NORMAL QUANTILE PLOTS

The purpose of a normal quantile plot is to verify if the data gathered comes from a normally distributed population. If the scatterplot generated is nearly linear then we can be fairly certain that the data indeed comes from a normally distributed population.

Generating a Normal Quantile Plot Using Excel

Excel does not automatically generate a normal quantile plot. However, the correct functions are available in Excel to create one.

1) The first step is to enter your data in a column. Type the data in column C, you will see why later, figure 1. After the data is entered, highlight the data. Then go to the toolbar and select the sort option, figure 2.

2) Once your data is sorted, next to the data, type the numbers 1, 2, 3 and so on, enumerating the data, to the left of the column containing the data. Make sure the number and the corresponding data value are on the same row. Leave a column empty between the two columns.

If there is a lot of data values you can automate the enumerating by typing 1 in the first row, see figure 3. Then, in the cell below type = a2 + 1, and press Enter (note that a2 corresponds to the cell location of the first data value).

Now click on cell a3, you will notice a box on the lower left corner of the highlighted rectangle, move your mouse to that position until the arrow turns into a cross, as shown in figure 4. Click down, on the right mouse button, hold it, then move the mouse down to highlight as many cells as you need below. When you let go of the button, all the cells highlighted will now have the command in cell a3.

In the column between the sorted data and the numbers that enumerate, type

\[ \text{=normsinv((a2-3/8)/(n+.25))}. \]

Note: the \( n \) refers to the total number of data values you have. So if you have 20 data values the n should be 20. If you see figure 5, you will notice I typed a 7, since I have seven data values.

Now, copy this function to the rest of the cells, using the same technique as used in figure 4. What this function does is give you the corresponding z-score for the particular position of the data value (figure 6).

Lastly, create a scatterplot using the columns B and C the z-score and data. The z-scores should come first then the actual data values. Refer to section 1.3, page 80, of your book for more explanation.

To create the scatterplot highlight the two columns containing the information; in my case B and C. Got to the tool bar and press the button with the bargraph, figure 7. Choose scatter. The first scatterplot highlighted on the right is the correct one. Click next, and type the appropriate titles for the axis and the chart. Click Finish.