

## CHECK-IN


**1.4 Apartment rentals.** A data set lists apartments available for students to rent. Information provided includes the monthly rent per person, whether cable is included free of charge, whether or not pets are allowed, the number of bedrooms, the number of bathrooms, and the distance to the campus. Describe the cases in the data set, give the number of variables, and specify whether each variable is categorical or quantitative.

## 1.8 Attending college in your state or in another state.

The U.S. Census Bureau collects a large amount of information concerning higher education.<sup>2</sup> For example, the bureau provides a table that includes the following variables: state, number of students from the state who attend college, and number of students who attend college in their home state.

- What are the cases for this set of data?
- Is there a label variable? If yes, what is it?
- Identify each variable as categorical or quantitative.
- Explain how you might use each of the quantitative variables to explain something about the states.
- Consider a variable computed as the number of students in each state who attend college in the state divided by the total number of students from the state who attend college. Explain how you would use this variable to explain something about the states.

## CHECK-IN

**1.8 Make a stemplot.** Here are the scores on the first exam in an introductory statistics course for 28 students in one section of the course: 

73	92	82	75	98	94	57	80	90	92	80	87	91	65
70	85	83	61	70	90	75	75	59	68	85	78	80	94

Use these data to make a stemplot. Then use the stemplot to describe the distribution of the first-exam scores for this course.

**1.47 Median versus mean for net worth.** A report on the assets of American households says that the median net worth of U.S. families is \$97,300. The mean net worth of these families is \$692,100.<sup>26</sup> What explains the difference between these two measures of center?

### 1.62 Means and medians.


- (a) Sketch a symmetric distribution that is *not* Normal. Mark the location of the mean and the median.
- (b) Sketch a distribution that is skewed to the right. Mark the location of the mean and the median.

### 1.63 The effect of changing the standard deviation.

- (a) Sketch a Normal curve that has mean 20 and standard deviation 2.
- (b) On the same  $x$  axis, sketch a Normal curve that has mean 20 and standard deviation 4.
- (c) How does the Normal curve change when the standard deviation is varied but the mean stays the same?

### 1.64 The effect of changing the mean.

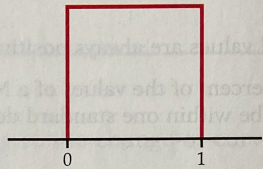
- (a) Sketch a Normal curve that has mean 20 and standard deviation 2.
- (b) On the same  $x$  axis, sketch a Normal curve that has mean 30 and standard deviation 2.
- (c) How does the Normal curve change when the mean is varied but the standard deviation stays the same?

**1.71 Do women talk more?** Conventional wisdom suggests that women are more talkative than men. One study designed to examine this stereotype collected data on the speech of 42 women and 37 men in the United States.<sup>32</sup>  TALK

- (a) The mean number of words spoken per day by the women was 14,297, with a standard deviation of 6441. Use the 68–95–99.7 rule to describe this distribution.
- (b) Do you think that applying the rule in this situation is reasonable? Explain your answer.
- (c) The men averaged 14,060 words per day, with a standard deviation of 9065. Answer the questions in parts (a) and (b) for the men.
- (d) Do you think that the data support the conventional wisdom? Explain your answer. Note that in Section 7.2 we will learn formal statistical methods to answer this type of question.

**1.73 A uniform distribution.** If you ask a computer to generate “random numbers” between 0 and 1, you will get observations from a **uniform distribution**. FIGURE 1.33 graphs the density curve for a uniform distribution. Use areas under this density curve to answer the following questions.

- (a) What proportion of the observations lie below 0.75?
- (b) What proportion of the observations lie below 0.50?
- (c) What proportion of the observations lie between 0.50 and 0.75?
- (d) Why is the total area under this curve equal to 1?



**FIGURE 1.33** The density curve of a uniform distribution, Exercise 1.73.

**1.74 Use a different range for the uniform distribution.** Many random number generators allow users to specify the range of the random numbers to be produced. Suppose that you specify that the outcomes are to be distributed uniformly between 0 and 5. Then the density curve of the outcomes has constant height between 0 and 5 and height 0 elsewhere.

- (a) What is the height of the density curve between 0 and 5? Draw a graph of the density curve.
- (b) Use your graph from part (a) and the fact that areas under the curve are proportions of outcomes to find the proportion of outcomes that are more than 2.
- (c) Find the proportion of outcomes that lie between 2.5 and 3.0.

**1.75 Find the mean, the median, and the quartiles.** What are the mean and the median of the uniform distribution in Figure 1.33? What are the quartiles?

**1.79 Find some proportions.** Using either software or Table A, find the proportion of observations from a standard Normal distribution that satisfies each of the following statements. In each case, sketch a standard Normal curve and shade the area under the curve that is the answer to the question.

- (a)  $Z > 1.85$
- (b)  $Z < 1.85$
- (c)  $Z > -0.90$
- (d)  $-0.90 < Z < 1.85$

**1.82 Find more values of  $z$  illustrate the result with a sketch.** The variable  $Z$  has a standard Normal distribution.

- (a) Find the number  $z$  that has cumulative proportion 0.68.
- (b) Find the number  $z$  such that the event  $Z > z$  has proportion 0.122.

**1.95 Do you have enough “good cholesterol”?** High-density lipoprotein (HDL) is sometimes called the “good cholesterol” because high values are associated with a reduced risk of heart disease. According to the American Heart Association, people over the age of 20 years should have at least 40 milligrams per deciliter (mg/dl) of HDL cholesterol.<sup>33</sup> U.S. women aged 20 and over have a mean HDL of 55 mg/dl with a standard deviation of 15.5 mg/dl. Assume that the distribution is Normal.

- (a) What percent of women have low values of HDL (40 mg/dl or less)?
- (b) HDL levels of 60 mg/dl and higher are believed to protect people from heart disease. What percent of women have protective levels of HDL?
- (c) Women with more than 40 mg/dl but less than 60 mg/dl of HDL are in the intermediate range, neither very good or very bad. What proportion are in this category?