

Wilcoxon test in SPSS (Practical)

Wilcoxon Sign Rank Test practical

In this practical we are going to investigate how to perform a Wilcoxon test using SPSS. A Wilcoxon test is used when we have two interval or ratio level variables measured for a set of observations and we want to test if the distribution is different for the two variables but we are unable to assume normality for one or both of the variables. It can also be to compare ordered categorical variables. It is the non-parametric equivalent of the paired t-test but unlike the t-test it tests differences in the median rather than the mean. The test does not assume any distribution for the two variables. To run a single test in SPSS requires that your dataset has two column containing the two variables to be compared.

The 2015 PISA survey collected information on students' degree of interest in a range of science-related topics. In this example, we will test whether 15 year olds in England are more interested in learning about topics related to the universe and its history (INT_UNIV) or related to how science can help prevent disease (INT_DISEASE). Students were asked to rate their interest in both topics on a four-point Likert scale, with responses ranging from Not interested to Highly interested. Both measures are ordinal variables and each participant contributes a pair of scores to the data, so a non-parametric test of difference is an appropriate method to use to explore differences in the distribution of responses on the two topics.

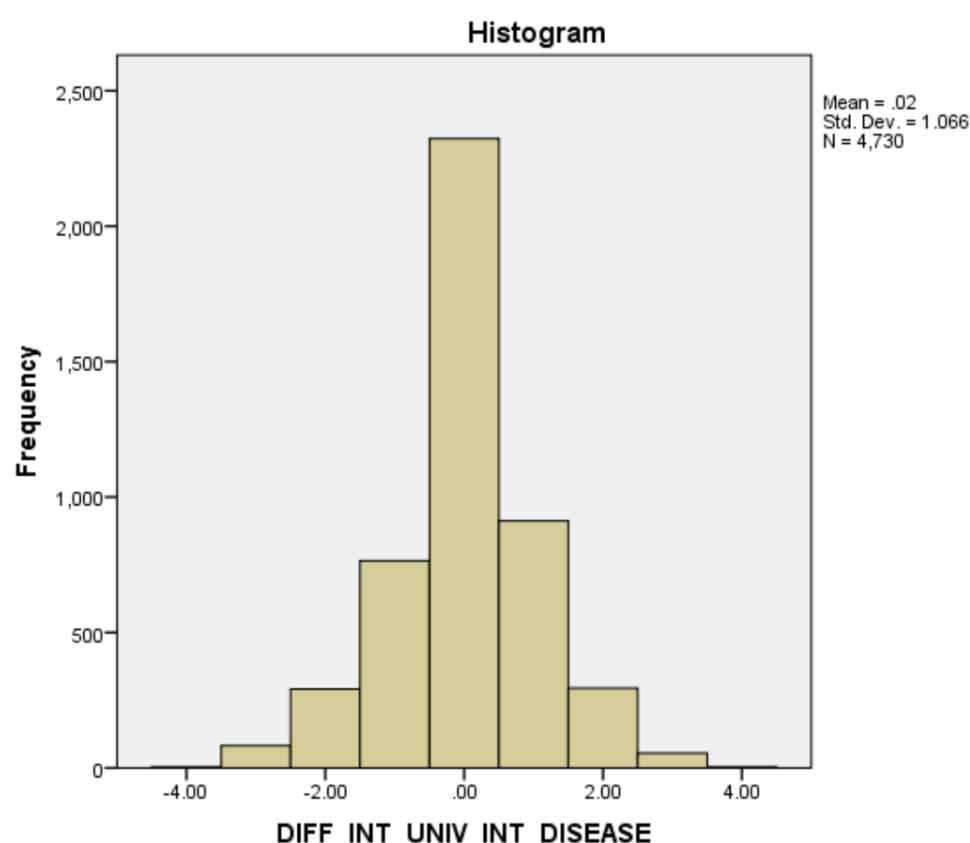
Before we can perform this test we need to check whether the differences between **INT_UNIV** and **INT_DISEASE** are normally distributed. First we need to create a difference variable as follows:

- Select **Compute** from the **Transform** menu.
- Type **DIFF_INT_UNIV_INT_DISEASE** into the **Target Variable** box.
- Type **INT_UNIV - INT_DISEASE** into the **Numeric Expression** box.
- Click on the **OK** button.

We can now use this new generated variable to perform normality checks. Do this as follows:

- Select **Descriptive Statistics** from the **Analyze** menu.
- Select **Explore** from the **Descriptive Statistics** sub-menu.
- Click on the **Reset** button.
- Copy the **DIFF_INT_UNIV_INT_DISEASE** variable into the **Dependent List:** box.
- Click on the **Plots...** button.
- On the screen that appears select the **Histogram** tick box.
- Unselect the **Stem and leaf** button.
- Select the **Normality plots with tests** button.
- Click on the **Continue** button.
- Click on the **OK** button.

We will first look at a histogram of the variable, **DIFF_INT_UNIV_INT_DISEASE**. This can be found in amongst the set of output objects and looks as follows:



Ideally for a normal distribution this histogram should look symmetric around the mean of the distribution, in this case .0152. This distribution appears to be significantly skewed to the left (negatively skewed).

We will next look at a statistical test to see if this backs up our visual impressions from the histogram.

The Kolmogorov-Smirnov test is used to test the null hypothesis that a set of data comes from a Normal distribution.

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
DIFF_INT_UNIV_INT_DISEASE	.253	4730	.000	.903	4730	.000

a. Lilliefors Significance Correction

The Kolmogorov Smirnov test produces test statistics that are used (along with a degrees of freedom parameter) to test for normality. Here we see that the Kolmogorov Smirnov statistic takes value .253. This has degrees of freedom which equals the number of data points, namely 4730.

Here we see that the p value (quoted under Sig. for Kolmogorov Smirnov) is .000 (reported as $p < .001$) which is less than 0.05. We therefore have significant evidence to reject the null hypothesis that the variable follows a normal distribution.

Although the Kolmogorov Smirnov statistic tells the researcher whether the distribution followed by a variable is statistically significantly different from a normal distribution one should take care in not overinterpreting such findings. Significance will be strongly effected by the number of observations and so only a small discrepancy from normality will be deemed significant for very large sample sizes whilst very large discrepancies will be required to reject the null hypothesis for small sample sizes.

SPSS also supplies QQ plots to assist in looking at normality but for brevity we do not show them here.

We will next move on to the Wilcoxon test itself and will test for a difference in distribution between the two variables, **INT_UNIV** and **INT_DISEASE**.

Below you will see instructions on how to perform the Wilcoxon test in SPSS. If you follow the instructions you will see the two tabular outputs that are embedded in the explanations below.

- Select **Non Parametric Tests** from the **Analyze** menu.
- Select **Legacy Dialogs** from the **Non Parametric Tests** sub-menu.
- Select **2 Related Samples...** from the **Legacy Dialogs** sub-menu.
- Click on the **Reset** button.
- Copy the **Interest in universe and its history[INT_UNIV]** variable into the **Test Pairs:** box.
- Copy the **Interest in how to prevent disease[INT_DISEASE]** variable into the **Test Pairs:** box.
- Click on the **Exact...** button.
- On the screen that appears select the **Exact** button.
- Click on the **Continue** button.
- Click on the **OK** button.

The first SPSS output table contains a summary of the rankings for the difference variable. Here observations are split into three types depending on whether the value of **INT_UNIV** is bigger than **INT_DISEASE** (negative ranks), the value of **INT_DISEASE** is bigger than **INT_UNIV** (positive ranks), and finally where both variables take the same value (ties). These can be seen below:

Ranks

		N	Mean Rank	Sum of Ranks
Interest in how to prevent disease - Interest in universe and its history	Negative Ranks	1265 ^a	1171.67	1482167.00
	Positive Ranks	1142 ^b	1239.81	1415861.00
	Ties	2323 ^c		
	Total	4730		

a. Interest in how to prevent disease < Interest in universe and its history

b. Interest in how to prevent disease > Interest in universe and its history

c. Interest in how to prevent disease = Interest in universe and its history

The Wilcoxon test works by firstly assigning a sign (or a tie) to the difference between each pair of observations. Here we have worked on **INT_DISEASE - INT_UNIV** so that positive ranks are when **INT_DISEASE > INT_UNIV**. Here we see that there are 1265 negative ranks, 1142 positive ranks and 2323 ties.

Having worked out which observed pairs result in which sign for their difference, the magnitude (excluding the sign) of these differences is calculated and these are then ranked in order (excluding ties). We now see that the total of the ranks for the negative differences is 1482167.00 resulting in a mean rank of 1171.67 whilst the total of the ranks for the positive differences is 1415861.00 resulting in a mean rank of 1239.81. Here the mean of the positive ranks is larger than that for negative ranks suggesting that values for **INT_DISEASE** are generally larger than for **INT_UNIV**.

The Wilcoxon test will now decide whether this difference in mean ranks is significant or not as is illustrated in the second table.

The second SPSS output table contains details of the test itself and can be seen below:

Test Statistics

Interest in how to prevent disease - Interest in universe and its history	
Z	-1.018 ^b
Asymp. Sig. (2-tailed)	.309
Exact Sig. (2-tailed)	.309
Exact Sig. (1-tailed)	.154
Point Probability	.000

b. Based on positive ranks.

The output here consists of test statistics and their significance as calculated in several ways. We are considering the Wilcoxon statistic which is calculated from the ranks and is not shown explicitly by SPSS but is used to calculate a Z score. Here we see that $Z = -1.018$ and this can be compared with a standard normal distribution to test whether there are significant differences between the groups.

Here we see that the p value (quoted next to Asymp. Sig. (2-tailed)) is .309 which is greater than 0.05 and therefore we cannot reject the null hypothesis that the medians of the two groups are the same. The normal approximation used above is only an approximation to the p value and it is possible to construct the exact p value. This is given in the next row and we see that the exact p value is .309 whilst the asymptotic p value is .309. The exact p value agrees with the asymptotic p value that the null hypothesis cannot be rejected. For completeness the table also gives a p value for a 1-sided test and a point probability but we will ignore these here.

In conclusion, we could report this to a reader as follows:

A comparison of the mean of the distribution of the variables **INT_UNIV** and **INT_DISEASE** was desired but due to the non-normality of the variables a Wilcoxon signed rank test was carried out. The mean of the positive ranks is larger than that for negative ranks suggesting that values for **INT_DISEASE** are generally larger than for **INT_UNIV**. The Wilcoxon Signed rank test results in a Z statistic of -1.018 which results in an exact p value of .309. This is not significant and we cannot reject the null hypothesis of equal medians for the 2 variables.

The results suggest that students are equally interested in learning about topics related to the universe and its history and those related to how science can help prevent disease. It does not appear that, overall, students in England are drawn more to one of these types of scientific inquiry.