7.1. *z* test with unmatched data

EXAMPLE 7.1. The evolution of *Littorina littoralis* at Aberystwyth, 2002

BOX 7.1. How to carry out an *F* test to check for homogeneous variances before carrying out a *z* test for unmatched data

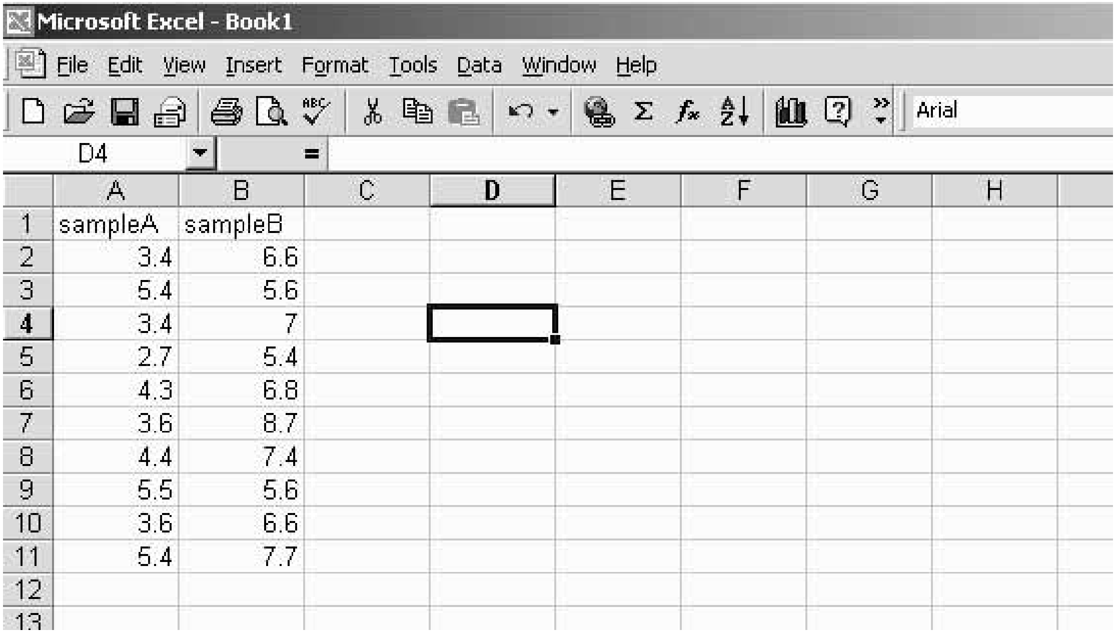
Step 1. Parametric statistical tests are accessed through the 'Analysis Tool Pack', which is activated via the add-in option under 'Tools' on the tool bar. Select 'Tools'. On the drop-down menu, select 'Add-ins' and from the dialogue box check the square for 'Analysis' toolkit. Returning to the 'Tools' option, you should now find that 'Data Analysis' is added to the drop-down menu. You are now ready to complete a parametric test.

Step 2. Entering data.
Start with a new spreadsheet: select 'File' and 'New' from the drop-down menu.

The data consist of two samples, each with a number of observations (more than 30 for a z test). The numbers of observations within each sample do not have to be the same.

Enter the data in columns. Select 'Format' and 'Cells' from the drop-down menu. In the dialogue box, select the 'Number' tab to select a category, e.g. 'Scientific', from which you specify the number of decimal places required. It is possible to use 'Labels'. These are the names (or labels) for your samples.

Start in cell 'A1' and label the first sample data set and continuing below, enter the numerical data. Enter the second sample data in the next column (B).



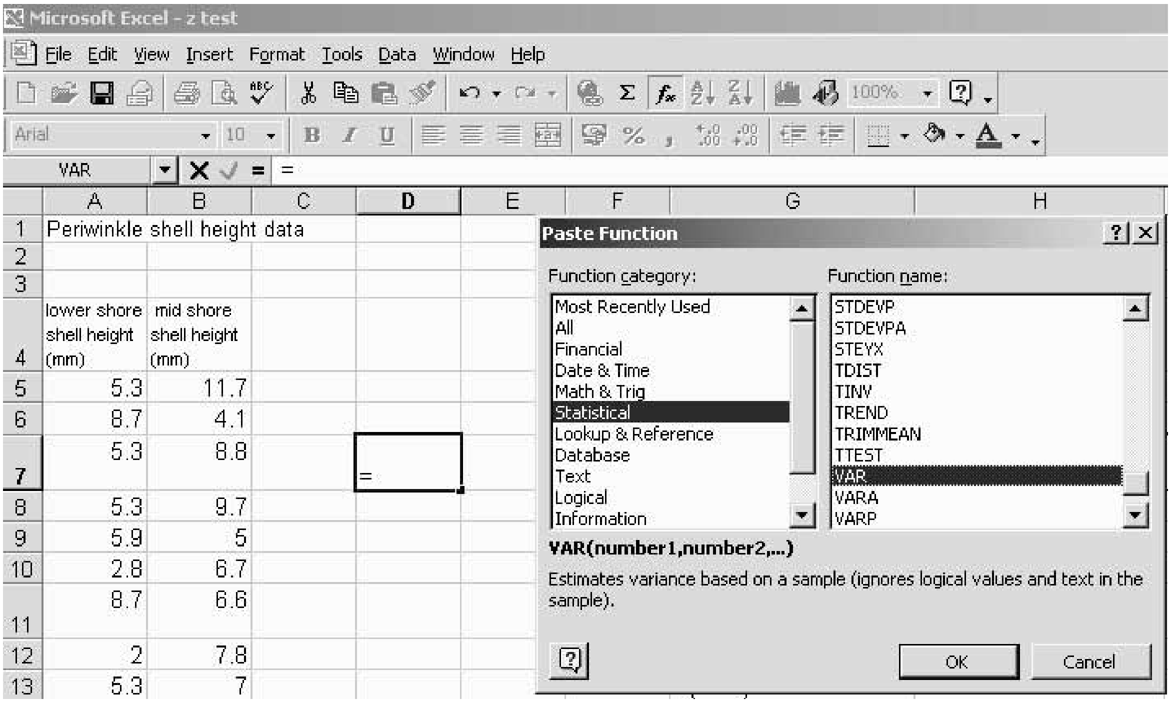
	A	B	C	D	E	F	G	H
1	sampleA	sampleB						
2	3.4	6.6						
3	5.4	5.6						
4	3.4	7						
5	2.7	5.4						
6	4.3	6.8						
7	3.6	8.7						
8	4.4	7.4						
9	5.5	5.6						
10	3.6	6.6						
11	5.4	7.7						
12								
13								

Next, select ‘Tools’, then ‘Data Analysis’ from the drop-down menu.

A dialogue box will open – here you select the test that you want to use, enter the data sets and select where the output from the test will be returned.

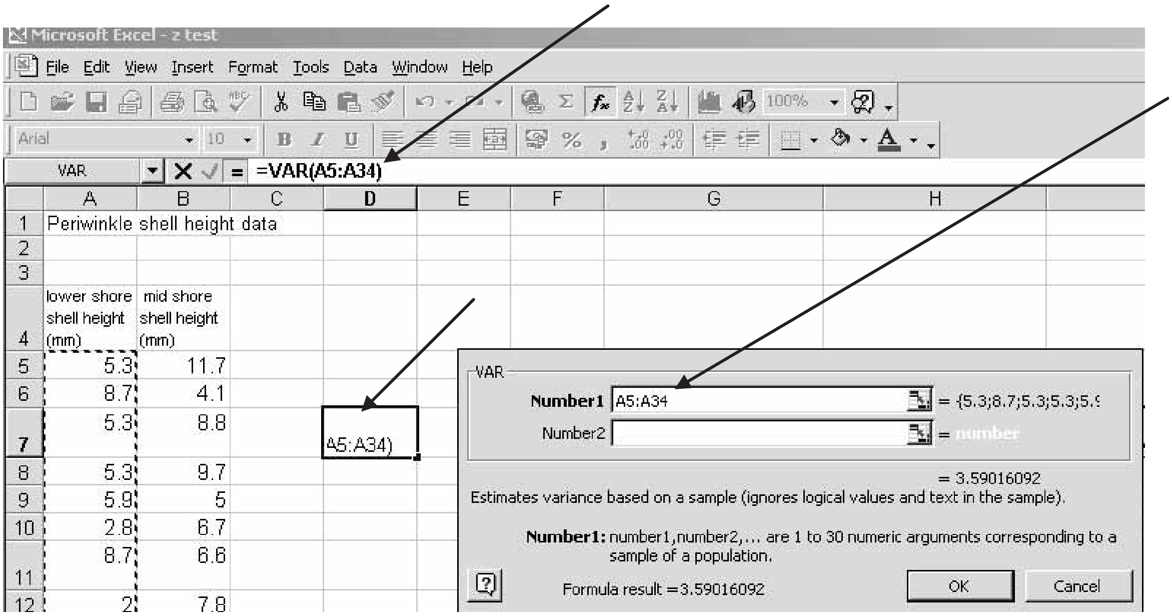
Step 3. The *F* test (BOX 7.1.) is used prior to the *z* test to ensure that the variances of the two data sets are similar. The *F* test can be performed indirectly by using the f_x paste function to find the variance for each sample.

Using Example 7.1. ‘The evolution of *Littorina littoralis* at Aberystwyth, 2002’, enter the data in columns, one for each sample. A title for the sheet is placed at the top in the figure to aid identification of the data but does not play any part in the calculation. Data labels are also entered in cells A4 and B4 to identify the samples. These are ignored when using the f_x paste function, but can be incorporated into output data when using the ‘Data Analysis Tools’.



Select an empty cell on the spreadsheet, say D7, in which the variance for the ‘lower shore shell height’ data will be returned. By pointing the cursor at the cell and clicking, the cell will become highlighted. Then click on f_x on the tool bar and the ‘Paste Function’ dialogue box will open. Note too that an equal sign (=) appears in cell D7.

In the dialogue box, select ‘Statistical’ from the ‘Function category’ list and ‘VAR’ from the ‘Function name’ list. Click on ‘OK’. A box will now open in which the cells containing the data can be identified.



The cells are identified by clicking on cell A5 and holding down the left mouse button, dragging down to cell A34 (the cell containing the last datum). The cells will become highlighted and notice that the cell locations now show in the cell D7 as well as in the dialogue box and on the formula bar. Clicking OK will return the variance value to the cell D7 (3.590). Repeat this process for the second set of data selecting a new cell location, say E7, in which to perform the calculation. The variance for this data set is 4.097.

To complete the *F* test, select a new cell and place a ‘ = ’ sign in it. This indicates to Excel that a calculation is to be performed. Next click on the cell with the higher variance value, cell E7, then enter ‘/’ to indicate that a division will be performed, then click on cell D7 (the value of the smaller variance), lastly press ‘Return’ and the *F* statistic will be given. The value of the *F* statistic is 1.1413. This then needs to be compared with the critical value for *F* at $p = 0.05$ in statistical tables. The value of the critical value of *F* is 2.10 at 29 degrees of freedom for both samples. The calculated value of *F* is less than the critical value, therefore the variances are similar. You are justified in using the *z* test to analyse these data.