Research Methods 2

Week 11: Exercise Sheet 1

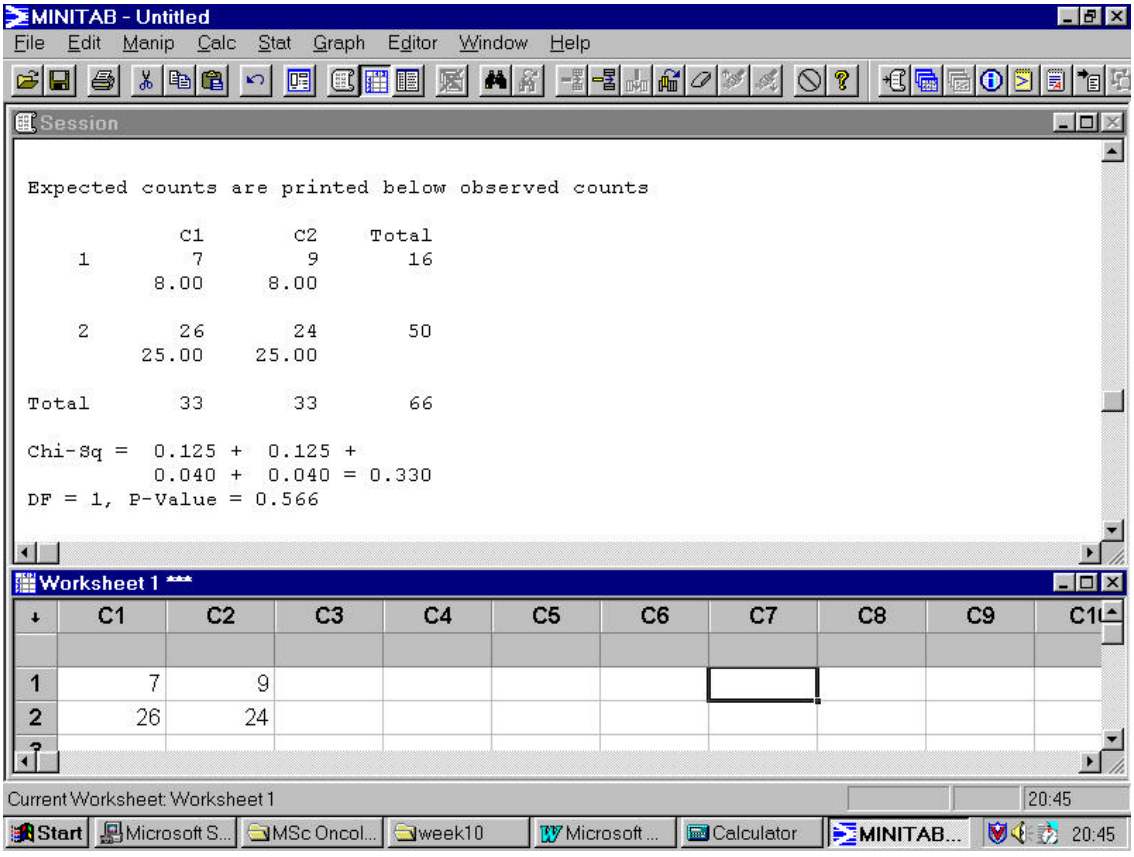
Solution sheet

Question 1

The table of observed values is as follows.

	Steroids	Placebo
Baby developed RDS	7	9
Baby did not develop RDS	26	24

The sample proportions developing RDS are 7/33 = 21% on steroids and 9/33 = 27% on placebo. To assess if these are indicative of a difference in the population proportions a χ^2 test is performed. To do this in Minitab, enter the data: the first column of the observed table is entered in one column, say C1, and the second column is entered in another column, say C2. The click on **Stat** -> **Tables** -> **Chi-Square Test...**. Then select the two columns into the **Columns containing the table:** box. The click on **OK**. The following screen is obtained.



The P value is 0.57, so the observed difference between the sample proportions could well occur by chance, even if the population proportions were identical. These data provide no evidence against the null hypothesis.

The table containing the percentages in each group is

	Steroids	Placebo
Baby developed RDS	21%	27%
Baby did not develop RDS	79%	73%

Question 2

The observed table is

	Steroids	Placebo
Baby developed RDS	700	900
Baby did not develop RDS	2600	2400

and the P-value obtained from the c^2 test is now 0.000 (this is an artifact of the way Minitab prints its P-values, it should be interpreted as P < 0.0005).

The table of percentages is

	Steroids	Placebo
Baby developed RDS	21%	27%
Baby did not develop RDS	79%	73%

Which is, of course, identical to the one obtained in question 1.

In question 1 the difference between the *sample* proportions is 0.21 –0.27 = 0.06. What the P-value tells us is that with samples of this size this difference can readily be due to chance. E.g. 7/33 and 9/33 are likely to vary substantially about the corresponding population proportions and the difference between them could have arisen even if the null hypothesis were true.

When the sample sizes are increased by a factor of 100 then the variation about the true value will be reduced and it is much less likely that the same difference in proportions can be ascribed to chance. Hence the smaller P-value. While the proportions are the same in the two questions, the SEs are much smaller in question 2.

If the analyses had been conducted on the percentages then, as the tables of percentages are the same in both questions, the results of the analyses in the two questions would have to be identical. This is clearly inappropriate. It is therefore essential that c^2 tests are performed on the actual numbers, not percentages.

End of solution sheet