Asian Journal of Applied Sciences (ISSN: 2321 – 0893)
Volume 04 – Issue 01, February 2016

All-pairwise Multiple Comparison for Normal Mean Vectors Based on Tukey-Welsch's Procedure

Tsunehisa Imada
Liberal Arts Education Center, Kumamoto Campus, Tokai University
Toroku, Higashi-ku, Kumamoto, 862-8652, Japan
Email: timada [AT] ktmail.tokai-u.jp

ABSTRACT---- ***In this study we consider all-pairwise multiple comparison for several normal mean vectors. Specifically, intended to more powerful procedure compared to the single step procedure we apply Tukey-Welsch's step down procedure to our problem. We give some simulation results regarding critical values and power of the test intended to compare procedures.***

Keyword--- Asymptotic distribution, Power of the test, Stepwise procedure

---

1. INTRODUCTION

There are independent p -dimensional normal random variable vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots,$ $\mathbf{X}_K$. Assume $\mathbf{X}_k \sim N_p(\boldsymbol{\mu}_k, \boldsymbol{\Sigma})$ $(k = 1,2,\ldots,K)$. If we want to test whether $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_K$ or not, we set up the null hypothesis and its alternative hypothesis as

$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_K$ vs. $H_1 : \boldsymbol{\mu}_i \neq \boldsymbol{\mu}_j$ for some $i, j\ (i < j)$.

For a sample $\mathbf{X}_{k1}, \mathbf{X}_{k2}, \ldots, \mathbf{X}_{kn_k}$ from $N_p(\boldsymbol{\mu}_k, \boldsymbol{\Sigma})$ $(k = 1,2,\ldots,K)$, let

$$\overline{\mathbf{X}}_k = \frac{1}{n_k}\sum_{i=1}^{n_k}\mathbf{X}_{ki}\ ,\ \overline{\mathbf{X}} = \frac{1}{N}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\mathbf{X}_{ki}$$

where $N = \sum_{k=1}^{K} n_k$. Letting

$$\mathbf{Q}_W = \sum_{k=1}^{K}\sum_{i=1}^{n_k}(\mathbf{X}_{ki} - \overline{\mathbf{X}}_k)(\mathbf{X}_{ki} - \overline{\mathbf{X}}_k)',\ \mathbf{Q}_B = \sum_{k=1}^{K} n_k(\overline{\mathbf{X}}_k - \overline{\mathbf{X}})(\overline{\mathbf{X}}_k - \overline{\mathbf{X}}),$$

by the likelihood ratio test criteria we reject H_0 when

$$\lambda = \frac{|\mathbf{Q}_W|^{N/2}}{|\mathbf{Q}_W + \mathbf{Q}_B|^{N/2}} > c$$

for a specified critical value c. We determine c so that $P(\lambda > c) = \alpha$ for a specified significance level α under H_0. Although it is difficult to determine the distribution of λ under H_0, $-2\log\lambda$ is asymptotically distributed according to χ^2-distribution with $f = p(K-1)$ degrees of freedom. Because $P(-2\log\lambda \le c) = P(\chi_f^2 \le c) + O(N^{-1})$ under H_0 (cf. [1]). Specifically, we obtain the asymptotic distribution

$$-2\log\lambda \sim \chi_f^2. \tag{1}$$

However, $P(-2r\log\lambda \le c) = P(\chi_f^2 \le c) + O(N^{-2})$ under H_0 gives the more precise asymptotic distribution

$$-2r\log\lambda \sim \chi_f^2. \tag{2}$$

Here

$$r = 1 - \frac{p+K+2}{2N}.$$

We compare the approximations (1) and (2). Letting $p = 2$ and $K = 4$, $f = 6$ and the upper 0.05-point of χ_6^2 is $c = 12.592$. Table 1 gives the probabilities $P(-2r\log\lambda \le c)$ and $P(-2\log\lambda \le c)$.

Table 1 : Comparisons of the closeness of the approximation
$(p = 2, K = 4, c = 12.592)$

N	$P(-2r\log\lambda > c)$	$P(-2\log\lambda > c)$
40	0.0501	0.0792
80	0.0499	0.0623
200	0.0499	0.0546

They are calculated by Monte Carlo simulation with 1,000,000 times of experiments.
Table 1 shows that the approximations (1) and (2) are closer to χ_f^2 as N is larger and the approximation (2) is closer to χ_f^2 compared to (1).

If H_0 is rejected, we occasionally want to find the pair $\boldsymbol{\mu}_i, \boldsymbol{\mu}_j$ $(i < j)$ satisfying $\boldsymbol{\mu}_i \ne \boldsymbol{\mu}_j$. Then we use multiple comparison procedures. Intended to compare $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ we set up a null hypothesis and its alternative hypothesis as

$$H_{ij} : \boldsymbol{\mu}_i = \boldsymbol{\mu}_j \text{ vs. } H_{ij}^A : \boldsymbol{\mu}_i \ne \boldsymbol{\mu}_j$$

and consider the simultaneous test of all H_{ij} s. Simple and basic procedure is the single step procedure (cf. [2]). Let

$$S_{ij} = \frac{n_i n_j}{n_i + n_j}(\overline{\mathbf{X}}_i - \overline{\mathbf{X}}_j)'\mathbf{S}^{-1}(\overline{\mathbf{X}}_i - \overline{\mathbf{X}}_j)$$

where

$$\mathbf{S} = \frac{1}{N-K}\sum_{k=1}^{K}\sum_{i=1}^{n_k}(\mathbf{X}_{ki} - \overline{\mathbf{X}}_k)(\mathbf{X}_{ki} - \overline{\mathbf{X}}_k)'.$$

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

$$P\left(\max_{i<j} S_{ij} > c\right) = \alpha \tag{3}$$

for a specified significance level α when all H_{ij} s are true. Since S_{ij} s are not independent and it is difficult to determine the distribution of $\max_{i<j} S_{ij}$, we can not obtain c satisfying (3). Under H_{ij}, each S_{ij} is distributed according to Hotelling's T^2-distribution with $(p, N-K)$ degrees of freedom denoted by $T^2_{p,N-K}$. If we determine c so that

$$P\left(T^2_{p,N-K} > c\right) = \frac{2\alpha}{K(K-1)}, \tag{4}$$

we obtain

$$P\left(\max_{i<j} S_{ij} > c\right) \le \alpha$$

by Bonferroni's inequality

$$P\left(\max_{i<j} S_{ij} > c\right) \le \sum_{i<j} P\left(S_{ij} > c\right).$$

Although c is determined easily by (4), it is conservative for the specified significance level α and the power of the test using it is lower compared to that using the exact critical value determined by (3). Less conservative critical values for the single step procedure were obtained by many researchers like [3] and [4]. However, their procedures are not remarkably more powerful. It is preferable to construct simple and more powerful procedures. It is known that the stepwize multiple comparison procedures are more powerful compared to the single step multiple comparison procedure. Although there exist various types of stepwize procedures, we focus on Tukey-Welsch's procedure (cf. [2], [5]). In this study we construct the all-pairwise multiple comparison procedure based on Tukey-Welsch's procedure using the asymptotic distribution (2). We compare our procedure with the single step procedure in terms of simulation results regarding the power of the test. In Section 1 we discuss the all-pairwise multiple comparison procedure based on Tukey-Welsch's procedure using the asymptotic distribution (2). In Section 2 we give some simulation results regarding critical values for a specified significance level and power of the test intended to compare the single step procedure and Tukey-Welsch's procedure. In Section 3 we give concluding remarks.

2. TUKEY-WELSCH'S PROCEDURE

Let $I_s = \{s_1, s_2, \ldots, s_k\}$ be an arbitrary subset of $I = \{1,2,\ldots,K\}$. $\#(I_s)$ denotes the number of elements of I_s. Defining the hypothesis H_{I_s} as $H_{I_s} : \boldsymbol{\mu}_{s_1} = \boldsymbol{\mu}_{s_2} = \cdots = \boldsymbol{\mu}_{s_k}$, we obtain $H_{I_s} = \bigcap_{s_i, s_j \in I_s, s_i < s_j} H_{s_i, s_j}$. Occasionally, H_{I_s} is denoted by $H_{s_1 s_2 \ldots s_k}$. Let F be the family consisting of all H_{I_s} s. We construct stepwise multiple comparison procedures for F applying Tukey-Welsch's procedure. For testing each H_{I_s} in F we use the statistic $S_{I_s} = -2 r_{I_s} \log \lambda_{I_s}$. Here

$$\lambda_{I_s} = \frac{|\mathbf{Q}_{W,I_s}|^{N_{I_s}/2}}{|\mathbf{Q}_{W,I_s} + \mathbf{Q}_{B,I_s}|^{N_{I_s}/2}},\ r_{I_s} = 1 - \frac{p + \#(I_s) + 2}{2N_{I_s}},$$

$$\mathbf{Q}_{W,I_s} = \sum_{j=1}^{\#(I_s)} \sum_{i=1}^{n_{s_j}} (\mathbf{X}_{s_j i} - \overline{\mathbf{X}}_{s_j})(\mathbf{X}_{s_j i} - \overline{\mathbf{X}}_{s_j})',\ \mathbf{Q}_B = \sum_{j=1}^{\#(I_s)} n_{s_j} (\overline{\mathbf{X}}_{s_j} - \overline{\mathbf{X}}_{I_s})(\overline{\mathbf{X}}_{s_j} - \overline{\mathbf{X}}_{I_s}),$$

$$\overline{\mathbf{X}}_{I_s} = \frac{1}{N_{I_s}} \sum_{j=1}^{\#(I_s)} \sum_{i=1}^{n_{s_j}} \mathbf{X}_{s_j i},\ N_{I_s} = \sum_{j=1}^{\#(I_s)} n_{s_j}.$$

We obtain the asymptotic distribution $S_{I_s} \sim \chi^2_{f_{I_s}}$ where $f_{I_s} = p(\#(I_s) - 1)$. Next, we discuss the determination of critical value $c_{\#(I_s)}$ for testing H_{I_s} using S_{I_s}. If $\#(I_s) > K - 2$, we determine $c_{\#(I_s)}$ so that $P\left(\chi^2_{f_{I_s}} > c_{\#(I_s)}\right) = \alpha$. If $\#(I_s) \le K - 2$, we determine $c_{\#(I_s)}$ so that

$$P\left(\chi^2_{f_{I_s}} > c_{\#(I_s)}\right) = \frac{\#(I_s)\alpha}{K}.$$

We test the hypotheses in F hierarchically as follows.
Step 1.
Case 1. If $S_I \le c_K$, we retain all hypotheses in F and stop the test.
Case 2. If $S_I > c_K$, we reject H_I and go to the next step.
Step 2.
We test all H_{I_s} s in F satisfying $\#(I_s) = K - 1$.
Case 1. If $S_{I_s} \le c_{K-1}$, we retain H_{I_s} and all hypotheses induced by H_{I_s}.
Case 2. If $S_{I_s} > c_{K-1}$, we reject H_{I_s}.
Step 3.
If all hypotheses satisfying $\#(I_s) = K - 2$ are retained at Step 2, we stop the test. Otherwise, we test all H_{I_s} s satisfying $\#(I_s) = K - 2$ which are not retained at Step 2.
Case 1. If $S_{I_s} \le c_{K-2}$, we retain H_{I_s} and all hypotheses induced by H_{I_s}.
Case 2. If $S_{I_s} > c_{K-2}$, we reject H_{I_s}.

We repeat similar judgments till up to Step $K - 1$. It is known that the maximum type I FWER (familywise error rate) of this procedure is not greater than α.

3. SIMULATION RESULTS

We discussed the all-pairwise multiple comparison for several normal mean vectors based on Tukey-Welsch's procedure in section 1. In this section we give some simulation results regarding critical values for a specified significance level α and power of the test intended to compare Tukey-Welsch's procedure and the single step procedure.

Let $\alpha = 0.05$ through this section. Letting $p = 2,3$ and $K = 4,5$, we set up the balanced sample size $n = 10,20$ for each population. Table 2 gives conservative critical values of the single step procedure determined by Bonferroni's inequality. Table 3 gives approximate critical values of the single step procedure determined by Monte Carlo simulation so that (3) may be satisfied. Table 4 gives Type I error obtained by using the critical value in Table 2. The results of Tables 3, 4 are obtained by Monte Carlo simulation with 1,000,000 times of experiments. Table 4 shows that the critical values for $K = 5$ are more conservative compared to those for $K = 4$. Tables 5,6 give critical values of Tukey-Welsch's procedure for $K = 4,5$, respectively.

Table 2 : Conservative critical values of the single step procedure determined by Bonferroni's inequality

n	10		20	
K	4	5	4	5
$p=2$	11.329	12.254	10.350	11.548
$p=3$	14.602	15.456	12.974	14.317

Table 3 : Approximate critical values of the single step procedure determined by Monte Carlo simulation

n	10		20	
K	4	5	4	5
$p=2$	10.750	11.565	9.900	10.800
$p=3$	14.000	14.710	12.500	13.450

Table 4 : Type I error by conservative critical values in Table 2

n	10		20	
K	4	5	4	5
$p=2$	0.0408	0.0391	0.0416	0.0370
$p=3$	0.0409	0.0400	0.0421	0.0365

Table 5 : Critical values of Tukey-Welsch's procedure $(K = 4)$

p	2	3
H_{1234}	12.592	16.919
$H_{123}, H_{124}, H_{134}, H_{234}$	9.488	12.592
$H_{12}, H_{13}, H_{14}, H_{23}, H_{24}, H_{34}$	7.378	9.349

Table 6 : Critical values of Tukey-Welsch's procedure $(K = 5)$

p	2	3
H_{12345}	15.508	21.027
H_{1234} etc.	12.592	16.919
H_{123} etc.	10.712	13.968
H_{12} etc.	7.825	9.838

Next, we consider the power of the test. Specify $p = 2$ and $K = 4$. Since we calculate the power by Monte Carlo simulation, we should specify $\mathbf{\Sigma}$ and $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, \boldsymbol{\mu}_4$. Let

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \text{ where } \rho = -0.8, -0.4, 0, 0.4, 0.8 .$$

We set up four types of $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, \boldsymbol{\mu}_4$ as follows.

Case 1. $\boldsymbol{\mu}_1 = (0,0)', \boldsymbol{\mu}_2 = (0,0)', \boldsymbol{\mu}_3 = (0,0)', \boldsymbol{\mu}_4 = (0,1)'$

Case 2. $\boldsymbol{\mu}_1 = (0,0)', \boldsymbol{\mu}_2 = (0,0)', \boldsymbol{\mu}_3 = (0,1)', \boldsymbol{\mu}_4 = (0,1)'$

Case 3. $\boldsymbol{\mu}_1 = (0,0)', \boldsymbol{\mu}_2 = (0,0)', \boldsymbol{\mu}_3 = (1,0)', \boldsymbol{\mu}_4 = (0,1)'$

Case 4. $\boldsymbol{\mu}_1 = (0,0)', \boldsymbol{\mu}_2 = (1,0)', \boldsymbol{\mu}_3 = (0,1)', \boldsymbol{\mu}_4 = (1,1)'$

In Case 1 the power is the probability that H_{14}, H_{24}, H_{34} are rejected. In Case 2 the power is the probability that H_{13} , H_{14}, H_{23}, H_{24} are rejected. In Case 3 the power is the probability that $H_{13}, H_{14}, H_{23}, H_{24}, H_{34}$ are rejected. In Case 4 the power is the probability that $H_{12}, H_{13}, H_{14}, H_{23}, H_{24}, H_{34}$ are rejected. We define abbreviated notations. SS-B denotes the single step procedure using the conservative critical value in Table 2. SS-A denotes the single step procedure using the critical value in Table 3. TW denotes Tukey-Welsch's procedure. Tables 7 to 10 give the power for each procedure in Cases 1 to 4, respectively. They are calculated by Monte Carlo simulation with 10,000 times of experiments. In each case the power decreases as ρ increases from -0.8 to 0, then the power increases as ρ increases from 0 to 0.8 for each procedure. TW are uniformly more powerful compared to SS-B and SS-A.

Table 7 : Power comparison in Case 1

n	10					20				
ρ	−0.8	−0.4	0	0.4	0.8	−0.8	−0.4	0	0.4	0.8
SS-B	0.487	0.090	0.062	0.088	0.487	0.957	0.413	0.298	0.416	0.960
SS-A	0.529	0.109	0.069	0.103	0.536	0.959	0.455	0.336	0.444	0.962
TW	0.589	0.142	0.100	0.143	0.579	0.978	0.560	0.433	0.556	0.980

Table 8 : Power comparison in Case 2

n	10					20				
ρ	−0.8	−0.4	0	0.4	0.8	−0.8	−0.4	0	0.4	0.8
SS-B	0.399	0.050	0.028	0.048	0.393	0.946	0.322	0.220	0.325	0.948
SS-A	0.447	0.058	0.034	0.056	0.435	0.954	0.368	0.241	0.361	0.956
TW	0.479	0.080	0.054	0.080	0.468	0.974	0.462	0.346	0.472	0.971

Table 9 : Power comparison in Case 3

n	10					20				
ρ	−0.8	−0.4	0	0.4	0.8	−0.8	−0.4	0	0.4	0.8
SS-B	0.102	0.015	0.010	0.015	0.316	0.585	0.230	0.145	0.252	0.940
SS-A	0.126	0.020	0.014	0.022	0.358	0.614	0.262	0.171	0.272	0.950
TW	0.171	0.034	0.023	0.040	0.397	0.714	0.385	0.273	0.402	0.974

Table 10 : Power comparison in Case 4

n	10					20				
ρ	−0.8	−0.4	0	0.4	0.8	−0.8	−0.4	0	0.4	0.8
SS-B	0.083	0.005	0.003	0.006	0.087	0.582	0.163	0.086	0.168	0.578
SS-A	0.101	0.007	0.005	0.008	0.102	0.614	0.199	0.109	0.196	0.601
TW	0.192	0.037	0.026	0.035	0.194	0.722	0.409	0.303	0.414	0.719

4. CONCLUSIONS

In this study we discussed Tukey-Welsch's procedure for all-pairwise multiple comparison for several normal mean vectors. We confirmed that Tukey-Welsch's procedure is uniformly more powerful compared to the single step procedure using the critical value which is close to the exact critical value for a specified significance level.

There exist other types of stepwise multiple comparison procedures like Peritz's procedure (cf. [6]) and the closed testing procedure (cf. [7]). We should develop all-pairwise multiple comparison procedures for several normal mean vectors using these procedures intended to obtain more powerful procedures in the future.

5. REFERENCES

[1] Anderson T W. An introduction to multivariate statistical analysis (Third ed.). Wiley, New York, 2003.
[2] Tukey J W. The problem of multiple comparisons, Unpublished manuscript, Princeton University, 1953.
[3] Seo T, Siotani M. The multivariate Studentized range and its upper percentiles.

Journal of the Japan Statistical Society, 1992, 22: 123-137.

[4] Fujikoshi Y, Seo T. Asymptotic expansions for the joint distribution of correlated Hotelling's T^2 statistics under normality. *Communications in Statistics, Theory and Methods*, 1999, 28: 773-788,.

[5] Welsch R E. A modification of the Newman-Keuls procedure for multiple comparisons, Working Paper Sloan School of Management, M.I.T., Boston, MA, 1972, pp. 612-672,.

[6] Peritz E. A note on multiple comparisons, Unpublished manuscript, Hebrew University, Israel, 1970.

[7] Hsu J C. Multiple comparisons. Boca Raton : Chapman&Hall, 1996.