



GLOBAL JOURNAL OF SCIENCE FRONTIER RESEARCH: F  
MATHEMATICS AND DECISION SCIENCES  
Volume 16 Issue 6 Version 1.0 Year 2016  
Type : Double Blind Peer Reviewed International Research Journal  
Publisher: Global Journals Inc. (USA)  
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

# A Study on Sensitivity and Robustness of One Sample Test Statistics to Outliers

By Kayode Ayinde, Taiwo Joel Adejumo & Gbenga Sunday Solomon

*Ladoke Akintola University of Technology*

**Abstract-** Outliers are observations that stand too different from others in a set of observations. When present in a data set, they affect both descriptive and inferential statistics. This work therefore, studies the sensitivity and robustness of one sample test statistics to outliers so as to know the appropriate one to test hypothesis about the population parameter when outliers are present. One sample test statistics considered are: parametrics test (Student t-test and z-test), non-parametric test (Wilcoxon Sign test (Distribution Sign test (DST), Asymptotic Sign test (AST)), Wilcoxon Signed rank test (Distribution Wilcoxon Signed rank test (DWST) and Asymptotic (AWST)), t-test for rank transformation (Rt-test) and Trimmed t-test statistics (Tt-test). Monte Carlo experiments, replicated five thousand (5000) times, were conducted at eight (8) sample sizes (10, 15, 20, 25, 30, 35, 40 and 50) by simulating data from normal distribution. At each of the sample sizes, 10% and 20% of the generated data were randomly selected and invoked with various magnitude of outliers (-10, -9, -8,... 8, 9, 10).

**Keywords:** outliers, type 1 error rate, sensitivity, robustness, inferential test statistics, levels of significance.

**GJSFR-F Classification:** MSC 2010: 13P25



*Strictly as per the compliance and regulations of :*





# A Study on Sensitivity and Robustness of One Sample Test Statistics to Outliers

Kayode Ayinde <sup>α</sup>, Taiwo Joel Adejumo <sup>σ</sup> & Gbenga Sunday Solomon <sup>ρ</sup>

**Abstract-** Outliers are observations that stand too different from others in a set of observations. When present in a data set, they affect both descriptive and inferential statistics. This work therefore, studies the sensitivity and robustness of one sample test statistics to outliers so as to know the appropriate one to test hypothesis about the population parameter when outliers are present. One sample test statistics considered are: parametrics test (Student t-test and z-test), non-parametric test (Wilcoxon Sign test (Distribution Sign test (DST), Asymptotic Sign test (AST)), Wilcoxon Signed rank test (Distribution Wilcoxon Signed rank test (DWST) and Asymptotic (AWST)), t-test for rank transformation (Rt-test) and Trimmed t-test statistics (Tt-test). Monte Carlo experiments, replicated five thousand (5000) times, were conducted at eight (8) sample sizes (10, 15, 20, 25, 30, 35, 40 and 50) by simulating data from normal distribution. At each of the sample sizes, 10% and 20% of the generated data were randomly selected and invoked with various magnitude of outliers (-10, -9, -8, ... 8, 9, 10). The test statistics were compared at three levels of significance, 0.1, 0.05 and 0.01. A test is considered robust if its estimated error rate approximates the true error rate and has the highest number of times it approximates the error rate when counted over the percentage (%) of outliers, magnitudes of outliers and levels of significance; and if the counts is minimum the test statistics is sensitive. At all the three (3) levels of significance, results revealed that Type 1 error rates of Student t-test, Rt-test and AWST statistics are good; and that z-test and Student t-test statistics are most sensitive to outliers. The statistics robustness is affected by the levels of significance in that the sign test (DST and AST) is robust at 0.1; Tt-test and Wilcoxon Sign Rank test (DWST and AWST) at 0.05; and DST, AWST, Tt-test and AST at 0.01 level of significance. Consequently, the Sign test and Tt-test statistics are recommended for hypothesis testing in the presence of outliers.

**Keywords:** outliers, type 1 error rate, sensitivity, robustness, inferential test statistics, levels of significance.

## I. INTRODUCTION

Edgeworth (1887) defined outliers as discordant observations which present the appearance of differing in respect of their law of frequency from other observations with which they are combined. Hawkins (1980) defined outlier as an observation which deviates so much from the other observations as to arouse suspicious that it was generated by a different mechanism. Also, Barnett, Lewis and John (1994) indicated outlier as outlying observation which appears to deviate markedly from other members of the sample in which it occurs. According to Osborne and Amy (2004), some of the major causes of outliers in the data set include: variability in the measurement or measurement error, data collection errors, data entry errors, invalidity of theory, intentional or motivated mis-reporting, standardization failure and faulty distributional assumptions. If outliers are present in the data and deleted, Osborne and Amy (2004) claimed that their deleterious effect may: alter the odds of making both Type 1 and Type 11 errors, seriously bias or influence estimates that may be of substantive interest,

*Author α σ ρ:* Department of Statistics, Ladoké Akintola University of Technology, P. M. B. 4000, Lautech, Ogbomoso, Oyo State, Nigeria.  
e-mails: kayinde@lautech.edu.ng, talk2adejumotj@gmail.com, gssolomon@outlook.com

generally serve to increase error variance and reduce the power of statistical tests. Some descriptive statistics of a data set such as arithmetic mean have been noted to be affected by outliers while the median is not being affected, other measures especially the arithmetic mean is affected. Most parametric test statistics including the Student t-test (Gosset, 1908), z-test (Gauss, 1809), etc. are directly meant for hypothesis testing about the mean while the non-parametric ones test hypothesis about the median. The non-parametric inferential test statistics for one sample problem include the Sign test (John, 1710) and Wilcoxon signed rank test (Wilcoxon, 1945). The t-test for rank transformation (Rt-test) by Conover and Ronald (1981) tried to bridge the gap between the parametric and the non-parametric while the trimmed t-test (Tt-test) by Yuen (1974) excluded outliers in its test procedure. The effect of outliers on these various test statistics needs to be investigated as this inevitably affects inference. Therefore, this research work examines the effect of outliers on these test statistics so as to determine the sensitive and robust ones to outliers.

## II. REVIEW ON SOME INFERENTIAL TEST STATISTICS

Many inferential test statistics which are meant for hypothesis testing about the mean and median of one sample problem have been discussed in literatures. Some of these test statistics are discussed as follows:

### a) Parametric statistics

Parametric tests are tests in which probability density function of the population in which sample is taken is known or assumed to be normal. The data used must at least be at interval scale and method does not require ranking of observations. However, it has long been established that moderate violations of parametric assumptions have little or no effect on substantive conclusions in most instances (Cohen, 1969).

#### i. One sample z-test

The z-test is one of the most popular techniques for statistical inference based on the assumption of normal distribution (Gauss, 1809; Agresti and Finlay, 1997). The test statistic requires the variance of the population. The test statistic for one sample z-test is defined as:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1) \quad (1)$$

where:  $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$  is the sample mean,  $\mu_0$  is the population mean or hypothesized mean,  $\sigma$  is the standard deviation and n is sample size.

#### ii. One sample student t-test

The student t-test was developed in the early 1900s by a statistician named (Gosset, 1900) who was working at the Guinness brewery and is called the student t-test because of proprietary rights. A single sample t-test is used to determine whether the mean of a sample is different from a known population mean. The t-test uses the mean, standard deviation and number of samples to calculate the test statistic. Plackett and Barnard, (1990).

One sample t-test has all the assumptions of z-test except that of small sample size. Its test statistic is as follows:

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \tag{2}$$

where  $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$  (the standard deviation),  $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$  is the sample mean,  $\mu_0$  is the hypothesized mean value and  $n$  is the sample size. This t-statistic is known to follow a t-distribution with  $n-1$  degree of freedom.

*b) Non-parametric statistics*

Non-parametric test is an alternative to parametric tests when the exact distributions are not known or when the level of measurement is weaker than the interval level. Moreover, non-parametric tests are generally designed to test hypothesis that do not concern population parameters, but are based on the shape of the population frequency distributions. Generally, if the data do not meet the criteria for a parametric test (normally distributed, equal variance and continuous), it must be analyzed with a nonparametric test.

*i. The sign test*

The sign test which was discovered by John (1710) is used to compare a single sample with some hypothesized value, and it is therefore of use in the situation in which the one sample or paired t-test might traditionally be applied. The sign test is so called because it allocates a sign, either positive (+) or negative (-) to each observation according to whether it is greater or less than some hypothesized median value and considers whether this is substantially different from what we would expect by chance. The test statistics are  $T^+$  or  $T^-$ . Under  $H_0$ , binomial distribution is used and  $H_0$  is rejected if  $P(T \leq T^+) < \alpha$  or  $P(T \leq T^-) < \alpha$ .  $\alpha/2$  is used instead of  $\alpha$  when the test is two-tail. Asymptotically, the sign test has its distribution following binomial  $(n, 1/2)$ . The asymptotic test statistic for sign test is:

$$z = \frac{T^+ - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \sim N(0,1) \tag{3}$$

*ii. Wilcoxon signed rank test*

Wilcoxon signed-rank test is named after Wilcoxon (1945) who in a single paper proposed both the test and rank-sum test for two independent samples. The test was further popularized by Siegel (1956) who used the symbol  $T$  for value related to, but not the same.

The asymptotic distribution of Wilcoxon signed rank test is:

$$T = \frac{T^+ - E_0(T^+)}{\sqrt{V_0(T^+)}} \sim N(0,1) \tag{4}$$

where  $E_0(T^+) = \frac{(n+1)}{4}$  and  $V_0(T^+) = \frac{\sqrt{n(n+1)(2n+1)}}{24}$

c) *T-test for Rank Transformation in one and matched pairs sample*

Conover and Ronald (1981) proposed t-test for rank transformation which bridged the gap between parametric and non-parametric test statistics. The test statistic is defined as follows: Let  $D_1, D_2, \dots, D_n$  represent independent random variables with a common mean where in the case of matched pairs  $(X_i, Y_i); D_i = X_i - Y_i$ . But, for one sample  $D_i = X_i - \mu_0$  where  $\mu_0$  is the hypothesized mean value and  $R_i = (\text{sign } D_i) \times (\text{rank of } |D_i|)$ . The test statistic is defined as:

$$T = \frac{\sum_{i=1}^n R_i}{\sqrt{\sum_{i=1}^n R_i^2}} \tag{5}$$

Year 2016

102 The alternative t-test statistic is computed on the signed ranks as;

$$t_R = \frac{\sum_{i=1}^n R_i}{\sqrt{\frac{n \sum_{i=1}^n R_i^2 - (\sum_{i=1}^n R_i)^2}{n-1}}} \tag{6}$$

which is compared with the t-distribution (n-1) degree of freedom. Also,  $t_R$  can be expressed as:

$$t_R = \frac{T}{\sqrt{\frac{n}{n-1} - \frac{T^2}{n-1}}} \tag{7}$$

which is a monotonic function of T.

d) *The Trimmed t-test*

Yuen (1974) proposed the Trimmed t-test for the independent two-sample case, under unequal population variances (Keselman, Wilcox, Algina, and Fradette, 2008).

In each sample, the trimmed mean is computed by removing g-observations from each tail of the distribution. The trimmed mean is computed as follows:

$$\bar{X}_t = \frac{X_{g+1} + X_{g+2} + \dots + X_{n-g}}{n - 2g} \tag{8}$$

Where  $x_1, \dots, x_n$  are the ordered values in a sample.

$g$  = observations trimmed from each tail of the sample distribution.

$n - 2g$  = the number of observations in the trimmed sample.

In addition to the trimmed mean, the Winsorized mean is required to compute the appropriate variance estimate. Instead of "trimming" this method replaces the most extreme  $g$ - observations by the next-most-extreme value. The Winsorized mean is computed as:

$$\bar{X}_w = \frac{([g + 1]X_{g+1} + X_{g+2} + \dots + [g + 1]X_{n-g})}{n} \tag{9}$$

Global Journal of Science Frontier Research (F) Volume XVI Issue VI Version I

The Winsorized sum-of-squared derivation is computed as:

$$SSD_w = [g + 1][x_{g+1} - \bar{X}_w]^2 + [x_{g+2} - \bar{X}_w]^2 + \dots + [g + 1][x_{n-g} - \bar{X}_w]^2 \quad (10)$$

The Winsorized variance is obtained as:

$$S_w^2 = \frac{SSD_w}{n - 2g - 1} \quad (11)$$

Also, the obtained value of the trimmed t-test for one sample is obtained by dividing the difference between the trimmed mean and the hypothesized mean by the estimated standard error.

Hence, the test statistic is defined as:

$$t_w = \frac{\bar{x}_t - \mu_t}{\frac{S_w}{\sqrt{n - 2g - 1}}} \quad \text{where } t_w \text{ follows a t-distribution} \quad (12)$$

### III. METHODOLOGY

#### a) The Monte Carlo Experiments

Data were generated from the univariate normal distribution using Monte Carlo simulation procedures with the aids of R-statistical programming codes.

$X_i$  is generated with mean ( $\mu$ ) = 10 and standard deviation ( $\sigma$ ) = 5. The percentage (%) of outliers to be invoked into the generated data were 10% and 20% while the magnitude of the outliers (f) were taken as: -10, -9, -8, ..., 8, 9, and 10. The experiments were conducted five thousand times (R=5000) at eight sample sizes namely: 10, 15, 20, 25, 30, 35, 40 and 50.

The procedures for invoking the outliers and estimation of the Type 1 errors are as follows:

- (i) Choose a percentage of the data to be replaced with outliers.
- (ii) Choose a particular magnitude of outlier to invoke into the generated data
- (iii) Choose a sample size to work with, say n.
- (iv) Generate random sample with size n from a normal distribution with  $\mu = 10$  and  $\sigma = 5$ ;  $X_i \sim N(10, 25)$
- (v) Randomly select those observations making up the percentage of the generated data to be replaced with outliers.
- (vi) Invoke outliers as follows:

$$X(i)_{\text{outlier}} = f * \text{Max}(X_i) + X_i \quad (13)$$

Where:  $X_i$  = selected generated observation i

$X(i)_{\text{outlier}}$  = Outlier to replace  $X_i$

f = Magnitude of outliers

Max( $X_i$ ) = Maximum of the generated data

- (vii) Replace the outliers in the data originally generated in (iv)
- (viii) Apply the various test statistics and keep their p-values.
- (ix) For each test statistics in (viii), define:



$$G_i = \begin{cases} 1, & \text{if } p\text{-value} < \alpha \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

Where  $\alpha$  is a preselected level of significance, say 0.1.

(x) Repeat steps (v) to (ix) five thousand (5000) times,  $R = 5000$ .

(xi) For each of the test statistics, sum the results obtained in step (x), i.e

$$G = \sum_{i=1}^R G_i \tag{16}$$

(xii) For each of the test statistics, divide the result in step (xi) by the number of replications to estimate the Type 1 error of each test statistics, i.e

$$Q_\alpha = \frac{\sum_{i=1}^R G_i}{R} = \frac{G}{R} \tag{17}$$

(xiii) Choose another magnitude of outlier to invoke into the generated data and repeat step (v) to (xii).

(xiv) Repeat step (v) to (xiii) until all the magnitudes of the outliers are exhausted.

(xv) Choose another sample size to work with and repeat step (v) to (ix)

(xvi) Repeat step (v) to (xv) until all the sample sizes are exhausted.

(xvii) Choose another percentage of the data to be replaced with outliers and repeat step (ii) to (xvi).

(xviii) Repeat step (ii) to (xvii) until all the levels of percentages are exhausted.

*b) Sensitivity and Robustness of the Test Statistics*

The test statistics were considered robust if their estimated Type 1 error rates at different % and magnitude of outliers are within the preferred interval of levels of significance suggested by Kuranga (2015) and used by Ayinde et.al (2016). This is presented in Table 1.

On the other hand, the test statistics were considered sensitive if as percentage of outliers and magnitude of outliers increase as the Type 1 error rates of the test statistics also increases. i.e many do not fall into the preferred intervals (the null hypothesis is rejected often).

*Table 1: The True level of significance and preferred interval*

Level of significance	Preferred interval
0.1	0.095 - 0.14
0.05	0.045 - 0.054
0.01	0.005 - 0.014

*Source: Kuranga (2015) and Ayinde et.al (2016)*

The number of times Type 1 error rates fell within the preferred interval was counted over the levels of percentage of outliers, magnitude of outliers and sample size. A test statistic that is robust is expected to have the highest number of counts, the mode; and that which is sensitive is to have the smallest.

IV. RESULTS AND DISCUSSION

The results of Type 1 error rates of one sample inferential test statistics as affected by outliers at 0.1, 0.05 and 0.01 levels of significance are respectively presented in Table 2, Table 3 and Table 4 and discussed.

a) *Results of Type 1 Error Rates on one sample test statistics at 0.1 level of significance*

From Table 2, it can be observed that all of the Type 1 error rates of Student t-test, Rt-test and AWST performed very well across all sample sizes. The z-test, DWST, AST and Tt-test, in this order also did well but not at all the sample sizes. However, the performance of DST is not good across all sample sizes.

*Table 2:* Results of Type 1 Error Rates on one sample test statistics at 0.1 level of significance

Test statistics	Sample size							
	10	15	20	25	30	35	40	50
Sd.t-test	0.0962	0.1026	0.1012	0.098	0.098	0.1	0.096	0.0974
z-test	0.0954	0.0978	0.1004	0.098	0.1	0.09	0.096	0.0948
Rt-test	0.105	0.0972	0.0952	0.099	0.101	0.1	0.097	0.0986
DWST	0.0828	0.0972	0.0952	0.094	0.095	0.09	0.093	0.0986
AWST	0.105	0.1098	0.0952	0.099	0.101	0.1	0.097	0.1004
DST	0.0208	0.0356	0.0366	0.043	0.091	0.09	0.075	0.062
AST	0.1082	0.119	0.1114	0.105	0.091	0.09	0.075	0.1152
Tt-test	0.0956	0.0978	0.1	0.099	0.094	0.09	0.094	0.0978

*Source: Simulation results*

b) *Results of Type 1 Error Rates on one sample test statistics at 0.05 level of significance*

From Table 3, it can be observed that all of the Type 1 error rates of the test statistics generally did well except that of the Sign test

*Table 3:* Results of Type 1 Error Rates on one sample test statistics at 0.05 level of significance

Test statistics	Sample size							
	10	15	20	25	30	35	40	50
Sd.t-test	0.0486	0.0504	0.046	0.05	0.05	0.045	0.05	0.0502
z-test	0.0492	0.0502	0.051	0.05	0.05	0.048	0.05	0.0482
Rt-test	0.0468	0.0552	0.05	0.05	0.05	0.048	0.05	0.0494
DWST	0.0468	0.048	0.046	0.05	0.05	0.046	0.05	0.0478
AWST	0.0468	0.048	0.046	0.05	0.05	0.046	0.05	0.0488
DST	0.0208	0.0356	0.037	0.04	0.04	0.041	0.04	0.033
AST	0.0208	0.0356	0.037	0.04	0.04	0.041	0.04	0.062
Tt-test	0.0488	0.0492	0.051	0.05	0.05	0.044	0.05	0.0546

*Source: Simulation results*



c) *Results of Type 1 Error Rates on one sample test statistics at 0.01 level of significance*

From Table 4, it can be seen that the Type 1 error rates of all test statistics did well across all sample sizes but the performance of DST and AST are only good in some instances.

**Table 4:** Results of Type 1 Error Rates on one sample test statistics at 0.01 level of significance  
Sample size

Test statistics	10	15	20	25	30	35	40	50
Sd.t-test	0.0082	0.0104	0.009	0.01	0.01	0.009	0.01	0.011
z-test	0.0098	0.0116	0.01	0.01	0.01	0.007	0.01	0.0094
Rt-test	0.0116	0.0126	0.011	0.01	0.01	0.009	0.01	0.011
DWST	0.009	0.0076	0.008	0.01	0.01	0.008	0.01	0.0104
AWST	0.0058	0.0062	0.007	0.01	0.01	0.008	0.01	0.0104
DST	0.0012	0.0074	0.003	0	0.01	0.004	0.01	0.007
AST	0.0012	0.0074	0.009	0.01	0.01	0.004	0.01	0.007
Tt-test	0.0076	0.0102	0.008	0.01	0.01	0.009	0.01	0.01

Source: Simulation results

Note: Estimated Type 1 error rates in the preferred interval are in bold form.

d) *Results of overall number of times Type 1 Error rates approximate true levels of significance when counted over levels of significance*

Results of counting the number of times the Type 1 error rates of the inferential test statistics approximate the true level of significance when counted over the three (3) levels of significance is presented in Table 5 and graphically represented in Figure 1.

From Table 5 and Figure 1, it can be seen that the Type 1 error rate of student t-test and AWST, Rt-test and z-test are generally very good while that of the Sign test not good.

**Table 5:** Overall Total number of Times Type 1 Error Rates approximates True levels of significance when counted over levels of significance  
Sample size

Test statistics	10	15	20	25	30	35	40	50	Total	Rank
Sd.t-test	3	3	3	3	3	3	3	3	24	1.5
z-test	3	3	3	3	3	2	3	2	22	4
Rt-test	3	2	3	3	3	3	3	3	23	3
DWST	2	3	3	2	3	2	2	3	20	5.5
AWST	3	3	3	3	3	3	3	3	24	1.5
DST	0	1	0	0	1	0	1	1	4	8
AST	1	2	2	2	1	0	1	2	11	7
Tt-test	3	2	3	3	2	2	2	3	20	5.5

Source: Counted from Table 2, 3 and 4

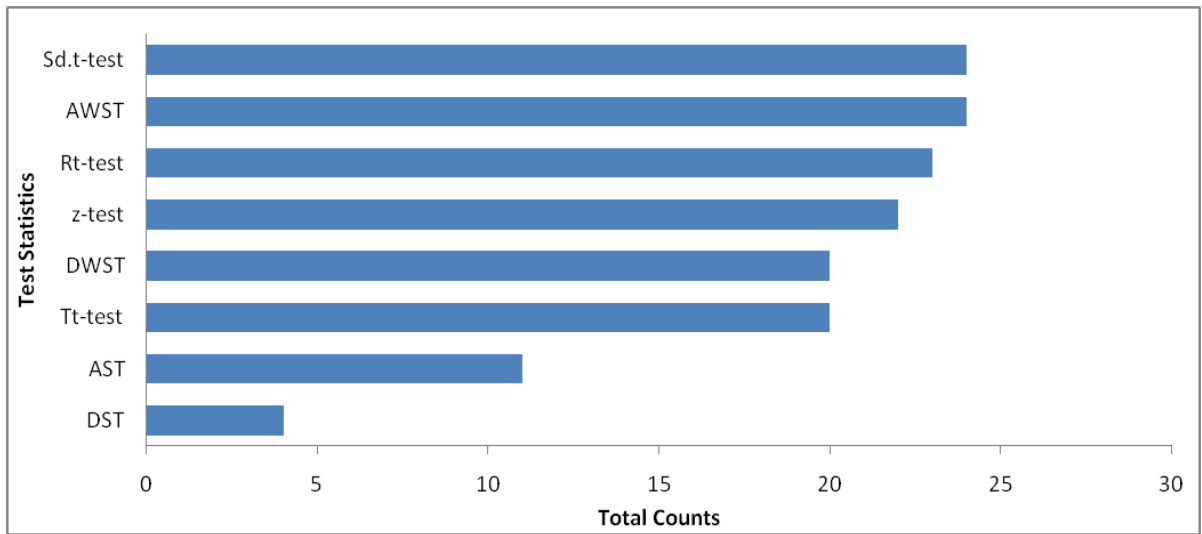


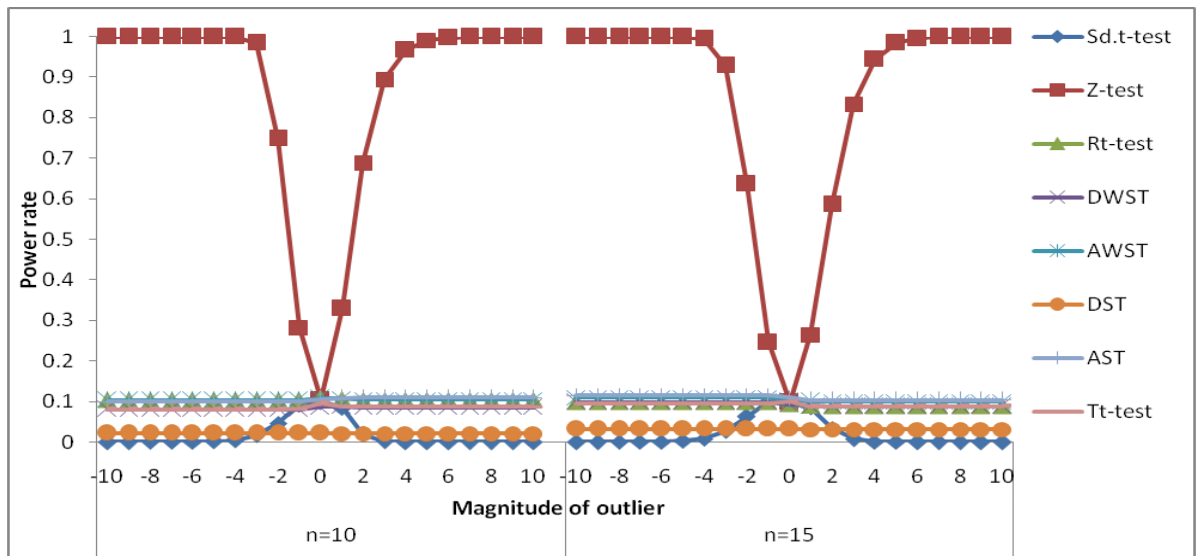
Figure 1: Bar chart showing overall total number of times Type 1 Error Rates approximate true levels of significance

e) Sensitivity and Robustness Investigation of One Sample test Statistics

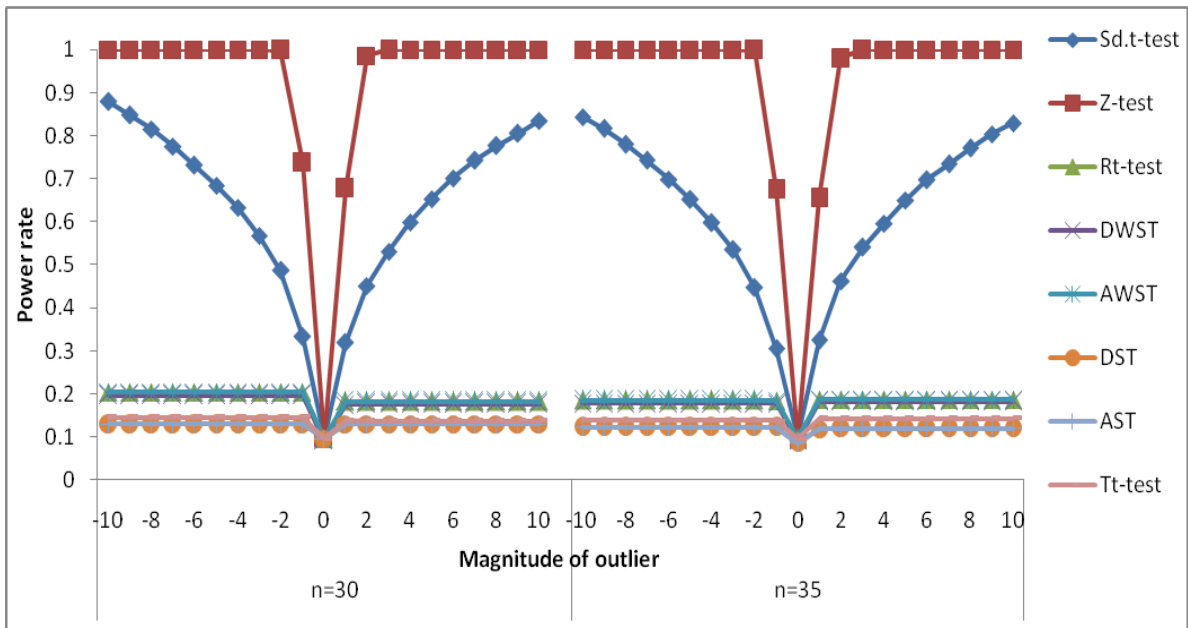
Sample graphs of Type 1 error rates of the inferential test statistics at different levels of significance, percentage of outliers and magnitude of outliers are presented and discussed. Adejumo (2016) provided the details of the graphs and the simulation results. However, Figure 2, Figure 3 and Figure 4 are sample graphs of the results for 0.1, 0.05 and 0.01 level of significance at 10% and 20% outliers respectively and discussed.

The test statistics are affected by the magnitude, percentage of outlier and levels of significance. From the graphical representations, it can be observed that as the sample size, magnitude and percentage of outliers are increasing the more the sensitivity of z-test and student t-test to outliers at all levels of significance.

Summarily, the overall results of the further counts over the levels of significance are presented in Table 6 and Figure 5. From Table 6 and Figure 5, it can be concluded that the z-test and student t- statistics are very sensitive and AST and Tt-test are robust to outliers.

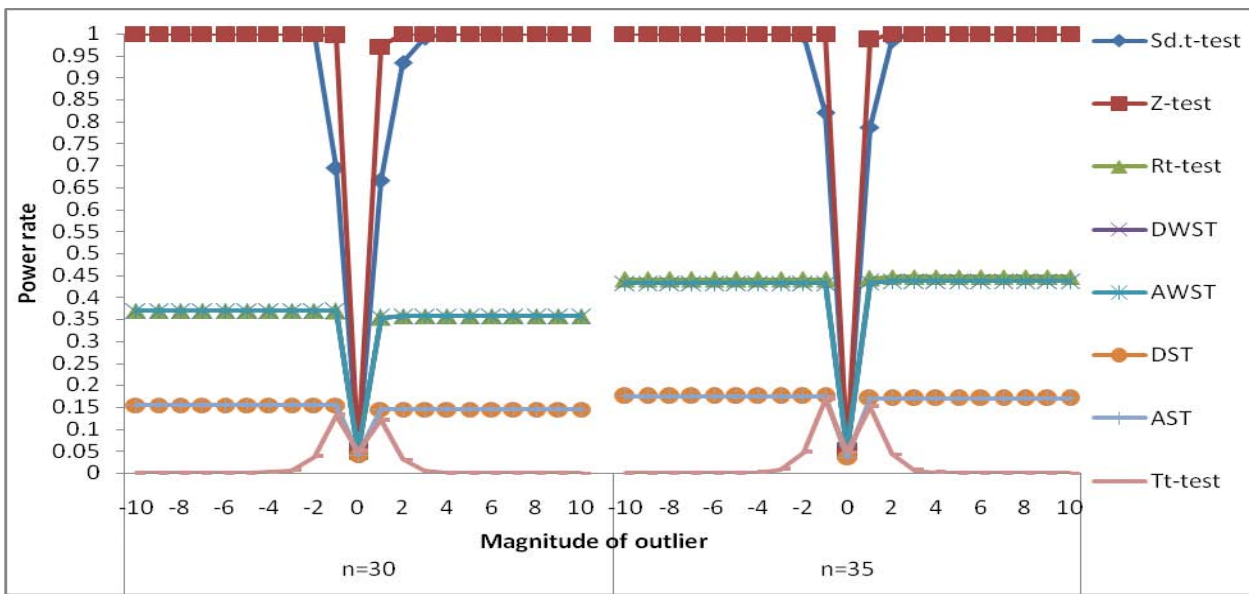


(a)

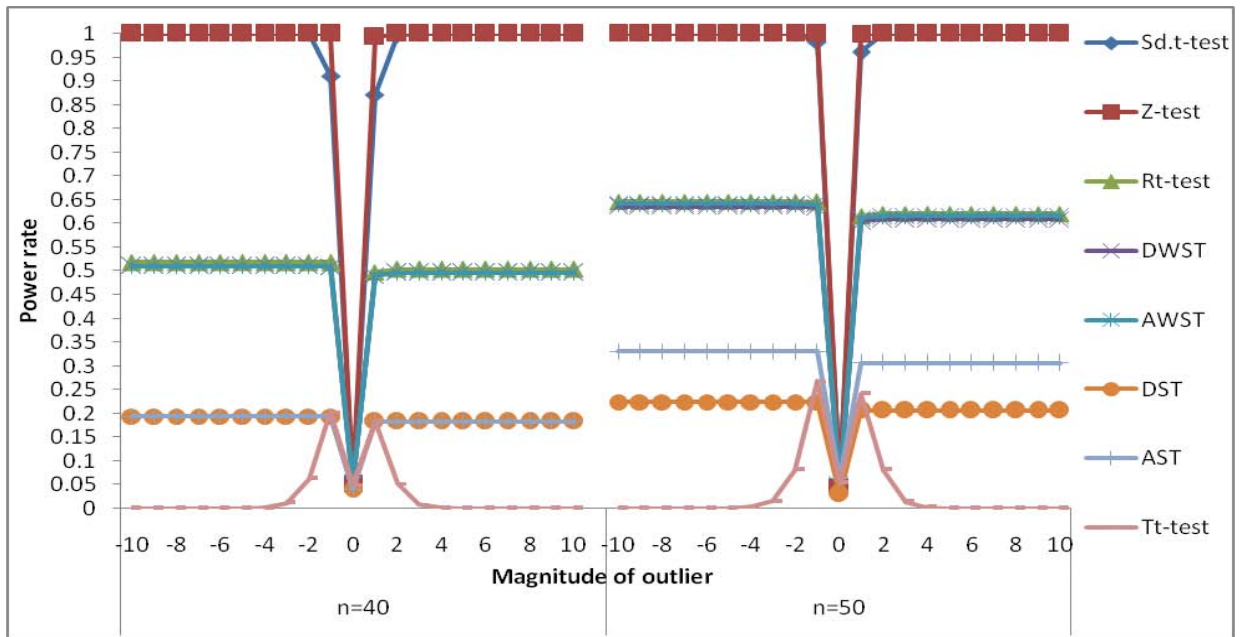


(b)

Figure 2: Graphical representation of power rate of one sample test statistics with 10% outlier at 0.1 level of significance

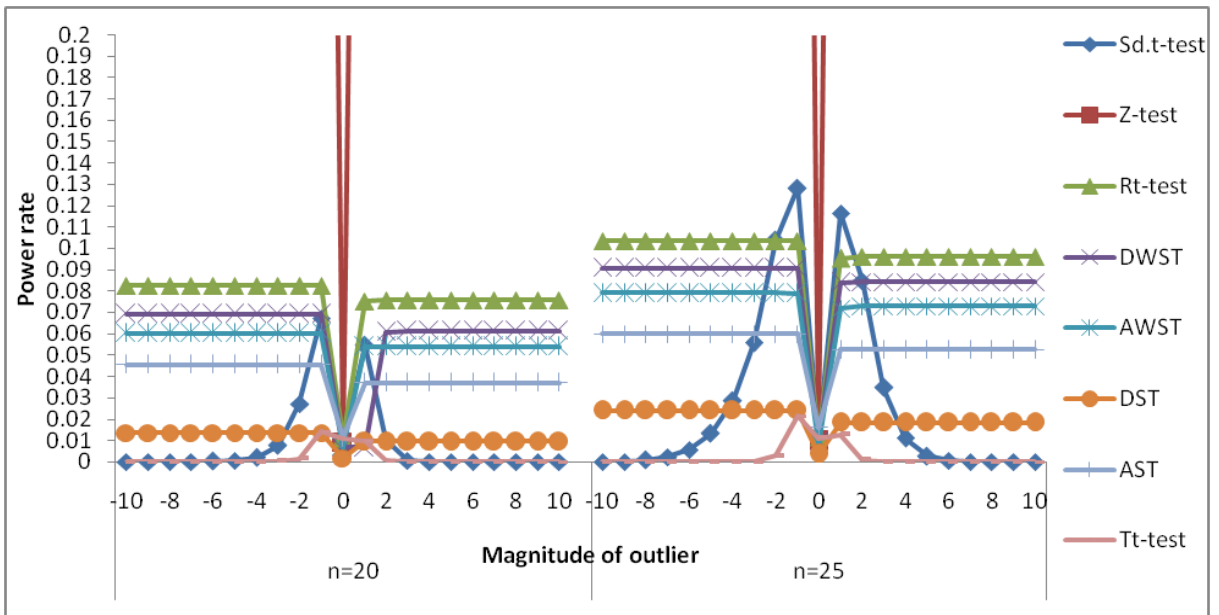


(a)

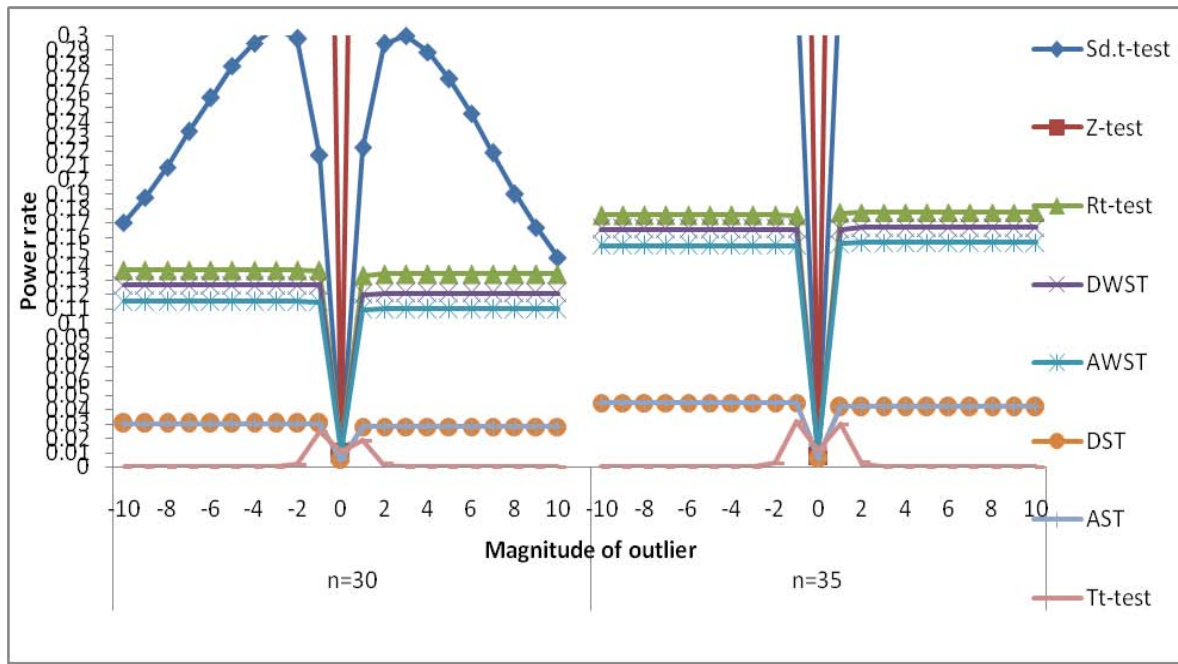


(b)

Figure 3: Graphical representation of power rate of one sample test statistics with 20% outlier at 0.05 level of significance



(a)



(b)

Figure 4: Graphical representation of power rate of one sample test statistics with 20% outlier at 0.01 level of significance

Table 6: Overall total number of times Power Rates approximates True levels of significance when counted over percentage of outliers and levels of significance for one sample problem

Test Statistics	Sample size									Total	Rank
	10	15	20	25	30	35	40	50			
Sd.t-test	10	9	12	11	8	8	7	8	73	7	
z-test	6	5	6	4	6	6	6	6	45	8	
Rt-test	66	41	3	14	6	6	6	6	148	5	
DWST	44	55	8	44	4	5	6	6	172	4	
AWST	86	66	15	33	6	5	6	6	223	3	
DST	0	32	19	0	20	22	12	2	107	6	
AST	22	54	45	34	41	42	32	4	274	1	
Tt-test	26	40	56	64	18	19	7	8	238	2	

Source: Counted from Simulation results

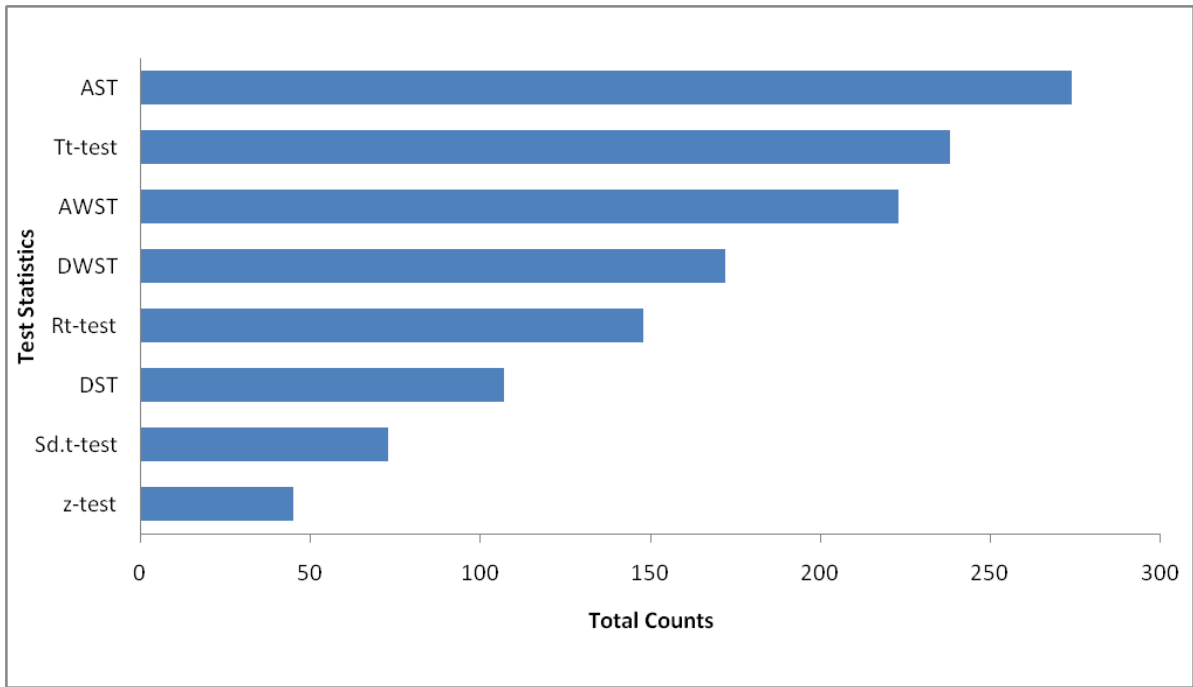


Figure 5: Bar chart showing overall total number of times Power Rates approximate true level of significance when counted over percentage of outlier and all levels of significance for one sample problem

V. SUMMARY AND CONCLUSION

Summary of findings and conclusions of the work is hereby presented in Table 6 as follows:

Table 6: Summary of findings on the One Sample Test Statistics

$\alpha$	One Sample Investigation		
	Type 1 Error	Robustness	Sensitivity
0.1	Sd.t-test, Rt-test, AWST	AST, DST	z-test, Sd.t-test
0.05	z-test, Sd.t-test, DWST, AWST	Tt-test, DWST	z-test, Sd.t-test
0.01	z-test, Sd.t-test, DWST, AWST, Rt-test, Tt-test	DST, AWST, Tt-test, AST	z-test, Sd.t-test
Overall	Sd.t-test, AWST	AST, Tt-test	z-test, Sd.t-test

From Table 6, the following can be observed:

The z-test and Student t- statistics are the most test statistics sensitive to outliers at all levels of significance.

At 0.1 level of significance, Student t-test, Rt-test and AWST have better Type 1 error while, the Sign tests, AST and DST in this order, are robust.

At 0.05 level of significance, the Type 1 error rate of z-test, student t-test DWST and AWST are good while, Tt-test and Wilcoxon Sign Rank test (DWST and AWST) in this order, are robust.



At 0.01 level of significance, the Type 1 error rate of all the test statistics are good except sign test meanwhile, DST, AWST, Tt-test and AST, in this order, are robust to outliers.

Summarily, over all levels of significance it can be concluded that, student t-test and AWST are the test statistics that have better Type 1 error rate, AST and Tt-test are the most robust test statistics to outliers whereas, z-test and Student t- statistics were identified to be the most test statistics sensitive to outliers.

#### REFERENCES RÉFÉRENCES REFERENCIAS

1. Agresti, N. and Finlay, B. (1997). *Statistical Method for the Social Sciences*, 3<sup>rd</sup> ed. Upper Saddle River, NJ., Prentice Hall.
2. Ayinde, K., Solomon G. S., Adejumo, A. O. (2016): A Comparative Study of the Type 1 Error Rates of Some Multivariate Tests of Normality. *Journal of the Nigerian Association of Mathematical Physics*. Vol. 34, pp 185 – 198
3. Barnett, V. and Lewis, T. (1994). *Outliers in Statistical Data* (3rd ed.), Chichester: John Wiley.
4. Cohen, J. (1969): A coefficient of agreement for nominal scales. *Education psychological measure*, 20, 37-46.
5. Conover, W.J. and Ronald, L.I. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statisticians*, 35, 125 – 129
6. Edgeworth, F. Y. (1887). On Discordant Observations. *Philosophical Magazine*, 23(5), 364-375.
7. Gauss, C. F. (1809). "Theoriamotus corporum coelestium in sectionibus conicis solem ambientium" Normal distribution-wikipedia, (2016); the free encyclopedia, <<http://www.encyclopedia.com>.
8. Gosset, W. S. (1908). The probable error of a mean. *Biometrika*-wikipedia (2016), the free encyclopedia, <<http://www.encyclopedia.com>.
9. Hawkins, D. (1980). Identification of outliers, Chapman and Hall, Reading, London. John Arbuthnot, (1710). Sign test-the free encyclopedia, <<http://www.encyclopedia.com>.
10. John A. (1710). Sign test-the free encyclopedia, <<http://www.encyclopedia.com> (2016).
11. Keselman, H. J., Wilcox F., Algina R.R. and Fradette, K. (2008): A comparative study of robust tests for spread: Asymmetric trimming strategies. *British Journal of Mathematical and Statistical Psychology*, 61, 235-253.
12. Kuranga, J. O. (2015): Type 1 Error Rate and Power comparison of some Normality test statistics. Unpublished M.Phil. Thesis, Department of Statistics, Ladoko Akintola University of Technonogy Ogbomoso, Nigeria.
13. Osborne, J. W. and Amy, O. (2004).The power of outliers (and why Researchers should always check for them). Practical Assessment, Research and Evaluation 9(6).North Carolina State University.
14. Plackett, R. L. and Barnard, G. A. (1990). "Student": A Statistical Biography of William Sealy Gosset. Based on the writings of E. S. Pearson. Oxford Clarendon Press.
15. Siegel Sidney, (1956). Non-parametric Statistics for the behavioral sciences.NewYork.McGraw-Hill. pp. 75-83.
16. Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin* 1 (6): 80-83.
17. Yuen, K. K. (1974). The two-sample trimmed *t* for unequal population variances. *Biometrika*, 61(1), 165-170.



# GLOBAL JOURNALS INC. (US) GUIDELINES HANDBOOK 2016

---

[WWW.GLOBALJOURNALS.ORG](http://WWW.GLOBALJOURNALS.ORG)