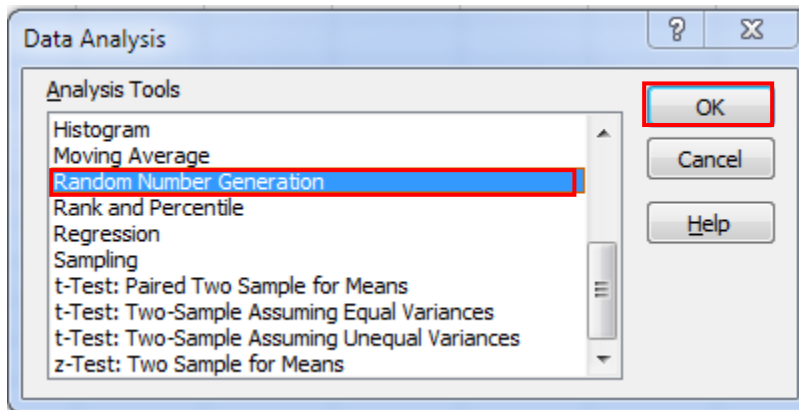


Probability Plot Tutorial

Probability plot (also known as normal probability plot, NPP) is a tool to determine whether the data follows normal distribution or not. It plots the measured values on the abscissa and the predicted values using normal distribution probability density function on the ordinate. It is quite simple and effective tool.

First, we need to create some measurement to demonstrate the procedure to make the plot. In this tutorial, we will use Random Number Generator to generate 75 random numbers that are normally distributed. You can also use Randbetween(bottom,top) function from worksheet functions library.

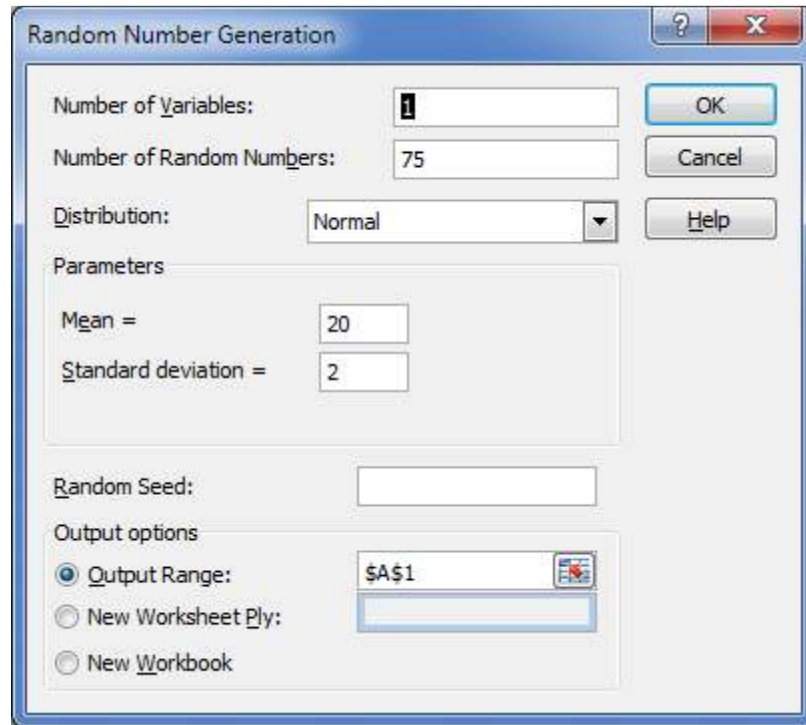
Click **Data/Data Analysis** and select **Random Number Generator** and click **OK**. If you cannot find the Data Analysis ToolPak, please refer to [Appendix 1](#) of this tutorial to enable the Analysis ToolPak.



Excel will prompt you with Random Number Generator dialog box. Input the following information into its perspective input boxes.

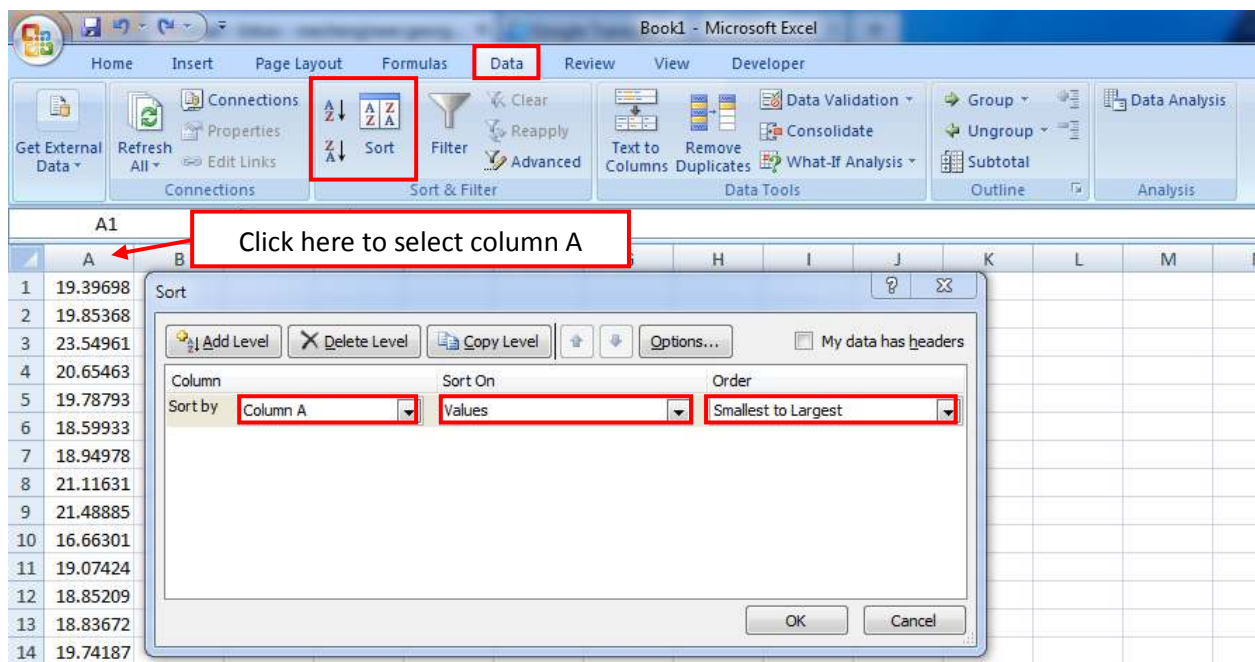
- Number of Variables = 1 “this refers to the number of columns”
- Number of Random Number = 75 “this refers to the number of measurements”
- Distribution = Normal “this refers to normal probability density function i.e. generated data will be normally distributed.”
- Mean = 20 “this refers to the sample mean or the central value of the bell-curve”
- Standard deviation = 2 “this refers to the variation in the generated data”
- Output Range = \$A\$1 “generated data will be stored in the active sheet starting at cell A1”

After entering all information above press OK. The generator will fill in A1:A75 with normally distributed random numbers. Now, we are ready to demonstrate NPP.



Probability plot procedure:

- 1- Sort data from ascending order “Smallest to Largest”. Click Data tab/Sort and set Sort by =Column A, Sort on = Values and Order Smallest to Largest

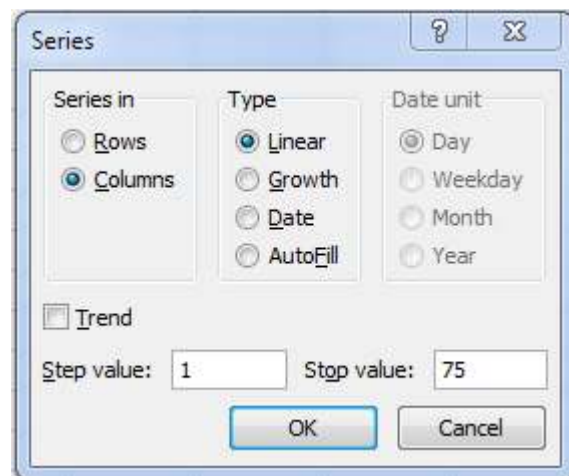


- 2- Fill in a series in column B. This series should represent the count, however if you use 1,2,3.....,75, you will get an error later when you are calculating the probability. For this reason, we will start our series at any number less than 1. In cell B1, insert a value of 0.9. While you are selecting cell B1, click Home tab/Fill/Series.



Excel will prompt you with Series dialog box. Input the following information into its perspective input boxes.

- Series in = Columns
- Type = Linear
- Step Value = 1
- Stop Value = 75



- 3- Calculate the percentile of each measurement by dividing the count by the total number of samples. To do so, in cell C1 insert the following formula, then double-click on the right-lower corner of C1 to populate all values in column C.

$$=B1/\text{count}(\$A\$1:\$A\$75)$$

Note: to fix the reference press F4.

	A	B	C	D	E	F	G
1	14.43229	0.9	0.012				
2	16.11682	1.9					
3	16.25443	2.9					
4	16.31456	3.9					
5	16.66301	4.9					

Formula bar: $=B1/\text{COUNT}(\$A\$1:\$A\$75)$

Note the formula

Double click to fill in values in column C

- 4- Calculate the predicted value. Recall the equation from lecture notes:

$$x_i = z\sigma + x'$$

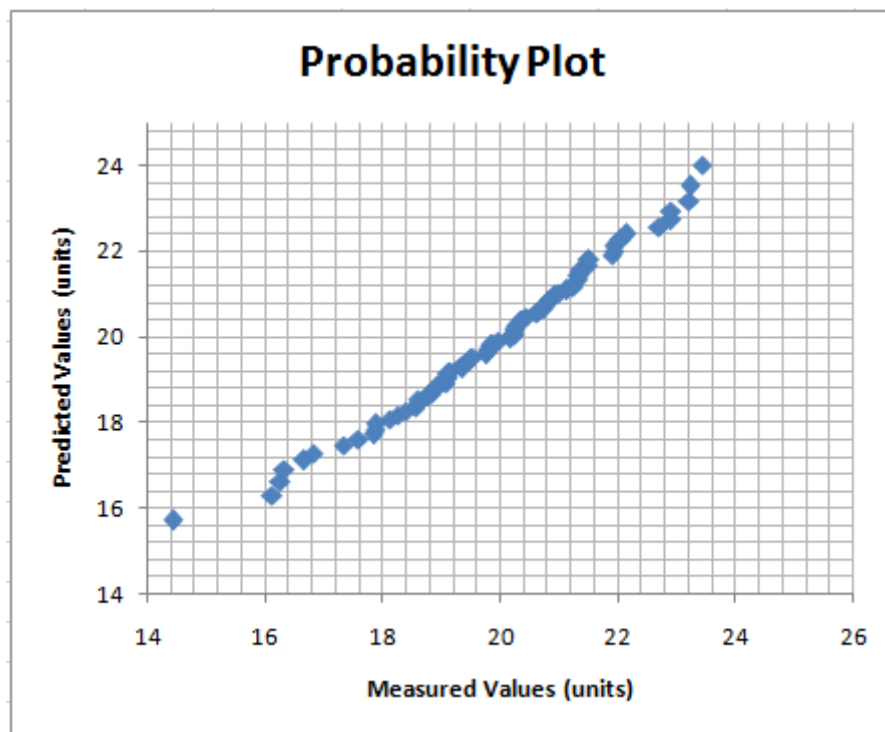
where, σ is the standard deviation which was set to 2 in Random Number generator and x' is the mean and was set to 20.

Note: you can calculate the mean and standard deviation in Excel using `=average(range)` and `=stdev(range)` functions, respectively.

To calculate the z values use `=NORMINV(cell,mean,standard_deviation)`, which calculate the inverse of the normal distribution. In cell D1 enter the following formula, then double-click on the right-lower corner of D1 to populate all values in column D.

`=NORMINV(C1,AVERAGE(A1:A75),STDEV(A1:A100))`

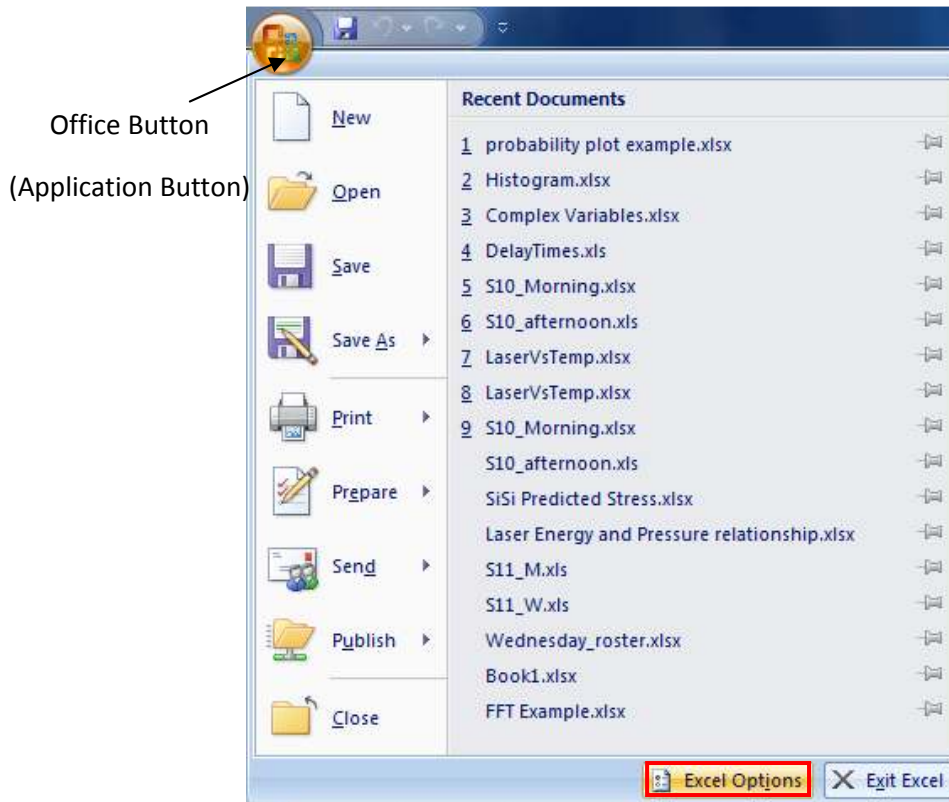
- 5- Create a Scatter plot (with only Markers) of values in column A (measured values) and column D (predicted values). Note that the plot lie almost on the diagonal because measured values (column A) were generated to be normally distributed.



Appendix 1: Data Analysis ToolPak

By default, this add-in is not activated. Therefore the first step is to activate the Analysis ToolPak.

Go to **Office Button** (application Button)/**Excel Options**



Select the **Add-Ins** Tab, then select **Excel Add-Ins** from the **Manage** dropdown menu located at the bottom of the dialog box and click **Go**.

Excel will prompt you with another dialog box “**Add-Ins**” to select the analysis tools you want to activate. Check the box beside **Analysis ToolPak** and **Analysis ToolPak-VBA** and click **OK**.

You don't have to perform these steps every time you use the data analysis tools, you follow these steps if you can't access the Data Analysis tools, which usually in the first time you use PC with freshly installed Office or the first time you use the ToolPak.

