ONE-WAY RANDOM EFFECTS MODEL FOR ASYMMETRIC TRIMMED MEANS

CHIN EE LAINE

UNIVERSITI SAINS MALAYSIA

2019

ONE-WAY RANDOM EFFECTS MODEL FOR ASYMMETRIC TRIMMED MEANS

by

CHIN EE LAINE

**Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science**

October 2019

ACKNOWLEDGEMENT

First and foremost, all praise and glory to my Lord and Saviour, Jesus Christ for his faithfulness, wisdom and strength in helping me through my years of postgraduate studies. Thank You, God for your grace and mercy in enabling me to successfully complete this Masters program in good health and mind.

I offer my sincerest gratitude to my supervisor, Professor Dr. Abdul Rahman bin Othman who has supported me throughout my thesis with his patience and knowledge whilst allowing me room to work in my own ways. I attribute the level of my Masters degree to his encouragement and effort for without him, this thesis would not have been completed. One simply could not wish for a better or friendlier supervisor. Thank you to my co-supervisors, Professor Dr. Sonia Aissa and Dr. Teh Sin Yin for their support and encouragement. I would also like to express my gratitude to the staff of the School of Distance Education and Institute of Postgraduate Studies in Universiti Sains Malaysia.

To my family; I thank my parents for supporting me throughout the duration of my studies; I am forever grateful for their unceasing prayers and love, providing a home that enables the completion of writing up of this thesis. To my brothers, sister-in-law and nephews, thank you for being there when I needed a listening ear and someone to talk and to laugh with. I would also like to acknowledge the many friends for their prayers and constant support. Finally, to the examiners of this thesis, thank you for taking the time and effort. May God bless each and everyone.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

MODEL KESAN RAWAK SEHALA UNTUK MIN TERPANGKAS TAK SIMETRI

ABSTRAK

Terdapat dua perkara yang perlu diberi perhatian untuk model kesan rawak. Yang pertama, anggapan kesamaan varians antara kumpulan dan yang kedua, anggapan kenormalan. Pelanggaran kedua-dua faktor ini akan menghasilkan Ralat Jenis I yang tidak memuaskan dan kehilangan kuasa yang agak ketara. Prosedur Jeyaratnam-Othman (J&O) (1985) berjaya menangani masalah pertama tersebut yang melibatkan ketaksamaan varians dengan kenormalan data. Wilcox (1994) meneruskan kajian dengan cadangan mengitlakkan prosedur J&O berdasarkan min terpangkas simetri. Prosedur ini berjaya menghasilkan nilai kuasa yang tinggi tetapi kurang memuaskan apabila taburan menjadi pencong dengan saiz sampel kumpulan tidak setara. Kajian ini menggantikan min terpangkas simetri di dalam prosedur Wilcox dengan min terpangkas tak simetri. Ralat Jenis I yang memuaskan ataupun lebih memuaskan daripada prosedur Wilcox menjadi sasaran. Dua penganggar engsel Q_1 dan Q_2 Reed dan Stark (1996) dipilih untuk mendapatkan min terpangkas tak simetri kajian ini. Simulasi dijalankan untuk prosedur J&O, procedure Wilcox dan kedua-dua prosedur kajian ini ke atas rekabentuk empat kumpulan dengan pelbagai taburan data. Rekabentuk seimbang kedua-dua prosedur yang dicadangkan menghasilkan Ralat Jenis I yang baik bernilai dari 0.026 sehingga 0.082. Begitu juga dengan kuasa dengan purata 0.782. Namun demikian, kuasa dan Ralat Jenis I yang tidak memuaskan diperolehi pada rekabentuk tak seimbang.

ONE-WAY RANDOM EFFECTS MODEL FOR ASYMMETRIC TRIMMED MEANS

ABSTRACT

There are two very important concerns for the random effects model. The first concern being the assumption of equal variances of groups and the second concern is assuming normality. Violations of these result in unsatisfactory Type I errors and considerable loss of power. Jeyaratnam-Othman (1985) addresses the first concern in dealing with unequal variances while assuming normality. Wilcox in 1994 continued the study by suggesting a generalization on Jeyaratnam-Othman's procedure based on symmetric trimmed means. The procedure resulted in significant gain in power but was unsatisfactory for skewed distributions with unequal group sizes. This research replaces Wilcox's symmetric trimmed means with asymmetric ones aiming to obtain good, if not, better Type I errors. Two hinge estimators by Reed and Stark (1996), Q_1 and Q_2, were employed to obtain the asymmetric trimmed means for this research. Simulations were carried out for Jeyaratnam-Othman (1985), Wilcox (1994) and the two proposed procedures for a four-group design subjected to different data distributions. Good control of Type I errors was evident for both proposed procedures for balanced designs with values ranging from 0.026 to 0.082. Good power averaging 0.782 was also obtained. However, power and Type I errors for the unbalanced design were very unsatisfactory.

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND

The random effects ANOVA model has been used extensively in fields such as psychological studies (Bowen & Huang, 1990; Mirman, Dixon & Magnuson, 2008), quantitative genetics (Snedecor & Cochran, 2014), health (Madden, Browne, Li, Kearney & Fitzgerald, 2018; Arku et al., 2018) and astronomy (Scheffe, 1999). The traditional random effects model for *J* randomly sampled groups is that of

$$Y_{ij} = \mu + a_j + e_{ij}$$

where Y_{ij} is the *i*-th observation from the *j*-th group $(i = 1, \ldots, n_j; j = 1, \ldots, J)$, n_j is the size of the group, μ is an unknown common constant; and a_j and e_{ij} are normally distributed random independent variables, i.e. $a_j \sim N\left(0, \sigma_a^2\right)$ and $e_{ij} \sim N\left(0, \sigma^2\right)$. The variance of a_j is σ_a^2 and since it is independent of e_{ij}, the variance of any observation is $Var\left(Y_{ij}\right) = \sigma_a^2 + \sigma^2$. For testing of hypotheses in this particular model, the testing of individual treatment effects is meaningless (Montgomery, 2017) but instead the test is about the variance component, σ_a^2.

The procedures of the ANOVA for the random effects model are identical to the way computations are done for the fixed effects cases. With that in mind, two fundamental concerns have been discussed in (Wilcox, 1994a). The first concern is the assumption of equal variances. Violation of this assumption will result in the traditional test of $H_0 : \sigma_a^2 = 0$ to be unsatisfactory in terms of Type I errors, even with data that are normally distributed. The second is departure from normality, regardless of a slight deviation or a major deviation from normality such as a heavy-tailed distribution, will lead to an increase in the standard error of the sample mean by a substantial amount thus resulting in a serious effect on power (Tukey, 1960).

Jeyaratnam and Othman (1985) derived a test of the null hypothesis as $H_0 : \sigma_a^2 = 0$ that shows unequal error variances whilst assuming normality. The assumption that the variance of $e_{ij} = \sigma_e^2$ (constant) was replaced with the assumption that the variance of $e_{ij} = \sigma_j^2$ now which can vary among the *J* groups.

Wilcox (1994a) continues to study a generalization of the Jeyaratnam and Othman method on the usual random effects model based on symmetric trimmed means. This symmetric trimmed mean procedure has error that is less affected by heavy-tailed distributions and outliers and can yield considerable gains in power. Work have been done in the fixed effects model using trimmed means and according to Wilcox, the use of the usual mean can possibly portray a distorted view of “how the typical individual in one group compares to the typical individual in another, and about accurate probability coverage, controlling the probability of a Type I error, and achieving relatively high power” (Wilcox, 1995, p.66).

Hence, as Keselman, Kowalchuk and Lix (1998b) notes, the trimmed mean is preferable as a robust estimator of location because of its computational simplicity and good theoretical properties (Wilcox, 1995), particularly when the standard error of the trimmed mean is less affected by departures from normality.

Although several results have proven that the trimmed mean is a well-accepted measure under the effects of nonnormality and variance heterogeneity, there is always the lingering question of the accuracy of trimming when distributions are skewed: should the data be trimmed symmetrically or asymmetrically and how much trimming should be done. Keselman, Wilcox, Othman and Fradette (2002) demonstrated the advantage of a prior test for symmetry in order to determine whether data should be trimmed from both tails (symmetric trimming) or just from one tail (asymmetric trimming).

Following the course of trimming on fixed effects models, Keselman, Wilcox, Lix, Algina and Fradette (2007) showed several methods that determine whether the data distribution should be trimmed and the quantity of trimming from the tails of the distribution. The two main adaptive trimming methods discussed by Keselman et al. (2007) are the Reed and Stark (1996) and the Tukey-McLaughlin-Jaeckel-Hogg methods. The Reed and Stark (1996) method is based on the work of Hogg (1974, 1982) whereby several adaptive location estimators were defined depending on measures of tail length and skewness and the Tukey-McLaughlin-Jaeckel-Hogg method suggests that the amount to be trimmed in each tail can be selected by adopting a trimming strategy that results in the smallest standard deviation of the sample trimmed mean.

Alkhazaleh and Razali (2010) adopted Hogg's tail weight measures in proposing a technique to estimate asymmetric trimmed means. In addition, Md. Yusof, Othman and Syed Yahaya (2010) compared Type I error rates between two statistics for unequal population variance by using variable trimming. These two statistics were modified using variable trimming with indeterminate percentages.

As extensive work has been done on fixed effects model, there has not been many follow ups on the random effects. This research proposes the use of asymmetric trimmed means in view of Wilcox' (1994a) measures on random effects model and how the asymmetric trimmed means compares with the symmetric trimmed means in terms of Type I error and power. In obtaining the asymmetric trimmed means, the hinge estimators were studied whereby this method of adaptive trimmed mean trims data using asymmetric trimming technique, where the tails of the distribution are trimmed based on the characteristics of a particular distribution. There are seven adaptive location estimators namely Q, Q_1, H_3, Q_2, H_1, SK_2 and SK_5 and the two that were chosen for this research are Q_1 and Q_2.

1.2 RESEARCH PROBLEM

Existing one-way random effects model are not robust to nonnormal data, existence of group variance heterogeneity and the combination of the unbalanced group sizes. Therefore, the one-way random effects model for asymmetric trimmed means was proposed to address the issue.

1.3 RESEARCH OBJECTIVES

The objectives for this research are as follows:

(i) To develop an asymmetric trimmed mean test statistic for REM

(ii) To compare the performance of (i) with the one-way random effects model with symmetric trimmed means, F_t.

1.4 SIGNIFICANCE OF RESEARCH

The significance of this research is to add to the knowledge of the behaviour of the one-way random effects model under nonnormality and to verify that the proposed measure of trimming is better performing than the existing symmetric measures.

1.5 THESIS ORGANIZATION

This first chapter briefly introduces the topic of interest for this research which includes some background literature and also the objectives, aims and hypotheses for this work of the one-way random effects model with asymmetric trimmed means. In Chapter Two, the review of literature will be expanded in detail and it will show how this research topic is built upon existing works. Chapter Three will show the methodology - the test procedures and study conditions - carried out for this research. Following that, Chapter Four is where the results and discussion are presented. Chapter Five concludes this research and suggests of future research work that can be done. The bibliography and appendices are also included after the five chapters.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

In the design of an experiment, systematic procedures are carried out to form or to test a hypothesis that will lead to the discovery of an unknown effect or to confirm known effects. In an experimental situation, the experimenter is interested in factors with large number of levels. When the experimenter randomly selects a number of these levels from the population of factor levels, the factor is said to be random. Because the levels of the factor used in the experiment were actually chosen at random, conclusions are made about the entire population of factor levels. These form the random effects model (REM) or also called the components of variance model. As stated in Chapter One, the random effects ANOVA model has been used extensively in fields such as psychological studies, quantitative genetics, health and astronomy.

2.2 RANDOM EFFECTS MODEL ANOVA TESTS

The traditional random effects ANOVA model for a factor with *J* randomly sampled levels is that of

$$Y_{ij} = \mu + a_j + e_{ij} \tag{2.1}$$

where Y_{ij} is the *i*-th observation from the *j*-th factor level $(i = 1, \ldots, n_j \text{ , } j = 1, \ldots, J)$, n_j is the size of the level, μ is the grand mean; and both a_j and e_{ij} are normally distributed random independent variables: $a_j \sim N\left(0, \sigma_a^2\right)$ and $e_{ij} \sim N\left(0, \sigma^2\right)$ respectively. The variance of a_j is σ_a^2 and since it is independent of e_{ij}, the variance of any observation is $Var\left(Y_{ij}\right) = \sigma_a^2 + \sigma^2$. Testing of hypotheses in this particular model is about the variance component, σ_a^2. (Montgomery, 2017, p. 506)

2.2.1 The Balanced One-Way REM with Equal Error Variances

For the balanced design model from Equation (2.1), let $n_j = n$ and $\sigma^2 = \sigma_e^2$ for all levels. Hence, let

$$\overline{Y_j} = \frac{1}{n}\sum_{i=1}^{n} Y_{ij} \; . \tag{2.2}$$

Therefore, the unbiased point estimators of the parameters are then given by

$$\mu = \overline{Y} = \frac{1}{J}\sum_{j=1}^{J} \overline{Y_j} \; , \tag{2.3}$$

$$\sigma_e^2 = \sum_{j=1}^{J}\sum_{i=1}^{n} \frac{\left(Y_{ij} - \overline{Y_j}\right)^2}{J\left(n-1\right)} \text{, and} \tag{2.4}$$

$$\sigma_a^2 = \sum_{j=1}^{J} \frac{\left(\overline{Y_j} - \overline{Y}\right)^2}{J-1} - \sum_{j=1}^{J}\sum_{i=1}^{n} \frac{\left(Y_{ij} - \overline{Y_j}\right)^2}{Jn\left(n-1\right)} \; . \tag{2.5}$$

Now, let

$$S_1 = \sum_{j=1}^{J}\sum_{i=1}^{n} \left(Y_{ij} - \overline{Y_j}\right)^2 \text{ and} \tag{2.6}$$

$$S_2 = \sum_{j=1}^{J} \left(\overline{Y_j} - \overline{Y}\right)^2 \; . \tag{2.7}$$

Therefore, the ratios of $\frac{S_1}{\sigma_e^2}$ and $\frac{S_2}{\sigma_e^2 + n\sigma_a^2}$ are independent chi-square variables with degrees of freedom, $J(n-1)$ and $J-1$, respectively.

Define

$$\overline{W} = \frac{\frac{S_2}{J-1}}{\frac{S_1}{J(n-1)}} \tag{2.8}$$

and multiplying both sides by $\frac{\sigma_e^2}{\sigma_e^2 + n\sigma_a^2}$, therefore

$$\frac{\sigma_e^2}{\sigma_e^2 + n\sigma_a^2}\overline{W} = \frac{\frac{S_2}{J-1}}{\frac{S_1}{J(n-1)}} \times \frac{\sigma_e^2}{\sigma_e^2 + n\sigma_a^2}$$

$$= \frac{\frac{S_2}{(J-1)\left(\sigma_e^2 + n\sigma_a^2\right)}}{\frac{S_1}{J(n-1)\sigma_e^2}}. \tag{2.9}$$

Hence,

$$\overline{W} = W\left(\frac{\sigma_e^2 + n\sigma_a^2}{\sigma_e^2}\right)$$

$$= W\left(1 + \frac{n\sigma_a^2}{\sigma_e^2}\right) \tag{2.10}$$

where

$$W = \frac{\left(\frac{S_2}{\sigma_e^2 + n\sigma_a^2}\right) \Big/ (J-1)}{\left(\frac{S_1}{\sigma_e^2}\right) \Big/ J(n-1)} \tag{2.11}$$

is distributed as $F_{(J-1),(J(n-1))}$. Under the null hypothesis of $H_0: \sigma_a^2 = 0$ the value $\overline{W}$ is equal to W. In this case, at a significant level of α, H_0 is rejected when the evaluated value $\overline{\omega}$ of $\overline{W}$ is greater than $F_{\alpha:(J-1),(J(n-1))}$.

2.2.2 The Unbalanced One-Way REM with Equal Error Variances

Reconsider Equation (2.1), with $\sigma^2 = \sigma_e^2$ for all treatment levels, let the sample sizes of the levels be different, n_j with $j = 1,...,J$ then

$$\overline{Y_j} = \frac{1}{n_j}\sum_{i=1}^{n_j} Y_{ij} \ . \tag{2.12}$$

Therefore, the unbiased point estimators of the parameters are then given by

$$\mu = \overline{Y} = \frac{\sum_{j=1}^{J} n_j \overline{Y_j}}{\sum_{j=1}^{J} n_j} \ , \tag{2.13}$$

$$\sigma_e^2 = \sum_{j=1}^{J}\sum_{i=1}^{n_j} \frac{\left(Y_{ij} - \overline{Y_j}\right)^2}{\sum_{j=1}^{J}\left(n_j - 1\right)} \text{, and} \tag{2.14}$$

$$\sigma_a^2 = \frac{1}{J_0}\left[\sum_{j=1}^{J} \frac{n_j\left(\overline{Y_j} - \overline{Y}\right)^2}{J-1} - \sum_{j=1}^{J}\sum_{i=1}^{n_j} \frac{\left(Y_{ij} - \overline{Y_j}\right)^2}{\sum_{j=1}^{J}\left(n_j - 1\right)}\right] \tag{2.15}$$

where

$$J_0 = \frac{\left(N^2 - \sum_{j=1}^{J} n_j^2\right)}{N\left(J-1\right)} \text{ and} \tag{2.16}$$

$$N = \sum_{j=1}^{J} n_j \ . \tag{2.17}$$

Now, let

$$S_i = \sum_{i=1}^{n_j} \left(Y_{ij} - \overline{Y_j} \right)^2 \text{ and} \tag{2.18}$$

$$S_o = \sum_{j=1}^{J} n_j \left(\overline{Y_j} - \overline{Y} \right)^2 . \tag{2.19}$$

Just as the case of the balanced model, define

$$\overline{W^*} = \frac{\dfrac{S_o}{J-1}}{\displaystyle\sum_{j=1}^{J} \frac{S_i}{N-J}} \tag{2.20}$$

and multiplying both sides by $\dfrac{\sigma_e^2}{\sigma_e^2 + J_0 \sigma_a^2}$, therefore

$$\frac{\sigma_e^2}{\sigma_e^2 + J_0 \sigma_a^2} \overline{W^*} = \frac{\dfrac{S_o}{J-1}}{\displaystyle\sum_{j=1}^{J} \frac{S_i}{N-J}} \times \frac{\sigma_e^2}{\sigma_e^2 + J_0 \sigma_a^2}$$

$$= \frac{\dfrac{S_o}{(J-1)\left(\sigma_e^2 + J_0 \sigma_a^2\right)}}{\displaystyle\sum_{j=1}^{J} \frac{S_i}{(N-J)\sigma_e^2}} . \tag{2.21}$$

Hence,

$$\overline{W^*} = \frac{\left(\dfrac{S_o}{\sigma_e^2 + J_0 \sigma_a^2} \right) \Big/ (J-1)}{\left(\displaystyle\sum_{j=1}^{J} \frac{S_i}{\sigma_e^2} \right) \Big/ (N-J)} . \tag{2.22}$$

Therefore, the ratios of $\frac{S_i}{\sigma_e^2}$ and $\frac{S_o}{\sigma_e^2 + J_0\sigma_a^2}$ are independent chi-square variables with degrees of freedom, $N-J$ and $J-1$, respectively. Under the null hypothesis of $H_0 : \sigma_a^2 = 0$, the value $\overline{W^*}$ has a distribution of $F_{(J-1),(N-J)}$. With that said, at a significant level of α, H_0 is rejected when the evaluated value $\overline{\omega^*}$ of $\overline{W^*}$ is greater than $F_{\alpha:(J-1),(N-J)}$.

While having important roles in various fields of study, the REM suffers from two very key concerns which are primarily addressed by Wilcox (1994a). The first concern is the assumption of equal variances of the groups. This assumption must be made as Wilcox (1994a) mentions that violation of this assumption will result with the traditional test of $H_0 : \sigma_a^2 = 0$ to be inadequate in terms of Type I error even under normality. The second concern is assuming normality. This assumption must be made because, regardless of a slight deviation or a heavy-tailed distribution, a slight departure from normality will lead to an increase in the standard error of the sample mean by a substantial amount thus resulting in a serious effect on power (Tukey, 1960).

2.3 THE JEYARATNAM-OTHMAN PROCEDURE

In addressing the first concern of the REM, Jeyaratnam and Othman (1985) did a study on the random effects ANOVA model by proposing an approximate procedure for testing $H_0 : \sigma_a^2 = 0$ versus $H_1 : \sigma_a^2 > 0$ with unequal variances of the groups under normality. That is, the assumption that the variance of e_{ij} is a constant

σ^2 is being replaced with the assumption that the variance is σ_j^2 which may vary among the J sampled treatment levels. The following model is considered:

$$Y_{ij} = \mu + a_j + \varepsilon_{ij} \tag{2.23}$$

for $i = 1,\ldots,n_j$ and $j = 1,\ldots,J$ where a_j and ε_{ij} are independently distributed with means zero and variances σ_a^2 and σ_j^2 respectively. The model in Equation (2.23) is referred to as the one-way REM with unequal variances by Jeyaratnam and Othman (1985). Jeyaratnam and Othman (1985) proposed a test statistic for testing $H_0 : \sigma_a^2 = 0$ without assuming equal variances as follows:

$$F = \frac{\dfrac{\sum_{j=1}^{J}\left(\overline{Y_j} - \overline{Y}^*\right)^2}{J-1}}{\sum_{j=1}^{J}\sum_{i=1}^{n_j}\dfrac{\left(Y_{ij} - \overline{Y_j}\right)^2}{Jn_j\left(n_j - 1\right)}} \tag{2.24}$$

where

$$\overline{Y_j} = \sum_{i=1}^{n_j}\frac{Y_{ij}}{n_j}, \tag{2.25}$$

$$\mu = \overline{Y}^* = \sum_{j=1}^{J}\frac{\overline{Y_j}}{J}, \tag{2.26}$$

$$\sigma_j^2 = \sum_{i=1}^{n_j}\frac{\left(Y_{ij} - \overline{Y_j}\right)^2}{n_j - 1} \text{ and} \tag{2.27}$$

$$\sigma_a^2 = \sum_{j=1}^{J}\frac{\left(\overline{Y_j} - \overline{Y}^*\right)^2}{J-1} - \frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}. \tag{2.28}$$

Now, let

$$Z_1 = \sum_{j=1}^{J} \left(\overline{Y_j} - \overline{Y}^* \right)^2 \tag{2.29}$$

$$Z_2 = \sum_{i=1}^{n_j} \left(\overline{Y_{ij}} - \overline{Y_j} \right)^2 \tag{2.30}$$

The proposed test statistic for testing $H_0 : \sigma_a^2 = 0$ as given by Equation (2.24) is as follows:

$$F = \frac{\dfrac{1}{J-1} \sum_{j=1}^{J} \left(\overline{Y_j} - \overline{Y}^* \right)^2}{\dfrac{1}{J} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \dfrac{\left(\overline{Y_{ij}} - \overline{Y_j} \right)^2}{n_j \left(n_j - 1 \right)}} \tag{2.31}$$

or in terms of Z_1 and Z_2,

$$F = \frac{\dfrac{1}{J-1} Z_1}{\dfrac{1}{J} \sum_{j=1}^{J} \dfrac{Z_2}{n_j \left(n_j - 1 \right)}} . \tag{2.32}$$

Following this, since the numerator F is a function of Z_1 and the denominator is a function of Z_2 accordingly, hence the numerator and denominator of F are independent.

Using the information from Equations (2.27) and (2.28), the following are obtained.

$$E\left(\frac{1}{J-1} Z_1 \right) = E\left[\frac{1}{J-1} \sum_{j=1}^{J} \left(\overline{Y_j} - \overline{Y}^* \right)^2 \right]$$

$$= \sigma_a^2 + \frac{1}{J} \sum_{j=1}^{J} \frac{\sigma_j^2}{n_j} \tag{2.33}$$

$$E\left(\frac{1}{J}\sum_{j=1}^{J}\frac{Z_2}{n_j\left(n_j-1\right)}\right)=E\left[\frac{1}{J}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\frac{\left(\overline{Y_{ij}}-\overline{Y_j}\right)^2}{n_j\left(n_j-1\right)}\right]$$

$$=\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j} \tag{2.34}$$

Now multiplying both sides of F by $\dfrac{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}{\sigma_a^2+\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}$, therefore

$$\frac{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}{\sigma_a^2+\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}F=\frac{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}{\sigma_a^2+\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}\times\frac{\frac{1}{J-1}\sum_{j=1}^{J}\left(\overline{Y_j}-\overline{Y}^*\right)^2}{\frac{1}{J}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\frac{\left(\overline{Y_{ij}}-\overline{Y_j}\right)^2}{n_j\left(n_j-1\right)}}$$

$$=\frac{\left(\frac{1}{J-1}\right)\left[\frac{\sum_{j=1}^{J}\left(\overline{Y_j}-\overline{Y}^*\right)^2}{\sigma_a^2+\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}\right]}{\frac{1}{J}\left[\frac{\sum_{j=1}^{J}\sum_{i=1}^{n_j}\frac{\left(\overline{Y_{ij}}-\overline{Y_j}\right)^2}{n_j\left(n_j-1\right)}}{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}\right]}. \tag{2.35}$$

From Equation (2.35), let $\dfrac{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}{\sigma_a^2+\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}F=\overline{F}$.

So
$$\overline{F}=\frac{\left(\frac{1}{J-1}\right)\left[\frac{\sum_{j=1}^{J}\left(\overline{Y_j}-\overline{Y}^{*}\right)^2}{\sigma_a^2+\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}\right]}{\frac{1}{J}\left[\frac{\sum_{j=1}^{J}\sum_{i=1}^{n_j}\frac{\left(\overline{Y_{ij}}-\overline{Y_j}\right)^2}{n_j\left(n_j-1\right)}}{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}\right]}. \tag{2.36}$$

The numerator and denominator of $\overline{F}$ can be defined such that

$$Num\left(\overline{F}\right)=\left(\frac{1}{J-1}\right)\left[\frac{\sum_{j=1}^{J}\left(\overline{Y_j}-\overline{Y}^{*}\right)^2}{\sigma_a^2+\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}\right] \tag{2.37}$$

$$Den\left(\overline{F}\right)=\frac{1}{J}\left[\frac{\sum_{j=1}^{J}\sum_{i=1}^{n_j}\frac{\left(\overline{Y_{ij}}-\overline{Y_j}\right)^2}{n_j\left(n_j-1\right)}}{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}\right] \tag{2.38}$$

and so $\overline{F}=\frac{Num\left(\overline{F}\right)}{Den\left(\overline{F}\right)}$. Under normality, $\overline{F}$ has an approximate F distribution with estimated degrees of freedom: (Jeyaratnam and Othman, 1985)

$$U=\frac{\left[\left(J-1\right)\sum_{j=1}^{J}\frac{\sigma_j^2}{Jn_j}\right]^2}{\left(\sum_{j=1}^{J}\frac{\sigma_j^2}{Jn_j}\right)^2+\left(J-2\right)\sum_{j=1}^{J}\frac{\sigma_j^4}{Jn_j^2}}\ \text{ and} \tag{2.39}$$

$$V = \frac{\left(\sum_{j=1}^{J} \frac{\sigma_j^2}{n_j} \right)^2}{\sum_{j=1}^{J} \frac{\sigma_j^4}{n_j^2 \left(n_j - 1 \right)}} . \tag{2.40}$$

Since we let $\dfrac{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}{\sigma_a^2 + \frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}} F = \overline{F}$, it can be also written as

$$F = \frac{\sigma_a^2 + \frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}}{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}} \overline{F}$$

$$= \left(\frac{\sigma_a^2}{\frac{1}{J}\sum_{j=1}^{J}\frac{\sigma_j^2}{n_j}} + 1 \right) \overline{F} . \tag{2.41}$$

The Jeyaratnam and Othman proposed test resulted in controlling the probability of the Type I error reasonably well. The estimated actual probabilities of Type I error were quite close to the specified significant level for number of groups $J = 2$ and $J = 3$ when both a_j and ε_{ij} are normally distributed (Jeyaratnam & Othman, 1985).

Wilcox (1994a) shows that his results are consistent with the results in Jeyaratnam and Othman (1985) for normal distributions. However, when distributions were just slightly nonnormal, the Jeyaratnam and Othman (1985) procedure gave Type I error probabilities that were unacceptable (≥ 0.1) in both situations of equal and unequal sample sizes (Wilcox, 1994a).

2.4 THE WILCOX 1994 PROCEDURE

In view of the methodology by Jeyaratnam and Othman in Section 2.3, Wilcox (1994a) continued this line of study by suggesting a generalization of that procedure on traditional REM based on trimmed means. The benefit of using the trimmed mean over the sample mean is such that the trimmed mean has a standard deviation that is unlikely to be affected by heavily skewed distributions. Wilcox (1994a) also shows that the symmetrically trimmed mean value employed in his model can result in a significant gain in power.

Let μ_{tj} be the population trimmed mean for the given *j*-th randomly sampled group. The REM is generalized by letting $\overline{\mu_w} = E_w\left(\mu_{tj}\right)$. In words, $\overline{\mu_w}$ is the Winsorized mean for the population of trimmed mean that is being sampled. It is important to note that μ_{tj} is conditional on the sampled group. This means, for a given *j*-th group, μ_{tj} is its trimmed mean. The model used for this part of study becomes

$$Y_{ij} = \overline{\mu_w} + b_j + \varepsilon_{ij} \tag{2.42}$$

where Y_{ij} is the *i*-th observation sampled from the *j*-th group: ($i = 1, \ldots, n_j$, $j = 1, \ldots, J$), $b_j = \mu_{tj} - \overline{\mu_w}$, $E_w\left(b_j\right) = 0$ and $E_w\left(\varepsilon_{ij}\right) = 0$. The Winsorized variance of b_j is $VAR_w\left(b_j\right)=\sigma_{wb}^2$ the Winsorized variance of ε_{ij} is $VAR_w\left(\varepsilon_{ij}\right)=\sigma_{wj}^2$. Also, $E_w\left(\sigma_{wj}^2\right)=\sigma_{we}^2$, where the Winsorized expectation is taken with respect to a randomly sampled group. When there are no differences among the trimmed means associated with the pool of treatment groups under investigation, $\sigma_{wb}^2=0$. Let $\overline{Y_{tj}}$ be the trimmed mean corresponding to the *j*-th group and so

$$\overline{Y_t}^* = \frac{1}{J}\sum_{j=1}^{J}\overline{Y_{tj}}. \tag{2.43}$$

Wilcox (1994a) then defines

$$Num\left(F_t\right)=\frac{1}{J-1}\sum_{j=1}^{J}\left(\overline{Y_{tj}}-\overline{Y_t}^*\right)^2 \tag{2.44}$$

$$Den\left(F_t\right)=\frac{1}{J}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\frac{\left(Y_{ij}-\overline{Y_{wj}}\right)^2}{\left(n_j-2k_j\right)\left(n_j-2k_j1\right)},\ k_j=\left[\gamma n_j\right] \tag{2.45}$$

where $\left[\gamma n_j\right]$ is the value of γn_j rounded down to the nearest integer. Y_{ij} is the value of Y_i based on the trimmed data in the *j*-th group and $\overline{Y_{wj}}$ is the resulting Winsorized mean with γ being the percentage of trimming. Hence,

$$F_t=\frac{\dfrac{1}{J-1}\displaystyle\sum_{j=1}^{J}\left(\overline{Y_{tj}}-\overline{Y_t}^*\right)^2}{\dfrac{1}{J}\displaystyle\sum_{j=1}^{J}\sum_{i=1}^{n_j}\frac{\left(Y_{ij}-\overline{Y_{wj}}\right)^2}{\left(n_j-2k_j\right)\left(n_j-2k_j1\right)}}. \tag{2.46}$$

Wilcox (1994a) also shows that

$$E\left[Num\left(F_t\right)\right]=\sigma_{wb}^2+\frac{1}{\left(1-2\gamma\right)^2 J}\sum_{j=1}^{J}\frac{\sigma_{wj}^2}{n_j}\ \text{and} \tag{2.47}$$

$$E\left[Den\left(F_t\right)\right] = \frac{1}{\left(1-2\gamma\right)^2 J} \sum_{j=1}^{J} \frac{\sigma_{wj}^2}{n_j}. \tag{2.48}$$

Therefore, when the test of null hypothesis $H_0 : \sigma_{wb}^2 = 0$ is true, the test statistic $F_t = \frac{Num\left(F_t\right)}{Den\left(F_t\right)}$ is considered and has an approximate F distribution with estimated degrees of freedom:

$$U_t = \frac{\left[\left(J-1\right)\sum_{j=1}^{J} q_j\right]^2}{\left(\sum_{j=1}^{J} q_j\right)^2 + \left(J-2\right)J\sum_{j=1}^{J} q_j^2} \text{ and} \tag{2.49}$$

$$V_t = \frac{\left(\sum_{j=1}^{J} Jq_j\right)^2}{\sum_{j=1}^{J} \frac{\left(Jq_j\right)^2}{n_j - 2k_j - 1}}. \tag{2.50}$$

where

$$q_j = \frac{\left(n_j - 1\right)s_{wj}^2}{J\left(n_j - 2k_j\right)\left(n_j - 2k_j - 1\right)}, \tag{2.51}$$

$$s_{wj}^2 = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \left(Y_{ij} - \overline{Y}_j\right)^2 \text{ and} \tag{2.52}$$

$$\overline{Y}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} Y_{ij}. \tag{2.53}$$

Wilcox (1994a) performed simulations of 10000 replications to study the sampling properties of the proposed test as well as the traditional F test and Jeyaratnam and Othman's test. Simulations were carried out for $J = 4$ groups because as J increases, the F test fails to perform fairly for the condition of unequal variances with respect to controlling the probability of the Type I error as reported by Wilcox, Charlin and Thompson (1986). Therefore, $J = 4$ is the ideal value for the F test to perform well.

Wilcox (1994a) reports that the simulations results for the procedures did not meet the criterion set by Bradley (1978). According to Bradley (1978), a test is considered robust if the observed rate of Type I error, α is within the interval 0.5α and 1.5α. Therefore, at the typical five percent level of significance $(\alpha = 0.05)$, a test is considered robust in a particular condition if the observed Type I error rates fall within the interval of 0.025 and 0.075.

Under the condition of normality, Wilcox (1994a) states that the conventional F test did not perform well with unequal sample sizes. As for equal sample sizes, the resulted Type I error of the F test has a significantly high value but this problem was corrected using the $\overline{F}$, as expected. The F_t considered by Wilcox resulted in an estimated probability of Type I error with value 0.079. Wilcox reported that F_t does not perform well in situations of skewed distributions and unequal sample sizes.

In terms of power gain, the conventional F and $\overline{F}$ tests resulted in poor power for heavy-tailed distributions compared to the F_t test that was satisfactory, although in some cases the $\overline{F}$ test did give more power but at the expense of really poor control over the probability of a Type I error. The F_t test performs really well for the situation of unequal sample sizes.

2.5 TRIMMED MEANS IN FIXED EFFECTS MODEL

2.5.1 Symmetric Trimming

The appropriateness of using the parameter of population mean as a measure of location is doubtful when the nature of the population distribution is skewed. According to Wilcox, the use of the usual mean can possibly portray a distorted view of "how the typical individual in one group compares to the typical individual in another, and about accurate probability coverage, controlling the probability of a Type I error, and achieving relatively high power" (Wilcox, 1995a, p. 66). By the substitution of robust measure of location, it becomes more likely to obtain test statistics which are insensitive to the combined effects of variance heterogeniety and nonnormality.

The trimmed mean is preferable because it is easy to obtain and has good theoretical properties (Wilcox, 1995a) as Keselman, Lix and Kowalchuk (1998a) notes, particularly when the standard error of the trimmed mean is less affected by departures from normality. This is because under extreme cases, observations in the heavy tails of a distribution are removed. Also, while the use of trimmed means can be effective, Keselman, Kowalchuk and Lix (1998b) advised that the measure should only be employed in testing for treatment effects across groups using a measure of location that will accurately reflect the typical score within a group when heavy-tailed distributions are involved.

For one-way designs, Lix and Keselman (1998) found that tests of mean equality (on fixed effects models) based on the usual mean and variance were affected by skewness when group sizes were unequal and this problem was improved upon when trimmed means and variances based on Winsorized data were used. Keselman et al. (1998b) states that tests of mean equality under variance heterogeneity and nonnormality have been conducted for nonorthogonal factorial designs which is the Welch-James (WJ) type statistic is able to provide robustness in unbalanced factorial designs when variances were heterogeneous. However, the WJ test had its own limitations when the assumptions of unequal variances and nonnormality were violated for unbalanced fixed-effects factorial designs.

The performance of the WJ test, however, could be improved upon by incorporating robust measures of location and variability instead of relying on the usual least squares estimators. Wilcox (1994b) showed results that using trimmed means could result in a more accurate solution when the distributions have heavy tails because "this type of nonnormality have smaller standard errors compared to least square means (Keselman et al., 1998b, p. 147). Keselman et al. (1998b) had also mentioned that Yuen (1974) suggested for trimmed means and variances based on Winsorized sums of squares to be used in conjunction with Welch (1938)'s statistic. For symmetric distributions that are heavily tailed, Yuen (1974) showed that the statistic based on these robust estimators was able to control the rate of Type I errors and generated greater power than a statistic based on the usual mean and variance.

Continuing with the investigation of the Welch's (1938) statistic, a research by Keselman, Wilcox, Kowalchuk and Olejnik (2002) investigated three tests to compare measure of locations across two groups: Welch (1938) test, the Zhou, Gao and Hui (1997) test and the Yuen (1974) procedure. In the study, the three tests were run under several cases of nonnormality and variance heterogeneity for balanced and unbalanced designs. The results proved the fact by Keselman et al. (1998b) and Lix and Keselman (1998) that Welch (1938)'s test was nonrobust. The study also shows that the procedure by Zhou et al. (1997) could not optimally perform for nonnormal skewed data of other forms such as the chi-square as the test was designed to be used with skewed lognormal data. Keselman, Wilcox, Kowalchuk and Olejnik (2002) notes that the Yuen (1974) statistic tests proved trimmed means - as opposed to the usual least squares means - provided better estimates of the typical individual in distributions that are skewed or have outliers. As pointed out by Zhou et al. (1997), distributions are generally skewed and the results by Keselman et al. (2002) proved that conforming to the use of trimmed means and Winsorized variances will provide enough control over the Type I error probability.

Although several results have proved that trimmed mean is a well accepted measure under the effects of nonnormality and variance heterogeneity, there is always the lingering question of the accuracy of trimming when distributions are skewed: should the data be trimmed symmetrically or asymmetrically and how much trimming should be done. Keselman, Wilcox, Othman and Fradette (2002) demonstrated the advantage of a prior test for symmetry in order to determine whether data should be trimmed from both tails (symmetric trimming) or just from one tail (asymmteric trimming). This approach of the preliminary test for symmetry was a modification by Babu, Padmanabhan and Puri (1999) due to Hogg, Fisher and Randles (1975).

Othman, Keselman, Wilcox, Fradette and Padmanabhan (2002) further proves that the test of symmetry is a powerful procedure as a test for treatment group equality when combined with a heteroscedastic statistic which compared means that are either symmetrically trimmed or asymmetrically trimmed because of the excellent control over Type I errors for very heterogenous distributions that are extremely nonnormal. Othman et al. (2002) also mentioned that Babu et al. (1999) had used the preliminary test for symmetry to establish "whether groups should be compared on their symmetrically determined trimmed means, when distributions were deemed symmetric, or on their medians, when distributions were deemed asymmetric" (Othman et al., 2002, p. 314).

2.5.2 Asymmetric Trimming

Following the course of trimming on fixed effects models, Keselman, Wilcox, Lix, Algina and Fradette (2007) showed several methods that determine if the data distribution should be trimmed and the quantity of trimming from the tails of the distribution. The two main adaptive trimming methods discussed by Keselman et al. (2007) are the Reed and Stark (1996) and the Tukey-McLaughlin-Jaeckel-Hogg methods.

The Reed and Stark (1996) method is based on the work of Hogg (1974, 1982) whereby several adaptive location estimators were defined depending on measures of tail-length and skewness for a set of observations.