ASSIGNMENT 01: NORMAL PROBABILITY PLOTS AND CENTRAL LIMIT THEOREM

Objectives:

1. To gain experience using the Excel Data Analysis Toolbox features
2. To gain experience constructing and interpreting Normal Probability Plots
3. To explore the theory underlying Normal Probability Plots, and the Central Limit Theorem

Practice:

Download the Excel file "Normal Plots"

1. Perform the following operations to use the Normal Probability Plot tabs (use the Working tab):
    a. Click on the following in the Menu toolbar:
        Tools > Data Analysis > Random Number Generation
    b. Enter the appropriate parameters, ex:
        i. *Number of variables* = 1
        ii. *Number of random numbers* = 20
        iii. *Distribution* = Normal
        iv. *Mean* = 12
        v. *Standard Deviation* = 6
        vi. *Random Seed* = 4321
        vii. *Output Range* (selected) = $A$6:$A$25
        viii. Select the *OK* button. It is also *OK* to overwrite the existing data.
    c. After the random numbers have been generated, select the region from A6 to A25, and click on the Copy icon (or press the keyboard combination *Ctrl-C*)
    d. Copy these numbers to the column for Sorted Data by clicking in cell *F6* and then going to the Menu toolbar and clicking *Edit > Paste Special > Values*.
    e. Using the top toolbar, click on the *Sort Ascending* icon to order the random numbers from lowest to highest. Excel will open a warning box because it finds more numbers adjacent to the data – select the button for *Continue with current selection* and then press the *Sort* button.
    f. Observe the adjacent plot window to interpret your Normal Probability Plot, as directed.

2. On the N(12,1) tab, generate a normal probability plot for 20 random numbers from the following distribution:
    a. N(12,1)
        i. *Number of variables* = 1
        ii. *Number of random numbers* = 20
        iii. *Distribution* = Normal
        iv. *Mean* = 12
        v. *Standard Deviation* = 1   (Note change!)
        vi. *Random Seed* = 4321
        vii. *Output Range* (selected) = $A$6:$A$25
    b. Observe your data, and drag the handles on the Pink Line to cross your data at the approximate 25%-ile and 75%-ile points. How does this plot compare to the dark line for

the N(12,1) plot? Comment on what is different, what is the same, and on the apparent outcome for the "Fat Pencil Test". Do these observations make sense?

3. On the N(14,1) tab, generate a normal probability plot for 20 random numbers from the following distribution:
   a. N(14,1)
      i. *Number of variables* = 1
      ii. *Number of random numbers* = 20
      iii. *Distribution* = Normal
      iv. *Mean* = 14   (Note change!)
      v. *Standard Deviation* = 1   (Note change!)
      vi. *Random Seed* = 4321
      vii. *Output Range* (selected) = $A$6:$A$25
   b. Observe your data, and drag the handles on the Pink Line to cross your data at the approximate 25%-ile and 75%-ile points. How does this plot compare to the dark line for the N(12,1) plot? Comment on what is different, what is the same, and on the apparent outcome for the "Fat Pencil Test". Do the slopes of the two lines appear similar? What does similar slope indicate? Do these observations make sense?

4. On the N(12,2) tab, generate a normal probability plot for 20 random numbers from the following distribution:
   a. N(12,2)
      i. *Number of variables* = 1
      ii. *Number of random numbers* = 20
      iii. *Distribution* = Normal
      iv. *Mean* = 12   (Note change!)
      v. *Standard Deviation* = 2   (Note change!)
      vi. *Random Seed* = 4321
      vii. *Output Range* (selected) = $A$6:$A$25
   b. Observe your data, and drag the handles on the Pink Line to cross your data at the approximate 25%-ile and 75%-ile points. How does this plot compare to the dark line for the N(12,1) plot? Comment on what is different, what is the same, and on the apparent outcome for the "Fat Pencil Test". Do the slopes of the two lines appear similar? Do these observations make sense for the specified parameters?

5. On the Binomial tab, generate a normal probability plot for 20 random numbers from the following distribution:
   a. Binomial(p=.6,trials=20)
      i. *Number of variables* = 1
      ii. *Number of random numbers* = 20
      iii. *Distribution* = Binomial   (Note change!)
      iv. *p-value* = 6   (Note: this is the probability of detection success, <u>not</u> p-value)
      v. *Number of Trials* = 20
      vi. *Random Seed* = 4321
      vii. *Output Range* (selected) = $A$6:$A$25

    b. Observe your data, and drag the handles on the Pink Line to cross your data at the approximate 25%-ile and 75%-ile points. How does this plot compare to your output for the N(12,2) plot? Comment on what is different, what is the same, and on the apparent outcome for the "Fat Pencil Test". Do the slopes of the two lines appear similar? Do these observations make sense?

    c. Create a histogram for the 20 data points. How close does the data appear to a normal distribution?

6. On the CLT-A tab, generate a normal probability plot for 20 random numbers from the following distributions:

    a. N(12,2)
  i. ***Number of variables*** = 1
  ii. ***Number of random numbers*** = 5   (Note change!)
  iii. ***Distribution*** = Normal
  iv. ***Mean*** = 12
  v. ***Standard Deviation*** = 2
  vi. ***Random Seed*** = 4321
  vii. ***Output Range*** (selected) = \$A\$6:\$A\$10   (Note change!)

    b. N(12,4)
  i. ***Number of variables*** = 1
  ii. ***Number of random numbers*** = 5
  iii. ***Distribution*** = Normal
  iv. ***Mean*** = 12
  v. ***Standard Deviation*** = 4   (Note change!)
  vi. ***Random Seed*** = 4321
  vii. ***Output Range*** (selected) = \$A\$11:\$A\$15   (Note change!)

    c. N(14,1)
  i. ***Number of variables*** = 1
  ii. ***Number of random numbers*** = 5
  iii. ***Distribution*** = Normal
  iv. ***Mean*** = 14   (Note change!)
  v. ***Standard Deviation*** = 4   (Note change!)
  vi. ***Random Seed*** = 4321
  vii. ***Output Range*** (selected) = \$A\$16:\$A\$20   (Note change!)

    d. N(14,6)
  i. ***Number of variables*** = 1
  ii. ***Number of random numbers*** = 5
  iii. ***Distribution*** = Normal
  iv. ***Mean*** = 14
  v. ***Standard Deviation*** = 6   (Note change!)
  vi. ***Random Seed*** = 4321
  vii. ***Output Range*** (selected) = \$A\$21:\$A\$25   (Note change!)

    e. Observe your data, and drag the handles on the Pink Line to cross your data at the approximate 25%-ile and 75%-ile points. How does this plot compare to the dark line for the N(12,2) plot? Comment on what is different, what is the same, and on the apparent

outcome for the "Fat Pencil Test". Do the slopes of the two lines appear similar? Does the Central Limit Theorem apply and does it appear to work?
    f. Create a histogram for the 20 data points. Does the data appear normally distributed?

7. On the CLT-B tab, generate a normal probability plot for 20 *__means__* of random numbers from the following distributions:
    a. N(12,2)
        i. *Number of variables* = 1
        ii. *Number of random numbers* = 20   (Note change!)
        iii. *Distribution* = Normal
        iv. *Mean* = 12
        v. *Standard Deviation* = 2
        vi. *Random Seed* = 4321
        vii. *Output Range* (selected) = $A$6:$A$25   (Note change!)
    b. N(12,4)
        i. *Number of variables* = 1
        ii. *Number of random numbers* = 20   (Note change!)
        iii. *Distribution* = Normal
        iv. *Mean* = 12
        v. *Standard Deviation* = 4   (Note change!)
        vi. *Random Seed* = 4321
        vii. *Output Range* (selected) = $B$6:$B$25   (Note change!)
    c. N(14,1)
        i. *Number of variables* = 1
        ii. *Number of random numbers* = 20   (Note change!)
        iii. *Distribution* = Normal
        iv. *Mean* = 14   (Note change!)
        v. *Standard Deviation* = 4   (Note change!)
        vi. *Random Seed* = 4321
        vii. *Output Range* (selected) = $C$6:$C$25   (Note change!)
    d. N(14,6)
        i. *Number of variables* = 1
        ii. *Number of random numbers* = 20   (Note change!)
        iii. *Distribution* = Normal
        iv. *Mean* = 14
        v. *Standard Deviation* = 6   (Note change!)
        vi. *Random Seed* = 4321
        vii. *Output Range* (selected) = $D$6:$D$25   (Note change!)
    e. Observe your data, and drag the handles on the Pink Line to cross your data at the approximate 25%-ile and 75%-ile points. How does this plot compare to the dark line for the N(12,2) plot? Comment on what is different, what is the same, and on the apparent outcome for the "Fat Pencil Test". Do the slopes of the two lines appear similar? Does the Central Limit Theorem apply and does it appear to work?
    f. Create a histogram for the 20 means. Do the means appear normally distributed?