1 Checking for Normality

Not all data sets are normal. Even for data sets that have a "bell shape" to them; there could be more area in the tails (fatter tails) than the normal distribution.

Have your data set in Excel in Column A where your first data point is in Cell A1. First, use the sort function to order you date from smallest to largest.

We use the formula $\frac{i}{n+1}$ for the percentile of the ith piece of data. In Cell B1, put $= \frac{1}{n+1}$, where n is the number of data values. In cell B2, put $=$ B2 $+ \frac{1}{n+1}$. Copy B2 down to the last row of the data.

The next step will convert the the numbers in Column B to standard units. Excel can do this for us. In C1, put = NORM.S.INV(B1). Copy this down to the last data point.

Now plot Column A against Column C to check for linearity. Highlight Column A and Column C. (You need to use either the control button or the command button.) Use the Scatterplot tool to make a scatterplot.

If your dots form a line, then your data is normal. If your dots are off the line, then your data is not normal.

Mail to rstephens@colgate.edu