b. A mean of 96 is observed based on a sample of 60 scores.

c. A mean of 100 is observed based on a sample of 29 scores.

Applying the Concepts

6.29 We asked 150 students (in our statistics classes) how long, in minutes, they typically spent getting ready for a date. The scores range from 1 minute to 120 minutes, and the mean is 51.52 minutes. Here are the data for 40 of those students:

30	90	60	60	5	90	30	40	45	60
60	30	90	60	25	10	90	20	15	60
60	75	45	60	30	75	15	30	45	1
20	25	45	60	90	10	105	90	30	60

a. Construct a histogram for the 10 scores in the first row.

b. Construct a histogram for all 40 of these scores.

c. What happened to the shape of the distribution as you increased the number of scores from 10 to 40? What do you think would happen if the data for all 150 students were included? What if we included 10,000 scores? Explain this phenomenon.

d. Are these distributions of scores or distributions of means? Explain.

e. The data here are self-reported. That is, our students wrote down how many minutes they believe that they typically take to get ready for a date. This accounts for the fact that the data include many "pretty" numbers, such as 30, 60, or 90 minutes. What might have been a better way to operationalize this variable?

f. Do these data suggest any hypotheses that you might like to study? List at least one.

6.30 The verbal subtest of the GRE has a population mean of 500 and a population standard deviation of 100 by design (the quantitative subtest has the same mean and standard deviation).

a. Use symbolic notation to state the mean and standard deviation of the GRE verbal test.

700-500 / 100

b. Convert a GRE score of 700 to a z score *without* using a formula.

550-500 / 100

c. Convert a GRE score of 550 to a z score *without* using a formula.

400-500 / 100

d. Convert a GRE score of 400 to a z score *without* using a formula.

6.31 A sample of 150 statistics students reported the typical number of hours that they sleep on a weeknight. The mean number of hours was 6.65, and the standard deviation was 1.24. (For this exercise, treat this sample as the entire population of interest.)

a. What is *always* the mean of the z distribution?

b. Using the sleep data, demonstrate that your answer to part (a) is the mean of the z distribution. (*Hint:* Calculate the z score for a student who is exactly at the mean.)

c. What is *always* the standard deviation of the z distribution?

d. Using the sleep data, demonstrate that your answer to part (c) is the standard deviation of the z distribution. (*Hint:* Calculate the z score for a student who is exactly 1 standard deviation above or below the mean.)

e. How many hours of sleep do you typically get on a weeknight? What would your z score be compared with this population?

6.32 A sample of 148 of our statistics students rated their level of admiration for Hillary Rodham Clinton on a scale of 1 to 7. The mean rating was 4.06, and the standard deviation was 1.70. (For this exercise, treat this sample as the entire population of interest.)

a. Use these data to demonstrate that the mean of the z distribution is always 0.

b. Use these data to demonstrate that the standard deviation of the z distribution is always 1.

c. Calculate the z score for a student who rated his admiration of Hillary Rodham Clinton as 6.1.

d. A student had a z score of -0.55. What rating did she give for her admiration of Hillary Rodham Clinton?

6.33 We have already discussed summary parameters for CFC scores for the population of participants in a study by Petrocelli (2003). The mean CFC score was 3.51, with a standard deviation of 0.61. (Remember that even though this was a sample, we treated the sample of 664 participants as the entire population.) Imagine that you randomly selected 40 people from this population and had them watch a series of videos on financial planning after graduation. The mean CFC score after watching the video was 3.62. We want to know whether watching these videos might change CFC scores in the population.

a. Why would it not make sense to compare this sample with the distribution of scores? Be sure to discuss the spread of distributions in your answer.

b. Using symbolic notation and formulas, what are the appropriate measures of central tendency and variability for the distribution from which this sample comes?

c. Using symbolic notation and the formula, what is the z statistic for this sample mean?

d. Roughly, to what percentile does that z statistic correspond?

6.34 A CFC study found a mean CFC score of 3.51, with a standard deviation of 0.61, for the 664 students in the sample (Petrocelli, 2003).

a. Imagine that your z score on the CFC score is -1.2. What is your raw score? Use symbolic notation and the formula. Explain why this answer makes sense.

b. Imagine that your z score on the CFC score is 0.66. What is your raw score? Use symbolic notation and the formula. Explain why this answer makes sense.

6.35 For each of the following variables, would the distribution of scores likely approximate a normal curve? Explain your answer.

a. Number of movies that college students watch in a year

b. Number of full-page advertisements in magazines

c. Human birth weight in Canada

6.36 Georgiou and colleagues (1997) reported that college students had healthier eating habits, on average, than did those who were neither college students nor college graduates. The 412 students in the study ate breakfast a mean of 4.1 times per week with a standard deviation of 2.4. For this exercise, imagine that this is the entire population of interest; thus, these numbers can be treated as parameters.

a. Roughly, what is the percentile for a student who eats breakfast four times per week?

b. Roughly, what is the percentile for a student who eats breakfast six times per week?

c. Roughly, what is the percentile for a student who eats breakfast twice a week?

6.37 A common quandary faces sports fans who live in the same city but avidly follow different sports. How does one determine whose team did better with respect to its league division? In 2004, the Boston Red Sox won the World Series; just months later, their local football counterparts, the New England Patriots, won the Super Bowl. In 2005, both teams made the playoffs but lost early on. Which team was better in 2005? The question, then, is: Were the Red Sox better, as compared to other teams in the American League of Major League Baseball, than the Patriots, as compared to the other teams in the American Football Conference of the National Football League? Some of us could debate it for hours, but it's better to examine some statistics. Let's operationalize performance over the season as the number of wins during regular season play.

a. In 2005, the mean number of wins for baseball teams in the American League was 81.71, with a standard deviation of 13.07. Because all teams were included, these are population parameters. The Red Sox won 95 games. What is their z score?

b. In 2005, the mean number of wins for football teams in the American Football Conference was 8.13, with a standard deviation of 3.70. The Patriots won 10 games. What is their z score?

c. Which team did better, according to these data?

d. How many games would the team with the lower z score have had to win to beat the team with the higher z score?

e. List at least one other way we could have operationalized the outcome variable (i.e., team performance).

6.38 Our statistics students, as noted in Exercise 6.32, were asked to rate their admiration of Hillary Rodham Clinton on a scale of 1 to 7. They also were asked to rate their admiration of Jennifer Lopez and Venus Williams on a scale of 1 to 7. As noted earlier, the mean rating of Clinton was 4.06 with a standard deviation of 1.70. The mean rating of Lopez was 3.72 with a standard deviation of 1.90. The mean rating of Williams was 4.58 with a standard deviation of 1.46. One of our students rated her admiration of Clinton and Williams at 5 and her admiration of Lopez at 4.

a. What is her z score for her admiration rating of Clinton?

b. What is her z score for her admiration rating of Williams?

c. What is her z score for her admiration rating of Lopez?

d. Compared to the other statistics students in our sample, which celebrity does this student most admire? (We can tell by her raw scores that she prefers Clinton and Williams to Lopez, but when we take into account the general perception of these celebrities, how does this student feel about them?)

e. How do z scores allow us to make comparisons that we cannot make with raw scores? That is, describe the benefits of standardization.

6.39 Let's look at baseball and football again, but this time we'll look at data for all of the teams in Major League Baseball (MLB) and the National Football League (NFL), respectively.

a. In 2005, the mean number of wins for MLB teams was 81.00, with a standard deviation of 10.83. The perennial underdogs, the Chicago Cubs, had a z score of -0.18. How many games did they win?

b. In 2005, the mean number of wins for all NFL teams was 8.00, with a standard deviation of 3.39. The New Orleans Saints had a z score of -1.475. How many games did they win?

c. The Pittsburgh Steelers were just below the 84th percentile in terms of NFL wins. How many games did they win? Explain how you obtained your answer.

d. Explain how you can examine your answers in parts (a), (b), and (c) to determine if the numbers make sense.

6.40 Researchers have reported that the projected life expectancy for people diagnosed with human immunodeficiency virus (HIV) and receiving antiretroviral therapy (ART) is 24.2 years (Schackman et al, 2006). Imagine that the researchers determined this by following 250 people with HIV who were receiving ART and calculating the mean. (The 24.2 is actually a projected number rather than a mean for a sample.)

a. What is the variable of interest?

b. What is the population?

c. What is the sample?

d. For the population, describe what the distribution of *scores* would be.

e. For the population, describe what the distribution of *means* would be.

f. If the distribution of the population were skewed, would the distribution of scores likely be skewed or approximately normal? Explain your answer.

g. Would the distribution of means be skewed or approximately normal? Explain your answer.

6.41 The revised version of the Minnesota Multiphasic Personality Inventory (MMPI-2) is the most frequently administered self-report personality measure. Test-takers respond to more than 500 true/false statements, and their responses are scored, typically by a computer, on a number of scales (e.g., hypochondriasis, depression, psychopathic deviation). Respondents receive a *T* score on each scale that can be compared to norms. (It is important to note that *T* scores are different from the *t* statistic we will learn about in a few chapters; you're likely to encounter *T* scores if you take psychology classes, and it's good to be aware that they're different from the *t* statistic.) *T* scores are another way to standardize scores so that percentiles and cutoffs can be determined. The mean *T* score is always 50, and the standard deviation is always 10. Imagine that you administer the MMPI-2 to 95 respondents who have recently lost a parent; you wonder whether their scores on the depression scale will be, on average, higher than the norms. You find a mean score on the depression scale of 55 in your sample.

a. Using symbolic notation, report the mean and standard deviation of the population.

b. Using symbolic notation and formulas (where appropriate), report the mean and standard error for the distribution of means to which your sample will be compared.

c. In your own words, explain why it makes sense that the standard error is smaller than the standard deviation.

6.42 You may need to find an apartment to rent upon graduation. The Internet is a valuable source of data to aid you in your search. From neighborhood safety to available transportation to housing costs, recent data can steer you in the right direction. On a Web site, San Mateo County in California published extensive descriptive statistics from its 1998 Quality of Life Survey. The county reported that the mean house payment (mortgage or rent) was \$1225.15, with a standard deviation of \$777.50. It also reported that the mean cost of an apartment rental, rather than a house rental or a mortgage, was \$868.86. For this exercise, treat the overall mean housing payment as a parameter, and treat the mean apartment rental cost as a statistic based on a sample of 100.

a. Using symbolic notation and formulas (where appropriate), determine the mean and the standard error for the distribution of means for the overall housing payment data.

b. Using symbolic notation and the formula, calculate the *z* statistic for the cost of an apartment rental.

c. Why is it likely that the *z* statistic is so large? (*Hint:* Is this distribution likely to be normal? Explain.)

d. Why is it permissible to use the normal curve percentages associated with the *z* distribution even though the data are not likely normally distributed?

6.43 The GSS is a survey of approximately 2000 adults conducted each year since 1972, for a total of more than 38,000 people. During several years of the GSS, participants were asked how many close friends they have. The mean for this variable is 7.44 friends, with a standard deviation of 10.98. The median is 5.00 and the mode is 4.00.

a. Are these data for a distribution of scores or a distribution of means? Explain.

b. What do the mean and standard deviation suggest about the shape of the distribution? (*Hint:* Compare the sizes of the mean and the standard deviation.)

c. What do the three measures of central tendency suggest about the shape of the distribution?

d. Let's say that these data represent the entire population. Pretend that you randomly selected a person from this population and asked how many close friends she or he had. Would you compare this person to a distribution of scores or a distribution of means? Explain your answer.

e. Now pretend that you randomly selected a sample of 80 people from this population. Would you compare this sample to a distribution of scores or a distribution of means? Explain your answer.