

## A Robust Alternative to the $t$ - Test

Zahayu Md Yusof<sup>1</sup>, Suhaida Abdullah<sup>1</sup>, Sharipah Soaad Syed Yahaya<sup>1</sup> & Abdul Rahman Othman<sup>2</sup>

<sup>1</sup> UUM College of Arts and Sciences, Universiti Utara Malaysia, Malaysia

<sup>2</sup> Universiti Sains Malaysia, Pulau Pinang, Malaysia

Correspondence: Abdul Rahman Othman, Pusat Pengajian Pendidikan Jarak Jauh, Universiti Sains Malaysia, USM Penang 11800, Pulau Pinang, Malaysia. E-mail: oarahman@usm.my

Received: February 27, 2012

Accepted: March 21, 2012

Online Published: May 1, 2012

doi:10.5539/mas.v6n5p27

URL: <http://dx.doi.org/10.5539/mas.v6n5p27>

### Abstract

$t$ -test is a classical test statistics for testing the equality of two groups. However, this test is very sensitive to non-normality as well as variance heterogeneity. To overcome these problems, robust method such as  $F_t$  and  $S_t$  tests statistics can be used. This study proposed the use of a robust estimator that is trimmed mean as the central tendency measure in  $F_t$  test and median as the central tendency measure in  $S_t$  test when comparing the equality of two groups. The performance of the  $S_t$  test with  $MAD_n$  was able to give the most convincing result than the other methods. The  $F_t$  with  $MAD_n$  showed comparable results with the conventional methods. This study has shown some improvement in the statistical solution of detecting differences between location parameters. These modified methods may serve as alternatives to some other robust statistical methods which are unable to handle either the problem of non-normality, variance heterogeneity or unbalanced design.

**Keywords:** trimmed mean, robust, Type I error rate, p-value

### 1. Introduction

In recent years, numerous methods for locating treatment effects or testing the equality of central tendency (location) parameters by simultaneously controlling the Type I error and the power to detect treatment effects are being studied. Progress has been made in terms of finding better methods for controlling the Type I error and the power of the test that detects treatment effects in one-way independent group designs (Babu, Padmanabhan & Puri, 1999; Othman et al., 2004; Wilcox & Keselman, 2003). Through a combination of impressive theoretical developments, more flexible statistical methods, and faster computers, serious practical problems that seemed insurmountable only a few years ago can now be addressed. These developments are important to applied researchers because they greatly enhance the ability to discover true differences between groups while maximizing the chance of detecting a genuine positive effect.

One way to overcome the problems of controlling Type I error rates is by using robust statistics. There are several definitions of robust statistics that have been found in the literature and these unfortunately lead to the inconsistency of its meaning. Most of the definitions are based on the objective of the particular study by different researchers (Huber, 1981).

A statistical method is considered robust if the inferences are not seriously invalidated by the violation of such assumptions, for instance non-normality and variance heterogeneity (Scheffe, 1959). Huber (1981) defined robustness as a situation which is not sensitive to small changes in assumptions while Brownlee (1965) reported slight effects on a procedure when appreciable departures from the assumptions were observed.

The theory of robust statistics deals with deviations from the assumptions on the model and is concerned with the construction of statistical procedures which is still reliable and reasonably efficient in a neighborhood of the model (Ronchetti, 2006). Hampel, Ronchetti, Rousseeuw and Stahel (1986), stated that in a broad informal sense, robust statistics is a body of knowledge, partly formalized into "theories of robustness" relating to deviations from idealized assumptions in statistics. As mentioned by Hoel, Port and Stone (1971), a test that is reliable under rather strong modifications of the assumptions on which it was based is said to be robust. Hence in this thesis, a statistical method is considered robust when it has estimators which cannot be influenced by the deviations from the given assumptions when hypothesis testing is being conducted.

Robust statistics has widely been used for many years now. Ronchetti (2006) reported that research in robust statistics has been conducted since 40 years ago and this area of research is still being actively studied today. In

Ronchetti’s (2006) quick search in the Current Index of Statistics, 1617 papers on robust statistics were found between 1987 and 2001 in statistics journals and related fields.

The goal of this study is to search for alternative methods in testing for the equality of central tendency measures by simultaneously controlling Type I error and improving power rates in the one-way independent group design under skewed distributions. The proposed procedures to be adopted in this study are among the latest procedures in robust statistics. The procedures are modified Ft and modified S1 which were proposed by Md Yusof et al. (2007) and Syed Yahaya (2005) respectively. These two procedures are for testing the equality of the central tendency measures. The Ft uses trimmed mean while S1 uses median as the central tendency measures. The performance of these methods in terms of type I error rates for the case of two groups are determined and compared. The performance of the methods was further demonstrated on real education data.

**2. Method**

This paper focuses on the modified  $F_t$  and  $S_t$  methods, which combines  $F_t$  and  $S_t$  statistics with one of the scale estimators suggested by Rousseuw and Croux (1993).

These methods were compared in terms of Type I error under conditions of normality and non-normality which will be represented by skewed  $g$ - and  $h$ - distributions.

*2.1  $F_t$  Statistic*

Lee and Fung (1985) introduced a statistical procedure that is able to handle problems with sample locations when non-normality occurs but the homogeneity of variances assumption still applies. This statistic was named trimmed  $F$  statistic,  $F_t$ . Their work focused on the best trimming percentages used to produce trimmed means which are able to control Type I error and to provide good power rates of the statistical procedure.

They recommended the trimmed  $F$  statistic with 15% symmetric trimming as an alternative to the usual  $F$  test especially when the distribution is long tailed symmetric. This method is simple and easy to program.

To further understand the  $F_t$  method, let

$X_{(1)j}, X_{(2)j}, \dots, X_{(n)j}$  be an ordered sample of group  $j$  with size  $n_j$ .

Hence the  $g$ -trimmed  $F$  is defined as

$$F_t(g) = \frac{\sum_{j=1}^J (\bar{X}_{tj} - \bar{X}_t)^2 / (J - 1)}{\sum_{j=1}^J SSD_{tj} / (H - J)} \tag{1}$$

where

$J$  = number of groups

$$h_j = n_j - g_{1j} - g_{2j}$$

$g_j$  represents the proportion of observations in  $j^{\text{th}}$  group that are to be trimmed in each tail of the distribution.

$$H = \sum_{j=1}^J h_j$$

$$\bar{X}_t = \sum_{j=1}^J h_j \bar{X}_{tj} / H$$

$\bar{X}_{tj}$  = the  $j^{\text{th}}$  group trimmed mean, and  $SSD_{tj}$  = the  $g$ -Winsorized sum of squared deviations.  $F_t(g)$  will follow approximately an  $F$  distribution with  $(J - 1, H - J)$  degrees of freedom. Modification on  $F_t$  was done on the calculation of trimmed mean.

**2.1.1 Trimmed Mean**

Let  $X_{(1)j}, X_{(2)j}, \dots, X_{(n)j}$  be an ordered sample of group  $j$  with size  $n_j$ . MOM trimmed mean of group  $j$  is calculated by using:

$$\bar{\theta}_{tj} = \frac{1}{n_j - g_{1j} - g_{2j}} \left[ \sum_{i=g_{1j}+1}^{n_j-g_{2j}} X_{(i)j} \right] \tag{2}$$

where

$g_{1j}$  = number of observations  $X_{(i)j}$  such that  $(X_{(i)j} - \hat{M}_j) < -K (MAD_n)$

$g_{2j}$  = number of observations  $X_{(i)j}$  such that  $(X_{(i)j} - \hat{M}_j) > K (MAD_n)$

$\hat{M}_j$  = median of group  $j$ ,  $K = 2.24$  (multiplier of scale estimator),  $n_j$  = group sample sizes

For the equal amounts of trimming in each tail of the distribution, the Winsorized sum of squared deviations is defined as

$$SSD_{ij} = (g_j + I) \left( X_{(g_j+1)j} - \bar{X}_{ij} \right)^2 + \left( X_{(g_j+2)j} - \bar{X}_{ij} \right)^2 + \dots + \left( X_{(n_j-g_j-1)j} - \bar{X}_{ij} \right)^2 + (g_j + I) \left( X_{(n_j-g_j)j} - \bar{X}_{ij} \right)^2 \quad (3)$$

When allowing different amounts of trimming in each tail of the distribution, the Winsorized sum of squared deviations is then defined as,

$$\begin{aligned} SSD_{ij} = & (g_{1j} + I) \left( X_{(g_{1j}+1)j} - \bar{\theta}_{ij} \right)^2 + \left( X_{(g_{1j}+2)j} - \bar{\theta}_{ij} \right)^2 + \dots \\ & + \left( X_{(n_j-g_{2j}-1)j} - \bar{\theta}_{ij} \right)^2 + (g_{2j} + I) \left( X_{(n_j-g_{2j})j} - \bar{\theta}_{ij} \right)^2 \\ & - \left\{ (g_{1j}) \left[ X_{(g_{1j}+1)j} - \bar{\theta}_{ij} \right] + (g_{2j}) \left[ X_{(n_j-g_{2j})j} - \bar{\theta}_{ij} \right] \right\}^2 / n_j \end{aligned} \quad (4)$$

## 2.2 $S_1$ Statistic

To understand  $S_1$ , consider the problem of comparing location parameters for skewed distributions. Let  $Y_{ij} = (Y_{1j}, Y_{2j}, \dots, Y_{n_jj})$  be a sample from an unknown distribution  $F_j$  and let  $M_j$  be the population median  $F_j$ ;  $j = 1, 2, \dots, J$ . For testing  $H_0: M_1 = M_2 = \dots = M_J$  versus  $H_1: M_i \neq M_j$  for at least one pair of  $(i, j)$ , the  $S_1$  statistic is defined as

$$S_1 = \sum_{1 \leq i < j \leq J} |s_{ij}| \quad (5)$$

$$\text{where } s_{ij} = \frac{(\hat{M}_i - \hat{M}_j)}{\sqrt{(\hat{\omega}_i + \hat{\omega}_j)}}$$

$$\omega_j = \left( \frac{1}{n_j} \sum |Y_{ij} - \hat{M}_j| \right)^2, \quad \hat{\omega}_j = \frac{\omega_j}{n_j}$$

$\hat{M}_j$  is the sample median from the  $j$ th group, of group  $j$

$\omega_j$  is the squared mean absolute deviation from sample median  $\hat{M}_j$

$n_j$  is the sample size for group  $j$ .

Modification on  $S_1$  was done by substituting the default scale estimator,  $\hat{\omega}_j$  with the well known robust scale estimator,  $MAD_n$ .

## 2.3 $MAD_n$

$MAD_n$  is the median absolute deviation about the median. It demonstrates the best possible breakdown value of 50%, twice as much as the interquartile range and its influence function is bounded with the sharpest possible bound among all scale estimators (Rousseeuw & Croux, 1993).

This robust scale estimator is given by

$$MAD_n = b \text{ med } |x_i - \text{med } x_j| \quad (6)$$

where the constant  $b = 1.4826$  is needed to make the estimator consistent for the parameter of interest,  $x_i = x_1, x_2, \dots, x_n$  and  $i > j$

However, there are drawbacks in this scale estimator. The efficiency of  $MAD_n$  is very low with only 37% at Gaussian distribution. Rousseeuw and Croux (1993) carried out a simulation on 10,000 batches of Gaussian observations to verify the efficiency gain at finite samples. They compared the variance of the standard deviation with the variance of  $MAD_n$  based on the finite samples.  $MAD_n$  also takes a symmetric view on dispersion and does not seem to be a natural approach for problems with asymmetric distributions.

### 2.4 Bootstrap Method

The bootstrap is a Monte Carlo method that can be used to estimate the standard error of any estimator  $\hat{\theta}$  and was introduced by Efron (1979). The advantage of bootstrapping is its simplicity. This method is straightforward to apply to derive estimates of standard errors and confidence intervals for complex estimators of complex parameters of the distribution, such as percentile points, proportions, odds ratio, and correlation coefficients. Staudte and Sheather (1990) in their study stated that bootstrap is used to indicate that the observed data are used not only to obtain an estimate of the parameter but also to generate new samples. Bootstrap can routinely answer questions far too complicated for traditional statistical analysis. *They work the same way (without formulae) for many different statistics in many different settings.* In addition, bootstrapping can help in increasing accuracy of the test statistic.

When the sampling distribution of the estimator of interest is unknown, a pseudo sampling distribution of the estimator can be estimated using bootstrap. With the establishment of the pseudo sampling distribution, we can now access variability of an estimator, bias of an estimator and significance of a test involving the estimator (Efron, 1979).

Bootstrap method is known to yield a better approximation than the one based on the normal approximation theory (Babu & Padmanabhan, 1996; Babu et al., 1999). Othman, Keselman, Padmanabhan, Wilcox and Fradette (2003) listed out two practical advantages of using bootstrap methods as detailed below:

- i) Theory and empirical findings indicate that they can result in better Type I error control than non-bootstrap methods.
- ii) There are some bootstrap methods which do not require the knowledge of the sampling distribution of the test statistic. This makes hypothesis testing quite flexible.

Westfall and Young (1993) suggested that Type I error control could be improved by combining bootstrap methods with methods based on trimmed means. The bootstrap seems preferable for general use if the goal is to avoid Type I error probability greater than the nominal level (Wilcox, 1998). The strategy behind the bootstrap is to use the shifted empirical distributions to estimate an appropriate critical value (Othman et al., 2003). Keselman, Wilcox and Lix (2003) stated that, further improvement in Type I error control is often possible by obtaining critical values for test statistic through bootstrap methods.

In this study, the methods of bootstrap and non-bootstrap are compared. The bootstrap procedures on  $S_j$  statistic and approximate procedures on  $F_t$  statistic

### 3. Empirical Investigation

This paper only focused on unequal sample sizes and heterogeneous variances for two groups with small samples. A group of size  $N = 40$  was chosen. The sample were set at  $n_1 = 15$  and  $n_2 = 25$ .

Each method will be tested under two types of distributions with  $g = 0.0$  and  $h = 0.0$  (normal) and  $g = 0.5$  and  $h = 0.5$  (skewed leptokurtic). The  $g$ - and  $h$ - distributions were first proposed by Hoaglin (1985). These distributions are transformations of the standard normal distribution. By manipulating the  $g$ - parameter one can transform the standard normal distribution into a skewed distribution. In addition to this, one can also transform the standard normal distribution into a heavy tailed distribution by changing the  $h$ - parameter. For this study, 5000 datasets were simulated for each of the procedure. The random samples were drawn using SAS generator RANNOR (SAS Institute Inc, 1999).

To test the Type I error, the group means were (0, 0, 0, and 0). For each design, 5000 datasets were simulated. For  $S_j$  statistic 599 bootstrap samples were generated.

Table 1. Design specification

Pairing	Group Sizes		Group Variances	
	1	2	1	2
Positive	15	25	1	36
Negative	15	25	36	1

### 4. Simulation Results

The robustness of a method is determined by its ability in controlling the Type I error. By adopting Bradley's liberal criterion of robustness (Bradley, 1978), a test can be considered robust if its empirical rate of Type I error, is

within the interval  $0.5\alpha$  and  $1.5\alpha$ . If the nominal level is  $\alpha = 0.05$ , the empirical Type I error rate should be in between 0.025 and 0.075. Correspondingly, a test is considered to be non-robust if, for any particular condition, its Type I error rate is not within this interval. We chose this criterion since it was widely used by most robust statistic researchers (e.g. Keselman et al., 2000; Othman et al., 2004; Syed Yahaya et al., 2004; Wilcox et al., 2000) to judge robustness. Nevertheless, for Guo and Luh (2000), if the empirical Type I error rate do not exceed the 0.075 level, it is considered robust. The best procedures are those procedures that can produce Type I error rates closest to the nominal (significance) level.

The Type I error rates for two groups case is presented in Table 2. The second column of the table displays the pairing categories. Positive pairing refers to the case in which the largest sample size is associated with population having the largest variance and the smallest sample size is associated with the population having the smallest variance. While negative pairing refers to the case in which the smallest sample size is associated with the population having the largest variance, and the largest sample size is associated with the population having the smallest variance.

Table 2. Type I error rates

Distribution	Pairing	Methods			
		$S_1$ with $MAD_n$	$F_t$ with $MAD_n$	$t$ -test	Mann Whitney
N(0.1) - normal	+ve	<b>0.0524</b>	<b>0.0460</b>	<b>0.0270</b>	<b>0.0508</b>
	-ve	<b>0.0564</b>	0.2068	0.1290	0.1244
	Average	<b>0.0544</b>	0.1264	0.0780	0.0876
$g = 0.5$ and $h = 0.5$ - extremely skewed	+ve	<b>0.0294</b>	0.0142	0.0140	<b>0.0420</b>
	-ve	<b>0.0314</b>	0.1446	0.1020	0.1080
	Average	<b>0.0304</b>	0.0794	<b>0.0580</b>	<b>0.0750</b>

Discription: As shown in Table 2, the Type I error rates produced by  $S_1$  with  $MAD_n$  are robust across the two types of distributions, while Mann-Whitney produced robust value for positive pairing only. However, for the  $F_t$  with  $MAD_n$  and  $t$ -test the Type I error rates are robust only under one condition i.e. positive pairing under normal distribution.

## 5. Analysis on Real Data

The performance of the modified  $S_1$  and  $F_t$  methods were demonstrated on real data. The following sections discuss the performance of both methods. We also compare the methods with the parametric and non parametric methods.

Two classes (groups) of Statistical Distribution Theory (2<sup>nd</sup> Semester 2010/2011) were randomly chosen. The final marks were recorded and tested for the equality between the classes. The sample sizes for Class 1 and 2 were 48 and 46 respectively. The descriptive statistics for each of the groups and the results of the test in the form of  $p$ -values are given in Table 3 and Table 4 respectively.

Table 3. Descriptive statistics for each group

Group	Sample size (N)	Mean of the marks	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	48	62.89	20.258	2.924	-14.249	1.0845	21	96
2	46	69.47	16.941	2.497	-14.222	1.0572	34	98

Table 4. Results of the test using different methods

Methods	<i>p</i> -value
$S_I$ with $MAD_n$	0.0133
$F_t$ with $MAD_n$	0.0915
<i>t</i> -test	0.0920
Mann-Whitney	0.1430

Description: Table 4 shows that only  $S_I$  with  $MAD_n$  method produce significant result (reject the null hypothesis). This indicate that this method able to detect the difference which exists between the groups. When testing using *ANOVA*, Mann-Whitney and  $F_t$  with  $MAD_n$  methods the result fails to reject the null hypothesis such that the performance for all groups is equal. Based on the *p*-values, the Mann Whitney procedure can least detect the difference (large *p*-value). The results (real data) are consistent with the simulation data result with  $S_I$  produced consistent and robust result as compared to the other three methods.

## 6. Discussion and Conclusions

The goal of this paper is to find the alternative procedures in testing location parameter for skewed distribution by simultaneously controlling the Type I error and power rates. Classical method such as *t* - test and *ANOVA* is not robust to non-normality and heteroscedasticity. When these problems occur at the same time, the Type I error will increase causing wary rejection of the null hypothesis and power of test can be substantially reduced from theoretical values, which will result in differences going undetected. Realizing the need of a good statistic in addressing these problems, we integrate the  $S_I$  statistic by Babu et al. (1999) and  $F_t$  statistic introduced by Lee and Fung (1985) with the high breakdown scale estimators of Rousseeuw and Croux (1993) and these new methods are known as the modified  $S_I$  and  $F_t$  methods. This study has shown some improvement in the statistical solution of detecting differences between location parameters.

In controlling the Type I error rate, the study reported in this study leads us to formulate the following conclusions and recommendations. For both distributions (normal and extremely skewed), the robust methods ( $S_I$  with  $MAD_n$  and  $F_t$  with  $MAD_n$ ) showed comparable results with the conventional methods (*t* - test and Mann-Whitney). However, the demonstration on real data showed that  $S_I$  with  $MAD_n$  method was able to give the most convincing result than the other methods.

It is our impression that applied researchers would prefer a method that compared treatment performance across groups with a measure for the typical score which was based on as much as the original data as possible. Modified  $S_I$  will be the best choice for this purpose because when working with two groups' case, the researchers can work with the original data without having to worry about shape of the distribution.

## Acknowledgments

The authors would like to acknowledge the work that led to this paper, which was partially funded by the University Research Grant Scheme of the Universiti Utara Malaysia.

## References

- Babu, J. G., & Padmanabhan, A. R. (1996). A robust test for omnibus alternatives. *Madan Puri Festschrift*, 319-327.
- Babu, J. G., Padmanabhan, A. R., & Puri, M. P. (1999). Robust one-way ANOVA under possibly non-regular conditions. *Biometrical Journal*, 41(3), 321-339.
- Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31, 144-152. <http://dx.doi.org/10.1111/j.2044-8317.1978.tb00581>
- Brownlee, K. A. (1965). *Statistical theory and methodology in science and engineering* (2nd ed.). New York: Wiley.
- Efron. (1979). Bootstrap methods: Another look at the Jackknife. *Annals of Statistics*, 7, 1-26. <http://dx.doi.org/10.1214/aos/1176344552>
- Guo, J. H., & Luh, W. M. (2000). An invertible transformation two-sample trimmed *t*-statistic under heterogeneity and nonnormality. *Statistic & Probability letters*, 49, 1-7.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (1986). *Robust statistics*. New York: Wiley.

- Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (1983). *Understanding robust and exploratory data analysis*. New York: Wiley.
- Hoel, P. G., Port, S. C., & Stone, C. J. (1971). *Introduction to statistical theory*. Boston: Houghton Mifflin.
- Huber, P. J. (1981). *Robust statistics*. New York: Wiley.
- 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. <http://dx.doi.org/10.1348/000711000159286>
- Keselman, H. J., Wilcox, R. R., & Lix, L. M. (2003). A generally robust approach to hypothesis testing in independent and correlated group designs. *Psychophysiology*, *40*, 586-596. <http://dx.doi.org/10.1111/1469-8986.00060>
- Lee, H., & Fung, K. Y. (1985). Behaviour of trimmed  $F$  and sine-wave  $F$  statistics in one-way ANOVA. *Sankhya: The Indian Journal of Statistics*, *47*(Series B), 186-201.
- Md Yusof, Z., Othman, A. R., & Syed Yahaya, S. S. (2007). Type I error rates of trimmed  $F$  statistic. *In proceeding of the 56th Session of the International Statistical Institute (ISI 2007)*, 22-29. Lisbon Portugal. (In CD).
- Othman, A. R., Keselman, H. J., Padmanabhan, A. R., Wilcox, R. R., & Fradette, K. (2003). An improved Welch-James test statistic. *In proceeding of the Regional Conference on Integrating Technology in the Mathematical Sciences 2003*, Universiti Sains Malaysia, Pulau Pinang, Malaysia.
- Othman, A. R., Keselman, H. J., Padmanabhan, A. R., Wilcox, R. R., & Fradette, K. (2004). Comparing measures of the 'typical' score across treatment groups. *British Journal of Mathematical and Statistical Psychology*, 215-234. <http://dx.doi.org/10.1348/0007110042307159>
- Ronchetti, E. M. (2006). The historical development of robust statistics. In Proceedings of the 7<sup>th</sup> International Conference on Teaching Statistics (ICOTS-7). Salvador, Brazil. International Statistical Institute: Working Cooperatively in Statistics Education. Retrieve from [http://www.stat.auckland.ac.nz/~iase/publications/17/3B1\\_RONC.pdf](http://www.stat.auckland.ac.nz/~iase/publications/17/3B1_RONC.pdf)
- Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, *88*, 1273-1283.
- SAS Institute Inc. (1999). *SAS/IML User's Guide version 8*. Cary, NC: SAS Institute Inc.
- Staudte, R. G., & Sheather, S. J. (1990). *Robust estimation and testing*. New York: Wiley.
- Scheffe, H. (1959). *The analysis of variance*. New York: Wiley.
- Syed Yahaya, S. S., Othman, A. R., & Keselman, H. J. (2004). Testing the equality of location parameters for skewed distributions using  $S_l$  with high breakdown robust scale estimators. In M. Hubert, G. Pison, A. Struyf and S. Van Aelst (Eds.). *Theory and Applications of Recent Robust Methods, Series: Statistics for Industry and Technology*, Birkhauser, Basel. 319 – 328.
- Syed Yahaya, S. S. (2005). *Robust statistical procedures for testing the equality of central tendency parameters under skewed distributions*. Unpublished Ph.D. thesis, Universiti Sains Malaysia.
- Westfall, P. H., & Young, S. S. (1993). *Resampling-based multiple testing*. New York: Wiley.
- Wilcox, R. R. (1998). The goal and strategies of robust methods. *British Journal of mathematical and Statistical Psychology*, *51*, 1-39. <http://dx.doi.org/10.1111/j.2044-8317.1998.tb00659>
- Wilcox, R. R., Keselman, H. J., Muska, J., & Cribbie, R. (2000). Repeated measures ANOVA: Some new results on comparing trimmed means and means. *British Journal of Mathematical and Statistical Psychology*, *53*, 69-82. <http://dx.doi.org/10.1037/1082-989X.13.2.110>
- Wilcox, R. R., & Keselman, H. J. (2003). Modern robust data analysis methods: Measures of central tendency. *Psychological Methods*, *8*, 254-274. <http://dx.doi.org/10.1037/1082-989X.8.3.254>