

Mann-Whitney tests in SPSS (Practical)

Mann Whitney test practical

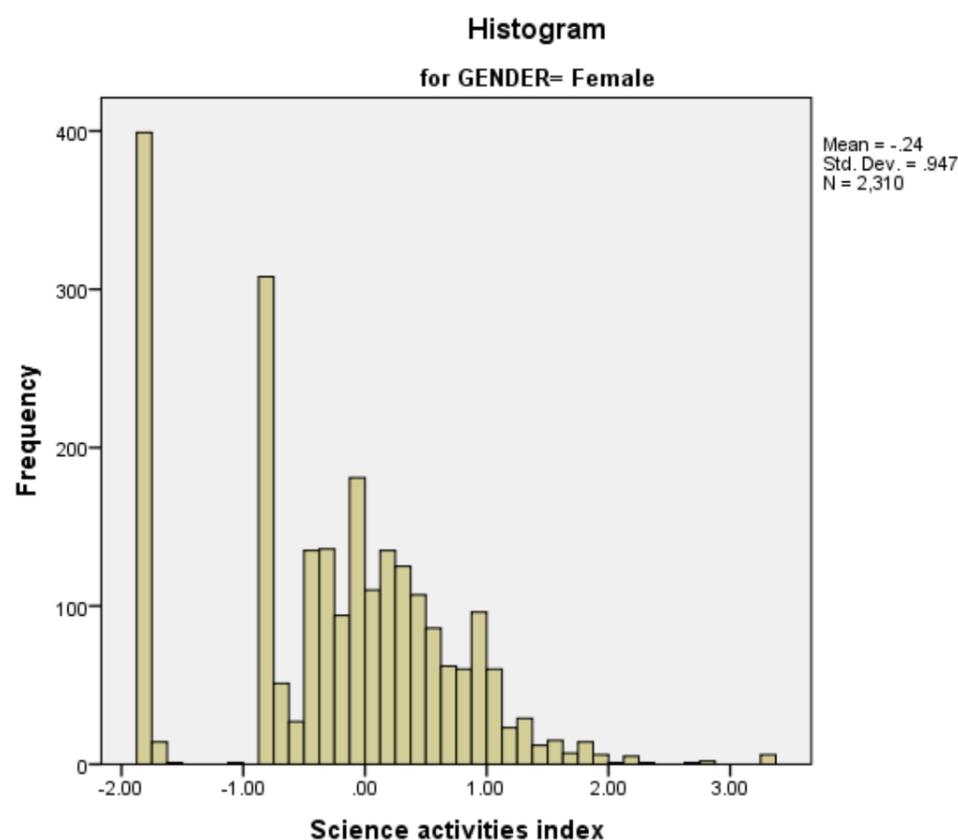
In this practical we are going to investigate how to perform a Mann Whitney test using SPSS. A Mann-Whitney test is used when we have a continuous level variable measured for all observations in two groups and we want to test if the distribution of this variable is different in the two groups but we are unable to assume normality in both groups. It can also be used to compare an ordered categorical variable measured on two groups. It is the non-parametric equivalent of the independent-samples t-test but unlike the t-test it tests for differences in the overall distribution across groups rather than for differences in the mean. The test does not assume any particular distribution for the variable in either of the groups. To run a single test in SPSS requires that your dataset has one column containing the variable to be tested and another column of the same length with the groups identified.

In the example we will explore whether there are gender differences in engagement in out-of-school science activities. The variable **SCIEACT** is a score derived from responses to nine items on how often the student engages in a particular science activity, such as watching TV programmes about science and attending a science club (see PISA datafile description for further details).

As a first step we should generally test for the normality of the variable of interest, **SCIEACT** in each of the two groups that are indicated by the grouping variable, **GENDER**. (In practice you might skip this step if your dependent variable is ordinal rather than continuous, but let's work through it here for completeness.)

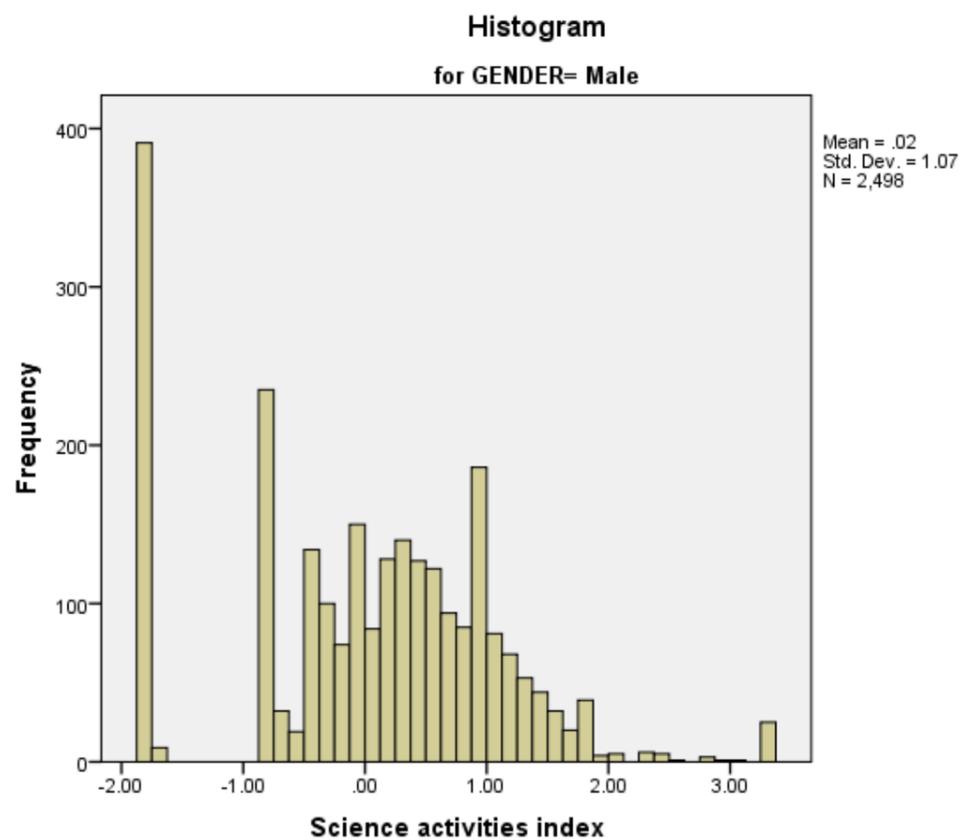
- Select **Descriptive Statistics** from the **Analyze** menu.
- Select **Explore** from the **Descriptive Statistics** sub-menu.
- Click on the **Reset** button.
- Copy the **Science activities index[SCIEACT]** variables into the **Dependent List:** box.
- Copy the **Student gender[GENDER]** variable into the **Factor 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, **GENDER = Female**. 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 $-.2397$. This distribution appears to be reasonably symmetric.

Next we will look at a histogram of the variable, **GENDER = Male**. 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 .0244. This distribution appears to be reasonably symmetric.

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. The Kolmogorov Smirnov test produces test statistics that are used (along with a degrees of freedom parameter) to test for normality.

Tests of Normality

	Student gender	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Science activities index	Female	.118	2310	.000	.953	2310	.000
	Male	.108	2498	.000	.953	2498	.000

a. Lilliefors Significance Correction

For the first group: The Kolmogorov-Smirnov test has degrees of freedom which equals the number of data points, namely 2310.

Here we see that the p value 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.

For the second group: The Kolmogorov-Smirnov test has degrees of freedom which equals the number of data points, namely 2498.

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. Many parametric statistics are robust to non-normality when sample sizes are very large, so non-parametric statistics like the Mann-Whitney are generally used for small samples of non-normal data or ordinal dependent variables.

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

We will now move on to performing the Mann Whitney test for **SCIEACT**.

Below you will see instructions to perform the Mann Whitney 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 Independent-Samples...** from the **Legacy Dialogs** sub-menu.
- Click on the **Reset** button.
- Copy the **Science activities index[SCIEACT]** variable into the **Test Variable List:** box.
- Copy the **Student gender[GENDER]** variable into the **Grouping Variable:** box.
- Click on the **Define Groups...** button.
- Type **1** into the **Group 1** box.
- Type **2** into the **Group 2** box.
- Click on the **Continue** button.

- 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 2 groups and can be seen below:

Ranks				
	Student gender	N	Mean Rank	Sum of Ranks
Science activities index	Female	2310	2208.48	5101596.00
	Male	2498	2585.76	6459240.00
	Total	4808		

The Mann Whitney test works by firstly constructing a ranked list of the observations labelled in their two groups. It will then work from the lowest observation and give that observation rank 1 and the next rank 2 and so on right up to the largest observation which in this case will have rank 4808. If there are observations with the same value then they are given the same rank that is an average of the ranks available (for example if three observations have the 9th smallest rank then rather than giving them ranks 9, 10 and 11 respectively they will each be given rank 10 $(9+10+11)/3 = 10$). The test works by comparing the sum of the ranks in the two groups.

The statistics required for the test are constructed from the ranks and shown in the table. Here we see that for **GENDER** category Female we have 2310 observations whose total sum of ranks is 5101596.00. This results in a mean rank of 2208.48. By contrast for **GENDER** category Male we have 2498 observations whose total sum of ranks is 6459240.00. This results in a mean rank of 2585.76. So **GENDER** category Male has a larger mean rank than **GENDER** category Female and thus tends to take larger values.

The Mann Whitney test will now decide on 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	
	Science activities index
Mann-Whitney U	2432391.000
Wilcoxon W	5101596.000
Z	-9.439
Asymp. Sig. (2-tailed)	.000

The output here consists of test statistics and their significance as calculated in several ways. We are considering the Mann Whitney U statistic and to calculate this we need to consider the sums of the rankings and compare them with what we would expect if these two groups came from the same distribution. We consider each group in turn and work out for each group a U statistic. The formula here is the sum of the ranks - $N \times (N+1)/2$ for each group. So for **GENDER** category Female we have $U1 = 5101596.00 - 2310 \times (2310+1)/2 = 2432391.0$ and for **GENDER** category Male we have $U2 = 6459240.00 - 2498 \times (2498+1)/2 = 3337989.0$. Here U1 is less than U2 and it is the lower of the two U statistics that we report when giving the results. So here the value 2432391.0 is the U test statistic as shown in the table.

A related approach that uses ranks is the Wilcoxon W statistic which is quoted here and is the maximum of the two rank sums but it is less used than the Mann-Whitney so we do not describe it further. One way to interpret the Mann-Whitney U statistic is to convert it to a normal score by subtracting its mean and dividing by its standard error and that is done in the Z row. Here we see that $Z = -9.439$ and this can be compared with a standard normal distribution to get a sense of the magnitude by which the groups differ.

Here we see that the p value, quoted next to Asymp. Sig. (2-tailed), is .000 (reported as $p < .001$) which is less than 0.05. We therefore have significant evidence to reject the null hypothesis that the distribution of **SCIEACT** is the same in the two groups. In this case **GENDER** category Male has a higher mean rank than **GENDER** category Female. The normal approximation used above is only an approximation to the p value and SPSS offers a computational procedure to construct an exact p value. However in this case the sample sizes are so large that the exact test was too computationally-intensive to complete, although the asymptotic p value should be accurate enough.

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

A comparison of the mean of the distribution of the variable SCIEACT was desired for **GENDER** categories Female and Male but due to the non-normality of the variable a Mann Whitney test was carried out. **GENDER** category Male (N= 2498) has a larger mean rank (2585.76) than **GENDER** category Female (N= 2310) with mean rank (2208.48) and thus tends to take larger values. A statistically significant difference was found ($U = 2432391.000$, $p < .001$).

The results indicate that boys tend to engage in more science-related out-of-school activities than girls, a differences that is statistically significant.