COMPARISON AND EVALUATION OF DIFFERENT POST-HOC TEST STATISTIC IN ENGINEERING AND EDUCATION USING RANDOMIZED COMPLETE BLOCK DESIGN UNDER ASSUMPTION OF EQUAL VARIANCE.

*[1]Bukar, S., [2]Ibrahim U.,[3] Dikko, H.G., [3]Tasiu, M., [3]Damisa A.S, [3]Falgore, J.Y.,

[1]Samaru College of Agriculture, Ahmadu Bello University, Zaria

[2]Department of Agronomy, Ahmadu Bello University, Zaria

[3]Department of Statistics, Ahmadu Bello University, Zaria

Correspondence Author: bukarsuleiman22@gmail.com 08065690543

ABSTRACT

The high percentage of error rate generated by post-hoc test is one of the major challenge of researchers across different fields of enquiry as well as the correct method to adopt, this has lead to under utilization of the numerous methods, this research, aimed at evaluating and comparing various methods to determine the most robust, convenient, optimal and most efficient in detecting the least percentage comparison wise and experimental wise error rates under the condition of homogeneity of variance, The research used secondary data from the field of Engineering and Education all in the form of Randomized Complete Block Design, the data were subjected to homogeneity of variance test and the result confirm that their variances were equal (α >0.05),On the same vein the result was able to find out that violation of the assumption for homogeneity of variance has no any significant effect on the comparison wise and experimental wise error rates,Similarly the EER generated by the various methods were all less than the alpha level of 10% (<0.10%) ,this shows that all the methods are best when the researcher is interested in EER, While for the CER, all the methods kept the error rate a little above ten percent (>10%), except for Tukey and SNK. Therefore, these methods can be adopted in the fields of Engineering and Education depending on the researcher's interest base on the error type, consequently when the researchers interest is to minimized both comparison and experimental wise error rate, the best methods across the research field is to adopt Tukey and SNK method as they kept the error rate below the chosen alpha level.

Key words: Comparison, Evaluation, Homoscedasticity, Post-Hoc, ANOVA, RCBD, Education and Engineering, Comparison and Experimental wise error rate.

1.0 INTRODUCTION

Researchers across different fields that perform analysis of variance (ANOVA) in their research may have heard of the term post-hoc test. It refers to "the analysis after the fact" and it is derived from the Latin word for "after that." Kim, (2017) The main reason behind performing a post-hoc test is that the conclusions derived from the ANOVA test have limitations. I.e, when the null hypothesis that declares the population means of three or more mutually independent groups are the same is rejected, the information obtained is not that the three groups are different from each other. It only provides information that the means of the three groups may differ and at least one group may show a difference. It does not provide information on which group differs from the other. As a result, the comparisons are made with different pairings of groups, undergoing an additional process of verifying which group differs from which other group. This process is referred to as the post-hoc test. Similarly, In an experiment with sequence of observations drawn from more than two groups independently e.g the variation of performance of different machines to execute a task within a specified period of time and the influence of weight to the distance covered by an athlete, the data collected from such experiment must satisfy the condition of assumption for normality and equal variance for an unbiased analysis i.e the observations on any particular treatment are independently selected from a normal distribution with variance σ^2 (the same variance for each treatment), and samples from different treatments are independent of one another Kiernan, (2014)., their differences can easily be obtain using students t-test when two samples are compared. However, when the groups are more than two (e.g. viability of six different seed and their yield per hectare) and the researcher may want to find their difference scientifically, one may not have any better option than to use the most commonly method known as analysis of variance (ANOVA) (Gelman *et al.,* 2012, Hedelsman, 2016). then followed by Post-Hoc method also known as multiple comparison. Researchers are not always interested in a single pair of comparisons per experiment but in the true picture of each and every variable to determine exactly the source and basis of the difference.

Sometimes a researcher may have to determine whether differences exist in the performance of for examples three different machines or more during a study, the null hypothesis (H_0) for the three machines and any analysis of variance (ANOVA) states that all means are equal while the alternative hypothesis (H_1) state otherwise (Ott, 1993, Lee & Lee, 2018). But the most common analytical method adopted by many areas of research is the one-way analysis of variance (ANOVA) instead of the two way. When the null hypothesis is rejected after the ANOVA, that is in the case of three groups, the null hypothesis declares that all the means are equal i.e, $H0 = A = B = C$, the researcher do not know how one group means differs from another. The result of ANOVA does not clearly show the detailed information about the difference among the different combinations of groups or means. Therefore, researchers usually perform additional analysis to find out the difference between particular pairs of experimental groups or means, which is known as multiple comparison or Post-Hoc (Abdi and Williams, 2010).

The Post-Hoc is an additional confirmatory tool that gives the exert differences among treatment which is only perform on the condition that the analysis of variance result have dictated a significant difference among the means (Srinivas et-al, 2015). There are several methods for multiple comparison procedure, but in our research we used the methods that are most appropriate to the field of Education and Engineering as well as their sensitivity to assumption violation over Bonferroni, Dunnett, Tukey, Newman keuls, Scheffes, Fisher least significant difference (LSD), Dunnett T3, Tamhane and Games Howel methods. However, each test has specific applications, advantages and disadvantages. LSD, Bornferroni, Scheffe, Tukey, DMRT and Newman Keuls, all these test are applied for independent comparison of sample means when condition of homogeneity of variance are fully satisfied, whereas

Dunnet`s, Tamhane and Games Howel are applied when the condition of homogeneity of variance are not satisfied or violated.

Empirically, several studies have been conducted in this concept by many authors, among which is Day and Quinn, (1989) studied Comparisons of Treatments after an ANOVA in Ecology. The assumption of the tests and variance equality for large sample sizes are crucial factors while utilizing this method. Scheffe's method is appropriate unless comparison could only have been pair wise. The issue of consistency in respect of Post-Hoc procedure was examined among the various Post-Hoc methods in the work of Saville, (2014), the concept of least significant procedure was recommended for general use with the condition that it should be view as hypothesis generator rather than as a method for simultaneous hypothesis generation and testing. Olorunju and Asiribo, (2003) and Sangseok and Dong, (2018) carried out a study in which they compared the effectiveness of some of the most frequently used methods for multiple treatment comparison in Agriculture, the methods considered were least significance difference (LSD), Duncan Multiple Range Test (DMRT), Tukey, Studentized-Newman keuls test (SNK), and scheffes Test. The research aimed at partitioning treatment into homogeneous groups and assess the performance of the methods. The study concluded that scheffe provide a better method of performing multiple comparison in term of partitioning treatment into homogeneous groups. However, this study only focuses on petitioning treatment into homogeneous group in the area of Agriculture only, instead of exploring other areas of research such as Education and Engineering. Ingersoll (2010) reviewed a limited set of Post-hoc techniques and general guidelines were provided that will accommodate equal and unequal sample sizes. Information regarding pair wise Post-hoc, Fisher test, Tukey test, Bonferroni test and Scheffe test were provided. For pair wise and preplanned pair wise comparisons Tukey and Dunn Bonferroni method is most appropriate respectively. McHugh (2011) provided information regarding Tukey, SNK, Scheffe, Bonferroni and Dunnett procedure which are mostly used. He found that some methods are best used for testing theory while others are useful in generating new theory. Selection of the appropriate post hoc test will provide researchers with the most detailed information while limiting Type I errors due to alpha inflation. Analysis of variance and post-hoc was used in medical research by kim, (2017) to show the effect of using t-test when comparison are more than two using conceptual figures, the figures clearly shows how ANOVA determines the mean difference problem by using between group and within group variance difference, it clearly show how the need for Analysis of variance arises from the error of alpha inflation, this increase the chance of type I error rates after the analysis, the findings clearly show that Post-Hoc is only possible after a significant f-value and they refer to it as "after that" which is after significant ANOVA test. Lee and Lee, (2018) studied four different Post-Hoc and the proper way of applying them in research, in their findings they test Tukey, Bonferroni, Dunnett and Scheffe methods, they only use Tukey method to demonstrate the practical application among the methods used in the research at 0.05 level of significance and examined how to maintain balance between error rate and power of the test when there are only three different hypothesis to be tested simultaneously in randomized complete design.

However, from the accessible studies, despite all efforts by this researcher but they could not make a practical application of all the methods used in their findings and most of the literature utilizes one way ANOVA which is not effective in reducing error rate. The literature also did not consider homogeneity of variance test and considered only few among the numerous post hoc methods. To fill in this gap, this study is informed. On this note, this research aimed to determine the effect of the error rate when condition of homogeneity of variance is violated and come up with a guide line that will help researchers in the various fields used, Academia, Statisticians, students of learning and indeed all other researchers to choose objectively from among the numerous multiple comparison procedure with the least percentage experimental and comparison wise error rate, so that there will be no basis for doubt about the appropriateness of the multiple comparison procedure adopted by any researcher, it's also our hope that this research will further create awareness and also motivate researchers on the frequent application of the various post-hoc methods in other not to make them disappear in literature completely due to lack of application in research fields.

2.0 MATERIALS AND METHODS

In this research we used randomized complete block design (CRBD) because it's the most commonly use types of design across different research field of inquiry which is due to its level of precision and numerous advantages Omer and Hussin, (2017) and it's one of the block design that the experimental elements are categorized into experimental units and each experimental units are known to be homogeneous in nature and further arrange into blocks, the condition of each experimental units within each block is ensured to be homogeneous as possible, in the same vein, variation may likely exist between block which is the basis for evaluation (Charyulu and Dharamyadav, 2013). This type of design is flexible with respect to different number of treatment and blocks and it provide more convincing result than that of complete randomized design (CRD) due to introduction of blocking which allows the computation of unbiased error for specific treatment, the choose of this type of design is also inline with the findings of (Omer & Hussin, 2017), The model for a completely randomized block design with i^{th} treatments and j^{th} block and n^{th} treatment in each Block can be written in the form of equation 2.1 (Einot et al., 1975).

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \qquad (2.1)$$

where the terms of the model are defined as follows:

y_{ij} : Observation on jth experimental unit receiving treatment i.

μ : Overall treatment means, an unknown constant.

α_i : An effect due to treatment i.

β_j : An effect due to j^{th} Block

e_{ij} : Random error associated with the response from the i^{th} and j^{th} experimental unit receiving treatment i.

We require that the errors have a normal distribution with mean 0 and a common variance, In addition, the errors must be independent.

Table: 2.1 The ANOVA table for two ways ANOVA is given below:

Source of variation	Degree of freedom	Sum of square	Means of square	F-ratio (F_i)
Treatments	$a-1$	SS_A	$SS_A/a-1$	$F_1=MS_A/MS_E$
Block	$b-1$	SS_B	$SS_B/b-1$	$F_2=MS_B/MS_E$
Error	$(a-1)(b-1)$	SS_E	$SS_E/(a-1)(b-1)$	
Total	$ab-1$	SS_T		

The sum of squares can be estimated using the computational formula below

SS(Total)= SS(Treatments) + SS(Blocks) + SS(Error)

$$CF=\frac{\Sigma Y^2..}{tr} \quad 2.2$$

$$SST=\sum_{k=1}^{r}\sum_{j=1}^{b}\sum_{i=1}^{a}\left(y_{ijk}-\bar{y}_{...}\right)^2 \quad 2.3$$

$$SSTR=rb\sum_{i=1}^{t}\left(\bar{y}_{i..}-\bar{y}_{...}\right)^2 \quad 2.4$$

$$SSB=ra\sum_{i=1}^{b}\left(\bar{y}_{.j.}-\bar{y}_{...}\right)^2 \quad 2.5$$

$$SSE=\sum_{k=1}^{r}\sum_{j=1}^{b}\sum_{i=1}^{a}\left(y_{ijk}-\bar{y}_{ij.}\right)^2 \quad 2.6$$

Test statistics for ANOVA (F-ratio) =

$$F_1=\frac{MS_A}{MSE} \quad 2.7$$

$$F_2=\frac{MS_B}{MSE} \quad 2.8$$

Decision rule = $\text{Re}\,ject\,H_0\,if\,F_i \succ F_{v1,v2},\ \forall i=1\,and\,2$

2.1 The basis for evaluation of the various Post-Hoc test procedures used are

The general acceptable method for evaluating the various utilities of Post-Hoc methods as adopted by many literature (Rodger and Roberts, 2013b) and Rodger and Roberts, 2013a) are fully utilized in this research and explained below; Conservativeness and Consistency, Simplicity and Convenience, Optimality and Flexibility, Robustness, Power Error rate

Conservativenes and Consistency: whenever analysis is carried out in whatever field of enquiry an inference is expected at the end of the findings, in such a situation an adopted procedure is known and is adopted as a criteria, once the process is violated, then the expected desired result may not be achieve, the conservativeness of a post-hoc method is the ability of the method to make strict statistical inference throughout an analysis, i.e the result of post-hoc method has significance result only when a certain level of control is available for the type I error, the method could produce reckless result when there are small difference among the means.

Optimality: the optimal statistic is the smallest confidence interval among the conservative statistic i.e the standard error is the smallest statistic among the conservative statistics.

Convenience: Literally this is considered as easy to use in estimating the statistic for the method as well as the procedures or steps involved, when a method have a complicated steps, it may tend to drive researchers away especially those in areas other than statistics.

Robustness: this has to do with the method being sensitive to assumption violation, some methods are very insensitive to assumption violation which may result in the increase in the probability of error during analysis of variation, because once assumption of equal variance is violated there are method that have been design to take care of such cases e.g Games Howell and Tamhani T3.

Power: the power of the any post-method is the ability of the procedure to produce the same result of the same data again and again i.e to produce minimal type I error and also a balance between the significant level and the error rate.

Error Rate: in the event of comparison of means using post-hoc, the probability of falsely declaring one pair of means different when in actual sense they are equal (Error of type-1) is substantially larger than the specified alpha level (Bender, & Lange, 2001). However, Selection of the most appropriate multiple comparison test is heavily influenced by the error rate, recall that a Type I error occurs when one incorrectly rejects a true H_0. The Type I error rate is the fraction of times a Type I error is made within a particular experiment. In a single comparison (image a simple t test) this is the value when comparing three or more treatment means, however, there are at least two different rates of Type I error: (Hayter,1986).

Comparison-wise Type I Error Rate (CER): The comparison-wise error rate (**∝c**) is the expected proportion of falsely rejected null hypothesis, in conducting all pairwise comparison in randomize block design, the comparison-wise error rate can be defined as: **∝c= ∝= P(Reject Hc |Hc is True).**

which can be obtain as the number of Type I errors divided by the total number of comparisons.

$$CER = \frac{number\ of\ errors}{total\ number\ of\ comparisons} \qquad 2.9$$

Experimental-wise Type I Error Rate (EER): its estimated as one minus the product of all paired wise significant values, I.e the product of probabilities of not committing type I error.

$$EER = 1 - \Pi(probability\ of\ not\ making\ type\ I\ error) \qquad 2.10$$

2.2 Data and method of Post-Hoc Used for the Study

The data used for this study was a research data (Secondary Data) collected from the field of Education and Engineering. The different data set was considered the best option because it is the one of the scientific way for evaluating the practical utilities of the various Multiple Comparison Procedures and their ability to control the error rate, Halldestam, (2016) each of the data set was subjected to analysis of variance and post-hoc test using the various methods in the research (Tukey, Scheffe, Duncan, Fisher LSD, Bonferroni, Dunnett T3, Games Howel and Tamhane) and all the criteria for evaluating them were examine to determine which among the methods gives the least error rate and the best among them.

2.3 Statistical Package and Level of Significance Used

SPSS version 25 was used for data analysis, this is because it's one of the Many statistical packages that offer most of the multiple comparison test as an option when conducting analysis of variance (ANOVA) (Olleveant *et al*., 1999). And the software can easily be adopted by researchers in different field of inquiry while the level of significance used is 10% as recommended by Lee and Lee (2018).

3.0 RESULTS AND DISCUSSIONS

Table 3.1: Levene's test result of data set collected from the field of Education.

Dependent Variable:weight			
F	df_1	df_2	Sig.
14.219	5	30	.061

Table 3.2: Results of analysis of variance for Education Data set.

Tests of Between-Subjects Effects (ANOVA TABLE)					
Dependent Variable:METERS					
Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	33976.991[a]	5	6795.398	1.225E2	.000
Intercept	911910.404	1	911910.404	1.643E4	.000
WEIGHT	33976.991	5	6795.398	1.225E2	.000
Error	1664.851	30	55.495		
Total	947552.245	36			
Corrected Total	35641.842	35			

The above tables (Tables 3.1 & 3.2), table 3.1 shows the result of the data set which test the effect of weight of athletes in relation to the distance they covered during a marathon, analysis was run to test for assumption of equal error variance, the significant value of 0.61 in the last column and degrees of freedom in column two shows that the data have satisfy the condition for normality of and variance equality, since the data have satisfy this condition the next is analysis of variance and the result is presented in table 3.2, the last column for significant value shows that all the different level of weight has a significant effect on the distance covered by all the athletes during the marathon because the significant value of 0.00 for both treatment, blocks and interaction has significant effect.

Table 3.3: Levene's test result of Engineering data set when subjected to test for assumption of equal error variance

Dependent Variable: METERS			
F	df_1	df_2	Sig.
.172	2	9	.845

Table 3.4: Results of analysis of variance result for Engineering Data Set.

Tests of Between-Subjects Effects					
Dependent Variable: Distance					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	484.607[a]	2	242.303	2.475E2	.000
Intercept	9622.003	1	9622.003	9.830E3	.000
speed	484.607	2	242.303	2.475E2	.000
Error	8.810	9	.979		
Total	10115.420	12			
Corrected Total	493.417	11			

The above tables (Table 3.3 & 3.4), table 3.3 shows the result of levenes test result of engineering data which test efficiency of irrigation machines at different pressure rate, when subjected to test for assumption of equal variance with degrees of freedom in column two while the significant value is 0.85 in the last column, this shows that the data have satisfy the condition for normality and the next table (table 3.4) shows the result for analysis of variance to further examines the mean differences from the significant value, base on the result it shows a clear difference in the efficiency and performance of all the machines

Table 3.5: Summary of percentage Experimental Wise and Comparison Wise Error Obtained from all the Comparison Methods when condition of normality are satisfied.

Post-Hoc Methods	EDUCATION		ENGINEERING	
% Error Rates	%CER	%EER	%CER	%EER
TUKEY	6.660	6.658	0	2.933
SCHEFFE	20.000	6.665	0	3.460
LSD	6.000	4.445	33.333	1.267
BONFERRONI	13.000	6.666	0	3.800
SNK	6.600	4.458	0	3.280
DUNCAN	6.600	4.537	33.300	3.060

Table 3.6: Percentage Error Rate obtained from the various Post-Hoc Methods When condition of normality are not satisfied.

Post-Hoc	EDUCATION		ENGINEERING	
% Error Rates	%CER	%EER	%CER	%EER
TAMHANE	6.600	6.661	33.333	5.533
DUNNETT T3	6.600	6.661	33.333	4.966
GAMES HOWEL	13.000	6.543	33.333	4.233

the above tables (table 3.5 and 3.6) shows the result of the of all percentage error rate generated by each of the comparison methods when equal variance are assumed and when not assumed respectively.

3.7 Result interpretation

3.8 Comparison and Experimental Wise Error Rate

From the result in (Table 3.5), this research can conclude that when experimental wise error rate is of concern then a researcher can adopt the use of any of the nine methods when number of treatment means to be compared are up to five (≥ 5) this is because the error rate is kept below the chosen alpha level of 0.01, and it's also within the error rate obtained in the findings of Lee and Lee (2018), but when the sample size are small (≤ 4) it was observed that Duncan and SNK perform best among other methods by keeping the error rate within alpha level of 10%. From the entire result we can deduce that when the sample size are large i.e above five treatment all the Post-Hoc method tends to give an error rate of less than 10% in the case where the researcher decide to use 0.01 level of significance and the result of this research have clearly shown that number of treatment has a significant effect on the experimental wise error rate as well as the Post-Hoc method, it was also observed that when the treatments number are less than or equal to four only Duncan and SNK keeps the error rate below 10% while other methods can only be adopted when chosen alpha is above 10% which is recommended in Education and Engineering literature.

Fisher least significant difference tends to give the least comparison wise error (0.33%) while Duncan and SNK gives error rate of (0.5%) when the treatments are four and five respectively, this also corroborate with the findings of Mohammad (2008) and Omer and Hussein (2017) that when treatments are less than six the comparison wise error rate tends to be higher than the experimental wise error rate.

3.9 EFFECT OF COMPARISON AND EXPERIMENTAL WISE ERROR RATE WHEN ASSUMPTION OF EQUAL VARIANCE ARE VIOLATED.

It was observed that the percentage EER and CER obtained by the three post-hoc methods when equal variance are not assumed (Table 3.6) Tamhane, Dunnett T3 and Games Howel are also a little above the chosen alpha level for experimental wise error rate while a small difference was observed in the comparison wise error rate, similarly, from the literature reviewed, there are no clear proof from any of the literature that differences exist in the percentage error rate when equal variance assumption are violated and this have further confirmed the findings of Jian and Richard (2011), that not all result from software are accurate that researchers should engage in finding out what the software does.

4.0 CONCLUSION AND RECOMMENDATION

The homogeneity of variance test is one of the major assumptions that need to carried out prior to analysis of variance test (Cohen, 2013), this assumption when violated has a significant effect on the error rate, fortunately all our data set satisfy this condition as the significant value from levene's test is greater than alpha value of (>0.05) (Table:3.1 and 3.3), the significant value also shows how close or farther the level of variability among the data set are and it has significant effect on the ANOVA and Post-hoc result, the closer the significant value to the chosen alpha level the higher the expected level of error (Saville, 2014).

The result showed that Tukey and SNK method has the smallest error rates within all the treatments number we therefore recommend that all researchers should adopt it especially when the chosen alpha level is 0.05 and 0.10, it was observed that Bonferroni and Tukey have the smallest confidence interval among all the Post-Hoc test, while all the test are considered to be convenient since the analysis are carried out using software, but for the sake of students of learning and using the principle of parsimony, LSD, Tukey and Duncan test statistics are less complex for manual computation, similarly the method with the less type I error rate at both treatments number (Tukey) is considered to be the most conservative methods (Lee and Lee, 2018).

Thereafter, we recommend that Tukey method should be use for both Education and Engineering research as the scientific method of examining means for better presentation and confirmation of group differences among the treatments means.

REFERENCES:

Abdi, H., & Williams, L. J. (2010). Tukey's Honestly Significant Difference (HSD) Test. Encyclopedia of Research Design. Thousand Oaks, CA: Sage, 1-5.

Bender, R., & Lange, S. (2001). Adjusting for multiple testing—when and how?. *Journal of clinical epidemiology*, *54*(4), 343-349.

Cohen J., (2013). *Statistical Power Analysis for the Behavioral Sciences. Academic Press New York, New York, USA*

Charyulu, B., & Dharamyadav, N. C. (2013). *Estimation of Missing Observation in Randomized Block Design*. ***1**:* 618–621.

Day, R. W., & Quinn, G. P. (1989). Comparisons of treatments after an analysis of variance in

Einot, I., & Gabriel, K. R. (1975). A study of the powers of several methods of multiple comparisons. *Journal of the American statistical association*, *70*(351a), 574-583.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of research on educational effectiveness*, *5*(2), 189-211.

Halldestam, B. M. (2016). ANOVA – The Effect of Outliers. *World Journal of Young Researchers*, 1(2): 28-31

Hayter, A. J. (1986). The maximum familywise error rate of Fisher's least significant difference test. *Journal of the american statistical association*, *81*(396), 1000-1004.

Hayter, A. J. (1989). Pairwise comparisons of generally correlated means. *Journal of the American Statistical Association*, *84*(405), 208-213.

Ingersoll R.M (2010), "Who controls Teachers Work ?" Cambridge, M.A : Harvard University Press.

Jian L., and Rechard G.(2011). What is your Statistical Software Actually Doing . *Journal of Experimental Education*, 79(3): 279-294

Kiernan, D. (2014). *Natural Resources Biometrics*. Open SUNY Textbooks, Milne Library, State University of New York at Geneseo.

Kim T.K (2017). Understanding One-Way ANOVA Using Conceptual Figures,. *Korean Journal of Anesthesiology*, 70: 21-26.

McHugh, M.L (2011) Multiple Comparison Testing in Anova*, Biochemia Medica:* 21(3);203-9

Muhammed, S.G (2008) *Evaluation of Multiple Comparison Procedures in Agricultural Experiment*. First Edition. Millenium Priniting and Publishing Company Limited

Olleveant, N.A. (1999) Multiple Comparison Test. *Blackwell Science Ltd, Journal of Clinical Nursing, 8: 299-304*

Olurunju, S.A.S and Asiribo O.E (2003). "Evaluation of Methods used for Multiple Comparisons of Means in Agricultural Experiments". MSc. Thesis, Department Of Mathematics Faculty Of Science Ahmadu Bello University

Omer, S. O., & Hussin, S. H. (2017). An Assessment of Multiple Comparisons for Single Variety Trial in Block Designs. *International Journal Of Multidisciplinary Research And Development* 2(1)197-200 .

Ott, L. (1993). An introduction to statistical methods and data analysis. Belmont, CA: Duxbury.

Rodger, R. S., & Roberts, M. (2013). Comparison of power for multiple comparison procedures. *Journal of Methods and Measurement in the Social Sciences*, *4*(1), 20-47.

Rodger, R. S., & Roberts, M. (2013b). *Comparison of Power for Multiple Comparison Procedures Contrasts and Alternatives*. *4*(1), 20–47.

Sangseok, L. And Dong, K.L (2018). What is the proper way to apply Multiple Comparison Test,. Korean Journal of Anesthesiology, 71(5), 353-360.

Saville, D.J. (2014) Multiple Comparison Procedures: The Practical Solution. The American Journal Statistics, 44, 174-180.

Srinivas, R, Bing Ho., Joseph Kang, Bruce Kaplan (2015). Post-Hoc Analysis After the Facts, *Journal of Transplantation Society*, 99 (1):17-20.

Tukey, J.W (1953). "*Problems of Multiple Comparisons*".Unpublished Manuscript, Princeton University, p 396.

Tukey, J.W (1962). "*The Future of Data Analysis*".Annals of Mathematical Statistics Vol. 33, pp1-67.