Scientific
Research

Statistical Comparison of Eight Alternative Methods for the Analysis of Paired Sample Data with Applications

Godday Uwawunkonye Ebuh*, Ikewelugo Cyprian Anaene Oyeka
Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University,
Awka, Nigeria
Email: *ablegod007@yahoo.com

ABSTRACT

This paper presents and statistically compares eight alternative methods that could possibly be used in the analysis of matched or paired sample data, including situations in which the data being analyzed satisfy the usual assumptions of normality and continuity necessary for the use of parametric tests as well as when the data are numeric and non-numeric measurements on as low as the ordinal scale. It is shown that only the modified sign tests based on only the raw observations or their assigned ranks may be used with non numeric measurement on the ordinal scale. If the ordinary sign test, the Wilcoxon signed rank sum test and the modified sign tests can be equally used in data analysis, then it is shown that the modified sign tests are more efficient and hence more powerful than the ordinary sign tests because the two test statistics are intrinsically and structurally modified for the possible presence of tied observations between the sampled populations for both using raw and simulated data. Of all the non-parametric methods presented, the modified Wilcoxon's signed rank sum test when applicable is the most efficient and powerful, followed in this order by the modified sign test by ranks and the modified sign test based on only raw scores for raw data while using simulation, modified sign test by ranks is the most efficient and powerful, followed in this order by modified Wilcoxon's signed rank sum test and modified sign test. Each of the non-parametric methods presented can be easily modified and re-specified for use with one sample data by simply re-designating the observations from one of the sampled populations to correspond with a hypothesized value of some measure of central tendency. The methods are illustrated with some raw data as well as simulated data and their relative performances compared.

Keywords: Normality; Continuity; Paired Sample; Parametric Test; Nonparametric; Numeric; Relative Performance; Tied Observation

1. Introduction

A clinician, medical researcher or research scientist may expose a random sample of subjects to some treatment or drug at two points in time or space, or expose two random samples of subjects matched on several characteristics, one to an active or new drug or treatment, and the other to a diluent, inactive placebo or control treatment and research interest is in comparing the responses after the exposure. A dietician may be interested in studying a random sample of subjects, treated with a regimen of diet or exercises and in measuring their responses in terms of the differences between body weights before and after the experiences. A panel of judges or examiners may be interested in comparing the performances of candidates in two tests or examinations taken at two points in time or space. A psychologist or psychiatrist may wish to compare the performance of two matched samples of subjects exposed to two experimental conditions. A beautician, marketing consultant or advertising agent, product promoter or investor may wish to compare the performance of a line of products in terms of their acceptability or sales, at two different points in time or space, etc.

In each of these and similar situations, the researcher may wish to select a statistical method often used in the analysis of matched or paired samples that is relatively efficient and powerful in terms of being able to more readily reject a false null hypothesis and accept a true one and hence be able to reach more reliable conclusions.

This paper presents, discusses and compares eight alternative statistical methods that may be used for this purpose.

2. The Proposed Methods

Let (x_{i1}, x_{i2}) be the ith pair randomly drawn from populations X_1 and X_2, for $i = 1,2, \cdots, n$. Populations X_1 and X_2

---

*Corresponding author.

may or may not be measurements that are continuous; normally distributed; numeric data; independent; but they should be measurements on at least the ordinal scale. Interest is in statistically comparing the following eight methods for analyzing paired samples. They include paired sample t test, ordinary sign test for two samples, exact binomial test, normal approximation to the ordinary two sample sign test, unmodified Wilcoxon signed rank sum test for paired samples, modified sign test, modified sign test by ranks, and modified Wilcoxon signed rank sum test for paired samples.

All modifications or adjustments of test statistics are aimed at adjusting and making provisions for the possibility of any ties, that is tied observations between sampled populations and hence obviate the need to require the sampled populations to be continuous or even numeric.

2.1. Paired Sample T Test

Required Assumptions

Populations continuous and normally distributed; or sample size sufficiently large [1]. This method can only be used for data that satisfy the required assumptions and are measurements on at least the interval scale.

Let

$$d_i = x_{i1} - x_{i2} \tag{1.1}$$

for $i = 1, 2, \cdots, n$.

Let $\bar{d} = \dfrac{\sum\limits_{i=1}^{n} di}{n}$ be the mean value of the differences

and

$$s_d^2 = \sum_{i=1}^{n} di^2 - \frac{\left(\dfrac{\left(\sum\limits_{i=1}^{n} di\right)^2}{n}\right)}{n-1} \tag{1.2}$$

Be the variance of the differences.

Let:

$$Se(\bar{d}) = \frac{S\bar{d}}{\sqrt{n}} \tag{1.3}$$

be the standard deviation of the mean difference $\bar{d}$. We want to test the null hypothesis

$$H_0 : \mu_d = d_0 \text{ vs } H_1 : \mu_d \neq d_0 \tag{1.4}$$

where d_0 is any real number including zero. The test statistic is [1].

$$t = \frac{\bar{d} - d_0}{S_d \sqrt{n}} \tag{1.5}$$

which has a t distribution with $n - 1$ degrees of freedom .We reject H_0 at the α level of significance if

$$|t| \geq t_{1-\alpha/2;\, n-1} \tag{1.6}$$

Otherwise H_0 is accepted.

2.2. Ordinary Sign Test for Two Samples

Required Assumptions

Populations continuous and numeric measurements.

The test statistic is based on the signs (+ sign, or − sign) of the differences between members of the paired sample observations.

Thus let

$$d_i = x_{i1} - x_{i2} \tag{2.1}$$

for $i = 1, 2, \cdots, n$.

Let

$$u_{i1} = \begin{cases} 1, & \text{if} \quad d_i > 0 \\ 0, & \text{if} \quad d_i < 0 \end{cases} \tag{2.2}$$

for $i = 1, 2, \cdots, n$.

Note that Equation (2.1) assumes that there are no ties that is d_i cannot be 0 and hence does not make any provisions for this possibility.

Let

$$p(u_i = 1) = \theta \tag{2.3}$$

Let

$$W = \sum_{i=1}^{n} u_i \tag{2.4}$$

It is easily shown [2] that

$$E(W) = n.\theta; \; Var(W) = n\theta(1-\theta) \tag{2.5}$$

The test statistics for the null hypothesis of equal population medians $(H_0: M_1 = M_2 = M_0)$, that is of the null hypothesis,

$$H_0 : \theta = \theta_0 = 0.50 \text{ vs } H_1 : \theta = \theta_0 \neq 0.50 \tag{2.6}$$

In general

$$\chi^2 = \frac{(w - n\theta_0)^2}{Var(w)} = \frac{(w - n\theta_0)^2}{n\theta_0(1-\theta_0)} \tag{2.7}$$

which has the chi-square distribution with 1 degree of freedom for sufficiently large n.

H_0 is rejected at the α level of significance if

$$\chi^2 \geq \chi_{1-\alpha:1}^2 \tag{2.8}$$

Otherwise H_0 is accepted.

Note that in particular under the null hypothesis usually tested in the sign test ($H_0: \theta = \theta_0 = 0.50$) Equation (2.7) reduces to

$$\chi^2 = \frac{(w - 0.5n)^2}{0.25n} \qquad (2.9)$$

which has the chi-square distribution with 1 degree of freedom, if n is sufficiently large.

2.3. Exact Binomial Test

Assumption: Data is discrete

As in 2 above, let

$$d_i = \begin{cases} 1\,(\text{or } +) & \text{if} \quad x_{i1} > x_{i2} \\ 0, & \text{if} \quad x_{i1} = x_{i2} \\ -1\,(\text{or } -) & \text{if} \quad x_{i1} < x_{i2} \end{cases} \qquad (3.1)$$

for $i = 1, 2, \cdots, n$

Under the null hypothesis of equal population medians, we would expect that 1's or +'s are as likely to occur as −1's or −'s. In other words we would expect that

$$p(1 \text{ or } +) = p(-1 \text{ or } -) = \theta_0 \quad \text{in general} \qquad (3.2)$$

Therefore too many of 1's (or + signs) or −1's (or − signs) will lead to a rejection of the null hypothesis.

If we let X be the number of plus signs (or minus signs, depending for simplicity on which one is smaller). Then the probability of obtaining at most $X = x$ plus signs is − [3] calculated from the binomial equation

$$p(X \le x) = \sum_{k=0}^{x} \binom{n}{k} \cdot \theta_0^k (1 - \theta_0)^{n-k} \qquad (3.3)$$

where n is the effective sample size (number of + signs plus number of minus signs, excluding all zero). In particular under the null hypothesis usually tested in paired sample tests (H_0: $\theta = \theta_0 = 0.50$), the null hypothesis of equal population medians is rejected at the α level of significance if

$$p(X \le x) = \sum_{k=0}^{x} \binom{n}{k} (0.5)^n \le \alpha/2 \qquad (3.4)$$

where α is the specified level of significance.

If the alternative hypothesis suggests a one-sided test, then H_0 is rejected at the α level of significance if

$$p(X \le x) = \sum_{k=0}^{x} \binom{n}{k} (0.5)^n \le \alpha \qquad (3.5)$$

otherwise H_0 is accepted.

Note that the exact binomial test leads to essentially the same conclusion as the ordinary sign test presented in Section 2.2 above.

2.4. Normal Approximation to the Ordinary Two Sample Sign Test

Assumption: Data is discreet.

The binomial test is usually used in the ordinary sign test to calculate the exact probability that is sufficiently satisfactory for most sample sizes encountered in practice.

In general where as in the usual sign test, the null hypothesis is

$$H_0 : \theta = \theta_0 = 0.50 \text{ vs } H_1 : \theta = \theta_0 \ne 0.50 \qquad (4.1)$$

Then using the notations of Section 2.2 the test statistic becomes

$$\chi^2 = \frac{\left((w \pm 0.5) - n\theta_0\right)^2}{n\theta_0(1 - \theta_0)} = \frac{\left((w \pm 0.5) - 0.5n\right)^2}{0.25n} \qquad (4.2)$$

where

$$w \pm 0.5 = \begin{cases} w + 0.5, & \text{if} \quad w \le 2/n \\ w - 0.5, & \text{if} \quad w > 2/n \end{cases} \qquad (4.3)$$

which has approximately the chi-square distribution with 1 degree of freedom.

However, for sufficiently large n the normal approximation can be used which then becomes

$$z = \frac{\left((w \pm 0.5) - n\theta_0\right)}{\sqrt{n\theta_0(1 - \theta_0)}} = \frac{\left((w \pm 0.5) - 0.5n\right)}{0.5\sqrt{n}} \qquad (4.4)$$

H_0 is rejected at the α level of significance if

$$|z| \ge z_{1-\alpha/2} \qquad (4.5)$$

Otherwise H_0 is accepted.

2.5. Unmodified Wilcoxon Signed Rank Sum Test for Paired Samples

This test is similar to the ordinary sign test except that it is based on the ranks of the absolute differences, $/d_i/$, of the differences, d_i between paired observations instead of only on the signs of the difference between the ith pair of sample observations, for $i = 1, 2, \ldots, n$

Let

$$T^+ = \sum_{i=1}^{n} r(|d_i|) u_i \qquad (5.1)$$

where "r" $(|d_i|)$ is the rank assigned to $|d_i|$, the absolute value of the differences $d_i = x_{i1} - x_{i2}$ without loss of generality we may assume that $r(|d_i|) = i$, so that

$$T^+ = \sum_{i=1}^{n} i \cdot u_i \qquad (5.2)$$

It is easily shown that

$$\text{E}(T^+) = \frac{n(n+1)}{2} \theta;$$
$$Var(T^+) = \frac{n(n+1)(2n+1)}{6} \theta(1 - \theta) \qquad (5.3)$$

The unmodified Wilcoxon's sign rank sum test statistic for the general null hypothesis of constant difference between population medians (H_0: $\theta = \theta_0$) is [4]

$$\chi_u^2 = \frac{\left(T^+ - \frac{n(n+1)}{2}\theta_0\right)^2}{\frac{n(n+1)(2n+1)}{6}\theta_0(1-\theta_0)} \qquad (5.4)$$

which under H_0 has approximately the chi-square distribution with 1 degree of freedom for sufficiently large n. H_0 is rejected at the α level of significance if Equation (2.8) is satisfied, otherwise H_0 is accepted.

In particular, under the null hypothesis usually tested in the Wilcoxon's signed rank sum test (H_0: $\theta = \theta_0 = 0.50$), Equation (5.4) becomes

$$\chi_u^2 = \frac{\left(T^+ - \frac{n(n+1)}{4}\right)^2}{\frac{n(n+1)(2n+1)}{24}} \qquad (5.5)$$

which has approximately a chi-square distribution with 1 degree of freedom for sufficiently large n.

2.6. Modified Sign Test

The ordinary sign test is modified for the possibility of tied observations between the two matched or paired observations and to also provide for the possibility that the ordinal scale data being analyzed may be non-numeric; we re-specify u_i as follows;

Let

$$u_i = \begin{cases} 1, & \text{if } x_{i1} \text{ is a higher or larger score or} \\ & \text{observation than } x_{i2} \\ 0, & \text{if } x_{i1} \text{ is the same score as, that is equal to } x_{i2} \\ -1, & \text{if } x_{i1} \text{ is a lower or smaller score or} \\ & \text{observation than } x_{i2} \end{cases}$$

for $i = 1, 2, \cdots, n$.

Let

$$P(u_i = 1) = \pi^+; P(u_i = 0) = \pi^0; P(u_i = -1) = \pi^- \qquad (6.2)$$

where

$$\pi^+ + \pi^0 + \pi^- = 1 \qquad (6.3)$$

Let

$$w = \sum_{i=1}^{n} u_i \qquad (6.4)$$

It is easily shown that [2]

$$E(w) = n(\pi^+ - \pi^-);$$
$$Var(w) = n\left(\pi^+ + \pi^- - (\pi^+ - \pi^-)^2\right) \qquad (6.5)$$

It can also be easily shown that the sample estimates of π^+, π^0 and π^- are respectively.

$$\hat{\pi}^+ = \frac{f^+}{n}; \hat{\pi}^0 = \frac{f^0}{n}; \hat{\pi}^- = \frac{f^-}{n} \qquad (6.6)$$

where f^+, f^0 and f^- are respectively the number of 1's, 0's and −1's in the frequency distribution of the n values of these numbers in u_i, for $i = 1, 2, \cdots, n$.

Also

$$w = f^+ - f^- = n(\hat{\pi}^+ - \hat{\pi}^-) \qquad (6.7)$$

The test statistic for the null hypothesis that the population medians differ by some constant β_0

$$H_0 : (\pi^+ - \pi^- = \beta_0)$$

Vs

$$H_1 : (\pi^+ - \pi^- \neq \beta_0)(0 \leq \beta_0 \leq 1)$$

is

$$\chi_m^2 = \frac{(w - n\beta_0)^2}{Var(w)} = \frac{(w - n\beta_0)^2}{n\left(\left(\hat{\pi}^+ + \hat{\pi}^-\right) - (\hat{\pi}^+ - \hat{\pi}^-)^2\right)} \qquad (6.8)$$

which under H_0 has approximately the chi-square distribution with 1 degree of freedom for sufficiently large n.

H_0 is rejected at the α level of significance if Equation (2.8) is satisfied otherwise H_0 is accepted. In particular the test statistic for the null hypothesis usually tested for paired samples $H_0 : (\pi^+ - \pi^- = \beta_0 = 0)$ is

$$\chi^2 = \frac{w^2}{n\left(\left(\hat{\pi}^+ + \hat{\pi}^-\right) - (\hat{\pi}^+ - \hat{\pi}^-)^2\right)} \qquad (6.9)$$

under which H_0 has approximately the chi-square distribution with 1 degree of freedom for sufficiently large n. As noted above this method may also be used with ordinal scale data that are non-numeric measurements.

2.7. Modified Paired Sample Test by Ranks

A rather noval and relatively more efficient and hence more powerful alternative method also exists. This method is however similar to the one discussed in six above and yields similar but often more powerful results because the paired raw scores or observations are first changed into ranks before use. Thus, let x_{i1} be assigned the rank $r_{i1} = k + 1$, k or $k − 1$ if x_{i1} is a higher or larger score or observation, the same or equal score, lower or smaller score than x_{i2}. Similarly, let x_{i2} be assigned the rank $r_{i2} = k + 1$, k, or $k − 1$, if x_{i2} is a larger or higher, the same or equal, or lower or smaller score than x_{i1}, for $i = 1, 2, \cdots, n$ where (x_{i1}, x_{i2}) is the ith pair of sample observations and k is any real number.

Let

$$r_i = r_{i1} - r_{i2} \qquad (7.1)$$

Also let

$$u_i = \begin{cases} 1, & \text{if} \quad r_i > 0 \\ 0, & \text{if} \quad r_i = 0 \\ -1, & \text{if} \quad r_i < 0 \end{cases} \qquad (7.2)$$

for $i = 1, 2, \cdots, n$.

Let

$$\pi^+ = P(u_i = 1); \pi^0 = P(u_i = 0); \pi^- = P(u_i = -1) \qquad (7.3)$$

where

$$\pi^+ + \pi^0 + \pi^- = 1 \qquad (7.4)$$

Define

$$w = \sum_{i=1}^{n} |r_i| u_i \qquad (7.5)$$

That is

$$W = R_{.1} - R_{.2} \qquad (7.6)$$

where $R_{.1}$ and $R_{.2}$ are respectively the sums of the ranks assigned to sample observations from populations X_1 and X_2.

$$Var(w) = \left(r_{.1}^2 + r_{.2}^2 - 2\left(n\left(k^2 - 1\right) + t\right)\right)$$
$$\left(\pi^+ + \pi^- - \left(\pi^+ - \pi^-\right)^2\right) \qquad (7.7)$$
$$= 4(n-t)\left(\pi^+ + \pi^- - \left(\pi^+ - \pi^-\right)^2\right)$$

which is independent of "k" since it is easily shown that

$$\left(r_{.1}^2 + r_{.2}^2 - 2\left(n\left(k^2 - 1\right) + t\right)\right) = 4(n-t) \qquad (7.8)$$

where t is the number of tied observations between populations X_1 and X_2 and $r_{.1}^2$ and $r_{.2}^2$ are respectively the sums of squares of the ranks assigned to sample observations from populations X_1 and X_2.

To test the general null hypothesis that the medians of the two sampled populations differ by some constant, that is the null hypothesis that the difference between the proportions of, or the probability that observations drawn from population X_1 are on the average higher (greater) than observations drawn from population X_2 and the probability that they are on the average lower (smaller) is some constant or notationally

$$H_0 : \left(\pi^+ - \pi^- = \beta_0\right)$$
$$\text{vs} \qquad\qquad (7.9)$$
$$H_1 : \left(\pi^+ - \pi^- \neq \beta_0\right)\left(0 \leq \beta_0 \leq 1\right)$$

is

$$\chi^2 = \frac{\left(w - \left(R_{.1} - R_{.2}\right)\beta_0\right)^2}{Var(w)}$$

$$= \frac{\left(w - \left(R_{.1} - R_{.2}\right)\beta_0\right)^2}{\left(r_{.1}^2 + r_{.2}^2 - 2\left(n\left(k^2-1\right)+t\right)\right)\left(\pi^+ + \pi^- - \left(\pi^+ - \pi^-\right)^2\right)}$$

$$= \frac{\left(w - \left(R_{.1} - R_{.2}\right)\beta_0\right)^2}{4(n-t)\left(\pi^+ + \pi^- - \left(\pi^+ - \pi^-\right)^2\right)}$$

$$(7.10)$$

which has approximately the chi-square distribution with 1 degree of freedom for sufficiently large "n".

The null hypothesis H_0 is rejected at the α level of significance if Equation (2.8) is satisfied otherwise H_0 is accepted.

In particular, under the null hypothesis usually tested in paired sample problems $H_0 : \left(\overset{+}{\pi} - \overset{-}{\pi} = \beta_0\right)$, Equation (7.10) reduces to

$$\chi^2 = \frac{w^2}{\left(r_{.1}^2 + r_{.2}^2 - 2\left(n\left(k^2-1\right)+t\right)\right)\left(\hat{\pi}^+ + \hat{\pi}^- - \left(\hat{\pi}^+ - \hat{\pi}^-\right)^2\right)}$$

$$\frac{w^2}{4(n-t)\left(\hat{\pi}^+ + \hat{\pi}^- - \left(\hat{\pi}^+ - \hat{\pi}^-\right)^2\right)}$$

$$(7.11)$$

These results are unaffected by any chosen real valued "k". However although the results obtained remain unchanged, it is often computationally easier and quicker if "k" is an integer.

The methods of Sections for 2.6 and 2.7 could be used alternatively to analyze the same types of data, although method 7 because it is based on ranks, is often more powerful than method 6 based on raw scores.

The two methods are nevertheless each more powerful than the unmodified Wilcoxon's signed rank sum test, because unlike the later, the former test statistics intrinsically adjust or make provisions for the possible presence of ties in the data. To show this, we note that the relative efficiency of W to T^+ is

$$RE\left(W; T^+\right) = \frac{Var\left(T^+\right)}{Var(w)} = \frac{n(n+1)(2n+1)/24}{n\left(\pi^+ + \pi^- - \left(\pi^+ - \pi^-\right)^2\right)}$$

$$\geq \frac{(n+1)(2n+1)}{24\left(1 - \overset{0}{\pi}\right)}$$

Since $\left(\pi^+ - \pi^-\right)^2 \geq 0$ and $\left(\pi^+ + \pi^-\right) = 1 - \pi^0$
Hence

$$RE\left(W;\ T^{+}\right) \geq 1 \qquad (7.12)$$

for all $n \geq 3$ and $0 \leq \overline{\lambda}^{0} < 1$ showing that W is more efficient and hence more powerful than T^{+} except for the very rare cases in which we have only one or two paired samples.

2.8. Modified Wilcoxon Signed Rank Sum Test for Paired Samples

This method is designed to correct for the shortfalls of the regular Wilcoxon Signed Rank Sum test T^{+} that does not intrinsically provide for the possibility of ties between the sampled populations. To do this, assuming d_i is as defined in Section 2.5, we let

$$u_i = \begin{cases} 1, & \text{if} \quad d_i > 0 \\ 0, & \text{if} \quad d_i = 0 \\ -1, & \text{if} \quad d_i < 0 \end{cases} \qquad (8.1)$$

Let

$$\pi^{+} = P\left(u_i = 1\right); \pi^{0} = P\left(u_i = 0\right); \pi^{-} = P\left(u_i = -1\right) \quad (8.2)$$

where

$$\pi^{+} + \pi^{0} + \pi^{-} = 1 \qquad (8.3)$$

Define

$$T = \sum_{i=1}^{n} r\left(\left|d_i\right|\right) u_i \qquad (8.4)$$

where $r\left(\left|d_i\right|\right) = i$ as defined in Section 2.5, the rank assigned to the absolute difference, $\left|d_i\right|$. Note that

$$T = T^{+} - T^{-} \qquad (8.5)$$

where T^{+} and T^{-} are respectively the sums of the ranks of absolute differences with positive and negative signs.

It is easily shown [5] that

$$E(T) = \frac{n(n+1)}{2}\left(\pi^{+} - \pi^{-}\right),$$

$$Var(T) = \frac{n(n+1)(2n+1)}{6}\left(\pi^{+} + \pi^{-} - \left(\pi^{+} - \pi^{-}\right)^2\right) \qquad (8.6)$$

The sample estimates of π^{+}, π^{0} and π^{-} are respectively

$$\hat{\pi}^{+} = \frac{f^{+}}{n}; \hat{\pi}^{0} = \frac{f^{0}}{n}; \hat{\pi}^{-} = \frac{f^{-}}{n} \qquad (8.7)$$

where f^{+}, f^{0}, and f^{-} are respectively the number of 1's 0's and -1's in the frequency distribution of the n values of these numbers in u_i, $i = 1, 2, \cdots, n$.

The corresponding test statistic for the general null hypothesis.

H_0: $\pi^{+} - \pi^{-} \geq \beta_0$ vs H_0: $\pi^{+} - \pi^{-} < \beta_0$ say $(0 \leq \beta_0 \leq 1)$

$$X_m^2 = \frac{\left(T - n\frac{(n+1)}{2} \cdot \beta_0\right)^2}{Var(T)} =$$

$$\frac{\left(T - n\frac{(n+1)}{2} \cdot \beta_0\right)^2}{\frac{n(n+1)(2n+1)}{6}\left(\hat{\pi}^{+} + \hat{\pi}^{-} - \left(\hat{\pi}^{+} - \hat{\pi}^{-}\right)^2\right)} \qquad (8.8)$$

which under H_0 has approximately the chi-square distribution with 1 degree of freedom for sufficiently large n. H_0 is rejected at the α level of significance if Equation (2.8) is satisfied otherwise H_0 is accepted.

In particular, the test statistic for the null hypothesis usually tested in paired sample problems

$\left(H_0 : \overset{+}{\pi} - \overset{-}{\pi} = \beta_0 = 0\right)$ reduces Equation (8.8) to simply

$$X_m^2 = \frac{T^2}{\frac{n(n+1)(2n+1)}{6}\left(\pi^{+} + \pi^{-} - \left(\pi^{+} - \pi^{-}\right)^2\right)} \qquad (8.9)$$

Note that the test statistic of Equation (5.4) could equivalently be expressed as

$$X_u^2 = \frac{3\left(4T^{+} - n(n+1)\right)^2}{2n(n+1)(2n+1)} \qquad (8.10)$$

while Equation (8.7) could equivalently be expressed as

$$X_m^2 = \frac{3\left(2T - n(n+1)\beta_0\right)^2}{2n(n+1)(2n+1)\left(\pi^{+} + \pi^{-} - \left(\pi^{+} - \pi^{-}\right)^2\right)} \qquad (8.11)$$

The relative efficiency of the test statistic given in Equation (8.11) for the modified Wilcoxon test statistic to Equation (8.10) for its unmodified counterpart may therefore be determined by comparing the variances of $4T^{+}$ and $2T$.

As

$$RE\left(2T; 4T^{+}\right) = \frac{Var\left(4T^{+}\right)}{Var\left(2T\right)} = \frac{4Var\left(T^{|+}\right)}{Var\left(T\right)} \qquad (8.12)$$

That is

$$RE\left(X_m^2; X_u^2\right) = \frac{4Var\left(T^{+}\right)}{Var\left(T\right)}$$

$$= \frac{n(n+1)(2n+1)}{n(n+1)(2n+1)\left(\pi^{+} + \pi^{-} - \left(\pi^{+} - \pi^{-}\right)^2\right)}$$

$$= \frac{1}{\pi^{+} + \pi^{-}} - \left(\pi^{+} - \pi^{-}\right)^2 \geq \frac{1}{1 - \pi^{0}}$$

Since $\left(\pi^{+} - \pi^{-}\right)^2 \geq 0$ and $\pi^{+} + \pi^{-} = 1 - \pi^{0}$.

Hence

$$RE\left(X_m^2; X_u^2\right) \ge 1 \qquad (8.13)$$

For all $0 \le \pi^0 < 1$.

Showing that the modified Wilcoxon test is more powerful than the unmodified test for all $\pi^0 \ne 1$ that is whenever there are tied observations between the sampled populations.

Finally, as an anecdote and for completeness it is necessary and instructive to add that correlation models may also be used to study the degree of association between paired or matched samples.

These include the Pearson's moment correlation coefficient used when the data being analysed are continuous and normally distributed and the Spearman's ranked correlation coefficient used when the data being analysed are measurements on at least the ordinal scale.

The corresponding test statistics which are fairly familiar are also available for use.

Again each of the proposed methods may be appropriately modified and used to analyse one sample data simply by setting values or scores from one of the sampled populations equal to a hypothesized value of some measure of central tendency.

An important advantage of the modified methods over the unmodified ones is that each of them intrinsically and structurally makes provisions for the possibility of tied observations in the sampled populations and hence makes it unnecessary to require the populations to be continuous. By making use of the information on all the observations instead of only on the non-zero ones, the modified methods are generally more efficient and hence more powerful than the unmodified methods. Their specifications also enable the researcher, policy maker or implementer determine or estimate the proportions of, or the probabilities that a randomly selected subject performs better, as well as, or worse at a given point in time or space than at another given point in time or space or under one condition compared with another condition, which are additional advantage that provides further useful information that may guide the introduction of any desired interventionist remedial measures.

3. Application

We here illustrate the application of these eight alternative methods for the analysis of paired (matched) sample data with two data sets as well as using simulation. The first are ordinal non-numeric score and the second are numeric scores as follows:

1) A health insurance company every year assesses the vital signs of its clients for the purpose of determining the annual insurance premium payable. In this process the company scores its clients from A^+ (excellent health) through C (fair health) down to F (poorest health-fail), persons with excellent health pay the lowest annual health premium while clients with very poor score pay the highest annual premium. A sample of the scores earned by a random sample of 15 of the clients of this health insurance company during the past two consecutive years are as follows:

Client No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Year 1 Score	A^-	A^+	D	B^-	A^-	B	F	A^-	A^-	C^+	A^+	E	F	B^+	A^+
Year 2 score	F	A^-	F	E	B	C^+	F	B^+	C	A	B^-	D	E	B^+	C^+

2) A random sample of members of each of 15 newly married couples (husband and wife) are asked to state their preferred family size (desired number of children) with the following results.

Couple No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Husband	4	1	6	1	7	1	4	2	8	5	4	4	5	5	4
Wife	5	5	5	6	5	9	4	6	8	5	4	5	6	6	4

As noted above the data of example (1) being ordinal non-numeric data may only be analysed using modified sign test, using either raw scores (method 6) or ranks (method 7) as shown in **Table 1**.

Interest is to determine whether the median scores by clients are the same for the two years, that is if clients are likely to pay equal insurance premium for each of the two years. To do this using method 6 we have from column 4 of **Table 1** ($u_i(6)$) that $f^+ = 10$, $f^0 = 2$; $f^- = 3$ and w $= 10 - 3 = 7$. Also

$$\hat{\pi}^+ = \frac{10}{15} = 0.667; \quad \hat{\pi}^0 = \frac{2}{15} = 0.133; \quad \hat{\pi}^- = \frac{3}{15} = 0.20$$

Hence

$$Var(w) = (15)(0.667 + 0.20 - (0.667 - 0.20)^2$$
$$= (15)(0.667 - 0.218) = 15(0.649) = 9.735$$

Table 1. Analysis of health Insurance data (ordinal non-numeric data) using methods 6 and 7.

Client No.	Year 1 score (x_{i1})	Year 2 score (x_{i2})	$U_i(6)$	Rank of x_{i1} (r_{i1})	Rank of x_{i2} (r_{i2})	Difference between rank $(r_i = r_{i1} - r_{i2})$	$U_i(7)$	$|r_i|\,u_i$	r_i^2
1	A^-	F	+1	K^{+1}	k^{-1}	2	1	2	4
2	A^+	A^-	+1	K^{+1}	k^{-1}	2	1	2	4
3	D	F	+1	K^{+1}	k^{-1}	2	1	2	4
4	B^-	E	+1	K^{+1}	k^{-1}	2	1	2	4
5	A^-	B	+1	K^{+1}	k^{-1}	2	1	2	4
6	B	C^+	+1	K^{+1}	k^{-1}	2	1	2	4
7	F	F	0	K	K	0	0	0	0
8	A^-	B^+	+1	K^{+1}	k^{-1}	2	1	2	4
9	A^-	C	+1	K^{+1}	k^{-1}	2	1	2	4
10	C^+	A	−1	K^{-1}	K^{+1}	-2	−1	−2	4
11	A^+	B^-	+1	K^{+1}	k^{-1}	2	1	2	4
12	E	D	−1	K^{-1}	K^{+1}	-2	−1	−2	4
13	F	E	−1	K^{-1}	K^{+1}	-2	−1	−2	4
14	B^+	B^+	0	K	K	0	0	0	0
15	A^+	C^+	+1	K^{+1}	k^{-1}	2	1	2	4
Total								14	52

$$\chi^2 = \frac{(7)^2}{9.735} = \frac{49}{9.735} = 5.033 \quad (\text{P-value} = 0.0249)$$

which with 1 degree of freedom is highly statistically significant.

Now using the modified sign test by ranks we have from column 9 of **Table 1** that $W = 20 - 6 = 14$; Also from column 10 we have that

$$Var(w) = 52\left(0.667 + 0.20 - (0.667 - 0.20)^2\right)$$
$$= 52(0.649) = 33.748$$

Hence the corresponding test statistic is

$$\chi^2 = \frac{(14)^2}{33.748} = \frac{196}{33.748} = 5.808 \quad (\text{P-value} = 0.0160)$$

which is statistically significant and infact more powerful than the modified sign test of Section 2.6 that depends only on raw scores but not on the ranks of these scores. To illustrate the application of each of the seven methods to numeric data we use example 2 first.

From **Table 2** we have that

$$\bar{d} = \frac{-22}{15} = -1.47; Var(d) = \frac{130}{14} = 9.286$$

$$\therefore Var(\bar{d}) = \frac{9.286}{15} = 0.619$$

$$Se(\bar{d}) = \sqrt{0.619} = 0.787$$

The test statistic for the null hypothesis of equal population medians (H_0: $d_0 = 0$) is

$$t = \frac{\bar{d} - 0}{Se(\bar{d})} = \frac{-1.47}{0.787} = -1.868 \quad (\text{P-value} = 0.0828)$$

which with 14 degrees of freedom is not statistically significant; showing that husbands and wives tend to prefer the same family sizes, that is, desire the same family sizes or number of children.

Analysis using the ordinary sign test, exact binomial test and its normal approximation are presented in **Table 3**.

Now note that the number of ties (0) is 5. Hence the effective sample size is $15 - 5 = 10$. Also the number of 1s(+) is $f^+ = w = 2$. Also $Var(w) = n\theta_0(1 - \theta_0)$ Hence under H_0 (H_0: $\theta = \theta_0 = 0.5$) we have

$$Var(w) = 10(0.5)(0.5) = 2.5$$

$$\therefore \chi^2 = \frac{(W - 0.5n)^2}{Var(w)} = \frac{(2-5)^2}{2.5} = \frac{9}{2.5} = 3.6$$

$$(\text{P-value} = 0.0578)$$

which with 1 degree of freedom is not statistically significant.

3.1. Exact Binomial Test

An equivalent approach to the ordinary sign test for these

OJS

Table 2. Paired sample "t" test for the analysis of family size differences by a random sample of husbands and wife couples.

Couple	Husband (x_{i1})	Wife (x_{i2})	Diff. $d_i = x_{i1} - x_{i2}$	D_i^2
1	4	5	−1	1
2	1	5	−4	16
3	6	5	1	1
4	1	6	−5	25
5	7	5	2	4
6	1	9	−8	64
7	4	4	0	0
8	2	6	−4	16
9	8	8	0	0
10	5	5	0	0
11	4	4	0	0
12	4	5	−1	1
13	5	6	−1	1
14	5	6	−1	1
15	4	4	0	0
Total			3 − 25 = −22	130

Table 3. Application of the ordinary sign test and other two to the data on family size preferences by husbands and wives.

Couple	Husband (x_{i1})	Wife (x_{i2})	Diff. $d_i = x_{i1} - x_{i2}$	u_i^2
1	4	5	−1	0
2	1	5	−4	0
3	6	5	1	1
4	1	6	−5	0
5	7	5	2	1
6	1	9	−8	0
7	4	4	0	-
8	2	6	−4	0
9	8	8	0	-
10	5	5	0	-
11	4	4	0	-
12	4	5	−1	0
13	5	6	−1	0
14	5	6	−1	0
15	4	4	0	-

data is the exact binomial test with $x = 2$, $n = 10$, and $\theta = \theta_0 = 0.5$. Hence

$$P(X \le 2) = \sum_{k=0}^{2} \binom{10}{k}(0.5)^{10} = (1 + 10 + 45)(0.000977)$$

$$= (56)(0.000977) = 0.0547$$

Since $P = 0.0547 > 0.05$, we do not reject the null hypothesis of equal population medians. That is with the exact method we may still conclude that newly married husbands and wives do not differ in their preferred or desired family sizes.

3.2. Normal Approximation

The normal approximation to the exact binomial test for the present data again with $x = 2$; $n = 10$ and $\theta = \theta_0 = 0.5$ is, with correction for continuity

$$\chi^2 = \frac{((2 + 0.5) - 10(0.5))^2}{10(0.5)(0.5)} = \frac{(2.5 - 5)^2}{2.5} = 2.50$$

Or in terms of the normal z-score we have

$$z = \frac{2.5 - 5}{0.5\sqrt{10}} = \frac{-2.5}{1.581} = -1.581 \quad (\text{P-value} = 0.1139)$$

which is also not statistically significant.

Analysis of example 2 using Wilcoxons signed rank sum test is presented in **Table 4**.

3.3. Unmodified Wilcoxons Signed Rank Sum Test

The sum of the ranks of absolute differences with positive signs ignoring zero differences is

$$3 + 6 = 9 = T^+$$

$$E(T^+) = \frac{10(10 + 1)}{4} = \frac{110}{4} = 27.5$$

and

$$Var(T^+) = \frac{10(11)(21)}{24} = \frac{2310}{24} = 96.25$$

Therefore (P-value = 0.0625)

With 1 degree of freedom which is not statistically significant, leading to an acceptance of the null hypothesis of equal family size desires by newly married husbands and wives.

Now from column 5 of **Table 5** (u_{i6}) we have that $f^+ = 2, f^0 = 5; f^- = 8$ and $w = f^+ - f^- = 2 - 8 = -6$. Also

$$\hat{\pi}^+ = \frac{2}{15} = 0.133; \quad \hat{\pi}^0 = \frac{5}{15} = 0.333; \quad \hat{\pi}^- = \frac{8}{15} = 0.533$$

Therefore

$$Var(w) = (15)(0.133 + 0.533 - (0.133 - 0.533)^2)$$
$$= (15)(0.666 - 0.160) = 15(0.506) = 7.59$$

Table 4. Analysis of data on family size, preferences by couples using Walloons signed rank sum test.

Couple	Husbands	Wife	$d_i = x_{i1} - x_{i2}$	$|d_i|$	Ranks Absolute Difference Omitting Zero	Rank of Absolute Differences Including Zeros
1	4	5	−1	1	3	8
2	1	5	−4	4	7.5	12.5
3	6	5	1	1	3	8
4	1	6	−5	5	9	14
5	7	5	2	2	6	11
6	1	9	−8	8	10	15
7	4	4	0	0	-	3
8	2	6	−4	4	7.5	12.5
9	8	8	0	0	-	3
10	5	5	0	0	-	3
11	4	4	0	0	-	3
12	4	5	−1	1	3	8
13	5	6	−1	1	3	8
14	5	6	−1	1	3	8
15	4	4	0	0	-	3

Table 5. Analysis of family size preferences by couples using modified sign test.

Couples	Husband (x_{i1})	Wife (x_{i2})	$d_i = x_{i1} - x_{i2}$	U_i (6)	Rank of x_{i1} (r_{i1})	Rank of x_{i2} (r_{i2})	Difference between rank $(r_i = r_{i1} - r_{i2})$	U_i (7)	$\lvert r_i \rvert u_i$	r_i^2
1	4	5	−1	−1	K^{-1}	K^{+1}	−2	−1	−2	4
2	1	5	−4	−1	K^{-1}	K^{+1}	−2	−1	−2	4
3	6	5	1	1	K^{+1}	k^{-1}	2	1	2	4
4	1	6	−5	−1	K^{-1}	K^{+1}	−2	−1	−2	4
5	7	5	2	1	K^{+1}	k^{-1}	2	1	2	4
6	1	9	−8	−1	K^{-1}	K^{+1}	−2	−1	−2	4
7	4	4	0	0	K	K	0	0	0	0
8	2	6	−4	−1	K^{-1}	K^{+1}	−2	−1	−2	4
9	8	8	0	0	K	K	0	0	0	0
10	5	5	0	0	K	K	0	0	0	0
11	4	4	0	0	K	K	0	0	0	0
12	4	5	−1	−1	K^{-1}	K^{+1}	−2	−1	−2	4
13	5	6	−1	−1	K^{-1}	K^{+1}	−2	−1	−2	4
14	5	6	−1	−1	K^{-1}	K^{+1}	−2	−1	−2	4
15	4	4	0	0	K	K	0	0	0	0
Total									**−12**	**40**

Therefore

$$\chi^2 = \frac{(-6)^2}{7.59} = \frac{36}{7.59} = 4.743$$

(P-value = 0.0294) or which with 1 degree of freedom is highly statistically significant now indicating that newly married husbands and wives do differ in their preferred or desired family sizes.

Now to apply the modified sign test by ranks to same data we have from column 10 of **Table 5** that $W = 4 - 16 = -12$; Also from column 11 of the table we have that

$$Var(w) = 40\left(0.133 + 0.533 - (0.133 - 0.533)^2\right)$$
$$= 40(0.506) = 20.240$$

Hence the corresponding test statistic is

$$\chi^2 = \frac{(-12)^2}{20.240} = \frac{144}{20.240} = 7.115 \quad \text{(P-value = 0.0076)}$$

which is highly significant.

Note that from the P-values and the associated chi-square values that the ordinary sign tests and the un-modified Wilcoxons sign rank sum test are likely to accept a false null hypothesis (Type II error) more frequently than the two type of modified signed tests (methods 6 and 7). The relative efficiency of the modified signed test w to the unmodified Wilcoxons signed

rank sum test T^+ for the present data is

$$RE(W : T^+) = \frac{96.25}{7.59} = 12.681 ,$$

showing that at least for the present data the modified sign tests are much more powerful than the unmodified Wilcoxon signed rank sum test.

The problem with the ordinary sign test and the un-modified Wilcoxon signed rank sum test is that non of the two adjusts or modifies the test statistics for the possible presence of tied observations between sampled populations, and simply ignores these ties if they occur, a procedure that because it uses less information tends to compromise the associated power of the test.

Now reanalyzing the data of example 2 using the modified Wilcoxon signed rank sum test of Section 2.8, we have from column 7 of **Table 4** that $T = 19 - 86 = -67$. Also $f^+ = 2, f^0 = 5; f^- = 8$

$$\hat{\pi}^+ = \frac{2}{15} = 0.133; \quad \hat{\pi}^0 = \frac{5}{15} = 0.333; \quad \hat{\pi}^- = \frac{8}{15} = 0.533$$

And

$$Var(T) = \frac{15(16)(31)}{6}\left(0.133 + 0.533 - (0.133 - 0.533)^2\right)$$
$$= 1240(0.666 - 0.160) = 1240(0.506) = 627.44$$

Now under the null hypothesis of equal population me-

dians $\left(H_0 : \overset{+}{\pi} - \overset{-}{\pi} = \beta_0 = 0 \right)$, the test statistic for the modified Wilcoxon signed rank sum test for the data becomes

$$\chi^2 = \frac{(-67)^2}{627.44} = \frac{4489}{627.44} = 7.154 \quad \text{(P-value = 0.0075)}$$

which with 1 degree of freedom is highly statistically significant now indicating that newly married husbands and wives differ significantly in their desired family size preferences.

Thus the modified Wilcoxon signed rank sum test is here shown at least for the present data to be the most powerful of the six non parametric statistical methods presented here for the analysis of paired or matched sample data. This is because this method uses all available information on the data being analyzed including direction and magnitude and also adjusts, that is makes provision, for the presence of any possible tied observations between the sampled populations.

Using simulation, the result is as shown in **Table 6**.

From **Table 6** we have that

$$\bar{d} = \frac{-33}{15} = -2.20; Var(d) = \frac{161}{14} = 11.50$$

$$\therefore Var(\bar{d}) = \frac{11.50}{15} = 0.767$$

$$Se(\bar{d}) = \sqrt{0.767} = 0.876$$

The test statistic for the null hypothesis of equal population medians ($H_0: d_0 = 0$) is

$$t = \frac{\bar{d} - 0}{Se(\bar{d})} = \frac{-2.20}{0.876} = -2.511 \quad \text{(P-value = 0.0249)}$$

which with 14 degrees of freedom is statistically significant; showing that wife and husband differ in their choices.

Analysis using the ordinary sign test, exact binomial test and its normal approximation are presented in **Table 7**.

Now note that the number of ties (0) is 1. Hence the effective sample size is $15 - 1 = 14$. Also the number of 1s(+) is $f^+ = w = 2$. Also $Var(w) = n\theta_0(1 - \theta_0)$. Hence under H_0 ($H_0: \theta = \theta_0 = 0.5$) we have

$$Var(w) = 14(0.5)(0.5) = 3.50$$

$$\therefore \chi^2 = \frac{(W - 0.5n)^2}{Var(w)} = \frac{(2 - 7)^2}{3.50} = \frac{25}{3.50} = 7.1429$$

(P-value = 0.0075) which with 1 degree of freedom is highly statistically significant.

3.4. Exact Binomial Test

An equivalent approach to the ordinary sign test for these data is the exact binomial test with $x = 2$, $n = 14$, and $\theta = \theta_0 = 0.5$. Hence

$$P(X \le 2) = \sum_{k=0}^{2} \binom{14}{k} (0.5)^{14} = (1 + 14 + 91)(0.000061)$$

$$= (106)(0.000061) = 0.0065$$

Since $P = 0.0065 < 0.05$, we therefore reject the null hypothesis of equal population medians. That is with the exact method we may still conclude that husbands and wives differ in their preferences.

3.5. Normal Approximation

The normal approximation to the exact binomial test for the present data again with $x = 2$; $n = 14$ and $\theta = \theta_0 = 0.5$ is, with correction for continuity

$$\chi^2 = \frac{((2 + 0.5) - 14(0.5))^2}{14(0.5)(0.5)} = \frac{(2.5 - 7.0)^2}{3.5} = 5.786$$

Or in terms of the normal z-score we have

$$z = \frac{2.5 - 7.0}{0.5\sqrt{14}} = \frac{-4.5}{1.871} = -2.405 \quad \text{(P-value = 0.0143)}$$

which is also statistically significant.

Analysis of simulated data using Wilcoxons signed rank sum test is presented in **Table 8**.

3.6. Unmodified Wilcoxons Signed Rank Sum Test

The sum of the ranks of absolute differences with positive signs ignoring zero differences is

$$10.5 + 1.5 = 12 = T^+$$

$$E(T^+) = \frac{14(14+1)}{4} = \frac{210}{4} = 52.5$$

and

$$Var(T^+) = \frac{14(15)(29)}{24} = \frac{6090}{24} = 253.75$$

$$\chi^2 = \frac{(-40.5)^2}{253.75} = \frac{1640.25}{253.75} = 6.464$$

Therefore (P-value = 0.0110)

With 1 degree of freedom which is statistically significant, leading to a rejection of the null hypothesis of equal preferences by husbands and wives.

Now from column 5 of **Table 9** (u_{i6}) we have that $f^+ = 2, f^0 = 1; f^- = 12$ and $w = f^+ - f^- = 2 - 12 = -10$. Also

$$\hat{\pi}^+ = \frac{2}{15} = 0.133; \; \hat{\pi}^0 = \frac{1}{15} = 0.667; \; \hat{\pi}^- = \frac{12}{15} = 0.800$$

Therefore

Table 6. Paired sample "t" test for the analysis of family size differences by a simulated random sample of husbands and wife couples.

Couple	Husband (x_{i1})	Wife (x_{i2})	Diff. $d_i = x_{i1} - x_{i2}$	D_i^2
1	1	4	−3	9
2	7	8	−1	1
3	4	6	−2	4
4	4	6	−2	4
5	4	6	−2	4
6	8	4	4	16
7	2	6	−4	16
8	2	5	−3	9
9	2	8	−6	36
10	1	5	−4	16
11	5	9	−4	16
12	4	6	−2	4
13	4	9	−5	25
14	4	4	0	0
15	5	4	1	1
Total			**5 − 38= −33**	**161**

Table 7. Application of the ordinary sign test and other two to the simulated data on family size preferences by husbands and wives.

Couple	Husband (x_{i1})	Wife (x_{i2})	Diff. $d_i = x_{i1} - x_{i2}$	u_i^2
1	1	4	−3	0
2	7	8	−1	0
3	4	6	−2	0
4	4	6	−2	0
5	4	6	−2	0
6	8	4	4	1
7	2	6	−4	0
8	2	5	−3	0
9	2	8	−6	0
10	1	5	−4	0
11	5	9	−4	0
12	4	6	−2	0
13	4	9	−5	0
14	4	4	0	-
15	5	4	1	1

Table 8. Analysis of simulated data on family size, preferences by couples using Walloons signed rank sum test.

Couple	Husbands	Wife	$d_i = x_{i1} - x_{i2}$	$|d_i|$	Ranks Absolute Difference Omitting Zero	Rank of Absolute Differences including Zeros
1	1	4	−3	3	7.5	8.5
2	7	8	−1	1	1.5	2.5
3	4	6	−2	2	4.5	5.5
4	4	6	−2	2	4.5	5.5
5	4	6	−2	2	4.5	5.5
6	8	4	4	4	10.5	11.5
7	2	6	−4	4	10.5	11.5
8	2	5	−3	3	7.5	8.5
9	2	8	−6	6	14	15
10	1	5	−4	4	10.5	11.5
11	5	9	−4	4	10.5	11.5
12	4	6	−2	2	4.5	5.5
13	4	9	−5	5	13	14
14	4	4	0	0	-	1
15	5	4	1	1	1.5	2.5

Table 9. Analysis of simulated family size preferences by couples using modified sign test.

Couples	Husb and (x_{i1})	Wife (x_{i2})	$d_i = x_{i1} - x_{i2}$	$U_i(6)$	Rank of x_{i1} (r_{i1})	Rank of x_{i2} (r_{i2})	Difference between rank ($r_i = r_{i1} - r_{i2}$)	$U_i(7)$	$|r_i| u_i$	r_i^2
1	1	4	−3	−1	K^{-1}	K^{+1}	−2	−1	−2	4
2	7	8	−1	−1	K^{-1}	K^{+1}	−2	−1	−2	4
3	4	6	−2	−1	K^{-1}	K^{+1}	−2	−1	−2	4
4	4	6	−2	−1	K^{-1}	K^{+1}	−2	−1	−2	4
5	4	6	−2	−1	K^{-1}	K^{+1}	−2	−1	−2	4
6	8	4	4	1	K^{+1}	K^{-1}	2	1	2	4
7	2	6	−4	−1	K^{-1}	K^{+1}	−2	−1	−2	4
8	2	5	−3	−1	K^{-1}	K^{+1}	−2	−1	−2	4
9	2	8	−6	−1	K^{-1}	K^{+1}	−2	−1	−2	4
10	1	5	−4	−1	K^{-1}	K^{+1}	−2	−1	−2	4
11	5	9	−4	−1	K^{-1}	K^{+1}	−2	−1	−2	4
12	4	6	−2	−1	K^{-1}	K^{+1}	−2	−1	−2	4
13	4	9	−5	−1	K^{-1}	K^{+1}	−2	−1	−2	4
14	4	4	0	0	K	K	0	0	0	0
15	5	4	1	1	K^{+1}	K^{-1}	2	1	2	4
Total									**−20**	**56**

$$Var(w) = (15)(0.133 + 0.800 - (0.133 - 0.800)^2$$
$$= (15)(0.933 - 0.4449) = 15(0.4881) = 7.3215$$

Therefore

$$\chi^2 = \frac{(-10)^2}{7.95} = \frac{100}{7.3215} = 13.6584 \quad \text{(P-value = 0.00020)}$$

or which with 1 degree of freedom is statistically significant now indicating that husbands and wives preferences differs.

Now to apply the modified sign test by ranks to same data we have from column 10 of **Table 9** that $W = 4 - 24 = -20$; Also from column 11 of the table we have that

$$Var(w) = (56)(0.133 + 0.800 - (0.133 - 0.800)^2$$
$$= (56)(0.933 - 0.4449) = 56(0.4881) = 27.334$$

Hence the corresponding test statistic is

$$\chi^2 = \frac{(-20)^2}{27.334} = \frac{400}{27.334} = 14.634 \quad \text{(P-value = 0.00010)}$$

which is highly statistically significant.

Note that from the P-values and the associated chi-square values that the ordinary sign test and unmodified wilcoxon sign rank sum test have lower P-values than the two type of modified signed tests (methods 6 and 7). The relative efficiency of the modified signed test w to the unmodified Wilcoxons signed rank sum test T^+ for the simulated data is

$$RE(W : T^+) = \frac{253.75}{27.334} = 9.283,$$

showing that also for the simulated data the modified sign tests are much more powerful than the unmodified Wilcoxon signed rank sum test.

The problem with the ordinary sign test and the unmodified Wilcoxon signed rank sum test is that non of the two adjusts or modifies the test statistics for the possible presence of tied observations between sampled populations, and simply ignores these ties if they occur, a procedure that because it uses less information tends to compromise the associated power of the test.

Now reanalyzing the simulated data using the modified Wilcoxon signed rank sum test of Section 2.8, we have from column 7 of **Table 8** that $T = 14 - 105 = -91$. Also $f^+ = 2, f^0 = 1; f^- = 12$

$$\hat{\pi}^+ = \frac{2}{15} = 0.133; \quad \hat{\pi}^0 = \frac{1}{15} = 0.667; \quad \hat{\pi}^- = \frac{12}{15} = 0.800$$

and

$$Var(T) = \frac{15(16)(31)}{6}\left(0.133 + 0.800 - (0.133 - 0.800)^2\right)$$
$$= 1240(0.933 - 0.4449) = 1240(0.4881)$$
$$= 605.244$$

Now under the null hypothesis of equal population medians $\left(H_0 : \overset{+}{\pi} - \overset{-}{\pi} = \beta_0 = 0\right)$, the test statistic for the modified Wilcoxon signed rank sum test for the data becomes

$$\chi^2 = \frac{(-91)^2}{605.244} = \frac{8281}{605.244} = 13.682 \quad \text{(P-value = 0.00019)}$$

which with 1 degree of freedom is highly statistically significant now indicating that couples differ significantly in their preferences.

Thus the modified Wilcoxon signed rank sum test is here shown again to be the second best using the simulated data being the second most powerful of the six non parametric statistical methods presented here for the analysis of paired or matched sample data. This is because this method uses all available information on the data being analyzed including direction and magnitude and also adjusts, that is makes provision, for the presence of any possible tied observations between the sampled populations.

4. Summary and Conclusion

We have in this paper presented and discussed eight alternative methods for the analysis of paired or matched sample data. If the sampled populations satisfy the necessary assumptions of continuity and normality, then the paired sample parametric "t" test becomes the method of choice and should be preferred since it is generally more powerful than most alternative non parametric methods. If however the data being analyzed are not continuous or are ordinal non-numeric measurements, then the modified sign tests using either the raw scores themselves (method 6) or their ranks (method 7) are the only available methods of analysis under the circumstance. If the data are numeric measurements on at least the ordinal scale but not appropriate for analysis using the parametric "t" test, then the modified Wilcoxon signed rank sum test, the modified sign tests by ranks, the modified sign test, the exact binomial or ordinary sign test and its normal approximation should be preferred and used in this order because of their relatively decreasing power, as shown by at least the illustrative examples used here and when reanalyzed using simulation, the modified sign test by ranks, modified Wilcoxon signed rank sum test, the modified sign test, the exact binomial or its ordinary sign test should be preferred and used in this order because of their relative decreasing power which is almost the same with the raw example except that modified sign tests by rank came first using simulation while second using raw data.

Finally each of the proposed methods may be appropriately modified and used to analyse one sample data

simply by setting values or scores from one of the sampled populations equal to some hypothesized value of a measure of central tendency.

REFERENCES

[1] I. C. A. Oyeka, "An Introduction to Applied Statistical Methods," 8th Edition, Nobern Avocation Publishing Company, Enugu, 2009.

[2] S. Siegel, "Non-Parametric Statistics for the Behavioral Sciences," McGraw-Hill Series in Psychology, New York, 1956.

[3] J D. Gibbons, "Non-Parametric Statistical Inference," McGraw Hill, New York, 1971.

[4] I. C. A. Oyeka, C. E. Utazi, C. R. Nwosu, P. A. Ikpegbu, G. U. Ebuh, H. O. Ilouno and C. C. Nwankwo, "Method of Analysing Paired Data Intrinsically Adjusted for Ties," *Global Journal of Mathematics*, Vol. 1, No. 1, 2009, pp. 1-6.

[5] I. C. A. Oyeka and G. U. Ebuh, "Modified Wilcoxon Signed Rank Sum Test," *Open Journal of Statistics*, Vol. 2, No. 2, 2012, pp. 172-176. doi:10.4236/ojs.2012.22019

Appendix 1: A Summary of Eight Alternative Test Statistics for the Analysis of Paired Samples

S/N	Method	Statistic	Variance (under H_0)	Test Statistic (under H_0)	Assumption	Comments
1	Parametric Test	$\bar{d}$ (mean of sample difference)	$\dfrac{s_{d2}}{n}$	$t = \dfrac{\bar{d} - d_0}{s\sqrt{n}}$	Populations continuous and normally distributed; measurements on the ratio scale	Most powerful if necessary assumptions are satisfied.
2	Ordinary Sign Test	W(=No. of 1's or −1's)	$n\theta_0(1 - \theta_0)$	$\chi^2 = \dfrac{(w - n\theta_0)^2}{n\theta_0(1 - \theta_0)}$	Population continuous numeric measurements on at least the ordinary scale	Usually ignores tied observations uses effective sample size (total number of non-zero differences).
3	Exact Binomial Test	$W = X' = x$ (=No. of 1's or −1's)		$P(x \leq 2) = \sum_{k=0}^{2} \binom{n}{k} \theta_0^k (1 - \theta_0)^{n-k}$	Population numeric and discrete	
4	Normal Approximation to the Sign Test	W(=No. of 1's or −1's)	$\sqrt{n\theta_0(1 - \theta_0)}$	$t = \dfrac{((w \pm 0.5) - n\theta_0)}{\sqrt{n\theta_0(1 - \theta_0)}}$	Population discrete "n" sufficiently large	Incorporates continuity correction.
5	Unmodified Wilcoxon Signed Rank Sum Test	T^+ = (sum of the ranks of absolute differences with positives sign).	$\dfrac{n(n+1)(2n+1)}{6}$ $\theta(1 - \theta)$	$\chi_u^2 = \dfrac{\left(T^+ - \dfrac{n(n+1)}{2}\theta_0\right)^2}{\dfrac{n(n+1)(2n+1)}{6}\theta_0(1 - \theta_0)}$	Population continuous numeric measurement on at least the ordinal scale	Ignores zero absolute differences; does not provide for tied observations between samples.
6	Modified Sign Test	W = (difference between total number of 1s and −1s	$n\left(\hat{\pi}^+ + \hat{\pi}^- - \left(\hat{\pi}^+ - \hat{\pi}^-\right)^2\right)$	$\chi^2 = \dfrac{\left(w - n\left(\hat{\pi}^+ - \hat{\pi}^-\right)\right)^2}{n\left(\hat{\pi}^+ + \hat{\pi}^- - \left(\hat{\pi}^+ - \hat{\pi}^-\right)^2\right)}$	Population may be non numeric measurement on at least the ordinal scale	May be used with both numeric and non numeric measurements on at least the ordinal scale. Intrinsically adjusted for any possible tied observations between sampled population more powerful than the unmodified Wilcoxon signed rank sum test.
7	Modified Sign Test by Ranks	$W = R_{.1} - R_{.2}$	$\left(r_1^2 + r_2^2 - 2\left(n\left(k^2 - 1\right) + t\right)\right)$ $\left(\hat{\pi}^+ + \hat{\pi}^- - \left(\hat{\pi}^+ - \hat{\pi}^-\right)^2\right)$ $= 4(n - t)$ $\left(\hat{\pi}^+ + \hat{\pi}^- - \left(\hat{\pi}^+ - \hat{\pi}^-\right)^2\right)$	χ^2 $= \dfrac{w^2}{\left(r_1^2 + r_2^2 - 2\left(n\left(k^2 - 1\right) + t\right)\right)\left(\hat{\pi}^+ + \hat{\pi}^- - \left(\hat{\pi}^+ - \hat{\pi}^-\right)^2\right)}$ $= \dfrac{w^2}{4(n - t)\left(\hat{\pi}^+ + \hat{\pi}^- - \left(\hat{\pi}^+ - \hat{\pi}^-\right)^2\right)}$	Population may be non numeric measurement on at least the ordinal scale	Same as No. 6 except that if uses the ranks of the paired observations rather than the observations themselves; may be used for numeric and non numeric measurements on at least the ordinal scale, intrinsically adjusted for ties, also more powerful than the unmodified Wilcoxon Signed Rank Sum Test.

Continued

| 8 | Modified Wilcoxon Signed Rank Sum Test | T (=difference between the sum of ranks of absolute differences with positive signs and the sum of the ranks of absolute differences with negative signs) | $$\frac{n(n+1)(2n+1)}{6}\left(\pi^{+}+\pi^{-}-\left(\pi^{+}-\pi^{-}\right)^{2}\right)$$ | $$\chi^{2}=\frac{\left(T-\frac{n(n+1)}{2}\cdot\left(\pi^{+}-\pi^{-}\right)\right)^{2}}{\frac{n(n+1)(2n+1)}{6}\left(\hat{\pi}^{+}+\hat{\pi}^{-}-\left(\hat{\pi}^{+}-\hat{\pi}^{-}\right)^{2}\right)}$$ | Population numeric measurement on at least the ordinal scale | Intrinsically adjusted for any possible tied observations between sampled populations, if applicable is the most powerful of all the non-parametric tests presented here. |