Ilorin Journal of Science
Volume 10, Number 2, 2023, pp. 87 – 102 (Printed in Nigeria)
ISSN: 2408 – 4840 © 2023 Faculty of Physical Sciences, University of Ilorin
https://doi.org/10.54908/iljs.2023.10.02.002

ILORIN
JOURNAL OF SCEINCE

ILJS-23-007

One–Sample Non-Parametric Location Test Statistic for Classified Normally Distributed Data

Ayinde, K.[1] and Owolabi, A. T.[2]

[1]Department of Mathematics and Statistics, Northwest Missouri State University, Maryville, Missouri, USA.
[2]Department of Statistics, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria.

Abstract

Hypothesis testing about the location parameter of normal distribution is often tested using parametric test statistics including t and Z statistics. In this research, the rank version of the one-sample parametric test statistics was obtained, and it resulted into the proposed test statistic for testing the hypothesis about the location parameter when normally distributed data are ranked or classified. The test statistic is the average rank of the first p observations closest to the hypothesized mean value. Monte Carlo experiments were conducted at eight (8) levels of sample sizes to ascertain the distribution of the proposed statistic, investigate its type 1 error and power rates, and examine its agreement with both existing parametric and non-parametric equivalent test statistics. The proposed test statistic is symmetric, and the values of p at which its type 1 error rate is not different from the pre-selected levels of significance were obtained, as well as their power rates and measures of agreement. The power and measures of agreement of the proposed statistic are better than that of the Sign test and compete favorably with that of the Wilcoxon Signed Rank test. A numerical example was used to illustrate the usage of the proposed statistic.

Keywords: Parametric, non-parametric, type 1 error, power rates, and agreement measures

1. Introduction

The methods of inferential statistics often require statistical hypothesis testing about the population parameters. There are several inference-making approaches, and these include the parametric and non-parametric test statistics. The term 'parametric' implies assumptions about parameters; hence in a parametric test, conditions about the population from which the sample is taken need to be specified, and information about the sample must be fully utilized. The robustness of the parametric tests is trusted by many researchers (Luepsen, 2017). A test is said to be robust when its significance level and power are insensitive to departures from the assumptions it is derived from (Ito, 1980). Some authors, including Faizi and Alvi (2023), Field (2009), Glass *et al.* (1972), Lindman (1974), Osborne (2008), and Wilcox (2005) did good reviews of the assumptions and robustness of parametric tests. To determine how sensitive and robust some inferential test statistics are to outliers, Ayinde *et al.* (2016) studied the student t-test, the z-test, and some other test statistics. Derrick *et al.* (2017) introduced two test statistics for partially overlapping samples with reference to the t-distribution.

Corresponding Author: Ayinde, K.
Email: ayindek@nwmissouri.edu

The non-parametric methods do not have the same restrictions as their parametric counterparts. It is a distribution-free test. They are always widely believed to protect the desired significance level of statistical tests, even under extreme violation of those assumptions (Zimmerman, 2004). Ayinde *et al.* (2016) examined the distributional and asymptotic distribution of the Sign and Wilcoxon Signed Rank test statistics as being affected by outliers. Derrick *et al*. (2019) proposed non-parametric statistics for the partially overlapping samples problem under normality and non-normality assumptions. Furthermore, the non-parametric tests including Mann-Kendall, Modified Mann-Kendall, and Kendall Rank Correlation were used to investigate the spatial and temporal patterns of trends and magnitude of rainfall (Malik and Kumar, 2020).

One of the challenges often faced in hypothesis testing is how to test a hypothesis when data are classified or ranked. Let us consider some examples. Suppose it is believed that the average grade of students who wrote a particular exam over the years is 55%, and after data classification, the score turned out to be C. If fifty (50) students after an exam were randomly selected and their grades were collected. Assuming the data are normally distributed, one may wish to test the claim that:

$$H_0: \mu = C \quad \text{vs.} \quad H_1: \mu \neq C \tag{1}$$

Also, suppose the blood pressure of patients with a particular disease is believed to be normally distributed and is categorized into hypo, normal, and hyper groups. One can also be interested in knowing whether the claim is sustained in a community where the disease is rampant, having examined twenty-five (25) randomly selected individuals with the disease. Assuming the data are normally distributed, the hypothesis of interest is stated as:

$$H_0: \mu = Normal \quad \text{vs.} \quad H_1: \mu \neq Normal \tag{2}$$

Thus, in this research, an attempt is made to provide test statistic that can be used to investigate this kind of hypotheses.

Various test statistics have been developed to test a hypothesis about the location parameter in one sample problem. These have been grouped and discussed under parametric, non-parametric, and rank transformation-based statistics as follows:

Parametric Test Statistics

One sample Z-test is one of the parametric tests for handling one sample problem. The sample forms a single treatment group, and the population variance is assumed to be known. Data points should be independent and have an equal chance of being selected. The test statistic, distributed normally with mean zero and variance one, is given as follows:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim \mathrm{N}\,(0,1) \tag{3}$$

where $\bar{X}$, μ_0, σ and n are sample mean, hypothesized mean, population standard deviation, and sample size respectively.

Z^2-test statistic is another parametric test that can handle one sample problem (Wallis, 2013). It is a chi-square distribution with one degree of freedom. The test statistic Z^2 is given as:

$$Z^2 = \left[\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right]^2 \sim \chi^2_{(1)} \tag{4}$$

One sample t-test is another parametric test for a one-sample problem. In early 1908, an English man named William Sealy Gosset discovered what is now called student t-distribution (Ayinde *et al.*, 2009; Fisher,1987). One sample t-test has all the assumptions of the Z-test except that of the small sample size. It follows a t distribution with (n-1) degree of freedom. Its test statistic is given as:

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t_{n-1} \tag{5}$$

where $\bar{X}$, μ_0, s and n are sample mean, hypothesized mean, sample standard deviation and size respectively.

Another test statistic for one sample problem is the chi-square test. It is used to test if the variance of a normally distributed population has a given value based on a sample variance. The properties of chi-square statistic were first investigated in 1900 by Karl Pearson (Pearson, 1900). Symbolically, it is written as:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)} \tag{6}$$

where s^2, σ^2 and n are sample variance, population variance, and sample size respectively.

Non-parametric Test Statistics

The sign test is a non-parametric equivalent to a one-sample t-test. John Arbuthnot was the first to use the sign test in 1710 (Conover, 1999). The test statistic is T^+ or T^- as the case may be. If values with negative signs are the least, then our test statistic is T^-, otherwise, the test statistic will be T^+. Asymptotically, the sign test is distributed binomial (n,$\frac{1}{2}$).

$$S = \frac{T - \frac{n}{2}}{\sqrt{n/4}} \sim N(0,1) \tag{7}$$

where T is the least of T^+ and T^- and n is the sample size.

The Wilcoxon Signed Rank Test is a non-parametric statistical hypothesis test used for one-sample and matched samples as an alternative to paired t-test or the t-test for dependent samples when the population cannot be assumed to be normally distributed. The test is named after Frank Wilcoxon, who, in a single paper, proposed both the Wilcoxon Signed Rank Test and the Wilcoxon Rank Sum Test for two independent samples (Adejumo *et al.*, 2020; Conover and Iman, 1981; Salzburg, 2001; Wilcoxon, 1945). The test was popularized by Siegel (1956). The asymptotic distribution of the Wilcoxon sign rank test is:

$$W = \frac{T - E_0(T)}{\sqrt{V_0(T)}} = \frac{T - \frac{n(n+1)}{4}}{\left[\frac{n(n+1)(2n+2)}{24}\right]^{\frac{1}{2}}} \sim N(0,1) \quad (8)$$

where T is the least of T^+ and T^- and n is the sample size.

2. Materials and Methods

2.1 The proposed test statistic

The summary of some parametric statistics and their rank equivalence is provided in Table 1.

Table 1: Table of some parametric summary statistics and their equivalence in rank forms

Summary Statistic	Parametric	Rank Equivalence
Population Sum	$\sum_{i=1}^{N} X_i$	$\frac{N(N+1)}{2}$
Sample Sum	$\sum_{i=1}^{n} X_i$	$\frac{n(n+1)}{2}$
Population Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i$	$\frac{(N+1)}{2}$
Sample Mean	$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$\frac{(n+1)}{2}$
Population Variance	$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N}$	$\frac{(N^2-1)}{12}$
Sample Variance	$s^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}$	$\frac{n(n+1)}{12}$
Standard error of the population mean	$\frac{\sigma}{\sqrt{n}}$	$\sqrt{\frac{N^2-1}{12n}}$
Standard error of the sample mean	$\frac{S}{\sqrt{n}}$	$\sqrt{\frac{n+1}{12}}$

Using the information in Table 1, the parametric statistics to test the hypothesis $H_0 : \mu = \mu_0$ is given, and its rank version containing the proposed statistics is obtained as follows:

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{\frac{n+1}{2} - \mu^{rp}}{\sqrt{\frac{N^2-1}{12n}}} \quad (9)$$

where μ^{r_p} is the proposed statistic, the average rank of the *p*-observations closest to the hypothesized mean value μ_0.
Similarly,

$$Z^2 = \left[\frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right]^2 = \frac{\left[\frac{n+1}{2}-\mu^{rp}\right]}{\frac{N^2-1}{12n}}, \text{ N=known} \tag{10}$$

where μ^{r_p} is as defined earlier.

Also,

$$t = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{\frac{n+1}{2}-\mu^{rp}}{\sqrt{\frac{n+1}{12}}} = \sqrt{3(n+1)} - 2\mu^{rp}\sqrt{\frac{3}{n+1}} \tag{11}$$

where μ^{r_p} is as defined earlier.

Furthermore, using the information in Table 1, the parametric statistic to test the hypothesis $H_0 : \sigma^2 = \sigma_0^2$ is given, and its rank version containing the proposed statistic is obtained as:

$$X^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{1}{2}\left[\frac{n(n^2-1)}{(N+1)(2N+1)-3\left(\mu^{r_p}\right)^2}\right], N = known \tag{12}$$

Notable Observations

There are three notable observations on the proposed statistic, μ^{r_p}:

i As p tends to n-1, the parametric test statistic in its rank form also tends to zero, and when p= n, the parametric test statistic in its rank form eventually becomes zero.

ii The distribution of the test statistics in their rank version now depends on the distribution of the proposed statistic. For instance, if n = 10 and N = 40, the results in Table 2 are produced each time μ^{r_p} = 5

Table 2: Results of the parametric statistics in their rank version when n = 10, N = 40, and μ^{r_p} = 5

Statistics	Z	Z^2	t	X^2
Value	0.0375	0.0188	0.5222	0.1525

Consequently, whether N is known or unknown, the distribution of the proposed statistic distribution is unaffected. Thus, all the parametric statistics in their rank form can now be seen in the light of the proposed statistic, μ^{r_p}. And so, the proposed test statistic can be referred to as a test statistic.

iii The hypothesis $H_0: \sigma^2 = \sigma_0^2$ is now the same or equivalent to the hypothesis.

$H_0: \mu = \mu_0$. This is an observed weakness of the proposed statistic, as it could not distinguish between a test of location and dispersion.

2.2 Mathematics of the proposed test statistic

Just like some other non-parametric test statistics where there are steps to follow before obtaining their test statistic, μ^{r_p} (which is the average rank of the first p observations closest to the hypothesized mean value, where p= 1, 2, …, n-1) also has some steps. The following steps are to be taken to obtain the proposed test statistic, μ^{r_p}:

i. Arrange the sample observations $X_1, X_2, \ldots, X_n$ in order of magnitude, say, $X^{(1)}, X^{(2)}, \ldots, X^{(n)}$.

ii. Assign the ordered sample observations their respective rank value, say, R_{X_i}= 1, 2,…, n.

iii. Find the average rank of the 1st p sample observations closest to the hypothesized mean value μ_0 and call it μ^{r_p}. This can be achieved by following these steps:

a. Obtain $D_i = |X_i - \mu_0|$.

b. Obtain the rank of D_i and call it R_{D_i}.

c. Define $W_i = \begin{cases} 1 \; if \; R_{D_i} \le p \\ 0 \; otherwise \end{cases}$

d. Obtain $$\mu^{r_p} = \frac{\sum_{i=1}^{n} R_{x_i} W_i}{\sum_{i=1}^{n} W_i} = \frac{\sum_{i=1}^{n} R_{x_i} W_i}{p} \quad (13)$$

2.3 Theoretical Probability Distribution of the proposed test statistic

The proposed test statistic, μ^{r_p}, which is the average rank of the first p observations closest to the hypothesized mean value, can be viewed in the light of the sampling distribution of N observations taken n, n $\le N$, at a time. In this case, there are n values (ranks), out of which p of them are to be taken at a time. Hence,

$$E(\mu^{r_p}) = \mu = \frac{\sum X_i}{n} = \frac{n+1}{2} \quad (14)$$

$$V(\mu^{r_p}) = \sigma^2_{\mu^{r_p}} = \frac{\sigma^2}{p}\frac{n-p}{n-1} = \frac{(n+1)(n-p)}{12p} \quad (15)$$

NOTE: When p=1, the distribution is uniform with mean = $\frac{n+1}{2}$ and variance = $\frac{n^2-1}{12}$.

A pictorial representation of the proposed statistic when n = 15 and p = 5 is given in Figure 1.

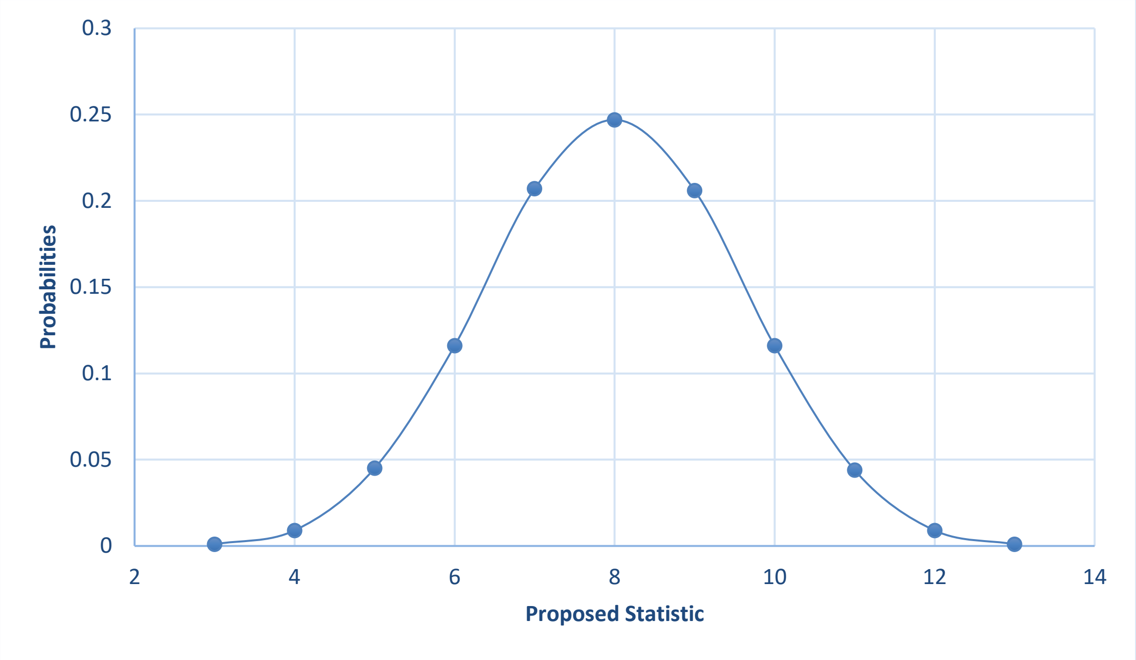

Figure 1: Graph of the theoretical probability distribution of the proposed statistic when n =15, p = 5

Figure 1 shows that the theoretical distribution of the proposed test statistic is symmetric and appears like a normal distribution. The descriptive statistics of the theoretical distribution of the proposed test statistic further reveals the minimum and maximum as:

$$\text{Min } (\mu^{r_p}) = \frac{p+1}{2} \tag{16}$$

$$\text{Max } (\mu^{r_p}) = \left[n - \frac{(p-1)}{2} \right] \tag{17}$$

Asymptotically,

$$Z = \frac{\mu^{r_p} - \frac{n+1}{2}}{\sqrt{\frac{(n+1)(n-p)}{12p}}} \sim N\,(0,1) \tag{18}$$

2.4 Monte Carlo Simulation Study

A Monte Carlo simulation study was carried out on the proposed test statistic to ascertain its sampling distribution. Random variable X was generated to be normally distributed, $X_i \sim N\ (10,2.5)$, at 8 levels of sample sizes (n = 10, 15, 20, 25, 30, 35, 40, and 50). The experiment was replicated 50,000 times, and a comparison of both the simulation and theoretical results was made. The type 1 error and power rates (with hypothesized values as 12, 14, 16, and 18) of the proposed test statistic and their equivalent parametric and non-parametric were examined and compared at three (3) levels of significance (0.1, 0.05 and 0.01). The value of p at the type I error rate of the proposed test statistic is closer to the pre-selected significance level is preferred, and the power rates at the preferred p were investigated. Furthermore, efforts were made to see how the proposed non-parametric and the existing equivalent ones (the Sign and Wilcoxon Signed Rank Tests) perform in terms of acceptance, rejection, and agreement with the equivalent parametric test statistics (the Z and t-test statistics). Sensitivity measures the agreement in terms of acceptance, while specificity measures the agreement in terms of rejection. The overall agreement between the non-parametric (proposed and existing) and parametric test statistics was measured using both Kappa and Tau Statistics (Cohen, 1960; Jolayemi,1990; Lawal, 2003; Tanner and Young, 1985;). The τ measure of agreement by Jolayemi (1990) has been demonstrated to be better than the Kappa measure of agreement for not too large sample sizes (Lawal, 2003).

3. Results and Discussion

The results from the research are now presented as follows. The theoretical and simulation results are compared in section 3.1, the results based on the type 1 error and power rates investigation are presented in sections 3.2 and 3.3 respectively, and the results on the agreement study between the non-parametric and parametric test statistics are presented in section 3.4.

3.1 Comparison of theoretical and simulation results

The results based on the theoretical distribution and simulation study of the proposed test statistic for some values of p (p=2,3,4, and 5) at various sample sizes are presented in Table 3. From the table, it was observed that the theoretical mean, minimum, and maximum agree with that of the simulation study, while the standard deviation, skewness, and kurtosis deviate a little, especially when the value of p is small. However, as p increases, the differences and coefficient of variation (C.V) reduce. Thus, the claims about the distribution of the proposed test statistic in (14), (15), (16) and (17) are ascertained.

3.2 Type 1 Error Investigation

The type 1 error rate of both the test statistics are provided in Table 4. A sample of that of the proposed is given in Table 4b, and the entire three levels of significance are graphically presented in Figures 2a, b, and c. The summary of the preferred ps is provided in Table 4c. Table 4a shows that the type 1 error of the parametric test statistics are good, and they conform to their pre-selected values of significance. Moreover, that of Wilcoxon Signed rank test is also good, except that it does not conform well at 0.01 level of significance. Furthermore, that of the Sign test only performs well at 0.1 level of significance.

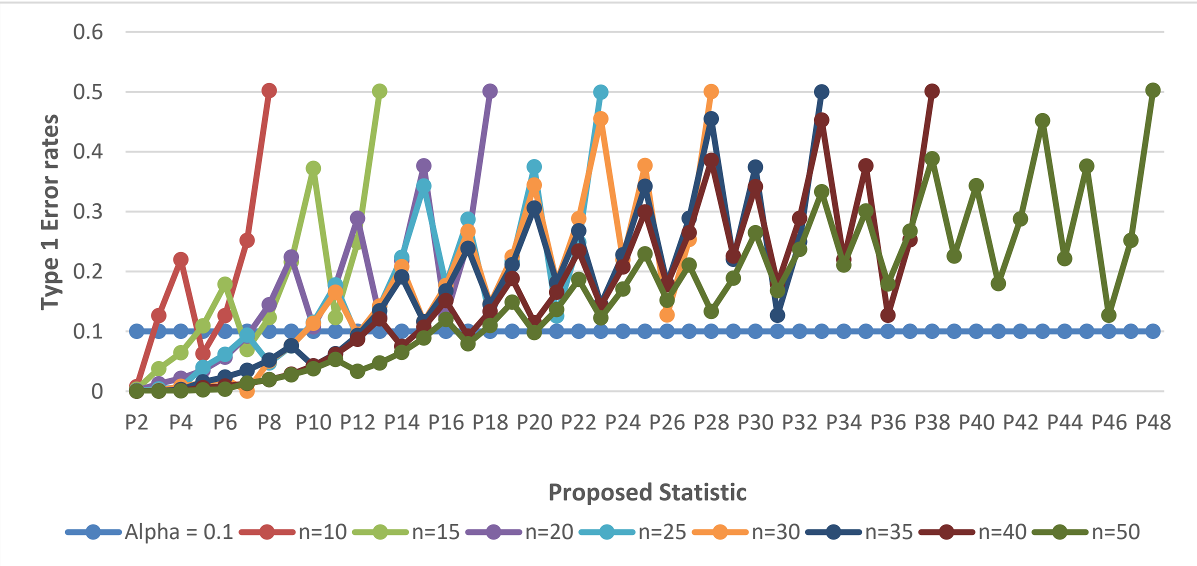

Figure 2a: Type 1 Error rates of the proposed non-parametric statistic at $\boldsymbol{\alpha = 0.1}$

At other tables and figures, it can be summarily observed that proposed statistic is only good at a 0.1 level of significance and at this instance, the preferred ps also have a pattern. In all the sample sizes, the last preferred p can be obtained with the formula n – 4 and so, none of the preferred ps exceed n – 4. Furthermore, the first preferred p is generally and approximately obtained with $\frac{n}{3}$. When n is odd, some of these can be obtained by $\frac{n-1}{2} - 2, \frac{n-1}{2} + 1$ and $\frac{n-1}{2} + 6$ and when n is even by $\frac{n}{2} - 2, \frac{n}{2} + 3$ and $\frac{n}{2} - 5$. At other levels of significance considered, no pattern is established but in most cases at least one of the preferred ps at 0.1 is among their preferred ps.

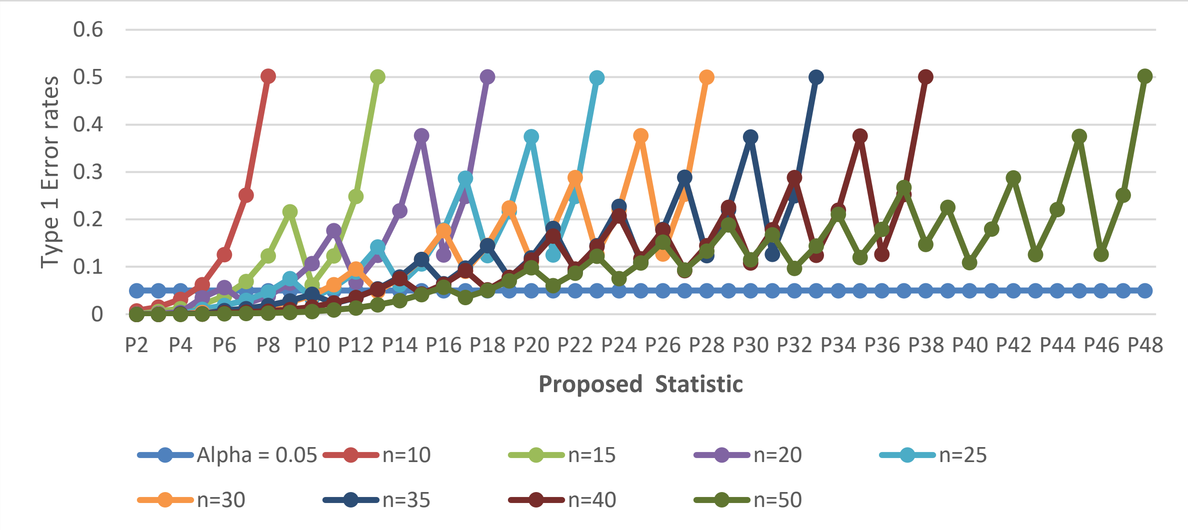

Figure 2b: Type 1 error of the proposed non-parametric statistic at $\alpha = 0.05$

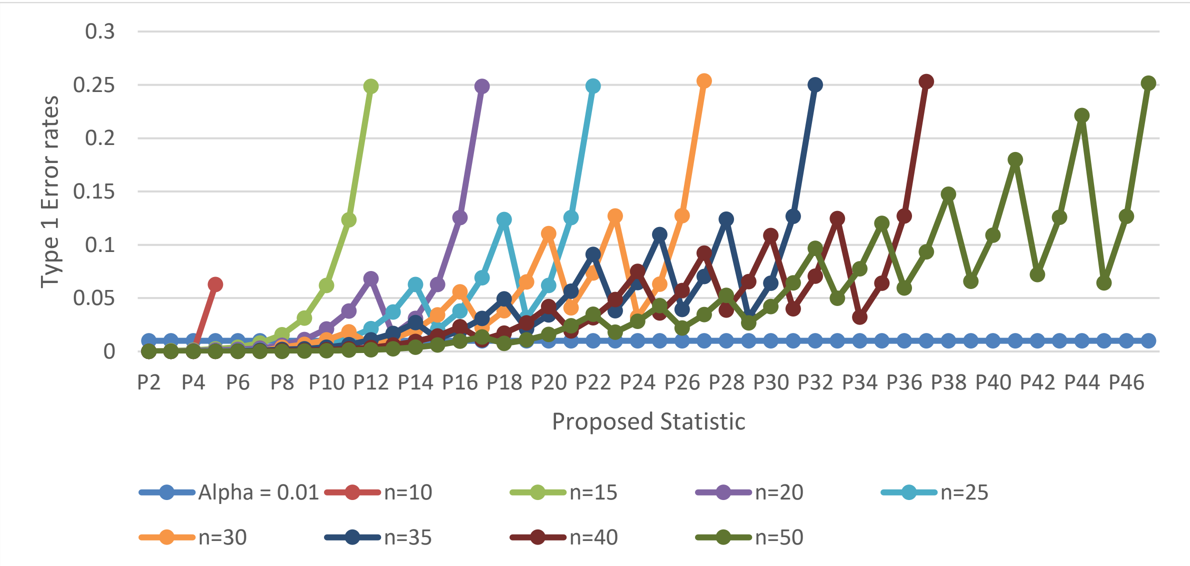

Figure 2c: Type 1 Error rates of the proposed non- parametric statistic at $\alpha = 0.01$

3.3 Power Rate Investigation

Power rate on the preferred ps at 0.1 level of significance was investigated at all the levels of sample sizes considered. A sample of the results with the power rates of other test statistics when n=15 and 20 is given in Table 5a and a pictorial representation of the same power results when n=10 is displayed in Figure 3. The summary of the preferred ps with the best power is provided in Table 5b.

Generally, from the tables and figure, in terms of power assessment, the proposed test statistic performs better than the Sign test in all cases. It competes favorably with the Wilcoxon Signed Rank Test in a few instances and on a rare occasion with the *t*-test. *Z*-test out-performs the proposed test statistic in all cases.

Table 3. Comparison of the Descriptive Statistics for Theoretical and Simulation Results

n	p		Mean	Standard Deviation	Skewness	Kurtosis	Minimum	Maximum	C.V
	2	T	5.5000	1.93649	0.000	-0.572	1.50	9.50	35.21
		S	5.5038	1.41406	0.006	-0.260	1.50	9.50	25.69
10	3	T	5.5000	1.46863	0.000	-0.430	2.00	9.00	26.70
		S	5.5037	1.32480	0.008	-0.292	2.00	9.00	24.07
	4	T	5.5000	1.17541	0.000	-0.371	2.50	8.50	21.37
		S	5.5058	1.22665	0.001	-0.345	2.50	8.50	22.28
	5	T	5.5000	0.95933	0.000	-0.354	3.00	8.00	17.44
		S	5.5038	1.12045	0.004	-0.411	3.00	8.00	20.36
	2	T	8.0000	2.95804	0.000	-0.588	1.50	14.50	36.98
		S	7.9939	1.79471	-0.003	-0.149	1.50	14.50	22.45
15	3	T	8.0000	2.31194	0.000	-0.416	2.00	14.00	28.90
		S	7.9950	1.72334	-0.008	-0.163	2.00	14.00	21.56
	4	T	8.0000	1.91556	0.000	-0.331	2.50	13.50	23.94
		S	7.9930	1.65236	-0.006	-0.164	2.50	13.50	20.67
	5	T	8.0000	1.63327	0.000	-0.283	3.00	13.00	20.42
		S	7.9949	1.57699	-0.003	-0.180	3.00	13.00	19.72
	2	T	10.5000	3.97911	0.000	-0.594	1.50	19.50	37.90
		S	10.4972	2.11868	0.008	-0.115	2.50	18.50	20.18
20	3	T	10.5000	3.15046	0.000	-0.410	2.00	19.00	30.00
		S	10.4965	2.05752	0.008	-0.123	3.00	19.00	19.60
	4	T	10.5000	2.64602	0.000	-0.317	2.50	18.50	25.20
		S	10.4995	2.00006	0.002	-0.120	3.50	18.50	19.05
	5	T	10.5000	2.29136	0.000	-0.263	3.00	18.00	21.82
		S	10.4983	1.93450	0.006	-0.127	3.00	18.00	18.43
	2	T	13.0000	5.00000	0.000	-0.596	1.50	24.50	38.46
		S	13.0012	2.38573	0.005	-0.109	3.50	21.50	18.35
25	3	T	13.0000	3.98695	0.000	-0.407	2.00	24.00	30.67
		S	12.9953	2.33119	-0.003	-0.102	4.00	22.00	17.94
	4	T	13.0000	3.37282	0.000	-0.311	2.50	23.50	25.94
		S	12.9954	2.27886	-0.003	-0.111	3.50	21.50	17.54
	5	T	13.0000	2.94395	0.000	-0.254	3.00	23.00	22.65
		S	12.9968	2.22446	-0.008	-0.115	4.00	21.00	17.12
	2	T	15.5000	6.02080	0.000	-0.597	1.50	29.50	38.84
		S	15.5101	2.63394	-0.008	-0.111	3.50	25.50	16.98
30	3	T	15.5000	4.82242	0.000	-0.405	2.00	29.00	31.11
		S	15.5118	2.58780	-0.005	-0.095	4.00	26.00	16.68
	4	T	15.5000	4.09784	0.000	-0.307	2.50	28.50	26.44
		S	15.5117	2.53847	-0.002	-0.105	4.50	25.50	16.36
	5	T	15.5000	3.59399	0.000	-0.249	3.00	28.00	23.19
		S	15.5095	2.48501	-0.004	-0.119	5.00	25.00	16.02
	2	T	18.0000	7.04154	0.000	-0.598	1.50	34.50	39.12
		S	18.0064	2.86579	0.003	-0.073	6.50	29.50	15.92
35	3	T	18.0000	5.65729	0.000	-0.404	2.00	34.00	31.43
		S	18.0057	2.82115	-0.002	-0.071	6.00	29.00	15.67
	4	T	18.0000	4.82187	0.000	--0.305	2.50	33.50	26.79
		S	18.0064	2.77793	0.000	-0.071	6.50	29.50	15.43
	5	T	18.0000	4.24265	0.000	-0.247	3.00	33.00	23.57
		S	18.0056	2.73183	0.004	-0.076	7.00	29.00	15.17
	2	T	20.5000	8.06226	0.000	-0.598	1.50	39.50	39.33
		S	20.5031	3.07546	0.023	-0.037	6.50	32.50	15.00
40	3	T	20.5000	6.49178	0.000	-0.403	2.00	39.00	31.67
		S	20.5045	3.03795	0.021	-0.031	6.00	33.00	14.82
	4	T	20.5000	5.54530	0.000	-0.304	2.50	38.50	27.05
		S	20.5024	2.99906	0.022	-0.033	5.50	33.50	14.63
	5	T	20.5000	4.89047	0.000	-0.245	3.00	38.00	23.86
		S	20.4977	2.95414	0.022	-0.032	5.00	33.00	14.41

Source: Computer output

Table 4a: Type 1 error for parametric and existing non – parametric test statistics

α	Stat	Sample size							
		10	15	20	25	30	35	40	50
0.1	Z	0.0997	0.0970	0.0989	0.0978	0.0995	0.0991	0.0995	0.1000
	t	0.1004	0.0990	0.0982	0.0985	0.0997	0.1004	0.1006	0.0999
	Z^2	0.1006	0.0998	0.1006	0.1001	0.0983	0.1002	0.1025	0.1020
	S	0.1095	0.1158	0.1154	0.1069	0.0977	0.0885	0.0790	0.1204
	W	0.1056	0.1051	0.0958	0.1001	0.0999	0.1000	0.1002	0.1008
0.05	Z	0.0484	0.0484	0.0484	0.0486	0.0496	0.0491	0.0511	0.0501
	t	0.0506	0.0489	0.0493	0.0493	0.0493	0.0483	0.0514	0.0502
	Z^2	0.0506	0.0508	0.0503	0.0491	0.0495	0.0498	0.0508	0.0509
	S	0.0210	0.0352	0.0415	0.0425	0.0420	0.0403	0.0375	0.0658
	W	0.0493	0.0466	0.0484	0.0468	0.0485	0.0483	0.0501	0.0497
0.01	Z	0.0096	0.0095	0.0096	0.0087	0.0092	0.0101	0.0100	0.0099
	t	0.0096	0.0099	0.0096	0.0093	0.0089	0.0101	0.0096	0.0101
	Z^2	0.0102	0.0104	0.0099	0.0095	0.0103	0.0097	0.0101	0.0100
	S	0.0018	0.0075	0.0117	0.0134	0.0043	0.0055	0.0062	0.0061
	W	0.0056	0.0065	0.0076	0.0073	0.0075	0.0084	0.0090	0.0091

Source: Computer output

Table 4b: Type 1 error of the proposed test statistic when n=20

ps	Distribution Investigation			Asymptotic investigation		
	$\propto = 0.1$	$\propto = 0.05$	$\propto = 0.01$	$\propto = 0.1$	$\propto = 0.05$	$\propto = 0.01$
p_1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
p_2	0.00116	0.00012	0.00000	0.00116	0.00012	0.00000
p_3	0.01224	0.00212	0.00002	0.01224	0.00212	0.00002
p_4	0.02138	0.00398	0.00058	0.02138	0.00398	0.00058
p_5	0.03472	0.00716	0.00112	0.03472	0.03472	0.00112
p_6	0.05672	0.01246	0.00176	0.05672	0.05672	0.00176
p_7	**0.09188**	0.02220	0.00288	**0.09188**	0.02220	0.00288
p_8	**0.14460**	0.03834	**0.00572**	**0.14460**	0.03834	**0.00572**
p_9	0.22382	0.06564	**0.01104**	0.22382	0.06564	**0.01104**
p_{10}	**0.10728**	0.10728	0.02112	**0.10728**	0.10728	0.02112
p_{11}				0.17636	0.17636	0.03770
p_{12}				0.28852	0.06818	0.06818
p_{13}				**0.12436**	0.12436	**0.01568**
p_{14}				0.21808	0.21808	0.03094
p_{15}				0.37652	0.37652	0.06284
p_{16}				**0.12530**	0.12530	0.12530
p_{17}				0.24852	0.24852	0.24852
p_{18}				0.50092	0.50092	0.00000

Note: Preferred ps are bolded

3.4 Agreement measures of non-parametric with parametric test statistics

The agreement of the parametric test statistics with the proposed test statistic having the best power rate and those of the other non-parametric statistics was examined, and the results are presented in Table 6. The higher the value of the agreement, the better the test statistic. From Table 6, it could be summarily observed that the proposed test statistic in terms of specificity (rejection), Kappa, and Tau agreements still generally performs better than the Signed test and competes favorably with the Wilcoxon Signed Rank test statistic.

Table 4c: Summary Table of Preferred ps of type 1 error of the proposed test statistic

n	$\propto = 0.1$	$\propto = 0.05$	$\propto = 0.01$
10	p_3, p_6	-	-
15	p_5, p_8, p_{11}	-	p_7
20	p_7, p_8, p_{10}, p_{13}, p_{16}	-	p_8, p_9, p_{13}
25	p_7, p_{10}, p_{13}, p_{18}, p_{21}	p_8	p_{10}, p_{11}
30	p_{10}, p_{12}, p_{13}, p_{15}, p_{18}, p_{26}	p_{13}	p_9, p_{10}, p_{13}
35	p_{12}, p_{13}, p_{15}, p_{18}, p_{23}, p_{31}	p_{13}	p_{11}, p_{12}, p_{15}
40	p_{13}, p_{15}, p_{17}, p_{18}, p_{20}, p_{23}, p_{36}	p_{13}, p_{18}	p_{13}, p_{14}, p_{15}, p_{17}
50	p_{16}, p_{18}, p_{19}, p_{20}, p_{21}, p_{23}, p_{28}, p_{46}	p_{18}	p_{15}, p_{16}, p_{17}, p_{18}, p_{19}

Table 5a: Power rates of the test statistics when n=15 and n=20 at $\boldsymbol{\alpha} = 0.1$

		Hypothesized values			
n		12	14	16	18
15	Z	0.928	1.000	1.000	1.000
	t	0.905	1.000	1.000	1.000
	S	0.806	0.999	1.000	1.000
	W	0.899	1.000	1.000	1.000
	P_5	0.864	1.000	1.000	1.000
	$\mathbf{P_8}$	0.876	1.000	1.000	1.000
	P_{11}	0.806	0.998	1.000	1.000
20	Z	0.974	1.000	1.000	1.000
	t	0.965	1.000	1.000	1.000
	S	0.890	1.000	1.000	1.000
	W	0.957	1.000	1.000	1.000
	P_7	0.927	1.000	1.000	1.000
	$\mathbf{P_8}$	0.957	1.000	1.000	1.000
	P_{10}	0.936	1.000	1.000	1.000
	P_{13}	0.925	1.000	1.000	1.000
	P_{16}	0.850	0.999	1.000	1.000

Table 5b: The preferred ps with the best power at various sample sizes

n	10	15	20	25	30	35	40	50
p	p_3	p_8	p_8	p_{13}	p_{13}	p_{18}	p_{18}	p_{19} and p_{21}

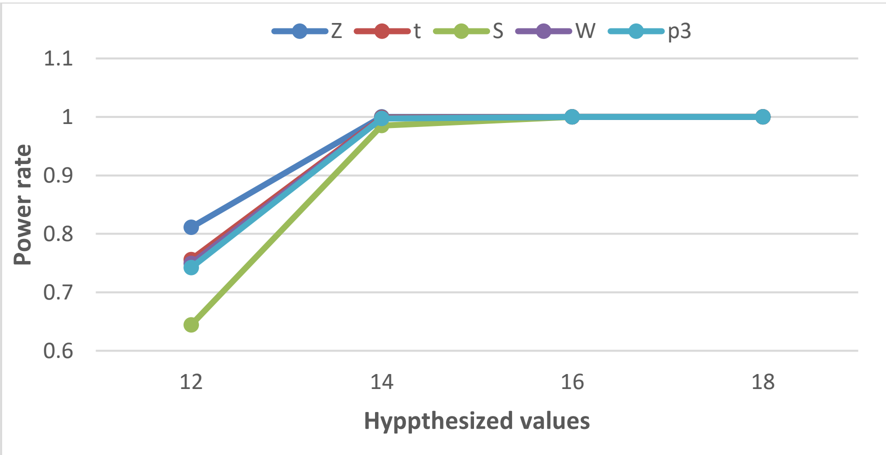

Figure 3: Pictorial Representation of power rates of the test statistics when n=15 and n=20 at $\boldsymbol{\alpha} = 0.1$

Table 6: Agreement measures of the non-parametric with parametric test statistics

		Sensitivity			**Specificity**			**Kappa**			**Tau**		
n		Z	t	Z^2	Z	t	Z^2	Z	t	Z^2	Z	T	Z^2
10	S	0.934	0.945	0.896	0.502	0.602	0.161	0.418	0.527	0.056	0.419	0.527	0.056
	W	0.956	0.982	0.902	0.662	0.887	0.176	0.603	0.849	0.077	0.603	0.849	0.077
	p3	0.933	0.95	0.881	0.656	0.811	0.19	0.527	0.684	0.064	0.527	0.684	0.064
15	S	0.929	0.935	0.888	0.531	0.582	0.152	0.424	0.481	0.038	0.426	0.483	0.038
	W	0.963	0.98	0.9	0.74	0.875	0.154	0.677	0.832	0.054	0.678	0.832	0.054
	p8	0.937	0.947	0.882	0.682	0.759	0.166	0.552	0.636	0.044	0.552	0.636	0.044
20	S	0.931	0.934	0.887	0.535	0.572	0.14	0.433	0.469	0.025	0.435	0.471	0.026
	W	0.972	0.984	0.908	0.717	0.831	0.128	0.699	0.825	0.037	0.699	0.825	0.037
	p8	0.921	0.928	0.859	0.747	0.81	0.178	0.554	0.61	0.031	0.567	0.625	0.031
25	S	0.936	0.94	0.895	0.505	0.535	0.124	0.424	0.458	0.019	0.425	0.458	0.019
	W	0.97	0.982	0.903	0.746	0.838	0.126	0.708	0.812	0.029	0.708	0.812	0.029
	p13	0.922	0.927	0.86	0.73	0.775	0.162	0.543	0.588	0.019	0.555	0.6	0.019
30	S	0.945	0.947	0.905	0.482	0.502	0.122	0.43	0.454	0.027	0.43	0.454	0.027
	W	0.973	0.981	0.903	0.759	0.827	0.13	0.73	0.807	0.033	0.73	0.807	0.033
	p13	0.924	0.928	0.861	0.742	0.779	0.169	0.559	0.595	0.025	0.571	0.607	0.025
35	S	0.952	0.954	0.913	0.453	0.469	0.103	0.425	0.446	0.017	0.426	0.447	0.017
	W	0.973	0.981	0.902	0.763	0.825	0.117	0.733	0.808	0.019	0.733	0.808	0.019
	p18	0.921	0.926	0.857	0.74	0.776	0.162	0.548	0.587	0.016	0.561	0.6	0.016
40	S	0.959	0.961	0.922	0.427	0.44	0.091	0.425	0.444	0.014	0.429	0.448	0.014
	W	0.975	0.982	0.902	0.778	0.831	0.122	0.75	0.814	0.024	0.75	0.814	0.024
	p18	0.934	0.938	0.869	0.736	0.765	0.154	0.582	0.614	0.02	0.59	0.622	0.02
50	S	0.927	0.928	0.882	0.55	0.559	0.141	0.437	0.447	0.021	0.44	0.449	0.021
	W	0.975	0.98	0.902	0.781	0.831	0.129	0.754	0.808	0.032	0.754	0.808	0.032
	p19	0.918	0.921	0.854	0.749	0.777	0.175	0.548	0.573	0.024	0.563	0.588	0.025
	P21	0.929	0.933	0.867	0.729	0.758	0.162	0.567	0.595	0.025	0.575	0.604	0.025

3.5 Illustration with Numerical Example

Below are the randomly selected scores of 25 LAUTECH students in the STA 509 examination: 40,73,40,45,51,53,55,40,60,76,50,64,50,41,67,77,72,50,64,66,65,66,41,53,60.
Test the claim that the mean score of students is 65 at a 10% significance level, assuming the scores are normally distributed.

Solution

(i) **Test of Normality Assumption**

Hypothesis: H_0: Scores are normally distributed. vs H_1: Scores are not normally distributed.

Decision: Since the p-value = 0.200 > $\propto$ = 0.1 for Kolmogorov-Smirnov test and p-value = 0.114 > $\propto$ = 0.1 for Shapiro-Wilk test, we do not reject H_0. Hence, we conclude the scores are normally distributed.

(ii) **Test of the location parameter value**

Hypothesis: $H_0: \mu = 65 \quad vs \quad H_1: \mu \neq 65$

Re-arranging scores in order of magnitude and providing their rank (R) and grade(G), Table 7 is obtained.

Table 7: Scores, ranks, and grades of the students.

X	40	40	40	41	41	45	50	50	50	51	53	53	55	60	60	64	64	65	66	66	67	72	73	76	77
R	1	2	3	4	5	6	7	8	9	10	11	12	13	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	22	23	24	25
G	E	E	E	E	E	D	C	C	C	C	C	C	C	**B**	**B**	**B**	**B**	**B**	**B**	**B**	**B**	A	A	A	A

NOTE:

(i) The ranks in bold form are the ranks of the first eight scores closest to the mean.
(ii) 0-39 = F, 40 – 44 = E, 44 - 49 = D, 50 – 59 = C, 60 – 69 = B, 70 and above = A

Decision: The t-test value obtained is -3.449 with a p-value of 0.00258. Since 0.00258<0.1, we reject H_0 and conclude that the claim is not sustained.

The new approach:

For n=25, the recommended p=8. The ranks of the first eight scores closest to the means are 14,15,16,17,18,19,20 and 21. $\mu^{r8} = \frac{14+15+16+17+18+19+20+21}{8} = 17.5$ with p-value =0.006 (distributionally). Asymptotically, the z value of the proposed statistic is 2.0972 with p-value=0.018.

Decision: Since the p-value in both cases (distributional and asymptotic) <0.1, we reject H_0 and conclude that the claim is not sustained.

(iii) **Test of the location grade**

Hypothesis: $H_0: \mu = B \quad vs \quad H_1: \mu \neq B$

The classification of the students' grades is summarized in Table 8.

Table 8: Frequency distribution of the student's grade

Grade	**E**	**D**	**C**	**B**	**A**
Frequency	5	1	7	8	4
Cumulative frequency	5	6	13	21	25

There are 2 middle observations in the grade B and 3 others on both sides making the 8 observations (p=8) prescribed by the results of the study. Thus, $\mu^{r8} = \frac{14+15+16+17+18+19+20+21}{8} = 17.5$ with p-value =0.006 (distributionally). Asymptotically, the z value of the proposed statistic is 2.0972 with p-value=0.018.

Decision: Since the p-value in both cases (distributional and asymptotic) <0.1, we reject H_0 and conclude that the claim is still not sustained.

4. Conclusion

This work proposed a one-sample non-parametric statistic for testing the hypothesis about the location parameter when normally distributed data are ranked or classified. The proposed test statistic, μ^{r_p}, is the average rank of the first p observations closest to the hypothesized mean value. Based on type I error rate and power rate investigation, it is an alternative non-parametric statistic that can handle normally distributed and classified data for one sample problem, and it is as good as the parametric test, especially at a 10% (0.1) significance level at a known value of p. The methodology of this newly proposed statistic is easy and straightforward, and test statistic is recommended for use even though it has a limitation on the choice of p.

References

Adejumo T. J., Omonijo D. O., Owolabi A. T., Okegbade A. I., Odukoya A. J. and Ayedun C. A. (2020). Comparison of Rank Transformation Test Statistics with Its Nonparametric Counterpart Using Real-Life Data. *Mathematics and Statistics*, **8**(6), 683-692.

Ayinde, K., Adejumo, T. J. and Solomon, S. G. (2016). A study on Sensitivity and Robustness of one sample test Statistics to outliers. *Global Journal of Science Frontier Research*, **16**(6), 99-112.

Ayinde. K., Sanni, O.O.M., and Oyejola, B.A. (2009). An alternative test statistic to the paired t- test statistic. Global Journal of Mathematics and Statistics, 2, 125-129.

Cohen, J. A. (1960). Coefficient of agreement for nominal scales. *Education psychological measure*, *20*, 37-46.

Conover, W. (1999). Practical Nonparametric Statistics. Wiley.

Conover, W.J., and Iman, R.L. (1981). Rank transformations as a bridge between and nonparametric statistics. *The American Statisticians*, **35**, 124 – 129.

Derrick, B., Russ, B., Toher, D., and White, P. (2017). Test statistics for the comparison of means for two samples that include both paired and independent observations. *Journal of Modern Applied Statistical Methods*, 137-157. doi:10.22237/jmasm/ 1493597280

Derrick, B., White, P., and Toher, D. (2019). Parametric and non-parametric tests for the comparison of two samples which both include paired and unpaired observations. *Journal of Modern Applied Statistical Methods*, **18**(1), eP2847. doi:10.22237/jmasm/1556669520

Field, A. (2009). Discovering Statistics using SPSS. London: Sage Publications.

Faizi,N. and Yasir Alvi, Y. (2023). Biostatistics Manual for Health Research: A Practical Guide to Guide to Data Analysis. Elsevier Inc. https://doi.org/10.1016/C2022-0-00374-3

Fisher Box, J. (1987). Guinness, Gosset, Fisher, and Small Samples. Statistical Science, 2, 45–52.

Glass, G. V., Peckham, P. D., and Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, **42**(3), 237-288.

Ito, P. (1980). Robustness of Anova and Manova Test Procedures. Handbook of Statistics. (P. Krishnaiah, Ed.) 1, 199-236.

Jolayemi, E. (1990). On the measure of agreement between two raters. *Biometric Journal*, **32**, 87 – 93.

Lawal, H.B. (2003) Categorical Data Analysis with SAS and SPSS Applications. Lawrence Erlbaum Associates, Inc., New Jersey, 83-128.

Lindman, H. R. (1974). Analysis of variance in complex experimental designs. San Francisco: W. H. Freeman & Co.

Luepsen, H. (2017). Comparison of nonparametric analysis of variance methods - a Vote for van der Waerden. *Communications in Statistics - Simulation and Computation*. doi:10.1080/03610918.2017.1353613

Malik, A. and Kumar, A. (2020). Spatio-temporal trend analysis of rainfall using parametric and non-parametric tests: case study in Uttarakhand, India. T*heoretical and Applied Climatology*, **140**, 183–207.

Osborne, J. (2008). Best Practices in Quantitative Methods. Los Angeles: Sage Publications.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philosophical Magazine Series 5, 5.

Salzburg, D. (2001). The Lady Tasting Tea. New York: W. H. Freeman and Company.

Siegel, S. (1956). Nonparametric statistics for the behavioral sciences. McGraw-Hill.

Tanner, M.A. and Young, M.A. (1985). Modelling agreement among raters. *Journal of American Statistical Association*, **80**, 175 – 180.

Wallis, S. (2013). z-squared: the origin and application of χ^2. *Journal of Quantitative Linguistics*, **20**, 350-378.

Wilcox, R. (2005). Introduction to robust estimation and hypothesis testing. Burlington MA: Elsevier.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, **1**, 80–83.

Zimmerman, D. (2004). Inflation of Type I Error Rates by Unequal Variances Associated with Parametric, Nonparametric, and Rank-Transformation Tests. *Psicológica*, **25**, 103-133