PAT June, 2021; 17(1): 60-69 ISSN: 0794-5213

Online copy available at *www.patnsukjournal.net/currentissue*

Publication of Nasarawa State University, Keffi

COMPARISON AND EVALUATION OF DIFFERENT POST-HOC TEST STATISTIC IN AGRICULTURE AND MEDICAL SCIENCE 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, and [3]Falgore, J.Y.

Samaru College of Agriculture, Ahmadu Bello University, Zaria
Department of Agronomy, Ahmadu Bello University, Zaria
Department of Statistics, Ahmadu Bello University, Zaria
Correspondence Author: bukarsuleiman22@gmail.com 08065690543

Abstract

The high percentage of error rate generated bypost-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 led to underutilization 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 Agriculture and Medical Sciences 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 percentage error rate generated by the various methods shows that Tukey is the best method that keeps the experimental wise error rate within the chosen alpha level of 0.05, While Duncan, SNK and Fisher LSD methods keeps the comparison wise error rate at a lower percentage error rate. Therefore, these methods can be adopted in all the fields of Agriculture and Medical Sciences depending on the researcher's interest base on the error type, consequently when the researcher interest is to minimized both comparison and experimental wise error rate, the best methods across the research field is to adopt Tukey method as they keep the error rate below the chosen alpha level of 0.05.

Key words: Comparison, Evaluation, Homoscedasticity, Post-Hoc, ANOVA, RCBD, Agriculture and Medical Science, Experimental wise error rate.

Introduction

In a research with sequence of observations drawn from two groups that satisfy the condition of assumption for normality and equal variance their differences can easily be obtain using students t-test. 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 difference exists in the means of three varieties of guinea corn at different agronomic practices or more varieties during a study, the null hypothesis (H_0) for any analysis of variance (ANOVA) clearly states that all means are equal or not equal (Ott, 1993), (Lee & Lee, 2018). But the most common analytical method for this is the one-way analysis of variance (ANOVA).When the null hypothesis is rejected after the one-way ANOVA, that is in the case of three groups $H_{o=}A = B = C$, the researcher do not know how one group means differs from another. The result of one –way 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 multiple comparison is an additional confirmatory tool which 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 Agriculture and Medical Science as well as their sensitivity to assumption violation (Bonferroni, Dunnett, Tukey, and Newman keuls, Scheffes, Fisher LSD, Dunnett T3, Tamhane and Games Howel). However, each test has specific applications, advantages and disadvantages. LSD, Bornferroni, Scheffe, Tukey, DMRT and newman Keuls, all this 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 satisfy or violated.

Empirically, several studies have been conducted in this concept by many authors, among which are Day and Quinn, (1989) studied Comparisons of Treatments after an ANOVA in Ecology. They recommended some important points like assumptions of the tests should be considered carefully, assumption of variance equality should be tested if sample sizes are large enough, unplanned comparison is made, 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, (2004), the concept of least significant procedure was recommended for general use with the condition that it should be view that as hypothesis generator rather than as a method for simultaneous hypothesis generation and testing. The findings also indicated the effect of scheffe test for general contrast.

Olorunju and Asiribo, (2003) 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 main consideration in the research was to find how the treatment could be partitioned into some homogeneous or near homogeneous groups that are very similar or not different from one another with respect to their performance. The study concluded that scheffe provide a better method of performing multiple comparison interm of partitioning treatment into homogenous groups which is also in line with the findings of (Sangseok and Dong, 2018). 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 Medical Sciences. Ingersoll (2010) reviewed a limited set of ANOVA Post-hoc techniques and general guidelines are provided that will accommodate most situations only i.e. equal and unequal sample sizes. He presented information regarding pair wise Post-hoc, Fisher test, Tukey test, Bonferroni test and Scheffe test. He gave general guideline for tests that for pair wise comparison Tukey method is appropriate, for pre-planned pair wise comparisons Dunn Bonferroni process is most appropriate, Weighted linear contrasts that are pre-planned should be tested using the Dunn Bonferroni method, Weighted linear contrasts that are unplanned should be tested using the Scheffe method.

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. Different tests can lead to different conclusions and careful consideration as to the appropriate test should be given in each circumstance, despite all efforts by this researcher but they could not make a practical application of all the methods used in their findings, in most cases researchers across different field of research as well as students of learning will always prefer to see practical application of methods to enable them understand it better. 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 clear 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 in spss to practically examine 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, the importance of effective presentation of research result of statistical analysis as a means of improving research quality and reducing error rate as aimed by much literature cannot be over emphasized. 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 (Tukey, Scheffe, Duncan, Fisher LSD, Bonferroni, Dunnett T3, Games Howel and Tamhane) 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

Materials and Methods

We decide to use RCBD because it's the most commonly use types of design across different research field of inquiry due to its numerous advantages (Omer and Hussin, 2017) and it's one of the complete block design that the experimental elements are categorized into experimental units and each experimental units are known to be homogeneous which are group in to blocks and the condition of each experimental units within each block is homogeneous as possible, but variation may 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 due to blocking it also allows the computation of unbiased error for specific treatment, this is also in line 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 (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 jth Block

e_{ij} : Random error associated with the response from the ith andjth 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 Block Error	$a-1$ $b-1$ $(a-1)(b-1)$	SS_A SS_B SS_E	$SS_A/a-1$ $SS_B/b-1$ $SS_E/(a-1)(b-1)$	$F_1=MS_A/MS_E$ $F_2=MS_B/MS_E$
Total	$ab-1$	SS_T		

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

$$SST = \sum_{i=1}^{t}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}\right)^2 \quad 2.3$$

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

$$SSB = \frac{\Sigma Y^2_{.j} - CF}{t} \quad 2.5$$

$$SSE = SST - SSTR \quad 2.6$$

Test statistics for ANOVA (F-ratio) =

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

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

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

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

Conservativeness 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

Literallythis 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 themethod beingsensitive 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 (Bandera and Langeb,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 obtained asthe number of Type I errors divided by the total number of comparisons.

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

Experimental-wise Type I Error Rate (EER)

it's estimated as one minus the product of all paired wise significant values, i.e. the product of probabilities of not committing type I error.

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

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 Agriculture and Medical Sciences. The different data set was considered the best option because we feel that it is the best 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 post-hoc test using the various methods used in the research (Tukey, Scheffe, Duncan, Fisher LSD, Bonferroni, Dunnett T3, Games Howel and Tamhane) and all the criteria shall be examine to determine which among the methods gives the least error rate and the best among the methods.

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 5% as recommended by (Lee and Lee. 2018)

Results and Discussions

Levene's test result of Agricultural data set when subjected to test for assumption of equal error variance

Dependent Variable:weight

F	df_1	df_2	Sig.
3.047	3	16	.059

Result of analysis of variance for Agricultural Data set.

Tests of Between-Subjects Effects

Dependent Variable:weight

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	64.714[a]	3	21.571	4.994E0	.012
Intercept	7189.632	1	7189.632	1.665E3	.000
Seed	64.714	3	21.571	4.994E0	.012
Error	69.110	16	4.319		
Total	7323.456	20			
Corrected Total	133.824	19			

Levene's test result of Medical Science data set when subjected to test for assumption of equal error variance

Dependent Variable: Delivery

F	df1	df2	Sig.
3.395	4	20	.082

Result of analysis of variance result for Medical Science Data Set.

Tests of Between-Subjects Effects

Dependent Variable: delivery

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5.781[a]	4	1.445	4.211	.012
Intercept	214.622	1	214.622	6.254E2	.000
Age	5.781	4	1.445	4.211	.012
Error	6.864	20	.343		
Total	227.268	25			
Corrected Total	12.645	24			

Summary Of Experimental Wise and Comparison Wise Error obtained from all the Comparison Methods

Post-Hoc Methods	AGRICULTURE		MEDICAL SCIENCES	
% Error Rates	%CER	%EER	%CER	%EER
TUKEY	66.600	16.256	70.000	9.990
SCHEFFE	66.600	16.290	70.000	9.999
LSD	33.330	10.860	70.000	9.774
BONFERRONI	66.600	16.490	70.000	2.028

SNK	83.300	7.295	50.000	5.098
DUNCAN	83.300	10.067	50.000	9.2418

(Percentage Error Rate obtained from the various Post-Hoc Methods when Equal Variance is not Assume)

Post-Hoc	AGRICULTURE		MEDICAL SCIENCES	
% Error Rates	%CER	%EER	%CER	%EER
TAMHANE	66.660	16.662	80.000	9.990
DUNNETT T3	66.660	16.648	70.000	4.991
GAMES HOWEL	66.660	16.480	70.000	9.999

Result interpretation

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 Bonferroni and SNK, when treatment means to be compared are up to five (≥ 5) this is because the error rate is kept below the chosen alpha level of 0.05, and it's also within the error rate obtained in the findings of (Mohammad, 2008), 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 not recommended in Agriculture and Medical Science 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), (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.

Effect of Comparison and Experimental Wise Error Rate when Assumption of Equal Variance is 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 of 0.05 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 has confirm the work of (Jian and Richard, 2011), that not all result from software are accurate that researchers should engage in finding out what the software does.

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, 2004).

The result showed that Tukey 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, 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 perssimony, 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 used for both Agricultural and Medical Science research as the scientific method for better presentation and confirmation of group differences among 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.

Bandera R, Langeb S. (2001) Adjusting for Multiple Testing When and How. "*Journal of Clinical Epidemiology* **54:** 343-349.

Cohen J., (2013). *Statistical Power Analysis for the Behavioural 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 Ecology. *Ecological Monographs*,***59:***(4), 433–463. https://doi.org/10.2307/1943075

Duncan, D. B. (1957). Multiple Range Test for Correlated and Heteroscedastic Means, Biometric s, 13: pp, 164-176

Dunnet, C.W (1980). "*Pairwise multiple Comparisons Homogeneous Variance and Unequal Sample Size Case*". JASA, 75, p 789.

Dunnet, C.W (1980). "*Pairwise Multiple Comparisons in the Unequal Variance Case*". J-Amar statistics Association, 75, p 769-800.

Einot, I., Gabriel, K. R., Einot, I., & Gabriel, K. R. (1975). *A Study of the Powers of Several Methods of Multiple Comparisons*. *Journal of the American statistical Association*, ***70**:* (351), 574–583.

Gelman A, Hill J, Yajima M, (2012). Why we (Usually) Don't have to worry About Multiple Comparison, Journal of research on Education Effectiveness, **5**: 189-211.

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

Hayter, A.J (1984). "*A Proof of the Conjecture that the Tukey-Kramer Method is Conservative*". The Annals of Statistics, **12**: 61-75.

Hayter, A.J. (1986). The Maximum Family Wise Error Rate of Fisher's Least Significant Difference Test. *Journal of the American Statistical Association*, **81**:1001-1004.

Hayter, A.J (1989). "*Pairwise Comparisons of Generally Correlated Means*".JASA, 84, pp 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

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*". In Press.

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. In press.

Rodger, R. S., & Roberts, M. (2013a). Comparison of Power for Multiple Comparison Procedures. *Journal of Methods and Measurement in the Social Sciences*, ***4*****(**1), 20. https://doi.org/10.2458/v4i1.17775

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. (2004) 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*, Vol.99, No. 1 (January, 2015), 17-20.

Steel, R.G.D and Torrie, J.H (1960)."*Principles and Procedures of Statistics*".McGraw Hill Books Co. New York.

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.www./t.nist.gov/div/898/hnadbook/prc/(Engineeringstatistichandbook).