A different view to two paired sample researches and an application

Çetin Ayhan Seyfullahoğulları
(Vocational School of Social Sciences, Marmara University, Istanbul 34722, Turkey)

Abstract: Hypothesis test is a function of decision making. With samples data, the author has got knowledge about population. It analysis the difference between two dependent or independent samples significance or not. Hypothesis test measures the difference of estimation from real values. This article resarches probability of these differences. In this study, difference and similarities between nonparametric Wilcoxon $t$ test and parametric $t$ test will be searched, and application is about basic rights and freedom of the students. These findings are in line, the normalization of the data due to dissolution of the test theory as non-parametric implementation, will be more even in terms of statistical theory. However, the implementation of parametric test does not show the absolutely wrong to deliver results.

Key words: $t$ test; Wilcoxon $t$ test

1. Introduction

Broad sense population, all units in a research framework defined in the cluster is formed. The numerical properties that set the characteristics of the population is called parameter. For example population, average of variance or standard deviation of population parameter is shown as an example. Variance of population, requires some assumptions that we should make relation with the hypothesis tests about the average, about parameters of population (Glass & Hopkins, 1984).

Some assumptions for parametric tests are applied (Kalaycı, 2008). These assumptions are often assumptions for population to know the distribution (normal, binomial, uniform distribution, etc.). Calculating confidence intervals about the average of population or hypothesis tests based on methods for making $t$ divided to normal or near normal as the division population is assumed. Similarly, the regression analysis might be able to generalize the normal distribution with two variables assumed need to have assumed (Mills, 1987). Appropriate and robust parametric techniques for data would be more accurate to use.

Nonparametric tests do not require such strict requirements, even for parameters, which do not reveal population assumptions about the distribution. Nonparametric tests, effective parametric tests are less sensitive and therefore the differences between existing groups may be insufficient to find (Kalaycı, 2008). Nonparametric test for nominal (categorical) and ordinal (ordered) scaled data are used. Parametric test is suitable for the continuous data.

For the purpose of this study, the conditions applied is not right nonparametric test (Wilcoxon test), instead, the $t$ test in parametric test is applied, to show similar results which can be obtained. In this study secondly, parametric $t$ test and $t$ test in which nonparametrik been described as the Wilcoxon test theory, and then thirdly,
similarities and differences between the two tests in the view of the fundamental rights and freedoms of students to address an application were examined by the expansion.

2. t Test

t distribution with zero average and the single-mode distribution is symmetrical. This distribution is similar to the shape of the normal distribution, while the shape variability is greater. The sample size (n), while the standard deviation is smaller and closer to the t distribution, is the standard normal distribution (Hamburg, 1987). There are three types of tests. The single-group t test (one-sample t test), an independent test of the differences between the two groups (independent samples t test), compared the two groups (paired-samples t test) between the t test of the difference is to be investigated (Altunışık, Coşkun, Yıldırım & Bayraktaroğlu, 2002).

This study examined the application of variable under different conditions determined by differences in the reactions used for the paired t test with two groups nonparametric that taken account of Wilcoxon t test, and thus only these two tests is not the theoretical content.

2.1 Parametric paired two-sample t test

Dependent sample t test is used to investigate whether there is or not difference between the average for two sample groups. The average is calculated before the events which occurred after the observation of an increase or decrease for the average which had been searched. There are no two different sample groups. Applying the same sample group (training, treatment, etc.), process analysis on the average before and after the hypothesis is tested.

Dependent t test, a group of the unit at different times, different modes of operation from two data sets the difference in action before and after the resulting values of the difference or the twins different from the values of the differences between zero, with the average of the random samples which will be test using a certain method (Özdamar, 2004).

Hypothesis two-way is established:

\[ H_0 : \mu_1 = \mu_2 \]
\[ H_1 : \mu_1 \neq \mu_2 \]

Usually, after the application process, an increase or decrease is expected on average. For example, a training class, the next test is expected to increase in the success, or a group of patients are treated, blood pressure and cholesterol are expected to fall. In this case, it will be appropriate to establish two-way hypothesis.

Growth is expected:

\[ H_0 : \mu_1 = \mu_2 \]
\[ H_1 : \mu_1 < \mu_2 \]

Decrease is expected:

\[ H_0 : \mu_1 = \mu_2 \]
\[ H_1 : \mu_1 > \mu_2 \]

should be established.

Test statistics: \[ t = \frac{\bar{X}_x - \bar{X}_y}{S_x} \]

In this formula, the average of the difference, S is the standard deviation of differences. This formula is derived from the value of t test statistics, t table, n-1 (observation-1) degrees of freedom and significance level of
$\alpha$ values is found by looking at the hypothesis against the acceptance or rejection of the given decision.

### 2.2 Wilcoxon test

Wilcoxon test, two-sample test is dependent. Nonparametric paired two-sample $t$ test is alternative. Unlike the parametric $t$ test, the comparison instead of average, Wilcoxon test, the value to sort and compare two different time zones and converts between the two time periods, for a change in whether the test (Heiman, 1996).

Testing will be used to perform the following steps to obtain the criteria is followed:
1. Each $D_i = X_i - Y_i$ difference values are calculated.
2. Take the absolute value of the difference from small to large is the right order.
3. Results rank each of the absolute value of the difference signal is assigned to this rank.
4. $T^+$: positive signs rank sum; $T^-$: negative signs are obtained as the total rank; $T^+$ or $T^-$, depending on the alternative hypothesis is the test criteria.

Hypothesis is two sided; $T^+$ or $T^-$, the value of which is a small test criterion ($T$) is used as (Turanlı, 1988).

One-sided hypothesis, and greater than (>) is, $T^-$ is used as criteria for testing.

One sided of hypothesis and less than (<) are $T^+$ value is used as the test criteria. Ordinal table is used critical values for the determination of the critical value of Wilcoxon sign.

### 3. Research and methods

#### 3.1 Research methodology

#### 3.1.1 Purpose of the study

The general, trend in the direction of the implementation of parametric methods are used in order to apply statistical methods. However, in some cases, certain assumptions and conditions are not right, alternative methods are applied. In this study, parametric statistical methods with the paired $t$ test, non-parametric Wilcoxon $t$ test statistical methods will be reviewed. Considering similarities and differences between the two tests, a sample application is to be revealed by investigation.

#### 3.1.2 Preparation of survey questions

Used for application data, fundamental rights and freedoms of students to address 21 questions of perspective that attitude in the scale consists of basic rights and freedoms. 21 questions were a variety of statistical and technical content, according to the filters which are divided into 4 factors:

1. Basic rights and freedoms awareness factor,
2. The use of fundamental rights and freedoms factor,
3. Limits of fundamental rights and freedoms,
4. Mutual respect.

#### 3.1.3 Scope of the survey

Questions, each of the SPSS environments as a variable range (scale) level is measured. Numeric variable types (numeric) is defined as value. The observations of 103 high school students are included. Students complete the first questionnaire of 21 questions and received answers are the first data was obtained. Then students were given training on basic rights and freedoms, and the various documents and sources of support and training will
increase the students’ rights to information about the topic. After 21 questions scale of this application is completed again and the received answer was obtained from the latest data. Reliability of 0.919 is applied before the scale; then, the reliability of the training scale was found to be 0.694.

3.1.4 Data analysis

This application of the same observation before and after the application was made in the form of two between the two applications and progresses an intervention such as the recorder, because it is perceived as paired samples. SPSS 15.0 was implemented to help the research on the process of statistical analysis. The results obtained from the table in the comments section. Research answers to all questions to the participants, missing or lost will not be observed. P = 0.05 significance level to investigate all the findings and has been tested with one-way. One-sided test is the reason, and education is expected to increase as the application, then the theory.

3.1.5 Previous studies

Parametric paired t test and non-parametric Wilcoxon test with the use of work in the field of medicine is focused in particularly. Because of the existing data set in this field, patient-control group (later) in the form of two pre-implementation requires the test. Coronary angiography in patients with the procedure before the substances captopril contrast effect on nefropaticy to investigate; coronary angiography before the captopril group 5 patients (8.3%), control group 1 patients (3%) nefropaticy contrast materials was developed, and this difference, both tests were used (t test and Wilcoxon test) in order to test. The result was significant and it is mentioned that (Toprak, Cirit, Bayata, Yeşil & Aslan, 2003). University of California Hodges, J. L. and Lehmann, E. L.—The efficiency of some of the t test, nonparametric competitors named in the article, t test results obtained with the Wilcoxon test will be approximately. However, in the field of medicine to Cazeau, Leclercq and others published in 2001 New England Journal of Medicine—Effects of multisite pacing in patients with biventricular heart failure and intraventricular, a t test or Wilcoxon test was used in the article studies.

3.2 Normal distribution test

Before applying the data analysis, the normal distribution should be checked. Kolmogorov-Smirnov Z statistics for testing the normal distribution is used for. Used for the normal distribution hypothesis:

\[ H_0 : \text{The data is normally dispersed.} \]
\[ H_1 : \text{The data is not normal dispersion.} \]

If the data is scattered normal parametric t test, nonparametric Wilcoxon signed rank test normally does not disintegrate that will be appropriate for application.

| Table 1 Kolmogorov-Smirnov normal distribution table |
|------------------------------------------|-----------------|-----------------|
| Awareness of fundamental rights and freedoms (before) | 103 | 1.590 | 0.013 |
| Use of fundamental rights and freedoms (before) | 103 | 1.479 | 0.025 |
| Limits of fundamental rights and freedoms (before) | 103 | 2.277 | 0.000 |
| Mutual respect factor (before) | 103 | 3.405 | 0.000 |
| Awareness of fundamental rights and freedoms (after) | 103 | 3.509 | 0.000 |
| Use of fundamental rights and freedoms (after) | 103 | 3.611 | 0.000 |
| Limits of fundamental rights and freedoms (after) | 103 | 2.252 | 0.000 |
| Mutual respect factor (after) | 103 | 4.403 | 0.000 |
Kolmogorov-Smirnov Z, as seen in Table 1 for statistical significance of 0.05 for all, because under the hypothesis $H_0$ is rejected. The data is not normal dispersion. However, this study analyzes the difference between the two assumptions of this review which like terms parametric, and parametric and nonparametric test was applied and differences in practice were examined.

Hypothesis 1

$H_0 : \mu_1 = \mu_2$

$H_0 : \text{Awareness of fundamental rights and freedoms after training has not changed.}$

$H_1 : \mu_1 < \mu_2$

$H_1 : \text{Awareness of fundamental rights and freedoms after training has increased.}$

Hypothesis 2

$H_0 : \mu_1 = \mu_2$

$H_0 : \text{Educational use of the fundamental rights and freedoms has not changed.}$

$H_1 : \mu_1 < \mu_2$

$H_1 : \text{Educational use of the fundamental rights and freedoms has increased.}$

Hypothesis 3

$H_0 : \mu_1 = \mu_2$

$H_0 : \text{Limits of fundamental rights and freedoms after training has not changed.}$

$H_1 : \mu_1 < \mu_2$

$H_1 : \text{Limits of fundamental rights and freedoms after training has increased.}$

Hypothesis 4

$H_0 : \mu_1 = \mu_2$

$H_0 : \text{Mutual respect has not changed after training.}$

$H_1 : \mu_1 < \mu_2$

$H_1 : \text{Mutual respect has increased after training.}$

3.3 Parametric paired two-sample $t$ test

| Table 2 | Paired two-sample $t$ test table |
|---------|----------------------------------|
|         | Average | Number of observations | Standard deviation | $t$ statistic | Degrees of freedom | Significance |
| Awareness of fundamental rights and freedoms (before) | 4.31 | 103 | 0.68 | 8.353 | 102 | 0.000 |
| Awareness of fundamental rights and freedoms (after) | 4.89 | 103 | 0.18 | 8.568 | 102 | 0.000 |
| Use of fundamental rights and freedoms (before) | 4.16 | 103 | 0.76 | 6.483 | 102 | 0.000 |
| Use of fundamental rights and freedoms (after) | 4.85 | 103 | 0.26 | 5.646 | 102 | 0.000 |
| Limits of fundamental rights and freedoms (before) | 4.77 | 103 | 0.78 | -6.483 | 102 | 0.000 |
| Limits of fundamental rights and freedoms (after) | 4.78 | 103 | 0.52 | -5.646 | 102 | 0.000 |

When looking at Table 2, we could see that it concerns the first factor in the scale of the fundamental rights, and freedoms awareness training is 4.31 from the previous average value with the standard deviation of 0.68 for education after the average value of 4.89, standard deviation value was found to be 0.18. Values of $t$ statistics were
A different view to two paired sample researches and an application

found to be \(-8.353\). \(t\) statistic at the significance value 0.000 <0.01 critical value, because \(H_0\) was rejected. After applying the training there has increased awareness of fundamental rights and freedoms.

The second factor in the scale of the fundamental rights and freedoms use the average value of 4.16 before training. The standard deviation is 0.76 for education after the average value of 4.85, and standard deviation value was found to be 0.26. Values of \(t\) statistics were found to be \(-8.568\). \(t\) statistic at the significance value 0.000 <0.01 critical value, because under \(H_0\) was rejected. After applying the training, there has increased use of fundamental rights and freedoms.

The third factor in the scale of fundamental rights and freedoms limits the average value of 4.24 before training. Standard deviation of 0.78 for education after the average value of 4.77, standard deviation value was found to be 0.26. Values of \(t\) statistics were found to be \(-6.483\). \(t\) statistic at the significance value 0.000 <0.01 critical value, \(H_0\) was rejected. Application after the training has increased boundaries of fundamental rights and freedoms.

The fourth factor scale mutual respect factor average value of 4.13 before training, the standard deviation of 1.07 for education after the average value of 4.78, standard deviation value was found to be 0.52. Values of \(t\) statistics were found to be \(-5.646\). \(t\) statistic at the significance value 0.000<0.01 critical value, \(H_0\) was rejected. Mutual respect grew after the applicable training.

3.4 Wilcoxon test

| Table 3 | Wilcoxon test table |
|---------|---------------------|
| Awareness of fundamental rights and freedoms (after)-awareness of fundamental rights and freedoms (before) | | | | | | |
| Number of observations | Average of ranks | Total of ranks | Wilcoxon Z statistic | Significance |
| Negative ranks | 12 | 16.46 | 197.50 | -7.089 | 0.000 |
| Positive ranks | 72 | 46.84 | 3,372.50 |
| Equal value | 19 | |
| Total | 103 | |
| Use of fundamental rights and freedoms (after)-use of fundamental rights and freedoms (before) | | | | | | |
| Number of observations | Average of ranks | Total of ranks | Wilcoxon Z statistic | Significance |
| Negative ranks | 12 | 23.71 | 284.50 | -6.905 | 0.000 |
| Positive ranks | 75 | 47.25 | 3,543.50 |
| Equal value | 16 | |
| Total | 103 | |
| Limits of fundamental rights and freedoms (after)-limits of fundamental rights and freedoms (before) | | | | | | |
| Number of observations | Average of ranks | Total of ranks | Wilcoxon Z statistic | Significance |
| Negative ranks | 23 | 28.09 | 646.00 | -5.650 | 0.000 |
| Positive ranks | 67 | 51.48 | 3,449.00 |
| Equal value | 13 | |
| Total | 103 | |
| Mutual respect factor (after)-mutual respect factor (before) | | | | | | |
| Number of observations | Average of ranks | Total of ranks | Wilcoxon Z statistic | Significance |
| Negative ranks | 12 | 17.00 | 204.00 | -4.973 | 0.000 |
| Positive ranks | 45 | 32.20 | 1,449.00 |
| Equal value | 46 | |
| Total | 103 | |

Table 3 is taken into account in the first scale factor, the fundamental rights and freedoms of the negative factors consciousness rank number 12, the negative rank average is 16.46, its total is 197.50 negative rank, positive rank number is 71, a positive average rank is 46.84, total is 3,372, positive rank as found to be 50. The number 19 has been found to be of equal value. Wilcoxon Z statistic value was found to be -7.089. Wilcoxon Z
statistic at the significance value 0.000 < 0.01 critical value, because under Ho was rejected. Apply after the training has increased awareness of fundamental rights and freedoms.

Scale is the first factor using fundamental rights and freedoms of the negative factors rank number 12, the negative rank average is 23.71, total is 284.50 negative rank, positive rank number is 75, the average rank of is 47.25 positive, positive rank sum was found to be 3,543.50. The number 16 has been found to be of equal value. Wilcoxon Z statistic value was found to be -6.905. Wilcoxon Z statistic at the significance value 0.000 < 0.01 critical value, H0 was rejected. Application after the training has increased using fundamental rights and freedoms.

Scale is the first factor of fundamental rights and freedoms of the negative factors limits rank number 23, the negative rank average is 28.09, total is 646.00 negative rank, positive rank number is 67, a positive average rank 51.48, positive rank sum was found to be 3,449.00. The number 13 has been found to be of equal value. Wilcoxon Z statistic value was found to be -5.650. Wilcoxon Z statistic at the significance value 0.000 < 0.01 critical value, H0 was rejected. Application after the training has increased boundaries of fundamental rights and freedoms.

Mutual respect is the first factor scale factor negative rank number 12, the negative rank average is 17.00, total is 204.00 negative rank, positive rank number is 45, a positive average rank is 32.20, 1,449.00, which sum has positive rank. The number 46 has been found to be of equal value. Wilcoxon Z statistic value was found to be -4.973. Wilcoxon Z statistic at the significance value 0.000 < 0.01 critical value, H0 was rejected. Mutual respect grew after the applicable training.

4. Results

The data does not show the normal distribution which is applied in both tests. Parametric paired t test results came out at the same 4 factors result in the application after the training as a statistical increase in the average value was determined. Similarly, nonparametric Wilcoxon test results at the same 4 factors out the results, after applying the training, the average increase in value has been determined statistically. Both had the same test results, the hypothesis were rejected in the same way. Fundamental rights and freedoms consciousness factor value for the t statistic -8.353, -7.089 value Wilcoxon statistics; the same direction and is very close to each other. Fundamental rights and the limits of freedom of factor values for the t statistic -8.568, -6.905 value Wilcoxon statistics, the same direction and is much close to each other. Fundamental rights and the limits of freedom of factor values for the t statistic -6.483, -5.650 value Wilcoxon statistics, the same direction and is very close to each other. Mutual respect factor value for the t statistic -5.646, -4.973 values Wilcoxon statistics, the same direction and is very close to each other. Although both the normal dispersion of the test results were very close to each other parametric t test and nonparametric Wilcoxon test of the theoretical formula, the test application form, although different hypotheses and decision criteria are the same. In practice, as seen in this study, two test cases of the same data in the analysis of the implementation are different, but both hypotheses about the test results in the same direction were decided. For other applications, in this regard, an analysis of data due to normal dispersion, the theoretical implementation of the test as a nonparametric statistics will be more even in terms of theory. However, the implementation of parametric test does not show the absolutely wrong to deliver results.

(to be continued on Page 52)