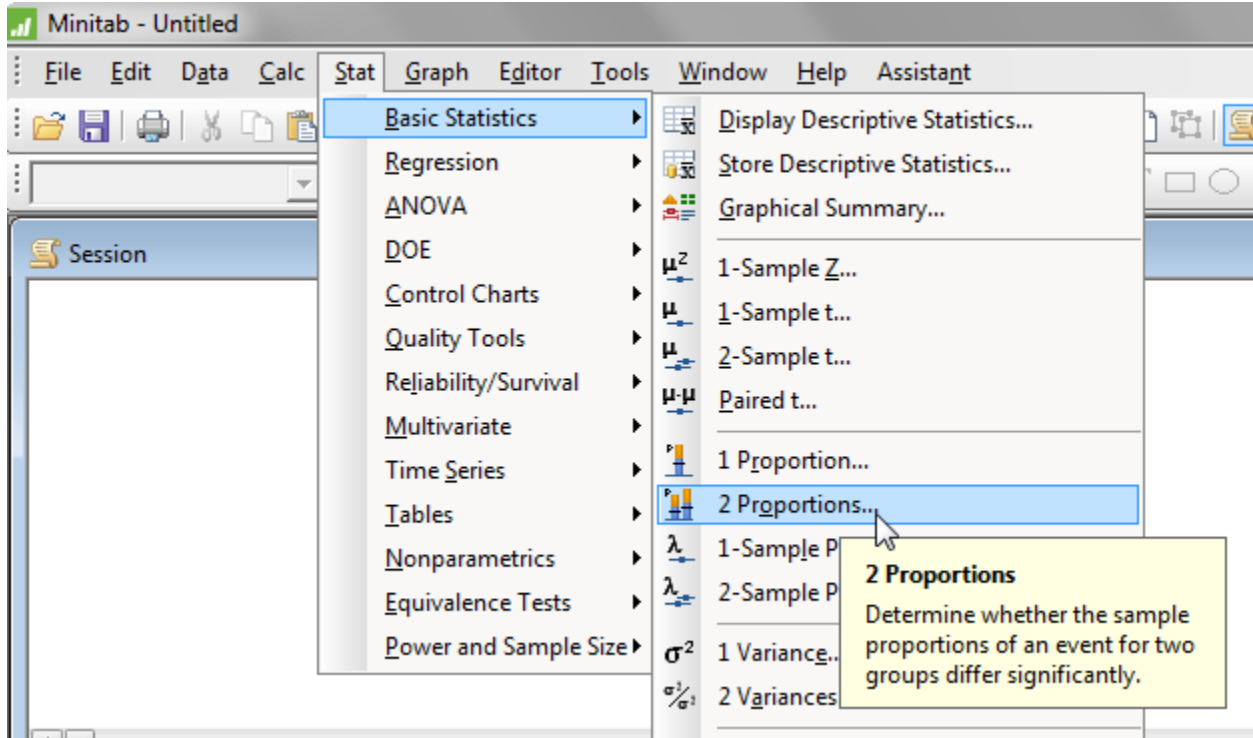# 2-Proportion command in Minitab 17

for confidence intervals and hypothesis tests on two proportions

Please read the material on the 2-sample-t command before reading this. There are many similarities.

There are three ways that the data can be read in:

- Both samples are in one column
- Each sample is in its own column
- Summarized data

Here is the worksheet with two different sets of data.
Columns 1 and 2 have both samples in one column. The data is in C1 and the Group is in C2.
Columns 4 and 5 have each sample in its own column.

↓	C1-T	C2-T	C3	C4-T	C5-T
	Responses	Group		Group A	Group B
4	No	Species 1		Yes	No
5	Yes	Species 1		No	No
6	No	Species 2		No	Yes
7	No	Species 3		Yes	Yes

In this handout, we will focus only on entering summarized data.   (The others are done in a similar way to the corresponding situations in 2-Sample-t.)

We will test Ho: $p_1 = p_2$ **versus Ha:** $p_1 \neq p_2$.

Since this is a hypothesis test, we use the pooled proportion to estimate the standard deviation, so choose that in the dialog box.

We aren't doing a confidence interval so the confidence level isn't relevant.   Our Ho says the two proportions are equal, so that says the "hypothesized difference" is 0. As that is the default value already in this box, we leave it.
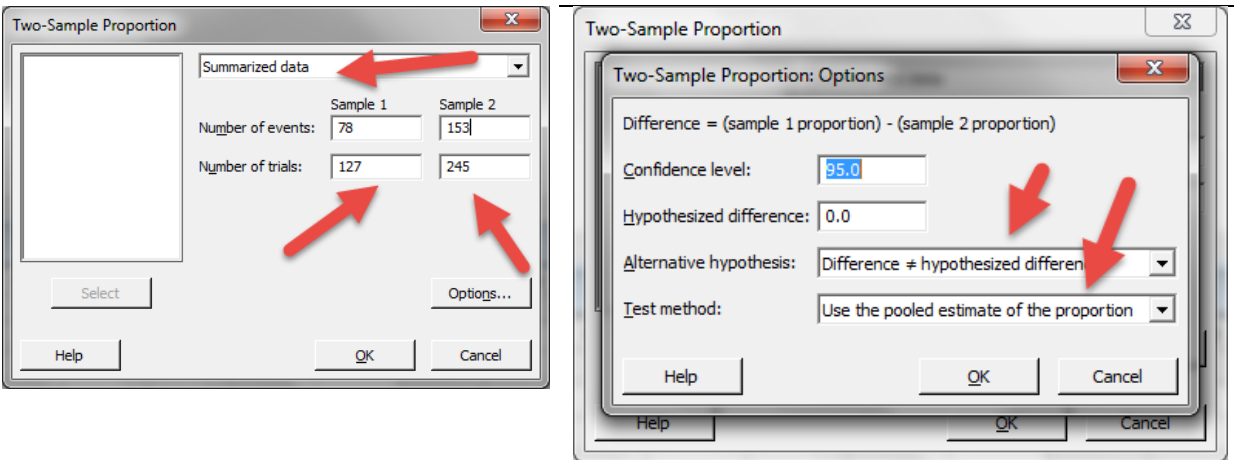
The output gives a very large p-value, which isn't a surprise because the two sample proportions are very close.   The output gives a confidence interval, but that is marked out here because you should not use it.   The method we chose, of pooling the data to estimate the proportion for the standard deviation, is appropriate for a hypothesis test, but NOT for a confidence interval.

```
Sample    X     N   Sample p
1        78   127   0.614173
2       153   245   0.624490


Difference = p (1) – p (2)
Estimate for difference:   -0.0103166
                                (-0.14......, .....
Test for difference = 0 (vs ≠ 0):   Z = -0.19   P-Value = 0.846

Fisher's exact test: P-Value = 0.910
```

 The output also gives a second p-value from a different test that could be used on these data.   We did not learn about Fisher's exact test in our course, so do not quote this result.

Find a 99% confidence interval for the difference of p_1 and p_2.

In the dialog box, we will put in the 99%, make sure the alternative hypothesis is "not equal", and change the option to "Estimate the proportions separately."
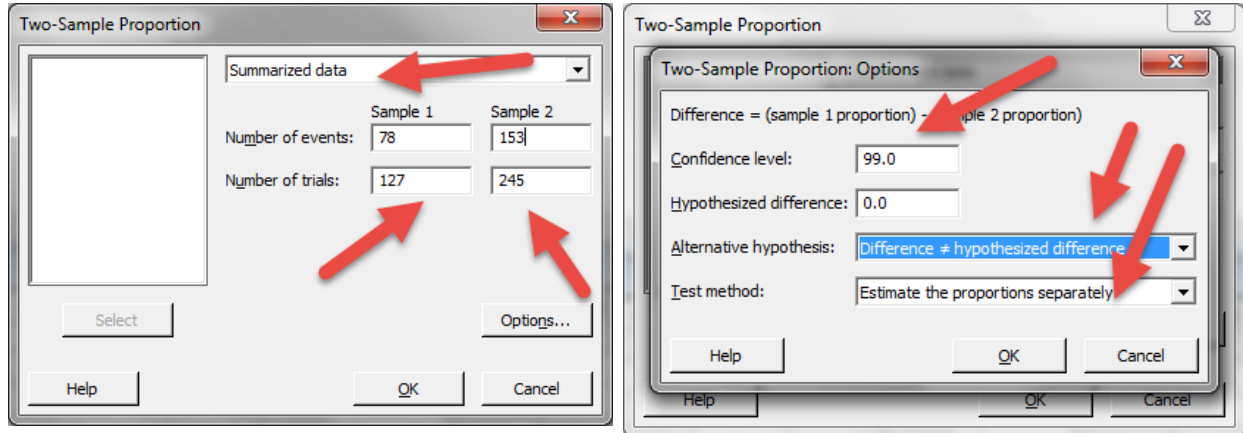
In the output, we see the confidence interval.   Since we didn't choose "pool" option, the hypothesis test conclusion here is not fully correct, so that is marked out.

Test and CI for Two Proportions

```
Sample    X    N   Sample p
1        78  127   0.614173
2       153  245   0.624490


Difference = p (1) - p (2)
Estimate for difference:   -0.0103166
99% CI for difference:   (-0.147176, 0.126543)
```