69. (2.1) Given the points $(-3, -4)$ and $(5, 2)$ find
   a. the distance between them
   b. the midpoint between them
   c. the slope of the line through them

70. (5.3) Use a calculator to find the value of each expression, then explain the results.
   a. $\log 2 + \log 5 =$ ________
   b. $\log 20 - \log 2 =$ ________

71. (9.4) Solve $2|x + 1| - 3 = 7$ two ways:
   a. using the definition of absolute value
   b. using a system

72. (4.2) Use the rational roots theorem to solve the equation completely, given $x = -3$ is one root.
$x^4 + x^3 - 3x^2 + 3x - 18 = 0$

11.9 Probability and the Normal Curve—Applications for Today

LEARNING OBJECTIVES

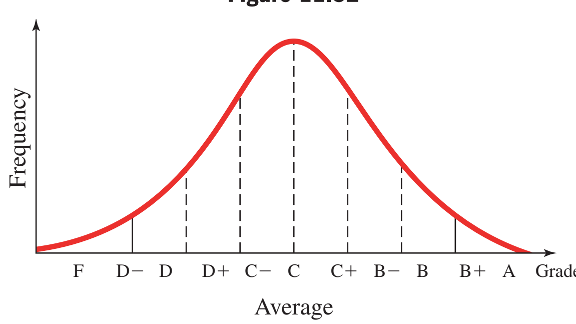
In Section 11.9 you will learn how to:

A. Find the mean and standard deviation for a set of data
B. Apply standard deviations to a normal curve
C. Use the normal curve to make probability statements
D. Use z-scores to make probability statements

INTRODUCTION

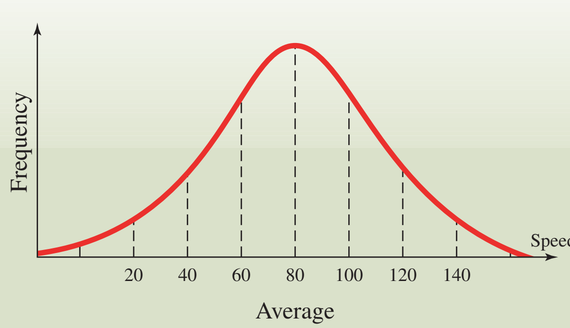
In previous sections, we've made probability statements using counting methods, simple games, formulas, information from tables, and other devices. In this section, we learn to make such statements using observations drawn from a large set of data. Specific characteristic of a large population tend to be **normally distributed,** meaning a large portion of the sample will be average, with decreasing portions tending to be below average and above average. One example might be the grade distribution for a large college, which might be represented by the graph shown in Figure 11.32.

Figure 11.32



POINT OF INTEREST

Many human characteristics and abilities have a normal distribution and the graph of any large sample would resemble that of Figure 11.32. For example, the typing speed of a human will have a like distribution, with a select few being extremely fast (150+ words per minute), and an equally small number being very slow.



A. The Mean and Standard Deviation of a Data Set

The graphs shown above are called **normal distributions** or **bell curves.** Both give a clear indication that for a large sample size, the majority of values are near the average and tend toward the center. These average values occur with the greatest frequency (creating the "hump" in the curve), and taper off as you deviate from the center. When studying normal distributions the average value of a data set is referred to as the **arithmetic mean** or simply the **mean.** It is just one of three common **measures of central tendency,** the others being the **median** (the center of an ordered list) and the **mode** (the value occurring most frequently). As the name implies, these are measures that help quantify the tendency of a data set to cluster about some center value. The mean is often denoted using the symbol $\bar{x}$, read "*x* bar," and is computed as a sum of the data values, divided by the number of values in the sum.

THE ARITHMETIC MEAN $\bar{x}$

The average value of a data set is the sum of all values divided by the number of values in the sum.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

EXAMPLE 1A Compute the mean temperature for the *lowest temperature of record by month* for (a) Honolulu, Hawaii, and (b) Saint Louis, Missouri.

Honolulu, Hawaii—Lowest Temperature of Record by Month (°F)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
53	53	55	57	60	65	66	67	66	61	57	54

St. Louis, Missouri—Lowest Temperature of Record by Month (°F)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
−18	−12	−5	22	31	43	51	47	36	23	1	−16

Source: *2004 Statistical Abstract of the United States*, Table 379]

Solution: **a.** The sum for all 12 months is 714, giving a mean of $\bar{x} = \frac{714}{12} \approx 60°$.

b. The sum for all 12 months is 203, giving a mean of $\bar{x} = \frac{203}{12} \approx 17°$.

While the mean values offer useful information (if you like warm weather, Honolulu is preferable to St. Louis), they tell us little else about the data. Just as the mean describes a tendency toward the center, we also find useful **measures of dispersion,** which describe how the data deviates from the center. In particular, note that the **range** of the Honolulu data (difference of the extreme values) is only $67 - 53 = 14°$, while the range for St. Louis is $51 - (-18) = 69°$, a significant difference! Another such measure is called the **standard deviation** and is denoted by the Greek letter σ (sigma). Since our concern is how much the data varies from center, calculation of the standard deviation begins with finding the mean x. We then find the difference or **deviation** between each data value x_i and the mean $x_i - x$. It seems reasonable that we would then find the average of these deviations, but since some of the results will be negative and others positive, averaging the deviations at this point would be misleading. To get around this, we first *square each deviation,* find the average value, and *then* compute the square root. We illustrate this process using the preceding data above, organizing calculations in a table.

EXAMPLE 1B Compute the standard deviation for the Hawaii and Missouri temperatures.

Solution: Calculations are shown in Tables 11.3 and 11.4. Results are rounded to the nearest unit.

Table 11.3

Hawaii: $\bar{x} \approx 60$

Ordered Data x_i	Deviation $x_i - \bar{x}$	Squared Deviation $(x_i - \bar{x})^2$
53	$53 - 60 = -7$	$(-7)^2 = 49$
53	$53 - 60 = -7$	$(-7)^2 = 49$
54	$54 - 60 = -6$	$(-6)^2 = 36$
55	$55 - 60 = -5$	$(-5)^2 = 25$
57	$57 - 60 = -3$	$(-3)^2 = 9$
57	$57 - 60 = -3$	$(-3)^2 = 9$
60	$60 - 60 = 0$	$0^2 = 0$
61	$61 - 60 = 1$	$1^2 = 1$
65	$65 - 60 = 5$	$5^2 = 25$
66	$66 - 60 = 6$	$6^2 = 36$
66	$66 - 60 = 6$	$6^2 = 36$
67	$67 - 60 = 7$	$7^2 = 49$
Sum = 714		Sum = 324

$$\sigma = \sqrt{\frac{324}{12}} = \sqrt{27} \approx 5.2$$

The standard deviation is $\sigma \approx 5.2$

Table 11.4

Missouri: $\bar{x} \approx 17$

Ordered Data x_i	Deviation $x_i - \bar{x}$	Squared Deviation $(x_i - \bar{x})^2$
−18	$-18 - 17 = -35$	$(-35)^2 = 1225$
−16	$-16 - 17 = -33$	$(-33)^2 = 1089$
−12	$-12 - 17 = -29$	$(-29)^2 = 841$
−5	$-5 - 17 = -22$	$(-22)^2 = 484$
1	$1 - 17 = -16$	$(-16)^2 = 256$
22	$22 - 17 = 5$	$5^2 = 25$
23	$23 - 17 = 6$	$6^2 = 36$
31	$31 - 17 = 14$	$14^2 = 196$
36	$36 - 17 = 19$	$19^2 = 361$
43	$43 - 17 = 26$	$26^2 = 676$
47	$47 - 17 = 30$	$30^2 = 900$
51	$51 - 17 = 34$	$34^2 = 1156$
Sum = 203		Sum = 7245

$$\sigma = \sqrt{\frac{7245}{12}} = \sqrt{603.75} \approx 24.6$$

The standard deviation is $\sigma \approx 24.6$

NOW TRY EXERCISES 7 THROUGH 10

Both the range calculation and the standard deviation indicate that the dispersion of St. Louis temperatures is much greater than that of the Honolulu temperatures.

When calculating standard deviations by hand, organizing your work in a table is a virtual necessity in order to prevent nagging errors. When standard deviations are done via calculating technology, the emphasis shifts to a careful input of the data, and a double-check that values obtained are reasonable. You should always guard against faulty data, faulty key strokes, and the like.

TECHNOLOGY HIGHLIGHT

Calculating the Mean and Standard Deviation

The keystrokes shown apply to a TI-84 Plus model. Please consult your manual or our Internet site for other models.

Virtually all graphing calculators have the ability to compute the mean and standard deviation from a list of data. On the TI-84 Plus, the **1-Var Stats** (single variable statistics) operation is used for this purpose. The operation is located on a submenu of the STAT key, and automatically computes the sum of the data entered, as well as the mean, median, standard deviation (and other measures) of the data set. We'll illustrate the process using the data from Example

1B. Begin by entering the Honolulu and St. Louis temperatures in L1 and L2, respectively. Then quit to the home screen, CLEAR the display, and press STAT ▶ to access the **CALC** submenu, noting that the first option is **1:1-Var Stats** (see Figure 11.33). Pressing ENTER at this point will place this operation on the home screen. Although L1 is the default list for this operation, we will need to distinguish between L1 and L2, so use 2nd 1 to give L1 as the argument. The screen now reads: **1:1-Var Stats L1**. Pressing ENTER will give the screen shown in Figure 11.34, which displays the desired information for the Honolulu data: $x \approx 60$ and $\sigma x \approx 5.2$. To find the related measures for another set of data, simply recall the last function (2nd ENTER) and overwrite L1.

Figure 11.33

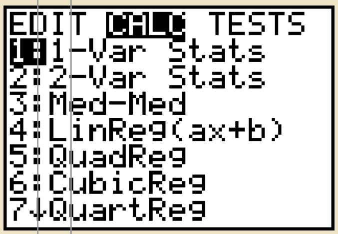
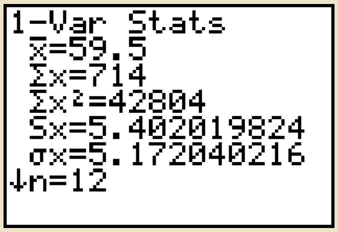


Figure 11.34



Exercise 1: Find the mean and standard deviation of the St. Louis data.

Exercise 2: Find the mean and standard deviation of the following data: $\{-45, -30, -27, -15, -7, 2, 15, 27, 32, 48\}$

Note: For Example 1, the TI-84 Plus returns a slightly different value of σ due to the calculating method programmed in.

B. Standard Deviation and the Normal Curve

In addition to quantifying how data is dispersed from center, standard deviations enable us to draw significant conclusions regarding the sample, and to make probability statements regarding a larger population. The graph in Figure 11.35 is a frequency distribution that illustrates how the heights of 1000 normal adult males are distributed. As you can see, there are few men who are shorter than Danny Devito (152 cm) and even fewer men with the stature of Shaquille O'Neal (over 216 cm). The majority of males seem to cluster around an average height of 178 cm.

Using some basic geometry and judging roughly from the area occupied by each bar, we might legitimately estimate that about 60% of all males in this age group are between 172 cm and 183.9 cm (shaded regions). By connecting the midpoint of each bar, the line graph in Figure 11.36 is obtained.

Figure 11.35

Frequency distribution for male heights.

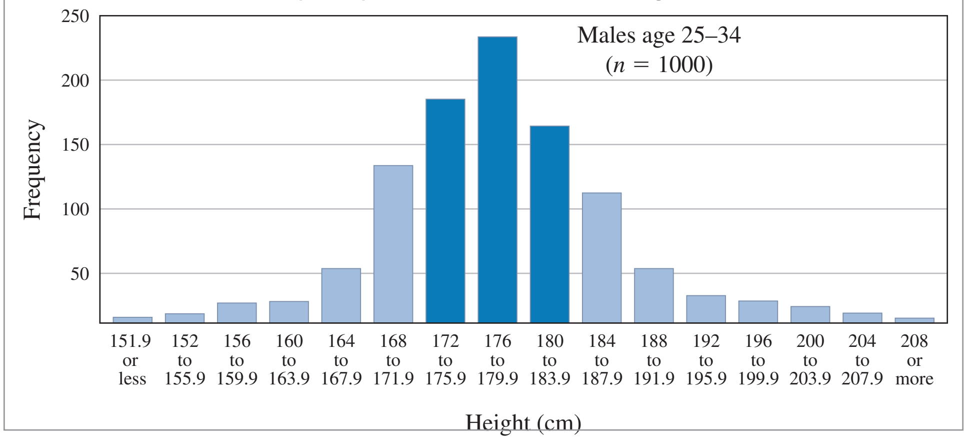


Figure 11.36

Corresponding line graph

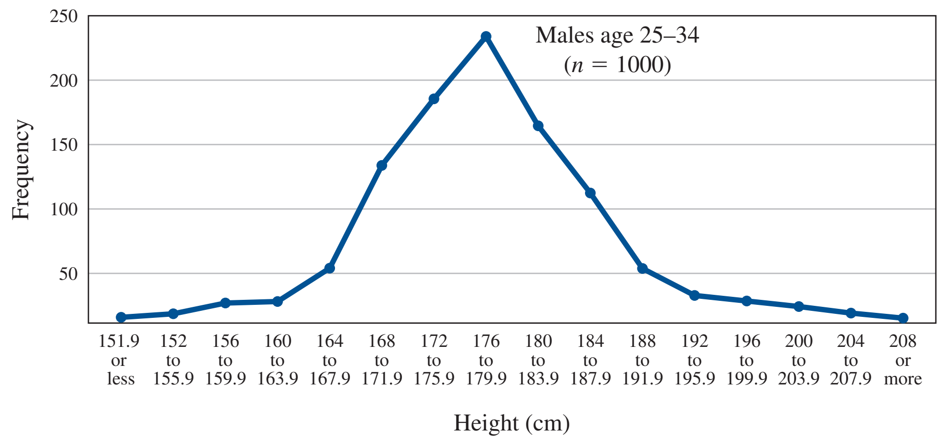
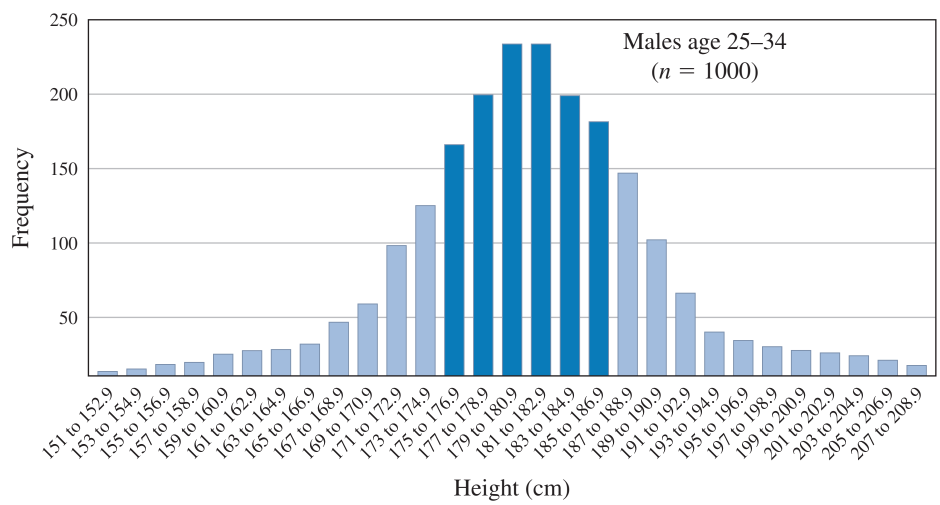


Figure 11.37

Frequency distribution (with smaller intervals)



We can refine the bar graph using twice as many intervals, which means each bar would be one-half as wide. The new graph is shown in Figure 11.37 and is very similar to the original. But in some sense, this new graph is more "accurate" in that we can more precisely describe the distribution of male heights. The line graph in Figure 11.38 was again formed by connecting the midpoint of each bar in the frequency distribution.

You might imagine what the graphs would look like if we refined them further, by taking even smaller intervals. In particular, the line graph would increasingly resemble a smooth, symmetric, bell-shaped curve. In the real world, many statistical distributions of moderate to large sample sizes have this shape, called a **normal curve.** Normal curves have several characteristics that make them indispensable as a mathematical tool. Since they are symmetric, the mean is located at the center of the curve. Since the curve has a definite peak, the mode also has this same value. But there are

Figure 11.38

Corresponding line graph

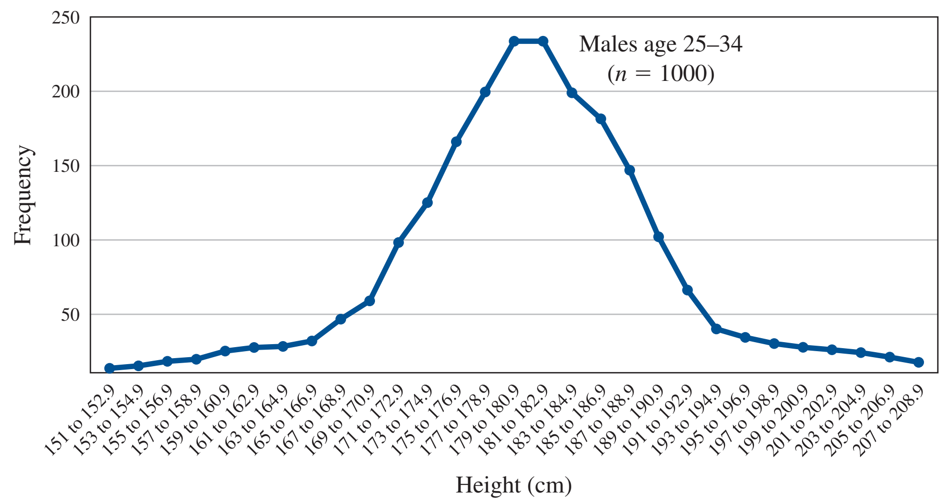
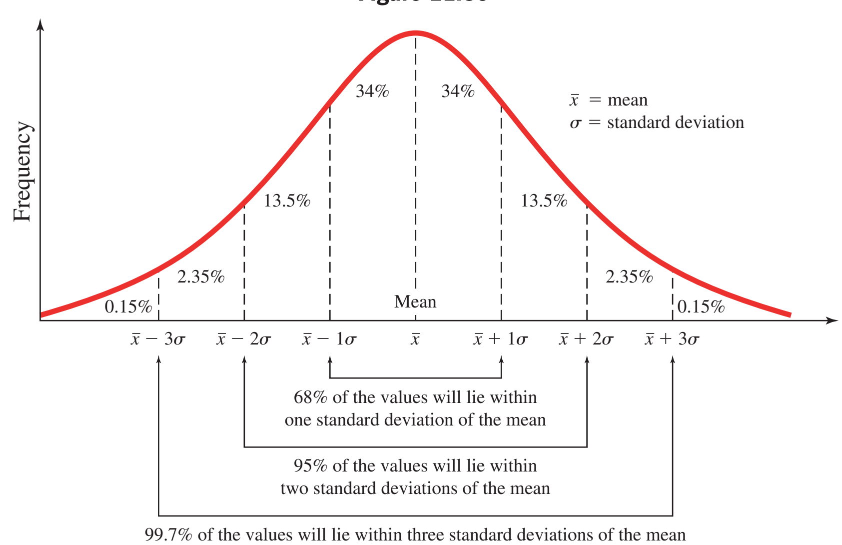


Figure 11.39



other very important features (see Figure 11.39)—the normal curve has the property that approximately 68% of all data values will lie within one standard deviation "1σ" of the mean, 95% of all values will lie within 2σ, and 99.7% of the values will lie within 3σ (figures have been rounded).

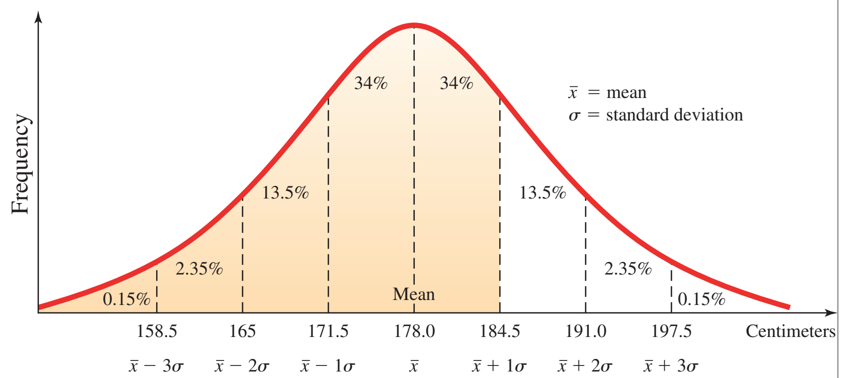
These figures not only represent the distribution of values, they're also a measure of the corresponding area under the curve partitioned off by each standard deviation. This means the area under the curve between $x - 1\sigma$ and $x + 1\sigma$ is 68% of the total area, the area between $x - 2\sigma$ and $x + 2\sigma$ is 95% of the total area, and so on. This property of standard deviations will enable us to "use the normal curve in reverse," by computing an area between standard deviations in order to draw conclusions about the sample.

For the 1000 male heights seen earlier, $\bar{x} = 178$ cm, with a standard deviation of $\sigma = 6.5$ cm. This says 68% of male heights are between $x - 1\sigma = 171.5$ cm and

$\bar{x} + 1\sigma = 184.5$ cm. The normal distribution shown in Figure 11.40 now reflects this information, allowing us to make a number of useful observations. It is important to note that the 0.15% actually represents the whole tail beyond $\bar{x} - 3\sigma$ and $\bar{x} + 3\sigma$. Thus, the values sum to 100% and events in each tail are extremely rare, but possible.

EXAMPLE 2 Suppose the distribution in Figure 11.40 is representative of the male population in the State of California, what percent of California's males are 184.5 cm or shorter?

Figure 11.40



Solution: The left-half of the normal distribution (to the left of the mean) represents 50% of area beneath the graph. Since 184.5 cm is exactly one standard deviation to the right of the mean, it represents an additional 34%. This shows that 50% + 34% = 84% of the male population is 184.5 cm or shorter.

NOW TRY EXERCISES 11 THROUGH 16

Actually, for large sample sizes, the normal curve is representative of the distribution for many characteristics of a population and the concepts illustrated here can be applied to any context where the mean and standard deviation are known. While we continue to use the data regarding male heights, there is a large variety of applications in the Exercise Set.

EXAMPLE 3A If this distribution is representative of the adult male population in the State of Florida, what percent of them are 191 cm or taller?

Solution: Now we are interested in the area under the curve and to the right of the second standard deviation (taller than). This means 2.35% + 0.15% = 2.5% of the male population is 191 cm or taller.

EXAMPLE 3B If there are 12,411,000 adult males in California, how many of them will be 158.5 cm or shorter?

Solution: This time we are interested in the extreme left-hand "tail" of the normal distribution, which represents 0.15% of the area under the normal curve. The number of males 158.5 cm or shorter is $0.0015 \times 12{,}411{,}000 \approx 18{,}617$.

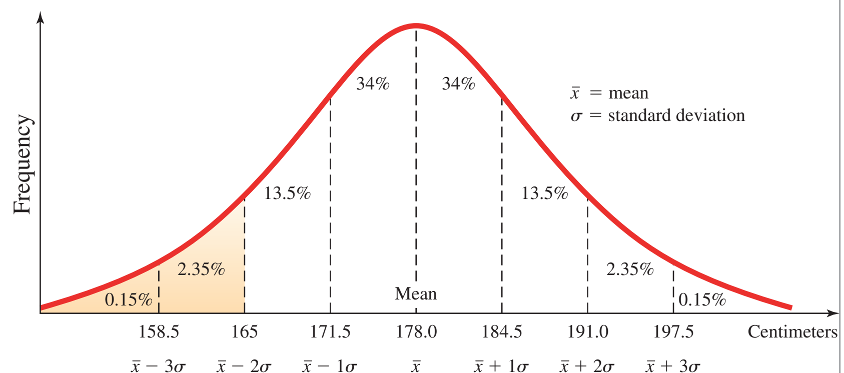
NOW TRY EXERCISES 17 THROUGH 20

C. The Normal Curve and Probability Statements

Recall that the basic definition of probability states the probability of an event E_1 is computed as the number of outcomes in E_1 divided by the number in the sample space: $P(E_1) = \frac{n(E_1)}{n(S)}$. Since this information can be extracted from the normal curve, we can now make probability statements regarding the population under study. Consider Example 4.

EXAMPLE 4 The adult male population of the State of Texas is approximately 7,543,000. Suppose the names of all 7,543,000 are in the state's computer, and the computer randomly picks one of them. See Figure 11.41. What is the probability the male chosen is taller than 165 cm but shorter than 184.5 cm?

Figure 11.41



Solution: We need to determine the area *between* two standard deviations. In Example 2 we found that 84% of the area was to the left of 184.5 cm. Example 3A showed that 2.5% of the area is to the left of 165 cm. Since the two overlap, the area between them is 84% − 2.5% = 81.5% of the total area. There is a 0.815 probability that the height of the male chosen at random falls within this range. As an alternative, we could have simply computed the sum of the percentages for bars in this range. This yields 13.5% + 34% + 34% = 81.5%.

NOW TRY EXERCISES 21 AND 22

D. *z*-Scores and the Normal Curve

Finding the area under the normal curve, and hence the related probabilities, is easily done when using integer multiples of the standard deviation. In reality, questions rarely hinge on these integer multiples and we need a more general method that can be applied in all cases. Consider Examples 5A and 5B.

EXAMPLE 5A Dynamic Power Company manufactures and sells car batteries. As a result of exhaustive testing, the company knows the average life of their battery is 4 yr, with a standard deviation of 6 months. Samantha purchases a battery and has it installed in her Corvette.

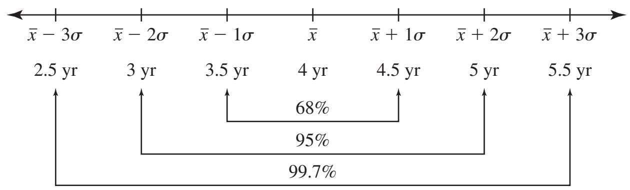
a. What is the probability her battery will last between 3.5 and 4.5 yr?

b. What is the probability her battery will last more than 5.5 yr?

c. What is the probability she will have to return the battery before 3 yr are up?

Solution: Use the information to construct a diagram such as the one in Figure 11.42 or refer to any of the previous normal curves. Since $x = 4$ and $\sigma = 0.5$ we have the following.

Figure 11.42



a. For 3.5 yr < battery life < 4.5 yr, the corresponding area under the normal curve is 68%. The probability the battery will last more than 3.5 yr, but less than 4.5 yr is 0.68.

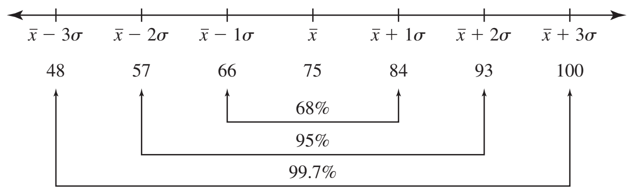
b. A battery life of more than 5.5 yr is beyond 3σ from the mean. The corresponding area is 0.15%. There is a 0.0015 probability the battery will last more than 5.5 yr (it's not very likely).

c. For battery life < 3 yr we use the "left-hand tail" of the normal curve, the area beyond 2σ from the mean. The corresponding area is 2.35%. The probability she will need to return the battery before 3 yr are up is 0.0235.

EXAMPLE 5B A mathematics entrance exam is given to 6500 freshmen who want to enter an engineering program. The scores have a normal distribution with a mean $\bar{x} = 75$ and a standard deviation of $\sigma = 9$. How many freshmen scored above a 93?

Solution: Use the information to construct a diagram as shown in Figure 11.43 or refer to any of the previous normal curves. Since $x = 75$ and $\sigma = 9$ we have the following. A score of 93 is two standard deviations from the mean: $75 + 2\sigma = 75 + 18 = 93$. The area under the curve and to the right of the second standard deviation is 2.5%. Hence $0.025 \times 6500 \approx 163$ students scored above a 93.

Figure 11.43



NOW TRY EXERCISES 25 AND 26

Of course, a more important question would concern the number of students scoring below a 60, which is traditionally the lowest passing score. Unfortunately, a score of 60 falls *between* two known deviations and our current methods cannot be used. But methods exist for computing a "nonstandard" deviation, and this will enable us to find the information needed. This is done by taking the difference between a given value x_i and the mean $\bar{x}$, then dividing by one standard deviation: $\frac{x_i - \bar{x}}{\sigma}$. This process is called *calculating* a ***z*-score.**

CALCULATING A *z*-SCORE

Given a data set with mean $\bar{x}$ and standard deviation σ. For known value x_i, $z = \frac{x_i - \bar{x}}{\sigma}$ is called the z-score relative to x_i, and represents the nonstandard deviation from $\bar{x}$.

EXAMPLE 6 Find the z-score corresponding to a test score of 60, given $\bar{x} = 75$ and $\sigma = 9$.

Solution:

$$z = \frac{x_i - \bar{x}}{\sigma} \quad \text{z-score formula}$$

$$= \frac{60 - 75}{9} \quad x = 75 \text{ and } \sigma = 9$$

$$\approx -1.67 \quad \text{result}$$

A score of 60 lies 1.67 standard deviations *to the left* of the mean.

NOW TRY EXERCISES 27 THROUGH 32

Now all that remains is to find the area under the normal curve corresponding to $z = 1.67\sigma$. Computing this area directly requires some very sophisticated mathematics, but fortunately this has already been done for all possible deviations correct to two decimal places. The results have been compiled in the ***z*-score table,** which appears in Appendix V. The table is read by locating the units place and the 10ths place digits of the deviation along the leftmost vertical column, then scanning horizontally along the top for the 100ths place digit. A small section of the table is reproduced in Table 11.5 to illustrate. Locating -1.6 along the left-hand column and 7 in the top row, we find the corresponding value in the table is 0.0475.

It's very important to note that z-scores are a *cumulative* value giving the *total area* under the curve and *to the left* of the given score (see Figure 11.44). The area under the

Table 11.5

Partial z-score table

z	0	1	2	3	4	5	6	7	8	9
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0238	0.0233
−1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0300	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0570	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681

curve and to the left of −1.67 is 0.0475 (4.75%) of the total area (entries in the table are correct to four decimal places). This means $0.0475 \times 6500 = 309$ freshmen did not pass the entrance exam.

Figure 11.44

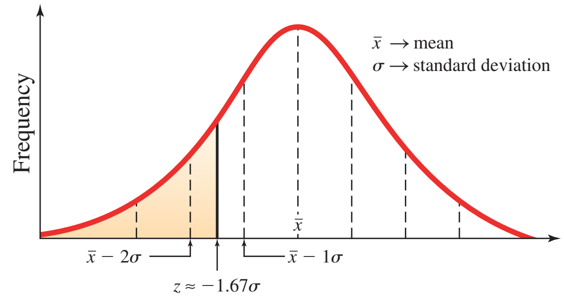
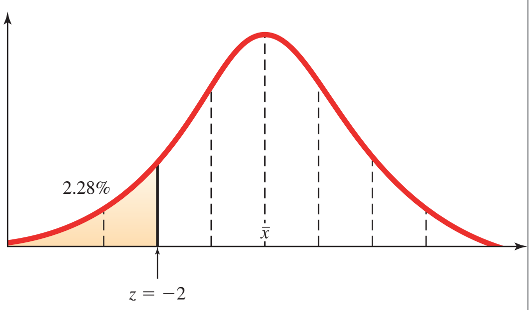


Figure 11.45



For a z-score of $z = -1.00$ (one standard deviation), the table gives a value of $0.1587 = 15.87\%$. This is very close to the 16% used earlier (values given in the table are actually more accurate). The normal curves shown in Figures 11.45, 11.46, and 11.47 have been re-marked using z-scores that coincide with the standard deviations. This will more clearly indicate the cumulative nature of a z-score. The values shown give the percentage of area under the normal curve that is shaded, as indicated by the z-score. The figures differ from those used earlier due to rounding.

Figure 11.46

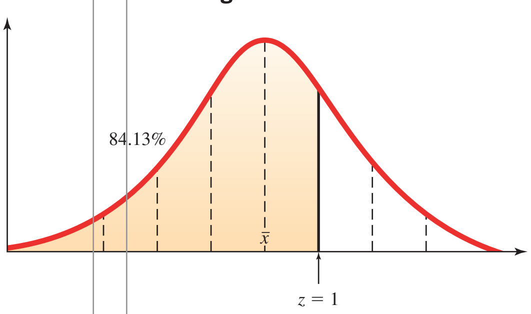
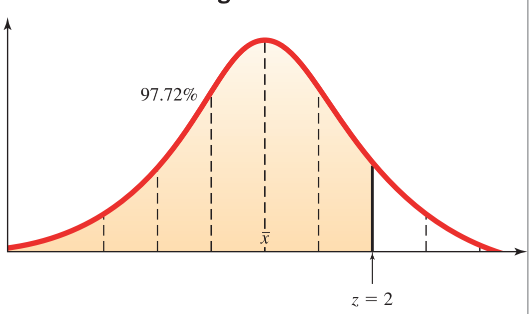


Figure 11.47



Mathematical resources are often used to make important decisions regarding efficiency, economy, safety, value, and decisions of other kinds. Examples 7 and 8 illustrate some of the various ways that properties of the normal curve can be used.

EXAMPLE 7 McClintock County needs to purchase some premium lightbulbs for the use in tunnels, dams, subway systems, and other specialized areas. Due to the expense, time, and difficulty involved in replacing these bulbs, the county requires manufacturers to guarantee that 93%

of all bulbs purchased will burn for at least 1400 hr. The mean lifetime for bulbs from Incandescent Inc. is $\bar{x} = 1510$ hr, with $\sigma = 74$ hr. Can the county purchase bulbs from this company?

Solution: We need to determine if $x_i \geq 1400$ hr represents 90% or more of the area under the normal curve for the values of $\bar{x}$ and σ given. The z-score calculation is $z = \frac{1400 - 1510}{74}$ or $z \approx -1.49$. The z-score table gives a value of 0.0681, indicating that 6.81% of the area is to the left, so 100% − 6.81% = 93.19% of the area lies to the right. It's a close call, but the company can legitimately claim that more than 93% of its bulbs will burn for more than 1400 hr.

NOW TRY EXERCISES 33 THROUGH 35

EXAMPLE 8 The two most widely known college placement exams are the SAT and the ACT. Each of them has a specific test for mathematics. Because of the way the tests are scored, the SAT math test has a mean of $\bar{x} = 500$ and a standard deviation of $\sigma = 100$, while the ACT math test has a mean of $\bar{x} = 18$ and a standard deviation of $\sigma = 6$. Laketa scores a 690 on the SAT while Matthew scores a 29 on the ACT. If the tests are roughly equivalent, who actually received a higher score relative to their peers?

Solution: We can answer this question in terms of the normal distribution for each test and the related z-scores.

For Laketa: $$z = \frac{(690 - 500)}{100} = \frac{190}{100} = 1.9$$

For Matthew: $$z = \frac{(29 - 18)}{6} = \frac{11}{6} = 1.8\overline{3}$$

The corresponding entries in the table for Laketa and Matthew are 97.13% and 96.64%, respectively, indicating that Laketa outperformed Matthew.

NOW TRY EXERCISES 36 THROUGH 38

11.9 EXERCISES

CONCEPTS AND VOCABULARY

Fill in the blank with the appropriate word or phrase. Carefully reread the section if needed.

1. The average difference between data points and $\bar{x}$ is called the __________ __________.

2. For normal distributions, about ______% of the values will be within one standard deviation.

3. The standard deviations applied to the normal curve represent the __________ of values and are also a measure of the corresponding __________ under the curve.

4. The normal curve enables us to make probability statements by using __________ __________ to find the __________ in an __________ E_1, then dividing by the outcomes in the sample space.

5. Describe how nonstandard deviations (z-scores) are calculated and what they represent in relation to the normal curve.

6. Using the concept of "area under the normal curve," give the percent of the population falling within one, two, and three deviations of the mean.

DEVELOPING YOUR SKILLS

7. The 30 final exam scores for an Educational Psychology course were: 74, 96, 88, 63, 99, 34, 37, 97, 99, 32, 51, 78, 81, 93, 91, 75, 75, 93, 69, 93, 84, 89, 93, 68, 90, 86, 92, 79, 89, and 94.

a. Find the mean and standard deviation. **b.** Verify (by counting) that roughly 68% (~20 of the scores) lie between one standard deviation of the mean.

8. The following list gives the number of home runs hit by the National League home run champion for the years 1971 to 2000: 48, 40, 44, 36, 38, 38, 52, 40, 48, 48, 31, 37, 40, 36, 37, 37, 49, 39, 47, 40, 38, 35, 46, 43, 40, 47, 49, 70, 65, and 50.

a. Find the mean and standard deviation. **b.** Verify (by counting) that roughly 68% (~20 of the data points) lie between one standard deviation of the mean.

Source: 2005 WorldAlmanac and Book of Facts.

9. The blood alcohol concentration of 15 drivers involved in fatal accidents and serving time in prison are given here. Using the data given to answer the following. (a) Compute $\bar{x}$ and σ. (b) Verify that roughly 68% of the levels fall within one standard deviation of the mean.

0.27	0.17	0.17	0.16	0.13	0.24	0.19	0.20
0.14	0.16	0.12	0.12	0.16	0.21	0.17	

10. As commercial planes age, several safety and economic concerns are raised. Suppose an airline with forty planes is trying to keep the mean age of their airplanes at $x = 12$ years. If x becomes greater than 12, the oldest aircraft are retired and new ones purchased. Using the ages (in years) of the 42 aircraft given to answer the following:

a. Compute $\bar{x}$ and σ.

b. Verify that roughly 68% of the ages fall within one standard deviation of the mean.

c. If $\bar{x} \geq 12$, how many of the oldest aircraft need to be retired and replaced with new ones?

3.2	22.6	23.1	16.9	0.4	6.6	12.5	22.8	26.3	8.1	13.6	17.0	21.3	15.2
18.7	11.5	4.9	5.3	5.8	20.6	23.1	24.7	3.6	12.4	27.3	22.5	3.9	7.0
16.2	24.1	0.1	2.1	7.7	10.5	23.4	0.7	15.8	6.3	11.9	16.8	16.2	8.7

For any population that is normally distributed, find the percent of the population that is

11. less than $x - 1\sigma$

12. greater than $x + 1\sigma$

13. less than $x + 2\sigma$

14. less than $x - 2\sigma$

15. between $x - 1\sigma$ and $x + 2\sigma$

16. between $x - 2\sigma$ and $x + 2\sigma$

17. A mathematics placement test is given to 6500 entering freshmen. The scores have a normal distribution with a mean of 75 and a standard deviation of 8. How many freshmen scored

a. between 67 and 83? **b.** between 59 and 91?

c. above 91? **d.** below 51?

e. between 75 and 83? **f.** between 75 and 91?

18. The mean score on a certain IQ test is 100 with a standard deviation of 15. There were 10,000 students who took this test, and their scores have a normal distribution. How many students have

a. an IQ between 85 and 115? **b.** an IQ between 70 and 85?

c. an IQ over 130? **d.** an IQ below 85?

e. an IQ between 115 and 130? **f.** an IQ over 145?

19. The mean weight of 2000 male students at a community college is 153 lb with a standard deviation of 15 lb. If the weights are normally distributed, how many students weigh

a. less than 153 lb? **b.** more than 183 lb?

c. between 138 and 168 lb? **d.** between 168 and 183 lb?

20. The mean amount of soft drink in a bottle is 2 L. The standard deviation is 25 mL and 100,000 bottles are produced each day. If the amount of liquid is normally distributed, how many bottles contain

a. between 1.975 L and 2.025 L? **b.** between 1.95 L and 2.05 L?

c. more than 2.05 L? **d.** less than 1.95 L?

21. The mean inside diameter of the bottle caps manufactured by a machine is 0.72 in. with a standard deviation of 0.005 in. A quality control manager picks one at random. What is the probability the cap's diameter is greater than 0.71 in., but less than 0.725 in., assuming diameters are normally distributed?

22. The mean length of copper water pipes made by a machine is 200 cm with a standard deviation of 0.25 cm. Assuming the lengths are normally distributed, what is the probability that a pipe randomly taken from the production line is longer than 199.5 cm but shorter than 200.25 cm?

WORKING WITH FORMULAS

The normal distribution function $f(x) = (2\pi e^{x^2})^{-\frac{1}{2}}$

The graph of the normal distribution function is given by the formula shown.

23. **a.** Verify that the function can be written in the form $y = \frac{1}{\sqrt{2\pi}} \cdot e^{\left(\frac{-x^2}{2}\right)}$

b. Compute the value of $f(2)$ and $f(-2)$. What do you notice? Compute $f(-1)$ and $f(1)$ to confirm.

c. Investigate what happens to y as x gets larger: $x = 0.5, 1, 1.5, 2, 2.5,$ and so on. How are these results reflected in the shape of the graph?

24. **a.** Show that the function can be in the form $y = \frac{1}{\sqrt{2\pi e^{x^2}}}$.

b. Use a table with ΔTbl $= 0.1$ to determine the point at which outputs are less than 0.10 and interpret the significance in terms of the total distribution.

c. What is the maximum value of this function? What is its significance?

APPLICATIONS

25. **Pencil manufacture:** A pencil manufacturer finds for all pencils produced, $\bar{x} = 150$ mm and of $\sigma = 2$ mm. One pencil is chosen at random. Determine the following probabilities (let L = length in millimeters):

a. $P(L > 154)$ **b.** $P(L < 144)$

c. $P(146 < L < 152)$ **d.** $P(L > 156)$

e. In a batch of 10,000 pencils, about how how many are less than 148 mm long?

f. In a batch of 15,000 pencils, about how many are greater than 146 mm but less than 150 mm?

26. **Exam times:** The average time required to complete an exam is $\bar{x} = 90$ min with a standard deviation of $\sigma = 10$ min. One student is selected at random. Determine the following probabilities (let T = time required to complete test in minutes):

a. $P(T > 110)$ **b.** $P(T < 70)$

c. $P(80 < T < 110)$ **d.** $P(70 < T < 80)$

e. How much time should be allowed to ensure that 99.7% test-takers can complete the test?

f. A person completes the test in 60 min. What percentage of test takers could be expected to finish faster?

Compute the z-scores using the information given, then use the table from Appendix V to determine what percent of the data falls *to the right* the computed z-score.

27. data value: 135, $\bar{x} = 152$, $\sigma = 12$

28. data value: 60, $\bar{x} = 73$, $\sigma = 9$

29. data value: 17, $\bar{x} = 15.2$, $\sigma = 0.9$

30. data value: 2.62, $\bar{x} = 2.5$, $\sigma = 0.05$

31. data value: 83, $\bar{x} = 75$, $\sigma = 12$

32. data value: 0.3, $\bar{x} = 0.25$, $\sigma = 0.04$

Use the z-table from Appendix V to complete these exercises. Assume all distributions are normal.

33. Appliance service life: The Cool-Blow Company manufactures household air conditioners with an average service life of $\bar{x} = 10$ yr with a standard deviation of $\sigma = 9$ months. One air conditioning unit is selected at random for testing. Determine the following probabilities (let S = service life in years):

a. $P(S < 9)$

b. $P(S > 12)$

c. $P(8.\overline{3} < S < 11.25)$

d. $P(S > 10.5)$

e. In March, the Cool-Blow Company produced 1500 air conditioners. How many can be expected to have a service life of over 11 yr and out last all manufacturer warranties?

f. In June production was 2500 units. The Quality Control Division has stated that a unit that lasts less than 9 yr is unfit for sale to the public. How many of them were unfit for sale?

34. Appliance service life: A company selling electric heaters finds that the mean lifetime of the heaters is $\bar{x} = 4000$ hr with a standard deviation of $\sigma = 250$ hr. One unit is picked at random for testing. Determine the following probabilities (let S = service life in hours):

a. $P(S < 3700)$

b. $P(S > 4350)$

c. $P(3820 < S < 4100)$

d. $P(S > 4700)$

e. In January, the company produced 5000 heaters. How many can be expected to have a service life of over 4750 hr and out last all manufacturer warranties?

f. In June production was 1500 units. The Quality Control Division has stated that a unit that lasts less than 3260 hr is unfit for sale to the public. How many of them were unfit for sale?

35. Average income: In a small underdeveloped country the average income per household is $\bar{x} = \$18{,}000$ with a standard deviation of $\sigma = \$1{,}125$. One family is chosen at random. Determine the probability (let I = income):

a. $P(I > \$19{,}000)$

b. $P(I < \$15{,}000)$

c. $P(\$16{,}000 < I < \$17{,}000)$

d. $P(I > \$20{,}500)$

e. The Department of Human Services considers families making less than \$16,000 to be living in poverty. If there are 5750 families in this country, how many of them are living in poverty?

f. Is it possible for there to be families making over \$25,000 living in this county? Discuss/explain.

36. Beam strength: Thick wooden beams are subjected to a stress test and found to have a mean breaking point of $\bar{x} = 1250$ lb, with a standard deviation of $\sigma = 150$ lb. One such beam is randomly chosen from inventory. Determine the following probabilities (let S = breaking point in pounds):

a. $P(S > 1300)$

b. $P(S < 1200)$

c. $P(1200 < S < 1300)$

d. $P(S > 1500)$

e. The manufacturer considers planks with a breaking strength of less than 875 lb to be defective. In a lot of 5000 of these planks, how many are defective?

f. Is it possible for a plank to have a breaking point of $S = 650$ lb? Discuss/explain.

37. Test scores: The following data gives the final exam scores for the 40 students in a general psychology course: 88, 67, 25, 99, 100, 72, 79, 89, 69, 99, 77, 42, 83, 75, 100, 88, 87, 53, 82, 91, 95, 92, 81, 76, 56, 69, 82, 95, 91, 91, 81, 57, 69, 95, 76, 88, 85, 89, 90, 77. One student is picked at random. Determine the probability that his/her score is at least a 64.

38. Batting average: The following data gives the batting averages of 25 players on the university softball team: 0.288, 0.267, 0.225, 0.299, 0.300, 0.272, 0.279, 0.289, 0.269, 0.299, 0.277, 0.242, 0.283, 0.275, 0.325, 0.288, 0.287, 0.253, 0.282, 0.291, 0.295, 0.292, 0.281, 0.276, 0.256. One player is picked at random. What is the probability her batting average is at least 0.280?

WRITING, RESEARCH, AND DECISION MAKING

39. In 1662, a London merchant name John Graunt wrote a paper called, "Natural and Political Observations Made upon Bills of Mortality." Many say this paper helped launch a more formal study of statistics. Do some research on John Graunt and why he wrote this paper, and include some of the conclusions he made from his research. In what way are his findings applied today?

40. A departmental exam is given to all students taking elementary applied mathematics. Two of the classes have the same mean, but one class has a standard deviation one and a half times as large as the other. Which class would be "more difficult" to teach? Why?

41. You and a friend recently took college algebra, but from different instructors. She begins "ribbing" you because she scored 84% while you only scored an 78% on the final exam. Later you find out that your test had a mean of $\bar{x} = 59$ with $\sigma = 20$, while her test had a mean of $\bar{x} = 76$ and $\sigma = 12$. Use a z-score to determine who actually got the "better" score.

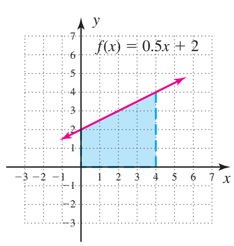
EXTENDING THE CONCEPT

42. Last semester an instructor had two sections of a college algebra class. On the final exam, there was a mean of 72 and a standard deviation of 10 for the first class, while the exam for the second class had a mean of 72 and a standard deviation of 5. What conclusions can be drawn?

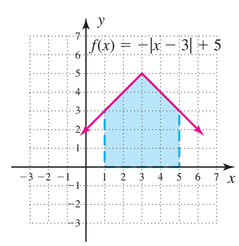
43. Standard IQ tests use a mean value of $\bar{x} = 100$ with a standard deviation of $\sigma = 20$. Assuming a normal distribution, what must a person score to be in the top 1% of the population?

44. The concept of "area under a graph" has many applications in additional to the normal curve. Find the shaded area under each graph using elementary geometry. For (d), see Section 3.3, Example 10.

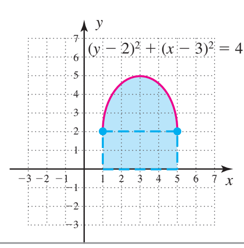
a.



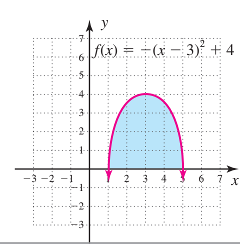
b.



c.



d.



Section 8.9 (shown in this document as 11.9) Student Solutions

1. standard deviation
3. distribution, area
5. the difference between a given value x_i and the average value $\bar{x}$, divided by the standard deviation σ. A *z*-score represents a non-standard deviation from the mean.
7. a) $\bar{x} = 79.4$; $\sigma \approx 18.73$ b) verified
9. a) $\bar{x} = 0.174$; $\sigma \approx 0.041$ b) verified
11. 16%
13. 97.5%
15. 81.5%
17. a) 4420 b) 6175 c) 163 d) ≈ 10 e) 2210 f) ≈ 3088
19. a) 1000 b) 50 c) 1360 d) 270
21. 0.815
23. a) verified b) $f(2) \approx 0.0540 = f(-2) \approx 0.0540$; They have the same value; $f(-1) = f(1)$ ✓; c) as x gets larger, $f(x) \to 0$. The graph is symmetric to the y-axis and asymptotic to the x-axis.
25. a) 0.025 b) 0.0015 c) 0.815 d) 0.0015 e) 1600 f) 7125
27. $z \approx 1.42$; 92.22%
29. $z \approx 2.00$; 97.72%
31. $z \approx 0.67$; 74.86%
33. a) 9.18% b) 0.39% c) 93.4% d) 25.46% e) 138 units f) 230 units
35. a) 18.67% b) 0.38% c) 14.92% d) 1.32% e) 216 f) possible/unlikely
37. about 0.84
39. Answers will vary.
41. Her *z*-score is 0.67 while your *z*-score is 0.95; clearly you did better on the test.
43. They must score at least 146.6.