THE POWER OF THE TEST FOR THE WINSORIZED MODIFIED ALEXANDER-GOVERN TEST

Tobi Kingsley Ochuko[1]; Suhaida Abdullah[2]; Zakiyah Zain[3]; Sharipah Syed Soaad Yahaya[4]

[1,2,3,4] College of Arts and Sciences, School of Quantitative Sciences,Universiti Utara Malaysia
06010 UUM Sintok, Kedah, Malaysia
[1] tobikingsley@rocketmail.com

ABSTRACT

This research examined the usage of the parametric method in comparing two or more means as independent group test, for instance, the Alexander-Govern (AG) test. The utilization of mean as the determinant for the center of distribution of variance diversity takes place in testing, and the test provides excellence in maintaining the amount of Type I error and giving immense sensitivity for a regular data. Unfortunately, it is ineffective on irregular data, leading to the application of trimmed mean upon testing as the determinant for the center of distribution under irregular data for two group condition. However, as the group quantity is more than two, the estimator unsuccessfully provides excellence in maintaining the amount of Type I error. Therefore, an estimator high in effectiveness called the MOM estimator was introduced for the testing as the determinant for the center of distribution. Group quantity in a test does not affect the estimator, but it unsuccessfully provides excellence in maintaining the amount of Type I error under intense asymmetry and unevenness. The application of Winsorized modified one-step M-estimator (WMOM) upon the Alexander-Govern testing takes place so that it can prevail against its drawbacks under irregular data in the presence of variance diversity, can eliminate the presence of the outside observation and can provide effectiveness for the testing on irregular data. Statistical Analysis Software (SAS) was used for the analysis of the tests. The results show that the AGWMOM test gave the most intense sensitivity under g = 0,5 and h = 0,5, for four group case and g = 0 and h = 0, under six group case, differing from three remaining tests and the sensitivity of the AG testing is said suffices and intense enough.

Keywords: *test power, Alexander-Govern (AG) test, the AGMOM test, AGWMOM test*

INTRODUCTION

In this study, the power of the test for the Alexander-Govern (AG) test, the modified one-step M-estimator in Alexander-Govern (AGMOM), the Winsorized modified one-step M-estimator in Alexander-Govern (AGWMOM), t-test and the ANOVA test for two, four and six group case with each of the g- and h- distribution is investigated.

The ANOVA has been applied in different fields of human endeavors, for instance in sociology, psychology, banking, marketing, medicine and agriculture as explained by Pardo *et al.* (1997). There are some hypotheses need to be considered for the ANOVA to perform properly, namely: normal distribution of the data, independent observations, and equality of the variance.

As discussed by Yusof, Abdullah, Yahaya, and Othman (2011), the ANOVA is seriously affected by heterogeneity of the variance and irregular data. Due to these, the amount of Type I error is seen to be increased, and the power of the test reduces.

The issue of variance diversity has been discussed by different researchers, and there has been an introduction of the alternatives to the ANOVA (Wilcox, 1988; Algina, Oshima & Algina, 1994; Lix, Keselman, & Keselman, 1996). Welch (1951) introduced the Welch test to put an experiment proving a hypothesis on two sample groups of equaling averages. This test has been mentioned in many kinds of literature as a better alternative to the ANOVA (Algina, Oshima & Lin, 1994; Lix, Keselman, & Keselman, 1996). For variance diversity, the Welch test provides excellence in maintaining the amount of Type I error.

It is advisable to use the parametric method that deals with heteroscedasticity. However, along with decreasing sample size and increasing group sizes, the Welch test unsuccessfully provides excellence in maintaining the amount of Type I error (Wilcox, 1988). James (1951) proposed a substitute for ANOVA, referred to as the James test. Sample means are weighed by this test which has been researched (Lix *et al.,* 1996; Oshima & Algina, 1992; Wilcox, 1988).

The James test cannot provide excellence in maintaining the amount of Type I error for a small sample size under irregular data. The Welch test and the James test are used for analyzing non-normal with variance diversity (Brunner, Dette & Munk, 1997; Krishnamoorthy, & Mathew, 2007; Wilcox & Keselman, 2003).

The Alexander-Govern (1994) discovered the Alexander-Govern test as a decent option for the Welch test, the James test, and the ANOVA because its test statistic is not complicated to obtain as described by Schneider and Penfield (1997). The usefulness of Alexander-Govern test is present when there is a violation on variance diversity in the hypothesis. Unfortunately, there are also some drawbacks. Lix and Keselman, (1998), Myers (1998), Schneider and Penfield (1997) discovered that the Alexander-Govern test is only effective for a regular data and is not for an irregular data. Their findings reveal that the test unsuccessfully provides excellence in maintaining the amount of Type I error for a regular data. It occurred that the test is ineffective on irregular data caused by using averages as the determinant for the center of distribution. The average is an extremely sensitive measurement with 0% breakdown point, such that if one data value is altered, the value of the average will be badly affected. Therefore, the mean cannot handle any occurrence of outliers and defies normality. To solve this problem, Lix and Keselman (1998) introduced the trimmed mean, that has been used in various statistical tests that base the average as the determinant for the center of distribution.

This shows that when trimmed mean is used, the problem of irregular data would be eliminated. Trimmed average replaces the usual average in the act of the determinant for the center of distribution in the Alexander-Govern test. Trimmed averages have been used by different researchers, because it is efficient and is reliable at providing excellence in maintaining the amount of Type I error (Keselman, Kowalchuk, Algina, Lix, & Wilcox, 2000; Luh, & Guo, 2005).

Trimmed average has drawbacks, namely: (1) the consideration of trimming percentage must be a priority, which would require an elimination process, (2) the trimming needs to be done properly, so it won't lose information, (3) trimmed mean can only handle the small size of values which are extreme (Yahaya, Othman, & Keselman, (2006). Researchers such as Abdullah, Yahaya, and Othman, (2007) provided a decent option to applying trimmed mean in Alexander-Govern test with an extraordinarily effective estimator, referred as the MOM. It was observed that for a skewed data, the MOM estimator provided excellence in maintaining the amount of Type I error. The MOM estimator is good at trimming data with extreme values with the consideration of characteristics of the distribution, whether it is slanted or not.

When it was introduced in the Alexander-Govern test, it provided excellence in maintaining the amount of Type I error, for a regular or greatly slanted data, but fails to do so under intense asymmetry and unevenness (Othman *et al.,* 2004).

The Winsorized MOM estimator was introduced in Alexander-Govern (AG) test to overcome the drawbacks of the test for irregular data, under variance diversity, in intense asymmetry and unevenness, to provide excellence in maintaining the amount of Type I error and to produce intensity in power for the test.

METHODS

The Alexander-Govern test was introduced by Alexander-Govern (1994). It serves a purpose for making a comparison of three or more groups where the utilization of the average as the determinant for the center of distribution for normal data under variance diversity takes place, but the test is ineffective on irregular data. The test statistic for the test is expressed with the use of the following procedures as listed below:

Firstly, the researchers order the data set with population sizes of j (j = 1, …, J). For each of the datasets, the mean is calculated by using the formula:

$$\bar{X} = \frac{\sum_j X_{ij}}{n_j}, \tag{1}$$

Where X_{ij} is defined as the known organized random sample with n_j as the sample size of the observations. The utilization of the average as the determinant for the center of distribution takes place in the Alexander-Govern test (1994). The usual unbiased estimate of the variance is defined using the formula:

$$s^2{}_j = \frac{\sum (X_{ij} - \bar{X}_j)^2}{n_j - 1}, \tag{2}$$

Where $\bar{X}_j$ is used for estimating μ_j with population j. The average's standard error is defined by:

$$S_{ej} = \sqrt{\frac{s^2{}_j}{n_j}}, \tag{3}$$

The weight (w_j) for the group sizes with population j of the known organized random sample is defined, where $\sum w_j$ must be equal to 1. The weight (w_j) for each of the independent groups is defined using the formula below:

$$w_j = \frac{1/S_e{}^2{}_j}{\sum_j 1/S_e{}^2{}_j}, \tag{4}$$

The null hypothesis testing by the Alexander-Govern (1994) for the equality of mean, with variance diversity is defined as:
H_o: $\mu_1 = \mu_2 = \ldots = \mu_j$
H_A: $\mu_1 \neq \mu_j$
For at least $i \neq j$

There is a contradiction between the statement of the alternative hypothesis with the null hypothesis. The variance impact is determined from the estimation of overall mean in the groups which belong in the organized data distribution, is described in the formula:

$$\hat{\mu} = \sum_{j=1}^{J} w_j \bar{X}_j \ , \tag{5}$$

Where, w_j it is the weight for each of the independent groups in the data distribution and $\bar{X}_j$ is the corresponding average in the independent groups in the known organized data sets. The t statistic for each of the independent groups is defined using:

$$t_j = \frac{\bar{X}_j - \hat{\mu}}{S_{ej}}, \tag{6}$$

Where $\bar{X}_j$ is the corresponding average in the independent group, $\hat{\mu}$ is defined as the overall grand average from each independent group with population j, the t statistic with $n_j - 1$ degree of freedom is obtained. Where ν is the degree of freedom for corresponding independent groups in the known organized data set. The t statistic defined for the corresponding groups are converted to standard normal deviates by using the Hill's (1970) normalization approximation in the Alexander-Govern (1994) technique.

The formula is defined using:

$$Z_j = c + \frac{[c^3 + 3c]}{b} - \frac{[4c^7 + 33c^5 + 240c^3 + 855c]}{[10b^2 + 8bc^4 + 1000b]}, \tag{7}$$

Where, $c = [a \times \log_e (1 + \frac{t^2{}_j}{\nu_j})]^{1/2}$, (8)

Where $\nu_j = n_j - 1$, $a = \nu_j - 0.5$, $b = 48a^2$ (9)

The test statistic for the AG test is defined using:

$$A = \sum_{j=1}^{J} Z^2{}_j \tag{10}$$

The test statistic for the AG test with a significance level of α = 0,05 at $(j-1)$ chi-square degree of freedom is chosen. When the p-value obtained for the AG test is > 0,05, the test is ineffective. Otherwise, the test would be effective.

Consider the known organized data sets to be defined as $X_1, X_2, ..., X_n$, with sample n and group sizes j. Then, the median is determined by designating the value in the middle of the observations. The MAD estimator sets the median of the absolute values of the differences between each of the score and the median. It is the median of $|X_j - M|$, …, $|X_n - M|$. Therefore, absolute deviation of the median (MAD_n) estimator is defined using the formula:

$$MAD_n = \frac{MAD}{0,6745} \ , \tag{11}$$

According to Wilcox and Keselman (2003), the constant value of 0,6745 is used for rescaling the MAD estimator with the aim of making the denominator to estimate σ when sampling from a normal distribution. Outliers in a data distribution can be detected using either:

$$\frac{|X_j - M|}{MAD_n} \succ K \ , \tag{12}$$

Alternatively, when $\frac{|X_j - M|}{MAD_n} \prec -K$, (13)

Where X_j is defined as the known organized random sample, M is the median of the ordered random samples and MAD_n is the median absolute deviation about the median. The value of K is 2,24. This value was proposed by Wilcox and Keselman (2003) for detecting the appearance of outliers in a data distribution because it has a very small standard error when the data sample is from a normal distribution.

Equation (12) and (13) are used for determining the appearance of outliers in a data distribution. In this research, there is a modification in which the average is utilized as the determinant for the center of distribution in the Alexander-Govern test, by replacing it with the Winsorized modified one-step M-estimator (WMOM) which utilizes mean as the determinant for the center of distribution of the test.

The WMOM estimator is applied to the data distribution, where the outlier value involved is replaced or exchanged with its predecessor most adjacent to where the point of the outlier is situated. The WMOM estimator is defined using the formula below:

$$WMOM = \bar{X}_{WMOMj} = \frac{\sum_{j=1}^{J} X_{WMOMj}}{n}, \tag{14}$$

The WMOM estimator is used as a substitute for the average as the determinant for the center of distribution in the Alexander-Govern test, because: (1) it eliminates the presence of outliers from the data distribution, (2) it makes the Alexander-Govern test effective on irregular data.

The Winsorized sample variance is defined by using:

$$S^2_{WMOMj} = \frac{\sum_{j=1}^{J} (X_j - \bar{X}_{WMOMj})^2}{n-1} \ , \tag{15}$$

Where $\bar{X}_j$ it is the known organized random sample and $\bar{X}_{WMOMj}$, is the Winsorized MOM estimator for the Winsorized data distribution. The standard error of the WMOM is determined by bootstrapping. The procedure for obtaining algorithm bootstrapped for the standard error estimation is defined as:

Firstly, selecting B independent bootstrap samples as defined below:
$x^{*1}, x^{*2}, ..., x^{*B}$, where each of these random samples having n data values with replacement from x as defined below:

$$x^* = (x_1, x_2, ..., x_n), \tag{16}$$

$$F \rightarrow (x_1^*, x_2^*, ..., x_n^*), \tag{17}$$

The symbol $(*)$ shows that x^* is not the actual data of x, but it is a randomized or resampled version of x. When estimating the standard error of the bootstrap samples, the number of B should be within the interval of (25 – 200). According to Efron and Tibshirani (1998), 50 samples of the bootstrap sample is sufficient to give a reasonable estimate of the standard error of the MOM estimator. In this research, 50 samples of the bootstrap samples were used for estimating the standard error of the MOM estimator.

Secondly, the copy of bootstrap which equals to each sampled bootstrap is expressed by using the formula below:

$$\hat{\theta}^*(b) = s(x^{*b})\, b = 1, 2, .., B. \tag{18}$$

Estimation of $t(\hat{F})$ use and the probability of $\frac{1}{n}$ distributes $\hat{F}$ empirically. For each of the known values is expressed as: $x_i\, , i = 1, 2, ..., n$.

Thirdly, the bootstrap estimate of $Se_F\ (\hat{\theta})$ is estimated from the sample standard deviation of the bootstrap replications that is expressed using:

$$\hat{se}_B = \{\sum_{b=1}^{B} [\hat{\theta}^*(b) - \theta^{/*}(\cdot)]^2 / (B-1)\}^{1/2}, \tag{19}$$

Where $\hat{\theta}^*(\cdot) = \sum_{b=1}^{B} \hat{\theta}^*(b) / B$ and $\hat{\theta}^* = s(x^*)$.

The weight w_j for the Winsorized data distribution for the corresponding independent group is defined as:

$$w_j = \frac{1 / S_e{}^2{}_{WMOMj}}{\sum_{j=1}^{J} 1 / S_e{}^2{}_{WMOMj}}, \tag{20}$$

Where $\sum_{j=1}^{J} 1 / S_e{}^2{}_{WMOMj}$ is the total of the squared standard error inversion for all the independent groups in the known organized random sample. Where $S_e{}^2{}_{WMOMj}$ is the standard error of the Winsorized data distribution and is defined as:

$$S_e{}^2{}_{WMOMj} = \frac{s_j{}^2{}_{WMOMj}}{n_j}, \tag{21}$$

The estimation of which the total mean in which the variance is weighted for the Winsorized data distribution for all the groups is defined by using:

$$\hat{\mu}_j = \sum_{j=1}^{J} w_j\, \bar{X}_{WMOMj}, \tag{22}$$

Where w_j is represented as the weight for the Winsorized data distribution and $\bar{X}_{WMOMj}$ is expressed as the mean of the Winsorized data distribution. The t statistic for each of the independent group is defined using the formula below:

$$t_j = \frac{\bar{X}_{WMOMj} - \hat{\mu}}{S_{eWMOMj}} \;, \tag{23}$$

Where $\bar{X}_{WMOMj}$ is the Winsorized MOM, $\hat{\mu}$ is the total mean for the Winsorized data distribution and lastly, S_e is the standard error of the Winsorized data distribution. In the Alexander-Govern technique, the t_j value is transformed to standard normal by using the Hill's (1970) normalization approximation and the hypothesis testing of the Winsorized sample variance of the WMOM estimator for μ_j is expressed as:

$H_O : \mu_1 = \mu_2 = ... = \mu_j$
$H_A : \mu_i \neq \mu_j$ For j = (j = 1, …., J)

The normalization approximation formula for the Alexander-Govern (AG) technique, with the use of the Winsorized Modified One Step M-estimator is expressed as:

$$Z_{WMOMj} = c + \frac{[c^3 + 3c]}{b} - \frac{[4c^7 + 33c^5 + 240c^3 + 855c]}{[10b^2 + 8bc^4 + 1000b]} \;,$$

Where $c = [a \times \log_e (1 + \frac{t^2{}_j}{v_j})]^{1/2}$,

$$v_j = n_j - 1 \;,\; a = v_j - 0{,}5 \;,\; b = 48a^2 \tag{24}$$

The test statistic of the Winsorized Modified One Step M-estimator in the Alexander-Govern test (AGWMOM) for all the independent groups in the known organized random data sample is expressed using the formula below:

$$AGWMOM = \sum_{j=1}^{J} Z^2{}_{WMOMj} \tag{25}$$

The test statistic for the AGWMOM test is obtainable using a chi-square distribution at $\alpha = 0.05$ the level of significance with J – 1 chi-square degree of freedom. The p-value is obtained from the standard chi-square distribution table. If the value of the test statistic for the AGWMOM is < 0,05, the test is considered to be effective. Otherwise, the test is regarded as ineffective.

In this research, five variables of different categories namely the balance condition of sample sizes, the equal of variance, group sizes, how they are paired and what kind the distribution is. Manipulation of variables is done to bring goodness and drawbacks of the AG test, the AGMOM test, the AGWMOM test, t-test and the ANOVA respectively.

Table 1 The Characteristics of the g- and h- Distribution

g- (Non-negative content)	h- (Non-negative content)	Skewness	Kurtosis	Types of Distribution
0	0	0	3	Standard normal
0	0,5	0	1198,20	Symmetric heavy tailed
0,5	0	1,81	18393,60	Skewed normal tailed
0,5	0,5	120,10	18393,60	Skewed heavy tailed

Source: Wilcox (1997)

For the AG test, the AGMOM test, the AGWMOM test, the t-test and the ANOVA, a testing was done to a good deal of considerable 5.000 data sets to give a satisfactory result for the effectiveness of the test of the five tests respectively. To obtain the pseudo-random variates, SAS generator RANNOR (SAS Institute, 1999) was used with a nominal level of α = 0n05 for the analysis of the tests in this research.

Table 2 The Research Design for Two Group Case for N = 40

The g- and h-distribution	Balanced and Unbalanced sample size	Variance ratio	Nature of Pairing	Notations for the Conditions
g = 0 and h = 0	20:20	1:1	Balanced condition	C1
		1:36	Positive Pairing	C2
	16:24	1:1		C3
		1:36	Positive Pairing	C4
		36:1	Negative Pairing	C5
g = 0 and h = 0.5	20:20	1:1	Balanced condition	C6
		1:36	Positive Pairing	C7
	16:24	1:1		C8
		1:36	Positive Pairing	C9
		36:1	Negative Pairing	C10
g = 0.5 and h = 0	20:20	1:1	Balanced condition	C11
		1:36	Positive Pairing	C12
	16:24	1:1		C13
		1:36	Positive Pairing	C14
		36:1	Negative Pairing	C15
g = 0.5 and h = 0.5	20:20	1:1	Balanced condition	C16
		1:36	Positive Pairing	C17
	16:24	1:1		C18
		1:36	Positive Pairing	C19
		36:1	Negative Pairing	C20

Table 3 Research Design for Four Groups Case for N = 80

The g- and h-distribution	Balanced and Unbalanced sample size	Variance ratio	Nature of Pairing	Notations for the Nature of Pairing
g = 0 and h = 0	20:20:20:20	1:1:1:1	Balanced condition	C21
		1:1:1:36	Positive Pairing	C22
		1:4:16:36	Positive Pairing	C22
	15:15:15:30	1:1:1:1		C24
		1:1:1:36	Positive Pairing	C25
		36:1:1:1	Negative Pairing	C26
		1:4:16:36	Positive Pairing	C27
		36:16:4:1	Negative Pairing	C28

Table 3 Research Design for Four Groups Case for N = 80 (Continued)

The g- and h-distribution	Balanced and Unbalanced sample size	Variance ratio	Nature of Pairing	Notations for the Nature of Pairing
g = 0 and h = 0.5	20:20:20:20	1:1:1:1	Balanced condition	C29
		1:1:1:36	Positive Pairing	C30
		1:4:16:36	Positive Pairing	C31
	15:15:20:30	1:1:1:1		C32
		1:1:1:36	Positive Pairing	C33
		36:1:1:1	Negative Pairing	C34
		1:4:16:36	Positive Pairing	C35
		36:16:4:1	Negative Pairing	C36
g = 0.5 and h = 0	20:20:20:20	1:1:1:1	Balanced condition	C37
		1:1:1:36	Positive Pairing	C38
		1:4:16:36	Positive Pairing	C39
	15:15:20:30	1:1:1:1		C40
		1:1:1:36	Positive Pairing	C41
		36:1:1:1	Negative Pairing	C42
		1:4:16:36	Positive Pairing	C43
		36:16:4:1	Negative Pairing	C44
g = 0.5 and h = 0.5	20:20:20:20	1:1:1:1	Balanced condition	C45
		1:1:1:36	Positive Pairing	C46
		1:4:16:36	Positive Pairing	C47
	15:15:20:30	1:1:1:1		C48
		1:1:1:36	Positive Pairing	C49
		36:1:1:1	Negative Pairing	C50
		1:4:16:36	Positive Pairing	C51
		36:16:4:1	Negative Pairing	C52

Table 4 Research Design for Six Groups Case for N = 120

The g- and h-distribution	Balanced and Unbalanced sample size	Variance ratio	Nature of Pairing	Notations for the Nature of Pairing
g = 0 and h = 0	20:20:20:20:20:20	1:1:1:1:1:1	Balanced condition	C53
		1:1:1:1:1:36	Positive Pairing	C54
		1:4:4:16:16:36	Positive Pairing	C55
g = 0 and h = 0	2:4:4:16:32:62	1:1:1:1:1:1		C56
		1:1:1:1:1:36	Positive Pairing	C57
		36:1:1:1:1:1	Negative Pairing	C58
		1:4:4:16:16:36	Positive Pairing	C59
		36:16:16:4:4:1	Negative Pairing	C60
g = 0 and h = 0.5	20:20:20:20:20:20	1:1:1:1:1:1		C61
		1:1:1:1:1:36	Positive Pairing	C62
		1:4:4:16:16:36	Positive Pairing	C63
	2:4:4:16:32:62	1:1:1:1:1:1		C64
		1:1:1:1:1:36	Positive Pairing	C65
		36:1:1:1:1:1	Negative Pairing	C66
		1:4:4:16:16:36	Positive Pairing	C67
		36:16:16:4:4:1	Negative Pairing	C68

Table 4 Research Design for Six Groups Case for N = 120 (Continued)

The g- and h-distribution	Balanced and Unbalanced sample size	Variance ratio	Nature of Pairing	Notations for the Nature of Pairing
g = 0.5 and h = 0	20:20:20:20:20:20	1:1:1:1:1:1	Balanced condition	C69
		1:1:1:1:1:36	Positive Pairing	C70
		1:4:4:16:16:36	Positive Pairing	C71
	2:4:4:16:32:62	1:1:1:1:1:1		C72
		1:1:1:1:1:36	Positive Pairing	C73
		36:1:1:1:1:1	Negative Pairing	C74
		1:4:4:16:16:36	Positive Pairing	C75
		36:16:16:4:4:1	Negative Pairing	C76
g = 0.5 and h = 0.5	20:20:20:20:20:20	1:1:1:1:1:1	Balanced condition	C77
		1:1:1:1:1:36	Positive Pairing	C78
		1:4:4:16:16:36	Positive Pairing	C79
		1:1:1:1:1:1		C80
		1:1:1:1:1:36	Positive Pairing	C81
	2:4:4:16:16:32:62	36:1:1:1:1:1	Negative Pairing	C82
		1:4:4:16:16:36	Positive Pairing	C83
		36:16:16:4:4:1	Negative Pairing	C84

Source: Ochuko, Abdullah, Zain & Yahaya (2015)

It is necessary that the power be at more than 0,5 and can be considered adequate when it stands at the point of 0,8 and above (Murphy & Myors, 1998). The likelihood of successfulness will be at the quadruple amount of certainty if the power is 0,8. However, if the power sits on 0,9, then the successfulness would be at nonuple of the certainty.

Table 5 Pattern of Variability for the Effect Size Index of 4 and 6 Groups

The Effect Size Index	For J = 4	For J = 6
Small	$-\frac{1}{2}d, 0, 0, \frac{1}{2}d$	$-\frac{1}{2}d, 0, 0, 0, 0, \frac{1}{2}d$
Medium	$-\frac{1}{2}d, -\frac{1}{4}d, \frac{1}{4}d, \frac{1}{2}d$	$-\frac{1}{2}d, -\frac{1}{3}d, -\frac{1}{6}d, \frac{1}{6}d, \frac{1}{3}d, \frac{1}{2}d$
Large	$-\frac{1}{2}d, -\frac{1}{2}d, \frac{1}{2}d, \frac{1}{2}d$	$-\frac{1}{2}d, -\frac{1}{2}d, -\frac{1}{2}d, \frac{1}{2}d, \frac{1}{2}d, \frac{1}{2}d$

Source: Cohen, (1988)

The power of the tests is represented graphically, where the y-axis corresponds to the power of the tests, and the horizontal axis represents the effect size index d for two groups condition and f for more than two group condition. The graph is used to show the trends of the power of the tests on the effect size index. According to scholars such as Murphy and Myors (1998) the power of a test must be above 0,5. It can be considered sufficient and high when its value is 0,8 and above.

The graph shows those tests that have low power, sufficient and high power on the effect size indexes (d and f). In this research, the effect size index was used for analyzing the power of the test of the five different tests accordingly.

RESULTS AND DISCUSSIONS

The Power of the Test for the AG Test, the AGMOM Test, the AGWMOM Test, the T-Test and the ANOVA

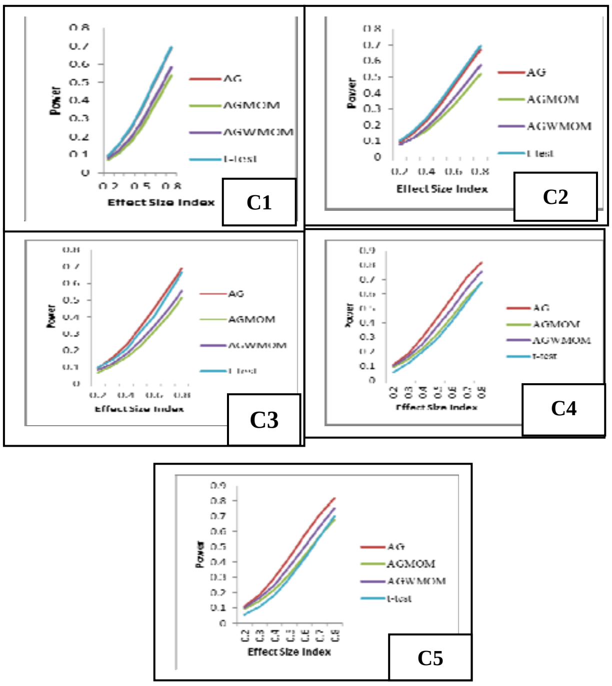

Figure 1 Graphical representation of g = 0 and h = 0, of Power against Effect Size Index, for Two Group Condition

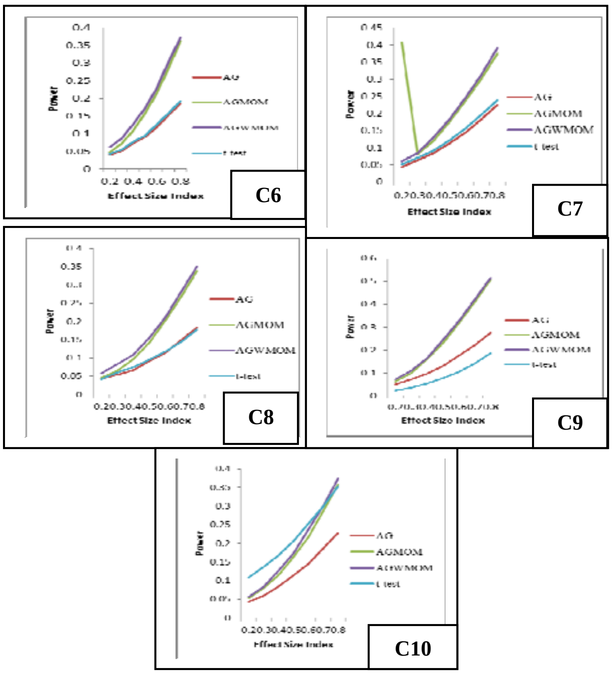

Figure 2 Graphical Representation of g = 0 and h = 0,5, of Power against the Effect Size Index, for Two Group Condition

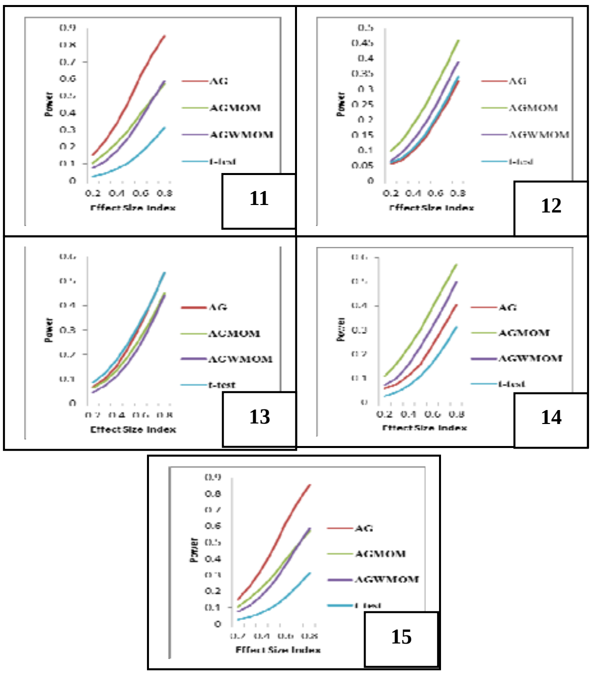

Figure 3 Graphical Representation of g = 0,5 and h = 0, of Power versus Effect Size Index, for Two Group Case

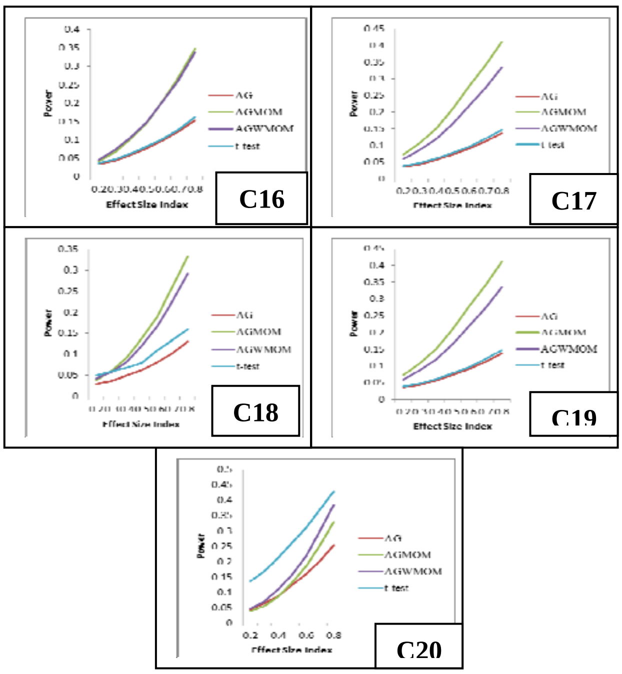

Figure 4 Power against Effect Size Index, for Two Groups Case, For g = 0 and h = 0

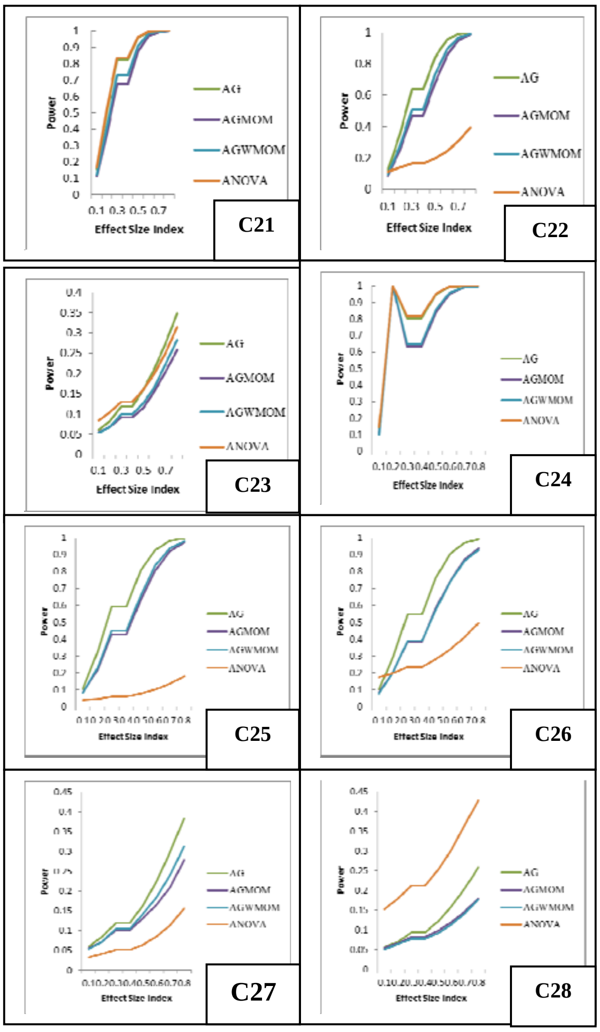

Figure 5 Power versus Effect Size Index, for Four Groups Condition, Under a Normal Distribution

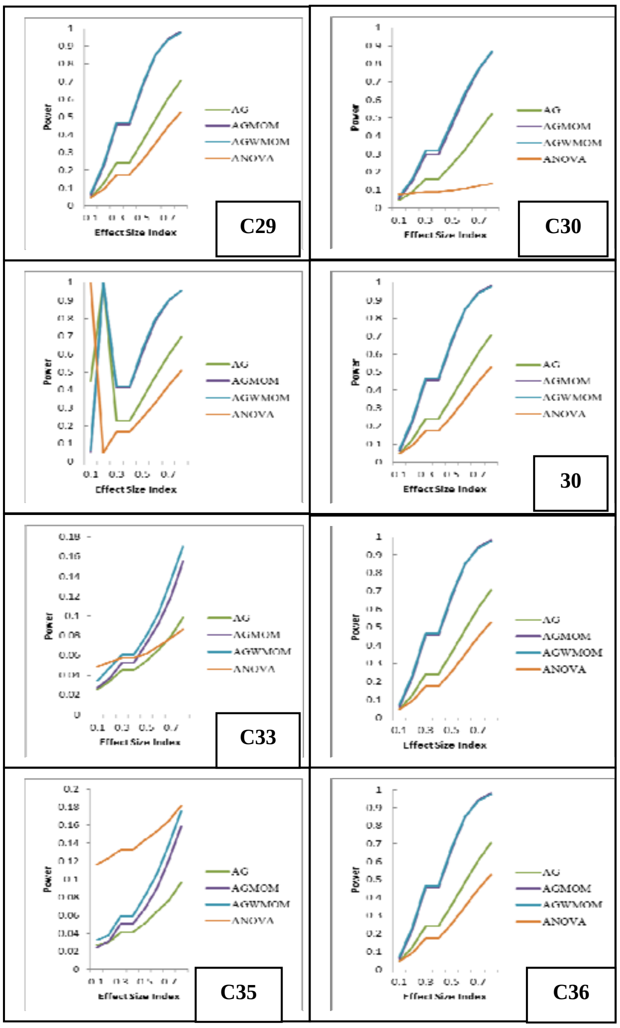

Figure 6 Power against Effect Size Index, for Four Group Case, For g = 0 and h = 0.5

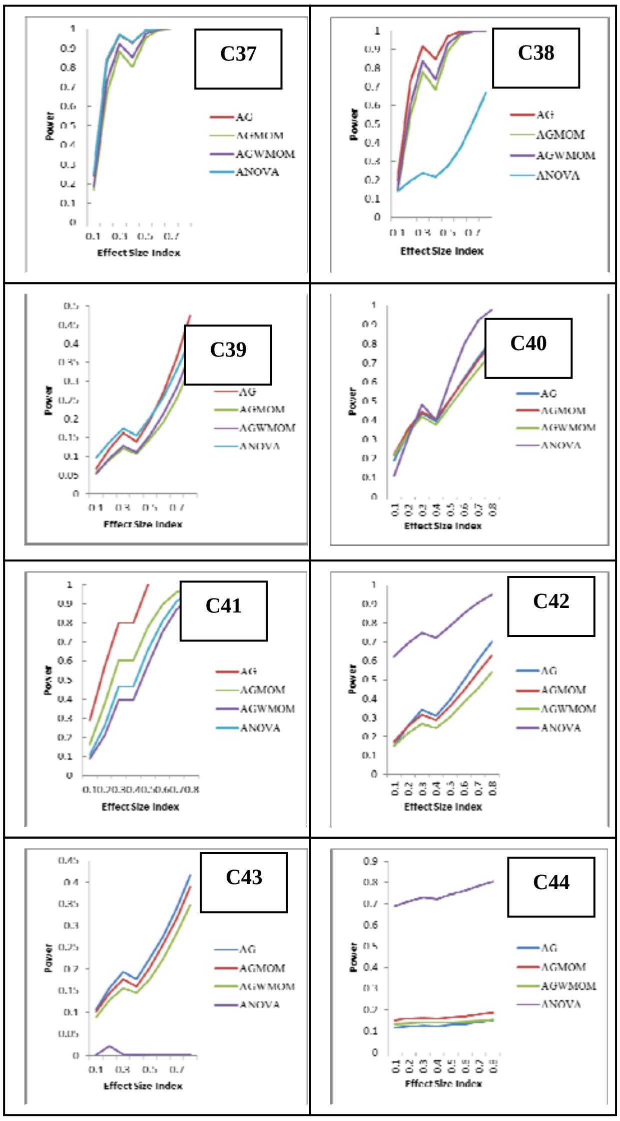

Figure 7 Graphical Representation of g = 0,5 and h = 0, of Power against Effect Size Index, for Four Groups Condition

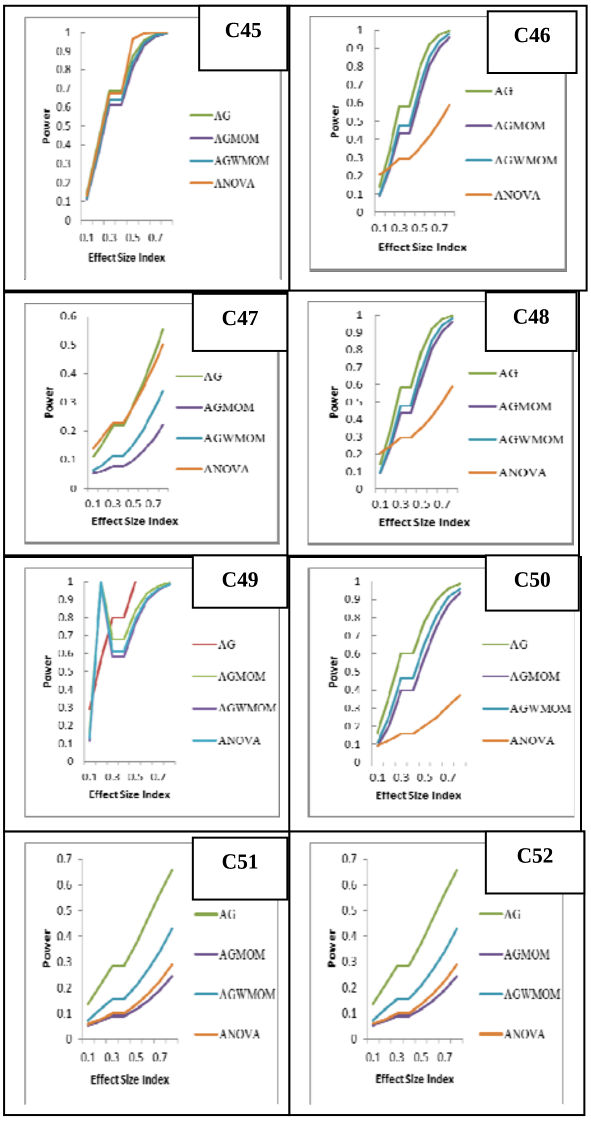

Figure 8 Graphical Representation of g = 0,5 and h = 0,5, of Power against Effect Size Index, for Four Groups Condition

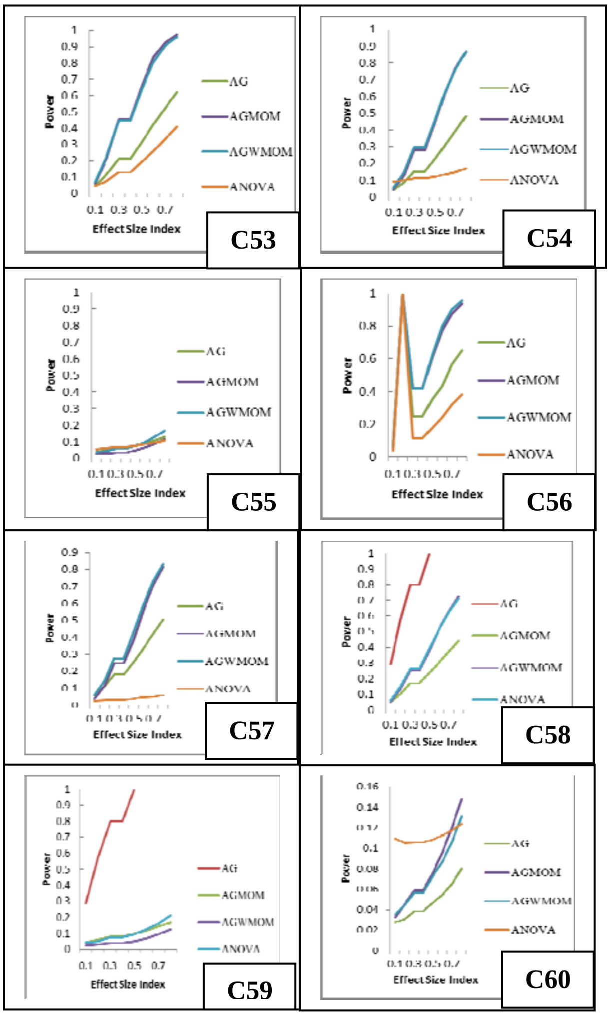

Figure 9 Graphical Representation of g = 0 and h = 0, of Power against Effect Size Index, for Six Group Condition

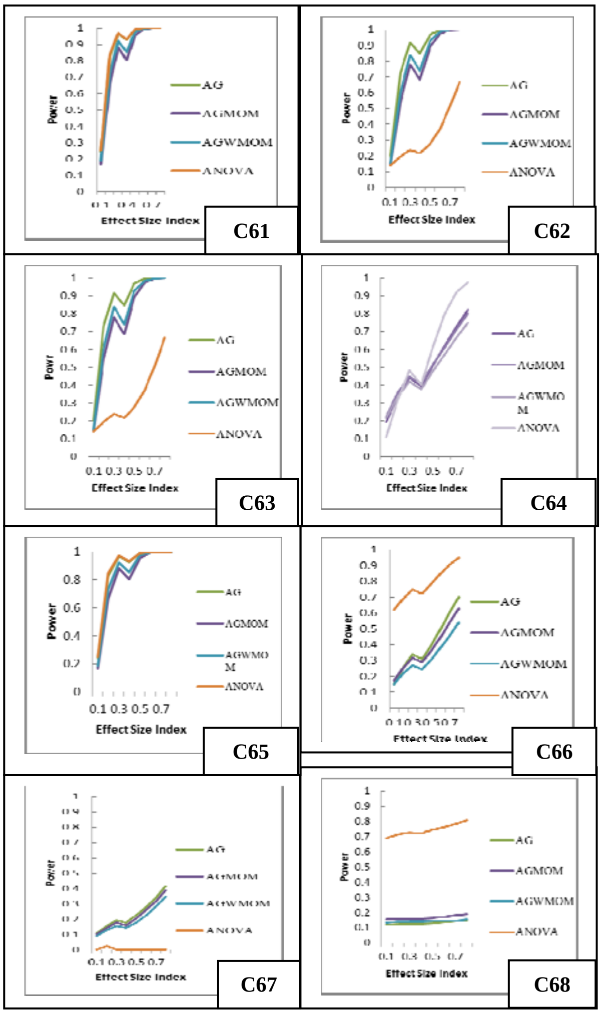

Figure 10 Graphical Representation for g = 0 and h = 0,5, of Power against Effect Size Index, for six Groups Conditions

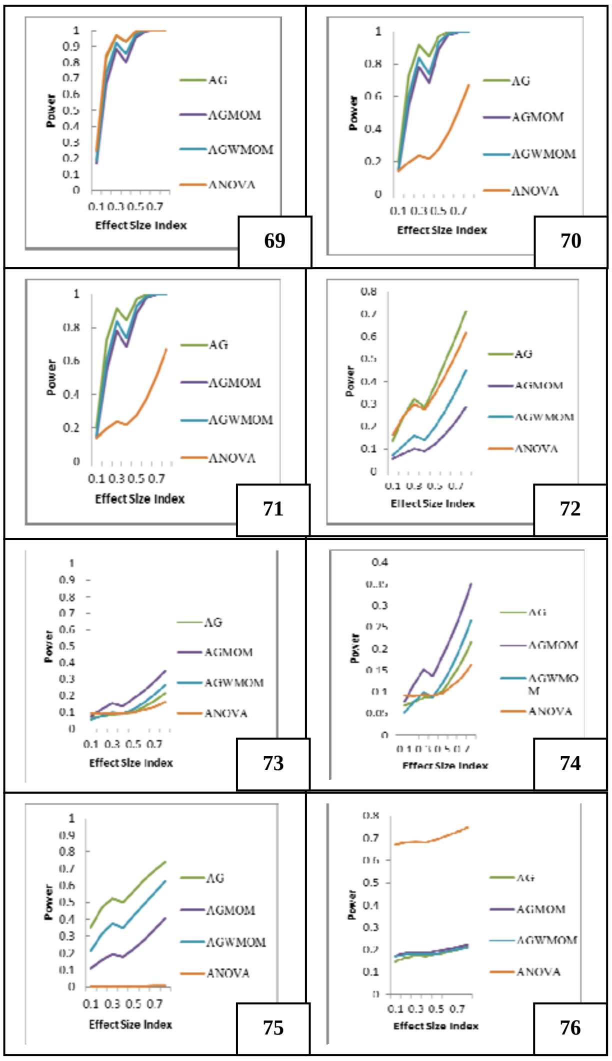

Figure 11 Graphical Representation for g = 0,5 and h = 0, of Power against Effect Size Index, for Six Groups Condition

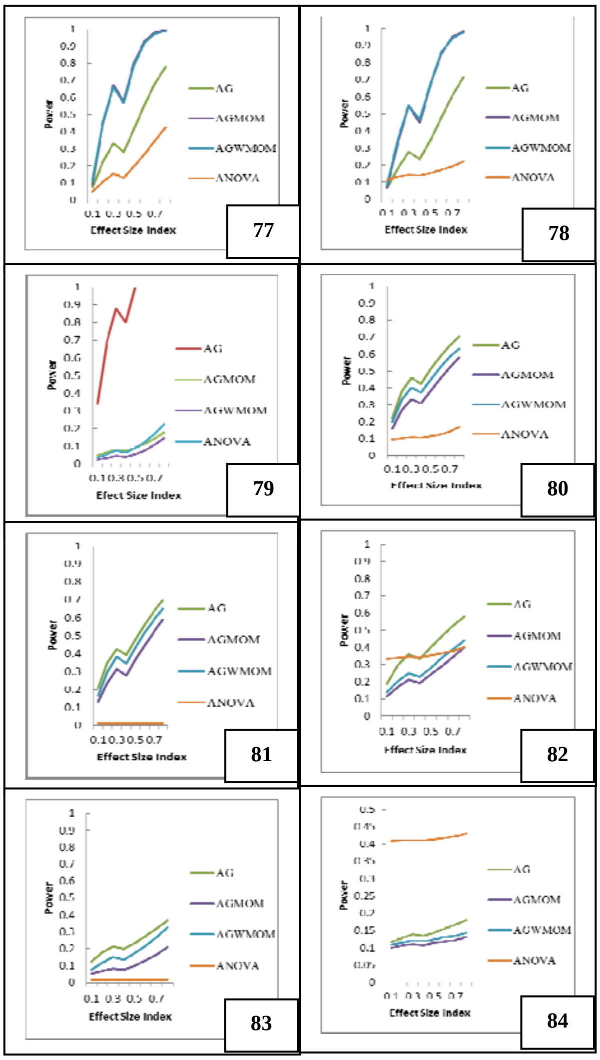

Figure 12 Graphical representation for g = 0.5 and h = 0.5, of Power against Effect Size Index, for Six Groups Condition

For g = 0 and h = 0, g = 0 and h = 0,5 and g = 0,5 and h = 0, under two group case, the power of the AG test, the AGMOM test, the AGWMOM test and the t-test is increasing as the effect size index is increasing. For g = 0,5 and h = 0,5, the AGWMOM test has the highest amount of power compared to the other three tests under this condition, the power of the AGWMOM test is above 0,8 and is regarded as high and sufficient. In C37, C38, C39, C40 and C41, the power of the four tests is above 0.5 and is considered to be sufficient. In C53, C54 and C57, under six group case, for g = 0 and h = 0, the AGWMOM test has the highest power compared to the other three tests and the power of the test is said to be sufficient and high. For g = 0,5 and h = 0,5, in C77 and C78, under six group case, the AGWMOM test produced the highest power compared to the other three tests and the power of the test is referred to as sufficient and high.

CONCLUSIONS

The AGWMOM test produced the highest power for g = 0,5 and h = 0,5, under four group case and g = 0 and h = 0, under six group case in comparison to the other three tests and the power of the test is referred to as sufficient and high.

REFERENCES

Abdullah, S., Yahaya, S. S. S., & Othman, A. R. (2007). Modified one step m-estimator as a central tendency measure for alexander-govern test. In *Proceedings of The 9th Islamic Countries Conference on Statistical Sciences,* 834-842.

Alexander, R. A., & Govern, D. M. (1994). A new and simpler approximation for ANOVA under variance heterogeneity. *Journal of Education Statistics, 19*(2), 91-101.

Algina, J., Oshima, T. C., & Lin, W-Y. (1994). Type I error rates for Welch's Test and James's second-order test under nonnormality and inequality of variance when there are two groups. *Journal of Educational and Behavioral Statistics, 19*(3), 275-291.

Brunner, E., Dette, H., & Munk, A. (1997). Box-type approximations in nonparametric factorial designs. *Journal of the American Statistical Association, 92*(440), 1494-1502.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* New York: Chapman & Hall.

Efron, B., & Tibshirani (1998). *An introduction to the bootstrap*. New York: Chapman & Hall.

Hill, G. W (1970). Algorithm 395 student's t-distribution. *Communications of the ACM, 13*, 617-619.

James, G. S. (1951). The comparison of several groups of observations when the ratios of the population variances are unknown. *Biometrika, 38*, 324-329.

Keselman, H. J., Kowalchuk, R. K., Algina, J., Lix, L. M., & Wilcox, R. R. (2000). Testing treatment effects in repeated measure designs: Trimmed means and bootstrapping. *British Journal of Mathematical and Statistical Psychology, 53,* 175-191.

Krishnamoorthy, K., F., & Matthew, T. (2007). A parametric bootstrap approach for ANOVA with unequal variances: Fixed and random models. *Computational Statistics & Data Analysis, 51*(12), 5731-5742.

Lix, Lisa, M., & Keselman, J.C., & Keselman, H. J (1996). Approximate degrees of freedom tests: A unified perspective on testing for mean equality. *Psychological Bulletin, 117*(3), 547-560.

Lix, L. M, & Keselman, H. J. (1998). To trim or not to trim. *Educational and Psychological Measurement, 58*(3), 409-429.

Luh, W. M., & Guo, J. H. (2005). Heteroscedastic test statistics for one-way analysis of variance: The trimmed means and Hall's transformation conjunction. *The Journal of Experimental Education, 74*(1), 75-100.

Myers, L. (1998). Comparability of the James' Second-Order approximation test and the Alexander and Govern: A statistic for Non-normal Heteroscedastic data. *Journal of Statistical Simulation Computation, 60*, 207-222.

Murphy, K. R., & Myors, B. (1998). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests*. Mahwah, NJ: Lawrence Erlbaum.

Ochuko, T. K., Abdullah S., Zain, Z., & Yahaya, S. S. S., (2015). The modification and evaluation of the Alexander-Govern test in terms of power. *Modern Applied Science, 9*(13), 1-21.

Oshima, T. C., & J. Algina (1992). Type I error rates for James's second-order test and Wilcoxon's Hm test under heteroscedasticity and non-normality. *British Journal of Mathematical and Statistical Psychology, 45*, 255-263.

Othman, A. R., Keselman, H. J., Padmanabban, A. R., Wilcox, R. R., Wilcox, R. R., & Fradette, K. (2004). Comparing measures of the "typical" score across treatment groups. *The British Journal of Mathematical and Statistical Psycholofy, 57*(2), 215-234.

Pardo, J. A, Pardo, M. C., Vincente, M. L., & Esteban, M. D. (1997). A statistical information theory approach to compare the homogeneity of several variances. *Computational Statistics & Data Analysis, 24*(4), 411-416.

SAS Institute Inc. (1999). *SAS/IML User's Guide Version 8. Cary.* NC: SAS Institute Inc.

Schneider, P. J., & Penfield, D. A. (1997). Alexander-Govern's approximation: Providing an alternative to ANOVA under variance heterogeneity. *Journal of Experimental Education, 65*(3), 271-287.

Welch, B. L. (1951). On the comparison of several means: An alternative approach. *Biometrica, 38*, 330-336.

Wilcox, R. R. (1988). A new alternative to the ANOVA F and new results on James's second-order method. *British Journal of Mathematical and Statistical Psychology, 41*, 109-117.

Wilcox, R. R. (1997). *Introduction to robust estimation and hypothesis testing*. San Diego, CA: Academic Press.

Wilcox, R. R., & Keselman, H. J. (2003). Modern robust data analysis methods: Measures of central tendency. *Psychological Methods, 8*(3), 254-274.

Yahaya, S. S. S., Othman, A. R., & Keselman, H. J. (2006). Comparing the "typical score" across independent groups based on different criteria for trimming, *Metodološki Zvezki, 3*(1), 49-62.

Yusof, Z., Abdullah, S. & Yahaya, S. S. S. (2011). Type I error rates of F_t statistic with different trimming strategies for two groups case. *Modern Applied Science, 5*(4), 1-7.