statstutor

community project

encouraging academics to share statistics support resources



stcp-gilchristsamuels-7

The following resources are associated:

Checking normality for parametric tests

# Paired Samples 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

## Example: Teaching Intervention

**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 | Difference |
|---|---|---|---|
| 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 mean and standard deviation, 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 with zero. It depends upon the mean difference, the standard deviation of the differences and the number of cases. Various assumptions also need to be made.




Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

## Hypotheses:

The null hypothesis is:

$H_0$: There is no difference in mean *Pre* and *Post* marks

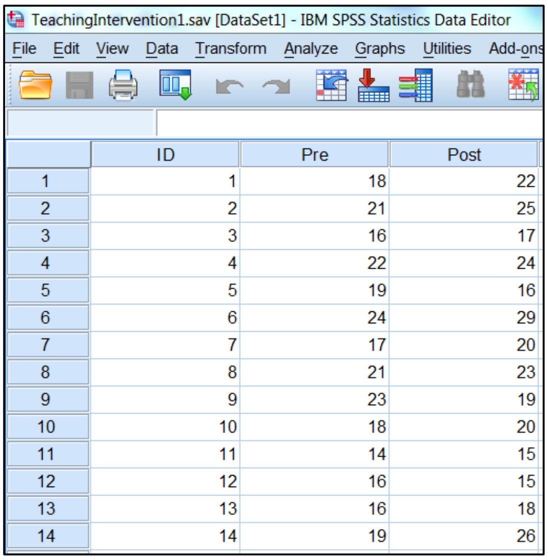
And the alternative hypothesis is:

$H_1$: There is a difference in mean *Pre* and *Post* marks

# Steps in SPSS

## Enter the data

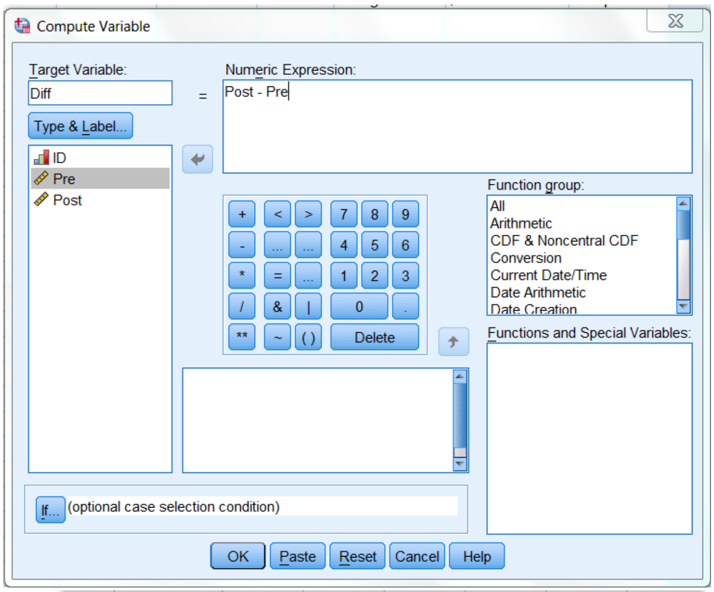
The data needs to be entered in SPSS in two columns, where one column indicates the pre-mark and the other has the post-mark – see right (a third column provides the case identity numbers).




| | ID | Pre | Post |
|---|---|---|---|
| 1 | 1 | 18 | 22 |
| 2 | 2 | 21 | 25 |
| 3 | 3 | 16 | 17 |
| 4 | 4 | 22 | 24 |
| 5 | 5 | 19 | 16 |
| 6 | 6 | 24 | 29 |
| 7 | 7 | 17 | 20 |
| 8 | 8 | 21 | 23 |
| 9 | 9 | 23 | 19 |
| 10 | 10 | 18 | 20 |
| 11 | 11 | 14 | 15 |
| 12 | 12 | 16 | 15 |
| 13 | 13 | 16 | 18 |
| 14 | 14 | 19 | 26 |

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, choose *Transform - Compute Variable* and complete the boxes as shown below right.



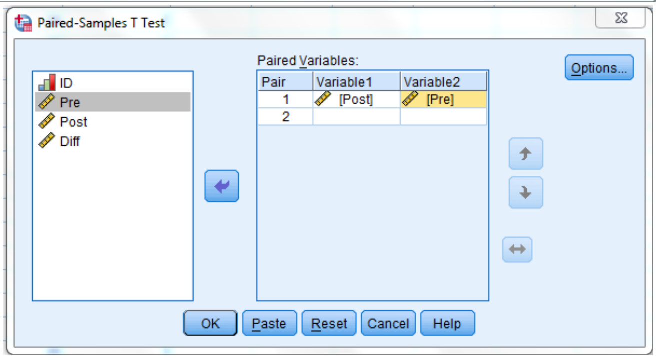

## Check the test assumptions

The normality of *Diff* should first be checked – see **Checking normality for parametric tests worksheet.**

There is no evidence for us to suspect that the data is not normally distributed.

## Running the paired samples t-test

- Select *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*) as shown **below** and click OK



## Results

The output should look like the table below:

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | 95% Confidence Interval of the Difference | | | | |
| | | Mean | Std. Deviation | Std. Error Mean | Lower | Upper | | | |
| Pair 1 | Post - Pre | 2.050 | 2.837 | .634 | .722 | 3.378 | 3.231 | 19 | .004 |

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

The t statistic (t) is 3.231, and p-value (Sig. (2-tailed)) is 0.004. Therefore we may reject the null hypothesis (that there was no difference between the means of the two groups) and conclude that there is strong evidence of a difference in the means.

### Conclusion

There is strong evidence ($t = 3.23$, $p < 0.01$) that the teaching intervention improves marks. The estimated improvement is approximately 2 marks. Alternatively this can be described as an **effect size** given by the absolute value of the difference in means (2.050) divided by the standard deviation (2.837) which is approximately 0.72 (this is classified as a large effect).

Of course, if we were to take other samples of marks, we might get a 'mean paired difference' in marks different from 2.05. The 'Lower' and 'Upper' limits of the **95% confidence interval** tell us that we can be 95% confident that the population mean difference between the *Pre* and *Post* marks is between 0.72 and 3.38 marks.

You would need to consider if this difference in marks is practically important, not just statistically significant.

## Notes

1. If the difference between the before and after values is not normally distributed then it should still be possible to run a paired t-test provided that the sample size is greater than 30 and the distribution is not very skewed: see (Rhiel and Chaffin, 1996) and (Zumbo and Jennings, 2002). This can be assessed by inspecting a histogram or the data or the value of the skewness statistic (both available in SPSS via Analyze > Descriptive Statistics > Explore).
2. If the paired t-test cannot be used the **Wilcoxon signed rank** test should be used instead.

## Reference

Rhiel, G. S. and Chaffin, W. W. (1996) An investigation of the large-sample/small-sample approach to the one-sample test for a mean (sigma unknown). *Journal of Statistics Education*, 4(3). Available at: http://www.amstat.org/publications/jse/v4n3/rhiel.html.

Zumbo, B. D. and Jennings, M. J. (2002) The robustness of validity and efficiency of the related samples t-test in the presence of outliers, *Psicológica*, 23(2), pp. 415-450.