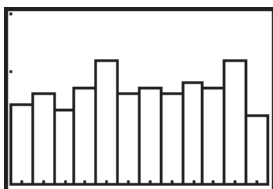


CHAPTER 2 EXPLORING DISTRIBUTIONS

Section 2.1 ■ The Shape of Things: Visualizing Distributions

Uniform or Rectangular Distribution



[0.5, 12.5, 1, 0, 30, 10]

As suggested on page 26 of the *Statistics in Action* student text, the TI-83 and TI-83 Plus have two commands that generate approximately uniform distributions. Under the **[MATH]** PRB submenu, `rand(n)` randomly selects *n* numbers from the open interval (0, 1), and `randInt(start, end, n)` selects *n* integers from the closed interval [*start*, *end*]. For example, if you enter `randInt(1,12,200) → L1` into the Home screen, a histogram of list L1 will look something like the screen at left. This particular example could represent the number of births per month for a sample of size 200. (For more information about histograms, see Section 2.2.)

Normal Distribution **[2nd]** **[DISTR]** **DISTR 1:normalpdf(**

To graph a normal curve, go to the Y= screen and define a function in the form $Y = \text{normalpdf}(X, \text{mean}, \text{standard deviation})$. You find the `normalpdf(` command by pressing **[2nd]** **[DISTR]** **DISTR 1:normalpdf(**. If you do not specify the mean and standard deviation, the calculator assumes they are 0 and 1. The pdf suffix stands for *probability density function*. Using `normalpdf(` as the definition of a function provides the *y*-coordinates of the normal curve.

There is no zoom command that gives a “friendly” graph of a normal curve. Instead, use these guidelines to set the Window screen:

$$X_{\min} = \text{mean} - 3 SD$$

$$X_{\max} = \text{mean} + 3 SD$$

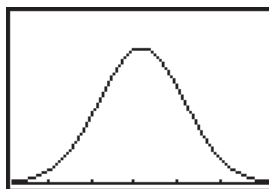
$$X_{\text{scl}} = SD$$

$$Y_{\min} = 0$$

$$Y_{\max} = \frac{1}{2 SD}$$

$$Y_{\text{scl}} = 0$$

```
Plot1 Plot2 Plot3
Y1=normalpdf(X,
50, 12)
Y2=
Y3=
Y4=
Y5=
Y6=
```



[14, 86, 12, 0, 0.042, 0]

(continued)

Section 2.1 ■ The Shape of Things: Visualizing Distributions (continued)

If you are interested in graphing more than one normal curve to show, for example, the effect of changes in standard deviation, you can enter two separate functions into the Y= screen. Or you can enclose the standard deviations in braces and use a single function. In either case, use the following “friendly” window:

$$Xmin = mean - 3(Largest\ SD)$$

$$Xmax = mean + 3(Largest\ SD)$$

$$Xscl = Largest\ SD$$

$$Ymin = 0$$

$$Ymax = \frac{1}{2\ Smallest\ SD}$$

$$Yscl = 0$$

```

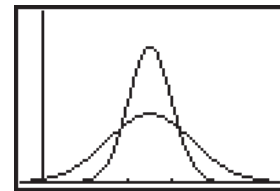
Plot1 Plot2 Plot3
Y1=normalpdf(X,
50,10)
Y2=normalpdf(X,
50,20)
Y3=
Y4=
Y5=

```

```

Plot1 Plot2 Plot3
Y1=normalpdf(X,
50,(10,20))
Y2=
Y3=
Y4=
Y5=
Y6=

```



[-10, 110, 20, 0, 0.05, 0]

Section 2.2 ■ Graphical Displays for Distributions

Histograms

The TI-83 and TI-83 Plus display frequency histograms and allow control over the attributes of the plot via the Window screen. First, enter the data values into a list, say, list L1. If you have frequencies associated with the data values, enter the frequencies into another list, say, list L2. (See “Lists” in the introduction for information about entering data into lists.) The examples below use the speeds of mammals from Display 2.24 on page 39 of the Statistics in Action student text. The data were entered in list L1.

Second, go to the Stat Plot screen, $\boxed{2nd}$ [Stat Plot], and define a histogram. Histograms are the third option under Type.

```

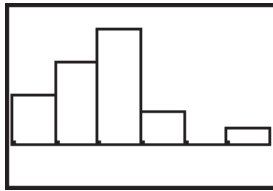
Plot1 Plot2 Plot3
Off Off
Type: L1 L2 L3
Freq: Off Off
Xlist: L1
Freq: 1

```

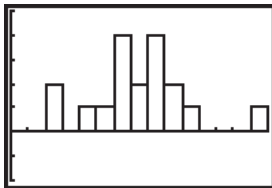
(continued)

Section 2.2 ■ Graphical Displays for Distributions (continued)

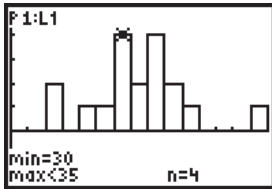
If you press **ZOOM** 9:ZoomStat, the histogram will fill the Graph window. The bars, however, may have unusual widths and dividing lines.



[11, 81.8, 11.8, -2.10, 8.19, 10]



[0, 75, 5, -2, 5, 1]



[0, 75, 5, -2, 5, 1]

The width of the bars is determined by Xscl, starting at Xmin. To change the width of the dividing lines of the bars, reset the Window screen. For example, the window [0, 75, 5, -2, 5, 1] starts the first bar at 0 and starts a new bar every 5 units. Needless to say, as the bar width changes and data values are redistributed, you may need to adjust Ymax, too.

When you trace the histogram, you see the lower and upper bounds of each bar and the number of data values (the frequency) of each bar. Please note that a value that falls at the dividing line between two bars is put in the bar on the right. Unfortunately, the TI-83 and TI-83 Plus do not provide relative frequency histograms. One way to work-around this is to divide each frequency by the total number of data values. For example, enter your data values into list L1 and enter your frequencies into list L2. Then arrow up and right to highlight the name of list L3 and enter $L2/\text{dim}(L2)$. Find the $\text{dim}()$ command by pressing **2nd** [LIST] OPS 3:dim(. (Note: If you enter the expression in quotation marks, **ALPHA** ["], the definition will be dynamic and the values will update whenever list L2 changes.) Then create a histogram using lists L1 and L3. You should remember, however, that the shape of the relative frequency histogram is identical to that of the frequency histogram.

Stem Plots

Stem plots are not a built-in feature of the TI-83 and TI-83 Plus. If desired, you could do an Internet search for programs that add this functionality.

Bar Graphs for Categorical Data

Bar graphs are not a built-in feature of the TI-83 and TI-83 Plus. Some bar graphs with numerical categories can be simulated with histograms, but using computer software may be more efficient.

Section 2.3 ■ Measures of Center and Spread

Numerical Summaries STAT CALC 1:1-Var Stats

```

EDIT  CALC  TESTS
1:1-Var Stats
2:2-Var Stats
3:Med-Med
4:LinReg(ax+b)
5:QuadReg
6:CubicReg
7:4QuartReg
    
```

Consistent with the discussion throughout Section 2.3 of the student text, the TI-83 and TI-83 Plus provide summary statistics. Using the 1-Var Stats command from the STAT CALC submenu, you can calculate a variety of summary statistics for any data set stored in a list. The summary statistics include mean, median, standard deviation, and quartiles.

In general, you enter 1-Var Stats into the Home screen followed by the name of the list.

```

1-Var Stats L1
    
```

```

1-Var Stats
x̄=49.43669621
Σx=2471.83481
Σx²=129252.405
Sx=11.99749813
σx=11.8769172
↓n=50
    
```

```

minX=23.291942
Q1=40.68246719
Med=49.5679037
Q3=57.51547317
maxX=75.355326
    
```

```

1-Var Stats L1,L2
2
    
```

If your data set is contained in a frequency table, you enter the names of two lists separated by a comma—first, the list that contains the data values, and second, the list that contains the frequencies.

Please note that the mean is listed as \bar{x} . The calculator does not distinguish between sample mean and population mean. The calculator does, however, provide two standard deviations. Sx is the sample standard deviation calculated with division by $(n - 1)$. σx is the population standard deviation calculated with division by n .

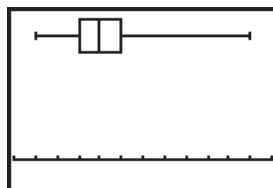
Also, please note the complete five-number summary is displayed on the lower portion of the 1-Var Stats screen.

Five-Number Summaries, Outliers, and Boxplots

The TI-83 and TI-83 Plus provide two types of boxplots: regular and modified. The regular boxplot (the fifth option for Type within a Stat Plot screen) does not indicate outliers, whereas the modified boxplot (sixth option) does. As with histograms, you specify a list for frequencies if your data is contained in a frequency table.

```

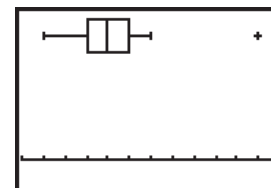
Plot1 Plot2 Plot3
Off Off Off
Type: [ ] [ ] [ ]
Xlist:L1
Freq:L2
    
```



[5, 65, 5, -1, 6, 0]

```

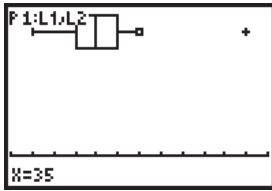
Plot1 Plot2 Plot3
Off Off Off
Type: [ ] [ ] [ ]
Xlist:L1
Freq:L2
Mark: [ ] [ ] [ ]
    
```



[5, 65, 5, -1, 6, 0]

(continued)

Section 2.3 ■ Measures of Center and Spread (continued)



[5, 65, 5, -1, 6, 0]

Pressing **TRACE** when a boxplot is displayed traces the values of the five-number summary. For modified boxplots, you can trace the values of the whiskers and the outliers.

Percentiles and Cumulative Frequency Plots

The TI-83 and TI-83 Plus can construct a cumulative frequency plot using the following procedure:

- Enter data values and frequencies into lists L1 and L2, respectively.

L1	L2	L3	3
10	3		
15	5		
20	8		
25	12		
30	9		
35	6		
60	2		
L3(1)=			

- Arrow up and right to highlight the name of list L3. Enter `cumSum(L2)`. The `cumSum(` command is found by pressing **2nd** [LIST] OPS 6:`cumSum(`. (Note: If you enter the expression in quotation marks, **ALPHA** ["], the definition will be dynamic and the values will update whenever list L2 changes.)

NAMES OPS MATH	
1:SortA(
2:SortD(
3:dim(
4:Fill(
5:seq(
6:cumSum(
7>List(

L1	L2	L3	3
10	3		
15	5		
20	8		
25	12		
30	9		
35	6		
60	2		
L3="cumSum(L2)"			

- Press **ENTER** to have list L3 calculate the cumulative sums.

L1	L2	L3	# 3
10	3	3	
15	5	8	
20	8	16	
25	12	28	
30	9	37	
35	6	43	
60	2	45	
L3(1)=3			

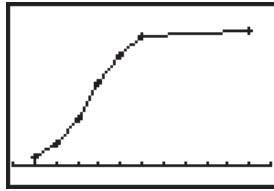
- Press **2nd** [STAT PLOT] to define an xyline plot that is a cumulative frequency plot. Use lists L1 and L3.

Plot1	Plot2	Plot3
Off	Off	Off
Type: L1		
Xlist: L1		
Ylist: L3		
Mark: .		

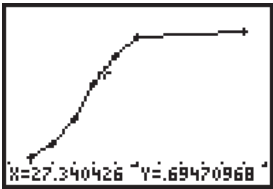
(continued)

Section 2.3 ■ Measures of Center and Spread (continued)

- e. Press $\boxed{\text{ZOOM}}$ 9:ZoomStat to display the cumulative frequency plot.



[5, 65, 5, -4.14, 52.14, 0]



[5, 65, 5, -0.92, 1.159, 0]

A cumulative relative frequency plot can be constructed by entering $\text{cumSum}(L_2)/\text{sum}(L_2)$ as the definition of list L_3 . Find the $\text{sum}()$ command by pressing $\boxed{2\text{nd}}$ [LIST] MATH 5:sum(. With a cumulative relative frequency plot, you can move the cursor to identify the data value for any percentile. For example, to find the value that is at the 70th percentile, move the cursor to the point on the graph whose y -coordinate is approximately .70. The x -coordinate is the data value for the 70th percentile.

Measuring Spread About the Mean: The Standard Deviation

Many calculators and computers have efficient, built-in functions for calculating the standard deviation. Nonetheless, performing the calculations by hand may help you understand the meaning in the formula. The TI-83 and TI-83 Plus can support hand calculations with the spreadsheet capabilities of the List Editor.

- Enter the data into list L_1 .
- Define list L_2 to be the deviations, $x - \bar{x}$. You do this with either the expression $L_1 - \text{mean}(L_1)$ or $L_1 - \text{sum}(L_1)/\text{dim}(L_1)$. Find $\text{mean}()$ and $\text{sum}()$ in the mathematical functions submenu, $\boxed{2\text{nd}}$ [LIST] MATH. Find $\text{dim}()$ in the list operations submenu, $\boxed{2\text{nd}}$ [LIST] OPS. (Note: If you enter the expression in quotation marks, $\boxed{\text{ALPHA}}$ ["], the definition will be dynamic and the values will update whenever list L_1 changes.)
- Define list L_3 to be the squared deviations, $(x - \bar{x})^2$. Use the expression L_2^2 .

L1	Mean	L3	Σ
10	-----	-----	
15			
20			
25			
30			
35			
60			
L2 = "L1 - mean(L1)"			

L1	L2	#	L3	#
10	-17.86		318.88	
15	-12.86		165.31	
20	-7.867		61.735	
25	-2.867		8.1633	
30	2.1428		4.5918	
35	7.1428		51.02	
60	32.143		1033.2	
L3 = "L2 ^ 2"				

$\sqrt{(\text{sum}(L_3)/(\text{dim}(L_3)-1))}$
16.54719081

- Complete the calculations on the Home screen. The expression $\sqrt{(\text{sum}(L_3)/(\text{dim}(L_3)-1))}$ calculates the sum of the square deviations, divides by one less than the number of data values, and takes the square root.

(continued)

Properties of Summary Statistics

The 1-Var Stats command makes it relatively easy for you to explore and recall the effects of recentering or rescaling. For example, to determine the effect of tripling each data value in any data set, first enter a small, hypothetical data set into list L1. Use 1-Var Stats on the original data set. Then define list L2 as the triple of each value, $3 \times L1$, and use 1-Var Stats on the rescaled data. You can see which of the summary statistics are likewise tripled.

L1	L2	L3	1
10	-----	-----	
20			
30			
40			
50			
L1(6)=			

1-Var Stats	
\bar{x}	=30
Σx	=150
Σx^2	=5500
Sx	=15.8113883
σx	=14.14213562
$\downarrow n$	=5

minX	=10
Q1	=15
Med	=30
Q3	=45
maxX	=50

L1	L2	L3	2
10	30	-----	
20	60		
30	90		
40	120		
50	150		
L2="3*L1"			

1-Var Stats	
\bar{x}	=30
Σx	=150
Σx^2	=5500
Sx	=15.8113883
σx	=14.14213562
$\downarrow n$	=5

minX	=10
Q1	=15
Med	=30
Q3	=45
maxX	=50

Summaries from a Frequency Table

See “Numerical Summaries” in this chapter for information about calculating summaries from a frequency table.

Section 2.4 ■ The Normal Distribution

Unknown Percentage and Unknown Value Problems

Normal Cumulative Distribution Function $\boxed{2\text{nd}}$ [DISTR] DISTR 2:normalcdf(

normalcdf(45,65,
50,12)
.5558890044

On the Home screen, a command in the form $\text{normalcdf}(x_1, x_2, \text{mean}, \text{standard deviation})$ returns the area under a normal curve and on the interval $[x_1, x_2]$ for the normal distribution specified by *mean* and *standard deviation*. You find the $\text{normalcdf}()$ command by pressing $\boxed{2\text{nd}}$ [DISTR] DISTR 2:normalcdf(. The *cdf* suffix stands for *cumulative distribution function*. Technically, a cumulative distribution function returns the percentage of area under a continuous distribution curve from negative infinity to the value of interest. The $\text{normalcdf}()$ command, however, is designed such that the lower bound of the interval can be specified as any value. For example, to find the percentage of area below a normal curve with mean 50 and standard deviation 12 over the interval $[45, 65]$, enter $\text{normalcdf}(45,65,50,12)$ into the Home screen.

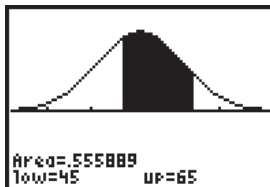
(continued)

Section 2.4 ■ The Normal Distribution (continued)

Shading a Normal Curve $\boxed{2\text{nd}}$ [DISTR] DRAW 1:ShadeNorm(

To see a graph of the shaded area, use the ShadeNorm(command. You find this command by pressing $\boxed{2\text{nd}}$ [DISTR] DRAW 1:ShadeNorm(. You enter ShadeNorm(into the Home screen.

```
ShadeNorm(45,65,
50,12)■
```



[14, 86, 12, -0.03, 0.042, 0]

Please note that the ShadeNorm(command does not adjust the window settings. Use these guidelines to set the Window screen:

$$X_{\min} = \text{mean} - 3 SD$$

$$X_{\max} = \text{mean} + 3 SD$$

$$X_{\text{scl}} = SD$$

$$Y_{\min} = -0.03$$

$$Y_{\max} = \frac{1}{2 SD}$$

$$Y_{\text{scl}} = 0$$

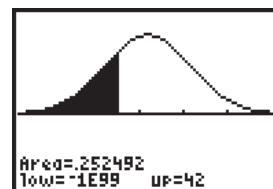
(Note: If you use ShadeNorm(repeatedly, press $\boxed{2\text{nd}}$ DRAW 1:ClrDraw to clear the drawing between each use.)

Using Infinity as a Bound

If one of the bounds of the interval is infinite, use 1×10^{99} or -1×10^{99} to represent positive or negative infinity, respectively. You enter scientific notation using $\boxed{2\text{nd}}$ [EE] to place E between the coefficient and the exponent on 10. You can use 1E99 and -1E99 with both normalcdf(and ShadeNorm(. For example, to find the area below a normal curve with mean 50 and standard deviation 12 over the interval $[-\infty, 42]$, use normalcdf(-1E99,42,50,12) or ShadeNorm(-1E99,42,50,12).

```
normalcdf(-1E99,
42,50,12)
.252492467
```

```
ShadeNorm(-1E99,
42,50,12)■
```



[14, 86, 12, -0.03, 0.042, 0]

The Standard Normal Curve

Defaults for the Normal Curve Commands

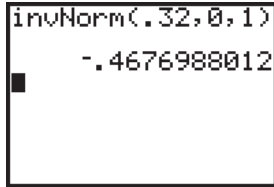
All of the normal curve commands—normalpdf(, normalcdf(, ShadeNorm(, and invNorm(—assume the standard normal distribution as a default. If you do not specify the mean and standard deviation, the calculator assumes 0 and 1,

(continued)

Section 2.4 ■ The Normal Distribution (continued)

respectively. However, it may be beneficial for you to always specify a mean and standard deviation, even if they are 0 and 1.

Z-scores $\boxed{2\text{nd}} \boxed{[\text{DISTR}]} \boxed{\text{DISTR 3:invNorm(}}$



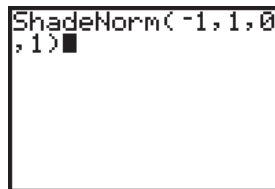
```
invNorm(.32,0,1)
-.4676988012
```

In order to find a z -score, the calculator includes the command `invNorm(`, which is found by pressing $\boxed{2\text{nd}} \boxed{[\text{DISTR}]} \boxed{\text{DISTR 3:invNorm(}}$. This command returns the value of z for the specified area to the right of z . The command is entered in the form `invNorm(area, mean, standard deviation)`. For example, to find the z -score that corresponds to a percentage of .32 (left tail area) under the standard normal curve, enter `invNorm(.32,0,1)` into the Home screen.

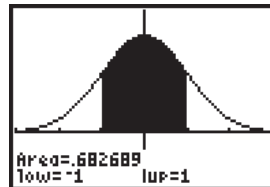
(Note: The `invNorm(` command can give the z -score that corresponds to the specified left tail area for *any* normal distribution, not just the standard normal distribution.)

Central Intervals for Normal Distributions

The central intervals shown on page 88 of the student text can be verified with great accuracy using the `ShadeNorm(` command. For example, to verify the 68% interval, enter `ShadeNorm(-1,1,0,1)`.



```
ShadeNorm(-1,1,0,1)
```



$[-3, 3, 1, -0.2, 0.5, 0]$