Multiple Comparison Procedures under Equal and unequal population Variances

Sanhita Sarmah[1] and Bipin Gogoi[2]

Research Scholar, Department of Statistics, Dibrugarh University[1]

Professor, Department of Statistics, Dibrugarh University[2]

Abstract: In many fields like, renewal process, life testing problem, stochastic modeling the assumption of exponentiality is heavily used. In many studies dealing with equipment of failure and repair time ,often these time are assumed to be exponentially distributed. However, considerable efforts have been dedicated to testing for exponentiality. Some of the workers in this fields are Bartholomew(1957),Epstein(1960),Gail and Gastwirth (1978), Kolmogorov-Smirnov(1975), Pyke(1967), Moran(1951), Spurior(1984), Doksum and Kendell(1984), Aschar(1990), D agostino and Stephen(1986) etc. In this paper we wish to study performance of some of these tests under different alternative hypotheses, viz. under lognormal distribution, Weibull distribution and gamma distributions etc. Results are obtain usingMonte Carlo simulation technique and displayed in different tables and graphs. Discussions are made based on simulated results and conclusion is drawn accordingly.

Keywords: Goodness of fit test, Exponential distribution, Lognormal, Weibull and gamma distribution, Monte Carlo technique.

1. INTRODUCTION

The term "Multiple Comparisons" refers to making several tests for statistical significance of differences between means (or proportions or variances, etc.) within a group. Statistical procedures that are designed to take into account and properly control for the multiplicity effect through some combined or joint measure of erroneous inferences are called multiple comparison procedures (MCPs). It is one of the fundamental problems of practical importance.

Multiple comparison procedures can be conducted in different ways. The following four types of multiple comparison procedures are seen in the literature based on the objective of the researcher:

(i) MCA (all-pairwise multiple comparisons) considers for all $\mu_i = \mu_j$ for all $i \neq j$ to be of primary interest.

(ii) MCB (multiple comparisons with the best) considers $\mu_i - max \, \mu_j$, $\neq i$, i =1,…, k to be of primary interest.

(iii) MCC (multiple comparisons with a control) considers for $\mu_i - \mu_k$

i = 1… k-1 to be of primary interest.

(iv) MCM (multiple comparisons with the mean) considers $\mu_i - \bar{\mu}$ or $\mu_i - \bar{\bar{\mu}}$ for all $i = 1,…, k$ to be of primary interest, where $\bar{\mu}$ and $\bar{\bar{\mu}}$ are the unweighted and the weighted means of the $\mu_i's$.

Except the MCA all other three types (MCB, MCC, and MCM) of multiple comparisons comes under the category many-to-one comparisons. Tukey (1993) recommends MCM over MCA for large k, because the result of k comparisons in MCM would be easier to comprehend than the result of comparisons in MCA. This advantage is shared by MCB and MCC, which make k and k-1 comparisons, respectively. In the quality control setting, MCM is usually known as analysis of means (ANOM).

The foundation of the subject of multiple comparisons was laid in the late 1940s and early 1950s, principally by David Duncan, S.N. Roy, R. C. Bose, Henry Scheffe and John W.Tukey, although some of the ideas appeared much earlier in the works of Fisher, Student, and others.

The MCPs have many applications in Pharmaceutical Companies, Clinical Research, Genomics, Education, Physiology, Data Mining in Market Research etc.. The following are some practical situations where MCPs are used:

(i) A medical research team conducts a clinical study comparing the success rates of different drug regimens for a particular disease.

(ii) Comparison of system designs via computer simulation.

(iii) In experiments of gain in animal weight effected by different feeding rations.

(iv) A polling service wishes to determine the most popular candidate before a certain election.

(v) A manufacturer would like to know which of three potential plant layouts will maximize expected revenues.

(vi) In a clinical trial a control group consists of patients treated with a standard existing therapy, and the treatment groups consist of patients treated with new therapies.

It is well known that many multiple comparison procedures have been developed and each has its advantages and disadvantages. Choosing from among the many alternatives procedures poses the problem for the users. Multiple comparison procedures when population variances are equal dealt by many authors. Several articles viewing multiple comparison procedures have commented on the variance heterogeneity problem and have recommended approximate solutions. Statisticians certainly have been developing new and improved approaches to this problem but we wondered whether the applied researchers are familiar with the issues and the

solutions. When population variances differ, several solutions have been suggested. Many of the proposed procedures control the overall risk of type I errors but have low statistical power. Three procedures that have often been recommended are those that have been developed by Game and Howell(1976) based on Welch's solution to the Behrens-Fisher problem, Dunnett C(1980) based on Cochran's(1964) solution to the Behrens-Fisher problem, and Dunnett T3(1980) based on Sidak's(1967) uncorrelated-t inequality. These procedures control the overall risk of a type I error experimentwise at approximately the nominal significance level and have the best statistical power among the alternative solutions.

With respect to the selection of a multiple comparison procedure, the researcher must be aware that his/her choice can often significantly affect the results of the experiment. For example, many multiple comparison procedure(e.g. , those that are based on traditional test statistics) are inappropriate(and may lead to incorrect decisions) when assumptions of the test statistics are not met(e.g. , normality,variance homogeneity). Furthermore, several multiple comparison procedure have recently been proposed that according to published results and/or statistical theory significantly improve on the properties(e.g., power) of existing procedures, while still maintaining the specified error rate at or below α.

Therefore , the goal of this article is to describe some of the newer multiple comparison procedures within the context of one-way completely randomized designs when validity assumptions are satisfied, as well as when the assumptions are not satisfied. That is, the goal is to help popularize newer procedures; procedures which should provide researchers with more robust and/or more powerful tests of their pairwise comparison null hypotheses.

In recent years considerable attention has been focused on problem of multiple comparisons among the μ_i when the σ_i^2 are unequal. Here are some multiple comparison methods that donot require equal population variance and are appropriate for unbalanced designs. The Game Howell procedure due to Game and Howell (1976) based on Studentized range distribution and Tamhane(1977,1979) T2 and T3 also due to Dunnet(1980) use the Studentized maximum modules distribution, where the degrees of freedom calculation includes sample variance. The Game Howell procedure is sometimes slightly liberal (the actual FWER is (family-wise error rate) higher than the α specified by the user). In other words, for some data sets, the Game Howell procedure does not protect the FWER. We have used six test procedures for our study. The are – Student t test, Scheffe Test, Game and Howell test, Tamhane two tests and the Dunn test based on ranks.

2. TEST PROCEDURES

Consider the one-way fixed effect ANOVA model:
$$y_{ij} = \mu_i + e_{ij} \quad \ldots \quad (1)$$
Where $e_{ij} \sim NID(0, \sigma_i^2)$ and μ_i and σ_i^2 are unknown. For i = 1,2, . . . , k and j = 1,2, . . . ,n_i . Let $\bar{y}_i$ be the ith sample mean and s_i^2 be the estimate of σ_i^2 based on v_i degrees freedom independent of $\bar{y}_i$. We shall take s_i^2 to

be usual sample variance based on $v_i = n_i -1$ defrees of freedom . The test statistics and the joint CI estimate for k* differences $\mu_i - \mu_j$ are described by the following method:

Unequal sample sizes and homogeneity of variance ($n_i \neq n_j, \sigma_i^2 = \sigma_j^2$)

2.1 Student's t-test :

A hypothesis for the pairwise comparison Hc : $\mu_i = \mu_j$, group sizes (n_1, n_2) are unequal , can be examined with the test statistic:

$$t_c = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{\frac{MSE}{n_i} + \frac{MSE}{n_j}}} \quad \ldots \quad (1)$$

where $\bar{y}_i$ is the ith group mean $(i \neq j)$ and MSE is the usual analysis of variance estimate of error variance. Note that this is the usual two-sample Student t-test, distributed as a t variate with $n_i + n_j -2$ degrees of freedom. When group sizes are equal, the statistic [with 2(n-1) degrees of freedom] would be

$$t_c = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{\frac{2MSE}{n}}} \quad \ldots \quad (2)$$

2.2 Scheffe (SK) Method:

$$SK = \bar{y}_i - \bar{y}_j \; / \; [s^2(\frac{1}{n_i} + \frac{1}{n_j})]^{1/2} \quad \ldots (3)$$

Where $s^2 = \frac{1}{n_i + n_j - 2}(\sum_{i=1}^{n_i}(y_i - \bar{y})^2 + \sum_{j=1}^{n_j}(y_j - \bar{y})^2)$

The statistic SK will be rejected if SK$\geq$ [(k-1)F(α, k, v)]$^{1/2}$

Where F($\alpha, k - 1, v$) $is\ the\ upper$ α point of the central F distribution with k-1 and v=$n_i + n_j - 2$ degrees of freedom.

Proposed 100(1-α)% joint CI for $\mu_i - \mu_j$ of it as

$$\bar{y}_i - \bar{y}_j \pm [(k-1)F(\alpha, k, v)]^{1/2} \; [s^2(\frac{1}{n_i} + \frac{1}{n_j})]^{1/2}$$

2.3 Hochberg GT2-Method:

The Hochberg statistics is given by

$$HB = (\bar{y}_i - \bar{y}_j) / s.[\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2} \quad \ldots \quad (4)$$

HB be will be rejected if HB$\geq m_{\alpha, k^*, v}$
Where $m_{\alpha, k^*, v}$, the upper α- point of the Studentized maximum modulus(SMM) distribution with parameter k* = k(k-1)/2 and v degrees of freedom.[Table Stoline and Uky(1979)], and s^2 is the usual pool estimate of σ^2 , based on $v = N - k$ degrees of freedom.

Proposed 100(1-α)% joint CI for $\mu_i - \mu_j$ of it as

$$\bar{y}_i - \bar{y}_j \pm s.(m_{\alpha, k^*, v})[\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2}$$

2.4 Spjotvoll-Stoline T' – Method:

The Spjotvoll-Stoline ST statistic is given by
$$ST = (\bar{y}_i - \bar{y}_j) / (s/min(\sqrt{n_i}, \sqrt{n_j})) \quad \ldots \quad (5)$$

The ST will be rejected if ST$\geq (q'_{\alpha;k,v})$
Where $(q'_{\alpha;k,v})$, the upper α- point of the studentized augmented range(SAR) distribution with parameter k and $v = N - k$ degrees of freedom.. [Table Stoline(1978)]
The Proposed 100(1-α)% joint CI for $\mu_i - \mu_j$ of it as

$$\bar{y}_i - \bar{y}_j \pm s.(q'_{\alpha;k,v})min(\sqrt{n_i}, \sqrt{n_j})$$

2.5 Tukey-Kramer (TK) Method:

The Tk method is known as Kramer's (1956) metod and was first discussed by Tukey(1953). This is an approximate extension of the T method for unequal sample size. The TK statistic may be defined as

$$TK = \bar{y}_i - \bar{y}_j \; / \; s.[\frac{1}{n_i} + \frac{1}{n_j}]^{1/2} \quad \dots (6)$$

The test TK is rejected if TK$\geq SR_{\alpha,k,v}$, where $SR_{\alpha,k,v}$ is the α point of the distribution of the Studentized Range(SR) of k normal variates,with v degrees of freedom.

The Proposed 100(1-α)% joint CI for $\mu_i - \mu_j$ of it as

$$\bar{y}_i - \bar{y}_j \; \pm SR_{\alpha,k,v} \; s.[\frac{1}{n_i} + \frac{1}{n_j}]^{1/2}$$

Unequal sample sizes and heterogeneity of variance ($n_i \neq n_j, \sigma_i^2 \neq \sigma_j^2$)

2.6 Game Howell (GH) Method:

Game and Howell(1976) proposed a pairwise comparison procedure which can be defined as

$$GH = \bar{y}_i - \bar{y}_j \; /[\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2} \quad \dots (7)$$

And the degrees of freedom v_{ij} are given by

$$v_{ij} = (\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j})^2 / (\frac{s_i^4}{n_i^2 v_i} + \frac{s_j^4}{n_j^2 v_j})$$

Denotes the Welch (1938) approximate formula for degrees of freedom.

The statistic GH will be rejected if GH$\geq A_{ij,\alpha,k} = SR_{\alpha,k,v_{ij}}/\sqrt{2}$

Here , $SR_{\alpha,k,v_{ij}}$ is the α point of the distribution of the Studentized range of k normal variates and v_{ij} the degrees of freedom.. The $A_{ij,\alpha,k}$ was chosen to achieve at least approximately the joint confidence coefficient 1-α.

It gives the 100(1-α)% joint Confidence Interval for $\mu_i - \mu_j$

$$\bar{y}_i - \bar{y}_j \; \pm A_{ij,\alpha,k}[\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2}$$

2.7 Cochran(C) Method:

Cochran statistic C is given by

$$C = [\bar{Y}_i - \bar{Y}_j] / [\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2} \quad \dots (8)$$

The test statistic C is rejected if C > $A_{ij,\alpha,k} = SR^*_{\alpha,k,v_{ij}}/\sqrt{2}$

where $SR^*_{\alpha,k,v_{ij}} = \dfrac{SR_{\alpha,k,v_i}\frac{s_i^2}{n_i} + SR_{\alpha,k,v_j} \; s_j^2/n_j}{\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}}$

The 100(1-α)% joint Confidence Interval for $\mu_i - \mu_j$ given by equation (1),

$$\bar{y}_i - \bar{y}_j \; \pm SR^*_{\alpha,k,v_{ij}}/\sqrt{2} \quad [\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2}$$

Corresponds to the weighted average of Student's t proposed by Cochran(1964) as an approximate solution to the Behrens_Fisher problem for k=2 groups.

2.8 Tamhane (T2) Method:

Tamhane (1977, 1979) developed two pairwise multiple comparison procedures for the case in which the variances in the groups are unequal and can be defined as follows:

$$T2 = [\bar{Y}_i - \bar{Y}_j] / [\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2} \quad \dots (9)$$

The statistics T2 will be rejected if T2 $\geq t_{\alpha^* v_{ij}}$, where $t_{\alpha^* v_{ij}}$ is the two-sided α^* point of student's t distribution with v_{ij} degrees of freedom and

$$\alpha^* = 1 - (1-\alpha)1/k^* ., k^* = k(k-1)/2$$

proposed 100(1-α)% joint CI for $\mu_i - \mu_j$

$$\bar{Y}_i - \bar{Y}_j \; \pm t_{\alpha^* v_{ij}}[\frac{s_1^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2}$$

2.9 Tamhane(T3) Method:

Tamhane(1977,1979) developed another procedure which can be defined as

$$T3 = \bar{Y}_i - \bar{Y}_j \; / [\frac{s_1^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2} \quad \dots (10)$$

The statistics T3 will be rejected if T3 $\geq SMM_{\alpha,k^* v_{ij}}$ where $SMM_{\alpha,k^*,v_{ij}}$ is the α point of the Studentized maximum modulas(SMM) distribution of of k^* uncorrelated normal variates with v_{ij} (given in (4)) degrees of freedom.

Proposed 100(1-α)% joint CI for $\mu_i - \mu_j$ of it as

$$\bar{Y}_i - \bar{Y}_j \; \pm SMM_{\alpha,k^* v_{ij}}[\frac{s_1^2}{n_i} + \frac{s_j^2}{n_j}]^{1/2}$$

2.10 Dunn (1964 Dn) Method:

Dunn(1964) developed a procedures based on mean rank sum of respective sample which ca be defined as:

$$Dn = \bar{R}_i - \bar{R}_j \; / \; [s^2(\frac{1}{n_i} + \frac{1}{n_j})]^{1/2} \quad \dots (11)$$

where $\bar{R}_i$ and $\bar{R}_j$ are the mean rank of ith and jth samples.

$$S^2 = \frac{N(N+1)}{12}(\frac{1}{n_i} + \frac{1}{n_j})$$

If tied rank are present , then the following equation to calculate the standard error

$$S^2 = \frac{N(N+1)}{12} - \frac{\sum_{l=1}^g t_i}{12(N-1)}(\frac{1}{n_i} + \frac{1}{n_j})$$

where g is the number of tied groups.

Dunn(1964) proposed 100(1-α)% joint CI for $\mu_i - \mu_j$ as

$$\bar{R}_i - \bar{R}_j \; \pm z_{\frac{\alpha}{2k}*}[s^2(\frac{1}{n_i} + \frac{1}{n_j})]^{1/2}$$

where k* = k(k+1)/2

2.11 Brown-Forsythe(BF) Method:

Brown and Forsythe(1974) proposed a statistic given by

$$BF = (\bar{Y}_i - \bar{Y}_j) / [(k-1)(\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j})]^{1/2} \quad \dots (12)$$

The test statistic BF will be rejected if BF> F($\alpha, k - 1, v_{ij}$)

The 100(1-α)% joint CI for $\mu_i - \mu_j$ given by BF statistic is

$$\bar{Y}_i - \bar{Y}_j \; \pm [(k-1)F(\alpha, k - 1, v_{ij})(\frac{s_1^2}{n_i} + \frac{s_j^2}{n_j})]^{1/2}$$

Comparing C method with T2 and T3 method, it is observed that C method is preferable for large degrees of freedom and T3 method is preferable for small degrees of freedom. The GH method is also recommended at the risk of being somewhat liberal.

3. MONTE CARLO STUDY

The purpose of the present study is to compare, via computer simulation techniques, eleven methods in analyzing pairwise treatment differences in a multi-treatment experiment. Out of these five methods for equal variances but equal and unequal sample sizes and the remaining six methods for unequal variances with equal and unequal sample sizes are considered under different alternatives.

Each simulated experiment used four independent random samples (k=4) . Pseudo-random numbers are generated by a computer using RND functions.

Then Box-Muller (1958) method is used to draw standard normal deviates. Necessary modification is made multiplying by scale parameter σ_i and adding location parameter μ_i in respective samples when required. For each set of combinations we generate 5000 normal deviates. For each set we calculate value of the test statistic and compare it with the table value for taking decision for reject and accept the null hypothesis. Number of time rejected the null hypothesis divided by the number of replication gives the empirical level of test statistic under null hypothesis and power for the alternative hypothesis.The recorded results are shown in Table 1(a) and Table 2(b) below.

Table 1(a): Empirical Levels of some selected multiple comparison tests under equal and unequal variances and Sample sizes at 0.05 levels.

Sample Sizes n_i	Sample No. n_i	Value of S.D. σ_i	Test Statistics				
			t	HB	TK	ST	SK
4	5 5 5 5	1 1 1	.0413	.0563	.0678	.0921	.0544
		1 1.5 2 2.5	.0488	.0800	.0950	.0275	.0634
		1 1 1 2	.0525	.0963	.1122	.0649	.0630
		1 1 2 2	.0495	.0801	.0966	.0432	.0636
		1 2 3 4	.0500	.0986	.1138	.0164	.0768
		1 3 5 7	.0581	.1141	.1320	.0083	.0916
4	10 10 10 10	1 1 1	.0517	.0052	.0612	.0577	.0408
		1 1.5 2 2.5	.0510	.0751	.0888	.0873	.0478
		1 1 1 2	.0534	.0868	.0992	.0362	.0490
		1 1 2 2	.0535	.0749	.0892	.0176	.0470
		1 2 3 4	.0550	.0893	.1030	.0372	.0540
		1 3 5 7	.0581	.0996	.1152	.0093	.0616
4	6 10 14 16	1 1 1	.0470	.0484	.0646	.0394	.0398
		1 1.5 2 2.5	.0273	.0476	.0378	.0238	.0200
		1 1 1 2	.0306	.0353	.0450	.0266	.0298
		1 1 2 2	.0273	.0296	.0410	.0272	.0226
		1 2 3 4	.0257	.0291	.0362	.0258	.0172
		1 3 5 7	.0249	.0310	.0382	.0272	.0160
4	16 14 10 6	1 1 1	.0436	.0570	.0620	.0290	.0380
		1 1.5 2 2.5	.0915	.1790	.2078	.1098	.1234
		1 1 1 2	.0957	.2180	.2490	.1306	.1100
		1 1 2 2	.0895	.1640	.1970	.1042	.1188
		1 2 3 4	.1314	.2680	.3000	.1662	.1810
		1 3 5 7	.1281	.3230	.3662	.2136	.2294

Table 1(b): Empirical Levels of some selected multiple comparison tests under equal and unequal variances and Sample sizes at 0.05 levels.

Sample Sizes n_i	Sample No. n_i	Value of S.D. σ_i	Test Statistics					
			GH	T2	T3	Dn	BF	C
4	5 5 5 5	1 1 1 1	.0658	.0490	.0496	.0410	.0442	.0438
		1 1.5 2 2.5	.0746	.0576	.0584	.0520	.0510	.0818
		1 1 1 2	.0760	.0566	.0566	.0516	.0500	.0538
		1 1 2 2	.0766	.0574	.0572	.0528	.0492	.0472
		1 2 3 4	.0834	.0690	.0654	.0568	.0578	.0580
		3 5 7	.0896	.0746	.0720	.0648	.0632	.0728
4	10 10 10 10	1 1 1 1	.0594	.0670	.0550	.0518	.0372	.0356
		1 1.5 2 2.5	.0700	.0654	.0600	.0632	.0416	.0424
		1 1 1 2	.0748	.0698	.0606	.0616	.0430	.0420

		1 1 2 2	.0686	.0628	.0506	.0634	.0406	.0404
		1 2 3 4	.0724	.0636	.0618	.0688	.0458	.0502
		3 5 7	.0736	.0662	.0624	.0738	.0450	.0566
4	6 10 14 16	1 1 1 1	.0686	.0742	.0576	.0440	.0575	.0400
		1 1.5 2 2.5	.0624	.0738	.0522	.0334	.0516	.0376
		1 1 1 2	.0658	.0752	.0548	.0402	.0494	.0408
		1 1 2 2	.0622	.0724	.0528	.0356	.0487	.0384
		1 2 3 4	.0634	.0766	.0506	.0344	.0498	.0406
		1 3 5 7	.0692	.0776	.0550	.0364	.0456	.0456
4	16 14 10 6	1 1 1 1	.0438	.0730	.0554	.0438	.0567	.0394
		1 1.5 2 2.5	.0808	.0760	.0634	.0920	.0725	.0518
		1 1 1 2	.0760	.0764	.0618	.0860	.0856	.0476
		1 1 2 2	.0782	.0688	.0630	.0874	.0898	.0506
		1 2 3 4	.0812	.0762	.0656	.1138	.0945	.0600
		1 3 5 7	.1822	.0698	.0658	.1306	.0988	.0636

Table 2(a): Empirical Power of some selected multiple comparison tests under equal and unequal variances and Sample sizes at 0.05 levels.

Sample Size n_i	Value of S.D. σ_i	Value of. Mean μ_i	Test Statistics				
			t	HB	TK	ST	SK
10 10 10 10	1 1 1 1	0 0 .5	.1372	.0651	.0763	.0792	.0505
		0 0 0 1	.4622	.3103	.3392	.3455	.2618
		0 0 0 1.5	.8472	.7162	.7443	.7502	.6570
		0 0 0 2	1.000	.9522	.9630	.9653	9306
		0 0 3	1.000	1.000	1.000	1.000	1.000
10 10 10 10	1 1 1 2	0 0 0 .5	.1062	.0666	.0753	.0399	.0522
		0 0 0 1	.3008	.1988	.2173	.1088	.1665
		0 0 0 1.5	.6672	.4313	.4576	.2546	.3824
		0 0 0 2	.8722	.6861	.7097	.4708	.6422
		0 0 3	1.000	.9654	.9712	.8676	.9533
10 12 14 16	1 1 1 1	0 0 0 .5	.2866	.0918	.1072	.0775	.0691
		0 0 0 1	.7570	.4527	.4851	.4122	.3916
		0 0 0 1.5	1.000	.8744	.8923	.8453	.8351
		0 0 0 2	1.000	.9976	1.000	.9910	.9899
		0 0 0 3	1.000	1.000	1.000	1.000	1.000
10 12 14 16	1 1 1 2	0 0 0 .5	.0732	.0597	.0673	.0506	.0450
		0 0 0 1	.1513	.2249	.2454	.1995	.1865
		0 0 0 1.5	.3624	.5240	.5526	.4859	.4680
		0 0 0 2	.6444	.8109	.8297	.7785	.7682
		0 0 3	1.000	1.000	.9960	.9915	.9910
10 12 14 16	1 1 1	0 0 0 .5	.2537	.0838	.0951	.0883	.0646
		0 0 0 1	.6156	.2938	.3240	.3633	.2400
		0 0 0 1.5	.9505	.6870	.7120	.6912	.9074
		0 0 2	1.000	.9387	.9531	.8407	1.000
		0 0 0 3	1.000	1.000	1.000	.9776	1.000
10 12 14 16	1 1 1 2	0 0 0 .5	.1224	.1018	.1143	.0428	.0833
		0 0 0 1	.2352	.2804	.3009	.1226	.2449
		0 0 0 1.5	.4405	.5433	.5669	.2840	.4976
		0 0 0 2	.6832	.7950	.8116	.5139	.7596
		0 0 0 3	.9853	.9885	.9903	.9016	.9827

Table 2(b): Empirical Power of some selected multiple comparison tests under equal and unequal variances and Sample sizes at 0.05 levels.

Sample Size n_i	Value of S.D. σ_i	Value of . Mean μ_i	GH	T2	T3	Dn	BF	C
10 10 10 10	1 1 1 1	0 0 0 .5	.0752	.0727	.0624	.0833	.0478	.1398
		0 0 0 1	.3200	.3167	.2886	.3560	.2387	.2367
		0 0 0 1.5	.6976	.6940	.6649	.6981	.6080	.6044
		0 0 0 2	.9459	.9405	.9313	.9118	.9006	.9018
		0 0 3	1.000	1.000	1.000	1.000	1.000	1.000
10 10 10 10	1 1 1 2	0 0 0 .5	.0399	.0399	.0342	.0670	.0679	.0265
		0 0 0 1	.1088	.1061	.1017	.1668	.0797	.0837
		0 0 0 1.5	.2546	.2499	.2463	.3326	.2064	.2167
		0 0 0 2	.4708	.4677	.4668	.5443	.4087	.4324
		0 0 3	.8676	.8672	.8695	.8655	.8295	.8522
10 12 14 16	1 1 1 1	0 0 0 .5	.0927	.1256	.0881	.1318	.0693	.0771
		0 0 0 1	.3908	.5001	.4187	.4968	.3695	.3914
		0 0 0 1.5	.7576	.8903	.8420	.8594	.8022	.8213
		0 0 0 2	.9312	1.000	.9898	.9819	.9816	.9851
		0 0 0 3	1.000	1.000	1.000	1.000	1.000	1.000
10 12 14 16	1 1 1 2	0 0 0 .5	.0513	.0667	.0433	.0677	.0576	.0382
		0 0 0 1	.1697	.2197	.1605	.2072	.1331	.1504
		0 0 0 1.5	.4073	.5027	.4131	.4471	.3651	.3953
		0 0 0 2	.6916	.7919	.7156	.7036	.6702	.7029
		0 0 3	.9812	1.000	.9912	.9619	.9838	.9887
10 12 14 16	2 1 1 1	0 0 0 .5	.0883	.1035	.0753	.1210	.0589	.0677
		0 0 1	.3633	.3966	.3340	.4211	.2959	.3185
		0 0 0 1.5	.6912	.6996	.6689	.7701	.6332	.6563
		0 0 0 2	.8407	.8299	.8269	.9310	.8050	.8238
		0 0 0 3	.9776	.9681	.9691	.9776	.9558	.9703
10 12 14 16	1 1 2	0 0 0 .5	.1224	.0429	.0360	.0808	.0427	.0545
		0 0 0 1	.2352	.1132	.1063	.2014	.0831	.2416
		0 0 0 1.5	.4405	.2636	.2583	.3955	.2177	.5857
		0 0 0 2	.6832	.4807	.4810	.6188	.4265	.8567
		0 0 0 3	.9853	.8769	.8812	.9142	.8453	.9995

4. DISCUSSION

From the Table 1(a) and Table1(b), it is seen that when sample sizes are equal and variance also equal ,all the test statistics satisfies the nominal levels except Game and Howell(GH) which is slightly liberal(higher) than other tests. In case of equal sample sizes but unequal variance, empirical levels of T2,T3 and GH tests are slightly increases and higher than the nominal levels. When both sample sizes and variances are unequal we have observed (i) when sample sizes are in increasing order and variance are also in increasing order , empirical level of t and Sk test going down and not maintain the nominal levels. On the other hand when sample sizes are in increasing order and variance are in decreasing order empirical level of t and Sk test going upward and not maintained the nominal levels. Other six tests viz. GH, T2, T3, DN, BF and C tests more or less satisfies the nominals level.

Table 2(a) and Table2(b) , show the empirical power of the eleven test statistics for both equal and unequal

sample sizes and variances. We have seen that when sample sizes and variance are equal power of t- test is more than the other tests. Powers of other tests slightly vary from each others. In case of unequal sample sizes and variances , power of t-test although high ,it not satisfy the nominal levels .Power of Sk test less than the other tests . It is obvious as they not satisfies the nominal levels. Tamhane(T2,T3) and Dunn(Dn) test found to be more powerful in this situation than all other tests considered.

5. SUMMARY

Selecting an appropriate multiple comparison procedure requires an extensive assessment of available information regarding the testing situation. Information about the importance of type I errors, power, computational simplicity, and so on, are extremely important to the selection process. In addition, the selection of a proper multiple comparison procedure is dependent on data

conforming to validity assumptions, such as normality and variance homogeneity. Routinely selecting a procedure without careful consideration of available information and alternatives can severely reduce the reliability and validity of the results.

REFERENCES

1. Alexader, R.A. and Govern, D.M.(1994):A New and Simple Approximation for ANOVA under variance heterogeneity, Jour. EDu. Stat., 19, 91-101.
2. Patel, C.M.( ) :Parametric and Nonparametric Mltiple Comparisons using SAS , Orho Phamaceutical Cooperation.
3. Brown, M.B. and Forsythe,A.B.(1974): Robust Tests for the Equality of Variances,Jour. Amer. Stat. Assoc.,69,364-367.
4. Cochran,W.G.(1964): Approximate Significance Levels of Behrens-Fisher
5. Tests, Biometrics, 20,191-195.
6. Dunnett, C.W.(1955): A Multiple Comparison Procedure for Comparing Several Treatments with a control. Jour.Amer. Stat. Assoc., 50, 1096-1121.
7. Dunnett,C.W.(1964): New Tables for Multiple Comparisons with a Control,Biometrika,20, 482-491. Jour. Amer. Stat. Assoc., 75,796-800.
8. Dunn,O.J.(1964): Multiple Comparisons Using Rank Sums, Technometrics, 6,241-252.
9. Dunn,O.J.(1974): On Multiple Tests and Confidence Intervals, Comm. In Statistics,3,101-103.
10. Chen,S.Y. and Lee,S.H.(2011): Multiple Comparison Procedures under Heteroscedasticity,Tamkarg Jour. Sc. And Eng.,14,293-302.
11. Game,P.A.(1977): An Improved t table for Simultaneous Control on g Contrasts, Jour. Amer. Stat. Assoc.72,531-534.
12. Game,P.A. and Howell,J.F.(1976): Pairwise Multiple Comparison Procedures with Unequal n's and/or Variances: A Monte-Carlo Study, Jour. Of Educational Statistics, 1,113-125.
13. Hong,Li.(2012): A Multiple Comparison Procedure for Populations with Unequal
14. Variances,11(2),165-181.
15. Horchberg,Y.(1974): Some Generalizations of the T-Method in Simultaneous Inference, Journal of Number of Replications, Biometrics,12,307-310.
16. Kowalchuk,R.K.; Keselman,H.J.;Wilcox,R.R. and Algina,J.(2006): Multiple Comparison
17. Procedures, Trimmed means and Transformed Statistics, Jour. Modern.
18. Applied Stat. Methods,5,44-65.
19. Kruskal,W.H. and Wallis,W.A.(1952): Use of Ranks in one-Criterion Variance
20. Analysis,Jour.Amer.Stat. Assoc.,47,583-621.
21. Rafter,J.A.;Abel,M.L. and Braselton,J.P.(2002): Multiple Comparison Methods for Means, Siam Review,44,259-278.
22. Ramsey, P.H. Ramsey, P.P. and Barrera, K.(2010): Choosing the Best Pairwise Comparisons of mean
23. From non-normal Populations, with uequal variances, but equal sample sizes, Jour. Stat. Comp. Simul. 80(6), 595-608.
24. Stoline,M.R.(1978):Tables of the Studentized Augmented Range and Application to Problems of Multiple Comparison, Jour. Amer. Stat. Assoc.,73,656-660.
25. Stoline,M.R. and Ury, H.K.(1979): Tables of the Studentized Maximum Modulas Distribution and an Application to Multiple Comparisons Among Means,
26. Technometrics, 21, 87-93.
27. Schefe, H.(1953): A Method for Judging all Contrasts in the Analysis of Variance,
28. Biometrika, 40, 87-104.
29. Tamhane, A.C.(1977): Multiple Comparisons in Model I One-Way ANOVA With Unequal
30. Equal Variances, Comm. in Statist. , A6(1),15-32.
31. Tamhane,A.C.(1979):A Comparison of Procedures for Multiple Comparisons of Means With Unequal Variances, Jour. Amer. Stat. Assoc.,74,471-480.
32. Tukey, J.W.(1953): The Problem of Multiple Comparisons, A Unpublished report, Princeton University, Dept. of Statistics.
33. Welch,B.L.(1938): The Significance of the Difference Between Two Means When the
34. Population Variances are Unequal, Biometrika, 29, 350-352.