Control of Maximum Type I Familywise Error Rate of Certain Multiple Comparison Procedures

by

Tsunehisa IMADA*

(Received October 30, 2020)

Abstract

The aim of this research report is to indicate that the maximum type I familywise error rate is strongly controlled for the single step procedure, the sequentially rejective step down procedure and the step-up procedure for multiple comparison for finding normal means which are not maximum among several normal means proposed by Imada (2020).

key words: Composite hypothesis; Single step procedure; Stepwise procedure

1. Introduction

When we test plural null hypotheses simultaneously, we use multiple comparison procedures. The type I error of multiple comparison procedures is called the maximum type I familywise error rate denoted by FWER as the abbreviated notation.

The maximum type I FWER is controlled weakly or strongly. Controlling the maximum type I FWER weakly means that the specified type I error is guaranteed only when all null hypotheses are true. Specifically, if false null hypotheses exist, the probability that at least one true null hypothesis is rejected may be greater than the specified Type I error. On the other hand, controlling the maximum Type I FWER strongly means that the probability that at least one true null hypothesis is rejected is not greater than the specified Type I error even when false null hypotheses exist. Controlling the maximum Type I FWER strongly is more preferable compared to controlling it weakly for multiple comparison procedures.

Assume there are K normal populations $N(\mu_k, \sigma^2)$ $(k = 1,2,\cdots,K)$. The multiple comparison for comparing μ_1 with $\mu_2, \mu_3, \cdots, \mu_K$ simultaneously is called multiple comparison procedure with a control. It is known that the maximum Type I FWER is strongly controlled for the single step procedure proposed by Dunnett (1955), the sequentially rejective step down procedure proposed by Dunnett and Tamhane (1991) and the step-up procedure proposed by Dunnett and Tamhane (1992).

On the other hand, Imada (2020) discussed multiple comparison procedures for finding normal means which are not maximum among several normal means. Specifically, Imada (2020) proposed the single step procedure, the sequentially rejective step-down procedure and the step-up procedure. The maximum Type I FWER is strongly controlled for these procedures. The aim of this research report is to give the indications for them.

*Department of Management, Professor

2. Multiple comparison procedures for finding non-maximum normal means

Assume there are K normal populations $N(\mu_k, \sigma^2)$ $(k = 1,2,\cdots,K)$. We occasionally want to find μ_k satisfying $\mu_k = \max_{1\le l\le K}\mu_l$ among $\mu_1, \mu_2, \cdots, \mu_K$. Imada (2020) discuss multiple comparison procedures for finding μ_k satisfying

$$\mu_k < \max_{1\le l\le K}\mu_l \ (k = 1,2,\cdots,K).$$

For example, assume there are several treatments evaluated by normal response. We can find treatments which are inferior to the best treatments by the procedures.

Imada (2020) set up a null hypothesis H_k and its alternative hypothesis H_k^A as

$$H_k: \mu_k = \max_{1\le l\le K}\mu_l \text{ vs. } H_k^A: \mu_k < \max_{1\le l\le K}\mu_l$$

for $k = 1,2,\cdots,K$ and test them simultaneously using a sample

$$X_{k1}, X_{k2}, \cdots, X_{kn_k}$$

from $N(\mu_k, \sigma^2)$ for $k = 1,2,\cdots,K$. Note each null hypothesis H_k is composite.

3. Single step procedure

First, we discuss the single step procedure for $H_1, H_2, \cdots, H_K$. Letting

$$\bar{X}_k = \frac{1}{n_k}\sum_{i=1}^{n_k} X_{ki} \ (k = 1,2,\cdots,K),\ N = \sum_{k=1}^{K} n_k,$$

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

where $\upsilon = \sum_{k=1}^{K} n_k - K$, we use the statistic

$$S_k = \frac{\sqrt{N}(\max_{1\le l\le K} X_l - \bar{X}_k)}{s}$$

for testing H_k. If $S_k > c$ for a specified critical value c (> 0), we reject H_k. Otherwise we retain H_k.

3.1. Determination of critical value

We determine c for a specified significance level α. if all H_ks are true, the probability measure is denoted by $P_0(\cdot)$. Otherwise it is denoted by $\mathrm{P}(\cdot)$. Specifically, we determine c so that

$$P_0(\max_{1\le k\le K} S_k > c) = \alpha \tag{3.1}$$

when all H_ks are true. Imada (2020) derived

$$\begin{aligned}&P_0(\max_{1\le k\le K} S_k > c)\\ &= 1 - P_0(S_1 \le c, S_2 \le c, \cdots, S_K \le c)\\ &= 1 - \sum_{k=1}^{K}\int_0^{\infty}\int_{-\infty}^{\infty}\prod_{i\ne k}\left\{\Phi\left(\sqrt{\frac{n_i}{n_k}}z\right)\right.\\ &\left.-\Phi\left(\sqrt{\frac{n_i}{n_k}}z - c\sqrt{\frac{n_i}{N}}v\right)\right\}\phi(z)g(v)dzdv.\end{aligned}$$

Here $\Phi(\cdot)$is the cumulative distribution function of $N(0,1)$, $\phi(\cdot)$ is the probability density function of $N(0,1)$ and $g(v)$ is the probability density function of v given by

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

3.2. Indicating that the specified Type I error is strongly controlled

We indicate that the single step procedure using c determined by (3.1) is strongly controlled at α. If we assume

$$\mu_1 = \mu_2 = \cdots = \mu_k = \delta > \mu_l = \delta_l \quad \text{for } l = k+1, k+2, \cdots, K, \tag{3.2}$$

$H_1, H_2, \cdots, H_k$ are true and $H_{k+1}, H_{k+2}, \cdots, H_K$ are false. Letting

$$Z_l = \bar{X}_l - \delta \text{ for } l = 1,2,\cdots,k,$$
$$Z_l = \bar{X}_l - \delta_l \text{ for } l = k+1, k+2, \cdots, K,$$

we obtain

$$Z_l \sim N(0, \sigma^2/n_l) \text{ for } l = 1,2,\cdots,K.$$

The probability that all $H_1, H_2, \cdots, H_k$ are retained is

$$\begin{aligned}&\mathrm{P}(S_1 \le c, S_2 \le c, \cdots, S_k \le c)\\ &= \mathrm{P}\left(\bar{X}_l \le \bar{X}_m + \mathrm{c}\frac{s}{\sqrt{N}} \text{ for } l = 1,2,\cdots,K \text{ and } m = 1,2,\cdots,k\right)\\ &= \mathrm{P}\left(Z_l \le Z_m + \mathrm{c}\frac{s}{\sqrt{N}} \text{ for } l = 1,2,\cdots,k \text{ and } m = 1,2,\cdots,k\right.\\ &\left. Z_l \le Z_m + \delta - \delta_l + \mathrm{c}\frac{s}{\sqrt{N}} \text{ for } l = k+1, k+2, \cdots, K \text{ and } m = 1,2,\cdots,k\right)\\ &> \mathrm{P}\left(Z_l \le Z_m + \mathrm{c}\frac{s}{\sqrt{N}} \text{ for } l = 1,2,\cdots,K \text{ and } m = 1,2,\cdots,k\right)\\ &= P_0(S_1 \le c, S_2 \le c, \cdots, S_k \le c)\\ &> P_0(S_1 \le c, S_2 \le c, \cdots, S_K \le c) = 1 - \alpha.\end{aligned}$$

Therefore

$$P_0(\max_{1\le l\le k} S_l > c) < \alpha,$$

which means that the single step procedure is strongly controlled at α.

4. Sequentially rejective step down procedure

The sequentially rejective step down procedure consists of $K-1$ steps of tests. Specify $c_K, c_{K-1}, \cdots, c_2$ satisfying

$$c_K > c_{K-1} > \cdots > c_2.$$

Arranging $S_1, S_2, \cdots, S_K$ in order of a size of value, assume

$$0 = S_{(1)} \le S_{(2)} \le \cdots \le S_{(K)}. \tag{4.1}$$

$H_{(1)}, H_{(2)}, \cdots, H_{(K)}$ denote hypotheses corresponding to $S_{(1)}, S_{(2)}, \cdots, S_{(K)}$. Then, we test $H_{(2)}, H_{(3)}, \cdots, H_{(K)}$ sequentially as follows. (Since $S_{(1)} = 0$, $H_{(1)}$ is automatically retained.)

Step 1.
Case 1. If $S_{(K)} \le c_K$, we retain $H_{(2)}, H_{(3)}, \cdots, H_{(K)}$ and stop the test.
Case 2. If $S_{(K)} > c_K$, we reject $H_{(K)}$ and go to the next step.
Step 2.
Case 1. If $S_{(K-1)} \le c_{K-1}$, we retain $H_{(2)}, H_{(3)}, \cdots, H_{(K-1)}$ 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 $\mathrm{K}-1$.
Case 1. If $S_{(2)} \le c_2$, we retain $H_{(2)}$ and stop the test.
Case 2. If $S_{(2)} > c_2$, we reject $H_{(2)}$ and stop the test.

4.1. Determination of the critical value at each step of the test

Let c_K be the critical value of the single step procedure for α. Assuming all $H_{(1)}, H_{(2)}, \cdots, H_{(K)}$ are true, we determine c_m $(m = 2,3,\cdots,K-1)$ as the minimum c satisfying

$$P_0\left(\max_{k=l_1,l_2,\cdots,l_m} S_k > c\right) \le \alpha \tag{4.2}$$

for $l_1, l_2, \cdots, l_m$ chosen from $2,3,\cdots,K$ arbitrarily. Here the formulation of $P_0\left(\max_{k=l_1,l_2,\cdots,l_m} S_k > c\right)$ is obtained similarly. Trivially $c_K > c_{K-1} > \cdots > c_2$.

4.2. Indicating that the specified Type I error is strongly controlled

We indicate that the step-down procedure using $c_K, c_{K-1}, \cdots, c_2$ is strongly controlled at α. Assume (3.2). Defining the event E by

$$E: S_1 \leq c_k, S_2 \leq c_k, \cdots, S_k \leq c_k,$$

we obtain $P_0(E) \geq 1-\alpha$ by (4.2). Then we obtain $P(E) \geq 1-\alpha$ by the similar derivation to Subsection 3.2. We assume that for $S_1, S_2, \cdots, S_k$ the relation

$$S_{(i_1)} \leq S_{(i_2)} \leq \cdots \leq S_{(i_k)} \tag{4.3}$$

is derived from (4.1). Then $c_k \leq c_{i_k}$, since $k \leq i_k$. Therefore under E

$$S_{(i_1)} \leq S_{(i_2)} \leq \cdots \leq S_{(i_k)} \leq c_{i_k},$$

which means that $H_1, H_2, \cdots, H_k$ are retained under E. Therefore, the probability that $H_1, H_2, \cdots, H_k$ are retained is not less than $1-\alpha$. It means that the probability that at least one hypothesis among $H_1, H_2, \cdots, H_k$ is rejected is not greater than α. Specifically, the sequentially rejective step down procedure is strongly controlled at α.

5. Step up procedure

The step-up procedure consists of $K-1$ steps of tests. Specify $c_K, c_{K-1}, \cdots, c_2$ satisfying $c_K > c_{K-1} > \cdots > c_2$. Arranging $S_1, S_2, \cdots, S_K$ in order of a size of value, assume (4.1). $H_{(1)}, H_{(2)}, \cdots, H_{(K)}$ denote hypotheses corresponding to $S_{(1)}, S_{(2)}, \cdots, S_{(K)}$. Then, we test $H_{(2)}, H_{(3)}, \cdots, H_{(K)}$ by the step-up procedure as follows. (Since $S_{(1)} = 0$, $H_{(1)}$ is automatically retained.)

Step 1.
Case 1. If $S_{(2)} \leq c_2$, we retain $H_{(2)}$ and go to the next step.
Case 2.
If $S_{(2)} > c_2$, we reject $H_{(2)}, H_{(3)}, \cdots, H_{(K)}$ and stop the test.
Step 2.
Case 1. If $S_{(3)} \leq c_3$, we retain $H_{(3)}$ and go to the next step.
Case 2. If $S_{(3)} > c_3$, we reject $H_{(3)}, H_{(4)}, \cdots, H_{(K)}$ and stop the test.
⋮
Step $\mathrm{K}-1$.
Case 1. If $S_{(K)} \leq c_K$, we retain $H_{(K)}$ and stop the test.
Case 2.
If $S_{(K)} > c_K$, we reject $H_{(K)}$ and stop the test.

5.1. Determination of the critical value at each step of the test

Assuming all H_ks are true, we determine the critical values $c_2, c_3, \cdots, c_K$ recursively as follows. First, we determine c_1 as the minimum c satisfying

$$P_0(S_k \leq c) \geq 1-\alpha,$$

for $k = 1,2,\cdots,K$. Although c_1 is not used for the test, it is necessary for determining $c_2, c_3, \cdots, c_K$ recursively. Next, we determine c_2 as the minimum c satisfying

$$P_0\left((S_{l_1}, S_{l_2}) \leq (c_1, c)\right) \geq 1-\alpha$$

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

$$P_0\left((S_{l_1}, S_{l_2}, \cdots, S_{l_m}) \leq (c_1, c_2, \cdots, c_{m-1}, c)\right) \geq 1-\alpha$$

for $l_1, l_2, \cdots, l_m$ chosen from $1,2,\cdots,K$ arbitrarily. The definition of the formula

$$(S_{l_1}, S_{l_2}, \cdots, S_{l_m}) \leq (c_1, c_2, \cdots, c_{m-1}, c_m)$$

is given in Imada (2020).

$$P_0\left((S_{l_1}, S_{l_2}, \cdots, S_{l_m}) \leq (c_1, c_2, \cdots, c_{m-1}, c_m)\right)$$

is expressed by the sum of plural probabilities. Imada (2020) gave the specific formula for $m = 2,3$. The condition

$$c_K > c_{K-1} > \cdots > c_2 \tag{5.1}$$

is necessary for constructing the step-up procedure. (5.1) can be mathematically proved only for $\mathrm{K} = 2,3$. However, (5.1) is true for $\mathrm{K} \leq 5$ in the numerical results in Imada (2020).

5.2. Indicating that the specified Type I error is strongly controlled

We indicate that the step-up procedure using $c_2, c_3, \cdots, c_K$ is strongly controlled at α. Assume (3.2) and (4.2). We assume that for $S_1, S_2, \cdots, S_k$ (4.3) is derived from (4.2). Defining the event E by

$$E: S_{(i_1)} \leq c_1, S_{(i_2)} \leq c_2, \cdots, S_{(i_k)} \leq c_k,$$

we obtain $P_0(E) \geq 1-\alpha$. Then we obtain

$$P(E) \geq 1-\alpha$$

similarly. We show $S_{(m)} \leq c_m$ for $1 \leq m \leq i_k$ under E. If $m \leq i_1$, $S_{(m)} \leq S_{(i_1)} \leq c_1 \leq c_m$. Next assume $i_h < m \leq i_{h+1}$. Then we obtain

$$S_{(m)} \leq S_{(i_{h+1})} \leq c_{h+1}.$$

Since $h \leq i_h < m$, $h+1 \leq m$. $S_{(m)} \leq c_m$, since $c_{h+1} \leq c_m$. This means that $H_1, H_2, \cdots, H_k$ are retained till Step i_k. Therefore $H_1, H_2, \cdots, H_k$ are retained under E. It means that the probability that $H_1, H_2, \cdots, H_k$ are retained is not less than $1-\alpha$. It means that the probability that at least one hypothesis among $H_1, H_2, \cdots, H_k$ is rejected is not greater than α. Specifically, the step-up procedure is strongly controlled at α.

6. Conclusions

In this study we indicated that the maximum Type I error is strongly controlled for the single step procedure, the sequentially rejective step down procedure and the step-up procedure for multiple comparison for finding normal means which are not maximum among several normal means proposed by Imada (2020).

Although controlling the maximum Type I FWER strongly is the preferable property for multiple comparison procedures, some researchers point out that it is the severe restriction. Specifically, the power of the test tends to be lower.

Recently, other types of concepts regarding the control of the maximum Type I FWER are proposed. For example, the false discovery rate denoted by FDR is proposed by Benjamini and Hochberg (1995), which controls the expected proportion of falsely rejected hypotheses and is the weaker restriction regarding the maximum Type I error compared to controlling the maximum Type I FWER strongly. We want to construct the multiple comparison procedures for finding normal means which are not maximum among several normal means controlling FDR.

References

[1] Benjamini, Y. and Hochberg, Y. (1995) "Controlling the False Discovery Rate: a Practical and powerful Approach to Multiple Testing." J. R. Statist. Soc.} ser.B, 57(1), 289-300.

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

[3] 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.

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

[5] Imada, T. (2020). "Multiple Comparison Procedures for Finding Non-maximum Normal Means." Communications in Statistics - Theory and Methods 49, No. 16, 4073--4090.