

## 8.1 Hypothesis Testing Using Excel (see below for LibreOffice)

### The Related Samples T Test

#### The Two -Tailed Test

##### Example 8.4

Consider the container design data in Data Set F (see the Data Annexe). Notice that the two variables Con1 and Con 2 indeed measure the same characteristic (the number of items sold), but under two different “conditions” (the two different container designs).

We conduct a two-tailed related samples t test of whether the underlying (population) mean number of items sold differs between the two container designs.

Strictly speaking, before undertaking the test we should calculate the differences

$$D = \text{Con1} - \text{Con 2}$$

for each observation. A normal plot of these differences (i.e. of the values of the variable D) should then be constructed in order to check whether the data are acceptably near-normally distributed.

We will assume for now that the data are indeed so distributed so that the resulting t test is valid. You might want to construct the normal plot as an additional exercise. (Ensure you save your answers in the Exercise sheets for your submission).

1. Open the Excel workbook **Exa 8.4F.xlsx** from the Examples folder. This contains the relevant data.
2. From the **Data** menu bar tab, select **Data Analysis** from the **Analysis** group, and from the ensuing dialogue box, select **t test: Paired Two Sample for Means**. A new dialogue box appears.
3. In the **Variable 1 Range** box, enter the cell range where the data for the first variable (Con1) can be found, including the variable name, that is, the range B1:B11. In the **Variable 2 Range** box, enter the cell range where the data for the second variable (Con2) can be found, including the variable name, that is, the range C1:C11. Ensure that the **Labels** box is checked.
4. Type:0 in the **Hypothesised Mean Difference** box. This represents the null hypothesis of no difference between the treatment means.
5. Ensure that the **Alpha** box contains the value 0.05. This is only of marginal relevance, as we shall make direct use of the p-value that will be output!
6. Select the **Output Range** button, and in the corresponding box, enter the cell reference E1. Click the **OK** button. Some output appears in your spreadsheet. Widen columns E, F and G so that all the text becomes readable.

7. In cell E16, type: Difference in Means, and in cell F16, enter the formula =F4-G4.

The resulting output is presented below.

Not all this output is relevant, so it need not all be discussed.

The obtained related samples  $t = 2.875$  with 9 degrees of freedom.

The associated two-tailed p-value is  $p = 0.018$ , so the observed  $t$  is significant at the 5% level (two-tailed).

t-Test: Paired Two Sample for Means		
	Con1	Con2
Mean	172.6	159.4
Variance	750.2666667	789.3777778
Observations	10	10
Pearson Correlation	0.863335004	
Hypothesized Mean Difference	0	
df	9	
t Stat	2.874702125	
P(T<=t) one-tail	0.009167817	
t Critical one-tail	1.833112923	
P(T<=t) two-tail	0.018335635	
t Critical two-tail	2.262157158	
Difference in Means	13.2	

The sample mean numbers of items sold for Container Designs 1 and 2 were, respectively 172.6 and 159.4. The data therefore constitute significant evidence that the underlying mean number of containers sold was greater for Design 1, by an estimated  $172.6 - 159.4 = 13.2$  items per store. The results suggest that Design 1 should be preferred.

#### Exercise 8.4

Consider the filtration data of Data Set G. Open the Excel workbook **Exe8.4G.xlsx** which contains these data from the Exercises folder.

Assuming the data to be suitably distributed, complete a two-tailed test of whether the population mean impurity differs between the two filtration agents, and interpret your findings.

## Answer

Skewness and kurtosis tests were performed using the Excel data analysis tool to double-check the normality of the data. Results showed acceptable skewness and kurtosis of the two sample distributions.

The p-value of the two-tailed t-test was 0.027, which means that the mean difference was statistically significant. Hence, Agent 1 resulted in less impurity (mean = 8.15 parts per 1000 by weight) compared to Agent 2 (mean = 8.54 parts per 1000 by weight) with an estimated 0.39 parts per 1000 by weight difference.

## The One-Tailed Test

### Example 8.5

Recall that in Example 8.4, we conducted a two-tailed related samples t test of whether the underlying (population) mean number of items sold differs between the two container designs of data Set F.

However, now suppose that Container Design 1 is a new, hopefully more attractive design, whereas Container Design 2 is the design in current use. Presumably, the company will only go to the expense of implementing the new design if it can be shown to lead to higher sales than the current design. Thus, the investigators seek evidence that  $\mu_1 > \mu_2$ , so wish to test:

$$H_0: \mu_1 \leq \mu_2 \quad \text{against} \quad H_1: \mu_1 > \mu_2$$

The relevant t test is conducted exactly as before. However, this time, the results are interpreted a little differently.

We first of all check whether the data are consistent with the one-tailed alternative hypothesis. As before, the sample mean numbers of items sold for Container Designs 1 and 2 were, respectively 172.6 and 159.4, so that the data are indeed consistent with  $H_1$ .

As before, the obtained related samples  $t = 2.875$  with 9 degrees of freedom.

The associated one-tailed p-value is  $p = 0.009$ , so the observed t is significant at the 1% level (one-tailed).

t-Test: Paired Two Sample for Means		
	<i>Con1</i>	<i>Con2</i>
Mean	172.6	159.4
Variance	750.2666667	789.3777778
Observations	10	10
Pearson Correlation	0.863335004	
Hypothesized Mean Difference	0	
df	9	
t Stat	2.874702125	

P(T<=t) one-tail	0.009167817	
t Critical one-tail	1.833112923	
P(T<=t) two-tail	0.018335635	
t Critical two-tail	2.262157158	
Difference in Means	13.2	

The data therefore constitute strong evidence (on a one-tailed test) that the underlying mean number of containers sold was greater for Design 1, by an estimated  $172.6 - 159.4 = 13.2$  items per store. The results continue to suggest that Design 1 should be preferred.

Although broadly similar conclusions were reached as before, a higher level of significance was obtained with the one-tailed test.

Notice that if we had sought to test the alternative pair of one-tailed hypotheses

$$H_0: \mu_1 \geq \mu_2 \quad \text{against} \quad H_1: \mu_1 < \mu_2$$

we would have found the difference in sample means to be consistent with the *null hypothesis* that the population mean sales for Design 2 was no greater than that for Design 1. We would thus have declared the result to be not significant without even bothering to inspect the p-value.

### Exercise 8.5

Recall that in Exercise 8.4, a two-tailed test was undertaken of whether the population mean impurity differs between the two filtration agents in Data Set G.

Suppose instead a one-tailed test had been conducted to determine whether Filter Agent 1 was the more effective. What would your conclusions have been?

### Answer

The one-tailed t-Test resulted in a p-value of 0.014 which is still significant. This means Agent1 is more efficient in filtration than Agent B. So, the conclusion is the same whether using a two-tailed or one-tailed paired t-Test.

## The INDEPENDENT Samples T Test

### Example 8.6

Consider again Data Set B, the dietary data. Not unreasonably, we wish to test whether the population mean weight loss differs between the two diets. Since completely separate samples of individuals undertook the two diets (i.e. no-one underwent both diets), the independent samples t test is appropriate here.

1. Open the Excel workbook **Exa 8.6B.xlsx** from the Examples folder. This contains the relevant data, together with some of the previously calculated summary statistics for the weight loss on each diet.

We begin by performing the F test of variances.

2. From the **Data** menu bar tab, select **Data Analysis** from the **Analysis** group, and from the ensuing dialogue box, choose **F-test Two-Sample for Variances** and click **OK**. A further dialogue box opens.
3. In the **Variable 1 Range** box, enter the cell range where the Diet A weight losses can be found (B2:B51), and in the **Variable 2 Range** box, enter the cell range where the Diet B weight losses can be found (B52:B101). Ensure that the **Labels** option is unchecked.
4. In the **Alpha** box, ensure that 0.05 is entered (although this is relatively unimportant as we are going to use p-values). Click the **Output Range** button and enter the cell reference H3 in the corresponding box. Then click **OK**.
5. Some output appears. Widen columns H to J to render it legible. In cell H14, type: p2, and in cell I14, enter the formula: =2\*I11 to obtain the required two-tailed p-value.

The relevant output is as follows:

F-Test Two-Sample for Variances		
	<i>Variable 1</i>	<i>Variable 2</i>
Mean	5.3412	3.70996
Variance	6.429280612	7.66759359
Observations	50	50
df	49	49
F	0.838500442	
P(F<=f) one-tail	0.269951479	
F Critical one-tail	0.622165467	
p2	0.5399	

The sample variances for the two diets are, respectively

$$s_1^2 = 6.429 \text{ and } s_2^2 = 7.668$$

The observed F test statistic is  $F = 0.839$  with 49 and 49 associated degrees of freedom, giving a two tailed p-value of  $p = 0.5399^{\text{NS}}$

The observed F ratio is thus *not significant*. The data are consistent with the assumption that the population variances underlying the weight losses under the two diets do not differ, and we therefore proceed to use the *equal variances* form of the unrelated samples t test.

Since we wish to test if the population mean weight losses differ between the two diets, a two-tailed t test is appropriate here.

6. From the **Data** menu bar tab, select **Data Analysis** from the **Analysis** group, and from the ensuing dialogue box, choose **t-test: Two-Sample Assuming Equal Variances** and click **OK**. A further dialogue box opens.
7. In the **Variable 1 Range** box, enter the cell range where the Diet A weight losses can be found (B2:B51), and in the **Variable 2 Range** box, enter the cell range where the Diet B weight losses can be found (B52:B101). Ensure that the **Labels** option is unchecked.
8. Type: 0 in the **Hypothesised Difference** box. In the **Alpha** box, ensure that 0.05 is entered (although this is relatively unimportant as we are going to use p-values). Click the **Output Range** button, and enter the cell reference H17 in the corresponding box. Then click **OK**.
9. Some output appears. Widen columns H to J to render it legible.
10. In cell H32, type: Difference in Means, and in cell I32, enter the formula =I20-J20.

The output is as follows:

t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	Variable 2
Mean	5.3412	3.70996
Variance	6.429280612	7.66759359
Observations	50	50
Pooled Variance	7.048437101	
Hypothesized Mean Difference	0	
df	98	
t Stat	3.072143179	
P(T<=t) one-tail	0.001375772	
t Critical one-tail	1.660551218	
P(T<=t) two-tail	0.002751544	
t Critical two-tail	1.984467404	
Difference in means	1.63124	

The obtained independent samples  $t = 3.072$  with 98 degrees of freedom.

The associated two-tailed p-value is  $p = 0.0028$ , so the observed  $t$  is significant at the 1% level (two-tailed).

The sample mean weight losses for Diets A and B were, respectively, 5.341 kg and 3.710 kg.

The data therefore constitute strong evidence that the underlying mean weight loss was greater for Diet A, by an estimated  $5.314 - 3.710 = 1.631$  kg. The results strongly suggest that Diet A is more effective in producing a weight loss.

### Exercise 8.6

Consider the bank cardholder data of Data Set C. Open the Excel workbook **Exe8.6C.xlsx** which contains this data from the Exercises folder.

Assuming the data to be suitably distributed, complete an appropriate test of whether the population mean income for males exceeds that of females and interpret your findings. What assumptions underpin the validity of your analysis, and how could you validate them?

### Answer

I will test the following hypothesis:

$$H_0: \mu_1 \leq \mu_2 \text{ against } H_1: \mu_1 > \mu_2$$

I we assume that the two sample distributions are normal and that their variances are equal. To verify that, first, I tested the skewness and kurtosis of the two samples, and the result showed reasonable skewness and kurtosis of the two samples assuming normality. Second, I did Two-Sample F-Test for the variances ratio, and the result was not significant (p-value = 0.22). This means we can assume that the variances of the populations are equal.

After that one-tailed t-Test was performed. The result was significant (p-value = 0.04). This means that the null hypothesis should be rejected. So, males have an average income that is greater than females by an estimated £ 5250.