SUPPLEMENTARY EXERCISES

for

INTRODUCTION TO THE PRACTICE OF STATISTICS

Fifth Edition

David S. Moore

and

George P. McCabe
These exercises appeared in the third or fourth editions of *Introduction to the Practice of Statistics*. They were replaced in the fifth edition in the interest of freshness, but they remain high-quality exercises that supplement those in the text.
CHAPTER 1

Section 1.1

1.1 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 Political party preference in the United States depends in part on the age, income, and gender of the voter. A political scientist selects a large sample of registered voters. For each voter, she records gender, age, household income, and whether they voted for the Democratic or for the Republican candidate in the last congressional election. Which of these variables are categorical and which are quantitative?

1.3 Here is a small part of a data set that describes mutual funds available to the public:

<table>
<thead>
<tr>
<th>Fund</th>
<th>Category</th>
<th>Net assets (millions of $)</th>
<th>Year-to-date return</th>
<th>Largest holding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fidelity Low-Priced Stock</td>
<td>Small value</td>
<td>6,189</td>
<td>4.56%</td>
<td>Dallas Semiconductor</td>
</tr>
<tr>
<td>Price International Stock</td>
<td>International stock</td>
<td>9,745</td>
<td>−0.45%</td>
<td>Vodafone</td>
</tr>
<tr>
<td>Vanguard 500 Index</td>
<td>Large blend</td>
<td>89,394</td>
<td>3.45%</td>
<td>General Electric</td>
</tr>
</tbody>
</table>


What individuals does this data set describe? In addition to the fund’s name, how many variables does the data set contain? Which of these variables are categorical and which are quantitative?

1.4 Congress wants the medical establishment to show that progress is being made in fighting cancer. Some variables that might be used are: (a) Total deaths from cancer. These have risen over time, from 331,000 in 1970 to 505,000 in 1980 to 550,000 in 1999.
(b) The percent of all Americans who die from cancer. The percent of deaths due to cancer has also risen steadily, from 17.2% in 1970 to 20.9% in 1980 to 23.0% in 1999.
(c) The percent of cancer patients who survive for five years from the time the disease is discovered. These rates are rising slowly. For whites, the five-year survival rate was 50.8% in the 1974 to 1979 period and 60.9% from 1989 to 1995.
Discuss the usefulness of each of these variables as a measure of the effectiveness of cancer treatment. In particular, explain why both (a) and (b) could increase even if treatment is getting more effective, and why (c) could increase even if treatment is getting less effective.

1.5 The National Highway Traffic Safety Administration says that an average of 11 children die each year in school bus accidents, and an average of 600 school-age children die each year in auto accidents during school hours. These numbers suggest that riding the bus is safer than driving to school with a parent. The counts aren’t fully convincing, however. What rates would you like to know to compare the safety of bus and private auto?

1.6 Popular magazines often rank cities in terms of how desirable it is to live and work in each city. Describe five variables that you would measure for each city if you were designing such a study. Give reasons for each of your choices.
1.7 All the members of a physical education class are asked to measure their pulse rate as they sit in the classroom. The students use a variety of methods. Method 1: count heart beats for 6 seconds and multiply by 10 to get beats per minute. Method 2: count heart beats for 30 seconds and multiply by 2 to get beats per minute.

(a) Which method do you prefer? Why?
(b) One student proposes a third method: starting exactly on a heart beat, measure the time needed for 50 beats and convert this time into beats per minute. This method is more accurate than either method in (a). Why?

1.8 Each year *Fortune* magazine lists the top 500 companies in the United States, ranked according to their total annual sales in dollars. Describe three other variables that could reasonably be used to measure the “size” of a company.

1.9 You are writing an article for a consumer magazine based on a survey of the magazine’s readers on the reliability of their household appliances. Of 13,376 readers who reported owning Brand A dishwashers, 2942 required a service call during the past year. Only 192 service calls were reported by the 480 readers who owned Brand B dishwashers. Describe an appropriate variable to measure the reliability of a make of dishwasher, and compute the values of this variable for Brand A and for Brand B.

1.10 In 1997, there were 12,298,000 undergraduate students in U.S. colleges. According to the U.S. Department of Education, there were 127,000 American Indian or Alaskan Native students, 737,000 Asian or Pacific Islander, 1,380,000 non-Hispanic black, 1,108,000 Hispanic, and 8,682,000 non-Hispanic white students. In addition, 265,000 foreign undergraduates were enrolled in U.S. colleges.

(a) Each number, including the total, is rounded to the nearest thousand. Separate rounding may cause roundoff errors, so that the sum of the counts does not equal
the total given. Are roundoff errors present in these data?

(b) Present the data in a graph.

1.11 The number of deaths among persons aged 15 to 24 years in the United States in 1999 due to the eight leading causes of death for this age group were: accidents, 13,602; homicide, 4989; suicide, 3885; cancer, 1724; heart disease, 1048; congenital defects, 430; respiratory disease, 208; AIDS, 197.

(a) Make a bar graph to display these data.

(b) What additional information do you need to make a pie chart?

1.12 According to the 2000 census, there are 105.5 million households in the United States. A household consists of people living together in the same residence, regardless of their relationship to each other. Of these, 71.8 million were “family households” in which at least one other person was related to the householder by blood, marriage, or adoption. The family households include 54.5 million headed by a married couple and 17.3 million other families (for example, a single parent with children). The other 33.7 million households are “nonfamily households.” Of these, 27.2 million contain a person living alone, 5.5 million are unmarried couples living together, and 1 million consist of other unrelated people living together. Be creative: make a bar graph that displays these facts, including the distinction between family and nonfamily households.

1.13 Here are the percents of doctoral degrees in each of several subjects that were earned by women in 1997–98: psychology, 67.5%; education, 63.2%; life sciences, 42.5%; business, 31.4%; physical sciences, 25.2%; engineering, 12.2%.

(a) Explain clearly why we cannot use a pie chart to display these data, even if we knew the percent female for every academic subject.

(b) Make a bar graph of the data. (Comparisons are easier if you order the bars by height, which is the order in which we give the percents.)
1.14 Outliers are sometimes the most interesting feature of a distribution. Figure 1.13 displays the distribution of batting averages for all 167 American League baseball players who batted at least 200 times in the 1980 season. The outlier is the .390 batting average of George Brett, the highest batting average in the major leagues since Ted Williams hit .406 in 1941. (See Exercise 1.87 on page xxx for a comparison of Brett and Williams.) Is the overall shape (ignoring the outlier) roughly symmetric or clearly skewed? What is the approximate midpoint of American League batting averages? What is the range if we ignore the outlier?

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

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

The college also administers the test to a sample of 20 first-year college men. Their scores are

\[
108 \ 140 \ 114 \ 91 \ 180 \ 115 \ 126 \ 92 \ 169 \ 146 \\
109 \ 132 \ 75 \ 88 \ 113 \ 151 \ 70 \ 115 \ 187 \ 104
\]

(a) Make a back-to-back stemplot of the men’s and women’s scores. The overall shapes of the distributions are indistinct, as often happens when only a few observations are available. Are there any outliers?

(b) Compare the midpoints and the ranges of the two distributions. What is the most noticeable contrast between the female and male scores?

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

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

Make a back-to-back stemplot of these data. Report the approximate midpoints of both groups. Does it appear that the chicks fed high-lysine corn grew faster? Are there any outliers or other problems?

1.17 There is some evidence that increasing the amount of calcium in the diet can lower blood pressure. In a medical experiment one group of men was given a daily calcium supplement, while a control group received a placebo (a dummy pill). The seated systolic blood pressure of all the men was measured before the treatments began and again after 12 weeks. The blood pressure distributions in the two groups should have been similar at the beginning of the experiment. Here are the initial blood pressure readings for the two groups:

**Calcium group**

107 110 123 129 112 111 107 112 136 102

**Placebo group**

123 109 112 102 98 114 119 112 110 117 130

Make a back-to-back stemplot of these data. Does your plot show any major differences in the two groups before the treatments began? In particular, are the centers of the two blood pressure distributions close together?
1.18 The Degree of Reading Power (DRP) test is often used to measure the reading ability of children. Here are the DRP scores of 44 third-grade students, measured during research on ways to improve reading performance. (Data provided by Mari-beth Cassidy Schmitt, from her Ph.D dissertation, “The effects of an elaborated directed reading activity on the metacomprehension skills of third graders,” Purdue University, 1987.)

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

Make a stemplot of these data. Then make a histogram. Which display do you prefer, and why? Describe the main features of the distribution.

1.19 The table below gives the number of medical doctors per 100,000 people in each state.
<table>
<thead>
<tr>
<th>State</th>
<th>Doctors</th>
<th>State</th>
<th>Doctors</th>
<th>State</th>
<th>Doctors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>198</td>
<td>Louisiana</td>
<td>246</td>
<td>Ohio</td>
<td>235</td>
</tr>
<tr>
<td>Alaska</td>
<td>167</td>
<td>Maine</td>
<td>223</td>
<td>Oklahoma</td>
<td>169</td>
</tr>
<tr>
<td>Arizona</td>
<td>202</td>
<td>Maryland</td>
<td>374</td>
<td>Oregon</td>
<td>225</td>
</tr>
<tr>
<td>Arkansas</td>
<td>190</td>
<td>Massachusetts</td>
<td>412</td>
<td>Pennsylvania</td>
<td>291</td>
</tr>
<tr>
<td>California</td>
<td>247</td>
<td>Michigan</td>
<td>224</td>
<td>Rhode Island</td>
<td>338</td>
</tr>
<tr>
<td>Colorado</td>
<td>238</td>
<td>Minnesota</td>
<td>249</td>
<td>South Carolina</td>
<td>207</td>
</tr>
<tr>
<td>Connecticut</td>
<td>354</td>
<td>Mississippi</td>
<td>163</td>
<td>South Dakota</td>
<td>184</td>
</tr>
<tr>
<td>Delaware</td>
<td>234</td>
<td>Missouri</td>
<td>230</td>
<td>Tennessee</td>
<td>246</td>
</tr>
<tr>
<td>Florida</td>
<td>238</td>
<td>Montana</td>
<td>190</td>
<td>Texas</td>
<td>203</td>
</tr>
<tr>
<td>Georgia</td>
<td>211</td>
<td>Nebraska</td>
<td>218</td>
<td>Utah</td>
<td>200</td>
</tr>
<tr>
<td>Hawaii</td>
<td>265</td>
<td>Nevada</td>
<td>173</td>
<td>Vermont</td>
<td>305</td>
</tr>
<tr>
<td>Idaho</td>
<td>154</td>
<td>New Hampshire</td>
<td>237</td>
<td>Virginia</td>
<td>241</td>
</tr>
<tr>
<td>Illinois</td>
<td>260</td>
<td>New Jersey</td>
<td>295</td>
<td>Washington</td>
<td>235</td>
</tr>
<tr>
<td>Indiana</td>
<td>195</td>
<td>New Mexico</td>
<td>212</td>
<td>West Virginia</td>
<td>215</td>
</tr>
<tr>
<td>Iowa</td>
<td>173</td>
<td>New York</td>
<td>387</td>
<td>Wisconsin</td>
<td>227</td>
</tr>
<tr>
<td>Kansas</td>
<td>203</td>
<td>North Carolina</td>
<td>232</td>
<td>Wyoming</td>
<td>171</td>
</tr>
<tr>
<td>Kentucky</td>
<td>209</td>
<td>North Dakota</td>
<td>222</td>
<td>D.C.</td>
<td>737</td>
</tr>
</tbody>
</table>

(a) Why is the number of doctors per 100,000 people a better measure of the availability of health care than a simple count of the number of doctors in a state?

(b) Make a graph to display the distribution of doctors per 100,000 people. Write a brief description of the distribution. Are there any outliers? If so, can you explain them?

1.20 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) The data appear in time order reading from left to right across each row in turn, beginning with the −5.7% return in July 1990. Make a time plot of the data. This was a period of increasing action against smoking, so we might expect a trend toward lower returns. But it was also a period in which stocks in general rose sharply, which would produce an increasing trend. What does your time plot show?

1.21 The distribution of the ages of a nation’s population has a strong influence on economic and social conditions. The following table shows the age distribution of U.S. residents in 1950 and 2050, in millions of people. The 1950 data come from that year's census, while the 2050 data are projections made by the Census Bureau.
Age group | 1950 | 2050
---------- | ---- | ----
Under 10 years | 29.3 | 53.3
10 to 19 years | 21.8 | 53.2
20 to 29 years | 24.0 | 51.2
30 to 39 years | 22.8 | 50.5
40 to 49 years | 19.3 | 47.5
50 to 59 years | 15.5 | 44.8
60 to 69 years | 11.0 | 40.7
70 to 79 years | 5.5 | 30.9
80 to 89 years | 1.6 | 21.7
90 to 99 years | 0.1 | 8.8
100 to 109 years | – | 1.1
Total | 151.1 | 403.7

(a) Because the total population in 2050 is much larger than the 1950 population, comparing percents (relative frequencies) 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 2050.

(b) Make a relative frequency histogram of the 1950 age distribution. 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 relative frequency histogram of the projected age distribution for the year 2050. 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 century between 1950 and 2050?

1.22 (Optional) Sometimes you want to make a histogram from data that are already grouped into classes of unequal width. A report on the recent graduates of a large state university includes the following relative frequency table of the first-year
salaries of last year’s graduates. Salaries are in $1000 units, and it is understood that each class includes its left endpoint but not its right endpoint—for example, a salary of exactly $20,000 belongs in the second class.

<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>Percent</td>
<td>7</td>
<td>14</td>
<td>29</td>
<td>23</td>
<td>13</td>
<td>9</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

The last three classes are wider than the others. An accurate histogram must take this into account. If the base of each bar in the histogram covers a class and the height is the percent of graduates with salaries in that class, the areas of the three rightmost bars will overstate the percent who have salaries in these classes. To make a correct histogram, the area of each bar must be proportional to the percent in that class. Most classes are $5000 wide. A class twice as wide ($10,000) should have a bar half as tall as the percent in that class. This keeps the area proportional to the percent. How should you treat the height of the bar for a class $20,000 wide? Make a correct histogram with the heights of the bars for the last three classes adjusted so that the areas of the bars reflect the percent in each class.

1.23 “Major hurricanes account for just over 20% of the tropical storms and hurricanes that strike the United States but cause more than 80% of the damage.” So say investigators who have shown that major hurricanes (with sustained wind speeds at least 50 meters per second) are tied to ocean temperature and other variables. These variables change slowly, so the high level of hurricane activity that began in 1995 “is likely to persist for an additional 10 to 40 years.” This is bad news for people with beach houses on the Atlantic coast. Here are the counts of major hurricanes for each year between 1944 and 2000. (Stanley B. Goldenberg et al., “The recent increase in Atlantic hurricane activity: causes and implications,” *Science*, 293 (2001), pp. 474–479.)
(a) What is the average number of major hurricanes per year during the period 1944 to 2000?

(b) Make a time plot of the count of major hurricanes. Draw a line across your plot at the average number of hurricanes per year. This helps divide the plot into three periods. Describe the pattern you see.

1.24 Treasury bills are short-term borrowing by the U. S. government. They are important in financial theory because the interest rate for Treasury bills is a “risk-free rate” that says what return investors can get while taking (almost) no risk. More risky investments should in theory offer higher returns in the long run. The table below gives the annual returns on Treasury bills from 1970 to 2000.

(a) Make a time plot of the returns paid by Treasury bills in these years.

(b) Interest rates, like many economic variables, show cycles, clear but irregular up-and-down movements. In which years did the interest rate cycle reach temporary peaks?

<table>
<thead>
<tr>
<th>Year</th>
<th>Count</th>
<th>Year</th>
<th>Count</th>
<th>Year</th>
<th>Count</th>
<th>Year</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1944</td>
<td>3</td>
<td>1956</td>
<td>2</td>
<td>1968</td>
<td>0</td>
<td>1980</td>
<td>2</td>
</tr>
<tr>
<td>1945</td>
<td>2</td>
<td>1957</td>
<td>2</td>
<td>1969</td>
<td>3</td>
<td>1981</td>
<td>3</td>
</tr>
<tr>
<td>1946</td>
<td>1</td>
<td>1958</td>
<td>4</td>
<td>1970</td>
<td>2</td>
<td>1982</td>
<td>1</td>
</tr>
<tr>
<td>1948</td>
<td>4</td>
<td>1960</td>
<td>2</td>
<td>1972</td>
<td>0</td>
<td>1984</td>
<td>1</td>
</tr>
<tr>
<td>1949</td>
<td>3</td>
<td>1961</td>
<td>6</td>
<td>1973</td>
<td>1</td>
<td>1985</td>
<td>3</td>
</tr>
<tr>
<td>1950</td>
<td>7</td>
<td>1962</td>
<td>0</td>
<td>1974</td>
<td>2</td>
<td>1986</td>
<td>0</td>
</tr>
<tr>
<td>1951</td>
<td>2</td>
<td>1963</td>
<td>2</td>
<td>1975</td>
<td>3</td>
<td>1987</td>
<td>1</td>
</tr>
<tr>
<td>1953</td>
<td>3</td>
<td>1965</td>
<td>1</td>
<td>1977</td>
<td>1</td>
<td>1989</td>
<td>2</td>
</tr>
<tr>
<td>1954</td>
<td>2</td>
<td>1966</td>
<td>3</td>
<td>1978</td>
<td>2</td>
<td>1990</td>
<td>1</td>
</tr>
<tr>
<td>1955</td>
<td>5</td>
<td>1967</td>
<td>1</td>
<td>1979</td>
<td>2</td>
<td>1991</td>
<td>2</td>
</tr>
</tbody>
</table>
(c) A time plot may show a consistent trend underneath cycles. When did interest rates reach their overall peak during these years? Has there been a general trend downward since that year?

<table>
<thead>
<tr>
<th>Year</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>6.52</td>
</tr>
<tr>
<td>1971</td>
<td>4.39</td>
</tr>
<tr>
<td>1972</td>
<td>3.84</td>
</tr>
<tr>
<td>1973</td>
<td>6.93</td>
</tr>
<tr>
<td>1974</td>
<td>8.01</td>
</tr>
<tr>
<td>1975</td>
<td>5.80</td>
</tr>
<tr>
<td>1976</td>
<td>5.08</td>
</tr>
<tr>
<td>1977</td>
<td>5.13</td>
</tr>
<tr>
<td>1978</td>
<td>7.19</td>
</tr>
<tr>
<td>1979</td>
<td>10.38</td>
</tr>
<tr>
<td>1980</td>
<td>11.26</td>
</tr>
<tr>
<td>1981</td>
<td>14.72</td>
</tr>
<tr>
<td>1982</td>
<td>10.53</td>
</tr>
<tr>
<td>1983</td>
<td>8.80</td>
</tr>
<tr>
<td>1984</td>
<td>9.84</td>
</tr>
<tr>
<td>1985</td>
<td>7.72</td>
</tr>
<tr>
<td>1986</td>
<td>6.16</td>
</tr>
<tr>
<td>1987</td>
<td>5.47</td>
</tr>
<tr>
<td>1988</td>
<td>6.36</td>
</tr>
<tr>
<td>1989</td>
<td>8.38</td>
</tr>
<tr>
<td>1990</td>
<td>7.84</td>
</tr>
<tr>
<td>1991</td>
<td>5.60</td>
</tr>
<tr>
<td>1992</td>
<td>3.50</td>
</tr>
<tr>
<td>1993</td>
<td>2.90</td>
</tr>
<tr>
<td>1994</td>
<td>3.91</td>
</tr>
<tr>
<td>1995</td>
<td>5.60</td>
</tr>
<tr>
<td>1996</td>
<td>5.20</td>
</tr>
<tr>
<td>1997</td>
<td>5.25</td>
</tr>
<tr>
<td>1998</td>
<td>4.85</td>
</tr>
<tr>
<td>1999</td>
<td>4.69</td>
</tr>
<tr>
<td>2000</td>
<td>5.69</td>
</tr>
</tbody>
</table>

1.25 Time series data often display the effects of changes in policy. Here are data on motor vehicle deaths in the United States. For proper year-to-year comparison, we look at the death rate per 100 million miles driven.

<table>
<thead>
<tr>
<th>Year</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>5.1</td>
</tr>
<tr>
<td>1962</td>
<td>5.1</td>
</tr>
<tr>
<td>1964</td>
<td>5.4</td>
</tr>
<tr>
<td>1966</td>
<td>5.5</td>
</tr>
<tr>
<td>1968</td>
<td>5.2</td>
</tr>
<tr>
<td>1970</td>
<td>4.7</td>
</tr>
<tr>
<td>1972</td>
<td>4.3</td>
</tr>
<tr>
<td>1974</td>
<td>3.5</td>
</tr>
<tr>
<td>1976</td>
<td>3.2</td>
</tr>
<tr>
<td>1978</td>
<td>3.3</td>
</tr>
<tr>
<td>1980</td>
<td>3.3</td>
</tr>
<tr>
<td>1982</td>
<td>2.8</td>
</tr>
<tr>
<td>1984</td>
<td>2.6</td>
</tr>
<tr>
<td>1986</td>
<td>2.5</td>
</tr>
<tr>
<td>1988</td>
<td>2.3</td>
</tr>
<tr>
<td>1990</td>
<td>2.1</td>
</tr>
<tr>
<td>1992</td>
<td>1.7</td>
</tr>
<tr>
<td>1994</td>
<td>1.7</td>
</tr>
<tr>
<td>1996</td>
<td>1.7</td>
</tr>
<tr>
<td>1998</td>
<td>1.6</td>
</tr>
</tbody>
</table>

(a) Make a time plot of these death rates. During these years, safety requirements for motor vehicles became stricter and interstate highways replaced older roads. How does the pattern of your plot reflect these changes?

(b) In 1974 the national speed limit was lowered to 55 miles per hour in an attempt to conserve gasoline after the 1973 Arab-Israeli War. In the mid-1980s most states raised speed limits on interstate highways to 65 miles per hour. Some said that the
lower speed limit saved lives. Is the effect of lower speed limits between 1974 and the mid-1980s visible in your plot?

(c) Does it make sense to make a histogram of these 20 death rates? Explain your answer.

1.26 The impression that a time plot gives depends on the scales you use on the two axes. If you stretch the vertical axis and compress the time axis, change appears to be more rapid. Compressing the vertical axis and stretching the time axis make change appear slower. Make two more time plots of the data in Exercise 1.25, one that makes motor vehicle death rates appear to decrease very rapidly and one that shows only a slow decrease. The moral of this exercise is: pay close attention to the scales when you look at a time plot.

1.27 Babe Ruth was a pitcher for the Boston Red Sox in the years 1914 to 1917. In 1918 and 1919 he played some games as a pitcher and some as an outfielder. From 1920 to 1934 Ruth was an outfielder for the New York Yankees. He ended his career in 1935 with the Boston Braves. The table below gives the number of home runs Ruth hit in each year. Make a time plot and describe its main features.

<table>
<thead>
<tr>
<th>Year</th>
<th>HRs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1914</td>
<td>0</td>
</tr>
<tr>
<td>1915</td>
<td>4</td>
</tr>
<tr>
<td>1916</td>
<td>3</td>
</tr>
<tr>
<td>1917</td>
<td>2</td>
</tr>
<tr>
<td>1918</td>
<td>11</td>
</tr>
<tr>
<td>1919</td>
<td>29</td>
</tr>
<tr>
<td>1920</td>
<td>54</td>
</tr>
<tr>
<td>1921</td>
<td>59</td>
</tr>
<tr>
<td>1922</td>
<td>35</td>
</tr>
<tr>
<td>1923</td>
<td>41</td>
</tr>
<tr>
<td>1924</td>
<td>46</td>
</tr>
<tr>
<td>1925</td>
<td>25</td>
</tr>
<tr>
<td>1926</td>
<td>47</td>
</tr>
<tr>
<td>1927</td>
<td>60</td>
</tr>
<tr>
<td>1928</td>
<td>54</td>
</tr>
<tr>
<td>1929</td>
<td>46</td>
</tr>
<tr>
<td>1930</td>
<td>49</td>
</tr>
<tr>
<td>1931</td>
<td>46</td>
</tr>
<tr>
<td>1932</td>
<td>41</td>
</tr>
<tr>
<td>1933</td>
<td>34</td>
</tr>
<tr>
<td>1934</td>
<td>22</td>
</tr>
<tr>
<td>1935</td>
<td>6</td>
</tr>
</tbody>
</table>

1.28 The following table gives the times (in minutes, rounded to the nearest minute) for the winning man in the Boston Marathon in the years 1959 to 2004:
Display these data in an appropriate graph. Describe the pattern that you see. Have times stopped improving in recent years? If so, when did improvement end?

Section 1.2

1.29 The mean and median salaries paid to major league baseball players in 1993 were $490,000 and $1,160,000. Which of these numbers is the mean, and which is the median? Explain your answer.

1.30 The NASDAQ Composite Index describes the average price of common stock traded over the counter, that is, not on one of the stock exchanges. In 1991, the mean capitalization of the companies in the NASDAQ index was $80 million and the median capitalization was $20 million. (A company’s capitalization is the total market value of its stock.) Explain why the mean capitalization is much higher than the median.
1.31 A college rowing coach tests the 10 members of the women's varsity rowing team on a Stanford Rowing Ergometer (a stationary rowing machine). The variable measured is revolutions of the ergometer's flywheel in a 1-minute session. The data are

\[446 \quad 552 \quad 527 \quad 504 \quad 450 \quad 583 \quad 501 \quad 545 \quad 549 \quad 506\]

(a) Make a stemplot of these data after rounding to two digits. Then find the mean and the median of the original, unrounded ergometer scores. Explain the similarity or difference in these two measures in terms of the symmetry or skewness of the distribution.

(b) The coach used $\bar{x}$ and $s$ to summarize these data. Find the standard deviation $s$. Do you agree that this is a suitable summary?

1.32 Here are the scores on the Survey of Study Habits and Attitudes (SSHA) for 18 first-year college women:

\[154 \quad 109 \quad 137 \quad 115 \quad 152 \quad 140 \quad 154 \quad 178 \quad 101 \]
\[103 \quad 126 \quad 126 \quad 137 \quad 165 \quad 165 \quad 129 \quad 200 \quad 148\]

and for 20 first-year college men:

\[108 \quad 140 \quad 114 \quad 91 \quad 180 \quad 115 \quad 126 \quad 92 \quad 169 \quad 146 \]
\[109 \quad 132 \quad 75 \quad 88 \quad 113 \quad 151 \quad 70 \quad 115 \quad 187 \quad 104\]

(a) Make a back-to-back stemplot of these data, or use your result from Exercise 1.15.

(b) Find the mean $\bar{x}$ and the median $M$ for both sets of SSHA scores. What feature of each distribution explains the fact that $\bar{x} > M$?

(c) Find the five-number summaries for both sets of SSHA scores. Your plot in (a) suggests that there is an outlier among the women's scores. Does the $1.5 \times R$ rule flag this observation? Make side-by-side modified boxplots for the two distributions.

(d) Use your results to write a brief comparison of the two groups. Do women as a group score higher than men? Which of your descriptions (stemplots, boxplots,
numerical measures) show this? Which group of scores is more spread out when we ignore outliers? Which of your descriptions shows this most clearly?

1.33 The SSHA data for women given in the previous exercise contain one high outlier. Calculate the mean $\bar{x}$ and the median $M$ for these data with and without the outlier. How does removing the outlier affect $\bar{x}$? How does it affect $M$? Your results illustrate the greater resistance of the median.

1.34 Exercise 1.19 gives the number of medical doctors per 100,000 people in each state. Your graph of the distribution in Exercise 1.19 shows that the District of Columbia (D.C.) is a high outlier. Because D.C. is a city rather than a state, we will omit it here.

(a) Calculate both the five-number summary and $\bar{x}$ and $s$ for the number of doctors per 100,000 people in the 50 states. Based on your graph, which description do you prefer?
(b) What facts about the distribution can you see in the graph that the numerical summaries don’t reveal? Remember that measures of center and spread are not complete descriptions of a distribution.

1.35 It is usual in the study of investments to use the mean and standard deviation to summarize and compare investment returns. Exercise 1.20 gives the monthly returns on one company’s stock for 82 consecutive months.

(a) 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?
(b) The distribution can be described as “symmetric and unimodal, with one low outlier.” 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? (Returns over longer periods of time, or returns on portfolios containing several investments, tend to follow a normal distribution more closely than these monthly returns do. So use of the mean and standard deviation is better justified for such data.)

1.36 Find the 10th and 90th percentiles of the distribution of doctors per 100,000 population in the states, from Exercise 1.19. Which states are in the top 10%? In the bottom 10%?

1.37 Here are the percents of the popular vote won by the successful candidate in each U.S. presidential election from 1948 to 2000:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent</td>
<td>49.6</td>
<td>55.1</td>
<td>57.4</td>
<td>49.7</td>
<td>61.1</td>
<td>43.4</td>
<td>60.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent</td>
<td>50.1</td>
<td>50.7</td>
<td>58.8</td>
<td>53.9</td>
<td>43.2</td>
<td>49.2</td>
<td>47.9</td>
</tr>
</tbody>
</table>

(a) Make a graph to display the distribution of winners’ percents. What are the main features of this distribution?

(b) What is the median percent of the vote won by the successful candidate in presidential elections?

(c) Call an election a landslide if the winner’s percent falls at or above the third quartile. Which elections were landslides?

1.38 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:
(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?

(b) Which observations are suspected outliers by the $1.5 \times IQR$ rule? Which observations would you call outliers based on the stemplot? (Data from the August 2000 supplement to the Current Population Survey, from the Census Bureau Web site.)

The table below shows Consumer Reports magazine’s laboratory test results for calories and milligrams of sodium (mostly due to salt) in a number of major brands of hot dogs. There are three types: all beef, “meat” (mainly pork and beef, but government regulations allow up to 15% poultry meat), and poultry. Exercises 1.39 to 1.41 analyze these data. (Consumer Reports, June 1986, pp. 366–367.)
<table>
<thead>
<tr>
<th>Calories</th>
<th>Sodium</th>
<th>Calories</th>
<th>Sodium</th>
<th>Calories</th>
<th>Sodium</th>
</tr>
</thead>
<tbody>
<tr>
<td>186</td>
<td>495</td>
<td>173</td>
<td>458</td>
<td>129</td>
<td>430</td>
</tr>
<tr>
<td>181</td>
<td>477</td>
<td>191</td>
<td>506</td>
<td>132</td>
<td>375</td>
</tr>
<tr>
<td>176</td>
<td>425</td>
<td>182</td>
<td>473</td>
<td>102</td>
<td>396</td>
</tr>
<tr>
<td>149</td>
<td>322</td>
<td>190</td>
<td>545</td>
<td>106</td>
<td>383</td>
</tr>
<tr>
<td>184</td>
<td>482</td>
<td>172</td>
<td>496</td>
<td>94</td>
<td>387</td>
</tr>
<tr>
<td>190</td>
<td>587</td>
<td>147</td>
<td>360</td>
<td>102</td>
<td>542</td>
</tr>
<tr>
<td>158</td>
<td>370</td>
<td>146</td>
<td>387</td>
<td>87</td>
<td>359</td>
</tr>
<tr>
<td>139</td>
<td>322</td>
<td>139</td>
<td>386</td>
<td>99</td>
<td>357</td>
</tr>
<tr>
<td>175</td>
<td>479</td>
<td>175</td>
<td>507</td>
<td>170</td>
<td>528</td>
</tr>
<tr>
<td>148</td>
<td>375</td>
<td>136</td>
<td>393</td>
<td>113</td>
<td>513</td>
</tr>
<tr>
<td>152</td>
<td>330</td>
<td>179</td>
<td>405</td>
<td>135</td>
<td>426</td>
</tr>
<tr>
<td>111</td>
<td>300</td>
<td>153</td>
<td>372</td>
<td>142</td>
<td>513</td>
</tr>
<tr>
<td>141</td>
<td>386</td>
<td>107</td>
<td>144</td>
<td>86</td>
<td>358</td>
</tr>
<tr>
<td>153</td>
<td>401</td>
<td>195</td>
<td>511</td>
<td>143</td>
<td>581</td>
</tr>
<tr>
<td>190</td>
<td>645</td>
<td>135</td>
<td>405</td>
<td>152</td>
<td>588</td>
</tr>
<tr>
<td>157</td>
<td>440</td>
<td>140</td>
<td>428</td>
<td>146</td>
<td>522</td>
</tr>
<tr>
<td>131</td>
<td>317</td>
<td>138</td>
<td>339</td>
<td>144</td>
<td>545</td>
</tr>
<tr>
<td>149</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>319</td>
</tr>
<tr>
<td>135</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>298</td>
</tr>
<tr>
<td>132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>253</td>
</tr>
</tbody>
</table>

1.39 Find the five-number summaries of the calorie content of the three types of hot dogs. Then use the $1.5 \times IQR$ rule to check for suspected outliers. Make modified boxplots to compare the three distributions. Write a brief discussion of your findings.
1.40 Make a stemplot of the calorie content of the 17 brands of meat hot dogs. What is the most important feature of the overall pattern of the distribution? Are there any outliers? Note that the five-number summary misses the big feature of this distribution. Routine numerical summaries are never a substitute for looking at the data.

1.41 Use graphs and numerical summaries to compare the sodium content of the three types of hot dogs. Write a summary of your findings suitable for readers who know no statistics. Can we hold down our sodium intake by buying poultry hot dogs?

1.42 Exercise 1.16 presents data on the growth of chicks fed normal corn (the control group) and a new variety with better protein quality (the experimental group). (a) The researchers used $\bar{x}$ and $s$ to summarize the data and as a basis for further statistical analysis. Find these measures for both groups. (b) What kinds of distributions are best summarized by $\bar{x}$ and $s$? Do these distributions seem to fit the criteria?

1.43 The weights in the previous exercise are given in grams. There are 28.35 grams in an ounce. Use the results of part (a) of the previous exercise to find the mean and standard deviation of the weight gains measured in ounces.

1.44 In each of the following settings, give the values of $a$ and $b$ for the linear transformation $x_{\text{new}} = a + bx$ that expresses the change in units of measurement. (a) Change a speed $x$ measured in miles per hour into the metric system value $x_{\text{new}}$ in kilometers per hour. (A kilometer is 0.62 mile.) What is 65 miles per hour in metric units? (b) You are writing a report on the power of car engines. Your sources use horsepower $x$. Reexpress power in watts $x_{\text{new}}$. (One horsepower is 746 watts.) What is the power in watts of a 140-horsepower engine?
1.45 In each of the following settings, give the values of $a$ and $b$ for the linear transformation $x_{\text{new}} = a + bx$ that expresses the change in units of measurement.

(a) You want to restate water temperature $x$ in a swimming pool, measured in degrees Fahrenheit, as the difference $x_{\text{new}}$ between $x$ and the “normal” body temperature of 98.6 degrees.

(b) The recommended daily allowance (RDA) for vitamin C was recently increased to 120 milligrams. You measure milligrams of vitamin C in foods and want to convert your results to percent of RDA.

Section 1.3

1.46 The Environmental Protection Agency requires that the exhaust of each model of motor vehicle be tested for the level of several pollutants. The level of oxides of nitrogen (NOX) in the exhaust of one light truck model was found to vary among individual trucks according to a normal distribution with mean $\mu = 1.45$ grams per mile driven and standard deviation $\sigma = 0.40$ grams per mile. Sketch the density curve of this normal distribution, with the scale of grams per mile marked on the horizontal axis.

1.47 A study of elite distance runners found a mean weight of 63.1 kilograms (kg), with a standard deviation of 4.8 kg. Assuming that the distribution of weights is normal, sketch the density curve of the weight distribution with the horizontal axis marked in kilograms. (Based on M. L. Pollock et al., “Body composition of elite class distance runners,” in P. Milvy (ed.), The Marathon: Physiological, Medical, Epidemiological, and Psychological Studies, New York Academy of Sciences, 1977.)

1.48 Give an interval that contains the middle 95% of NOX levels in the exhaust of trucks using the model described in Exercise 1.46.
1.49 Use the 68–95–99.7 rule to find intervals centered at the mean that will include 68%, 95%, and 99.7% of the weights of the elite runners described in Exercise 1.47.

1.50 Eleanor scores 680 on the mathematics part of the SAT examination. The distribution of SAT scores in a reference population is normal with mean 500 and standard deviation 100. Gerald takes the ACT mathematics test and scores 27. ACT scores are normally distributed with mean 18 and standard deviation 6. Find the z-scores for both students. Assuming that both tests measure the same kind of ability, who has the higher score?

1.51 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 decades. 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. (Stephen Jay Gould, “Entropic homogeneity isn’t why no one hits .400 anymore,” Discover, August 1986, pp. 60–66.)

<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.

1.52 It is possible to score higher than 800 on the SAT, but scores above 800 are reported as 800. (That is, a student can get a reported score of 800 without a perfect paper.) In 2000, the scores of men on the math part of the SAT approximately
followed a normal distribution with mean 533 and standard deviation 115. What percent of scores were above 800 (and so reported as 800)?

1.53 Scores on the Wechsler Adult Intelligence Scale for the 20 to 34 age group are approximately normally distributed with mean 110 and standard deviation 25. Scores for the 60 to 64 age group are approximately normally distributed with mean 90 and standard deviation 25.

Sarah, who is 30, scores 135 on this test. Sarah’s mother, who is 60, also takes the test and scores 120. Who scored higher relative to her age group, Sarah or her mother? Who has the higher absolute level of the variable measured by the test? At what percentile of their age groups are Sarah and her mother? (That is, what percent of the age group has lower scores?)

1.54 The Graduate Record Examinations (GRE) are widely used to help predict the performance of applicants to graduate schools. The range of possible scores on a GRE is 200 to 900. The psychology department at a university finds that the scores of its applicants on the quantitative GRE are approximately normal with mean \( \mu = 544 \) and standard deviation \( \sigma = 103 \). Find the proportion of applicants whose score \( X \) satisfies each of the following conditions:

(a) \( X > 700 \)

(b) \( X < 500 \)

(c) \( 500 < X < 800 \)

1.55 Using either Table A or your calculator or software, find the proportion of observations from a standard normal distribution that satisfies each of the following statements. In each case, sketch a standard normal curve and shade the area under the curve that is the answer to the question.

(a) \( Z < 2.85 \)

(b) \( Z > 2.85 \)
(c) $Z > -1.66$
(d) $-1.66 < Z < 2.85$

1.56 Using either Table A or your calculator or software, find the proportion of observations on a standard normal distribution for each of the following events. In each case, sketch a standard normal curve with the area representing the proportion shaded.
(a) $Z \leq -2.25$
(b) $Z \geq -2.25$
(c) $Z > 1.77$
(d) $-2.25 < Z < 1.77$

1.57 Find the value $z$ of a standard normal variable $Z$ that satisfies each of the following conditions. (If you use Table A, report the value of $z$ that comes closest to satisfying the condition.) In each case, sketch a standard normal curve with your value of $z$ marked on the axis.
(a) The point $z$ with 25% of the observations falling below it.
(b) The point $z$ with 40% of the observations falling above it.

1.58 The variable $Z$ has a standard normal distribution.
(a) Find the number $z$ such that the event $Z < z$ has proportion 0.8.
(b) Find the number $z$ such that the event $Z > z$ has proportion 0.35.

1.59 The scores of a reference population on the Wechsler Intelligence Scale for Children (WISC) are normally distributed with $\mu = 100$ and $\sigma = 15$.
(a) What percent of this population have WISC scores below 100?
(b) Below 80?
(c) Above 140?
(d) Between 100 and 120?
1.60 The distribution of scores on the WISC is described in the previous exercise. What score will place a child in the top 5% of the population? In the top 1%?

1.61 In 2003, scores on the math part of the SAT approximately followed a normal distribution with mean 519 and standard deviation 115.
(a) What proportion of students scored above 500?
(b) What proportion scored between 400 and 600?

1.62 Some companies “grade on a bell curve” to compare the performance of their managers and professional workers. This forces the use of some low performance ratings, so that not all workers are graded “above average.” Until the threat of lawsuits forced a change, Ford Motor Company’s “performance management process” assigned 10% A grades, 80% B grades, and 10% C grades to the company’s 18,000 managers. It isn’t clear that the “bell curve” of ratings is really a normal distribution. Nonetheless, suppose that Ford’s performance scores are normally distributed. One year, managers with scores less than 25 received C’s and those with scores above 475 received A’s. What are the mean and standard deviation of the scores?

1.63 The Survey of Study Habits and Attitudes (SSHA) is a common psychological instrument to evaluate the attitudes of students. The SSHA is used for subjects from seventh grade through college. Different groups have different distributions. To prepare to use the SSHA to evaluate future teachers, researchers gave the test to 238 college juniors majoring in elementary education. Their scores were roughly normal with mean 114 and standard deviation 30. Take this as the distribution of SSHA scores in the population of future elementary school teachers.

A study of Native American education students in Canada found that this relatively disadvantaged group had mean SSHA score 99. (Graham Hurlbut, Eldon Glade, and John McLaughlin, “Teaching attitudes and study attitudes of Indian
education students,” *Journal of American Indian Education*, 29, No. 3, (1990), pp. 12–18.) What percentile of the overall distribution is this?

1.64 How high a score on the SSHA test of the previous exercise (mean 114, standard deviation 30) must an elementary education student obtain to be among the highest-scoring 30% of the population? What scores make up the lowest 30%?

1.65 The median of any normal distribution is the same as its mean. We can use normal calculations to find the quartiles and related descriptive measures for normal distributions.

(a) What is the area under the standard normal curve to the left of the first quartile? Use this to find the value of the first quartile for a standard normal distribution. Find the third quartile similarly.

(b) Your work in (a) gives the z-scores for the quartiles of any normal distribution. Scores on the Wechsler Intelligence Scale for Children (WISC) are normally distributed with mean 100 and standard deviation 15. What are the quartiles of WISC scores?

(c) What is the value of the $IQR$ for the standard normal distribution?

(d) What percent of the observations in the standard normal distribution are suspected outliers according to the $1.5 \times IQR$ rule? (This percent is the same for any normal distribution.)

1.66 The distribution of Internet access costs in Exercise 1.38 has a compact center with a long tail on either side. Make a normal quantile plot of these data. Explain carefully why the pattern of this plot is typical of a “long-tailed” distribution.

1.67 Is the distribution of monthly returns on Philip Morris stock approximately normal with the exception of possible outliers? Make a normal quantile plot of the data in Exercise 1.20, and report your conclusions.
1.68 Exercise 1.16 presents data on the weight gains of chicks fed two types of corn. The researchers use $\bar{x}$ and $s$ to summarize each of the two distributions. Make a normal quantile plot for each group and report your findings. Is use of $\bar{x}$ and $s$ justified?

Chapter 1 Review Exercises

1.69 Here is a stemplot of the percents of residents aged 25 to 34 in each of the 50 states. The stems are whole percents and the leaves are tenths of a percent.

```
10 | 9
11 | 0
12 | 134677889
13 | 01245566789999
14 | 112234445789
15 | 24478999
```

(a) Montana and Wyoming have the smallest percents of young adults, perhaps because they lack job opportunities. What are the percents for these two states?

(b) Ignoring Montana and Wyoming, describe the shape, center, and spread of this distribution.

1.70 Stemplots help you find the five-number summary because they arrange the observations in increasing order. The previous exercise gives a stemplot of the percent of residents aged 25 to 34 in each of the 50 states.

(a) Find the five-number summary of this distribution.

(b) Does the $1.5 \times IQR$ criterion flag Montana and Wyoming as suspected outliers?

(c) How much does the median change if you omit Montana and Wyoming?

1.71 The American Housing Survey provides data on all housing units in the United States—houses, apartments, mobile homes, and so on. Here are the years in which a random sample of 100 housing units were built. The survey does not produce exact
dates for years before 1990. Years before 1920 are given as 1919. Dates between 1920 and 1970 are given in ten-year blocks, so that a unit built in 1956 appears as 1950. Dates between 1970 and 1990 are given in five-year blocks, so that 1987 appears as 1985.


(a) Make a histogram of these dates, using classes 10 years wide beginning with 1910 to 1919. The first class will contain all housing units built before 1920. In which decades after 1920 were most housing units that still exist built?

(b) Give the five-number summary of these data. Write a brief warning on how to interpret your results. For example, what does the fact that the median is 1960 tell us about the age of American housing?

1.72 The table below shows the salaries paid to the members of the New York Yankees baseball team as of opening day of the 2001 season. Display this distribution with a graph and describe its main features. Find the mean and median salary and explain how the pattern of the distribution explains the relationship between these two measures of center. Find the standard deviation and the quartiles. Do you prefer the five-number summary or $\bar{x}$ and $s$ as a quick description of this distribution?
<table>
<thead>
<tr>
<th>Player</th>
<th>Salary</th>
<th>Player</th>
<th>Salary</th>
<th>Player</th>
<th>Salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jeter</td>
<td>$12,600,000</td>
<td>Posada</td>
<td>$4,050,000</td>
<td>Spencer</td>
<td>$320,000</td>
</tr>
<tr>
<td>B Williams</td>
<td>12,357,143</td>
<td>Stanton</td>
<td>2,450,000</td>
<td>T Williams</td>
<td>320,000</td>
</tr>
<tr>
<td>Clemens</td>
<td>10,300,000</td>
<td>Hernandez</td>
<td>2,050,000</td>
<td>Almanzar</td>
<td>270,000</td>
</tr>
<tr>
<td>Mussina</td>
<td>10,000,000</td>
<td>Watson</td>
<td>1,700,000</td>
<td>Bellinger</td>
<td>230,000</td>
</tr>
<tr>
<td>Rivera</td>
<td>9,150,000</td>
<td>Mendoza</td>
<td>1,600,000</td>
<td>Einertson</td>
<td>206,000</td>
</tr>
<tr>
<td>Justice</td>
<td>7,000,000</td>
<td>Oliver</td>
<td>1,100,000</td>
<td>Choate</td>
<td>204,750</td>
</tr>
<tr>
<td>Pettitte</td>
<td>7,000,000</td>
<td>Rodriguez</td>
<td>850,000</td>
<td>Coleman</td>
<td>204,000</td>
</tr>
<tr>
<td>O’Neill</td>
<td>6,500,000</td>
<td>Soriano</td>
<td>630,000</td>
<td>Jimenez</td>
<td>200,000</td>
</tr>
<tr>
<td>Knoblauch</td>
<td>6,000,000</td>
<td>Sojo</td>
<td>500,000</td>
<td>Parker</td>
<td>200,000</td>
</tr>
<tr>
<td>Martinez</td>
<td>6,000,000</td>
<td>Boehringer</td>
<td>350,000</td>
<td>Seabol</td>
<td>200,000</td>
</tr>
<tr>
<td>Brosius</td>
<td>5,250,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.73 At the time the salaries in the previous exercise were announced, one dollar was worth 1.72 Swiss francs. Answer these questions without doing any calculations in addition to those you did in the previous exercise.

(a) What transformation converts a salary in dollars into the same salary in Swiss francs?

(b) What are the mean, median, and quartiles of the distribution in francs?

(c) What are the standard deviation and interquartile range of the distribution in francs?

1.74 The Internal Revenue Service reports that in 1998 about 124 million individual income tax returns showed adjusted gross income (AGI) greater than 0. The mean and median AGI on these tax returns were $25,491 and $44,186. Which of these numbers is the mean and which is the median? How do you know?

1.75 The Bureau of Justice Statistics says that in 1999, 51% of homicides were committed with handguns, 14% with other firearms, 13% with knives, and 6% with
blunt objects. Make a graph to display these data. Do you need an “other methods” category?

1.76 You are planning a sample survey of households in California. You decide to select households separately within each county and to choose more households from the more populous counties. To aid in the planning, the table below gives the populations of California counties from the 2000 census. Examine the distribution of county populations both graphically and numerically, using whatever tools are most suitable. Write a brief description of the main features of this distribution. Sample surveys often select households from all of the most populous counties but from only some of the less populous. How would you divide California counties into three groups according to population, with the intent of including all of the first group, half of the second, and a smaller fraction of the third in your survey?
1.77 The Chapin Social Insight Test evaluates how accurately the subject appraises other people. In the reference population used to develop the test, scores are approximately normally distributed with mean 25 and standard deviation 5. The range of possible scores is 0 to 41.

(a) What proportion of the population has scores below 20 on the Chapin test?
(b) What proportion has scores below 10?
(c) How high a score must you have in order to be in the top quarter of the population in social insight?

1.78 The Chapin Social Insight Test described in the previous exercise has a mean of 25 and a standard deviation of 5. You want to rescale the test using a linear transformation so that the mean is 100 and the standard deviation is 20. Let $x$ denote the score in the original scale and $x_{\text{new}}$ be the transformed score.

(a) Find the linear transformation required. That is, find the values of $a$ and $b$ in the equation $x_{\text{new}} = a + bx$.

(b) Give the rescaled score for someone who scores 30 in the original scale.

(c) What are the quartiles of the rescaled scores?

1.79 The Florida State University Seminoles have been among the more successful teams in college football. The table below gives the weights in pounds and positions of the players on the 2000–2001 football team, which was defeated in the national title game by the University of Oklahoma. The positions are quarterback (QB), running back (RB), offensive line (OL), receiver (R), tight end (TE), kicker (K), defensive back (DB), linebacker (LB), and defensive line (DL). (From the Florida State University athletics Web site, seminoles.fansonly.com.)

(a) Make side-by-side modified boxplots of the weights for running backs, receivers, offensive linemen, defensive linemen, linebackers, and defensive backs.

(b) Briefly compare the weight distributions. Which position has the heaviest players overall? Which has the lightest?

(c) Are any individual players outliers within their position?
<table>
<thead>
<tr>
<th>QB 235</th>
<th>QB 220</th>
<th>QB 215</th>
<th>QB 228</th>
<th>K 175</th>
<th>K 205</th>
</tr>
</thead>
<tbody>
<tr>
<td>K 220</td>
<td>K 185</td>
<td>RB 200</td>
<td>RB 188</td>
<td>RB 190</td>
<td>RB 215</td>
</tr>
<tr>
<td>RB 190</td>
<td>RB 225</td>
<td>RB 225</td>
<td>RB 240</td>
<td>RB 237</td>
<td>R 205</td>
</tr>
<tr>
<td>R 185</td>
<td>R 185</td>
<td>R 190</td>
<td>R 195</td>
<td>R 201</td>
<td>R 190</td>
</tr>
<tr>
<td>R 195</td>
<td>R 180</td>
<td>OL 291</td>
<td>OL 280</td>
<td>OL 300</td>
<td>OL 320</td>
</tr>
<tr>
<td>OL 325</td>
<td>OL 285</td>
<td>OL 305</td>
<td>OL 305</td>
<td>OL 290</td>
<td>OL 310</td>
</tr>
<tr>
<td>OL 315</td>
<td>OL 285</td>
<td>OL 290</td>
<td>OL 325</td>
<td>OL 310</td>
<td>OL 256</td>
</tr>
<tr>
<td>OL 305</td>
<td>OL 300</td>
<td>DB 170</td>
<td>DB 207</td>
<td>DB 185</td>
<td>DB 175</td>
</tr>
<tr>
<td>DB 180</td>
<td>DB 190</td>
<td>DB 210</td>
<td>DB 200</td>
<td>DB 180</td>
<td>DB 195</td>
</tr>
<tr>
<td>DB 185</td>
<td>DB 170</td>
<td>DB 180</td>
<td>DB 190</td>
<td>DB 190</td>
<td>LB 220</td>
</tr>
<tr>
<td>LB 212</td>
<td>LB 233</td>
<td>LB 190</td>
<td>LB 195</td>
<td>LB 215</td>
<td>LB 220</td>
</tr>
<tr>
<td>DL 240</td>
<td>DL 275</td>
<td>DL 255</td>
<td>DL 285</td>
<td>DL 245</td>
<td>DL 270</td>
</tr>
<tr>
<td>DL 250</td>
<td>DL 250</td>
<td>TE 255</td>
<td>TE 245</td>
<td>TE 260</td>
<td>TE 245</td>
</tr>
</tbody>
</table>

### Chapter 2

#### Section 2.1

2.1 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?

2.2 Here are the golf scores of 12 members of a college women’s golf team in two rounds of tournament play. (A golf score is the number of strokes required to complete the course, so that low scores are better.)
Section 2.1

<table>
<thead>
<tr>
<th>Player</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>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round 1</td>
<td>89</td>
<td>90</td>
<td>87</td>
<td>95</td>
<td>86</td>
<td>81</td>
<td>102</td>
<td>105</td>
<td>83</td>
<td>88</td>
<td>91</td>
<td>79</td>
</tr>
<tr>
<td>Round 2</td>
<td>94</td>
<td>85</td>
<td>89</td>
<td>89</td>
<td>81</td>
<td>76</td>
<td>107</td>
<td>89</td>
<td>87</td>
<td>91</td>
<td>88</td>
<td>80</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot of the data, taking the first-round score as the explanatory variable.

(b) Is there an association between the two scores? If so, is it positive or negative? Explain why you would expect scores in two rounds of a tournament to have an association like that you observed.

(c) The plot shows one outlier. Circle it. The outlier may occur because a good golfer had an unusually bad round or because a weaker golfer had an unusually good round. Can you tell from the data given whether the outlier is from a good player or from a poor player? Explain your answer.

2.3 There is some evidence that drinking moderate amounts of wine helps prevent heart attacks. Here are 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. (M. H. Criqui, University of California, San Diego, reported in the New York Times, December 28, 1994.)
Chapter 2 Exercises

(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 language 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?

2.4 The National Assessment of Educational Progress (NAEP) assesses what students know in several subject areas based on large representative samples. The table below reports some findings of the NAEP year 2000 Mathematics Assessment for fourth graders in the 40 states that participated. For each state we give the mean NAEP math score (out of 500) and also the percent of students who were at least “proficient” in the sense of being able to use math skills to solve real-world problems. Nationally, about 25% of students are “proficient” by NAEP’s standards. We expect that average performance and percent of proficient performers will be strongly related.
<table>
<thead>
<tr>
<th>State</th>
<th>Mean NAEP score</th>
<th>Percent proficient</th>
<th>Percent poverty</th>
<th>State</th>
<th>Mean NAEP score</th>
<th>Percent proficient</th>
<th>Percent poverty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>218</td>
<td>14</td>
<td>21.8</td>
<td>Missouri</td>
<td>229</td>
<td>24</td>
<td>14.4</td>
</tr>
<tr>
<td>Arizona</td>
<td>219</td>
<td>17</td>
<td>23.6</td>
<td>Montana</td>
<td>230</td>
<td>25</td>
<td>21.2</td>
</tr>
<tr>
<td>Arkansas</td>
<td>217</td>
<td>14</td>
<td>13.1</td>
<td>Nebraska</td>
<td>226</td>
<td>24</td>
<td>14.8</td>
</tr>
<tr>
<td>California</td>
<td>214</td>
<td>15</td>
<td>22.3</td>
<td>Nevada</td>
<td>220</td>
<td>16</td>
<td>12.8</td>
</tr>
<tr>
<td>Connecticut</td>
<td>234</td>
<td>32</td>
<td>13.4</td>
<td>New Mexico</td>
<td>214</td>
<td>12</td>
<td>23.5</td>
</tr>
<tr>
<td>Georgia</td>
<td>220</td>
<td>18</td>
<td>24.7</td>
<td>New York</td>
<td>227</td>
<td>22</td>
<td>28.9</td>
</tr>
<tr>
<td>Hawaii</td>
<td>216</td>
<td>14</td>
<td>14.5</td>
<td>North Carolina</td>
<td>232</td>
<td>28</td>
<td>21.3</td>
</tr>
<tr>
<td>Idaho</td>
<td>227</td>
<td>21</td>
<td>17.4</td>
<td>North Dakota</td>
<td>231</td>
<td>25</td>
<td>17.2</td>
</tr>
<tr>
<td>Illinois</td>
<td>225</td>
<td>22</td>
<td>12.2</td>
<td>Ohio</td>
<td>231</td>
<td>26</td>
<td>16.0</td>
</tr>
<tr>
<td>Indiana</td>
<td>234</td>
<td>31</td>
<td>12.6</td>
<td>Oklahoma</td>
<td>225</td>
<td>17</td>
<td>19.9</td>
</tr>
<tr>
<td>Iowa</td>
<td>233</td>
<td>28</td>
<td>14.2</td>
<td>Oregon</td>
<td>227</td>
<td>24</td>
<td>19.4</td>
</tr>
<tr>
<td>Kansas</td>
<td>232</td>
<td>30</td>
<td>13.3</td>
<td>Rhode Island</td>
<td>225</td>
<td>23</td>
<td>20.5</td>
</tr>
<tr>
<td>Kentucky</td>
<td>221</td>
<td>17</td>
<td>16.7</td>
<td>South Carolina</td>
<td>220</td>
<td>18</td>
<td>17.6</td>
</tr>
<tr>
<td>Louisiana</td>
<td>218</td>
<td>14</td>
<td>29.8</td>
<td>Tennessee</td>
<td>220</td>
<td>13</td>
<td>14.5</td>
</tr>
<tr>
<td>Maine</td>
<td>231</td>
<td>24</td>
<td>12.0</td>
<td>Texas</td>
<td>233</td>
<td>27</td>
<td>20.1</td>
</tr>
<tr>
<td>Maryland</td>
<td>222</td>
<td>22</td>
<td>8.1</td>
<td>Utah</td>
<td>227</td>
<td>24</td>
<td>11.8</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>235</td>
<td>33</td>
<td>15.0</td>
<td>Vermont</td>
<td>232</td>
<td>30</td>
<td>12.2</td>
</tr>
<tr>
<td>Michigan</td>
<td>231</td>
<td>29</td>
<td>14.8</td>
<td>Virginia</td>
<td>230</td>
<td>25</td>
<td>7.9</td>
</tr>
<tr>
<td>Minnesota</td>
<td>235</td>
<td>34</td>
<td>12.6</td>
<td>West Virginia</td>
<td>225</td>
<td>18</td>
<td>25.7</td>
</tr>
<tr>
<td>Mississippi</td>
<td>211</td>
<td>9</td>
<td>19.3</td>
<td>Wyoming</td>
<td>229</td>
<td>25</td>
<td>13.0</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot, using mean NAEP score as the explanatory variable. Notice that there are several pairs of states with identical values. Use a different symbol for points that represent two states.

(b) Describe the form, direction, and strength of the relationship.

(c) Circle your home state’s point in the scatterplot. Although there are no clear outliers, there are some points that you may consider interesting, perhaps because they are on the edge of the pattern. Choose one such point: which state is this, and in what way is it interesting?
2.5 Water flowing across farmland washes away soil. Researchers released water across a test bed at different flow rates and measured the amount of soil washed away. The following table gives the flow (in liters per second) and the weight (in kilograms) of eroded soil. (G. R. Foster, W. R. Ostercamp, and L. J. Lane, “Effect of discharge rate on rill erosion,” paper presented at the 1982 Winter Meeting of the American Society of Agricultural Engineers.)

<table>
<thead>
<tr>
<th>Flow rate</th>
<th>0.31</th>
<th>0.85</th>
<th>1.26</th>
<th>2.47</th>
<th>3.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eroded soil</td>
<td>0.82</td>
<td>1.95</td>
<td>2.18</td>
<td>3.01</td>
<td>6.07</td>
</tr>
</tbody>
</table>

(a) Plot the data. Which is the explanatory variable?

(b) Describe the pattern that you see. Would it be reasonable to describe the overall pattern by a straight line? Is the association positive or negative?

2.6 Here are data on a group of people who contracted botulism, a form of food poisoning that can be fatal. The variables recorded are the person’s age in years, the incubation period (the time in hours between eating the infected food and the first signs of illness), and whether the person survived (S) or died (D). (Modified from data provided by Dana Quade, University of North Carolina.)

<table>
<thead>
<tr>
<th>Person</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>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>29</td>
<td>39</td>
<td>44</td>
<td>37</td>
<td>42</td>
<td>17</td>
<td>38</td>
<td>43</td>
<td>51</td>
</tr>
<tr>
<td>Incubation</td>
<td>13</td>
<td>46</td>
<td>43</td>
<td>34</td>
<td>20</td>
<td>20</td>
<td>18</td>
<td>72</td>
<td>19</td>
</tr>
<tr>
<td>Outcome</td>
<td>D</td>
<td>S</td>
<td>S</td>
<td>D</td>
<td>S</td>
<td>D</td>
<td>S</td>
<td>D</td>
<td>D</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Person</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>30</td>
<td>32</td>
<td>59</td>
<td>33</td>
<td>31</td>
<td>32</td>
<td>32</td>
<td>36</td>
<td>50</td>
</tr>
<tr>
<td>Incubation</td>
<td>36</td>
<td>48</td>
<td>44</td>
<td>21</td>
<td>32</td>
<td>86</td>
<td>48</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>Outcome</td>
<td>D</td>
<td>D</td>
<td>S</td>
<td>D</td>
<td>D</td>
<td>S</td>
<td>D</td>
<td>S</td>
<td>D</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot of incubation period against age, using different symbols for people who survived and those who died.

(b) Is there an overall relationship between age and incubation period? If so, describe
it.

(c) More important, is there a relationship between either age or incubation period and whether the victim survived? Describe any relations that seem important here.

(d) Are there any unusual observations that may require individual investigation?

2.7 The presence of harmful insects in farm fields is detected by erecting boards covered with a sticky material and examining the insects trapped on the boards. Some colors are more attractive to insects than others. In an experiment aimed at determining the best color for attracting cereal leaf beetles, six boards of each of four colors were placed in a field of oats in July. The table below gives data on the number of cereal leaf beetles trapped. (M. C. Wilson and R. E. Shade, “Relative attractiveness of various luminescent colors to the cereal leaf beetle and the meadow spittlebug,” Journal of Economic Entomology, 60 (1967), pp. 578–580.)

<table>
<thead>
<tr>
<th>Board color</th>
<th>Insects trapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lemon yellow</td>
<td>45 59 48 46 38 47</td>
</tr>
<tr>
<td>White</td>
<td>21 12 14 17 13 17</td>
</tr>
<tr>
<td>Green</td>
<td>37 32 15 25 39 41</td>
</tr>
<tr>
<td>Blue</td>
<td>16 11 20 21 14  7</td>
</tr>
</tbody>
</table>

(a) Make a plot of the counts of insects trapped against board color (space the four colors equally on the horizontal axis). Compute the mean count for each color, add the means to your plot, and connect the means with line segments.

(b) Based on the data, state your conclusions about the attractiveness of these colors to the beetles.

(c) Does it make sense to speak of a positive or negative association between board color and insect count?
Section 2.2

2.8 *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. 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: (Marilyn A. Houck et al., “Allometric scaling in the earliest fossil bird, *Archaeopteryx lithographica,*” *Science*, 247 (1990), pp. 195–198.)

<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>

(a) Make a scatterplot with femur length on the horizontal axis. There is a strong positive linear relationship.

(b) Find the correlation $r$ step-by-step. That is, find the mean and standard deviation of the femur lengths and of the humerus lengths. Then find the five standardized values for each variable and use the formula for $r$.

(b) Now enter these data into your calculator or software and use the correlation function to find $r$. Check that you get the same result as in (a).

2.9 Here are the golf scores of 11 members of a women’s golf team in two rounds of college tournament play.

<table>
<thead>
<tr>
<th>Player</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>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round 1</td>
<td>89</td>
<td>90</td>
<td>87</td>
<td>95</td>
<td>86</td>
<td>81</td>
<td>105</td>
<td>83</td>
<td>88</td>
<td>91</td>
<td>79</td>
</tr>
<tr>
<td>Round 2</td>
<td>94</td>
<td>85</td>
<td>89</td>
<td>89</td>
<td>81</td>
<td>76</td>
<td>89</td>
<td>87</td>
<td>91</td>
<td>88</td>
<td>80</td>
</tr>
</tbody>
</table>

If you did not make a scatterplot in Exercise 2.6, do so now. Find the correlation between the Round 1 and Round 2 scores. Remove Player 7’s scores and find the correlation for the remaining 10 players. Explain carefully why removing this single case substantially increases the correlation.
2.10 Changing the units of measurement can dramatically alter the appearance of a scatterplot. Return to the fossil data from Exercise 2.8. The measurements are in centimeters. Suppose a deranged 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.

2.11 A mutual-fund company’s newsletter says, “A well-diversified portfolio includes assets with low correlations.” The newsletter includes a table of correlations between the annual returns on various classes of investments. For example, the correlation between municipal bonds and large-cap stocks is 0.50 and the correlation between municipal bonds and small-cap stocks is 0.21.

(a) Rachel invests heavily in municipal bonds. She wants to diversify by adding an investment whose returns do not closely follow the returns on her bonds. Should she choose large-cap stocks or small-cap stocks for this purpose? Explain your answer.

(b) If Rachel wants an investment that tends to increase when the return on her bonds drops, what kind of correlation should she look for?

2.12 Many mutual funds compare their performance with that of a benchmark, an index of the returns on all securities of the kind the fund buys. The Vanguard International Growth Fund, for example, takes as its benchmark the Morgan Stanley EAFE (Europe, Australasia, Far East) index of overseas stock market performance. Here are the percent returns for the fund and for the EAFE from 1982 (the first
full year of the fund’s existence) to 2000. (From the performance data for the fund presented at the Vanguard Group Web site, personal.vanguard.com.)

<table>
<thead>
<tr>
<th>Year</th>
<th>Fund</th>
<th>EAFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1982</td>
<td>5.27</td>
<td>−0.86</td>
</tr>
<tr>
<td>1983</td>
<td>43.08</td>
<td>24.61</td>
</tr>
<tr>
<td>1984</td>
<td>−1.02</td>
<td>7.86</td>
</tr>
<tr>
<td>1985</td>
<td>56.94</td>
<td>56.72</td>
</tr>
<tr>
<td>1986</td>
<td>56.71</td>
<td>69.94</td>
</tr>
<tr>
<td>1987</td>
<td>12.48</td>
<td>24.93</td>
</tr>
<tr>
<td>1988</td>
<td>11.61</td>
<td>28.59</td>
</tr>
<tr>
<td>1989</td>
<td>24.76</td>
<td>10.80</td>
</tr>
<tr>
<td>1990</td>
<td>−12.05</td>
<td>−23.20</td>
</tr>
<tr>
<td>1991</td>
<td>4.74</td>
<td>12.50</td>
</tr>
<tr>
<td>1992</td>
<td>−5.79</td>
<td>−11.85</td>
</tr>
<tr>
<td>1993</td>
<td>44.74</td>
<td>32.94</td>
</tr>
<tr>
<td>1994</td>
<td>0.76</td>
<td>8.06</td>
</tr>
<tr>
<td>1995</td>
<td>14.89</td>
<td>11.55</td>
</tr>
<tr>
<td>1996</td>
<td>14.65</td>
<td>6.36</td>
</tr>
<tr>
<td>1997</td>
<td>4.12</td>
<td>2.06</td>
</tr>
<tr>
<td>1998</td>
<td>16.93</td>
<td>20.33</td>
</tr>
<tr>
<td>1999</td>
<td>26.34</td>
<td>27.30</td>
</tr>
<tr>
<td>2000</td>
<td>−8.60</td>
<td>−13.96</td>
</tr>
<tr>
<td>2001</td>
<td>−18.92</td>
<td>−21.44</td>
</tr>
</tbody>
</table>

Make a scatterplot suitable for predicting fund returns from EAFE returns. Is there a clear straight-line pattern? How strong is this pattern? (Give a numerical measure.) Are there any extreme outliers from the straight-line pattern?

**Section 2.3**

2.13 (Review of straight lines) Fred keeps his savings in his mattress. He begins 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. (Hint: Choose two values of \( x \), such as 0 and 10. Find 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 adds $200 instead of $100 each year to his initial $500, what is the equation that describes his savings after $x$ years?

2.14 (Review of straight lines) Sound travels at a speed of 1500 meters per second in sea water. You dive into the sea from your yacht. Give an equation for the distance $y$ at which a shark can hear your splash in terms of the number of seconds $x$ since you hit the water.

2.15 (Review of straight lines) During the period after birth, a male white rat gains 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 g in a pound. To help you assess the result, note that a large cat weighs about 10 pounds.)

2.16 (Review of straight lines) A cellular telephone company offers two plans. Plan A charges $20 a month for up to 75 minutes of air time and $0.45 per minute above 75 minutes. Plan B charges $30 a month for up to 250 minutes and $0.40 per minute above 250 minutes.

(a) Draw a graph of the Plan A charge against minutes used from 0 to 250 minutes.

(b) How many minutes a month must the user talk in order for Plan B to be less expensive than Plan A?

2.17 Concrete road pavement gains strength over time as it cures. Highway builders use regression lines to predict the strength after 28 days (when curing is complete) from measurements made after 7 days. Let $x$ be strength after 7 days (in pounds per
square inch) and $y$ the strength after 28 days. One set of data gives this least-squares regression line:

$$\hat{y} = 1389 + 0.96x$$

(a) Draw a graph of this line, with $x$ running from 3000 to 4000 pounds per square inch.

(b) Explain what the slope $b = 0.96$ in this equation says about how concrete gains strength as it cures.

(c) A test of some new pavement after 7 days shows that its strength is 3300 pounds per square inch. Use the equation of the regression line to predict the strength of this pavement after 28 days. Also draw the “up and over” lines from $x = 3300$ on your graph, as in IPS Figure 2.12.

2.18 Researchers studying acid rain measured the acidity of precipitation in an isolated wilderness area in Colorado for 150 consecutive weeks. The acidity of a solution is measured by pH, with lower pH values indicating that the solution is more acid. The acid rain researchers observed a linear pattern over time. They reported that the least-squares line

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


(a) Draw a graph of this line. Note that the linear change is decreasing rather than increasing.

(b) According to the fitted 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 fitted line? Explain clearly what this slope says about the change in the pH of the precipitation in this wilderness area.
2.19 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 1990.

<table>
<thead>
<tr>
<th>Year</th>
<th>Boats (thousands)</th>
<th>Manatees killed</th>
<th>Year</th>
<th>Boats (thousands)</th>
<th>Manatees killed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>447</td>
<td>13</td>
<td>1984</td>
<td>559</td>
<td>34</td>
</tr>
<tr>
<td>1978</td>
<td>460</td>
<td>21</td>
<td>1985</td>
<td>585</td>
<td>33</td>
</tr>
<tr>
<td>1979</td>
<td>481</td>
<td>24</td>
<td>1986</td>
<td>614</td>
<td>33</td>
</tr>
<tr>
<td>1980</td>
<td>498</td>
<td>16</td>
<td>1987</td>
<td>645</td>
<td>39</td>
</tr>
<tr>
<td>1981</td>
<td>513</td>
<td>24</td>
<td>1988</td>
<td>675</td>
<td>43</td>
</tr>
<tr>
<td>1982</td>
<td>512</td>
<td>20</td>
<td>1989</td>
<td>711</td>
<td>50</td>
</tr>
<tr>
<td>1983</td>
<td>526</td>
<td>15</td>
<td>1990</td>
<td>719</td>
<td>47</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot of these data. Describe the form and direction of the relationship.

(b) Find the correlation. What fraction of the variation in manatee deaths can be explained by the number of boats registered? Does it appear that the number of manatees killed can be predicted accurately from power boat registrations?

(c) Find the least-squares regression line. Predict the number of manatees that will be killed by boats in a year when 716,000 powerboats are registered.

(d) Suppose that in some far future year 2 million powerboats are registered in Florida. Use the regression line to predict manatees killed. Explain why this prediction is very unreliable.

(e) Here are four more years of manatee data, in the same form as in Table S3:

<table>
<thead>
<tr>
<th>Year</th>
<th>Boats (thousands)</th>
<th>Manatees killed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>716</td>
<td>53</td>
</tr>
<tr>
<td>1992</td>
<td>716</td>
<td>38</td>
</tr>
<tr>
<td>1993</td>
<td>716</td>
<td>35</td>
</tr>
<tr>
<td>1994</td>
<td>735</td>
<td>49</td>
</tr>
</tbody>
</table>

Add these points to your scatterplot. Florida took stronger measures to protect manatees during these years. Do you see any evidence that these measures suc-
(f) In part (c) you predicted manatee deaths in a year with 716,000 power boat registrations. In fact, powerboat registrations remained at 716,000 for the next three years. Compare the mean manatee deaths in these years with your prediction from part (c). How accurate was your prediction?

2.20 The number of people living on American farms has declined steadily during the past century. Here are data on the farm population (millions of persons) from 1935 to 1980:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>32.1</td>
<td>30.5</td>
<td>24.4</td>
<td>23.0</td>
<td>19.1</td>
<td>15.6</td>
<td>12.4</td>
<td>9.7</td>
<td>8.9</td>
<td>7.2</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot of these data and find the least-squares regression line of farm population on year.

(b) According to the regression line, how much did the farm population decline each year on the average during this period? What percent of the observed variation in farm population is accounted for by linear change over time?

(c) Use the regression equation to predict the number of people living on farms in 1990. Is this result reasonable? Why?

2.21 Keeping water supplies clean requires regular measurement of levels of pollutants. The measurements are indirect—a typical analysis involves forming a dye by a chemical reaction with the dissolved pollutant, then passing light through the solution and measuring its “absorbance.” To calibrate such measurements, the laboratory measures known standard solutions and uses regression to relate absorbance to pollutant concentration. This is usually done every day. Here is one series of data on the absorbance for different levels of nitrates. Nitrates are measured in milligrams per liter of water. (From a presentation by Charles Knauf, Monroe County (New York) Environmental Health Laboratory.)

<table>
<thead>
<tr>
<th>Nitrates</th>
<th>50</th>
<th>50</th>
<th>100</th>
<th>200</th>
<th>400</th>
<th>800</th>
<th>1200</th>
<th>1600</th>
<th>2000</th>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absorbance</td>
<td>7.0</td>
<td>7.5</td>
<td>12.8</td>
<td>24.0</td>
<td>47.0</td>
<td>93.0</td>
<td>138.0</td>
<td>183.0</td>
<td>230.0</td>
<td>226.0</td>
</tr>
</tbody>
</table>
(a) Chemical theory says that these data should lie on a straight line. If the correlation is not at least 0.997, something went wrong and the calibration procedure is repeated. Plot the data and find the correlation. Must the calibration be done again?

(b) What is the equation of the least-squares line for predicting absorbance from concentration? If the lab analyzed a specimen with 500 milligrams of nitrates per liter, what do you expect the absorbance to be? Based on your plot and the correlation, do you expect your predicted absorbance to be very accurate?

2.22 You have data on an explanatory variable $x$ and a response variable $y$, and have found the least-squares regression line of $y$ on $x$. Add one more data point, with $x$ equal to the mean $\bar{x}$ for the existing points and $y$ greater than the mean $\bar{y}$ for the existing points. Show that the new least-squares line is parallel to the existing line. (To do this, you must show that the slope does not change when you add the new point. Start with the fact that the slope is $b = rs_{y}/s_{x}$ and substitute the definition of the correlation $r$.)

Studies of disease often ask people about their diet in years past in order to discover links between diet and disease. How well do people remember their past diet? Can we predict actual past diet as well or better from what subjects eat now as from their memory of past habits? Data on actual past diet are available for 91 people who were asked about their diet when they were 18 years old and again when they were 30. Researchers asked them at about age 55 to describe their eating habits at ages 18 and 30 and also their current diet. The study report says:

The first study aim, to determine how accurately this group of participants remembered past consumption, was addressed by correlations between recalled and historical consumption in each time period. To evaluate the second study aim, that is, whether recalled intake or current
intake more accurately predicts historical intake of food groups at age 30 years, we performed regression analysis.

Exercises 2.23 to 2.25 ask you to interpret the results of this paper to a group of people who know no statistics. (J. T. Dwyer et al., “Memory of food intake in the distant past,” *American Journal of Epidemiology*, 130 (1989), pp. 1033–1046.)

**2.23** Explain in nontechnical language what “correlation” means, why correlation suits the first aim of the study, what “regression” means, and why regression fits the second study aim. Be sure to point out the distinction between correlation and regression.

**2.24** The study looked at the correlations between actual intake of many foods at age 18 and the intake the subjects now remember for age 18. The median correlation was $r = 0.217$. The authors say, “We conclude that memory of food intake in the distant past is fair to poor.” Explain to your audience why $r = 0.217$ points to this conclusion.

**2.25** The authors used regression to predict the intake of a number of foods at age 30 from current intake of those foods and from what the subjects now remember about their intake at age 30. They conclude that “recalled intake more accurately predicted historical intake at age 30 years than did current diet.” As evidence, they present $r^2$-values for the regressions. Explain to your audience why comparing $r^2$-values is one way to compare how well different explanatory variables predict a response.

**2.26** The following table gives the U.S. resident population of voting age and the votes cast for president, both in thousands, for presidential elections between 1960 and 2000:
Section 2.4

<table>
<thead>
<tr>
<th>Year</th>
<th>Population</th>
<th>Votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>109,672</td>
<td>68,838</td>
</tr>
<tr>
<td>1964</td>
<td>114,090</td>
<td>70,645</td>
</tr>
<tr>
<td>1968</td>
<td>120,285</td>
<td>73,212</td>
</tr>
<tr>
<td>1972</td>
<td>140,777</td>
<td>77,719</td>
</tr>
<tr>
<td>1976</td>
<td>152,308</td>
<td>81,556</td>
</tr>
<tr>
<td>1980</td>
<td>163,945</td>
<td>86,515</td>
</tr>
<tr>
<td>1984</td>
<td>173,995</td>
<td>92,653</td>
</tr>
<tr>
<td>1988</td>
<td>181,956</td>
<td>91,595</td>
</tr>
<tr>
<td>1992</td>
<td>189,524</td>
<td>104,425</td>
</tr>
<tr>
<td>1996</td>
<td>196,511</td>
<td>96,456</td>
</tr>
<tr>
<td>2000</td>
<td>209,128</td>
<td>105,363</td>
</tr>
</tbody>
</table>

(a) For each year compute the percent of people who voted. Make a time plot of the percent who voted. Describe the change over time in participation in presidential elections.

(b) Before proposing political explanations for this change, we should examine possible lurking variables. The minimum voting age in presidential elections dropped from 21 to 18 years in 1970. Use this fact to propose a partial explanation for the trend you saw in (a).

Section 2.4

2.27 The table for Exercise 1.39 gives the calories and sodium content for each of 17 brands of meat hot dogs. “Eat Slim Veal Hot Dogs,” with just 107 calories, is a low outlier in the distribution of calories.

(a) Make a scatterplot of sodium content $y$ against calories $x$. Describe the main features of the relationship. Is “Eat Slim Veal Hot Dogs” an outlier in this plot? Circle its point.

(b) Calculate two least-squares regression lines, one using all of the observations and the other omitting “Eat Slim.” Draw both lines on your plot. Does a comparison of the two regression lines show that the “Eat Slim” brand is influential? Explain your answer.
(c) A new brand of meat hot dog (not made with veal) has 150 calories per frank.
How many milligrams of sodium do you estimate that one of these hot dogs contains?

2.28 Research on digestion requires accurate measurements of blood flow through the lining of the stomach. A promising way to make such measurements easily is to inject mildly radioactive microscopic spheres into the blood stream. The spheres lodge in tiny blood vessels at a rate proportional to blood flow; their radioactivity allows blood flow to be measured from outside the body. Medical researchers compared blood flow in the stomachs of dogs, measured by use of microspheres, with simultaneous measurements taken using a catheter inserted into a vein. The data, in milliliters of blood per minute (ml/minute), appear below. (Based on L. H. Archibald, F. G. Moody, and M. Simons, “Measurement of gastric blood flow with radioactive microspheres,” *Journal of Applied Physiology*, 38 (1975), pp. 1051–1056.)

<table>
<thead>
<tr>
<th>Spheres</th>
<th>4.0</th>
<th>4.7</th>
<th>6.3</th>
<th>8.2</th>
<th>12.0</th>
<th>15.9</th>
<th>17.4</th>
<th>18.1</th>
<th>20.2</th>
<th>23.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vein</td>
<td>3.3</td>
<td>8.3</td>
<td>4.5</td>
<td>9.3</td>
<td>10.7</td>
<td>16.4</td>
<td>15.4</td>
<td>17.6</td>
<td>21.0</td>
<td>21.7</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot of these data, with the microsphere measurement as the explanatory variable. There is a strongly linear pattern.

(b) Calculate the least-squares regression line of venous flow on microsphere flow. Draw your regression line on the scatterplot.

(c) Predict the venous measurement for microsphere measurements 6, 12, and 18 ml/minute. If the microsphere measurements are within about 10% to 15% of the predicted venous measurements, the researchers will use the microsphere measurements in future studies. Is this condition satisfied over this range of blood flow?

2.29 Exercise 2.4 gives information on states’ performance on the National Assessment of Educational Progress (NAEP) year 2000 Mathematics Assessment. The two measures of performance are closely related. (In fact, the correlation is about $r = 0.95$.) The table in Exercise 2.4 also gives the percent of children aged 5 to 17 years in each state who lived in households with incomes below the federal poverty
level in 1998. We expect that poverty among children will be related to NAEP performance.

(a) Make a scatterplot suitable for predicting mean NAEP score from poverty percent. Describe the relationship, using correlation as a measure to complement your verbal description.

(b) Use software to find the least-squares regression line for predicting mean NAEP score from poverty rate, and the residuals from this line. We might call states with large positive residuals overachievers, because their fourth graders do better than the state poverty rate would lead us to guess. Similarly, states with large negative residuals might be called underachievers. What are the three states with the largest positive residuals and the three states with the largest negative residuals?

2.30 We might expect states with more poverty to have fewer doctors. Here are data on the percent of each state’s residents living below the poverty line and on the number of M.D.’s per 100,000 residents in each state.
<table>
<thead>
<tr>
<th>State</th>
<th>Poverty percent</th>
<th>M.D.’s per 100,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>15.1</td>
<td>198</td>
</tr>
<tr>
<td>Alaska</td>
<td>8.6</td>
<td>167</td>
</tr>
<tr>
<td>Arizona</td>
<td>15.2</td>
<td>202</td>
</tr>
<tr>
<td>Arkansas</td>
<td>16.4</td>
<td>190</td>
</tr>
<tr>
<td>California</td>
<td>15.3</td>
<td>247</td>
</tr>
<tr>
<td>Colorado</td>
<td>8.6</td>
<td>238</td>
</tr>
<tr>
<td>Connecticut</td>
<td>8.4</td>
<td>354</td>
</tr>
<tr>
<td>Delaware</td>
<td>10.1</td>
<td>234</td>
</tr>
<tr>
<td>D.C.</td>
<td>19.7</td>
<td>737</td>
</tr>
<tr>
<td>Florida</td>
<td>13.3</td>
<td>238</td>
</tr>
<tr>
<td>Georgia</td>
<td>13.7</td>
<td>211</td>
</tr>
<tr>
<td>Hawaii</td>
<td>11.9</td>
<td>265</td>
</tr>
<tr>
<td>Idaho</td>
<td>13.9</td>
<td>154</td>
</tr>
<tr>
<td>Illinois</td>
<td>10.4</td>
<td>260</td>
</tr>
<tr>
<td>Indiana</td>
<td>8.3</td>
<td>195</td>
</tr>
<tr>
<td>Iowa</td>
<td>8.7</td>
<td>173</td>
</tr>
<tr>
<td>Kansas</td>
<td>10.5</td>
<td>203</td>
</tr>
<tr>
<td>Kentucky</td>
<td>13.8</td>
<td>209</td>
</tr>
<tr>
<td>Louisiana</td>
<td>18.2</td>
<td>246</td>
</tr>
<tr>
<td>Maine</td>
<td>10.4</td>
<td>223</td>
</tr>
<tr>
<td>Maryland</td>
<td>7.6</td>
<td>374</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>10.9</td>
<td>412</td>
</tr>
<tr>
<td>Michigan</td>
<td>10.3</td>
<td>224</td>
</tr>
<tr>
<td>Minnesota</td>
<td>9.1</td>
<td>249</td>
</tr>
<tr>
<td>Mississippi</td>
<td>16.8</td>
<td>163</td>
</tr>
<tr>
<td>Missouri</td>
<td>11.1</td>
<td>230</td>
</tr>
<tr>
<td>Montana</td>
<td>15.9</td>
<td>190</td>
</tr>
<tr>
<td>Nebraska</td>
<td>11.0</td>
<td>218</td>
</tr>
<tr>
<td>Nevada</td>
<td>11.0</td>
<td>173</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>8.9</td>
<td>237</td>
</tr>
<tr>
<td>New Jersey</td>
<td>8.5</td>
<td>295</td>
</tr>
<tr>
<td>New Mexico</td>
<td>20.8</td>
<td>212</td>
</tr>
<tr>
<td>New York</td>
<td>15.7</td>
<td>387</td>
</tr>
<tr>
<td>North Carolina</td>
<td>13.0</td>
<td>232</td>
</tr>
<tr>
<td>North Dakota</td>
<td>13.9</td>
<td>222</td>
</tr>
<tr>
<td>Ohio</td>
<td>11.4</td>
<td>235</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>13.5</td>
<td>169</td>
</tr>
<tr>
<td>Oregon</td>
<td>13.1</td>
<td>225</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>10.6</td>
<td>291</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>11.4</td>
<td>338</td>
</tr>
<tr>
<td>South Carolina</td>
<td>12.8</td>
<td>207</td>
</tr>
<tr>
<td>South Dakota</td>
<td>11.7</td>
<td>184</td>
</tr>
<tr>
<td>Tennessee</td>
<td>13.2</td>
<td>246</td>
</tr>
<tr>
<td>Texas</td>
<td>15.6</td>
<td>203</td>
</tr>
<tr>
<td>Utah</td>
<td>7.9</td>
<td>200</td>
</tr>
<tr>
<td>Vermont</td>
<td>9.6</td>
<td>305</td>
</tr>
<tr>
<td>Virginia</td>
<td>9.8</td>
<td>241</td>
</tr>
<tr>
<td>Washington</td>
<td>9.2</td>
<td>235</td>
</tr>
<tr>
<td>West Virginia</td>
<td>16.7</td>
<td>215</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>8.5</td>
<td>227</td>
</tr>
<tr>
<td>Wyoming</td>
<td>11.9</td>
<td>171</td>
</tr>
</tbody>
</table>
(a) Make a scatterplot and calculate a regression line suitable for predicting M.D.’s per 100,000 from poverty rate. Draw the line on your plot. Surprise: the slope is positive, so poverty and M.D.’s go up together.

(b) The District of Columbia is an outlier, with both very many M.D.’s and a high poverty rate. (D.C. is a city rather than a state.) Circle the point for D.C. on your plot and explain why this point may strongly influence the least-squares line.

(c) Calculate the regression line for the 50 states, omitting D.C. Add the new line to your scatterplot. Was this point highly influential? Does the number of doctors now go down with increasing poverty, as we initially expected?

2.31 The Standard & Poor’s 500-stock index is an average of the price of 500 stocks. There is a moderately strong correlation (roughly $r = 0.6$) between how much this index changes in January and how much it changes during the entire year. If we looked instead at data on all 500 individual stocks, we would find a quite different correlation. Would the correlation be higher or lower? Why?

2.32 Airborne particles such as dust and smoke are an important part of air pollution. To measure particulate pollution, a vacuum motor draws air through a filter for 24 hours. Weigh the filter at the beginning and end of the period. The weight gained is a measure of the concentration of particles in the air. A study of air pollution made measurements every 6 days with identical instruments in the center of a small city and at a rural location 10 miles southwest of the city. Because the prevailing winds blow from the west, we suspect that the rural readings will be generally lower than the city readings, but that the city readings can be predicted from the rural readings. Here are readings taken every 6 days over a 7-month period. The entry NA means that the reading for that date is not available, usually because of equipment failure. (Data provided by Matthew Moore.)
(a) We hope to use the rural particulate level to predict the city level on the same day. Make a graph to examine the relationship. Does the graph suggest that using the least-squares regression line for prediction will give approximately correct results over the range of values appearing in the data? Calculate the least-squares line for predicting city pollution from rural pollution. What percent of the observed variation in city pollution levels does this straight-line relationship account for?

(b) Find the residuals from your least-squares fit. Plot the residuals both against $x$ and against the time order of the observations, and comment on the results.

(c) Which observation appears to be the most influential? Circle this observation on your plot. Is it the observation with the largest residual?

(d) On the fourteenth date in the series, the rural reading was 88 and the city reading was not available. What do you estimate the city reading to be for that date?

(e) Make a normal quantile plot of the residuals. (Make a stemplot or histogram if your software does not make normal quantile plots.) Is the distribution of the residuals approximately normal?
Section 2.5

2.33 A study of grade-school children aged 6 to 11 years found a high positive correlation between reading ability \( y \) and shoe size \( x \). Explain why common response to a lurking variable \( z \) accounts for this correlation.

2.34 There is a negative correlation between the number of flu cases \( y \) reported each week through the year and the amount of ice cream \( x \) sold that week. It is unlikely that ice cream prevents flu. What is a more plausible explanation for this correlation?

2.35 Members of a high school language club believe that study of a foreign language improves a student’s command of English. From school records, they obtain the scores on an English achievement test given to all seniors. The mean score of seniors who had studied a foreign language for at least two years is much higher than the mean score of seniors who studied no foreign language. The club’s advisor says that these data are not good evidence that language study strengthens English skills. Identify the explanatory and response variables in this study. Then explain what lurking variable prevents the conclusion that language study improves students’ English scores.

Chapter 2 Review Exercises

Exercises 2.36 to 2.38 concern these data on the total returns on U.S. and overseas common stocks over a 30-year period. (The total return is change in price plus any dividends paid, converted into U.S. dollars. Both returns are averages over many individual stocks.)
Chapter 2 Exercises

2.36 (a) Make a scatterplot suitable for predicting overseas returns from U.S. returns.

(b) Find the correlation and $r^2$. Describe the relationship between U.S. and overseas returns in words, using $r$ and $r^2$ to make your description more precise.

(c) Find the least-squares regression line of overseas returns on U.S. returns. Draw the line on the scatterplot. What are the predicted return $\hat{y}$ and the observed return $y$ for 1993?

(d) Are you confident that predictions using the regression line will be quite accurate? Why?
2.37 Return to the scatterplot and regression line in the previous exercise.

(a) Circle the point that has the largest residual (either positive or negative). What year is this? Redo the regression without this point and add the new regression line to your plot. Was this observation very influential?

(b) Whenever we regress two variables that both change over time, we should plot the residuals against time as a check for time-related lurking variables. Make this plot for the stock returns data. Are there any suspicious patterns in the residuals?

2.38 Investors also want to know what typical returns are and how much year-to-year variability (called volatility in finance) there is. Regression and correlation don’t answer questions about center and spread.

(a) Find the five-number summaries for both U.S. and overseas returns, and make side-by-side boxplots to compare the two distributions.

(b) Were returns generally higher in the United States or overseas during this period? Explain your answer.

(c) Were returns more volatile (more variable) in the United States or overseas during this period? Explain your answer.

There are different ways to measure the amount of money spent on education. Average salary paid to teachers and expenditures per pupil are two possible measures. The table at the top of the next page gives the 1995 values for these variables by state. The states are classified according to region: NE (New England), MA (Middle Atlantic), ENC (East North Central), WNC (West North Central), SA (South Atlantic), ESC (East South Central), WSC (West South Central), MN (Mountain), and PA (Pacific). Exercises 2.39 to 2.41 are based on these data.

2.39 Make a stemplot or histogram for teachers’ pay. Is the distribution roughly symmetric or clearly skewed? Find the five-number summary. Are there any sus-
pected outliers by the $1.5 \times IQR$ criterion? Which states may be outliers? Do the same for spending per pupil. Are the same states outliers in both distributions?

2.40 (a) Make a scatterplot of teachers’ pay $y$ against spending $x$. Describe the pattern of the relationship between pay and spending. Is there a strong association? If so, is it positive or negative? Explain why you might expect to see an association of this kind.

(b) Find the least-squares regression line for predicting teachers’ pay from education spending and draw it on your scatterplot. How much on the average does mean teachers’ pay increase when spending increases by $1000$ per pupil from one state to another? Give a numerical measure of the success of overall spending on education in explaining variations in teachers’ pay among states.

(c) On your plot, circle any outlying points found in (a). Label the circled points with the state identifier. Do these points have large residuals? (You need not actually calculate the residuals.) The states you have identified lie close together on the plot. To see if they are influential as a group, find the regression line with all of these states removed from the calculation. Draw this new line on your plot. Was this group of states influential?

2.41 Continue the analysis of teachers’ pay and education spending by looking for regional effects. We will compare these three groups:

- Coastal: Middle Atlantic, New England, and Pacific
- South: South Atlantic, East South Central, and West South Central
- Midwest: East North Central and West North Central

Omit the District of Columbia, which is a city rather than a state.

(a) Make side-by-side boxplots for education spending in the three regions. For each region, label any outliers (points identified by the $1.5 \times IQR$ criterion) with the state identifier.

(b) Repeat part (a) for teachers’ pay.
(c) Do you see important differences in spending and pay by region? Are the differences consistent for the two variables? That is, are regions that are high in spending also high in pay and vice versa?

CHAPTER 3

Section 3.1

3.1 Yvette is a young banker. She and all her friends carry cell phones and use them heavily. Last year, two of Yvette’s acquaintances developed brain tumors. Yvette wonders if the tumors are related to use of cell phones. Explain briefly why the experience of Yvette’s friends does not provide good evidence that cell phones cause brain tumors.

3.2 There is strong public support for “term limits” that restrict the number of terms that legislators can serve. One possible explanation for this support is that voters are dissatisfied with the performance of Congress and other legislative bodies. A political scientist asks a sample of voters if they support term limits for members of Congress and also asks several questions that gauge their satisfaction with Congress. He finds no relationship between approval of Congress and support for term limits. Is this an observational study or an experiment? Why? What are the explanatory and response variables?

3.3 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 selects a large sample of registered voters, both men and women. She asks every voter whether they voted for the Democratic or the Republican candidate in
the last congressional election. Is this an observational study or an experiment? Why? What are the explanatory and response variables?

3.4 Many studies have found that people who drink alcohol in moderation have lower risk of heart attacks than either nondrinkers or heavy drinkers. Does alcohol consumption also improve survival after a heart attack? One study followed 1913 people who were hospitalized after severe heart attacks. In the year before their heart attack, 47% of these people did not drink, 36% drank moderately, and 17% drank heavily. After four years, fewer of the moderate drinkers had died. (K. J. Mukamal et al., “Prior alcohol consumption and mortality following acute myocardial infarction,” *Journal of the American Medical Association*, 285 (2001), pp. 1965–1970.) Is this an observational study or an experiment? Why? What are the explanatory and response variables?

3.5 A study of the effect of living in public housing on family stability and other variables in poverty-level households was carried out as follows. The 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 study an experiment? Why or why not? What are the explanatory and response variables in the study?

3.6 The National Halothane Study was a major investigation of the safety of anesthetics used in surgery. Records of over 850,000 operations performed in 34 major hospitals showed the following death rates for four common anesthetics. (L. E. Moses and F. Mosteller, “Safety of anesthetics,” in J. M. Tanur et al. (eds.), *Statistics: A Guide to the Unknown*, 3rd ed., Wadsworth, 1989, pp. 15–24.)

<table>
<thead>
<tr>
<th>Anesthetic</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death rate</td>
<td>1.7%</td>
<td>1.7%</td>
<td>3.4%</td>
<td>1.9%</td>
</tr>
</tbody>
</table>
There is a clear association between the anesthetic used and the death rate of patients. Anesthetic C appears dangerous.

(a) Explain why we call the National Halothane Study an observational study rather than an experiment, even though it compared the results of using different anesthetics in actual surgery.

(b) When the study looked at other variables that are confounded with a doctor’s choice of anesthetic, it found that Anesthetic C was not causing extra deaths. Suggest several variables that are mixed up with what anesthetic a patient receives.

3.7 Some people believe that exercise raises the body’s metabolic rate for as long as 12 to 24 hours, enabling us to continue to burn off fat after our workout has ended. In a study of this effect, subjects walked briskly on a treadmill for several hours. Their metabolic rates were measured before, immediately after, and 12 hours after the exercise. The study was criticized because eating raises the metabolic rate, and no record was kept of what the subjects ate after exercising. Was this study an experiment? Why or why not? What are the explanatory and response variables?

Section 3.2

3.8 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 required to peel the liner? To answer this question, the engineers prepare 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 peel strength of each seal. Identify the experimental units or subjects, the factors, the treatments, and the response variables.

3.9 Sickle cell disease is an inherited disorder of the red blood cells that in the United States affects mostly blacks. It can cause severe pain and many complications.
Can the drug hydroxyurea reduce the severe pain caused by sickle cell disease? A study by the National Institutes of Health gave the drug to 150 sickle cell sufferers and a placebo to another 150. The researchers then counted the episodes of pain reported by each subject. Identify the experimental units or subjects, the factors, the treatments, and the response variables.

3.10 People who eat lots of fruits and vegetables have lower rates of colon cancer than those who eat little of these foods. Fruits and vegetables are rich in “antioxidants” such as vitamins A, C, and E. Will taking antioxidants help prevent colon cancer? A clinical trial studied this question with 864 people who were at risk of colon cancer. The subjects were divided into four groups: daily beta carotene, daily vitamins C and E, all three vitamins every day, and daily placebo. After four years, the researchers were surprised to find no significant difference in colon cancer among the groups. (G. Kolata, “New study finds vitamins are not cancer preventers,” New York Times, July 21, 1994.)

(a) What are the explanatory and response variables in this experiment?

(b) Outline the design of the experiment. Use your judgment in choosing the group sizes.

(c) Assign labels to the 864 subjects and use Table B starting at line 118 to choose the first 5 subjects for the beta carotene group.

(d) The study was double-blind. What does this mean?

(e) What does “no significant difference” mean in describing the outcome of the study?

(f) Suggest some lurking variables that could explain why people who eat lots of fruits and vegetables have lower rates of colon cancer. The experiment suggests that these variables, rather than the antioxidants, may be responsible for the observed benefits of fruits and vegetables.
Exercise 3.9 describes a medical study of a new treatment for sickle cell disease.

(a) Outline the design of this experiment.

(b) Use of a placebo is considered ethical if there is no effective standard treatment to give the control group. It might seem humane to give all the subjects hydroxyurea in the hope that it will help them. Explain clearly why this would not provide information about the effectiveness of the drug. (In fact, the experiment was stopped ahead of schedule because the hydroxyurea group had only half as many pain episodes as the control group. Ethical standards required stopping the experiment as soon as significant evidence became available.)

Outline the design of the package liner experiment of Exercise 3.10. Label the pairs of liner pieces 01 to 20 and carry out the randomization that your design calls for. (If you use Table B, start at line 120.)

Surgery patients are often cold because the operating room is kept cool and the body’s temperature regulation is disturbed by anesthetics. Will warming patients to maintain normal body temperature reduce infections after surgery? In one experiment, patients undergoing colon surgery received intravenous fluids from a warming machine and were covered with a blanket through which air circulated. For some patients, the fluid and the air were warmed; for others, they were not. The patients received identical treatment in all other respects.

(a) Outline the design of a randomized comparative experiment for this study.

(b) The following subjects have given consent to participate in this study. Do the random assignment required by your design. (If you use Table B, begin at line 121.)
3.14 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 suspect that a higher percentage of those who read the description that includes child care will choose Company B.

(a) Outline the design of the experiment. Be sure to identify the response variable.

(b) The names of the subjects appear below. Do the randomization required by your design and list the subjects who will read the version that mentions child care. (If you use Table B, begin at line 121.)
3.15 A horticulturist is comparing two methods (call them A and B) of growing potatoes. Standard potato cuttings will be planted in small plots of ground. The response variables are number of tubers per plant and fresh weight (weight when just harvested) of vegetable growth per plant. There are 20 plots available for the experiment. Sketch the outline of a rectangular field divided into 5 rows of 4 plots each. Then outline the experimental design and do the required randomization. (If you use Table B, start at line 145.) Mark on your sketch which growing method you will use in each plot.

3.16 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.

3.17 Here are some questions about the study of heating surgery patients in Exercise 3.13.

(a) To simplify the setup of the study, we might warm the fluids and air blanket for one operating team and not for another doing the same kind of surgery. Why might this design result in bias?
(b) The operating team did not know whether fluids and air blanket were heated, nor did the doctors who followed the patients after surgery. What is this practice called? Why was it used here?

3.18 You want to determine the best color for attracting cereal leaf beetles to boards on which they will be trapped. You will compare four colors: blue, green, white, and yellow. The response variable is the count of beetles trapped. You will mount one board on each of 16 poles evenly spaced in a square field, with four poles in each of four rows. Sketch the field with the locations of the 16 poles. Outline the design of a completely randomized experiment to compare the colors. Randomly assign colors to the poles, and mark on your sketch the color assigned to each pole.

(If you use Table B, start at line 115.)

3.19 Continue the discussion of the experiment of the previous exercise. The researchers decide to use two oat fields in different locations and to space eight poles equally within each field. Outline a randomized block design using the fields as blocks. Then use Table B, beginning at line 105, to carry out the random assignment of colors to poles. Report your results by means of a sketch of the two fields with the color at each pole noted.

3.20 A mathematics education researcher is studying where in high school mathematics texts it is most effective to insert questions. She wants to know whether it is better to present questions as motivation before the text passage or as review after the passage. The result may depend on the type of question asked: simple fact, computation, or word problem.

(a) This experiment has two factors. What are they? How many treatments do all combinations of levels of these factors form? List the treatments.

(b) Because it is disruptive to assign high school students at random to the treatment groups, the researcher will assign two classes of the same grade level to each treatment. The response variable is score on a mathematics test taken by all the
students in these classes. Outline the design of the experiment. Carry out the random assignment required.

3.21 A study of the effects of running on personality involved 231 male runners who each ran about 20 miles a week. The runners were given the Cattell Sixteen Personality Factor Questionnaire, a 187-item multiple-choice test often used by psychologists. A news report (New York Times, February 15, 1988) stated, “The researchers found statistically significant personality differences between the runners and the 30-year-old male population as a whole.” A headline on the article said, “Research has shown that running can alter one’s moods.”

(a) Explain carefully, to someone who knows no statistics, what “statistically significant” means.

(b) Explain carefully, to someone who knows no statistics, why the headline is misleading.

3.22 Is the right hand generally stronger than the left in right-handed people? You can crudely measure hand strength by placing a bathroom scale on a shelf with the end protruding, then squeezing the scale between the thumb below and the four fingers above it. The reading of the scale shows the force exerted. Describe the design of a matched pairs experiment to compare the strength of the right and left hands, using 16 right-handed people as subjects. Use Table B at line 114 to choose which 8 subjects will try their right hands first.

3.23 Do consumers prefer the taste of Pepsi or Coke in a blind test in which neither cola is identified? Describe briefly the design of a matched pairs experiment to investigate this question. How will you use randomization?

3.24 There are several psychological tests available to 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.

3.25 Will people spend less on health care if their health insurance requires them to pay some part of the cost themselves? An experiment on this issue asked if the percent of medical costs that are paid by health insurance has an effect either on the amount of medical care that people use or on their health. The treatments were four insurance plans. Each plan paid all medical costs above a ceiling. Below the ceiling, the plans paid 100%, 75%, 50%, or 0% of costs incurred.

(a) Outline the design of a randomized comparative experiment suitable for this study.

(b) Describe briefly the practical and ethical difficulties that might arise in such an experiment.

3.26 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. Two batches of the feedstock will be processed at each combination of temperature and stirring rate.

(a) How many factors are there in this experiment? How many treatments? Identify each of the treatments. How many experimental units (batches of feedstock) does the experiment require?

(b) Outline in graphic form the design of an appropriate experiment.
(c) The randomization in this experiment determines the order in which batches of the feedstock will be processed according to each treatment. Use Table B starting at line 128 to carry out the randomization and state the result.

Section 3.3

3.27 A political scientist wants to know how college students feel about the Social Security system. She obtains a list of the 3456 undergraduates at her college and mails a questionnaire to 250 students selected at random. Only 104 questionnaires are returned. What is the population in this study? What is the sample from which information was actually obtained? What is the rate (percent) of nonresponse?

3.28 Different types of writing can sometimes be distinguished by the lengths of the words used. A student interested in this fact wants to study the lengths of words used by Tom Clancy in his novels. She opens a Clancy novel at random and records the lengths of each of the first 250 words on the page. What is the population in this study? What is the sample? What is the variable measured?

3.29 A newspaper article about an opinion poll says that “43% of Americans approve of the president’s overall job performance.” Toward the end of the article, you read: “The poll is based on telephone interviews with 1210 adults from around the United States, excluding Alaska and Hawaii.” What variable did this poll measure? What population do you think the newspaper wants information about? What was the sample? Are there any sources of bias in the sampling method used?

3.30 A newspaper advertisement for USA Today: The Television Show said:


Charge is 50 cents for the first minute.

Explain why this opinion poll is almost certainly biased.
3.31 The students listed below are enrolled in a statistics course taught on television. Choose an SRS of 6 students to be interviewed in detail about the quality of the course. (If you use Table B, start at line 139.)

<table>
<thead>
<tr>
<th>Abate</th>
<th>Dubois</th>
<th>Hixson</th>
<th>Putnam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson</td>
<td>Fernandez</td>
<td>Klassen</td>
<td>Rodriguez</td>
</tr>
<tr>
<td>Baxter</td>
<td>Frank</td>
<td>Liang</td>
<td>Rubin</td>
</tr>
<tr>
<td>Bowen</td>
<td>Fuhrmann</td>
<td>Moser</td>
<td>Santiago</td>
</tr>
<tr>
<td>Bruvold</td>
<td>Goel</td>
<td>Naber</td>
<td>Shen</td>
</tr>
<tr>
<td>Casella</td>
<td>Gupta</td>
<td>Petrucelli</td>
<td>Shyr</td>
</tr>
<tr>
<td>Choi</td>
<td>Hicks</td>
<td>Pliego</td>
<td>Sundheim</td>
</tr>
</tbody>
</table>

3.32 You want to choose an SRS of 25 of a city’s 440 voting precincts for special voting-fraud surveillance on election day. How will you label the 440 precincts? Choose the SRS, and list the precincts you selected. (Use the Simple Random Sample applet. If you use Table B, enter at line 117 and select only the first 5 precincts in the sample.)

3.33 A firm wants to understand the attitudes of its minority managers toward its system for assessing management performance. Below is a list of all the firm’s managers who are members of minority groups. Use Table B at line 139 to choose 6 to be interviewed in detail about the performance appraisal system.

<table>
<thead>
<tr>
<th>Acosta</th>
<th>Dewald</th>
<th>Huang</th>
<th>Puri</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ali</td>
<td>Fleming</td>
<td>Kim</td>
<td>Richards</td>
</tr>
<tr>
<td>Baxter</td>
<td>Fonseca</td>
<td>Lujan</td>
<td>Rodriguez</td>
</tr>
<tr>
<td>Bowman</td>
<td>Gates</td>
<td>Mourning</td>
<td>Santiago</td>
</tr>
<tr>
<td>Brams</td>
<td>Goel</td>
<td>Nunez</td>
<td>Shen</td>
</tr>
<tr>
<td>Cortez</td>
<td>Gomez</td>
<td>Peters</td>
<td>Vega</td>
</tr>
<tr>
<td>Cross</td>
<td>Hernandez</td>
<td>Pliego</td>
<td>Watanabe</td>
</tr>
</tbody>
</table>
3.34 An academic department wishes to choose a three-member advisory committee at random from the members of the department. Use Table B at line 140 to choose an SRS of size 3 from the 28 faculty listed below. (For convenience, they are labeled in alphabetical order.)

<table>
<thead>
<tr>
<th>Label</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Abate</td>
</tr>
<tr>
<td>01</td>
<td>Cicirelli</td>
</tr>
<tr>
<td>02</td>
<td>Cuellar</td>
</tr>
<tr>
<td>03</td>
<td>Dunsmore</td>
</tr>
<tr>
<td>04</td>
<td>Engle</td>
</tr>
<tr>
<td>05</td>
<td>Fitzpatrick</td>
</tr>
<tr>
<td>06</td>
<td>Garcia</td>
</tr>
<tr>
<td>07</td>
<td>Goodwin</td>
</tr>
<tr>
<td>08</td>
<td>Haglund</td>
</tr>
<tr>
<td>09</td>
<td>Johnson</td>
</tr>
<tr>
<td>10</td>
<td>Keegan</td>
</tr>
<tr>
<td>11</td>
<td>Luo</td>
</tr>
<tr>
<td>12</td>
<td>Martinez</td>
</tr>
<tr>
<td>13</td>
<td>Nguyen</td>
</tr>
<tr>
<td>14</td>
<td>Pilotte</td>
</tr>
<tr>
<td>15</td>
<td>Raman</td>
</tr>
<tr>
<td>16</td>
<td>Riemann</td>
</tr>
<tr>
<td>17</td>
<td>Rodriguez</td>
</tr>
<tr>
<td>18</td>
<td>Rowe</td>
</tr>
<tr>
<td>19</td>
<td>Salazar</td>
</tr>
<tr>
<td>20</td>
<td>Stone</td>
</tr>
<tr>
<td>21</td>
<td>Theobald</td>
</tr>
<tr>
<td>22</td>
<td>Vader</td>
</tr>
<tr>
<td>23</td>
<td>Wang</td>
</tr>
<tr>
<td>24</td>
<td>Wieczorek</td>
</tr>
<tr>
<td>25</td>
<td>Williams</td>
</tr>
<tr>
<td>26</td>
<td>Wilson</td>
</tr>
<tr>
<td>27</td>
<td>Wong</td>
</tr>
</tbody>
</table>

3.35 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.

3.36 A labor organization wants to study the attitudes of college faculty members toward collective bargaining. These attitudes appear to be different 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 faculty.

3.37 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. (W. Mitofsky, “Mr. Perot, you’re no pollster,” *New York Times*, March 27, 1993.)

Section 3.4

3.38 Voter registration records show that 68% of all voters in Indianapolis are registered as Republicans. To test a random digit dialing device, you use the device to call 150 randomly chosen residential telephones in Indianapolis. Of the registered voters contacted, 73% are registered Republicans.

3.39 A carload lot of ball bearings has a mean diameter of 2.503 centimeters (cm). This is within the specifications for acceptance of the lot by the purchaser. The inspector happens to inspect 100 bearings from the lot with a mean diameter of 2.515 cm. This is outside the specified limits, so the lot is mistakenly rejected.
3.40 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, 43 are unlisted. This is not surprising, because 52% of all Los Angeles residential phones are unlisted.

3.41 The Carolina Abecedarian Project investigated the effect of high-quality preschool programs on children from poor families. Children were randomly assigned to two groups. One group participated in a year-round preschool program from age three months. The control group received social services but no preschool. At age 21, 35% of the treatment group and 14% of the control group were attending a four-year college or had already graduated from college.

3.42 Just before a presidential election, a national opinion polling firm increases the size of its weekly sample from the usual 1500 people to 4000 people. Why do you think the firm does this?

3.43 A management student is planning to take a survey of student attitudes toward part-time work while attending college. He develops a questionnaire and plans to ask 25 randomly selected students to fill it out. His faculty advisor approves the questionnaire but urges that the sample size be increased to at least 100 students. Why is the larger sample helpful?

3.44 An entomologist samples a field for egg masses of a harmful insect by placing a yard-square frame at random locations and examining the ground within the frame carefully. She wishes to estimate the proportion of square yards in which egg masses are present. Suppose that in a large field egg masses are present in 20% of all possible yard-square areas—that is, $p = 0.2$ in this population.

(a) Use Table B to simulate the presence or absence of egg masses in each square yard of an SRS of 10 square yards from the field. Be sure to explain clearly which digits you used to represent the presence and the absence of egg masses. What
proportion of your 10 sample areas had egg masses?

(b) Repeat (a) with different lines from Table B, until you have simulated the results of 20 SRSs of size 10. What proportion of the square yards in each of your 20 samples had egg masses? Make a stemplot from these 20 values to display the sampling distribution of \( \hat{p} \) in this case. What is the mean of this distribution? What is its shape?

Chapter 3 Review Exercises

3.45 You read a news report of an experiment that claims to show that a meditation technique lowered the anxiety level of subjects. The experimenter interviewed the subjects and assessed their levels of anxiety. The subjects then learned how to meditate and did so regularly for a month. The experimenter reinterviewed them at the end of the month and assessed whether their anxiety levels had decreased or not.

(a) There was no control group in this experiment. Why is this a blunder? What lurking variables might be confounded with the effect of meditation?

(b) The experimenter who diagnosed the effect of the treatment knew that the subjects had been meditating. Explain how this knowledge could bias the experimental conclusions.

(c) Briefly discuss a proper experimental design, with controls and blind diagnosis, to assess the effect of meditation on anxiety level.

3.46 A 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 an elaborate 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.

3.47 The National Institutes of Mental Health (NIMH) wants to know whether intense education about the risks of AIDS will help change the behavior of people who now report sexual activities that put them at risk of infection. NIMH investigators screened 38,893 people to identify 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. (NIMH Multisite HIV Prevention Trial Group, “The NIMH multisite HIV prevention trial: reducing HIV sexual risk behavior,” *Science*, 280 (1998), pp. 1889–1894.)

(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.

3.48 It is possible to use a computer to make telephone calls over the Internet. How will the cost affect the behavior of users of this service? You will offer the service to all 200 rooms in a college dormitory. Some rooms will pay a flat rate.
Chapter 3 Exercises

Others will pay higher rates at peak periods and very low rates off-peak. You are interested in the amount and time of use and in the effect on the congestion of the network. Outline the design of an experiment to study the effect of rate structure. Use Table B, starting at line 125, to assign the first 5 rooms to the flat-rate group.

3.49 What are the most important goals of schoolchildren? Do girls and boys have different goals? Are goals different in urban, suburban, and rural areas? To find out, researchers wanted to ask children in the fourth, fifth, and sixth grades this question:

What would you most like to do at school?

A. Make good grades.
B. Be good at sports.
C. Be popular.

Because most children live in heavily populated urban and suburban areas, an SRS might contain few rural children. Moreover, it is too expensive to choose children at random—we must start by choosing schools rather than children. Describe a suitable sample design for this study, using the ideas of stratified and multistage samples. Explain the reasoning behind your choice.

3.50 Advice columnist Ann Landers once asked her female readers whether they would be content with affectionate treatment by men, with no sex ever. Over 90,000 women wrote in, with 72% answering “Yes.” Many of the letters described unfeeling treatment by men. Explain why this sample is certainly biased. What is the likely direction of the bias? That is, is that 72% probably higher or lower than the truth about the population of all adult women?

3.51 A national opinion poll recently estimated that 44% (\(\hat{p} = 0.44\)) of all American adults agree that parents of school-age children should be given vouchers good for
education at any public or private school of their choice. The polling organization used a probability sampling method for which the sample proportion has a normal distribution with standard deviation about 0.015. The poll therefore announced a “margin of error” of 0.03 (two standard deviations) for its result. If a sample were drawn by the same method from the state of New Jersey (population 8 million) instead of from the entire United States (population 270 million), would this margin of error be larger or smaller? Explain your answer.

CHAPTER 4

Section 4.1

4.1 Toss a thumbtack on a hard surface 100 times. How many times did it land with the point up? What is the approximate probability of landing point up?

4.2 In the game of Heads or Tails, Betty and Bob toss a coin four times. Betty wins a dollar from Bob for each head and pays Bob a dollar for each tail— that is, she wins or loses the difference between the number of heads and the number of tails. For example, if there are one head and three tails, Betty loses $2. You can check that Betty’s possible outcomes are

\[ \{-4, -2, 0, 2, 4\} \]

Assign probabilities to these outcomes by playing the game 20 times and using the proportions of the outcomes as estimates of the probabilities. If possible, combine your trials with those of other students to obtain long-run proportions that are closer to the probabilities.

4.3 You read in a book on poker that the probability of being dealt three of a kind in a five-card poker hand is 1/50. Explain in simple language what this means.
4.4 A recent opinion poll showed that about 73% of married women agree that their husbands do at least their fair share of household chores. Suppose that this is exactly true. Choosing a married woman at random then has probability 0.73 of getting one who agrees that her husband does his share. You can use the Probability applet or software to simulate choosing many women independently. (In most software, the key phrase to look for is “Bernoulli trials.” This is the technical term for independent trials with “Yes/No” outcomes. Our outcomes here are “Agree” and “Disagree.”)

(a) Simulate drawing 20 women, then 80 women, then 320 women. What proportion agree in each case? We expect (but because of chance variation we can’t be sure) that the proportion will be closer to 0.73 in longer runs of trials.

(b) Simulate drawing 20 women 10 times and record the percents in each trial who agree. Then simulate drawing 320 women 10 times and again record the 10 percents. Which set of 10 results is less variable? We expect the results of larger samples to be more predictable (less variable) than the results of smaller samples. That is “long-run regularity” showing itself.

4.5 Continue the exploration begun in the previous exercise. Software allows you to simulate many independent “Yes/No” trials more quickly if all you want to save is the count of “Yes” outcomes. The keyword “Binomial” simulates \( n \) independent Bernoulli trials, each with probability \( p \) of a “Yes,” and records just the count of “Yes” outcomes.

(a) Simulate 100 draws of 20 women from this population. Record the number who say “Agree” on each draw. What is the approximate probability that out of 20 women drawn at random at least 14 agree?

(b) Convert the counts who agree into percents of the 20 women in each trial who agree. Make a histogram of these 100 percents. Describe the shape, center, and spread of this distribution.

(c) Now simulate drawing 320 women. Do this 100 times and record the percent
who agree on each of the 100 draws. Make a histogram of the percents and describe
the shape, center, and spread of the distribution.
(d) In what ways are the distributions in parts (b) and (c) alike? In what ways do
they differ? (Because regularity emerges in the long run, we expect the results of
drawing 320 women to be less variable than the results of drawing 20 women.)

Section 4.2

4.6 The percent return on U.S. common stocks in the next year is random. The
table for Exercise 2.36 reports historical data for the years 1971 to 2000. Give a
reasonable sample space for the possible returns next year. Explain how you chose
this $S$.

4.7 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 student enrolls in a statistics course and at the end of the semester receives a
letter grade.
(d) A basketball player shoots two free throws.
(e) A year after knee surgery, a patient is asked to rate the amount of pain in the
knee. A seven-point scale is used, with 1 corresponding to no pain and 7 corre-
sponding to extreme discomfort.

4.8 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 smallest value in $S$.
(a) Choose a student in your class at random. Ask how much time that student
spent studying during the past 24 hours.
(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.
(d) Choose a student in your class at random. Ask how much cash that student is carrying.
(e) 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.

4.9 Probability is a measure of how likely an event is to occur. Match one of the probabilities that follow with each statement about an event. (The probability is usually a much more exact measure of likelihood than is the verbal statement.)

\[ 0, 0.01, 0.3, 0.6, 0.99, 1 \]

(a) This event is impossible. It can never occur.
(b) This event is certain. It will occur on every trial of the random phenomenon.
(c) This event is very unlikely, but it will occur once in a while in a long sequence of trials.
(d) This event will occur more often than not.

4.10 In each of the following situations, state whether or not the given assignment of probabilities to individual outcomes is legitimate, that is, satisfies the rules of probability. 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 \( P(6) = 1/6 \).

4.11 Here are several assignments of probabilities to the six faces of a die:
We can learn which assignment is actually *accurate* for a particular die only by rolling the die many times. However, some of the assignments are not *legitimate* assignments of probability. That is, they do not obey the rules. Which are legitimate and which are not? In the case of the illegitimate models, explain what is wrong.

**4.12** Chose a student in grades 9 to 12 at random and ask if he or she is studying a language other than English. Here is the distribution of results:

<table>
<thead>
<tr>
<th>Language</th>
<th>Spanish</th>
<th>French</th>
<th>German</th>
<th>All others</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.26</td>
<td>0.09</td>
<td>0.03</td>
<td>0.03</td>
<td>0.59</td>
</tr>
</tbody>
</table>

(a) Explain why this is a legitimate probability model.
(b) What is the probability that a randomly chosen student is studying a language other than English?
(c) What is the probability that a randomly chosen student is studying French, German, or Spanish?

**4.13** If you draw an M&M candy at random from a bag of the candies, the candy you draw will have one of six colors. The probability of drawing each color depends on the proportion of each color among all candies made.

(a) The table below gives the probability of each color for a randomly chosen plain M&M:

<table>
<thead>
<tr>
<th>Color</th>
<th>Brown</th>
<th>Red</th>
<th>Yellow</th>
<th>Green</th>
<th>Orange</th>
<th>Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>?</td>
</tr>
</tbody>
</table>

What must be the probability of drawing a blue candy?

(b) The probabilities for peanut M&M’s are a bit different. Here they are:
What is the probability that a peanut M&M chosen at random is blue?

(c) What is the probability that a plain M&M is any of red, yellow, or orange?
What is the probability that a peanut M&M has one of these colors?

4.14 Wabash Red, when asked to predict the Big Ten Conference men’s basketball
champion, follows the modern practice of giving probabilistic predictions. He says,
“Michigan State has probability 0.3 of winning. Michigan, Minnesota, Northwestern,
and Penn State have no chance. That leaves 6 teams. Iowa, Illinois, and Purdue
all have the same probability of winning. Indiana, Ohio State, and Wisconsin also
have the same probability, but that probability is one-half that of the first 3.” What
probability does Red give to each of the 11 teams?

4.15 Choose an acre of land in Canada at random. The probability is 0.35 that it
is forest and 0.03 that it is pasture.
(a) What is the probability that the acre chosen is not forested?
(b) What is the probability that it is either forest or pasture?
(c) What is the probability that a randomly chosen acre in Canada is something
other than forest or pasture?

4.16 Choose a new car or light truck at random and note its color. Here are the
probabilities of the most popular colors for vehicles made in North America in 2003.
(From the Dupont Automotive North America Color Popularity Survey, reported
at www.dupont.com/automotive/.)

<table>
<thead>
<tr>
<th>Color</th>
<th>Silver</th>
<th>White</th>
<th>Black</th>
<th>Gray</th>
<th>Blue</th>
<th>Medium red</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.201</td>
<td>0.184</td>
<td>0.116</td>
<td>0.088</td>
<td>0.085</td>
<td>0.069</td>
</tr>
</tbody>
</table>

(a) What is the probability that the vehicle you choose has any color other than the
six listed?
(b) What is the probability that a randomly chosen vehicle is either silver or white?
(c) Choose two vehicles at random. What is the probability that both are silver or white?

4.17 A company that offers courses to prepare would-be MBA students for the GMAT examination has the following information about its customers: 20% are currently undergraduate students in business; 15% are undergraduate students in other fields of study; 60% are college graduates who are currently employed; and 5% are college graduates who are not employed.
(a) Is this a legitimate assignment of probabilities to customer backgrounds? Why?
(b) What percent of customers are currently undergraduates?

4.18 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?

4.19 The “pick four” games in many state lotteries announce a four-digit winning number each day. The winning number is essentially a four-digit group from a table of random digits. You win if your choice matches the winning digits. Suppose your chosen number is 5974.

(a) What is the probability that your number matches the winning number exactly?
(b) What is the probability that your number matches the digits in the winning number in any order?

4.20 Abby, Deborah, Mei-Ling, Sam, and Roberto work in a firm’s public relations office. Their employer must choose two of them to attend a conference in Paris. To avoid unfairness, the choice will be made by drawing two names from a hat. (This is an SRS of size 2.)

(a) Write down all possible choices of two of the five names. This is the sample space.
(b) The random drawing makes all choices equally likely. What is the probability of each choice?
(c) What is the probability that Mei-Ling is chosen?
(d) What is the probability that neither of the two men (Sam and Roberto) is chosen?

4.21 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?
4.22 An automobile manufacturer buys computer chips from a supplier. The supplier sends a shipment containing 5% defective chips. Each chip chosen from this shipment has probability 0.05 of being defective, and each automobile uses 12 chips selected independently. What is the probability that all 12 chips in a car will work properly?

4.23 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 a 3-year period. The lights fail independently of each other. What is the probability that the string of lights will remain bright for 3 years?

4.24 The most popular game of chance in Roman times was tossing four astragali. An astragalus is a small six-sided bone from the heel of an animal that comes to rest on one of four sides when tossed. (The other two sides are rounded.) The table gives the probabilities of the outcomes for a single astragalus based on modern experiments. The names “broad convex” etc. describe the four sides of the heel bone. The best throw was the “Venus,” with all four uppermost sides different. What is the probability of rolling a Venus? (From Florence N. David, *Games, Gods and Gambling*, Charles Griffin, London, 1962, p. 7.)

<table>
<thead>
<tr>
<th>Side</th>
<th>broad convex</th>
<th>broad concave</th>
<th>narrow flat</th>
<th>narrow hollow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.4</td>
<td>0.4</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

4.25 Government data show that 27% of employed people have at least 4 years of college and that 14% of employed people work as laborers or operators of machines or vehicles. Can you conclude that because $(0.27)(0.14) = 0.038$ about 3.8% of employed people are college-educated laborers or operators? Explain your answer.
Section 4.3

4.26 If a carefully made die is rolled once, it is reasonable to assign probability 1/6 to each of the six faces. What is the probability of rolling a number less than 3?

4.27 A couple plans to have three children. There are 8 possible arrangements of girls and boys. For example, GGB means the first two children are girls and the third child is a boy. All 8 arrangements are (approximately) equally likely.
(a) Write down all 8 arrangements of the sexes of three children. What is the probability of any one of these arrangements?
(b) Let $X$ be the number of girls the couple has. What is the probability that $X = 2$?
(c) Starting from your work in (a), find the distribution of $X$. That is, what values can $X$ take, and what are the probabilities for each value?

4.28 Choose an American household at random and let the random variable $X$ be the number of cars (including SUVs and light trucks) they own. Here is the probability model if we ignore the few households that own more than 5 cars:

<table>
<thead>
<tr>
<th>Number of cars $X$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.09</td>
<td>0.36</td>
<td>0.35</td>
<td>0.13</td>
<td>0.05</td>
<td>0.02</td>
</tr>
</tbody>
</table>

(a) Verify that this is a legitimate discrete distribution. Display the distribution in a probability histogram.
(b) Say in words what the event $\{X \geq 1\}$ is. Find $P(X \geq 1)$.
(c) A housing company builds houses with two-car garages. What percent of households have more cars than the garage can hold?

4.29 A study of social mobility in England looked at the social class reached by the sons of lower-class fathers. Social classes are numbered from 1 (low) to 5 (high). Take the random variable $X$ to be the class of a randomly chosen son of a father in Class 1. The study found that the distribution of $X$ is
Son’s class | 1 | 2 | 3 | 4 | 5  
Probability | 0.48 | 0.38 | 0.08 | 0.05 | 0.01

(a) What percent of the sons of lower-class fathers reach the highest class, Class 5?
(b) Check that this distribution satisfies the two requirements for a discrete probability distribution.
(c) What is $P(X \leq 3)$? (Be careful: the event “$X \leq 3$” includes the value 3.)
(d) What is $P(X < 3)$?
(e) Write the event “a son of a lower-class father reaches one of the two highest classes” in terms of values of $X$. What is the probability of this event?

4.30 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 discrete 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?

4.31 An SRS of 400 American adults is asked, “What do you think is the most serious problem facing our schools?” Suppose that in fact 30% of all adults would answer “drugs” if asked this question. The proportion $\hat{p}$ of the sample who answer
“drugs” will vary in repeated sampling. In fact, we can assign probabilities to values of \( \hat{p} \) using the normal density curve with mean 0.3 and standard deviation 0.023. Use this density curve to find the probabilities of the following events:

(a) At least half of the sample believes that drugs are the schools’ most serious problem.
(b) Less than 25% of the sample believes that drugs are the most serious problem.
(c) The sample proportion is between 0.25 and 0.35.

4.32 An opinion poll asks an SRS of 1500 adults, “Do you happen to jog?” Suppose that the population proportion who jog (a parameter) is \( p = 0.15 \). To estimate \( p \), we use the proportion \( \hat{p} \) in the sample who answer “Yes.” The statistic \( \hat{p} \) is a random variable that is approximately normally distributed with mean \( \mu = 0.15 \) and standard deviation \( \sigma = 0.0092 \). Find the following probabilities:

(a) \( P(\hat{p} \geq 0.16) \)
(b) \( P(0.14 \leq \hat{p} \leq 0.16) \)

Section 4.4

4.33 Exercise 4.28 gives the distribution of the number \( X \) of cars (including SUVs and light trucks) owned by American households. What is the average (mean) number of vehicles owned?

4.34 Keno is a favorite game in casinos, and similar games are popular with the states that operate lotteries. Balls numbered 1 to 80 are tumbled in a machine as the bets are placed, then 20 of the balls are chosen at random. Players select numbers by marking a card. The simplest of the many wagers available is “Mark 1 Number.” Your payoff is $3 on a $1 bet if the number you select is one of those chosen. Because 20 of 80 numbers are chosen, your probability of winning is 20/80, or 0.25.

(a) What is the probability distribution (the outcomes and their probabilities) of
Section 4.4

the payoff $X$ on a single play?

(b) What is the mean payoff $\mu_X$?

(c) In the long run, how much does the casino keep from each dollar bet?

4.35 (a) A gambler knows that red and black are equally likely to occur on each spin of a roulette wheel. He observes five consecutive reds and bets heavily on red at the next spin. Asked why, he says that “red is hot” and that the run of reds is likely to continue. Explain to the gambler what is wrong with this reasoning.

(b) After hearing you explain why red and black remain equally probable after five reds on the roulette wheel, the gambler moves to a poker game. He is dealt five straight red cards. He remembers what you said and assumes that the next card dealt in the same hand is equally likely to be red or black. Is the gambler right or wrong? Why?

4.36 In an experiment on the behavior of young children, each subject is placed in an area with five toys. The response of interest is the number of toys that the child plays with. Past experiments with many subjects have shown that the probability distribution of the number $X$ of toys played with is as follows:

<table>
<thead>
<tr>
<th>Number of toys $x_i$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability $p_i$</td>
<td>0.03</td>
<td>0.16</td>
<td>0.30</td>
<td>0.23</td>
<td>0.17</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Calculate the mean $\mu_X$ and the standard deviation $\sigma_X$.

4.37 You have two balanced, six-sided dice. The first has 1, 3, 4, 5, 6, and 8 spots on its six faces. The second die has 1, 2, 2, 3, 3, and 4 spots on its faces.

(a) What is the mean number of spots on the up-face when you roll each of these dice?

(b) Write the probability model for the outcomes when you roll both dice independently. From this, find the probability distribution of the sum of the spots on the up-faces of the two dice.

(c) Find the mean number of spots on the two up faces in two ways: from the dis-
tribution you found in (b) and by applying the addition rule to your results in (a). You should of course get the same answer.

4.38 Laboratory data show that the time required to complete two chemical reactions in a production process varies. The first reaction has a mean time of 40 minutes and a standard deviation of 2 minutes; the second has a mean time of 25 minutes and a standard deviation of 1 minute. The two reactions are run in sequence during production. There is a fixed period of 5 minutes between them as the product of the first reaction is pumped into the vessel where the second reaction will take place. What is the mean time required for the entire process?

4.39 The times for the two reactions in the chemical production process described in the previous exercise are independent. Find the standard deviation of the time required to complete the process.

4.40 The academic motivation and study habits of female students as a group are better than those of males. The Survey of Study Habits and Attitudes (SSHA) is a psychological test that measures these factors. The distribution of SSHA scores among the women at a college has mean 120 and standard deviation 28, and the distribution of scores among men students has mean 105 and standard deviation 35. You select a single male student and a single female student at random and give them the SSHA test.

(a) Explain why it is reasonable to assume that the scores of the two students are independent.

(b) What are the mean and standard deviation of the difference (female minus male) of their scores?

(c) From the information given, can you find the probability that the woman chosen scores higher than the man? If so, find this probability. If not, explain why you cannot.
4.41 The number of offspring produced by a female Asian stochastic beetle is random, with this pattern: 20% of females die without female offspring, 30% have one female offspring, and 50% have two female offspring. Females of the benign boiler beetle have this reproductive pattern: 40% die without female offspring, 40% have one female offspring, and 20% have two female offspring.

(a) Find the mean number of female offspring for each species of beetles.

(b) Use the law of large numbers to explain why the population should grow if the expected number of female offspring is greater than 1 and die out if this expected value is less than 1.

4.42 A study of the weights of the brains of Swedish men found that the weight $X$ was a random variable with mean 1400 grams and standard deviation 20 grams. Find positive numbers $a$ and $b$ such that $Y = a + bX$ has mean 0 and standard deviation 1.

4.43 In a process for manufacturing glassware, glass stems are sealed by heating them in a flame. The temperature of the flame varies a bit. Here is the distribution of the temperature $X$ measured in degrees Celsius:

<table>
<thead>
<tr>
<th>Temperature</th>
<th>540$^\circ$</th>
<th>545$^\circ$</th>
<th>550$^\circ$</th>
<th>555$^\circ$</th>
<th>560$^\circ$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.1</td>
<td>0.25</td>
<td>0.3</td>
<td>0.25</td>
<td>0.1</td>
</tr>
</tbody>
</table>

(a) Find the mean temperature $\mu_X$ and the standard deviation $\sigma_X$.

(b) The target temperature is 550$^\circ$ C. What are the mean and standard deviation of the number of degrees off target, $X - 550$?

(c) A manager asks for results in degrees Fahrenheit. The conversion of $X$ into degrees Fahrenheit is given by

$$Y = \frac{9}{5}X + 32$$

What are the mean $\mu_Y$ and standard deviation $\sigma_Y$ of the temperature of the flame in the Fahrenheit scale?
4.44 One consequence of the law of large numbers is that once we have a probability distribution for a random variable, we can find its mean by simulating many outcomes and averaging them. The law of large numbers says that if we take enough outcomes, their average value is sure to approach the mean of the distribution.

I have a little bet to offer you. Toss a coin ten times. If there is no run of three or more straight heads or tails in the ten outcomes, I’ll pay you $2. If there is a run of three or more, you pay me just $1. Surely you will want to take advantage of me and play this game?

Simulate enough plays of this game (the outcomes are +$2 if you win and −$1 if you lose) to estimate the mean outcome. Is it to your advantage to play?

4.45 You have two scales for measuring weights in a chemistry lab. Both scales give answers that vary a bit in repeated weighings of the same item. If the true weight of a compound is 2 grams (g), the first scale produces readings $X$ that have mean 2.000 g and standard deviation 0.002 g. The second scale’s readings $Y$ have mean 2.001 g and standard deviation 0.001 g.

(a) What are the mean and standard deviation of the difference $Y - X$ between the readings? (The readings $X$ and $Y$ are independent.)

(b) You measure once with each scale and average the readings. Your result is $Z = (X + Y)/2$. What are $\mu_Z$ and $\sigma_Z$? Is the average $Z$ more or less variable than the reading $Y$ of the less variable scale?

Section 4.5

4.46 Here is a two-way table of all suicides committed in a recent year by sex of the victim and method used.
<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firearms</td>
<td>15,802</td>
<td>2,367</td>
</tr>
<tr>
<td>Poison</td>
<td>3,262</td>
<td>2,233</td>
</tr>
<tr>
<td>Hanging</td>
<td>3,822</td>
<td>856</td>
</tr>
<tr>
<td>Other</td>
<td>1,571</td>
<td>571</td>
</tr>
<tr>
<td>Total</td>
<td>24,457</td>
<td>6,027</td>
</tr>
</tbody>
</table>

(a) What is the probability that a randomly selected suicide victim is male?

(b) What is the probability that the suicide victim used a firearm?

(c) What is the conditional probability that a suicide used a firearm, given that it was a man? Given that it was a woman?

(d) Describe in simple language (don’t use the word “probability”) what your results in (a) tell you about the difference between men and women with respect to suicide.

4.47 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 0.5, and that the probability of winning both jobs (event \{A and B\}) is 0.3. What is the probability of the event \{A or B\} that Consolidated will win at least one of the jobs?

4.48 In the setting of the previous exercise, are events $A$ and $B$ independent? Do a calculation that proves your answer.

4.49 Draw a Venn diagram that illustrates the relation between events $A$ and $B$ in the previous exercise. Write each of the following events in terms of $A$, $B$, $A^c$, and $B^c$. Indicate the events on your diagram and use the information in the previous exercise to calculate the probability of each.

(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.

4.50 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 $P(A) = 0.46$ and the probability of managerial and professional jobs among women is $P(B | A) = 0.32$. Find the probability that a randomly chosen employed person is a woman holding a managerial or professional position.

4.51 Common sources of caffeine in the diet 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?

4.52 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 probability that the supplier will demand that the contract be renegotiated is 0.8 if the dollar falls,
and 0.2 if the dollar does not fall. What is the probability that the supplier will demand renegotiation? (Use a tree diagram to organize the information given.)

*Here are data on the age and marital status of adult American women. Exercises 4.53 and 4.54 use this information.*

<table>
<thead>
<tr>
<th>Age</th>
<th>18 to 29</th>
<th>30 to 64</th>
<th>65 and over</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married</td>
<td>7,842</td>
<td>43,808</td>
<td>8,270</td>
<td>59,920</td>
</tr>
<tr>
<td>Never married</td>
<td>13,930</td>
<td>7,184</td>
<td>751</td>
<td>21,865</td>
</tr>
<tr>
<td>Widowed</td>
<td>36</td>
<td>2,523</td>
<td>8,385</td>
<td>10,944</td>
</tr>
<tr>
<td>Divorced</td>
<td>704</td>
<td>9,174</td>
<td>1,263</td>
<td>11,141</td>
</tr>
<tr>
<td>Total</td>
<td>22,512</td>
<td>62,689</td>
<td>18,669</td>
<td>103,870</td>
</tr>
</tbody>
</table>

4.53 Choose an adult American woman at random.

(a) What is the probability that the woman chosen is 65 years old or older?

(b) What is the conditional probability that the woman chosen is married, given that she is 65 or over?

(c) How many women are both married and in the 65 and over age group? What is the probability that the woman we choose is a married woman at least 65 years old?

(d) Verify that the three probabilities you found in (a), (b), and (c) satisfy the multiplication rule.

4.54 Choose an adult American woman at random.

(a) What is the conditional probability that the woman chosen is 18 to 29 years old, given that she is married?

(b) Verify that \( P(\text{married} \mid \text{age 18 to 29}) = 0.348 \). Complete this sentence: 0.348 is the proportion of women who are ________ among those women who are ________.
(c) In (a), you found \( P(\text{age 18 to 29 } | \text{ married}) \). Write a sentence of the form given in (b) that describes the meaning of this result. The two conditional probabilities give us very different information.

**4.55** A telemarketing company calls telephone numbers chosen at random. It finds that 70% of calls are not completed (the party does not answer or refuses to talk), that 20% result in talking to a woman, and that 10% result in talking to a man. After that point, 30% of the women and 20% of the men actually buy something. What percent of calls result in a sale?

**4.56** An examination consists of multiple-choice questions, each having five possible answers. Linda estimates that she has probability 0.75 of knowing the answer to any question that may be asked. If she does not know the answer, she will guess, with conditional probability 1/5 of being correct. What is the probability that Linda gives the correct answer to a question? (Draw a tree diagram to guide the calculation.)

**4.57** In the setting of Exercise 4.55, what percent of sales are made to women? (Write this as a conditional probability.)

**4.58** In the setting of Exercise 4.56, find the conditional probability that Linda knows the answer, given that she supplies the correct answer. (You can use the result of the previous Exercise and the definition of conditional probability, or you can use Bayes’s rule.)

**4.59** Zipdrive, Inc., has developed a new disk drive for small computers. The demand for the new product is uncertain but can be described as “high” or “low” in any one year. After 4 years, the product is expected to be obsolete. Management must decide whether to build a plant or to contract with a factory in Hong Kong to manufacture the new drive. Building a plant will be profitable if demand remains high but could lead to a loss if demand drops in future years.
After careful study of the market and of all relevant costs, Zipdrive’s planning office provides the following information. Let $A$ be the event that the first year’s demand is high, and $B$ be the event that the following 3 years’ demand is high. The marketing division’s best estimate of the probabilities is

\[ P(A) = 0.9 \]
\[ P(B \mid A) = 0.36 \]
\[ P(B \mid A^c) = 0 \]

The probability that building a plant is more profitable than contracting the production to Hong Kong is 0.95 if demand is high all 4 years, 0.3 if demand is high only in the first year, and 0.1 if demand is low all 4 years.

Draw a tree diagram that organizes this information. The tree will have three stages: first year’s demand, next 3 years’ demand, and whether building or contracting is more profitable. Which decision has the higher probability of being more profitable? (When decision analysis is used for investment decisions like this, firms in fact compare the mean profits rather than the probability of a profit. We ignore this complication.)

4.60 John has coronary artery disease. He and his doctor must decide between medical management of the disease and coronary bypass surgery. Because John has been quite active, he is concerned about his quality of life as well as the length of life. He wants to make the decision that will maximize the probability of the event $A$ that he survives for 5 years and is able to carry on moderate activity during that time. The doctor makes the following probability estimates for patients of John’s age and condition. (Based loosely on M. C. Weinstein, J. S. Pliskin, and W. B. Stason, “Coronary artery bypass surgery: decision and policy analysis,” in J. P. Bunker, B. A. Barnes, and F. W. Mosteller (eds.), Costs, Risks and Benefits of Surgery, Oxford University Press, 1977, pp. 342–371.)

- Under medical management, $P(A) = 0.7$. 
• There is probability 0.05 that John will not survive bypass surgery, probability 0.10 that he will survive with serious complications, and probability 0.85 that he will survive the surgery without complications.

• If he survives with complications, the conditional probability of the desired outcome $A$ is 0.73. If there are no serious complications, the conditional probability of $A$ is 0.76.

Draw a tree diagram that summarizes this information. Then calculate $P(A)$ assuming that John chooses the surgery. Does surgery or medical management offer him the better chance of achieving his goal?

Chapter 4 Review Exercises

4.61 Deal a five-card poker hand from a shuffled deck. The probabilities of several types of hand are approximately as follows:

<table>
<thead>
<tr>
<th>Hand</th>
<th>Worthless</th>
<th>One pair</th>
<th>Two pairs</th>
<th>Better hands</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.50</td>
<td>0.42</td>
<td>0.05</td>
<td>?</td>
</tr>
</tbody>
</table>

What must be the probability of getting a hand better than two pairs? What is the probability of getting a hand that is not worthless?

4.62 You have torn a tendon and are facing surgery to repair it. The orthopedic surgeon explains the risks to you. Infection occurs in 3% of such operations, the repair fails in 14%, and both infection and failure occur together in 1%. What percent of these operations succeed and are free from infection?

4.63 You are playing a board game in which the severity of a penalty is determined by rolling three dice and adding the spots on the up-faces. The dice are all balanced so that each face is equally likely, and the three dice fall independently. If $X_1$, $X_2$, and $X_3$ are the number of spots on the up-faces of the three dice, then $X = X_1 + X_2 + X_3$. Use this fact to find the mean $\mu_X$ and the standard deviation $\sigma_X$. 

without finding the distribution of \( X \). (Start with the distribution of each of the \( X_i \).)

4.64 Enzyme immunoassay (EIA) tests are used to screen blood specimens for the presence of antibodies to HIV, the virus that causes AIDS. Antibodies indicate the presence of the virus. The test is quite accurate but is not always correct. Here are approximate probabilities of positive and negative EIA outcomes when the blood tested does and does not actually contain antibodies to HIV. (J. Richard George, “Alternative specimen sources: methods for confirming positives,” 1998 Conference on the Laboratory Science of HIV, found online at the Centers for Disease Control and Prevention, www.cdc.gov.)

<table>
<thead>
<tr>
<th>Test result</th>
<th>Antibodies present</th>
<th>Antibodies absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>0.9985</td>
<td>0.0060</td>
</tr>
<tr>
<td>−</td>
<td>0.0015</td>
<td>0.9940</td>
</tr>
</tbody>
</table>

Suppose that 1% of a large population carries antibodies to HIV in their blood.

(a) Draw a tree diagram for selecting a person from this population (outcomes: antibodies present or absent) and for testing his or her blood (outcomes: EIA positive or negative).

(b) What is the probability that the EIA is positive for a randomly chosen person from this population?

(c) What is the probability that a person has the antibody, given that the EIA test is positive?

(This exercise illustrates a fact that is important when considering proposals for widespread testing for HIV, illegal drugs, or agents of biological warfare: if the condition being tested is uncommon in the population, many positives will be false positives.)
4.65 The previous exercise gives data on the results of EIA tests for the presence of antibodies to HIV. Repeat part (c) of that exercise for two different populations:

(a) Blood donors are prescreened for HIV risk factors, so perhaps only 0.1% (0.001) of this population carries HIV antibodies.

(b) Clients of a drug rehab clinic are a high-risk group, so perhaps 10% of this population carries HIV antibodies.

(c) What general lesson do your calculations illustrate?

CHAPTER 5

Section 5.1

5.1 For each of the following situations, indicate whether a binomial distribution is a reasonable probability model for the random variable $X$. Give your reasons in each case.

(a) You observe the sex of the next 50 children born at a local hospital; $X$ is the number of girls among them.

(b) A couple decides to continue to have children until their first girl is born; $X$ is the total number of children the couple has.

(c) You want to know what percent of married people believe that mothers of young children should not be employed outside the home. You plan to interview 50 people, and for the sake of convenience you decide to interview both the husband and the wife in 25 married couples. The random variable $X$ is the number among the 50 persons interviewed who think mothers should not be employed.

5.2 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 student studies binomial distributions using computer-assisted instruction. After the initial instruction is completed, the computer presents 10 problems. The student solves each problem and enters the answer; the computer gives additional instruction between problems if the student’s answer is wrong. The number of problems that the student solves correctly is counted.

(c) A chemist repeats a solubility test 10 times on the same substance. Each test is conducted at a temperature 10° higher than the previous test. She counts the number of times that the substance dissolves completely.

5.3 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?

5.4 A university that is better known for its basketball program than for its academic strength 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 6 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 who graduate among the 20 studied should have the $B(20, 0.8)$ distribution.

(a) Find the probability that exactly 11 of the 20 players graduate.

(b) Find the probability that 11 or fewer players graduate. This probability is so small that it casts doubt on the university’s claim.

5.5 (a) What is the mean number of Hispanics on randomly chosen committees of 15 workers in Exercise 5.3?
(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?

5.6 (a) Find the mean number of graduates out of 20 players in the setting of Exercise 5.4 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?

5.7 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 $B(150, 0.5)$ distribution for the number $X$ who respond. Explain why.

(b) What is the expected number (the mean) who will respond?

(c) What is the probability that 70 or fewer will respond? (Use software or the normal approximation.)

(d) How large a sample must you take to increase the mean number of respondents to 100?

5.8 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) What is the sample proportion of orders shipped on time?

(b) If the company really ships 90% of its orders on time, what is the probability
that the proportion in an SRS of 100 orders is as small as the proportion in your sample or smaller? (Use software or the normal approximation.)

(c) A critic says, “Aha! You claim 90%, but in your sample the on-time percentage is lower than that. So the 90% claim is wrong.” Explain in simple language why your probability calculation in (b) shows that the result of the sample does not refute the 90% claim.

5.9 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, it is reasonable to use the $B(200, 0.4)$ distribution to describe the count $X$ of respondents who seek nutritious food when eating out. Explain why.

(b) What is the mean number of nutrition-conscious people in your sample if $p = 0.4$ is true? What is the probability that $X$ lies between 75 and 85?

(c) 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.

5.10 “How would you describe your own physical health at this time? Would you say your physical health is—excellent, good, only fair, or poor?” The Gallup Poll asked this question of 1005 randomly selected adults, of whom 29% said “excellent.” (David W. Moore and Joseph Carroll, “Most Americans call their physical and mental health ‘good’ or ‘excellent,’” www.gallup.com/poll/releases/, November 28, 2001.) Suppose that in fact the proportion of the adult population who say their health is excellent is $p = 0.29$. 
(a) What is the probability that the sample proportion \( \hat{p} \) of an SRS of size \( n = 1000 \) who say their health is excellent lies between 26% and 32%? (That is, within ±3% of the truth about the population.)

(b) Repeat the probability calculation of (a) for SRSs of sizes \( n = 250 \) and \( n = 4000 \). What general conclusion can you draw from your calculations?

5.11 “How would you describe your own personal weight situation right now—very overweight, somewhat overweight, about right, somewhat underweight, or very underweight?” When the Gallup Poll asked an SRS of 1005 adults this question, 51% answered “about right.” (Lydia Saad, “Hold the gravy! Six in ten Americans want to lose weight,” www.gallup.com/poll/releases/, November 21, 2001.) Suppose that in fact 51% of the entire adult population think their weight is about right.

(a) Many opinion polls announce a “margin of error” of about ±3%. What is the probability that an SRS of size 1005 has a sample proportion \( \hat{p} \) that is within ±3% (±0.03) of the population proportion \( p = 0.51 \)?

(b) Answer the same question if the population proportion is \( p = 0.06 \). (This is the proportion who told Gallup that they were “very overweight.”) How does the probability change as \( p \) moves from near 0.5 to near zero?

5.12 A student organization is planning to ask a sample of 50 students if they have noticed alcohol abuse brochures on campus. The sample percentage who say “Yes” will be reported. The organization’s statistical advisor says that the standard deviation of this percentage will be about 7%.

(a) What would the standard deviation be if the sample contained 100 students rather than 50?

(b) How large a sample is required to reduce the standard deviation of the percentage who say “Yes” from 7% to 3.5%? Explain to someone who knows no statistics the advantage of taking a larger sample in a survey of opinion.
5.13 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. Suppose that in fact the population proportion who feel this way is \( p = 0.44 \).

(a) Many opinion polls have a “margin of error” of about ±3%. What is the probability that an SRS of size 300 has a sample proportion \( \hat{p} \) that is within ±3% of the population proportion \( p = 0.44 \)? (Use software or the normal approximation.)

(b) Answer the same question for SRSs of sizes 600 and 1200. What is the effect of increasing the size of the sample?

5.14 A sociology professor asks her class to observe cars having a man and a woman in the front seat and record which of the two is the driver.

(a) Explain why it is reasonable to use the binomial distribution for the number of male drivers in \( n \) cars if all observations are made in the same location at the same time of day.

(b) Explain why the binomial model may not apply if half the observations are made outside a church on Sunday morning and half are made on campus after a dance.

(c) The professor requires students to observe 10 cars during business hours in a retail district close to campus. Past observations have shown that the man is driving in about 85% of cars in this location. What is the probability that the man is driving in 8 or fewer of the 10 cars?

(d) The class has 10 students, who will observe 100 cars in all. What is the probability that the man is driving in 80 or fewer of these?

5.15 A study by a federal agency concludes that polygraph (lie detector) tests given to truthful persons have probability about 0.2 of suggesting that the person is deceptive. (Office of Technology Assessment, *Scientific Validity of Polygraph Testing: A Research Review and Evaluation*, Government Printing Office, 1983.)

(a) A firm asks 12 job applicants about thefts from previous employers, using a
polygraph to assess their truthfulness. Suppose that all 12 answer truthfully. What is the probability that the polygraph 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?

Section 5.2

5.16 A company that owns and services a fleet of cars for its sales force has found that the service lifetime of disc brake pads varies from car to car according to a normal distribution with mean $\mu = 55,000$ miles and standard deviation $\sigma = 4500$ miles. The company installs a new brand of brake pads on 8 cars.

(a) If the new brand has the same lifetime distribution as the previous type, what is the distribution of the sample mean lifetime for the 8 cars?

(b) The average life of the pads on these 8 cars turns out to be $\bar{x} = 51,800$ miles. What is the probability that the sample mean lifetime is 51,800 miles or less if the lifetime distribution is unchanged? The company takes this probability as evidence that the average lifetime of the new brand of pads is less than 55,000 miles.

5.17 Investors remember 1987 as the year stocks lost 22% of their value in a single day. For 1987 as a whole, the mean return of all common stocks on the New York Stock Exchange was $\mu = -3.5\%$. (That is, these stocks lost an average of 3.5% of their value in 1987.) The standard deviation of the returns was about $\sigma = 26\%$. The distribution of annual returns for stocks is roughly normal.

(a) What percent of stocks lost money? (That is the same as the probability that a stock chosen at random has a return less than 0.)

(b) Suppose that you held a portfolio of 5 stocks chosen at random from New York Stock Exchange stocks. What are the mean and standard deviation of the returns
of randomly chosen portfolios of 5 stocks?

(c) What percent of such portfolios lost money? Explain the difference between this result and the result of (a).

5.18 Newly manufactured automobile radiators may have small leaks. Most have no leaks, but some have 1, 2, or more. The number of leaks in radiators made by one supplier has mean 0.15 and standard deviation 0.4. The distribution of number of leaks cannot be normal because only whole-number counts are possible. The supplier ships 400 radiators per day to an auto assembly plant. Take \( \bar{x} \) to be the mean number of leaks in these 400 radiators. Over several years of daily shipments, what range of values will contain the middle 95% of the many \( \bar{x} \)'s?

5.19 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 \( L \)?

5.20 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?

5.21 The scores of 12th-grade students on the National Assessment of Educational Progress year 2000 mathematics test have a distribution that is approximately normal with mean $\mu = 300$ and standard deviation $\sigma = 35$.

(a) Choose one 12th-grader at random. What is the probability that his or her score is higher than 300? Higher than 335?

(b) Now choose an SRS of four 12th-graders. What is the probability that their mean score is higher than 300? Higher than 335?

5.22 The number of accidents per week at a hazardous intersection varies with mean 2.2 and standard deviation 1.4. This distribution takes only whole-number values, so it is certainly not normal.

(a) Let $\bar{x}$ be the mean number of accidents per week at the intersection during a year (52 weeks). What is the approximate distribution of $\bar{x}$ according to the central limit theorem?

(b) What is the approximate probability that $\bar{x}$ is less than 2?

(c) What is the approximate probability that there are fewer than 100 accidents at the intersection in a year? (*Hint: Restate this event in terms of $\bar{x}$.*)

5.23 A roulette wheel has 38 slots, of which 18 are black, 18 are red, and 2 are green. When the wheel is spun, the ball is equally likely to come to rest in any of the slots. Gamblers can place a number of different bets in roulette. One of the simplest wagers chooses red or black. A bet of $1 on red will pay off an additional dollar if the ball lands in a red slot. Otherwise, the player loses his dollar. When gamblers bet on red or black, the two green slots belong to the house.

(a) A gambler’s winnings on a $1 bet are either $1 or $-1. Give the probabilities of these outcomes. Find the mean and standard deviation of the gambler’s winnings.

(b) Explain briefly what the law of large numbers tells us about what will happen
if the gambler makes a large number of bets on red.

(c) The central limit theorem tells us the approximate distribution of the gambler’s mean winnings in 50 bets. What is this distribution? Use the 68–95–99.7 rule to give the range in which the mean winnings will fall 95% of the time. Multiply by 50 to get the middle 95% of the distribution of the gambler’s winnings on nights when he places 50 bets.

(d) What is the probability that the gambler will lose money if he makes 50 bets? (This is the probability that the mean is less than 0.)

(e) The casino takes the other side of these bets. If 100,000 bets on red are placed in a week at the casino, what is the distribution of the mean winnings of gamblers on these bets? What range covers the middle 95% of mean winnings in 100,000 bets? Multiply by 100,000 to get the range of gamblers’ losses. (Gamblers’ losses are the casino’s winnings. Part (c) shows that a gambler gets excitement. Now we see that the casino has a business.)

5.24 An experiment to compare the nutritive value of normal corn and high-lysine corn divides 40 chicks at random into two groups of 20. One group is fed a diet based on normal corn while the other receives high-lysine corn. At the end of the experiment, inference about which diet is superior is based on the difference $\bar{y} - \bar{x}$ between the mean weight gain $\bar{y}$ of the 20 chicks in the high-lysine group and the mean weight gain $\bar{x}$ of the 20 in the normal-corn group. Because of the randomization, the two sample means are independent.

(a) Suppose that $\mu_X = 360$ grams (g) and $\sigma_X = 55$ g in the population of all chicks fed normal corn, and that $\mu_Y = 385$ g and $\sigma_Y = 50$ g in the high-lysine population. What are the mean and standard deviation of $\bar{y} - \bar{x}$?

(b) The weight gains are normally distributed in both populations. What is the distribution of $\bar{x}$? Of $\bar{y}$? What is the distribution of $\bar{y} - \bar{x}$?
Chapter 5 Exercises

(c) What is the probability that the mean weight gain in the high-lysine group exceeds the mean weight gain in the normal-corn group by 25 g or more?

5.25 An experiment on the teaching of reading compares two methods, A and B. The response variable is the Degree of Reading Power (DRP) score. The experimenter uses Method A in a class of 26 students and Method B in a comparable class of 24 students. The classes are assigned to the teaching methods at random. Suppose that in the population of all children of this age the DRP score has the $N(34, 12)$ distribution if Method A is used and the $N(37, 11)$ distribution if Method B is used.

(a) What is the distribution of the mean DRP score $\bar{x}$ for the 26 students in the A group? (Assume that this group can be regarded as an SRS from the population of all children of this age.)

(b) What is the distribution of the mean score $\bar{y}$ for the 24 students in the B group?

(c) Use the results of (a) and (b), keeping in mind that $\bar{x}$ and $\bar{y}$ are independent, to find the distribution of the difference $y - x$ between the mean scores in the two groups.

(d) What is the probability that the mean score for the B group will be at least 4 points higher than the mean score for the A group?

5.26 Leona and Fred are friendly competitors in high school. Both are about to take the ACT college entrance examination. They agree that if one of them scores 5 or more points better than the other, the loser will buy the winner a pizza. Suppose that in fact Fred and Leona have equal ability, so that each score varies normally with mean 24 and standard deviation 2. (The variation is due to luck in guessing and the accident of the specific questions being familiar to the student.) The two scores are independent. What is the probability that the scores differ by 5 or more points in either direction?

5.27 The design of an electronic circuit calls for a 100-ohm resistor and a 250-ohm resistor connected in series so that their resistances add. The components used are
not perfectly uniform, so that the actual resistances vary independently according
to normal distributions. The resistance of 100-ohm resistors has mean 100 ohms and
standard deviation 2.5 ohms, while that of 250-ohm resistors has mean 250 ohms
and standard deviation 2.8 ohms.
(a) What is the distribution of the total resistance of the two components in series?
(b) What is the probability that the total resistance lies between 345 and 355 ohms?

5.28 ACT, Inc., the producer of the ACT test of readiness for college work, also
produces tests for 8th- and 9th-grade students designed to help them plan for the
future. Two of these tests measure reading and mathematics achievement. Each
has scores that range from 1 to 25. For students tested in the fall of their 8th-
grade year, the reading test has mean 13.9 and standard deviation 3.63. The mean
score on the math test is 14.4 and the standard deviation is 3.46. Scores roughly
follow a normal distribution. (Based on data from over 7000 students, reported at
www.act.org/explore/newscale/summary.html.)
(a) If a student’s reading score $X$ and mathematics score $Y$ were independent, what
would be the distribution of total $X + Y$?
(b) Using the distribution you found in (a), what percent of 8th-graders have a total
score of 30 or higher?
(c) In fact, the $X$ and $Y$ scores are strongly correlated. In this case, does the mean
of $X + Y$ still have the value you found in (a)? Does the standard deviation still
have the value you found in (a)?

Chapter 5 Review Exercises

5.29 An opinion poll asks a sample 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 at least half of the sample are in favor? (Assume that the sample is an SRS.)

5.30 A political activist is gathering signatures on a petition by going door-to-door asking citizens to sign. She wants 100 signatures. Suppose that the probability of getting a signature at each household is 1/10 and let the random variable $X$ be the number of households visited to collect exactly 100 signatures. Is it reasonable to use a binomial distribution for $X$? If so, give $n$ and $p$. If not, explain why not.

5.31 High school dropouts make up 12.1% of all Americans aged 18 to 24. A vocational school that wants to attract dropouts mails an advertising flyer to 25,000 persons between the ages of 18 and 24.

(a) If the mailing list can be considered a random sample of the population, what is the mean number of high school dropouts who will receive the flyer? What is the standard deviation of this number?

(b) What is the probability that at least 3500 dropouts will receive the flyer?

5.32 According to a market research firm, 52% of all residential telephone numbers in Los Angeles are unlisted. A telemarketing company uses random digit dialing equipment that dials residential numbers at random, regardless of whether they are listed in the telephone directory. The firm calls 500 numbers in Los Angeles.

(a) What is the exact distribution of the number $X$ of unlisted numbers that are called?

(b) Use a suitable approximation to calculate the probability that at least half of the numbers called are unlisted.

5.33 The level of nitrogen oxides (NOX) in the exhaust of a particular car model varies with mean 0.9 grams per mile (g/mi) and standard deviation 0.15 g/mi. A company has 125 cars of this model in its fleet.

(a) What is the approximate distribution of the mean NOX emission level $\overline{x}$ for
these cars?

(b) What is the level \( L \) such that the probability that \( \bar{\pi} \) is greater than \( L \) is only 0.01?

5.34 The World Health Organization MONICA Project collected health information from random samples of adults in many nations. One question asked was, “Have you had your blood pressure measured in the past year?” Of 186 American males aged 25 to 34, 153 said “Yes.” For females in the same age group, 235 of the sample of size 248 said “Yes.” (From the Web site of the Finnish National Public Health Institute, www.ktl.fi/monica/.) Let us suppose that for the entire population in this age group, 82% of men and 95% of women have had their blood pressure measured in the past year.

(a) What is the approximate distribution of the proportion \( \hat{p}_1 \) of “Yes” responses in an SRS of 186 men? Of the corresponding proportion \( \hat{p}_2 \) for an SRS of 248 women?

(b) The samples of women and men are of course independent. What is the approximate distribution of the difference \( \hat{p}_2 - \hat{p}_1 \)?

(c) What is the approximate probability that the female proportion exceeds the male proportion by at least 10 percentage points?

5.35 The distribution of scores for persons over 16 years of age on the Wechsler Adult Intelligence Scale (WAIS) is approximately normal with mean 100 and standard deviation 15. The WAIS is one of the most common “IQ tests” for adults.

(a) What is the probability that a randomly chosen individual has a WAIS score of 105 or higher?

(b) What are the mean and standard deviation of the average WAIS score \( \bar{x} \) for an SRS of 60 people?

(c) What is the probability that the average WAIS score of an SRS of 60 people is 105 or higher?
(d) Would your answers to any of (a), (b), or (c) be affected if the distribution of WAIS scores in the adult population were distinctly nonnormal?

5.36 The study habits portion of the Survey of Study Habits and Attitudes (SSHA) psychological test consists of two sets of questions. One set measures “delay avoidance” and the other measures “work methods.” A subject’s study habits score is the sum $X + Y$ of the delay avoidance score $X$ and the work methods score $Y$. The distribution of $X$ in a broad population of first-year college students is close to $N(25, 10)$, while the distribution of $Y$ in the same population is close to $N(25, 9)$.

(a) If a subject’s $X$ and $Y$ scores were independent, what would be the distribution of the study habits score $X + Y$?

(b) Using the distribution you found in (a), what percent of the population have a study habits score of 60 or higher?

(c) In fact, the $X$ and $Y$ scores are strongly correlated. In this case, does the mean of $X + Y$ still have the value you found in (a)? Does the standard deviation still have the value you found in (a)?

CHAPTER 6

Section 6.1

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

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

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

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

(b) The poll shows Ringel leading. But the newspaper article said that the election was too close to call. Explain why.

6.5 A student reads that a 95% confidence interval for the mean SAT math score of California high school seniors is 452 to 470. Asked to explain the meaning of this interval, the student says, “95% of California high school seniors have SAT math scores between 452 and 470.” Is the student right? Justify your answer.
6.6 As we prepare to take a sample and compute a 95% confidence interval, we know that the probability that the interval we compute will cover the parameter is 0.95. That’s the meaning of 95% confidence. If we use several such intervals, however, our confidence that all give correct results is less than 95%.

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

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

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

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

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

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

(a) What was the response rate for this survey? (The response rate is the percent of the planned sample that responded.)

(b) Do you think that the small margin of error is a good measure of the accuracy of the survey’s results? Explain your answer.

6.12 The mean amount $\mu$ for all of the invoices for your company last month is not known. Based on your past experience, you are willing to assume that the standard deviation of invoice amounts is about $200.

(a) If you take a random sample of 100 invoices, what is the value of the standard deviation for $\bar{x}$?

(b) The 68–95–99.7 rule says that the probability is about 0.95 that $\bar{x}$ is within _____ of the population mean $\mu$. Fill in the blank.

(c) About 95% of all samples will capture the true mean of all of the invoices in the interval $\bar{x}$ plus or minus ______. Fill in the blank.

6.13 You measure the weights of 24 male runners. You do not actually choose an SRS, but you are willing to assume that these runners are a random sample from the population of male runners in your town or city. Here are their weights in kilograms:
Suppose that the standard deviation of the population is known to be \( \sigma = 4.5 \text{ kg} \).

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

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

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

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

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

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

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

\[
\begin{array}{cccccccccc}
138.0 & 139.1 & 113.0 & 132.5 & 140.7 & 109.7 & 118.9 & 134.8, \\
109.6 & 127.3 & 115.6 & 130.4 & 130.2 & 111.7 & 105.5 \\
\end{array}
\]

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

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

(b) Find the 95% confidence interval.

(c) Find the 99% confidence interval.

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

6.16 Suppose that the crop researchers in the previous exercise obtained the same value of \( \bar{x} \) from a sample of 50 plots rather than 15.
Section 6.1

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

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

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

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

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

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

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

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

6.20 A study of the career paths of hotel general managers sent questionnaires to an SRS of 160 hotels belonging to major U.S. hotel chains. There were 114 responses. The average time these 114 general managers had spent with their current company was 11.78 years. Give a 99% confidence interval for the mean number of years general managers of major-chain hotels have spent with their current company. (Take it as known that the standard deviation of time with the company for all general managers is 3.2 years.)
6.21 Researchers planning a study of the reading ability of third-grade children want to obtain a 95% confidence interval for the population mean score on a reading test, with margin of error no greater than 3 points. They carry out a small pilot study to estimate the variability of test scores. The sample standard deviation is $s = 12$ points in the pilot study, so in preliminary calculations the researchers take the population standard deviation to be $\sigma = 12$.

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

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

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

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

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

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

(b) Explain to someone who knows no statistics what the announced result 30\% \pm 3\% means.

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

$$\text{estimate} \pm z^* \sigma_{\text{estimate}}$$

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

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

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

$$\text{estimate} \pm z^* \sigma_{\text{estimate}}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

6.29 A radio talk show invites listeners to enter a dispute about a proposed pay increase for city council members. “What yearly pay do you think council members should get? Call us with your number.” In all, 958 people call. The mean pay they suggest is $\bar{x} = 9740$ 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 $9669$ to $9811$. Is this result trustworthy? Explain your answer.
Section 6.2

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

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

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

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

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

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

(b) The examinations in a large accounting class are scaled after grading so that the mean score is 50. A self-confident teaching assistant thinks that his students have a higher mean score than the class as a whole. His students this semester can be considered a sample from the population of all students he might teach, so he
Section 6.2

compares their mean score with 50.

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

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

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

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

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

(b) Would the hypothesis that \( \mu = 5 \) mg be rejected at the 5% level in favor of a two-sided alternative?
6.35 Market pioneers, companies that are among the first to develop a new product or service, tend to have higher market shares than latecomers to the market. What accounts for this advantage? Here is an excerpt from the conclusions of a study of a sample of 1209 manufacturers of industrial goods:

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

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

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

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

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

(c) You purchase a shipment of 60-watt light bulbs to be used in a variety of your products. If the wattage is too low or too high, your product will not look good. You measure the wattage of a random sample of 20 bulbs.

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

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

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

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

6.38 In each of the following situations, state an appropriate null hypothesis $H_0$ and alternative hypothesis $H_a$. Be sure to identify the parameters that you use to state the hypotheses. (We have not yet learned how to test these hypotheses.)
(a) A sociologist asks a large sample of high school students which academic subject they like best. She suspects that a higher percent of males than of females will name mathematics as their favorite subject.

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

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

6.39 A test of the null hypothesis $H_0: \mu = \mu_0$ gives test statistic $z = 1.8$.

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

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

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

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

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

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

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

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

(b) Can you reject the null hypothesis that $\mu = 36$ at the 5% significance level? Why?
6.42 A new supplier offers a good price on a catalyst used in your production process. You compare the purity of this catalyst with that from your current supplier. The $P$-value for a test of “no difference” is 0.15. Can you be confident that the purity of the new product is the same as the purity of the product that you have been using? Discuss.

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

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

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

6.44 A randomized comparative experiment examined whether a calcium supplement in the diet reduces the blood pressure of healthy men. The subjects received either a calcium supplement or a placebo for 12 weeks. The statistical analysis was quite complex, but one conclusion was that “the calcium group had lower seated systolic blood pressure ($P = 0.008$) compared with the placebo group.” (R. M. Lyle et al., “Blood pressure and metabolic effects of calcium supplementation in normotensive white and black men,” *Journal of the American Medical Association*, 257 (1987), pp. 1772–1776.) Explain this conclusion, especially the $P$-value, as if you were speaking to a doctor who knows no statistics.
6.45 A social psychologist reports that “ethnocentrism was significantly higher ($P < 0.05$) among church attenders than among nonattenders.” Explain what this means in language understandable to someone who knows no statistics. Do not use the word “significance” in your answer.

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

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

(b) Do you reject the null hypothesis?

(c) What is your conclusion?

(d) What other facts about the study would you like to know for a proper interpretation of the results?

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

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

\[ H_0: \mu = 120 \]

\[ H_a: \mu \neq 120 \]

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

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

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

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

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

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

(b) Is the result significant at the 1% level?
6.51 You measure the weights of 24 male runners. These runners are not a random sample from a population, but you are willing to assume that their weights represent the weights of similar runners. Here are their weights in kilograms:

\[
\begin{array}{cccccccc}
67.8 & 61.9 & 63.0 & 53.1 & 62.3 & 59.7 & 55.4 & 58.9 \\
60.9 & 69.2 & 63.7 & 68.3 & 64.7 & 65.6 & 56.0 & 57.8 \\
66.0 & 62.9 & 53.6 & 65.0 & 55.8 & 60.4 & 69.3 & 61.7 \\
\end{array}
\]

Exercise 6.13 asks you to find a 95% confidence interval for the mean weight of the population of all such runners, assuming that the population standard deviation is \( \sigma = 4.5 \) kg.

(a) Give the confidence interval from that exercise, or calculate the interval if you did not do the exercise.

(b) Based on this confidence interval, does a test of

\[ H_0: \mu = 61.3 \text{ kg} \]

\[ H_a: \mu \neq 61.3 \text{ kg} \]

reject \( H_0 \) at the 5% significance level?

(c) Would \( H_0: \mu = 63 \) be rejected at the 5% level if tested against a two-sided alternative?

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

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

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

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

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

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

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

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

(b) If the farmer is simply guessing, what is the distribution of the number $X$ of
correct answers in 5 tries?
(c) The observed outcome is \( X = 4 \). What is the \( P \)-value of the test that takes large values of \( X \) to be evidence against \( H_0 \)?

Section 6.3

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

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

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

in each of the following situations:
(a) The students’ average score is \( \bar{x} = 541.4 \). Is this result significant at the 5% level?

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

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

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

(b) Why is it important that the effects were small in size as well as insignificant?

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

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

(a) Is it proper to conclude that these four people have ESP? Explain your answer.

(b) What should the researcher now do to test whether any of these four subjects have ESP?

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

(a) What is the distribution of the number $X$ of tests that are significant?

(b) Find the probability that 2 or more of the tests are significant.
Section 6.4

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

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

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

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

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

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

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

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

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

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

Power calculations help us see how large a shortfall in the bottle contents the test can be expected to detect.
(a) Find the power of this test against the alternative $\mu = 299$.

(b) Find the power against the alternative $\mu = 295$.

(c) Is the power against $\mu = 290$ higher or lower than the value you found in (b)? Explain why this result makes sense.

**6.64** You have an SRS of size $n = 9$ from a normal distribution with $\sigma = 1$. You wish to test

$$H_0: \mu = 0$$

$$H_a: \mu > 0$$

You decide to reject $H_0$ if $x > 0$ and to accept $H_0$ otherwise.

(a) Find the probability of a Type I error, that is, the probability that your test rejects $H_0$ when in fact $\mu = 0$.

(b) Find the probability of a Type II error when $\mu = 0$. This is the probability that your test accepts $H_0$ when in fact $\mu = 0$.

(c) Find the probability of a Type II error when $\mu = 1$.

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

**6.66 (Optional)** An acceptance sampling test has probability 0.05 of rejecting a good lot of bearings and probability 0.08 of accepting a bad lot. The consumer of the bearings may imagine that acceptance sampling guarantees that most accepted lots are good. Alas, it is not so. Suppose that 90% of all lots shipped by the producer are bad.

(a) Draw a tree diagram for shipping a lot (the branches are “bad” and “good”) and then inspecting it (the branches at this stage are “accept” and “reject”).

(b) Write the appropriate probabilities on the branches, and find the probability that a lot shipped is accepted.

(c) Use the definition of conditional probability or Bayes’s formula to find the prob-
ability that a lot is bad, given that the lot is accepted. This is the proportion of 
bad lots among the lots that the sampling plan accepts.

Chapter 6 Review Exercises

6.67 A study compares two groups of mothers with young children who were on 
welfare two years ago. One group attended a voluntary training program offered 
free of charge at a local vocational school and advertised in the local news media. 
The other group did not choose to attend the training program. The study finds 
a significant difference \( (P < 0.01) \) between the proportions of the mothers in the 
two groups who are still on welfare. The difference is not only significant but quite 
large. The report says that with 95% confidence the percent of the nonattending 
group still on welfare is 21% \( \pm \) 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 briefly and in nontechnical language what “a significant difference \( (P < 0.01) \)” means.

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

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

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

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

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

(a) Simulate the drawing of 25 SRSs of size $n = 100$ from this population.

(b) The 95% confidence interval for the population mean $\mu$ has the form $\bar{x} \pm m$. What is the margin of error $m$? (Remember that we know $\sigma = 100$.)

(c) Use your software to calculate the 95% confidence interval for $\mu$ when $\sigma = 100$ for each of your 25 samples. Verify the computer’s calculations by checking the interval given for the first sample against your result in (b). Use the $\bar{x}$ reported by the software.

(d) How many of the 25 confidence intervals contain the true mean $\mu = 460$? If you repeated the simulation, would you expect exactly the same number of intervals to contain $\mu$? In a very large number of samples, what percent of the confidence intervals would contain $\mu$?

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

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

$$H_0: \mu = 460$$

$$H_a: \mu \neq 460$$
for each of the 25 samples.

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

(c) How many of your 25 tests reject the null hypothesis at the $\alpha = 0.05$ significance level? (That is, how many have $P$-value 0.05 or smaller?) Because the simulation was done with $\mu = 460$, samples that lead to rejecting $H_0$ produce the wrong conclusion. In a very large number of samples, what percent would falsely reject the hypothesis?

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

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

(b) Repeat on these new data the test of

\[ H_0: \mu = 460 \]

\[ H_a: \mu \neq 460 \]

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

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

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

6.75 Statisticians prefer large samples. Describe briefly the effect of increasing the size of a sample (or the number of subjects in an experiment) on each of the following:

(a) The width of a level $C$ confidence interval.

(b) The $P$-value of a test, when $H_0$ is false and all facts about the population remain unchanged as $n$ increases.

(c) The power of a fixed level $\alpha$ test, when $\alpha$, the alternative hypothesis, and all facts about the population remain unchanged.

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

6.77 The text demonstrates the behavior of a confidence interval in repeated sampling by showing the results of 25 samples from the same population. Now you will do a similar demonstration. Suppose that (unknown to the researcher) the mean SATM score of all California high school seniors is $\mu = 475$, and that the standard deviation is known to be $\sigma = 100$. The scores vary normally.

(a) Simulate the drawing of 50 SRSs of size $n = 100$ from this population.

(b) The 95% confidence interval for the population mean $\mu$ has the form $\bar{x} \pm m$. What is the margin of error $m$? (Remember that we know $\sigma = 100$.)
(c) Use your software to calculate the 95% confidence interval for $\mu$ when $\sigma = 100$ for each of your 50 samples. Verify the computer’s calculations by checking the interval given for the first sample against your result in (b). Use the $\bar{x}$ reported by the software.

(d) How many of the 50 confidence intervals contain the true mean $\mu = 475$? If you repeated the simulation, would you expect exactly the same number of intervals to contain $\mu$? In a very large number of samples, what percent of the confidence intervals would contain $\mu$?

6.78 In the previous exercise you simulated the SATM scores of 50 SRSs of 100 California seniors. Now use these samples to demonstrate the behavior of a significance test. We know that the population of all SATM scores is normal with standard deviation $\sigma = 100$.

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

$$H_0: \mu = 475$$

$$H_a: \mu \neq 475$$

for each of the 50 samples.

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

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

6.79 Suppose that in fact the mean SATM score of California high school seniors is $\mu = 500$. Would the test in the previous exercise usually detect a mean this far from the hypothesized value? This is a question about the power of the test.
(a) Simulate the drawing of 50 SRSs from a normal population with mean $\mu = 500$ and $\sigma = 100$. These represent the results of sampling when in fact the alternative $\mu = 500$ is true.

(b) Repeat on these new data the test of

$$H_0: \mu = 475$$

$$H_a: \mu \neq 475$$

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

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

**CHAPTER 7**

**Section 7.1**

7.1 What critical value $t^*$ from Table D should be used for a confidence interval for the mean of the population in each of the following situations?

(a) A 90% confidence interval based on $n = 12$ observations.

(b) A 95% confidence interval from an SRS of 30 observations.

(c) An 80% confidence interval from a sample of size 18.

7.2 Use software to find the critical values $t^*$ that you would use for each of the following confidence intervals for the mean.

(a) A 99% confidence interval based on $n = 55$ observations.
(b) A 90% confidence interval from an SRS of 35 observations.
(c) An 95% confidence interval from a sample of size 90.

7.3 The one-sample t statistic for testing

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

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

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

\[ H_0: \mu = 64 \]
\[ H_a: \mu \neq 64 \]

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

7.5 The one-sample t statistic for a test of

\[ H_0: \mu = 20 \]
Section 7.1

$H_0: \mu < 20$

based on $n = 12$ observations has the value $t = -2.45$.

(a) What are the degrees of freedom for this statistic?

(b) How would you report the $P$-value based on Table D?

(c) If you have software available, find the exact $P$-value.

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

(a) Is there significant evidence at the 1% level that the mean amount charged increases under the no-fee offer? State $H_0$ and $H_a$ and carry out a $t$ test.

(b) Give a 95% confidence interval for the mean amount of the increase.

(c) The distribution of the amount charged is skewed to the right, but outliers are prevented by the credit limit that the bank enforces on each card. Use of the $t$ procedures is justified in this case even though the population distribution is not normal. Explain why.

(d) A critic points out that the customers would probably have charged more this year than last even without the new offer because the economy is more prosperous and interest rates are lower. Briefly describe the design of an experiment to study the effect of the no-fee offer that would avoid this criticism.

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

(b) Write the criterion for rejecting \( H_0: \mu = 0 \) in terms of the \( t \) statistic. Then take \( s = 108 \) (an estimate based on the data in Exercise 7.21) and state the rejection criterion in terms of \( \bar{x} \).

(c) Assume that \( \mu = 100 \) (the given alternative) and that \( \sigma = 108 \) (an estimate from the data in the previous exercise). The approximate power is the probability of the event you found in (b), calculated under these assumptions. Find the power. Would you recommend that the bank do a test on 50 customers, or should more customers be included?

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

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

Find a 95% confidence interval for the mean amount of D-glucose in cockroach hindguts under these conditions. (Based on D. L. Shankland et al., “The effect of 5-thio-D-glucose on insect development and its absorption by insects,” Journal of Insect Physiology, 14 (1968), pp. 63–72.)

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

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.

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

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

(a) Give a 95% confidence interval for the mean ARSMA score of Mexicans.

(b) What assumptions does your confidence interval require? Which of these assumptions is most important in this case?

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

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

(b) Carry out a significance test for the hypothesis that the two tests have the same population mean. Give the \( P \)-value and state your conclusion.

(c) Give a 95% confidence interval for the difference between the two population mean scores.

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

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

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

against the alternative \( \mu = 0.2 \). Note that this is a two-sided test.

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

(b) Write the criterion for rejecting \( H_0 \) at the \( \alpha = 0.05 \) level. Then take \( s = 0.3 \), the approximate value observed in Exercise 7.36, and restate the rejection criterion in terms of \( \bar{x} \).
(c) Find the probability of this event when \( \mu = 0.2 \) (the alternative given) and \( \sigma = 0.3 \) (estimated from the data in Exercise 7.36) by a normal probability calculation. This is the approximate power.

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

21.6  20.0  25.0  21.9

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

7.16 Your local newspaper contains a large number of advertisements for unfurnished one-bedroom apartments. You choose 10 at random and calculate that their mean monthly rent is $540 and that the standard deviation of their rents is $80.

(a) What is the standard error of the mean?

(b) What are the degrees of freedom for a one-sample \( t \) statistic?

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

500, 650, 600, 505, 450, 550, 515, 495, 650, 395
Chapter 7 Exercises

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

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

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

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

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

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

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

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

7.21 A manufacturer of small appliances employs a market research firm to estimate retail sales of its products. Here are last month’s sales of electric can openers from an SRS of 50 stores in the Midwest sales region:
(a) Make a stemplot of the data. The distribution is skewed to the right and has several high outliers. The bootstrap (page xxx) is a modern computer-intensive tool for getting accurate confidence intervals without the normality condition. Three bootstrap simulations, each with 10,000 repetitions, give these 95% confidence intervals for mean sales in the entire region: (20.42, 27.26), (20.40, 27.18), and (20.48, 27.28).

(b) Find the 95% \( t \) confidence interval for the mean. It is essentially the same as the bootstrap intervals. The lesson is that for sample sizes as large as \( n = 50 \), \( t \) procedures are very robust.

7.22 Refer to the previous exercise. Each electric can opener sold generates a profit of $2.50 for the manufacturer.

(a) What is the mean profit per store in the Midwest sales region?

(b) Transform the confidence interval you found in the previous exercise into an interval for the mean profit for stores in the Midwest region.

7.23 Refer to the previous two exercises. There are 4325 stores that sell can openers manufactured by this company.

(a) Estimate the total profit for sales last month in the Midwest region.

(b) Give a 95% confidence interval for the total profit for sales last month in the Midwest region.

7.24 The scores of four roommates on the Law School Admission Test (LSAT) are

\[ 628, 593, 455, 503 \]
Find the mean, the standard deviation, and the standard error of the mean. Is it appropriate to calculate a confidence interval for these data? Explain why or why not.

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

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

(a) Display the data using a stemplot and make a normal quantile plot. Describe the distribution of calcium intakes for these women.

(b) Calculate the mean, the standard deviation, and the standard error.

(c) Find a 95% confidence interval for the mean.

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

7.27 Refer to Exercise 7.25. Suppose that the recommended daily allowance (RDA) of calcium for women in this age range is 1300 milligrams (this value is changed from time to time on the basis of the statistical analysis of new data). We want to express the results in terms of percent of the RDA.

(a) Divide each intake by the RDA, multiply by 100, and compute the 95% confidence interval from the transformed data.

(b) Verify that you can obtain the same result by similarly transforming the interval you calculated in Exercise 7.25.

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

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

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

```
20  40  22  22  21  21  20  10  20  20
20  13  18  50  20  18  15  8  22  25
22  10  20  22  22  21  15  23  30  12
  9  20  40  22  29  19  15  20  20  20
20  15  19  21  14  22  21  35  20  22
```

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

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

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

**7.33** Refer to the three previous exercises. Suppose you are interested in the cost per year rather than the cost per month. Find a 95% confidence interval for the mean yearly cost of Internet access. How does this interval relate to the one that you found in Exercise 7.20?

**7.34** The cost of health care is the subject of many studies that use statistical methods. One such study estimated that the average length of service for home health care among people over the age of 65 who use this type of service is 96.0 days with a standard error of 5.1 days. Assuming that the degrees of freedom are large, calculate a 90% confidence interval for the true mean length of service. (A. N. Dey, “Characteristics of elderly home health care users,” National Center for Health Statistics, 1996.)

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

(a) What does the abbreviation “SEM” stand for?

(b) The researchers made three measurements of ATP, which had \( \bar{x} = 0.84 \). What was the sample standard deviation \( s \) for these measurements?

(c) Give a 90% confidence interval for the mean ATP level in dormant brine shrimp embryos.

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

(b) The project hoped to show that right-handed people find right-hand threads easier to use. State the appropriate $H_0$ and $H_a$ about the mean time required to complete the task.

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

7.37 Refer to the previous exercise. Give a 90% confidence interval for the mean time advantage of right-hand over left-hand threads in the setting of the previous exercise. Do you think that the time saved would be of practical importance if the task were performed many times—for example, by an assembly-line worker? To help answer this question, find the mean time for right-hand threads as a percent of the mean time for left-hand threads.
7.38 An agricultural field trial compares the yield of two varieties of tomatoes for commercial use. The researchers divide in half each of 10 small plots of land in different locations and plant each tomato variety on one half of each plot. After harvest, they compare the yields in pounds per plant at each location. The 10 differences (Variety A – Variety B) give the following statistics: $\bar{x} = 0.46$ and $s = 0.92$. Is there convincing evidence that Variety A has the higher mean yield? State $H_0$ and $H_a$, and give a $P$-value to answer this question.

7.39 The tomato experts who carried out the field trial described in the previous exercise suspect that the relative lack of significance there is due to low power. They would like to be able to detect a mean difference in yields of 0.6 pound per plant at the 0.05 significance level. Based on the previous study, use 0.92 as an estimate of both the population $\sigma$ and the value of $s$ in future samples.
(a) What is the power of the test from Exercise 7.43 with $n = 12$ against the alternative $\mu = 0.6$?
(b) If the sample size is increased to $n = 30$ plots of land, what will be the power against the same alternative?

7.40 The following situations all require inference about a mean or means. Identify each as (1) a single sample, (2) matched pairs, or (3) two independent samples. The procedures of this section apply to cases (1) and (2). We will learn procedures for (3) in the next section.
(a) An education researcher wants to learn whether inserting questions before or after introducing a new concept in an elementary school mathematics text is more effective. He prepares two text segments that teach the concept, one with motivating questions before and the other with review questions after. Each text segment is used to teach a different group of children, and their scores on a test over the material are compared.
(b) Another researcher approaches the same problem differently. She prepares text
segments on two unrelated topics. Each segment comes in two versions, one with questions before and the other with questions after. Each of a group of children is taught both topics, one topic (chosen at random) with questions before and the other with questions after. Each child’s test scores on the two topics are compared to see which topic he or she learned better.

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

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

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

Section 7.2

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

- **Control:** 2 4 3 4 2 3 3 5 3 2 6 3 4
- **Treatment:** 0 1 1 2 1 2 1 1 2 1 1 1
(Based on M. C. Wilson et al., “Impact of cereal leaf beetle larvae on yields of oats,” *Journal of Economic Entomology*, 62 (1969), pp. 699–702.) Is there significant evidence at the 1% level that the mean number of larvae per stem is reduced by Malathion? Be sure to state $H_0$ and $H_a$.

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

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

(a) Do the data show a significant difference between the mean amounts charged by customers offered the two plans? Give the null and alternative hypotheses, and calculate the two-sample $t$ statistic. Obtain the $P$-value (either approximately from Table D or more accurately from software). State your practical conclusions.

(b) The distributions of amounts charged are skewed to the right, but outliers are prevented by the limits that the bank imposes on credit balances. Do you think that skewness threatens the validity of the test that you used in (a)? Explain your answer.

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

TTEST PROCEDURE

Variable: KNEE

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

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

(a) The researchers believed that the knee velocity would be higher for skilled rowers. State $H_0$ and $H_a$.

(b) Give the value of the two-sample $t$ statistic and its $P$-value (note that SAS provides two-sided $P$-values). What do you conclude?

(c) Give a 90% confidence interval for the mean difference between the knee velocities of skilled and novice female rowers.
The novice and skilled rowers in the previous exercise were also compared with respect to several physical variables. Here is the SAS computer output for weight in kilograms:

**TTEST PROCEDURE**

Variable: WEIGHT

<table>
<thead>
<tr>
<th>GROUP</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKILLED</td>
<td>10</td>
<td>70.3700</td>
<td>6.1003</td>
<td>1.9291</td>
</tr>
<tr>
<td>NOVICE</td>
<td>8</td>
<td>68.4500</td>
<td>9.0399</td>
<td>3.1961</td>
</tr>
</tbody>
</table>

Variances T DF Prob>|T|

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

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

The Johns Hopkins Regional Talent Searches give the SAT (intended for high school juniors and seniors) to 13-year-olds. In all, 19,883 males and 19,937 females took the tests between 1980 and 1982. The mean scores of males and females on the verbal test are nearly equal, but there is a clear difference between the sexes on the mathematics test. The reason for this difference is not understood. Here are the data (from a news article in *Science*, 224 (1983), pp. 1029–1031):
Give a 99% confidence interval for the difference between the mean score for males and the mean score for females in the population that Johns Hopkins searches.

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

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>380 321 366 356</td>
<td>361 447 401 375</td>
</tr>
<tr>
<td></td>
<td>283 349 402 462</td>
<td>434 403 393 426</td>
</tr>
<tr>
<td></td>
<td>356 410 329 399</td>
<td>406 318 467 407</td>
</tr>
<tr>
<td></td>
<td>350 384 316 272</td>
<td>427 420 477 392</td>
</tr>
<tr>
<td></td>
<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) State the hypotheses for a statistical test of the claim that chicks fed high-lysine corn gain weight faster. Carry out the test. Is the result significant at the 10% level? At the 5% level? At the 1% level?

(c) Give a 95% confidence interval for the mean extra weight gain in chicks fed high-lysine corn.
7.48 The data on weights of skilled and novice rowers in Exercise S7.19 can be analyzed by the pooled \( t \) procedures, which assume equal population variances. Report the value of the \( t \) statistic, its degrees of freedom, and its \( P \)-value, and then state your conclusion. (The pooled procedures should not be used for the comparison of knee velocities in Exercise S7.18, because the sample standard deviations in the two groups are different enough to cast doubt on the assumption of a common population standard deviation.)

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

\[
595, 500, 580, 650, 675, 675, 750, 500, 495, 670
\]

Here are the rents for the one-bedroom apartments:

\[
500, 650, 600, 505, 450, 550, 515, 495, 650; 395
\]

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

7.50 Pat wonders if two-bedroom apartments rent for significantly more than one-bedroom apartments. Use the data in the previous exercise to find out.

(a) State appropriate null and alternative hypotheses.

(b) Report the test statistic, its degrees of freedom, and the \( P \)-value. What do you conclude?

(c) Can you conclude that every one-bedroom apartment costs less than every two-bedroom apartment?

(d) In the previous exercise you found a confidence interval. In this exercise you performed a significance test. Which do you think is more useful to someone planning to rent an apartment? Why?
7.51 Physical fitness is related to personality characteristics. In one study of this relationship, middle-aged college faculty who had volunteered for a fitness program were divided into low-fitness and high-fitness groups based on a physical examination. The subjects then took the Cattell Sixteen Personality Factor Questionnaire. (A. H. Ismail and R. J. Young, “The effect of chronic exercise on the personality of middle-aged men,” *Journal of Human Ergology*, 2 (1973), pp. 47–57.) Here are the data for the “ego strength” personality factor:

<table>
<thead>
<tr>
<th>Low fitness</th>
<th>High fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.99 5.53 3.12</td>
<td>6.68 5.93 5.71</td>
</tr>
<tr>
<td>4.24 4.12 3.77</td>
<td>6.42 7.08 6.20</td>
</tr>
<tr>
<td>4.74 5.10 5.09</td>
<td>7.32 6.37 6.04</td>
</tr>
<tr>
<td>4.93 4.47 5.40</td>
<td>6.38 6.53 6.51</td>
</tr>
<tr>
<td>4.16 5.30</td>
<td>6.16 6.68</td>
</tr>
</tbody>
</table>

(a) Is the difference in mean ego strength significant at the 5% level? At the 1% level? Be sure to state $H_0$ and $H_a$.

(b) You should be hesitant to generalize these results to the population of all middle-aged men. Explain why.

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

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

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

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

7.54 A market research firm supplies manufacturers with estimates of the retail sales of their products from samples of retail stores. Marketing managers are prone to look at the estimate and ignore sampling error. Suppose that an SRS of 75 stores this month shows mean sales of 52 units of a small appliance, with standard deviation 13 units. During the same month last year, an SRS of 53 stores gave mean sales of 49 units, with standard deviation 11 units. An increase from 49 to 52 is a rise of 6%. The marketing manager is happy, because sales are up 6%.

(a) Use the two-sample \( t \) procedure to give a 95% confidence interval for the difference in mean number of units sold at all retail stores.

(b) Explain in language that the manager can understand why he cannot be certain that sales rose by 6%, and that in fact sales may even have dropped.

7.55 In a study of heart surgery, one issue was the effect of drugs called beta-blockers on the pulse rate of patients during surgery. The available subjects were divided at random into two groups of 30 patients each. One group received a beta-blocker; the other, a placebo. The pulse rate of each patient at a critical point during the operation was recorded. The treatment group had mean 65.2 and standard deviation 7.8. For the control group, the mean was 70.3 and the standard deviation was 8.3.

(a) Do beta-blockers reduce the pulse rate? State the hypotheses and do a \( t \) test. Is the result significant at the 5% level? At the 1% level?

(b) Give a 99% confidence interval for the difference in mean pulse rates.

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

<table>
<thead>
<tr>
<th>Beef hot dogs</th>
<th>Meat hot dogs</th>
<th>Poultry hot dogs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calories</td>
<td>Sodium</td>
<td>Calories</td>
</tr>
<tr>
<td>186</td>
<td>495</td>
<td>173</td>
</tr>
<tr>
<td>181</td>
<td>477</td>
<td>191</td>
</tr>
<tr>
<td>176</td>
<td>425</td>
<td>182</td>
</tr>
<tr>
<td>149</td>
<td>322</td>
<td>190</td>
</tr>
<tr>
<td>184</td>
<td>482</td>
<td>172</td>
</tr>
<tr>
<td>190</td>
<td>587</td>
<td>147</td>
</tr>
<tr>
<td>158</td>
<td>370</td>
<td>146</td>
</tr>
<tr>
<td>139</td>
<td>322</td>
<td>139</td>
</tr>
<tr>
<td>175</td>
<td>479</td>
<td>175</td>
</tr>
<tr>
<td>148</td>
<td>375</td>
<td>136</td>
</tr>
<tr>
<td>152</td>
<td>330</td>
<td>179</td>
</tr>
<tr>
<td>111</td>
<td>300</td>
<td>153</td>
</tr>
<tr>
<td>141</td>
<td>386</td>
<td>107</td>
</tr>
<tr>
<td>153</td>
<td>401</td>
<td>195</td>
</tr>
<tr>
<td>190</td>
<td>645</td>
<td>135</td>
</tr>
<tr>
<td>157</td>
<td>440</td>
<td>140</td>
</tr>
<tr>
<td>131</td>
<td>317</td>
<td>138</td>
</tr>
<tr>
<td>149</td>
<td>319</td>
<td></td>
</tr>
<tr>
<td>135</td>
<td>298</td>
<td></td>
</tr>
<tr>
<td>132</td>
<td>253</td>
<td></td>
</tr>
</tbody>
</table>

(a) Give a 95% confidence interval for the difference in mean calorie content between beef and poultry hot dogs.

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

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

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

<table>
<thead>
<tr>
<th>Calcium group</th>
<th>Placebo group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begin End Decrease</td>
<td>Begin End Decrease</td>
</tr>
<tr>
<td>107 100 7</td>
<td>123 124 -1</td>
</tr>
<tr>
<td>110 114 -4</td>
<td>109 97 12</td>
</tr>
<tr>
<td>123 105 18</td>
<td>112 113 -1</td>
</tr>
<tr>
<td>129 112 17</td>
<td>102 105 -3</td>
</tr>
<tr>
<td>112 115 -3</td>
<td>98 95 3</td>
</tr>
<tr>
<td>111 116 -5</td>
<td>114 119 -5</td>
</tr>
<tr>
<td>107 106 1</td>
<td>119 114 5</td>
</tr>
<tr>
<td>112 102 10</td>
<td>114 112 2</td>
</tr>
<tr>
<td>136 125 11</td>
<td>110 121 -11</td>
</tr>
<tr>
<td>102 104 -2</td>
<td>117 118 -1</td>
</tr>
<tr>
<td>130 133 -3</td>
<td></td>
</tr>
</tbody>
</table>

One group took a calcium supplement, and the other group received a placebo. Example 7.20 compares the decrease in blood pressure in the two groups using pooled two-sample \( t \) procedures.

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

(b) Give a 90% confidence interval for the difference in means, again using a proce-
Chapter 7 Exercises

dure that does not require equal standard deviations. How does the margin of error of your interval compare with that in Example 7.21?

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

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

(a) What is the standard error of the sample mean VOT for the 20 adult subjects? What is the standard error of the difference $x_1 - x_2$ between the mean VOT for children and adults?

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

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

7.59 The researchers in the study discussed in the previous exercise looked at VOTs for adults and children pronouncing several different words. Explain why they should not perform a separate two-sample $t$ test for each word and conclude that the words with a significant difference (say, $P < 0.05$) distinguish children from adults. (The researchers did not make this mistake.)
7.60 Repeat the comparison of mean VOTs for children and adults in Exercise 7.58 using a pooled \( t \) procedure. (In practice, we would not pool in this case, because the data suggest some difference in the population standard deviations.)

(a) Carry out the significance test, and give a \( P \)-value.
(b) Give a 95% confidence interval for the difference in population means.
(c) How similar are your results to those you obtained in Exercise 7.79 from the two-sample \( t \) procedures?

7.61 College financial aid offices expect students to use summer earnings to help pay for college. But how large are these earnings? One college studied this question by asking a sample of students how much they earned. Omitting students who were not employed, 1296 responses were received. (Based on studies conducted by Marvin Schlatter, Division of Financial Aid, Purdue University.) Here are the data in summary form:

<table>
<thead>
<tr>
<th>Group</th>
<th>( n )</th>
<th>( \bar{x} )</th>
<th>( s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>675</td>
<td>$3297.91</td>
<td>$2394.65</td>
</tr>
<tr>
<td>Females</td>
<td>621</td>
<td>$2380.68</td>
<td>$1815.55</td>
</tr>
</tbody>
</table>

(a) Use the two-sample \( t \) procedures to give a 90% confidence interval for the difference between the mean summer earnings of male and female students.
(b) The distribution of earnings is strongly skewed to the right. Nevertheless, use of \( t \) procedures is justified. Why?
(c) Once the sample size was decided, the sample was chosen by taking every \( k \)th name from an alphabetical list of undergraduates. Is it reasonable to consider the sample as two SRSs chosen from the male and female undergraduate populations?
(d) What other information about the study would you request before accepting the results as describing all undergraduates?

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

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


The control group data were


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

\[ H_0: \mu_1 = \mu_2 \]

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

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

TTEST PROCEDURE
Variable: SPIKE

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

Variances

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

(a) Interpret the output.

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

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

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

Do these data support the contention that female and male students differ in average social insight? Use the pooled two-sample procedure and the procedure that does not assume that the standard deviations are the same. Compare the results.
Section 7.3

7.64 The $F$ statistic $F = \frac{s_1^2}{s_2^2}$ is calculated from samples of size $n_1 = 10$ and $n_2 = 21$. (Remember that $n_1$ is the numerator sample size.)

(a) What is the upper 5% critical value for this $F$?

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

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

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

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

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

(a) State $H_0$ and $H_a$.

(b) Calculate the $F$ statistic. Between which two levels does the $P$-value lie?

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

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

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

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

(a) What is the upper 5% critical value for this $F$?

(b) In a test of equality of standard deviations against the two-sided alternative, this statistic has the value $F = 2.71$. Is this value significant at the 5% level? Is it significant at the 1% level?

**7.71** The $F$ statistic for equality of standard deviations based on samples of sizes $n_1 = 31$ and $n_2 = 28$ takes the value $F = 1.72$.

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

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

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

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

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

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

Chapter 7 Review Exercises

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

**7.77** In a study of the effectiveness of weight-loss programs, 47 subjects who were at least 20% overweight took part in a group support program for 10 weeks. Private weighings determined each subject’s weight at the beginning of the program and 6 months after the program’s end. The matched pairs $t$ test was used to assess the significance of the average weight loss. The paper reporting the study said, “The subjects lost a significant amount of weight over time, $t(46) = 4.68, p < 0.01.$” It is common to report the results of statistical tests in this abbreviated style. (Based loosely on D. R. Black et al., “Minimal interventions for weight control: a cost-effective alternative,” *Addictive Behaviors*, 9 (1984), pp. 279–285.)

(a) Why was the matched pairs statistic appropriate?
(b) Explain to someone who knows no statistics but is interested in weight-loss programs what the practical conclusion is.

(c) The paper follows the tradition of reporting significance only at fixed levels such as $\alpha = 0.01$. In fact, the results are more significant than “$p < 0.01$” suggests. What can you say about the $P$-value of the $t$ test?

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

<table>
<thead>
<tr>
<th>Group</th>
<th>$n$</th>
<th>$\bar{x}$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrite</td>
<td>30</td>
<td>7880</td>
<td>1115</td>
</tr>
<tr>
<td>Control</td>
<td>30</td>
<td>8112</td>
<td>1250</td>
</tr>
</tbody>
</table>

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

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

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

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

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

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

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

$$23, 33.5, 37, 40.5, 46$$

for the first minute and

$$32, 39, 44, 49, 59$$

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

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

$$63.4, 65.0, 64.4, 63.3, 54.8, 64.5, 60.8, 49.1, 51.0$$

These values are quite different from the present 78.1% of nitrogen in the atmosphere. Assume (this is not yet agreed on by experts) that these observations are an SRS from the late Cretaceous atmosphere. (Data from R. A. Berner and G. P.

(a) Graph the data, and comment on skewness and outliers.

(b) The *t* procedures will be only approximate in this case. Give a 90% *t* confidence interval for the mean percent of nitrogen in ancient air.

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

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

(a) Is there significant evidence at the 5% level that the new method frees less lead from the soil? What about the 1% level?

(b) A critic argues that because of variations in the soil, the effectiveness of the new method is confounded with characteristics of the particular soil specimens used. Briefly describe a better data production design that avoids this criticism.

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

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

(a) Is there strong evidence that pets have a higher mean cholesterol level than clinic dogs? State the \( H_0 \) and \( H_a \) and carry out an appropriate test. Give the \( P \)-value and state your conclusion.

(b) Give a 95% confidence interval for the difference in mean cholesterol levels between pets and clinic dogs.

(c) Give a 95% confidence interval for the mean cholesterol level in clinic dogs.

(d) What assumptions must be satisfied to justify the procedures you used in (a), (b), and (c)? Assuming that the cholesterol measurements have no outliers and are not strongly skewed, what is the chief threat to the validity of the results of this study?

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

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

Use confidence intervals to describe the difference between runners and typical young men.
CHAPTER 8

Section 8.1

8.1 In each of the following cases state whether or not the normal approximation to the binomial should be used for a significance test on the population proportion $p$.

(a) $n = 10$ and $H_0: p = 0.4$.
(b) $n = 100$ and $H_0: p = 0.6$.
(c) $n = 1000$ and $H_0: p = 0.996$.
(d) $n = 500$ and $H_0: p = 0.3$.

8.2 The Gallup Poll asked a sample of 1785 U.S. adults, “Did you, yourself, happen to attend church or synagogue in the last 7 days?” Of the respondents, 750 said “Yes.” Suppose (it is not, in fact, true) that Gallup’s sample was an SRS.

(a) Give a 99% confidence interval for the proportion of all U.S. adults who attended church or synagogue during the week preceding the poll.
(b) Do the results provide good evidence that less than half of the population attended church or synagogue?
(c) How large a sample would be required to obtain a margin of error of $\pm 0.01$ in a 99% confidence interval for the proportion who attend church or synagogue? (Use Gallup’s result as the guessed value of $p$.)

8.3 Leroy, a starting player for a major college basketball team, made only 38.4% of his free throws last season. During the summer he worked on developing a softer shot in the hope of improving his free-throw accuracy. In the first eight games of this season Leroy made 25 free throws in 40 attempts. Let $p$ be his probability of making each free throw he shoots this season.

(a) State the null hypothesis $H_0$ that Leroy’s free-throw probability has remained the same as last year and the alternative $H_a$ that his work in the summer resulted
in a higher probability of success.

(b) Calculate the $z$ statistic for testing $H_0$ versus $H_a$.

(c) Do you accept or reject $H_0$ for $\alpha = 0.05$? Find the $P$-value.

(d) Give a 90% confidence interval for Leroy’s free-throw success probability for the new season. Are you convinced that he is now a better free-throw shooter than last season?

(e) What assumptions are needed for the validity of the test and confidence interval calculations that you performed?

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

8.5 In the previous exercise we found that 11 customers from a random sample of 40 would be willing to buy a software upgrade that costs $100. If the upgrade is to be profitable, you will need to sell it to more than 20% of your customers. Do the sample data give good evidence that more than 20% are willing to buy?

(a) Formulate this problem as a hypothesis test. Give the null and alternative hypotheses. Will you use a one-sided or a two-sided alternative? Why?

(b) Carry out the significance test. Report the test statistic and the $P$-value.

(c) Should you proceed with plans to produce and market the upgrade?

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

8.7 When trying to hire managers and executives, companies sometimes verify the academic credentials described by the applicants. One company that performs these checks summarized its findings for a six-month period. Of the 84 applicants whose credentials were checked, 15 lied about having a degree. (Data provided by Jude M. Werra & Associates, Brookfield, Wisconsin.)

(a) Find the proportion of applicants who lied about having a degree and the standard error.

(b) Consider these data to be a random sample of credentials from a large collection of similar applicants. Give a 95% confidence interval for the true proportion of applicants who lie about having a degree.

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

8.9 A question in a Christmas tree market survey was “Did you have a Christmas tree last year?” Of the 500 respondents, 421 answered “Yes.”

(a) Find the sample proportion and its standard error.

(b) Give a 90% confidence interval for the proportion of Indiana households who had a Christmas tree this year.

8.10 Of the 500 respondents in the Christmas tree market survey, 44% had no children at home and 56% had at least one child at home. The corresponding figures for the most recent census are 48% with no children and 52% with at least one child. Test the null hypothesis that the telephone survey technique has a probability of selecting a household with no children that is equal to the value obtained by the census. Give the $z$ statistic and the $P$-value. What do you conclude?
8.11 Refer to the previous exercise. There we arbitrarily chose to state the hypotheses in terms of the proportion of rural respondents. We could as easily have used the proportion of urban respondents.

(a) Write hypotheses in terms of the proportion of urban residents to examine how well the sample represents the state in regard to rural versus urban residence.

(b) Perform the test of significance and summarize the results.

(c) Compare your results with the results of the previous exercise. Summarize and generalize your conclusion.

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

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

8.14 In recent years over 70% of first-year college students responding to a national survey have identified “being well-off financially” as an important personal goal. A state university finds that 103 of an SRS of 150 of its first-year students say that this goal is important. Give a 95% confidence interval for the proportion of all first-
Section 8.1

year students at the university who would identify being well-off as an important personal goal.

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

8.16 Shereka, a starting player for a major college basketball team, made only 36.2% of her free throws last season. During the summer she worked on developing a softer shot in the hope of improving her free-throw accuracy. In the first eight games of this season Shereka made 22 free throws in 42 attempts. Let \( p \) be her probability of making each free throw she shoots this season.

(a) State the null hypothesis \( H_0 \) that Shereka’s free-throw probability has remained the same as last year and the alternative \( H_a \) that her work in the summer resulted in a higher probability of success.

(b) Calculate the \( z \) statistic for testing \( H_0 \) versus \( H_a \).

(c) Do you accept or reject \( H_0 \) for \( \alpha = 0.05 \)? Find the \( P \)-value.

(d) Give a 90% confidence interval for Shereka’s free-throw success probability for the new season. Are you convinced that she is now a better free-throw shooter than last season?

(e) What assumptions are needed for the validity of the test and confidence interval calculations that you performed?

8.17 Land’s Beginning is a company that sells its merchandise through the mail. It is considering buying a list of addresses from a magazine. The magazine claims that at least 25% of its subscribers have high incomes (they define this to be household income in excess of \$100,000). Land’s Beginning would like to estimate the proportion of high-income people on the list. Checking income is very difficult and
expensive but another company offers this service. Land’s Beginning will pay to find incomes for an SRS of people on the magazine’s list. They would like the margin of error of the 95% confidence interval for the proportion to be 0.05 or less. Use the guessed value \( p^* = 0.25 \) to find the required sample size.

8.18 Refer to the previous exercise. For each of the following variations on the design specifications, state whether the required sample size will be higher, lower, or the same as that found above.

(a) Use a 90% confidence interval.
(b) Change the allowable margin of error to 0.10.
(c) Use a planning value of \( p^* = 0.30 \).
(d) Use a different company to do the income checks.

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

Section 8.2

8.20 In the 1996 regular baseball season, the World Series Champion New York Yankees played 80 games at home and 82 games away. They won 49 of their home games and 43 of the games played away. We can consider these games as samples from potentially large populations of games played at home and away. How much advantage does the Yankee home field provide?

(a) Find the proportion of wins for the home games. Do the same for the away games.
Section 8.2

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

(c) Compute a 90% confidence interval for the difference between the probability that the Yankees win at home and the probability that they win when on the road. Are you convinced that the 1996 Yankees were more likely to win at home?

8.21 Return to the New York Yankees baseball data in the previous exercise.

(a) Combining all of the games played, what proportion did the Yankees win?

(b) Find the standard error needed for testing that the probability of winning is the same at home and away.

(c) Most people think that it is easier to win at home than away. Formulate null and alternative hypotheses to examine this idea.

(d) Compute the $z$ statistic and its $P$-value. What conclusion do you draw?

8.22 The 1958 Detroit Area Study was an important sociological investigation of the influence of religion on everyday life. It is described in Gerhard Lenski, *The Religious Factor*, Doubleday, New York, 1961. The sample “was basically a simple random sample of the population of the metropolitan area.” Of the 656 respondents, 267 were white Protestants and 230 were white Catholics. One question asked whether the government was doing enough in areas such as housing, unemployment, and education; 161 of the Protestants and 136 of the Catholics said “No.” Is there evidence that white Protestants and white Catholics differed on this issue?

8.23 The respondents in the Detroit Area Study (see the previous exercise) were also asked whether they believed that the right of free speech included the right to make speeches in favor of communism. Of the white Protestants, 104 said “Yes,” while 75 of the white Catholics said “Yes.” Give a 95% confidence interval for the amount by which the proportion of Protestants who agreed that communist speeches are protected exceeds the proportion of Catholics who held this opinion.
8.24 A university financial aid office polled an SRS of undergraduate students to study their summer employment. Not all students were employed the previous summer. Here are the results for men and women:

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

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

(b) Give a 99% confidence interval for the difference between the proportions of male and female students who were employed during the summer. Does the difference seem practically important to you?

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

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

(b) Compare the results of this significance test with your results in Exercise 8.42. What do you observe about the effect of the sample size on the results of these significance tests?

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

(a) Test the null hypothesis that there is no difference between the proportions of the two types of shields removed. Give the $z$ statistic and the $P$-value. State your conclusion in words.

(b) Give a 90% confidence interval for the difference in the proportions of removed shields for the bolt-on and the flip-up types. Based on the data, what recommendation would you make about the type of shield to be used on new tractors?

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

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

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

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

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

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

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

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

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

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

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

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

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

8.31 Refer to the New York Yankees baseball data in the previous exercise.

(a) Combining all of the games played, what proportion did the Yankees win?
(b) Find the standard error needed for testing that the probability of winning is the same at home and away.

(c) Most people think that it is easier to win at home than away. Formulate null and alternative hypotheses to examine this idea.

(d) Compute the $z$ statistic and its $P$-value. What conclusion do you draw?

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

8.33 The state agriculture department asked random samples of Indiana farmers in each county whether they favored a mandatory corn checkoff program to pay for corn product marketing and research. In Tippecanoe County, 263 farmers were in favor of the program and 252 were not. In neighboring Benton County, 260 were in favor and 377 were not.

(a) Find the proportions of farmers in favor of the program in each of the two counties.

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

(c) Compute a 95% confidence interval for the difference between the proportions of farmers favoring the program in Tippecanoe County and in Benton County. Do you think opinions differed in the two counties?

8.34 Return to the survey of farmers described in the previous exercise.

(a) Combine the two samples and find the overall proportion of farmers who favor the corn checkoff program.

(b) Find the standard error needed for testing that the population proportions of
farmers favoring the program are the same in the two counties.

(c) Formulate null and alternative hypotheses for comparing the two counties.

(d) Compute the $z$ statistic and its $P$-value. What conclusion do you draw?

8.35 A study of chromosome abnormalities and criminality examined data on 4124 Danish males born in Copenhagen. (H. A. Witkin et al., “Criminality in XYY and XXY men,” Science, 193 (1976), pp. 547–555.) The study used the penal registers maintained in the offices of the local police chiefs and classified each man as having a criminal record or not. Each was also classified as having the normal male XY chromosome pair or one of the abnormalities XYY or XXY. Of the 4096 men with normal chromosomes, 381 had criminal records, while 8 of the 28 men with chromosome abnormalities had criminal records. Some experts believe that chromosome abnormalities are associated with increased criminality. Do these data lend support to this belief? Report your analysis and draw a conclusion.

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

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

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

(b) Give a 95% confidence interval for the difference between the proportions of male and female students who were employed during the summer. Does the difference seem practically important to you?
8.37 Refer to the study of undergraduate student summer employment described in the previous exercise. Similar results from a smaller number of students may not have the same statistical significance. Specifically, suppose that 73 of 82 men surveyed were employed and 60 of 75 women surveyed were employed. The sample proportions are essentially the same as in the earlier exercise.

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

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

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

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

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

(c) The physicians conducting the study had concluded from previous research that aspirin was likely to increase the chance of a favorable outcome. Carry out a significance test to confirm this conclusion. State hypotheses, find the $P$-value, and write a summary of your results.
**Exercise 8.39** The pesticide diazinon is in common use to treat infestations of the German cockroach, *Blattella germanica*. A study investigated the persistence of this pesticide on various types of surfaces. (Elray M. Roper and Charles G. Wright, “German cockroach (Orthoptera: Blatellidae) mortality on various surfaces following application of diazinon,” *Journal of Economic Entomology*, 78 (1985), pp. 733–737.) Researchers applied a 0.5% emulsion of diazinon to glass and plasterboard. After 14 days, they placed 18 cockroaches on each surface and recorded the number that died within 48 hours. On glass, 9 cockroaches died, while on plasterboard, 13 died.

(a) Calculate the mortality rates (sample proportion that died) for the two surfaces.

(b) Find a 95% confidence interval for the difference in the two population proportions.

(c) Chemical analysis of the residues of diazinon suggests that it may persist longer on plasterboard than on glass because it binds to the paper covering on the plasterboard. The researchers therefore expected the mortality rate to be greater on plasterboard than on glass. Conduct a significance test to assess the evidence that this is true.

**Exercise 8.40** Suppose that the experiment of the previous exercise placed more cockroaches on each surface and observed similar mortality rates. Specifically, suppose that 36 cockroaches were placed on each surface and that 26 died on the plasterboard, while 18 died on the glass.

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

(b) Compare the results of this significance test with those you gave in Exercise 8.51. What do you observe about the effect of the sample size on the results of these significance tests?
Chapter 8 Review Exercises

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

(a) What is the proportion of women in the class? Call this value $p_0$.

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

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

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

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

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

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

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

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

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

8.50 Eleven percent of the products produced by an industrial process over the past several months fail to conform to the specifications. The company modifies the process in an attempt to reduce the rate of nonconformities. In a trial run, the modified process produces 16 nonconforming items out of a total of 300 produced. Do these results demonstrate that the modification is effective? Support your conclusion with a clear statement of your assumptions and the results of your statistical calculations.
8.51 In the setting of the previous exercise, give a 95% confidence interval for the proportion of nonconforming items for the modified process. Then, taking $p_0 = 0.11$ to be the old proportion and $p$ the proportion for the modified process, give a 95% confidence interval for $p - p_0$.

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

<table>
<thead>
<tr>
<th></th>
<th>Black</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetarian</td>
<td>42</td>
<td>135</td>
</tr>
<tr>
<td>Not vegetarian</td>
<td>47</td>
<td>88</td>
</tr>
</tbody>
</table>

Are the proportions of vegetarians the same among all black and white Seventh-Day Adventists who attended this meeting? Analyze the data, paying particular attention to this question. Summarize your analysis and conclusions. What can you infer about the proportions of vegetarians among black and white Seventh-Day Adventists in general? What about blacks and whites in general?

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

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

\textbf{8.54} An experiment designed to assess the effects of aspirin on cardiovascular disease studied 5139 male British medical doctors. The doctors were randomly assigned to two groups. One group of 3429 doctors took one aspirin daily, and the other group did not take aspirin. After 6 years, there were 148 deaths from heart attack or stroke in the first group and 79 in the second group. A similar experiment used male American medical doctors as subjects. These doctors were also randomly assigned to one of two groups. The 11,037 doctors in the first group took one aspirin every other day, and the 11,034 doctors in the second group took no aspirin. After nearly 5 years, there were 104 deaths from heart attacks in the first group and 189 in the second. (The first study is reported in an article in the \textit{New York Times} of January 30, 1988; the second was described in the \textit{New York Times} on January 27, 1988.) Analyze the data from these two studies and summarize the results. How do the conclusions of the two studies differ, and why?

\textbf{CHAPTER 9}

\textbf{Chapter 9 Exercises}

\textbf{9.1} Investors use many “indicators” in their attempts to predict the behavior of the stock market. One of these is the “January indicator.” Some investors believe that if the market is up in January, then it will be up for the rest of the year. On the other hand, if it is down in January, then it will be down for the rest of the year.
The following table gives data for the Standard & Poor’s 500 stock index for the 75 years from 1916 to 1990:

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

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

(a) Calculate the column percents for this table. Explain briefly what they express.

(b) Do the same for the row percents.

(c) State appropriate null and alternative hypotheses for this problem. Use words rather than symbols.

(d) Find the table of expected counts under the null hypothesis. In which cells do the expected counts exceed the observed counts? In what cells are they less than the observed counts? Explain why the pattern suggests that the January indicator is valid.

(e) Give the value of the $X^2$ statistic, its degrees of freedom, and the $P$-value. What do you conclude?

(f) Write a short discussion of the evidence for the January indicator, referring to your analysis for substantiation.

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

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

(a) Find the row and column sums and put them in the margins of the table.

(b) For each type of county find the percent of physicians who prescribed tetracycline and the percent of those who did not. Do the same for the combined sample. Display the percents in a table and describe briefly what they show.

(c) Write null and alternative hypotheses to assess whether county type and prescription practices are unrelated.

(d) Carry out a significance test, give a full report of the results, and interpret them in plain language.

9.3 Alcohol and nicotine consumption during pregnancy may harm children. Because drinking and smoking behaviors may be related, it is important to understand the nature of this relationship when assessing the possible effects on children. One study classified 452 mothers according to their alcohol intake prior to pregnancy recognition and their nicotine intake during pregnancy. The data are summarized in the following table (from Ann P. Streissguth et al., “Intrauterine alcohol and nicotine exposure: attention and reaction time in 4-year-old children,” *Developmental Psychology*, 20 (1984), pp. 533–541):
Carry out a complete analysis of the association between alcohol and nicotine consumption. That is, describe the nature and strength of this association and assess its statistical significance. Include charts or figures to display the association.

9.4 Nutrition and illness are related in a complex way. If the diet is inadequate, the ability to resist infections can be impaired and illness results. On the other hand, some illnesses cause lack of appetite, so that poor nutrition can be the result of illness. In a study of morbidity and nutritional status in 1165 preschool children living in poor conditions in Delhi, India, data were obtained on nutrition and illness. Nutrition was described by a standard method as normal or as one of four levels of inadequate: I, II, III, and IV. For the purpose of analysis, the two most severely undernourished groups, III and IV, were combined. One part of the study examined four categories of illness during the past year: upper respiratory infection (URI), diarrhea, URI and diarrhea, and none. The following table gives the data. (Data from Vimlesh Seth et al., “Profile of morbidity and nutritional status and their effect on the growth potentials in preschool children in Delhi, India,” *Tropical Pediatrics and Environmental Health*, 25 (1979), pp. 23–29.)

<table>
<thead>
<tr>
<th>Alcohol (ounces/day)</th>
<th>Nicotine (milligrams/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>None</td>
<td>105</td>
</tr>
<tr>
<td>0.01–0.10</td>
<td>58</td>
</tr>
<tr>
<td>0.11–0.99</td>
<td>84</td>
</tr>
<tr>
<td>1.00 or more</td>
<td>57</td>
</tr>
</tbody>
</table>
Carry out a complete analysis of the association between nutritional status and type of illness. That is, describe the association numerically, assess its significance, and write a brief summary of your findings that refers to your analysis for substantiation.

9.5 Aluminum is suspected as a factor in the development of Alzheimer’s disease. In one study, researchers compared a group of Alzheimer’s patients with a carefully selected control group of people who did not have Alzheimer’s but were similar in other ways. (Selection of a matching control group is a difficult task. In epidemiological studies such as this, however, experiments are not possible.) The focus of the study was on the use of antacids that contain aluminum. Each subject was classified according to the use of these antacids. The two-way table below gives the data. (Data from Amy Borenstein Graves et al., “The association between aluminum-containing products and Alzheimer’s disease,” *Journal of Clinical Epidemiology, 43* (1990), pp. 35–44.)

<table>
<thead>
<tr>
<th>Aluminum-containing antacid use</th>
<th>None</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s patients</td>
<td>112</td>
<td>3</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Control group</td>
<td>114</td>
<td>9</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Analyze the data and summarize your results. Does the use of aluminum-containing antacids appear to be associated with Alzheimer’s disease?
9.6 Are there gender differences in the progress of students in doctoral programs?
A major university classified all students entering Ph.D. programs in a given year by their status 6 years later. The categories used were as follows: completed the degree, still enrolled, and dropped out. Here are the data:

<table>
<thead>
<tr>
<th>Status</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed</td>
<td>423</td>
<td>98</td>
</tr>
<tr>
<td>Still enrolled</td>
<td>134</td>
<td>33</td>
</tr>
<tr>
<td>Dropped out</td>
<td>238</td>
<td>98</td>
</tr>
</tbody>
</table>

Assume that these data can be viewed as a random sample giving us information on student progress. Describe the data using whatever percents are appropriate. State and test a null hypothesis and alternative that address the question of gender differences. Summarize your conclusions. What factors not given might be relevant to this study?

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

9.8 An article in the *New York Times* of April 24, 1991, discussed data from the Centers for Disease Control that showed an increase in cases of measles in the
United States. Of particular concern are complications from measles that can lead to death. The article noted that young children, who do not have fully developed immune systems, face an increased risk of death from complications of measles. Here are data on the 23,067 cases of measles reported in 1990. For each age group, the probability of death from measles is a parameter of interest. A comparison of the estimates of these parameters across age groups will provide information about the relationship between age and survival of an attack of measles.

<table>
<thead>
<tr>
<th>Age group</th>
<th>Dead</th>
<th>Survived</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 1 year</td>
<td>17</td>
<td>3806</td>
</tr>
<tr>
<td>1–4 years</td>
<td>37</td>
<td>7113</td>
</tr>
<tr>
<td>5–9 years</td>
<td>3</td>
<td>2208</td>
</tr>
<tr>
<td>10–14 years</td>
<td>3</td>
<td>1888</td>
</tr>
<tr>
<td>15–19 years</td>
<td>8</td>
<td>2715</td>
</tr>
<tr>
<td>20–24 years</td>
<td>6</td>
<td>2209</td>
</tr>
<tr>
<td>25–29 years</td>
<td>9</td>
<td>1492</td>
</tr>
<tr>
<td>30 years and over</td>
<td>14</td>
<td>1636</td>
</tr>
</tbody>
</table>

Summarize the death rates by age group. Prepare a plot to illustrate the pattern. Test the hypothesis that survival and age are related, report the results, and summarize your conclusion. From the data given, is it possible to study the association between catching measles and age? Explain why or why not.

9.9 Refer to Exercise 9.5, where we examined the relationship between use of aluminum-containing antacid and Alzheimer’s disease. In that exercise the $P$-value was 0.068, failing to achieve the traditional standard for statistical significance (0.05). Suppose that we did a similar study with more data. In particular, let’s double each of the counts in the original table. Perform the analysis on these counts and summarize the effect of increasing the sample size.
9.10 The Census Bureau collects data on years of school completed by Americans of different ages. The following table gives the years of education for three different age groups. People under the age of 25 are not included because many have not yet completed their education. Note that the unit of measure for each entry in the table is thousands of persons.

<table>
<thead>
<tr>
<th>Age group</th>
<th>Education</th>
<th>25 to 34</th>
<th>35 to 54</th>
<th>55 and over</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did not complete high school</td>
<td>5,325</td>
<td>9,152</td>
<td>16,035</td>
<td>30,512</td>
<td></td>
</tr>
<tr>
<td>Completed high school</td>
<td>14,061</td>
<td>24,070</td>
<td>18,320</td>
<td>56,451</td>
<td></td>
</tr>
<tr>
<td>College, 1 to 3 years</td>
<td>11,659</td>
<td>19,926</td>
<td>9,662</td>
<td>41,247</td>
<td></td>
</tr>
<tr>
<td>College, 4 or more years</td>
<td>10,342</td>
<td>19,878</td>
<td>8,005</td>
<td>38,225</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>41,388</td>
<td>73,028</td>
<td>52,022</td>
<td>166,438</td>
<td></td>
</tr>
</tbody>
</table>

(a) Give the joint distribution of education and age for this table.

(a) What is the marginal distribution of age?

(c) What is the marginal distribution of education?

9.11 Refer to the previous exercise. Find the conditional distribution of education for each of the three age categories. Make a bar graph for each distribution and summarize their differences and similarities.

9.12 Refer to the previous exercise. Compute the conditional distribution of age for each of the four education categories. Summarize the distributions graphically and write a short paragraph describing the distributions and how they differ.

9.13 The National Center for Education Statistics collects data on undergraduate students enrolled in U.S. colleges and universities. The following table gives counts of undergraduate enrollment in four different classifications:
(a) How many undergraduate students were enrolled in colleges and universities?
(b) What percent of all undergraduate students were under 18 years old?
(c) Find the percent of the undergraduates enrolled in each of the four types of program who were 22 to 34 years old. Make a bar graph to compare these percents.
(d) The 18 to 21 group is the traditional age group for college students. Briefly summarize what you have learned from the data about the extent to which this group predominates in different kinds of college programs.

9.14 Refer to the previous exercise. Find the marginal distribution of college type. Describe the distribution graphically and write a short summary.

9.15 Take the counts for the four college types that you used in the previous exercise and write these in a $2 \times 2$ table with year (2-year or 4-year) as the columns and time (full-time or part-time) as the rows. The marginal distribution that you calculated for the previous exercise is the joint distribution for this exercise. Compute the conditional distribution of time for each year category. Describe and contrast these distributions. Does there appear to be a relationship? If so, describe it.
9.16 For each type of college, find the conditional distributions of age for the \(4 \times 3\) table in Exercise 9.13. Display the distributions with bar graphs and describe how the age profile of the students varies with the type of college.

9.17 The following two-way table describes the age and marital status of American women. 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>18–24</td>
<td>9,289</td>
<td>3,046</td>
<td>19</td>
<td>260</td>
<td>12,613</td>
</tr>
<tr>
<td>25–39</td>
<td>6,948</td>
<td>21,437</td>
<td>206</td>
<td>3,408</td>
<td>32,000</td>
</tr>
<tr>
<td>40–64</td>
<td>2,307</td>
<td>26,679</td>
<td>2,219</td>
<td>5,508</td>
<td>36,713</td>
</tr>
<tr>
<td>≥ 65</td>
<td>768</td>
<td>7,767</td>
<td>8,636</td>
<td>1,091</td>
<td>18,264</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>19,312</strong></td>
<td><strong>58,931</strong></td>
<td><strong>11,080</strong></td>
<td><strong>10,266</strong></td>
<td><strong>99,588</strong></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?

9.18 Here is a two-way table of suicides committed, categorized by the gender of the victim and the method used. (“Hanging” also includes suffocation.) Write a brief account of differences in suicide between males and females. Use calculations and a graph to justify your statements.
9.19 Here are the numbers of flights on time and delayed for two airlines at five airports. (These data, from reports submitted by airlines to the Department of Transportation, appear in A. Barnett, “How numbers can trick you,” Technology Review, October 1994, pp. 38–45.) Overall on-time percents for each airline are often reported in the news. Lurking variables can make such reports misleading.

<table>
<thead>
<tr>
<th>Method</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firearms</td>
<td>16,381</td>
<td>2,559</td>
</tr>
<tr>
<td>Poison</td>
<td>3,569</td>
<td>2,110</td>
</tr>
<tr>
<td>Hanging</td>
<td>3,824</td>
<td>803</td>
</tr>
<tr>
<td>Other</td>
<td>1,641</td>
<td>623</td>
</tr>
</tbody>
</table>

(a) What percent of all Alaska Airlines flights were delayed? What percent of all America West flights were delayed? These are the numbers usually reported.

(b) Now find the percent of delayed flights for Alaska Airlines at each of the five airports. Do the same for America West.

(c) America West does worse at every one of the five airports, yet does better overall. That sounds impossible. Explain carefully, referring to the data, how this can happen. (The weather in Phoenix and Seattle lies behind this example of Simpson’s paradox.)
9.20 Psychological and social factors can influence the survival of patients with serious diseases. One study examined the relationship between survival of patients with coronary heart disease (CHD) and pet ownership. (Erika Friedmann et al., “Animal companions and one-year survival of patients after discharge from a coronary care unit,” *Public Health Reports*, 96 (1980), pp. 307–312.) Each of 92 patients was classified as having a pet or not and by whether they survived for one year. Here are the data:

<table>
<thead>
<tr>
<th>Patient status</th>
<th>Pet ownership</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Alive</td>
<td>28</td>
</tr>
<tr>
<td>Dead</td>
<td>11</td>
</tr>
</tbody>
</table>

(a) Was this study an experiment? Why or why not?

(b) The researchers thought that having a pet might improve survival, so pet ownership is the explanatory variable. Compute appropriate percents to describe the data and state your preliminary findings.

(c) State in words the null hypothesis for this problem. What is the alternative hypothesis?

(d) Find the $X^2$ statistic, its degrees of freedom, and the $P$-value.

(e) What do you conclude? Do the data give convincing evidence that owning a pet is an effective treatment for increasing the survival of CHD patients?

9.21 The baseball player Reggie Jackson had a reputation for hitting better in the World Series than during the regular season. In his 21-year career, Jackson was at bat 9864 times in regular-season play and had 2584 hits. During World Series games, he was at bat 98 times and had 35 hits. We can view Jackson’s regular-season at-bats as a random sample from a population of potential at-bats (he might have batted many more times if the season were longer, for example), and his World Series at-bats as a sample from a second population.

(a) Display the data in a $2 \times 2$ table of counts with “regular season” and “World
Series” as the column headings, and fill in the marginal sums.

(b) Calculate appropriate percents to compare Jackson’s regular-season and World Series performances. Did he hit better in World Series games?

(c) Is there a significant difference between Jackson’s regular-season and World Series performances? State hypotheses (in words), and then calculate the $X^2$ statistic, its degrees of freedom, and its $P$-value. What is your conclusion?

9.22 If the performance of a stock fund is due to the skill of the manager, then we would expect a fund that does well this year to perform well next year also. This is called persistence of fund performance. One study classified funds as losers or winners depending on whether their rate of return was less than or greater than the median of all funds. (Burton G. Malkiel, “Returns from investing in equity mutual funds, 1971 to 1991,” *Journal of Finance*, 50 (1995), pp. 549–572.) To examine the question of interest we form a two-way table that classifies each fund as a loser or winner in each of two successive years. Here are the data for one such table:

<table>
<thead>
<tr>
<th></th>
<th>Next year</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>This year</td>
<td></td>
<td>Winner</td>
</tr>
<tr>
<td>Winner</td>
<td>85</td>
<td>35</td>
</tr>
<tr>
<td>Loser</td>
<td>37</td>
<td>83</td>
</tr>
</tbody>
</table>

Is there evidence in favor of persistence of fund performance in this table? Support your conclusion with a complete analysis of the data.

9.23 Refer to the previous exercise. Rerun the analysis using the method for comparing two proportions of Chapter 8. Verify that the $X^2$ statistic is the square of the $z$ statistic and that the $P$-values for both analyses are the same.

9.24 In the previous exercise, it is natural to use “this year” to define the two samples. If we drew separate random samples of winners and losers this year and we recorded the outcome next year, we would call this a prospective study (forward looking). On the other hand, if we drew separate random samples of winners and
losers “next year” and looked back historically to determine if they were winners or losers in the previous year, we would have a **retrospective study** (backward looking). Verify that you get the same value of $z$ (and therefore the same $P$-value) using these two different approaches.

9.25 If we find evidence in favor of an effect in one set of circumstances, it is natural to want to conclude that it holds in many others. Unfortunately, this reasoning can sometimes lead us to incorrect conclusions. For example, here is another table from the study described in Exercise 9.22:

<table>
<thead>
<tr>
<th></th>
<th>Next year</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>This year</td>
<td>Winner</td>
<td>Loser</td>
</tr>
<tr>
<td>Winner</td>
<td>96</td>
<td>148</td>
</tr>
<tr>
<td>Loser</td>
<td>145</td>
<td>99</td>
</tr>
</tbody>
</table>

Analyze these data in the same way. What do you conclude?

9.26 There is much evidence that high blood pressure is associated with increased risk of death from cardiovascular disease. A major study of this association examined 2676 men with low blood pressure and 3338 men with high blood pressure. (J. Stamler, “The mass treatment of hypertensive disease: defining the problem,” *Mild Hypertension: To Treat or Not to Treat*, New York Academy of Sciences, 1978, pp. 333–358.) During the period of the study, 21 men in the low-blood-pressure and 55 in the high-blood-pressure group died from cardiovascular disease.

(a) What is the explanatory variable? Describe the association in these data numerically and in words.

(b) Do the study data confirm that death rates are higher among men with high blood pressure? State hypotheses, carry out a significance test, and give your conclusions.

(c) Present the data in a two-way table. Is the chi-square test appropriate for the hypotheses you stated in (b)?
(d) Give a 95% confidence interval for the difference between the death rates for the low- and high-blood-pressure groups.

9.27 It is traditional practice in Egypt to withhold food from children with diarrhea. Because it is known that feeding children with this illness reduces mortality, medical authorities undertook a nationwide program designed to promote feeding sick children. To evaluate the impact of the program, surveys were taken before and after the program was implemented. (O. M. Galal et al., “Feeding the child with diarrhea: a strategy for testing a health education message within the primary health care system in Egypt,” *Socio-economic Planning Sciences*, 21 (1987), pp. 139–147.) In the first survey, 457 of 1003 surveyed mothers followed the practice of feeding children with diarrhea. For the second survey, 437 of 620 surveyed followed this practice.

(a) Assume that the data come from two independent samples. Test the hypothesis that the program was effective, that is, that the practice of feeding children with diarrhea increased between the time of the first study and the time of the second. State $H_0$ and $H_a$, give the test statistic and its $P$-value, and summarize your conclusion.

(b) Present the data in a two-way table. Can the $X^2$ statistic test your hypotheses?

(c) Describe the results using a 95% confidence interval for the difference in proportions.

9.28 PTC is a compound that has a strong bitter taste for some people and is tasteless for others. The ability to taste this compound is an inherited trait. Many studies have assessed the proportions of people in different populations who can taste PTC. The following table gives results for samples from several countries: (A. E. Mourant et al., *The Distribution of Human Blood Groups and Other Polymorphisms*, Oxford University Press, 1976.)
Complete the table and describe the data. Do they provide evidence that the proportion of PTC tasters varies among the four countries? Give a complete summary of your analysis.

9.29 There are four major blood types in humans: O, A, B, and AB. In a study conducted using blood specimens from the Blood Bank of Hawaii, individuals were classified according to blood type and ethnic group. The ethnic groups were Hawaiian, Hawaiian-white, Hawaiian-Chinese, and white. (A. E. Mourant et al., The Distribution of Human Blood Groups and Other Polymorphisms, Oxford University Press, 1976.) Assume that the blood bank specimens are random samples from the Hawaiian populations of these ethnic groups.

<table>
<thead>
<tr>
<th>Ethnic group</th>
<th>Hawaiian-</th>
<th>Hawaiian-</th>
<th>Hawaiian-</th>
<th>Hawaiian-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood type</td>
<td>Hawaiian</td>
<td>white</td>
<td>Chinese</td>
<td>White</td>
</tr>
<tr>
<td>O</td>
<td>1,903</td>
<td>4,469</td>
<td>2,206</td>
<td>53,759</td>
</tr>
<tr>
<td>A</td>
<td>2,490</td>
<td>4,671</td>
<td>2,368</td>
<td>50,008</td>
</tr>
<tr>
<td>B</td>
<td>178</td>
<td>606</td>
<td>568</td>
<td>16,252</td>
</tr>
<tr>
<td>AB</td>
<td>99</td>
<td>236</td>
<td>243</td>
<td>5,001</td>
</tr>
</tbody>
</table>

Summarize the data numerically and with a graph. Is there evidence to conclude that blood type and ethnic group are related? Explain how you arrived at your conclusion.

9.30 In healthy individuals the concentration of various substances in the blood remains within relatively narrow bounds. One such substance is potassium. A person is said to be hypokalemic if the potassium level is too low (less than 3.5 mil-
Hypokalemia is associated with a variety of symptoms such as excessive tiredness, while hyperkalemia is generally an indication of a serious problem. Patients being treated with diuretics (pharmaceuticals that help the body to eliminate water) sometimes have abnormal potassium concentrations. In a large study of patients on chronic diuretic therapy, several risk factors were studied to see if they were associated with abnormal potassium levels. (William M. Tierney, Clement J. McDonald, and George P. McCabe, “Serum potassium testing in diuretic-treated outpatients,” Medical Decision Making, 5 (1985), pp. 91–104.) Of the 5810 patients studied, 1094 were hypokalemic, 4689 had normal potassium levels, and 27 were hyperkalemic. The following table gives the percents of patients having each of four risk factors in the three potassium groups:

<table>
<thead>
<tr>
<th>Potassium group</th>
<th>Hypokalemic</th>
<th>Normal</th>
<th>Hyperkalemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1094</td>
<td>4689</td>
<td>27</td>
</tr>
<tr>
<td>Hypertension</td>
<td>88.3%</td>
<td>78.1%</td>
<td>40.7%</td>
</tr>
<tr>
<td>Heart failure</td>
<td>16.5%</td>
<td>24.7%</td>
<td>55.6%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>20.6%</td>
<td>25.5%</td>
<td>29.6%</td>
</tr>
<tr>
<td>Gender (% female)</td>
<td>72.5%</td>
<td>68.0%</td>
<td>48.1%</td>
</tr>
</tbody>
</table>

For example, 88.3% of the 1094 hypokalemic patients had hypertension, and 78.1% of the 4689 normal patients had hypertension. For each of the four risk factors, use the percents and n’s given to compute the counts for the $2 \times 3$ table needed to study the association between the factor and potassium. Then analyze each table using the methods presented in this chapter. Note that there are very few patients in the hyperkalemic group. Therefore, reanalyze the data dropping this category from the tables. Write a short summary explaining what you have found.
9.31 The proportion of women entering many professions has undergone considerable change in recent years. A study of students enrolled in pharmacy programs describes the changes in this field. A random sample of 700 students in their third or higher year of study at colleges of pharmacy was taken in each of nine years. (The data are based on *Seventh Report to the President and Congress on the Status of Health Personnel in the United States*, Public Health Service, 1990.) The following table gives the numbers of women in each of these samples:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>164</td>
<td>195</td>
<td>226</td>
<td>283</td>
<td>302</td>
<td>342</td>
<td>369</td>
<td>385</td>
<td>412</td>
</tr>
</tbody>
</table>

Use the chi-square test to assess the change in the percent of women pharmacy students over time, and summarize your results. (You will need to calculate the number of male students for each year using the fact that the sample size each year is 700.) Plot the percent of women versus year. Describe the plot. Is it roughly linear? Find the least-squares line that summarizes the relation between time and the percent of women pharmacy students.

9.32 Refer to the previous exercise. Here are the percents of women pharmacy students for the years 1987 to 2000: Data provided by Dr. Susan Meyer, Senior Vice President of the American Association of Colleges of Pharmacy.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>60.0%</td>
<td>60.6%</td>
<td>61.6%</td>
<td>62.4%</td>
<td>63.0%</td>
<td>63.4%</td>
<td>63.2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>63.3%</td>
<td>63.4%</td>
<td>63.8%</td>
<td>64.2%</td>
<td>64.4%</td>
<td>64.9%</td>
<td>65.9%</td>
</tr>
</tbody>
</table>

Plot these percents versus year and summarize the pattern. Using your analysis of the data in this and the previous exercise, write a report summarizing the changes that have occurred in the percent of women pharmacy students from 1970 to 2000. Include an estimate of the percent for the year 2010 with an explanation of why you chose this estimate.
Chapter 10 Exercises

10.1 Manatees are large sea creatures that live in the shallow water along the coast of Florida. Many manatees are injured or killed each year by powerboats. Exercise S2.11 gives data on manatees killed and powerboat registrations (in thousands of boats) in Florida for the period 1977 to 1990.
(a) Make a scatterplot of boats registered and manatees killed. Is there a strong straight-line pattern?
(b) Find the equation of the least-squares regression line. Draw this line on your scatterplot.
(c) Is there strong evidence that the mean number of manatees killed increases as the number of powerboats increases? State this question as null and alternative hypotheses about the slope of the population regression line, obtain the $t$ statistic, and give your conclusion.
(d) Predict the number of manatees that will be killed if there are 716,000 powerboats registered. In 1991, 1992, and 1993, the number of powerboats remained at 716,000. The numbers of manatees killed were 53, 38, and 35. Compare your prediction with these data. Does the comparison suggest that measures taken to protect the manatees in these years were effective?

10.2 Can a pretest on mathematics skills predict success in a statistics course? The 55 students in an introductory statistics class took a pretest at the beginning of the semester. The least-squares regression line for predicting the score $y$ on the final exam from the pretest score $x$ was $\hat{y} = 10.5 + 0.82x$. The standard error of $b_1$ was 0.38. Test the null hypothesis that there is no linear relationship between the pretest score and the score on the final exam against the two-sided alternative.
10.3 Exercise 2.56 (page xxx) gives the following data from a study of two methods for measuring the blood flow in the stomachs of dogs:

<table>
<thead>
<tr>
<th>Spheres</th>
<th>4.0</th>
<th>4.7</th>
<th>6.3</th>
<th>8.2</th>
<th>12.0</th>
<th>15.9</th>
<th>17.4</th>
<th>18.1</th>
<th>20.2</th>
<th>23.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vein</td>
<td>3.3</td>
<td>8.3</td>
<td>4.5</td>
<td>9.3</td>
<td>10.7</td>
<td>16.4</td>
<td>15.4</td>
<td>17.6</td>
<td>21.0</td>
<td>21.7</td>
</tr>
</tbody>
</table>

“Spheres” is an experimental method that the researchers hope will predict “Vein,” the standard but difficult method. Examination of the data gives no reason to doubt the validity of the simple linear regression model. The estimated regression line is \( \hat{y} = 1.031 + 0.902x \), where \( y \) is the response variable Vein and \( x \) is the explanatory variable Spheres. The estimate of \( \sigma \) is \( s = 1.757 \).

(a) Find \( \bar{x} \) and \( \sum (x_i - \bar{x})^2 \) from the data.

(b) We expect \( x \) and \( y \) to be positively associated. State hypotheses in terms of the slope of the population regression line that express this expectation, and carry out a significance test. What conclusion do you draw?

(c) Find a 99% confidence interval for the slope.

(d) Suppose that we observe a value of Spheres equal to 15.0 for one dog. Give a 90% interval for predicting the variable Vein for that dog.

10.4 Ohm’s law \( I = V/R \) states that the current \( I \) in a metal wire is proportional to the voltage \( V \) applied to its ends and is inversely proportional to the resistance \( R \) in the wire. Students in a physics lab performed experiments to study Ohm’s law. They varied the voltage and measured the current at each voltage with an ammeter. The goal was to determine the resistance \( R \) of the wire. We can rewrite Ohm’s law in the form of a linear regression as \( I = \beta_0 + \beta_1 V \), where \( \beta_0 = 0 \) and \( \beta_1 = 1/R \). Because voltage is set by the experimenter, we think of \( V \) as the explanatory variable.

The current \( I \) is the response. Here are the data for one experiment (data provided by Sara McCabe):
Exercises

<table>
<thead>
<tr>
<th>$V$</th>
<th>0.50</th>
<th>1.00</th>
<th>1.50</th>
<th>1.80</th>
<th>2.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I$</td>
<td>0.52</td>
<td>1.19</td>
<td>1.62</td>
<td>2.00</td>
<td>2.40</td>
</tr>
</tbody>
</table>

(a) Plot the data. Are there any outliers or unusual points?

(b) Find the least-squares fit to the data, and estimate $1/R$ for this wire. Then give a 95% confidence interval for $1/R$.

(c) If $b_1$ estimates $1/R$, then $1/b_1$ estimates $R$. Estimate the resistance $R$. Similarly, if $L$ and $U$ represent the lower and upper confidence limits for $1/R$, then the corresponding limits for $R$ are given by $1/U$ and $1/L$, as long as $L$ and $U$ are positive. Use this fact and your answer to (b) to find a 95% confidence interval for $R$.

(d) Ohm’s law states that $\beta_0$ in the model is 0. Calculate the test statistic for this hypothesis and give an approximate $P$-value.

10.5 Most statistical software systems have an option for doing regressions in which the intercept is set in advance at 0. If you have access to such software, reanalyze the Ohm’s law data given in the previous exercise with this option and report the estimate of $R$. The output should also include an estimated standard error for $1/R$. Use this to calculate the 95% confidence interval for $R$. Note: With this option the degrees of freedom for $t^*$ will be 1 greater than for the model with the intercept.

10.6 Return to the data on current versus voltage given in the Ohm’s law experiment of Exercise S10.4.

(a) Compute all values for the ANOVA table.

(b) State the null hypothesis tested by the ANOVA $F$ statistic, and explain in plain language what this hypothesis says.

(c) What is the distribution of this $F$ statistic when $H_0$ is true? Find an approximate $P$-value for the test of $H_0$. 


10.7 Here are the golf scores of 12 members of a college women’s golf team in two rounds of tournament play. (A golf score is the number of strokes required to complete the course, so that low scores are better.)

<table>
<thead>
<tr>
<th>Player</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>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round 1</td>
<td>89</td>
<td>90</td>
<td>87</td>
<td>95</td>
<td>86</td>
<td>81</td>
<td>102</td>
<td>105</td>
<td>83</td>
<td>88</td>
<td>91</td>
<td>79</td>
</tr>
<tr>
<td>Round 2</td>
<td>94</td>
<td>85</td>
<td>89</td>
<td>89</td>
<td>81</td>
<td>76</td>
<td>107</td>
<td>89</td>
<td>87</td>
<td>91</td>
<td>88</td>
<td>80</td>
</tr>
</tbody>
</table>

(a) Plot the data and describe the relationship between the two scores.

(b) Find the correlation between the two scores and test the null hypothesis that the population correlation is 0. Summarize your results.

(c) The plot shows one outlier. Recompute the correlation and redo the significance test without this observation. Write a short summary explaining the effect of the outlier on the correlation and significance test in (b).

10.8 A study reported a correlation \( r = 0.5 \) based on a sample size of \( n = 20 \); another reported the same correlation based on a sample size of \( n = 10 \). For each, perform the test of the null hypothesis that \( \rho = 0 \). Describe the results and explain why the conclusions are different.

10.9 Returns on common stocks in the United States and overseas appear to be growing more closely correlated as economies become more interdependent. Suppose that this population regression line connects the total annual returns (in percent) of two indexes of stock prices:

\[
\text{MEAN OVERSEAS RETURN} = 4.7 + 0.66 \times \text{U.S. RETURN}
\]

(a) What is \( \beta_0 \) in this line? What does this number say about overseas returns when the U.S. market is flat (0% return)?

(b) What is \( \beta_1 \) in this line? What does this number say about the relationship between U.S. and overseas returns?

(c) We know that overseas returns will vary during years that have the same return
on U.S. common stocks. Write the regression model based on the population regression line given above. What part of this model allows overseas returns to vary when U.S. returns remain the same?

10.10 How well does the number of beers a student drinks predict his or her blood alcohol content? Sixteen student volunteers at Ohio State University drank a randomly assigned number of cans of beer. Thirty minutes later, a police officer measured their blood alcohol content (BAC). Here are the data: (These are part of the data from the EESEE story “Blood Alcohol Content,” found on the IPS Web site www.whfreeman.com/ips.)

<table>
<thead>
<tr>
<th>Student</th>
<th>Beers</th>
<th>BAC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>0.10</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.03</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>0.19</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>0.12</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>0.04</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>0.095</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>0.07</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>0.06</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>0.02</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>0.05</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>0.07</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>0.10</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>0.085</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>0.09</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0.01</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>0.05</td>
</tr>
</tbody>
</table>

The students were equally divided between men and women and differed in weight and usual drinking habits. Because of this variation, many students don’t believe that number of drinks predicts blood alcohol well.

(a) Make a scatterplot of the data. Find the equation of the least-squares regression line for predicting blood alcohol from number of beers and add this line to your plot. What is \( r^2 \) for these data? Briefly summarize what your data analysis shows.

(b) Is there significant evidence that drinking more beers increases blood alcohol on the average in the population of all students? State hypotheses, give a test statistic and \( P \)-value, and state your conclusion.

10.11 Your scatterplot in the previous exercise shows one unusual point: student number 3, who drank 9 beers.
(a) Does student 3 have the largest residual from the fitted line? (You can use the scatterplot to see this.) Is this observation extreme in the $x$ direction, so that it may be influential?

(b) Do the regression again, omitting student 3. Add the new regression line to your scatterplot. Does removing this observation greatly change predicted BAC? Does $r^2$ change greatly? Does the $P$-value of your test change greatly? What do you conclude: did your work in the previous problem depend heavily on this one student?

10.12 Utility companies need to estimate the amount of energy that will be used by their customers. The consumption of natural gas required for heating homes depends on the outdoor temperature. When the weather is cold, more gas will be consumed. A study of one home recorded the average daily gas consumption $y$ (in hundreds of cubic feet) for each month during one heating season. The explanatory variable $x$ is the average number of heating degree-days per day during the month. One heating degree-day is accumulated for each degree a day’s average temperature falls below 65° F. An average temperature of 50°, for example, corresponds to 15 degree-days. The data for October through June are given in the following table:

(Data were provided by Professor Robert Dale of the Purdue University Agronomy Department.)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree-days</td>
<td>15.6</td>
<td>26.8</td>
<td>37.8</td>
<td>36.4</td>
<td>35.5</td>
<td>18.6</td>
<td>15.3</td>
<td>7.9</td>
<td>0.0</td>
</tr>
<tr>
<td>Gas consumption</td>
<td>5.2</td>
<td>6.1</td>
<td>8.7</td>
<td>8.5</td>
<td>8.8</td>
<td>4.9</td>
<td>4.5</td>
<td>2.5</td>
<td>1.1</td>
</tr>
</tbody>
</table>

(a) Find the equation of the least-squares line.

(b) Test the null hypothesis that the slope is zero and describe your conclusion.

(c) Give a 90% confidence interval for the slope.

(d) The parameter $\beta_0$ corresponds to natural gas consumption for cooking, hot
water, and other uses when there is no demand for heating. Give a 90% confidence interval for this parameter.

10.13 The previous exercise demonstrates that there is a strong linear relationship between household consumption of natural gas and outdoor temperature, measured by heating degree-days. The slope and intercept depend on the particular house and on the habits of the household living there. Data for two heating seasons (18 months) for another household produce the least-squares line $\hat{y} = 2.405 + 0.26896x$ for predicting average daily gas consumption $y$ from average degree-days per day $x$.

The standard error of the slope is $SE_{b_1} = 0.00815$.

(a) Explain briefly what the slope $\beta_1$ of the population regression line represents. Then give a 90% confidence interval for $\beta_1$.

(b) This interval is based on twice as many observations as the one calculated in the previous exercise for a different household, and the two standard errors are of similar size. How would you expect the margins of error of the two intervals to be related? Check your answer by comparing the two margins of error.

10.14 The standard error of the intercept in the regression of gas consumption on degree-days for the household in the previous exercise is $SE_{b_0} = 0.20351$.

(a) Explain briefly what the intercept represents in this setting. Find a 90% confidence interval for the intercept.

(b) Compare the width of your interval with the one calculated for a different household in Exercise 10.12. Explain why it is narrower.

10.15 Exercise 10.12 gives information about the regression of natural gas consumption on degree-days for a particular household.

(a) What is the $t$ statistic for testing $H_0: \beta_1 = 0$?

(b) For the alternative $H_a: \beta_1 > 0$, what critical value would you use for a test at the $\alpha = 0.05$ significance level? Do you reject $H_0$ at this level?

(c) How would you report the $P$-value for this test?
10.16 Can a pretest on mathematics skills predict success in a statistics course? The 102 students in an introductory statistics class took a pretest at the beginning of the semester. The least-squares regression line for predicting the score $y$ on the final exam from the pretest score $x$ was $\hat{y} = 10.5 + 0.73x$. The standard error of $b_1$ was 0.42.

(a) Test the null hypothesis that there is no linear relationship between the pretest score and the score on the final exam against the two-sided alternative.

(b) Would you reject this null hypothesis versus a two-sided alternative? Explain your answer.

10.17 The human body takes in more oxygen when exercising than when it is at rest. To deliver the oxygen to the muscles, the heart must beat faster. Heart rate is easy to measure, but measuring oxygen uptake requires elaborate equipment. If oxygen uptake (VO2) can be accurately predicted from heart rate (HR), the predicted values can replace actually measured values for various research purposes. Unfortunately, not all human bodies are the same, so no single prediction equation works for all people. Researchers can, however, measure both HR and VO2 for one person under varying sets of exercise conditions and calculate a regression equation for predicting that person’s oxygen uptake from heart rate. They can then use predicted oxygen uptakes in place of measured uptakes for this individual in later experiments. (These data are from experiments conducted in Don Corrigan’s laboratory at Purdue University and were provided by Paul Waldsmith.) Here are data for one individual:

<table>
<thead>
<tr>
<th>HR</th>
<th>94</th>
<th>96</th>
<th>95</th>
<th>95</th>
<th>94</th>
<th>95</th>
<th>94</th>
<th>104</th>
<th>104</th>
<th>106</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO2</td>
<td>0.473</td>
<td>0.753</td>
<td>0.929</td>
<td>0.939</td>
<td>0.832</td>
<td>0.983</td>
<td>1.049</td>
<td>1.178</td>
<td>1.176</td>
<td>1.292</td>
</tr>
<tr>
<td>HR</td>
<td>108</td>
<td>110</td>
<td>113</td>
<td>113</td>
<td>118</td>
<td>115</td>
<td>121</td>
<td>127</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>VO2</td>
<td>1.403</td>
<td>1.499</td>
<td>1.529</td>
<td>1.599</td>
<td>1.749</td>
<td>1.746</td>
<td>1.897</td>
<td>2.040</td>
<td>2.231</td>
<td></td>
</tr>
</tbody>
</table>
Exercises

(a) Plot the data. Are there any outliers or unusual points?

(b) Compute the least-squares regression line for predicting oxygen uptake from heart rate for this individual.

(c) Test the null hypothesis that the slope of the regression line is 0. Explain in words the meaning of your conclusion from this test.

(d) Calculate a 95% interval for the oxygen uptake of this individual on a future occasion when his heart rate is 96. Repeat the calculation for a heart rate of 115.

(e) From what you have learned in (a), (b), (c), and (d) of this exercise, do you think that the researchers should use predicted VO2 in place of measured VO2 for this individual under similar experimental conditions? Explain your answer.

10.18 Premature infants are often kept in intensive-care nurseries after they are born. It is common practice to measure their blood pressure frequently. The oscillometric method of measuring blood pressure is noninvasive and easy to use. The traditional procedure, called the direct intra-arterial method, is believed to be more accurate but is invasive and more difficult to perform. Several studies have reported high correlations between measurements made by the two methods, ranging from \( r = 0.49 \) to \( r = 0.98 \). These correlations are statistically significant. One study that investigated the relation between the two methods reported the regression equation \( \hat{y} = 15 + 0.83x \). Here \( x \) represents the easy method and \( y \) represents the difficult one. (John A. Wareham et al., “Prediction of arterial blood pressure on the premature neonate using the oscillometric method,” *American Journal of Diseases of Children*, 141 (1987), pp. 1108–1110.) The standard error of the slope is 0.065 and the sample size is 81. Calculate the \( t \) statistic for testing \( H_0: \beta_1 = 0 \). Specify an appropriate alternative hypothesis for this problem, and give an approximate \( P \)-value for the test. Then explain your conclusion in words a physician can understand. (The authors of the study calculated and plotted prediction intervals. They found the widths to be
unacceptably large and concluded that statistical significance does not imply that results are clinically useful.)

10.19 Soil aeration and soil water evaporation involve the exchange of gases between the soil and the atmosphere. Experimenters have investigated the effect of the airflow above the soil on this process. One such experiment varied the speed of the air $x$ and measured the rate of evaporation $y$. The fitted regression equation based on 18 observations was $\hat{y} = 5.0 + 0.00665x$. The standard error of the slope was reported to be 0.00182.

(a) It is reasonable to suppose that greater airflow will cause more evaporation. State hypotheses to test this belief and calculate the test statistic. Find an approximate $P$-value for the significance test and report your conclusion.

(b) Construct a 90% confidence interval for the additional evaporation experienced when airflow increases by 1 unit.

10.20 Here are data on the average yield in bushels per acre for corn in the United States.

<table>
<thead>
<tr>
<th>Year</th>
<th>Yield</th>
<th>Year</th>
<th>Yield</th>
<th>Year</th>
<th>Yield</th>
<th>Year</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>1957</td>
<td>48.3</td>
<td>1968</td>
<td>79.5</td>
<td>1979</td>
<td>109.5</td>
<td>1990</td>
<td>118.5</td>
</tr>
<tr>
<td>1958</td>
<td>52.8</td>
<td>1969</td>
<td>85.9</td>
<td>1980</td>
<td>91.0</td>
<td>1991</td>
<td>108.6</td>
</tr>
<tr>
<td>1960</td>
<td>54.7</td>
<td>1971</td>
<td>88.1</td>
<td>1982</td>
<td>113.2</td>
<td>1993</td>
<td>100.7</td>
</tr>
<tr>
<td>1961</td>
<td>62.4</td>
<td>1972</td>
<td>97.0</td>
<td>1983</td>
<td>81.1</td>
<td>1994</td>
<td>138.6</td>
</tr>
<tr>
<td>1962</td>
<td>64.7</td>
<td>1973</td>
<td>91.3</td>
<td>1984</td>
<td>106.7</td>
<td>1995</td>
<td>113.5</td>
</tr>
<tr>
<td>1963</td>
<td>67.9</td>
<td>1974</td>
<td>71.9</td>
<td>1985</td>
<td>118.0</td>
<td>1996</td>
<td>127.1</td>
</tr>
<tr>
<td>1964</td>
<td>62.9</td>
<td>1975</td>
<td>86.4</td>
<td>1986</td>
<td>119.4</td>
<td>1997</td>
<td>126.7</td>
</tr>
<tr>
<td>1965</td>
<td>74.1</td>
<td>1976</td>
<td>88.0</td>
<td>1987</td>
<td>119.8</td>
<td>1998</td>
<td>134.4</td>
</tr>
<tr>
<td>1966</td>
<td>73.1</td>
<td>1977</td>
<td>90.8</td>
<td>1988</td>
<td>84.6</td>
<td>1999</td>
<td>133.8</td>
</tr>
<tr>
<td>1967</td>
<td>80.1</td>
<td>1978</td>
<td>101.0</td>
<td>1989</td>
<td>116.3</td>
<td>2000</td>
<td>136.9</td>
</tr>
</tbody>
</table>
(a) Plot the yield versus year. Describe the relationship. Are there any outliers or unusual years?

(b) Perform the regression analysis and summarize the results. How rapidly has yield increased over time?

10.21 Refer to the previous exercise. Give a 95% prediction interval for the yield in the year 2006.

10.22 In the previous exercise you examined the relationship between time and the yield of corn in the United States. Here are similar data for the yield of soybeans (in bushels per acre).

<table>
<thead>
<tr>
<th>Year</th>
<th>Yield</th>
<th>Year</th>
<th>Yield</th>
<th>Year</th>
<th>Yield</th>
<th>Year</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>1957</td>
<td>23.2</td>
<td>1968</td>
<td>26.7</td>
<td>1979</td>
<td>32.1</td>
<td>1990</td>
<td>34.1</td>
</tr>
<tr>
<td>1958</td>
<td>24.2</td>
<td>1969</td>
<td>27.4</td>
<td>1980</td>
<td>26.5</td>
<td>1991</td>
<td>34.2</td>
</tr>
<tr>
<td>1959</td>
<td>23.5</td>
<td>1970</td>
<td>26.7</td>
<td>1981</td>
<td>30.1</td>
<td>1992</td>
<td>37.6</td>
</tr>
<tr>
<td>1960</td>
<td>23.5</td>
<td>1971</td>
<td>27.5</td>
<td>1982</td>
<td>31.5</td>
<td>1993</td>
<td>32.6</td>
</tr>
<tr>
<td>1961</td>
<td>25.1</td>
<td>1972</td>
<td>27.8</td>
<td>1983</td>
<td>26.2</td>
<td>1994</td>
<td>41.4</td>
</tr>
<tr>
<td>1963</td>
<td>24.4</td>
<td>1974</td>
<td>23.7</td>
<td>1985</td>
<td>34.1</td>
<td>1996</td>
<td>37.6</td>
</tr>
<tr>
<td>1964</td>
<td>22.8</td>
<td>1975</td>
<td>28.9</td>
<td>1986</td>
<td>33.3</td>
<td>1997</td>
<td>38.9</td>
</tr>
<tr>
<td>1965</td>
<td>24.5</td>
<td>1976</td>
<td>26.1</td>
<td>1987</td>
<td>33.9</td>
<td>1998</td>
<td>38.9</td>
</tr>
<tr>
<td>1966</td>
<td>25.4</td>
<td>1977</td>
<td>30.6</td>
<td>1988</td>
<td>27.0</td>
<td>1999</td>
<td>36.6</td>
</tr>
<tr>
<td>1967</td>
<td>24.5</td>
<td>1978</td>
<td>29.4</td>
<td>1989</td>
<td>32.3</td>
<td>2000</td>
<td>38.1</td>
</tr>
</tbody>
</table>

Give a complete analysis of these data. Include a plot of the data, significance test results, examination of the residuals, and your conclusions.

10.23 The corn yields of Exercise 10.20 and the soybean yields of Exercise 10.2 both vary over time for similar reasons, including improved technology and weather conditions. Let’s examine the relationship between the two yields.
(a) Plot the two yields with corn on the $x$ axis and soybeans on the $y$ axis. Describe the relationship.

(b) Find the correlation. How well does it summarize the relation?

(c) Use corn yield to predict soybean yield. Give the equation and the results of the significance test for the slope. This test also tests the null hypothesis that the two yields are uncorrelated.

(d) Obtain the residuals from the model in part (c) and plot them versus time. Describe the pattern.

10.24 The corn yield data of Exercise 10.20 show a larger amount of scatter about the least-squares line for the later years, when the yields are higher. This may be an indication that the standard deviation $\sigma$ of our model is not a constant but is increasing with time. Take logs of the yields and rerun the analyses. Prepare a short report comparing the two analyses. Include plots, a comparison of the significance test results, and the percent of variation explained by each model.

10.25 Exercise 10.10 gives data from measuring the blood alcohol content (BAC) of students 30 minutes after they drank an assigned number of cans of beer. Steve thinks he can drive legally 30 minutes after he drinks 5 beers. The legal limit is $\text{BAC} = 0.08$. Give a 95% confidence interval for Steve’s BAC. Can he be confident he won’t be arrested if he drives and is stopped?

10.26 Return to the oxygen uptake and heart rate data given in Exercise 10.17.

(a) Construct the ANOVA table.

(b) What null hypothesis is tested by the ANOVA $F$ statistic? What does this hypothesis say in practical terms?

(c) Give the degrees of freedom for the $F$ statistic and an approximate $P$-value for the test of $H_0$.

(d) Verify that the square of the $t$ statistic that you calculated in Exercise 10.20 is equal to the $F$ statistic in your ANOVA table. (Any difference found is due to
(e) What proportion of the variation in oxygen uptake is explained by heart rate for this set of data?

10.27 A study conducted in the Egyptian village of Kalama examined the relationship between the birth weights of 40 infants and various socioeconomic variables. (M. El-Kholy, F. Shaheen, and W. Mahmoud, “Relationship between socioeconomic status and birth weight, a field study in a rural community in Egypt,” *Journal of the Egyptian Public Health Association*, 61 (1986), pp. 349–358.)

(a) The correlation between monthly income and birth weight was $r = 0.39$. Calculate the $t$ statistic for testing the null hypothesis that the correlation is 0 in the entire population of infants.

(b) The researchers expected that higher birth weights would be associated with higher incomes. Express this expectation as an alternative hypothesis for the population correlation.

(c) Determine a $P$-value for $H_0$ versus the alternative that you specified in (b). What conclusion does your test suggest?

10.28 Chinese students from public schools in Hong Kong were the subjects of a study designed to investigate the relationship between various measures of parental behavior and other variables. The sample size was 713. The data were obtained from questionnaires filled in by the students. One of the variables examined was parental control, an indication of the amount of control that the parents exercised over the behavior of the students. Another was the self-esteem of the students. (S. Lau and P. C. Cheung, “Relations between Chinese adolescents’ perception of parental control and organization and their perception of parental warmth,” *Developmental Psychology*, 23 (1987), pp. 726–729.)

(a) The correlation between parental control and self-esteem was $r = -0.19$. Calculate the $t$ statistic for testing the null hypothesis that the population correlation
is 0.

(b) Find an approximate $P$-value for testing $H_0$ versus the two-sided alternative and report your conclusion.

10.29 The data on gas consumption and degree-days from Exercise 10.12 are as follows:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree-days</td>
<td>15.6</td>
<td>26.8</td>
<td>37.8</td>
<td>36.4</td>
<td>35.5</td>
<td>18.6</td>
<td>15.3</td>
<td>7.9</td>
<td>0.0</td>
</tr>
<tr>
<td>Gas consumption</td>
<td>5.2</td>
<td>6.1</td>
<td>8.7</td>
<td>8.5</td>
<td>8.8</td>
<td>4.9</td>
<td>4.5</td>
<td>2.5</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Suppose that the gas consumption for January was incorrectly recorded as 85 instead of 8.5.

(a) Calculate the least-squares regression line for the incorrect set of data.

(b) Find the standard error of $b_1$.

(c) Compute the test statistic for $H_0: \beta_1 = 0$ and find the $P$-value. How do the results compare with those for the correct set of data?

10.30 *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 are different species rather than individuals from the same 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: (Marilyn A. Houck et al., “Allometric scaling in the earliest fossil bird, *Archaeopteryx lithographica*,” *Science*, 247 (1990), pp. 195–198. The authors conclude from a variety of evidence that all specimens represent the same species.)

<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>

(a) Plot the data and describe the pattern. Is it reasonable to summarize this kind of relationship with a correlation?
(b) Find the correlation and perform the significance test. Summarize the results and report your conclusion.

**10.31** How does the fuel consumption of a car change as its speed increases? Here are data for a British Ford Escort. Speed is measured in kilometers per hour, and fuel consumption is measured in liters of gasoline used per 100 kilometers traveled. (Based on T. N. Lam, “Estimating fuel consumption from engine size,” *Journal of Transportation Engineering, 111* (1985), pp. 339–357. The data for 10 to 50 km/h are measured; those for 60 and higher are calculated from a model given in the paper and are therefore smoothed.)

<table>
<thead>
<tr>
<th>Speed (km/h)</th>
<th>Fuel used (liters/100 km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>21.00</td>
</tr>
<tr>
<td>20</td>
<td>13.00</td>
</tr>
<tr>
<td>30</td>
<td>10.00</td>
</tr>
<tr>
<td>40</td>
<td>8.00</td>
</tr>
<tr>
<td>50</td>
<td>7.00</td>
</tr>
<tr>
<td>60</td>
<td>5.90</td>
</tr>
<tr>
<td>70</td>
<td>6.30</td>
</tr>
<tr>
<td>80</td>
<td>6.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Speed (km/h)</th>
<th>Fuel used (liters/100 km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>7.57</td>
</tr>
<tr>
<td>100</td>
<td>8.27</td>
</tr>
<tr>
<td>110</td>
<td>9.03</td>
</tr>
<tr>
<td>120</td>
<td>9.87</td>
</tr>
<tr>
<td>130</td>
<td>10.79</td>
</tr>
<tr>
<td>140</td>
<td>11.77</td>
</tr>
<tr>
<td>150</td>
<td>12.83</td>
</tr>
</tbody>
</table>

You were assigned to analyze these data for a team project. (The other two members of your team did not take a statistics course based on this text.) The first team member prepared a draft with the following summary:

Fuel consumption does not depend on speed for this vehicle \( (t = -0.63, df = 13, P = 0.54) \).

The second team member had some trouble with the statistical software. When the data were read, only the first two columns of the data table were used by the
software. These are the data for speeds between 10 and 80 km/h. This team member prepared the following draft:

There is a strong relationship between fuel consumption and speed for this vehicle ($t = -3.63$, df = 6, $P = 0.0109$). Speed explains 68.8% of the variation in fuel consumption.

First, verify that your two teammates have computed the quantities that they reported correctly. Then analyze the data and write a summary of your analysis of the relationship between fuel consumption and speed.

**CHAPTER 11**

**Chapter 11 Exercises**

11.1 One model for subpopulation means for the computer science study is described in Example 11.1 as

$$
\mu_{\text{GPA}} = \beta_0 + \beta_1 HSM + \beta_2 HSS + \beta_3 HSE
$$

(a) Give the model for the subpopulation mean GPA for students having high school grade scores HSM = 9 (A−), HSS = 8 (B+), and HSE = 7 (B).

(b) Using the parameter estimates given in Figure 11.4, calculate the estimate of this subpopulation mean. Then briefly explain in words what your numerical answer means.

11.2 Use the model given in the previous exercise to do the following:

(a) For students having high school grade scores HSM = 6 (B−), HSS = 7 (B), and HSE = 8 (B+), express the subpopulation mean in terms of the parameters $\beta_j$. 

(b) Calculate the estimate of this subpopulation mean using the \( b_j \) given in Figure 11.4. Briefly explain the meaning of the number you obtain.

11.3 A multiple regression is used to relate a response variable to a set of 7 explanatory variables. There are 120 observations. Outline the analysis of variance table for this analysis giving the sources of variation and the degrees of freedom.

11.4 A multiple regression analysis of 105 cases was performed with 4 explanatory variables. Suppose that SSM = 20 and SSE = 200.
(a) Find the value of the \( F \) statistic for testing the null hypothesis that the coefficients of all of the explanatory variables are zero.
(b) Use Table E to determine if the result is significant at the 5% level. Is it significant at the 1% level?

11.5 Refer to the previous exercise. What proportion of the variation in the response variable is explained by the explanatory variables?

11.6 An instructor in an introductory statistics class used multiple regression to predict the score on the final exam using the results of 10 quizzes that were given during the semester. The \( F \) statistic for this analysis was highly significant (\( P < 0.0001 \)) but none of the \( t \) tests for the individual coefficients were significant. Explain this apparent inconsistency.

The following eight exercises are related to the case study described in this chapter. They require use of the CSDATA data set described in the Data Appendix.

11.7 Use software to make a plot of GPA versus SATM. Do the same for GPA versus SATV. Describe the general patterns. Are there any unusual values?

11.8 Make a plot of GPA versus HSM. Do the same for the other two high school grade variables. Describe the three plots. Are there any outliers or influential points?
11.9 Regress GPA on the three high school grade variables. Calculate and store the residuals from this regression. Plot the residuals versus each of the three predictors and versus the predicted value of GPA. Are there any unusual points or patterns in these four plots?

11.10 Use the two SAT scores in a multiple regression to predict GPA. Calculate and store the residuals. Plot the residuals versus each of the explanatory variables and versus the predicted GPA. Describe the plots.

11.11 It appears that the mathematics explanatory variables are strong predictors of GPA in the computer science study. Run a multiple regression using HSM and SATM to predict GPA.
(a) Give the fitted regression equation.
(b) State the $H_0$ and $H_a$ tested by the ANOVA $F$ statistic, and explain their meaning in plain language. Report the value of the $F$ statistic, its $P$-value, and your conclusion.
(c) Give 95% confidence intervals for the regression coefficients of HSM and SATM. Do either of these include the point 0?
(d) Report the $t$ statistics and $P$-values for the tests of the regression coefficients of HSM and SATM. What conclusions do you draw from these tests?
(e) What is the value of $s$, the estimate of $\sigma$?
(f) What percent of the variation in GPA is explained by HSM and SATM in your model?

11.12 How well do verbal variables predict the performance of computer science students? Perform a multiple regression analysis to predict GPA from HSE and SATV. Summarize the results and compare them with those obtained in the previous exercise. In what ways do the regression results indicate that the mathematics variables are better predictors?
11.13 The variable SEX has the value 1 for males and 2 for females. Create a data set containing the values for males only. Run a multiple regression analysis for predicting GPA from the three high school grade variables for this group. Using the case study in the text as a guide, interpret the results and state what conclusions can be drawn from this analysis. In what way (if any) do the results for males alone differ from those for all students?

11.14 Refer to the previous exercise. Perform the analysis using the data for females only. Are there any important differences between female and male students in predicting GPA?

The following three exercises use the CONCEPT data set described in the Data Appendix.

11.15 Find the correlations between the response variable GPA and each of the explanatory variables IQ, AGE, SEX, SC, and C1 to C6. Of all the explanatory variables, IQ does the best job of explaining GPA in a simple linear regression with only one variable. How do you know this without doing all of the regressions? What percent of the variation in GPA can be explained by the straight-line relationship between GPA and IQ?

11.16 Let us look at the role of positive self-concept about one’s physical appearance (variable C3). We include IQ in the model because it is a known important predictor of GPA.

(a) Report the fitted regression model

\[ \hat{GPA} = b_0 + b_1 \text{IQ} + b_2 \text{C3} \]

with its \( R^2 \), and report the \( t \) statistic for significance of self-concept about one’s physical appearance with its \( P \)-value. Does C3 contribute significantly to explaining GPA when added to IQ? How much does adding C3 to the model raise \( R^2 \)?
(b) Now start with the model that includes overall self-concept SC along with IQ. Does C3 help explain GPA significantly better? Does it raise $R^2$ enough to be of practical value? In answering these questions, report the fitted regression model

$$\hat{GPA} = b_0 + b_1IQ + b_2C3 + b_3SC$$

with its $R^2$ and the $t$ statistic for significance of C3 with its $P$-value.

(c) Explain carefully in words why the coefficient $b_2$ for the variable C3 takes quite different values in the regressions of parts (a) and (b). Then explain simply how it can happen that C3 is a useful explanatory variable in part (a) but worthless in part (b).

11.17 A reasonable model explains GPA using IQ, C1 (behavior self-concept), and C5 (popularity self-concept). These three explanatory variables are all significant in the presence of the other two, and no other explanatory variable is significant when added to these three. (You do not have to verify these statements.) Let’s use this model.

(a) What is the fitted regression model, its $R^2$, and the standard deviation $s$ about the fitted model? What GPA does this model predict for a student with IQ 109, C1 score 13, and C5 score 8?

(b) If both C1 and C5 could be held constant, how much would GPA increase for each additional point of IQ score, according to the fitted model? Give a 95% confidence interval for the mean increase in GPA in the entire population if IQ could be increased by 1 point while holding C1 and C5 constant.

(c) Compute the residuals for this model. Plot the residuals against the predicted values and also make a normal quantile plot. Which observation (value of OBS) produces the most extreme residual? Circle this observation in both of your plots of the residuals. For which individual variables does this student have very unusual values? What are these values? (Look at all the variables, not just those in the model.)
(d) Repeat part (a) with this one observation removed. How much did this one student affect the fitted model and the prediction?

The following 7 exercises use the corn and soybean data given in Examples 10.20 and 10.22.

11.18 Run the simple linear regression using year to predict corn yield.
(a) Summarize the results of your analysis, including the significance test results for the slope and $R^2$ for this model.
(b) Analyze the residuals with a normal quantile plot. Is there any indication in the plot that the residuals are not normal?
(c) Plot the residuals versus soybean yield. Does the plot indicate that soybean yield might be useful in a multiple linear regression with year to predict corn yield?

11.19 Run the simple linear regression using soybean yield to predict corn yield.
(a) Summarize the results of your analysis, including the significance test results for the slope and $R^2$ for this model.
(b) Analyze the residuals with a normal quantile plot. Is there any indication in the plot that the residuals are not normal?
(c) Plot the residuals versus year. Does the plot indicate that year might be useful in a multiple linear regression with soybean yield to predict corn yield?

11.20 From the previous two exercises, we conclude that year and soybean yield may be useful together in a model for predicting corn yield. Run this multiple regression.
(a) Explain the results of the ANOVA $F$ test. Give the null and alternative hypotheses, the test statistic with degrees of freedom, and the $P$-value. What do you conclude?
(b) What percent of the variation in corn yield is explained by these two variables? Compare it with the percent explained in the simple linear regression models of the
previous two exercises. 
(c) Give the fitted model. Why do the coefficients of year and soybean yield differ from those in the previous two exercises? 
(d) Summarize the significance test results for the regression coefficients for year and soybean yield. 
(e) Give a 95% confidence interval for each of these coefficients. 
(f) Plot the residuals versus year and versus soybean yield. What do you conclude? 

11.21 We need a new variable to model the curved relation that we see between corn yield and year in the residual plot of the last exercise. Let year2 = (year − 1978.5)². 
(When adding a squared term to a multiple regression model, it is a good idea to subtract the mean of the variable being squared before squaring. This avoids all sorts of messy problems that we cannot discuss here.) 
(a) Run the multiple linear regression using year, year2, and soybean yield to predict corn yield. Give the fitted regression equation. 
(b) Give the null and alternative hypotheses for the ANOVA F test. Report the results of this test, giving the test statistic, degrees of freedom, P-value, and conclusion. 
(c) What percent of the variation in corn yield is explained by this multiple regression? Compare this with the model in the previous exercise. 
(d) Summarize the results of the significance tests for the individual regression coefficients. 
(e) Analyze the residuals and summarize your conclusions. 

11.22 Run the model to predict corn yield using year and the squared term year2 defined in the previous exercise. 
(a) Summarize the significance test results. 
(b) The coefficient of year2 is not statistically significant in this run, but it was highly significant in the model analyzed in the previous exercise. Explain how this
(c) Obtain the fitted values for each year in the data set and use these to sketch the curve on a plot of the data. Plot the least-squares line on this graph for comparison. Describe the differences between the two regression functions. For what years do they give very similar fitted values; for what years are the differences between the two relatively large?

11.23 Use the simple linear regression model with corn yield as the response variable and year as the explanatory variable to predict the corn yield for the year 2001, and give the 95% prediction interval. Also, use the multiple regression model where year and year2 are both explanatory variables to find another predicted value with the 95% interval. Explain why these two predicted values are so different. The actual yield for 2001 is 138.2 bushels per acre. How well did your models predict this value?

11.24 Repeat the previous exercise doing the prediction for 2002. Compare the results of this exercise with the previous one. Why are they different?

The following three exercises use the bank wages data that are used in Exercises for Chapter 10. Here is a brief description of the setting: We assume that our wages will increase as we gain experience and become more valuable to our employers. Wages also increase because of inflation. By examining a sample of employees at a given point in time, we can look at part of the picture. How does length of service (LOS) relate to wages? The following table gives data on the LOS in months and wages for 60 women who work in Indiana banks. Wages are yearly total income divided by the number of weeks worked. We have multiplied wages by a constant for reasons of confidentiality. (The data were provided by Professor Shelly MacDermid, Department of Child Development and Family Studies, Purdue University, from a study reported in S. M. MacDermid et al., “Is small beautiful? Work-family tension,
work conditions, and organizational size,” Family Relations, 44 (1994), pp. 159–167.)

<table>
<thead>
<tr>
<th>Wages</th>
<th>LOS</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>48.3355</td>
<td>94</td>
<td>Large</td>
</tr>
<tr>
<td>49.0279</td>
<td>48</td>
<td>Small</td>
</tr>
<tr>
<td>40.8817</td>
<td>102</td>
<td>Small</td>
</tr>
<tr>
<td>36.5854</td>
<td>20</td>
<td>Small</td>
</tr>
<tr>
<td>46.7596</td>
<td>60</td>
<td>Large</td>
</tr>
<tr>
<td>59.5238</td>
<td>78</td>
<td>Small</td>
</tr>
<tr>
<td>39.1304</td>
<td>45</td>
<td>Large</td>
</tr>
<tr>
<td>39.2465</td>
<td>39</td>
<td>Large</td>
</tr>
<tr>
<td>40.2037</td>
<td>20</td>
<td>Large</td>
</tr>
<tr>
<td>38.1563</td>
<td>65</td>
<td>Small</td>
</tr>
<tr>
<td>50.0905</td>
<td>76</td>
<td>Large</td>
</tr>
<tr>
<td>46.9043</td>
<td>48</td>
<td>Small</td>
</tr>
<tr>
<td>43.1894</td>
<td>61</td>
<td>Small</td>
</tr>
<tr>
<td>60.5637</td>
<td>30</td>
<td>Large</td>
</tr>
<tr>
<td>97.6801</td>
<td>70</td>
<td>Large</td>
</tr>
<tr>
<td>48.5795</td>
<td>108</td>
<td>Large</td>
</tr>
<tr>
<td>67.1551</td>
<td>61</td>
<td>Large</td>
</tr>
<tr>
<td>38.7847</td>
<td>10</td>
<td>Small</td>
</tr>
<tr>
<td>51.8926</td>
<td>68</td>
<td>Large</td>
</tr>
<tr>
<td>51.8326</td>
<td>54</td>
<td>Large</td>
</tr>
</tbody>
</table>

For these exercises we code the size of the bank as 1 if it is large and 0 if it is small. There is one outlier in this data set. Delete it and use the remaining 59 observations in these exercises.
11.25 Use length of service (LOS) to predict wages with a simple linear regression. Write a short summary of your results and conclusions.

11.26 Predict wages using the size of the bank as the explanatory variable. Use the coded values 0 and 1 for this model.

(a) Summarize the results of your analysis. Include a statement of all hypotheses, test statistics with degrees of freedom, \( P \)-values, and conclusions.

(b) Calculate the \( t \) statistic for comparing the mean wages for the large and the small banks assuming equal standard deviations. Give the degrees of freedom. Verify that this \( t \) is the same as the \( t \) statistic for the coefficient of size in the regression. Explain why this makes sense.

(c) Plot the residuals versus LOS. What do you conclude?

11.27 Use a multiple linear regression to predict wages from LOS and the size of the bank. Write a report summarizing your work. Include graphs and the results of significance tests.

---

**CHAPTER 12**

**Chapter 12 Exercises**

12.1 For each of the following situations, identify the response variable and the populations to be compared, and give \( I \), the \( n_i \), and \( N \).

(a) To compare four varieties of tomato plants, 12 plants of each variety are grown and the yield in pounds of tomatoes is recorded.

(b) A marketing experiment compares five different types of packaging for a laundry detergent. Each package is shown to 40 different potential consumers, who rate the attractiveness of the product on a 1 to 10 scale.
(c) To compare the effectiveness of three different weight-loss programs, 20 people are randomly assigned to each. At the end of the program, the weight loss for each of the participants is recorded.

12.2 For each of the following situations, identify the response variable and the populations to be compared, and give $I$, the $n_i$, and $N$.

(a) In a study on smoking, subjects are classified as nonsmokers, moderate smokers, or heavy smokers. A sample of size 100 is drawn from each group. Each person is asked to report the number of hours of sleep he or she gets on a typical night.

(b) The strength of concrete depends upon the formula used to prepare it. One study compared four different mixtures. Five batches of each mixture were prepared, and the strength of the concrete made from each batch was measured.

(c) Which of three methods of teaching sign language is most effective? Twenty students are randomly assigned to each of the methods, and their scores on a final exam are recorded.

12.3 How do nematodes (microscopic worms) affect plant growth? A botanist prepares 16 identical planting pots and then introduces different numbers of nematodes into the pots. A tomato seedling is transplanted into each plot. Here are data on the increase in height of the seedlings (in centimeters) 16 days after planting (data provided by Matthew Moore):

<table>
<thead>
<tr>
<th>Nematodes</th>
<th>Seedling growth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>10.8 9.1 13.5 9.2</td>
</tr>
<tr>
<td>1,000</td>
<td>11.1 11.1 8.2 11.3</td>
</tr>
<tr>
<td>5,000</td>
<td>5.4 4.6 7.4 5.0</td>
</tr>
<tr>
<td>10,000</td>
<td>5.8 5.3 3.2 7.5</td>
</tr>
</tbody>
</table>

(a) Make a table of means and standard deviations for the four treatments, and plot the means. What does the plot of the means show?

(b) State $H_0$ and $H_a$ for an ANOVA on these data, and explain in words what
ANOVA tests in this setting.
(c) Using computer software, run the ANOVA. What are the $F$ statistic and its $P$-value? Give the values of $s_p$ and $R^2$. Report your conclusion.

12.4 Refer to the previous exercise.
(a) Define the contrast that compares the 0 treatment (the control group) with the average of the other three.
(b) State $H_0$ and $H_a$ for using this contrast to test whether or not the presence of nematodes causes decreased growth in tomato seedlings.
(c) Perform the significance test and give the $P$-value. Do you reject $H_0$?
(d) Define the contrast that compares the 0 treatment with the treatment with 10,000 nematodes. This contrast is a measure of the decrease in growth due to having a very large nematode infestation. Give a 95% confidence interval for this decrease in growth.

12.5 In large classes instructors sometimes use different forms of an examination. When average scores for the different forms are calculated, students who received the form with the lowest average score may complain that their examination was more difficult than the others. Analysis of variance can help determine whether the variation in mean scores is larger than would be expected by chance. One such class used three forms. Summary statistics were as follows. (Data provided by Peter Georgeoff of the Purdue University Department of Educational Studies.)

<table>
<thead>
<tr>
<th>Form</th>
<th>n</th>
<th>$\bar{x}$</th>
<th>s</th>
<th>Min.</th>
<th>$Q_1$</th>
<th>Median</th>
<th>$Q_3$</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>79</td>
<td>31.78</td>
<td>4.45</td>
<td>18</td>
<td>29</td>
<td>32</td>
<td>35</td>
<td>42</td>
</tr>
<tr>
<td>2</td>
<td>81</td>
<td>32.88</td>
<td>4.40</td>
<td>20</td>
<td>30</td>
<td>33</td>
<td>36</td>
<td>42</td>
</tr>
<tr>
<td>3</td>
<td>81</td>
<td>34.47</td>
<td>4.29</td>
<td>24</td>
<td>32</td>
<td>35</td>
<td>38</td>
<td>46</td>
</tr>
</tbody>
</table>

Here is the SAS output for a one-way ANOVA run on the exam scores:

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Squares</th>
<th>$\text{Square}$</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of</td>
<td>2</td>
<td>317.80</td>
<td>1.235</td>
<td>0.248</td>
<td>0.630</td>
</tr>
<tr>
<td>Mean</td>
<td>2</td>
<td>317.80</td>
<td>1.235</td>
<td>0.248</td>
<td>0.630</td>
</tr>
</tbody>
</table>
Model               2  292.01871  146.00936  7.61  0.0006
Error               238  4566.28004  19.18605
Corrected Total     240  4858.29876

R-Square  C.V.  Root MSE  SCORE Mean
0.060107  13.25164  4.3802  33.054

Bonferroni (Dunn) T tests for variable: SCORE

NOTE: This test controls the type I experimentwise error rate but generally has a higher type II error rate than Tukey's for all pairwise comparisons.

Alpha= 0.05  Confidence= 0.95  df= 238  MSE= 19.18605

Critical value of T= 2.41102

Comparisons significant at the 0.05 level are indicated by '***'.

Simultaneous  Simultaneous
          Lower   Difference   Upper
FORM       Confidence  Between  Confidence
Comparison  Limit     Means    Limit

3  - 2   -0.0669   1.5926    3.2521
3  - 1   1.0144    2.6843    4.3543  ***
2  - 3   -3.2521   -1.5926   0.0669
2  - 1   -0.5782   1.0917    2.7617
Exercises

(a) Compare the distributions of exam scores for the three forms with side-by-side boxplots. Give a short summary of the information contained in these plots.
(b) Summarize and interpret the results of the ANOVA, including the multiple comparisons procedure.

12.6 The presence of lead in the soil of forests is an important ecological concern. One source of lead contamination is the exhaust from automobiles. In recent years this source has been greatly reduced by the elimination of lead from gasoline. Can the effects be seen in our forests? The Hubbard Brook Experimental Forest in West Thornton, New Hampshire, is the site of an ongoing study of the forest floor. Lead measurements of samples taken from this forest are available for several years. The variable of interest is lead concentration recorded as milligrams per square meter. Because the data are strongly skewed to the right, logarithms of the concentrations were analyzed. Here are some summary statistics for 5 years (data provided by Tom Siccama of the Yale University School of Forestry and Environmental Studies):

<table>
<thead>
<tr>
<th>Year</th>
<th>n</th>
<th>( \bar{x} )</th>
<th>s</th>
<th>Min.</th>
<th>( Q_1 )</th>
<th>Median</th>
<th>( Q_3 )</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>76</td>
<td>59</td>
<td>6.80</td>
<td>.58</td>
<td>5.74</td>
<td>6.33</td>
<td>6.73</td>
<td>7.32</td>
<td>8.05</td>
</tr>
<tr>
<td>77</td>
<td>58</td>
<td>6.75</td>
<td>.68</td>
<td>3.95</td>
<td>6.39</td>
<td>6.80</td>
<td>7.23</td>
<td>8.10</td>
</tr>
<tr>
<td>78</td>
<td>58</td>
<td>6.76</td>
<td>.50</td>
<td>5.01</td>
<td>6.50</td>
<td>6.78</td>
<td>7.10</td>
<td>7.66</td>
</tr>
<tr>
<td>82</td>
<td>68</td>
<td>6.50</td>
<td>.55</td>
<td>5.15</td>
<td>6.11</td>
<td>6.53</td>
<td>6.83</td>
<td>7.82</td>
</tr>
<tr>
<td>87</td>
<td>70</td>
<td>6.40</td>
<td>.68</td>
<td>4.38</td>
<td>6.09</td>
<td>6.46</td>
<td>6.85</td>
<td>8.15</td>
</tr>
</tbody>
</table>

Here is the SAS output for a one-way ANOVA run on the logs of the lead concentrations:

\[
\begin{array}{llllll}
\text{Sum of} & \text{Mean} \\
\end{array}
\]
Source | DF | Squares | Square | F Value | Pr > F
---|---|---|---|---|---
Model | 4 | 8.4437799 | 2.1109450 | 5.75 | 0.0002
Error | 308 | 113.1440666 | 0.3673509 |
Corrected Total | 312 | 121.5878465 |

R-Square | C.V. | Root MSE | LLEAD Mean
---|---|---|---
0.069446 | 9.143762 | 0.6061 | 6.6285

Bonferroni (Dunn) T tests for variable: LLEAD

NOTE: This test controls the type I experimentwise error rate but generally has a higher type II error rate than Tukey’s for all pairwise comparisons.

Alpha= 0.05  Confidence= 0.95  df= 308  MSE= 0.367351

Critical value of T= 2.82740

Comparisons significant at the 0.05 level are indicated by '***'.

<table>
<thead>
<tr>
<th>YEAR Comparison</th>
<th>Lower</th>
<th>Difference</th>
<th>Upper</th>
<th>Lower</th>
<th>Difference</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>76 - 78</td>
<td>-0.2745</td>
<td>0.0424</td>
<td>0.3592</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76 - 77</td>
<td>-0.2687</td>
<td>0.0482</td>
<td>0.3650</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76 - 82</td>
<td>-0.0046</td>
<td>0.3003</td>
<td>0.6052</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76 - 87</td>
<td>0.0995</td>
<td>0.4024</td>
<td>0.7052  ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>78</td>
<td>76</td>
<td>78 - 76</td>
<td>-0.3592</td>
<td>-0.0424</td>
<td>0.2745</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>78</td>
<td>77</td>
<td>78 - 77</td>
<td>-0.3124</td>
<td>0.0058</td>
<td>0.3240</td>
</tr>
<tr>
<td></td>
<td>78</td>
<td>82</td>
<td>78 - 82</td>
<td>-0.0484</td>
<td>0.2579</td>
<td>0.5642</td>
</tr>
<tr>
<td></td>
<td>78</td>
<td>87</td>
<td>78 - 87</td>
<td>0.0557</td>
<td>0.3600</td>
<td>0.6643</td>
</tr>
<tr>
<td></td>
<td>77</td>
<td>76</td>
<td>77 - 76</td>
<td>-0.3650</td>
<td>-0.0482</td>
<td>0.2687</td>
</tr>
<tr>
<td></td>
<td>77</td>
<td>78</td>
<td>77 - 78</td>
<td>-0.3240</td>
<td>-0.0058</td>
<td>0.3124</td>
</tr>
<tr>
<td></td>
<td>77</td>
<td>82</td>
<td>77 - 82</td>
<td>-0.0542</td>
<td>0.2521</td>
<td>0.5584</td>
</tr>
<tr>
<td></td>
<td>77</td>
<td>87</td>
<td>77 - 87</td>
<td>0.0499</td>
<td>0.3542</td>
<td>0.6585</td>
</tr>
<tr>
<td></td>
<td>82</td>
<td>76</td>
<td>82 - 76</td>
<td>-0.6052</td>
<td>-0.3003</td>
<td>0.0046</td>
</tr>
<tr>
<td></td>
<td>82</td>
<td>78</td>
<td>82 - 78</td>
<td>-0.5642</td>
<td>-0.2579</td>
<td>0.0484</td>
</tr>
<tr>
<td></td>
<td>82</td>
<td>77</td>
<td>82 - 77</td>
<td>-0.5584</td>
<td>-0.2521</td>
<td>0.0542</td>
</tr>
<tr>
<td></td>
<td>82</td>
<td>87</td>
<td>82 - 87</td>
<td>-0.1897</td>
<td>0.1021</td>
<td>0.3938</td>
</tr>
<tr>
<td></td>
<td>87</td>
<td>76</td>
<td>87 - 76</td>
<td>-0.7052</td>
<td>-0.4024</td>
<td>-0.0995</td>
</tr>
<tr>
<td></td>
<td>87</td>
<td>78</td>
<td>87 - 78</td>
<td>-0.6643</td>
<td>-0.3600</td>
<td>-0.0557</td>
</tr>
<tr>
<td></td>
<td>87</td>
<td>77</td>
<td>87 - 77</td>
<td>-0.6585</td>
<td>-0.3542</td>
<td>-0.0499</td>
</tr>
<tr>
<td></td>
<td>87</td>
<td>82</td>
<td>87 - 82</td>
<td>-0.3938</td>
<td>-0.1021</td>
<td>0.1897</td>
</tr>
</tbody>
</table>

(a) Display the data with side-by-side boxplots. Describe the major features of the data.

(b) Summarize the ANOVA results. Do the data suggest that the (log) concentration of lead in the Hubbard Forest floor is decreasing?

12.7 A randomized comparative experiment compares three programs designed to help people lose weight. There are 20 subjects in each program. The sample standard deviations for the amount of weight lost (in pounds) are 5.2, 8.9, and 10.1.
Can you use the assumption of equal standard deviations to analyze these data? Compute the pooled variance and find $s_p$.

12.8 A study of physical fitness collected data on the weight (in kilograms) of men in four different age groups. The sample sizes for the groups were 92, 34, 35, and 24. The sample standard deviations for the groups were 12.2, 10.4, 9.2, and 11.7. Can you use the assumption of equal standard deviations to analyze these data? Compute the pooled variance and find $s_p$.

12.9 For each part of Exercise S12.1, outline the ANOVA table, giving the sources of variation and the degrees of freedom. (Do not compute the numerical values for the sums of squares and mean squares.)

12.10 For each part of Exercise S12.2, outline the ANOVA table, giving the sources of variation and the degrees of freedom. (Do not compute the numerical values for the sums of squares and mean squares.)

12.11 Return to the change-of-majors study described in Example 12.3.
(a) State $H_0$ and $H_a$ for ANOVA.
(b) Outline the ANOVA table, giving the sources of variation and the degrees of freedom.
(c) What is the distribution of the $F$ statistic under the assumption that $H_0$ is true?
(d) Using Table E, find the critical value for an $\alpha = 0.05$ test.

12.12 Return to the survey of college students described in Example 12.4.
(a) State $H_0$ and $H_a$ for ANOVA.
(b) Outline the ANOVA table, giving the sources of variation and the degrees of freedom.
(c) What is the distribution of the $F$ statistic under the assumption that $H_0$ is true?
(d) Using Table E, find the critical value for an $\alpha = 0.05$ test.
12.13 Return to the nematode experiment described in Exercise S12.3. Suppose that when entering the data into the computer, you accidentally entered the first observation as 108 rather than 10.8.

(a) Run the ANOVA with the incorrect observation. Summarize the results.

(b) Compare this run with the results obtained with the correct data set. What does this illustrate about the effect of outliers in an ANOVA?

(c) Compute a table of means and standard deviations for each of the four treatments using the incorrect data. How would this table have helped you to detect the incorrect observation?

12.14 With small numbers of observations in each group, it can be difficult to detect deviations from normality and violations of the equal standard deviations assumption for ANOVA. Return to the nematode experiment described in Exercise S12.3. The log transformation is often used for variables such as the growth of plants. In many cases this will tend to make the standard deviations more similar across groups and to make the data within each group look more normal. Rerun the ANOVA using the logarithms of the recorded values. Answer the questions given in Exercise S12.3. Compare these results with those obtained by analyzing the raw data.

12.15 You are planning a study of the SAT mathematics scores of four groups of students. From Example 12.3, we know that the standard deviations of the three groups considered in that study were 86, 67, and 83. In Example 12.5, we found the pooled standard deviation to be 82.5. Since the power of the $F$ test decreases as the standard deviation increases, use $\sigma = 90$ for the calculations in this exercise. This choice will lead to sample sizes that are perhaps a little larger than we need but will prevent us from choosing sample sizes that are too small to detect the effects of interest. You would like to conclude that the population means are different when $\mu_1 = 620$, $\mu_2 = 600$, $\mu_3 = 580$, and $\mu_4 = 560$.

(a) Pick several values for $n$ (the number of students that you will select from each
Chapter 12 Exercises

12.15 (a) Group the data and calculate the power of the ANOVA F test for each of your choices.

(b) Plot the power versus the sample size. Describe the general shape of the plot.

(c) What choice of n would you choose for your study? Give reasons for your answer.

12.16 Refer to the previous exercise. Repeat all parts for the alternative \( \mu_1 = 610, \mu_2 = 600, \mu_3 = 590, \) and \( \mu_4 = 580. \)

12.17 For each of the following situations, identify the response variable and the populations to be compared, and give \( I, n_i, \) and \( N. \)

(a) A company wants to compare three different training programs for its new employees. In a one-month period there are 90 new hires. One-third of these are randomly assigned to each of the three training programs. At the end of the program the employees are asked to rate the effectiveness of the program on a 7-point scale.

(b) A marketing experiment compares six different types of packaging for computer disks. Each package is shown to 50 different potential consumers, who rate the attractiveness of the product on a 1 to 10 scale.

(c) Four different new formulations for a hand lotion have been produced by your research and development group, and you want to decide which of these, if any, to market. Samples of the first new lotion are sent to 100 randomly selected customers who use your regular product. The same procedure is followed for each of the other three new lotions. You ask each customer to compare the new lotion sent to them with the regular product by rating it on a 7-point scale. The middle point of the scale corresponds to no preference, while higher values indicate that the new product is preferred and lower values indicate that the regular product is better.

12.18 Refer to the previous exercise. For each situation, give the following:

(a) Degrees of freedom for the model, for error, and for the total.

(b) Null and alternative hypotheses.

(c) Numerator and denominator degrees of freedom for the F statistic.
12.19 For each of the following situations, identify the response variable and the populations to be compared, and give $I$, the $n_i$, and $N$.

(a) A company wants to compare three different water treatment devices that can be attached to a kitchen faucet. From a list of potential customers, they select 225 households who will receive free samples. One-third of the households will receive each of the devices. The household is asked to rate the likelihood that they would buy this kind of device on a 5-point scale.

(b) The strength of concrete depends upon the formula used to prepare it. One study compared five different mixtures. Six batches of each mixture were prepared, and the strength of the concrete made from each batch was measured.

(c) Which of three methods of teaching statistics is most effective? Twenty students are randomly assigned to each of the methods, and their scores on a final exam are recorded.

12.20 Refer to the previous exercise. For each situation, give the following:

(a) Degrees of freedom for the model, for error, and for the total.

(b) Null and alternative hypotheses.

(c) Numerator and denominator degrees of freedom for the $F$ statistic.

12.21 An experiment was run to compare three groups. The sample sizes were 10, 12, and 14, and the corresponding estimated standard deviations were 18, 24, and 20.

(a) Is it reasonable to use the assumption of equal standard deviations when we analyze these data?

(b) Give the values of the variances for the three groups.

(c) Find the pooled variance.

(d) What is the value of the pooled standard deviation?

12.22 An experiment was run to compare four groups. The sample sizes were 20, 220, 18, and 15, and the corresponding estimated standard deviations were 62, 40,
52, and 48.

(a) Is it reasonable to use the assumption of equal standard deviations when we analyze these data?

(b) Give the values of the variances for the four groups.

(c) Find the pooled variance.

(d) What is the value of the pooled standard deviation?

(e) Explain why your answer in part (c) is much closer to the standard deviation for the second group than to any of the other standard deviations.

12.23 For each of the following situations find the degrees of freedom for the $F$ statistic and then use Table E to approximate the $P$-value or use computer software to obtain an exact value.

(a) Three groups are being compared, with 8 observations per group. The value of the $F$ statistic is 5.82.

(b) Six groups are being compared, with 11 observations per group. The value of the $F$ statistic is 2.16.

12.24 For each of the following situations find the degrees of freedom for the $F$ statistic and then use Table E to approximate the $P$-value or use computer software to obtain an exact value.

(a) Five groups are being compared, with 13 observations per group. The value of the $F$ statistic is 1.61.

(b) Ten groups are being compared, with 4 observations per group. The value of the $F$ statistic is 4.68.

12.25 The presence of harmful insects in farm fields is detected by erecting boards covered with a sticky material and then examining the insects trapped on the board. To investigate which colors are most attractive to cereal leaf beetles, researchers placed six boards of each of four colors in a field of oats in July. (M. C. Wilson and R. E. Shade, “Relative attractiveness of various luminescent colors to the cereal leaf
beetle and the meadow spittlebug,” *Journal of Economic Entomology*, 60 (1967), pp. 578–580.) The table below gives data on the number of cereal leaf beetles trapped:

<table>
<thead>
<tr>
<th>Color</th>
<th>Insects trapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lemon yellow</td>
<td>45 59 48 46 38 47</td>
</tr>
<tr>
<td>White</td>
<td>21 12 14 17 13 17</td>
</tr>
<tr>
<td>Green</td>
<td>37 32 15 25 39 41</td>
</tr>
<tr>
<td>Blue</td>
<td>16 11 20 21 14  7</td>
</tr>
</tbody>
</table>

(a) Make a table of means and standard deviations for the four colors, and plot the means.

(b) State \( H_0 \) and \( H_a \) for an ANOVA on these data, and explain in words what ANOVA tests in this setting.

(c) Using computer software, run the ANOVA. What are the \( F \) statistic and its \( P \)-value? Give the values of \( s_p \) and \( R^2 \). What do you conclude?

12.26 Return to the previous exercise. For the Bonferroni procedure with \( \alpha = 0.05 \), the value of \( t^{**} \) is 2.61. Use this multiple comparisons procedure to decide which pairs of colors are significantly different. Summarize your results. Which color would you recommend for attracting cereal leaf beetles?

12.27 A study of the effects of exercise on physiological and psychological variables compared four groups of male subjects. The treatment group (T) consisted of 10 participants in an exercise program. A control group (C) of 5 subjects volunteered for the program but were unable to attend for various reasons. Subjects in the other two groups were selected to be similar to those in the first two groups in age and other characteristics. These were 11 joggers (J) and 10 sedentary people (S) who did not regularly exercise. (Data provided by Dennis Lobstein, from his Ph.D. dissertation, “A multivariate study of exercise training effects on beta-endorphin and emotionality in psychologically normal, medically healthy men,” Purdue University,
One of the variables measured at the end of the program was a physical fitness score. Part of the ANOVA table used to analyze these data is given below:

<table>
<thead>
<tr>
<th>Degrees</th>
<th>Source</th>
<th>of freedom</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Groups</td>
<td>3</td>
<td>104,855.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>32</td>
<td>70,500.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>175,356.46</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Fill in the missing entries in the ANOVA table.

(b) State $H_0$ and $H_a$ for this experiment.

(c) What is the distribution of the $F$ statistic under the assumption that $H_0$ is true? Using Table E, give an approximate $P$-value for the ANOVA test. Write a brief conclusion.

(d) What is $s_p^2$, the estimate of the within-group variance? What is $s_p$?

Another variable measured in the experiment described in the previous exercise was a depression score. Higher values of this score indicate more depression. Part of the ANOVA table for these data appears below:

<table>
<thead>
<tr>
<th>Degrees</th>
<th>Source</th>
<th>of freedom</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Groups</td>
<td>3</td>
<td>158.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>32</td>
<td>62.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>75.77</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Fill in the missing entries in the ANOVA table.

(b) State $H_0$ and $H_a$ for this experiment.

(c) What is the distribution of the $F$ statistic under the assumption that $H_0$ is true? Using Table E, give an approximate $P$-value for the ANOVA test. What do you conclude?

(d) What is $s_p^2$, the estimate of the within-group variance? What is $s_p$?
The weight gain of women during pregnancy has an important effect on the birth weight of their children. If the weight gain is not adequate, the infant is more likely to be small and will tend to be less healthy. In a study conducted in three countries, weight gains (in kilograms) of women during the third trimester of pregnancy were measured. (These data were taken from Collaborative Research Support Program in Food Intake and Human Function, *Management Entity Final Report*, University of California, Berkeley, 1988.) The results are summarized in the following table:

<table>
<thead>
<tr>
<th>Country</th>
<th>n</th>
<th>$\bar{x}$</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egypt</td>
<td>46</td>
<td>3.7</td>
<td>2.5</td>
</tr>
<tr>
<td>Kenya</td>
<td>111</td>
<td>3.1</td>
<td>1.8</td>
</tr>
<tr>
<td>Mexico</td>
<td>52</td>
<td>2.9</td>
<td>1.8</td>
</tr>
</tbody>
</table>

(a) Find the pooled estimate of the within-country variance $s_p^2$. What entry in the ANOVA table gives this quantity?

(b) The sum of squares for countries (groups) is 17.22. Use this information and that given above to complete all the entries in an ANOVA table.

(c) State $H_0$ and $H_a$ for this study.

(d) What is the distribution of the $F$ statistic under the assumption that $H_0$ is true? Use Table E to find an approximate $P$-value for the significance test. Report your conclusion.

(e) Calculate $R^2$, the coefficient of determination.

The previous exercise gives data on the weight gains of pregnant women in Egypt, Kenya, and Mexico. Computer software gives the critical value for the Bonferroni multiple comparisons procedure with $\alpha = 0.05$ as $t^{**} = 2.41$. Explain in plain language what $\alpha = 0.05$ means in the Bonferroni procedure. Use this procedure to compare the mean weight gains for the three countries. Summarize your conclusions.
12.31 In another part of the study described in the previous exercise, measurements of food intake in kilocalories were taken on many individuals several times during the period of a year. From these data, average daily food intake values were computed for each individual. The results for toddlers aged 18 to 30 months are summarized in the following table:

<table>
<thead>
<tr>
<th>Country</th>
<th>n</th>
<th>( \bar{x} )</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egypt</td>
<td>88</td>
<td>1217</td>
<td>327</td>
</tr>
<tr>
<td>Kenya</td>
<td>91</td>
<td>844</td>
<td>184</td>
</tr>
<tr>
<td>Mexico</td>
<td>54</td>
<td>1119</td>
<td>285</td>
</tr>
</tbody>
</table>

(a) Find the pooled estimate of the within-country variance \( s_p^2 \). What entry in the ANOVA table gives this quantity?

(b) The sum of squares for countries (groups) is 6,572,551. Use this information and that given above to complete all the entries in an ANOVA table.

(c) State \( H_0 \) and \( H_a \) for this study.

(d) What is the distribution of the \( F \) statistic under the assumption that \( H_0 \) is true? Use Table E to find an approximate \( P \)-value for the significance test. Report your conclusion.

(e) Calculate \( R^2 \), the coefficient of determination.

12.32 The previous exercise gives summary statistics for the food intake values for toddlers in Egypt, Kenya, and Mexico. Computer software gives the critical value for the Bonferroni multiple comparisons procedure with \( \alpha = 0.05 \) as \( t^{**} = 2.41 \). Explain in plain language what \( \alpha = 0.05 \) means in the Bonferroni procedure. Use this procedure to compare the toddler food intake means for the three countries. What do you conclude?

12.33 In the exercise program study described in Exercise 12.27, the summary statistics for physical fitness scores are as follows:
The researchers wanted to address the following questions for the physical fitness scores. In these questions “better” means a higher fitness score. (1) Is T better than C? (2) Is T better than the average of C and S? (3) Is J better than the average of the other three groups?

(a) For each of the three questions, define an appropriate contrast. Translate the questions into null and alternative hypotheses about these contrasts.

(b) Test your hypotheses and give approximate $P$-values. Summarize your conclusions. Do you think that the use of contrasts in this way gives an adequate summary of the results?

(c) You found that the groups differ significantly in the physical fitness scores. Does this study allow conclusions about causation—for example, that a sedentary lifestyle causes people to be less physically fit? Explain your answer.

12.34 Refer to the physical fitness scores for the four groups in the exercise program study discussed in the previous exercise. Computer software gives the critical value for the Bonferroni multiple comparisons procedure with $\alpha = 0.05$ as $t^{**} = 2.81$. Use this procedure to compare the mean fitness scores for the four groups. Summarize your conclusions.

12.35 Exercise 12.28 gives the ANOVA table for depression scores from the exercise program study described in Exercise 12.27. Here are the summary statistics for the depression scores:
In planning the experiment, the researchers wanted to address the following questions for the depression scores. In these questions “better” means a lower depression score. (1) Is T better than C? (2) Is T better than the average of C and S? (3) Is J better than the average of the other three groups?

(a) For each of the three questions, define an appropriate contrast. Translate the questions into null and alternative hypotheses about these contrasts.

(b) Test your hypotheses and give approximate $P$-values. Summarize your conclusions. Do you think that the use of contrasts in this way gives an adequate summary of the results?

(c) You found that the groups differ significantly in the depression scores. Does this study allow conclusions about causation—for example, that a sedentary lifestyle causes people to be more depressed? Explain your answer.

12.36 Refer to the depression scores for the four groups in the exercise program study discussed in the previous exercises. Computer software gives the critical value for the Bonferroni multiple comparisons procedure with $\alpha = 0.05$ as $t^{**} = 2.81$. Use this procedure to compare the mean depression scores for the four groups. Summarize your conclusions.

12.37 You are planning a study of the weight gains of pregnant women during the third trimester of pregnancy similar to that described in Exercise 12.29. The standard deviations given in that exercise range from 1.8 to 2.5. To perform power calculations, assume that the standard deviation is $\sigma = 2.4$. You have three groups, each with $n$ subjects, and you would like to reject the ANOVA $H_0$ when the alter-
native $\mu_1 = 2.6$, $\mu_2 = 3.0$, and $\mu_3 = 3.4$ is true. Use software to make a table of powers against this alternative for the following numbers of women in each group: $n = 50, 100, 150, 175, \text{ and } 200$. What sample size would you choose for your study?

**12.38** Repeat the previous exercise for the alternative $\mu_1 = 2.7$, $\mu_2 = 3.1$, and $\mu_3 = 3.5$. Why are the results the same?

**12.39** Refer to the color attractiveness experiment described in Exercise 12.25. Suppose that when entering the data into the computer, you accidentally entered the first observation as 450 rather than 45.

(a) Run the ANOVA with the incorrect observation. Summarize the results.

(b) Compare this run with the results obtained with the correct data set. What does this illustrate about the effect of outliers in an ANOVA?

(c) Compute a table of means and standard deviations for each of the four treatments using the incorrect data. How would this table have helped you to detect the incorrect observation?

**12.40** Refer to the color attractiveness experiment described in Exercise 12.25. The square root transformation is often used for variables that are counts, such as the number of insects trapped in this example. In many cases data transformed in this way will conform more closely to the assumptions of normality and equal standard deviations. Rerun the ANOVA using the square roots of the original counts of insects. Answer the questions given in Exercise 12.25. Compare these results with those obtained by analyzing the raw data.

**CHAPTER 13**
Chapter 13 Exercises

13.1 Each of the following situations is a two-way study design. For each case, identify the response variable and both factors, and state the number of levels for each factor ($I$ and $J$) and the total number of observations ($N$).

(a) A study of the productivity of tomato plants compares five varieties of tomatoes and two types of fertilizer. Four plants of each variety are grown with each type of fertilizer. The yield in pounds of tomatoes is recorded for each plant.

(b) A marketing experiment compares six different types of packaging for a laundry detergent. A survey is conducted to determine the attractiveness of the packaging in six U.S. cities. Each type of packaging is shown to 50 different consumers in each city, who rate the attractiveness of the product on a 1 to 10 scale.

(c) To compare the effectiveness of four different weight-loss programs, 10 men and 10 women are randomly assigned to each. At the end of the program, the weight loss for each of the participants is recorded.

13.2 For each part of the previous exercise, outline the ANOVA table, giving the sources of variation and the degrees of freedom. (Do not compute the numerical values for the sums of squares and mean squares.)

13.3 Each of the following situations is a two-way study design. For each case, identify the response variable and both factors, and state the number of levels for each factor ($I$ and $J$) and the total number of observations ($N$).

(a) A study of smoking classifies subjects as nonsmokers, moderate smokers, or heavy smokers. Samples of 120 men and 120 women are drawn from each group. Each person reports the number of hours of sleep he or she gets on a typical night.

(b) The strength of concrete depends upon the formula used to prepare it. An experiment compares four different mixtures. Six specimens of concrete are poured from each mixture. Two of these specimens are subjected to 0 cycles of freezing and
thawing, two are subjected to 100 cycles, and two specimens are subjected to 500 cycles. The strength of each specimen is then measured.

(c) Three methods for teaching sign language are to be compared. Seven students in special education and seven students in linguistics are randomly assigned to each of the methods and the scores on a final exam are recorded.

13.4 For each part of the previous exercise, outline the ANOVA table, giving the sources of variation and the degrees of freedom. (Do not compute the numerical values for the sums of squares and mean squares.)

13.5 A large research project studied the physical properties of wood materials constructed by bonding together small flakes of wood. Different species of trees were used, and the flakes were made in different sizes. One of the physical properties measured was the tension modulus of elasticity in the direction perpendicular to the alignment of the flakes, in pounds per square inch (psi). Some of the data are given in the following table. The sizes of the flakes are $S_1 = 0.015$ inches by 2 inches and $S_2 = 0.025$ inches by 2 inches. (Data provided by Michael Hunt and Bob Lattanzi of the Purdue University Forestry Department.)

<table>
<thead>
<tr>
<th>Species</th>
<th>Size of flakes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S1</td>
</tr>
<tr>
<td>Aspen</td>
<td>308</td>
</tr>
<tr>
<td></td>
<td>428</td>
</tr>
<tr>
<td></td>
<td>426</td>
</tr>
<tr>
<td>Birch</td>
<td>214</td>
</tr>
<tr>
<td></td>
<td>433</td>
</tr>
<tr>
<td></td>
<td>231</td>
</tr>
<tr>
<td>Maple</td>
<td>272</td>
</tr>
<tr>
<td></td>
<td>376</td>
</tr>
<tr>
<td></td>
<td>322</td>
</tr>
</tbody>
</table>
(a) Compute means and standard deviations for the three observations in each species-size group. Find the marginal mean for each species and for each size of flakes. Display the means and marginal means in a table.

(b) Plot the means of the six groups. Put species on the $x$ axis and modulus of elasticity on the $y$ axis. For each size connect the three points corresponding to the different species. Describe the patterns you see. Do the species appear to be different? What about the sizes? Does there appear to be an interaction?

(c) Run a two-way ANOVA on these data. Summarize the results of the significance tests. What do these results say about the impressions that you described in part (b) of this exercise?

13.6 Refer to the previous exercise. Another of the physical properties measured was the strength, in kilopounds per square inch (ksi), in the direction perpendicular to the alignment of the flakes. Some of the data are given in the following table. The sizes of the flakes are $S_1 = 0.015$ inches by 2 inches and $S_2 = 0.025$ inches by 2 inches.

<table>
<thead>
<tr>
<th>Species</th>
<th>S1</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspen</td>
<td>1296</td>
<td>1472</td>
</tr>
<tr>
<td></td>
<td>1997</td>
<td>1441</td>
</tr>
<tr>
<td></td>
<td>1686</td>
<td>1051</td>
</tr>
<tr>
<td>Birch</td>
<td>903</td>
<td>1422</td>
</tr>
<tr>
<td></td>
<td>1246</td>
<td>1376</td>
</tr>
<tr>
<td></td>
<td>1355</td>
<td>1238</td>
</tr>
<tr>
<td>Maple</td>
<td>1211</td>
<td>1440</td>
</tr>
<tr>
<td></td>
<td>1827</td>
<td>1238</td>
</tr>
<tr>
<td></td>
<td>1541</td>
<td>748</td>
</tr>
</tbody>
</table>

(a) Compute means and standard deviations for the three observations in each species-size group. Find the marginal means for the species and for the flake sizes.
Display the means and marginal means in a table.
(b) Plot the means of the six groups. Put species on the $x$ axis and strength on the $y$ axis. For each size connect the three points corresponding to the different species. Describe the patterns you see. Do the species appear to be different? What about the sizes? Does there appear to be an interaction?
(c) Run a two-way ANOVA on these data. Summarize the results of the significance tests. What do these results say about the impressions that you described in part (b) of this exercise?

13.7 Each of the following situations is a two-way study design. For each case, identify the response variable and both factors, and state the number of levels for each factor ($I$ and $J$) and the total number of observations ($N$).
(a) A company wants to compare three different training programs for its new employees. Each of these programs takes 8 hours to complete. The training can be given for 8 hours on one day or for 4 hours on two consecutive days. The next 120 employees that the company hires will be the subjects for this study. After the training is completed, the employees are asked to evaluate the effectiveness of the program on a 7-point scale.
(b) A marketing experiment compares four different types of packaging for computer disks. Each type of packaging can be presented in three different colors. Each combination of package type with a particular color is shown to 40 different potential customers, who rate the attractiveness of the product on a 1 to 10 scale.
(c) Five different formulations for your hand lotion product have been produced by your research and marketing group, and you want to decide which of these, if any, to market. The lotions can be made with three different fragrances. Samples of each formulation-by-fragrance lotion are sent to 120 randomly selected customers who use your regular product. You ask each customer to compare the new lotion with the regular product by rating it on a 7-point scale. The middle point of the scale
corresponds to no preference, while higher values indicate that the new product is preferred and lower values indicate that the regular product is better.

13.8 For each part of the previous exercise, outline the ANOVA table, giving the sources of variation and the degrees of freedom. (Do not compute the numerical values for the sums of squares and mean squares.)

13.9 Each of the following situations is a two-way study design. For each case, identify the response variable and both factors, and state the number of levels for each factor \((I\) and \(J))\) and the total number of observations \((N)\).

(a) A study of smoking classifies subjects as nonsmokers, moderate smokers, or heavy smokers. Samples of 80 men and 80 women are drawn from each group. Each person reports the number of hours of sleep he or she gets on a typical night.

(b) The strength of concrete depends upon the formula used to prepare it. An experiment compares six different mixtures. Nine specimens of concrete are poured from each mixture. Three of these specimens are subjected to 0 cycles of freezing and thawing, three are subjected to 100 cycles, and three specimens are subjected to 500 cycles. The strength of each specimen is then measured.

(c) Four methods for teaching sign language are to be compared. Sixteen students in special education and sixteen students majoring in other areas are the subjects for the study. Within each group they are randomly assigned to the methods. Scores on a final exam are compared.

13.10 For each part of the previous exercise, outline the ANOVA table, giving the sources of variation and the degrees of freedom. (Do not compute the numerical values for the sums of squares and mean squares.)

13.11 A two-way ANOVA model was used to analyze an experiment with three levels of one factor, four levels of a second factor, and 6 observations per treatment combination.
Exercises

(a) For each of the main effects and the interaction, give the degrees of freedom for the corresponding \( F \) statistic.

(b) Using Table E or statistical software, find the value that each of these \( F \) statistics must exceed for the result to be significant at the 5% level.

(c) Answer part (b) for the 1% level.

13.12 A two-way ANOVA model was used to analyze an experiment with two levels of one factor, three levels of a second factor, and 6 observations per treatment combination.

(a) For each of the main effects and the interaction, give the degrees of freedom for the corresponding \( F \) statistic.

(b) Using Table E or statistical software, find the value that each of these \( F \) statistics must exceed for the result to be significant at the 5% level.

(c) Answer part (b) for the 1% level.

13.13 In the course of a clinical trial of measures to prevent coronary heart disease, blood pressure measurements were taken on 12,866 men. Individuals were classified by age group and race. (W. M. Smith et al., “The multiple risk factor intervention trial,” in H. M. Perry, Jr., and W. M. Smith (eds.), Mild Hypertension: To Treat or Not to Treat, New York Academy of Sciences, 1978, pp. 293–308.) The means for systolic blood pressure are given in the following table:

<table>
<thead>
<tr>
<th></th>
<th>35–39</th>
<th>40–44</th>
<th>45–49</th>
<th>50–54</th>
<th>55–59</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>131.0</td>
<td>132.3</td>
<td>135.2</td>
<td>139.4</td>
<td>142.0</td>
</tr>
<tr>
<td>Nonwhite</td>
<td>132.3</td>
<td>134.2</td>
<td>137.2</td>
<td>141.3</td>
<td>144.1</td>
</tr>
</tbody>
</table>

(a) Plot the group means, with age on the \( x \) axis and blood pressure on the \( y \) axis. For each racial group connect the points for the different ages.

(b) Describe the patterns you see. Does there appear to be a difference between the two racial groups? Does systolic blood pressure appear to vary with age? If so, how
does it vary? Is there an interaction?
(c) Compute the marginal means. Then find the differences between the white and nonwhite mean blood pressures for each age group. Use this information to summarize numerically the patterns in the plot.

13.14 The means for diastolic blood pressure recorded in the clinical trial described in the previous exercise are:

<table>
<thead>
<tr>
<th></th>
<th>35–39</th>
<th>40–44</th>
<th>45–49</th>
<th>50–54</th>
<th>55–59</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>89.4</td>
<td>90.2</td>
<td>90.9</td>
<td>91.6</td>
<td>91.4</td>
</tr>
<tr>
<td>Nonwhite</td>
<td>91.2</td>
<td>93.1</td>
<td>93.3</td>
<td>94.5</td>
<td>93.5</td>
</tr>
</tbody>
</table>

(a) Plot the group means with age on the x axis and blood pressure on the y axis. For each racial group connect the points for the different ages.
(b) Describe the patterns you see. Does there appear to be a difference between the two racial groups? Does diastolic blood pressure appear to vary with age? If so, how does it vary? Is there an interaction between race and age?
(c) Compute the marginal means. Find the differences between the white and nonwhite mean blood pressures for each age group. Use this information to summarize numerically the patterns in the plot.

13.15 The Chapin Social Insight Test measures how well people can appraise others and predict what they may say or do. A study administered this test to different groups of people and compared the mean scores. (This exercise is based on results reported in H. G. Gough, *The Chapin Social Insight Test*, Consulting Psychologists Press, 1968.) Some of the results are given in the table below. Means for males and females who were psychology graduate students (PG) and liberal arts undergraduates (LU) are presented. The two factors are labeled Gender and Group.
Plot the means and describe the essential features of the data in terms of main effects and interactions.

13.16 Refer to the previous exercise. Part of the ANOVA table for these data is given below:

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (Gender)</td>
<td>62.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B (Group)</td>
<td>1,599.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB</td>
<td></td>
<td>13,633.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>15,458.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>15,458.52</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) There were 150 individuals tested in each of the groups. Fill in the missing values in the ANOVA table.

(b) What is the value of the $F$ statistic to test the null hypothesis that there is no interaction? What is its distribution when the null hypothesis is true? Using Table E, find an approximate $P$-value for this test.

(c) Answer the questions in part (b) for the main effect of Gender and the main effect of Group.

(d) What is $s_p^2$, the within-group variance? What is $s_p$?

(e) Using what you have learned in this exercise and your answer to Exercise 13.10, summarize the results of this study.
<table>
<thead>
<tr>
<th>State</th>
<th>Region</th>
<th>Pay</th>
<th>Spend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Me.</td>
<td>NE</td>
<td>32.0</td>
<td>6.41</td>
</tr>
<tr>
<td>Vt.</td>
<td>NE</td>
<td>35.4</td>
<td>7.37</td>
</tr>
<tr>
<td>R.I.</td>
<td>NE</td>
<td>40.7</td>
<td>7.36</td>
</tr>
<tr>
<td>N.Y.</td>
<td>MA</td>
<td>47.6</td>
<td>9.45</td>
</tr>
<tr>
<td>Pa.</td>
<td>MA</td>
<td>44.5</td>
<td>7.20</td>
</tr>
<tr>
<td>Ind.</td>
<td>ENC</td>
<td>36.8</td>
<td>6.00</td>
</tr>
<tr>
<td>Mich.</td>
<td>ENC</td>
<td>47.4</td>
<td>6.93</td>
</tr>
<tr>
<td>Minn.</td>
<td>ENC</td>
<td>35.9</td>
<td>5.11</td>
</tr>
<tr>
<td>Mo.</td>
<td>ENC</td>
<td>31.2</td>
<td>4.97</td>
</tr>
<tr>
<td>S.Dak.</td>
<td>WNC</td>
<td>26.0</td>
<td>4.84</td>
</tr>
<tr>
<td>Kans.</td>
<td>WNC</td>
<td>34.7</td>
<td>5.76</td>
</tr>
<tr>
<td>Md.</td>
<td>SA</td>
<td>40.7</td>
<td>6.72</td>
</tr>
<tr>
<td>Va.</td>
<td>SA</td>
<td>34.0</td>
<td>5.66</td>
</tr>
<tr>
<td>N.C.</td>
<td>SA</td>
<td>30.8</td>
<td>4.95</td>
</tr>
<tr>
<td>Ga.</td>
<td>SA</td>
<td>32.6</td>
<td>5.40</td>
</tr>
<tr>
<td>Ky.</td>
<td>ESC</td>
<td>32.3</td>
<td>5.61</td>
</tr>
<tr>
<td>Ala.</td>
<td>ESC</td>
<td>31.1</td>
<td>4.46</td>
</tr>
<tr>
<td>Ark.</td>
<td>WSC</td>
<td>28.9</td>
<td>4.26</td>
</tr>
<tr>
<td>Okla.</td>
<td>WSC</td>
<td>28.2</td>
<td>4.38</td>
</tr>
<tr>
<td>Mont.</td>
<td>MN</td>
<td>28.8</td>
<td>5.83</td>
</tr>
<tr>
<td>Wyo.</td>
<td>MN</td>
<td>31.3</td>
<td>6.07</td>
</tr>
<tr>
<td>N.Mex.</td>
<td>MN</td>
<td>28.5</td>
<td>5.42</td>
</tr>
<tr>
<td>Utah</td>
<td>MN</td>
<td>29.1</td>
<td>3.67</td>
</tr>
<tr>
<td>Wash.</td>
<td>PA</td>
<td>36.2</td>
<td>5.81</td>
</tr>
<tr>
<td>Calif.</td>
<td>PA</td>
<td>41.1</td>
<td>4.73</td>
</tr>
<tr>
<td>Hawaii</td>
<td>PA</td>
<td>38.5</td>
<td>6.16</td>
</tr>
</tbody>
</table>