Modern Applied Science; Vol. 9, No. 12; 2015
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Modifying and Evaluating the Alexander-Govern Test Using Real Data

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

[1, 2, 3, 4] College of Arts and Sciences, Universiti Utara Malaysia, Malaysia

Correspondence: Tobi Kingsley Ochuko, College of Arts and Sciences, Universiti Utara Malaysia, Malaysia. E-mail: tobikingsley@rocketmail.com

Received: July 8, 2015 Accepted: August 19, 2015 Online Published: October 12, 2015

doi:10.5539/mas.v9n12p1 URL: http://dx.doi.org/10.5539/mas.v9n12p1

Abstract

This study examines the use of independent group test of comparing two or more means by using parametric method, such as the Alexander-Govern (*AG*) test. The Alexander-Govern test is used for comparing two or more groups and is a better alternative compared to the James test, the Welch test and the *ANOVA*. This test has a good control of Type I error rates and gives a high power under variance heterogeneity for a normal data, but it is not robust for non-normal data. As a result, trimmed mean was applied on the test under non-normal data for two group condition. But this test could not control the Type I error rates, when the number of groups exceed two groups. As a result, the *MOM* estimator was introduced on the test, as its central tendency measure and is not influenced by the number of groups. But this estimator fails to give a good control of Type I error rates, under skewed heavy tailed distribution. In this study, the *AGWMOM* test was applied in Alexander-Govern test as its central tendency measure. To evaluate the capacity of the test, a real life data was used. Descriptive statistics, Tests of Normality and boxplots were used to determine the normality and non-normality of the independent groups. The results show that only the group middle is not normally distributed due extreme value in the data distribution. The results from the test statistic show that the *AGWMOM* test has a smaller p-value of 0.0000002869 that is less than 0.05, compared to the *AG* test that produced a p-value of 0.06982, that is greater than 0.05. Therefore, the *AGWMOM* test is considered to be significant, compared to the *AG* test.

Keywords: Alexander-Govern test, *MOM* estimator, *AGWMOM* test, descriptive statistics, tests of normality, boxplots and test statistic

1. Introduction

In comparing independent group means, the analysis of variance is applicable in different aspects of life, such as in sociology, agriculture, economics and in medicine, as explained by Pardo, Pardo, Vincente and Esteban (1997). The three main assumptions that must be fulfilled before the *ANOVA* can perform effectively are: (i) homogeneity of the variance (ii) normality of the data and (iii) independent observations.

The *ANOVA* is a classical group test that is used for comparing three or more means. The *ANOVA* is very sensitive to the assumptions of homogeneity of the variance. In a situation where there is a violation in the assumptions, it would affect the authenticity of the test and thereby the p-value may become too conservative or large (Brown & Forsythe, 1974; Wilcox, Charlin & Thompson, 1986).

Welch (1951) proposed the Welch test to solve the problem of heterogeneity of variance. This test modifies the calculation of the degree of freedom in the common *F* test. For unequal variance, the Welch test gives a good control of Type I error rates, but fails to control the Type I error rates when the group sizes increases (Wilcox, 1988). James (1951) introduced the James test. According to Lix et al. (1996), Oshima and Algina (1992) and Wilcox (1988) stated that the James test is used for weighting sample means and is a better alternative to the *ANOVA* under variance heterogeneity. But this test fails to give a good control of Type I error rates for a small sample size. The Welch test and the James test are used in analyzing a data distribution that is non-normal under variance heterogeneity (Brunner, Dette, & Munk, 1997; Kohr & Games, 1974; Krishnamoorthy, Lu, & Matthew, 2007; Wilcox & Keselma, 2003).

The Alexander-Govern test was introduced by Alexander-Govern (1994) to handle the problem of variance heterogeneity, for a normal data, but this test is not robust to non-normal data. The Alexander-Govern test was

compared with the James test and the Welch test and it was admitted by Schneider and Penfield (1997) and Myers (1988) that the Alexander-Govern test is a better alternative compared to the James test and the Welch test. Myers (1998) admitted that the Alexander-Govern test gives a good solution to the problem of variance heterogeneity. Although, the Alexander-Govern test is a better alternative to the *ANOVA* under variance heterogeneity, the test suffers some disadvantages. As stated by Myers (1998) the major weakness of the test is that it cannot handle any deviation from normality. The test performs excellently well in the control of Type I error for a normal data.

It is an established fact that the common mean is a very good estimator for a normal data, but it is extremely sensitive to the appearance of outliers. The Alexander-Govern test uses mean as its central tendency measure. But the test fails to give a high power and good control of Type I error rates, for a non-normal data. A non-normal data is a situation whereby a given data set is not normally distributed. Investigation under the empirical test shows that the Alexander-Govern test performed remarkably well in the control of Type I error rates and power under variance heterogeneity, compared to the *ANOVA* for a normal data (Alexander & Govern, 1994). In addition, Schneider and Penfield (1997) observed that the Alexander-Govern test is a better alternative to the *ANOVA* under variance heterogeneity compared to the James test and the Welch test. It is due to the fact the Alexander-Govern test is simple in its calculation, and possesses good control of Type I error rates and high power for a normal data distribution. But this test fails to give a good control of Type I error rates under non-normal data.

According to Myers (1998) the Alexander-Govern test is only recommended for a normal data and not for a non-normal data in the control of Type I error rates. In dealing with non-normal data, transformation becomes a better solution. Transformation is a technique that is used in transforming a data distribution that is non-normal, under variance heterogeneity. As a result, the present scores in the data distribution become normal with equal variance. Despite the fact that transformation possesses the ability of transforming a skewed data, but it possesses some disadvantages in its usage. According to Wilcox (2002) in applying transformation on the square root of the mean and also the log of the mean, removes the influences on a real data set. Transformation cannot eliminate the influence of outliers in a data set. In a condition, where the extent of transformation is complex in a given data set, it fails to normalize a skewed data. A better alternative in handling non-normal data is by using non-parametric approach.

Marascuilo and McSweeney (1977) explained that non-parametric test makes no definite assumption in association with one or more of the population parameters that describes the given data distribution that is to be used. It is used to remove a nominal and a ranked order data and is also referred to as a distribution free test. Non-parametric test are not as sensitive as using parametric test, when the assumptions in using the parametric test are accomplished. Therefore, larger differences are needed before a rejection of the null hypothesis is carried out.

Also, non-parametric methods require a large sample size to prevent loss of information. In examining the weakness in using non-parametric technique, scholars have discovered the use of robust estimator as a better alternative in handling non-normal data. Robust estimator that is frequently used in improving the independent group test is the use of trimmed mean. The trimmed mean has been successfully used to improve the Alexander-Govern test, under variance heterogeneity for a non-normal data distribution (Guo & Luh, 2000; Lix & Keselman, 1995; Luh, 1999).

Lix and Keselman (1998) introduced the trimmed mean as a better alternative to the common mean as its central tendency measure under non-normal data. The Trimmed mean is calculated by averaging only the middle data value after removing a certain percentage of the largest and smallest data value, while its variance is evaluated using the Winsorized variance. Trimming is defined as the process of eliminating extreme values in percentage from both tails of the data distribution, in the process of analyzing the data.

Generally, the percentage of trimming is carried out irrespective of the type of the distribution. It will be a very costly mistake, to remove a data distribution where outliers are not found, mainly in a normal distribution, because in doing so, it will lead to loss of information. For a skewed data distribution, the trimming process is not done equally on the tails of the data distribution. Another weakness in using the trimmed mean is that it could not give a good control of Type I error rates, when the number of groups is more than two, especially when it was applied in Alexander-Govern test (Lix & Keselman, 1995).

Another technique in handling the influence of outliers in a data distribution is with the use of Winsorization approach. According to Hasings, Monsteler, Tukey and Winsor (1947) describes the winsorization process as an exchange or replacement for an outlier detected value with the value closest to the outlier. In Winsorization

process, the sample size of the data distribution remains the same. While in the trimming process, the outliers detected are removed from both the upper and lower tail of the data distribution. The trimming process results to loss of information, while the Winsorization process, helps to preserve the data.

According to Abdullah, Yahaya and Othman (2007) a better alternative to the use of trimmed mean in the Alexander-Govern test, is a highly robust estimator known as the modified one step *M*-estimator (*MOM*). It was discovered by these researchers that when the distribution of the data is skewed, the *MOM* estimator, gave a good control of Type I error rates and it empirically trims extreme data values only, depending on the nature of the distribution, be it normal or skewed data distribution. The *MOM* estimator gave an excellent control of Type I error rates, when it was applied in Alexander-Govern test, under normal or highly skewed data distribution, but fails to give a remarkable control of Type I error rates under skewed heavy tailed distribution (Othman et al., 2004).

In this study, to evaluate the capacity of the test using real data, for the *AG* test and the *AGWMOM* test, Descriptive statistics, Test of Normality, Box plots and the Test statistic were employed. The results from the Test statistic, shows that the *AGWMOM* test is very significant compared to the *AG* test.

2. Method

2.1 The Alexander-Govern Test

The Alexander-Govern test is a test introduced by Alexander-Govern (1994) that uses mean as a measure of its central tendency and is also used in comparing two or more groups. This test gives a good control of Type I error rates, and provides high power under variance heterogeneity for a normal data but it is not robust to non-normal data. The test statistic for the Alexander-Govern test is obtained by using the procedures below:

Firstly, the mean of the test is calculated using:

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

where

X_{ij} denotes the observed ordered random samples and n_j represent the sample sizes of the observations. The mean is used as a measure of the central tendency in the Alexander-Govern (1994) method. After the mean is obtained, the estimate of the usual unbiased variance is obtained by using:

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

where

$\bar{X}_j$ *is* used for estimating μ_j for the population *j*. The standard error of the mean is obtained for each of the groups, using:

$$S_{ej} = [\frac{s^2{}_j}{n_j}]^{1/2} \tag{3}$$

The weight (w_j) for the group sizes with *j* population of the ordered sample data is defined, where $\sum w_j$ must be equivalent to 1. So, the weight (w_j) for each of the independent groups is obtained using the formula:

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

Note: The weight (w_j) for the group sizes with *j* population of the ordered sample data is defined such that, $\sum_{j=1}^{J} w_j$ must be equal to 1 (Alexander & Govern, 1994).

The null hypothesis testing for the Alexander-Govern (1994) technique, for the equality of the mean, under variance heterogeneity is defined as:

$$\begin{aligned} &H_o : \mu_1 = \ldots \mu_J \\ &H_A : \mu_1 \neq \ldots \mu_j,\ j = 1, \ldots, J \end{aligned} \tag{5}$$

The alternative hypothesis negates the claim or statement made by the null hypothesis. The variance weighted estimate of the total mean for all the groups in the data distribution, is obtained using:

$$\hat{\mu} = \sum_{j=1}^{J} w_j \bar{X}_j \tag{6}$$

where

w_j is the weight for each of the group in the data distribution and $\bar{X}_j$ is the mean of each of the groups in the ordered sample. The t statistic for each of the groups is obtained using:

$$t_j = \frac{\bar{X}_j - \hat{\mu}}{S_{ej}} \tag{7}$$

where

$\bar{X}_j$ represent the mean for each of the independent group, $\hat{\mu}$ represent the grand mean for all the groups under analysis and S_{ej} denotes the standard error for each of the independent groups with population j.

The t statistic is distributed as a t variable, with $n_j - 1$ degrees of freedom for ν.

where

ν is the degree of freedom for each of the independent groups in the ordered data distribution. The t statistic is obtained for each group and is transformed to a standard normal deviates by using the Hill's (1970) normalization approximation in the Alexander-Govern (1994) technique. The *AG* formula is defined as:

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

where

$$c = [a * \log_e (1 + \frac{t^2_j}{\nu_j})]^{1/2} \tag{9}$$

and where

$$\nu_j = n_j - 1, a = \nu_j - 0.5, b = 48a^2 \tag{10}$$

The test statistic for the Alexander-Govern approach is expressed as:

$$A = \sum_{j=1}^{J} Z^2_j \tag{11}$$

The test statistic value of the Alexander-Govern (*AG*) test at $\alpha = 0.05$ level of significance is denoted by *A*. The p-value of the test is obtained from the standard chi-square distribution table, with $J - 1$ degree of freedom. If the p-value obtained for the *AG* test is > 0.05, then we say the *AG* test is not significant, otherwise, the test is regarded as significant (when the p-value of *AG* test is < 0.05).

3. The Modified Alexander-Govern Test

Given an ordered data sample of $X_1, X_2, \ldots, X_n$, with sample size *n* and group sizes *j*. Firstly, the median of the data set is obtained by selecting the middle value from the observations. The *MAD* estimator is the median of the set of absolute values of the differences between each of the score and the median. It is the median of: $/X_j - M/, \ldots, /X_n - M/$. Thereafter, the median absolute deviation about the median (MAD_n) estimator is obtained by using:

$$MAD_n = \frac{MAD}{0.6745}, \tag{12}$$

According to Wilcox and Keselman (2003) the constant value of 0.6745 is used to rescale the *MAD* estimator, with the aim of making the denominator estimates σ when sampling from a normal distribution.

Outliers in a data distribution can be detected by using the formula below:

$$\frac{/X_j - M/}{MAD_n} \succ K, \tag{13}$$

$$\text{Or when } \frac{/X_j - M/}{MAD_n} \prec -K, \tag{14}$$

where

X_j represents the observed ordered 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 presence of outliers in a data set, because it has a very small standard error, when sampling from a normal distribution.

Equation (13) and (14) is also referred to as the *MOM* estimator that is used for detecting the presence of outliers in a data distribution. The Winsorized *MOM* estimator would be applied in the data distribution when the outlier detected value is replaced or exchanged with a preceding value closest to where the outlier is located.

The *WMOM* estimator, which replaces the mean as a measure of the central tendency is calculated by using the formula below:

$$WMOM = \frac{\sum_{j=1}^{J} X_{WMOMj}}{n}, \tag{15}$$

where

$\bar{X}_{WMOMj}$ is the mean of the Winsorized data distribution, X_{WMOMj} is the ordered sample data of the Winsorized data distribution and n is the sample size of the Winsorized data distribution. The *WMOM* estimator becomes a replacement for the common mean as a measure of the central tendency in the Alexander-Govern test, for the following reasons:

1. To remove the presence of outliers from the data distribution.
2. To make the Alexander-Govern test to be robust to non-normal data.

The Winsorized sample variance is expressed as:

$$S^2{}_{WMOMj} = \frac{\sum_{j=1}^{J} (X_j - \bar{X}_{WMOMj})^2}{n-1}, \tag{16}$$

where

$S^2{}_{WMOMj}$ is the Winsorized sample variance for the Winsorized data distribution, X_j is the observed random sample, $\bar{X}_{WMOMj}$ is the Winsorized *MOM* estimator for the Winsorized data distribution and n is the sample size of the ordered data set.

The standard error of the Winsorized *MOM* is obtained by using the bootstrapping algorithm, for estimating standard errors, which is defined using the following procedure. Firstly, we select B independent bootstrap samples which are expressed as: $x^{*1}, x^{*2}, ..., x^{*B}$, where each of the random samples consists of n data sets chosen by making replacement from x, which is defined as:

$$\begin{aligned} x^* &= (x_1^*, x_2^*, ..., x_n^*) \\ F &\rightarrow (x_1^*, x_2^*, ..., x_n^*) \end{aligned} \tag{17}$$

The symbol ($*$) indicates that x^* is not the actual data set of x, but it is referring to a randomized or resampled version of x. To estimate the standard error of the bootstrap samples, the number of B must be within the interval of (25 – 200). As stated by Efron and Tibshira (1998) 50 amounts of the bootstrap sample is a reasonable amount to give a sufficient estimate of the standard error of the *MOM* estimator. In this study, 50 amounts of the bootstrap sample were used to give a reasonable estimate of the standard error of the *MOM* estimator.

Secondly, we estimated the bootstrap replications equating to each of the bootstrap samples expressed as:

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

Thirdly, we estimate the standard error of $se_F(\hat{\theta})$ from the sample standard deviation of the bootstrap (B) replications as defined as:

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

where

$$\hat{\theta}^*(\cdot) = \sum_{b=1}^{B} \hat{\theta}^*(b) / B. \tag{20}$$

The weight w_j for the Winsorized data distribution for each of the independent groups is obtained using:

$$w_j = \frac{1/S_{e\,WMOMj}^{2}}{\sum_{j=1}^{J} 1/S_{e\,WMOMj}^{2}}, \tag{21}$$

Note: The weight (w_j) for the group sizes with j population is defined such that $\sum_{j=1}^{J} w_j$ must be equivalent to 1 (Alexander & Govern, 1994).

$\sum_{j=1}^{J} 1/S^2_{eWMOMj}$, is the sum of the reciprocal of the square of the standard error for all the independent groups in the ordered data set from the real life data distribution.

The variance weighted estimate of the total mean for the Winsorized data distribution for all the independent groups is defined as:

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

The t statistic for the Winsorized data distribution for each of the independent groups is expressed as:

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

where

S_{eWMOMj} is the Winsorized sample standard error from the Winsorized data distribution, for each of the independent groups of $\bar{X}_{WMOMj}$.

In the Alexander-Govern (1994) approach, the t_j value is transformed to standard normal by using the Hill's (1970) normalization approximation and the hypothesis testing of the Winsorized data distribution, where S^2_{WMOMj} is the usual estimate of the Winsorized sample variance of the *WMOM* estimator for the μ_j is defined as:

$$\begin{aligned} &H_o: \mu_1 = \ldots = \mu_J \\ &H_A: \mu_1 \neq \ldots \neq \mu_J, \; for\; j = (j = 1, \ldots, J). \end{aligned} \tag{24}$$

Therefore, the normalization approximation formula for the Alexander-Govern approach, using the Winsorized data distribution is defined as:

$$Z_j = c + \frac{[c^3 + 3c]}{b} - \frac{[4c^7 + 33c^5 + 240c^3 + 855c]}{[10b^2 + 8bc^4 + 1000b]}, \tag{25}$$

where

$$c = [a * \log_e (1 + \frac{t^2_j}{\nu_j})]^{1/2}, \tag{26}$$

and where

$$\nu_j = n_j - 1, a = \nu_j - 0.5, b = 48a^2, \tag{27}$$

The test statistic for the *AGWMOM* for all the independent groups in the ordered sample data is defined as:

$$AGWMOM = \sum_{j=1}^{J} Z^2_j \ , \tag{28}$$

The test statistic of the *AGWMOM* test follows a chi-square distribution at $\alpha = 0.05$ level of significance, with *J* – 1 chi-square degrees of freedom. The p-value of the *AGWMOM* test is obtained from the standard chi-square distribution table. If the p-value is < 0.05, it shows that the *AGWMOM* test is significant, otherwise it is not.

4. To Evaluate the Capacity of the Test Using Real Data

Table 1. Given an Ordered Data set with Observations as follows

Young (Y)	Middle (M)	Old (O)
482.43	335.59	519.01
484.36	338.43	528.50
488.84	353.54	530.23
495.15	404.27	536.03
495.24	437.50	538.56
502.69	469.01	538.83
504.62	485.85	557.24
518.29	487.30	558.61
519.10	493.08	558.95
524.10	494.31	565.43
524.12	499.10	586.39
531.18	886.41	594.69
548.42	-	629.22
572.10	-	645.69
584.68	-	691.84
609.09	-	-
609.53	-	-
666.63	-	-
676.40	-	-

Source: Keselman et al. (2007).

Table 2. The Winsorized data distribution from the real life data

Winsorized Young	Winsorized Middle	Winsorized Old
482.43	404.27	519.01
484.36	404.27	528.50
488.84	404.27	530.23
495.15	404.27	536.03
495.24	437.50	538.56
502.69	469.01	538.83
504.62	485.85	557.24
518.29	487.30	558.61
519.10	493.08	558.95
524.10	494.31	565.43
524.12	499.10	586.39
531.18	499.10	594.69
548.42	-	629.22
572.10	-	645.69
584.68	-	645.69
609.09	-	-
609.53	-	-
609.53	-	-
609.53	-	-

Table 3. Descriptive Statistics for the bootstrap sample (n = 50), for the *AG* test

Descriptive Statistics	**Young (Y)**	**Middle (M)**	**Old (O)**
Usual mean	544.0511	299.1784	451.5379
Standard Error of Mean	13.7022	59.7782	55.8659
Skewness	1.100	0.210	-0.409
Kurtosis	0.248	-0.488	0.258

In Table 3, the values of the skewness for the group young, middle and old are all greater than zero. Hence, the distribution of the data is said to be right skewed. The values of the kurtosis for the group young and old is less than three and it indicates that the values of the data are majorly distributed around the mean in the distribution. While the kurtosis value for the group middle is greater than three and hence, it shows that the distribution of the data is centred about the mean, with a thicker tail. As a result, there is high probability for extreme values.

Table 4. Descriptive Statistics for the bootstrap sample (n = 50), for the *AGWMOM* test

Descriptive Statistics	Winsorized Young	Winsorized Middle	Winsorized Old
Usual Mean	505.8433	456.8608	551.0392
Standard Error of Mean	4.9059	12.1963	6.7518
Skewness	0.058	-0.411	0.606
Kurtosis	-1.554	-1.937	-0.432

Table 5. Tests of Normality for the real life data

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
Young	.185	18	.200*	.924	18	.319
Middle	.347	11	.000	.721	11	.001
Old	.199	14	.200*	.935	14	.431

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction.

5. Discussion and Conclusion

Shapiro-Wilk Test is a test that is most suitable for sample sizes that is less than 50. This test is also suitable to handle sample sizes that is as large as 2000. As a result of this, the Shapiro-Wilk Test is used to test for the normality of the three independent groups, namely, the group young, middle and old respectively. If the significant values of any of the three tests is greater than 0.05, then the data is said to be normally distributed. Otherwise, if the significant value is less than 0.05, then the data distribution is said to be non-normal. The results from Table 5, show that the p-value for the group young and old is greater than 0.05, hence both groups are normally distributed. The group middle has a p-value of 0.001 which is less than 0.05 and is considered to be non-normally distributed.

In Figure 1, shows the boxplots of the reaction time against the group young, middle and old respectively. It can be observed very clearly from the plots that there is no extreme value seen in the group young and old and hence, the data distribution for both groups are said to be normally distributed. There is an extreme value observed in the group middle, and this indicates that the data distribution for the group middle is non-normal.

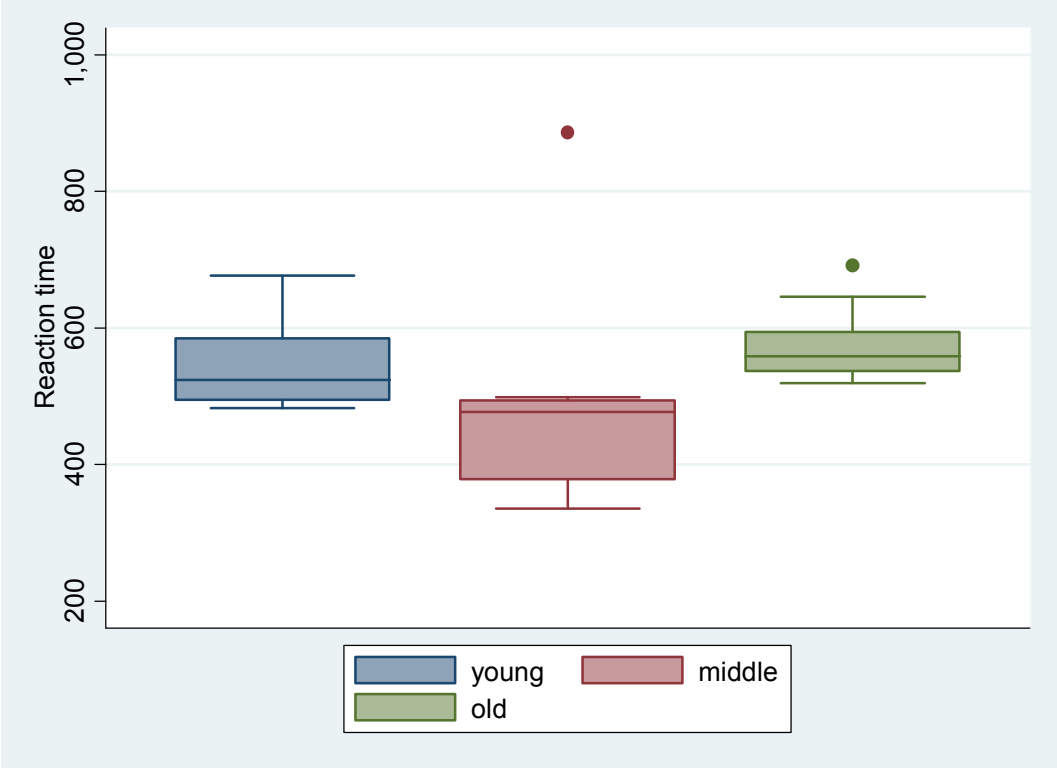

Figure 1. Boxplots on reaction time versus group young, middle and old

Table 6. The test statistic for the *AG* test and *AGWMOM* test

Test	Test Statistic	p-Value
Original *AG*	5.3237	0.06982
AGWMOM	30.1280	0.0000002869

$\alpha = 0.05$

In Table 6, the test statistic for the *AG* test has a value of 5.3237, with a p-value of 0.06982 at $\alpha = 0.05$ level of significant. This implies that the p-value of the *AG* test is regarded as not significant, since the p-value of the *AG* test is > 0.05. While the test statistic value of the *AGWMOM* test is six times more than that of the original *AG* test. The test statistic value of the *AGWMOM* test produced a value of 30.1280, with a p-value of 0.0000002869 at 0.05 level of significant. As a result, the p-value of the *AGWMOM* test is regarded as significant, since its value is < 0.05, compare to the *AG* test. The standard error of the Winsorized *AGMOM* from the real life data for the group young, middle and old is far smaller compared to the standard error of the *AG* test from the original real life data.

Therefore, the *AGWMOM* test has assisted to minimize error as much as possible from the real life data compared to the *AG* test.

Acknowledgements

I give God Almighty all the thanks, praises, adoration, honour, power, majesty, worship and glory for everything. "He is the author of wisdom, knowledge and understanding", The Everlasting Father, the beginning and ending of everything.

I also want to acknowledge and thank my blessed, beloved, special, wonderful, and very caring parents, in person of Mr. and Mrs. D. K.O. Tobi, for their constant encouragement, love, support, backings, sacrifice and goodwill. I love and appreciate them very greatly.

References

Abdullah, S., Yahaya, S. S. S., & Othman, A. R. (2007). Proceedings of The 9th Islamic Countries Conference on

Statistical Sciences 2007. In Modified One Step M-Estimator as a Central Tendency Measure for Alexander-Govern Test (pp. 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.

Brown, M. B., & Forsythe, A. B. (1974). The small sample behavior of some statistics which test the equality of several means. *Technomtrics*, *16*, 129-132.

Brunner, E., Dette, H., & Munk, A. (1997). Box-Type Approximations in Nonparametric Factorial Designs. *Journal of the American Statistical Association, 92*(440), 1494-1502.

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

Guo, J., & Luh, W. (2000). An invertible transformation two-sample trimmed t-statistic under heterogeneity and nonnormality, 49, 1-7.

Hasings, C., Monsteller, F., Tukey, J. W., & Winsor, C. P. (1947). Low moments for small sample: a comparative study of order statistics. *Annals of Mathematical Statistics, 18,* 413-426.

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., Wilcox, R., Lix, L. M., Algina, J., & Fradette, K. (2007). Adaptive robust estimation and testing. *British Journal of Mathematical Psychology, 60*, 267-293.

Kohr, R. L., & Games, P. A. (1974). Robustness of the analysis of variance, the Welch procedure, and a Box procedure to heterogeneous variances. *Journal of Experimental Education, 43,* 61- 69.

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. http://dx.doi.org/10.1016/j.csda.2006.09.039

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

Lix, L. M, Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational Research, 66*, 579-619.

Lix, Lisa, M., & Keslman, H. J. (1995). Approximate degrees of freedom tests: A unified perspective on testing for mean equality. *Psychological Bulletin, 117*(3), 547-560. http://dx.doi.org/10.1037//0033-2909.117.3.547

Luh, W. M. (1999). Developing trimmed mean test statistics for two-way fixed-effects ANOVA models under variance heterogeneity and nonnormality. *Journal of Experimental Education, 67*(3), 243-265.

Marascuilo, L. A., & McSweeney, M. (1977). *Nonparametric and distribution-free methods for the social sciences.* Monterey, CA: Brooks/Cole Publishing Company.

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 Computational, 60*, 207-222.

Oshima, T. C., & Algina, J. (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., Padmanabhan, A. R., Wilcox, R. R.., & Fradette, K. (2004). Comparing measures of the "typical" score across treatment groups. *The British Journal of Mathematical and Statistical Psychology, 57*(Pt 2), 215-234. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15511305

Pardo, J. A, Pardo, M. C., Vicente, 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. doi: 10.1016/S0167-9473(96)00080-1Scheffe, H. (1959). *The analysis of variance.* New York: Wiley.

Schneider, P. J., & Penfield, D. A. (1997). Alexander and 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. *Biometrika,* 38, 330-336.sults on the robustness of the ANOVA F, W, and F statistics. *Communications in Statistics-Simulation, 15,* 933-943.

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., Charlin, V. L., & Thompson, K. L. (1986). New Monte Carlor.

Wilcox, R. R., & Keselman, H. J. (2002). Power analysis when comparing trimmed means. *J. Modern Appl. Stat. Methods, 1*(1), 24-31.

Wilcox, R. R., & Keselman, H. J. (2003). Modern Robust Data Analysis Methods: Measures of Central Tendency. *Psychological Methods, 8*(3), 254-274.