Far East Journal of Theoretical Statistics
Volume 34, Number 1, 2011, Pages 1-16
Published Online: March 12, 2011
This paper is available online at http://pphmj.com/journals/fjts.htm

ALL-PAIRWISE COMPARISONS FOR POPULATIONS WITH UNEQUAL VARIANCES

HONG LI

Department of Mathematical Sciences
Cameron University
Lawton, OK 73505, U. S. A.
e-mail: lhong@cameron.edu

Abstract

Tukey's method provided simultaneous inference for all-pairwise comparisons (MCA) under balanced design, usual normality and equality of variances assumptions. Under the unbalanced design, the Tukey-Kramer method assumes the variances are equal across all treatment groups. It provides a set of conservative simultaneous confidence intervals for all-pairwise differences and has been widely used. In practice, however, homogeneity of variances is seldom satisfied. In this article, an approximate approach is proposed when the equality of variances cannot be assumed and the ratios of population variances among treatments are known from previous experience. The results from a simulation study show that the error rate of the Tukey-Kramer method is excessive, while the error rate of the proposed method is within the nominal level when the variances are different. In addition, an approximate approach is proposed to provide the simultaneous confidence intervals for all-pairwise differences when the ratios of variances are unknown.

1. Introduction

For the problem of comparing means of two normal populations with unequal variances, a large number of approximate tests and exact tests are available in the

2010 Mathematics Subject Classification: 97K70.

Keywords and phrases: error rate, simultaneous confidence intervals, unequal variances.

Received November 30, 2010

literature. Welch [13, 14] proposed an approximate *t* test when the variances of two samples are different. Tsui and Weerahandi [12] considered a generalized test to the Behrens-Fisher problem. There have been only a few attempts to extend these results to the problem of testing the equality of a number of means when the population variances are not homogeneous. For instance, Tamhane [10] proposed two approximate approaches for multiple comparisons with a control and all-pairwise comparisons when the variances are unequal; Games and Howell [4] provided the approximate simultaneous confidence intervals for all-pairwise differences under heteroscedasticity.

In some circumstances, when the assumption of homogeneity of variances is unreasonable, while knowledge of ratios of population variances is available from prior experience, we propose an approximate simultaneous confidence interval method for MCA in Section 2. An example to illustrate our method is given in Section 3. A simulation study is preformed to compare the error rate of the proposed method and the Tukey-Kramer method, the result is given in Section 4. In Section 5, we propose an approximate approach to estimate the critical value that provides the simultaneous confidence intervals for all-pairwise differences when the ratios of variances are unknown. We conduct Monte Carlo studies to compare the error rate between the proposed approach and Tamhane's [10] approach, the results are given in Section 6. Finally, a discussion is given in Section 7.

2. All-pairwise Comparisons when the Ratios of Variances are Known

Many well established methods on all-pairwise comparisons assume homogeneity of variances across all treatment groups, such as Tukey's method and Tukey-Kramer method [5, 6]. In practice, however, equality of variances is seldom satisfied. We are motivated to consider a simultaneous confidence interval method for all-pairwise comparisons when the condition of equal variances is not satisfied. In this section, we assume the ratios of population variances among treatments are known from previous experience.

Suppose the Y_{ij} is the observed measurement of the *j*th subject in the *i*th treatment, and the Y_{ij}'s are independently distributed as $N(\mu_i, \sigma_i^2)$. Consider the unbalanced one-way model,

$$Y_{ij} = \mu_i + \varepsilon_{ij}$$

with $i = 1, \ldots, k,\;\; j = 1, 2, \ldots, n_i,$ where ε_{ij}'s follow $N(0, \sigma_i^2)$. Denote the sample mean measurement by $\overline{Y}_i$, it is a least square estimate for μ_i. Denote the sample variance by S_i^2, it is known as an unbiased estimate of σ_i^2, and given by

$$S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_i)^2,$$

where $S_i^2 \sim \frac{\sigma_i^2}{v_i} \chi_{v_i}^2$, $(v_i = n_i - 1)$, and S_i^2 is independent of $\overline{Y}_i$. Denote the ratios of population variance of responses from the first treatment group to that of the ith group by λ_{1i}, for $i = 1, \ldots, k$, i.e., $\lambda_{1i} = \sigma_1^2 / \sigma_i^2$, and $\lambda_{11} = 1$. They are assumed known from prior experience or pilot studies.

Consider testing the pairwise differences among k population means, $H_{0i} : \mu_i - \mu_j = \delta$ for all $i \neq j$ vs. $H_{ai} : \mu_i - \mu_j \neq \delta$ for at least one $i \neq j$, with $i, j = 1, \ldots, k$, where δ is a prespecified threshold. The test statistic is considered as

$$z_{ij} = \frac{\bar{y}_i - \bar{y}_j - \delta}{\sqrt{\frac{\sigma_i^2}{n_i} + \frac{\sigma_j^2}{n_j}}} = \frac{\bar{y}_i - \bar{y}_j - \delta}{\sigma_1 \sqrt{\frac{1}{\lambda_{1i} n_i} + \frac{1}{\lambda_{1j} n_j}}}, \quad i, j = 1, \ldots, k, \quad i \neq j,$$

where z_{ij} has a standard normal distribution, σ_1 is the population standard deviation for the first treatment group. We derive a pooled estimate for σ_1, which is given by

$$\hat{\sigma}_1 = \sqrt{\frac{1}{v} \left[\sum_{i=1}^{k} \lambda_{1i} S_i^2 (n_i - 1) \right]},$$

where

$$v = \sum_{i=1}^{k} (n_i - 1).$$

Let $r = \hat{\sigma}_1 / \sigma_1$. Then it is a $\sqrt{\frac{\chi_v^2}{v}}$ random variable with density

$$p(r) = \frac{2}{\Gamma\left(\frac{v}{2}\right)2^{\frac{v}{2}}} r^{v-1} e^{-\frac{v}{2}r^2} v^{\frac{v}{2}}.$$

Hence, the test statistic can be written as

$$t_{ij} = \frac{\overline{y}_i - \overline{y}_0 - \delta}{\hat{\sigma}_1\sqrt{\frac{1}{\lambda_{1i}n_i} + \frac{1}{\lambda_{1j}n_j}}}, \quad i \neq j,$$

where t_{ij} has a student t distribution with v degrees of freedom.

Therefore, the $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_i - \mu_j$ are given by

$$\mu_i - \mu_j \in \hat{\mu}_i - \hat{\mu}_j \pm |\, h^* \,|\hat{\sigma}_1\sqrt{1/(\lambda_{1i}n_i) + 1/(\lambda_{1j}n_j)} \quad \text{for all } i \neq j,$$

where $|\, h^* \,|$ is the solution to the equation

$$P\left\{\frac{|\,\hat{\mu}_i - \mu_i - (\hat{\mu}_j - \mu_j)\,|}{\hat{\sigma}_1\sqrt{1/(\lambda_{1i}n_i) + 1/(\lambda_{1j}n_j)}} \leq |\, h^* \,| \text{ for all } i < j\right\} = 1 - \alpha.$$

In general, it is difficult to compute the exact value of $|\, h^* \,|$ especially when the number of treatments is large. We suggest an approximate solution based on Sidak's inequality [6, 8], which is stated in the following theorem:

Theorem 2.1 [6, 8]**.** *Let* $X = (X_1, X_2, ..., X_k)$ *be the vector of random variables having the k-dimensional normal distribution with zero means, arbitrary variances* $\sigma_1^2, \sigma_2^2, ..., \sigma_k^2$, *and an arbitrary correlation matrix* $R = \rho_{ij}$. *Then, for any positive numbers* $c_1, ..., c_k$,

$$P\{|\, X_i \,| \leq c_i \;\; \textit{for } i = 1, ..., k\} \geq \prod_{i=1}^{k} P\{|\, X_i \,| \leq c_i\}.$$

Applying the inequality, an approximate solution of $|\, h^* \,|$ can be obtained by solving the following equation:

$$\prod_{1 \le i < j \le k} \int_0^{\infty} \int_{-\infty}^{\infty} \left[\Phi\left(|\, h^* \,| r \sqrt{1 + \frac{\lambda_{1i} n_i}{\lambda_{1j} n_j}} + \sqrt{\frac{\lambda_{1i} n_i}{\lambda_{1j} n_j}}\, z \right) - \Phi\left(-|\, h^* \,| r \sqrt{1 + \frac{\lambda_{1i} n_i}{\lambda_{1j} n_j}} + \sqrt{\frac{\lambda_{1i} n_i}{\lambda_{1j} n_j}}\, z \right) \right] \cdot p(r)\phi(z)\, dr dz = 1 - \alpha, \qquad (1)$$

where Φ is the c.d.f. of the standard normal random variable, $\phi(z)$ is the density function of standard normal distribution, $r = \hat{\sigma}_1 / \sigma_1$ is a $\sqrt{\frac{\chi_\nu^2}{\nu}}$ random variable.

Another approximation is based on the Bonferroni's procedure. The critical value can be replaced by $t_\nu^{\alpha/2q}$, where $q = k(k-1)/2$, and $\nu = \sum_{i=1}^{k} n_i - k$. Thus, the conservative $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_i - \mu_j$ are given by

$$\mu_i - \mu_j \in \hat{\mu}_i - \hat{\mu}_j \pm t_{\frac{\alpha}{k(k-1)}, \nu} \hat{\sigma}_1 \sqrt{1/(\lambda_{1i} n_i) + 1/(\lambda_{1j} n_j)} \text{ for all } i \ne j.$$

Sidak's inequality is known to be less conservative and yields sharper inference than the Bonferroni's procedure.

3. An Example

We give an example to illustrate the method, we proposed for all-pairwise comparisons under unequal variances. This is an example from the textbook of Hsu ([6, Subsection 4.1.2, pp. 86-87]). The presence of harmful insects in farm fields can be detected by examining insects trapped in boards covered with a sticky material and erected in the fields. According to [6], Wilson and Shade reported on the number of cereal leaf beetles trapped when six boards of each of three colors were placed in a field of oats in July 1967. A hypothetical data set, patterned after their experiment, was used to illustrate the multiple comparisons with the best methods in [6]. Summary statistics of the number of beetles trapped are given in Table 1. The Levenue test for equality of variances yields a significant result $(F = 21.79,$ $p\text{-value} < 0.0001)$, it indicates that the variances of numbers of beetles trapped among colors are different. The test can be set up as $H_{0i}: \mu_i - \mu_j = 0$ vs. $H_{ai}: \mu_i - \mu_j \neq 0$, for $i, j = 1, 2, 3$. Assume the ratios of variances are known from prior experience, and given by $\lambda_{12} = 0.5$, $\lambda_{13} = 1.6$. Let us first consider the method we proposed in the previous section. The pooled estimate for the variance of the first group $\hat{\sigma}_1^2 = 46.944$. Based on the proposed approach, a set of 95% simultaneous confidence intervals on the difference of mean numbers is given by

$$\bar{y}_i - \bar{y}_j - |h^*| \hat{\sigma}_1 \sqrt{1/\lambda_{1i} n_i + 1/\lambda_{1j} n_j}, \text{ for } i, j = 1, 2, 3.$$

For $\alpha = 0.05$, $v = 15$, the critical value $|h^*|$ is 2.687 by solving equation (1), it infers

$$2.655 < \mu_1 - \mu_2 < 28.685,$$

$$22.761 < \mu_1 - \mu_3 < 41.919,$$

$$4.496 < \mu_2 - \mu_3 < 28.844.$$

It indicates that yellow is more attractive than all other colors; red is more attractive than blue. Assume the variances are equal, the common standard deviation is estimated as $\hat{\sigma} = 7.597$ based on Tukey's method. The 95% confidence intervals are given as

$$\bar{y}_i - \bar{y}_j - |q^*| \hat{\sigma} \sqrt{2/n}, \text{ for } i, j = 1, 2, 3.$$

For 3 treatments and 15 degrees of freedom, we have $|q^*| = 2.598$, it infers

$$4.275 < \mu_1 - \mu_2 < 27.065,$$

$$20.945 < \mu_1 - \mu_3 < 43.735,$$

$$5.275 < \mu_2 - \mu_3 < 28.065.$$

Tukey's method gives the same conclusion as above, while the simultaneous confidence lower limits and upper limits are different from those of the proposed approach. Without loss of generality, suppose $\delta = 5$. Then we conduct the test again and conclude that red is not more attractive than blue based on the proposed approach, while the opposite conclusion is obtained using Tukey's method. It indicates that Tukey's method may provide erroneous inference when the equality of variance assumption is not satisfied.

Table 1. Summary statistics of the number of beetles trapped

Color	Treatment label	Sample size	Sample mean	Sample Std. dev.
Yellow	1	6	47.17	6.79
Red	2	6	31.5	9.91
Blue	3	6	14.83	5.34

4. A Simulation Study on Error Rate

Under unbalanced design and unequal variances, Tukey and Kramer proposed what we call the *Tukey-Kramer method* [6]. It gives approximate simultaneous confidence intervals for all-pairwise differences:

$$\mu_i - \mu_j \in \hat{\mu}_i - \hat{\mu}_j \pm |q^*|\hat{\sigma}\sqrt{1/n_i + 1/n_j} \text{ for all } i \neq j,$$

where $|q^*|$ is the same critical value as given in Tukey's method. According to [6], Tukey stated that 'the approximation ... is apparently in the conservative direction'.

For all-pairwise comparisons, the Tukey-Kramer procedure has been widely used when the sample size is unequal for the convenience of the available tables of $|q^*|$ and the availability in some statistical packages. We perform a simulation study in this section to show that the Tukey-Kramer method can have inflated error rate under certain conditions. Three groups of data with the same mean are generated from $N(2, \sigma_i)$, with $n_1 = 40$, $n_2 = 50$, $n_3 = 10$. $\sigma_1^2 = 2$, $\sigma_2^2 = 4$, and let σ_3^2 increase from 2 to 15. Thus, the variances (σ_i^2) and sample size (n_i) are inversely paired. The larger the variance of the third groups, the larger the ratio of (σ_3^2/n_3), the larger the imbalance among (σ_i^2/n_i)'s. We compute the error rate, the probability of rejecting at least one null hypothesis given all the hypotheses are true under significance level 0.05. We can see from Table 2 that the error rate of the Tukey-Kramer method is within 0.05 only when $\sigma_3^2/n_3 = 0.2$, it increases to 0.2228 as the ratio increases to 1.5, it is more than four times the nominal level α. In contrast, all the values of the error rate for the proposed approach are within 0.05, and a little conservative as we expect. When the heterogeneous variances are present and the imbalance among the (σ_i^2/n_i)'s is large, the Tukey-Kramer method is found to have excessive error rate, and it is well known that the Tukey-Kramer procedure is inherently conservative. The proposed method is a little conservative, while it controls the family-wise error rate. When the prior knowledge of the ratios of variances is available and equality of variances cannot be assumed, the proposed approach may be considered for inference on all-pairwise differences.

Table 2. Estimated error rate for $k = 3$, $\alpha = 0.05$

σ_3^2/n_3	Tukey-Kramer method error rate (standard error)	Proposed method error rate (standard error)
0.2	0.0241 (0.00153)	0.0472 (0.00213)
0.4	0.0596 (0.00240)	0.0405 (0.00197)
0.6	0.0992 (0.00291)	0.0409 (0.00198)
0.8	0.1333 (0.00342)	0.0408 (0.00197)
1.0	0.1644 (0.00375)	0.0401 (0.00196)
1.5	0.2228 (0.00420)	0.0380 (0.00191)

5. All-pairwise Comparisons when the Ratios of Variances are Unknown

When prior knowledge on the ratios of variances is unavailable, we propose an approximate procedure to estimate the critical value that provides the confidence intervals for all-pairwise differences in this section. Under the unbalanced one-way model, consider testing the pairwise differences among k population means when the ratios of variance are unknown, $H_{0i}: \mu_i - \mu_j = \delta$ for all $i \neq j$ vs. $H_{ai}: \mu_i - \mu_j \neq \delta$ for at least one $i \neq j$, with $i, j = 1, \ldots, k$, where δ is a given threshold constant. Let λ_{i1} denote the ratio of variances of the ith treatment to that of the first group, that is, $\lambda_{i1} = \sigma_i^2/\sigma_1^2$, which are unknown. The ratios can be estimated by $\hat{\lambda}_{i1} = cS_i^2/S_1^2$, $i = 1, \ldots, k$, where $c = (n_1 - 3)/(n_1 - 1)$ and $\hat{\lambda}_{i1} = 1$ for $i = 1$. $\hat{\lambda}_{i1}$ is an unbiased estimator of λ_{i1} according to Lehmann and Casella [7]. The test statistic is considered as

$$t_{ij} = \frac{\bar{y}_i - \bar{y}_j - \delta}{\hat{\sigma}_1 \sqrt{\frac{\hat{\lambda}_{i1}}{n_i c} + \frac{\hat{\lambda}_{j1}}{n_j c}}} \tag{2}$$

$$= \frac{\bar{y}_i - \bar{y}_j - \delta}{\sqrt{\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}}}, \quad i \neq j, \tag{3}$$

where $\hat{\sigma}_1 = S_1,$ the sample standard deviation of the first treatment. S_i^2 and S_j^2 are the sample estimates of variances of the *i*th and the *j*th treatment groups.

The $100(1 - \alpha)\%$ simultaneous confidence intervals for $\mu_i - \mu_j$ are given by

$$\mu_i - \mu_j \in \bar{y}_i - \bar{y}_j \pm |h_1| \hat{\sigma}_1 \sqrt{\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}} \text{ for all } i \neq j,$$

where $|h_1|$ is the solution to the equation

$$P\left\{ \frac{|(\bar{y}_i - \mu_i) - (\bar{y}_j - \mu_j)|}{\hat{\sigma}_1 \sqrt{\frac{\hat{\lambda}_{i1}}{n_i c} + \frac{\hat{\lambda}_{j1}}{n_j c}}} \leq |h_1| \text{ for all } i < j \right\} = 1 - \alpha, \tag{4}$$

in order to guarantee the overall coverage probability to be $1 - \alpha.$

Based on Welch's [13] method and the test statistic given in (3), Tamhane's [10] method provides the approximate $100(1 - \alpha)\%$ simultaneous confidence intervals for $\mu_i - \mu_j, \ i \neq j,$ which are given by

$$\mu_i - \mu_j \in \bar{y}_i - \bar{y}_j - t_{\hat{v}_{ij}, \beta} \left(\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j} \right)^{1/2} \text{ for all } i \neq j,$$

where

$$\beta = 1 - (1 - \alpha)^{2/k(k-1)},$$

$\hat{v}_{ij}$ is given by

$$\hat{v}_{ij} = \frac{\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j}\right)^2}{\frac{\hat{\sigma}_i^4}{n_i^2(n_i - 1)} + \frac{\hat{\sigma}_j^4}{n_j^2(n_j - 1)}}.$$

Tamhane's method has joint confidence level less than $1 - \alpha$ in some cases, it will result in inflated family-wise error rate. We estimate the error rate by Monte Carlo studies in the next section. Here, we propose to estimate the critical value based on the test statistic given in (2). The probability given in (4) can be written as

$$E_{\hat{\underset{\sim}{\lambda}}} E_{\hat{\sigma}_1}\left[P\left\{|\, z_{ij}\,| < |\, h_1\,| \left(\frac{\hat{\sigma}_1}{\sigma_1}\right)\sqrt{b_{ij}},\ i < j\right\} |\, \hat{\sigma}_1,\ \hat{\underset{\sim}{\lambda}}\right], \tag{5}$$

where $\hat{\underset{\sim}{\lambda}} = (\hat{\lambda}_{i1}, \hat{\lambda}_{j1}, i < j)$, $(|\, z_{ij}\,|, i < j)$ are multivariate normal random variables. $b_{ij} = \left(\frac{\hat{\lambda}_{i1}}{n_i c} + \frac{\hat{\lambda}_{j1}}{n_j c}\right) \Big/ \left(\frac{\lambda_{i1}}{n_i} + \frac{\lambda_{j1}}{n_j}\right)$. It can be shown that b_{ij} approximately follow F distribution with $\hat{f}_{ij}$ and v_1 degrees of freedom based on Welch's [13] method and $\hat{f}_{ij}$ is given by

$$\hat{f}_{ij} = \frac{\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j}\right)^2}{\frac{\hat{\sigma}_i^4}{n_i^2(n_i - 1)} + \frac{\hat{\sigma}_j^4}{n_j^2(n_j - 1)}}.$$

Since the correlation among z_{ij} for $i < j$ depends on the unknown σ_i^2, we propose an approximate solution for the critical value by applying Sidak's inequality introduced in Section 2. The approximate solution for $|\, h_1\,|$ can be obtained by solving the following equation:

$$\prod_{1 \le i < j \le k} \int_0^\infty \int_0^\infty [1 - 2\Phi(-|\, h_1\,|(r)\sqrt{b_{ij}})]\, p(r)\, f(b_{ij})\, dr\, db_{ij} = 1 - \alpha, \tag{6}$$

where $r = \hat{\sigma}_1/\sigma_1$, Φ is the c.d.f of the standard normal distribution. $p(r)$ is the density function of r, and $f(b_{ij})$ is the density function of b_{ij}.

6. Monte Carlo Studies on Error Rate

We carry out two simulation studies to compare the error rate for the proposed method based on Sidak's inequality and Tamhane's [10] method. We generate three random samples from a normal distribution with mean equal 2. The variances are different and given as $\sigma_1^2 = 2$, $\sigma_2^2 = 4$, and σ_3^2 increases from 2 to 8. We compute the error rate under significance level 0.05. The results are given in Tables 3 and 4.

Table 3. Estimated error rate for $k = 3$, $\alpha = 0.05$ $(n_1 = n_2 = n_3 = 20)$

$(\sigma_1^2, \sigma_2^2, \sigma_3^2)$	Proposed method error rate (standard error)	Tamhane's method error rate (standard error)
(2, 4, 2)	0.041 (0.0032)	0.091 (0.0064)
(2, 4, 4)	0.045 (0.0032)	0.079 (0.0060)
(2, 4, 6)	0.038 (0.0032)	0.093 (0.0064)
(2, 4, 8)	0.039 (0.0038)	0.096 (0.0065)

Table 4. Estimated error rate for $k = 3$, $\alpha = 0.05$ $(n_1 = 30,\ n_2 = 40,\ n_3 = 10)$

$(\sigma_1^2, \sigma_2^2, \sigma_3^2)$	Proposed method error rate (standard error)	Tamhane's method error rate (standard error)
(2, 4, 2)	0.037 (0.0046)	0.091 (0.0064)
(2, 4, 4)	0.043 (0.0038)	0.084 (0.0082)
(2, 4, 6)	0.039 (0.0041)	0.076 (0.0059)
(2, 4, 8)	0.028 (0.0042)	0.081 (0.0086)

We can see that all the values of error rate for the proposed approach are within nominal level 0.05. It is relatively conservative since some confidence intervals for the estimated error rate fall below 0.05. While all the error rates for Tamhane's method are beyond 0.05, and all the confidence intervals for the estimated error rate fall above 0.05. It indicates that Tamhane's method cannot control the family-wise error rate.

7. Discussion

In this article, we proposed an approximate method for the all-pairwise comparisons without the equal variance assumption. It provides approximate simultaneous confidence intervals for all-pairwise differences when the ratios of variances are known. Simulation results indicate that the proposed method always controls the family-wise error rate, while the Tukey-Kramer method has inflated error rate especially when the small size is paired with large variance. Thus, it may lead to erroneous inference when the equal variance assumption is not satisfied.

In practice, the ratios of the variances may not be known sometimes. To handle such a situation, we provide another approximate approach for MCA to estimate the critical value. Compared to Tamhane’s approximation, our approximation with Sidak’s inequality controls the family-wise error rate for different sample sizes and variances but slightly conservative. Instead, Tamhane’s method has excessive error rate for different sizes and variances, therefore cannot control the error rate. The advantage of Tamhane’s method is that it is easy to apply. The proposed procedure using Sidak’s inequality to find an approximated critical value through the numerical integration, which may be involved with a large amount of calculations. Therefore, to improve our methods to get a set of sharper confidence intervals will be the topic in our future research.

In summary, how to control the family-wise error rate is a central issue in the area of multiple comparisons. The plausibility of equal variance condition should always be considered and verified. When the assumption of the equal variances is not satisfied, the methods with more flexible restrictions, such as the approximate approaches proposed in the article, may be considered as a more reasonable candidate for MCA.

References

[1] P. Bauer, A note on multiple testing procedures in dose finding, Biometrics 53 (1997), 1125-1128.

[2] C. W. Dunnett, A multiple comparison procedure for comparing several treatments with a control, J. Amer. Statist. Assoc. 50 (1955), 482-491.

[3] C. W. Dunnett and A. C. Tamhane, A step-up multiple test procedure, J. Amer. Statist. Assoc. 87 (1992), 162-170.

[4] P. A. Games and J. F. Howell, Pairwise multiple comparison procedures with unequal *N*'s and/or variances, J. Edu. Statist. 1 (1976), 113-125.

[5] Y. Hochberg and A. C. Tamhane, Multiple Comparison Procedure, Wiley, New York, 1987.

[6] J. C. Hsu, Multiple Comparisons: Theory and Methods, Chapman and Hall, London, 1996.

[7] E. L. Lehmann and G. C. Casella, Theory of Point Estimation, Springer, 1998.

[8] Z. Sidak, Rectangular confidence regions for the means of multivariate normal distributions, J. Amer. Statist. Assoc. 62 (1967), 626-633.

[9] W. Y. Tan and M. A. Tabatabi, A robust procedure for comparing several means under heteroscedasticity and nonnormality, Comm. Statist. Simulation Comput. 15 (1986), 733-745.

[10] A. C. Tamhane, Multiple comparisons in model I one-way Anova with unequal variances, Comm. Statist. Theory Methods A 6(1) (1977), 15-32.

[11] A. C. Tamhane, A comparison of procedures for multiple comparisons of means with unequal variances, J. Amer. Statist. Assoc. 74 (1979), 471-480.

[12] K. W. Tsui and S. Weerahandi, Generalized *p* values in significance testing of hypotheses in the presence of nuisance parameters, J. Amer. Statist. Assoc. 84 (1989), 602-607.

[13] B. L. Welch, The significance of the difference between two means when the population variance are unequal, Biometrika 29 (1938), 350-362.

[14] B. L. Welch, The generalization of student's' problem when several population variance are involved, Biometrika 34 (1947), 28-35.

Appendix

In the appendix, we give the details of derivations of equations (5) and (6), and the distribution of b_{ij} in Section 5.

Derivation of equations (5) and (6).

Proof.

$$P\left\{\frac{|(\bar{y}_i-\mu_i)-(\bar{y}_j-\mu_j)|}{\hat{\sigma}_1\sqrt{\frac{\hat{\lambda}_{i1}}{n_ic}+\frac{\hat{\lambda}_{j1}}{n_jc}}}\le |h_1| \text{ for all } i<j\right\}$$

$$=P\left\{\frac{|\bar{y}_i-\bar{y}_j-(\mu_i-\mu_j)|/\sigma_1\sqrt{\frac{\lambda_{i1}}{n_i}+\frac{\lambda_{j1}}{n_j}}}{\frac{\hat{\sigma}_1}{\sigma_1}\sqrt{\frac{\hat{\lambda}_{i1}}{n_ic}+\frac{\hat{\lambda}_{j1}}{n_jc}}\Big/\sqrt{\frac{\lambda_{i1}}{n_i}+\frac{\lambda_{j1}}{n_j}}}<|h_1| \text{ for all } i<j\right\}$$

$$=E_{\hat{\sigma}_1}\left[P\left\{\frac{|z_{ij}|}{\sqrt{\left(\frac{\hat{\lambda}_{i1}}{n_ic}+\frac{\hat{\lambda}_{j1}}{n_jc}\right)\Big/\left(\frac{\lambda_{i1}}{n_i}+\frac{\lambda_{j1}}{n_j}\right)}}<|h_1|\left(\frac{\hat{\sigma}_1}{\sigma_1}\right) \text{ for all } i<j\right\}\Bigg|\,\hat{\sigma}_1\right]$$

$$=E_{\underset{\sim}{\hat{\lambda}}}E_{\hat{\sigma}_1}\left[P\left\{|z_{ij}|<|h_1|\left(\frac{\hat{\sigma}_1}{\sigma_1}\right)\sqrt{\left(\frac{\hat{\lambda}_{i1}}{n_ic}+\frac{\hat{\lambda}_{j1}}{n_jc}\right)\Big/\left(\frac{\lambda_{i1}}{n_i}+\frac{\lambda_{j1}}{n_j}\right)},\ i<j\right\}\Bigg|\,\hat{\sigma}_1,\ \underset{\sim}{\hat{\lambda}}\right],$$

where $\underset{\sim}{\hat{\lambda}}=(\hat{\lambda}_{i1},\hat{\lambda}_{j1},\ i<j)$, and $(|z_{ij}|,\ i<j)$ are multivariate normal random variables. Let

$$b_{ij}=\left(\frac{\hat{\lambda}_{i1}}{n_ic}+\frac{\hat{\lambda}_{j1}}{n_jc}\right)\Big/\left(\frac{\lambda_{i1}}{n_i}+\frac{\lambda_{j1}}{n_j}\right).$$

Then we have

$$P\left\{\frac{|(\bar{y}_i - \mu_i) - (\bar{y}_j - \mu_j)|}{\hat{\sigma}_1\sqrt{\frac{\hat{\lambda}_{i1}}{n_i c} + \frac{\hat{\lambda}_{j1}}{n_j c}}} \le |h_1| \text{ for all } i < j\right\}$$

$$= P\left\{\frac{|z_{ij}|}{\frac{\hat{\sigma}_1}{\sigma_1}\sqrt{b_{ij}}} < |h_1| \text{ for all } i < j\right\}$$

$$\ge \prod_{1 \le i < j \le k} P\left\{\frac{|z_{ij}|}{\frac{\hat{\sigma}_1}{\sigma_1}\sqrt{b_{ij}}} < |h_1|\right\}$$

(by Sidak's inequality)

$$= \prod_{1 \le i < j \le k} E_{(b_{ij},\, i<j)} E_{\hat{\sigma}_1}\left[P\left\{|z_{ij}| < |h_1|\left(\frac{\hat{\sigma}_1}{\sigma_1}\right)\sqrt{b_{ij}}\right\} \Big| \hat{\sigma}_1, b_{ij}\right]$$

$$= \prod_{1 \le i < j \le k} \int_0^\infty \int_0^\infty \left[1 - 2\Phi\left(-|h_1|\left(\frac{\hat{\sigma}_1}{\sigma_1}\right)\sqrt{b_{ij}}\right)\right] p\left(\frac{\hat{\sigma}_1}{\sigma_1}\right) f(b_{ij})\, d\left(\frac{\hat{\sigma}_1}{\sigma_1}\right) db_{ij}$$

$\left(\text{let } r = \hat{\sigma}_1/\sigma_1\right)$

$$= \prod_{1 \le i < j \le k} \int_0^\infty \int_0^\infty [1 - 2\Phi(-|h_1|(r)\sqrt{b_{ij}}]\, p(r)\, f(b_{ij})\, dr db_{ij}. \qquad \square$$

Derivation of the distribution of b_{ij}.

Proof.

$$b_{ij} = \frac{\frac{\hat{\lambda}_{i1}}{n_i c} + \frac{\hat{\lambda}_{j1}}{n_j c}}{\frac{\lambda_{i1}}{n_i} + \frac{\lambda_{j1}}{n_j}} = \frac{\left(\frac{S_i^2}{n_i S_1^2} + \frac{S_j^2}{n_j S_1^2}\right) S_1^2}{\left(\frac{\sigma_i^2}{n_i \sigma_1^2} + \frac{\sigma_j^2}{n_j \sigma_1^2}\right) S_1^2} = \frac{\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}}{\frac{S_1^2}{\sigma_1^2}\left(\frac{\sigma_i^2}{n_i} + \frac{\sigma_j^2}{n_j}\right)}$$

and

$$\frac{\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}}{\frac{\sigma_i^2}{n_i} + \frac{\sigma_j^2}{n_j}}$$

is approximately $\chi^2_{f_{ij}} / f_{ij}$ random variable based on Welch [13]. f_{ij} can be approximated by

$$\hat{f}_{ij} = \frac{\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_j^2}{n_j} \right)^2}{\frac{\hat{\sigma}_i^4}{n_i^2 (n_i - 1)} + \frac{\hat{\sigma}_j^4}{n_j^2 (n_j - 1)}},$$

$\frac{S_1^2}{\sigma_1^2} \sim \chi^2_{v_1} / v_1$, with $v_1 = n_1 - 1$. Hence, b_{ij} has asymptotic F distribution with $\hat{f}_{ij}$ and v_1 degrees of freedom. □