EXPLORATIONS

Chapter 8

The Normal Distribution

Topic 18 discusses the graphing of normal distributions, shading desired areas, and finding related probabilities. Normal probability plots are introduced to check on the normal shape of a distribution of data.

Topic 18—The Normal Distribution

Graphing Normal Distributions

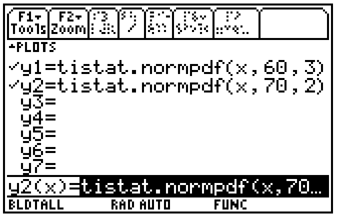
For this topic, use folder **BLDTALL**.

1. Press [MODE].
2. Press ▼ to select Current Folder, and then press ▶.
3. Highlight **BLDTALL** and press [ENTER].

Example: Two populations of heights are normally distributed $\mu_1 = 60$ inches, $\sigma_1 = 3$ inches, and $\mu_2 = 70$ inches, $\sigma_2 = 2$ inches.

Set up the Plots function.

1. Press [♦] [Y=].
2. Select **y1** and press [ENTER] to highlight the input line. Press [CATALOG] [F3] **Flash Apps** and select **normPdf(... tistat**.
3. Enter **x, 60, 3)**. The **60** and **3** represent the mean and standard deviation of the first distribution, respectively.

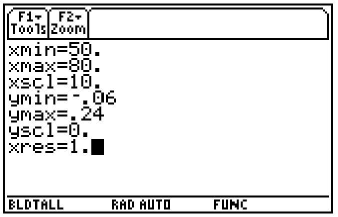
(1)


4. Set up **y2** the same as above, using the μ_2 and σ_2 values (screen 1).

5. Set up the window using [◆] [WINDOW] with the following entries:
   - **xmin = 50**
   - **xmax = 80**
   - **xscl = 10**
   - **ymin = -.06**
   - **ymax = .24**
   - **yscl = 0**
   - **xres = 1**

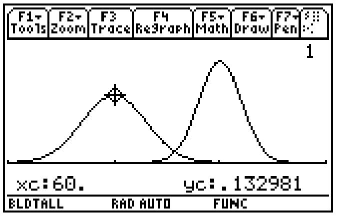
   (See screen 2.)

(2)



6. Press [◆] [GRAPH], and then press [F3] **Trace**.
7. Type **60** and press [ENTER] to display screen 3.

(3)



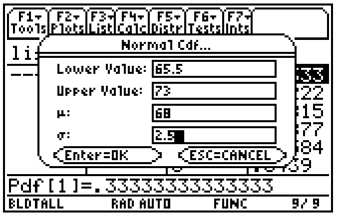
Finding Probabilities

Example: A population with heights normally distributed with $\mu = 68$ inches and $\sigma = 2.5$ inches.

What proportions of people in the population are between 65.5 inches and 73 inches tall? In other words, what is the probability that a person picked at random from the population is between 65.5 inches and 73 inches tall?

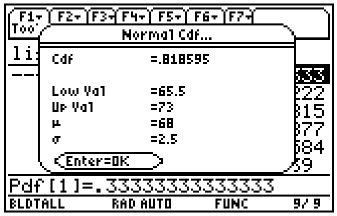
1. From the Stats/List Editor, press [F5] **Distr**, **4:Normal Cdf**. Enter the data given above (screen 4).

(4)

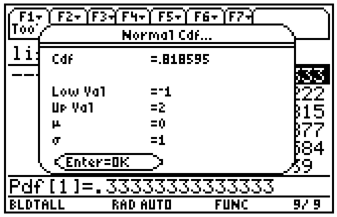


2. Press [ENTER] to display screen 5 with Cdf **= 0.818595 = p(65.5 < x < 73)**.

   You can repeat steps 1 and 2 by using the z values associated with the lower value: -1 or $\left(\dfrac{65.5 - 68}{2.5}\right)$ and upper value: 2 or $\left(\dfrac{73 - 68}{2.5}\right)$ with μ and σ cleared or with $\mu = 0$, $\sigma = 1$ of a standard normal curve to get the same results as in screen 6.



(5)



(6)

Shade Option

Example: As in the previous example, what is **p(65.5 < x < 73)**?

1. From the Stats/List Editor, turn off all functions and plots with [F2] **Plots**, **4:FnOff** and [F2] **Plots**, **3:PlotsOff**.
2. Press [F5] **Distr**, **1:Shade**, **1:Shade Normal**.
3. Re-enter the values:

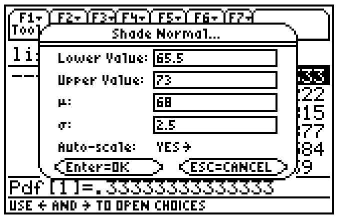
   Lower Value: **65.5**

   Upper Value: **73**

   µ: **68**

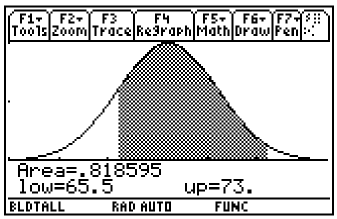
   σ: **2.5**

   Auto-scale: **YES**. (This is a required field. See screen 7.)



(7)

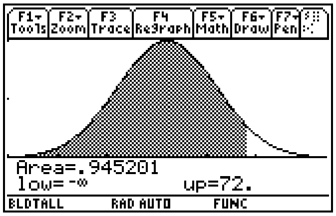
4. Press [ENTER] to display screen 8 with **p(65.5 < x < 73) = 0.818595**.



(8)

Example: What proportion of people in the population are less than 6 ft. tall (72 inches)?

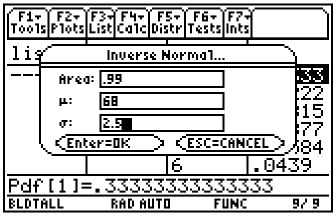
1. Press [2nd] [⊞] to return to the Stats/List Editor.
2. Set up Shade Normal as in the previous example, but with Lower Value: [(-)] [◆] [∞], Upper Value: **72**, μ: **68**, σ: **2.5**, and Auto-scale: **YES**.
3. Press [ENTER]. Screen 9 shows that nearly 95% of the population is less than 6 ft. tall.

(9) 

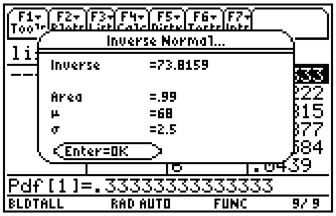
Finding a Value from a Given Population

Example: What height separates the tallest 1% from the other 99%, or what is the 99th percentile?

1. Press [2nd] [⊞] to return to the Stats/List Editor.
2. Press [F5] **Distr**, **2:Inverse**, and then press **1:Inverse Normal**, with Area: **0.99**, μ: **68**, and σ: **2.5** (screen 10).

(10) 

3. Press [ENTER] to display screen 11, which shows a height of **73.8159** inches or **6 ft 1.8 inches = p_{99}**.

(11) 

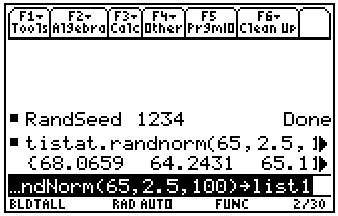
Simulating a Sample from a Normal Distribution

From the Home screen, generate a sample of size 100 from a normal population of $\mu = 68$, $\sigma = 2.5$ inches.

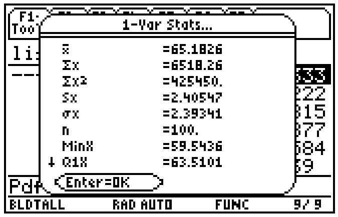
1. Set **RandSeed 1234**, as in Topic 15, if you want to repeat these results.
2. Paste **randnorm(…tistat** from [CATALOG], [F3] **Flash Apps**.

3. Complete for **tistat.randnorm(65,2.5,100)→list1** for a set of 100 generated heights stored in **list1**, starting with **{68.0659, 64.2431, …}** (screen 12).

4. Return to the Stats/List Editor and press F4 **Calc**, **1:1-Var Stats**, with List: **list1** and Freq: **1**.

(12)


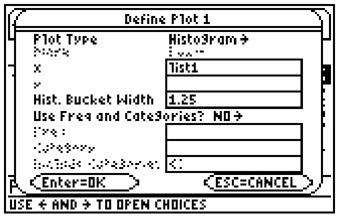
5. Press ENTER for the results $\bar{x}$ = **65.1826 ≈ 65**, Sx = **2.40547 ≈ 2.5** inches, minX = **59.5436**, and maxX = **70.2831** (screen 13).

(13)


From Topic 2, use about eight cells of width **(70.28 - 59.54)/8 = 1.34**, so try **1.25** as the histogram bucket width with 10 buckets.

6. Press F2 **Plots**, **1:Plot Setup**.

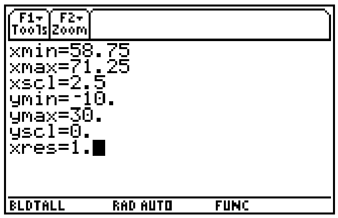
7. Set up and define **Plot 1** as Plot Type: **Histogram** with x: **list1**, Hist. Bucket Width: **1.25**, and Use Freq and Categories?: **NO** (screen 14).

(14)


8. Set up the window using ◆ [WINDOW] with the following entries:

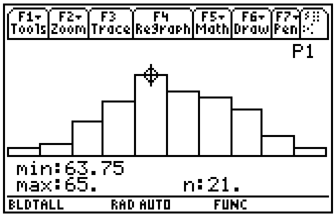
   - **xmin = 58.75**
   - **xmax = 71.25**
   - **xscl = 2.5**
   - **ymin = -10**
   - **ymax = 30**
   - **yscl = 0**
   - **xres = 1**

   (See screen 15.)

(15)


9. Press ◆ [GRAPH], and then press [F3] **Trace** and ▶ a few times. The graph in screen 16 looks somewhat normal, or at least mound-shaped.

How do you check to see if it is normal enough?



(16)

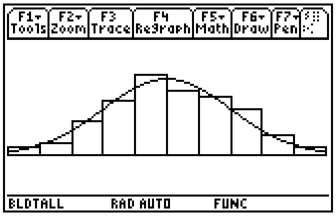
Checking Normality with Normal Probability Plots

The histogram in screen 16 can be considered to be made up of 100 blocks of width 1.25 units and height 1 unit for a total Area = 100 ∗ 1.25 ∗ 1 = 125. The area under a normal PDF (x, 65, 2.5) has the same base, and area = 1, so multiply the height by 125 so it will fit the histogram.

1. Let **y1 = 125 ∗ tistat.normpdf(x, 65, 2.5)**, similar to what you did in screen 1 at the beginning of Topic 18.

2. With **Plot 1** still set up (as in screen 16), press ◆ [GRAPH] to view screen 17.

   This helps with your visualization of normality, but it is an "eyeball" estimation.

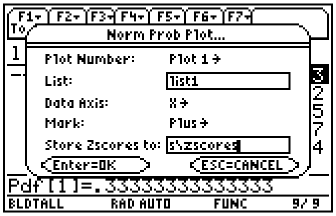


(17)

3. From the Stats/List Editor, turn off all functions with [F2] **Plots**, **4:FnOff**.

4. Press [F2] **Plots**, **1:Plot Setup**, highlight **Plot 1**, press [F3] **Clear**, and then press [ENTER].

5. Press [F2] **Plots**, **2:Norm Prob Plot**, with Plot Number: **Plot1**, List: **list1**, Data Axis: **X**, Mark: **Plus**, Store Zscores to: **statvars\zscores** (screen 18).

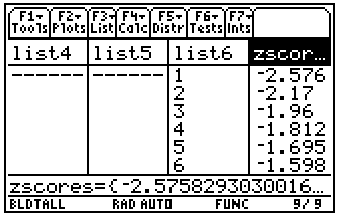


(18)

Note: *If* **Plot 1** *was not cleared in step 4, it could not be used here.*

6. Press [ENTER] to return to the Stats/List Editor that now has **List zscores** pasted to the end of the list (screen 19).

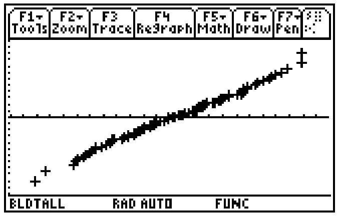
7. Press [F2] **Plots**, **1:Plot Setup** for the Plot Setup screen (not shown) and observe that **Plot 1** has been automatically set up with Plot Type: **Scatter**, Mark: **Plus**, X List: **npplist**, and Y List: **zscores**.

(19)


Note: *This is a scatterplot with list* **npplist** (**list1**) *sorted in ascending order. List* **zscores** *is also a list, in order from low to high. If you wish to make a second normal probability plot but need to save the above results, you must store lists* **npplist** *and* **zscores** *to other list names.*

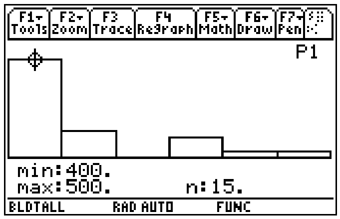
8. Press [F5] **ZoomData** (screen 20).

   The data are close to lying on a straight line, which is easier to eyeball than normality in screen 17. Linearity in a normal probability plot is an indication that the data come from a normal distribution.

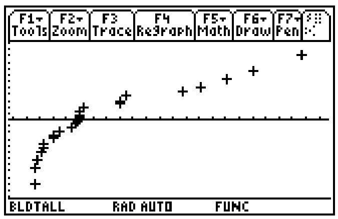
(20)


Skewed Distribution and a Normal Probability Plot

Example: In Topic 3 screen 16, the heights of tall buildings in Philadelphia, PA were skewed to the right, with most of the building heights between 400 and 500 ft., but a few were over 700 ft. tall. Topic 3, screen 16 is repeated in screen 21.

(21)


Screen 22 is the normal probability plot for the data in list **phily** (in folder **BLDTALL**). Notice that the plotted data do not lie on a straight line. This indicates that these data are not normally distributed.

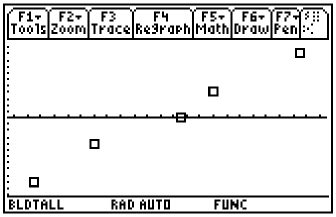
(22)


Small Samples and Normal Probability Plots

Example: Five measurements were made of the thickness of paper by using Vernier calibers on a thin stack of paper and then dividing this value by the number of sheets in the stack, with the following results: 0.09302, 0.09293, 0.09315, 0.09333, and 0.09320. (Source: Reprinted from *Experimentation and Measurements*, W. J. Youden, U.S. Department of Commerce, National Institute of Standards and Technology, N. B. S. Special Publication 672, 1984. Not copyrightable in the United States.)

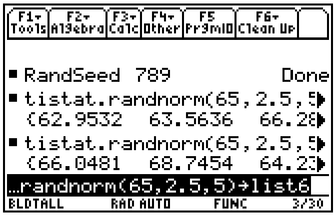
Screen 23 gives the normal probability plot for these data. These measurements can be thought of as coming from a normally distributed population of measurement.

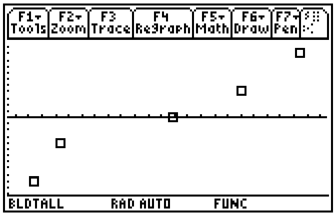
You should not expect all small samples from a normal population to look this good.

(23) 

Note: *The data appear to lie, approximately, on a straight line.*

1. From the Home screen, set **RandSeed 789** and store **tistat.randnorm(65,2.5,5)→list 5** and **tistat.randnorm(65,2.5,5)→list 6** (screen 24).
2. Repeat steps 4 through 8 corresponding to screens 18 through 20, except use **list5** instead of **list1** and use Mark: **Square** instead of **Plus**.

(24) 

(25) 

3. Repeat for **list6** to get screen 26.

These plots illustrate the types of variations one might expect. Examining plots of different sample sizes is good practice for using this tool effectively. Other examples will be given at the end of Topics 19 and 20.

(26)