**Worksheet: Unpaired t-test in EXCEL**

**The unpaired t-test is used to compare the means of two independent groups for a continuous outcome**

**Example: People with diabetes – comparison of HbA1c values between two groups randomised to either an educational health intervention or care as usual.**

Data need to be arranged in columns such that the observations for the two groups are in separate columns:



To check the assumption that the data in the groups is approximately normally distributed, construct histograms of the data separately for each group, as described in worksheet 1

To check that the variability in the two groups are similar, use the following formula to calculate the standard deviations:

*=STDEV.S(A:B)*

Where A= start of the data range

 B= end of the data range

For this example in the ‘*Care-as-usual*’ group there are 74 observations in rows 2 to 75 of column A and 109 observations in rows 2 to 110 of column B for the *Intervention* group in thus the functions are:

*=STDEV.S(A2:A75)*

*=STDEV.S(B2:B110)*

The standard deviation for the Care-as-Usual group is 1.53 and for the Education group is 1.37. As the larger one is less than twice the value of the smaller one, it is ok to use the equal variances t-test.

To do a t-test you can either use the T-Test formula:

*=T.TEST(A, B, C, D)*

Where A= data range for first occasion

 B= data range for second occasion

 C= 1 if one-sided test

 2 if two sided test \*\*\*Recommended \*\*\*

 D= 1 paired t-test

 2 unpaired t-test, equal variances assumed

 3 unpaired t-test, equal variances not assumed

For this example, in column 1 there are 74 observations in rows 2 to 75 and in column 2 109 observations in rows 2 to 110, I want to do a 2 sided test assuming equal variances and the data are not paired thus the function is:

*=T.TEST(A2:A75,B2:B110,2,2)*

You get a single value, the two-sided p-value for an unpaired t-test = 0.0303. As this is < 0.05 we say that the result is statistically significant at the 5% level and there is evidence to suggest that the difference between the population means is not = 0.

1. Alternatively you can use the data analysis toolpak:



This will open up the Data Analysis dialogue box. Select either *t-Test: Two-sample Assuming Equal Variances* or *t-Test: Two-Sample Assuming Unequal Variances* depending on whether the variances were similar or not. In our case they were and we select the first option:



Then fill in the options:



And here’s the output:



The table shows you various summary statistics including the mean in each group and its variance. You can convert the variances into standard deviations by taking the square root. It also gives you the p-value for the comparison between the two groups. If it is significant, this tells you that there are differences in the means between the groups. As with the t-test formula you can see that the two-sided p-value = 0.0303. As this is < 0.05 we say that the result is statistically significant at the 5% level and there is evidence to suggest that the difference between the population means is not = 0.

Having obtained the p-value for the difference, you can use EXCEL to calculate the mean difference and its 95% confidence interval using the following formulae:



And you get the following results:



To summarise:

*Two groups of adults with type 1 diabetes were randomised to either take part in an educational training programme aimed at giving them more freedom and control over what they could eat, or to receive care as usual. At the end of 12 months follow-up an unpaired t-test was conducted to examine whether there was a difference in the control of their diabetes as measured by HbA1c. The result was significant (P=0.03) indicating that there was a significant difference in their HbA1c levels, with a positive difference in favour of the intervention group of 0.47% points (95% confidence interval: 0.05 to 0.90)*