} \le c)$$

by the inequality given by Hsu (1996), we obtain

$$P(\max\{S_{I_1}, S_{I_2}, \ldots, S_{I_q}\} > c) \le 1 - \prod_{i=1}^q P(S_{I_i} \le c)$$

by (11). Letting

$$P_{\sharp(I_i)}(c) = \sharp(I_1) \int_0^\infty \left[\int_{-\infty}^\infty \{\Phi(z) - \Phi(z - \sqrt{2} c s_0)\}^{\sharp(I_i)-1} \phi(z) dz \right] g(s_0) ds_0$$

for $i = 1, 2, \ldots, q$, we obtain

$$P(\max\{S_{I_1}, S_{I_2}, \ldots, S_{I_q}\} > c) \le 1 - \prod_{i=1}^q P_{\sharp(I_i)}(c). \tag{12}$$

If we determine $c_{\sharp(I_1),\sharp(I_2),\ldots,\sharp(I_q)}$ so that

$$\prod_{i=1}^q P_{\sharp(I_i)}(c_{\sharp(I_1),\sharp(I_2),\ldots,\sharp(I_q)}) = 1 - \alpha, \tag{13}$$

we obtain

$$P(\max\{S_{I_1}, S_{I_2}, \ldots, S_{I_q}\} > c_{\sharp(I_1),\sharp(I_2),\ldots,\sharp(I_q)}) \le \alpha$$

by (12). Letting

$$M_2 = \frac{\sharp(I_1)(\sharp(I_1) - 1)}{2} + \frac{\sharp(I_2)(\sharp(I_2) - 1)}{2} + \cdots + \frac{\sharp(I_q)(\sharp(I_q) - 1)}{2},$$

we obtain

$$\prod_{i=1}^{q} P_{\sharp(I_i)}(c) \geq P(|t_\psi| \leq c)^{M_2}$$

by (10). Therefore

$$\prod_{i=1}^{q} P_{\sharp(I_i)}(c_{M_2}^{(2)}) \geq 1-\alpha$$

which means

$$c_{\sharp(I_1),\sharp(I_2),\ldots,\sharp(I_q)} \leq c_{M_2}^{(2)}$$

by (13).

3.4. Holm's procedure

We discuss Holm's procedure which is the sequentially rejective step down procedure for all-pairwise comparison. It consists of $K(K-1)/2$ steps of tests. Arranging $|S_{ij}|$s in order of a size of value, assume

$$|S_{(1)}| \leq |S_{(2)}| \leq \cdots \leq |S_{(K(K-1)/2)}|.$$

$H_{(1)}, H_{(2)}, \ldots, H_{(K(K-1)/2)}$ denote hypotheses corresponding to $S_{(1)}, S_{(2)}, \ldots, S_{(K(K-1)/2)}$. Then, we test $H_{(1)}, H_{(2)}, \ldots, H_{(K(K-1)/2)}$ sequentially as follows.

Step 1.

Case 1. If $|S_{(K(K-1)/2)}| \leq c_{K(K-1)/2}^{(1)}$, we retain $H_{(1)}, H_{(2)}, \ldots, H_{(K(K-1)/2)}$ and stop the test.

Case 2. If $|S_{(K(K-1)/2)}| > c_{K(K-1)/2}^{(1)}$, we reject $H_{(K(K-1)/2)}$ and go to the next step.

Step 2.

Case 1. If $|S_{(K(K-1)/2-1)}| \leq c_{K(K-1)/2-1}^{(1)}$, we retain $H_{(1)}, H_{(2)}, \ldots, H_{(K(K-1)/2-1)}$ and stop the test.

Case 2. If $|S_{(K(K-1)/2-1)}| > c_{K(K-1)/2-1}^{(1)}$, we reject $H_{(K(K-1)/2-1)}$ and go to the next step.

We repeat similar judgments till up to Step $K(K-1)/2$.

3.5. Shaffer's procedure and Holland and Copenhaver's procedure

We discuss Shaffer's procedure and Holland and Copenhaver's procedure improving Holm's procedure. First, we discuss Shaffer's procedure. We test $H_{(1)}$, $H_{(2)}$, …, $H_{(K(K-1)/2)}$ sequentially as follows.

Step 1.

Case 1. If $|S_{(K(K-1)/2)}| \leq c_{K(K-1)/2}^{(1)}$, we retain $H_{(1)}, H_{(2)}, \ldots, H_{(K(K-1)/2)}$ and stop the test.

Case 2. If $|S_{(K(K-1)/2)}| > c_{K(K-1)/2}^{(1)}$, we reject $H_{(K(K-1)/2)}$ and go to the next step.

Step 2.

When $H_{(K(K-1)/2)}$ is not true, let $m(K(K-1)/2-1)$ be the maximum number of hypotheses which can be simultaneously true among $H_{(1)}, H_{(2)}, \ldots, H_{(K(K-1)/2-1)}$.

Case 1. If $|S_{(K(K-1)/2-1)}| \leq c^{(1)}_{m(K(K-1)/2-1)}$, we retain $H_{(1)}, H_{(2)}, \ldots, H_{(K(K-1)/2-1)}$ and stop the test.
Case 2. If $|S_{(K(K-1)/2-1)}| > c^{(1)}_{m(K(K-1)/2-1)}$, we reject $H_{(K(K-1)/2-1)}$ and go to the next step.
Step 3.
When $H_{(K(K-1)/2)}$ and $H_{(K(K-1)/2-1)}$ are not true, let $m(K(K-1)/2-2)$ be the maximum number of hypotheses which can be simultaneously true among $H_{(1)}$, $H_{(2)}$, …, $H_{(K(K-1)/2-2)}$.
Case 1. If $|S_{(K(K-1)/2-2)}| \leq c^{(1)}_{m(K(K-1)/2-2)}$, we retain $H_{(1)}, H_{(2)}, \ldots, H_{(K(K-1)/2-2)}$ and stop the test.
Case 2. If $|S_{(K(K-1)/2-2)}| > c^{(1)}_{m(K(K-1)/2-2)}$, we reject $H_{(K(K-1)/2-2)}$ and go to the next step.

We repeat similar judgments till up to Step $K(K-1)/2$.

$c^{(1)}_l$ for $l = m(K(K-1)/2-1), m(K(K-1)/2-2), \ldots, m(1)$ is determined depending on the process of the test. However, Holland and Copenhaver (1987) set up the critical values conservatively in advance of the test considering all sorts of cases and tabulated them for $3 \leq K \leq 10$ using $c^{(2)}_m$ instead of $c^{(1)}_m$. Shaffer's procedure and Holland and Copenhaver's procedure are more powerful compared to Holm's procedure.

3.6. Relation between two stepwise procedures

In this Subsection we discuss the relation between the closed testing procedure and the sequentially rejective step down procedure. Here, we focus on Holland and Copenhaver's procedure among three types of sequentially rejective step down procedures. Since it is difficult to clarify the theoretical relation regarding power of the test between Ryan-Einot-Gabriel-Welsch's procedure and Holland and Copenhaver's procedure, we focus on the closed testing procedure discussed in Subsection 3.3. Although we expect that the power of Holland and Copenhaver's procedure is not higher than that of the proposed closed testing procedure, it is difficult to indicate it in general situation. Here, we indicate it specifying K. We give the indications only for $K = 3, 4, 5$, because the indications need many pages for $K \geq 6$.

We define abbreviated notations. HC and CT mean Holland and Copenhaver's procedure and the closed testing procedure discussed in Subsection 3.3, respectively.

I. $K = 3$
The critical values of HC are $c^{(2)}_3$, $c^{(2)}_1$, $c^{(2)}_1$. The critical values of CT are c_3, c_2. Assume H_{ij} is rejected by HC.
Case 1. If $|S_{ij}| > c^{(2)}_3$, H_{ij} is rejected by CT, because $c^{(2)}_3 \geq c_3$.
Case 2. If $c^{(2)}_3 \geq |S_{ij}| > c^{(2)}_1$, $\{i', j'\}$ exists satisfying $|S_{i'j'}| > c^{(2)}_3$. Therefore, H_{ij} is rejected by CT, because $S_{\{1,2,3\}} \geq c^{(2)}_3 \geq c_3$ and $|S_{ij}| > c^{(2)}_1 = c_2$.

II. $K = 4$
The critical values of HC are $c^{(2)}_6$, $c^{(2)}_3$, $c^{(2)}_3$, $c^{(2)}_3$, $c^{(2)}_2$, $c^{(2)}_1$. The critical values of CT are c_4, c_3, $c_{2,2}$, c_2. Assume H_{ij} is rejected by HC.

Case 1. If $|S_{ij}| > c_6^{(2)}$, H_{ij} is rejected by CT, because $c_6^{(2)} \geq c_4$.
Case 2. If $c_6^{(2)} \geq |S_{ij}| > c_3^{(2)}$, $\{i', j'\}$ exists satisfying $|S_{i'j'}| > c_6^{(2)}$. Therefore, H_{ij} is also rejected by CT, because $S_{\{1,2,3,4\}} \geq c_6^{(2)} \geq c_4$ and $|S_{ij}| > c_3^{(2)} \geq c_3$.
Case 3. If $c_3^{(2)} \geq |S_{ij}| > c_2^{(2)}$, $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$ satisfying

$$|S_{i_1j_1}| > c_6^{(2)},\ |S_{i_2j_2}| > c_3^{(2)},\ |S_{i_3j_3}| > c_3^{(2)},\ |S_{i_4j_4}| > c_3^{(2)}$$

exist. Then $S_{\{1,2,3,4\}} > c_6^{(2)} \geq c_4$. When $\{s_1, s_2, s_3\}$ includes $\{i, j\}$, $\{s_1, s_2, s_3\}$ includes at least one of $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$. This means $S_{\{s_1,s_2,s_3\}} > c_3^{(2)} \geq c_3$. Therefore, H_{ij} is rejected by CT, because $|S_{ij}| > c_2^{(2)} = c_{2,2}$.
Case 4. If $c_2^{(2)} \geq |S_{ij}| > c_1^{(2)}$, $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$ satisfying

$$|S_{i_1j_1}| > c_6^{(2)},\ |S_{i_2j_2}| > c_3^{(2)},\ |S_{i_3j_3}| > c_3^{(2)},\ |S_{i_4j_4}| > c_3^{(2)},\ |S_{i_5j_5}| > c_2^{(2)}$$

exist. Then $S_{\{1,2,3,4\}} > c_6^{(2)} \geq c_4$. When $\{s_1, s_2, s_3\}$ includes $\{i, j\}$, $H_{\{s_1,s_2,s_3\}}$ rejected similarly as Case 3. Assuming that i', j' are obtained by excluding i, j from 1,2,3,4, $\{i', j'\}$ is equal to one of $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$. Therefore $|S_{i'j'}| > c_2^{(2)} = c_{2,2}$. This means $H_{ij} \cap H_{i'j'}$ is rejected. Therefore, H_{ij} is rejected by CT, because $|S_{ij}| > c_1^{(2)} = c_2$.

III. $K = 5$
The critical values of HC are $c_{10}^{(2)}$, $c_6^{(2)}$, $c_6^{(2)}$, $c_6^{(2)}$, $c_6^{(2)}$, $c_4^{(2)}$, $c_4^{(2)}$, $c_3^{(2)}$, $c_2^{(2)}$, $c_1^{(2)}$. The critical values of CT are c_5, c_4, $c_{3,2}$, c_3, $c_{2,2}$, c_2. Assume H_{ij} is rejected by HC.
Case 1. If $|S_{ij}| > c_{10}^{(2)}$, H_{ij} is also rejected by CT, because $c_{10}^{(2)} \geq c_5$.
Case 2. If $c_{10}^{(2)} \geq |S_{ij}| > c_6^{(2)}$, $\{i', j'\}$ exists satisfying $|S_{i'j'}| > c_{10}^{(2)}$. Then $S_{\{1,2,3,4,5\}} \geq c_{10}^{(2)} \geq c_5$. Therefore, H_{ij} is rejected by CT, because $|S_{ij}| > c_6^{(2)} \geq c_4$.
Case 3. If $c_6^{(2)} \geq |S_{ij}| > c_4^{(2)}$, $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$ satisfying

$$|S_{i_1j_1}| > c_{10}^{(2)},\ |S_{i_2j_2}| > c_6^{(2)},\ |S_{i_3j_3}| > c_6^{(2)},\ |S_{i_4j_4}| > c_6^{(2)},\ |S_{i_5j_5}| > c_6^{(2)}$$

exist. Then $S_{\{1,2,3,4,5\}} \geq c_{10}^{(2)} \geq c_5$. If $\{s_1, s_2, s_3, s_4\}$ includes $\{i, j\}$, $\{s_1, s_2, s_3, s_4\}$ includes at least one of $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$. This means $S_{\{s_1,s_2,s_3,s_4\}} > c_6^{(2)} \geq c_4$. Therefore, H_{ij} is also rejected by CT, because $|S_{ij}| > c_4^{(2)} \geq c_{3,2}$.
Case 4. If $c_4^{(2)} \geq |S_{ij}| > c_3^{(2)}$, $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$, $|i_6, j_6\}$, $\{i_7, j_7\}$ satisfying

$$|S_{i_1j_1}| > c_{10}^{(2)},\ |S_{i_2j_2}| > c_6^{(2)},\ |S_{i_3j_3}| > c_6^{(2)},\ |S_{i_4j_4}| > c_6^{(2)},\ |S_{i_5j_5}| > c_6^{(2)},$$

$$|S_{i_6j_6}| > c_4^{(2)},\ |S_{i_7j_7}| > c_4^{(2)}$$

exist. Then $S_{\{1,2,3,4,5\}} \geq c_{10}^{(2)} \geq c_5$. If $\{s_1, s_2, s_3, s_4\}$ includes $\{i, j\}$, we obtain $S_{\{s_1,s_2,s_3,s_4\}} > c_6^{(2)} \geq c_4$ similarly as Case 3. Assume $\{s_1, s_2, s_3\}$ includes $\{i, j\}$ and $\{s_4, s_5\}$ is obtained by excluding s_1, s_2, s_3 from 1,2,3,4,5. If $\{s_1, s_2, s_3\}$ includes at least one of $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$, $|i_6, j_6\}$, $\{i_7, j_7\}$, $S_{\{s_1,s_2,s_3\}} > c_4^{(2)} \geq c_{3,2}$. Otherwise $\{s_4, s_5\}$ is equal to one of $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$, $|i_6, j_6\}$, $\{i_7, j_7\}$. Then

$S_{s_4 s_5} > c_4^{(2)} \geq c_{3,2}$. These mean $H_{\{s_1,s_2,s_3\}} \cap H_{s_4 s_5}$ is rejected. Assuming $\{s_1, s_2, s_3\}$ is obtained by excluding $\{i, j\}$ from 1,2,3,4,5, $\{s_1, s_2, s_3\}$ includes at least one of $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$, $|i_6, j_6\}$, $\{i_7, j_7\}$. Then $S_{\{s_1,s_2,s_3\}} > c_4^{(2)} \geq c_{3,2}$. These mean $H_{\{s_1,s_2,s_3\}} \cap H_{ij}$ is rejected. Therefore, H_{ij} is also rejected by CT, because $|S_{ij}| > c_3^{(2)} \geq c_3$.

Case 5. If $c_3^{(2)} \geq |S_{ij}| > c_2^{(2)}$, $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$, $|i_6, j_6\}$, $\{i_7, j_7\}$, $\{i_8, j_8\}$ satisfying

$$|S_{i_1 j_1}| > c_{10}^{(2)},\ |S_{i_2 j_2}| > c_6^{(2)},\ |S_{i_3 j_3}| > c_6^{(2)},\ |S_{i_4 j_4}| > c_6^{(2)},\ |S_{i_5 j_5}| > c_6^{(2)},$$

$$|S_{i_6 j_6}| > c_4^{(2)},\ |S_{i_7 j_7}| > c_4^{(2)},\ |S_{i_8 j_8}| > c_3^{(2)}$$

exist. Then $S_{\{1,2,3,4,5\}} \geq c_{10}^{(2)} \geq c_5$. If $\{s_1, s_2, s_3, s_4\}$ includes $\{i, j\}$, we obtain $S_{\{s_1,s_2,s_3,s_4\}} > c_6^{(2)} \geq c_4$ similarly as Case 3. Assuming $\{s_1, s_2, s_3\}$ includes $\{i, j\}$ and $\{s_4, s_5\}$ is obtained by excluding s_1, s_2, s_3 from 1,2,3,4,5, $H_{\{s_1,s_2,s_3\}} \cap H_{s_4 s_5}$ is rejected similarly as Case 4. Assuming $\{s_1, s_2, s_3\}$ is obtained by excluding $\{i, j\}$ from 1,2,3,4,5, $H_{\{s_1,s_2,s_3\}} \cap H_{ij}$ is rejected similarly as Case 4. If $\{s_1, s_2, s_3\}$ includes $\{i, j\}$, $\{s_1, s_2, s_3\}$ includes at least one of $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$, $|i_6, j_6\}$, $\{i_7, j_7\}$, $\{i_8, j_8\}$. Then $S_{\{s_1,s_2,s_3\}} > c_3^{(2)} \geq c_3$. Therefore, H_{ij} is also rejected by CT, because $|S_{ij}| > c_2^{(2)} = c_{2,2}$.

Case 6. If $c_2^{(2)} \geq |S_{ij}| > c_1^{(2)}$, $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$, $|i_6, j_6\}$, $\{i_7, j_7\}$, $\{i_8, j_8\}$, $\{i_9, j_9\}$ satisfying

$$|S_{i_1 j_1}| > c_{10}^{(2)},\ |S_{i_2 j_2}| > c_6^{(2)},\ |S_{i_3 j_3}| > c_6^{(2)},\ |S_{i_4 j_4}| > c_6^{(2)},\ |S_{i_5 j_5}| > c_6^{(2)},$$

$$|S_{i_6 j_6}| > c_4^{(2)},\ |S_{i_7 j_7}| > c_4^{(2)},\ |S_{i_8 j_8}| > c_3^{(2)},\ |S_{i_9 j_9}| > c_2^{(2)}$$

exist. Then $S_{\{1,2,3,4,5\}} \geq c_{10}^{(2)} \geq c_5$. If $\{s_1, s_2, s_3, s_4\}$ includes $\{i, j\}$, we obtain $S_{\{s_1,s_2,s_3,s_4\}} > c_6^{(2)} \geq c_4$ similarly as Case 3. Assuming $\{s_1, s_2, s_3\}$ includes $\{i, j\}$ and $\{s_4, s_5\}$ is obtained by excluding s_1, s_2, s_3 from 1,2,3,4,5, $H_{\{s_1,s_2,s_3\}} \cap H_{s_4 s_5}$ is rejected similarly as Case 4. Assuming $\{s_1, s_2, s_3\}$ is obtained by excluding $\{i, j\}$ from 1,2,3,4,5, $H_{\{s_1,s_2,s_3\}} \cap H_{ij}$ is rejected similarly as Case 4. Assuming $\{s_1, s_2, s_3\}$ includes $\{i, j\}$, $H_{\{s_1,s_2,s_3\}}$ is rejected similarly as Case 5. Assuming $\{i, j\}$ and $\{i', j'\}$ are disjoint, $\{i', j'\}$ is the one of $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\{i_3, j_3\}$, $|i_4, j_4\}$, $\{i_5, j_5\}$, $|i_6, j_6\}$, $\{i_7, j_7\}$, $\{i_8, j_8\}$, $\{i_9, j_9\}$. $H_{ij} \cap H_{i'j'}$ is rejected, because $|S_{i'j'}| > c_2^{(2)} = c_{2,2}$. Therefore, H_{ij} is also rejected by CT, because $|S_{ij}| > c_1^{(2)} = c_2$.

It is possible to indicate that the power of Holland and Copenhaver's procedure is not higher than that of Ryan-Einot-Gabriel-Welsch's procedure for $K = 3, 4$. However, it is difficult for $K \geq 5$.

3.7. Simulation results

We discussed Ryan-Einot-Gabriel-Welsch's procedure, another type of closed testing procedure and three types of sequentially rejective step down procedures. Focusing on Holland-Copenhaver's procedure among three types of sequentially rejective step down procedures, we indicated that the power of Holland-Copenhaver's procedure is not higher than that of the proposed closed testing procedure specifying the total number of means. In this Subsection we give simulation results regarding the critical values and the power of the test intended to compare two types of closed testing procedures and Holland-Copenhaver's procedure. CT1 and CT2 denote the closed testing procedures discussed in Subsections 3.2 and 3.3, respectively. HC denotes Holland and Copenhaver's procedure. Let $K = 5$ and $\alpha = 0.05$. Since critical values are determined by $N = n_1 + n_2 + n_3 + n_4 + n_5$, let $N = 75$. Tables 4 to 6 give critical values of CT1, CT2 and HC, respectively.

Table 4 : Critical values of CT1

c_5	c_4	$c_{3,5}$	$c_{2,5}$	c_3	$c_{2,4}$	c_2
2.800	2.632	2.599	2.375	2.395	2.286	1.995

Table 5 : Critical values of CT2

c_5	c_4	$c_{3,2}$	c_3	$c_{2,2}$	c_2
2.800	2.632	2.523	2.395	2.286	1.995

Table 6 : Critical values of HC

$c_{10}^{(2)}$	$c_6^{(2)}$	$c_4^{(2)}$	$c_3^{(2)}$	$c_2^{(2)}$	$c_1^{(2)}$
2.891	2.708	2.557	2.447	2.286	1.995

Next, we consider the power of the test. Since the power depends on unknown σ^2, we specify $\sigma^2 = 1$. We set up four types of $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$ as follows.

Case 1 : $(0, \delta, 2\delta, 3\delta, 4\delta)$, Case 2 : $(0, \delta, 2\delta, 3\delta, 3\delta)$,

Case 3 : $(0, \delta, 2\delta, 2\delta, 2\delta)$, Case 4 : $(0, \delta, \delta, \delta, \delta)$.

Here $\delta = 1.0, 1.5$. We set up two types of arrangements of $(n_1, n_2, n_3, n_4, n_5)$ satisfying $n_1 + n_2 + n_3 + n_4 + n_5 = 75$ as

$$\text{Sam.1} : (15, 15, 15, 15, 15),\ \text{Sam.2} : (10, 20, 15, 20, 10).$$

Table 7 gives the power of three procedures. The power is calculated by Monte Carlo simulation with 1,000,000 times of experiments in each case. CT1 and CT2 are uniformly more powerful compared to HC. Although the differences of the power between CT1 and CT2 are uniformly small in Cases 1,2, the power of CT1 is higher than that of CT2 in Cases 3,4. The differences of the power among CT1, CT2 and HC are larger as the number of the pairs consisting of different means is smaller.

Table 7 : Power comparison

		Case 1		Case 2		Case 3		Case 4	
		Sam.1	Sam.2	Sam.1	Sam.2	Sam.1	Sam.2	Sam.1	Sam.2
$\delta = 1.0$	CT1	0.266	0.288	0.202	0.203	0.202	0.213	0.329	0.249
	CT2	0.266	0.287	0.201	0.203	0.170	0.181	0.275	0.204
	HC	0.248	0.273	0.150	0.152	0.116	0.127	0.221	0.162
$\delta = 1.5$	CT1	0.927	0.919	0.894	0.870	0.868	0.855	0.867	0.739
	CT2	0.927	0.919	0.893	0.870	0.850	0.828	0.826	0.684
	HC	0.927	0.920	0.868	0.834	0.794	0.779	0.779	0.624

4. Conclusions

In this study we discussed the closed testing procedures and the sequentially rejective step down procedures for the multiple comparison with a control and the all-pairwise multiple comparison. For the multiple comparison with a control we indicate that the power of the sequentially rejective step down procedure is not higher than that of the closed testing procedure and two procedures are equivalent when we use same critical values for them. We gave simulation results regarding the power of the test intended to compare two stepwise procedures for unbalanced sample sizes. From the simulation results we confirmed that the difference of the power between two procedures is fairly small. The closed testing procedure is accompanied with computational complications compared to the sequentially rejective step down procedure when the number of populations is large. Since the difference of the power between two procedures is small, it seems more appropriate to use the sequentially rejective step down procedure in such cases.

For the all-pairwise multiple comparison we constructed another type of closed testing procedure which enables us to test the intersection of plural mutually disjoint hypotheses at a time and indicated that the power of Holland-Copenhaver's procedure is not higher than that of the proposed closed testing procedure specifying the total number of populations. We gave simulation results regarding the power of the test intended to compare the procedures. Two types of closed testing procedures are uniformly more powerful compared to Holland-Copenhaver's procedure. The power of the proposed closed testing procedure is not higher than that of Ryan-Einot-Gabriel-Welsch's procedure. Although it was difficult to indicate that the power of Holland-Copenhaver's procedure is not higher than that of Ryan-Einot-Gabriel-Welsch's procedure, it was most powerful among three procedures. However, the proposed closed testing procedure is simpler for practical use when the number of populations is large.

There exist other types of stepwise procedures. Dunnett and Tamhane (1992) proposed the step up procedure for the multiple comparison with a control. Dunnett *et al.* (2001) compared the step up procedure and the sequentially rejective step down procedure in terms of the power of the test through simulation. We should clarify theoretical relations regarding power of the test between the step up procedure and the closed testing procedure in the future.

Acknowledgement

The author is deeply grateful to the referee and the editors for their valuable comments and suggestions.

References

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association*, **50**, 1096-1121.

Dunnett, C. W. and Tamhane, A.C. (1991). Step Down Multiple Tests for Comparing Treatments with a Control in Unbalanced One-way Layouts. *Statistics in Medicine*, **10**, 939-947.

Dunnett, C. W. and Tamhane, A.C. (1992). A step-up multiple test procedure. *Journal of the American Statistical Association*, **87**, 162-170.

Dunnett, C. W., Horn, M. and Vollandt, R. (2001). Sample Size Determination in Step-down and Step-up Multiple Tests for Comparing Treatments with a Control. *Journal of Statistical Planning and Inference*, **97**, 367-384.

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.

Hayter, A. J. (1984). A Proof of the Conjecture that the Tukey-Kramer Multiple Comparisons Procedure is Conservative *Annals of Statistics*, **12** No. 1, 61-75.

Holland, B. S. and Copenhaver, M. D. (1987). An Improved Sequentially Rejective Bonferroni Test Procedure. *Biometrics*, **43** No. 2, 417-423.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6** (2), 65-70. *Biometrical Journal*, **49**, No.1, 18-29.

Hsu, J. C. (1996). *Multiple Comparisons.* Boca Raton : Chapman & Hall.

Marcus, R., Peritz, E. and Gabriel, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, **63** (3), 655-660.

Ramsey, P. H. (1978). Power differences between pairwise multiple comparisons. *Journal of the American Statistical Association*, **73**, 479-485.

Ryan, T. A. (1960). Significance tests for multiple comparison of proportions, variances, and other statistics. *Psychological Bulletin*, **57**, 318-328.

Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association*, **81**, 826- 831.

Tukey, J. W. (1953). The problem of multiple comparisons. Unpublished manuscript, Princeton University.

Welsch, R. E. (1977). Stepwise multiple comparison procedures. *Journal of the American Statistical Association*, **72**, 566-575.

Received September 12, 2016
Revised August 7, 2017