STEPWISE MULTIPLE COMPARISON PROCEDURES FOR NORMAL VARIANCES

Imada, Tsunehisa
Department of Management, Tokai University

https://doi.org/10.5109/2547370

出版情報：Bulletin of informatics and cybernetics. 51, pp.1-20, 2019-12. Research Association of Statistical Sciences
バージョン：
権利関係：

STEPWISE MULTIPLE COMPARISON PROCEDURES FOR NORMAL VARIANCES

by

Tsunehisa IMADA

Reprinted from the Bulletin of Informatics and Cybernetics
Research Association of Statistical Sciences, Vol.51

FUKUOKA, JAPAN
2019

Bulletin of Informatics and Cybernetics, Vol. 51, 2019

STEPWISE MULTIPLE COMPARISON PROCEDURES FOR NORMAL VARIANCES

By

Tsunehisa Imada*

Abstract

In this study we discuss stepwise multiple comparison procedures for normal variances intended to obtain higher power compared to the single step procedures proposed by Imada (2018A, 2018B). Specifically, we construct the sequentially rejective step down procedure and the step up procedure for the multiple comparison with a control. Furthermore, we construct the closed testing procedure called Ryan-Einot-Gabriel-Welsch's procedure for the all-pairwise multiple comparison. Finally, we give some numerical results regarding critical values and power of the test intended to compare the procedures.

Key Words and Phrases: Closed testing procedure, Sequentially rejective step down procedure, Step up procedure.

1. Introduction

Assume there are independent normal random variables $X_1, X_2, \ldots, X_K$ and X_k is distributed according to normal $N(\mu_k, \sigma_k^2)$ for $k = 1, 2, \ldots, K$. For testing whether $\mu_1 = \mu_2 = \cdots = \mu_K$ or not by the analysis of variance the assumption $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2$ is necessary. The assumption is also necessary for multiple comparison procedures proposed by Dunnett (1955) and Tukey (1953) for checking specific differences among $\mu_1, \mu_2, \ldots, \mu_K$. When the hypothesis $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2$ is rejected, we occasionally want to find the pair σ_i^2, σ_j^2 satisfying $\sigma_i^2 \neq \sigma_j^2$. Imada (2018A) discussed the multiple comparison with a control for comparing σ_1^2 with $\sigma_2^2, \sigma_3^2, \ldots, \sigma_K^2$ simultaneously and the all-pairwise multiple comparison for $\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2$ based on the single step procedures (cf. Dunnett (1955) and Tukey (1953)). For the multiple comparison with a control Imada (2018A) determined the critical value for pairwise comparison satisfying a specified significance level exactly and formulated the power of the test under a specified alternative hypothesis. For the all-pairwise multiple comparison Imada (2018A) determined two kinds of conservative critical values for pairwise comparison for a specified significance level using Bonferroni's inequality and the improved Bonferroni's inequality respectively and calculated the power of the test by Monte Carlo simulation. Furthermore, Imada (2018B) determined the critical value for pairwise comparison of the all-pairwise multiple comparison satisfying a specified significance level exactly.

In this study we discuss stepwise multiple comparison procedures for normal variances intended to obtain higher power. There are various types of stepwise multiple

* 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

comparison procedures. For the multiple comparison with a control Imada (2017) indicated that the power of the sequentially rejective step down procedure is not higher than that of the closed testing procedure and confirmed that the difference of the power between the two stepwise procedures is fairly small through the simulation results. For the all-pairwise multiple comparison Ryan-Einot-Gabriel-Welsch's procedure (cf. Ryan (1960), Einot and Gabriel (1975) and Welsch (1977)) is the well known closed testing procedure. Imada (2017) 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 (1987)'s sequentially rejective step down procedure is not higher than that of the proposed closed testing procedure specifying the total number of populations. Imada (2017) confirmed that the power of the proposed closed testing procedure is uniformly higher than that of Holland-Copenhaver's procedure and is not higher than that of Ryan-Einot-Gabriel-Welsch's procedure through the simulation results.

In this study we focus on the sequentially rejective step down procedure and the step up procedure for the multiple comparison with a control. For these procedures we determine the critical value at each step of the test for a specified significance level and formulate the power of the test under a specified alternative hypothesis. Next, we construct the closed testing procedure called Ryan-Einot-Gabriel-Welsch's procedure for the all-pairwise multiple comparison. Finally, we give some numerical results regarding critical values and power of the test intended to compare the procedures.

2. Multiple comparison with a control

First, we consider the multiple comparison with a control for comparing σ_1^2 with $\sigma_2^2, \sigma_3^2, \ldots, \sigma_K^2$ simultaneously. For pairwise comparison we consider the one-sided test and the two-sided test. For the one-sided test we set up a null hypothesis and its alternative hypothesis as

$$H_{1,k}^{(1)} : \sigma_1^2 = \sigma_k^2 \quad \text{vs.} \quad H_{1,k}^{(1)A} : \sigma_1^2 < \sigma_k^2 \quad \text{for } k = 2, 3, \ldots, K. \tag{1}$$

For the two-sided test we set up a null hypothesis and its alternative hypothesis as

$$H_{1,k}^{(2)} : \sigma_1^2 = \sigma_k^2 \quad \text{vs.} \quad H_{1,k}^{(2)A} : \sigma_1^2 \neq \sigma_k^2 \quad \text{for } k = 2, 3, \ldots, K. \tag{2}$$

We consider the simultaneous test of $H_{1,2}^{(i)}, H_{1,3}^{(i)}, \ldots, H_{1,K}^{(i)}$ for $i = 1, 2$ using a sample x_{k1}, x_{k2}, …, x_{kn_k} from $N(\mu_k, \sigma_k^2)$ for $k = 1, 2, \ldots, K$.

2.1. Single step procedure

First, we discuss the single step procedure for $H_{1,2}^{(i)}, H_{1,3}^{(i)}, \ldots, H_{1,K}^{(i)}$ for $i = 1, 2$ proposed by Imada (2018A). Letting

$$\bar{x}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_{ki}, \quad \nu_k^2 = \frac{\sum_{i=1}^{n_k} (x_{ki} - \bar{x}_k)^2}{n_k - 1}$$

for $k = 1, 2, \ldots, K$, we use the statistic

$$F_{1,k} = \frac{\nu_k^2}{\nu_1^2}$$

for testing $H_{1,k}^{(i)}$ for $i = 1, 2$. If $n_2, n_3, \ldots, n_K$ are unbalanced, it is preferable to set up an appropriate critical value for each $H_{1,k}^{(i)}$. However, we set up a common critical value for all $H_{1,k}^{(i)}$s for simplicity. First, we consider (1). If $F_{1,k} > c$ for a specified positive critical value c, we reject $H_{1,k}^{(1)}$. Otherwise, we retain $H_{1,k}^{(1)}$. The probability that at least one hypothesis among $H_{1,2}^{(1)}$, $H_{1,3}^{(1)}$, …, $H_{1,K}^{(1)}$ is rejected is

$$P(\max_{2 \leq k \leq K} F_{1,k} > c).$$

We determine c so that

$$P(\max_{2 \leq k \leq K} F_{1,k} > c) = \alpha \tag{3}$$

for a specified significance level α under the assumption that $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2$. (3) is equivalent to

$$P(F_{1,2} \leq c, F_{1,3} \leq c, \ldots, F_{1,K} \leq c) = 1 - \alpha.$$

Imada (2018A) derived

$$P(F_{1,2} \leq c, F_{1,3} \leq c, \ldots, F_{1,K} \leq c) = \int_0^\infty f_1(x_1) \left\{ \prod_{k=2}^{K} \int_0^{c\lambda_{1,k}x_1} f_k(x_k) dx_k \right\} dx_1$$

where $f_k(x_k)$ denotes the probability density function of χ^2-distribution with degrees of freedom $n_k - 1$ for $k = 1, 2, \ldots, K$ and

$$\lambda_{1,k} = \frac{n_k - 1}{n_1 - 1}$$

for $k = 2, 3, \ldots, K$. Next, we consider (2). If $F_{1,k} < c_1$ or $c_2 < F_{1,k}$ for specified critical values c_1, c_2 satisfying $0 < c_1 < c_2$, we reject $H_{1,k}^{(2)}$. Otherwise, we retain $H_{1,k}^{(2)}$. Since

$$F_{1,k} < c_1 \text{ or } c_2 < F_{1,k} \ \Leftrightarrow \ F_{1,k}^{-1} < c_2^{-1} \text{ or } c_1^{-1} < F_{1,k}^{-1},$$

we restrict c_1, c_2 as

$$c_2 = c_1^{-1} = c > 1.$$

Then, we obtain

$$F_{1,k} < c^{-1} \text{ or } c < F_{1,k} \ \Leftrightarrow \ F_{1,k}^{-1} < c^{-1} \text{ or } c < F_{1,k}^{-1}.$$

Furthermore, letting

$$G_{1,k} = \max\{F_{1,k}, F_{1,k}^{-1}\},$$

we obtain

$$F_{1,k} < c^{-1} \text{ or } c < F_{1,k} \ \Leftrightarrow \ G_{1,k} > c.$$

The probability that at least one hypothesis among $H_{1,2}^{(2)}$, $H_{1,3}^{(2)}$, …, $H_{1,K}^{(2)}$ is rejected is

$$1 - P(G_{1,2} \leq c, G_{1,3} \leq c, \ldots, G_{1,K} \leq c).$$

We determine c so that

$$P(G_{1,2} \leq c, G_{1,3} \leq c, \ldots, G_{1,K} \leq c) = 1 - \alpha$$

for a specified significance level α under the assumption that $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2$. Imada (2018A) derived

$$P(G_{1,2} \le c, G_{1,3} \le c, \ldots, G_{1,K} \le c) = \int_0^\infty f_1(x_1) \left\{ \prod_{k=2}^{K} \int_{c^{-1}\lambda_{1,k}x_1}^{c\lambda_{1,k}x_1} f_k(x_k)dx_k \right\} dx_1.$$

Next, we consider the power of the test. First, we consider the power of the test for (1). Assume

$$\sigma_1^2 = \gamma_{1,2}\sigma_2^2 = \gamma_{1,3}\sigma_3^2 = \cdots = \gamma_{1,l}\sigma_l^2 \text{ and } \sigma_1^2 = \sigma_m^2 \text{ for } m = l+1, l+2, \ldots, K \tag{4}$$

where $0 < \gamma_{1,2} < 1, 0 < \gamma_{1,3} < 1, \ldots, 0 < \gamma_{1,l} < 1$. We focus on the all pairs power defined by Ramsey (1978). If $l = 2$, the power of the test under (4) is

$$P(F_{1,2} > c) = \int_{c\gamma_{1,2}}^\infty f_{1,2}(x)dx$$

where $f_{1,2}(v)$ is the probability density function of F-distribution with degrees of freedom $(n_2 - 1, n_1 - 1)$. If $l > 2$, the power of the test under (4) is

$$P(F_{1,i} > c \text{ for } i = 2, 3, \ldots, l) = \int_0^\infty f_1(x_1) \left\{ \prod_{i=2}^{l} \int_{c\lambda_{1,i}\gamma_{1,i}x_1}^\infty f_i(x_i)dx_i \right\} dx_1$$

by Imada (2018A). Next, we consider the power of the test for (2). Assume

$$\sigma_1^2 = \gamma_{1,2}\sigma_2^2 = \gamma_{1,3}\sigma_3^2 = \cdots = \gamma_{1,l}\sigma_l^2 \text{ and } \sigma_1^2 = \sigma_m^2 \text{ for } m = l+1, l+2, \ldots, K \tag{5}$$

where $\gamma_{1,2} \ne 1, \gamma_{1,3} \ne 1, \ldots, \gamma_{1,l} \ne 1$. If $l = 2$, the power of the test under (5) is

$$P(G_{1,2} > c) = 1 - \int_{c^{-1}\gamma_{1,2}}^{c\gamma_{1,2}} f_{1,2}(v)dv.$$

If $l > 2$, the power of the test under (5) is

$$P(G_{1,i} > c \text{ for } i = 2, 3, \ldots, l) = \int_0^\infty f_1(x_1) \prod_{i=2}^{l} \left\{ 1 - \int_{c^{-1}\lambda_{1,i}\gamma_{1,i}x_1}^{c\lambda_{1,i}\gamma_{1,i}x_1} f_i(x_i)dx_i \right\} dx_1$$

by Imada (2018A).

2.2. Sequentially rejective step down procedure

Dunnett and Tamhane (1991) discussed a step down procedure for the multiple comparison with a control for normal means. It is called the sequentially rejective step down procedure. In this Section we construct the sequentially rejective step down procedure for (1) and (2). First, we consider (1). We determine c_1 as the minimum c satisfying

$$P(F_{1,k} > c) \le \alpha$$

for all $k = 2, 3, \ldots, K$ under the assumption that $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2$. Here

$$P(F_{1,k} > c) = \int_c^\infty f_{1,k}(v)dv$$

where $f_{1,k}(v)$ is the probability density function of F-distribution with degrees of freedom (n_k-1, n_1-1). Next, we determine c_m $(m = 2, 3, \ldots, K-1)$ as the minimum c satisfying

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

for $l_1, l_2, \ldots, l_m$ chosen from $2, 3, \ldots, K$ arbitrarily under the assumption that $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2$. Here

$$P(\max_{k=l_1,l_2,\ldots,l_m} F_{1,k} > c) = 1 - \int_0^\infty f_1(x_1) \left\{ \prod_{j=1}^m \int_0^{c\lambda_{1,l_j} x_1} f_{l_j}(x_{l_j}) dx_{l_j} \right\} dx_1.$$

Apparently $c_{K-1} > c_{K-2} > \cdots > c_1$. Arranging $F_{1,2}, F_{1,3}, \ldots, F_{1,K}$ in order of a size of value, assume

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

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

Step 1.

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

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

Step 2.

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

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

⋮

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

The sequentially rejective step down procedure for (2) is similarly constructed using $G_{1,2}, G_{1,3}, \ldots, G_{1,K}$ instead of $F_{1,2}, F_{1,3}, \ldots, F_{1,K}$.

Next, we consider the power of the test. First, we introduce notations which were used by Hayter and Tamhane (1991) and Dunnett *et al.* (2001). Let $W_1, W_2, \ldots, W_l$ be statistics. Let $b_1, b_2, \ldots, b_l$ be constants satisfying $b_1 < b_2 < \cdots < b_l$. Calculating $W_1, W_2, \ldots, W_l$ based on observations, we assume $W_{(1)} \leq W_{(2)} \leq \cdots \leq W_{(l)}$. If $W_{(1)} > b_1, W_{(2)} > b_2, \ldots, W_{(l)} > b_l$, we denote

$$(W_1, W_2, \ldots, W_l) > (b_1, b_2, \ldots, b_l). \tag{6}$$

If $W_{(1)} \leq b_1, W_{(2)} \leq b_2, \ldots, W_{(l)} \leq b_l$, we denote

$$(W_1, W_2, \ldots, W_l) \leq (b_1, b_2, \ldots, b_l). \tag{7}$$

The events (6) and (7) are recursively divided into plural disjoint events. We discuss the process only for (6). The process for (7) is similar. Specifically under (6) there are l kinds of ranges regarding the value of W_l as follows.

$$W_l > b_l,\ b_l > W_l > b_{l-1},\ b_{l-1} > W_l > b_{l-2},\ \ldots,\ b_2 > W_l > b_1.$$

Corresponding to each range of W_l the ranges of $W_1, W_2, \ldots, W_{l-1}$ are determined as follows.

$$
\begin{array}{lll}
W_l > b_l & \Rightarrow & (W_1, W_2, \ldots, W_{l-1}) > (b_1, b_2, \ldots, b_{l-1}), \\
b_l > W_l > b_{l-1} & \Rightarrow & (W_1, W_2, \ldots, W_{l-1}) > (b_1, b_2, \ldots, b_{l-2}, b_l), \\
b_{l-1} > W_l > b_{l-2} & \Rightarrow & (W_1, W_2, \ldots, W_{l-1}) > (b_1, b_2, \ldots, b_{l-3}, b_{l-1}, b_l), \\
 & \vdots & \\
b_2 > W_l > b_1 & \Rightarrow & (W_1, W_2, \ldots, W_{l-1}) > (b_2, b_3, \ldots, b_l).
\end{array}
$$

By repeating the similar step each event is divided into plural disjoint events. Finally the range of each of $W_1, W_2, \ldots, W_l$ is determined in each event.

We consider the power of the test for (1). The all-pairs power of (4) by the step down procedure is the probability that $H_{12}^{(1)}$, $H_{13}^{(1)}$, …, $H_{1l}^{(1)}$ are rejected till up to Step $K-1$. Therefore, if $l = K$, the power is given by

$$
P((F_{1,2}, F_{1,3}, \ldots, F_{1,K}) > (c_1, c_2, \ldots, c_{K-1})). \tag{8}
$$

Next we consider the power for $l < K$. When $H_{1,2}$, $H_{1,3}$, …, $H_{1,l}$ are rejected till up to Step $K-1$, other hypotheses $H_{1,l'}(l' \geq l+1)$ also may be rejected. Specifically following disjoint events E_0,E_1,E_2,…,E_{K-l-1} can occur.

E_0 : Non of $H_{1,l+1}^{(1)}, H_{1,l+2}^{(1)}, \ldots, H_{1,K}^{(1)}$ is rejected.

E_1 : One of $H_{1,l+1}^{(1)}, H_{1,l+2}^{(1)}, \ldots, H_{1,K}^{(1)}$ is rejected.

E_2 : Two of $H_{1,l+1}^{(1)}, H_{1,l+2}^{(1)}, \ldots, H_{1,K}^{(1)}$ are rejected.

E_3 : Three of $H_{1,l+1}^{(1)}, H_{1,l+2}^{(1)}, \ldots, H_{1,K}^{(1)}$ are rejected.

$\vdots$

E_{K-l} : All $H_{1,l+1}^{(1)}, H_{1,l+2}^{(1)}, \ldots, H_{1,K}^{(1)}$ are rejected.

If $H_{1,i_1}^{(1)}, H_{1,i_2}^{(1)}, \ldots, H_{1,i_m}^{(1)}$ are rejected and other hypotheses are retained in the step down test, $H_{1,i_1}^{(1)}, H_{1,i_2}^{(1)}, \ldots, H_{1,i_m}^{(1)}$ are rejected till Step m and other hypotheses are retained at Step $m+1$. Therefore the event is expressed by

$$
(F_{1,i_1}, F_{1,i_2}, \ldots, F_{1,i_m}) > (c_{K-m}, c_{K-m+1}, \ldots, c_{K-1})
$$

and

$$
F_{1,j} \leq c_{K-m-1} \text{ for all } j \neq i_1, i_2, \ldots, i_m.
$$

Therefore the power is given by

$$
\begin{aligned}
& P((F_{1,2}, F_{1,3}, \ldots, F_{1,l}) > (c_{K-l+1}, c_{K-l+2}, \ldots, c_{K-1}), \\
& \qquad F_{1,m} \leq c_{K-l} \text{ for all } m \neq 2, 3, \ldots, l) \\
& + \sum_{m_1 \neq 2,3,\ldots,l} P((F_{1,2}, F_{1,3}, \ldots, F_{1,l}, F_{1m_1}) > (c_{K-l}, c_{K-l+1}, \ldots, c_{K-1}), \\
& \qquad F_{1,m} \leq c_{K-l-1} \text{ for all } m \neq 2, 3, \ldots, l, m_1) \\
& + \sum_{m_1, m_2 \neq 2,3,\ldots,l} P((F_{1,2}, F_{1,3}, \ldots, F_{1,l}, F_{1,m_1}, F_{1,m_2}) > (c_{K-l-1}, c_{K-l}, \ldots, c_{K-1}), \\
& \qquad F_{1,m} \leq c_{K-l-2} \text{ for all } m \neq 2, 3, \ldots, l, m_1, m_2) \\
& + \cdots\cdots + P((F_{1,2}, F_{1,3}, \ldots, F_{1,K}) > (c_1, c_2, \ldots, c_{K-1})).
\end{aligned} \tag{9}
$$

Each probability in (9) is expressed as the sum of multiple integration. The specific expressions of (9) for $K = 3, 4$ are given in Appendix. (Since we give numerical results for $K = 3, 4, 5$ in Section 4, we should also give the specific formulae for $K = 5$. However, they need many pages.) The power of the test for (2) under (5) is similarly expressed using $G_{1,2}, G_{1,3}, \ldots, G_{1,K}$ instead of $F_{1,2}, F_{1,3}, \ldots, F_{1,K}$.

2.3. Step up procedure

Dunnett and Tamhane (1992) discussed a step up procedure for the multiple comparison with a control for normal means. In this Section we construct the step up procedure for (1) and (2). First, we consider (1). Assuming $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2$, we determine the critical values $c_1, c_2, \ldots, c_{K-1}$ recursively as follows. First, we determine c_1 as the minimum c satisfying $P(F_{1,k} \leq c) \geq 1 - \alpha$ for $k = 2, 3, \ldots, K$. Next, we determine c_2 as the minimum c satisfying

$$P((F_{1,l_1}, F_{1,l_2}) \leq (c_1, c)) \geq 1 - \alpha$$

for l_1, l_2 chosen from $2, 3, \ldots, K$ arbitrarily. We repeat similar steps. Specifically, we determine c_m $(m = 2, 3, \ldots, K-1)$ as the minimum c satisfying

$$P((F_{1,l_1}, F_{1,l_2}, \ldots, F_{1,l_m}) \leq (c_1, c_2, c_3, \ldots, c_{m-1}, c)) \geq 1 - \alpha$$

for $l_1, l_2, \ldots, l_m$ chosen from $1, 2, \ldots, K$ arbitrarily. The condition

$$c_1 < c_2 < c_3 < \cdots < c_{K-1} \tag{10}$$

is necessary for constructing the step up procedure. (10) can be mathematically proved only for $K = 2, 3$. However, (10) is true for $K \leq 5$ in the numerical results in Section 5. We give the specific formulae of $P((F_{1,2}, F_{1,3}) \leq (c_1, c_2))$ and $P((F_{1,2}, F_{1,3}, F_{1,4}) \leq (c_1, c_2, c_3))$ in the Appendix. Arranging $F_{1,2}, F_{1,3}, \ldots, F_{1,K}$ in order of a size of value, assume

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

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

Step 1.

Case 1. If $F_{(1)} > c_1$, we reject $H_{(1)}^{(1)}, H_{(2)}^{(1)}, \ldots, H_{(K-1)}^{(1)}$ and stop the test.

Case 2. If $F_{(1)} \leq c_1$, we retain $H_{(1)}^{(1)}$ and go to the next step.

Step 2.

Case 1. If $F_{(2)} > c_2$, we reject $H_{(2)}^{(1)}, H_{(3)}^{(1)}, \ldots, H_{(K-1)}^{(1)}$ and stop the test.

Case 2. If $F_{(2)} \leq c_2$, we retain $H_{(2)}^{(1)}$ and go to the next step.

⋮

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

The step up procedure for $H_{1,2}^{(2)}, H_{1,3}^{(2)}, \ldots, H_{1,K}^{(2)}$ is similarly constructed using $G_{1,2}$, $G_{1,3}$, …, $G_{1,K}$ instead of $F_{1,2}, F_{1,3}, \ldots, F_{1,K}$.

Next, we consider the power of the test. First, we consider the power of the test for (1). The all-pairs power of (4) by the step up procedure is the probability that $H_{12}^{(1)}$,

$H_{13}^{(1)}$, …, $H_{1l}^{(1)}$ are rejected till up to Step $K-1$. Therefore, if $l = K$, the power of (4) is given by $P(\min\{F_{1,2}, F_{1,3}, \ldots, F_{1,K}\} > c_1)$. Next, we assume $l < K$. When $H_{1,2}^{(1)}$, $H_{1,3}^{(1)}$, …, $H_{1,l}^{(1)}$ are rejected till up to Step $K-1$, other hypotheses $H_{1,l'}(l' \geq l+1)$ also may be rejected. Specifically, the disjoint events E_0,E_1,E_2,…,E_{K-l} defined in Subsection 2.2 can occur. If $H_{1,i_1}^{(1)}, H_{1,i_2}^{(1)}, \ldots, H_{1,i_m}^{(1)}$ are retained and other hypotheses are rejected in the step up test, $H_{1,i_1}^{(1)}, H_{1,i_2}^{(1)}, \ldots, H_{1,i_m}^{(1)}$ are retained till Step m and other hypotheses are rejected at Step $m+1$. Therefore, the power of (4) is given by

$$
\begin{aligned}
&P((F_{1,l+1}, \ldots, F_{1,K}) \leq (c_1, c_2, \ldots, c_{K-l}), \min\{F_{1,2}, F_{1,3}, \ldots, F_{1,l}\} > c_{K-l+1}) \\
&+ \sum_{k_1=1}^{K-l} P((F_{1,l+1}, \ldots, \check{F}_{1,l+k_1}, \ldots, F_{1,K}) \leq (c_1, c_2, \ldots, c_{K-l-1}), \\
&\min\{F_{1,2}, F_{1,3}, \ldots, F_{1,l}, F_{1,l+k_1}\} > c_{K-l}) \\
&+ \sum_{l+1 \leq k_1 < k_2 \leq K} P((F_{1,l+1}, \ldots, \check{F}_{1,l+k_1}, \ldots, \check{F}_{1,l+k_2}, \ldots, F_{1,K}) \leq (c_1, c_2, \ldots, c_{K-l-2}), \\
&\min\{F_{1,2}, F_{1,3}, \ldots, F_{1,l}, F_{1,l+k_1}, F_{1,l+k_2}\} > c_{K-l-1})) \\
&+ \cdots\cdots + \sum_{l=1}^{K-l} P(\min\{F_{1,2}, F_{1,3}, \ldots, F_{1,K}\} > c_1).
\end{aligned} \tag{11}
$$

Here the notation ˇ means omitting. Each probability in (11) is expressed as the sum of multiple integration. The specific formulae for $K = 3, 4$ are given in Appendix. The power of the test for (2) under (5) is similarly expressed using $G_{1,2}, G_{1,3}, \ldots, G_{1,K}$ instead of $F_{1,2}, F_{1,3}, \ldots, F_{1,K}$.

3. All-pairwise multiple comparison for normal variances

We consider the all-pairwise multiple comparison for $\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2$. Intended to compare σ_k^2 and σ_l^2 for $1 \leq k < l \leq K$ we set up a null hypothesis and its alternative hypothesis as

$$H_{k,l} : \sigma_k^2 = \sigma_l^2 \quad \text{vs.} \quad H_{k.l}^A : \sigma_k^2 \neq \sigma_l^2$$

and consider the simultaneous test of all $H_{k,l}$s. We use the statistic

$$F_{k,l} = \frac{\nu_l^2}{\nu_k^2}$$

for testing $H_{k,l}$.

3.1. Single step procedure

We consider the single step procedure for $H_{k,l}$s discussed by Imada (2018A, 2018B). We specify a critical value $c(> 1)$. If $F_{k,l} < c^{-1}$ or $c < F_{k,l}$, we reject $H_{k,l}$. Otherwise, we retain $H_{k,l}$. Letting

$$G_{k,l} = \max\{F_{k,l}, F_{k,l}^{-1}\},$$

we obtain

$$F_{k,l} < c^{-1} \text{ or } c < F_{k,l} \;\Leftrightarrow\; G_{k,l} > c.$$

The probability that at least one hypothesis among $H_{k,l}$s is rejected is

$$P(\max_{1\le k<l\le K} G_{k,l} > c).$$

We want to determine c so that

$$P(\max_{1\le k<l\le K} G_{k,l} > c) = \alpha$$

for a specified significance level α under the assumption that $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2$. Letting

$$\lambda_{k_1,k_2} = \frac{n_{k_2}-1}{n_{k_1}-1}$$

for each pair (k_1, k_2) chosen from $1, 2, \ldots, K$, Imada (2018B) derived

$$P(\max_{1\le k<l\le K} G_{k,l} > c)$$

$$= 1 - \sum_{k_1,k_2} \int_0^\infty \int_{\lambda_{k_1,k_2}x_1}^{c\lambda_{k_1,k_2}x_1} \left\{ \prod_{l\neq k_1,k_2} \int_{\lambda_{k_1,l}x_1}^{\lambda_{k_2,l}x_2} f_l(x)dx \right\} f_{k_2}(x_2)dx_2 f_{k_1}(x_1)dx_1.$$

It is difficult to formulate the power of the single step procedure under a specified alternative hypothesis. We calculate the power using Monte Carlo simulation.

3.2. Closed testing procedure called Ryan-Einot-Gabriel-Welsch's procedure

Let I_s be an arbitrary subset of $I = \{1, 2, \ldots, K\}$ with the cardinal number $\sharp(I_s) \ge 2$. Letting $I_s = \{s_1, s_2, \ldots, s_k\}$ $(s_1 < s_2 < \cdots < s_k)$, define the hypothesis H_{I_s} as

$$H_{I_s} : \sigma_{s_1}^2 = \sigma_{s_2}^2 = \cdots = \sigma_{s_k}^2.$$

We obtain

$$H_{I_s} = \cap_{s_i,s_j\in I_s, s_i<s_j} H_{s_i,s_j}$$

using the notation defined in Subsection 3.1. Letting H be the family of hypotheses consisting of all H_{I_s}s and all sorts of intersections of plural hypotheses H_{I_s}s, H is closed. Each hypothesis in H is equal to single H_{I_s} or $H_{I_{s_1}} \cap H_{I_{s_2}} \cap \cdots \cap H_{I_{s_k}}$ where $I_{s_1}, I_{s_2}, \ldots, I_{s_k}$ are disjoint. We construct the closed testing procedure called Ryan-Einot-Gabriel-Welsch's procedure for H. For testing H_{I_s} we use the statistic

$$G_{I_s} = \max_{s_i,s_j\in I_s, s_i<s_j} G_{s_i,s_j}$$

and determine the critical value c_{I_s} so that

$$P(G_{I_s} > c_{I_s}) = \alpha. \tag{12}$$

If $G_{I_s} > c_{I_s}$, we reject H_{I_s}. Otherwise, we retain H_{I_s}. If $n_1 = n_2 = \cdots = n_K$ and $\sharp(I_{s_1}) = \sharp(I_{s_2})$, $c_{I_{s_1}} = c_{I_{s_2}}$. Therefore, If $n_1 = n_2 = \cdots = n_K$, c_{I_s} satisfying (12) is denoted by $c_{\sharp(I_s)}$.

Next, we discuss how to test $H_{I_{s_1}} \cap H_{I_{s_2}} \cap \cdots \cap H_{I_{s_k}}$ where $I_{s_1}, I_{s_2}, \ldots, I_{s_k}$ are disjoint. Let $M = \sharp(I_{s_1}) + \sharp(I_{s_2}) + \cdots + \sharp(I_{s_k})$. For $l = 1, 2, \ldots, k$ we determine $c_{I_{s_l},M}$ so that

$$P(F_{I_{s_l}} > c_{I_{s_l},M}) = 1 - (1-\alpha)^{\frac{\sharp(I_{s_l})}{M}}.$$

If $n_1 = n_2 = \cdots = n_K$, $c_{I_{s_l},M}$ is denoted by $c_{\sharp(I_{s_l}),M}$. Intended to test $H_{I_{s_1}} \cap H_{I_{s_2}} \cap \cdots \cap H_{I_{s_k}}$ we set up the critical value $c_{I_{s_l},M}$ for testing $H_{I_{s_l}}$ for $l = 1, 2, \ldots, k$. If $F_{I_{s_l}} > c_{I_{s_l},M}$ for at least one l, $H_{I_{s_1}} \cap H_{I_{s_2}} \cap \cdots \cap H_{I_{s_k}}$ is rejected. Otherwise, it is retained. Then, the probability that $H_{I_{s_1}} \cap H_{I_{s_2}} \cap \cdots \cap H_{I_{s_k}}$ is rejected when $H_{I_{s_1}} \cap H_{I_{s_2}} \cap \cdots \cap H_{I_{s_k}}$ is true is not greater than α. Because the probability that $H_{I_{s_1}} \cap H_{I_{s_2}} \cap \cdots \cap H_{I_{s_k}}$ is rejected is

$$\begin{aligned} P(F_{I_{s_l}} > c_{I_{s_l},M} \text{ for some } l) &= 1 - P(F_{I_{s_l}} \le c_{I_{s_l},M} \text{ for } l = 1, 2, \ldots, k) \\ &= 1 - \prod_{l=1}^{k} P(F_{I_{s_l}} \le c_{I_{s_l},M}) \\ &= \alpha. \end{aligned}$$

We specified the way to test each hypothesis in H satisfying the specified significance level α. We test the hypotheses in H hierarchically. Specifically, if a hypothesis and all hypotheses inducing it are rejected, we reject the hypothesis. Otherwise we retain it.

It is difficult to formulate the power of the closed testing procedure under a specified alternative hypothesis. We calculate the power using Monte Carlo simulation.

4. Numerical examples

In this Section we give some numerical examples regarding critical values and power of the test intended to compare the procedures.

Let $K = 3, 4, 5$ and $\alpha = 0.05$. We set up two types of sample sizes for $K = 3, 4, 5$, respectively. Specifically, if $K = 3$,

$$\text{Sam.1} : (20, 20, 20), \ \text{Sam.2} : (15, 30, 15).$$

If $K = 4$,

$$\text{Sam.1} : (20, 20, 20, 20), \ \text{Sam.2} : (15, 25, 15, 25).$$

If $K = 5$,

$$\text{Sam.1} : (20, 20, 20, 20, 20), \ \text{Sam.2} : (15, 25, 20, 25, 15).$$

First, we consider the multiple comparison with a control. O-S means the one-sided test (1) and T-S means the two-sided test (2). Furthermore, SS, SD and SU mean the single step procedure, the sequentially rejective step down procedure and the step up procedure, respectively. Table 1 gives critical values of SD and SU for O-S and T-S. The critical value of SS is equal to c_{K-1} of SD.

Table 1: Critical values of the sequentially rejective step down procedure and the step up procedure

			O-S				T-S			
			c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4
$K=3$	Sam.1	SD	2.169	2.444	–	–	2.527	2.854	–	–
		SU	2.169	2.465	–	–	2.527	2.875	–	–
	Sam.2	SD	2.484	2.725	–	–	2.979	3.171	–	–
		SU	2.484	2.740	–	–	2.979	3.180	–	–
$K=4$	Sam.1	SD	2.169	2.444	2.602	–	2.527	2.854	3.052	–
		SU	2.169	2.465	2.620	–	2.527	2.875	3.060	–
	Sam.2	SD	2.484	2.753	2.899	–	2.979	3.216	3.362	–
		SU	2.484	2.770	2.905	–	2.979	3.218	3.364	–
$K=5$	Sam.1	SD	2.169	2.444	2.602	2.713	2.527	2.854	3.052	3.194
		SU	2.169	2.465	2.620	2.718	2.527	2.875	3.060	3.205
	Sam.2	SD	2.484	2.790	2.934	3.032	2.979	3.293	3.425	3.523
		SU	2.484	2.815	2.942	3.034	2.979	3.300	3.427	3.525

Next, we consider the power of the test. Let γ be a positive constant which is less than 1. For $K = 3$ we set up two cases of alternative hypotheses as follows.

$$\begin{array}{ll} \text{Case 1.} & \sigma_1^2 = \gamma\sigma_2^2 = \sigma_3^2, \\ \text{Case 2.} & \sigma_1^2 = \gamma\sigma_2^2 = \gamma\sigma_3^2. \end{array}$$

For $K = 4$ we set up three cases of alternative hypotheses as follows.

$$\begin{array}{ll} \text{Case 1.} & \sigma_1^2 = \gamma\sigma_2^2 = \sigma_3^2 = \sigma_4^2, \\ \text{Case 2.} & \sigma_1^2 = \gamma\sigma_2^2 = \gamma\sigma_3^2 = \sigma_4^2, \\ \text{Case 3.} & \sigma_1^2 = \gamma\sigma_2^2 = \gamma\sigma_3^2 = \gamma\sigma_4^2. \end{array}$$

For $K = 5$ we set up four cases of alternative hypotheses as follows.

$$\begin{array}{ll} \text{Case 1.} & \sigma_1^2 = \gamma\sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2, \\ \text{Case 2.} & \sigma_1^2 = \gamma\sigma_2^2 = \gamma\sigma_3^2 = \sigma_4^2 = \sigma_5^2, \\ \text{Case 3.} & \sigma_1^2 = \gamma\sigma_2^2 = \gamma\sigma_3^2 = \gamma\sigma_4^2 = \sigma_5^2, \\ \text{Case 4.} & \sigma_1^2 = \gamma\sigma_2^2 = \gamma\sigma_3^2 = \gamma\sigma_4^2 = \gamma\sigma_5^2. \end{array}$$

In Case 1 the power is the probability that H_{12} is rejected. In Case 2 the power is the probability that H_{12},H_{13} are rejected. In Case 3 the power is the probability that H_{12},H_{13},H_{14} are rejected. In Case 4 the power is the probability that H_{12},H_{13},H_{14},H_{15} are rejected. Tables 2 to 4 give the power of SS, SD and SU for O-S and T-S when $\gamma = 0.75, 0.50, 0.25$. SD and SU are uniformly more powerful compared to SS. Although the power of SS remarkably decreases as the number of hypotheses which should be rejected increases for each γ, the power of SD and SU is comparatively stable independently of the number of hypotheses which should be rejected. In each case the differences of the power between SD and SU are not remarkably large.

Table 2: Power for $K = 3$

		γ	O-S SS	SD	SU	T-S SS	SD	SU
Case 1	Sam.1	0.75	0.098	0.249	0.246	0.056	0.156	0.155
		0.50	0.333	0.854	0.851	0.223	0.768	0.764
		0.25	0.854	1.000	1.000	0.765	1.000	1.000
	Sam.2	0.75	0.079	0.203	0.201	0.045	0.128	0.127
		0.50	0.275	0.814	0.812	0.183	0.712	0.711
		0.25	0.814	1.000	1.000	0.712	1.000	1.000
Case 2	Sam.1	0.75	0.038	0.179	0.189	0.018	0.104	0.103
		0.50	0.189	0.829	0.833	0.109	0.725	0.724
		0.25	0.757	1.000	1.000	0.634	1.000	1.000
	Sam.2	0.75	0.039	0.145	0.150	0.020	0.080	0.081
		0.50	0.162	0.733	0.736	0.099	0.597	0.599
		0.25	0.669	0.999	0.999	0.548	0.998	0.998

Table 3: Power for $K = 4$

		γ	O-S SS	SD	SU	T-S SS	SD	SU
Case 1	Sam.1	0.75	0.077	0.209	0.206	0.041	0.125	0.125
		0.50	0.286	0.822	0.818	0.183	0.721	0.720
		0.25	0.822	1.000	1.000	0.719	1.000	1.000
	Sam.2	0.75	0.067	0.173	0.173	0.039	0.109	0.109
		0.50	0.237	0.764	0.762	0.157	0.658	0.658
		0.25	0.763	1.000	1.000	0.658	1.000	1.000
Case 2	Sam.1	0.75	0.028	0.126	0.128	0.012	0.068	0.069
		0.50	0.153	0.755	0.752	0.084	0.632	0.631
		0.25	0.710	1.000	1.000	0.575	1.000	1.000
	Sam.2	0.75	0.030	0.107	0.108	0.016	0.061	0.062
		0.50	0.132	0.651	0.648	0.081	0.527	0.528
		0.25	0.612	0.999	0.999	0.492	0.997	0.997
Case 3	Sam.1	0.75	0.015	0.117	0.129	0.006	0.062	0.070
		0.50	0.101	0.768	0.778	0.051	0.642	0.657
		0.25	0.631	1.000	1.000	0.484	1.000	1.000
	Sam.2	0.75	0.018	0.099	0.106	0.009	0.052	0.055
		0.50	0.093	0.661	0.671	0.054	0.515	0.522
		0.25	0.544	0.999	0.999	0.419	0.998	0.998

Table 4: Power for $K = 5$

		γ	O-S SS	SD	SU	T-S SS	SD	SU
Case 1	Sam.1	0.75	0.065	0.184	0.203	0.033	0.106	0.123
		0.50	0.256	0.798	0.818	0.158	0.686	0.710
		0.25	0.798	1.000	1.000	0.685	1.000	1.000
	Sam.2	0.75	0.057	0.152	0.165	0.032	0.093	0.102
		0.50	0.211	0.733	0.753	0.137	0.621	0.642
		0.25	0.733	1.000	1.000	0.620	0.999	1.000
Case 2	Sam.1	0.75	0.023	0.101	0.101	0.010	0.051	0.052
		0.50	0.132	0.709	0.705	0.070	0.574	0.574
		0.25	0.676	1.000	1.000	0.533	0.999	0.999
	Sam.2	0.75	0.025	0.090	0.090	0.013	0.050	0.051
		0.50	0.118	0.624	0.623	0.070	0.495	0.494
		0.25	0.597	0.999	0.999	0.470	0.998	0.998
Case 3	Sam.1	0.75	0.012	0.079	0.084	0.005	0.040	0.042
		0.50	0.085	0.676	0.679	0.042	0.543	0.543
		0.25	0.594	1.000	1.000	0.441	0.999	0.999
	Sam.2	0.75	0.015	0.074	0.076	0.008	0.039	0.040
		0.50	0.083	0.594	0.590	0.047	0.451	0.453
		0.25	0.525	0.999	0.999	0.396	0.999	0.999
Case 4	Sam.1	0.75	0.008	0.084	0.096	0.003	0.042	0.050
		0.50	0.061	0.716	0.732	0.029	0.579	0.600
		0.25	0.532	1.000	1.000	0.379	1.000	1.000
	Sam.2	0.75	0.010	0.073	0.082	0.005	0.037	0.040
		0.50	0.060	0.602	0.618	0.033	0.449	0.463
		0.25	0.448	0.999	0.999	0.327	0.998	0.998

Next, we consider the all-pairwise comparison. SS and CT mean the single step procedure and the closed testing procedure, respectively. Tables 5 to 10 give critical values of CT for $K = 3,4,5$. The critical value of SS is equal to c_K for balanced sample sizes and is equal to $c_{\{1,2,\ldots,K\}}$ for unbalanced sample sizes.

Table 5 : Critical values of CT for $K = 3$ and Sam.1

c_3	c_2
3.037	2.527

Table 6 : Critical values of CT for $K = 3$ and Sam.2

$c_{\{1,2,3\}}$	$c_{\{1,2\}}$	$c_{\{1,3\}}$
3.317	2.741	2.979

($c_{\{1,2\}} = c_{\{2,3\}}$)

Table 7 : Critical values of CT for $K = 4$ and Sam.1

c_4	c_3	$c_{2,4}$	c_2
3.393	3.037	2.895	2.527

Table 8 : Critical values of CT for $K = 4$ and Sam.2

$c_{\{1,2,3,4\}}$	$c_{\{1,2,3\}}$	$c_{\{1,2,4\}}$	$c_{\{1,2\},4}$	$c_{\{1,3\},4}$	$c_{\{2,4\},4}$	$c_{\{1,2\}}$	$c_{\{1,3\}}$	$c_{\{2,4\}}$
3.614	3.376	3.030	3.257	3.505	2.558	2.789	2.979	2.270

$(c_{\{1,2,3\}} = c_{\{1,3,4\}},\ c_{\{1,2,4\}} = c_{\{2,3,4\}},\ c_{\{1,2\},4} = c_{\{1,4\},4} = c_{\{2,3\},4} = c_{\{3,4\},4},$
$c_{\{1,2\}} = c_{\{1,4\}} = c_{\{2,3\}} = c_{\{3,4\}})$

Table 9 : Critical values of CT for $K = 5$ and Sam.1

c_5	c_4	$c_{3,5}$	$c_{2,5}$	c_3	$c_{2,4}$	c_2
3.659	3.393	3.340	3.019	3.037	2.895	2.527

Table 10 : Critical values of CT for $K = 5$ and Sam.2

$c_{\{1,2,3,4,5\}}$	$c_{\{1,2,3,4\}}$	$c_{\{1,2,3,5\}}$	$c_{\{1,2,4,5\}}$
3.881	3.396	3.716	3.614

$(c_{\{1,2,3,4\}} = c_{\{2,3,4,5\}},\ c_{\{1,2,3,5\}} = c_{\{1,3,4,5\}})$

$c_{\{1,2,3\},5}$	$c_{\{1,2,4\},5}$	$c_{\{1,2,5\},5}$	$c_{\{1,3,5\},5}$	$c_{\{2,3,4\},5}$
3.484	3.345	3.758	3.892	3.054

$(c_{\{1,2,3\},5} = c_{\{1,3,4\},5} = c_{\{2,3,5\},5} = c_{\{3,4,5\},5},\ c_{\{1,2,4\},5} = c_{\{2,4,5\},5},\ c_{\{1,2,5\},5} = c_{\{1,4,5\},5})$

$c_{\{1,2\},5}$	$c_{\{1,3\},5}$	$c_{\{1,5\},5}$	$c_{\{2,3\},5}$	$c_{\{2,4\},5}$
3.417	3.518	3.685	2.756	2.653

$(c_{\{1,2\},5} = c_{\{1,4\},5} = c_{\{2,5\},5} = c_{\{4,5\},5},\ c_{\{1,3\},5} = c_{\{3,5\},5},\ c_{\{2,3\},5} = c_{\{3,4\},5})$

$c_{\{1,2,3\}}$	$c_{\{1,2,4\}}$	$c_{\{1,2,5\}}$	$c_{\{1,3,5\}}$	$c_{\{2,3,4\}}$
3.151	3.031	3.376	3.493	2.794

$(c_{\{1,2,3\}} = c_{\{1,3,4\}} = c_{\{2,3,5\}} = c_{\{3,4,5\}},\ c_{\{1,2,4\}} = c_{\{2,4,5\}},\ c_{\{1,2,5\}} = c_{\{1,4,5\}})$

$c_{\{1,2\},4}$	$c_{\{1,3\},4}$	$c_{\{1,5\},4}$	$c_{\{2,3\},4}$	$c_{\{2,4\},4}$
3.257	3.351	3.505	2.654	2.558

$(c_{\{1,2\},4} = c_{\{1,4\},4} = c_{\{2,5\},4} = c_{\{4,5\},4},\ c_{\{1,3\},4} = c_{\{3,5\},4},\ c_{\{2,3\},4} = c_{\{3,4\},4})$

$c_{\{1,2\}}$	$c_{\{1,3\}}$	$c_{\{1,5\}}$	$c_{\{2,3\}}$	$c_{\{2,4\}}$
2.789	2.861	2.979	2.346	2.270

$(c_{\{1,2\}} = c_{\{1,4\}} = c_{\{2,5\}} = c_{\{4,5\}},\ c_{\{1,3\}} = c_{\{3,5\}},\ c_{\{2,3\}} = c_{\{3,4\}})$

Next, we consider the power of the test. For $K = 3$ we set up two cases of alternative hypotheses as follows.

$$\text{Case 1.}\quad \sigma_1^2 = \gamma\sigma_2^2 = \sigma_3^2 = 1,$$
$$\text{Case 2.}\quad \sigma_1^2 = \gamma\sigma_2^2 = \gamma^2\sigma_3^2 = 1.$$

In Case 1 the power is the probability that H_{12} and H_{23} are rejected. In Case 2 the power is the probability that H_{12}, H_{13}, H_{23} are rejected. For $K = 4$ we set up three cases of alternative hypotheses as follows.

$$\text{Case 1.}\quad \sigma_1^2 = \gamma\sigma_2^2 = \sigma_3^2 = \sigma_4^2 = 1,$$
$$\text{Case 2.}\quad \sigma_1^2 = \gamma\sigma_2^2 = \gamma\sigma_3^2 = \sigma_4^2 = 1,$$
$$\text{Case 3.}\quad \sigma_1^2 = \gamma\sigma_2^2 = \gamma^2\sigma_3^2 = \sigma_4^2 = 1.$$

In Case 1 the power is the probability that H_{12},H_{23},H_{24} are rejected. In Case 2 the power is the probability that H_{12},H_{13},H_{24},H_{34} are rejected. In Case 3 the power is the probability that H_{12},H_{13},H_{23},H_{24},H_{34} are rejected. For $K = 5$ we set up four cases of alternative hypotheses as follows.

Case 1. $\sigma_1^2 = \gamma\sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2 = 1$,
Case 2. $\sigma_1^2 = \gamma\sigma_2^2 = \gamma\sigma_3^2 = \sigma_4^2 = \sigma_5^2 = 1$,
Case 3. $\sigma_1^2 = \gamma\sigma_2^2 = \gamma^2\sigma_3^2 = \sigma_4^2 = \sigma_5^2 = 1$,
Case 4. $\sigma_1^2 = \gamma\sigma_2^2 = \gamma^2\sigma_3^2 = \gamma^3\sigma_4^2 = \sigma_5^2 = 1$.

In Case 1 the power is the probability that H_{12},H_{23},H_{24},H_{25} are rejected. In Case 2 the power is the probability that H_{12},H_{13},H_{24},H_{25},H_{34},H_{35} are rejected. In Case 3 the power is the probability that H_{12},H_{13},H_{23},H_{24},H_{25},H_{34},H_{35} are rejected. In Case 4 the power is the probability that H_{12},H_{13},H_{14},H_{23},H_{24},H_{25},H_{34},H_{35},H_{45} are rejected. Tables 11 to 13 give the power of SS and CT when $\gamma = 0.75, 0.50, 0.25$. CT is uniformly more powerful compared to SS. Although the power of SS remarkably decreases as the number of hypotheses which should be rejected increases for each n and γ, the power of CT is comparatively stable independently of the number of hypotheses which should be rejected.

Table 11: Power for $K = 3$

		Case 1		Case 2	
	γ	SS	CT	SS	CT
Sam.1	0.75	0.039	0.082	0.002	0.013
	0.50	0.592	0.729	0.472	0.679
	0.25	0.999	1.000	0.999	1.000
Sam.2	0.75	0.022	0.051	0.002	0.013
	0.50	0.507	0.668	0.387	0.600
	0.25	1.000	1.000	0.998	0.999

Table 12: Power for $K = 4$

		Case 1		Case 2		Case 3	
	γ	SS	CT	SS	CT	SS	CT
Sam.1	0.75	0.017	0.022	0.003	0.016	0.000	0.000
	0.50	0.385	0.554	0.299	0.544	0.218	0.502
	0.25	0.998	0.999	0.997	1.000	0.998	1.000
Sam.2	0.75	0.006	0.014	0.002	0.013	0.000	0.001
	0.50	0.295	0.492	0.240	0.470	0.136	0.429
	0.25	0.994	0.999	0.993	0.998	0.992	0.998

Table 13: Power for $K = 5$

		Case 1		Case 2		Case 3		Case 4	
	γ	SS	CT	SS	CT	SS	CT	SS	CT
Sam.1	0.75	0.002	0.008	0.000	0.002	0.000	0.000	0.000	0.000
	0.50	0.273	0.446	0.153	0.336	0.103	0.330	0.049	0.341
	0.25	0.996	0.999	0.994	0.999	0.996	0.999	0.996	0.999
Sam.2	0.75	0.002	0.005	0.000	0.002	0.000	0.000	0.000	0.000
	0.50	0.242	0.407	0.115	0.286	0.072	0.339	0.028	0.378
	0.25	0.987	0.999	0.986	0.997	0.994	0.999	0.992	1.000

5. Conclusions

In this study we discussed stepwise multiple comparison procedures for normal variances. Specifically, we constructed the sequentially rejective step down procedure and the step up procedure for multiple comparison with a control and constructed the closed testing procedure called Ryan-Einot-Gabriel-Welsch's procedure for all-pairwise multiple comparison. We confirmed that our proposed stepwise procedures are uniformly more powerful compared to the single step procedures proposed by Imada (2018A, 2018B) through the numerical results.

Although we focused on the multiple comparison with a control and the all-pairwise multiple comparison, we should also discuss other types of multiple comparisons for normal variances. For example, we want to construct the multiple comparisons for finding minimum variances. Among several treatments evaluated by normal response, it enables us to find treatments having minimum variance. We want to discuss them referring to the multiple comparisons with the best discussed by Hsu (1981, 1982, 1984, 1985) and Hsu and Edwards (1983).

Appendix

In this Appendix we give specific formulae regarding the sequential rejective step down procedure and the step up procedure for $K = 3, 4$.

A.1. Specific formulae of the power of the sequential rejective step down procedure

First, let $K = 3$. We give the specific formulae of the power of the test for two cases of alternative hypotheses.

Case 1. $\sigma_1^2 = \gamma_{1,2}\sigma_2^2 = \sigma_3^2$

$$P(F_{1,2} > c_2, F_{1,3} \le c_1) + P((F_{1,2}, F_{1,3}) > (c_1, c_2))$$

$$= P(F_{1,2} > c_2, F_{1,3} \le c_1) + P(F_{1,2} > c_2, F_{1,3} > c_1) + P(c_2 > F_{1,2} > c_1, F_{1,3} > c_2)$$

$$= P(F_{1,2} > c_2) + P(c_2 > F_{1,2} > c_1, F_{1,3} > c_2)$$

$$= \int_{c_2\gamma_{1,2}}^{\infty} f_{1,2}(x)dx + \int_0^{\infty} f_1(x_1)\left\{\int_{c_1\lambda_{1,2}\gamma_{1,2}x_1}^{c_2\lambda_{1,2}\gamma_{1,2}x_1} f_2(x_2)dx_2\right\}\left\{\int_{c_2\lambda_{1,3}x_1}^{\infty} f_3(x_3)dx_3\right\}dx_1.$$

Case 2. $\sigma_1^2 = \gamma_{1,2}\sigma_2^2 = \gamma_{1,3}\sigma_3^2$

$$P((F_{1,2}, F_{1,3}) > (c_1, c_2))$$

$$= P(F_{1,2} > c_2, F_{1,3} > c_1) + P(c_2 > F_{1,2} > c_1, F_{1,3} > c_2)$$

$$= \int_0^{\infty} f_1(x_1)\left\{\int_{c_2\lambda_{1,2}\gamma_{1,2}x_1}^{\infty} f_2(x_2)dx_2\right\}\left\{\int_{c_1\lambda_{1,3}\gamma_{1,3}x_1}^{\infty} f_3(x_3)dx_3\right\}dx_1$$

$$+ \int_0^{\infty} f_1(x_1)\left\{\int_{c_1\lambda_{1,2}\gamma_{1,2}x_1}^{c_2\lambda_{1,2}\gamma_{1,2}x_1} f_2(x_2)dx_2\right\}\left\{\int_{c_2\lambda_{1,3}\gamma_{1,3}x_1}^{\infty} f_3(x_3)dx_3\right\}dx_1.$$

Next, let $K = 4$. We give the specific formulae of the power of the test for three cases of alternative hypotheses. Although each probability in the following formulae is expressed by the multiple integration, we omit them.

Case 1. $\sigma_1^2 = \gamma_{1,2}\sigma_2^2 = \sigma_3^2 = \sigma_4^2$

$$P(F_{1,2} > c_3, F_{1,3} \le c_2, F_{1,4} \le c_2) + P((F_{1,2}, F_{1,3}) > (c_2, c_3), F_{1,4} \le c_1)$$
$$+P((F_{1,2}, F_{1,4}) > (c_2, c_3), F_{1,3} \le c_1) + P((F_{1,2}, F_{1,3}, F_{1,4}) > (c_1, c_2, c_3))$$
$$= P(F_{1,2} > c_3, F_{1,3} \le c_2, F_{1,4} \le c_2) + P(F_{1,2} > c_3, F_{1,3} > c_2, F_{1,4} \le c_1)$$
$$+P(c_3 > F_{1,2} > c_2, F_{1,3} > c_3, F_{1,4} \le c_1) + P(F_{1,2} > c_3, F_{1,3} \le c_1, F_{1,4} > c_2)$$
$$+P(c_3 > F_{1,2} > c_2, F_{1,3} \le c_1, F_{1,4} > c_3) + P(F_{1,2} > c_3, F_{1,3} > c_2, F_{1,4} > c_1)$$
$$+P(F_{1,2} > c_3, c_2 > F_{1,3} > c_1, F_{1,4} > c_2) + P(c_3 > F_{1,2} > c_2, F_{1,3} > c_3, F_{1,4} > c_1)$$
$$+P(c_3 > F_{1,2} > c_2, c_3 > F_{1,3} > c_1, F_{1,4} > c_3) + P(c_2 > F_{1,2} > c_1, F_{1,3} > c_3, F_{1,4} > c_2)$$
$$+P(c_2 > F_{1,2} > c_1, c_3 > F_{1,3} > c_2, F_{1,4} > c_3).$$

Case 2. $\sigma_1^2 = \gamma_{1,2}\sigma_2^2 = \gamma_{1,3}\sigma_3^2 = \sigma_4^2$

$$P((F_{1,2}, F_{1,3}) > (c_2, c_3), F_{1,4} \le c_1) + P((F_{1,2}, F_{1,3}, F_{1,4}) > (c_1, c_2, c_3))$$
$$= P(F_{1,2} > c_3, F_{1,3} > c_2, F_{1,4} \le c_1) + P(c_3 > F_{1,2} > c_2, F_{1,3} > c_3, F_{1,4} \le c_1)$$
$$+P(F_{1,2} > c_3, F_{1,3} > c_2, F_{1,4} > c_1) + P(F_{1,2} > c_3, c_2 > F_{1,3} > c_1, F_{1,4} > c_2)$$
$$+P(c_3 > F_{1,2} > c_2, F_{1,3} > c_3, F_{1,4} > c_1) + P(c_3 > F_{1,2} > c_2, c_3 > F_{1,3} > c_1, F_{1,4} > c_3)$$
$$+P(c_2 > F_{1,2} > c_1, F_{1,3} > c_3, F_{1,4} > c_2) + P(c_2 > F_{1,2} > c_1, c_3 > F_{1,3} > c_2, F_{1,4} > c_3).$$

Case 3. $\sigma_1^2 = \gamma_{1,2}\sigma_2^2 = \gamma_{1,3}\sigma_3^2 = \gamma_{1,4}\sigma_4^2$

$$P((F_{1,2}, F_{1,3}, F_{1,4}) > (c_1, c_2, c_3))$$
$$= P(F_{1,2} > c_3, F_{1,3} > c_2, F_{1,4} > c_1) + P(F_{1,2} > c_3, c_2 > F_{1,3} > c_1, F_{1,4} > c_2)$$
$$+P(c_3 > F_{1,2} > c_2, F_{1,3} > c_3, F_{1,4} > c_1) + P(c_3 > F_{1,2} > c_2, c_3 > F_{1,3} > c_1, F_{1,4} > c_3)$$
$$+P(c_2 > F_{1,2} > c_1, F_{1,3} > c_3, F_{1,4} > c_2) + P(c_2 > F_{1,2} > c_1, c_3 > F_{1,3} > c_2, F_{1,4} > c_3).$$

A.2. Specific formulae regarding the step up procedure

A.2.1. Specific formulae of the probabilities used for determining critical values

We give the specific formulae of $P((F_{1,2}, F_{1,3}) \le (c_1, c_2))$ and $P((F_{1,2}, F_{1,3}, F_{1,4}) \le (c_1, c_2, c_3))$ under the assumption that $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$. Although each probability in the following formulae is expressed by the multiple integration, we omit those of $P((F_{1,2}, F_{1,3}, F_{1,4}) \le (c_1, c_2, c_3))$.

$$P((F_{1,2}, F_{1,3}) \le (c_1, c_2))$$
$$= P(F_{1,2} \le c_1, F_{1,3} \le c_2) + P(c_1 \le F_{1,2} \le c_2, F_{1,3} \le c_1)$$
$$= \int_0^\infty f_1(x_1) \left\{ \int_0^{c_1\lambda_{1,2}x_1} f_2(x_2)dx_2 \right\} \left\{ \int_0^{c_2\lambda_{1,3}x_1} f_3(x_3)dx_3 \right\} dx_1$$

$$+\int_0^\infty f_1(x_1)\left\{\int_{c_1\lambda_{1,2}x_1}^{c_2\lambda_{1,2}x_1} f_2(x_2)dx_2\right\}\left\{\int_0^{c_1\lambda_{1,3}x_1} f_3(x_3)dx_3\right\}dx_1.$$

$$\begin{aligned}
&P((F_{1,2},F_{1,3},F_{1,4})\le(c_1,c_2,c_3))\\
&=P(F_{1,2}\le c_1,(F_{1,3},F_{1,4})\le(c_2,c_3))+P(c_1\le F_{1,2}\le c_2,(F_{1,3},F_{1,4})\le(c_1,c_3))\\
&\quad+P(c_2\le F_{1,2}\le c_3,(F_{1,3},F_{1,4})\le(c_1,c_2))\\
&=P(F_{1,2}\le c_1,F_{1,3}\le c_2,F_{1,4}\le c_3)+P(F_{1,2}\le c_1,c_2\le F_{1,3}\le c_3,F_{1,4}\le c_2)\\
&\quad+P(c_1\le F_{1,2}\le c_2,F_{1,3}\le c_1,F_{1,4}\le c_3)+P(c_1\le F_{1,2}\le c_2,c_1\le F_{1,3}\le c_3,F_{1,4}\le c_1)\\
&\quad+P(c_2\le F_{1,2}\le c_3,F_{1,3}\le c_1,F_{1,4}\le c_2)+P(c_2\le F_{1,2}\le c_3,c_1\le F_{1,3}\le c_2,F_{1,4}\le c_1).
\end{aligned}$$

A.2.2. Specific formulae of the power of the step up procedure

First, let $K=3$. We give the specific formulae of the power of the test for two cases of alternative hypotheses.

Case 1. $\sigma_1^2=\gamma_{1,2}\sigma_2^2=\sigma_3^2$

$$\begin{aligned}
&P(F_{1,2}>c_2,F_{1,3}\le c_1)+P(F_{1,2}>c_1,F_{1,3}>c_1)\\
&=\int_0^\infty f_1(x_1)\left\{\int_{c_2\lambda_{1,2}\gamma_{1,2}x_1}^{\infty} f_2(x_2)dx_2\right\}\left\{\int_0^{c_1\lambda_{1,3}x_1} f_3(x_3)dx_3\right\}dx_1\\
&\quad+\int_0^\infty f_1(x_1)\left\{\int_{c_1\lambda_{1,2}\gamma_{1,2}x_1}^{\infty} f_2(x_2)dx_2\right\}\left\{\int_{c_1\lambda_{1,3}x_1}^{\infty} f_3(x_3)dx_3\right\}dx_1.
\end{aligned}$$

Case 2. $\sigma_1^2=\gamma_{1,2}\sigma_2^2=\gamma_{1,3}\sigma_3^2$

$$\begin{aligned}
&P(F_{1,2}>c_1,F_{1,3}>c_1)\\
&=\int_0^\infty f_1(x_1)\left\{\int_{c_1\lambda_{1,2}\gamma_{1,2}x_1}^{\infty} f_2(x_2)dx_2\right\}\left\{\int_{c_1\lambda_{1,3}\gamma_{1,3}x_1}^{\infty} f_3(x_3)dx_3\right\}dx_1.
\end{aligned}$$

Next, let $K=4$. We give the specific formulae of the power of the test for three cases of alternative hypotheses. Although each probability in the following formulae is expressed by the multiple integration, we omit them.

Case 1. $\sigma_1^2=\gamma_{1,2}\sigma_2^2=\sigma_3^2=\sigma_4^2$

$$\begin{aligned}
&P(F_{1,2}>c_3,(F_{1,3},F_{1,4})\le(c_1,c_2))+P(F_{1,2}>c_2,F_{1,3}\le c_1,F_{1,4}>c_2)\\
&\quad+P(F_{1,2}>c_2,F_{1,3}>c_2,F_{1,4}\le c_1)+P(F_{1,2}>c_1,F_{1,3}>c_1,F_{1,4}>c_1)\\
&=P(F_{1,2}>c_3,F_{1,3}\le c_1,F_{1,4}\le c_2)+P(F_{1,2}>c_3,c_1\le F_{1,3}\le c_2,F_{1,4}\le c_1)\\
&\quad+P(F_{1,2}>c_2,F_{1,3}\le c_1,F_{1,4}>c_2)+P(F_{1,2}>c_2,F_{1,3}>c_2,F_{1,4}\le c_1)\\
&\quad+P(F_{1,2}>c_1,F_{1,3}>c_1,F_{1,4}>c_1).
\end{aligned}$$

Case 2. $\sigma_1^2=\gamma_{1,2}\sigma_2^2=\gamma_{1,3}\sigma_3^2=\sigma_4^2$

$$P(F_{1,2}>c_2,F_{1,3}>c_2,F_{1,4}\le c_1)+P(F_{1,2}>c_1,F_{1,3}>c_1,F_{1,4}>c_1).$$

Case 3. $\sigma_1^2 = \gamma_{1,2}\sigma_2^2 = \gamma_{1,3}\sigma_3^2 = \gamma_{1,4}\sigma_4^2$

$$P(F_{1,2} > c_1, F_{1,3} > c_1, F_{1,4} > c_1).$$

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. and Tamhane, A. C. (1991). Sample size determination for step-down multiple test procedures; orthogonal contrasts and comparisons with a control. *Journal of Statistical Planning and Inference* **27**, 271-290.

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

Hsu, J. C. (1981). Simultaneous confidence intervals for all distances from the "best". *The Annals of Statistics*, **9** No. 5, 1026-1034.

Hsu, J. C. (1982). Simultaneous inference with respect to the best treatment in block designs. *Journal of the American Statistical Association*, **77** No. 378, 461-467.

Hsu, J. C. (1984). Constrained simultaneous confidence intervals for multiple comparisons with the best. *The Annals of Statistics*, **12** No. 3, 1136-1144.

Hsu, J. C. (1985). A note on multiple comparisons with the best. *In 45th Session of the International Statistical Institute, Book 2*, 445-446.

Hsu, J. C. and Edwards, D. G. (1983). Sequential multiple comparisons with the best. *Journal of the American Statistical Association*, **78** No. 384, 958-964.

Imada, T. (2017). Note on Closed Testing Procedure and Sequentially Rejective Step Down Procedure. *Bulletin of Informatics and Cybernetics* , **49**, 35-51 .

Imada, T. (2018A). Multiple comparison for checking differences among normal variances. *Communications in Statistics-Theory and Method*, **47**, Issue 18, 4402-4414.

Imada, T. (2018B). All-Pairwise Multiple Comparison for Normal Variances *Bulletin of Informatics and Cybernetics* , **50**, 19-26.

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.

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 February 5, 2019
Revised September 4, 2019