Paired t test in SPSS (Practical)

Paired t test practical

In this practical we are going to investigate how to perform a paired t-test using SPSS. A paired t-test is used when we have two continuous variables measured for all observations in a dataset and we want to test if the means of these variables are different. The test assumes that both the variables are normally distributed. To run a single test in SPSS requires that your dataset has two separate columns containing the two variables to be tested. In this situation we could perform a standard 2 sample t-test by reshaping the two variables into one long variable with an accompanying indicator column that defines which original variable each observation refers to but this would be a less efficient test as it does not take account of the paired nature of the data.

The 2015 version of PISA focused on science, and produced separate scales to measure understanding of different content areas. Here we will explore in which domain students scored better, on average, on a test of scientific knowledge of physical systems (SCI_PHYS) and of living systems (SCI_LIVING). Both scores are available for every student in the sample, so a paired test of difference is appropriate here.

Before we can perform this test we need to check whether the differences between **SCI_PHYS** and **SCI_LIVING** are normally distributed. First we need to create a difference variable which can be done as follows:

- Select **Compute** from the **Transform** menu.
- Type **DIFF_SCI_PHYS_SCI_LIVING** into the **Target Variable** box.
- Type **SCI_PHYS - SCI_LIVING** 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_SCI_PHYS_SCI_LIVING** 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_SCI_PHYS_SCI_LIVING**. This can be found in amongst the set of output objects and looks as follows:
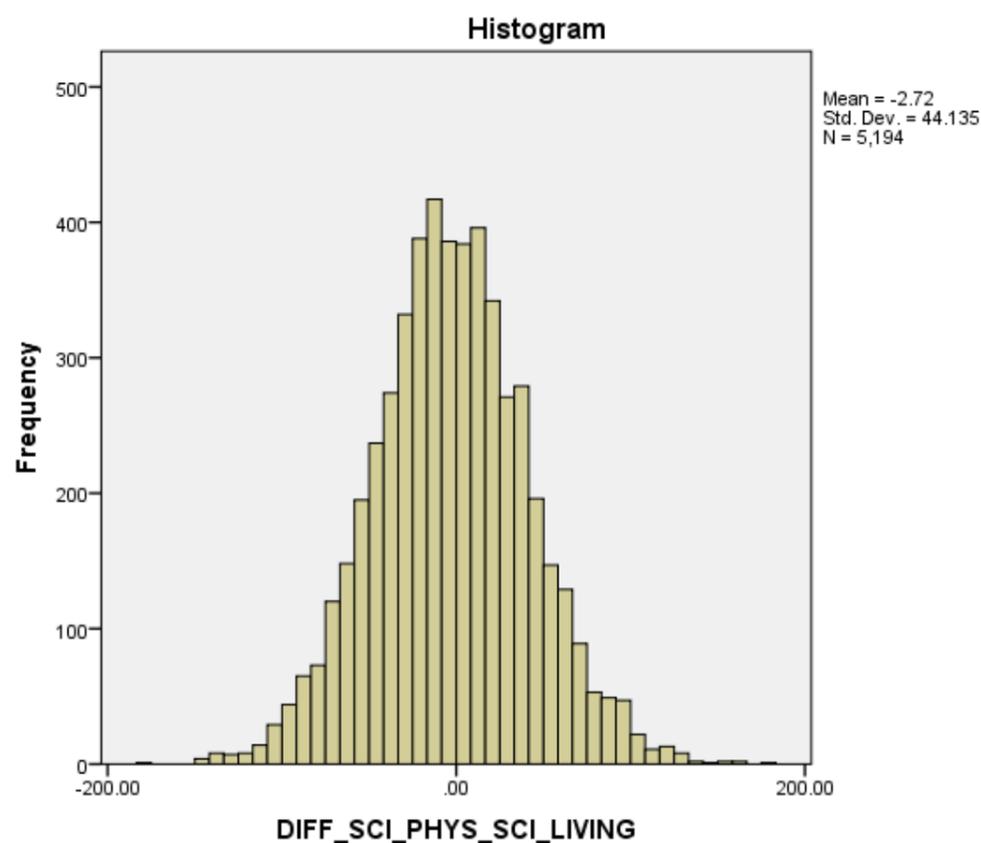
Ideally for a normal distribution this histogram should look symmetric around the mean of the distribution, in this case -2.71597. This distribution appears to be reasonably symmetric.

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.

Tests of Normality

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
DIFF_SCI_PHYS_SCI_LIVING	.013	5194	.062

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 .013. This has degrees of freedom which equals the number of data points, namely 5194.

Here we see that the p value (quoted under Sig. for Kolmogorov Smirnov) is .062 which is greater than 0.05 and therefore we cannot reject the null hypothesis that the distribution is normal.

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 paired t test itself and will test the two variables, **SCI_PHYS** and **SCI_LIVING** for differences.

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

- Select **Compare Means** from the **Analyze** menu.
- Select **Paired-Samples T Test...** from the **Compare Means** sub-menu.
- Click on the **Reset** button.
- Copy the **Physical systems sub-score[SCI_PHYS]** variable into the **Variable1:** box for Pair 1.
- Copy the **Living systems sub-score[SCI_LIVING]** variable into the **Variable2:** box for Pair 1.
- Click on the **OK** button.

The first SPSS output table contains summary statistics for the two variables to be compared and can be seen below:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Physical systems sub-score	520.3541	5194	106.85597	1.48268
	Living systems sub-score	523.0700	5194	106.28832	1.47480

The summary statistics table contains 5 columns and 1 row for each of the two variable to be tested. After the first column which contains the name of each variable, next we see that the mean of variable **SCI_PHYS** is 520.3541 whilst the mean of variable **SCI_LIVING** is 523.0700. Hence the variable **SCI_LIVING** has the bigger mean and the t test will now establish if this difference is statistically significant. We next see the number of valid observations for each variable, i.e. cases with valid values for both **SCI_PHYS** and **SCI_LIVING**. Here we have 5194 valid observations for both variables.

In the next column we see the standard deviations for **SCI_PHYS** and **SCI_LIVING**. In this case the standard deviation of **SCI_PHYS** is 106.85597 whilst for **SCI_LIVING** it is 106.28832. So there is slightly more variability for **SCI_PHYS** than **SCI_LIVING**. In the final column are the standard errors of the means for each group. Whilst the standard deviations measure the variability in the data the standard errors of the means measures how confident we are in the estimates of the means. As we collect more data the standard error of the mean gets smaller as we get more confident in the mean estimate and in fact the formula for the standard error of the mean = standard deviation / square root of N. In this case the standard error of the mean for **SCI_PHYS** is 1.48268 whilst for **SCI_LIVING** it is 1.47480.
The second SPSS output table contains information on the correlation between the two variables to be compared and can be seen below:

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Physical systems sub-score & Living systems sub-score	5194	.914	.000

The correlation between two variables is a single number that describes how related they are to each other. It is represented by a correlation coefficient which is a numerical value to describe the correlation. Correlations lie between -1 and +1 with a positive value meaning that in general that large values of the first variable are more likely to be observed with large values of the second variable and conversely small values of the first variable are more likely to be observed with small values of the second variable. In the case of a negative correlation the opposite is true and large values of the first variable are more likely to be observed with small values of the second variable and conversely small values of the first variable are more likely to be observed with large values of the second variable. A correlation of 0 means there is no (linear) relationship between the variables. Here SPSS is giving out a form of correlation known as a Pearson correlation and we see that the correlation between **SCI_PHYS** and **SCI_LIVING** is .914. It is helpful to look at the correlation between the two variables here as typically a paired t-test is more useful than a 2-sample t-test when there is a positive correlation between the two variables as is the case here. SPSS also gives out a p value which describes whether the correlation is statistically significantly different from zero. Here we see that the p value is less than 0.05 and therefore we can reject the null hypothesis that the correlation is zero.
The third SPSS output table contains details of the t test itself and can be seen below:

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	
Pair 1	Physical systems sub-score - Living systems sub-score	-2.71597	44.13499	.61240	-3.91653	-1.51542	-4.435	5193	.000

The above table describes the paired t test. If we now look at the row of numbers we will start with the column headed Mean (underneath paired differences). Here we see the value -2.71597. If you look back at the summary statistics table this value is calculated by subtracting one mean from the other. Next to the mean is the standard deviation (of the differences) which has value 44.13499. If we have two positively correlated variables then this standard deviation will typically be smaller than the standard deviations of the two variables which it is. Next up is the standard error of the mean (of the differences). This here has the value .61240 and is simply the standard deviation divided by the square root of the sample size. Moving forwards two columns, the column entitled t is the statistic used in the t test and t is a standard statistical distribution. The t statistic is calculated by dividing the mean difference by its standard error so -2.71597 / .61240 = -4.435. Next to t is a column labelled df which stands for degrees of freedom and is a parameter used to choose the correct t distribution for the statistic. Here the degrees of freedom equal the number of observations - 1 (5193) as we have used 1 degrees of freedom in estimating the mean difference.

The column labelled "Sig (2-tailed)" contains a test of the null hypothesis that the means of the two variables (**SCI_PHYS** and **SCI_LIVING**) are the same. By default, the two-tailed test reported uses a non-directional alternative hypothesis. It gives the probability that the data in the sample came from a population in which the variable means are truly equal, when either a positive or a negative difference between sample means is evidence against that null hypothesis. To conduct a one-tailed test, in which the alternative hypothesis specifies a particular direction to the difference, we would simply halve the p-value provided by SPSS.

We can reject the null hypothesis if there is sufficient evidence that the mean of **SCI_PHYS** is either higher or lower than the mean of **SCI_LIVING**. SPSS looks up the t statistic in the appropriate table for the degrees of freedom and in this case the corresponding p value is .000. Here we see that the p value is less than 0.05 and therefore we can reject the null hypothesis that the two groups have the same means. Finally we can see the 95% confidence interval for the difference which runs from -3.91653 to -1.51542. Here we see it does not contain the value 0 backing up our rejection of the null hypothesis.

In conclusion, we could report this to a reader as follows: Mean values were compared for 2 variables with sample size 5194. The mean was higher for variable **SCI_LIVING** (M=523.0700, SD=106.28832) than for variable **SCI_PHYS** (M=520.3541, SD=106.85597). The difference in means (difference = -2.71597) was statistically significant, t(5193) = -4.435, p=.000.

The results suggest that students abilities were significantly stronger in the living systems domain than in the physical systems domain. The advantage of a paired research design is that the same people were tested in both domains, so that subject-specific variables are all held constant. If different participants had taken the physical systems test from those who had taken the living systems test, some of the variability in the two sets of scores would be attributable to individual differences in factors like aptitude for science and general cognitive ability. The paired design helps to isolate only what is different between the two test conditions.