ADDITIONAL EXERCISES
for
THE BASIC PRACTICE OF STATISTICS
Fourth Edition

David S. Moore

The 421 exercises in this supplement appeared in the second or third edition of BPS. They were replaced in later editions in the interest of freshness, but they remain high-quality exercises that supplement those in the fourth edition. Some have been updated using more recent data.

W. H. Freeman and Company
Chapter 1

1.1 A medical study. Data from a medical study contain values of many variables for each of the people who were the subjects of the study. Which of the following variables are categorical and which are quantitative?
(a) Gender (female or male)
(b) Age (years)
(c) Race (Asian, black, white, or other)
(d) Smoker (yes or no)
(e) Systolic blood pressure (millimeters of mercury)
(f) Level of calcium in the blood (micrograms per milliliter)

1.2 Mutual funds. Here is information on several Vanguard Group mutual funds:

<table>
<thead>
<tr>
<th>Fund</th>
<th>Number of stocks held</th>
<th>Largest holding</th>
<th>Annual return (10 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 Index Fund</td>
<td>508</td>
<td>General Electric</td>
<td>10.01%</td>
</tr>
<tr>
<td>Equity Income Fund</td>
<td>167</td>
<td>ExxonMobil</td>
<td>11.96%</td>
</tr>
<tr>
<td>Health Care Fund</td>
<td>128</td>
<td>Pharmacia</td>
<td>20.27%</td>
</tr>
<tr>
<td>International Value Fund</td>
<td>84</td>
<td>Mazda Motor</td>
<td>5.04%</td>
</tr>
<tr>
<td>Precious Metals Fund</td>
<td>26</td>
<td>Barrick Gold</td>
<td>2.50%</td>
</tr>
</tbody>
</table>

In addition to the fund name, how many variables are recorded for each fund? Which variables are categorical and which are quantitative?

1.3 Sports car fuel economy. Interested in a sports car? The Environmental Protection Agency lists most such vehicles in its “two-seater” category. The table below gives the city and highway mileages (miles per gallon) for the 22 two-seaters listed for the 2002 model year. Make a histogram of the highway mileages of these cars using classes with width 5 miles per gallon.

<table>
<thead>
<tr>
<th>Model</th>
<th>City</th>
<th>Highway</th>
<th>Model</th>
<th>City</th>
<th>Highway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acura NSX</td>
<td>17</td>
<td>24</td>
<td>Honda Insight</td>
<td>57</td>
<td>56</td>
</tr>
<tr>
<td>Audi TT Quattro</td>
<td>20</td>
<td>28</td>
<td>Honda S2000</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>Audi TT Roadster</td>
<td>22</td>
<td>31</td>
<td>Lamborghini Murcielago</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>BMW M Coupe</td>
<td>17</td>
<td>25</td>
<td>Mazda Miata</td>
<td>22</td>
<td>28</td>
</tr>
<tr>
<td>BMW Z3 Coupe</td>
<td>19</td>
<td>27</td>
<td>Mercedes-Benz SL500</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>BMW Z3 Roadster</td>
<td>20</td>
<td>27</td>
<td>Mercedes-Benz SL600</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>BMW Z8</td>
<td>13</td>
<td>21</td>
<td>Mercedes-Benz SLK230</td>
<td>23</td>
<td>30</td>
</tr>
<tr>
<td>Chevrolet Corvette</td>
<td>18</td>
<td>25</td>
<td>Mercedes-Benz SLK320</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>Chrysler Prowler</td>
<td>18</td>
<td>23</td>
<td>Porsche 911 GT2</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>Ferrari 360 Modena</td>
<td>11</td>
<td>16</td>
<td>Porsche Boxster</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td>Ford Thunderbird</td>
<td>17</td>
<td>23</td>
<td>Toyota MR2</td>
<td>25</td>
<td>30</td>
</tr>
</tbody>
</table>

1.4 The statistics of writing style. Numerical data can distinguish different types of writing, and sometimes even individual authors. Here are data on the percentages of words of 1 to 15 letters used in articles in Popular Science magazine.
(a) Make a histogram of this distribution. Describe its shape, center, and spread.
(b) How does the distribution of lengths of words used in *Popular Science* compare with the similar distribution in Figure 1.12(a) for Shakespeare’s plays? Look in particular at short words (2, 3, and 4 letters) and very long words (more than 10 letters).

1.5 Sports car fuel economy. Exercise 1.3 gives data on the fuel economy of 2002 model sports cars. Your histogram shows an extreme high outlier. This is the Honda Insight, a hybrid gas-electric car that is quite different from the others listed. Make a new histogram of highway mileage, leaving out the Insight. Classes that are about 2 miles per gallon wide work well.
(a) Describe the main features (shape, center, spread, outliers) of the distribution of highway mileage.
(b) The government imposes a “gas guzzler” tax on cars with low gas mileage. Which of these cars do you think may be subject to the gas guzzler tax?

1.6 Weight of newborns. The table below gives the distribution of the weight at birth for all babies born in the United States in 1999:

(a) For comparison with other years and with other countries, we prefer a histogram of the percents in each weight class rather than the counts. Explain why.
(b) Make a histogram of the distribution, using percents on the vertical scale.
(c) A “low birth weight” baby is one weighing less than 2,500 grams. Low birth weight is tied to many health problems. What percent of all births were low birth weight babies?

<table>
<thead>
<tr>
<th>Weight</th>
<th>Count</th>
<th>Weight</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 500 grams</td>
<td>5,912</td>
<td>3,000 to 3,499 grams</td>
<td>1,470,019</td>
</tr>
<tr>
<td>500 to 999 grams</td>
<td>22,815</td>
<td>3,500 to 3,999 grams</td>
<td>1,137,401</td>
</tr>
<tr>
<td>1,000 to 1,499 grams</td>
<td>28,750</td>
<td>4,000 to 4,499 grams</td>
<td>332,863</td>
</tr>
<tr>
<td>1,500 to 1,999 grams</td>
<td>59,531</td>
<td>4,500 to 4,999 grams</td>
<td>53,751</td>
</tr>
<tr>
<td>2,000 to 2,499 grams</td>
<td>184,175</td>
<td>5,000 to 5,499 grams</td>
<td>6,069</td>
</tr>
<tr>
<td>2,500 to 2,999 grams</td>
<td>653,327</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.7 The changing age distribution of the United States. The distribution of the ages of a nation’s population has a strong influence on economic and social conditions. The table on the next page shows the age distribution of U.S. residents in 1950 and 2075, in millions of persons. The 1950 data come from that year’s census. The 2075 data are projections made by the Census Bureau.
(a) Because the total population in 2075 is much larger than the 1950 population, comparing percents in each age group is clearer than comparing counts. Make a table of the percent of the total population in each age group for both 1950 and 2075.
(b) Make a histogram of the 1950 age distribution (in percents). Then describe the main features of the distribution. In particular, look at the percent of children relative to the rest of the population.
(c) Make a histogram of the projected age distribution for the year 2075. Use the same scales as in (b) for easy comparison. What are the most important changes in the U.S. age distribution projected for the 125-year period between 1950 and 2075?

<table>
<thead>
<tr>
<th>Age group</th>
<th>1950</th>
<th>2075</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 10 years</td>
<td>29.3</td>
<td>34.9</td>
</tr>
<tr>
<td>10 to 19 years</td>
<td>21.8</td>
<td>35.7</td>
</tr>
<tr>
<td>20 to 29 years</td>
<td>24.0</td>
<td>36.8</td>
</tr>
<tr>
<td>30 to 39 years</td>
<td>22.8</td>
<td>38.1</td>
</tr>
<tr>
<td>40 to 49 years</td>
<td>19.3</td>
<td>37.8</td>
</tr>
<tr>
<td>50 to 59 years</td>
<td>15.5</td>
<td>37.5</td>
</tr>
<tr>
<td>60 to 69 years</td>
<td>11.0</td>
<td>34.5</td>
</tr>
<tr>
<td>70 to 79 years</td>
<td>5.5</td>
<td>27.2</td>
</tr>
<tr>
<td>80 to 89 years</td>
<td>1.6</td>
<td>18.8</td>
</tr>
<tr>
<td>90 to 99 years</td>
<td>0.1</td>
<td>7.7</td>
</tr>
<tr>
<td>100 to 109 years</td>
<td>–</td>
<td>1.7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>151.1</td>
<td>310.6</td>
</tr>
</tbody>
</table>

1.8 Students’ attitudes. The Survey of Study Habits and Attitudes (SSHA) is a psychological test that evaluates college students’ motivation, study habits, and attitudes toward school. A private college gives the SSHA to 18 of its incoming first-year women students. Their scores are

154 109 137 115 152 140 154 178 101
103 126 126 137 165 165 129 200 148

Make a stemplot of these data. The overall shape of the distribution is irregular, as often happens when only a few observations are available. Are there any outliers? About where is the center of the distribution (the score with half the scores above it and half below)? What is the spread of the scores (ignoring any outliers)?

1.9 Bear markets. Investors speak of a “bear market” when stock prices drop substantially. Here are data on all declines of at least 10% in the Standard & Poor’s 500-stock index between 1940 and 2005. The data show how far the index fell from its peak and how long the decline in stock prices lasted.

<table>
<thead>
<tr>
<th>Year</th>
<th>Decline (percent)</th>
<th>Duration (months)</th>
<th>Year</th>
<th>Decline (percent)</th>
<th>Duration (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940–1942</td>
<td>42</td>
<td>28</td>
<td>1966</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>1946</td>
<td>27</td>
<td>5</td>
<td>1968–1970</td>
<td>36</td>
<td>18</td>
</tr>
<tr>
<td>1950</td>
<td>14</td>
<td>1</td>
<td>1973–1974</td>
<td>48</td>
<td>21</td>
</tr>
<tr>
<td>1952</td>
<td>15</td>
<td>8</td>
<td>1981–1982</td>
<td>26</td>
<td>19</td>
</tr>
<tr>
<td>1955</td>
<td>10</td>
<td>1</td>
<td>1983–1984</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>1956–1957</td>
<td>22</td>
<td>15</td>
<td>1987</td>
<td>34</td>
<td>3</td>
</tr>
<tr>
<td>1962</td>
<td>26</td>
<td>6</td>
<td>1998</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>2000–2003</td>
<td>31</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(a) Make a stemplot of the percent declines in stock prices during these bear markets. Make a second stemplot, splitting the stems. Which graph do you prefer? Why?
(b) The shape of this distribution is irregular, but we could describe it as somewhat skewed. Is the distribution skewed to the right or to the left? Are there any outliers?
(c) Describe the center and spread of the data. What would you tell an investor about how far stocks fall in a bear market?

1.10 Babe Ruth’s home runs. Here are the numbers of home runs that Babe Ruth hit in his 15 years with the New York Yankees, 1920 to 1934:

\[ 54\ 59\ 35\ 41\ 46\ 25\ 47\ 60\ 54\ 46\ 49\ 46\ 34\ 22 \]

Make a stemplot for these data. Is the distribution roughly symmetric, clearly skewed, or neither? About how many home runs did Ruth hit in a typical year? Is his famous 60 home runs in 1927 an outlier?

1.11 Back-to-back stemplot. A leading recent home run hitter was Mark McGwire, who retired after the 2001 season. Here are McGwire’s home run counts for 1987 to 2001:

\[ 49\ 32\ 33\ 39\ 22\ 42\ 9\ 9\ 39\ 52\ 58\ 70\ 65\ 32\ 29 \]

A back-to-back stemplot helps us compare two distributions. Write the stems as usual, but with a vertical line both to their left and to their right. On the right, put leaves for Babe Ruth (see the previous exercise). On the left, put leaves for McGwire. Arrange the leaves on each stem in increasing order out from the stem. Now write a brief comparison of Ruth and McGwire as home run hitters. McGwire was injured in 1993 and there was a baseball strike in 1994. How do these events appear in the data?

1.12 Lions feeding. Feeding at a carcass leads to competition among lions. Ecologists collected data on feeding contests in Serengeti National Park, Tanzania. In each contest, a lion feeding at a carcass is challenged by another lion seeking to take its place. Who wins these contests tells us something about lion society. Here are data on contests between an adult lion (female or male) and an opponent of a different class:

<table>
<thead>
<tr>
<th>Opponent</th>
<th>Adult male</th>
<th>Adult female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult female</td>
<td>136</td>
<td>53</td>
</tr>
<tr>
<td>Subadult</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>Yearling</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Cub</td>
<td>25</td>
<td>115</td>
</tr>
</tbody>
</table>

Make separate graphs for males and females that compare their success against different classes of opponent. Explain why you decided to make the type of graph you chose. Then describe the most important differences between the behavior of female and male lions in feeding contests.

1.13 Vanishing landfills. Garbage that is not recycled is buried in landfills. Here are time series data that emphasize the need for recycling: the number of landfills operating in the United States in the years 1988 to 2002.
Make a time plot of these data. Describe the trend that your plot shows. Why does the trend emphasize the need for recycling?

1.14 The influenza epidemic of 1918 (EESEE). In 1918 and 1919 a worldwide outbreak of influenza killed more than 25 million people. Here are data on the number of new influenza cases and the number of deaths from the epidemic in San Francisco week by week from October 5, 1918, to January 25, 1919. The date given is the last day of the week.  

<table>
<thead>
<tr>
<th>Date</th>
<th>Oct. 5</th>
<th>Oct. 12</th>
<th>Oct. 19</th>
<th>Oct. 26</th>
<th>Nov. 2</th>
<th>Nov. 9</th>
<th>Nov. 16</th>
<th>Nov. 23</th>
<th>Nov. 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>36</td>
<td>531</td>
<td>4233</td>
<td>8682</td>
<td>7164</td>
<td>2229</td>
<td>600</td>
<td>164</td>
<td>57</td>
</tr>
<tr>
<td>Deaths</td>
<td>0</td>
<td>0</td>
<td>130</td>
<td>552</td>
<td>738</td>
<td>414</td>
<td>198</td>
<td>90</td>
<td>56</td>
</tr>
</tbody>
</table>

(a) Make a time plot of weekly new cases. Based on your plot, describe the progress of the epidemic.

(b) We would like to compare the patterns over time of number of new cases and number of deaths. To make the two variables similar in size for easier comparison, plot the number of deaths against time for October 5 to January 25, then plot the number of cases divided by 10 on the same graph using a different color. What do you see? In particular, about how long is the lag between changes in the number of cases and corresponding changes in deaths?

1.15 Poverty in the states. The table below gives the percents of people living below the poverty line in 2000 in the 26 states east of the Mississippi River. Make a stemplot of these data. Is the distribution roughly symmetric, skewed to the right, or skewed to the left? Which states (if any) are outliers?

<table>
<thead>
<tr>
<th>State</th>
<th>Percent</th>
<th>State</th>
<th>Percent</th>
<th>State</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>14.6</td>
<td>Maryland</td>
<td>7.3</td>
<td>Pennsylvania</td>
<td>9.9</td>
</tr>
<tr>
<td>Connecticut</td>
<td>7.6</td>
<td>Massachusetts</td>
<td>10.2</td>
<td>Rhode Island</td>
<td>10.0</td>
</tr>
<tr>
<td>Delaware</td>
<td>9.8</td>
<td>Michigan</td>
<td>10.2</td>
<td>South Carolina</td>
<td>11.9</td>
</tr>
<tr>
<td>Florida</td>
<td>12.1</td>
<td>Mississippi</td>
<td>15.5</td>
<td>Tennessee</td>
<td>13.3</td>
</tr>
<tr>
<td>Georgia</td>
<td>12.6</td>
<td>New Hampshire</td>
<td>7.4</td>
<td>Vermont</td>
<td>10.1</td>
</tr>
<tr>
<td>Illinois</td>
<td>10.5</td>
<td>New Jersey</td>
<td>8.1</td>
<td>Virginia</td>
<td>8.1</td>
</tr>
<tr>
<td>Indiana</td>
<td>8.2</td>
<td>New York</td>
<td>14.7</td>
<td>West Virginia</td>
<td>15.8</td>
</tr>
<tr>
<td>Kentucky</td>
<td>12.5</td>
<td>North Carolina</td>
<td>13.2</td>
<td>Wisconsin</td>
<td>8.8</td>
</tr>
<tr>
<td>Maine</td>
<td>9.8</td>
<td>Ohio</td>
<td>11.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.16 Split the stems. Make another stemplot of the poverty data in the previous exercise, splitting the stems to double the number of classes. Do you prefer this stemplot or that from the previous exercise? Why?
1.17 The Boston Marathon. Women were allowed to enter the Boston Marathon in 1972. The times (in minutes, rounded to the nearest minute) for the winning woman from 1972 to 2006 appear below. The fastest time to date is 2 hours, 20 minutes, and 43 seconds, by Margaret Okayo of Kenya in 2002.

(a) Make a time plot of the winning times.
(b) Give a brief description of the pattern of Boston Marathon winning times over these years. Has the rate of improvement in times slowed in recent years?

<table>
<thead>
<tr>
<th>Year</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972</td>
<td>190</td>
</tr>
<tr>
<td>1973</td>
<td>186</td>
</tr>
<tr>
<td>1974</td>
<td>167</td>
</tr>
<tr>
<td>1975</td>
<td>162</td>
</tr>
<tr>
<td>1976</td>
<td>167</td>
</tr>
<tr>
<td>1977</td>
<td>168</td>
</tr>
<tr>
<td>1978</td>
<td>165</td>
</tr>
<tr>
<td>1979</td>
<td>155</td>
</tr>
<tr>
<td>1980</td>
<td>154</td>
</tr>
<tr>
<td>1981</td>
<td>147</td>
</tr>
<tr>
<td>1982</td>
<td>150</td>
</tr>
<tr>
<td>1983</td>
<td>143</td>
</tr>
<tr>
<td>1984</td>
<td>149</td>
</tr>
<tr>
<td>1985</td>
<td>154</td>
</tr>
<tr>
<td>1986</td>
<td>145</td>
</tr>
<tr>
<td>1987</td>
<td>146</td>
</tr>
<tr>
<td>1988</td>
<td>145</td>
</tr>
<tr>
<td>1989</td>
<td>144</td>
</tr>
<tr>
<td>1990</td>
<td>145</td>
</tr>
<tr>
<td>1991</td>
<td>144</td>
</tr>
<tr>
<td>1992</td>
<td>144</td>
</tr>
<tr>
<td>1993</td>
<td>145</td>
</tr>
<tr>
<td>1994</td>
<td>142</td>
</tr>
<tr>
<td>1995</td>
<td>145</td>
</tr>
<tr>
<td>1996</td>
<td>147</td>
</tr>
<tr>
<td>1997</td>
<td>146</td>
</tr>
<tr>
<td>1998</td>
<td>143</td>
</tr>
<tr>
<td>1999</td>
<td>143</td>
</tr>
<tr>
<td>2000</td>
<td>146</td>
</tr>
<tr>
<td>2001</td>
<td>144</td>
</tr>
<tr>
<td>2002</td>
<td>141</td>
</tr>
<tr>
<td>2003</td>
<td>145</td>
</tr>
<tr>
<td>2004</td>
<td>144</td>
</tr>
<tr>
<td>2005</td>
<td>145</td>
</tr>
<tr>
<td>2006</td>
<td>144</td>
</tr>
</tbody>
</table>

Chapter 2

2.1 Sports car gas mileage. Exercise 1.3 gives the gas mileages for the 22 two-seater cars listed in the government’s fuel economy guide.

(a) Find the mean highway gas mileage from the formula for the mean. Then enter the data into your calculator and use the calculator’s $\bar{x}$ button to obtain the mean. Verify that you get the same result.
(b) The Honda Insight is an outlier that doesn’t belong with the other cars. Use your calculator to find the mean of the 21 cars that remain if we leave out the Insight. How does the outlier change the mean?

2.2 Sports car gas mileage. What is the median highway mileage for the 22 two-seater cars listed in Exercise 1.3? What is the median of the 21 cars that remain if we remove the Honda Insight? Compare the effect of the Insight on mean mileage (Exercise 2.1) and on the median mileage. What general fact about the mean and median does this comparison illustrate?

(a) Find the five-number summaries for both city and highway mileage for the midsize cars in the table above and for the two-seater cars in Exercise 1.3. (Leave out the Honda Insight.)
(b) Make four side-by-side boxplots to display the summaries. Write a brief description of city versus highway mileage and two-seaters versus midsize cars.

2.3 Midsize car gas mileage. The table on the next page gives are the city and highway gas mileage for 36 midsize cars from the 2002 model year. There is one low outlier, the 12-cylinder Rolls-Royce. We wonder if midsize sedans get better mileage than sports cars.
(a) Find the five-number summaries for both city and highway mileage for these midsize cars and for the two-seater cars in Exercise 1.3. (Leave out the Honda Insight.)

(b) Make four side-by-side boxplots to display the summaries. Write a brief description of city versus highway mileage and two-seaters versus midsize cars.

### 2.4 How old are presidents?

How old are presidents at their inauguration? Was Bill Clinton, at age 46, unusually young? Here are the ages of all U.S. presidents when they took office.
(a) Make a stemplot of the distribution of ages. From the shape of the distribution, do you expect the median to be much less than the mean, about the same as the mean, or much greater than the mean?
(b) Find the mean and the five-number summary. Verify your expectation about the median.
(c) What is the range of the middle half of the ages of new presidents? Was Bill Clinton in the youngest 25%?

2.5 Internet access. How much do users pay for Internet access? Here are the monthly fees (in dollars) paid by a random sample of 50 users of commercial Internet service providers in August 2000:9

<table>
<thead>
<tr>
<th>20</th>
<th>40</th>
<th>22</th>
<th>22</th>
<th>21</th>
<th>21</th>
<th>20</th>
<th>10</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>13</td>
<td>18</td>
<td>50</td>
<td>20</td>
<td>18</td>
<td>15</td>
<td>8</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>22</td>
<td>10</td>
<td>20</td>
<td>22</td>
<td>22</td>
<td>21</td>
<td>15</td>
<td>23</td>
<td>30</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
<td>40</td>
<td>22</td>
<td>29</td>
<td>19</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>15</td>
<td>19</td>
<td>21</td>
<td>14</td>
<td>22</td>
<td>21</td>
<td>35</td>
<td>20</td>
<td>22</td>
</tr>
</tbody>
</table>

(a) Make a stemplot of these data. Briefly describe the pattern you see. About how much do you think America Online and its larger competitors were charging in August 2000? Are there any outliers?
(b) To report a quick summary of how much people pay for Internet service, do you prefer \( \bar{x} \) and \( s \) or the five-number summary? Why? Calculate your preferred summary.

2.6 McGwire versus Ruth. Exercises 1.10 and 1.11 give the numbers of home runs hit each season by Babe Ruth and Mark McGwire. Find the five-number summaries and make side-by-side boxplots to compare these two home run hitters. What do your plots show? Why are the boxplots less informative than the back-to-back stemplot you made in Exercise 1.11?

2.7 Students’ attitudes. Here are the scores of 18 first-year college women on the Survey of Study Habits and Attitudes (SSHA):

<table>
<thead>
<tr>
<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>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>

(a) Find the mean score from the formula for the mean. Then enter the data into your calculator and use the calculator’s \( \bar{x} \) button to obtain the mean. Verify that you get the same result.
(b) A stemplot (Exercise 1.8) suggests that the score 200 is an outlier. Use your calculator to find the mean for the 17 observations that remain when you drop the outlier. How does the outlier change the mean?

2.8 The density of the earth. In 1798 the English scientist Henry Cavendish measured the density of the earth with great care. It is common practice to repeat measurements several times and use the mean as the final result. Cavendish repeated his work 29 times. Here are his results (the data give the density of the earth as a multiple of the density of water):10
5.50 5.61 4.88 5.07 5.26 5.55 5.36 5.29 5.58 5.65
5.57 5.53 5.62 5.29 5.44 5.34 5.79 5.10 5.27 5.39
5.42 5.47 5.63 5.34 5.46 5.30 5.75 5.68 5.85

Present these measurements with a graph of your choice. Scientists usually give the mean and standard deviation to summarize a set of measurements. Does the shape of this distribution suggest that $\bar{x}$ and $s$ are adequate summaries? Calculate $\bar{x}$ and $s$.

2.9 Students’ attitudes. In Exercise 2.7 you found the mean of the SSHA scores of 18 first-year college women. Now find the median of these scores. Is the median smaller or larger than the mean? Explain why this is so.

2.10 Gas guzzlers? The BMW Z8, Ferrari, and Lamborghini have notably low gas mileage among the two-seater cars in Exercise 1.3. Find the five-number summary of the city gas mileages (omitting the Honda Insight). Are any of these cars suspected low outliers by the $1.5IQR$ rule?

2.11 College students studying. Do women study more than men? We asked the students in a large first-year college class how many minutes they studied on a typical weeknight. Here are the responses of random samples of 30 women and 30 men from the class:

<table>
<thead>
<tr>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>90</td>
</tr>
<tr>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td>150</td>
<td>240</td>
</tr>
<tr>
<td>200</td>
<td>215</td>
</tr>
<tr>
<td>120</td>
<td>60</td>
</tr>
<tr>
<td>90</td>
<td>0</td>
</tr>
</tbody>
</table>

(a) Examine the data. Why are you not surprised that most responses are multiples of 10 minutes? We eliminated one student who claimed to study 30,000 minutes per night. Are there any other responses you consider suspicious?
(b) Make a back-to-back stemplot of these data. Does it appear that women study more than men (or at least claim that they do)? Give numerical summaries that back up your conclusion.

2.12 More on study times. In the previous exercise you examined the nightly study time claimed by first-year college men and women. The most common methods for formal comparison of two groups use $\bar{x}$ and $s$ to summarize the data.
(a) What kinds of distributions are best summarized by $\bar{x}$ and $s$?
(b) Each set of study times appears to contain a high outlier. How much does removing the outlier change $\bar{x}$ and $s$ for each group? The presence of outliers makes us reluctant to use the mean and standard deviation for these data unless we remove the outliers on the grounds that these students were exaggerating.

2.13 How much oil? How much oil wells in a given field will ultimately produce is key information in deciding whether to drill more wells. Here are the estimated total amounts of oil recovered from 64 wells in the Devonian Richmond Dolomite area of the Michigan basin, in thousands of barrels:11
Describe the distribution with both a graph and suitable numerical summaries. What are the main features of the distribution that might be of interest to a landowner in this area thinking about drilling for oil? Follow the four-step process in your answer.

2.14 A hot stock? Here are the monthly percent returns on Philip Morris stock for the period from July 1990 to May 1997. (The return on an investment consists of the change in its price plus any cash payments made, given here as a percent of its price at the start of each month.)

(a) Make either a histogram or a stemplot of these data. How did you decide which graph to make?
(b) There is one clear outlier. What is the value of this observation? (It is explained by news of action against smoking, which depressed this tobacco company stock.) Describe the shape, center, and spread of the data after you omit the outlier.
(c) It is usual in the study of investments to use the mean and standard deviation to summarize and compare investment returns. Find the mean monthly return and the standard deviation of the returns. If you invested $100 in this stock at the beginning of a month and got the mean return, how much would you have at the end of the month?
(d) If you invested $100 in this stock at the beginning of the worst month in the data (the outlier), how much would you have at the end of the month? Find the mean and standard deviation again, this time leaving out the low outlier. How much did this one observation affect the summary measures? Would leaving out this one observation change the median? The quartiles? How do you know, without actual calculation?

2.15 Household net worth. A household’s “net worth” is the total value of the household’s possessions and investments less the total of its debts. In 2000, the mean and median net worth of American households were $55,000 and $182,000. Which of these numbers is the mean and which is the median? Explain your reasoning.
2.16 Highly paid athletes. A news article reports that of the 411 players on National Basketball Association rosters in February 1998, only 139 “made more than the league average salary” of $2.36 million. Is $2.36 million the mean or median salary for NBA players? How do you know?

2.17 Mean or median? Which measure of center, the mean or the median, should you use in each of the following situations?
(a) Middletown is considering imposing an income tax on citizens. The city government wants to know the average income of citizens so that it can estimate the total tax base.
(b) In a study of the standard of living of typical families in Middletown, a sociologist estimates the average family income in that city.

Chapter 3

3.1 How big are soldiers’ heads? The army reports that the distribution of head circumference among male soldiers is approximately Normal with mean 22.8 inches and standard deviation 1.1 inches. Use the 68–95–99.7 rule to answer the following questions.
(a) What percent of soldiers have head circumference greater than 23.9 inches?
(b) What percent of soldiers have head circumference between 21.7 inches and 23.9 inches?

3.2 Three great hitters. Three landmarks of baseball achievement are Ty Cobb’s batting average of .420 in 1911, Ted Williams’s .406 in 1941, and George Brett’s .390 in 1980. These batting averages cannot be compared directly because the distribution of major league batting averages has changed over the years. The distributions are quite symmetric and (except for outliers such as Cobb, Williams, and Brett) reasonably Normal. While the mean batting average has been held roughly constant by rule changes and the balance between hitting and pitching, the standard deviation has dropped over time. Here are the facts:

<table>
<thead>
<tr>
<th>Decade</th>
<th>Mean</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1910s</td>
<td>.266</td>
<td>.0371</td>
</tr>
<tr>
<td>1940s</td>
<td>.267</td>
<td>.0326</td>
</tr>
<tr>
<td>1970s</td>
<td>.261</td>
<td>.0317</td>
</tr>
</tbody>
</table>

Compute the standardized batting averages for Cobb, Williams, and Brett to compare how far each stood above his peers.

3.3 Length of pregnancies. The length of human pregnancies from conception to birth varies according to a distribution that is approximately Normal with mean 266 days and standard deviation 16 days.
(a) What percent of pregnancies last less than 240 days (that’s about 8 months)?
(b) What percent of pregnancies last between 240 and 270 days (roughly between 8 months and 9 months)?
(c) How long do the longest 20% of pregnancies last?
3.4 The stock market. The annual rate of return on stock indexes (which combine many individual stocks) is approximately Normal. Since 1945, the Standard & Poor’s 500 index has had a mean yearly return of 12%, with a standard deviation of 16.5%. Take this Normal distribution to be the distribution of yearly returns over a long period.

(a) In what range do the middle 95% of all yearly returns lie?
(b) The market is down for the year if the return on the index is less than zero. In what proportion of years is the market down?
(c) In what proportion of years does the index gain 25% or more?

3.5 Are we getting smarter? When the Stanford-Binet “IQ test” came into use in 1932, it was adjusted so that scores for each age group of children followed roughly the Normal distribution with mean $\mu = 100$ and standard deviation $\sigma = 15$. The test is readjusted from time to time to keep the mean at 100. If present-day American children took the 1932 Stanford-Binet test, their mean score would be about 120. The reasons for the increase in IQ over time are not known but probably include better childhood nutrition and more experience in taking tests.\textsuperscript{13}

(a) IQ scores above 130 are often called “very superior.” What percent of children had very superior scores in 1932?
(b) If present-day children took the 1932 test, what percent would have very superior scores? (Assume that the standard deviation $\sigma = 15$ does not change.)

3.6 Deciles. The deciles of any distribution are the points that mark off the lowest 10% and the highest 10%. On a density curve, these are the points with areas 0.1 and 0.9 to their left under the curve.

(a) What are the deciles of the standard Normal distribution?
(b) The heights of young women are approximately Normal with mean 64 inches and standard deviation 2.7 inches. What are the deciles of this distribution?

Chapter 4

4.1 Treating breast cancer. The most common treatment for breast cancer was once removal of the breast. It is now usual to remove only the tumor and nearby lymph nodes, followed by radiation. The change in policy was due to a large medical experiment that compared the two treatments. Each treatment was given to a separate group of breast cancer patients, chosen at random. The patients were closely followed to see how long they lived following surgery. What are the explanatory and response variables? Are they categorical or quantitative variables?

4.2 The risks of obesity. A study observes a large group of people over a 10-year period. The goal is to see if overweight and obese people are more likely to die during the decade than people who weigh less. Such studies can be misleading, because obese people are more likely to be inactive and to be poor. What is the explanatory variable and the response variable? What other variables are mentioned that may influence the relationship between the explanatory variable and the response variable?
4.3 Predicting height. How well does a child’s height at age 6 predict height at age 16? To find out, measure the heights of a large group of children at age 6, wait until they reach age 16, then measure their heights again. What are the explanatory and response variables here? Are these variables categorical or quantitative?

4.4 The endangered manatee. Manatees are large, gentle sea creatures that live along the Florida coast. Many manatees are killed or injured by powerboats. Here are data on powerboat registrations (in thousands) and the number of manatees killed by boats in Florida in the years 1977 to 2005. (Total vessel registrations for the years 1995 to 1997 are not available–Florida gives only recreational vessel registrations for these years.)

<table>
<thead>
<tr>
<th>Year</th>
<th>Powerboat registrations (1000)</th>
<th>Manatees killed</th>
<th>Year</th>
<th>Powerboat registrations (1000)</th>
<th>Manatees killed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>447</td>
<td>13</td>
<td>1992</td>
<td>716</td>
<td>38</td>
</tr>
<tr>
<td>1978</td>
<td>460</td>
<td>21</td>
<td>1993</td>
<td>716</td>
<td>35</td>
</tr>
<tr>
<td>1979</td>
<td>481</td>
<td>24</td>
<td>1994</td>
<td>735</td>
<td>49</td>
</tr>
<tr>
<td>1980</td>
<td>498</td>
<td>16</td>
<td>1995</td>
<td>NA</td>
<td>42</td>
</tr>
<tr>
<td>1981</td>
<td>513</td>
<td>24</td>
<td>1996</td>
<td>NA</td>
<td>60</td>
</tr>
<tr>
<td>1982</td>
<td>512</td>
<td>20</td>
<td>1997</td>
<td>NA</td>
<td>54</td>
</tr>
<tr>
<td>1983</td>
<td>526</td>
<td>15</td>
<td>1998</td>
<td>809</td>
<td>66</td>
</tr>
<tr>
<td>1984</td>
<td>559</td>
<td>34</td>
<td>1999</td>
<td>830</td>
<td>82</td>
</tr>
<tr>
<td>1985</td>
<td>585</td>
<td>33</td>
<td>2000</td>
<td>880</td>
<td>78</td>
</tr>
<tr>
<td>1986</td>
<td>614</td>
<td>33</td>
<td>2001</td>
<td>944</td>
<td>81</td>
</tr>
<tr>
<td>1987</td>
<td>645</td>
<td>39</td>
<td>2002</td>
<td>962</td>
<td>95</td>
</tr>
<tr>
<td>1988</td>
<td>675</td>
<td>43</td>
<td>2003</td>
<td>978</td>
<td>73</td>
</tr>
<tr>
<td>1989</td>
<td>711</td>
<td>50</td>
<td>2004</td>
<td>983</td>
<td>69</td>
</tr>
<tr>
<td>1990</td>
<td>719</td>
<td>47</td>
<td>2005</td>
<td>1010</td>
<td>79</td>
</tr>
<tr>
<td>1991</td>
<td>716</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) We want to examine the relationship between number of powerboats and number of manatees killed by boats. Which is the explanatory variable?
(b) Make a scatterplot of these data. (Be sure to label the axes with the variable names, not just x and y.) What does the scatterplot show about the relationship between these variables?

4.5 Calories and salt in hot dogs. Are hot dogs that are high in calories also high in salt? The scatterplot on the next page plots salt content (milligrams of sodium) against calories for 17 brands of meat hot dogs.

(a) Roughly what are the lowest and highest calorie counts among these brands? Roughly what is the sodium level in the brands with the fewest and with the most calories?
(b) Does the scatterplot show a clear positive or negative association? Say in words what this association means about calories and salt in hot dogs.
(c) Are there any outliers? Is the relationship (ignoring any outliers) roughly linear in form? Still ignoring any outliers, how strong would you say the relationship between calories and sodium is?
4.6 Is wine good for your heart? There is some evidence that drinking moderate amounts of wine helps prevent heart attacks. The table on the next page gives data on yearly wine consumption (liters of alcohol from drinking wine, per person) and yearly deaths from heart disease (deaths per 100,000 people) in 19 developed nations.\textsuperscript{16}

<table>
<thead>
<tr>
<th>Country</th>
<th>Alcohol from wine</th>
<th>Heart disease deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>2.5</td>
<td>211</td>
</tr>
<tr>
<td>Austria</td>
<td>3.9</td>
<td>167</td>
</tr>
<tr>
<td>Belgium</td>
<td>2.9</td>
<td>131</td>
</tr>
<tr>
<td>Canada</td>
<td>2.4</td>
<td>191</td>
</tr>
<tr>
<td>Denmark</td>
<td>2.9</td>
<td>220</td>
</tr>
<tr>
<td>Finland</td>
<td>0.8</td>
<td>297</td>
</tr>
<tr>
<td>France</td>
<td>9.1</td>
<td>71</td>
</tr>
<tr>
<td>Iceland</td>
<td>0.8</td>
<td>211</td>
</tr>
<tr>
<td>Ireland</td>
<td>0.7</td>
<td>300</td>
</tr>
<tr>
<td>Italy</td>
<td>7.9</td>
<td>107</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Country</th>
<th>Alcohol from wine</th>
<th>Heart disease deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netherlands</td>
<td>1.8</td>
<td>167</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1.9</td>
<td>266</td>
</tr>
<tr>
<td>Norway</td>
<td>0.8</td>
<td>227</td>
</tr>
<tr>
<td>Spain</td>
<td>6.5</td>
<td>86</td>
</tr>
<tr>
<td>Sweden</td>
<td>1.6</td>
<td>207</td>
</tr>
<tr>
<td>Switzerland</td>
<td>5.8</td>
<td>115</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1.3</td>
<td>285</td>
</tr>
<tr>
<td>United States</td>
<td>1.2</td>
<td>199</td>
</tr>
<tr>
<td>West Germany</td>
<td>2.7</td>
<td>172</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot that shows how national wine consumption helps explain heart disease death rates.
(b) Describe the form of the relationship. Is there a linear pattern? How strong is the relationship?
(c) Is the direction of the association positive or negative? Explain in simple lan-
guage what this says about wine and heart disease. Do you think these data give
good evidence that drinking wine causes a reduction in heart disease deaths? Why?

4.7 Do heavier people burn more energy? Metabolic rate, the rate at which
the body consumes energy, is important in studies of weight gain, dieting, and
exercise. Here are data on the lean body mass and resting metabolic rate for 12
women and 7 men who are subjects in a study of dieting. Lean body mass, given
in kilograms, is a person’s weight leaving out all fat. Metabolic rate is measured in
calories burned per 24 hours, the same calories used to describe the energy content
of foods. The researchers believe that lean body mass is an important influence on
metabolic rate.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>Mass (kg)</th>
<th>Rate (cal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>62.0</td>
<td>1792</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>62.9</td>
<td>1666</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>36.1</td>
<td>995</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>54.6</td>
<td>1425</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>48.5</td>
<td>1396</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>42.0</td>
<td>1418</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>47.4</td>
<td>1362</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>50.6</td>
<td>1502</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>42.0</td>
<td>1256</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>48.7</td>
<td>1614</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>40.3</td>
<td>1189</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>33.1</td>
<td>913</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>51.9</td>
<td>1460</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>42.4</td>
<td>1124</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>34.5</td>
<td>1052</td>
</tr>
<tr>
<td>16</td>
<td>F</td>
<td>41.2</td>
<td>1204</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>51.9</td>
<td>1867</td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>46.9</td>
<td>1439</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot of the data for the female subjects. Which is the explanatory
variable?

(b) Is the association between these variables positive or negative? What is the form
of the relationship? How strong is the relationship?

(c) Now add the data for the male subjects to your graph, using a different color or
a different plotting symbol. Does the pattern of relationship that you observed in
(b) hold for men also? How do the male subjects as a group differ from the female
subjects as a group?

4.8 Classifying fossils. Archaeopteryx is an extinct beast having feathers like a
bird but teeth and a long bony tail like a reptile. Only six fossil specimens are
known. Because these specimens differ greatly in size, some scientists think they
belong to different species. We will examine some data. If the specimens belong
to the same species and differ in size because some are younger than others, there
should be a positive linear relationship between the lengths of a pair of bones from all
individuals. An outlier from this relationship would suggest a different species. Here
are data on the lengths in centimeters of the femur (a leg bone) and the humerus
(a bone in the upper arm) for the five specimens that preserve both bones:\n
<table>
<thead>
<tr>
<th>Femur</th>
<th>38 56 59 64 74</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humerus</td>
<td>41 63 70 72 84</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot. Do you think that all five specimens come from the same
species?

(b) Find the correlation $r$ step-by-step. That is, find the mean and standard devi-
ation of the femur lengths and of the humerus lengths. (Use your calculator for
means and standard deviations.) Then find the five standardized values for each
variable and use the formula for \( r \).
(c) Now enter these data into your calculator and use the calculator’s correlation
function to find \( r \). Check that you get the same result as in (b).

### 4.9 Body mass and metabolic rate, continued.

Exercise 4.7 gives data on the
lean body mass and metabolic rate for 12 women and 7 men.
(a) Based on your scatterplot, do you think the correlation will be about the same
for men and women or quite different for the two groups? Why?
(b) Calculate \( r \) for women alone and also for men alone. (Use your calculator.)
(c) Calculate the mean body mass for the women and for the men. Does the fact
that the men are heavier than the women on the average influence the correlations?
If so, in what way?
(d) Lean body mass was measured in kilograms. How would the correlations change
if we measured body mass in pounds? (There are about 2.2 pounds in a kilogram.)

### 4.10 The effect of changing units.

Changing the units of measurement can
dramatically alter the appearance of a scatterplot. Return to the fossil data from
Exercise 4.8. These measurements are in centimeters. Suppose a mad scientist
measured the femur in meters and the humerus in millimeters. The data would
then be

<table>
<thead>
<tr>
<th>Femur</th>
<th>0.38</th>
<th>0.56</th>
<th>0.59</th>
<th>0.64</th>
<th>0.74</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humerus</td>
<td>410</td>
<td>630</td>
<td>700</td>
<td>720</td>
<td>840</td>
</tr>
</tbody>
</table>

(a) Draw an \( x \) axis extending from 0 to 75 and a \( y \) axis extending from 0 to 850.
Plot the original data on these axes. Then plot the new data on the same axes in a
different color. The two plots look very different.
(b) Nonetheless, the correlation is exactly the same for the two sets of measurements.
Why do you know that this is true without doing any calculations? Find the two
correlations to verify that they are the same.

### Brains and bodies

*The scatterplot on the next page plots the average brain weight
in grams versus average body weight in kilograms for 96 species of mammals. There
are many small mammals whose points at the lower left overlap. Exercises 4.11 to
4.13 are based on this scatterplot.*

### 4.11 Dolphins and hippos.

The points for the dolphin and hippopotamus are
labeled in the scatterplot. Read from the graph the approximate body weight and
brain weight for these two species.

### 4.12 Dolphins and hippos, continued.

One reaction to this scatterplot is “Dolphins are smart, hippos are dumb.” What feature of the plot lies behind this
reaction?

### 4.13 Outliers.

The African elephant is much larger than any other mammal in
the data set but lies roughly in the overall straight-line pattern. Dolphins, humans,
and hippos lie outside the overall pattern. The correlation between body weight
and brain weight for the entire data set is \( r = 0.86 \).
(a) If we removed the elephant, would this correlation increase or decrease or not
change much? Explain your answer.

(b) If we removed dolphins, hippos, and humans, would this correlation increase or decrease or not change much? Explain your answer.

4.14 Mice. For a biology project, you measure the tail length (centimeters) and weight (grams) of 12 mice of the same variety.

(a) Explain why you expect the correlation between tail length and weight to be positive.

(b) The mean tail length turns out to be 9.8 centimeters. What is the mean length in inches? (There are 2.54 centimeters in an inch.)

(c) The correlation between tail length and weight turns out to be $r = 0.6$. If you measured length in inches instead of centimeters, what would be the new value of $r$?

4.15 How many calories? A food industry group asked 3368 people to guess the number of calories in each of several common foods. The table on the next page gives the averages of their guesses and the correct number of calories.19

(a) We think that how many calories a food actually has helps explain people’s guesses of how many calories it has. With this in mind, make a scatterplot of these data.

(b) Find the correlation $r$ (use your calculator). Explain why your $r$ is reasonable based on the scatterplot.

(c) The guesses are all higher than the true calorie counts. Does this fact influence the correlation in any way? How would $r$ change if every guess were 100 calories higher?

(d) The guesses are much too high for spaghetti and snack cake. Circle these points
on your scatterplot. Calculate $r$ for the other eight foods, leaving out these two points. Explain why $r$ changed in the direction that it did.

<table>
<thead>
<tr>
<th>Food</th>
<th>Guessed calories</th>
<th>Correct calories</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 oz. whole milk</td>
<td>196</td>
<td>159</td>
</tr>
<tr>
<td>5 oz. spaghetti with tomato sauce</td>
<td>394</td>
<td>163</td>
</tr>
<tr>
<td>5 oz. macaroni with cheese</td>
<td>350</td>
<td>269</td>
</tr>
<tr>
<td>One slice wheat bread</td>
<td>117</td>
<td>61</td>
</tr>
<tr>
<td>One slice white bread</td>
<td>136</td>
<td>76</td>
</tr>
<tr>
<td>2-oz. candy bar</td>
<td>364</td>
<td>260</td>
</tr>
<tr>
<td>Saltine cracker</td>
<td>74</td>
<td>12</td>
</tr>
<tr>
<td>Medium-size apple</td>
<td>107</td>
<td>80</td>
</tr>
<tr>
<td>Medium-size potato</td>
<td>160</td>
<td>88</td>
</tr>
<tr>
<td>Cream-filled snack cake</td>
<td>419</td>
<td>160</td>
</tr>
</tbody>
</table>

4.16 **Sports car gas mileage.** Exercise 1.3 gives the city and highway gas mileages for two-seater cars. We expect a positive association between the city and highway mileages of a group of vehicles.

(a) Make a scatterplot that shows the relationship between city and highway mileage, using city mileage as the explanatory variable. Describe the overall pattern. Does the outlier (the Honda Insight) extend the pattern of the other cars or is it far from the line they form?

(b) Find the correlation between city and highway mileage, leaving out the Insight. Explain how the correlation matches the pattern of the plot. Based on your plot, will adding the Insight make the correlation stronger (closer to 1) or weaker? Verify your guess by calculating the correlation for all 22 cars

4.17 **Transforming data.** Data analysts often look for a transformation of data that simplifies the overall pattern. Here is an example of how transforming the response variable can simplify the pattern of a scatterplot. The data show the growth of Europe between 1750 and 1950.

<table>
<thead>
<tr>
<th>Year</th>
<th>1750</th>
<th>1800</th>
<th>1850</th>
<th>1900</th>
<th>1950</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population (millions)</td>
<td>125</td>
<td>187</td>
<td>274</td>
<td>423</td>
<td>594</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot of population against year. Briefly describe the pattern of Europe’s growth.

(b) Now take the logarithm of the population in each year (use the log button on your calculator). Plot the logarithms against year. What is the overall pattern on this plot?
Chapter 5

5.1 Review of straight lines. Fred keeps his savings in his mattress. He began with $500 from his mother and adds $100 each year. His total savings $y$ after $x$ years are given by the equation

$$y = 500 + 100x$$

(a) Draw a graph of this equation. (Choose two values of $x$, such as 0 and 10. Compute the corresponding values of $y$ from the equation. Plot these two points on graph paper and draw the straight line joining them.)

(b) After 20 years, how much will Fred have in his mattress?

(c) If Fred had added $200 instead of $100 each year to his initial $500, what is the equation that describes his savings after $x$ years?

5.2 Review of straight lines. During the period after birth, a male white rat gains exactly 40 grams (g) per week. (This rat is unusually regular in his growth, but 40 g per week is a realistic rate.)

(a) If the rat weighed 100 g at birth, give an equation for his weight after $x$ weeks. What is the slope of this line?

(b) Draw a graph of this line between birth and 10 weeks of age.

(c) Would you be willing to use this line to predict the rat’s weight at age 2 years? Do the prediction and think about the reasonableness of the result. (There are 454 grams in a pound. To help you assess the result, note that a large cat weighs about 10 pounds.)

5.3 Acid rain. Researchers studying acid rain measured the acidity of precipitation in a Colorado wilderness area for 150 consecutive weeks. Acidity is measured by pH. Lower pH values show higher acidity. The acid rain researchers observed a linear pattern over time. They reported that the least-squares regression line

$$\text{pH} = 5.43 - (0.0053 \times \text{weeks})$$

fit the data well.20

(a) Draw a graph of this line. Is the association positive or negative? Explain in plain language what this association means.

(b) According to the regression line, what was the pH at the beginning of the study (weeks = 1)? At the end (weeks = 150)?

(c) What is the slope of the regression line? Explain clearly what this slope says about the change in the pH of the precipitation in this wilderness area.

5.4 Is wine good for your heart? Exercise 4.6 gives data on wine consumption and heart disease death rates in 19 countries. Your scatterplot from that exercise shows a moderately strong relationship.

(a) The correlation for these variables is $r = -0.843$. What does a negative correlation say about wine consumption and heart disease deaths? About what percent of the variation among countries in heart disease death rates is explained by the straight-line relationship with wine consumption?
(b) The least-squares regression line for predicting heart disease death rate from wine consumption is
\[
\hat{y} = 260.56 - 22.969x
\]
Use this equation to predict the heart disease death rate in another country where adults average 4 liters of alcohol from wine each year.
(c) The correlation in (a) and the slope of the least-squares line in (b) are both negative. Is it possible for these two quantities to have opposite signs? Explain your answer.

5.5 Lots of wine. The previous exercise gives the least-squares line for predicting a nation's heart disease death rate from its wine consumption. What is the predicted heart disease death rate for a country that drinks enough wine to supply 150 liters of alcohol per person? Explain why this result can't be true. Explain why using the regression line for this prediction is not intelligent.

5.6 Brain and body: regression. The line on the scatterplot on page 18 is the least-squares regression line for predicting brain weight from body weight. Suppose that a new mammal species is discovered hidden in the rain forest with body weight 600 kilograms. Predict the brain weight for this species.

5.7 Slope. The line on the scatterplot on page 18 is the least-squares regression line for predicting brain weight from body weight. The slope of this line is one of the numbers below. Which number is the slope? Why?
(a) \(b = 0.5\)  \hspace{1cm} (b) \(b = 1.3\)  \hspace{1cm} (c) \(b = 3.2\).

5.8 Brain and body: correlation. The correlation between body weight and brain weight for the data plotted on page 18 is \(r = 0.86\). How well does body weight explain brain weight for mammals? Give a number to answer this question, and briefly explain what the number tells us.

5.9 The endangered manatee. Exercise 4.4 gives data on the number of powerboats registered in Florida and the number of manatees killed by boats in the years from 1977 to 2005. Your scatterplot shows a positive linear relationship between these variables.
(a) Find the least-squares regression line of manatees killed on boats registered and add this line to your scatterplot from Exercise 4.4.
(b) Manatee kills vary from 13 to 95. What percent of this variation is explained by straight-line dependence of kills on boat registrations?
(c) Predict the number of manatees killed by boats in a year in which 1 million powerboats are registered in Florida.

5.10 Sports car gas mileage. Exercise 1.3 gives the city and highway gas mileages for two-seater cars. A scatterplot (Exercise 4.16) shows a strong positive linear relationship.
(a) Find the least-squares regression line for predicting highway mileage from city mileage, using data from all 22 car models. Make a scatterplot and plot the regression line.
(b) What is the slope of the regression line? Explain in words what the slope says
about gas mileage for two-seater cars.
(c) Another two-seater is rated at 20 miles per gallon in the city. Predict its highway mileage.

5.11 **Sports car gas mileage.** In the previous exercise you found the least-squares regression line for predicting highway mileage from city mileage for the 22 two-seater car models in Exercise 1.3. Find the mean city mileage and mean highway mileage for these cars. Use your regression line to predict the highway mileage for a car with city mileage equal to the mean for the group. Explain why you knew the answer before doing the prediction.

5.12 **Take me out to the ball game.** What is the relationship between the price charged for a hot dog and the price charged for a 16-ounce soda in major league baseball stadiums? Here are some data:

<table>
<thead>
<tr>
<th>Team</th>
<th>Hot dog</th>
<th>Soda</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angels</td>
<td>2.50</td>
<td>1.75</td>
</tr>
<tr>
<td>Astros</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Braves</td>
<td>2.50</td>
<td>1.79</td>
</tr>
<tr>
<td>Brewers</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Cardinals</td>
<td>3.50</td>
<td>2.00</td>
</tr>
<tr>
<td>Dodgers</td>
<td>2.75</td>
<td>2.00</td>
</tr>
<tr>
<td>Expos</td>
<td>1.75</td>
<td>2.00</td>
</tr>
<tr>
<td>Giants</td>
<td>2.75</td>
<td>2.17</td>
</tr>
<tr>
<td>Indians</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Marlins</td>
<td>2.25</td>
<td>1.80</td>
</tr>
<tr>
<td>Mets</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>Padres</td>
<td>1.75</td>
<td>2.25</td>
</tr>
<tr>
<td>Phillies</td>
<td>2.75</td>
<td>2.20</td>
</tr>
<tr>
<td>Pirates</td>
<td>1.75</td>
<td>1.75</td>
</tr>
<tr>
<td>Red Sox</td>
<td>2.25</td>
<td>2.29</td>
</tr>
<tr>
<td>Rockies</td>
<td>2.25</td>
<td>2.25</td>
</tr>
<tr>
<td>Royals</td>
<td>1.75</td>
<td>1.99</td>
</tr>
<tr>
<td>Tigers</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Twins</td>
<td>2.50</td>
<td>2.22</td>
</tr>
<tr>
<td>White Sox</td>
<td>2.00</td>
<td>2.00</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot appropriate for predicting soda price from hot dog price. Describe the relationship that you see. Are there any outliers?
(b) Find the correlation between hot dog price and soda price. What percent of the variation in soda price does a linear relationship account for?
(c) Find the equation of the least-squares line for predicting soda price from hot dog price. Draw the line on your scatterplot. Based on your findings in (b), explain why it is not surprising that the line is nearly horizontal (slope near zero).
(d) Circle the observation that is potentially the most influential. What team is this? Find the least-squares line without this one observation and draw it on your scatterplot. Was the observation in fact influential?

5.13 **How many calories: influence?** Exercise 4.15 gives data on the true calories in ten foods and the average guesses made by a large group of people. That exercise explored the influence of two outlying observations on the correlation.
(a) Make a scatterplot suitable for predicting guessed calories from true calories. Circle the points for spaghetti and snack cake on your plot. These points lie outside the linear pattern of the other eight points.
(b) Use your calculator to find the least-squares regression line of guessed calories on true calories. Do this twice, first for all ten data points and then leaving out spaghetti and snack cake.
(c) Plot both lines on your graph. (Make one dashed so you can tell them apart.) Are spaghetti and snack cake, taken together, influential observations? Explain your answer.
5.14 **Sports car gas mileage: influence.** The data on gas mileage of two-seater cars (Exercise 1.3) contain an outlier, the Honda Insight. When we predict highway mileage from city mileage, this point is an outlier in both the $x$ and $y$ directions. We wonder if it influences the least-squares line.

(a) Make a scatterplot and draw (again) the least-squares line from all 22 car models.
(b) Find the least-squares line when the Insight is left out of the calculation and draw this line on your plot.
(c) Influence is a matter of degree, not a yes-or-no question. Use both regression lines to predict highway mileages for city mileages of 10, 20, and 25 MPG. (These city mileage values span the range of car models other than the Insight.) Do you think the Insight changes the predictions enough to be important to a car buyer?

5.15 **TV watching and school grades.** Children who watch many hours of television get lower grades in school on the average than those who watch less TV. Suggest some lurking variables that may explain this relationship because they contribute to both heavy TV viewing and poor grades.

5.16 **Do power lines cause cancer?** It has been suggested that electromagnetic fields of the kind present near power lines can cause leukemia in children. Experiments with children and power lines are not ethical. Careful studies have found no association between exposure to electromagnetic fields and childhood leukemia. Suggest several lurking variables that you would want information about in order to investigate the claim that living near power lines is associated with cancer.

5.17 **Do firefighters make fires worse?** Someone says, “There is a strong positive correlation between the number of firefighters at a fire and the amount of damage the fire does. So sending lots of firefighters just causes more damage.” Explain why this reasoning is wrong.

5.18 **Height and reading score.** A study of elementary school children, ages 6 to 11, finds a high positive correlation between height $x$ and score $y$ on a test of reading comprehension. What explains this correlation?

5.19 **How’s your self-esteem?** People who do well tend to feel good about themselves. Perhaps helping people feel good about themselves will help them do better in school and life. Raising self-esteem became for a time a goal in many schools. California even created a state commission to advance the cause. Can you think of explanations for the association between high self-esteem and good school performance other than “Self-esteem causes better work in school”?

5.20 **Using residuals.** Return to the regression of highway mileage on city mileage in Exercise 5.10. Use your calculator or software to obtain the residuals. Make a residual plot (residuals against city mileage) and add a horizontal line at $y = 0$ (the mean of the residuals).

(a) Which car has the largest positive residual? The largest negative residual?
(b) The Honda Insight, an extreme outlier, does not have the largest residual in either direction. Why is this not surprising?
(c) Explain briefly what a large positive residual says about a car. What does a large negative residual say?
5.21 Predicting enrollments. The mathematics department of a large state university would like to use the number of freshmen entering the university $x$ to predict the number of students $y$ who will sign up for freshman-level math courses in the fall semester. Here are data for several years:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>4595</td>
<td>4827</td>
<td>4427</td>
<td>4258</td>
<td>3995</td>
<td>4330</td>
<td>4265</td>
<td>4351</td>
</tr>
<tr>
<td>$y$</td>
<td>7364</td>
<td>7547</td>
<td>7099</td>
<td>6894</td>
<td>6572</td>
<td>7156</td>
<td>7232</td>
<td>7450</td>
</tr>
</tbody>
</table>

Software gives the correlation $r = 0.8333$ and the least-squares regression line

$$\hat{y} = 2492.69 + 1.0663x$$

The software also gives a table of the residuals:

|------|------|------|------|------|------|------|------|------|

(a) Make a scatterplot and draw the regression line on it. The regression line does not predict very accurately. What percent of variation in class enrollment is explained by the linear relationship with the count of freshmen?

(b) Check that the residuals have sum zero (at least up to roundoff error).

(c) Plots of the residuals against other variables are often revealing. Plot the residuals against year. One of the schools in the university recently changed its program to require that entering students take another mathematics course. How does the residual plot show the change? In what year was the change effective?

5.22 A hot stock? It is usual in finance to describe the returns from investing in a single stock by regressing the stock’s returns on the returns from the stock market as a whole. This helps us see how closely the stock follows the market. We analyzed the monthly percent total return $y$ on Philip Morris common stock and the monthly return $x$ on the Standard & Poor’s 500-stock index, which represents the market, for the period between July 1990 and May 1997. Here are the results:

$$\bar{x} = 1.304 \quad s_x = 3.392 \quad r = 0.5251$$

$$\bar{y} = 1.878 \quad s_y = 7.554$$

A scatterplot shows no very influential observations.

(a) Find the equation of the least-squares line from this information. What percent of the variation in Philip Morris stock is explained by the linear relationship with the market as a whole?

(b) Explain carefully what the slope of the line tells us about how Philip Morris stock responds to changes in the market. This slope is called “beta” in investment theory.

(c) Returns on most individual stocks have a positive correlation with returns on the entire market. That is, when the market goes up, an individual stock tends to also go up. Explain why an investor should prefer stocks with beta > 1 when the market is rising and stocks with beta < 1 when the market is falling.
5.23 Education and income. There is a strong positive correlation between years of education and income for economists employed by business firms. In particular, economists with doctorates earn more than economists with only a bachelor's degree. There is also a strong positive correlation between years of education and income for economists employed by colleges and universities. But when all economists are considered, there is a negative correlation between education and income. The explanation for this is that business pays high salaries and employs mostly economists with bachelor's degrees, while colleges pay lower salaries and employ mostly economists with doctorates. Sketch a scatterplot with two groups of cases (business and academic) illustrating how a strong positive correlation within each group and a negative overall correlation can occur together.

5.24 Women scientists. A study by the National Science Foundation found that the median salary of newly graduated female engineers and scientists was only 73% of the median salary for males. When the new graduates were broken down by field, however, the picture changed. Women’s median salaries as a percent of the male median in the 16 fields studied were

<table>
<thead>
<tr>
<th>Field</th>
<th>94%</th>
<th>96%</th>
<th>98%</th>
<th>95%</th>
<th>85%</th>
<th>85%</th>
<th>84%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>103%</td>
<td>100%</td>
<td>107%</td>
<td>93%</td>
<td>104%</td>
<td>93%</td>
<td>106%</td>
<td>100%</td>
</tr>
</tbody>
</table>

How can women do nearly as well as men in every field yet fall far behind men when we look at all young engineers and scientists?

Chapter 6

6.1 Does herbal tea help nursing home residents? A group of college students believes that herbal tea has remarkable powers. To test this belief, they make weekly visits to a local nursing home, where they visit with the residents and serve them herbal tea. The nursing home staff reports that after several months many of the residents are healthier and more cheerful. We should commend the students for their good deeds but doubt that herbal tea helped the residents. Identify the explanatory and response variables in this informal study. Then explain what lurking variables account for the observed association.

6.2 Smoking by students and their parents. Here are data from eight high schools on smoking among students and among their parents:

<table>
<thead>
<tr>
<th></th>
<th>Neither parent smokes</th>
<th>One parent smokes</th>
<th>Both parents smoke</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student does not smoke</td>
<td>1168</td>
<td>1823</td>
<td>1380</td>
</tr>
<tr>
<td>Student smokes</td>
<td>188</td>
<td>416</td>
<td>400</td>
</tr>
</tbody>
</table>

(a) How many students do these data describe?
(b) What percent of these students smoke?
(c) Calculate and compare percents to show how parents’ smoking influences students’ smoking. Briefly state your conclusions about the relationship.
6.3 Firearm deaths. Firearms are second to motor vehicles as a cause of non-disease deaths in the United States. Here are counts from a study of all firearm-related deaths in Milwaukee, Wisconsin, between 1990 and 1994.\textsuperscript{25} We want to compare the types of firearms used in homicides and in suicides. We suspect that long guns (shotguns and rifles) will more often be used in suicides because many people keep them at home for hunting. Make a careful comparison of homicides and suicides with a bar graph. What do you find about long guns versus handguns?

<table>
<thead>
<tr>
<th></th>
<th>Handgun</th>
<th>Shotgun</th>
<th>Rifle</th>
<th>Unknown</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homicides</td>
<td>468</td>
<td>28</td>
<td>15</td>
<td>13</td>
<td>524</td>
</tr>
<tr>
<td>Suicides</td>
<td>124</td>
<td>22</td>
<td>24</td>
<td>5</td>
<td>175</td>
</tr>
</tbody>
</table>

6.4 Smoking and staying alive. In the mid-1970s, a medical study contacted randomly chosen people in a district in England. Here are data on the 1314 women contacted who were either current smokers or who had never smoked. The table classifies these women by their smoking status and age at the time of the survey and whether they were still alive 20 years later.\textsuperscript{26}

<table>
<thead>
<tr>
<th>Age 18 to 44</th>
<th>Age 45 to 64</th>
<th>Age 65+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker</td>
<td>Not</td>
<td>Smoker</td>
</tr>
<tr>
<td>Dead</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>Alive</td>
<td>269</td>
<td>327</td>
</tr>
</tbody>
</table>

(a) Make from these data a two-way table of smoking (yes or no) by dead or alive. What percent of the smokers stayed alive for 20 years? What percent of the non-smokers survived? It seems surprising that a higher percent of smokers stayed alive.
(b) The age of the women at the time of the study is a lurking variable. Show that within each of the three age groups in the data, a higher percent of non-smokers remained alive 20 years later. This is another example of Simpson’s paradox.
(c) The study authors give this explanation: “Few of the older women (over 65 at the original survey) were smokers, but many of them had died by the time of follow-up.” Compare the percent of smokers in the three age groups to verify the explanation.

6.5 Age and marital status of women. The following two-way table describes the age and marital status of American women in 2004. The table entries are in thousands of women.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Never married</th>
<th>Married</th>
<th>Widowed</th>
<th>Divorced</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 to 24</td>
<td>10,972</td>
<td>2,473</td>
<td>11</td>
<td>156</td>
<td>13,612</td>
</tr>
<tr>
<td>25 to 39</td>
<td>7,789</td>
<td>19,209</td>
<td>244</td>
<td>2,822</td>
<td>30,064</td>
</tr>
<tr>
<td>40 to 64</td>
<td>4,151</td>
<td>32,891</td>
<td>2,253</td>
<td>8,062</td>
<td>47,357</td>
</tr>
<tr>
<td>≥ 65</td>
<td>743</td>
<td>8,710</td>
<td>8,633</td>
<td>1,764</td>
<td>19,850</td>
</tr>
<tr>
<td>Total</td>
<td>23,655</td>
<td>63,282</td>
<td>11,141</td>
<td>12,804</td>
<td>110,883</td>
</tr>
</tbody>
</table>

(a) Find the sum of the entries in the “Married” column. Why does this sum differ from the “Total” entry for that column?
(b) Give the marginal distribution of marital status for all adult women (use percents). Draw a bar graph to display this distribution.

(c) Compare the conditional distributions of marital status for women aged 18 to 24 and women aged 40 to 64. Briefly describe the most important differences between the two groups of women, and back up your description with percents.

(d) You are planning a magazine aimed at women who have never been married. Find the conditional distribution of age among single women and display it in a bar graph. What age group or groups should your magazine aim to attract?

6.6 Regulating guns. The 1998 National Gun Policy Survey, conducted by the National Opinion Research Center at the University of Chicago, asked many questions about regulation of guns in the United States. One of the questions was “Do you think there should be a law that would ban possession of handguns except for the police and other authorized persons?” Here are the responses, broken down by the respondents’ level of education:

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than high school</td>
<td>58</td>
<td>58</td>
</tr>
<tr>
<td>High school graduate</td>
<td>84</td>
<td>129</td>
</tr>
<tr>
<td>Some college</td>
<td>169</td>
<td>294</td>
</tr>
<tr>
<td>College graduate</td>
<td>98</td>
<td>135</td>
</tr>
<tr>
<td>Postgraduate degree</td>
<td>77</td>
<td>99</td>
</tr>
</tbody>
</table>

How does the proportion of the sample who favor banning possession of handguns differ among people with different levels of education? Make a bar graph that compares the proportions and briefly describe the relationship between education and opinion about a handgun ban.

6.7 Nonresponse in a survey. A business school conducted a survey of companies in its state. They mailed a questionnaire to 200 small companies, 200 medium-sized companies, and 200 large companies. The rate of nonresponse is important in deciding how reliable survey results are. Here are the data on response to this survey:

<table>
<thead>
<tr>
<th></th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>125</td>
<td>81</td>
<td>40</td>
</tr>
<tr>
<td>No response</td>
<td>75</td>
<td>119</td>
<td>160</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
</tbody>
</table>

(a) What was the overall percent of nonresponse?
(b) Describe how nonresponse is related to the size of the business. (Use percents to make your statements precise.)
(c) Draw a bar graph to compare the nonresponse percents for the three size categories.

6.8 Aspirin and heart attacks. Does taking aspirin regularly help prevent heart attacks? The Physicians’ Health Study tried to find out. The subjects were 22,071 healthy male doctors at least 40 years old. Half the subjects, chosen at random, took aspirin every other day. The other half took a placebo, a dummy pill that
looked and tasted like aspirin. Here are the results.28 (The row for “None of these” is left out of the two-way table.)

<table>
<thead>
<tr>
<th></th>
<th>Aspirin group</th>
<th>Placebo group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal heart attacks</td>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td>Other heart attacks</td>
<td>129</td>
<td>213</td>
</tr>
<tr>
<td>Strokes</td>
<td>119</td>
<td>98</td>
</tr>
<tr>
<td>Total</td>
<td>11,037</td>
<td>11,034</td>
</tr>
</tbody>
</table>

What do the data show about the association between taking aspirin and heart attacks and stroke? Use percents to make your statements precise. Do you think the study provides evidence that aspirin actually reduces heart attacks (cause and effect)?

**Chapter 8**

**8.1 The political gender gap.** There may be a “gender gap” in political party preference in the United States, with women more likely than men to prefer Democratic candidates. A political scientist interviews a large sample of registered voters, both men and women. She asks each voter whether they voted for the Democratic or the Republican candidate in the last congressional election. Is this study an experiment? Why or why not? What are the explanatory and response variables?

**8.2 Physical fitness and leadership.** A study of the relationship between physical fitness and leadership uses as subjects middle-aged executives who have volunteered for an exercise program. The executives are divided into a low-fitness group and a high-fitness group on the basis of a physical examination. All subjects then take a psychological test designed to measure leadership, and the results for the two groups are compared. Is this an observational study or an experiment? Explain your answer.

**8.3 Teaching reading.** An educator wants to compare the effectiveness of computer software that teaches reading with that of a standard reading curriculum. He tests the reading ability of each student in a class of fourth graders, then divides them into two groups. One group uses the computer regularly, while the other studies a standard curriculum. At the end of the year, he retests all the students and compares the increase in reading ability in the two groups. Is this an experiment? Why or why not? What are the explanatory and response variables?

**8.4 Public housing.** To study the effect of living in public housing on family stability in poverty-level households, researchers obtained a list of all applicants for public housing during the previous year. Some applicants had been accepted, while others had been turned down by the housing authority. Both groups were interviewed and compared. Was this an observational study or an experiment? Explain your answer. What are the explanatory and response variables?
8.5 Rating the police. The Miami Police Department wants to know how black residents of Miami feel about police service. A sociologist prepares several questions about the police. A sample of 300 mailing addresses in predominantly black neighborhoods is chosen, and a uniformed black police officer goes to each address to ask the questions of an adult living there. What are the population and the sample? Why are the results likely to be biased?

8.6 The effects of propaganda. In 1940, a psychologist conducted an experiment to study the effect of propaganda on attitude toward a foreign government. He administered a test of attitude toward the German government to a group of American students. After the students read German propaganda for several months, he tested them again to see if their attitudes had changed.

Unfortunately, Germany attacked and conquered France while the experiment was in progress. Explain clearly why confounding makes it impossible to determine the effect of reading the propaganda.

8.7 Treating breast cancer. What is the preferred treatment for breast cancer that is detected in its early stages? The most common treatment was once removal of the breast. It is now usual to remove only the tumor and nearby lymph nodes, followed by radiation. To study whether these treatments differ in their effectiveness, a medical team examines the records of 25 large hospitals and compares the survival times after surgery of all women who have had either treatment.

(a) What are the explanatory and response variables?
(b) Explain carefully why this study is not an experiment.
(c) Explain why confounding will prevent this study from discovering which treatment is more effective. (The current treatment was in fact recommended after a large randomized comparative experiment.)

8.8 What’s the population? Identify the population in each of the following sample surveys. You will have to use common sense to fill in missing details.
(a) An opinion poll contacts 1161 adults and asks, “Which political party do you think has better ideas for leading the country in the twenty-first century?”
(b) A machinery manufacturer purchases voltage regulators from a supplier. There are reports that variation in the output voltage of the regulators is affecting the performance of the finished products. To assess the quality of the supplier’s production, the manufacturer sends a sample of 5 regulators from the last shipment to a laboratory for study.

8.9 Letters to Congress. You are on the staff of a member of Congress who is considering a bill that would provide government-sponsored insurance for nursing home care. You report that 1128 letters have been received on the issue, of which 871 oppose the legislation. "I'm surprised that most of my constituents oppose the bill. I thought it would be quite popular," says the congresswoman. Are you convinced that a majority of the voters oppose the bill? How would you explain the statistical issue to the congresswoman?

8.10 Choose an SRS. Your class in ancient Ugaritic religion is poorly taught and wants to complain to the dean. The class decides to choose 4 of its members at
random to carry the complaint. The class list appears below. Choose an SRS of 4 using the table of random digits, beginning at line 145.

<table>
<thead>
<tr>
<th>Anderson</th>
<th>Gupta</th>
<th>Patnaik</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspin</td>
<td>Gutierrez</td>
<td>Pirelli</td>
</tr>
<tr>
<td>Bennett</td>
<td>Harter</td>
<td>Rao</td>
</tr>
<tr>
<td>Bock</td>
<td>Henderson</td>
<td>Rider</td>
</tr>
<tr>
<td>Breiman</td>
<td>Hughes</td>
<td>Robertson</td>
</tr>
<tr>
<td>Castillo</td>
<td>Johnson</td>
<td>Rodriguez</td>
</tr>
<tr>
<td>Dixon</td>
<td>Kempthorne</td>
<td>Sosa</td>
</tr>
<tr>
<td>Edwards</td>
<td>Laskowsky</td>
<td>Tran</td>
</tr>
<tr>
<td>Gonzalez</td>
<td>Liang</td>
<td>Trevino</td>
</tr>
<tr>
<td>Green</td>
<td>Olds</td>
<td>Wang</td>
</tr>
</tbody>
</table>

8.11 Choose an SRS. A manufacturer of chemicals chooses 3 containers from each lot of 25 containers of a reagent to test for purity and potency. Below are the control numbers stamped on the bottles in the current lot. Use Table B at line 111 to choose an SRS of 3 of these bottles.

A1096 A1097 A1098 A1101 A1108
A1112 A1113 A1117 A2109 A2211
A2220 B0986 B1011 B1096 B1101
B1102 B1103 B1110 B1119 B1137
B1189 B1223 B1277 B1286 B1299

8.12 Campaign contributions. Here are two wordings for the same question. The first question was asked by presidential candidate Ross Perot, and the second by a Time/CNN poll, both in March 1993.

A. Should laws be passed to eliminate all possibilities of special interests giving huge sums of money to candidates?

B. Should laws be passed to prohibit interest groups from contributing to campaigns, or do groups have a right to contribute to the candidates they support?

One of these questions drew 40% favoring banning contributions; the other drew 80% with this opinion. Which question produced the 40% and which got 80%? Explain why the results were so different.

8.13 Wording survey questions. Comment on each of the following as a potential sample survey question. Is the question clear? Is it slanted toward a desired response?

(a) Which of the following best represents your opinion on gun control?

1. The government should confiscate our guns.

2. We have the right to keep and bear arms.
(b) A freeze in nuclear weapons should be favored because it would begin a much-needed process to stop everyone in the world from building nuclear weapons now and reduce the possibility of nuclear war in the future. Do you agree or disagree?
(c) In view of escalating environmental degradation and incipient resource depletion, would you favor economic incentives for recycling of resource-intensive consumer goods?

8.14 A university has 2000 male and 500 female faculty members. A stratified random sample of 50 female and 200 male faculty members gives each member of the faculty 1 chance in 10 to be chosen. This sample design gives every individual in the population the same chance to be chosen for the sample. Is it an SRS? Explain your answer.

8.15 A university has 2000 male and 500 female faculty members. The equal opportunity employment officer wants to poll the opinions of a random sample of faculty members. In order to give adequate attention to female faculty opinion, he decides to choose a stratified random sample of 200 males and 200 females. He has alphabetized lists of female and male faculty members. Explain how you would assign labels and use random digits to choose the desired sample. Enter Table B at line 122 and give the labels of the first 5 females and the first 5 males in the sample.

8.16 Sampling college faculty. A labor organization wants to study the attitudes of college faculty members toward collective bargaining. These attitudes appear to differ depending on the type of college. The American Association of University Professors classifies colleges as follows:

- **Class I.** Offer doctorate degrees and award at least 15 per year.
- **Class IIA.** Award degrees above the bachelor’s but are not in Class I.
- **Class IIB.** Award no degrees beyond the bachelor’s.
- **Class III.** Two-year colleges.

Discuss the design of a sample of faculty from colleges in your state, with total sample size about 200.

8.17 Sampling employed women. A sociologist wants to know the opinions of employed adult women about government funding for day care. She obtains a list of the 520 members of a local business and professional women’s club and mails a questionnaire to 100 of these women selected at random. Only 48 questionnaires are returned. What is the population in this study? What is the sample from whom information is actually obtained? What is the rate (percent) of nonresponse?

8.18 You call the shots. A newspaper advertisement for USA Today: The Television Show once said:

> Should handgun control be tougher? You call the shots in a special call-in poll tonight. If yes, call 1-900-720-6181. If no, call 1-900-720-6182. Charge is 50 cents for the first minute.

Explain why this opinion poll is almost certainly biased.
8.19 Errors in a poll. A *New York Times* opinion poll on women’s issues contacted a sample of 1025 women and 472 men by randomly selecting telephone numbers. The *Times* publishes descriptions of its polling methods. Here is part of the description for this poll:

*In theory, in 19 cases out of 20 the results based on the entire sample will differ by no more than three percentage points in either direction from what would have been obtained by seeking out all adult Americans.*

*The potential sampling error for smaller subgroups is larger. For example, for men it is plus or minus five percentage points.*

Explain why the margin of error is larger for conclusions about men alone than for conclusions about all adults.

Chapter 9

9.1 Effects of herbal tea. Exercise 6.1 describes a study of the effects of herbal tea on the attitude of nursing home residents. Is this study an experiment? Why?

9.2 Treating drunk drivers. Once a person has been convicted of drunk driving, one purpose of court-mandated treatment or punishment is to prevent future offenses of the same kind. Suggest three different treatments that a court might require. Then outline the design of an experiment to compare their effectiveness. Be sure to specify the response variables you will measure.

9.3 Calcium and blood pressure. Some medical researchers suspect that added calcium in the diet reduces blood pressure. You have available 40 men with high blood pressure who are willing to serve as subjects.

(a) Outline an appropriate design for the experiment, taking the placebo effect into account.

(b) The names of the subjects appear below. Use Table B, beginning at line 119, to do the randomization required by your design, and list the subjects to whom you will give the drug.

<table>
<thead>
<tr>
<th>Alomar</th>
<th>Denman</th>
<th>Han</th>
<th>Liang</th>
<th>Rosen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asihiro</td>
<td>Durr</td>
<td>Howard</td>
<td>Maldonado</td>
<td>Solomon</td>
</tr>
<tr>
<td>Bennett</td>
<td>Edwards</td>
<td>Hruska</td>
<td>Marsden</td>
<td>Tompkins</td>
</tr>
<tr>
<td>Bikalis</td>
<td>Farouk</td>
<td>Imrani</td>
<td>Moore</td>
<td>Townsend</td>
</tr>
<tr>
<td>Chen</td>
<td>Fratianna</td>
<td>James</td>
<td>O’Brian</td>
<td>Tullock</td>
</tr>
<tr>
<td>Clemente</td>
<td>George</td>
<td>Kaplan</td>
<td>Ogle</td>
<td>Underwood</td>
</tr>
<tr>
<td>Cranston</td>
<td>Green</td>
<td>Krushchev</td>
<td>Plochman</td>
<td>Willis</td>
</tr>
<tr>
<td>Curtis</td>
<td>Guileen</td>
<td>Lawless</td>
<td>Rodriguez</td>
<td>Zhang</td>
</tr>
</tbody>
</table>

9.4 Does child care help recruit employees? Will providing child care for employees make a company more attractive to women, even those who are unmarried? You are designing an experiment to answer this question. You prepare recruiting
material for two fictitious companies, both in similar businesses in the same location. Company A’s brochure does not mention child care. There are two versions of Company B’s material, identical except that one describes the company’s on-site child-care facility. Your subjects are 40 unmarried women who are college seniors seeking employment. Each subject will read recruiting material for both companies and choose the one she would prefer to work for. You will give each version of Company B’s brochure to half the women. You expect that a higher percentage of those who read the description that includes child care will choose Company B.

(a) Outline an appropriate design for the experiment.

(b) The names of the subjects appear below. Use Table B, beginning at line 131, to do the randomization required by your design. List the subjects who will read the version that mentions child care.

<table>
<thead>
<tr>
<th>Abrams</th>
<th>Danielson</th>
<th>Gutierrez</th>
<th>Lippman</th>
<th>Rosen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adamson</td>
<td>Durr</td>
<td>Howard</td>
<td>Martinez</td>
<td>Sugiwara</td>
</tr>
<tr>
<td>Afifi</td>
<td>Edwards</td>
<td>Hwang</td>
<td>McNeill</td>
<td>Thompson</td>
</tr>
<tr>
<td>Brown</td>
<td>Fluharty</td>
<td>Iselin</td>
<td>Morse</td>
<td>Travers</td>
</tr>
<tr>
<td>Cansico</td>
<td>Garcia</td>
<td>Janle</td>
<td>Ng</td>
<td>Turing</td>
</tr>
<tr>
<td>Chen</td>
<td>Gerson</td>
<td>Kaplan</td>
<td>Quinones</td>
<td>Ullmann</td>
</tr>
<tr>
<td>Cortez</td>
<td>Green</td>
<td>Kim</td>
<td>Rivera</td>
<td>Williams</td>
</tr>
<tr>
<td>Curzakis</td>
<td>Gupta</td>
<td>Lattimore</td>
<td>Roberts</td>
<td>Wong</td>
</tr>
</tbody>
</table>

9.5 Sealing food packages. A manufacturer of food products uses package liners that are sealed at the top by applying heated jaws after the package is filled. The customer peels the sealed pieces apart to open the package. What effect does the temperature of the jaws have on the force needed to peel the liner? To answer this question, engineers obtain 20 pairs of pieces of package liner. They seal five pairs at each of 250° F, 275° F, 300° F, and 325° F. Then they measure the force needed to peel each seal.

(a) What are the individuals?

(b) There is one factor (explanatory variable). What is it, and what are its values?

(c) What is the response variable?

9.6 An industrial experiment. A chemical engineer is designing the production process for a new product. The chemical reaction that produces the product may have higher or lower yield, depending on the temperature and the stirring rate in the vessel in which the reaction takes place. The engineer decides to investigate the effects of combinations of two temperatures (50° C and 60° C) and three stirring rates (60 rpm, 90 rpm, and 120 rpm) on the yield of the process. She will process two batches of the product at each combination of temperature and stirring rate.

(a) What are the individuals and the response variable in this experiment?

(b) How many factors are there? How many treatments? Use a diagram like that in text Figure 9.1 to lay out the treatments.

(c) How many individuals are required for the experiment?

9.7 Sealing food packages, continued. Use a diagram to describe a completely randomized experimental design for the package liner experiment of Exercise 9.5.
Use software or Table B (starting at line 120) to do the randomization required by your design.

9.8 An industrial experiment, continued. Exercise 9.6 describes an industrial experiment and asks you to identify the treatments.
(a) Outline in graphic form the design of an appropriate experiment.
(b) The randomization in this experiment determines the order in which batches of the product will be processed according to each treatment. Use Table B, starting at line 128, to carry out the randomization and state the result.

9.9 Comparing corn varieties. New varieties of corn with altered amino acid content may have higher nutritional value than standard corn, which is low in the amino acid lysine. An experiment compares two new varieties, called opaque-2 and floury-2, with normal corn. The researchers mix corn-soybean meal diets using each type of corn at each of three protein levels: 12% protein, 16% protein, and 20% protein. They feed each diet to 10 one-day-old male chicks and record their weight gains after 21 days. The weight gain of the chicks is a measure of the nutritional value of their diet.
(a) What are the individuals and the response variable in this experiment?
(b) How many factors are there? How many treatments? Use a diagram like text Figure 9.1 to describe the treatments. How many individuals does the experiment require?
(c) Use a diagram to describe a completely randomized design for this experiment. (You do not need to actually do the randomization.)

9.10 Reducing risky sex. The National Institute of Mental Health (NIMH) wants to know whether intense education about the risks of AIDS will help change the behavior of people who report sexual activities that put them at risk of infection. NIMH investigators screened 38,893 people and identified 3706 suitable subjects. The subjects were assigned to a control group (1855 people) or an intervention group (1851 people). The control group attended a one-hour AIDS education session; the intervention group attended seven single-sex discussion sessions, each lasting 90 to 120 minutes. After 12 months, 64% of the intervention group and 52% of the control group said they used condoms. (None of the subjects used condoms regularly before the study began.)
(a) Because none of the subjects used condoms when the study started, we might just offer the intervention sessions and find that 64% used condoms 12 months after the sessions. Explain why this greatly overstates the effectiveness of the intervention.
(b) Outline the design of this experiment.
(c) You must randomly assign 3706 subjects. How would you label them? Use line 119 of Table B to choose the first 5 subjects for the intervention group.

9.11 Treating prostate disease. A large study used records from Canada’s national health care system to compare the effectiveness of two ways to treat prostate disease. The two treatments are traditional surgery and a new method that does not require surgery. The records described many patients whose doctors had chosen each method. The study found that patients treated by the new method were significantly more likely to die within 8 years.
(a) Further study of the data showed that this conclusion was wrong. The extra
deaths among patients who got the new method could be explained by lurking vari-
ables. What lurking variables might be confounded with a doctor's choice of surgical
or nonsurgical treatment?
(b) You have 300 prostate patients who are willing to serve as subjects in an ex-
periment to compare the two methods. Use a diagram to outline the design of a
randomized comparative experiment. (When using a diagram to outline the de-
sign of an experiment, be sure to indicate the size of the treatment groups and the
response variable.)

9.12 Marketing to children. If children are given more choices within a class of
products, will they tend to prefer that product to a competing product that offers
fewer choices? Marketers want to know. An experiment prepared three sets of
beverages. Set 1 contained two milk drinks and two fruit drinks. Set 2 had two fruit
drinks and four milk drinks. Set 3 contained four fruit drinks but only two milk
drinks. The researchers divided 210 children aged 4 to 12 years into 3 groups at
random. They offered each group one of the sets. As each child chose a beverage to
drink from the set presented, the researchers noted whether the choice was a milk
drink or a fruit drink.
(a) What are the experimental subjects?
(b) What is the factor and what are its values? What is the response variable?
(c) Use a diagram to outline a completely randomized design for the study.
(d) Explain how you would assign labels to the subjects. Use Table B at line 125 to
choose only the first 5 subjects assigned to the first treatment.

9.13 Reading a medical journal. The article in the New England Journal of
Medicine that presents the final results of the Physicians’ Health Study begins
with these words: “The Physicians’ Health Study is a randomized, double-blind,
placebo-controlled trial designed to determine whether low-dose aspirin (325 mg ev-
every other day) decreases cardiovascular mortality and whether beta carotene reduces
the incidence of cancer.” Doctors are expected to understand this. Explain to a
doctor who knows no statistics what “randomized,” “double-blind,” and “placebo-
controlled” mean.

9.14 Fizz Labs need help. Fizz Laboratories, a pharmaceutical company, has
developed a new pain-relief medication. Sixty patients suffering from arthritis and
needing pain relief are available. Each patient will be treated and asked an hour
later, “About what percentage of pain relief did you experience?”
(a) Why should Fizz not simply administer the new drug and record the patients’
responses?
(b) Outline the design of an experiment to compare the drug’s effectiveness with
that of aspirin and of a placebo.
(c) Should patients be told which drug they are receiving? How would this knowledge
probably affect their reactions?
(d) If patients are not told which treatment they are receiving, the experiment is
single-blind. Should this experiment be double-blind also? Explain.
9.15 **Temperature and work performance.** An industrial psychologist is interested in the effect of room temperature on the performance of tasks requiring manual dexterity. She chooses temperatures of 70° F and 90° F as treatments. The response variable is the number of correct insertions, during a 30-minute period, in a peg-and-hole apparatus that requires the use of both hands simultaneously. Each subject is trained on the apparatus and then asked to make as many insertions as possible in 30 minutes of continuous effort.

(a) Outline a completely randomized design to compare dexterity at 70° and 90°. Twenty subjects are available.

(b) Because individuals differ greatly in dexterity, the wide variation in individual scores may hide the systematic effect of temperature unless there are many subjects in each group. Describe in detail the design of a matched pairs experiment in which each subject serves as his or her own control.

9.16 **The culture of Mexican Americans.** There are several psychological tests that measure the extent to which Mexican Americans are oriented toward Mexican/Spanish or Anglo/English culture. Two such tests are the Bicultural Inventory (BI) and the Acculturation Rating Scale for Mexican Americans (ARSMA). To study the correlation between the scores on these two tests, researchers will give both tests to a group of 22 Mexican Americans.

(a) Briefly describe a matched pairs design for this study. In particular, how will you use randomization in your design?

(b) You have an alphabetized list of the subjects (numbered 1 to 22). Carry out the randomization required by your design and report the result.

9.17 **McDonald’s versus Wendy’s.** Do consumers prefer the taste of a cheeseburger from McDonald’s or from Wendy’s in a blind test in which neither burger is identified? Describe briefly the design of a matched pairs experiment to investigate this question.

9.18 **Comparing weight-loss treatments.** Twenty overweight females have agreed to participate in a study of the effectiveness of 4 weight-loss treatments: A, B, C, and D. The researcher first calculates how overweight each subject is by comparing the subject’s actual weight with her “ideal” weight. The subjects and their excess weights in pounds are

<table>
<thead>
<tr>
<th>Subject</th>
<th>Excess Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birnbaum</td>
<td>35</td>
</tr>
<tr>
<td>Brown</td>
<td>34</td>
</tr>
<tr>
<td>Brunk</td>
<td>30</td>
</tr>
<tr>
<td>Cruz</td>
<td>34</td>
</tr>
<tr>
<td>Deng</td>
<td>24</td>
</tr>
<tr>
<td>Hernandez</td>
<td>25</td>
</tr>
<tr>
<td>Jackson</td>
<td>33</td>
</tr>
<tr>
<td>Kendall</td>
<td>30</td>
</tr>
<tr>
<td>Loren</td>
<td>32</td>
</tr>
<tr>
<td>Mann</td>
<td>28</td>
</tr>
<tr>
<td>Moses</td>
<td>25</td>
</tr>
<tr>
<td>Nevesky</td>
<td>39</td>
</tr>
<tr>
<td>Obrach</td>
<td>30</td>
</tr>
<tr>
<td>Rodriguez</td>
<td>30</td>
</tr>
<tr>
<td>Santiago</td>
<td>28</td>
</tr>
<tr>
<td>Smith</td>
<td>25</td>
</tr>
<tr>
<td>Stall</td>
<td>39</td>
</tr>
<tr>
<td>Tran</td>
<td>35</td>
</tr>
<tr>
<td>Wilansky</td>
<td>42</td>
</tr>
<tr>
<td>Williams</td>
<td>27</td>
</tr>
<tr>
<td>Williams</td>
<td>22</td>
</tr>
</tbody>
</table>

The response variable is the weight lost after 8 weeks of treatment. Because a subject’s excess weight will influence the response, a block design is appropriate.

(a) Arrange the subjects in order of increasing excess weight. Form 5 blocks of 4 subjects each by grouping the 4 least overweight, then the next 4, and so on.

(b) Use Table B to randomly assign the 4 subjects in each block to the 4 weight-loss treatments. Be sure to explain exactly how you used the table.
9.19 **Calcium and blood pressure.** You are participating in the design of a medical experiment to investigate whether a calcium supplement in the diet will reduce the blood pressure of middle-aged men. Preliminary work suggests that calcium may be effective and that the effect may be greater for black men than for white men.

(a) Outline in graphical form a block design for this experiment.
(b) Choosing the sizes of the treatment groups requires more statistical expertise. We will learn more about this aspect of design in later chapters. Explain in plain language the advantage of using larger groups of subjects.

9.20 **Statistical significance.** 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 was found between the sexes, with men earning more on the average. No significant difference was found between the earnings of black and white students.” Explain the meaning of “a significant difference” and “no significant difference” in plain language.

9.21 **Statistical significance.** A randomized comparative experiment examines whether a calcium supplement in the diet reduces the blood pressure of healthy men. The subjects receive either a calcium supplement or a placebo for 12 weeks. The researchers conclude that “the blood pressure of the calcium group was significantly lower than that of the placebo group.” “Significant” in this conclusion means statistically significant. Explain what statistically significant means in the context of this experiment, as if you were speaking to a doctor who knows no statistics.

---

**Chapter 10**

10.1 **Sample space.** In each of the following situations, describe a sample space $S$ for the random phenomenon. In some cases you have some freedom in specifying $S$, especially in setting the largest and the smallest value in $S$.

(a) A study of healthy adults measures the weights of adult women.
(b) The Physicians’ Health Study asked 11,000 physicians to take an aspirin every other day and observed how many of them had a heart attack in a five-year period.
(c) In a test of a new package design, you drop a carton of a dozen eggs from a height of 1 foot and count the number of broken eggs.

10.2 **Sample space.** In each of the following situations, describe a sample space $S$ for the random phenomenon. In some cases, you have some freedom in your choice of $S$.

(a) A seed is planted in the ground. It either germinates or fails to grow.
(b) A patient with a usually fatal form of cancer is given a new treatment. The response variable is the length of time that the patient lives after treatment.
(c) A nutrition researcher feeds a new diet to a young male white rat. The response variable is the weight (in grams) that the rat gains in 8 weeks.
10.3 Got money? Choose a student at random and record the number of dollars in bills (ignore change) that he or she is carrying. Give a reasonable sample space $S$ for this random phenomenon. (We don’t know the largest amount that a student could reasonably carry, so you will have to make a choice in stating the sample space.)

10.4 Forests in Missouri. What happens to trees over a five-year period? A study lasting more than 30 years found these probabilities for a randomly chosen 12-inch-diameter tree in the Ozark Highlands of Missouri: stay in the 12-inch class, 0.686; move to the 14-inch class, 0.256. The remaining trees die during the five-year period. What is the probability that a tree dies?

10.5 Moving up. A sociologist studying social mobility finds that the probability that the son of a lower-class father remains in the lower class is 0.46. What is the probability that the son moves to one of the higher classes?

10.6 Causes of death. Government data assign a single cause for each death that occurs in the United States. The data show that the probability is 0.45 that a randomly chosen death was due to cardiovascular (mainly heart) disease, and 0.22 that it was due to cancer. What is the probability that a death was due either to cardiovascular disease or to cancer? What is the probability that the death was due to some other cause?

10.7 High school academic rank. Select a first-year college student at random and ask what his or her academic rank was in high school. Here are the probabilities, based on proportions from a large sample survey of first-year students:

<table>
<thead>
<tr>
<th>Rank</th>
<th>Top 20%</th>
<th>Second 20%</th>
<th>Third 20%</th>
<th>Fourth 20%</th>
<th>Lowest 20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.41</td>
<td>0.23</td>
<td>0.29</td>
<td>0.06</td>
<td>0.01</td>
</tr>
</tbody>
</table>

(a) What is the sum of these probabilities? Why do you expect the sum to have this value?
(b) What is the probability that a randomly chosen first-year college student was not in the top 20% of his or her high school class?
(c) What is the probability that a first-year student was in the top 40% in high school?

10.8 Legitimate probabilities? In each of the following situations, state whether or not the given assignment of probabilities to individual outcomes is a legitimate discrete probability model. If not, give specific reasons for your answer.
(a) When a coin is spun, $P(H) = 0.55$ and $P(T) = 0.45$.
(b) When two coins are tossed, $P(HH) = 0.4$, $P(HT) = 0.4$, $P(TH) = 0.4$, and $P(TT) = 0.4$.
(c) When a die is rolled, the number of spots on the up-face has $P(1) = 1/2$, $P(4) = 1/6$, $P(5) = 1/6$, and $P(6) = 1/6$

10.9 Rolling a soft 4. A “soft 4” in rolling two dice is a roll of 1 on one die and 3 on the other. If you roll two dice, what is the probability of rolling a soft 4? Of rolling a 4?
10.10 Classifying occupations. Choose an American worker at random and classify his or her occupation into one of the following classes. These classes are used in government employment data.

A Managerial and professional
B Technical, sales, administrative support
C Service occupations
D Precision production, craft, and repair
E Operators, fabricators, and laborers
F Farming, forestry, and fishing

The table below gives the probabilities that a randomly chosen worker falls into each of 12 sex-by-occupation classes:

<table>
<thead>
<tr>
<th>Class</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.14</td>
<td>0.11</td>
<td>0.06</td>
<td>0.11</td>
<td>0.12</td>
<td>0.03</td>
</tr>
<tr>
<td>Female</td>
<td>0.09</td>
<td>0.20</td>
<td>0.08</td>
<td>0.01</td>
<td>0.04</td>
<td>0.01</td>
</tr>
</tbody>
</table>

(a) Verify that this is a legitimate assignment of probabilities to these outcomes.
(b) What is the probability that the worker is female?
(c) What is the probability that the worker is not engaged in farming, forestry, or fishing?
(d) Classes D and E include most mechanical and factory jobs. What is the probability that the worker holds a job in one of these classes?
(e) What is the probability that the worker does not hold a job in Classes D or E?

10.11 Predicting the ACC champion. Las Vegas Zeke, when asked to predict the Atlantic Coast Conference basketball champion, follows the modern practice of giving probabilistic predictions. He says, “North Carolina’s probability of winning is twice Duke’s. North Carolina State and Virginia each have probability 0.1 of winning, but Duke’s probability is three times that. Nobody else has a chance.” Has Zeke given a legitimate assignment of probabilities to the eight teams in the conference? Explain your answer.

10.12 Living in San Jose. Let the random variable $X$ be the number of rooms in a randomly chosen owner-occupied housing unit in San Jose, California. Here is the distribution of $X$:

<table>
<thead>
<tr>
<th>Rooms $X$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.003</td>
<td>0.002</td>
<td>0.023</td>
<td>0.104</td>
<td>0.210</td>
<td>0.224</td>
<td>0.197</td>
<td>0.149</td>
<td>0.053</td>
<td>0.035</td>
</tr>
</tbody>
</table>

(a) Is the random variable $X$ discrete or continuous? Why?
(b) Express “the unit has no more than 2 rooms” in terms of $X$. What is the probability of this event?
(c) Express the event $\{X > 5\}$ in words. What is its probability?

10.13 Living in San Jose, continued. The previous exercise gives the distribution of the number of rooms in owner-occupied housing in San Jose, California. Here is the distribution for rented housing:
What are the most important differences between the distributions of the random variables $X$ and $Y$ (owner-occupied and rented housing)? Compare at least two probabilities for $X$ and $Y$ to justify your answer.

**10.14 How big are farms?** Choose an American farm at random and measure its size in acres. Here are the probabilities that the farm chosen falls in several acreage categories:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.09</td>
<td>0.20</td>
<td>0.15</td>
<td>0.16</td>
<td>0.22</td>
<td>0.09</td>
<td>0.05</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Let $A$ be the event that the farm is less than 50 acres in size, and let $B$ be the event that it is 500 acres or more.

(a) Find $P(A)$ and $P(B)$.

(b) Describe the event “$A$ does not occur” in words and find its probability by Rule 3.

(c) Describe the event “$A$ or $B$” in words and find its probability by the addition rule.

**10.15 How large are households?** Choose an American household at random and let the random variable $X$ be the number of persons living in the household. If we ignore the few households with more than seven inhabitants, the probability distribution of $X$ is as follows:

<table>
<thead>
<tr>
<th>Inhabitants</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.25</td>
<td>0.32</td>
<td>0.17</td>
<td>0.15</td>
<td>0.07</td>
<td>0.03</td>
<td>0.01</td>
</tr>
</tbody>
</table>

(a) Verify that this is a legitimate probability distribution.

(b) What is $P(X \geq 5)$?

(c) What is $P(X > 5)$?

(d) What is $P(2 < X \leq 4)$?

(e) What is $P(X \neq 1)$?

(f) Write the event that a randomly chosen household contains more than two persons in terms of the random variable $X$. What is the probability of this event?

**10.16 How far do fifth graders go in school?** A study of education followed a large group of fifth-grade children to see how many years of school they eventually completed. Let $X$ be the highest year of school that a randomly chosen fifth grader completes. (Students who go on to college are included in the outcome $X = 12$.) The study found this probability distribution for $X$:

<table>
<thead>
<tr>
<th>Years</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.010</td>
<td>0.007</td>
<td>0.007</td>
<td>0.013</td>
<td>0.032</td>
<td>0.068</td>
<td>0.070</td>
<td>0.041</td>
<td>0.752</td>
</tr>
</tbody>
</table>

(a) What percent of fifth graders eventually finished twelfth grade?

(b) Check that this is a legitimate probability distribution.
(c) Find $P(X \geq 6)$. (Be careful: the event “$X \geq 6$” includes the value 6.)
(d) Find $P(X > 6)$.
(e) What values of $X$ make up the event “the student completed at least one year of high school”? (High school begins with the ninth grade.) What is the probability of this event?

10.17 Polling women. Suppose that 47% of all adult women think they do not get enough time for themselves. An opinion poll interviews 1025 randomly chosen women and records the sample proportion who don’t feel they get enough time for themselves. Call this proportion $V$. It will vary from sample to sample if the poll is repeated. The distribution of the random variable $V$ is approximately Normal with mean 0.47 and standard deviation about 0.016. Sketch this Normal curve and use it to answer the following questions.
(a) The truth about the population is 0.47. In what range will the middle 95% of all sample results fall?
(b) What is the probability that the poll gets a sample in which fewer than 45% say they do not get enough time for themselves?

Chapter 11

11.1 Apartment rents. 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. Is each of the boldface numbers a parameter or a statistic?

11.2 Rat weights. A researcher carries out a randomized comparative experiment with young rats to investigate the effects of a toxic compound in food. She feeds the control group a normal diet. The experimental group receives a diet with 2500 parts per million of the toxic material. After 8 weeks, the mean weight gain is 335 grams for the control group and 289 grams for the experimental group. Is each of the boldface numbers a parameter or a statistic?

11.3 Unlisted numbers. A telemarketing firm in Los Angeles uses a device that dials residential telephone numbers in that city at random. Of the first 100 numbers dialed, 48% are unlisted. This is not surprising because 52% of all Los Angeles residential phones are unlisted. Is each of the boldface numbers a parameter or a statistic?

11.4 Measuring blood cholesterol. The distribution of blood cholesterol level in the population of young men aged 20 to 34 years is close to Normal, with mean $\mu = 188$ milligrams per deciliter (mg/dl) and standard deviation $\sigma = 41$ mg/dl. You measure the cholesterol level of 100 young men chosen at random and calculate the mean $\bar{x}$.
(a) If you did this many times, what would be the mean and standard deviation of the distribution of all the $\bar{x}$-values?
(b) What is the probability that your sample has mean $\bar{x}$ less than 180?
11.5 Dust in coal mines. A laboratory weighs filters from a coal mine to measure the amount of dust in the mine atmosphere. Repeated measurements of the weight of dust on the same filter vary Normally with standard deviation $\sigma = 0.08$ milligram (mg) because the weighing is not perfectly precise. The dust on a particular filter actually weighs 123 mg. Repeated weighings will then have the Normal distribution with mean 123 mg and standard deviation 0.08 mg.

(a) The laboratory reports the mean of 3 weighings. What is the distribution of this mean?

(b) What is the probability that the laboratory reports a weight of 124 mg or higher for this filter?

11.6 Potassium in the blood. Judy’s doctor is concerned that she may suffer from hypokalemia (low potassium in the blood). There is variation both in the actual potassium level and in the blood test that measures the level. Judy’s measured potassium level varies according to the Normal distribution with $\mu = 3.8$ and $\sigma = 0.2$. A patient is classified as hypokalemic if the potassium level is below 3.5.

(a) If a single potassium measurement is made, what is the probability that Judy is diagnosed as hypokalemic?

(b) If measurements are made instead on 4 separate days and the mean result is compared with the criterion 3.5, what is the probability that Judy is diagnosed as hypokalemic?

11.7 ACT scores. The scores of students on the ACT college entrance examination in 2001 had mean $\mu = 21.0$ and standard deviation $\sigma = 4.7$. The distribution of scores is only roughly Normal.

(a) What is the approximate probability that a single student randomly chosen from all those taking the test scores 23 or higher?

(b) Now take an SRS of 50 students who took the test. What are the mean and standard deviation of the sample mean score $\bar{x}$ of these 50 students? What is the approximate probability that the mean score $\bar{x}$ of these students is 23 or higher?

(c) Which of your two Normal probability calculations in (a) and (b) is more accurate? Why?

11.8 Flaws in carpets. The number of flaws per square yard in a type of carpet material varies with mean 1.6 flaws per square yard and standard deviation 1.2 flaws per square yard. The population distribution cannot be Normal, because a count takes only whole-number values. An inspector samples 200 square yards of the material, records the number of flaws found in each square yard, and calculates $\bar{x}$, the mean number of flaws per square yard inspected. Use the central limit theorem to find the approximate probability that the mean number of flaws exceeds 2 per square yard.

11.9 The cost of Internet access. The amount that households pay service providers for access to the Internet varies quite a bit, but the mean monthly fee is $28 and the standard deviation is $10. The distribution is not Normal: many households pay about $10 for limited dial-up access or about $25 for unlimited dial-up access, but some pay much more for faster connections. A sample survey asks
an SRS of 500 households with Internet access how much they pay. What is the probability that the average fee paid by the sample households exceeds $29?

11.10 How many people in a car? A study of rush-hour traffic in San Francisco counts the number of people in each car entering a freeway at a suburban interchange. Suppose that this count has mean 1.5 and standard deviation 0.75 in the population of all cars that enter at this interchange during rush hours.
(a) Could the exact distribution of the count be Normal? Why or why not?
(b) Traffic engineers estimate that the capacity of the interchange is 700 cars per hour. According to the central limit theorem, what is the approximate distribution of the mean number of persons $\bar{x}$ in 700 randomly selected cars at this interchange?
(c) What is the probability that 700 cars will carry more than 1075 people? (Hint: Restate this event in terms of the mean number of people $\bar{x}$ per car.)

11.11 Testing kindergarten children. Children in kindergarten are sometimes given the Ravin Progressive Matrices Test (RPMT) to assess their readiness for learning. Experience at Southwark Elementary School suggests that the RPMT scores for its kindergarten pupils have mean 13.6 and standard deviation 3.1. The distribution is close to Normal. Mr. Lavin has 22 children in his kindergarten class this year. He suspects that their RPMT scores will be unusually low because the test was interrupted by a fire drill. To check this suspicion, he wants to find the level $L$ such that there is probability only 0.05 that the mean score of 22 children falls below $L$ when the usual Southwark distribution remains true. What is the value of $L$? (Hint: This requires a backward Normal calculation.)

Chapter 12

12.1 Bright lights? A string of holiday lights contains 20 lights. The lights are wired in series, so that if any light fails, the whole string will go dark. Each light has probability 0.02 of failing during the holiday season. The lights fail independently of each other. What is the probability that the string of lights will remain bright?

12.2 Nonconforming chips. An automobile manufacturer buys computer chips from a supplier. The supplier sends a shipment of which 5% fail to conform to performance specifications. Each chip chosen from this shipment has probability 0.05 of being nonconforming, and each automobile uses 12 chips selected independently. What is the probability that all 12 chips in a car will work properly?

12.3 Albinism. The gene for albinism in humans is recessive. That is, carriers of this gene have probability $1/2$ of passing it to a child, and the child is albino only if both parents pass the albinism gene. Parents pass their genes independently of each other. If both parents carry the albinism gene, what is the probability that their first child is albino? If they have two children (who inherit independently of each other), what is the probability that both are albino? That neither is albino?

12.4 Detecting steroids. An athlete suspected of having used steroids is given two tests that operate independently of each other. Test A has probability 0.9 of
being positive if steroids have been used. Test B has probability 0.8 of being positive if steroids have been used. What is the probability that neither test is positive if steroids have been used?

12.5 Military strategy. A general can plan a campaign to fight one major battle or three small battles. He believes that he has probability 0.6 of winning the large battle and probability 0.8 of winning each of the small battles. Victories or defeats in the small battles are independent. The general must win either the large battle or all three small battles to win the campaign. Which strategy should he choose?

12.6 Psychologists roll the dice. A six-sided die has four green and two red faces and is balanced so that each face is equally likely to come up. The die will be rolled several times. You must choose one of the following three sequences of colors:

- RGRRR
- RGRRRG
- GRRRRR

You will win $25 if the first rolls of the die give the sequence that you have chosen.

(a) What is the probability that one roll gives green? That it gives red?
(b) What is the probability of each of the sequences of colors given above? Which sequence would you choose in an attempt to win the $25? Why? (In a psychological experiment, 63% of 260 students who had not studied probability chose the second sequence. This is evidence that our intuitive understanding of probability is not very accurate.)

12.7 High school rank. Select a first-year college student at random and ask what his or her academic rank was in high school. Here are the probabilities, based on proportions from a large sample survey of first-year students:

<table>
<thead>
<tr>
<th>Rank</th>
<th>Top 20%</th>
<th>Second 20%</th>
<th>Third 20%</th>
<th>Fourth 20%</th>
<th>Lowest 20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.41</td>
<td>0.23</td>
<td>0.29</td>
<td>0.06</td>
<td>0.01</td>
</tr>
</tbody>
</table>

(a) Choose two first-year college students at random. Why is it reasonable to assume that their high school ranks are independent?
(b) What is the probability that both were in the top 20% of their high school classes?
(c) What is the probability that the first was in the top 20% and the second was in the lowest 20%?

12.8 Construction bidding. Consolidated Builders has bid on two large construction projects. The company president believes that the probability of winning the first contract (event $A$) is 0.6, that the probability of winning the second (event $B$) is 0.4, and that the probability of winning both jobs (event $\{A \text{ and } B\}$) is 0.2.

(a) What is the probability of the event $\{A \text{ or } B\}$ that Consolidated will win at least one of the jobs?
(b) You hear that Consolidated won the second job. Given this information, what is the conditional probability that Consolidated won the first job?

12.9 Construction bidding, continued. In the setting of the previous exercise, are events $A$ and $B$ independent? Do a calculation that proves your answer.
12.10 Construction bidding, continued. Draw a Venn diagram that illustrates the relation between events $A$ and $B$ in Exercise 12.8. Indicate the events below on your diagram and use the information in Exercise 12.8 to calculate their probabilities.
(a) Consolidated wins both jobs.
(b) Consolidated wins the first job but not the second.
(c) Consolidated does not win the first job but does win the second.
(d) Consolidated does not win either job.

12.11 Classifying occupations. Exercise 10.10 gives the probability distribution of the gender and occupation of a randomly chosen American worker. Use this distribution to answer the following questions:
(a) Given that the worker chosen holds a managerial (Class A) job, what is the conditional probability that the worker is female?
(b) Classes D and E include most mechanical and factory jobs. What is the conditional probability that a worker is female, given that he or she holds a job in one of these classes?
(c) Are gender and job type independent? How do you know?

12.12 Inspecting switches. A shipment contains 10,000 switches. Of these, 1000 are bad. An inspector draws switches at random, so that each switch has the same chance to be drawn.
(a) Draw one switch. What is the probability that the switch you draw is bad? What is the probability that it is not bad?
(b) Suppose the first switch drawn is bad. How many switches remain? How many of them are bad? Draw a second switch at random. What is the conditional probability that this switch is bad?
(c) Answer the questions in (b) again, but now suppose that the first switch drawn is not bad.
Comment: Knowing the result of the first trial changes the conditional probability for the second trial, so the trials are not independent. But because the shipment is large, the probabilities change very little. The trials are almost independent.

12.13 Employment data. In the language of government statistics, the “labor force” includes all civilians at least 16 years of age who are working or looking for work. Select a member of the U.S. labor force at random. Let $A$ be the event that the person selected is white, and $B$ the event that he or she is employed. Suppose that 84.6% of the labor force is white. Of the whites in the labor force, 95.1% are employed. Among nonwhite members of the labor force, 91.9% are employed.
(a) Express each of the percents given as a probability involving the events $A$ and $B$; for example, $P(A) = 0.846$.
(b) Are the events “an employed person is chosen” and “a white is chosen” independent? How do you know?
(c) Find the probability that the person chosen is an employed white.
(d) Find the probability that an employed nonwhite is chosen.
(e) Find the probability that the person chosen is employed. (Hint: An employed person is either white or nonwhite.)
12.14 Woman managers. Choose an employed person at random. Let \( A \) be the event that the person chosen is a woman, and \( B \) the event that the person holds a managerial or professional job. Government data tell us that \( P(A) = 0.47 \) and the probability of managerial and professional jobs among women is \( P(B \mid A) = 0.32 \). Find the probability that a randomly chosen employed person is a woman holding a managerial or professional position.

12.15 Buying from Japan. Functional Robotics Corporation buys electrical controllers from a Japanese supplier. The company’s treasurer thinks that there is probability 0.4 that the dollar will fall in value against the Japanese yen in the next month. The treasurer also believes that if the dollar falls there is probability 0.8 that the supplier will demand renegotiation of the contract. What personal probability has the treasurer assigned to the event that the dollar falls and the supplier demands renegotiation?

12.16 Tastes in music. Musical styles other than rock and pop are becoming more popular. A survey of college students finds that 40% like country music, 30% like gospel music, and 10% like both.
(a) What is the conditional probability that a student likes gospel music if we know that he or she likes country music?
(b) What is the conditional probability that a student who does not like country music likes gospel music? (A Venn diagram may help you.)

12.17 Caffeine in the diet. Common sources of caffeine are coffee, tea, and cola drinks. Suppose that

- 55% of adults drink coffee
- 25% of adults drink tea
- 45% of adults drink cola

and also that

- 15% drink both coffee and tea
- 5% drink all three beverages
- 25% drink both coffee and cola
- 5% drink only tea

Draw a Venn diagram marked with this information. Use it along with the addition rules to answer the following questions.
(a) What percent of adults drink only cola?
(b) What percent drink none of these beverages?

12.18 Prosperity and education. Call a household prosperous if its income exceeds $100,000. Call the household educated if the householder completed college. Select an American household at random, and let \( A \) be the event that the selected household is prosperous and \( B \) the event that it is educated. According to the Current Population Survey, \( P(A) = 0.138 \), \( P(B) = 0.261 \), and the probability that a household is both prosperous and educated is \( P(A \text{ and } B) = 0.082 \).
(a) Draw a Venn diagram that shows the relation between the events \( A \) and \( B \).
What is the probability \( P(A \text{ or } B) \) that the household selected is either prosperous
or educated? 
(b) In your diagram, shade the event that the household is educated but not prosperous. What is the probability of this event?

12.19 Prosperity and education. Call a household prosperous if its income exceeds $100,000. Call the household educated if the householder completed college. Select an American household at random, and let $A$ be the event that the selected household is prosperous and $B$ the event that it is educated. According to the Current Population Survey, $P(A) = 0.134$, $P(B) = 0.254$, and the probability that a household is both prosperous and educated is $P(A \text{ and } B) = 0.080$.
(a) Find the conditional probability that a household is educated, given that it is prosperous.
(b) Find the conditional probability that a household is prosperous, given that it is educated.
(c) Are events $A$ and $B$ independent? How do you know?

12.20 Preparing for the GMAT. A company that offers courses to prepare would-be MBA students for the Graduate Management Admission Test (GMAT) finds that 40% of its customers are currently undergraduate students and 60% are college graduates. After completing the course, 50% of the undergraduates and 70% of the graduates achieve scores of at least 600 on the GMAT. Draw a tree diagram that organizes this information. What percent of all customers score at least 600 on the GMAT?

12.21 Buying from Japan. In the setting of Exercise 12.15, the treasurer also thinks that if the dollar does not fall, there is probability 0.2 that the supplier will demand that the contract be renegotiated. What is the probability that the supplier will demand renegotiation? (Use a tree diagram to organize the information given.)

12.22 Albinism. People with albinism have little pigment in their skin, hair, and eyes. The gene that governs albinism has two forms (called alleles), which we denote by $a$ and $A$. Each person has a pair of these genes, one inherited from each parent. A child inherits one of each parent’s two alleles, independently with probability 0.5. Albinism is a recessive trait, so a person is albino only if the inherited pair is $aa$.
(a) Beth’s parents are not albino but she has an albino brother. This implies that both of Beth’s parents have type $Aa$. Why?
(b) Which of the types $aa$, $Aa$, and $AA$ could a child of Beth’s parents have? What is the probability of each type?

12.23 Albinism, continued. Beth is not albino. What are the conditional probabilities for Beth’s possible genetic types, given this fact? (Use the definition of conditional probability and results from the previous exercise.)
Chapter 13

13.1 Binomial setting? In each of the following cases, decide whether or not a binomial distribution is an appropriate model, and give your reasons.
(a) Fifty students are taught about binomial distributions by a television program. After completing their study, all students take the same examination. The number of students who pass is counted.
(b) A chemist repeats a solubility test 10 times on the same substance. Each test is conducted at a temperature $10^\circ$ higher than the previous test. She counts the number of times that the substance dissolves completely.

13.2 Binomial? You observe the sex of the next 20 children born at a local hospital; $X$ is the number of girls among them. Does $X$ have a binomial distribution?

13.3 Binomial? A couple decides to continue to have children until their first girl is born; $X$ is the total number of children the couple has. Does $X$ have a binomial distribution?

13.4 Random digits. Each entry in a table of random digits like Table B has probability 0.1 of being a 0, and digits are independent of each other.
(a) What is the probability that a group of five digits from the table will contain at least one 0?
(b) What is the mean number of 0s in lines 40 digits long?

13.5 Universal donors. People with type O-negative blood are universal donors whose blood can be safely given to anyone. Only 7.2% of the population have O-negative blood. A blood center is visited by 20 donors in an afternoon. What is the probability that there are at least 2 universal donors among them?

13.6 Are we shipping on time? Your mail-order company advertises that it ships 90% of its orders within three working days. You select an SRS of 100 of the 5000 orders received in the past week for an audit. The audit reveals that 86 of these orders were shipped on time.
(a) If the company really ships 90% of its orders on time, what is the probability that 86 or fewer in an SRS of 100 orders are shipped on time?
(b) A critic says, “Aha! You claim 90%, but in your sample the on-time percentage is only 86%. So the 90% claim is wrong.” Explain in simple language why your probability calculation in (a) shows that the result of the sample does not refute the 90% claim.

13.7 Hispanic representation. A factory employs several thousand workers, of whom 30% are Hispanic. If the 15 members of the union executive committee were chosen from the workers at random, the number of Hispanics on the committee would have the binomial distribution with $n = 15$ and $p = 0.3$.
(a) What is the probability that exactly 3 members of the committee are Hispanic?
(b) What is the probability that 3 or fewer members of the committee are Hispanic?

13.8 Do our athletes graduate? A university claims that 80% of its basketball players get degrees. An investigation examines the fate of all 20 players who entered
the program over a period of several years that ended six years ago. Of these players, 11 graduated and the remaining 9 are no longer in school. If the university’s claim is true, the number of players among the 20 who graduate should have the binomial distribution with \( n = 20 \) and \( p = 0.8 \). What is the probability that exactly 11 out of 20 players graduate?

13.9 Hispanic representation, continued.
(a) What is the mean number of Hispanics on randomly chosen committees of 15 workers in Exercise 13.7?
(b) What is the standard deviation \( \sigma \) of the count \( X \) of Hispanic members?
(c) Suppose that 10% of the factory workers were Hispanic. Then \( p = 0.1 \). What is \( \sigma \) in this case? What is \( \sigma \) if \( p = 0.01 \)? What does your work show about the behavior of the standard deviation of a binomial distribution as the probability of a success gets closer to 0?

13.10 Do our athletes graduate, continued.
(a) Find the mean number of graduates out of 20 players in the setting of Exercise 13.8 if the university’s claim is true.
(b) Find the standard deviation \( \sigma \) of the count \( X \).
(c) Suppose that the 20 players came from a population of which \( p = 0.9 \) graduated. What is the standard deviation \( \sigma \) of the count of graduates? If \( p = 0.99 \), what is \( \sigma \)? What does your work show about the behavior of the standard deviation of a binomial distribution as the probability \( p \) of success gets closer to 1?

13.11 Unmarried women. Among employed women, 25% have never been married. Select 10 employed women at random.
(a) The number in your sample who have never been married has a binomial distribution. What are \( n \) and \( p \)?
(b) What is the probability that exactly 2 of the 10 women in your sample have never been married?
(c) What is the probability that 2 or fewer have never been married?
(d) What is the mean number of women in such samples who have never been married? What is the standard deviation?

13.12 Lie detectors. A federal report finds that lie detector tests given to truthful persons have probability about 0.2 of suggesting that the person is deceptive.\(^{37}\)
(a) A company asks 12 job applicants about thefts from previous employers, using a lie detector to assess their truthfulness. Suppose that all 12 answer truthfully. What is the probability that the lie detector says all 12 are truthful? What is the probability that the lie detector says at least 1 is deceptive?
(b) What is the mean number among 12 truthful persons who will be classified as deceptive? What is the standard deviation of this number?
(c) What is the probability that the number classified as deceptive is less than the mean?

13.13 A market research survey. You operate a restaurant. You read that a sample survey by the National Restaurant Association shows that 40% of adults are committed to eating nutritious food when eating away from home. To help plan
your menu, you decide to conduct a sample survey in your own area. You will use random digit dialing to contact an SRS of 200 households by telephone.
(a) If the national result holds in your area, what distribution describes the count $X$ of respondents who seek nutritious food when eating out?
(b) What is the probability that $X$ lies between 75 and 85? (Use the Normal approximation.)

13.14 Planning a survey. You are planning a sample survey of small businesses in your area. You will choose an SRS of businesses listed in the telephone book’s Yellow Pages. Experience shows that only about half the businesses you contact will respond.
(a) If you contact 150 businesses, it is reasonable to use the binomial distribution with $n = 150$ and $p = 0.5$ for the number $X$ who respond. Explain why.
(b) What is the mean number who respond to surveys like yours?
(c) What is the probability that 70 or fewer will respond? (Use the Normal approximation.)
(d) How large a sample must you take to increase the mean number of respondents to 100?

13.15 A market research survey, continued. Return to the restaurant sample described in Exercise 13.13. You find 100 of your 200 respondents concerned about nutrition. Is this reason to believe that the percent in your area is higher than the national 40%? To answer this question, find the probability that $X$ is 100 or larger if $p = 0.4$ is true. If this probability is very small, that is reason to think that $p$ is actually greater than 0.4.

13.16 Bomber crews. During World War II, the British estimated that each of their bombers had probability 0.05 of being lost due to enemy action on a mission over occupied Europe. A tour of duty for a member of the crew was 30 missions.
(a) What is the probability that a bomber survived 30 missions?
(b) What is the mean number of missions that a bomber survived?
(c) What is the probability that a bomber survived 5 or fewer missions?

13.17 School vouchers. An opinion poll asks an SRS of 500 adults whether they favor giving parents of school-age children vouchers that can be exchanged for education at any public or private school of their choice. Each school would be paid by the government on the basis of how many vouchers it collected. Suppose that in fact 45% of the population favor this idea. What is the probability that more than half of the sample are in favor?

13.18 Survival of trees. A study of trees with diameter 12 inches in the Ozark Highlands area of Missouri found that the probability that such a tree survives the next five years is 0.94. A lot contains 345 trees in this class.
(a) What is the mean number that survive for five years?
(b) What is the probability that between 325 and 335 of these trees survive for five years?
Chapter 14

14.1 Who should get welfare? A news article on a Gallup Poll noted that “28 percent of the 1548 adults questioned felt that those who were able to work should be taken off welfare.” The article also said, “The margin of error for a sample size of 1548 is plus or minus three percentage points.” Opinion polls usually announce margins of error for 95% confidence. Using this fact, explain to someone who knows no statistics what “margin of error plus or minus three percentage points” means.

14.2 Would women govern better? A Gallup Poll asked, “Do you think this country would be governed better or governed worse if more women were in political office?” Of the 1026 adults in the sample, 57% said “better.” Gallup added, “For results based on the total sample of National Adults, one can say with 95% confidence that the margin of sampling error is 3 percentage points.” Explain to someone who knows no statistics what the phrase “95% confidence” means here.

14.3 Polling women. A New York Times poll on women’s issues interviewed 1025 women 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 its conclusions. What is the 95% confidence interval for the percent of all adult women who think they do not get enough time for themselves?
(b) 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.
(c) Then explain clearly what “95% confidence” means.

14.4 Surveying hotel managers. 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.)

14.5 Blood tests. 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 $x = 3.2$. Give a 90% confidence interval for her mean potassium level.
(b) If three measurements were taken on different days and the mean result is $\bar{x} = 3.2$, what is a 90% confidence interval for Julie’s mean blood potassium level?

14.6 Iron deficiency in infants. Researchers studying iron deficiency in infants examined infants who were following different feeding patterns. One group of 26 infants was being breast-fed. At 6 months of age, these children had mean hemoglobin level $\bar{x} = 12.9$ grams per 100 milliliters of blood. Assume that the population standard deviation is $\sigma = 1.6$. Give a 95% confidence interval for the mean hemoglobin level.
level of breast-fed infants. What assumptions (other than the unrealistic assumption that we know $\sigma$) does the method you used to get the confidence interval require?

14.7 Confidence level and margin of error. The National Assessment of Educational Progress (NAEP) gave its test of quantitative skills to a sample of 1077 women of ages 21 to 25 years. Their mean quantitative score was 275. Individual NAEP scores have a Normal distribution with standard deviation $\sigma = 60$.

(a) Give a 95% confidence interval for the mean score $\mu$ in the population of all young women.
(b) Give the 90% and 99% confidence intervals for $\mu$.
(c) What are the margins of error for 90%, 95%, and 99% confidence? How does increasing the confidence level affect the margin of error of a confidence interval?

14.8 Sample size and margin of error. The NAEP sample of 1077 young women had mean quantitative score $\bar{x} = 275$. Individual NAEP scores have a Normal distribution with standard deviation $\sigma = 60$.

(a) Give a 95% confidence interval for the mean score $\mu$ in the population of all young women.
(b) Suppose that the same result, $\bar{x} = 275$, had come from a sample of 250 women. Give the 95% confidence interval for the population mean $\mu$ in this case.
(c) Then suppose that a sample of 4000 women had produced the sample mean $\bar{x} = 275$, and again give the 95% confidence interval for $\mu$.
(d) What are the margins of error for samples of size 250, 1077, and 4000? How does increasing the sample size affect the margin of error of a confidence interval?

14.9 Cellulose in hay. An agronomist examines the cellulose content of alfalfa hay. Suppose that the cellulose content in the population has standard deviation $\sigma = 8$ mg/g. A sample of 15 cuttings has mean cellulose content $\bar{x} = 145$ mg/g.

(a) Give a 90% confidence interval for the mean cellulose content in the population.
(b) A previous study claimed that the mean cellulose content was $\mu = 140$ 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?

14.10 How large a sample of the hotel managers in Exercise 14.4 would be needed to estimate the mean $\mu$ within $\pm 1$ year with 99% confidence?

15.1 Gas mileage. Larry’s car averages 26 miles per gallon on the highway. He 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 average gas mileage has increased. What are the null and alternative hypotheses? Explain briefly what the parameter $\mu$ in your hypotheses represents.
15.2 **Diameter of a part.** 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 work properly. The manufacturer measures the diameter in a sample of motors to determine whether the mean diameter has moved away from the target. What are the null and alternative hypotheses? Explain briefly the distinction between the mean $\mu$ in your hypotheses and the mean $\overline{x}$ of the spindles the manufacturer measures.

15.3 **Mice in a maze.** Experiments on learning in animals sometimes measure how long it takes mice to find their 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. What are the null hypothesis $H_0$ and alternative hypothesis $H_a$?

15.4 **Service calls.** Last year, your company’s service technicians took an average of 2.6 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? State $H_0$ and $H_a$.

15.5 **Shoppers’ incomes.** 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. The researchers suspect the mean household income of mall shoppers is higher than that of the general population. State $H_0$ and $H_a$.

15.6 **Grading a teaching assistant.** The examinations in a large accounting class are scaled after grading so that the mean score is 50. The professor thinks that one teaching assistant is a poor teacher and suspects that his students have a lower mean score than the class as a whole. The TA’s students this semester can be considered a sample from the population of all students in the course, so the professor compares their mean score with 50. State the hypotheses $H_0$ and $H_a$.

15.7 **Spending on housing.** The Census Bureau reports that households spend an average of 31% of their total spending on housing. A home builders association in Cleveland interviews an SRS of 40 households in the Cleveland metropolitan area to learn what percent of their spending goes toward housing. Suppose we know that spending on housing in Cleveland follows a Normal distribution with standard deviation $\sigma = 9.6\%$.

(a) What is the sampling distribution of the mean percent $\overline{x}$ that the sample spends on housing if the null hypothesis is true? Sketch the density curve of the sampling distribution. (Hint: Sketch a Normal curve first, then mark the axis using what you know about locating $\mu$ and $\sigma$ on a Normal curve.)

(b) Suppose that the study finds $\overline{x} = 30.2\%$ for the 40 households in the sample. Mark this point on the axis in your sketch. Then suppose that the study result is $\overline{x} = 27.6\%$. Mark this point on your sketch. Referring to your sketch, explain in simple language why one result is good evidence that average Cleveland spending on housing is less than 31% and the other result is not.
15.8 What’s the *P*-value? A test of the null hypothesis $H_0: \mu = 0$ gives test statistic $z = 1.8$.
(a) What is the $P$-value if the alternative is $H_a: \mu > 0$?
(b) What is the $P$-value if the alternative is $H_a: \mu < 0$?
(c) What is the $P$-value if the alternative is $H_a: \mu \neq 0$?

15.9 *P* and significance. The $P$-value for a significance test is 0.033.
(a) Do you reject the null hypothesis at level $\alpha = 0.05$? Explain your answer.
(b) Do you reject the null hypothesis at level $\alpha = 0.01$? Explain your answer.

15.10 *P* and significance. The $P$-value for a significance test is 0.078.
(a) Do you reject the null hypothesis at level $\alpha = 0.05$? Explain your answer.
(b) Do you reject the null hypothesis at level $\alpha = 0.01$? Explain your answer.

15.11 CEO pay. 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 $x = 6.9\%$ and the standard deviation of the increases was $s = 55\%$. 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 $\sigma = 55\%$.
(a) Sketch the Normal curve for the sampling distribution of $\bar{x}$ when $H_0$ is true. Shade the area that represents the $P$-value for the observed outcome $\bar{x} = 6.9\%$.
(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?

15.12 Engine crankshafts. Here are measurements (in millimeters) of a critical dimension on a sample of automobile engine crankshafts:

| 224.120 | 224.001 | 224.017 | 223.982 | 223.989 | 223.961 |
| 223.960 | 224.089 | 223.987 | 223.976 | 223.902 | 223.980 |
| 224.098 | 224.057 | 223.913 | 223.999 |

The manufacturing process is known to vary Normally with standard deviation $\sigma = 0.060$ mm. The process mean is supposed to be 224 mm. Do these data give evidence that the process mean is not equal to the target value 224 mm? State hypotheses and calculate a test statistic and its $P$-value. Are you convinced that the process mean is not 224 mm?

15.13 Student study times. A student group claims that first-year students at a university must study 2.5 hours per night during the school week. A skeptic suspects that they study less than that on the average. A class survey finds that the average study time claimed by 269 students is $\bar{x} = 137$ minutes. Regard these students as a random sample of all first-year students and suppose we know that study times
Additional Exercises

follow a Normal distribution with standard deviation 65 minutes. Carry out a test of \( H_0: \mu = 150 \) against \( H_a: \mu < 150 \). What do you conclude?

15.14 Filling cola bottles. Bottles of a popular cola are supposed to contain 300 milliliters (ml) of cola. There is some variation from bottle to bottle because the filling machinery is not perfectly precise. The distribution of the contents is Normal with standard deviation \( \sigma = 3 \) ml. An inspector who suspects that the bottler is underfilling measures the contents of six bottles. The results are

\[
299.4 \quad 297.7 \quad 301.0 \quad 298.9 \quad 300.2 \quad 297.0
\]

Is this convincing evidence that the mean content of cola bottles is less than the advertised 300 ml?
(a) State the hypotheses that you will test.
(b) Calculate the test statistic.
(c) Find the \( P \)-value and state your conclusion.

15.15 Nicotine in cigarettes. 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 \( z = 2.42 \).
(a) Is the result significant at the 5% level?
(b) Is the result significant at the 1% level?
(c) By comparing \( z \) with the critical values in the bottom row of Table C, give two numbers that catch the \( P \)-value between them.

15.16 Sales of coffee. Weekly sales of regular ground coffee at a supermarket have in the recent past varied according to a Normal distribution with mean \( \mu = 354 \) units per week and standard deviation \( \sigma = 33 \) units. The store reduces the price by 5%. Sales in the next three weeks are 405, 378, and 411 units. Is this good evidence that average sales are now higher? The hypotheses are

\[
H_0: \mu = 354 \\
H_a: \mu > 354
\]

Assume that the standard deviation of the population of weekly sales remains \( \sigma = 33 \).
(a) Find the sample mean \( \bar{x} \) and the value of the one-sample \( z \) test statistic.
(b) Calculate the \( P \)-value. Sketch a standard Normal curve with the area corresponding to the \( P \)-value shaded.
(c) Is the result statistically significant at the \( \alpha = 0.05 \) level? Is it significant at the \( \alpha = 0.01 \) level? Do you think there is convincing evidence that mean sales are higher?

15.17 Is this milk watered down? Cobra Cheese Company buys milk from several suppliers. Cobra suspects that some producers are adding water to their milk
to increase their profits. Excess water can be detected by measuring the freezing point of the milk. The freezing temperature of natural milk varies Normally, with mean \( \mu = -0.545^\circ \text{Celsius (C)} \) and standard deviation \( \sigma = 0.008^\circ \text{C} \). Added water raises the freezing temperature toward 0\(^\circ\) C, the freezing point of water. Cobra’s laboratory manager measures the freezing temperature of five consecutive lots of milk from one producer. The mean measurement is \( \bar{x} = -0.538^\circ \text{C} \). Is this good evidence that the producer is adding water to the milk? State hypotheses, carry out the test, give the \( P \)-value, and state your conclusion.

15.18 Is it significant? There are other \( z \) statistics that we have not yet studied. You can use Table C to assess the significance of any \( z \) statistic. A study compares American-Japanese joint ventures in which the U.S. company is larger than its Japanese partner with joint ventures in which the U.S. company is smaller. One variable measured is the excess returns earned by shareholders in the American company. 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 this result significant at the 5% level?
(b) Is the result significant at the 10% level?

15.19 Benefits of patent protection? 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?

15.20 Students’ earnings. 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.” 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.

15.21 Explaining significance. A social psychologist reports: “In our sample, 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.

15.22 Diet and diabetes. Does eating more fiber reduce the blood cholesterol level of patients with diabetes? A randomized clinical trial compared normal and high-fiber diets. Here is part of the researchers’ conclusion:

The high-fiber diet reduced plasma total cholesterol concentrations by 6.7 percent ($P = 0.02$), triglyceride concentrations by 10.2 percent ($P = 0.02$), and very-low-density lipoprotein cholesterol concentrations by 12.5 percent ($P = 0.01$).

A doctor who knows no statistics says that a drop of 6.7% in cholesterol isn’t a lot—maybe it’s just an accident due to the chance assignment of patients to the two diets. Explain in simple language how “$P = 0.02$” answers this objection.

15.23 Does she look 25? The cigarette industry has adopted a voluntary code requiring that models appearing in its advertising must appear to be at least 25 years old. Studies have shown, however, that consumers think many of the models are younger. Here is a quote from a study that asked whether different brands of cigarettes use models that appear to be of different ages.

The ANCOVA revealed that the brand variable is highly significant ($P < 0.001$), indicating that the average perceived age of the models is not equal across the 12 brands. As discussed previously, certain brands such as Lucky Strike Lights, Kool Milds, and Virginia Slims tended to have younger models . . .

ANCOVA is an advanced statistical technique, but significance and $P$-values have their usual meaning. Explain to someone who knows no statistics what “highly significant ($P < 0.001$)” means and why this is good evidence of differences among all advertisements of these brands even though the subjects saw only a sample of ads.

15.24 Workers’ earnings. The Bureau of Labor Statistics generally uses 90% confidence in its reports. One report gives a 90% confidence interval for the mean hourly earnings of American workers in 2000 as $15.49$ to $16.11$. This result was calculated from the National Compensation Survey, a multistage probability sample of businesses.

(a) Would a 95% confidence interval be wider or narrower?
(b) Would the null hypothesis that the 2000 mean hourly earnings of all workers was $16$ be rejected at the 10% significance level in favor of the two-sided alternative? What about the null hypothesis that the mean was $15$?

15.25 IQ test scores. Here are the IQ test scores of 31 seventh-grade girls in a Midwest school district:

<table>
<thead>
<tr>
<th>114</th>
<th>100</th>
<th>104</th>
<th>89</th>
<th>102</th>
<th>91</th>
<th>114</th>
<th>114</th>
<th>103</th>
<th>105</th>
</tr>
</thead>
<tbody>
<tr>
<td>108</td>
<td>130</td>
<td>120</td>
<td>132</td>
<td>111</td>
<td>128</td>
<td>118</td>
<td>119</td>
<td>86</td>
<td>72</td>
</tr>
<tr>
<td>111</td>
<td>103</td>
<td>74</td>
<td>112</td>
<td>107</td>
<td>103</td>
<td>98</td>
<td>96</td>
<td>112</td>
<td>112</td>
</tr>
</tbody>
</table>
Treat the 31 girls as an SRS of all seventh-grade girls in the school district. Suppose that the standard deviation of IQ scores in this population is known to be $\sigma = 15$.

(a) Give a 95% confidence interval for the mean IQ score $\mu$ in the population.

(b) Is there significant evidence at the 5% level that the mean IQ score in the population differs from 100? State hypotheses and use your confidence interval to answer the question without more calculations.

Chapter 16

16.1 Plagiarizing online. The Excite Poll can be found online at poll.excite.com. The question appears on the screen, and you click to choose a response. On October 14, 2002, the question was

*It is now possible for school students to log on to Internet sites and download homework. Everything from book reports to doctoral dissertations can be downloaded free or for a fee. Do you believe giving a student who is caught plagiarizing an “F” for their assignment is the right punishment?*

Of the 20,125 people who responded, 14,793 clicked “Yes.” That’s 73.5% of the sample. Based on this sample, a 95% confidence interval for the percent of the population who would say “Yes” is $73.5\% \pm 0.61\%$ (details in Chapter 20). Why is this confidence interval worthless?

16.2 Ages of presidents. Joe is writing a report on the backgrounds of American presidents. He looks up the ages of all 43 presidents when they entered office. Because Joe took a statistics course, he uses these 43 numbers to get a 95% confidence interval for the mean age of all men who have been president. This makes no sense. Why not?

16.3 Helping welfare mothers. A study compares two groups of mothers with young children who were on welfare two years ago. One group attended a voluntary training program that was offered free of charge at a local vocational school and was 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 $21\% \pm 4\%$ 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 in simple 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?
16.4 Prayer in the schools? A New York Times/CBS News poll asked the question “Do you favor an amendment to the Constitution that would permit organized prayer in public schools?” Sixty-six percent of the sample answered “Yes.” The article describing the poll says that it “is based on telephone interviews conducted from Sept. 13 to Sept. 18 with 1,664 adults around the United States, excluding Alaska and Hawaii . . . the telephone numbers were formed by random digits, thus permitting access to both listed and unlisted residential numbers.” The article gives the margin of error as 3 percentage points. Opinion polls customarily announce margins of error for 95% confidence, so we are 95% confident that the percent of all adults who favor prayer in the schools lies in the interval 66% ± 3%.

The news article goes on to say: “The theoretical errors do not take into account a margin of additional error resulting from the various practical difficulties in taking any survey of public opinion.” List some of the “practical difficulties” that may cause errors in addition to the ±3% margin of error. Pay particular attention to the news article’s description of the sampling method.

16.5 A talk show opinion poll. 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 $x = 8740 per year, and the standard deviation of the responses is $s = 1125$. 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 $8669$ to $8811$.
(a) Is the station’s calculation correct?
(b) Does their conclusion describe the population of all the city’s citizens? Explain your answer.

16.6 What’s a gift worth? Do people value gifts from others more highly than they value the money it would take to buy the gift? We would like to think so, because we hope that “the thought counts.” A survey of 209 adults asked them to list three recent gifts and then asked, “Aside from any sentimental value, if, without the giver ever knowing, you could receive an amount of money instead of the gift, what is the minimum amount of money that would make you equally happy?” It turned out that most people would need more money than the gift cost to be equally happy. The magic words “significant ($P < 0.01$)” appear in the report of this finding.42
(a) The sample consisted of students and staff in a graduate program and of “members of the general public at train stations and airports in Boston and Philadelphia.” The report says this sample is “not ideal.” What’s wrong with the sample?
(b) In simple language, what does it mean to say that the sample thought their gifts were worth “significantly more” than their actual cost?
(c) Now be more specific: what does “significant ($P < 0.01$)” mean?
Chapter 18

18.1 Measuring blood pressure. A medical study finds that \( \bar{x} = 114.9 \) and \( s = 9.3 \) for the seated systolic blood pressure of the 27 members of one treatment group. What is the standard error of the mean?

18.2 Shrimp embryos. Biologists studying the levels of several compounds in shrimp embryos reported their results 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. Readers are supposed to understand that \( \bar{x} = 0.84 \) based on \( n = 3 \) measurements and that the standard error of the mean (SEM) is 0.01. What was the sample standard deviation \( s \) for these measurements?

18.3 Significance. You are testing \( H_0: \mu = 0 \) against \( H_a: \mu \neq 0 \) based on an SRS of 20 observations from a Normal population. What values of the \( t \) statistic are statistically significant at the \( \alpha = 0.005 \) level?

18.4 What critical value? You have an SRS of 15 observations from a Normally distributed population. What critical value would you use to obtain a 98% confidence interval for the mean \( \mu \) of the population?

18.5 Measuring acculturation. 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 a symmetric distribution with \( \bar{x} = 1.67 \) and \( s = 0.25 \). Because low scores should indicate a Mexican cultural orientation, these results helped to establish the validity of the test.\(^{43}\)

(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?

18.6 The cost of Internet access. 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:\(^{44}\)

<table>
<thead>
<tr>
<th>20</th>
<th>40</th>
<th>22</th>
<th>22</th>
<th>21</th>
<th>21</th>
<th>20</th>
<th>10</th>
<th>20</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>13</td>
<td>18</td>
<td>50</td>
<td>20</td>
<td>18</td>
<td>15</td>
<td>8</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>22</td>
<td>10</td>
<td>20</td>
<td>22</td>
<td>22</td>
<td>21</td>
<td>15</td>
<td>23</td>
<td>30</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
<td>40</td>
<td>22</td>
<td>29</td>
<td>19</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>15</td>
<td>19</td>
<td>21</td>
<td>14</td>
<td>22</td>
<td>21</td>
<td>35</td>
<td>20</td>
<td>22</td>
</tr>
</tbody>
</table>

(a) Make a stemplot of the data. 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.
18.7 Red wine is good for the heart. Observational studies suggest that moderate use of alcohol reduces heart attacks, and that red wine may have special benefits. One reason may be that red wine contains polyphenols, substances that do good things to cholesterol in the blood and so may reduce the risk of heart attacks. In an experiment, healthy men were assigned at random to several groups. One group of 9 men drank half a bottle of red wine each day for two weeks. The level of polyphenols in their blood was measured before and after the two-week period. Here are the percent changes in level:

3.5  8.1  7.4  4.0  0.7  4.9  8.4  7.0  5.5

Make a stemplot of the data. It is difficult to assess Normality from just 9 observations. Give a 90% $t$ confidence interval for the mean percent change in blood polyphenols among all healthy men if all drank this amount of red wine.

18.8 The cost of Internet access, continued. The data in Exercise 18.6 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? Follow the four-step process in your work.

18.9 DDT poisoning. 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 made measurements on 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):

1.6  1.7  1.8  1.9

(a) Find the mean refractory period $\bar{x}$ 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.
(c) 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 give good evidence for this claim? State $H_0$ and $H_a$ and do a $t$ test. Between what levels from Table C does the $P$-value lie? What do you conclude from the test?

18.10 Growing tomatoes. 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 $\bar{x} = 0.34$ and $s = 0.83$. Is there convincing evidence that Variety A has the higher mean yield?
(a) Describe in words what the parameter $\mu$ is in this setting.
(b) Carry out the test, following the four-step process.

18.11 Mutual-fund performance. Do “index funds” that simply buy and hold all the stocks in one of the stock market indexes, such as the Standard & Poor’s 500
Index, perform better than actively managed mutual funds? Compare the percent total return (price change plus dividends) of a large actively managed fund with that of the Vanguard Index 500 fund for the 24 years from 1977 to 2000. Vanguard did better by an average of 2.83% per year, and the standard deviation of the 24 annual differences was 11.65%. Is there convincing evidence that the index fund does better?

(a) Describe in words the parameter \( \mu \) for this comparison.

(b) State the hypotheses \( H_0 \) and \( H_a \).

(c) Find the matched pairs \( t \) statistic and its \( P \)-value. What do you conclude?

**18.12 Weight-loss programs.** In a study of the effectiveness of a weight-loss program, 47 subjects who were at least 20% overweight took part in the 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.\(^{46}\)

(a) Why was the matched pairs \( t \) test 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. Use Table C to say more about the \( P \)-value of the \( t \) test.

**18.13 ARSMA versus BI.** The ARSMA test (Exercise 18.5) 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 are the same. The differences in scores (ARSMA – BI) for the 22 subjects had \( \bar{x} = 0.2519 \) and \( s = 0.2767 \).

(a) Describe briefly how to arrange the administration of the two tests to the subjects, 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.

**18.14 Market research.** A manufacturer of small appliances employs a market research firm to estimate retail sales of its products by gathering information from a sample of retail stores. This month an SRS of 75 stores in the Midwest sales region finds that these stores sold an average of 24 of the manufacturer’s hand mixers, with standard deviation 11.

(a) Give a 95% confidence interval for the mean number of mixers sold by all stores in the region.

(b) The distribution of sales is strongly right-skewed because there are many smaller
stores and a few very large stores. The use of \( t \) in (a) is reasonably safe despite this violation of the Normality condition. Why?

18.15 Will they charge more? A bank wonders whether omitting the annual credit card fee for customers who charge at least $2400 in a year will increase the amount charged on its credit cards. The bank makes this offer to an SRS of 200 of its credit card customers. It then compares how much these customers charge this year with the amount that they charged last year. The mean increase in the sample is $332, 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 99% 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.

18.16 CEO pay. A study of the pay of corporate CEOs (chief executive officers) examined the cash compensation, adjusted for inflation, of the CEOs of 104 corporations over the period 1977 to 1988. Among the data are the average annual pay increases for each of the 104 CEOs. The mean percent increase in pay was 6.9%. The data showed great variation, with a standard deviation of 17.4%. The distribution was strongly skewed to the right.
(a) Despite the skewness of the distribution, there were no extreme outliers. Explain why we can use \( t \) procedures for these data.
(b) What are the degrees of freedom? When the exact degrees of freedom do not appear in Table C, use the next lower degrees of freedom in the table.
(c) Give a 99% confidence interval for the mean increase in pay for all corporate CEOs. What essential condition must the data satisfy if we are to trust your result?

18.17 Calibrating an instrument. Gas chromatography is a sensitive technique used 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 (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 nanogram of this substance, and give the margin of error for your choice of confidence level. Follow the four-step process. Then explain to a chemist who knows no statistics what your margin of error means.
18.18 The density of the earth. Exercise 2.8 gives 29 measurements of the density of the earth, made in 1798 by Henry Cavendish. Cavendish reported the mean of these measurements, but confidence intervals had not yet been invented. Display the data graphically to check for skewness and outliers. Then give an estimate for the density of the earth from Cavendish’s data and a margin of error for your estimate. Follow the four-step process.

18.19 Way down under. A remarkable discovery: large lakes exist deep under the Antarctic ice cap, kept liquid by the enormous pressure of the ice above. The largest is Lake Vostok. Do these lakes contain populations of ancient bacteria adapted to the dark, cold, high-pressure environment? Drilling over 3600 meters (over 11,000 feet) down to a depth where the ice consists of water frozen from Lake Vostok did indeed find bacteria. The researchers estimated “mean bacterial biomass” in nanograms of carbon per liter of melted water. They had 5 specimen ice cores. They split each specimen into “top” and “bottom” to get two samples of size 5, which they processed separately. Separate study of the two samples is a check on the complicated process of preparing the cores for analysis. The results “are mean estimates ±SD.” Here is an extract:  

<table>
<thead>
<tr>
<th></th>
<th>Top melt</th>
<th>Bottom melt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean bacterial biomass (ng of C liter(^{-1}))</td>
<td>2.9 ± 0.4</td>
<td>2.8 ± 0.4</td>
</tr>
</tbody>
</table>

Use the results for the 5 top-melt specimens to give a 90% confidence interval for the mean bacterial biomass in the ice above Lake Vostok.

18.20 The power of a t test (Optional). Exercise 18.13 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 \(\alpha = 0.05\)) of

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

against the alternative \(\mu = 0.2\). We do this by acting as if \(\sigma\) were known.

(a) From Table C, what is the critical value for \(\alpha = 0.05\)?

(b) Write the rule for rejecting \(H_0\) at the \(\alpha = 0.05\) level. Then take \(s = 0.3\), the approximate value observed in Exercise 18.13, and restate the rejection criterion in terms of \(\bar{x}\). Note that this is a two-sided test.

(c) Find the probability of this event when \(\mu = 0.2\) (the alternative given) and \(\sigma = 0.3\) (estimated from the data in Exercise 18.13) by a Normal probability calculation. This is the approximate power.
Chapter 19

19.1 Which data design? Is each of these designs (1) single sample, (2) matched pairs, or (3) two independent samples? Explain your choices.

(a) An education researcher wants to learn whether it is more effective to put questions before or after introducing a new concept in an elementary school mathematics text. He prepares two text segments that teach the concept, one with motivating questions before and the other with review questions after. He uses each text segment to teach a separate group of children. The researcher compares the scores of the groups on a test over the material.

(b) Another researcher approaches the same issue 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. The subjects are a single group of children. Each child studies both topics, one (chosen at random) with questions before and the other with questions after. The researcher compares test scores for each child on the two topics to see which topic he or she learned better.

19.2 Beetles in oats. 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. The data appear roughly Normal. Here are the summary statistics:

<table>
<thead>
<tr>
<th>Group</th>
<th>Treatment</th>
<th>n</th>
<th>x̄</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Control</td>
<td>13</td>
<td>3.47</td>
<td>1.21</td>
</tr>
<tr>
<td>2</td>
<td>Malathion</td>
<td>14</td>
<td>1.36</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Is there significant evidence at the 1% level that malathion reduces the mean number of larvae per stem? Follow the four-step process in your work.

19.3 Road rage. “The phenomenon of road rage has been frequently discussed but infrequently examined.” So begins a report based on interviews with randomly selected drivers. The respondents’ answers to interview questions produced scores on an “angry/threatening driving scale” with values between 0 and 19. Here are summaries of the scores:

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>x̄</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>596</td>
<td>1.78</td>
<td>2.79</td>
</tr>
<tr>
<td>Female</td>
<td>769</td>
<td>0.97</td>
<td>1.84</td>
</tr>
</tbody>
</table>

(a) We suspect that men are more susceptible to road rage than women. Carry out a test of that hypothesis. (State hypotheses, find the test statistic and P-value, and state your conclusion.)

(b) The subjects were selected using random-digit dialing. The large sample sizes make the Normality condition unnecessary. There is one aspect of the data production that might reduce the validity of the data. What is it?

19.4 Treating scrapie in hamsters. Scrapie is a degenerative disease of the nervous system. A study of the substance IDX as a treatment for scrapie used as subjects 20 infected hamsters. Ten, chosen at random, were injected with IDX. The
other 10 were untreated. The researchers recorded how long each hamster lived. They reported, “Thus, although all infected control hamsters had died by 94 days after infection (mean ± SEM = 88.5 ± 1.9 days), IDX-treated hamsters lived up to 128 days (mean ± SEM = 116 ± 5.6 days).” Readers are supposed to know that SEM stands for “standard error of the mean.”

(a) Fill in the values in this summary table:

<table>
<thead>
<tr>
<th>Group</th>
<th>Treatment</th>
<th>n</th>
<th>( \bar{x} )</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IDX</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>Untreated</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

(b) What degrees of freedom would you use in the conservative two-sample \( t \) procedures to compare the two treatments?

19.5 Social insight among men and women. The Chapin Social Insight Test is a psychological test designed to measure how accurately a person 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? Follow the four-step process.

19.6 Treating scrapie. Exercise 19.4 reports the results of a study to determine whether IDX is an effective treatment for scrapie.

(a) Is there good evidence that hamsters treated with IDX live longer on the average?
(b) Give a 95% confidence interval for the mean amount by which IDX prolongs life.

19.7 More nutritious corn. Ordinary corn doesn’t have as much of the amino acid lysine as animals need in their feed. Plant scientists have developed varieties of corn that have increased amounts of lysine. In a test of the quality of high-lysine corn as animal feed, an experimental group of 20 one-day-old male chicks ate 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:

<table>
<thead>
<tr>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>380 321 366 356</td>
<td>361 447 401 375</td>
</tr>
<tr>
<td>283 349 402 462</td>
<td>434 403 393 426</td>
</tr>
<tr>
<td>356 410 329 399</td>
<td>406 318 467 407</td>
</tr>
<tr>
<td>350 384 316 272</td>
<td>427 420 477 392</td>
</tr>
<tr>
<td>345 455 360 431</td>
<td>430 339 410 326</td>
</tr>
</tbody>
</table>

(a) Present the data graphically. Are there outliers or strong skewness that might prevent the use of \( t \) procedures?
(b) Is there good evidence that chicks fed high-lysine corn gain weight faster? Carry
out a test and report your conclusions.
(c) Give a 95% confidence interval for the mean extra weight gain in chicks fed high-lysine corn.

19.8 Is red wine better than white wine? Observational studies suggest that moderate use of alcohol reduces heart attacks, and that red wine may have special benefits. One reason may be that red wine contains polyphenols, substances that do good things to cholesterol in the blood and so may reduce the risk of heart attacks. In an experiment, healthy men were assigned at random to drink half a bottle of either red or white wine each day for two weeks. The level of polyphenols in their blood was measured before and after the two-week period. Here are the percent changes in level for the subjects in both groups:

<table>
<thead>
<tr>
<th></th>
<th>Red wine</th>
<th>White wine</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.5 8.1 7.4 4.0 0.7 4.9 8.4 7.0 5.5</td>
<td>3.1 0.5 −3.8 4.1 −0.6 2.7 1.9 −5.9 0.1</td>
</tr>
</tbody>
</table>

(a) Is there good evidence that red wine drinkers gain more polyphenols on the average than white wine drinkers? Do a complete analysis, following the four-step process.
(b) Does this study give reason to think that it is drinking red wine, rather than some lurking variable, that causes the increase in blood polyphenols?

19.9 Fitness and personality. 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. Here are the data for the “ego strength” personality factor:

<table>
<thead>
<tr>
<th>Group</th>
<th>Fitness</th>
<th>n</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low</td>
<td>14</td>
<td>4.64</td>
<td>0.69</td>
</tr>
<tr>
<td>2</td>
<td>High</td>
<td>14</td>
<td>6.43</td>
<td>0.43</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.

19.10 Red wine versus white wine, continued. Using the data in Exercise 19.8, give a 95% confidence interval for the difference (red minus white) in mean change in blood polyphenol levels.

19.11 Market research. 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. 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 between this year and last year in the mean number of units sold at all retail
stores.
(b) Explain in language that the manager can understand why he cannot be confident that sales rose by 6%, and that in fact sales may even have dropped.

19.12 Will they charge more? 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 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>$x$</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. 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. (c) Is the bank’s study an experiment? Why? How does this affect the conclusions the bank can draw from the study?

19.13 Depressed teens. To study depression among adolescents, investigators administered the Children’s Depression Inventory (CDI) to teenagers in rural Newfoundland, Canada. As is often the case in social science studies, there is some question about whether the subjects can be considered a random sample from an interesting population. We will ignore this issue. One finding was that “older adolescents scored significantly higher on the CDI.” Higher scores indicate symptoms of depression. Here are summary data for two grades:

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>$\bar{x}$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 9</td>
<td>84</td>
<td>6.94</td>
<td>6.03</td>
</tr>
<tr>
<td>Grade 11</td>
<td>70</td>
<td>8.98</td>
<td>7.08</td>
</tr>
</tbody>
</table>

Do an analysis to verify the quoted conclusion. Follow the four-step process in your work.

19.14 Depressed teens, continued. The study in the previous exercise also concluded that there were no sex differences in depression. Following the four-step process, use these summary data to verify this finding:

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>$\bar{x}$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>112</td>
<td>7.38</td>
<td>6.95</td>
</tr>
<tr>
<td>Females</td>
<td>104</td>
<td>7.15</td>
<td>6.31</td>
</tr>
</tbody>
</table>
19.15 Studying speech. 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.58

<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) 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.

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

19.16 New welfare programs. A major study of alternative welfare programs randomly assigned women on welfare to one of two programs, called “WIN” and “Options.” WIN was the existing program. The new Options program gave more incentives to work. An important question was how much more (on the average) women in Options earned than those in WIN. Here is Minitab output for earnings in dollars over a three-year period.59

\[
\text{TWOSAMPLE T FOR 'OPT' VS 'WIN'}
\]

<table>
<thead>
<tr>
<th>N</th>
<th>MEAN</th>
<th>STDEV</th>
<th>SE MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPT</td>
<td>1362</td>
<td>7638</td>
<td>289</td>
</tr>
<tr>
<td>WIN</td>
<td>1395</td>
<td>6595</td>
<td>247</td>
</tr>
</tbody>
</table>

95 PCT CI FOR MU OPT - MU WIN: (1022.90, 1063.10)

(a) Give a 99% confidence interval for the amount by which the mean earnings of Options participants exceeded the mean earnings of WIN subjects. (Minitab will give a 99% confidence interval if you instruct it to do so. Here we have only the basic output, which includes the 95% confidence interval.)

(b) The distribution of incomes is strongly skewed to the right but includes no extreme outliers because all the subjects were on welfare. What fact about these data allows us to use \(t\) procedures despite the strong skewness?

19.17 Studying speech, continued. The researchers in the study discussed in Exercise 19.15 looked at VOTs for adults and children pronouncing many different words. Explain why they should not do a separate two-sample \(t\) test for each word and conclude that those words with a significant difference (say \(P < 0.05\)) distinguish children from adults. (The researchers did not make this mistake.)

19.18 Testing pharmaceuticals. A pharmaceutical manufacturer does a chemical analysis to check the potency of products. The standard release potency for cephalothin crystals is 910. An assay of 16 lots gives the following potency data:
(a) Check the data for outliers or strong skewness that might threaten the validity of the \( t \) procedures.
(b) Give a 95\% confidence interval for the mean potency.
(c) Is there significant evidence at the 5\% level that the mean potency is not equal to the standard release potency?

19.19 Way down under, continued. Do the results reported in Exercise 18.19 give any reason to think that the population mean biomasses as estimated from top-melt and bottom-melt samples are different? Use the four-step process to guide your work. (A difference would suggest that the preparation process influences the findings.)

19.20 Lead in soil. The amount of lead in a type of soil, measured by a standard method, averages 86 parts per million (ppm). A new 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 1\% level that the new method frees less lead from the soil?
(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.

19.21 Fertilizing pasture. An agricultural researcher thinks that applying potassium fertilizer to grasslands at several times during the growing season may give higher yields than applying only in the spring. He therefore compares two treatments: 100 pounds per acre of potassium in the spring (Treatment 1) and 50, 25, and 25 pounds per acre applied in the spring, early summer, and late summer (Treatment 2). He continues the experiment over several years because growing conditions vary from year to year.

The table below gives the yields, in pounds of dry matter per acre. It is known from long experience that yields vary roughly Normally.\(^6\)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Year 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3902</td>
<td>4281</td>
<td>5135</td>
<td>5350</td>
<td>5746</td>
</tr>
<tr>
<td>2</td>
<td>3970</td>
<td>4271</td>
<td>5440</td>
<td>5490</td>
<td>6028</td>
</tr>
</tbody>
</table>

(a) Do the data give good evidence that Treatment 2 leads to higher average yields? (State hypotheses, carry out a test, give a \( P \)-value as exact as the tables in the text allow, and state your conclusions in words.)
(b) Give a 98\% confidence interval for the mean increase in yield due to spreading potassium applications over the growing season.

19.22 Treating scrapie (Optional). The report in Exercise 19.4 suggests that hamsters in the longer-lived group also had more variation in their length of life. It is common to see \( s \) increase along with \( \bar{x} \) when we compare groups. Do the data give significant evidence of unequal standard deviations?

19.23 Social insight among men and women (Optional). The data in Exercise 19.5 show that scores of men and women on the Chapin Social Insight Test have
similar sample standard deviations. The samples are large, however, so we might nonetheless find evidence of a significant difference between the population standard deviations. Is this the case? (Use the $F(120, 100)$ critical values from Table D.)

19.24 Red and white wine (Optional). Is there a statistically significant difference between the standard deviations of blood polyphenol level change in the red and white wine groups in Exercise 19.8?

19.25 Studying speech (Optional). The data for VOTs of children and adults in Exercise 19.15 show quite different sample standard deviations. How statistically significant is the observed inequality?

19.26 The power of a two-sample $t$ test (Optional). A bank asks you to compare two ways to increase the use of their credit cards. Plan A would offer customers a cash-back rebate based on their total amount charged. Plan B would reduce the interest rate charged on card balances. The bank thinks that Plan B will be more effective. The response variable is the total amount a customer charges during the test period. You decide to offer each of Plan A and Plan B to a separate SRS of the bank’s credit card customers. In the past, the mean amount charged in a six-month period has been about $1100, with a standard deviation of $400. Will a two-sample $t$ test based on SRSs of 350 customers in each group detect a difference of $100 in the mean amounts charged under the two plans? We will compute the approximate power of the two-sample $t$ test of

$$H_0: \mu_B = \mu_A$$
$$H_a: \mu_B > \mu_A$$

against the specific alternative $\mu_B - \mu_A = 100$. We will use the past value $400$ as a rough estimate of both the population $\sigma$’s and future sample $s$’s.

(a) What is the approximate value of the $\alpha = 0.05$ critical value $t^*$ for the two-sample $t$ statistic when $n_1 = n_2 = 350$?

(b) **Step 1.** Write the rule for rejecting $H_0$ in terms of $\bar{x}_B - \bar{x}_A$. The test rejects $H_0$ when

$$\frac{\bar{x}_B - \bar{x}_A}{\sqrt{\frac{s_B^2}{n_B} + \frac{s_A^2}{n_A}}} \geq t^*$$

Take both $s_B$ and $s_A$ to be 400, and $n_B$ and $n_A$ to be 350. Find the number $c$ such that the test rejects $H_0$ when $\bar{x}_B - \bar{x}_A \geq c$.

(c) **Step 2.** The power is the probability of rejecting $H_0$ when the alternative is true. Suppose that $\mu_B - \mu_A = 100$ and that both $\sigma_B$ and $\sigma_A$ are 400. The power we seek is the probability that $\bar{x}_B - \bar{x}_A \geq c$ under these assumptions. Calculate the power.
Chapter 20

In each of Exercises 20.1 to 20.3:
(a) Describe the population and explain in words what the parameter $p$ is.
(b) Give the numerical value of the statistic $\hat{p}$ that estimates $p$.

20.1 Tonya wants to estimate what proportion of the students in her dormitory like the dorm food. She interviews an SRS of 50 of the 175 students living in the dormitory. She finds that 14 think the dorm food is good.

20.2 Glenn wonders what proportion of the students at his school think that tuition is too high. He interviews an SRS of 50 of the 2400 students at his college. Thirty-eight of those interviewed think tuition is too high.

20.3 A college president says, “99% of the alumni support my firing of Coach Boggs.” You contact an SRS of 200 of the college’s 15,000 living alumni and find that 76 of them support firing the coach.

20.4 Do you jog? The Gallup Poll once asked a random sample of 1540 adults, “Do you happen to jog?” Suppose that in fact 15% of all adults jog.
(a) Find the mean and standard deviation of the proportion $\hat{p}$ of the sample who jog. (Assume the sample is an SRS.)
(b) What sample size would be required to reduce the standard deviation of the sample proportion to one-half the value you found in (a)?
(c) Use the Normal approximation to find the probability that between 13% and 17% of the sample jog.

20.5 Harley motorcycles. Harley-Davidson motorcycles make up 14% of all the motorcycles registered in the United States. You plan to interview an SRS of 500 motorcycle owners.
(a) What is the approximate distribution of the proportion of your sample who own Harleys?
(b) Is your sample likely to contain 20% or more who own Harleys? Is it likely to contain at least 15% Harley owners? Do Normal probability calculations to answer these questions.

20.6 Do you jog, continued. Suppose that 15% of all adults jog. Exercise 20.4 asks the probability that the sample proportion $\hat{p}$ from an SRS estimates $p = 0.15$ within $\pm 2$ percentage points. Find this probability for SRSs of sizes 200, 800, and 3200. What general conclusion can you draw from your calculations?

20.7 Information online. A random sample of 1318 Internet users was asked where they will go for information the next time they need information about health or medicine; 606 said that they would use the Internet. Give a 99% large-sample confidence interval for the proportion of all Internet users who feel this way. Be sure to check that the conditions for this method are met.

20.8 Equality for women? Have efforts to promote equality for women gone far enough in the United States? A poll on this issue by the cable network MSNBC contacted 1019 adults. A newspaper article about the poll said, “Results have a
Additional Exercises

margin of sampling error of plus or minus 3 percentage points.”

(a) Overall, 54% of the sample (550 of 1019 people) answered “Yes.” Find the large-sample 95% confidence interval for the proportion in the adult population who would say “Yes” if asked. Is the report’s claim about the margin of error roughly right? (Assume that the sample is an SRS.)

(b) The news article said that 65% of men, but only 43% of women, think that efforts to promote equality have gone far enough. Explain why we do not have enough information to give confidence intervals for men and women separately.

(c) Would a 95% confidence interval for women alone have a margin of error less than 0.03, about equal to 0.03, or greater than 0.03? Why? You see that the news article’s statement about the margin of error for poll results is a bit misleading.

20.9 Unhappy HMO patients. How likely are patients who file complaints with a health maintenance organization (HMO) to leave the HMO? In one recent year, 639 of the more than 400,000 members of a large New England HMO filed complaints. Fifty-four of the complainers left the HMO voluntarily. (That is, they were not forced to leave by a move or a job change.) Consider this year’s complainers as an SRS of all patients who will complain in the future. Give a 90% confidence interval for the proportion of complainers who voluntarily leave the HMO. Can you use the large-sample method?

20.10 Fear of crime among older black women. The elderly fear crime more than younger people, even though they are less likely to be victims of crime. One of the few studies that looked at older blacks recruited a random sample of 56 black women over the age of 65 from Atlantic City, New Jersey. Of these women, 27 said that they “felt vulnerable” to crime.

(a) Give the two estimates \( \hat{p} \) and \( \tilde{p} \) of the proportion \( p \) of all elderly black women in Atlantic City who feel vulnerable to crime. There is little difference between them. This is generally true when \( \hat{p} \) is not close to either 0 or 1.

(b) Give both the large-sample 95% confidence interval and the plus four 95% confidence interval for \( p \). The plus four interval is a bit narrower. This is generally true when \( \hat{p} \) is not close to either 0 or 1.

20.11 Polling students. Exercise 20.2 describes an SRS of 50 college students, of whom 38 think tuition at their school is too high. Give the plus four 95% confidence interval for the proportion of all students who share this opinion. Explain why the large-sample confidence interval should not be used for these data.

20.12 The millennium begins with optimism. In January of the year 2000, a Gallup Poll asked a random sample of 1633 adults, “In general, are you satisfied or dissatisfied with the way things are going in the United States at this time?” It found that 1127 said that they were satisfied. Write a short report of this finding, as if you were writing for a newspaper. Be sure to include a margin of error. Use the large-sample confidence interval.

20.13 Alternative medicine. A nationwide random survey of 1500 adults asked about attitudes toward “alternative medicine” such as acupuncture, massage therapy, and herbal therapy. Among the respondents, 660 said they would use alternative
(a) Give a 95% confidence interval for the proportion of all adults who would use alternative medicine.

(b) Write a short paragraph for a news report based on the survey.

20.14 Going to church. A Gallup Poll asked a sample of 1785 adults, “Did you, yourself, happen to attend church or synagogue in the last 7 days?” Of the respondents, 750 said “Yes.” Treat Gallup’s sample as an SRS of all American adults. Following the four-step process, estimate with 99% confidence the proportion of all adults who claim that they attended church or synagogue during the week preceding the poll. (The proportion who actually attended is no doubt lower—some people say “Yes” if they usually attend, often attend, or sometimes attend.)

20.15 School vouchers. A national opinion poll found that 44% of all American adults agree that parents should be given vouchers good for education at any public or private school of their choice. The result was based on a small sample. How large an SRS is required to obtain a margin of error of 0.03 (that is, ±3%) in a 95% confidence interval?

(a) Answer this question using the previous poll’s result as the guessed value \( p^* \).

(b) Do the problem again using the conservative guess \( p^* = 0.5 \). By how much do the two sample sizes differ?

20.16 Going to church, continued. 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 claim to have attended church or synagogue (see Exercise 20.14)? (Use the conservative guess \( p^* = 0.5 \), and explain why this method is reasonable in this situation.)

20.17 Side effects. An experiment on the side effects of pain relievers assigned arthritis patients to one of several over-the-counter pain medications. Of the 440 patients who took one brand of pain reliever, 23 suffered some “adverse symptom.”

(a) If 10% of all patients suffer adverse symptoms, what would be the sampling distribution of the proportion with adverse symptoms in a sample of 440 patients?

(b) Does the experiment provide strong evidence that fewer than 10% of patients who take this medication have adverse symptoms?

20.18 Stolen Harleys. Harley-Davidson motorcycles make up 14% of all motorcycles registered in the United States. In 1995, 9224 motorcycles were reported stolen; 2490 of these were Harleys. We can think of motorcycles stolen in 1995 as an SRS of motorcycles stolen in recent years.

(a) If Harleys made up 14% of motorcycles stolen, what would be the sampling distribution of the proportion of Harleys in a sample of 9224 stolen motorcycles?

(b) Is the proportion of Harleys among stolen bikes significantly higher than their share of all motorcycles?

20.19 Attitudes toward nuclear power. A Gallup Poll on energy use asked 512 randomly selected adults if they favored “increasing the use of nuclear power as a major source of energy.” Gallup reported that 225 said “Yes.” Does this poll
give good evidence that fewer than half of all adults favor increased use of nuclear power? Use the four-step process to guide your work.

20.20 Teens and their TV sets. The *New York Times* and CBS News conducted a nationwide survey of 1048 randomly selected 13- to 17-year-olds. Of these teenagers, 692 had a television in their room.\(^6\)

(a) Check that we can use the large-sample confidence interval.

(b) Give a 95% confidence interval for the proportion of all teens who have a TV set in their room.

(c) The news article says, “In theory, in 19 cases out of 20, the survey results will differ by no more than three percentage points in either direction from what would have been obtained by seeking out all American teenagers.” Explain how your results agree with this statement.

20.21 Do chemists have more girls? Some people think that chemists are more likely than other parents to have female children. (Perhaps chemists are exposed to something in their laboratories that affects the sex of their children.) The Washington State Department of Health lists the parents’ occupations on birth certificates. Between 1980 and 1990, 555 children were born to fathers who were chemists. Of these births, 273 were girls. During this period, 48.8% of all births in Washington State were girls.\(^6\) Is there evidence that the proportion of girls born to chemists is higher than the state proportion? Follow the four-step process.

20.22 Teens and their TV sets, continued. A random sample of 1048 13- to 17-year-olds found that 692 had a television set in their room. Is this good evidence that more than half of all teens have a TV in their room? State hypotheses, give the test statistic, use Table A to find its \(P\)-value, and state your conclusion.

20.23 We want to be rich. In a recent year, 73% of first-year college students responding to a national survey identified “being very well-off financially” as an important personal goal. A state university finds that 132 of an SRS of 200 of its first-year students say that this goal is important.

(a) 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.

(b) Is there good evidence that the proportion of first-year students at this university who think being very well-off is important differs from the national value, 73%? (Be sure to state hypotheses, give the \(P\)-value, and state your conclusion.)

(c) Check that you can safely use the methods of this chapter in both (a) and (b).

20.24 Matched pairs. One-sample procedures for proportions, like those for means, are used to analyze data from matched pairs designs. Here is an example.

Each of 50 subjects tastes two unmarked cups of coffee and says which he or she prefers. One cup in each pair contains instant coffee; the other, fresh-brewed coffee. Thirty-one of the subjects prefer the fresh-brewed coffee. Take \(p\) to be the proportion of the population who would prefer fresh-brewed coffee in a blind tasting.

(a) Test the claim that a majority of people prefer the taste of fresh-brewed coffee. State hypotheses and report the \(z\) statistic and its \(P\)-value. Is your result significant
20.25 Do our athletes graduate? The National Collegiate Athletic Association (NCAA) requires colleges to report the graduation rates of their athletes. Here are data from a Big Ten university’s report.\(^68\)

(a) Ninety-five of the 147 athletes admitted in 1989–1991 graduated within six years. Does the proportion of athletes who graduate differ significantly from the all-university proportion, which was 70%?

(b) The graduation rates were 37 of 45 female athletes and 58 of 102 male athletes. Is there evidence that a smaller proportion of male athletes than of female athletes graduate within six years?

(c) We are willing to regard athletes admitted in this period as an SRS from the large population of athletes the university will admit under its present standards. Explain why you can safely use the \(z\) procedures in parts (a) and (b). Then explain why you cannot use these procedures to test whether male basketball players (4 out of 8 admitted graduated) differ from other athletes.
complaints. 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.

(a) The HMO has more than 400,000 members. Check that you can safely use the large-sample confidence interval.

(b) How much higher is the proportion of complainers who leave? Give a 90% confidence interval.

21.4 Fear of crime among older blacks. The elderly fear crime more than younger people, even though they are less likely to be victims of crime. One of the few studies that looked at older blacks recruited random samples of 56 black women and 63 black men over the age of 65 from Atlantic City, New Jersey. Of the women, 27 said they “felt vulnerable” to crime; 46 of the men said this.\textsuperscript{71}

(a) What proportion of women in the sample feel vulnerable? Of men? Men are victims of crime more often than women, so we expect a higher proportion of men to feel vulnerable.

(b) Give the large-sample 95% confidence interval for the difference (men minus women).

(c) Although the sample sizes do satisfy our rule of thumb for using the large-sample interval, they are not large. Give the plus four interval and compare it with the large-sample interval.

21.5 Are genetically modified foods risky? Europe and the United States differ considerably in their attitudes toward food made from crops that have been genetically modified (GM) to, for example, resist pests or contain more protein. A random sample of 12,178 European adults found that 63% thought such foods were risky. In the United States, a random sample of 863 adults who were asked the same questions found that 46% considered GM foods risky.\textsuperscript{72}

(a) What are the counts of people in each sample who thought GM foods were risky?

(b) Give a 95% confidence interval to compare Europe and the United States.

21.6 Aspirin and heart attacks. The Physicians’ Health Study examined the effects of taking an aspirin every other day. Earlier studies suggested that aspirin might reduce the risk of heart attacks. The subjects were 22,071 healthy male physicians at least 40 years old. The study assigned 11,037 of the subjects at random to take aspirin. The others took a placebo pill. The study was double-blind. Here are the counts for some of the outcomes of interest to the researchers:

<table>
<thead>
<tr>
<th></th>
<th>Aspirin group</th>
<th>Placebo group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal heart attacks</td>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td>Non-fatal heart attacks</td>
<td>129</td>
<td>213</td>
</tr>
<tr>
<td>Strokes</td>
<td>119</td>
<td>98</td>
</tr>
</tbody>
</table>

For which outcomes is the difference between the aspirin and placebo groups significant? (Use two-sided alternatives and follow the four-step process.)

21.7 Treating AIDS. The drug AZT was the first drug that seemed effective in delaying the onset of AIDS. Evidence for AZT’s effectiveness came from a large randomized comparative experiment. The subjects were 1300 volunteers who were
infected with HIV, the virus that causes AIDS, but did not yet have AIDS. The study assigned 435 of the subjects at random to take 500 milligrams of AZT each day, and another 435 to take a placebo. (The others were assigned to a third treatment, a higher dose of AZT. We will compare only two groups.) At the end of the study, 38 of the placebo subjects and 17 of the AZT subjects had developed AIDS. We want to test the claim that taking AZT lowers the proportion of infected people who will develop AIDS in a given period of time.

(a) State hypotheses, and check that you can safely use the z procedures.
(b) How significant is the evidence that AZT is effective?
(c) The experiment was double-blind. Explain what this means.

**Comment:** Medical experiments on treatments for AIDS and other fatal diseases raise hard ethical questions. Some people argue that because AIDS is always fatal, infected people should get any drug that has any hope of helping them. The counter-argument is that we will then never find out which drugs really work. The placebo patients in this study were given AZT as soon as the results were in.

### 21.8 Who gets stock options?
Different kinds of companies compensate their key employees in different ways. Established companies may pay higher salaries, while new companies may offer stock options that will be valuable if the company succeeds. Do high-tech companies tend to offer stock options more often than other companies? One study looked at a random sample of 200 companies. Of these, 91 were listed in the *Directory of Public High Technology Corporations* and 109 were not listed. Treat these two groups as SRSs of high-tech and non-high-tech companies. Seventy-three of the high-tech companies and 75 of the non-high-tech companies offered incentive stock options to key employees.73

(a) Is there evidence that a higher proportion of high-tech companies offer stock options?
(b) Give the plus four 95% confidence interval for the difference in the proportions of the two types of companies that offer stock options.

### 21.9 Nobody is home in July.
Nonresponse to sample surveys may differ with the season of the year. In Italy, for example, many people leave town during the summer. The Italian National Statistical Institute called random samples of telephone numbers between 7 p.m. and 10 p.m. at several seasons of the year. Here are the results for two seasons:74

<table>
<thead>
<tr>
<th>Dates</th>
<th>Number of calls</th>
<th>No answer</th>
<th>Total nonresponse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan. 1 to Apr. 13</td>
<td>1558</td>
<td>333</td>
<td>491</td>
</tr>
<tr>
<td>July 1 to Aug. 31</td>
<td>2075</td>
<td>861</td>
<td>1174</td>
</tr>
</tbody>
</table>

(a) How much higher is the proportion of “no answers” in July and August compared with the early part of the year? Give a 99% confidence interval.
(b) The difference between the proportions of “no answers” is so large that it is clearly statistically significant. How can you tell from your work in (a) that the difference is significant at the $\alpha = 0.01$ level?
(c) Use the information given to find the counts of calls that had nonresponse for some reason other than “no answer.” Do the rates of nonresponse due to other causes also differ significantly for the two seasons?
21.10 Preventing strokes. Aspirin prevents blood from clotting and so helps prevent strokes. The Second European Stroke Prevention Study asked whether adding another anticlotting drug named dipyridamole would be more effective for patients who had already had a stroke. Here are the data on strokes and deaths during the two years of the study.\textsuperscript{75}

<table>
<thead>
<tr>
<th></th>
<th>Number of patients</th>
<th>Number of strokes</th>
<th>Number of deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin alone</td>
<td>1649</td>
<td>206</td>
<td>182</td>
</tr>
<tr>
<td>Aspirin + dipyridamole</td>
<td>1650</td>
<td>157</td>
<td>185</td>
</tr>
</tbody>
</table>

(a) The study was a randomized comparative experiment. Outline the design of the study.
(b) Is there a significant difference in the proportion of strokes in the two groups? State hypotheses, find the $P$-value, and state your conclusion.
(c) Is there a significant difference in death rates for the two groups?

21.11 Access to computers. A sample survey by Nielsen Media Research looked at computer access and use of the Internet. Whites were significantly more likely than blacks to own a home computer, but the black-white difference in computer access at work was not significant. The study team then looked separately at the households with at least $40,000$ income. The sample contained 1916 white and 131 black households in this class. Here are the sample counts for these households.\textsuperscript{76}

<table>
<thead>
<tr>
<th></th>
<th>Blacks</th>
<th>Whites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own home computer</td>
<td>86</td>
<td>1173</td>
</tr>
<tr>
<td>PC access at work</td>
<td>100</td>
<td>1132</td>
</tr>
</tbody>
</table>

Do higher-income blacks and whites differ significantly at the 5% level in the proportion who own home computers? Do they differ significantly in the proportion who have PC access at work? Use the four-step process to organize your answers.

21.12 Sickle-cell and malaria. Sickle-cell trait is a hereditary condition that is common among blacks and can cause medical problems. Some biologists suggest that sickle-cell trait protects against malaria. That would explain why it is found in people who originally came from Africa, where malaria is common. A study in Africa tested 543 children for the sickle-cell trait and also for malaria. In all, 136 of the children had the sickle-cell trait, and 36 of these had heavy malaria infections. The other 407 children lacked the sickle-cell trait, and 152 of them had heavy malaria infections.\textsuperscript{77}

(a) Give a 95% confidence interval for the proportion of all children in the population studied who have the sickle-cell trait.
(b) Is there good evidence that the proportion of heavy malaria infections is lower among children with the sickle-cell trait?

21.13 Side effects of medication. A study of “adverse symptoms” in users of over-the-counter pain relief medications assigned subjects at random to one of two common pain relievers: acetaminophen and ibuprofen. (Both of these pain relievers are sold under various brand names, sometimes combined with other ingredients.) In all, 650 subjects took acetaminophen, and 44 experienced some adverse symptom.
Of the 347 subjects who took ibuprofen, 49 had an adverse symptom. How strong is the evidence that the two pain relievers differ in the proportion of people who experience an adverse symptom?
(a) State hypotheses and check that you can use the z test.
(b) Find the P-value of the test and give your conclusion.

21.14 Australia versus the U.S. A study comparing American and Australian corporations examined a sample of 133 American and 63 Australian corporations. There are the usual practical difficulties involving nonresponse and the question of what population the samples represent. Ignore these issues and treat the samples as SRSs from the United States and Australia. The average percent of revenues from “highly regulated businesses” was 27% among the Australian companies and 41% among the American companies. 

(a) The data are given as percents. Explain carefully why comparing the percent of revenues from highly regulated businesses for U.S. and Australian corporations is not a comparison of two population proportions.
(b) What test would you use to make the comparison? (Don’t try to carry out a test.)

21.15 Exercising to lose weight. Experience shows that overweight people find it tough to keep exercising. Perhaps they will do better with several short sessions each day rather than one longer session. Perhaps having exercise equipment at home will help. An experiment looked at these issues. The subjects were women aged 25 to 45 whose weights were 20% to 75% higher than ideal. The study report says:

Subjects were randomly assigned to 1 of 3 groups. All groups were prescribed a similar volume of exercise. The 3 groups differed in the way the exercise was prescribed. 

Long-Bout Exercise Group Forty-nine subjects were instructed to exercise 5 d/wk; duration progressed from 20 min/d ...to 40 min/d ...Participants performed the exercise in one long bout.

Short-Bout Exercise Group Fifty-one subjects were instructed to exercise 5 d/wk ...However, rather than exercising continuously for the prescribed duration, subjects were instructed to divide the exercise into multiple 10-minute bouts that were performed at convenient times throughout the day.

Short-Bout Plus Exercise Equipment Group The exercise prescription was identical to the exercise prescribed for the short-bout group ...The 48 subjects in this group were also provided with motorized home treadmills.

The researchers recorded weight, fitness, and whether the subject continued the exercise program.
(a) Use a diagram to outline the design of this experiment.
(b) How many subjects are there in all? Use Table B starting at line 114 to choose the first 10 subjects for the long-bout group.
21.16 Exercising to lose weight. Here is part of the summary of the results of the study described in the previous exercise:

There was no significant difference between the LB and SB groups for mean (SD) weight loss at 18 months (LB, −5.8 [7.1] kg; SB, −3.7 [6.6] kg.)

That’s pretty terse, but it gives the mean and standard deviation of the weight loss (in kilograms) for the long bout (LB) and short bout (SB) groups. The data come from the 37 LB subjects and 36 SB subjects who completed the study. Do an appropriate analysis, following the four-step process, to confirm the report that there is not a significant difference in the mean weight loss in the two groups.

21.17 Child-care workers. The Current Population Survey (CPS) is the monthly government sample survey of 60,000 households that provides data on employment in the United States. A study of child-care workers drew a sample from the CPS data tapes. We can consider this sample to be an SRS from the population of child-care workers. Out of 2455 child-care workers in private households, 7% were black. Of 1191 nonhousehold child-care workers, 14% were black.
(a) Give a 99% confidence interval for the proportion of nonhousehold child-care workers who are black.
(b) Give a 99% confidence interval for the difference in the percents of these groups of workers who are black. Is the difference statistically significant at the $\alpha = 0.01$ level?

21.18 Child-care workers, continued. The study described in the previous exercise also examined how many years of school child-care workers had. For household workers, the mean and standard deviation were $\bar{x}_1 = 11.6$ years and $s_1 = 2.2$ years. For nonhousehold workers, $\bar{x}_2 = 12.2$ years and $s_2 = 2.1$ years.
(a) Give a 99% confidence interval for the mean years of education of nonhousehold childcare workers.
(b) Give a 99% confidence interval for the difference in mean years of education for the two groups. Is the difference significant at the $\alpha = 0.01$ level?

Chapter 23

23.1 Treating cocaine addiction. Cocaine addicts need the drug to feel pleasure. Perhaps giving them a medication that fights depression will help them stay off cocaine. A three-year study compared an antidepressant called desipramine with lithium (a standard treatment for cocaine addiction) and a placebo. The subjects were 72 chronic users of cocaine who wanted to break their drug habit. Twenty-four of the subjects were randomly assigned to each treatment. Here are the counts and proportions of the subjects who avoided relapse into cocaine use during the study:

<table>
<thead>
<tr>
<th>Group</th>
<th>Treatment</th>
<th>Subjects</th>
<th>No relapse</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Desipramine</td>
<td>24</td>
<td>14</td>
<td>0.583</td>
</tr>
<tr>
<td>2</td>
<td>Lithium</td>
<td>24</td>
<td>6</td>
<td>0.250</td>
</tr>
<tr>
<td>3</td>
<td>Placebo</td>
<td>24</td>
<td>4</td>
<td>0.167</td>
</tr>
</tbody>
</table>
Chapter 23

Does data analysis suggest that desipramine is more successful than the other two treatments? Are there significant differences among the outcomes for the treatments?

23.2 Smoking by students and their parents. Exercise 6.2 gives these data on the smoking habits of students and of their parents.

<table>
<thead>
<tr>
<th></th>
<th>Student smokes</th>
<th>Student does not smoke</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both parents smoke</td>
<td>400</td>
<td>1380</td>
</tr>
<tr>
<td>One parent smokes</td>
<td>416</td>
<td>1823</td>
</tr>
<tr>
<td>Neither parent smokes</td>
<td>188</td>
<td>1168</td>
</tr>
</tbody>
</table>

(a) Find the percent of students who smoke in each of the three parent groups. Make a graph to compare these percents. Describe the association between parent smoking and student smoking.

(b) Explain in words what the null hypothesis for the chi-square test says about student smoking.

(c) Find the expected counts if H₀ is true, and display them in a two-way table similar to the table of observed counts.

(d) Compare the tables of observed and expected counts. Explain how the comparison expresses the same association you saw in (a).

(e) Give the chi-square statistic and its P-value. Examine the terms of chi-square to confirm the pattern you saw in (a) and (d). What is your overall conclusion?

23.3 Unhappy HMO patients. Exercise 21.3 compared HMO members who filed complaints with an SRS of members who did not complain. The study actually broke the complainers into two subgroups: those who filed complaints about medical treatment and those who filed nonmedical complaints. Here are the data on the total number in each group and the number who voluntarily left the HMO:

<table>
<thead>
<tr>
<th></th>
<th>No Medical complaint</th>
<th>Medical complaint</th>
<th>Nonmedical complaint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>743</td>
<td>199</td>
<td>440</td>
</tr>
<tr>
<td>Left</td>
<td>22</td>
<td>26</td>
<td>28</td>
</tr>
</tbody>
</table>

(a) Find the percent of each group who left.

(b) Make a two-way table of complaint status by left or not.

(c) Find the expected counts and check that you can safely use the chi-square test.

(d) The chi-square statistic for this table is \( X^2 = 31.765 \). What null and alternative hypotheses does this statistic test? What are its degrees of freedom? Why is it clear without looking at a table that this value of \( X^2 \) is highly significant?

(e) Use Table E to approximate the P-value. What do you conclude from these data?

23.4 Stress and heart attacks. You read a newspaper article that describes a study of whether stress management can help reduce heart attacks. The 107 subjects all had reduced blood flow to the heart and so were at risk of a heart attack. They were assigned at random to three groups. The article goes on to say:
One group took a four-month stress management program, another underwent a four-month exercise program and the third received usual heart care from their personal physicians.

In the next three years, only three of the 33 people in the stress management group suffered “cardiac events,” defined as a fatal or non-fatal heart attack or a surgical procedure such as a bypass or angioplasty. In the same period, seven of the 34 people in the exercise group and 12 out of the 40 patients in usual care suffered such events.\(^2\)

(a) Use the information in the news article to make a two-way table that describes the study results.

(b) What are the success rates of the three treatments in avoiding cardiac events?

(c) Find the expected cell counts under the null hypothesis that there is no difference among the treatments. Verify that the expected counts meet our guideline for use of the chi-square test.

(d) Is there a significant difference among the success rates for the three treatments?

### 23.5 Nobody is home in July.

Exercise 21.9 describes a study of nonresponse to a national telephone survey in Italy. Here is a table of the percents of responses and of three types of nonresponse at different seasons. The percents in each row add to 100%.

<table>
<thead>
<tr>
<th>Season</th>
<th>Calls made</th>
<th>Successful interviews</th>
<th>No answer</th>
<th>Busy signal</th>
<th>Refusal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan. 1 to Apr. 13</td>
<td>1558</td>
<td>68.5%</td>
<td>21.4%</td>
<td>5.8%</td>
<td>4.3%</td>
</tr>
<tr>
<td>Apr. 21 to June 20</td>
<td>1589</td>
<td>52.4%</td>
<td>35.8%</td>
<td>6.4%</td>
<td>5.4%</td>
</tr>
<tr>
<td>July 1 to Aug. 31</td>
<td>2075</td>
<td>43.4%</td>
<td>41.5%</td>
<td>8.6%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Sept. 1 to Dec. 15</td>
<td>2638</td>
<td>60.0%</td>
<td>30.0%</td>
<td>5.3%</td>
<td>4.7%</td>
</tr>
</tbody>
</table>

(a) What are the degrees of freedom for the chi-square test of the hypothesis that the distribution of responses varies with the season? (Don’t do the test. The sample sizes are so large that the results are sure to be highly significant.)

(b) Consider just the proportion of successful interviews. Describe how this proportion varies with the seasons, and assess the statistical significance of the changes. What do you think explains the changes? (Look at the full table for ideas.)

### 23.6 Incorrect use of chi-square.

It is incorrect to apply the chi-square test to percents rather than to counts. If you enter the 4 × 4 table of percents in the previous exercise into statistical software and ask for a chi-square test, well-written software should give an error message. (Counts must be whole numbers, so the software should check that.) Try this using your software or calculator, and report the result.

### 23.7 Nobody is home in July, continued.

Continue the analysis of the data in Exercise 23.5 by considering just the proportion of people called who refused to participate. We might think that the refusal rate changes less with the season than, for example, the rate of “no answer.” State the hypothesis that the refusal rate does not change with the season. Check that you can safely use the chi-square test. Carry out the test. What do you conclude?
23.8 Preventing strokes. Exercise 21.10 compared aspirin plus another drug with aspirin alone as treatments for patients who had suffered a stroke. The study actually assigned stroke patients at random to four treatments. Here are the data:

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Number of patients</th>
<th>Number of strokes</th>
<th>Number of deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo</td>
<td>1649</td>
<td>250</td>
<td>202</td>
</tr>
<tr>
<td>Aspirin</td>
<td>1649</td>
<td>206</td>
<td>182</td>
</tr>
<tr>
<td>Dipyridamole</td>
<td>1654</td>
<td>211</td>
<td>188</td>
</tr>
<tr>
<td>Both</td>
<td>1650</td>
<td>157</td>
<td>185</td>
</tr>
</tbody>
</table>

(a) Make a two-way table of treatment by whether or not a patient had a stroke during the two-year study period. Compare the rates of strokes for the four treatments. Which treatment appears most effective in preventing strokes? Is there a significant difference among the four rates of strokes? Which terms of chi-square account for most of the total?

(b) The data report two response variables: whether the patient had a stroke and whether the patient died. Repeat your analysis for patient deaths.

(c) Write a careful summary of your overall findings.

23.9 Ebonics awareness. Ebonics, often called “black English,” is a variety of English common among blacks in the United States. How aware are college students of the existence of Ebonics? Here are data from a sample of students at a racially diverse college in the South:

<table>
<thead>
<tr>
<th></th>
<th>Aware</th>
<th>Not aware</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black students</td>
<td>121</td>
<td>11</td>
</tr>
<tr>
<td>White students</td>
<td>159</td>
<td>21</td>
</tr>
<tr>
<td>Other students</td>
<td>75</td>
<td>28</td>
</tr>
</tbody>
</table>

Do both data analysis and a formal test to compare awareness in the three groups of students. Write a clear summary of your findings. Follow the four-step process.

23.10 Survey response rates. To study the export activity of manufacturing firms on Taiwan, researchers mailed questionnaires to an SRS of firms in each of five industries that export many of their products. The response rate was only 12.5%, because private companies don’t like to fill out long questionnaires from academic researchers. Here are data on the planned sample sizes and the actual number of responses received from each industry:

<table>
<thead>
<tr>
<th></th>
<th>Sample size</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metal products</td>
<td>185</td>
<td>17</td>
</tr>
<tr>
<td>Machinery</td>
<td>301</td>
<td>35</td>
</tr>
<tr>
<td>Electrical equipment</td>
<td>552</td>
<td>75</td>
</tr>
<tr>
<td>Transportation equipment</td>
<td>100</td>
<td>15</td>
</tr>
<tr>
<td>Precision instruments</td>
<td>90</td>
<td>12</td>
</tr>
</tbody>
</table>

If the response rates differ greatly, comparisons among the industries may be difficult. Is there good evidence of unequal response rates among the five industries? (Start by creating a two-way table of response or nonresponse by industry.)
23.11 **Python eggs.** How is the hatching of water python eggs influenced by the temperature of the snake’s nest? Researchers assigned newly laid eggs to one of three temperatures: hot, neutral, or cold. Hot duplicates the extra warmth provided by the mother python, and cold duplicates the absence of the mother. Here are the data on the number of eggs and the number that hatched:

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Eggs</th>
<th>Hatched</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cold</td>
<td>27</td>
<td>16</td>
</tr>
<tr>
<td>Neutral</td>
<td>56</td>
<td>38</td>
</tr>
<tr>
<td>Hot</td>
<td>104</td>
<td>75</td>
</tr>
</tbody>
</table>

(a) Make a two-way table of temperature by outcome (hatched or not).
(b) Calculate the percent of eggs in each group that hatched. The researchers anticipated that eggs would not hatch at cold temperatures. Do the data support that anticipation?
(c) Are there significant differences among the proportions of eggs that hatched in the three groups?

23.12 **Standards for child care.** Do unregulated providers of child care in their homes follow different health and safety practices in different cities? A study looked at people who regularly provided care for someone else’s children in poor areas of three cities. The numbers who required medical releases from parents to allow medical care in an emergency were 42 of 73 providers in Newark, N.J., 29 of 101 in Camden, N.J., and 48 of 107 in South Chicago, Ill.

(a) Use the chi-square test to see if there are significant differences among the proportions of child-care providers who require medical releases in the three cities.
(b) How should the data be produced in order for your test to be valid? (In fact, the samples came in part from asking parents who were subjects in another study who provided their child care. The author of the study wisely did not use a statistical test. He wrote: “Application of conventional statistical procedures appropriate for random samples may produce biased and misleading results.”)

23.13 **Colors of M&M’s.** The M&M’s candies Web site [www.mms.com](http://www.mms.com) says that the distribution of colors for milk chocolate M&M’s is

<table>
<thead>
<tr>
<th>Color</th>
<th>Purple</th>
<th>Yellow</th>
<th>Red</th>
<th>Orange</th>
<th>Brown</th>
<th>Green</th>
<th>Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Open a package of M&M’s: out spill 57 candies. (The count varies slightly from package to package.) The color counts are

<table>
<thead>
<tr>
<th>Color</th>
<th>Purple</th>
<th>Yellow</th>
<th>Red</th>
<th>Orange</th>
<th>Brown</th>
<th>Green</th>
<th>Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>11</td>
<td>13</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>3</td>
</tr>
</tbody>
</table>

How well do the counts from this package fit the claimed distribution? Do a chi-square test of fit, report the $P$-value, and give a conclusion.
24.1 Natural gas consumption. The table below contains data on the natural gas consumption of the Sanchez household for 16 months. Gas consumption is higher in cold weather. Here are the average amount of natural gas consumed each day during the month, in hundreds of cubic feet, and the average number of heating degree-days each day during the month. (Degree-days are the number of degrees the average daily temperature falls below 65°F.)

<table>
<thead>
<tr>
<th>Month</th>
<th>Degree-days</th>
<th>Gas (100 cu. ft.)</th>
<th>Month</th>
<th>Degree-days</th>
<th>Gas (100 cu. ft.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov.</td>
<td>24</td>
<td>6.3</td>
<td>July</td>
<td>0</td>
<td>1.2</td>
</tr>
<tr>
<td>Dec.</td>
<td>51</td>
<td>10.9</td>
<td>Aug.</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td>Jan.</td>
<td>43</td>
<td>8.9</td>
<td>Sept.</td>
<td>6</td>
<td>2.1</td>
</tr>
<tr>
<td>Feb.</td>
<td>33</td>
<td>7.5</td>
<td>Oct.</td>
<td>12</td>
<td>3.1</td>
</tr>
<tr>
<td>Mar.</td>
<td>26</td>
<td>5.3</td>
<td>Nov.</td>
<td>30</td>
<td>6.4</td>
</tr>
<tr>
<td>Apr.</td>
<td>13</td>
<td>4.0</td>
<td>Dec.</td>
<td>32</td>
<td>7.2</td>
</tr>
<tr>
<td>May</td>
<td>4</td>
<td>1.7</td>
<td>Jan.</td>
<td>52</td>
<td>11.0</td>
</tr>
<tr>
<td>June</td>
<td>0</td>
<td>1.2</td>
<td>Feb.</td>
<td>30</td>
<td>6.9</td>
</tr>
</tbody>
</table>

(a) We want to predict gas used from degree-days. Make a scatterplot of the data with this goal in mind. Use your calculator to find the correlation \( r \) and the equation of the least-squares regression line. Describe the form and strength of the relationship.

(b) The model for regression inference has three parameters, which we call \( \alpha \), \( \beta \), and \( \sigma \). Which of these parameters does the least-squares line estimate? What are the estimated values?

24.2 An extinct beast. Exercise 4.8 gives these measurements on the lengths in centimeters of the femur (a leg bone) and the humerus (a bone in the upper arm) for the five fossil specimens of the extinct beast *Archaeopteryx* that preserve both bones:

<table>
<thead>
<tr>
<th>Femur</th>
<th>38</th>
<th>56</th>
<th>59</th>
<th>64</th>
<th>74</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humerus</td>
<td>41</td>
<td>63</td>
<td>70</td>
<td>72</td>
<td>84</td>
</tr>
</tbody>
</table>

The strong linear relationship between the lengths of the two bones helped persuade scientists that all five specimens belong to the same species.

(a) Examine the data. Make a scatterplot with femur length as the explanatory variable. Use your calculator to obtain the correlation \( r \) and the equation of the least-squares regression line. Do you think that femur length will allow good prediction of humerus length?

(b) Explain in words what the slope \( \beta \) of the true regression line says about *Archaeopteryx*. Based on the data, what are the estimates of \( \beta \) and the intercept \( \alpha \) of the true regression line?

(c) Calculate the residuals for the five data points. Check that their sum is 0 (up to roundoff error). Use the residuals to estimate the standard deviation \( \sigma \) in the regression model. You have now estimated all three parameters in the model.
24.3 Natural gas consumption, continued. Here is Minitab output for the regression of natural gas consumed by the Sanchez household on degree-days:

The regression equation is
Gas = 1.09 + 0.189 D-days

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>StDev</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.0892</td>
<td>0.1389</td>
<td>7.84</td>
<td>0.000</td>
</tr>
<tr>
<td>D-days</td>
<td>0.188999</td>
<td>0.004934</td>
<td>38.31</td>
<td>0.000</td>
</tr>
</tbody>
</table>

S = 0.3389  R-Sq = 99.1%

Use the information in this output to give a 90% confidence interval for the slope $\beta$ of the true regression line. Explain clearly what your result tells us about how gas usage responds to falling temperatures.

24.4 An extinct beast, continued. Exercise 24.2 presents data on the lengths of two bones in five fossil specimens of the extinct beast *Archaeopteryx*. In that exercise, you estimated the parameters using only a basic calculator. Software tells us that the least-squares slope is $b = 1.1969$ with standard error $SE_b = 0.0751$.
(a) What is the $t$ statistic for testing $H_0: \beta = 0$?
(b) How many degrees of freedom does $t$ have? Use Table C to approximate the $P$-value of $t$ against the one-sided alternative $H_a: \beta > 0$. What do you conclude?

24.5 Natural gas consumption: prediction. After the data in Exercise 24.1 were collected, the Sanchez family installed solar panels.
(a) In the month of January after solar panels were installed, there were 40 degree-days per day. How much gas do you predict the Sanchez household would have used per day without the solar panels? They actually used 7.5 hundred cubic feet per day. How much gas per day did the solar panels save?
(b) Here is the prediction output from Minitab for 40 degree-days per day. Give a 95% interval for the amount of gas that would have been used this January without the solar panels.

<table>
<thead>
<tr>
<th>Fit</th>
<th>StDev Fit</th>
<th>95.0% CI</th>
<th>95.0% PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.6492</td>
<td>0.1216</td>
<td>(8.3883, 8.9100)</td>
<td>(7.8768, 9.4215)</td>
</tr>
</tbody>
</table>

(c) Give a 95% interval for the mean gas consumption per day in all months with 40 degree-days per day.
(d) Minitab gives only one of the two standard errors used in prediction. It is $SE_\hat{\mu}$, the standard error for estimating the mean response. Use Table C and this standard error to give a 90% confidence interval for mean gas consumption per day in all months with 40 degree-days per day.

24.6 Natural gas consumption: residuals. Find the residuals for the Sanchez household gas consumption data in Exercise 24.1.
(a) Display the distribution of the residuals in a plot. It is hard to assess the shape of a distribution from only 16 observations. Do the residuals appear roughly symmetric? Are there any outliers?
(b) Plot the residuals against the explanatory variable, degree-days. Draw a horizontal line at height 0 on the plot. Is there clear evidence of a nonlinear relationship? Does the variation about the line appear roughly the same as the number of degree-days changes?

24.7 The endangered manatee. Exercise 4.4 gives the number of Florida power-boat registrations (in thousands) and the number of manatees killed by boats in Florida in the years 1977 to 2005.
(a) Make a scatterplot showing the relationship between powerboats registered and manatees killed. Is the overall pattern roughly linear? Are there clear outliers or strongly influential data points?
(b) Regression software ignores the three years for which boat registrations are not available. Here is part of the output from the CrunchIt! regression command:

```
Estimate of error standard deviation: 8.04709
R-Sq = 0.8943
```

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std.Err</th>
<th>DF</th>
<th>T-Stat</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-43.3046</td>
<td>6.49909</td>
<td>24</td>
<td>-6.663</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Boats</td>
<td>0.128178</td>
<td>0.00899</td>
<td>24</td>
<td>14.251</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Why are there 24 degrees of freedom? What does $r^2$ tell you about the relationship between boats and manatees killed?
(c) Explain what the slope $\beta$ of the true regression line means in this setting. Give a 90% confidence interval for $\beta$.

24.8 Is wine good for your heart? There is some evidence that drinking moderate amounts of wine helps prevent heart attacks. Exercise 4.6 gives data on yearly wine consumption (liters of alcohol from drinking wine, per person) and yearly deaths from heart disease (deaths per 100,000 people) in 19 developed nations. Is there statistically significant evidence that the correlation between wine consumption and heart disease deaths is negative? Use Table F.

(a) Based on these data, you want to predict the number of manatees killed in a year when 1 million powerboats are registered. Use the regression line from the software output in Exercise 24.7 to give a prediction.
(b) Here is the result of asking CrunchIt! to do prediction for $x^* = 1000$:

<table>
<thead>
<tr>
<th>X value</th>
<th>Pred. Y</th>
<th>s.e.(Pred.y)</th>
<th>95% C.I.</th>
<th>95% P.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>84.8736</td>
<td>3.11885</td>
<td>(78.437, 91.311)</td>
<td>(67.061, 102.686)</td>
</tr>
</tbody>
</table>

Check that the prediction 84.8736 agrees with your result in (a). Then give a 95% interval for the number of manatees that would be killed in a future year when 1 million boats are registered.

24.10 City and highway gas mileage. Exercise 1.3 gives the city and highway gas mileages for 22 models of two-seater cars. The Honda Insight, a gas-electric hybrid car, is an outlier in both the $x$ and $y$ directions. Exercise 5.14 asks you to
investigate the influence of the Insight on the least-squares line. The influence is not large because the Insight does not lie far from the least-squares line calculated from the other 21 cars. Now you will investigate the influence of the Insight on inference. Carry out regression both with and without the Insight.

(a) How does the Insight influence the value of the $t$ statistic for the regression slope and its $P$-value? Is the influence practically important?
(b) How does the Insight influence the 95% confidence interval for the regression slope? Is the influence practically important?

24.11 More manatee predictions. Exercise 24.9 gives 95% intervals for predicting manatee deaths when 1 million powerboats are registered. The output gives only one of the two standard errors used in prediction. It is $\text{SE}_\mu$, the standard error for estimating the mean response. Use Table C and this standard error to give a 90% confidence interval for the mean number of manatees killed in all years when 1 million boats are registered.

24.12 Do heavy people burn more energy? Metabolic rate, the rate at which the body consumes energy, is important in studies of weight gain, dieting, and exercise. Lean body mass is an important influence on metabolic rate. Exercise 4.7 gives data for 19 people. Because men and women showed a similar pattern, we will now ignore gender. Here are the data on mass (in kilograms) and metabolic rate (in calories):

<table>
<thead>
<tr>
<th>Mass</th>
<th>62.0</th>
<th>62.9</th>
<th>36.1</th>
<th>54.6</th>
<th>48.5</th>
<th>42.0</th>
<th>47.4</th>
<th>50.6</th>
<th>42.0</th>
<th>48.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate</td>
<td>1792</td>
<td>1666</td>
<td>995</td>
<td>1425</td>
<td>1396</td>
<td>1418</td>
<td>1362</td>
<td>1502</td>
<td>1256</td>
<td>1614</td>
</tr>
<tr>
<td>Mass</td>
<td>40.3</td>
<td>33.1</td>
<td>51.9</td>
<td>42.4</td>
<td>34.5</td>
<td>51.1</td>
<td>41.2</td>
<td>51.9</td>
<td>46.9</td>
<td></td>
</tr>
<tr>
<td>Rate</td>
<td>1189</td>
<td>913</td>
<td>1460</td>
<td>1124</td>
<td>1052</td>
<td>1347</td>
<td>1204</td>
<td>1867</td>
<td>1439</td>
<td></td>
</tr>
</tbody>
</table>

Use software to analyze these data. Make a scatterplot and find the least-squares line. Give a 90% confidence interval for the slope $\beta$ and explain clearly what your interval says about the relationship between lean body mass and metabolic rate. Find the residuals and examine them. Are the assumptions for regression inference met?

24.13 Confidence interval for the intercept (Optional). Exercise 24.3 gives Minitab output for the regression of household natural gas consumption on degree-days. You used the output to find a 90% confidence interval for the slope $\beta$ of the true regression line. The intercept $\alpha$ of the true regression line is the average amount of gas the household consumes when there are zero degree-days. (There are zero degree-days when the average temperature is 65$^\circ$F or above.) Confidence intervals for $\alpha$ have the form

$$a \pm t^\star \text{SE}_\alpha$$

Use the intercept $a$ of the least-squares line and its standard error, given in the computer output, to find a 95% confidence interval for $\alpha$. (The degrees of freedom are $n - 2$ once again.)
Chapter 25

For each setting described in Exercises 25.1 to 25.3, identify the populations and the response variable. Then give \( I \), the \( n_i \), and \( N \). Finally, give the degrees of freedom of the ANOVA \( F \) test.

25.1 **Strong concrete.** The strength of concrete depends on the mixture of sand, gravel, and cement used to prepare it. A study compares five different mixtures. Workers prepare six batches of each mixture and measure the strength of the concrete made from each batch.

25.2 **Which package design is best?** A maker of detergents wants to compare the attractiveness to consumers of six package designs. Each package is shown to 120 different consumers who rate the attractiveness of the design on a 1 to 10 scale.

25.3 **Teaching sign language.** Which of four methods of teaching American Sign Language is most effective? Assign 10 of the 42 students in a class at random to each of three methods. Teach the remaining 12 students by the fourth method. Record the students’ scores on a standard test of sign language after a semester’s study.

25.4 **Plant diversity.** Does the presence of more species of plants increase the productivity of a natural area, as measured by the total mass of plant material? An experiment to investigate this question started as follows:

> In a 7-year experiment, we controlled one component of diversity, the number of plant species, in 168 plots, each 9 m by 9 m. We seeded the plots, in May 1994, to have 1, 2, 4, 8, or 16 species, with 39, 35, 29, 30, and 35 replicates, respectively.\(^{88}\)

Total plant mass in each plot was measured in 2000.

(a) What null and alternative hypotheses does ANOVA test here?

(b) What are the values of \( N \), \( I \), and the \( n_i \)?

(c) What are the degrees of freedom of the ANOVA \( F \) statistic?

25.5 **Smoking among French men.** A study of smoking categorized a sample of French men aged 20 to 60 years as nonsmokers (146 men), former smokers (125 men), moderate smokers (104 men), or heavy smokers (84 men). One set of response variables was the amount of several types of food consumed, in grams.\(^{89}\)

(a) What are the degrees of freedom for an ANOVA \( F \) statistic that compares the mean weights of a food consumed in the 4 populations? Use Table D with 3 and 200 degrees of freedom to give an approximate \( P \)-value for each of the following foods:

(b) For milk, \( F = 3.58 \).

(c) For eggs, \( F = 1.08 \).

(d) For vegetables, \( F = 6.02 \).

25.6 **Who succeeds in college?** What factors influence the success of college students? Look at all 256 students who entered a university planning to study computer science (CS) in a specific year. We are willing to regard these students as a random sample of the students the university CS program will attract in subsequent years. After three semesters of study, some of these students were CS majors, some
were majors in another field of science or engineering, and some had left science and engineering or left the university. The table below gives the sample means and standard deviations and the ANOVA $F$ statistics for three variables that describe the students’ high school performance. These are three separate ANOVA $F$ tests.\footnote{90}

The first variable is a student’s rank in the high school class, given as a percentile (so rank 50 is the middle of the class and rank 100 is the top). The next variable is the number of semester courses in mathematics the student took in high school. The third variable is the student’s average grade in high school mathematics. The mean and standard deviation appear in a form common in published reports, with the standard deviation in parentheses following the mean.

<table>
<thead>
<tr>
<th>Group</th>
<th>$n$</th>
<th>Mean (Standard Deviation)</th>
<th>Mean (Standard Deviation)</th>
<th>Mean (Standard Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High school rank</td>
<td>HS math</td>
<td>in HS math</td>
</tr>
<tr>
<td>CS majors</td>
<td>103</td>
<td>88.0 (10.5)</td>
<td>8.74 (1.28)</td>
<td>3.61 (0.46)</td>
</tr>
<tr>
<td>Sci./eng. majors</td>
<td>31</td>
<td>89.2 (10.8)</td>
<td>8.65 (1.31)</td>
<td>3.62 (0.40)</td>
</tr>
<tr>
<td>Other</td>
<td>122</td>
<td>85.8 (10.8)</td>
<td>8.25 (1.17)</td>
<td>3.35 (0.55)</td>
</tr>
<tr>
<td>$F$ statistic</td>
<td></td>
<td>1.95</td>
<td>4.56</td>
<td>9.38</td>
</tr>
</tbody>
</table>

(a) What null and alternative hypotheses does $F$ test for rank in the high school class? Express the hypotheses both in symbols and in words. The hypotheses are similar for the other two variables.

(b) What are the degrees of freedom for each $F$?

(c) Check that the standard deviations allow use of all three $F$ tests. The shapes of the distributions also allow use of $F$. How significant is $F$ for each of these variables?

(d) Write a brief summary of the differences among the three groups of students, taking into account both the significance of the $F$ tests and the values of the means.

25.7 How much corn should I plant? How much corn per acre should a farmer plant to obtain the highest yield? Too few plants will give a low yield. On the other hand, if there are too many plants, they will compete with each other for moisture and nutrients, and yields will fall. To find out, plant at different rates on several plots of ground and measure the harvest. (Treat all the plots the same except for the planting rate.) Use software to analyze these data from such an experiment.\footnote{91}

<table>
<thead>
<tr>
<th>Plants per acre</th>
<th>Yield (bushels per acre)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12,000</td>
<td>150.1 113.0 118.4 142.6</td>
</tr>
<tr>
<td>16,000</td>
<td>166.9 120.7 135.2 149.8</td>
</tr>
<tr>
<td>20,000</td>
<td>165.3 130.1 139.6 149.9</td>
</tr>
<tr>
<td>24,000</td>
<td>134.7 138.4 156.1</td>
</tr>
<tr>
<td>28,000</td>
<td>119.0 150.5</td>
</tr>
</tbody>
</table>

(a) Do data analysis to see what the data appear to show about the influence of plants per acre on yield and also to check the conditions for ANOVA.

(b) Carry out the ANOVA $F$ test. State hypotheses; give $F$ and its $P$-value. What do you conclude?

(c) The observed differences among the mean yields in the samples are quite large. Why are they not statistically significant?
25.8 Weights of newly hatched pythons. A study of the effect of nest temperature on the development of water pythons separated python eggs at random into nests at three temperatures: cold, neutral, and hot. Exercise 23.11 shows that the proportions of eggs that hatched at each temperature did not differ significantly. Now we will examine the little pythons. In all, 16 eggs hatched at the cold temperature, 38 at the neutral temperature, and 75 at the hot temperature. The report of the study summarizes the data in the common form “mean ± standard error” as follows:

<table>
<thead>
<tr>
<th>Temperature</th>
<th>n</th>
<th>Weight (grams) at hatching</th>
<th>Propensity to strike</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cold</td>
<td>16</td>
<td>28.89 ± 8.08</td>
<td>6.40 ± 5.67</td>
</tr>
<tr>
<td>Neutral</td>
<td>38</td>
<td>32.93 ± 5.61</td>
<td>5.82 ± 4.24</td>
</tr>
<tr>
<td>Hot</td>
<td>75</td>
<td>32.27 ± 4.10</td>
<td>4.30 ± 2.70</td>
</tr>
</tbody>
</table>

(a) We will compare the mean weights at hatching. Recall that the standard error of the mean is \( s/\sqrt{n} \). Find the standard deviations of the weights in the three groups and verify that they satisfy our rule of thumb for using ANOVA.

(b) Starting from the sample sizes \( n_i \), the means \( \bar{x}_i \), and the standard deviations \( s_i \), carry out an ANOVA. That is, find MSG, MSE, and the \( F \) statistic, and use Table D to approximate the \( P \)-value. Is there evidence that nest temperature affects the mean weight of newly hatched pythons?

25.9 Python strikes. The data in the previous exercise also describe the “propensity to strike” of the hatched pythons at 30 days of age. This is the number of taps on the head with a small brush until the python launches a strike. (Don’t try this with adult pythons.) The data are again summarized in the form “sample mean ± standard error of the mean.” Follow the outline in (a) and (b) of the previous exercise for propensity to strike. Does nest temperature appear to influence propensity to strike?

25.10 Earnings of athletic trainers. How much do newly hired athletic trainers earn? Earnings depend on the educational and other qualifications of the trainer and on the type of job. Here are summaries from a sample survey that contacted institutions advertising for trainers and asked about the people they hired:

<table>
<thead>
<tr>
<th>Type of institution</th>
<th>n</th>
<th>Mean salary</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>High school</td>
<td>57</td>
<td>$30,625</td>
<td>$10,999</td>
</tr>
<tr>
<td>Clinic</td>
<td>108</td>
<td>$35,415</td>
<td>$5211</td>
</tr>
<tr>
<td>College</td>
<td>94</td>
<td>$34,731</td>
<td>$8282</td>
</tr>
</tbody>
</table>

Calculate the ANOVA table and the \( F \) statistic. Are there significant differences among the mean salaries for the three types of job in the population of all newly hired trainers? (Comments: The study authors did an ANOVA. As is often the case, the published work does not answer questions about the suitability of the analysis. We wonder if responses from 60% of employers who advertised positions generate a random sample of new hires. The actual data do not appear in the report. The salary distributions are no doubt right-skewed; we hope that strong skewness and outliers are absent because the subjects in each group hold similar jobs. The large
samples will then justify ANOVA. Also, the sample standard deviations do not quite satisfy our rule of thumb for safe use of ANOVA.)

**Notes and Data Sources**


2. These data were collected by students as a class project.


8. See Note 1.


19. From a survey by the Wheat Industry Council reported in *USA Today*, October 20, 1983.


22. Peter Cook, Purdue University.


27. Based closely on Susan B. Sorenson, “Regulating firearms as a consumer product,” *Science*, 286 (1999), pp. 1481–1482. Because the results in the paper were “weighted to the U.S. population,” I have changed some counts slightly for consistency.


35. From the Bureau of the Census’s 1998 American Housing Survey.


41. Darlene Gordon, Purdue University.


44. See Note 9.


70. See Note 60.

71. See Note 63.


87. Robert Dale, Purdue University.


