

6.3. Pearson's product moment correlation

EXAMPLE 6.1. The hardness of egg shells in pullets

BOX 6.2. How to carry out a Pearson's product moment correlation

Step 1. Where you have two data sets that are likely to be correlated, a scatter plot should first be drawn to check the general distribution of data. This is done using the chart wizard as previously described (see 'Getting Started with Excel').

The data below relate to the hardness of egg shells produced by 13 Maren pullets and their consumption of a food supplement (g) (Fig. W6.1.). The test will derive a value for the strength of the association between these two variables.

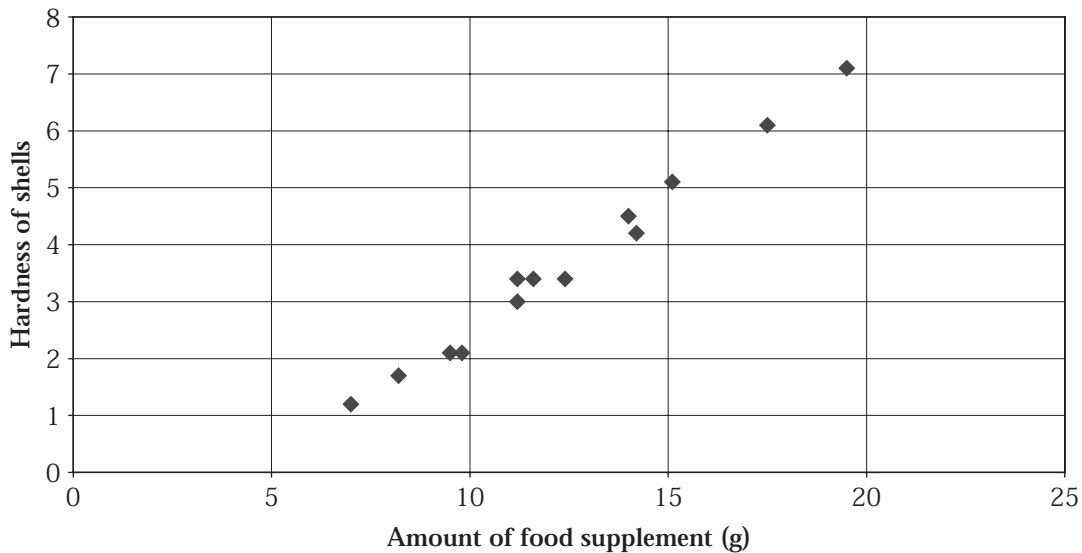


Fig. W6.1. Association between the hardness of egg shells produced by 13 Maren pullets and their consumption of a food supplement (g)

On inspection of the plot, there appears to be a positive linear association in the data, so it is worth continuing with the investigation and calculating the correlation coefficient.

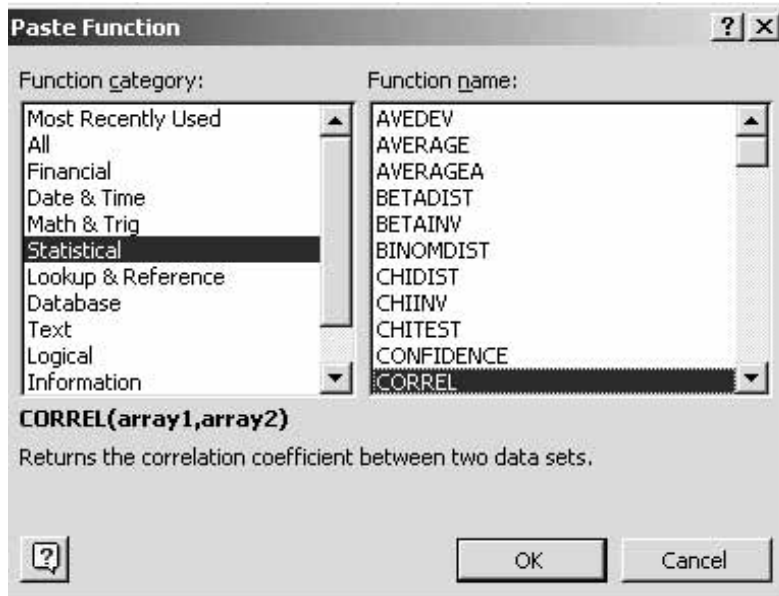
Step 2. In Excel, data must be entered in columns. The first column labelled 'Pullet' is present only for reference and takes no part in the calculation, so could be omitted.

Table 6.1. The hardness of egg shells produced by 13 Maren pullets and their consumption of a food supplement (g)

Pullet	Amount of food supplement (g)	Hardness of shells
1	19.5	7.1
2	11.2	3.4
3	14	4.5
4	15.1	5.1
5	9.5	2.1
6	7	1.2
7	9.8	2.1
8	11.6	3.4
9	17.5	6.1
10	11.2	3
11	8.2	1.7
12	12.4	3.4
13	14.2	4.2

Step 3. There are two methods for deriving the correlation coefficient on Excel: one uses the f_x paste function and the other, the Data Analysis Tools.

Using ' f_x Paste Function' from the tool bar, first, select a cell on the spreadsheet (by clicking on it) where you want the result to be returned. Select 'Statistical' from the 'Function category' and 'CORREL' from the 'Function name', as shown below. Click on 'OK'.



A dialogue box opens in which the cell locations of the data can be entered. Note that when using the f_x paste function, the cell locations for the data labels must not be entered. This function only works with numerical data.

Microsoft Excel - Book3

File Edit View Insert Format Tools Data Window Help

CORREL X ✓ = =CORREL(B3:C14+B3:B15)

	A	B	C	D	E
1	Table 6.1. The hardness of egg shells produced by 13 Maren pullets and their consumption of a food supplement (g)				
2	Pullet	Amount of food supplement (g)	Hardness of shells		
3	1	19.5	7.1		
4	2	11.2	3.4		
5	3	14	4.5		
6	4	15.1	5.1		
7	5	9.5	2.1		
8	6	7	1.2		
9	7	9.8	2.1		
10	8	11.6	3.4		
11	9	17.5	6.1		
12	10	11.2	3		
13	11	8.2	1.7		
14	12	12.4	3.4		
15	13	14.2	4.2		

CORREL

Array1 B3:C14+B3:B15 = {39,26.6;22.4,14.6;

Array2 = array

=

Returns the correlation coefficient between two data sets.

Array1 is a cell range of values. The values should be numbers, names, arrays, or references that contain numbers.

Formula result =

OK Cancel

In the dialogue box, ensure that the cursor is flashing in the box marked 'Array 1'. Enter the cell locations for the first set of data in column B by clicking on the first cell with data (B3) and, holding down the left mouse button, drag down the column. The cell locations will be entered into the

array box. Click in the box marked 'Array 2' to bring the cursor to this location. Then enter the second set of data in column C in the same way.

After the second set of data is inputted, the correlation coefficient is shown on the dialogue box (and in this case has the value of 0.993608288).

CORREL

Array1 B3:B15 = {19.5;11.2;14;15.1;
Array2 C3:C15 = {7.1;3.4;4.5;5.1;2.1

= 0.993608288

Returns the correlation coefficient between two data sets.

Array2 is a second cell range of values. The values should be numbers, names, arrays, or references that contain numbers.

Formula result = 0.993608288

OK Cancel

Click on 'OK' and the value will be returned to the selected cell on the spreadsheet.

Using the Data Analysis Tool Pack: the data are entered onto the spreadsheet in the same manner as described above. Select 'Tools' from the tool bar, then 'Data Analysis'. From the dialogue box, select 'Correlation' and click OK. A dialogue box will open.

In the 'Input' box, enter the cell locations for the data for both variables, including the data labels. This is done by clicking on the first cell B2 and dragging across the entire set of data. Data are grouped by columns so ensure that the circle is checked by clicking on it. Also click in the box for 'Labels in First Row' and the tick will appear.

Correlation ? X

Input

Input Range: \$B\$2:\$C\$15

Grouped By: Columns Rows

Labels in First Row

Output options

Output Range: \$B\$17:\$C\$28

New Worksheet Ply:

New Workbook

OK Cancel Help

In 'Output Options', select the option you require. For 'Output Range', remember to click in the box so that the cursor appears here before dragging across the area where you wish the output to be located. You can click on just one cell as Excel will determine the amount of space it requires. Click 'OK'.

	<i>Amount of food supplement (g)</i>	<i>Hardness of shells</i>
Amount of food supplement (g)	1	
Hardness of shells	0.993608288	1

The correlation coefficient is given in the cell relating to the Hardness of shells. Note where the same variable coincides a value of 1 is given.

Step 4. The significance of the correlation coefficient must be checked using the appropriate statistical table.