CHAPTER 6

Section 6.1

6.1 The Acculturation Rating Scale for Mexican Americans (ARSMA) is a psychological test developed to measure the degree of Mexican/Spanish versus Anglo/English acculturation of Mexican Americans. The distribution of ARSMA scores in a population used to develop the test was approximately normal, with mean 3.0 and standard deviation 0.8. A further study gave ARSMA to 50 first-generation Mexican Americans. The mean of their scores is $\bar{x} = 2.36$. If the standard deviation for the first-generation population is also $\sigma = 0.8$, give a 95% confidence interval for the mean ARSMA score for first-generation Mexican Americans.

6.2 The 2000 census “long form” asked the total 1999 income of the householder, the person in whose name the dwelling unit was owned or rented. This census form was sent to a random sample of 17% of the nation’s households. Suppose that the households that returned the long form are an SRS of the population of all households in each district. In Middletown, a city of 40,000 persons, 2621 householders reported their income. The mean of the responses was $\bar{x} = $33,453, and the standard deviation was $s = $8721. The sample standard deviation for so large a sample will be very close to the population standard deviation $\sigma$. Use these facts to give an approximate 99% confidence interval for the 1999 mean income of Middletown householders who reported income.

6.3 Refer to the previous problem. Give a 99% confidence interval for the total 1999 income of the households that reported income in Middletown.

6.4 A newspaper headline describing a poll of registered voters taken two weeks before a recent election read “Ringel leads with 52%.” The accompanying article describing the poll stated that the margin of error was 2% with 95% confidence.
(a) Explain in plain language to someone who knows no statistics what “95% confidence” means.
(b) The poll shows Ringel leading. But the newspaper article said that the election was too close to call. Explain why.

6.5 A student reads that a 95% confidence interval for the mean SAT math score of California high school seniors is 452 to 470. Asked to explain the meaning of this interval, the student says, “95% of California high school seniors have SAT math scores between 452 and 470.” Is the student right? Justify your answer.

6.6 As we prepare to take a sample and compute a 95% confidence interval, we know that the probability that the interval we compute will cover the parameter is 0.95. That’s the meaning of 95% confidence. If we use several such intervals, however, our confidence that all give correct results is less than 95%.

6.7 In an agricultural field trial a corn variety is planted in seven separate locations, which may have different mean yields due to differences in soil and climate. At the end of the experiment, seven independent 95% confidence intervals will be calculated,
one for the mean yield at each location.
(a) What is the probability that every one of the seven intervals covers the true mean yield at its location? This probability (expressed as a percent) is our overall confidence level for the seven simultaneous statements.
(b) What is the probability that at least six of the seven intervals cover the true mean yields?

6.8 A newspaper headline describing a poll of registered voters taken two weeks before a recent election read “Ringel leads with 52%.” The accompanying article describing the poll stated that the margin of error was 2% with 95% confidence.
(a) Explain in plain language to someone who knows no statistics what “95% confidence” means.
(b) The poll shows Ringel leading. But the newspaper article said that the election was too close to call. Explain why.

6.9 A newspaper ad for a manager trainee position contained the statement “Our manager trainees have a first-year earnings average of $20,000 to $24,000.” Do you think that the ad is describing a confidence interval? Explain your answer.

6.10 A student reads that a 95% confidence interval for the mean SAT math score of California high school seniors is 452 to 470. Asked to explain the meaning of this interval, the student says, “95% of California high school seniors have SAT math scores between 452 and 470.” Is the student right? Justify your answer.

6.11 A survey of users of the Internet found that males outnumbered females by nearly 2 to 1. This was a surprise, because earlier surveys had put the ratio of men to women closer to 9 to 1. Later in the article we find this information:

Detailed surveys were sent to more than 13,000 organizations on the Internet; 1,468 usable responses were received. According to Mr. Quartermaw, the margin of error is 2.8 percent, with a confidence level of 95 percent.11

(a) What was the response rate for this survey? (The response rate is the percent of the planned sample that responded.)
(b) Do you think that the small margin of error is a good measure of the accuracy of the survey’s results? Explain your answer.

6.12 The mean amount $\mu$ for all of the invoices for your company last month is not known. Based on your past experience, you are willing to assume that the standard deviation of invoice amounts is about $200.
(a) If you take a random sample of 100 invoices, what is the value of the standard deviation for $\bar{x}$?
(b) The $68$–$95$–$99.7$ rule says that the probability is about 0.95 that $\bar{x}$ is within ______ of the population mean $\mu$. Fill in the blank.
(c) About 95% of all samples will capture the true mean of all of the invoices in the interval $\bar{x}$ plus or minus ______. Fill in the blank.
6.13 You measure the weights of 24 male runners. You do not actually choose an SRS, but you are willing to assume that these runners are a random sample from the population of male runners in your town or city. Here are their weights in kilograms:

<table>
<thead>
<tr>
<th>67.8</th>
<th>61.9</th>
<th>63.0</th>
<th>53.1</th>
<th>62.3</th>
<th>59.7</th>
<th>55.4</th>
<th>58.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>60.9</td>
<td>69.2</td>
<td>63.7</td>
<td>68.3</td>
<td>64.7</td>
<td>65.6</td>
<td>65.6</td>
<td>57.8</td>
</tr>
<tr>
<td>66.0</td>
<td>62.9</td>
<td>53.6</td>
<td>65.0</td>
<td>55.8</td>
<td>60.4</td>
<td>69.3</td>
<td>61.7</td>
</tr>
</tbody>
</table>

Suppose that the standard deviation of the population is known to be \( \sigma = 4.5 \) kg.

(a) What is \( \sigma_{\bar{x}} \), the standard deviation of \( \bar{x} \)?

(b) Give a 95% confidence interval for \( \mu \), the mean of the population from which the sample is drawn. Are you quite sure that the average weight of the population of runners is less than 65 kg?

6.14 Suppose that you had measured the weights of the runners in the previous exercise in pounds rather than kilograms. Use your answers to the previous exercise and the fact that 1 kilogram equals 2.2 pounds to answer these questions.

(a) What is the mean weight of these runners?

(b) What is the standard deviation of the mean weight?

(c) Give a 95% confidence interval for the mean weight of the population of runners that these runners represent.

6.15 Crop researchers plant 15 plots with a new variety of corn. The yields in bushels per acre are

\[
138.0 \quad 139.1 \quad 113.0 \quad 132.5 \quad 140.7 \quad 109.7 \quad 118.9 \quad 134.8, \\
109.6 \quad 127.3 \quad 115.6 \quad 130.4 \quad 130.2 \quad 111.7 \quad 105.5
\]

Assume that \( \sigma = 10 \) bushels per acre.

(a) Find the 90% confidence interval for the mean yield \( \mu \) for this variety of corn.

(b) Find the 95% confidence interval.

(c) Find the 99% confidence interval.

(d) How do the margins of error in (a), (b), and (c) change as the confidence level increases?

6.16 Suppose that the crop researchers in the previous exercise obtained the same value of \( \bar{x} \) from a sample of 50 plots rather than 15.

(a) Compute the 95% confidence interval for the mean yield \( \mu \).

(b) Is the margin of error larger or smaller than the margin of error found for the sample of 15 plots in the previous exercise? Explain in plain language why the change occurs.

(c) Will the 90% and 99% intervals for a sample of size 50 be wider or narrower than those for \( n = 15 \)? (You need not actually calculate these intervals.)

6.17 In the two previous exercises, we compared confidence intervals based on corn yields from 15 and 50 small plots of ground. How large a sample is required to estimate the mean yield within ±6 bushels per acre with 90% confidence?

6.18 A test for the level of potassium in the blood is not perfectly precise. Moreover, the actual level of potassium in a person’s blood varies slightly from day to day.
Suppose that repeated measurements for the same person on different days vary
normally with $\sigma = 0.2$.

(a) Julie’s potassium level is measured once. The result is $x = 3.4$. Give a 90%
confidence interval for her mean potassium level.

(b) If three measurements were taken on different days and the mean result is
$\bar{x} = 3.4$, what is a 90% confidence interval for Julie’s mean blood potassium level?

6.19 How large a sample of Julie’s potassium levels in the previous exercise would
be needed to estimate her mean $\mu$ within $\pm 0.06$ with 95% confidence?

6.20 A study of the career paths of hotel general managers sent questionnaires to an
SRS of 160 hotels belonging to major U.S. hotel chains. There were 114 responses.
The average time these 114 general managers had spent with their current company
was 11.78 years. Give a 99% confidence interval for the mean number of years
general managers of major-chain hotels have spent with their current company.
(Take it as known that the standard deviation of time with the company for all
general managers is 3.2 years.)

6.21 Researchers planning a study of the reading ability of third-grade children want
to obtain a 95% confidence interval for the population mean score on a reading test,
with margin of error no greater than 3 points. They carry out a small pilot study
to estimate the variability of test scores. The sample standard deviation is $s = 12$
points in the pilot study, so in preliminary calculations the researchers take the
population standard deviation to be $\sigma = 12$.

(a) The study budget will allow as many as 100 students. Calculate the margin of
error of the 95% confidence interval for the population mean based on $n = 100$.

(b) There are many other demands on the research budget. If all of these demands
were met, there would be funds to measure only 10 children. What is the margin of
error of the confidence interval based on $n = 10$ measurements?

(c) Find the smallest value of $n$ that would satisfy the goal of a 95% confidence
interval with margin of error 3 or less. Is this sample size within the limits of the
budget?

6.22 The Gallup Poll asked 1571 adults what they considered to be the most serious
problem facing the nation’s public schools; 30% said drugs. This sample percent
is an estimate of the percent of all adults who think that drugs are the schools’
most serious problem. The news article reporting the poll result adds, “The poll
has a margin of error—the measure of its statistical accuracy—of three percentage
points in either direction; aside from this imprecision inherent in using a sample to
represent the whole, such practical factors as the wording of questions can affect
how closely a poll reflects the opinion of the public in general.”

The Gallup Poll uses a complex multistage sample design, but the sample per-
cent has approximately a normal distribution. Moreover, it is standard practice to
announce the margin of error for a 95% confidence interval unless a different confi-
dence level is stated.

(a) The announced poll result was $30\% \pm 3\%$. Can we be certain that the true
population percent falls in this interval?

(b) Explain to someone who knows no statistics what the announced result $30\% \pm 3\%$
Section 6.1

(c) This confidence interval has the same form we have met earlier:

\[
\text{estimate} \pm z^* \sigma_{\text{estimate}}
\]

(Actually \(\sigma\) is estimated from the data, but we ignore this for now.) What is the standard deviation \(\sigma_{\text{estimate}}\) of the estimated percent?

(d) Does the announced margin of error include errors due to practical problems such as undercoverage and nonresponse?

6.23 When the statistic that estimates an unknown parameter has a normal distribution, a confidence interval for the parameter has the form

\[
\text{estimate} \pm z^* \sigma_{\text{estimate}}
\]

In a complex sample survey design, the appropriate unbiased estimate of the population mean and the standard deviation of this estimate may require elaborate computations. But when the estimate is known to have a normal distribution and its standard deviation is given, we can calculate a confidence interval for \(\mu\) from complex sample designs without knowing the formulas that led to the numbers given.

A report based on the Current Population Survey estimates the 1999 median annual earnings of households as $40,816 and also estimates that the standard deviation of this estimate is $191. The Current Population Survey uses an elaborate multistage sampling design to select a sample of about 50,000 households. The sampling distribution of the estimated median income is approximately normal. Give a 95% confidence interval for the 1999 median annual earnings of households.

6.24 The previous problem reports data on the median household income for the entire United States. In a detailed report based on the same sample survey, you find that the estimated median income for four-person families in Michigan is $65,467. Is the margin of error for this estimate with 95% confidence greater or less than the margin of error for the national median. Why?

6.25 As we prepare to take a sample and compute a 95% confidence interval, we know that the probability that the interval we compute will cover the parameter is 0.95. That’s the meaning of 95% confidence. If we use several such intervals, however, our confidence that all give correct results is less than 95%.

Suppose we are interested in confidence intervals for the median household incomes for three states. We compute a 95% interval for each of the three, based on independent samples in the three states.

(a) What is the probability that all three intervals cover the true median incomes? This probability (expressed as a percent) is our overall confidence level for the three simultaneous statements.

(b) What is the probability that at least two of the three intervals cover the true median incomes?

6.26 The Bowl Championship Series (BCS) was designed to select the top two teams in college football for a final championship game. The teams are selected by a complicated formula. In 2001, the University of Miami Hurricanes and the University
of Nebraska Cornhuskers played for the championship. However, many football fans thought that Nebraska should not have played in the game because it was rated only fourth in both major opinion polls. Third-ranked University of Colorado fans were particularly upset because Colorado soundly beat the Cornhuskers late in the season. A new CNN/USA Today/Gallup poll reports that a majority of fans would prefer a national championship play-off as an alternative to the BCS. The news media polled a random sample of 1019 adults 18 years of age or older. A summary of the results states that 54% prefer the play-off, and the margin of error is 3% for 95% confidence.

(a) Give the 95% confidence interval.

(b) Do you think that a newspaper headline stating that a majority of fans prefer a play-off is justified by the results of this study? Explain your answer.

6.27 An advertisement in the student newspaper asks you to consider working for a telemarketing company. The ad states, “Earn between $500 and $1000 per week.” Do you think that the ad is describing a confidence interval? Explain your answer.

6.28 A *New York Times* poll on women’s issues interviewed 1025 women and 472 men randomly selected from the United States, excluding Alaska and Hawaii. The poll found that 47% of the women said they do not get enough time for themselves.

(a) The poll announced a margin of error of ±3 percentage points for 95% confidence in conclusions about women. Explain to someone who knows no statistics why we can’t just say that 47% of all adult women do not get enough time for themselves.

(b) Then explain clearly what “95% confidence” means.

(c) The margin of error for results concerning men was ±4 percentage points. Why is this larger than the margin of error for women?

6.29 A radio talk show invites listeners to enter a dispute about a proposed pay increase for city council members. “What yearly pay do you think council members should get? Call us with your number.” In all, 958 people call. The mean pay they suggest is $9740 per year, and the standard deviation of the responses is $1125. For a large sample such as this, $s$ is very close to the unknown population $\sigma$. The station calculates the 95% confidence interval for the mean pay $\mu$ that all citizens would propose for council members to be $9669$ to $9811$. Is this result trustworthy? Explain your answer.

6.30 A study based on a sample of size 25 reported a mean of 76 with a margin of error of 12 for 95% confidence. Give the 95% confidence interval.

6.31 Refer to the previous exercise. If you wanted 99% confidence for the same study, would your margin of error be greater than, equal to, or less than 12? Explain your answer.

6.32 Suppose that the sample mean is 50 and the standard deviation is assumed to be 5. Make a diagram similar to Figure 6.5 (page xxx) that illustrates the effect of sample size on the width of a 95% interval. Use the following sample sizes: 10, 20, 40, and 100. Summarize what the diagram shows.
6.33 A study with 25 observations gave a mean of 70. Assume that the standard deviation is 15. Make a diagram similar to Figure 6.6 (page xxx) that illustrates the effect of the confidence level on the width of the interval. Use 80%, 90%, 95%, and 99%. Summarize what the diagram shows.

6.34 Consider the following two scenarios. (A) Take a simple random sample of 100 students from an elementary school with children in grades kindergarten through fifth grade; (B) Take a simple random sample of 100 third-graders from the same school. For each of these samples you will measure the height of each child in the sample. Which sample should have the smaller margin of error for 95% confidence? Explain your answer.

6.35 You want to rent an unfurnished one-bedroom apartment for next semester. The mean monthly rent for a random sample of 10 apartments advertised in the local newspaper is $580. Assume that the standard deviation is $90. Find a 95% confidence interval for the mean monthly rent for unfurnished one-bedroom apartments available for rent in this community.

6.36 A questionnaire about study habits was given to a random sample of students taking a large introductory statistics class. The sample of 25 students reported that they spent an average of 80 minutes per week studying statistics. Assume that the standard deviation is 35 minutes.
(a) Give a 95% confidence interval for the mean time spent studying statistics by students in this class.
(b) Is it true that 95% of the students in the class have weekly study times that lie in the interval you found in part (a)? Explain your answer.

6.37 Refer to the previous exercise.
(a) Give the mean and standard deviation in hours.
(b) Calculate the 95% confidence interval in hours from your answer to part (a).
(c) Explain how you could have directly calculated this interval from the 95% interval that you calculated in the previous exercise.

6.38 Computers in some vehicles calculate various quantities related to performance. One of these is the fuel efficiency, or gas mileage, usually expressed as miles per gallon (MPG). For one vehicle equipped in this way, the MPG were recorded each time the gas tank was filled, and the computer was then reset. Here are the MPG values for a random sample of 20 of these records:

<table>
<thead>
<tr>
<th>15.8</th>
<th>13.6</th>
<th>15.6</th>
<th>19.1</th>
<th>22.4</th>
<th>15.6</th>
<th>22.5</th>
<th>17.2</th>
<th>19.4</th>
<th>22.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.4</td>
<td>18.0</td>
<td>14.6</td>
<td>18.7</td>
<td>21.0</td>
<td>14.8</td>
<td>22.6</td>
<td>21.5</td>
<td>14.3</td>
<td>20.9</td>
</tr>
</tbody>
</table>

Suppose that the standard deviation of the population is known to be $\sigma = 2.9$ MPG.
(a) What is $\sigma_x$, the standard deviation of $\bar{x}$?
(b) Give a 95% confidence interval for $\mu$, the mean MPG for this vehicle.

6.39 Refer to the previous exercise. Here are the values of the average speed in miles per hour (MPH) for the same sample:

<table>
<thead>
<tr>
<th>21.0</th>
<th>19.0</th>
<th>18.7</th>
<th>39.2</th>
<th>45.8</th>
<th>19.8</th>
<th>48.4</th>
<th>21.0</th>
<th>29.1</th>
<th>35.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>31.6</td>
<td>49.0</td>
<td>16.0</td>
<td>34.6</td>
<td>36.3</td>
<td>19.0</td>
<td>43.3</td>
<td>37.5</td>
<td>16.5</td>
<td>34.5</td>
</tr>
</tbody>
</table>
Assume that the standard deviation is 10.3 MPH. Estimate the mean speed at which this vehicle was driven with a margin of error for 95% confidence.

6.40 Here are the Degree of Reading Power (DRP) scores for a sample of 44 third-grade students:

<table>
<thead>
<tr>
<th>40</th>
<th>26</th>
<th>39</th>
<th>14</th>
<th>42</th>
<th>18</th>
<th>25</th>
<th>43</th>
<th>46</th>
<th>27</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>19</td>
<td>26</td>
<td>35</td>
<td>34</td>
<td>15</td>
<td>44</td>
<td>40</td>
<td>38</td>
<td>31</td>
<td>46</td>
</tr>
<tr>
<td>52</td>
<td>25</td>
<td>35</td>
<td>35</td>
<td>33</td>
<td>29</td>
<td>34</td>
<td>41</td>
<td>49</td>
<td>28</td>
<td>52</td>
</tr>
<tr>
<td>47</td>
<td>35</td>
<td>48</td>
<td>22</td>
<td>33</td>
<td>41</td>
<td>51</td>
<td>27</td>
<td>14</td>
<td>54</td>
<td>45</td>
</tr>
</tbody>
</table>

Suppose that the standard deviation of the population of DRP scores is known to be $\sigma = 11$. Give a 95% confidence interval for the population mean score.

6.41 You are planning a survey of starting salaries for recent liberal arts major graduates from your college. From a pilot study you estimate that the standard deviation is about $9000. What sample size do you need to have a margin of error equal to $400 with 95% confidence?

6.42 Suppose that in the setting of the previous exercise you are willing to settle for a margin of error of $800. Will the required sample size be larger or smaller? Verify your answer by performing the calculations.

6.43 How large a sample of one-bedroom apartments in Exercise 6.10 would be needed to estimate the mean $\mu$ within $\pm$20 with 90% confidence?

6.44 A newspaper invites readers to send email stating whether or not they are in favor of making full-day kindergarten available to all students in the state. A total of 320 responses are received and, of these, 80% are in favor of the new program. In an article describing the results, the authors state that the margin of error is 4% for 95% confidence. Assume that they have computed this number correctly.
(a) Use the sample proportion and the margin of error to compute the 95% confidence interval.
(b) Do you think that these results are trustworthy? Discuss your answer.

6.45 A recent Gallup poll conducted telephone interviews with a random sample of adults aged 18 and older. Data were obtained for 1011 people. Of these, 37% said that football is their favorite sport to watch on television.
(a) The poll announced a margin of error of $\pm 3$ percentage points for 95% confidence. Explain to someone who knows no statistics why we can’t just say that 37% of Americans would say that football is their favorite sport to watch on television.
(b) Then explain clearly what “95% confidence” means.
(c) Give the 95% confidence interval.
(d) The poll was taken in December, an exciting part of the football season. Do you think that a similar poll conducted in June might produce different results? Explain why or why not.
Section 6.2

6.46 Each of the following situations requires a significance test about a population mean $\mu$. State the appropriate null hypothesis $H_0$ and alternative hypothesis $H_a$ in each case.

(a) The mean area of the several thousand apartments in a new development is advertised to be 1250 square feet. A tenant group thinks that the apartments are smaller than advertised. They hire an engineer to measure a sample of apartments to test their suspicion.

(b) Larry’s car averages 32 miles per gallon on the highway. He now switches to a new motor oil that is advertised as increasing gas mileage. After driving 3000 highway miles with the new oil, he wants to determine if his gas mileage actually has increased.

(c) The diameter of a spindle in a small motor is supposed to be 5 millimeters. If the spindle is either too small or too large, the motor will not perform properly. The manufacturer measures the diameter in a sample of motors to determine whether the mean diameter has moved away from the target.

6.47 In each of the following situations, a significance test for a population mean $\mu$ is called for. State the null hypothesis $H_0$ and the alternative hypothesis $H_a$ in each case.

(a) Experiments on learning in animals sometimes measure how long it takes a mouse to find its way through a maze. The mean time is 18 seconds for one particular maze. A researcher thinks that a loud noise will cause the mice to complete the maze faster. She measures how long each of 10 mice takes with a noise as stimulus.

(b) The examinations in a large accounting class are scaled after grading so that the mean score is 50. A self-confident teaching assistant thinks that his students have a higher mean score than the class as a whole. His students this semester can be considered a sample from the population of all students he might teach, so he compares their mean score with 50.

(c) A university gives credit in French language courses to students who pass a placement test. The language department wants to know if students who get credit in this way differ in their understanding of spoken French from students who actually take the French courses. Some faculty think the students who test out of the courses are better, but others argue that they are weaker in oral comprehension. Experience has shown that the mean score of students in the courses on a standard listening test is 24. The language department gives the same listening test to a sample of 40 students who passed the credit examination to see if their performance is different.

6.48 You have performed a two-sided test of significance and obtained a value of $z = 3.3$. Use Table D to find the approximate $P$-value for this test.

6.49 You have performed a one-sided test of significance and obtained a value of $z = 0.215$. Use Table D to find the approximate $P$-value for this test.

6.50 An understanding of cockroach biology may lead to an effective control strategy for these annoying insects. Researchers studying the absorption of sugar by insects feed cockroaches a diet containing measured amounts of a particular sugar. After 10
hours, the cockroaches are killed and the concentration of the sugar in various body parts is determined by a chemical analysis. The paper that reports the research states that a 95% confidence interval for the mean amount (in milligrams) of the sugar in the hindguts of the cockroaches is $4.2 \pm 2.3$. (From D. L. Shankland et al., “The effect of 5-thio-D-glucose on insect development and its absorption by insects,” *Journal of Insect Physiology*, 14 (1968), pp. 63–72.)

(a) Does this paper give evidence that the mean amount of sugar in the hindguts under these conditions is not equal to 7 mg? State $H_0$ and $H_a$ and base a test on the confidence interval.

(b) Would the hypothesis that $\mu = 5$ mg be rejected at the 5% level in favor of a two-sided alternative?

6.51 Market pioneers, companies that are among the first to develop a new product or service, tend to have higher market shares than latecomers to the market. What accounts for this advantage? Here is an excerpt from the conclusions of a study of a sample of 1209 manufacturers of industrial goods:

Can patent protection explain pioneer share advantages? Only 21% of the pioneers claim a significant benefit from either a product patent or a trade secret. Though their average share is two points higher than that of pioneers without this benefit, the increase is not statistically significant ($z = 1.13$). Thus, at least in mature industrial markets, product patents and trade secrets have little connection to pioneer share advantages.

Find the $P$-value for the given $z$. Then explain to someone who knows no statistics what “not statistically significant” in the study’s conclusion means. Why does the author conclude that patents and trade secrets don’t help, even though they contributed 2 percentage points to average market share? (From William T. Robinson, “Sources of market pioneer advantages: the case of industrial goods industries,” *Journal of Marketing Research*, 25 (1988), pp. 87–94.)

6.52 Each of the following situations requires a significance test about a population mean $\mu$. State the appropriate null hypothesis $H_0$ and alternative hypothesis $H_a$ in each case.

(a) A dual X-ray absorptiometry (DXA) scanner is used to measure bone mineral density for people who may be at risk for osteoporosis. To be sure that the measurements are accurate, an object called a “phantom” that has known mineral density $\mu = 1.4$ grams per square centimeter is measured. This phantom is scanned 10 times.

(b) Feedback from your customers shows that many think it takes too long to fill out the online order form for your products. You redesign the form and survey a random sample of customers to determine whether or not they think that the new form is actually an improvement. The response uses a five-point scale: $-2$ if the new form takes much less time than the old form; $-1$ if the new form takes a little less time; 0 if the new form takes about the same time; $+1$ if the new form takes a little more time; and $+2$ if the new form takes much more time.

(c) You purchase a shipment of 60-watt light bulbs to be used in a variety of your products. If the wattage is too low or too high, your product will not look good. You measure the wattage of a random sample of 20 bulbs.
6.53 In each of the following situations, a significance test for a population mean \( \mu \) is called for. State the null hypothesis \( H_0 \) and the alternative hypothesis \( H_a \) in each case.

(a) Experiments on learning in animals sometimes measure how long it takes a mouse to find its way through a maze. The mean time is 18 seconds for one particular maze. A researcher thinks that a loud noise will cause the mice to complete the maze faster. She measures how long each of 10 mice takes with a noise as stimulus.

(b) The examinations in a large history class are scaled after grading so that the mean score is 50. A self-confident teaching assistant thinks that his students have a higher mean score than the class as a whole. His students this semester can be considered a sample from the population of all students he might teach, so he compares their mean score with 50.

(c) The Census Bureau reports that households spend an average of 31% of their total spending on housing. A homebuilders association in Cleveland wonders if the national finding applies in their area. They interview a sample of 40 households in the Cleveland metropolitan area to learn what percent of their spending goes toward housing.

6.54 In each of the following situations, state an appropriate null hypothesis \( H_0 \) and alternative hypothesis \( H_a \). Be sure to identify the parameters that you use to state the hypotheses. (We have not yet learned how to test these hypotheses.)

(a) A sociologist asks a large sample of high school students which academic subject they like best. She suspects that a higher percent of males than of females will name mathematics as their favorite subject.

(b) An education researcher randomly divides sixth-grade students into two groups for physical education class. He teaches both groups basketball skills, using the same methods of instruction in both classes. He encourages Group A with compliments and other positive behavior but acts cool and neutral toward Group B. He hopes to show that positive teacher attitudes result in a higher mean score on a test of basketball skills than do neutral attitudes.

(c) An economist believes that among employed young adults there is a positive correlation between income and the percent of disposable income that is saved. To test this, she gathers income and savings data from a sample of employed persons in her city aged 25 to 34.

6.55 A test of the null hypothesis \( H_0: \mu = \mu_0 \) gives test statistic \( z = 1.8 \).

(a) What is the \( P \)-value if the alternative is \( H_a: \mu > \mu_0 \) ?

(b) What is the \( P \)-value if the alternative is \( H_a: \mu < \mu_0 \) ?

(c) What is the \( P \)-value if the alternative is \( H_a: \mu \neq \mu_0 \) ?

6.56 The \( P \)-value for a two-sided test of the null hypothesis \( H_0: \mu = 10 \) is 0.06.

(a) Does the 95% confidence interval include the value 10? Why?

(b) Does the 90% confidence interval include the value 10? Why?

6.57 A 95% confidence interval for a population mean is (28, 35).

(a) Can you reject the null hypothesis that \( \mu = 34 \) at the 5% significance level? Why?
(b) Can you reject the null hypothesis that \( \mu = 36 \) at the 5% significance level? Why?

6.58 A new supplier offers a good price on a catalyst used in your production process. You compare the purity of this catalyst with that from your current supplier. The \( P \)-value for a test of “no difference” is 0.15. Can you be confident that the purity of the new product is the same as the purity of the product that you have been using? Discuss.

6.59 We often see televised reports of brushfires threatening homes in California. Some people argue that the modern practice of quickly putting out small fires allows fuel to accumulate and so increases the damage done by large fires. A detailed study of historical data suggests that this is wrong—the damage has risen simply because there are more houses in risky areas. (Jon E. Keeley, C. J. Fotheringham, and Marco Morais, “Reexamining fire suppression impacts on brushland fire regimes,” *Science*, 284 (1999), pp. 1829–1831.) As usual, the study report gives statistical information tersely. Here is the summary of a regression of number of fires on decade (9 data points, for the 1910s to the 1990s):

Collectively, since 1910, there has been a highly significant increase \( (r^2 = 0.61, P < 0.01) \) in the number of fires per decade.

How would you explain this statement to someone who knows no statistics? Include an explanation of both the description given by \( r^2 \) and the statistical significance.

6.60 A randomized comparative experiment examined whether a calcium supplement in the diet reduces the blood pressure of healthy men. The subjects received either a calcium supplement or a placebo for 12 weeks. The statistical analysis was quite complex, but one conclusion was that “the calcium group had lower seated systolic blood pressure \( (P = 0.008) \) compared with the placebo group.” (R. M. Lyle et al., “Blood pressure and metabolic effects of calcium supplementation in normotensive white and black men,” *Journal of the American Medical Association*, 257 (1987), pp. 1772–1776.) Explain this conclusion, especially the \( P \)-value, as if you were speaking to a doctor who knows no statistics.

6.61 A social psychologist reports that “ethnocentrism was significantly higher \( (P < 0.05) \) among church attenders than among nonattenders.” Explain what this means in language understandable to someone who knows no statistics. Do not use the word “significance” in your answer.

6.62 A study examined the effect of exercise on how students perform on their final exam in statistics. The \( P \)-value was given as 0.87.

(a) State null and alternative hypotheses that could have been used for this study. (Note that there is more than one correct answer.)

(b) Do you reject the null hypothesis?

(c) What is your conclusion?

(d) What other facts about the study would you like to know for a proper interpretation of the results?
6.63 The financial aid office of a university asks a sample of students about their employment and earnings. The report says that “for academic year earnings, a significant difference \( (P = 0.038) \) was found between the sexes, with men earning more on the average. No difference \( (P = 0.476) \) was found between the earnings of black and white students.” (From a study by M. R. Schlatter et al., Division of Financial Aid, Purdue University.) Explain both of these conclusions, for the effects of sex and of race on mean earnings, in language understandable to someone who knows no statistics.

6.64 The mean yield of corn in the United States is about 120 bushels per acre. A survey of 40 farmers this year gives a sample mean yield of \( \bar{x} = 123.8 \) bushels per acre. We want to know whether this is good evidence that the national mean this year is not 120 bushels per acre. Assume that the farmers surveyed are an SRS from the population of all commercial corn growers and that the standard deviation of the yield in this population is \( \sigma = 10 \) bushels per acre. Give the \( P \)-value for the test of

\[
H_0: \mu = 120 \\
H_a: \mu \neq 120
\]

Are you convinced that the population mean is not 120 bushels per acre? Is your conclusion correct if the distribution of corn yields is somewhat nonnormal? Why?

6.65 In the past, the mean score of the seniors at South High on the American College Testing (ACT) college entrance examination has been 20. This year a special preparation course is offered, and all 53 seniors planning to take the ACT test enroll in the course. The mean of their 53 ACT scores is 22.1. The principal believes that the new course has improved the students’ ACT scores.

(a) Assume that ACT scores vary normally with standard deviation 6. Is the outcome \( \bar{x} = 22.1 \) good evidence that the population mean score is greater than 20? State \( H_0 \) and \( H_a \), compute the test statistic and the \( P \)-value, and answer the question by interpreting your result.

(b) The results are in any case inconclusive because of the design of the study. The effects of the new course are confounded with any change from past years, such as other new courses or higher standards. Briefly outline the design of a better study of the effect of the new course on ACT scores.

6.66 There are other \( z \) statistics that we have not yet studied. We can use Table D to assess the significance of any \( z \) statistic. A study compares the habits of students who are on academic probation with students whose grades are satisfactory. One variable measured is the hours spent watching television last week. The null hypothesis is “no difference” between the means for the two populations. The alternative hypothesis is two-sided. The value of the test statistic is \( z = -1.37 \).

(a) Is the result significant at the 5% level?
(b) Is the result significant at the 1% level?

6.67 You measure the weights of 24 male runners. These runners are not a random sample from a population, but you are willing to assume that their weights represent the weights of similar runners. Here are their weights in kilograms:
Exercise 6.13 asks you to find a 95% confidence interval for the mean weight of
the population of all such runners, assuming that the population standard deviation
is \( \sigma = 4.5 \) kg.
(a) Give the confidence interval from that exercise, or calculate the interval if you
did not do the exercise.
(b) Based on this confidence interval, does a test of
\[
H_0: \mu = 61.3 \text{ kg} \\
H_a: \mu \neq 61.3 \text{ kg}
\]
reject \( H_0 \) at the 5% significance level?
(c) Would \( H_0: \mu = 63 \) be rejected at the 5% level if tested against a two-sided
alternative?

6.68 An understanding of cockroach biology may lead to an effective control strategy
for these annoying insects. Researchers studying the absorption of sugar by insects
feed cockroaches a diet containing measured amounts of a particular sugar. After 10
hours, the cockroaches are killed and the concentration of the sugar in various body
parts is determined by a chemical analysis. The paper that reports the research
states that a 95% confidence interval for the mean amount (in milligrams) of the
sugar in the hindguts of the cockroaches is \( 4.2 \pm 2.3 \). (D. L. Shankland et al.,
“The effect of 5-thio-D-glucose on insect development and its absorption by insects,”
Journal of Insect Physiology, 14 (1968), pp. 63–72.) (a) Does this paper give evidence
that the mean amount of sugar in the hindguts under these conditions is not equal
to 7 mg? State \( H_0 \) and \( H_a \) and base a test on the confidence interval.
(b) Would the hypothesis that \( \mu = 5 \) mg be rejected at the 5% level in favor of a
two-sided alternative?

6.69 Market pioneers, companies that are among the first to develop a new product
or service, tend to have higher market shares than latecomers to the market. What
accounts for this advantage? Here is an excerpt from the conclusions of a study of
a sample of 1209 manufacturers of industrial goods:

Can patent protection explain pioneer share advantages? Only 21% of
the pioneers claim a significant benefit from either a product patent or a
trade secret. Though their average share is two points higher than that of
pioneers without this benefit, the increase is not statistically significant
\( (z = 1.13) \). Thus, at least in mature industrial markets, product patents
and trade secrets have little connection to pioneer share advantages.

(William T. Robinson, “Sources of market pioneer advantages: the case of industrial
goods industries,” Journal of Marketing Research, 25 (1988), pp. 87–94.) Find the
\( P \)-value for the given \( z \). Then explain to someone who knows no statistics what
“not statistically significant” in the study’s conclusion means. Why does the author
conclude that patents and trade secrets don’t help, even though they contributed 2 percentage points to average market share?

**6.70** An old farmer claims to be able to detect the presence of water with a forked stick. In a test of this claim, he is presented with 5 identical barrels, some containing water and some not. He is right in 4 of the 5 cases.

(a) Suppose the farmer has probability $p$ of being correct. If he is just guessing, $p = 0.5$. State an appropriate $H_0$ and $H_a$ in terms of $p$ for a test of whether he does better than guessing.

(b) If the farmer is simply guessing, what is the distribution of the number $X$ of correct answers in 5 tries?

(c) The observed outcome is $X = 4$. What is the $P$-value of the test that takes large values of $X$ to be evidence against $H_0$?

**6.71** Here are several situations where there is an incorrect application of the ideas presented in this section. Write a short paragraph explaining what is wrong in each situation and why it is wrong.

(a) A climatologist wants to test the null hypothesis that it will rain tomorrow.

(b) A random sample of size 20 is taken from a population that is assumed to have a standard deviation of 15. The standard deviation of the sample mean is $15/20$.

(c) A researcher tests the following null hypothesis: $H_0: \bar{x} = 10$.

**6.72** Here are several situations where there is an incorrect application of the ideas presented in this section. Write a short paragraph explaining what is wrong in each situation and why it is wrong.

(a) A change is made that should improve student satisfaction with the way grades are processed at your college. The null hypothesis, that there is an improvement, is tested versus the alternative, that there is no change.

(b) A significance test rejected the null hypothesis that the sample mean is 25.

(c) A report on a study says that the results are statistically significant and the $P$-value is 0.95.

**6.73** State the appropriate null hypothesis $H_0$ and alternative hypothesis $H_a$ in each of the following cases.

(a) An experiment is designed to examine the effect of a diet high in soy products on the bone density of adult rats.

(b) The student newspaper at your college recently changed the format for their news stories. You take a random sample of students and select those who regularly read the newspaper. These are asked to indicate their opinions on the changes using a five-point scale: $-2$ if the new format is much worse than the old, $-1$ if the new format is somewhat worse than the old, 0 if the new format is the same as the old, $+1$ if the new format is somewhat better than the old, and $+2$ if the new format is much better than the old.

(c) The examinations in a large history class are scaled after grading so that the mean score is 75. A self-confident teaching assistant thinks that his students have a higher mean score than the class as a whole. His students this semester can be considered a sample from the population of all students he might teach, so he compares their mean score with 75.
6.74 State the null hypothesis $H_0$ and the alternative hypothesis $H_a$ in each case. 
(a) A national study reports that households spend an average of 30% of their food expenditures in restaurants. A restaurant association in your area wonders if the national finding applies locally. They interview a sample of 40 households and ask about their total food budget and the amount spent in restaurants.
(b) Experiments on learning in animals sometimes measure how long it takes a mouse to find its way through a maze. The mean time is 20 seconds for one particular maze. A researcher thinks that playing rap music will cause the mice to complete the maze faster. She measures how long each of 12 mice takes with the rap music as a stimulus.
(c) A dual X-ray absorptiometry (DXA) scanner is used to measure bone mineral density for people who may be at risk for osteoporosis. To be sure that the measurements are accurate, an object called a “phantom” that has known mineral density $\mu = 1.3$ grams per square centimeter is measured. This phantom is scanned 8 times.

6.75 In each of the following situations, state an appropriate null hypothesis $H_0$ and alternative hypothesis $H_a$. Be sure to identify the parameters that you use to state the hypotheses. (We have not yet learned how to test these hypotheses.)
(a) An education researcher randomly divides sixth-grade students into two groups for physical education class. He teaches both groups volleyball skills, using the same methods of instruction in both classes. He encourages Group A with compliments and other positive behavior but acts cool and neutral toward Group B. He hopes to show that positive teacher attitudes result in a higher mean score on a test of volleyball skills than do neutral attitudes.
(b) An education researcher believes that among college students there is a positive correlation between grade point average and self-esteem. To test this, she gathers grade point average and self-esteem data from a sample of students at your college.
(c) A sociologist asks a large sample of high school students which academic subject they like best. She suspects that a higher percent of females than of males will name English as their favorite subject.

6.76 Translate each of the following research questions into appropriate $H_0$ and $H_a$.
(a) Census Bureau data show that the mean household income in the area served by a shopping mall is $72,500 per year. A market research firm questions shoppers at the mall to find out whether the mean household income of mall shoppers is higher than that of the general population.
(b) Last year, your company’s service technicians took an average of 1.8 hours to respond to trouble calls from business customers who had purchased service contracts. Do this year’s data show a different average response time?

6.77 A test statistic for a two-sided significance test for a population mean is $z = 2.3$. Sketch a standard normal curve and mark this value of $z$ on it. Find the $P$-value and shade the appropriate areas under the curve to illustrate your calculations.

6.78 A test statistic for a two-sided significance test for a population mean is $z = -1.4$. Sketch a standard normal curve and mark this value of $z$ on it. Find the $P$-value and shade the appropriate areas under the curve to illustrate your calculations.
6.79 The $P$-value for a significance test is 0.082.
(a) Do you reject the null hypothesis at level $\alpha = 0.05$?
(b) Do you reject the null hypothesis at level $\alpha = 0.01$?
(c) Explain your answers.

6.80 The $P$-value for a significance test is 0.032.
(a) Do you reject the null hypothesis at level $\alpha = 0.05$?
(b) Do you reject the null hypothesis at level $\alpha = 0.01$?
(c) Explain your answers.

6.81 A test of the null hypothesis $H_0: \mu = \mu_0$ gives test statistic $z = -1.6$.
(a) What is the $P$-value if the alternative is $H_a: \mu > \mu_0$?
(b) What is the $P$-value if the alternative is $H_a: \mu < \mu_0$?
(c) What is the $P$-value if the alternative is $H_a: \mu \neq \mu_0$?

6.82 The $P$-value for a two-sided test of the null hypothesis $H_0: \mu = 30$ is 0.09.
(a) Does the 95% confidence interval include the value 30? Why?
(b) Does the 90% confidence interval include the value 30? Why?

6.83 The $P$-value for a two-sided test of the null hypothesis $H_0: \mu = 30$ is 0.04.
(a) Does the 95% confidence interval include the value 30? Why?
(b) Does the 90% confidence interval include the value 30? Why?

6.84 A 95% confidence interval for a population mean is (57, 65).
(a) Can you reject the null hypothesis that $\mu = 68$ at the 5% significance level? Why?
(b) Can you reject the null hypothesis that $\mu = 62$ at the 5% significance level? Why?

6.85 A 90% confidence interval for a population mean is (12, 15).
(a) Can you reject the null hypothesis that $\mu = 13$ at the 10% significance level? Why?
(b) Can you reject the null hypothesis that $\mu = 10$ at the 10% significance level? Why?

6.86 A report based on the National Assessment of Educational Progress (NAEP)\textsuperscript{13} states that the average score on their mathematics test for eighth-grade students in the District of Columbia was 243 in 2003, which was 235. The report then says that this value is higher than the average in 2000. A footnote states that comparisons (higher/lower/different) are determined by statistical tests with 0.05 as the level of significance. Explain what this means in language understandable to someone who knows no statistics. Do not use the word “significance” in your answer.

6.87 An NAEP report\textsuperscript{14} similar to the one described in the previous exercise states that the average score on their mathematics test for eighth-grade students in Boston was 262. The report then says that this value was not significantly different from 287, the average score for eighth-grade students in U.S. public schools that are located in large central cities. A footnote states that comparisons (higher/lower/different) are determined by statistical tests with 0.05 as the level of significance. Explain what
this means in language understandable to someone who knows no statistics. Do not use the word “significance” in your answer.

6.88 Here are the Degree of Reading Power (DRP) scores for a sample of 44 third-grade students:

| 40 | 26 | 39 | 14 | 42 | 18 | 25 | 43 | 46 | 27 | 19 |
| 47 | 19 | 26 | 35 | 34 | 15 | 44 | 40 | 38 | 31 | 46 |
| 52 | 25 | 35 | 35 | 33 | 29 | 34 | 41 | 49 | 28 | 52 |
| 47 | 35 | 48 | 22 | 33 | 41 | 51 | 27 | 14 | 54 | 45 |

These students can be considered to be an SRS of the third-graders in a suburban school district. DRP scores are approximately normal. Suppose that the standard deviation of scores in this school district is known to be $\sigma = 11$. The researcher believes that the mean score $\mu$ of all third-graders in this district is higher than the national mean, which is 32.

(a) State the appropriate $H_0$ and $H_a$ to test this suspicion.
(b) Carry out the test. Give the $P$-value, and then interpret the result in plain language.

6.89 To determine whether the mean nicotine content of a brand of cigarettes is greater than the advertised value of 1.4 milligrams, a health advocacy group tests

$$H_0: \mu = 1.4$$
$$H_a: \mu > 1.4$$

The calculated value of the test statistic is $z = 1.75$.

(a) Is the result significant at the 5% level?
(b) Is the result significant at the 1% level?

6.90 A computer has a random number generator designed to produce random numbers that are uniformly distributed on the interval from 0 to 1. If this is true, the numbers generated come from a population with $\mu = 0.5$ and $\sigma = 0.2887$. A command to generate 100 random numbers gives outcomes with mean $\bar{x} = 0.4365$. Assume that the population $\sigma$ remains fixed. We want to test

$$H_0: \mu = 0.5$$
$$H_a: \mu \neq 0.5$$

(a) Calculate the value $z$ of the $z$ test statistic.
(b) Is the result significant at the 5% level ($\alpha = 0.05$)?
(c) Is the result significant at the 1% level ($\alpha = 0.01$)?

6.91 Consider a significance test for a null hypothesis versus a two-sided alternative with a $z$ test statistic. Give a value of $z$ that will give a result significant at the 1% level but not at the 0.5% level.

6.92 You have performed a two-sided test of significance and obtained a value of $z = -4.3$. Use Table D to find the approximate $P$-value for this test.
6.93 You have performed a one-sided test of significance and obtained a value of \( z = 0.22 \). Use Table D to find the approximate \( P \)-value for this test.

6.94 You will perform a significance test of

\[
H_0: \mu = 0
\]

versus

\[
H_a: \mu > 0
\]

(a) What values of \( z \) would lead you to reject \( H_0 \) at the 5% level?

(b) If the alternative hypothesis was

\[
H_a: \mu \neq 0
\]

what values of \( z \) would lead you to reject \( H_0 \) at the 5% level?

(c) Explain why your answers to parts (a) and (b) are different.

6.95 Consider a significance test for a null hypothesis versus a two-sided alternative. Between what values from Table D does the \( P \)-value for an outcome \( z = 1.34 \) lie? Calculate the \( P \)-value using Table A, and verify that it lies between the values you found from Table D.

6.96 Refer to the previous exercise. Find the \( P \)-value for \( z = -1.34 \).

6.97 Radon is a colorless, odorless gas that is naturally released by rocks and soils and may concentrate in tightly closed houses. Because radon is slightly radioactive, there is some concern that it may be a health hazard. Radon detectors are sold to homeowners worried about this risk, but the detectors may be inaccurate. University researchers placed 12 detectors in a chamber where they were exposed to 105 picocuries per liter (pCi/l) of radon over 3 days. Here are the readings given by the detectors:

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>91.9</td>
<td>97.8</td>
<td>111.4</td>
<td>122.3</td>
<td>105.4</td>
<td>95.0</td>
</tr>
<tr>
<td>103.8</td>
<td>99.6</td>
<td>96.6</td>
<td>119.3</td>
<td>104.8</td>
<td>101.7</td>
</tr>
</tbody>
</table>

Assume (unrealistically) that you know that the standard deviation of readings for all detectors of this type is \( \sigma = 9 \).

(a) Give a 95% confidence interval for the mean reading \( \mu \) for this type of detector.

(b) Is there significant evidence at the 5% level that the mean reading differs from the true value 105? State hypotheses and base a test on your confidence interval from (a).

Section 6.3

6.98 Give an example of a set of data for which statistical inference is not valid.

6.99 More than 200,000 people worldwide take the GMAT examination each year as they apply for MBA programs. Their scores vary normally with mean about \( \mu = 525 \).
and standard deviation about \( \sigma = 100 \). One hundred students go through a rigorous training program designed to raise their GMAT scores. Test the hypotheses

\[
H_0 : \mu = 525 \\
H_a : \mu > 525
\]

in each of the following situations:

(a) The students’ average score is \( \bar{x} = 541.4 \). Is this result significant at the 5% level?
(b) The average score is \( \bar{x} = 541.5 \). Is this result significant at the 5% level?

The difference between the two outcomes in (a) and (b) is of no importance. Beware attempts to treat \( \alpha = 0.05 \) as sacred.

6.100 How much education children get is strongly associated with the wealth and social status of their parents. In social science jargon, this is “socioeconomic status,” or SES. But the SES of parents has little influence on whether children who have graduated from college go on to yet more education. One study looked at whether college graduates took the graduate admissions tests for business, law, and other graduate programs. The effects of the parents’ SES on taking the LSAT test for law school were “both statistically insignificant and small.”

(a) What does “statistically insignificant” mean?
(b) Why is it important that the effects were small in size as well as insignificant?

6.101 A local television station announces a question for a call-in opinion poll on the six o’clock news and then gives the response on the eleven o’clock news. Today’s question concerns a proposed increase in funds for student loans. Of the 2372 calls received, 1921 oppose the increase. The station, following standard statistical practice, makes a confidence statement: “81% of the Channel 13 Pulse Poll sample oppose the increase. We can be 95% confident that the proportion of all viewers who oppose the increase is within 1.6% of the sample result.” Is the station’s conclusion justified? Explain your answer.

6.102 A researcher looking for evidence of extrasensory perception (ESP) tests 500 subjects. Four of these subjects do significantly better (\( P < 0.01 \)) than random guessing.

(a) Is it proper to conclude that these four people have ESP? Explain your answer.
(b) What should the researcher now do to test whether any of these four subjects have ESP?

6.103 The text cites an example in which researchers carried out 77 separate significance tests, of which 2 were significant at the 5% level. Suppose that these tests are independent of each other. (In fact they were not independent, because all involved the same subjects.) If all of the null hypotheses are true, each test has probability 0.05 of being significant at the 5% level.

(a) What is the distribution of the number \( X \) of tests that are significant?
(b) Find the probability that 2 or more of the tests are significant.

6.104 You are the statistical expert on a team that is planning a study. After you have made a careful presentation of the mechanics of significance testing, one of the
Section 6.3

team members suggests using $\alpha = 0.50$ for the study because you would be more likely to obtain statistically significant results with this choice. Explain in simple terms why this would not be a good use of statistical methods.

6.105 A study with 5000 subjects reported a result that was statistically significant at the 5% level. Explain why this result might not be particularly large or important.

6.106 A study with 12 subjects reported a result that failed to achieve statistical significance at the 5% level. The $P$-value was 0.052. Write a short summary of how you would interpret these findings.

6.107 A $P$-value of 0.95 is reported for a significance test for a population mean. Interpret this result.

6.108 Every user of statistics should understand the distinction between statistical significance and practical importance. A sufficiently large sample will declare very small effects statistically significant. Let us suppose that SAT mathematics (SATM) scores in the absence of coaching vary normally with mean $\mu = 480$ and $\sigma = 100$. Suppose further that coaching may change $\mu$ but does not change $\sigma$. An increase in the SATM score from 480 to 483 is of no importance in seeking admission to college, but this unimportant change can be statistically very significant. To see this, calculate the $P$-value for the test of

$$H_0: \mu = 480$$

$$H_a: \mu > 480$$

in each of the following situations:
(a) A coaching service coaches 100 students; their SATM scores average $x = 483$.
(b) By the next year, the service has coached 1000 students; their SATM scores average $x = 483$.
(c) An advertising campaign brings the number of students coached to 10,000; their average score is still $x = 483$.

6.109 Give a 99% confidence interval for the mean SATM score $\mu$ after coaching in each part of the previous exercise. For large samples, the confidence interval says, “Yes, the mean score is higher after coaching, but only by a small amount.”

6.110 As in the previous exercises, suppose that SATM scores vary normally with $\sigma = 100$. One hundred students go through a rigorous training program designed to raise their SATM scores by improving their mathematics skills. Carry out a test of

$$H_0: \mu = 480$$

$$H_a: \mu > 480$$

in each of the following situations:
(a) The students’ average score is $x = 496.4$. Is this result significant at the 5% level?
(b) The average score is $x = 496.5$. Is this result significant at the 5% level?
The difference between the two outcomes in (a) and (b) is of no importance. Beware attempts to treat \(\alpha = 0.05\) as sacred.

6.111 Refer to the previous problem. A researcher has performed 10 tests of significance and wants to apply the Bonferroni procedure with \(\alpha = 0.05\). The calculated \(P\)-values are 0.045, 0.888, 0.050, 0.004, 0.001, 0.041, 0.888, 0.010, 0.002, 0.223. Which of the null hypotheses are rejected with this procedure?

6.112 Table 1.5 (page xx) gives average property damage per year due to tornadoes for each of the states. Is it appropriate to use the statistical methods we discussed in this chapter for these data? Explain why or why not.

6.113 Give an example of an interesting set of data for which statistical inference is valid. Explain your answer.

Section 6.4

6.114 A previous example gives a test of a hypothesis about the SAT scores of California high school students based on an SRS of 500 students. The hypotheses are

\[ H_0: \mu = 450 \]
\[ H_a: \mu > 450 \]

Assume that the population standard deviation is \(\sigma = 100\). The test rejects \(H_0\) at the 1% level of significance when \(z \geq 2.326\), where

\[ z = \frac{\bar{x} - 450}{100/\sqrt{500}} \]

Is this test sufficiently sensitive to usually detect an increase of 10 points in the population mean SAT score? Answer this question by calculating the power of the test against the alternative \(\mu = 460\).

6.115 Use the result of the previous exercise to give the probability of a Type I error and the probability of a Type II error for the test in that exercise when the alternative is \(\mu = 462\).

6.116 A previous example discusses a test about the mean contents of cola bottles. The hypotheses are

\[ H_0: \mu = 300 \]
\[ H_a: \mu < 300 \]

The sample size is \(n = 6\), and the population is assumed to have a normal distribution with \(\sigma = 3\). A 5% significance test rejects \(H_0\) if \(z \leq -1.645\), where the test statistic \(z\) is

\[ z = \frac{\bar{x} - 300}{3/\sqrt{6}} \]

Power calculations help us see how large a shortfall in the bottle contents the test can be expected to detect.
(a) Find the power of this test against the alternative \( \mu = 299 \).
(b) Find the power against the alternative \( \mu = 295 \).
(c) Is the power against \( \mu = 290 \) higher or lower than the value you found in (b)?
Explain why this result makes sense.

6.117 You have an SRS of size \( n = 9 \) from a normal distribution with \( \sigma = 1 \). You wish to test
\[
H_0: \mu = 0 \\
H_a: \mu > 0
\]
You decide to reject \( H_0 \) if \( \bar{x} > 0 \) and to accept \( H_0 \) otherwise.
(a) Find the probability of a Type I error, that is, the probability that your test rejects \( H_0 \) when in fact \( \mu = 0 \).
(b) Find the probability of a Type II error when \( \mu = 0.3 \). This is the probability that your test accepts \( H_0 \) when in fact \( \mu = 0.3 \).
(c) Find the probability of a Type II error when \( \mu = 1 \).

6.118 Use the result of previous exercise to give the probabilities of Type I and Type II errors for the test discussed there. Take the alternative hypothesis to be \( \mu = 294 \).

6.119 (Optional) An acceptance sampling test has probability 0.05 of rejecting a good lot of bearings and probability 0.08 of accepting a bad lot. The consumer of the bearings may imagine that acceptance sampling guarantees that most accepted lots are good. Alas, it is not so. Suppose that 90% of all lots shipped by the producer are bad.
(a) Draw a tree diagram for shipping a lot (the branches are “bad” and “good”) and then inspecting it (the branches at this stage are “accept” and “reject”).
(b) Write the appropriate probabilities on the branches, and find the probability that a lot shipped is accepted.
(c) Use the definition of conditional probability or Bayes’s formula to find the probability that a lot is bad, given that the lot is accepted. This is the proportion of bad lots among the lots that the sampling plan accepts.

6.120 You want to see if a redesign of the cover of a mail-order catalog will increase sales. A very large number of customers will receive the original catalog, and a random sample of customers will receive the one with the new cover. For planning purposes, you are willing to assume that the sales from the new catalog will be approximately normal with \( \sigma = 60 \) dollars and that the mean for the original catalog will be \( \mu = 40 \) dollars. You decide to use a sample size of \( n = 1000 \). You wish to test
\[
H_0: \mu = 40 \\
H_a: \mu > 40
\]
You decide to reject \( H_0 \) if \( \bar{x} > 43.12 \) and to accept \( H_0 \) otherwise.
(a) Find the probability of a Type I error, that is, the probability that your test rejects \( H_0 \) when in fact \( \mu = 40 \) dollars.
(b) Find the probability of a Type II error when \( \mu = 45 \) dollars. This is the
probability that your test accepts $H_0$ when in fact $\mu = 45$.

(c) Find the probability of a Type II error when $\mu = 50$.

(d) The distribution of sales is not normal, because many customers buy nothing. Why is it nonetheless reasonable in this circumstance to assume that the mean will be approximately normal?

6.121 Example 6.15 (page xxx) gives a test of a hypothesis about the SAT scores of California high school students based on an SRS of 500 students. The hypotheses are

$$H_0: \mu = 450$$
$$H_a: \mu > 450$$

Assume that the population standard deviation is $\sigma = 100$. The test rejects $H_0$ at the 1% level of significance when $z \geq 2.326$, where

$$z = \frac{\bar{x} - 450}{100/\sqrt{500}}$$

Is this test sufficiently sensitive to usually detect an increase of 12 points in the population mean SAT score? Answer this question by calculating the power of the test against the alternative $\mu = 462$.

Chapter 6 Review Exercises

6.122 A study compares two groups of mothers with young children who were on welfare two years ago. One group attended a voluntary training program offered free of charge at a local vocational school and advertised in the local news media. The other group did not choose to attend the training program. The study finds a significant difference ($P < 0.01$) between the proportions of the mothers in the two groups who are still on welfare. The difference is not only significant but quite large. The report says that with 95% confidence the percent of the nonattending group still on welfare is 21% ± 4% higher than that of the group who attended the program. You are on the staff of a member of Congress who is interested in the plight of welfare mothers and who asks you about the report.

(a) Explain briefly and in nontechnical language what “a significant difference ($P < 0.01$)” means.

(b) Explain clearly and briefly what “95% confidence” means.

(c) Is this study good evidence that requiring job training of all welfare mothers would greatly reduce the percent who remain on welfare for several years?

6.123 Use a computer to generate $n = 5$ observations from a normal distribution with mean 20 and standard deviation 5—$N(20, 5)$. Find the 95% confidence interval for $\mu$. Repeat this process 100 times and then count the number of times that the confidence interval includes the value $\mu = 20$. Explain your results.

6.124 Use a computer to generate $n = 5$ observations from a normal distribution with mean 20 and standard deviation 5—$N(20, 5)$. Test the null hypothesis that $\mu = 20$ using a two-sided significance test. Repeat this process 100 times and then count the number of times that you reject $H_0$. Explain your results.
6.125 Use the same procedure for generating data as in the previous exercise. Now test the null hypothesis that $\mu = 22.5$. Explain your results.

6.126 Figure 6.2 demonstrates the behavior of a confidence interval in repeated sampling by showing the results of 25 samples from the same population. Now you will do a similar demonstration. Suppose that (unknown to the researcher) the mean SAT-M score of all California high school seniors is $\mu = 460$, and that the standard deviation is known to be $\sigma = 100$. The scores vary normally.

(a) Simulate the drawing of 25 SRSs of size $n = 100$ from this population.
(b) The 95% confidence interval for the population mean $\mu$ has the form $\bar{x} \pm m$. What is the margin of error $m$? (Remember that we know $\sigma = 100$.)
(c) Use your software to calculate the 95% confidence interval for $\mu$ when $\sigma = 100$ for each of your 25 samples. Verify the computer's calculations by checking the interval given for the first sample against your result in (b). Use the $\bar{x}$ reported by the software.
(d) How many of the 25 confidence intervals contain the true mean $\mu = 460$? If you repeated the simulation, would you expect exactly the same number of intervals to contain $\mu$? In a very large number of samples, what percent of the confidence intervals would contain $\mu$?

6.127 In the previous exercise you simulated the SAT-M scores of 25 SRSs of 100 California seniors. Now use these samples to demonstrate the behavior of a significance test. We know that the population of all SAT-M scores is normal with standard deviation $\sigma = 100$.

(a) Use your software to carry out a test of

$$H_0: \mu = 460$$
$$H_a: \mu \neq 460$$

for each of the 25 samples.
(b) Verify the computer's calculations by using Table A to find the $P$-value of the test for the first of your samples. Use the $\bar{x}$ reported by your software.
(c) How many of your 25 tests reject the null hypothesis at the $\alpha = 0.05$ significance level? (That is, how many have $P$-value 0.05 or smaller?) Because the simulation was done with $\mu = 460$, samples that lead to rejecting $H_0$ produce the wrong conclusion. In a very large number of samples, what percent would falsely reject the hypothesis?

6.128 Suppose that in fact the mean SAT-M score of California high school seniors is $\mu = 480$. Would the test in the previous exercise usually detect a mean this far from the hypothesized value? This is a question about the power of the test.

(a) Simulate the drawing of 25 SRSs from a normal population with mean $\mu = 480$ and $\sigma = 100$. These represent the results of sampling when in fact the alternative $\mu = 480$ is true.
(b) Repeat on these new data the test of

$$H_0: \mu = 460$$
$$H_a: \mu \neq 460$$
that you did in the previous exercise. How many of the 25 tests have \( P \)-values 0.05 or smaller? These tests reject the null hypothesis at the \( \alpha = 0.05 \) significance level, which is the correct conclusion.

(c) The power of the test against the alternative \( \mu = 480 \) is the probability that the test will reject \( H_0 : \mu = 460 \) when in fact \( \mu = 480 \). Calculate this power. In a very large number of samples from a population with mean 480, what percent would reject \( H_0 \)?

6.129 In a study of possible iron deficiency in infants, researchers compared several groups of infants who were following different feeding patterns. One group of 26 infants was being breast-fed. At 6 months of age, these children had a mean hemoglobin level of \( \bar{x} = 12.9 \) grams per 100 milliliters of blood and a standard deviation of 1.6. Taking the standard deviation to be the population value \( \sigma \), give a 95% confidence interval for the mean hemoglobin level of breast-fed infants. What assumptions are required for the validity of the method you used to get the confidence interval?

6.130 Statisticians prefer large samples. Describe briefly the effect of increasing the size of a sample (or the number of subjects in an experiment) on each of the following:
(a) The width of a level \( C \) confidence interval.
(b) The \( P \)-value of a test, when \( H_0 \) is false and all facts about the population remain unchanged as \( n \) increases.
(c) The power of a fixed level \( \alpha \) test, when \( \alpha \), the alternative hypothesis, and all facts about the population remain unchanged.

6.131 A roulette wheel has 18 red slots among its 38 slots. You observe many spins and record the number of times that red occurs. Now you want to use these data to test whether the probability of a red has the value that is correct for a fair roulette wheel. State the hypotheses \( H_0 \) and \( H_a \) that you will test. (The test for this situation is discussed in Chapter 8.)

6.132 The text demonstrates the behavior of a confidence interval in repeated sampling by showing the results of 25 samples from the same population. Now you will do a similar demonstration. Suppose that (unknown to the researcher) the mean SATM score of all California high school seniors is \( \mu = 475 \), and that the standard deviation is known to be \( \sigma = 100 \). The scores vary normally.
(a) Simulate the drawing of 50 SRSs of size \( n = 100 \) from this population.
(b) The 95% confidence interval for the population mean \( \mu \) has the form \( \bar{x} \pm m \). What is the margin of error \( m \)? (Remember that we know \( \sigma = 100 \).)
(c) Use your software to calculate the 95% confidence interval for \( \mu \) when \( \sigma = 100 \) for each of your 50 samples. Verify the computer’s calculations by checking the interval given for the first sample against your result in (b). Use the \( \bar{x} \) reported by the software.
(d) How many of the 50 confidence intervals contain the true mean \( \mu = 475 \)? If you repeated the simulation, would you expect exactly the same number of intervals to contain \( \mu \)? In a very large number of samples, what percent of the confidence intervals would contain \( \mu \)?
6.133 In the previous exercise you simulated the SATM scores of 50 SRSs of 100 California seniors. Now use these samples to demonstrate the behavior of a significance test. We know that the population of all SATM scores is normal with standard deviation $\sigma = 100$.

(a) Use your software to carry out a test of

$$H_0: \mu = 475$$

$$H_a: \mu \neq 475$$

for each of the 50 samples.

(b) Verify the computer’s calculations by using Table A to find the $P$-value of the test for the first of your samples. Use the $\bar{x}$ reported by your software.

(c) How many of your 50 tests reject the null hypothesis at the $\alpha = 0.05$ significance level? (That is, how many have $P$-value 0.05 or smaller?) Because the simulation was done with $\mu = 475$, samples that lead to rejecting $H_0$ produce the wrong conclusion. In a very large number of samples, what percent would falsely reject the hypothesis?

6.134 Suppose that in fact the mean SATM score of California high school seniors is $\mu = 500$. Would the test in the previous exercise usually detect a mean this far from the hypothesized value? This is a question about the power of the test.

(a) Simulate the drawing of 50 SRSs from a normal population with mean $\mu = 500$ and $\sigma = 100$. These represent the results of sampling when in fact the alternative $\mu = 500$ is true.

(b) Repeat on these new data the test of

$$H_0: \mu = 475$$

$$H_a: \mu \neq 475$$

that you did in the previous exercise. How many of the 50 tests have $P$-values 0.05 or smaller? These tests reject the null hypothesis at the $\alpha = 0.05$ significance level, which is the correct conclusion.

(c) The power of the test against the alternative $\mu = 500$ is the probability that the test will reject $H_0 : \mu = 475$ when in fact $\mu = 500$. Calculate this power. In a very large number of samples from a population with mean 500, what percent would reject $H_0$?

6.135 You are testing the null hypothesis that $\mu = 0$ versus the alternative $\mu > 0$ using $\alpha = 0.05$. Assume $\sigma = 17$. Suppose $\bar{x} = 5$ and $n = 10$. Calculate the test statistic and its $P$-value. Repeat assuming the same value of $\bar{x}$ but with $n = 20$. Do the same for sample sizes of 30, 40, and 50. Plot the values of the test statistic versus the sample size. Do the same for the $P$-values. Summarize what this demonstration shows about the effect of the sample size on significance testing.

6.136 An agronomist examines the cellulose content of a variety of alfalfa hay. Suppose that the cellulose content in the population has standard deviation $\sigma = 8$ milligrams per gram (mg/g). A sample of 16 cuttings has mean cellulose content $\bar{x} = 140$ mg/g.
(a) Give a 95% confidence interval for the mean cellulose content in the population.
(b) A previous study claimed that the mean cellulose content was $\mu = 135$ mg/g, but the agronomist believes that the mean is higher than that figure. State $H_0$ and $H_a$ and carry out a significance test to see if the new data support this belief.
(c) The statistical procedures used in (a) and (b) are valid when several assumptions are met. What are these assumptions?

6.137 Because sulfur compounds cause “off-odors” in wine, oenologists (wine experts) have determined the odor threshold, the lowest concentration of a compound that the human nose can detect. For example, the odor threshold for dimethyl sulfide (DMS) is given in the oenology literature as 25 micrograms per liter of wine ($\mu g/l$). Untrained noses may be less sensitive, however. Here are the DMS odor thresholds for 10 beginning students of oenology:

\[32 \quad 33 \quad 40 \quad 35 \quad 24 \quad 36 \quad 31 \quad 30 \quad 20 \quad 25\]

Assume (this is not realistic) that the standard deviation of the odor threshold for untrained noses is known to be $\sigma = 7 \mu g/l$.
(a) Make a stemplot to verify that the distribution is roughly symmetric with no outliers. (A normal quantile plot confirms that there are no systematic departures from normality.)
(b) Give a 95% confidence interval for the mean DMS odor threshold among all beginning oenology students.
(c) Are you convinced that the mean odor threshold for beginning students is higher than the published threshold, 25 $\mu g/l$? Carry out a significance test to justify your answer.

6.138 A study of the pay of corporate chief executive officers (CEOs) examined the increase in cash compensation of the CEOs of 104 companies, adjusted for inflation, in a recent year. The mean increase in real compensation was $\overline{x} = 6.8\%$, and the standard deviation of the increases was $s = 53\%$. Is this good evidence that the mean real compensation $\mu$ of all CEOs increased that year? The hypotheses are

\[H_0: \mu = 0 \quad \text{(no increase)}\]
\[H_a: \mu > 0 \quad \text{(an increase)}\]

Because the sample size is large, the sample $s$ is close to the population $\sigma$, so take $\sigma = 53\%$.
(a) Sketch the normal curve for the sampling distribution of $\overline{x}$ when $H_0$ is true. Shade the area that represents the $P$-value for the observed outcome $\overline{x} = 6.8\%$.
(b) Calculate the $P$-value.
(c) Is the result significant at the $\alpha = 0.05$ level? Do you think the study gives strong evidence that the mean compensation of all CEOs went up?

6.139 When asked to explain the meaning of “statistically significant at the $\alpha = 0.05$ level,” a student says, “This means there is only probability 0.05 that the null hypothesis is true.” Is this an essentially correct explanation of statistical significance? Explain your answer.
6.140 Use a computer to generate \( n = 12 \) observations from a normal distribution with mean 20 and standard deviation 5: \( N(20, 5) \). Find the 95% confidence interval for \( \mu \). Repeat this process 100 times and then count the number of times that the confidence interval includes the value \( \mu = 20 \). Explain your results.

6.141 Use a computer to generate \( n = 12 \) observations from a normal distribution with mean 20 and standard deviation 5: \( N(20, 5) \). Test the null hypothesis that \( \mu = 20 \) using a two-sided significance test. Repeat this process 100 times and then count the number of times that you reject \( H_0 \). Explain your results.

6.142 Use the same procedure for generating data as in the previous exercise. Now test the null hypothesis that \( \mu = 18 \). Explain your results.
CHAPTER 7

Section 7.1

7.1 What critical value $t^*$ from Table D should be used for a confidence interval for the mean of the population in each of the following situations?
(a) A 90% confidence interval based on $n = 12$ observations.
(b) A 95% confidence interval from an SRS of 30 observations.
(c) An 80% confidence interval from a sample of size 18.

7.2 Use software to find the critical values $t^*$ that you would use for each of the following confidence intervals for the mean.
(a) A 99% confidence interval based on $n = 55$ observations.
(b) A 90% confidence interval from an SRS of 35 observations.
(c) An 95% confidence interval from a sample of size 90.

7.3 The one-sample $t$ statistic for testing

$H_0$: $\mu = 0$

$H_a$: $\mu > 0$

from a sample of $n = 15$ observations has the value $t = 1.97$.
(a) What are the degrees of freedom for this statistic?
(b) Give the two critical values $t^*$ from Table D that bracket $t$.
(c) What are the right-tail probabilities $p$ for these two entries?
(d) Between what two values does the $P$-value of the test fall?
(e) Is the value $t = 1.97$ significant at the 5% level? Is it significant at the 1% level?
(f) If you have software available, find the exact $P$-value.

7.4 The one-sample $t$ statistic from a sample of $n = 30$ observations for the two-sided test of

$H_0$: $\mu = 64$

$H_a$: $\mu \neq 64$

has the value $t = 1.12$.
(a) What are the degrees of freedom for $t$?
(b) Locate the two critical values $t^*$ from Table D that bracket $t$. What are the right-tail probabilities $p$ for these two values?
(c) How would you report the $P$-value for this test?
(d) Is the value $t = 1.12$ statistically significant at the 10% level? At the 5% level?
(e) If you have software available, find the exact $P$-value.

7.5 The one-sample $t$ statistic for a test of

$H_0$: $\mu = 20$

$H_a$: $\mu < 20$

132
based on \( n = 12 \) observations has the value \( t = -2.45 \).

(a) What are the degrees of freedom for this statistic?
(b) How would you report the \( P \)-value based on Table D?
(c) If you have software available, find the exact \( P \)-value.

7.6 A bank wonders whether omitting the annual credit card fee for customers who charge at least $2400 in a year would increase the amount charged on its credit card. The bank makes this offer to an SRS of 250 of its existing credit card customers. It then compares how much these customers charge this year with the amount that they charged last year. The mean increase is $342, and the standard deviation is $108.

(a) Is there significant evidence at the 1% level that the mean amount charged increases under the no-fee offer? State \( H_0 \) and \( H_a \) and carry out a \( t \) test.
(b) Give a 95% confidence interval for the mean amount of the increase.
(c) The distribution of the amount charged is skewed to the right, but outliers are prevented by the credit limit that the bank enforces on each card. Use of the \( t \) procedures is justified in this case even though the population distribution is not normal. Explain why.
(d) A critic points out that the customers would probably have charged more this year than last even without the new offer because the economy is more prosperous and interest rates are lower. Briefly describe the design of an experiment to study the effect of the no-fee offer that would avoid this criticism.

7.7 The bank in the previous exercise tested a new idea on a sample of 250 customers. Suppose that the bank wanted to be quite certain of detecting a mean increase of \( \mu = 100 \) in the amount charged, at the \( \alpha = 0.01 \) significance level. Perhaps a sample of only \( n = 50 \) customers would accomplish this. Find the approximate power of the test with \( n = 50 \) against the alternative \( \mu = 100 \) as follows:

(a) What is the \( t \) critical value for the one-sided test with \( \alpha = 0.01 \) and \( n = 50 \)?
(b) Write the criterion for rejecting \( H_0: \mu = 0 \) in terms of the \( t \) statistic. Then take \( s = 108 \) (an estimate based on the data in Exercise 7.21) and state the rejection criterion in terms of \( \tau \).
(c) Assume that \( \mu = 100 \) (the given alternative) and that \( \sigma = 108 \) (an estimate from the data in the previous exercise). The approximate power is the probability of the event you found in (b), calculated under these assumptions. Find the power. Would you recommend that the bank do a test on 50 customers, or should more customers be included?

7.8 In an experiment on the metabolism of insects, American cockroaches were fed measured amounts of a sugar solution after being deprived of food for a week and of water for 3 days. After 2, 5, and 10 hours, the researchers dissected some of the cockroaches and measured the amount of sugar in various tissues. Five cockroaches fed the sugar D-glucose and dissected after 10 hours had the following amounts (in micrograms) of D-glucose in their hindguts:

\[
55.95 \quad 68.24 \quad 52.73 \quad 21.50 \quad 23.78
\]

Find a 95% confidence interval for the mean amount of D-glucose in cockroach hindguts under these conditions. (Based on D. L. Shankland et al., “The effect of

**7.9** Poisoning by the pesticide DDT causes tremors and convulsions. In a study of DDT poisoning, researchers fed several rats a measured amount of DDT. They then measured electrical characteristics of the rats’ nervous systems that might explain how DDT poisoning causes tremors. One important variable was the “absolutely refractory period,” the time required for a nerve to recover after a stimulus. This period varies normally. Measurements on four rats gave the data below (in milliseconds). (Data from D. L. Shankland, “Involvement of spinal cord and peripheral nerves in DDT-poisoning syndrome in albino rats,” *Toxicology and Applied Pharmacology*, 6 (1964), pp. 97–213.)

<table>
<thead>
<tr>
<th>Refractory Period (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6</td>
</tr>
<tr>
<td>1.7</td>
</tr>
<tr>
<td>1.8</td>
</tr>
<tr>
<td>1.9</td>
</tr>
</tbody>
</table>

(a) Find the mean refractory period $\mu$ and the standard error of the mean.  
(b) Give a 90% confidence interval for the mean “absolutely refractory period” for all rats of this strain when subjected to the same treatment.

**7.10** Suppose that the mean “absolutely refractory period” for unpoisoned rats is known to be 1.3 milliseconds. DDT poisoning should slow nerve recovery and so increase this period. Do the data in the previous exercise give good evidence for this supposition? State $H_0$ and $H_a$ and do a $t$ test. Between what levels from Table D does the $P$-value lie? What do you conclude from the test?

**7.11** The Acculturation Rating Scale for Mexican Americans (ARSMA) measures the extent to which Mexican Americans have adopted Anglo/English culture. During the development of ARSMA, the test was given to a group of 17 Mexicans. Their scores, from a possible range of 1.00 to 5.00, had $\mu = 1.67$ and $s = 0.25$. Because low scores should indicate a Mexican cultural orientation, these results helped to establish the validity of the test. (Based on I. Cuellar, L. C. Harris, and R. Jasso, “An acculturation scale for Mexican American normal and clinical populations,” *Hispanic Journal of Behavioral Sciences*, 2 (1980), pp. 199–217.)

(a) Give a 95% confidence interval for the mean ARSMA score of Mexicans.  
(b) What assumptions does your confidence interval require? Which of these assumptions is most important in this case?

**7.12** The ARSMA test discussed in the previous exercise was compared with a similar test, the Bicultural Inventory (BI), by administering both tests to 22 Mexican Americans. Both tests have the same range of scores (1.00 to 5.00) and are scaled to have similar means for the groups used to develop them. There was a high correlation between the two scores, giving evidence that both are measuring the same characteristics. The researchers wanted to know whether the population mean scores for the two tests were the same. The differences in scores (ARSMA − BI) for the 22 subjects had $\mu = 0.2519$ and $s = 0.2767$.

(a) Describe briefly how the administration of the two tests to the subjects should be conducted, including randomization.

(b) Carry out a significance test for the hypothesis that the two tests have the same population mean. Give the $P$-value and state your conclusion.
(c) Give a 95% confidence interval for the difference between the two population mean scores.

7.13 The paper reporting the results on ARSMA used in Exercise S7.11 does not give the raw data or any discussion of normality. You would like to replace the *t* procedure used in Exercise 7.36 by a sign test. Can you do this from the available information? Carry out the sign test and state your conclusion, or explain why you are unable to carry out the test.

7.14 Exercise S7.12 reports a small study comparing ARSMA and BI, two tests of the acculturation of Mexican Americans. Would this study usually detect a difference in mean scores of 0.2? To answer this question, calculate the approximate power of the test (with *n* = 22 subjects and *α* = 0.05) of

\[ H_0: \mu = 0 \]
\[ H_a: \mu \neq 0 \]

against the alternative *µ* = 0.2. Note that this is a two-sided test.
(a) From Table D, what is the critical value for *α* = 0.05?
(b) Write the criterion for rejecting *H*₀ at the *α* = 0.05 level. Then take *s* = 0.3, the approximate value observed in Exercise 7.36, and restate the rejection criterion in terms of *π*.
(c) Find the probability of this event when *µ* = 0.2 (the alternative given) and *σ* = 0.3 (estimated from the data in Exercise 7.36) by a normal probability calculation. This is the approximate power.

7.15 Gas chromatography is a sensitive technique used by chemists to measure small amounts of compounds. The response of a gas chromatograph is calibrated by repeatedly testing specimens containing a known amount of the compound to be measured. A calibration study for a specimen containing 1 nanogram (ng) (that’s \(10^{-9}\) gram) of a compound gave the following response readings:

\[ 21.6 \quad 20.0 \quad 25.0 \quad 21.9 \]

The response is known from experience to vary according to a normal distribution unless an outlier indicates an error in the analysis. Estimate the mean response to 1 ng of this substance, and give the margin of error for your choice of confidence level. Then explain to a chemist who knows no statistics what your margin of error means.
(Data from the appendix of D. A. Kurtz (ed.), *Trace Residue Analysis*, American Chemical Society Symposium Series, no. 284, 1985.)

7.16 Your local newspaper contains a large number of advertisements for unfurnished one-bedroom apartments. You choose 10 at random and calculate that their mean monthly rent is $540 and that the standard deviation of their rents is $80.
(a) What is the standard error of the mean?
(b) What are the degrees of freedom for a one-sample *t* statistic?

7.17 You want to rent an unfurnished one-bedroom apartment for next semester. You take a random sample of 10 apartments advertised in the local newspaper and
record the rental rates. Here are the rents (in dollars per month):

500, 650, 600, 505, 450, 550, 515, 495, 650, 395

Find a 95% confidence interval for the mean monthly rent for unfurnished one-bedroom apartments available for rent in this community.

7.18 If you chose 99% rather than 95% confidence, would your margin of error in the previous exercise be larger or smaller? Explain your answer and verify it by doing the calculations.

7.19 A random sample of 10 one-bedroom apartments from your local newspaper has these monthly rents (dollars):

500, 650, 600, 505, 450, 550, 515, 495, 650, 395

Do these data give good reason to believe that the mean rent of all advertised apartments is greater than $500 per month? State hypotheses, find the $t$ statistic and its $P$-value, and state your conclusion.

7.20 National Fuelsaver Corporation manufactures the Platinum Gasaver, a device they claim “may increase gas mileage by 22%.” Here are the percent changes in gas mileage for 15 identical vehicles, as presented in one of the company’s advertisements:

48.3 46.9 46.8 44.6 40.2 38.5 34.6 33.7
28.7 28.7 24.8 10.8 10.4 6.9 −12.4

Would you recommend use of a $t$ confidence interval to estimate the mean fuel savings in the population of all such vehicles? Explain your answer.

7.21 A manufacturer of small appliances employs a market research firm to estimate retail sales of its products. Here are last month’s sales of electric can openers from an SRS of 50 stores in the Midwest sales region:

19 19 16 19 25 26 24 63 22 16
13 26 34 10 48 16 20 14 13 24
34 14 25 16 26 25 25 26 11 79
17 25 18 15 13 35 17 15 21 12
19 20 32 19 24 19 17 41 24 27

(a) Make a stemplot of the data. The distribution is skewed to the right and has several high outliers. The bootstrap (page xxx) is a modern computer-intensive tool for getting accurate confidence intervals without the normality condition. Three bootstrap simulations, each with 10,000 repetitions, give these 95% confidence intervals for mean sales in the entire region: (20.42, 27.26), (20.40, 27.18), and (20.48, 27.28).
(b) Find the 95% $t$ confidence interval for the mean. It is essentially the same as the bootstrap intervals. The lesson is that for sample sizes as large as $n = 50$, $t$ procedures are very robust.
7.22 Refer to the previous exercise. Each electric can opener sold generates a profit of $2.50 for the manufacturer.
(a) What is the mean profit per store in the Midwest sales region?
(b) Transform the confidence interval you found in the previous exercise into an interval for the mean profit for stores in the Midwest region.

7.23 Refer to the previous two exercises. There are 4325 stores that sell can openers manufactured by this company.
(a) Estimate the total profit for sales last month in the Midwest region.
(b) Give a 95% confidence interval for the total profit for sales last month in the Midwest region.

7.24 The scores of four roommates on the Law School Admission Test (LSAT) are 628, 593, 455, 503
Find the mean, the standard deviation, and the standard error of the mean. Is it appropriate to calculate a confidence interval for these data? Explain why or why not.

7.25 Here are estimates of the daily intakes of calcium (in milligrams) for 38 women between the ages of 18 and 24 years who participated in a study of women’s bone health:

<table>
<thead>
<tr>
<th>808</th>
<th>882</th>
<th>1062</th>
<th>970</th>
<th>909</th>
<th>802</th>
<th>374</th>
<th>416</th>
<th>784</th>
<th>997</th>
</tr>
</thead>
<tbody>
<tr>
<td>651</td>
<td>716</td>
<td>438</td>
<td>1420</td>
<td>1425</td>
<td>948</td>
<td>1050</td>
<td>976</td>
<td>572</td>
<td>403</td>
</tr>
<tr>
<td>626</td>
<td>774</td>
<td>1253</td>
<td>549</td>
<td>1325</td>
<td>446</td>
<td>465</td>
<td>1269</td>
<td>671</td>
<td>696</td>
</tr>
<tr>
<td>1156</td>
<td>684</td>
<td>1933</td>
<td>748</td>
<td>1203</td>
<td>2433</td>
<td>1255</td>
<td>1100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Display the data using a stemplot and make a normal quantile plot. Describe the distribution of calcium intakes for these women.
(b) Calculate the mean, the standard deviation, and the standard error.
(c) Find a 95% confidence interval for the mean.

7.26 Refer to the previous exercise. Eliminate the two largest values and answer parts (a), (b), and (c).

7.27 Refer to Exercise 7.25. Suppose that the recommended daily allowance (RDA) of calcium for women in this age range is 1300 milligrams (this value is changed from time to time on the basis of the statistical analysis of new data). We want to express the results in terms of percent of the RDA.
(a) Divide each intake by the RDA, multiply by 100, and compute the 95% confidence interval from the transformed data.
(b) Verify that you can obtain the same result by similarly transforming the interval you calculated in Exercise 7.25.

7.28 Refer to Exercises 7.25 and 7.27. You want to compare the average calcium intake of these women with the RDA using a significance test.
(a) State appropriate null and alternative hypotheses.
(b) Give the test statistic, the degrees of freedom, and the P-value.
(c) State your conclusion.
(d) Repeat the calculations without the two largest values. Does your conclusion depend on whether or not these observations are included in the analysis?

7.29 The calcium intake data used in Exercise 7.25 contain two large observations and we have some concern about the use of the \( t \) procedures because of this. In Exercise 7.27 we compared the mean of the data with 1300 milligrams, the RDA. We can use a version of the sign test to compare the median intake with the RDA. First subtract 1300 from each intake. If the population median is 1300, we expect approximately half of the observations to be above the median and half to be below it. The number of observations that will be above the median is binomial with \( n = 38 \) and \( p = 0.5 \). Carry out the sign test and summarize your results.

7.30 How much do users pay for Internet service? Here are the monthly fees (in dollars) paid by a random sample of 50 users of commercial Internet service providers in August 2000: (Data from the August 2000 supplement to the Current Population Survey, from the Census Bureau Web site, www.census.gov.)

\[
\begin{array}{cccccccccccc}
20 & 40 & 22 & 22 & 21 & 21 & 20 & 10 & 20 & 20 \\
20 & 13 & 18 & 50 & 20 & 18 & 15 & 8 & 22 & 25 \\
22 & 10 & 20 & 22 & 22 & 21 & 15 & 23 & 30 & 12 \\
9 & 20 & 40 & 22 & 29 & 19 & 15 & 20 & 20 & 20 \\
20 & 15 & 19 & 21 & 14 & 22 & 21 & 35 & 20 & 22 \\
\end{array}
\]

(a) Make a stemplot of the data. Also make a normal quantile plot if your software permits. The data are not normal: there are stacks of observations taking the same values, and the distribution is more spread out in both directions and somewhat skewed to the right. The \( t \) procedures are nonetheless approximately correct because \( n = 50 \) and there are no extreme outliers.
(b) Give a 95% confidence interval for the mean monthly cost of Internet access in August 2000.

7.31 The data in the previous exercise show that many people paid $20 per month for Internet access, presumably because major providers such as AOL charged this amount. Do the data give good reason to think that the mean cost for all Internet users differs from $20 per month?

7.32 Refer to the two previous exercises concerning fees paid for Internet access by a national random sample of clients of Internet service providers in 2000. The Census Bureau estimates that 44 million households had Internet access in 2000. Use the confidence interval that you found to give a 95% confidence interval for the total amount these households paid in Internet access fees. This is one aspect of the national economic impact of the Internet.

7.33 Refer to the three previous exercises. Suppose you are interested in the cost per year rather than the cost per month. Find a 95% confidence interval for the mean yearly cost of Internet access. How does this interval relate to the one that you found in Exercise 7.20?
7.34 The cost of health care is the subject of many studies that use statistical methods. One such study estimated that the average length of service for home health care among people over the age of 65 who use this type of service is 96.0 days with a standard error of 5.1 days. Assuming that the degrees of freedom are large, calculate a 90% confidence interval for the true mean length of service. (A. N. Dey, “Characteristics of elderly home health care users,” National Center for Health Statistics, 1996.)

7.35 The embryos of brine shrimp can enter a dormant phase in which metabolic activity drops to a low level. Researchers studying this dormant phase measured the level of several compounds important to normal metabolism. The results were reported in a table, with the note, “Values are means ± SEM for three independent samples.” The table entry for the compound ATP was 0.84 ± 0.01. Biologists reading the article are presumed to be able to decipher this. (S. C. Hand and E. Gnaiger, “Anaerobic dormancy quantified in *Artemia* embryos,” *Science*, 239 (1988), pp. 1425–1427.)

(a) What does the abbreviation “SEM” stand for?
(b) The researchers made three measurements of ATP, which had \( \bar{x} = 0.84 \). What was the sample standard deviation \( s \) for these measurements?
(c) Give a 90% confidence interval for the mean ATP level in dormant brine shrimp embryos.

7.36 The design of controls and instruments has a large effect on how easily people can use them. A student project investigated this effect by asking 25 right-handed students to turn a knob (with their right hands) that moved an indicator by screw action. There were two identical instruments, one with a right-hand thread (the knob turns clockwise) and the other with a left-hand thread (the knob turns counterclockwise). The table below gives the times required (in seconds) to move the indicator a fixed distance: (Data provided by Timothy Sturm, Purdue University.)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Right thread</th>
<th>Left thread</th>
<th>Subject</th>
<th>Right thread</th>
<th>Left thread</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>113</td>
<td>137</td>
<td>14</td>
<td>107</td>
<td>87</td>
</tr>
<tr>
<td>2</td>
<td>105</td>
<td>105</td>
<td>15</td>
<td>118</td>
<td>166</td>
</tr>
<tr>
<td>3</td>
<td>130</td>
<td>133</td>
<td>16</td>
<td>103</td>
<td>146</td>
</tr>
<tr>
<td>4</td>
<td>101</td>
<td>108</td>
<td>17</td>
<td>111</td>
<td>123</td>
</tr>
<tr>
<td>5</td>
<td>138</td>
<td>115</td>
<td>18</td>
<td>104</td>
<td>135</td>
</tr>
<tr>
<td>6</td>
<td>118</td>
<td>170</td>
<td>19</td>
<td>111</td>
<td>112</td>
</tr>
<tr>
<td>7</td>
<td>87</td>
<td>103</td>
<td>20</td>
<td>89</td>
<td>93</td>
</tr>
<tr>
<td>8</td>
<td>116</td>
<td>145</td>
<td>21</td>
<td>78</td>
<td>76</td>
</tr>
<tr>
<td>9</td>
<td>75</td>
<td>78</td>
<td>22</td>
<td>100</td>
<td>116</td>
</tr>
<tr>
<td>10</td>
<td>96</td>
<td>107</td>
<td>23</td>
<td>89</td>
<td>78</td>
</tr>
<tr>
<td>11</td>
<td>122</td>
<td>84</td>
<td>24</td>
<td>85</td>
<td>101</td>
</tr>
<tr>
<td>12</td>
<td>103</td>
<td>148</td>
<td>25</td>
<td>88</td>
<td>123</td>
</tr>
<tr>
<td>13</td>
<td>116</td>
<td>147</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Each of the 25 students used both instruments. Discuss briefly how the experiment should be arranged and how randomization should be used.
(b) The project hoped to show that right-handed people find right-hand threads
easier to use. State the appropriate $H_0$ and $H_a$ about the mean time required to complete the task.

(c) Carry out a test of your hypotheses. Give the $P$-value and report your conclusions.

7.37 Refer to the previous exercise. Give a 90% confidence interval for the mean time advantage of right-hand over left-hand threads in the setting of the previous exercise. Do you think that the time saved would be of practical importance if the task were performed many times—for example, by an assembly-line worker? To help answer this question, find the mean time for right-hand threads as a percent of the mean time for left-hand threads.

7.38 An agricultural field trial compares the yield of two varieties of tomatoes for commercial use. The researchers divide in half each of 10 small plots of land in different locations and plant each tomato variety on one half of each plot. After harvest, they compare the yields in pounds per plant at each location. The 10 differences (Variety A $-$ Variety B) give the following statistics: $\overline{x} = 0.46$ and $s = 0.92$. Is there convincing evidence that Variety A has the higher mean yield? State $H_0$ and $H_a$, and give a $P$-value to answer this question.

7.39 The tomato experts who carried out the field trial described in the previous exercise suspect that the relative lack of significance there is due to low power. They would like to be able to detect a mean difference in yields of 0.6 pound per plant at the 0.05 significance level. Based on the previous study, use 0.92 as an estimate of both the population $\sigma$ and the value of $s$ in future samples.

(a) What is the power of the test from Exercise 7.43 with $n = 12$ against the alternative $\mu = 0.6$?

(b) If the sample size is increased to $n = 30$ plots of land, what will be the power against the same alternative?

7.40 The following situations all require inference about a mean or means. Identify each as (1) a single sample, (2) matched pairs, or (3) two independent samples. The procedures of this section apply to cases (1) and (2). We will learn procedures for (3) in the next section.

(a) An education researcher wants to learn whether inserting questions before or after introducing a new concept in an elementary school mathematics text is more effective. He prepares two text segments that teach the concept, one with motivating questions before and the other with review questions after. Each text segment is used to teach a different group of children, and their scores on a test over the material are compared.

(b) Another researcher approaches the same problem differently. She prepares text segments on two unrelated topics. Each segment comes in two versions, one with questions before and the other with questions after. Each of a group of children is taught both topics, one topic (chosen at random) with questions before and the other with questions after. Each child’s test scores on the two topics are compared to see which topic he or she learned better.

(c) To evaluate a new analytical method, a chemist obtains a reference specimen of known concentration from the National Institute of Standards and Technology. She
then makes 20 measurements of the concentration of this specimen with the new method and checks for bias by comparing the mean result with the known concentration.

(d) Another chemist is evaluating the same new method. He has no reference specimen, but a familiar analytic method is available. He wants to know if the new and old methods agree. He takes a specimen of unknown concentration and measures the concentration 10 times with the new method and 10 times with the old method.

7.41 A table gives the number of medical doctors per 100,000 people for each of the 50 states. It does not make sense to use the $t$ procedures (or any other statistical procedures) to give a 95% confidence interval for the mean number of medical doctors per 100,000 people in the population of the American states. Explain why not.

7.42 Computers in some vehicles calculate various quantities related to the performance. One of these is the fuel efficiency, or gas mileage, usually expressed as miles per gallon (MPG). For one vehicle equipped in this way, the MPG was recorded each time the gas tank was filled and the computer was then reset.\footnote{16} Here are the MPG values for a random sample of 20 of these records:

<p>| | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15.8</td>
<td>13.6</td>
<td>15.6</td>
<td>19.1</td>
<td>22.4</td>
<td>15.6</td>
<td>22.5</td>
<td>17.2</td>
<td>19.4</td>
<td>22.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.4</td>
<td>18.0</td>
<td>14.6</td>
<td>18.7</td>
<td>21.0</td>
<td>14.8</td>
<td>22.6</td>
<td>21.5</td>
<td>14.3</td>
<td>20.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Describe the distribution using graphical methods and summarize the results.
(b) Is it appropriate to use methods based on normal distributions to analyze these data? Explain why or why not.
(c) Find the mean, the standard deviation, the standard error, and the margin of error for 95% confidence. Report the 95% confidence interval for $\mu$, the mean MPG for this vehicle based on these data.
(d) Do you think that this interval would apply to other similar vehicles? Give reasons why and why not.

7.43 Refer to the previous exercise. Here are the values of the average speed in miles per hour (MPH) for the same sample:

|   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|
| 21.0 | 19.0 | 18.7 | 39.2 | 45.8 | 19.8 | 48.4 | 21.0 | 29.1 |
| 31.6 | 49.0 | 16.0 | 34.6 | 36.3 | 19.0 | 43.3 | 37.5 | 16.5 |

Answer the questions given in the previous exercise.

7.44 You will have complete sales information for last month in a week, but right now you have data from a random sample of 40 stores. The mean change in sales in the sample is +3.8% and the standard deviation of the changes is 12%. Are average sales for all stores different from last month?
(a) State appropriate null and alternative hypotheses. Explain how you decided between the one- and two-sided alternatives.
(b) Find the $t$ statistic and its $P$-value. State your conclusion.
(c) If the test gives strong evidence against the null hypothesis, would you conclude that sales are up in every one of your stores? Explain your answer.

7.45 For a sample of size 5, a test of a null hypothesis versus a two-sided alternative gives $t = 2.45$. 
(a) Is the test result significant at the 5% level? Draw a sketch of the appropriate t
distribution and illustrate your calculation with this sketch.

(b) Now assume that the same statistic was obtained for a sample size of \( n = 10 \).
Assess the statistical significance of the result and illustrate the calculation with a
sketch. How did the statistical significance change with the sample size? Explain
your answer.

7.46 Assume a sample size of \( n = 2000 \). Draw a picture of the distribution of the \( t \)
statistic under the null hypothesis. Use your picture to illustrate the values of the
test statistic that would lead to rejection of the null hypothesis at the 1% level for
a two-sided alternative.

7.47 Repeat the previous exercise for the two situations where the alternative is
one-sided.

7.48 Computer software reports \( \bar{x} = 12.3 \) and \( P = 0.08 \) for a \( t \) test of \( H_0: \mu = 0 \)
versus \( H_a: \mu \neq 0 \). Base on prior knowledge, you can justify testing the alternative
\( H_a: \mu > 0 \). What is the \( P \)-value for your significance test?

7.49 Suppose that \( \bar{x} = -12.3 \) in the setting of the previous exercise. Would this
change your answer? Use a sketch of the distribution of the test statistic under the
null hypothesis to illustrate and explain your answer.

7.50 Use Table D to find the critical value \( t^* \) to be used for a confidence interval
for the mean of the population in each of the following situations.
(a) A 95% confidence interval based on \( n = 20 \) observations.
(b) A 90% confidence interval from an SRS of 30 observations.
(c) An 80% confidence interval from a sample of size 50.

7.51 Use software to find the critical values \( t^* \) that you would use for 95% confidence
intervals for sample sizes of 10, 20, 30, 40, 50, 100, 200, and 500. Plot the values of
\( t^* \) versus the sample size and describe the relationship.

7.52 A sample of size \( n = 15 \) is used to perform a significance test for \( H_0: \mu = 0 \)
versus \( H_a: \mu > 0 \). The test statistic is \( t = 2.15 \).
(a) What are the degrees of freedom for this statistic?
(b) Give the two critical values \( t^* \) from Table D that bracket \( t \).
(c) What are the right-tail probabilities \( p \) for these two entries?
(d) Between what two values does the \( P \)-value of the test fall?
(e) Sketch the \( t \) distribution for this exercise and illustrate your answers to parts
(c) and (d) with the sketch.
(f) Is the value \( t = 2.15 \) significant at the 5% level? Is it significant at the 1% level?
(g) If you have software available, find the exact \( P \)-value.

7.53 The hypotheses \( H_0: \mu = 50 \) and \( H_a: \mu \neq 50 \) are examined using a sample of
size \( n = 30 \). The one-sample \( t \) statistic has the value \( t = 1.35 \).
(a) Give the degrees of freedom for the test statistic.
(b) Locate the two critical values \( t^* \) from Table D that bracket \( t \). What are the
right-tail probabilities \( p \) for these two values?
(c) How would you report the $P$-value for this test?
(d) Is the value $t = 1.35$ statistically significant at the 10% level? At the 5% level?
(e) Illustrate your answers to the previous parts of this exercise with a sketch of the $t$ distribution.
(f) If you have software available, find the exact $P$-value.

7.54 The one-sample $t$ statistic for a test of $H_0: \mu = 30$ versus $H_a: \mu < 30$ based on
$n = 12$ observations has the value $t = -3.21$.
(a) Find the degrees of freedom for this statistic.
(b) Use Table D to find an approximate $P$-value. Use a sketch of the $t$ distribution
to illustrate your work.
(c) Find the exact $P$-value if you have software available.

7.55 Here are estimates of the daily intakes of calcium (in milligrams) for 40 women
between the ages of 18 and 24 years who participated in a study of women’s bone
health:

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>725</td>
<td>764</td>
<td>853</td>
<td>559</td>
<td>1225</td>
<td>456</td>
<td>555</td>
<td>1168</td>
</tr>
<tr>
<td>808</td>
<td>882</td>
<td>1062</td>
<td>970</td>
<td>909</td>
<td>802</td>
<td>574</td>
<td>576</td>
</tr>
<tr>
<td>976</td>
<td>848</td>
<td>1050</td>
<td>976</td>
<td>572</td>
<td>483</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1256</td>
<td>685</td>
<td>1144</td>
<td>848</td>
<td>1213</td>
<td>3223</td>
<td>1270</td>
<td>1108</td>
</tr>
<tr>
<td>350</td>
<td>421</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Use a stemplot or other graphical summary to describe the distribution of in-
takes. If you have software available, make a normal quantile plot. Write a short
paragraph describing the distribution. Be sure to refer to your graphics in your
summary.
(b) Find the mean and the standard deviation. Sketch a normal curve with this
mean and standard deviation.
(c) Write a sentence in which you give the 95% confidence interval for the mean and
an explanation of how it should be interpreted.
(d) There is an outlier. Eliminate it and answer parts (a), (b), and (c) again. How
do the results change?
(e) Take one side of the following issue and present reasons for your views. “These
results do not apply to women aged 18 to 24 years from the same community who
did not volunteer to participate in the study.”

7.56 Refer to the previous exercise. For some nutrients including calcium, recom-
mendations are expressed as adequate intakes (AI). The AI for women in this age
range is 1000 milligrams. Let’s compare the intakes of the women in this sample
with the AI using a significance test. Use the data for all 40 women.
(a) What null and alternative hypotheses would you use for this problem?
(b) Give the test statistic, the degrees of freedom, the $P$-value, and a sketch illus-
trating the $P$-value.
(c) Write a short paragraph giving the descriptive statistics and the significance test
results for this problem.
(d) (Optional) Use the sign test for this problem. Compare the two approaches to
the analysis of these data.

7.57 Many organizations are doing surveys to determine the satisfaction of their
customers. Attitudes toward various aspects of campus life were the subject of one
such study conducted at Purdue University. Each item was rated on a 1 to 5 scale, with 5 being the highest rating. The average response of 1406 first-year students to “Feeling welcomed at Purdue” was 3.9 with a standard deviation of 0.98. Assuming that the respondents are an SRS, give a 99% confidence interval for the mean of all first-year students.

7.58 Do piano lessons improve the spatial-temporal reasoning of preschool children? Neurobiological arguments suggest that this may be true. A study designed to test this hypothesis measured the spatial-temporal reasoning of 34 preschool children before and after six months of piano lessons. (The study also included children who took computer lessons and a control group; but we are not concerned with those here.) The changes in the reasoning scores are

\[ \begin{array}{ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
started with. For example, a loss of 50 mg/100 g would probably not be of much concern if we started with 5000 mg/100 g. The specifications call for the blend to contain 98 mg/100 g (dry basis). The difference between this specification and the “before” values above is due to sample variation in the manufacturing process and the handling of the product from the time it was manufactured until it was used to prepare gruel in these Haitian homes. Express the “after” data as percent of specification and give a 95% confidence interval for the mean percent.

7.62 A bank wonders whether omitting the annual credit card fee for customers who charge at least $3000 in a year would increase the amount charged on its credit card. The bank makes this offer to an SRS of 500 of its existing credit card customers. It then compares how much these customers charge this year with the amount that they charged last year. The mean increase is $565, and the standard deviation is $267.

(a) Is there significant evidence at the 1% level that the mean amount charged increases under the no-fee offer? State $H_0$ and $H_a$ and carry out a $t$ test.
(b) Give a 95% confidence interval for the mean amount of the increase.
(c) The distribution of the amount charged is skewed to the right, but outliers are prevented by the credit limit that the bank enforces on each card. Use of the $t$ procedures is justified in this case even though the population distribution is not normal. Explain why.
(d) A critic points out that the customers would probably have charged more this year than last even without the new offer because the economy is more prosperous and interest rates are lower. Briefly describe the design of an experiment to study the effect of the no-fee offer that would avoid this criticism.

7.63 In a randomized comparative experiment on the effect of dietary calcium on blood pressure, 54 healthy white males were divided at random into two groups. One group received calcium; the other, a placebo. At the beginning of the study, the researchers measured many variables on the subjects. The paper reporting the study gives $\overline{x} = 114.9$ and $s = 9.3$ for the seated systolic blood pressure of the 27 members of the placebo group.

(a) Give a 95% confidence interval for the mean blood pressure of the population from which the subjects were recruited.
(b) What assumptions about the population and the study design are required by the procedure you used in (a)? Which of these assumptions are important for the validity of the procedure in this case?

7.64 How accurate are radon detectors of a type sold to homeowners? To answer this question, university researchers placed 12 detectors in a chamber that exposed them to 105 picocuries per liter (pCi/l) of radon. The detector readings were as follows:

<table>
<thead>
<tr>
<th>91.9</th>
<th>97.8</th>
<th>111.4</th>
<th>122.3</th>
<th>105.4</th>
<th>95.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>103.8</td>
<td>99.6</td>
<td>96.6</td>
<td>119.3</td>
<td>104.8</td>
<td>101.7</td>
</tr>
</tbody>
</table>

(a) Make a stemplot of the data. The distribution is somewhat skewed to the right, but not strongly enough to forbid use of the $t$ procedures.
(b) Is there convincing evidence that the mean reading of all detectors of this type
differs from the true value of 105? Carry out a test in detail and write a brief conclusion.

**7.65** The researchers studying vitamin C in CSB in Example 7.1 were also interested in a similar commodity called wheat soy blend (WSB). A major concern was the possibility that some of the vitamin C content would be destroyed as a result of storage and shipment of the commodity to its final destination. The researchers specially marked a collection of bags at the factory and took a sample from each of these to determine the vitamin C content. Five months later in Haiti they found the specially marked bags and took samples. The data consist of two vitamin C measures for each bag, one at the time of production in the factory and the other five months later in Haiti. The units are mg/100 g as in Example 7.1. Here are the data:

<table>
<thead>
<tr>
<th>Factory</th>
<th>Haiti</th>
<th>Factory</th>
<th>Haiti</th>
<th>Factory</th>
<th>Haiti</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>40</td>
<td>45</td>
<td>38</td>
<td>39</td>
<td>43</td>
</tr>
<tr>
<td>50</td>
<td>37</td>
<td>32</td>
<td>40</td>
<td>52</td>
<td>38</td>
</tr>
<tr>
<td>48</td>
<td>39</td>
<td>47</td>
<td>35</td>
<td>45</td>
<td>38</td>
</tr>
<tr>
<td>44</td>
<td>35</td>
<td>40</td>
<td>38</td>
<td>37</td>
<td>38</td>
</tr>
<tr>
<td>42</td>
<td>35</td>
<td>38</td>
<td>34</td>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td>47</td>
<td>41</td>
<td>41</td>
<td>35</td>
<td>44</td>
<td>40</td>
</tr>
<tr>
<td>49</td>
<td>37</td>
<td>43</td>
<td>37</td>
<td>43</td>
<td>35</td>
</tr>
<tr>
<td>50</td>
<td>37</td>
<td>40</td>
<td>34</td>
<td>39</td>
<td>38</td>
</tr>
<tr>
<td>39</td>
<td>34</td>
<td>37</td>
<td>40</td>
<td>44</td>
<td>36</td>
</tr>
</tbody>
</table>

(a) Describe the data graphically and numerically. Summarize your results.
(b) Set up hypotheses to examine the question of interest to these researchers.
(c) Perform the significance test and summarize your results.
(d) Find 95% confidence intervals for the mean at the factory, the mean five months later in Haiti, and for the change.

**7.66** The table below gives the pretest and posttest scores on the MLA listening test in Spanish for 20 high school Spanish teachers who attended an intensive summer course in Spanish. The setting is identical to the one described in the previous exercise.

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Teacher</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>29</td>
<td>11</td>
<td>30</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>30</td>
<td>12</td>
<td>29</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>32</td>
<td>13</td>
<td>31</td>
<td>34</td>
</tr>
<tr>
<td>4</td>
<td>26</td>
<td>30</td>
<td>14</td>
<td>29</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>16</td>
<td>15</td>
<td>34</td>
<td>32</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>25</td>
<td>16</td>
<td>20</td>
<td>27</td>
</tr>
<tr>
<td>7</td>
<td>34</td>
<td>31</td>
<td>17</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>18</td>
<td>18</td>
<td>25</td>
<td>29</td>
</tr>
<tr>
<td>9</td>
<td>28</td>
<td>33</td>
<td>19</td>
<td>31</td>
<td>32</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>25</td>
<td>20</td>
<td>29</td>
<td>32</td>
</tr>
</tbody>
</table>

Summarize the data graphically and numerically. Then analyze the data using a significance test and a confidence interval. Write a short report summarizing your results.
Exercise 7.31 gives data on the amount of vitamin C in gruel made from wheat soy blend in 5 Haitian households before and after cooking. Is there evidence that the median amount of vitamin C is less after cooking? State hypotheses, carry out a sign test, and report your conclusion.

Apply the sign test to the data in Exercise 7.29 to assess the effects of piano lessons on spatial-temporal reasoning.
(a) State the hypotheses two ways: in terms of a population median and in terms of the probability of an improvement in the test score.
(b) Carry out the sign test. Find the approximate \( P \)-value using the normal approximation to the binomial distributions, and report your conclusion.

Use the sign test to assess whether the summer institute of Exercise 7.42 improves Spanish listening skills. State the hypotheses, give the \( P \)-value using the binomial table (Table C), and report your conclusion.

In a study of cereal leaf beetle damage on oats, researchers measured the number of beetle larvae per stem in small plots of oats after randomly applying one of two treatments: no pesticide or Malathion at the rate of 0.25 pound per acre. Here are the data:

| Control: | 2 4 3 4 2 3 3 5 3 2 6 3 4 |
| Treatment: | 0 1 1 2 1 2 1 1 2 1 1 1 |

(Based on M. C. Wilson et al., “Impact of cereal leaf beetle larvae on yields of oats,” *Journal of Economic Entomology*, 62 (1969), pp. 699–702.) Is there significant evidence at the 1% level that the mean number of larvae per stem is reduced by Malathion? Be sure to state \( H_0 \) and \( H_a \).

A bank compares two proposals to increase the amount that its credit card customers charge on their cards. (The bank earns a percentage of the amount charged, paid by the stores that accept the card.) Proposal A offers to eliminate the annual fee for customers who charge $2400 or more during the year. Proposal B offers a small percent of the total amount charged as a cash rebate at the end of the year. The bank offers each proposal to an SRS of 150 of its existing credit card customers. At the end of the year, the total amount charged by each customer is recorded. Here are the summary statistics:

<table>
<thead>
<tr>
<th>Group</th>
<th>( n )</th>
<th>( \bar{x} )</th>
<th>( s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>150</td>
<td>$1987$</td>
<td>$392$</td>
</tr>
<tr>
<td>B</td>
<td>150</td>
<td>$2056$</td>
<td>$413$</td>
</tr>
</tbody>
</table>

(a) Do the data show a significant difference between the mean amounts charged by customers offered the two plans? Give the null and alternative hypotheses, and calculate the two-sample \( t \) statistic. Obtain the \( P \)-value (either approximately from Table D or more accurately from software). State your practical conclusions.
(b) The distributions of amounts charged are skewed to the right, but outliers are
prevented by the limits that the bank imposes on credit balances. Do you think that skewness threatens the validity of the test that you used in (a)? Explain your answer.

7.72 What aspects of rowing technique distinguish between novice and skilled competitive rowers? Researchers compared two groups of female competitive rowers: a group of skilled rowers and a group of novices. The researchers measured many mechanical aspects of rowing style as the subjects rowed on a Stanford Rowing Ergometer. One important variable is the angular velocity of the knee (roughly, the rate at which the knee joint opens as the legs push the body back on the sliding seat). This variable was measured when the oar was at right angles to the machine. (Based on W. N. Nelson and C. J. Widule, “Kinematic analysis and efficiency estimate of intercollegiate female rowers,” unpublished manuscript, 1983.) The data show no outliers or strong skewness. Here is the SAS computer output:

TTEST PROCEDURE

Variable: KNEE

<table>
<thead>
<tr>
<th>GROUP</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKILLED</td>
<td>10</td>
<td>4.18283335</td>
<td>0.47905935</td>
<td>0.15149187</td>
</tr>
<tr>
<td>NOVICE</td>
<td>8</td>
<td>3.01000000</td>
<td>0.95894830</td>
<td>0.33903942</td>
</tr>
</tbody>
</table>

Variances

|          | T    | DF  | Prob>|T| |
|----------|------|-----|------|---|
| Unequal  | 3.1583 | 9.8 | 0.0104 |
| Equal    | 3.3918 | 16.0 | 0.0037 |

(a) The researchers believed that the knee velocity would be higher for skilled rowers. State $H_0$ and $H_a$.
(b) Give the value of the two-sample $t$ statistic and its $P$-value (note that SAS provides two-sided $P$-values). What do you conclude?
(c) Give a 90% confidence interval for the mean difference between the knee velocities of skilled and novice female rowers.

7.73 The novice and skilled rowers in the previous exercise were also compared with respect to several physical variables. Here is the SAS computer output for weight in kilograms:

TTEST PROCEDURE

Variable: WEIGHT

<table>
<thead>
<tr>
<th>GROUP</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKILLED</td>
<td>10</td>
<td>70.3700000</td>
<td>6.10034898</td>
<td>1.92909973</td>
</tr>
<tr>
<td>NOVICE</td>
<td>8</td>
<td>68.4500000</td>
<td>9.03999930</td>
<td>3.19612240</td>
</tr>
</tbody>
</table>
Section 7.2

| Variances | T      | DF | Prob>|T| |
|-----------|--------|----|----------------|
| Unequal   | 0.5143 | 11.8| 0.6165        |
| Equal     | 0.5376 | 16.0| 0.5982        |

Is there significant evidence of a difference in the mean weights of skilled and novice rowers? State $H_0$ and $H_a$, report the two-sample $t$ statistic and its $P$-value, and state your conclusion.

**7.74** The Johns Hopkins Regional Talent Searches give the SAT (intended for high school juniors and seniors) to 13-year-olds. In all, 19,883 males and 19,937 females took the tests between 1980 and 1982. The mean scores of males and females on the verbal test are nearly equal, but there is a clear difference between the sexes on the mathematics test. The reason for this difference is not understood. Here are the data (from a news article in *Science*, 224 (1983), pp. 1029–1031):

<table>
<thead>
<tr>
<th>Group</th>
<th>$\bar{x}$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>416</td>
<td>87</td>
</tr>
<tr>
<td>Females</td>
<td>386</td>
<td>74</td>
</tr>
</tbody>
</table>

Give a 99% confidence interval for the difference between the mean score for males and the mean score for females in the population that Johns Hopkins searches.

**7.75** Plant scientists have developed varieties of corn that have increased amounts of the essential amino acid lysine. In a test of the protein quality of this corn, an experimental group of 20 one-day-old male chicks was fed a ration containing the new corn. A control group of another 20 chicks received a ration that was identical except that it contained normal corn. Here are the weight gains (in grams) after 21 days. (Based on G. L. Cromwell et al., “A comparison of the nutritive value of opaque-2, floury-2 and normal corn for the chick,” *Poultry Science*, 47 (1968), pp. 840–847.)

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>380</td>
<td>361</td>
</tr>
<tr>
<td></td>
<td>321</td>
<td>447</td>
</tr>
<tr>
<td></td>
<td>366</td>
<td>401</td>
</tr>
<tr>
<td></td>
<td>356</td>
<td>375</td>
</tr>
<tr>
<td></td>
<td>410</td>
<td>406</td>
</tr>
<tr>
<td></td>
<td>329</td>
<td>318</td>
</tr>
<tr>
<td></td>
<td>399</td>
<td>467</td>
</tr>
<tr>
<td></td>
<td>272</td>
<td>407</td>
</tr>
<tr>
<td></td>
<td>345</td>
<td>430</td>
</tr>
<tr>
<td></td>
<td>455</td>
<td>439</td>
</tr>
<tr>
<td></td>
<td>360</td>
<td>410</td>
</tr>
<tr>
<td></td>
<td>431</td>
<td>326</td>
</tr>
</tbody>
</table>

(a) Present the data graphically. Are there outliers or strong skewness that might prevent the use of $t$ procedures?
(b) State the hypotheses for a statistical test of the claim that chicks fed high-lysine corn gain weight faster. Carry out the test. Is the result significant at the 10% level? At the 5% level? At the 1% level?
(c) Give a 95% confidence interval for the mean extra weight gain in chicks fed high-lysine corn.

**7.76** The data on weights of skilled and novice rowers in Exercise S7.19 can be analyzed by the pooled $t$ procedures, which assume equal population variances. Report the value of the $t$ statistic, its degrees of freedom, and its $P$-value, and then
state your conclusion. (The pooled procedures should not be used for the comparison of knee velocities in Exercise S7.18, because the sample standard deviations in the two groups are different enough to cast doubt on the assumption of a common population standard deviation.)

7.77 Pat wants to compare the cost of one- and two-bedroom apartments in the area of your campus. She collects data for a random sample of 10 advertisements of each type. Here are the rents for the two-bedroom apartments (in dollars per month):

595, 500, 580, 650, 675, 750, 500, 495, 670

Here are the rents for the one-bedroom apartments:

500, 650, 600, 505, 450, 550, 495, 650; 395

Find a 95% confidence interval for the additional cost of a second bedroom.

7.78 Pat wonders if two-bedroom apartments rent for significantly more than one-bedroom apartments. Use the data in the previous exercise to find out.
(a) State appropriate null and alternative hypotheses.
(b) Report the test statistic, its degrees of freedom, and the P-value. What do you conclude?
(c) Can you conclude that every one-bedroom apartment costs less than every two-bedroom apartment?
(d) In the previous exercise you found a confidence interval. In this exercise you performed a significance test. Which do you think is more useful to someone planning to rent an apartment? Why?

7.79 Physical fitness is related to personality characteristics. In one study of this relationship, middle-aged college faculty who had volunteered for a fitness program were divided into low-fitness and high-fitness groups based on a physical examination. The subjects then took the Cattell Sixteen Personality Factor Questionnaire. (A. H. Ismail and R. J. Young, “The effect of chronic exercise on the personality of middle-aged men,” *Journal of Human Ergology*, 2 (1973), pp. 47–57.) Here are the data for the “ego strength” personality factor:

<table>
<thead>
<tr>
<th>Low fitness</th>
<th>High fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.99</td>
<td>5.53</td>
</tr>
<tr>
<td>4.24</td>
<td>4.12</td>
</tr>
<tr>
<td>4.74</td>
<td>5.10</td>
</tr>
<tr>
<td>4.93</td>
<td>4.47</td>
</tr>
<tr>
<td>4.16</td>
<td>5.30</td>
</tr>
</tbody>
</table>

(a) Is the difference in mean ego strength significant at the 5% level? At the 1% level? Be sure to state $H_0$ and $H_a$.
(b) You should be hesitant to generalize these results to the population of all middle-aged men. Explain why.

7.80 The U.S. Department of Agriculture (USDA) uses many types of surveys to obtain important economic estimates. In one pilot study they estimated wheat prices
in July and in September using independent samples. Here is a brief summary from the report:\[19\]

<table>
<thead>
<tr>
<th>Month</th>
<th>( n )</th>
<th>( \bar{x} )</th>
<th>( s/\sqrt{n} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>July</td>
<td>90</td>
<td>$2.95$</td>
<td>$0.023$</td>
</tr>
<tr>
<td>September</td>
<td>45</td>
<td>$3.61$</td>
<td>$0.029$</td>
</tr>
</tbody>
</table>

(a) Note that the report gave standard errors. Find the standard deviation for each of the samples.
(b) Use a significance test to examine whether or not the price of wheat was the same in July and September. Be sure to give details and carefully state your conclusion.

7.81 Refer to the previous exercise. Give a 95% confidence interval for the increase in price between July and September.

7.82 A market research firm supplies manufacturers with estimates of the retail sales of their products from samples of retail stores. Marketing managers are prone to look at the estimate and ignore sampling error. Suppose that an SRS of 75 stores this month shows mean sales of 52 units of a small appliance, with standard deviation 13 units. During the same month last year, an SRS of 53 stores gave mean sales of 49 units, with standard deviation 11 units. An increase from 49 to 52 is a rise of 6%. The marketing manager is happy, because sales are up 6%.
(a) Use the two-sample \( t \) procedure to give a 95% confidence interval for the difference in mean number of units sold at all retail stores.
(b) Explain in language that the manager can understand why he cannot be certain that sales rose by 6%, and that in fact sales may even have dropped.

7.83 In a study of heart surgery, one issue was the effect of drugs called beta-blockers on the pulse rate of patients during surgery. The available subjects were divided at random into two groups of 30 patients each. One group received a beta-blocker; the other, a placebo. The pulse rate of each patient at a critical point during the operation was recorded. The treatment group had mean 65.2 and standard deviation 7.8. For the control group, the mean was 70.3 and the standard deviation was 8.3.
(a) Do beta-blockers reduce the pulse rate? State the hypotheses and do a \( t \) test. Is the result significant at the 5% level? At the 1% level?
(b) Give a 99% confidence interval for the difference in mean pulse rates.

7.84 The table below shows Consumer Reports magazine’s laboratory test results for calories and milligrams of sodium (mostly due to salt) in a number of major brands of hot dogs. There are three types: all beef, “meat” (mainly pork and beef, but government regulations allow up to 15% poultry meat), and poultry. (Consumer Reports, June 1986, pp. 366–367.)
(a) Give a 95% confidence interval for the difference in mean calorie content between beef and poultry hot dogs.

(b) Based on your confidence interval, can the hypothesis that the population means are equal be rejected at the 5% significance level? Explain your answer.

(c) What assumptions does your statistical procedure in (a) require? Which of these assumptions are justified or not important in this case? Are any of the assumptions doubtful in this case?

7.85 The following table gives data on the blood pressure before and after treatment for two groups of black males.

<table>
<thead>
<tr>
<th>Calcium group</th>
<th>Placebo group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begin</td>
<td>End</td>
</tr>
<tr>
<td>107</td>
<td>100</td>
</tr>
<tr>
<td>110</td>
<td>114</td>
</tr>
<tr>
<td>123</td>
<td>105</td>
</tr>
<tr>
<td>129</td>
<td>112</td>
</tr>
<tr>
<td>112</td>
<td>115</td>
</tr>
<tr>
<td>111</td>
<td>116</td>
</tr>
<tr>
<td>107</td>
<td>106</td>
</tr>
<tr>
<td>112</td>
<td>102</td>
</tr>
<tr>
<td>136</td>
<td>125</td>
</tr>
<tr>
<td>102</td>
<td>104</td>
</tr>
<tr>
<td>130</td>
<td>133</td>
</tr>
</tbody>
</table>
One group took a calcium supplement, and the other group received a placebo. Example 7.20 compares the decrease in blood pressure in the two groups using pooled two-sample \( t \) procedures.

(a) Repeat the significance test using a two-sample \( t \) test that does not require equal population standard deviations. Compare your \( P \)-value with the result \( P = 0.059 \) for the pooled \( t \) test.

(b) Give a 90% confidence interval for the difference in means, again using a procedure that does not require equal standard deviations. How does the margin of error of your interval compare with that in Example 7.21?

### 7.86

Researchers studying the learning of speech often compare measurements made on the recorded speech of adults and children. One variable of interest is called the voice onset time (VOT). Here are the results for 6-year-old children and adults asked to pronounce the word “bees.” The VOT is measured in milliseconds and can be either positive or negative. (M. A. Zlatin and R. A. Koenigsknecht, “Development of the voicing contrast: a comparison of voice onset time in stop perception and production,” *Journal of Speech and Hearing Research*, 19 (1976), pp. 93–111.)

<table>
<thead>
<tr>
<th>Group</th>
<th>( n )</th>
<th>( \bar{x} )</th>
<th>( s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children</td>
<td>10</td>
<td>-3.67</td>
<td>33.89</td>
</tr>
<tr>
<td>Adults</td>
<td>20</td>
<td>-23.17</td>
<td>50.74</td>
</tr>
</tbody>
</table>

(a) What is the standard error of the sample mean VOT for the 20 adult subjects? What is the standard error of the difference \( \bar{x}_1 - \bar{x}_2 \) between the mean VOT for children and adults?

(b) The researchers were investigating whether VOT distinguishes adults from children. State \( H_0 \) and \( H_a \) and carry out a two-sample \( t \) test. Give a \( P \)-value and report your conclusions.

(c) Give a 95% confidence interval for the difference in mean VOTs when pronouncing the word “bees.” Explain why you knew from your result in (b) that this interval would contain 0 (no difference).

### 7.87

The researchers in the study discussed in the previous exercise looked at VOTs for adults and children pronouncing several different words. Explain why they should not perform a separate two-sample \( t \) test for each word and conclude that the words with a significant difference (say, \( P < 0.05 \)) distinguish children from adults. (The researchers did not make this mistake.)

### 7.88

Repeat the comparison of mean VOTs for children and adults in Exercise 7.58 using a pooled \( t \) procedure. (In practice, we would not pool in this case, because the data suggest some difference in the population standard deviations.)

(a) Carry out the significance test, and give a \( P \)-value.

(b) Give a 95% confidence interval for the difference in population means.

(c) How similar are your results to those you obtained in Exercise 7.79 from the two-sample \( t \) procedures?

### 7.89

College financial aid offices expect students to use summer earnings to help pay for college. But how large are these earnings? One college studied this question
154 Chapter 7 Exercises

by asking a sample of students how much they earned. Omitting students who were not employed, 1296 responses were received. (Based on studies conducted by Marvin Schlatter, Division of Financial Aid, Purdue University.) Here are the data in summary form:

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>x̄</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>675</td>
<td>$3297.91</td>
<td>$2394.65</td>
</tr>
<tr>
<td>Females</td>
<td>621</td>
<td>$2380.68</td>
<td>$1815.55</td>
</tr>
</tbody>
</table>

(a) Use the two-sample t procedures to give a 90% confidence interval for the difference between the mean summer earnings of male and female students.

(b) The distribution of earnings is strongly skewed to the right. Nevertheless, use of t procedures is justified. Why?

(c) Once the sample size was decided, the sample was chosen by taking every kth name from an alphabetical list of undergraduates. Is it reasonable to consider the sample as two SRSs chosen from the male and female undergraduate populations?

(d) What other information about the study would you request before accepting the results as describing all undergraduates?

7.90 The pesticide DDT causes tremors and convulsions if it is ingested by humans or other mammals. Researchers seek to understand how the convulsions are caused. In a randomized comparative experiment, 6 white rats poisoned with DDT were compared with a control group of 6 unpoisoned rats. Electrical measurements of nerve activity are the main clue to the nature of DDT poisoning. When a nerve is stimulated, its electrical response shows a sharp spike followed by a much smaller second spike. Researchers found that the second spike is larger in rats fed DDT than in normal rats. This observation helps biologists understand how DDT causes tremors. (This example is loosely based on D. L. Shankland, “Involvement of spinal cord and peripheral nerves in DDT-poisoning syndrome in albino rats,” Toxicology and Applied Pharmacology, 6 (1964), pp. 197–213.)

The researchers measured the amplitude of the second spike as a percentage of the first spike when a nerve in the rat’s leg was stimulated. For the poisoned rats the results were

\[
\begin{align*}
12.207 & \quad 16.869 & \quad 25.050 & \quad 22.429 & \quad 8.456 & \quad 20.589 \\
\end{align*}
\]

The control group data were

\[
\begin{align*}
11.074 & \quad 9.686 & \quad 12.064 & \quad 9.351 & \quad 8.182 & \quad 6.642 \\
\end{align*}
\]

Normal quantile plots (Figure 7.13) show no evidence of outliers or strong skewness. Both populations are reasonably normal, as far as can be judged from 6 observations. The difference in means is quite large, but in such small samples the sample mean is highly variable. A significance test can help confirm that we are seeing a real effect. Because the researchers did not conjecture in advance that the size of the second spike would increase in rats fed DDT, we test

\[
H_0: \mu_1 = \mu_2 \\
H_a: \mu_1 \neq \mu_2
\]

Here is the output from a statistical software system for these data:
TTEST PROCEDURE

Variable: SPIKE

<table>
<thead>
<tr>
<th>GROUP</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDT</td>
<td>6</td>
<td>17.60000000</td>
<td>6.34014839</td>
<td>2.58835474</td>
</tr>
<tr>
<td>CONTROL</td>
<td>6</td>
<td>9.49983333</td>
<td>1.95005932</td>
<td>0.79610839</td>
</tr>
</tbody>
</table>

Variances

|         | T     | DF  | Prob>|T| |
|---------|-------|-----|-----|----|
| Unequal | 2.9912 | 5.9 | 0.0247 |
| Equal   | 2.9912 | 10.0| 0.0135 |

(a) Interpret the output.

(b) Starting from the computer’s results for $\bar{x}_i$ and $s_i$, verify the values given for the test statistic $t = 2.99$ and the degrees of freedom $df = 5.9$.

7.91 The Chapin Social Insight Test is a psychological test designed to measure how accurately the subject appraises other people. The possible scores on the test range from 0 to 41. During the development of the Chapin test, it was given to several different groups of people. Here are the results for male and female college students majoring in the liberal arts:

<table>
<thead>
<tr>
<th>Group</th>
<th>Sex</th>
<th>n</th>
<th>$\bar{x}$</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>133</td>
<td>25.34</td>
<td>5.05</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>162</td>
<td>24.94</td>
<td>5.44</td>
</tr>
</tbody>
</table>

Do these data support the contention that female and male students differ in average social insight? Use the pooled two-sample procedure and the procedure that does not assume that the standard deviations are the same. Compare the results.

7.92 In each of the following situations explain what is wrong and why.

(a) A researcher wants to test $H_0: \bar{x}_1 = \bar{x}_2$ versus the two-sided alternative $H_1: \bar{x}_1 \neq \bar{x}_2$.

(b) A study recorded the scores of 20 children who were similar in age. The scores of the 10 boys in the study were compared with the scores of all 20 children using the two-sample methods of this section.

(c) A two-sample $t$ statistic gave a $P$-value of 0.96. From this you can reject the null hypothesis with 95% confidence.

7.93 For each of the following, answer the question and give a short explanation of your reasoning.

(a) A 95% confidence interval for the difference between two means is reported as (1.6, 2.3). What can you conclude about the results of a significance test of the null hypothesis that the population means are equal versus the two-sided alternative?

(b) Will larger samples generally give a larger or smaller margin of error for the difference between two sample means?
7.94 For each of the following, answer the question and give a short explanation of your reasoning.
(a) A significance test for comparing two means gave \( t = -3.11 \) with 23 degrees of freedom. Can you reject the null hypothesis that the \( \mu \)'s are equal versus the two-sided alternative at the 5% significance level?
(b) Answer part (a) for the one-sided alternative that the difference in means is positive.

7.95 You want to compare the daily sales for two different designs of Web pages for your Internet business. You assign the next 60 days to either Design A or Design B, 30 days to each.
(a) Would you use a one-sided or two-sided significance test for this problem? Explain your choice.
(b) If you use Table D to find the critical value, what are the degrees of freedom?
(c) The \( t \) statistic for comparing the mean sales is 2.06. If you use Table D, what \( P \)-value would you report? What would you conclude?

7.96 If you perform the significance test in the previous exercise using level \( \alpha = 0.05 \), how large (positive or negative) must the \( t \) statistic be to reject the null hypothesis that the two designs give the same average sales?

7.97 Do piano lessons improve the spatial-temporal reasoning of preschool children? We examined this question in Exercises 7.29 and 7.30 (page xxx) by analyzing the change in spatial-temporal reasoning of 34 preschool children after six months of piano lessons. Here we examine the same question by comparing the changes of those students with the changes of 44 children in a control group. Here are the data for the children who took piano lessons:

| 2 5 7 -2 2 7 4 1 0 7 3 4 3 4 9 4 5 |
| 2 9 6 0 3 6 -1 3 4 6 7 -2 7 -3 3 4 4 |

The control group scores are

| 1 -1 0 1 -4 0 0 1 0 -1 0 1 1 -3 -2 |
| 4 -1 2 4 2 2 2 -3 -3 0 2 0 -1 3 -1 |
| 5 -1 7 0 4 0 2 1 -6 0 2 -1 0 -2 |

(a) Display the data and summarize the distributions.
(b) Make a table with the sample size, the mean, the standard deviation, and the standard error of the mean for each of the two groups.
(c) Translate the question of interest into hypotheses, test them, and summarize your conclusions.

7.98 Refer to the previous exercise. Give a 95% confidence interval that describes the comparison between the children who took piano lessons and the controls.

7.99 Refer to Exercises 7.29 and 7.30 (page xxx) and the previous two exercises. We have used four ways to address the question of interest. Discuss the relative merits of each approach.
7.100 In what ways are companies that fail different from those that continue to do business? A study compared various characteristics of 68 healthy and 33 failed firms. One of the variables was the ratio of current assets to current liabilities. Roughly speaking, this is the amount that the firm is worth divided by what it owes.22 The data are given in Table 7.6.

(a) Display the data so that the two distributions can be compared. Describe the shapes of the distributions and any important characteristics.
(b) We expect that failed firms will have a lower ratio. Describe and test appropriate hypotheses for these data. What do you conclude?
(c) It is not possible to do a randomized experiment for this kind of question. Explain why.

7.101 Does cocaine use by pregnant women cause their babies to have low birth weight? To study this question, birth weights of babies of women who tested positive for cocaine/crack during a drug-screening test were compared with the birth weights of babies whose mothers either tested negative or were not tested, a group we call “other.” Here are the summary statistics. The birth weights are measured in grams.23

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>x</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive test</td>
<td>134</td>
<td>2733</td>
<td>599</td>
</tr>
<tr>
<td>Other</td>
<td>5974</td>
<td>3118</td>
<td>672</td>
</tr>
</tbody>
</table>

(a) Formulate appropriate hypotheses and carry out the test of significance for these data.
(b) Give a 95% confidence interval for the mean difference in birth weights.
(c) Discuss the limitations of the study design. What do you believe can be concluded from this study?

7.102 The Survey of Study Habits and Attitudes (SSHA) is a psychological test designed to measure the motivation, study habits, and attitudes toward learning of college students. These factors, along with ability, are important in explaining success in school. Scores on the SSHA range from 0 to 200. A selective private college gives the SSHA to an SRS of both male and female first-year students. The data for the women are as follows:

<table>
<thead>
<tr>
<th></th>
<th>154</th>
<th>109</th>
<th>137</th>
<th>115</th>
<th>152</th>
<th>140</th>
<th>154</th>
<th>178</th>
<th>101</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>103</td>
<td>126</td>
<td>126</td>
<td>137</td>
<td>165</td>
<td>165</td>
<td>129</td>
<td>200</td>
<td>148</td>
</tr>
</tbody>
</table>

Here are the scores of the men:

<table>
<thead>
<tr>
<th></th>
<th>108</th>
<th>140</th>
<th>114</th>
<th>91</th>
<th>180</th>
<th>115</th>
<th>126</th>
<th>92</th>
<th>169</th>
<th>146</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>109</td>
<td>132</td>
<td>75</td>
<td>88</td>
<td>113</td>
<td>151</td>
<td>70</td>
<td>115</td>
<td>187</td>
<td>104</td>
</tr>
</tbody>
</table>

(a) Examine each sample graphically, with special attention to outliers and skewness. Is use of a t procedure acceptable for these data?
(b) Most studies have found that the mean SSHA score for men is lower than the mean score in a comparable group of women. Test this supposition here. That is, state hypotheses, carry out the test and obtain a P-value, and give your conclusions.
(c) Give a 90% confidence interval for the mean difference between the SSHA scores of male and female first-year students at this college.

**7.103** Exercise 7.76 (page xxx) compares the DBHs of samples of trees from the eastern and western halves of the Wade Tract. Is it reasonable to use the two-sample pooled procedures with these data? Analyze the data with these methods and compare your results with those you found in Exercise 7.76.

**Section 7.3**

**7.104** The $F$ statistic $F = s_1^2 / s_2^2$ is calculated from samples of size $n_1 = 10$ and $n_2 = 21$. (Remember that $n_1$ is the numerator sample size.)

(a) What is the upper 5% critical value for this $F$?
(b) In a test of equality of standard deviations against the two-sided alternative, this statistic has the value $F = 2.45$. Is this value significant at the 10% level? Is it significant at the 5% level?

**7.105** The $F$ statistic for equality of standard deviations based on samples of sizes $n_1 = 21$ and $n_2 = 26$ takes the value $F = 2.88$.

(a) Is this significant evidence of unequal population standard deviations at the 5% level?
(b) Use Table E to give an upper and a lower bound for the $P$-value.

**7.106** Exercise S7.18 records the results of comparing a measure of rowing style for skilled and novice female competitive rowers. Is there significant evidence of inequality between the standard deviations of the two populations?

(a) State $H_0$ and $H_a$.
(b) Calculate the $F$ statistic. Between which two levels does the $P$-value lie?

**7.107** Answer the same questions for the weights of the two groups, recorded in Exercise S7.19.

**7.108** The observed inequality between the sample standard deviations of male and female SAT mathematics scores in Exercise S7.20 is clearly significant. You can say this without doing any calculations. Find $F$ and look in Table E. Then explain why the significance of $F$ could be seen without arithmetic.

**7.109** An $F$ statistic will be used to compare two variances. The sample sizes are both 20. How large does the ratio of the largest to the smallest variance need to be for the significance test to reject the null hypothesis that the population variances are the same?

**7.110** The $F$ statistic $F = s_1^2 / s_2^2$ is calculated from samples of size $n_1 = 16$ and $n_2 = 20$. (Remember that $n_1$ is the numerator sample size.)

(a) What is the upper 5% critical value for this $F$?
(b) In a test of equality of standard deviations against the two-sided alternative, this statistic has the value $F = 2.71$. Is this value significant at the 5% level? Is it significant at the 1% level?
The $F$ statistic for equality of standard deviations based on samples of sizes $n_1 = 31$ and $n_2 = 28$ takes the value $F = 1.72$.

(a) Is this significant evidence of unequal population standard deviations at the 5\% level?

(b) Use Table E to give an upper and a lower bound for the $P$-value.

Exercise 7.49 compares the rents of one-bedroom and two-bedroom apartments. Is there any evidence in the data that would lead us to conclude that the standard deviations are different? State the appropriate hypotheses, calculate the test statistic, and write a short summary of the results.

A USDA survey used to estimate wheat prices in July and September is described in Exercise 7.52. Using the standard deviations you calculated there, perform the test for equality of standard deviations and summarize your conclusion.

The data for VOTs of children and adults in Exercise 7.58 show quite different sample standard deviations. How statistically significant is the observed inequality?

Suppose that you wanted to compare intramural basketball players and intramural soccer players on the “ego strength” personality factor described in Exercise 7.51. With the data from that exercise, you will use $\sigma = 0.7$ for planning purposes. The pooled two-sample $t$ test with $\alpha = 0.05$ will be used to make the comparison. Based on Exercise 7.69, you judge a difference of 0.5 points to be of interest. Pick several values of $n$ and find the power. Plot the power versus $n$ and use the plot to find a value of $n$ that will give approximately 80\% power. Calculate the power for the value of $n$ that you found.

An $F$ statistic will be used to compare two variances. How large does the ratio of the largest to the smallest variance need to be for the significance test to reject the null hypothesis that the population variances are the same in the following settings? Use the 5\% level of significance.

(a) The two sample sizes are 5.

(b) The two sample sizes are 10.

(c) The two sample sizes are 26.

(d) What do you conclude?

Return to the SSHA data in Exercise 7.84 (page xxx). SSHA scores are generally less variable among women than among men. We want to know whether this is true for this college.

(a) State $H_0$ and $H_a$. Note that $H_a$ is one-sided in this case.

(b) Because Table E contains only upper critical values for $F$, a one-sided test requires that in calculating $F$ the numerator $s^2$ belongs to the group that $H_a$ claims to have the larger $\sigma$. Calculate this $F$.

(c) Compare $F$ to the entries in Table E (no doubling of $p$) to obtain the $P$-value. Be sure the degrees of freedom are in the proper order. What do you conclude about the variation in SSHA scores?

In Exercise 7.82 (page xxx) data on cocaine use and birth weight are summarized. The study has been criticized because of several design problems. Suppose
that you are designing a new study. Based on the results in Exercise 7.82, you think that the true difference in mean birth weights may be about 350 grams (g); a difference this large is clinically important. For planning purposes assume that you will have 75 women in each group and that the common standard deviation is 650 g, a guess that is between the two standard deviations in Exercise 7.82. If you use a pooled two-sample \( t \) test with a Type I error of 0.05, what is the power of the test for this design?

7.119 Refer to the previous exercise. Repeat the power calculation for 20, 40, 60, 80, 100, and 120 women in each group. Plot the power versus the sample size and write a short summary of the results.

7.120 Refer to the previous two exercises. For each of the sample sizes considered, what is your guess at the margin of error for the 95% confidence interval for the difference in mean weights? Display these results with a graph or a sketch.

Chapter 7 Review Exercises

7.121 Data on the numbers of manatees killed by boats each year are given in Exercise S1.11. After a long period of increasing numbers of deaths, the pattern flattens somewhat. In fact, the total for 1990 is 47, less than the total of 50 for 1989. Perhaps the trend has now reversed. We would like to do a significance test to compare these two counts. Theoretical considerations suggest that the standard errors \( \sigma / \sqrt{n} \) for these types of counts can be approximated by the square root of the count. So, for example, the 1990 count, 47, has a standard error that is approximately \( \sqrt{47} \). Use this approximation to perform an approximate two-sample \( z \) test for the difference between the 1989 and 1990 deaths. Find an approximate 95% confidence interval for the difference. What do you conclude?

7.122 In a study of the effectiveness of weight-loss programs, 47 subjects who were at least 20% overweight took part in a group support program for 10 weeks. Private weighings determined each subject’s weight at the beginning of the program and 6 months after the program’s end. The matched pairs \( t \) test was used to assess the significance of the average weight loss. The paper reporting the study said, “The subjects lost a significant amount of weight over time, \( t(46) = 4.68, p < 0.01 \).” It is common to report the results of statistical tests in this abbreviated style. (Based loosely on D. R. Black et al., “Minimal interventions for weight control: a cost-effective alternative,” Addictive Behaviors, 9 (1984), pp. 279–285.)
(a) Why was the matched pairs statistic appropriate?
(b) Explain to someone who knows no statistics but is interested in weight-loss programs what the practical conclusion is.
(c) The paper follows the tradition of reporting significance only at fixed levels such as \( \alpha = 0.01 \). In fact, the results are more significant than “\( p < 0.01 \)” suggests. What can you say about the \( P \)-value of the \( t \) test?

7.123 Nitrites are often added to meat products as preservatives. In a study of the effect of these chemicals on bacteria, the rate of uptake of a radiolabeled amino acid
was measured for a number of cultures of bacteria, some growing in a medium to which nitrites had been added. Here are the summary statistics from this study:

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>$\bar{x}$</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrite</td>
<td>30</td>
<td>7880</td>
<td>115</td>
</tr>
<tr>
<td>Control</td>
<td>30</td>
<td>8112</td>
<td>1250</td>
</tr>
</tbody>
</table>

Carry out a test of the research hypothesis that nitrites decrease amino acid uptake, and report your results.

7.124 The one-hole test is used to test the manipulative skill of job applicants. This test requires subjects to grasp a pin, move it to a hole, insert it, and return for another pin. The score on the test is the number of pins inserted in a fixed time interval. In one study, male college students were compared with experienced female industrial workers. Here are the data for the first minute of the test: (G. Salvendy, “Selection of industrial operators: the one-hole test,” International Journal of Production Research, 13 (1973), pp. 303–321.)

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>$\bar{x}$</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>750</td>
<td>35.12</td>
<td>4.31</td>
</tr>
<tr>
<td>Workers</td>
<td>412</td>
<td>37.32</td>
<td>3.83</td>
</tr>
</tbody>
</table>

(a) It was expected that the experienced workers would outperform the students, at least during the first minute, before learning occurs. State the hypotheses for a statistical test of this expectation and perform the test. Give a $P$-value and state your conclusions.

(b) The distribution of scores is slightly skewed to the left. Explain why the procedure you used in (a) is nonetheless acceptable.

(c) One purpose of the study was to develop performance norms for job applicants. Based on the data above, what is the range that covers the middle 95% of experienced workers? (Be careful! This is not the same as a 95% confidence interval for the mean score of experienced workers.)

(d) The five-number summary of the distribution of scores among the workers is

$$23 \quad 33.5 \quad 37 \quad 40.5 \quad 46$$

for the first minute and

$$32 \quad 39 \quad 44 \quad 49 \quad 59$$

for the fifteenth minute of the test. Display these facts graphically, and describe briefly the differences between the distributions of scores in the first and fifteenth minute.

7.125 The composition of the earth’s atmosphere may have changed over time. One attempt to discover the nature of the atmosphere long ago studies the gas trapped in bubbles inside ancient amber. Amber is tree resin that has hardened and been trapped in rocks. The gas in bubbles within amber should be a sample of the atmosphere at the time the amber was formed. Measurements on specimens of amber from the late Cretaceous era (75 to 95 million years ago) give these percents of nitrogen:

$$63.4 \quad 65.0 \quad 64.4 \quad 63.3 \quad 54.8 \quad 64.5 \quad 60.8 \quad 49.1 \quad 51.0$$
These values are quite different from the present 78.1% of nitrogen in the atmosphere. Assume (this is not yet agreed on by experts) that these observations are an SRS from the late Cretaceous atmosphere. (Data from R. A. Berner and G. P. Landis, "Gas bubbles in fossil amber as possible indicators of the major gas composition of ancient air," Science, 239 (1988), pp. 1406–1409.)

(a) Graph the data, and comment on skewness and outliers.
(b) The \( t \) procedures will be only approximate in this case. Give a 90% \( t \) confidence interval for the mean percent of nitrogen in ancient air.

7.126 Table 1.3 (page xx) gives the number of medical doctors per 100,000 population by state. Is it proper to apply the one-sample \( t \) method to these data to give a 95% confidence interval for the mean number of medical doctors per 100,000 population per state? Explain your answer.

7.127 The amount of lead in a certain type of soil, when released by a standard extraction method, averages 86 parts per million (ppm). A new extraction method is tried on 40 specimens of the soil, yielding a mean of 83 ppm lead and a standard deviation of 10 ppm.

(a) Is there significant evidence at the 5% level that the new method frees less lead from the soil? What about the 1% level?
(b) A critic argues that because of variations in the soil, the effectiveness of the new method is confounded with characteristics of the particular soil specimens used. Briefly describe a better data production design that avoids this criticism.

7.128 High levels of cholesterol in the blood are not healthy in either humans or dogs. Because a diet rich in saturated fats raises the cholesterol level, it is plausible that dogs owned as pets have higher cholesterol levels than dogs owned by a veterinary research clinic. “Normal” levels of cholesterol based on the clinic’s dogs would then be misleading. A clinic compared healthy dogs it owned with healthy pets brought to the clinic to be neutered. (V. D. Bass, W. E. Hoffmann, and J. L. Dorner, “Normal canine lipid profiles and effects of experimentally induced pancreatitis and hepatic necrosis on lipids,” American Journal of Veterinary Research, 37 (1976), pp. 1355–1357.) The summary statistics for blood cholesterol levels (milligrams per deciliter of blood) appear below:

<table>
<thead>
<tr>
<th>Group</th>
<th>( n )</th>
<th>( \bar{x} )</th>
<th>( s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pets</td>
<td>26</td>
<td>193</td>
<td>68</td>
</tr>
<tr>
<td>Clinic</td>
<td>23</td>
<td>174</td>
<td>44</td>
</tr>
</tbody>
</table>

(a) Is there strong evidence that pets have a higher mean cholesterol level than clinic dogs? State the \( H_0 \) and \( H_a \) and carry out an appropriate test. Give the \( P \)-value and state your conclusion.
(b) Give a 95% confidence interval for the difference in mean cholesterol levels between pets and clinic dogs.
(c) Give a 95% confidence interval for the mean cholesterol level in clinic dogs.
(d) What assumptions must be satisfied to justify the procedures you used in (a), (b), and (c)? Assuming that the cholesterol measurements have no outliers and are
not strongly skewed, what is the chief threat to the validity of the results of this study?

7.129 Elite distance runners are thinner than the rest of us. Here are data on skinfold thickness, which indirectly measures body fat, for 20 elite runners and 95 ordinary men in the same age group. (M. L. Pollock et al., “Body composition of elite class distance runners,” in P. Milvey (ed.), The Marathon: Physiological, Medical, Epidemiological, and Psychological Studies, New York Academy of Sciences, 1977, p. 366.) The data are in millimeters and are given in the form “mean (standard deviation).”

<table>
<thead>
<tr>
<th></th>
<th>Runners</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdomen</td>
<td>7.1 (1.0)</td>
<td>20.6 (9.0)</td>
</tr>
<tr>
<td>Thigh</td>
<td>6.1 (1.8)</td>
<td>17.4 (6.6)</td>
</tr>
</tbody>
</table>

Use confidence intervals to describe the difference between runners and typical young men.

7.130 Table 7.3 (page xxx) gives the levels of three pollutants in the exhaust of 46 randomly selected vehicles of the same type. You will investigate emissions of nitrogen oxides (NOX).

(a) Make a stemplot and, if your software allows, a normal quantile plot of the NOX levels. Do the plots suggest that the distribution of NOX emissions is approximately normal? Can you safely employ $t$ procedures to analyze these data?

(b) Give a 95% confidence interval for the mean NOX level in vehicles of this type.

(c) Your supervisor hopes the average NOX level is less than 1 gram per mile. You will have to tell him that it’s not so. Carry out a significance test to assess the strength of the evidence that the mean NOX level is greater than 1, and then write a short report to your supervisor based on your work in (b) and (c). (Your supervisor has never heard of $P$-values, so you must use plain language.)

7.131 Refer to the previous exercise. Take a simple random sample of one-half of the data. Analyze the data for these 112 computer science majors. Compare your results with those you obtained in the previous exercise and comment on the effect of the sample size on these procedures.

7.132 (Optional) In Exercise 7.105 (page xxx) you found the power for a study designed to compare birth weights of children born to cocaine users with those born to controls. Fix the sample size at 50 in each group and assume the standard deviation is 650 grams and the significance level is 0.05. Pick a set of alternatives that will give values of power ranging from fairly low values to fairly high values. Plot your results versus that alternative and give a short summary of what you have found.
CHAPTER 8

Section 8.1

8.1 In each of the following cases state whether or not the normal approximation to the binomial should be used for a significance test on the population proportion $p$.  
(a) $n = 10$ and $H_0: p = 0.4$.  
(b) $n = 100$ and $H_0: p = 0.6$.  
(c) $n = 1000$ and $H_0: p = 0.996$.  
(d) $n = 500$ and $H_0: p = 0.3$.

8.2 The Gallup Poll asked a sample of 1785 U.S. adults, “Did you, yourself, happen to attend church or synagogue in the last 7 days?” Of the respondents, 750 said “Yes.” Suppose (it is not, in fact, true) that Gallup’s sample was an SRS.  
(a) Give a 99% confidence interval for the proportion of all U.S. adults who attended church or synagogue during the week preceding the poll.  
(b) Do the results provide good evidence that less than half of the population attended church or synagogue?  
(c) How large a sample would be required to obtain a margin of error of ±0.01 in a 99% confidence interval for the proportion who attend church or synagogue? (Use Gallup’s result as the guessed value of $p$.)

8.3 Leroy, a starting player for a major college basketball team, made only 38.4% of his free throws last season. During the summer he worked on developing a softer shot in the hope of improving his free-throw accuracy. In the first eight games of this season Leroy made 25 free throws in 40 attempts. Let $p$ be his probability of making each free throw he shoots this season.  
(a) State the null hypothesis $H_0$ that Leroy’s free-throw probability has remained the same as last year and the alternative $H_a$ that his work in the summer resulted in a higher probability of success.  
(b) Calculate the $z$ statistic for testing $H_0$ versus $H_a$.  
(c) Do you accept or reject $H_0$ for $\alpha = 0.05$? Find the $P$-value.  
(d) Give a 90% confidence interval for Leroy’s free-throw success probability for the new season. Are you convinced that he is now a better free-throw shooter than last season?  
(e) What assumptions are needed for the validity of the test and confidence interval calculations that you performed?

8.4 To profitably produce a planned upgrade of a software product you make, you must charge customers $100. Are your customers willing to pay this much? You contact a random sample of 40 customers and find that 11 would pay $100 for the upgrade. Find a 95% confidence interval for the proportion of all of your customers (the population) who would be willing to buy the upgrade for $100.

8.5 In the previous exercise we found that 11 customers from a random sample of 40 would be willing to buy a software upgrade that costs $100. If the upgrade is to be profitable, you will need to sell it to more than 20% of your customers. Do the
sample data give good evidence that more than 20% are willing to buy?
(a) Formulate this problem as a hypothesis test. Give the null and alternative hypotheses. Will you use a one-sided or a two-sided alternative? Why?
(b) Carry out the significance test. Report the test statistic and the $P$-value.
(c) Should you proceed with plans to produce and market the upgrade?

8.6 A poll of 811 adults aged 18 or older asked about purchases that they intended to make for the upcoming holiday season. (The poll is part of the “American Express Retail Index Project” and is reported in *Stores*, December 2000, pp. 38–40.) One of the questions asked about what kind of gift they intended to buy for the person on whom they will spend the most. Clothing was the first choice of 487 people. Give a 99% confidence interval for the proportion of people in this population who intend to buy clothing as their first choice.

8.7 When trying to hire managers and executives, companies sometimes verify the academic credentials described by the applicants. One company that performs these checks summarized its findings for a six-month period. Of the 84 applicants whose credentials were checked, 15 lied about having a degree. (Data provided by Jude M. Werra & Associates, Brookfield, Wisconsin.)
(a) Find the proportion of applicants who lied about having a degree and the standard error.
(b) Consider these data to be a random sample of credentials from a large collection of similar applicants. Give a 95% confidence interval for the true proportion of applicants who lie about having a degree.

8.8 Refer to the previous exercise. Suppose that 10 applicants lied about their major. Can we conclude that a total of 25 = 15 + 10 applicants lied about having a degree or about their major? Explain your answer.

8.9 A question in a Christmas tree market survey was “Did you have a Christmas tree last year?” Of the 500 respondents, 421 answered “Yes.”
(a) Find the sample proportion and its standard error.
(b) Give a 90% confidence interval for the proportion of Indiana households who had a Christmas tree this year.

8.10 Of the 500 respondents in the Christmas tree market survey, 44% had no children at home and 56% had at least one child at home. The corresponding figures for the most recent census are 48% with no children and 52% with at least one child. Test the null hypothesis that the telephone survey technique has a probability of selecting a household with no children that is equal to the value obtained by the census. Give the $z$ statistic and the $P$-value. What do you conclude?

8.11 Refer to the previous exercise. There we arbitrarily chose to state the hypotheses in terms of the proportion of rural respondents. We could as easily have used the proportion of urban respondents.
(a) Write hypotheses in terms of the proportion of urban residents to examine how well the sample represents the state in regard to rural versus urban residence.
(b) Perform the test of significance and summarize the results.
(c) Compare your results with the results of the previous exercise. Summarize and
generalize your conclusion.

**8.12** As part of a quality improvement program, your mail-order company is study-
ing the process of filling customer orders. According to company standards, an order
is shipped on time if it is sent within 3 working days of the time it is received. You
select an SRS of 200 of the 5000 orders received in the past month for an audit. The
audit reveals that 185 of these orders were shipped on time. Find a 95% confidence
interval for the true proportion of the month’s orders that were shipped on time.

**8.13** Large trees growing near power lines can cause power failures during storms
when their branches fall on the lines. Power companies spend a great deal of time
and money trimming and removing trees to prevent this problem. Researchers are
developing hormone and chemical treatments that will stunt or slow tree growth. If
the treatment is too severe, however, the tree will die. In one series of laboratory
experiments on 216 sycamore trees, 41 trees died. Give a 95% confidence interval for
the proportion of sycamore trees that would be expected to die from this particular
treatment.

**8.14** In recent years over 70% of first-year college students responding to a national
survey have identified “being well-off financially” as an important personal goal. A
state university finds that 103 of an SRS of 150 of its first-year students say that
this goal is important. Give a 95% confidence interval for the proportion of all first-
year students at the university who would identify being well-off as an important
personal goal.

**8.15** An entomologist samples a field for egg masses of a harmful insect by placing
a yard-square frame at random locations and carefully examining the ground within
the frame. An SRS of 75 locations selected from a county’s pastureland found egg
masses in 13 locations. Give a 90% confidence interval for the proportion of all
possible locations that are infested.

**8.16** Shereka, a starting player for a major college basketball team, made only 36.2%
of her free throws last season. During the summer she worked on developing a softer
shot in the hope of improving her free-throw accuracy. In the first eight games of
this season Shereka made 22 free throws in 42 attempts. Let \( p \) be her probability of
making each free throw she shoots this season.
(a) State the null hypothesis \( H_0 \) that Shereka’s free-throw probability has remained
the same as last year and the alternative \( H_a \) that her work in the summer resulted
in a higher probability of success.
(b) Calculate the \( z \) statistic for testing \( H_0 \) versus \( H_a \).
(c) Do you accept or reject \( H_0 \) for \( \alpha = 0.05 \)? Find the \( P \)-value.
(d) Give a 90% confidence interval for Shereka’s free-throw success probability for
the new season. Are you convinced that she is now a better free-throw shooter than
last season?
(e) What assumptions are needed for the validity of the test and confidence interval
calculations that you performed?
8.17 Land’s Beginning is a company that sells its merchandise through the mail. It is considering buying a list of addresses from a magazine. The magazine claims that at least 25% of its subscribers have high incomes (they define this to be household income in excess of $100,000). Land’s Beginning would like to estimate the proportion of high-income people on the list. Checking income is very difficult and expensive but another company offers this service. Land’s Beginning will pay to find incomes for an SRS of people on the magazine’s list. They would like the margin of error of the 95% confidence interval for the proportion to be 0.05 or less. Use the guessed value \( p^* = 0.25 \) to find the required sample size.

8.18 Refer to the previous exercise. For each of the following variations on the design specifications, state whether the required sample size will be higher, lower, or the same as that found above.
(a) Use a 90% confidence interval.
(b) Change the allowable margin of error to 0.10.
(c) Use a planning value of \( p^* = 0.30 \).
(d) Use a different company to do the income checks.

8.19 A student organization wants to start a nightclub for students under the age of 21. To assess support for this proposal, they will select an SRS of students and ask each respondent if he or she would patronize this type of establishment. They expect that about 60% of the student body would respond favorably. What sample size is required to obtain a 95% confidence interval with an approximate margin of error of 0.08? Suppose that 50% of the sample responds favorably. Calculate the margin of error of the 95% confidence interval.

8.20 In each of the following circumstances state whether you would use the large-sample confidence interval, the plus four method, or neither for a 95% confidence interval.
(a) \( n = 20, X = 15 \).
(b) \( n = 100, X = 15 \).
(c) \( n = 10, X = 2 \).
(d) \( n = 5, X = 2 \).
(e) \( n = 50, X = 20 \).

8.21 In each of the following circumstances state whether you would use the large-sample confidence interval, the plus four method, or neither for a 95% confidence interval.
(a) \( n = 8, X = 4 \).
(b) \( n = 1000, X = 12 \).
(c) \( n = 40, X = 18 \).
(d) \( n = 15, X = 2 \).
(e) \( n = 500, X = 225 \).

8.22 Explain what is wrong with each of the following:
(a) An approximate 95% confidence interval for an unknown proportion \( p \) is \( \hat{p} \) plus or minus its standard error.
(b) You can use a significance test to evaluate the hypothesis \( H_0: \hat{p} = 0.3 \) versus the
two-sided alternative.
(c) The large-sample significance test for a population proportion is based on a \( t \) statistic.

8.23 Dogs are big and expensive. Rats are small and cheap. Can rats be trained to replace dogs in sniffing out illegal drugs? One study trained six male albino Sprague-Dawley rats to rear up on their hind legs in response to the smell of cocaine. After training, each rat was tested 80 times. In the test a rat was presented with a large number of cups, one of which smelled like cocaine. A success was recorded if the rat correctly identified the cup containing cocaine by rearing up in front of it. The numbers of successes for the six rats were 80, 80, 73, 80, 74, and 80. You want to estimate the success rate in the future for each of the six rats. Compare the use of the large-sample estimates with the plus four estimates for this problem and make a recommendation concerning which is better. Write a short summary giving reasons for your recommendation.

8.24 The National Congregations Study collected data in a one-hour interview with a key informant—that is, a minister, priest, rabbi, or other staff person or leader. One question asked concerned the length of the typical sermon. For this question 390 out of 1191 congregations reported that the typical sermon lasted more than 30 minutes.
(a) Use the large-sample inference procedures to estimate the true proportion for this question with a 95% confidence interval.
(b) Compute the interval using the plus four method. Compare these results with those from part (a) and summarize what this example tells you about the two methods.
(c) There were 1236 congregations surveyed in this study. Calculate the nonresponse rate for this question. Does this influence how you interpret the results? Write a short discussion of this issue.
(d) The respondents to this question were not asked to use a stopwatch to record the lengths of a random sample of sermons at their congregations. They responded based on their impressions of the sermons. Do you think that ministers, priests, rabbis, or other staff persons or leaders might perceive sermon lengths differently from the people listening to the sermons? Discuss how your ideas would influence your interpretation of the results of this study.

8.25 The study described in the previous exercise also asked each respondent to classify his or her congregation according to theological orientation. For this question, 707 out of 1191 congregations were classified as “more conservative.” Using the questions in the previous exercise as a guide, analyze and interpret these data. Compare your answers to parts (c) and (d) and discuss reasons why you think the answers should be similar or different.

8.26 A survey of 1280 student loan borrowers found that 448 had loans totaling more than $20,000 for their undergraduate education. Give a 95% confidence interval for the proportion of all student loan borrowers who have loans of $20,000 or more for their undergraduate education.
8.27 In the survey described in the previous exercise, there were 1050 borrowers whose total debt was $10,000 or more. Of these, 192 left school without completing a degree. Consider the population to be borrowers whose total debt was $10,000 or more. Find a 95% confidence interval for the proportion of borrowers who left school without completing a degree in this population.

8.28 Refer to Exercise 8.11. Would a 99% confidence interval be wider or narrower than the one that you found in that exercise? Verify your results by computing the interval.

8.29 Refer to Exercise 8.12. Would a 90% confidence interval be wider or narrower than the one that you found in that exercise? Verify your results by computing the interval.

8.30 Yesterday, your top salesperson called on 10 customers and obtained orders for your new product from all 10. Suppose that it is reasonable to view these 10 customers as a random sample of all of her customers.
(a) Give the plus four estimate of the proportion of her customers who would buy the new product. Notice that we don’t estimate that all customers will buy, even though all 10 in the sample did.
(b) Give the margin of error for 95% confidence. (You may see that the upper endpoint of the confidence interval is greater than 1. In that case, take the upper endpoint to be 1.)
(c) Do the results apply to all of your sales force? Explain why or why not.

8.31 In each of the following cases state whether or not the normal approximation to the binomial should be used for a significance test on the population proportion \( p \).
(a) \( n = 30 \) and \( H_0: p = 0.3 \).
(b) \( n = 30 \) and \( H_0: p = 0.6 \).
(c) \( n = 1000 \) and \( H_0: p = 0.5 \).
(d) \( n = 500 \) and \( H_0: p = 0.01 \).

8.32 You are planning an evaluation of an alcohol awareness program at your college that will take place six months after the program. Previous evaluations indicate that about 30% of the students who participate will respond “Yes” to the question “Do you think your behavior toward alcohol consumption has changed since the program?” How large a sample should you take if you want the margin of error for 95% confidence to be about 0.1?

8.33 An automobile manufacturer would like to know what proportion of its customers are dissatisfied with the service received from their local dealer. The customer relations department will survey a random sample of customers and compute a 95% confidence interval for the proportion that are dissatisfied. From past studies, they believe that this proportion will be about 0.25. Find the sample size needed if the margin of error of the confidence interval is to be about 0.02. Suppose 15% of the sample say that they are dissatisfied. What is the margin of error of the 95% confidence interval?
8.34 You have been asked to survey students at a large college to determine the proportion who favor an increase in student fees to support an expansion of the student newspaper. Each student will be asked whether he or she is in favor of the proposed increase. Using records provided by the registrar you can select a random sample of students from the college. After careful consideration of your resources, you decide that it is reasonable to conduct a study with a sample of 10 students. For this sample size, construct a table of the margins of error for 95% confidence intervals when \( \hat{p} \) takes the values 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9.

8.35 A former editor of the student newspaper agrees to underwrite the study in the previous exercise because she believes the results will demonstrate that most students support an increase in fees. She is willing to provide funds for a sample of size 400. Write a short summary for your benefactor of why the increased sample size will provide better results.

Section 8.2

8.36 In the 1996 regular baseball season, the World Series Champion New York Yankees played 80 games at home and 82 games away. They won 49 of their home games and 43 of the games played away. We can consider these games as samples from potentially large populations of games played at home and away. How much advantage does the Yankee home field provide?
(a) Find the proportion of wins for the home games. Do the same for the away games.
(b) Find the standard error needed to compute a confidence interval for the difference in the proportions.
(c) Compute a 90% confidence interval for the difference between the probability that the Yankees win at home and the probability that they win when on the road. Are you convinced that the 1996 Yankees were more likely to win at home?

8.37 Return to the New York Yankees baseball data in the previous exercise.
(a) Combining all of the games played, what proportion did the Yankees win?
(b) Find the standard error needed for testing that the probability of winning is the same at home and away.
(c) Most people think that it is easier to win at home than away. Formulate null and alternative hypotheses to examine this idea.
(d) Compute the \( z \) statistic and its \( P \)-value. What conclusion do you draw?

8.38 The 1958 Detroit Area Study was an important sociological investigation of the influence of religion on everyday life. It is described in Gerhard Lenski, *The Religious Factor*, Doubleday, New York, 1961. The sample “was basically a simple random sample of the population of the metropolitan area.” Of the 656 respondents, 267 were white Protestants and 230 were white Catholics. One question asked whether the government was doing enough in areas such as housing, unemployment, and education; 161 of the Protestants and 136 of the Catholics said “No.” Is there evidence that white Protestants and white Catholics differed on this issue?
8.39 The respondents in the Detroit Area Study (see the previous exercise) were also asked whether they believed that the right of free speech included the right to make speeches in favor of communism. Of the white Protestants, 104 said “Yes,” while 75 of the white Catholics said “Yes.” Give a 95% confidence interval for the amount by which the proportion of Protestants who agreed that communist speeches are protected exceeds the proportion of Catholics who held this opinion.

8.40 A university financial aid office polled an SRS of undergraduate students to study their summer employment. Not all students were employed the previous summer. Here are the results for men and women:

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed</td>
<td>718</td>
<td>593</td>
</tr>
<tr>
<td>Not employed</td>
<td>79</td>
<td>139</td>
</tr>
<tr>
<td>Total</td>
<td>797</td>
<td>732</td>
</tr>
</tbody>
</table>

(a) Is there evidence that the proportion of male students employed during the summer differs from the proportion of female students who were employed? State $H_0$ and $H_a$, compute the test statistic, and give its $P$-value.
(b) Give a 99% confidence interval for the difference between the proportions of male and female students who were employed during the summer. Does the difference seem practically important to you?

8.41 Refer to the study of undergraduate student summer employment described in the previous exercise. Similar results from a smaller number of students may not have the same statistical significance. Specifically, suppose that 72 of 80 men surveyed were employed and 59 of 73 women surveyed were employed. The sample proportions are essentially the same as in the earlier exercise.

(a) Compute the $z$ statistic for these data and report the $P$-value. What do you conclude?
(b) Compare the results of this significance test with your results in Exercise 8.42. What do you observe about the effect of the sample size on the results of these significance tests?

8.42 The power takeoff driveline on farm tractors is a potentially serious hazard to farmers. A shield covers the driveline on new tractors, but for a variety of reasons, the shield is often missing on older tractors. Two types of shield are the bolt-on and the flip-up. A study initiated by the National Safety Council took a sample of older tractors to examine the proportions of shields removed. The study found that 35 shields had been removed from the 83 tractors having bolt-on shields and that 15 had been removed from the 136 tractors with flip-up shields. (Data from W. E. Sell and W. E. Field, “Evaluation of PTO master shield usage on John Deere tractors,” paper presented at the American Society of Agricultural Engineers 1984 Summer Meeting.)

(a) Test the null hypothesis that there is no difference between the proportions of the two types of shields removed. Give the $z$ statistic and the $P$-value. State your conclusion in words.
(b) Give a 90% confidence interval for the difference in the proportions of removed
shields for the bolt-on and the flip-up types. Based on the data, what recommendation would you make about the type of shield to be used on new tractors?

8.43 Is lying about credentials by job applicants changing? In Exercise 8.7 we looked at the proportion of applicants who lied about having a degree in a six-month period. To see if there is a change over time, we can compare that period with the following six months. Here are the data:

<table>
<thead>
<tr>
<th>Period</th>
<th>n</th>
<th>X (lied)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>84</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>106</td>
<td>21</td>
</tr>
</tbody>
</table>

Use a 95% confidence interval to address the question of interest.

8.44 Data on the proportion of applicants who lied about having a degree in two consecutive six-month periods are given in the previous exercise. Formulate appropriate null and alternative hypotheses that can be addressed with these data, carry out the significance test, and summarize the results.

8.45 In a Christmas tree market survey, respondents who had a tree during the holiday season were asked whether the tree was natural or artificial. Respondents were also asked if they lived in an urban area or in a rural area. Of the 421 households displaying a Christmas tree, 160 lived in rural areas and 261 were urban residents. The tree growers want to know if there is a difference in preference for natural trees versus artificial trees between urban and rural households. Here are the data:

<table>
<thead>
<tr>
<th>Population</th>
<th>n</th>
<th>X (natural)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (rural)</td>
<td>160</td>
<td>64</td>
</tr>
<tr>
<td>2 (urban)</td>
<td>261</td>
<td>89</td>
</tr>
</tbody>
</table>

(a) Give the null and alternative hypotheses that are appropriate for this problem assuming that we have no prior information suggesting that one population would have a higher preference than the other.

(b) Test the null hypothesis. Give the test statistic and the P-value, and summarize the results.

(c) Give a 95% confidence interval for the difference in proportions.

8.46 In the 2000 regular baseball season, the World Series Champion New York Yankees played 80 games at home and 81 games away. They won 44 of their home games and 43 of the games played away. We can consider these games as samples from potentially large populations of games played at home and away. How much advantage does the Yankee home field provide?

(a) Find the Wilson estimate of proportion of wins for all home games. Do the same for away games.

(b) Find the standard error needed to compute a confidence interval for the difference in the proportions.

(c) Compute a 90% confidence interval for the difference between the probability that the Yankees win at home and the probability that they win when on the road. Are you convinced that the Yankees were more likely to win at home in 2000?
Section 8.2

8.47 Refer to the New York Yankees baseball data in the previous exercise.
(a) Combining all of the games played, what proportion did the Yankees win?
(b) Find the standard error needed for testing that the probability of winning is the same at home and away.
(c) Most people think that it is easier to win at home than away. Formulate null and alternative hypotheses to examine this idea.
(d) Compute the $z$ statistic and its $P$-value. What conclusion do you draw?

8.48 In the 2000 World Series the New York Yankees played the New York Mets. The previous two exercises examine the Yankees’ home and away victories. During the regular season the Mets won 55 of the 84 home games that they played and 39 of the 81 games that they played away. Perform the same analyses for the Mets and write a short summary comparing these results with those you found for the Yankees.

8.49 The state agriculture department asked random samples of Indiana farmers in each county whether they favored a mandatory corn checkoff program to pay for corn product marketing and research. In Tippecanoe County, 263 farmers were in favor of the program and 252 were not. In neighboring Benton County, 260 were in favor and 377 were not.
(a) Find the proportions of farmers in favor of the program in each of the two counties.
(b) Find the standard error needed to compute a confidence interval for the difference in the proportions.
(c) Compute a 95% confidence interval for the difference between the proportions of farmers favoring the program in Tippecanoe County and in Benton County. Do you think opinions differed in the two counties?

8.50 Return to the survey of farmers described in the previous exercise.
(a) Combine the two samples and find the overall proportion of farmers who favor the corn checkoff program.
(b) Find the standard error needed for testing that the population proportions of farmers favoring the program are the same in the two counties.
(c) Formulate null and alternative hypotheses for comparing the two counties.
(d) Compute the $z$ statistic and its $P$-value. What conclusion do you draw?

8.51 A study of chromosome abnormalities and criminality examined data on 4124 Danish males born in Copenhagen. (H. A. Witkin et al., “Criminality in XYY and XXY men,” Science, 193 (1976), pp. 547–555.) The study used the penal registers maintained in the offices of the local police chiefs and classified each man as having a criminal record or not. Each was also classified as having the normal male XY chromosome pair or one of the abnormalities XYY or XXY. Of the 4096 men with normal chromosomes, 381 had criminal records, while 8 of the 28 men with chromosome abnormalities had criminal records. Some experts believe that chromosome abnormalities are associated with increased criminality. Do these data lend support to this belief? Report your analysis and draw a conclusion.
8.52 A university financial aid office polled an SRS of undergraduate students to study their summer employment. Not all students were employed the previous summer. Here are the results for men and women:

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed</td>
<td>728</td>
<td>603</td>
</tr>
<tr>
<td>Not employed</td>
<td>89</td>
<td>149</td>
</tr>
<tr>
<td>Total</td>
<td>817</td>
<td>752</td>
</tr>
</tbody>
</table>

(a) Is there evidence that the proportion of male students employed during the summer differs from the proportion of female students who were employed? State $H_0$ and $H_a$, compute the test statistic, and give its $P$-value.

(b) Give a 95% confidence interval for the difference between the proportions of male and female students who were employed during the summer. Does the difference seem practically important to you?

8.53 Refer to the study of undergraduate student summer employment described in the previous exercise. Similar results from a smaller number of students may not have the same statistical significance. Specifically, suppose that 73 of 82 men surveyed were employed and 60 of 75 women surveyed were employed. The sample proportions are essentially the same as in the earlier exercise.

(a) Compute the $z$ statistic for these data and report the $P$-value. What do you conclude?

(b) Compare the results of this significance test with your results in Exercise 8.49. What do you observe about the effect of the sample size on the results of these significance tests?

8.54 A clinical trial examined the effectiveness of aspirin in the treatment of cerebral ischemia (stroke). Patients were randomized into treatment and control groups. The study was double-blind in the sense that neither the patients nor the physicians who evaluated the patients knew which patients received aspirin and which the placebo tablet. (William S. Fields et al., “Controlled trial of aspirin in cerebral ischemia,” *Stroke*, 8 (1977), pp. 301–315.) After six months of treatment, the attending physicians evaluated each patient’s progress as either favorable or unfavorable. Of the 78 patients in the aspirin group, 63 had favorable outcomes; 43 of the 77 control patients had favorable outcomes.

(a) Compute the sample proportions of patients having favorable outcomes in the two groups.

(b) Give a 90% confidence interval for the difference between the favorable proportions in the treatment and control groups.

(c) The physicians conducting the study had concluded from previous research that aspirin was likely to increase the chance of a favorable outcome. Carry out a significance test to confirm this conclusion. State hypotheses, find the $P$-value, and write a summary of your results.

8.55 The pesticide diazinon is in common use to treat infestations of the German cockroach, *Blattella germanica*. A study investigated the persistence of this pesticide on various types of surfaces. (Elray M. Roper and Charles G. Wright, “German cockroach (Orthoptera: Blatellidae) mortality on various surfaces following appli-
cation of diazinon,” *Journal of Economic Entomology*, 78 (1985), pp. 733–737.) Researchers applied a 0.5% emulsion of diazinon to glass and plasterboard. After 14 days, they placed 18 cockroaches on each surface and recorded the number that died within 48 hours. On glass, 9 cockroaches died, while on plasterboard, 13 died.

(a) Calculate the mortality rates (sample proportion that died) for the two surfaces.
(b) Find a 95% confidence interval for the difference in the two population proportions.
(c) Chemical analysis of the residues of diazinon suggests that it may persist longer on plasterboard than on glass because it binds to the paper covering on the plasterboard. The researchers therefore expected the mortality rate to be greater on plasterboard than on glass. Conduct a significance test to assess the evidence that this is true.

8.56 Suppose that the experiment of the previous exercise placed more cockroaches on each surface and observed similar mortality rates. Specifically, suppose that 36 cockroaches were placed on each surface and that 26 died on the plasterboard, while 18 died on the glass.

(a) Compute the $z$ statistic for these data and report its $P$-value. What do you conclude?
(b) Compare the results of this significance test with those you gave in Exercise 8.51. What do you observe about the effect of the sample size on the results of these significance tests?

8.57 In each of the following circumstances state whether you would use the large-sample confidence interval, the plus four method, or neither for a 95% confidence interval.

(a) $n_1 = 25$, $n_2 = 25$, $X_1 = 10$, and $X_2 = 15$.
(b) $n_1 = 5$, $n_2 = 5$, $X_1 = 2$, and $X_2 = 5$.
(c) $n_1 = 25$, $n_2 = 25$, $X_1 = 8$, and $X_2 = 20$.
(d) $n_1 = 4$, $n_2 = 8$, $X_1 = 2$, and $X_2 = 7$.
(e) $n_1 = 100$, $n_2 = 10$, $X_1 = 40$, and $X_2 = 2$.

8.58 In each of the following circumstances state whether you would use the large-sample confidence interval, the plus four method, or neither for a 95% confidence interval.

(a) $n_1 = 4$, $n_2 = 100$, $X_1 = 1$, and $X_2 = 65$.
(b) $n_1 = 500$, $n_2 = 300$, $X_1 = 175$, and $X_2 = 208$.
(c) $n_1 = 6$, $n_2 = 10$, $X_1 = 4$, and $X_2 = 2$.
(d) $n_1 = 60$, $n_2 = 55$, $X_1 = 24$, and $X_2 = 37$.
(e) $n_1 = 200$, $n_2 = 100$, $X_1 = 128$, and $X_2 = 94$.

8.59 Suppose there are two binomial populations. For the first, the true proportion of successes is 0.4; for the second, it is 0.5. Consider taking independent samples from these populations, 50 from the first and 60 from the second.

(a) Find the mean and the standard deviation of the distribution of $\hat{p}_1 - \hat{p}_2$.
(b) This distribution is approximately normal. Sketch this normal distribution and mark the location of the mean.
(c) Find a value $d$ for which the probability is 0.95 that the difference in sample proportions is within $\pm d$. Mark these values on your sketch.

8.60 Refer to Exercise 8.35. Redo the exercise in terms of the proportions of men in each classification. Explain how you could have obtained these results from the calculations you did in Exercise 8.35.

8.61 Refer to Exercise 8.37. Redo the exercise in terms of the proportions of men in each classification. Explain how you could have obtained these results from the calculations you did in Exercise 8.37.

8.62 A study was designed to find reasons why patients leave a health maintenance organization (HMO). Patients were classified as to whether or not they had filed a complaint with the HMO. We want to compare the proportion of complainers who leave the HMO with the proportion of those who do not file complaints but who also leave the HMO. In the year of the study, 639 patients filed complaints, and 54 of these patients left the HMO voluntarily. For comparison, the HMO chose an SRS of 743 patients who had not filed complaints. Twenty-two of these patients left voluntarily. Give an estimate of the difference in the two proportions with a 95% confidence interval.

8.63 In the previous exercise you examined data from a study designed to find reasons why patients leave an HMO. There you compared the proportion of complainers who leave the HMO with the proportion of noncomplainers who leave. In the year of the study, 639 patients filed complaints and 54 of these patients left the HMO voluntarily. For comparison, the HMO chose an SRS of 743 patients who had not filed complaints. Twenty-two of those patients left voluntarily. We expect a higher proportion of complainers to leave. Do the data support this expectation? State hypotheses, find the test statistic and its $P$-value, and state your conclusion.

8.64 Exercise 8.47 addresses a question about gender bias with a confidence interval. Set up the problem as a significance test. Carry out the test and summarize the results.

8.65 The proportions of female and male fatally injured bicyclists were compared with a confidence interval in Exercise 8.48. Examine the same data with a test of significance.

Chapter 8 Review Exercises

8.66 Many colleges that once enrolled only male or only female students have become coeducational. Some administrators and alumni were concerned that the academic standards of the institutions would decrease with the change. One formerly all-male college undertook a study of the first class to contain women. The class consisted of 851 students, 214 of whom were women. An examination of first-semester grades revealed that 15 of the top 30 students were female.

(a) What is the proportion of women in the class? Call this value $p_0$.
(b) Assume that the number of females in the top 30 is approximately a binomial
random variable with \( n = 30 \) and unknown probability \( p \) of success. In this case success corresponds to the student being female. What is the value of \( \hat{p} \)?

(c) Are women more likely to be top students than their proportion in the class would suggest? State hypotheses that ask this question, carry out a significance test, and report your conclusion.

8.67 In the Section 6.1 we studied the effect of the sample size on the margin of error of the confidence interval for a single proportion. In this exercise we perform some calculations to observe this effect for the two-sample problem. As in the exercise above, suppose that \( \hat{p}_1 = 0.6 \), \( \hat{p}_2 = 0.4 \), and \( n \) represents the common value of \( n_1 \) and \( n_2 \). Compute the 95% confidence intervals for the difference in the two proportions for \( n = 15, 25, 50, 75, 100, \) and 500. For each interval calculate the margin of error. Summarize and explain your results.

8.68 For a single proportion the margin of error of a confidence interval is largest for any given sample size \( n \) and confidence level \( C \) when \( \hat{p} = 0.5 \). This led us to use \( p^* = 0.5 \) for planning purposes. The same kind of result is true for the two-sample problem. The margin of error of the confidence interval for the difference between two proportions is largest when \( \hat{p}_1 = \hat{p}_2 = 0.5 \). Use these conservative values in the following calculations, and assume that the sample sizes \( n_1 \) and \( n_2 \) have the common value \( n \). Calculate the margins of error of the 99% confidence intervals for the difference in two proportions for the following choices of \( n \): 10, 30, 50, 100, 200, and 500. Present the results in a table or with a graph. Summarize your conclusions.

8.69 You are planning a survey in which a 90% confidence interval for the difference between two proportions will present the results. You will use the conservative guessed value 0.5 for \( \hat{p}_1 \) and \( \hat{p}_2 \) in your planning. You would like the margin of error of the confidence interval to be less than or equal to 0.1. It is very difficult to sample from the first population, so that it will be impossible for you to obtain more than 20 observations from this population. Taking \( n_1 = 20 \), can you find a value of \( n_2 \) that will guarantee the desired margin of error? If so, report the value; if not, explain why not.

8.70 “The nature of work is changing at whirlwind speed. Perhaps now more than ever before, job stress poses a threat to the health of workers and, in turn, to the health of organizations.” (National Institute for Occupational Safety and Health, *Stress at Work*, 2000, [www.cdc.gov/niosh/stresswk.html](http://www.cdc.gov/niosh/stresswk.html).) So says the National Institute for Occupational Safety and Health. Employers are concerned about the effect of stress on their employees. Stress can lower morale and efficiency and increase medical costs. A large survey of restaurant employees found that 75% reported that work stress had a negative impact on their personal lives. (Results of this survey were reported in *Restaurant Business*, September 15, 1999, pp. 45–49.) The human resources manager of a chain of restaurants is concerned that work stress may be affecting the chain’s employees. She asks a random sample of 100 employees to respond Yes or No to the question “Does work stress have a negative impact on your personal life?” Of these, 68 say “Yes.” Give a 95% confidence interval for the proportion of employees who work for this chain of restaurants who believe that work stress has a negative impact on their personal lives.
8.71 Refer to the previous exercise. Is there evidence to conclude that the proportion for this chain of restaurants differs from the value given for the national survey? For this exercise, assume that there is no error associated with the estimate for the national survey.

8.72 A Gallup Poll used telephone interviews to survey a sample of 1025 U.S. residents over the age of 18 regarding their use of credit cards. (Based on a Gallup poll conducted April 6–8, 2001.) The poll reported that 76% of Americans said that they had at least one credit card. Give the 95% margin of error for this estimate.

8.73 The Gallup Poll in the previous exercise reported that 41% of those who have credit cards do not pay the full balance each month. Find the number of people in the survey who said that they had at least one credit card, using the information in the previous exercise. Combine this number with the reported 41% to give a margin of error for the proportion of credit card owners who do not pay their full balance.

8.74 A television news program conducts a call-in poll about a proposed city ban on handgun ownership. Of the 2372 calls, 1921 oppose the ban. The station, following recommended practice, makes a confidence statement: “81% of the Channel 13 Pulse Poll sample opposed the ban. We can be 95% confident that the true proportion of citizens opposing a handgun ban is within 1.6% of the sample result.” Is this conclusion justified?

8.75 Eleven percent of the products produced by an industrial process over the past several months fail to conform to the specifications. The company modifies the process in an attempt to reduce the rate of nonconformities. In a trial run, the modified process produces 16 nonconforming items out of a total of 300 produced. Do these results demonstrate that the modification is effective? Support your conclusion with a clear statement of your assumptions and the results of your statistical calculations.

8.76 In the setting of the previous exercise, give a 95% confidence interval for the proportion of nonconforming items for the modified process. Then, taking \( p_0 = 0.11 \) to be the old proportion and \( p \) the proportion for the modified process, give a 95% confidence interval for \( p - p_0 \).

8.77 In a study on blood pressure and diet, a random sample of Seventh-Day Adventists were interviewed at a national meeting. Because many people who belong to this denomination are vegetarians, they are a very useful group for studying the effects of a meatless diet. (Data provided by Chris Melby and David Goldflies, Department of Physical Education, Health, and Recreation Studies, Purdue University.) Blacks in the population as a whole have a higher average blood pressure than whites. A study of this type should therefore take race into account in the analysis. The 312 people in the sample were categorized by race and whether or not they were vegetarians. The data are given in the following table:

<table>
<thead>
<tr>
<th></th>
<th>Black</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetarian</td>
<td>42</td>
<td>135</td>
</tr>
<tr>
<td>Not vegetarian</td>
<td>47</td>
<td>88</td>
</tr>
</tbody>
</table>
Are the proportions of vegetarians the same among all black and white Seventh-Day Adventists who attended this meeting? Analyze the data, paying particular attention to this question. Summarize your analysis and conclusions. What can you infer about the proportions of vegetarians among black and white Seventh-Day Adventists in general? What about blacks and whites in general?

8.78 A study that examined the association between high blood pressure and increased risk of death from cardiovascular disease. There were 2676 men with low blood pressure and 3338 men with high blood pressure. In the low-blood-pressure group, 21 men died from cardiovascular disease; in the high-blood-pressure group, 55 died.

(a) Compute the 95% confidence interval for the difference in proportions.
(b) Do the study data confirm that death rates are higher among men with high blood pressure? State hypotheses, carry out a significance test, and give your conclusions.

8.79 An experiment designed to assess the effects of aspirin on cardiovascular disease studied 5139 male British medical doctors. The doctors were randomly assigned to two groups. One group of 3429 doctors took one aspirin daily, and the other group did not take aspirin. After 6 years, there were 148 deaths from heart attack or stroke in the first group and 79 in the second group. A similar experiment used male American medical doctors as subjects. These doctors were also randomly assigned to one of two groups. The 11,037 doctors in the first group took one aspirin every other day, and the 11,034 doctors in the second group took no aspirin. After nearly 5 years, there were 104 deaths from heart attacks in the first group and 189 in the second. (The first study is reported in an article in the New York Times of January 30, 1988; the second was described in the New York Times on January 27, 1988.) Analyze the data from these two studies and summarize the results. How do the conclusions of the two studies differ, and why?

8.80 Refer to Exercise 8.43. If you have not already done so, give a 95% confidence interval for the difference in the proportions of the two types of companies that offer stock options. Then compare the two types of companies using a significance test. Be sure to state your hypotheses, the test statistic and the P-value. Write a short summary of your conclusions.

8.81 A Gallup poll used telephone interviews to survey a sample of 1006 U.S. residents over the age of 18 regarding their ideal family size. The poll reported that 38% of Americans said that their ideal family would include three or more children. Assuming that this is an SRS of U.S. residents over the age of 18, give the 95% margin of error for this estimate.

8.82 Many new products introduced into the market are targeted toward children. The choice behavior of children with regard to new products is of particular interest to companies that design marketing strategies for these products. As part of one study, children in different age groups were compared on their ability to sort new products into the correct product category (milk or juice). Here are some of the data:
<table>
<thead>
<tr>
<th>Age group</th>
<th>n</th>
<th>Number who sorted correctly</th>
</tr>
</thead>
<tbody>
<tr>
<td>4- to 5-year-olds</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>6- to 7-year-olds</td>
<td>53</td>
<td>28</td>
</tr>
</tbody>
</table>

Test the null hypothesis that the two age groups are equally skilled at sorting. Justify your choice of an alternative hypothesis. Also, give a 90% confidence interval for the difference. Summarize your results in a short paragraph.

**8.83** The Gallup poll in Exercise 8.63 reported that in a similar poll in 1973, 43% of Americans said that their ideal family would include three or more children. Give the margin of error for this poll, assuming that the sample size was the same. Then compare the proportions with a significance test and give a 95% confidence interval for the difference. Write a summary of your results.

**8.84** In a random sample of 875 students from a large public university, it was found that 411 of the students changed majors during their college years.
(a) Give a 95% confidence interval for the proportion of students at this university who change majors.
(b) Express your results from (a) in terms of the percent of students who change majors.
(c) University officials concerned with counseling students are interested in the number of students who change majors rather than the proportion. The university has 37,000 undergraduate students. Convert the confidence interval you found in (a) to a confidence interval for the number of students who change majors during their college years.

**8.85** Gastric freezing was once a recommended treatment for ulcers in the upper intestine. A randomized comparative experiment found that 28 of the 82 patients who were subjected to gastric freezing improved, while 30 of the 78 patients in the control group improved.
(a) State the appropriate null hypothesis and a two-sided alternative. Carry out a z test. What is the P-value?
(b) What do you conclude about the effectiveness of gastric freezing as a treatment for ulcers? (See Example 3.5 on page xxx for a discussion of gastric freezing.)

**8.86** In this exercise we examine the effect of the sample size on the significance test for comparing two proportions. In each case suppose that \( \hat{p}_1 = 0.6 \) and \( \hat{p}_2 = 0.5 \), and take \( n \) to be the common value of \( n_1 \) and \( n_2 \). Use the \( z \) statistic to test \( H_0: p_1 = p_2 \) versus the alternative \( H_a: p_1 \neq p_2 \). Compute the statistic and the associated P-value for the following values of \( n \): 10, 20, 40, 50, 80, 100, 500, and 1000. Summarize the results in a table. Explain what you observe about the effect of the sample size on statistical significance when the sample proportions \( \hat{p}_1 \) and \( \hat{p}_2 \) are unchanged.

**8.87** In the first section of this chapter, we studied the effect of the sample size on the margin of error of the confidence interval for a single proportion. In this exercise we perform some calculations to observe this effect for the two-sample problem. Suppose that \( \hat{p}_1 = 0.6 \) and \( \hat{p}_2 = 0.5 \), and \( n \) represents the common value of \( n_1 \) and \( n_2 \). Compute the 95% margins of error for the difference in the two proportions for \( n = 10, 20, 40, 50, 80, 100, 500, \) and 1000. Be sure to use the plus four method.
where appropriate. Present the results in a table and with a graph. Write a short summary of your findings.

8.88 For a single proportion the margin of error of a confidence interval is largest for any given sample size \( n \) and confidence level \( C \) when \( \hat{p} = 0.5 \). This led us to use \( p^* = 0.5 \) for planning purposes. The same kind of result is true for the two-sample problem. The margin of error of the confidence interval for the difference between two proportions is largest when \( \hat{p}_1 = \hat{p}_2 = 0.5 \). Use these conservative values in the following calculations, and assume that the sample sizes \( n_1 \) and \( n_2 \) have the common value \( n \). Calculate the margins of error of the 95% confidence intervals for the difference in two proportions for the following choices of \( n \): 10, 20, 40, 50, 80, 100, 500, and 1000. Be sure to use the plus four method where appropriate. Present the results in a table and with a graph. Summarize your conclusions.

8.89 As the previous problem noted, using the guessed value 0.5 for both \( \hat{p}_1 \) and \( \hat{p}_2 \) gives a conservative margin of error in confidence intervals for the difference between two population proportions. You are planning a survey and will calculate a 95% confidence interval for the difference in two proportions when the data are collected. You would like the margin of error of the interval to be less than or equal to 0.10. You will use the same sample size \( n \) for both populations.

(a) How large a value of \( n \) is needed?

(b) Give a general formula for \( n \) in terms of the desired margin of error \( m \) and the critical value \( z^* \).
9.1 Investors use many “indicators” in their attempts to predict the behavior of the stock market. One of these is the “January indicator.” Some investors believe that if the market is up in January, then it will be up for the rest of the year. On the other hand, if it is down in January, then it will be down for the rest of the year. The following table gives data for the Standard & Poor’s 500 stock index for the 75 years from 1916 to 1990:

<table>
<thead>
<tr>
<th>Rest of year</th>
<th>January Up</th>
<th>January Down</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up</td>
<td>35</td>
<td>13</td>
</tr>
<tr>
<td>Down</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>

A chi-square analysis is valid for this problem if we assume that the yearly data are independent observations of a process that generates either an “up” or a “down” both in January and for the rest of the year.

(a) Calculate the column percents for this table. Explain briefly what they express.
(b) Do the same for the row percents.
(c) State appropriate null and alternative hypotheses for this problem. Use words rather than symbols.
(d) Find the table of expected counts under the null hypothesis. In which cells do the expected counts exceed the observed counts? In what cells are they less than the observed counts? Explain why the pattern suggests that the January indicator is valid.
(e) Give the value of the $X^2$ statistic, its degrees of freedom, and the $P$-value. What do you conclude?
(f) Write a short discussion of the evidence for the January indicator, referring to your analysis for substantiation.

9.2 In January 1975, the Committee on Drugs of the American Academy of Pediatrics recommended that tetracycline drugs not be given to children under the age of 8. A two-year study conducted in Tennessee investigated the extent to which physicians had prescribed these drugs between 1973 and 1975. The study categorized family practice physicians according to whether the county of their practice was urban, intermediate, or rural. The researchers examined how many doctors in each of these categories prescribed tetracycline to at least one patient under the age of 8. Here is the table of observed counts (data from Wayne A. Ray et al., “Prescribing of tetracycline to children less than 8 years old,” Journal of the American Medical Association, 237 (1977), pp. 2069–2074):

<table>
<thead>
<tr>
<th>County type</th>
<th>Urban</th>
<th>Intermediate</th>
<th>Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tetracycline</td>
<td>65</td>
<td>90</td>
<td>172</td>
</tr>
<tr>
<td>No tetracycline</td>
<td>149</td>
<td>136</td>
<td>158</td>
</tr>
</tbody>
</table>