Journal of Physical Science and Innovation **Volume 5, No. 2, 2013**
ISSN: 2277-0119

ON THE ANALYTICAL EFFICIENCY OF THE EXTENDED WILCOXON MATCHED PAIRS SIGNED RANK TEST

[1]S.A. Abdulazeez and [2]Lasisi A.R.
[1]Department of Mathematics, Statistics and Computer Science, Kaduna Polytechnic, Kaduna State, Nigeria.
[2]Department of Physics, Federal College of Education, Kontagora, Niger State, Nigeria.
E-mail: mailto:yinkasikiruabdul@yahoo.comlasisiar@yahoo.com

ABSTRACT

The Extended Wilcoxons' Matched Pairs Sign Rank Test is a non-parametric test which is equivalent to the t-test. This new method provides for an adjustment by the constant 0.5 in the conventional Wilcoxons' matched pairs sign rank test. This adjustment is justified because we are approximating a discrete distribution by a continuous distribution. 0.5 is added when $T \leq k/2$ and it is subtracted if $T > k/2$. k is the effective sample size. More efficient results are obtained with this method as illustrated by two data sets. This method should be viewed as a useful tool for solving new problems.

Keywords: Parametric Method, Non-parametric Method, Analytical Efficiency.

INTRODUCTION

Statistical procedures in which assumptions are required about the value of a parameter and the nature of the distributions from which the samples are drawn are known as parametric methods. Non-parametric methods do not depend on any parameter for its computation. It requires no assumption concerning the distribution from which the sample is drawn; hence it is sometimes referred to as distribution free methods. Non-parametric methods can be applied to nominal and ordinal data which includes ranks, medians and frequencies rather than parameters.

In situations where both tests, parametric and non-parametric are equally applicable, then the parametric test will be more powerful than the non- parametric test because the latter has a greater probability of accepting a false hypothesis, which is, committing a Type II Error. Examples of non-parametric statistics include Sign Test, Sign Rank Test, Median Test, Mann-Whitney U Test, Kolmogorov Smirnov Test, Spearman's Test and Wilcoxons' Test.

Advantages of non-parametric statistical methods include:

i. They provide short cut test that involve less mathematical details.
ii. They are simpler and less complex than their parametric parallels.
iii. They may be used to test data that are not exact in numerical sense but which in effect are simply ranking.
iv. It makes it possible to analyze samples whose population distribution is not known exactly.
v. It requires fewer assumptions about the shape of the population distribution or sample characteristics so that we may not worry about violating the assumptions.

Disadvantages of non-parametric statistical methods are:

i. It is less efficient than the parametric equivalence.
ii. It ignores much sample information

iii. Lack of universal dissemination in probability tables which are also often woefully inadequate in their coverage of the breadth of which some statistics might otherwise be employed.

MATERIAL AND METHOD

The Extended Wilcoxon Matched-Pairs Signed-Rank Test

The extended Wilcoxons' test is a more powerful test than the sign test since it gives more weight to a large difference than a small one. It considers both the sign and the difference between pairs of observation and the magnitude of that difference.

If d_i represents the absolute difference between x_i and y_i, then each pair of observations will have a value for $d_{i..}$

If $d_i = 0$ that pair of observations will be dropped from the study. When all values of d_i. are computed, they are ranked in order of magnitude from the smallest to the largest without regard to sign.

If there are ties in ranks, all the tied values of d_i are given the same rank which is obtained by obtaining the average of the ranks that would have been allocated to each tied observation.
After the values of d_i are ranked, each rank is then given the sign of the original difference between x_i and y_i. The value T is the sum either of the positive ranks or the negative ranks, whichever sum is smaller.

The null hypothesis, assumes that for the universe, the sum of the like-sign ranks is equal.
For samples of more than 25 observations, the theoretical sampling distribution of T is approximately normal with

$$\mu_T = \frac{k(k+1)}{4} \quad and \quad \sigma_T = \sqrt{\frac{k(k+1)(2k+1)}{24}}$$

thus,

,

$$Z_c = \frac{T - \mu_T \pm 0.5}{\sigma_T}$$

The adjustment by the constant 0.5 is necessary because we are approximating a discrete distribution by a continuous distribution. 0.5 is added when $T \leq k/2$ and it is subtracted if $T > k/2$. k is the effective sample size.

Illustration 1

Table 1: The Table Below is on the Sales of Frozen Food in Cartons at Two Different Location in Each of 30 Different Stores

No. of Cartons Sold			No. of Cartons Sold		
Store	**Frozen Food**	**Paper Food**	**Store**	**Frozen Food**	**Paper Food**
1	40	65	16	20	29
2	75	60	17	49	60
3	24	36	18	32	22
4	21	15	19	15	32
5	8	12	20	80	120
6	10	15	21	30	30
7	15	12	22	16	32
8	30	48	23	41	20
9	22	21	24	35	11
10	16	16	25	24	50
11	15	8	26	18	49
12	56	85	27	16	58
13	12	20	28	10	8
14	4	18	29	37	18
15	32	45	30	50	70

Steps

1. **Hypothesis**

   H_0: The universe sum of the positive sign ranks is the same as the sum of negative sign ranks. The two locations are equally good.

   H_1: The universe sum of the positive sign ranks is less than the sum of negative sign ranks the paper food location is the better.

2. Significance level

   An $\alpha = 0.05$ require a $Z_c = 1.64$.

   Test Statistic

   $$\mu_T = \frac{k(k+1)}{4} \quad and \quad \sigma_T = \sqrt{\frac{k(k+1)(2k+1)}{24}}$$

3. *thus,*

   $$Z_c = \frac{T - \mu_T \pm 0.5}{\sigma_T}$$

4. Decision Criterion: Reject H_0 if $Z_c \leq 1.64$

5. The computation of T is shown in the table below. The figures shown in the column headed d_i are the absolute differences between the values of X_i and Y_i.

The d_i values are ranked after which the initial sign is being considered to classify each rank as positive or negative. We denote the rank corresponding to positive difference as T_+ and that corresponding to negative difference as T_-.

Also, let

$\sum T_-$ = Sum of the negative ranks and,

$\sum T_+$ = Sum of the positive ranks

k = Effective sample size for which $d_i \neq 0$

Table 2: Computation of d_i for the Sales of Frozen Food Cartons at Two Different Location in Each of 30 Different Stores

No. of Cartons Sold				No. of Cartons Sold			
Store	**Frozen Food (X_i)**	**Paper Foods (Y_i)**	**(d_i)**	**Store**	**Frozen Food (X_i)**	**Paper Food (Y_i)**	**(d_i)**
1	40	65	-25	16	20	29	-9
2	75	60	+15	17	49	60	-11
3	24	36	-12	18	32	22	+10
4	21	15	+6	19	15	32	-17
5	8	12	-4	20	80	120	-40
6	10	15	-5	21	30	30	0
7	15	12	+3	22	16	32	-16
8	30	48	-18	23	41	20	+21
9	22	21	+1	24	35	11	+24
10	16	16	0	25	24	50	-26
11	15	8	+7	26	18	49	-31
12	56	85	-29	27	16	58	-42
13	12	20	-8	28	10	8	+2
14	4	18	-14	29	37	18	+19
15	32	45	-13	30	50	70	-20

Table 3: Computation of T for an Extended Wilcoxon Matched-Pairs Signed–Ranked Test

Store	**Di**	**Rank of di**	**T_-**	**T_+**
		-23	23	
1	25	+15		15
2	15	-12	12	
3	12	+6		6
4	6	-4	4	
5	4	-5	5	
6	5	+3		3
7	3	-18	18	
8	18	+1		1
9	1	discard	-	-
10	0	+7		7
11	7	-25	25	
12	29	-8	8	
13	8	-14	14	
14	14	-13	13	
15	13	-9	9	
16	9	-11	11	
17	11	+10		10
18	10	-17	17	
19	17	-27	27	
20	40	discard	-	-
21	0	-16	16	
22	16	+21		21
23	21	+22		22
24	24	-24	24	
25	26	-26	26	
26	31	-28	28	
27	42	+2		2
28	2	+19		19
29	19	-20	20	
30	20			
			$\sum T_- = 300$	$\sum T_+ = 106$

$$
\begin{aligned}
T &= \min(\textstyle\sum T_-, \sum T_+) \\
&= \min(300, 106) \\
&= 106
\end{aligned}
$$

$T = 106$, $k = 28$

Since $T > k/2$

Then

$$\mu_T = \frac{k(k+1)}{4}$$
$$= \frac{28 \text{ x } 29}{4}$$
$$= 203$$

$$\sigma_T = \sqrt{\frac{k(k+1)(2k+1)}{24}}$$
$$= \sqrt{\frac{28(29)(57)}{24}}$$
$$= 43.915$$
$$Z_c = \frac{T - \mu_T - 0.5}{\sigma_T}$$
$$= \frac{106 - 203 - 0.5}{43.915}$$
$$= -2.2202$$

Since Z_c = -2.2202 < -1.64 reject H_O. the location of the frozen food cartons with the paper goods will lead to greater sales than in a location mean the frozen foods.

Illustration II
Use the Wilcoxon's test to analyze the paired data below at α = 5% taking

Ho $\tilde{\mu}_x = \tilde{\mu}_Y$ vs H_1 $\tilde{\mu}_x \rangle \tilde{\mu}_Y$

Table 4

X	Y	di	$R_{/di/}$	T_+	T_-
9	5	+4	2.5	2.5	
7	3	+4	2.5	2.5	
3	4	-1	1		1
16	11	+5	5	5	
12	7	+5	5	5	
12	5	+7	7	7	
5	5	0	-		
6	1	+5	5	5	
				27	1

$T = \min(\sum T_-, \sum T_+)$
$= \min(1, 27)$
$= 1$
$T = 1$, $k = n - 1 = 7$

$$\mu_T \frac{K(k+1)}{4} = \frac{7(8)}{4} = 14,$$

$$\sigma_T = \sqrt{\frac{k(k+1)(2k+1)}{24}} = \sqrt{\frac{7x8x15}{24}}$$
$$= 5.92$$

Test statistic

Since $T \leq k/2$

then

$$Z_c = \frac{T - \mu_T + 0.5}{\sigma_T} = \frac{1 - 14 + 0.5}{5.92} = -2.111$$

From table $Z_{0.05} = 1.64$

Since $|Z_C| = 2.111 > 1.64$ reject H_O.

DISCUSSION AND CONCLUSION

The results above show that the introduction of an extended Wilcoxons matched pairs Sign rank test provides a modest difference in the decisions based on the null hypothesis and the probability of committing a type II error is thereby minimized. The model offers a flexible tool for the study of non-parametric paired data for large and small samples. This method should not be seen as merely an alternative non-parametric test but as a supplement that improves the efficiency of the test of hypothesis.

REFERENCES

Chernoff H. (1952), A Measure of Asymptotic Efficiency for Tests of Hypothesis Based on the Sum of Observations. *Ann. Math Stat 23*, 493 – 507.

Dixon W.J. (1954), Power Under Normality of Several Non-Parametric Tests. *Ann. Math Stat 25*, 610 – 614.

Fallik F. & Brown B. (1983), *Statistics for Behavioural Sciences The Dorsey Press. New York* Pp. 413 – 438.

Freund J.E. & Walpole R.E. (1987), *Mathematical Statistics*. Fourth Edition. Prentice Hall. New Delhi Pp. 521 -545.

Hodges J.L. & Lehman E.L. (1956), The Efficiency of Some Non-Parametric Competitors of the t- test. *Ann. Math Stat* 27, 324 – 335.

Hoeffding W. & Rosenblatt J. (1955). The Efficiency of Tests. *Ann. Math Stat 26*, 53 – 63

Journal Amer. Stat Assoc. 47, 583 – 621.

Kruskal W.H. & Wallis W.A. (1952), Use of Ranks in One Criterion Analysis of Variance.

Mood A. (1955) On the Asymptotic Efficiency of Certain Non-Parametric Two Sample Test. *Ann. Math Stat 25*, 514 – 522.

Reference to this paper should be made as follows**:** S.A. Abdulazeez and Lasisi A.R. (2013), On the Analytical Efficiency of the Extended Wilcoxon Matched Pairs Signed Rank Test. *J. of Physical Science and Innovation*, Vol. 5, No. 2, Pp. 120 – 127.