Research Methods 2

Week 10: Document 2

Comparing two groups: II the paired *t*-test

The structure of data : an example

Patients undergoing a thoracotomy received morphine via a patient-controlled device (PCA) for post-operative analgesia. With this type of operation, breathing out is often limited by pain, so the peak expiratory flow rate (in l/min) (PEFR) is a useful indirect measurement of pain, with higher values corresponding to lower pain.

The PEFR recorded on admission to the high dependency unit HDU for ten patients undergoing thoracotomy is given in the second column of table 1. The PEFR recorded one hour later is in the third column. (Data from Sudarshan *et al*., Br J. Anaesthesia, 1995, 75, 19-22, by kind permission of Dr I.D. Conacher).

Patient	PEFR on admission to HDU (I)	PEFR one hour post admission (II)	(II) randomly re-ordered
1	100	110	90
2	80	60	80
3	180	160	160
4	60	80	110
5	210	200	280
6	130	80	60
7	80	90	200
8	80	60	60
9	120	80	80
10	250	280	80

Table 1

A question that might well be asked is whether or not the mean PEFR is lower one hour later than on admission to HDU. You might think you could do this by applying the method outlined in the previous document, with the admission PEFR forming one group and the one hour values another group.

Unfortunately there is a problem with this. In the final column of table 1 the one-hour values have been randomly rearranged. If you put the three columns of PEFR into Minitab in exactly the same way as in table 1, then comparing column 1 with column 2 using an unpaired *t*-test gives you the same result as comparing column 1 with column 3 (cf. last question on previous worksheet). The value of 110 in the first row of column 3 is the PEFR recorded on the patient who had a PEFR of 100 on admission , i.e. in the corresponding row of column 2. The value in the corresponding row of column 4 is the PEFR from an entirely different person. Surely it is important to know how each patient's peak flow changed to over the hour? The unpaired *t*-test will happily give you the same answer whatever the re-ordering of the third column, and quite blithely ignores this aspect of the data.

The important distinction between this example and that for the retention indices is that in the latter the two groups being compared were defined by who responded and who did not, i.e. by two quite separate and independent groups of patients. There is no sense in which any member of one group can be paired with a member of the other. On the other hand, the PEFR values in the one hour group can be paired in a sensible way with a PEFR value in the admission group, by pairing the values from the same patient.

This distinction between the two examples is an instance of what can be called the *structure* of the data. A fuller exposure to this notion is beyond the scope of the present course, but it is of great importance in statistics. If a method of analysis ignores an important aspect of the structure of the data, as did the unpaired *t*-test when applied to the PEFR data, it is a good indication that you are using the wrong method.

The paired* t*-test

So what method should be applied to compare the PEFR values on admission with those at one hour?

Not surprisingly the answer is to use the paired *t*-test. As with the unpaired version, we will avoid the details (which can be found in section 10.2 of Bland) but it is useful to appreciate one methodological aspect.

Since the data in the two groups are paired, we can form the pairwise differences. For example, in the PEFR data, the first thing the paired *t*-test does is to form the column of differences between columns 2 and 3 of table 1[†]. These are the changes in PEFR over the first hour in HDU. Once these differences have been formed then, as far as the paired *t*-test is concerned, the admission and one-hour values can be ignored, only the differences are analysed. *Note that you could not do this with the example on retention indices, as the pairing is vital to defining the differences.*

As the analysis effectively considers only the one sample of differences, the paired *t*-test is often referred to as the 1-sample *t*-test (and the unpaired version as the 2-sample *t*-test).

Assumptions of the paired *t*-test

The assumptions focus on the properties of the sample of differences. The differences assumed to follow a Normal distribution, as in assumption i) for the unpaired test. There is no restriction on the form of the distribution of the values that make up the difference. So, e.g., the distribution of the admission values of PEFR and of the one hour values can be anything you like, just so long as the difference between the two is, at least approximately, Normally distributed.

As there is only one sample being analysed, there is no assumption corresponding to assumption ii) (about equal population SDs) for the paired test.

Application

The paired *t*-test comparing columns 2 and 3 of table 1 is performed as follows. In Minitab, click on **<u>S</u>tat** -> **<u>B</u>asic statistics** -> **<u>P</u>aired t…**. The column containing the admission values should be entered in the **<u>F</u>irst sample:** box and that for the one hour values in the **<u>S</u>econd sample:** box. You then click on **<u>O</u>K** and the results in the Session window are shown below.

Paired T-Test and CI: Admission, one hour

```
Paired T for Admission - one hour

             N     Mean     StDev  SE Mean
Admission   10  129.000    63.849   20.191
one hour    10  120.000    71.957   22.755
Difference  10  9.00000  26.01282  8.22598


95% CI for mean difference: (-9.60845, 27.60845)
T-Test of mean difference = 0 (vs not = 0): T-Value = 1.09  P-Value = 0.302
```

The output and its interpretation are virtually identical to that for the unpaired version of the test. The important values are: the mean change in PEFR, which is the mean (`9.00`) in the row labelled '`Difference`'; the P-value of 0.302; and the 95% confidence interval for the mean change in PEFR.

The P-value indicates that if there was no evidence of a difference in mean PEFR at the two times: a difference of at least 9 l/min would arise by chance on about 30% of occasions, so the present data does not discredit the null hypothesis of no mean change. However, the confidence interval indicates that

---

[†] note when you come to do a paired *t*-test using Minitab that the program will form this difference internally – the user does not have to subtract the two columns explicitly.

the mean change (admission – one hour) could be between –9.6 l/min (i.e. lower at one hour) up to 27.6 l/min higher at one hour.