Statistical Analysis 3: Paired t-test

Research question type: Difference between (comparison of) two **related** (**paired**, **repeated** or **matched**) variables

What kind of variables? Continuous (**scale/interval/ratio**)

Common Applications: Comparing the means of data from two related samples; say, observations before and after an intervention on the same participant; comparison of measurements from the same participant using 2 measurement techniques

Example 1:

Research question: Is there a **difference** in marks following a teaching intervention?

The marks for a group of students before (pre) and after (post) a teaching intervention are recorded below:

Student	Before mark	After mark	Diff
1	18	22	4
2	21	25	4
3	16	17	1
4	22	24	2
5	19	16	-3
6	24	29	5
7	17	20	3
8	21	23	2
9	23	19	-4
10	18	20	2
11	14	15	1
12	16	15	-1
13	16	18	2
14	19	26	7
15	18	18	0
16	20	24	4
17	12	18	6
18	22	25	3
19	15	19	4
20	17	16	-1
Mean	18.40	20.45	2.05

Marks are **continuous (scale) data**. Continuous data are often summarised by giving their average and standard deviation (SD), and the **paired t-test** is used to compare the means of the two samples of related data.

The paired t-test compares the mean difference of the values to zero. It depends on the mean difference, the variability of the differences and the number of data.

Various assumptions also need to hold – see validity section below.

You should practise entering the data into SPSS (PASW), but the data are available on W:\EC\STUDENT\ MATHS SUPPORT CENTRE STATS WORKSHEETS\marks.sav

[NB The Diff column is given here for illustration purposes; it does not have to be entered to SPSS]

Hypotheses:

The 'null hypothesis' might be:

H_0: There is no difference in mean pre- and post-marks

And an 'alternative hypothesis' might be:

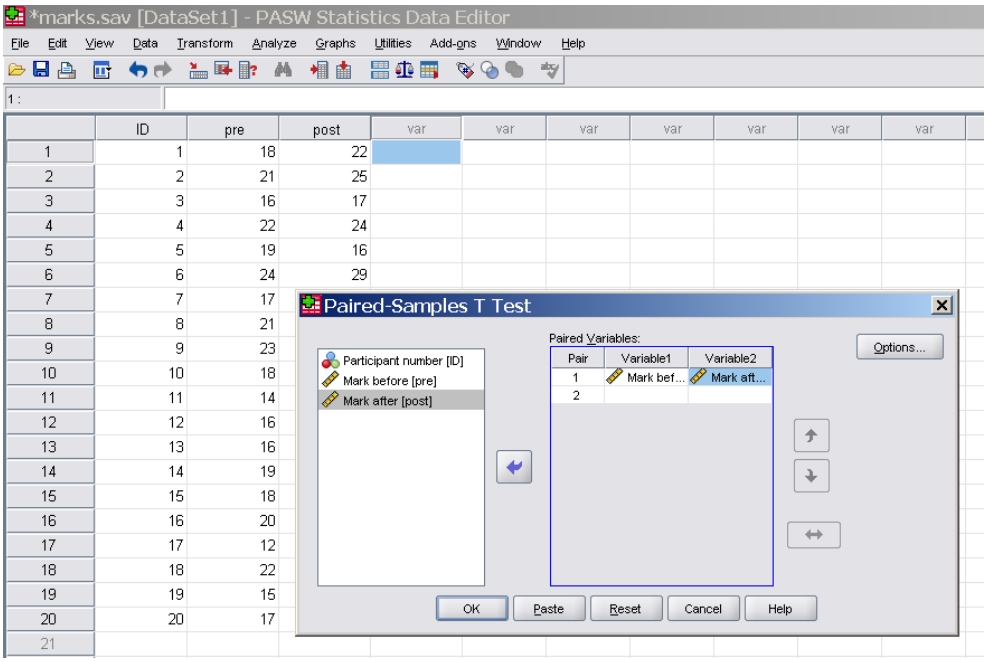
H_1: There is a difference in mean pre- and post-marks

Steps in SPSS (PASW):

The data need to be entered in SPSS in 2 columns, where one column indicates the pre-mark and the other has the post-mark – see over. [A third column could include participant numbers].

Analyze > Compare Means > Paired Samples T-test

- Select the two paired variables as the Paired Variables, selecting the after variable first (post), followed by the before variable (pre) – see below
- Click OK



Output should look something like below:

[There is another table showing correlation – not needed for this purpose]

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mark after	20.45	20	4.058	.907
	Mark before	18.40	20	3.152	.705

***p*-value**

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Mark after - Mark before	**2.050**	2.837	.634	.722	3.378	3.231	19	**.004**

Results:

Notice that this option automatically gives you the sample summary data.

The relevant results for the paired t-test are in **bold**.

From this row observe the t statistic, t = 3.231, and p = 0.004; ie, a very small probability of this result occurring by chance, under the null hypothesis of no difference.

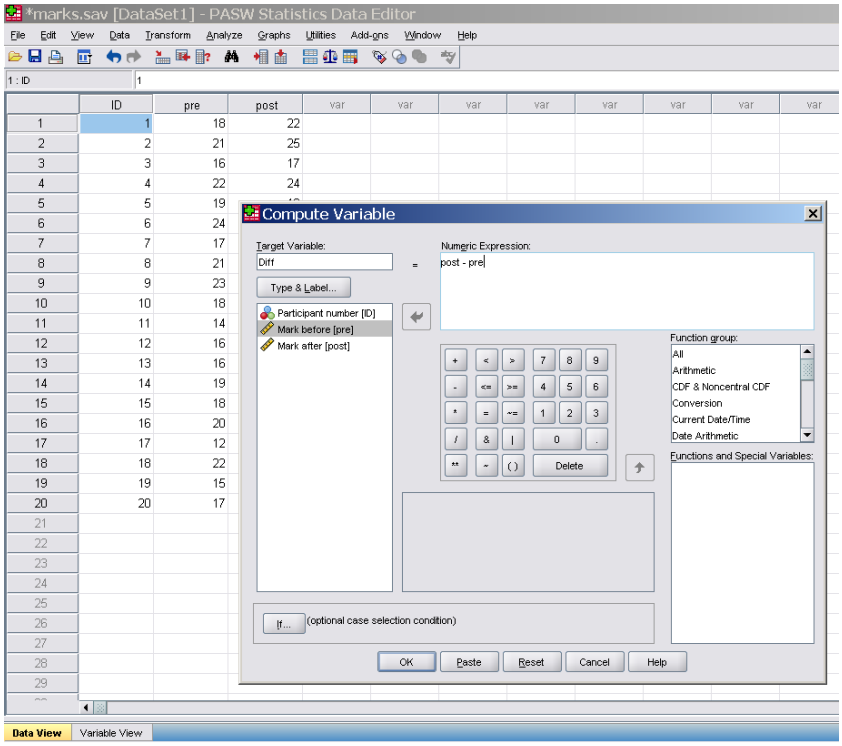
The null hypothesis is rejected, since $p < 0.05$ (in fact p = 0.004).

Conclusion:

There is strong evidence (t = 3.23, p = 0.004) that the teaching intervention improves marks. In this data set, it improved marks, on average, by approximately 2 points. Of course, if we were to take other samples of marks, we could get a 'mean paired difference' in marks different from 2.05. This is why it is important to look at the 95% Confidence Interval (95% CI).

If we were to do this experiment 100 times, 95 times the **true value** for the difference would lie in the 95% confidence interval. In our case, the 95% CI is from 0.7 to 3.4. This confirms that, although the difference in marks is statistically significant, it is actually relatively small. You would need to consider if this difference in marks is **practically important**, not just **statistically significant**.

Validity of paired (related) t-tests:

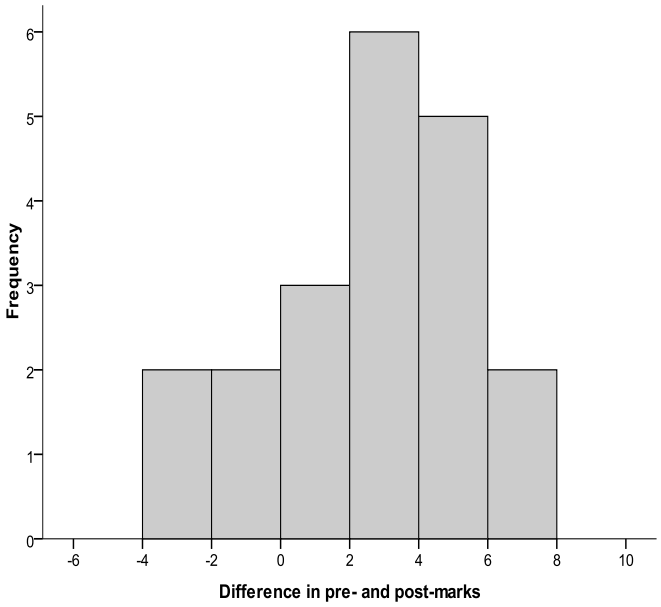


For the paired samples t-test to be valid the differences between the paired values should be approximately **normally distributed**.

To calculate the differences between pre- and post-marks, from the Data Editor in SPSS (PASW), choose:

Transform>Compute Variable and complete the boxes as shown on the left:

***Histogram of differences in marks**



Normal distribution can be checked by:

- looking at a histogram of the 'Diff' data*,
- looking at a normal probability (QQ) plot**
- doing a simple Kolomogorov-Smirnov test***

[NB this chart has been edited in SPSS (PASW) Chart Editor

**Analyse>Descriptive Statistics>Q-Q Plots…

*** Analyse>Nonparametric tests>1-Sample K-S… You require a **non-significant** result (ie p > 0.05: in this example, p = 0.808)

Example 2:

In an experiment to compare anxiety levels induced between looking at real spiders and pictures of spiders, the following data was collected from 12 people with a fear of spiders (arachnophobia):

	Anxiety score		
Participant	Picture	Real	Diff
1	30	40	10
2	35	35	0
3	45	50	5
4	40	55	15
5	50	65	15
6	35	55	20
7	55	50	-5
8	25	35	10
9	30	30	0
10	45	50	5
11	40	60	20
12	50	39	11

Suitable null and alternate hypotheses could be:
H_0: There is no difference in mean anxiety scores between looking at real or pictures of spiders, and

H_1: There is a difference in mean anxiety scores between looking at real or pictures of spiders

You should practise entering the data into SPSS (PASW), but the data are available on W:\EC\STUDENT\ MATHS SUPPORT CENTRE STATS WORKSHEETS\spiders.sav

Results:

Following the steps in SPSS (PASW) outlined previously, you should get the following output:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Anxiety score when looking at a spider picture	40.00	12	9.293	2.683
	Anxiety score when looking at a real spider	47.00	12	11.029	3.184

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	Std. Error Mean	95% CI of the Difference Lower	Upper			
Pair 1	Anxiety score when looking at a spider picture - Anxiety score when looking at a real spider	-7.000	9.807	2.831	-13.231	-.769	-2.473	11	.031

Conclusion:

You would report something along the lines that there is evidence to suggest that participants experienced statistically significantly greater anxiety (p = 0.031) when exposed to real spiders (mean = 47.0 units, SD = 9.3) than to pictures of spiders (mean = 40.0 units, SD = 11.0).`The 95% confidence interval for the difference is (-13.2,-0.77).

[#]Example taken from Field A (2009) *Discovering statistics using SPSS* 2nd ed. Sage Publications