Research Methods 2

Week 10: Document 1

Comparing two groups: I the unpaired *t*-test

Introductory example, continued

Last week we considered the example from Takeuchi *et al.* (*Int. J. Oncology*, 2002, **20**, 53-58) where the retention index for thallium was compared between a group of patients with breast cancer who responded to treatment and a group who did not respond. The emphasis last week was on the principles of hypothesis testing and various approaches followed there, e.g. postulating plausible values for population parameters were adopted to clarify the exposition. It is clear that in practice such assumptions are not tenable in practice and the emphasis this week will be more practical.

As an example we will compare the retention indices for thallium between the two groups in the way it would be done in practice, namely with an *unpaired* t-*test.*

The unpaired t*-test.*

This is the method that is used to compare the means of two independent groups. It tests the null hypothesis of the equality of the means of the populations from which the samples have been drawn. The method applies when the data are continuous, such as haemoglobin concentrations, heights, radiotherapy doses. In fact there are two assumptions that the method makes:

i) the data have a Normal distribution.
ii) the two *population* SDs are the same.

Before using these methods it is useful to check if these assumptions are reasonable.

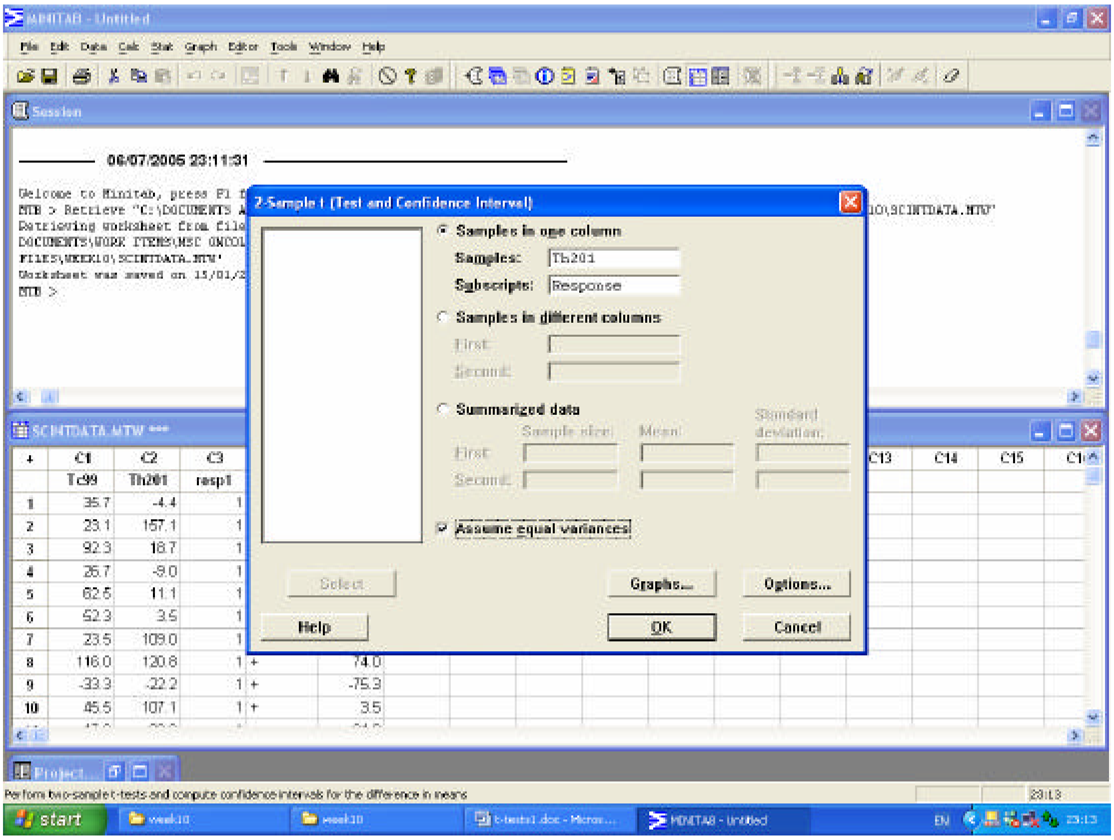
To assess i) it is sensible to plot a couple of histograms. If the data are markedly non-Normal then you may need to use a different method. However, the technique is not particularly sensitive to departures from this assumption, so moderate departures from the ideal bell-shape, particularly for small samples, are of little importance.

Assumption ii) is assessed by informally comparing the ratio of the *sample* SDs. It is important to note that the assumption specifies the equality of population SDs, which we do not know. Inequality in the sample SDs might reflect differences in the population SDs, but they might also reflect chance differences. For small samples the sample SDs are very variable, and quite large discrepancies in the sample quantities are acceptable. It is difficult to give precise guidance on the degree of discrepancy, as this is related to the many things, including sample size.

To perform the test in Minitab, suppose that all the retention indices for thallium are held in a column called Th201. Suppose further that a further column, called response, contains the symbols + and -, indicating whether the retention index in the corresponding row of Th201 was from a patient who responded (+) or did not respond (-). Now click on **Stat** -> **Basic Statistics** -> **2-Sample t…**

You will be presented with the screen shown below. The retention indices are all in one column, with a second column indicating which group the values come from. Therefore the button **Samples in one column** has been selected (if the indices had been in two columns, one for responders and one for non-responders, then the **Samples in different columns** button would have had to be checked). The column containing the retention indices is placed in **Samples:** and the column indicating group membership is in **Subscripts:** . It is also necessary to check the **Assume equal variances box**†. The command is run by clicking on **OK**.

† Minitab can apply a form of *t*-test which does not use assumption ii). This sounds attractive but there are reasons, beyond our present scope, why it is not as attractive as it sounds. Checking this box ensures that the mainstream version of the *t*-test is applied.



The results appear in the Session window as follows.

Two-Sample T-Test and CI: Th201, Response

```
Two-sample T for Th201

                               SE
Response    N  Mean  StDev  Mean
-          14  55.8   46.7    12
+          11  47.5   63.1    19


Difference = mu (-) - mu (+)
Estimate for difference:  8.33117
95% CI for difference:  (-37.04946, 53.71180)
T-Test of difference = 0 (vs not =): T-Value = 0.38  P-Value = 0.708  DF =
23
Both use Pooled StDev = 54.4468
```

The first part is the summary giving the sample sizes and the mean and SD in each group. This is useful for making an informal assessment of the SDs. For samples of these sizes the difference in the sample SDs provides no cause for concern apropos assumption ii).

Perhaps the most important single quantity in this analysis is the difference in samples means, which is our best estimate of the difference in population means. This is the quantity labelled `Estimate of difference`.

A very useful quantity is the 95% confidence interval for the difference in population means and this is given next.

The line starting `T-Test of`... gives several quantities which we have not discussed directly. The main quantity we need is the P-value, which here is 0.71 (to two d.p.)

The interpretation of these results is as follows.

1. The difference between the sample means, namely 8.3, is the sort of difference that would frequently occur by chance if the population means were the same.
2. Statement 1 is based on the P-value of 0.71, which means that the differences of 8.3 or more occur 71% of the time if the null hypothesis (equal population means) is true.
3. The above can be equivalently expressed by saying that the data are compatible with the difference in population means, m_R-m_{NR} , having the value zero.
4. In fact, more can be said. The 95% confidence interval for the difference in population means, is given as -37.0 to 53.7. With 95% confidence we can say that the data are compatible with m_R-m_{NR} having any value in this range.
5. Note that the conclusion in 3 is consistent with that in 4 because 0 is in the confidence interval. There is an intimate connection between hypothesis tests and confidence intervals and those interested can consult the Appendix.