SUPPLEMENTARY EXERCISES

for

INTRODUCTION TO THE PRACTICE OF STATISTICS

Seventh Edition

David S. Moore
George P. McCabe
and
Bruce A. Craig

These exercises appeared in the third, fourth, fifth, or sixth editions of *Introduction to the Practice of Statistics*. They do not appear in the seventh edition, 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:

Fund	Category	Net assets (millions of $)	Year-to-date return	Largest holding
⋮				
Fidelity Low-Priced Stock	Small value	6,189	4.56%	Dallas Semiconductor
Price International Stock	International stock	9,745	−0.45%	Vodafone
Vanguard 500 Index	Large blend	89,394	;45%	General Electric
⋮				

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 and 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 1997–1998: 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 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. Although 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:

Decade	Mean	Std Dev
1910s	.266	.0371
1940s	.267	.0326
1970s	.261	.0317

Compute the standardized batting averages for Cobb, Williams, and Brett to compare how far each stood above his peers. (Data from Stephen Jay Gould, "Entropic homogeneity isn't why no one hits .400 any more," *Discover*, August 1986, pp. 60–66. Gould does not standardize but gives a speculative discussion instead.)

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):

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

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 Maribeth Cassidy Schmitt, from her PhD 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 following table gives the number of medical doctors per 100,000 people in each state:

State	Doctors	State	Doctors	State	Doctors
Alabama	198	Louisiana	246	Ohio	235
Alaska	167	Maine	223	Oklahoma	169
Arizona	202	Maryland	374	Oregon	225
Arkansas	190	Massachusetts	412	Pennsylvania	291
California	247	Michigan	224	Rhode Island	338
Colorado	238	Minnesota	249	South Carolina	207
Connecticut	354	Mississippi	163	South Dakota	184
Delaware	234	Missouri	230	Tennessee	246
Florida	238	Montana	190	Texas	203
Georgia	211	Nebraska	218	Utah	200
Hawaii	265	Nevada	173	Vermont	305
Idaho	154	New Hampshire	237	Virginia	241
Illinois	260	New Jersey	295	Washington	235
Indiana	195	New Mexico	212	West Virginia	215
Iowa	173	New York	387	Wisconsin	227
Kansas	203	North Carolina	232	Wyoming	171
Kentucky	209	North Dakota	222	D.C.	737

(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):

−5.7	1.2	4.1	3.2	7.3	7.5	18.6	3.7	−1.8	2.4
−6.5	6.7	9.4	−2.0	−2.8	−3.4	19.2	−4.8	0.5	−0.6
2.8	−0.5	−4.5	8.7	2.7	4.1	−10.3	4.8	−2.3	−3.1
−10.2	−3.7	−26.6	7.2	−2.9	−2.3	3.5	−4.6	17.2	4.2
0.5	8.3	−7.1	−8.4	7.7	−9.6	6.0	6.8	10.9	1.6
0.2	−2.4	−2.4	3.9	1.7	9.0	3.6	7.6	3.2	−3.7
4.2	13.2	0.9	4.2	4.0	2.8	6.7	−10.4	2.7	10.3
5.7	0.6	−14.2	1.3	2.9	11.8	10.6	5.2	13.8	−14.7
3.5	11.7	1.3							

(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.

Salary	15–20	20–25	25–30	30–35	35–40	40–50	50–60	60–80
Percent	7	14	29	23	13	9	4	1

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):

Year	Count	Year	Count	Year	Count	Year	Count	Year	Count
1944	3	1956	2	1968	0	1980	2	1992	1
1945	2	1957	2	1969	3	1981	3	1993	1
1946	1	1958	4	1970	2	1982	1	1994	0
1947	2	1959	2	1971	1	1983	1	1995	5
1948	4	1960	2	1972	0	1984	1	1996	6
1949	3	1961	6	1973	1	1985	3	1997	1
1950	7	1962	0	1974	2	1986	0	1998	3
1951	2	1963	2	1975	3	1987	1	1999	5
1952	3	1964	5	1976	2	1988	3	2000	3
1953	3	1965	1	1977	1	1989	2		
1954	2	1966	3	1978	2	1990	1		
1955	5	1967	1	1979	2	1991	2		

(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 following table 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?
(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?

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

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.

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

(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 following table gives the number of home runs Ruth hit in each year. Make a time plot and describe its main features.

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

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:

Year	Time	Year	Time	Year	Time	Year	Time	Year	Time
1959	143	1970	131	1981	129	1992	128	2003	130
1960	141	1971	139	1982	129	1993	130	2004	131
1961	144	1972	136	1983	129	1994	127		
1962	144	1973	136	1984	131	1995	129		
1963	139	1974	134	1985	134	1996	129		
1964	140	1975	130	1986	128	1997	131		
1965	137	1976	140	1987	132	1998	128		
1966	137	1977	135	1988	129	1999	130		
1967	136	1978	130	1989	129	2000	130		
1968	142	1979	129	1990	128	2001	130		
1969	134	1980	132	1991	131	2002	129		

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?

1.29 Here is a small part of a data set that describes the fuel economy (in miles per gallon) of 2004 model motor vehicles:

Make and model	Vehicle type	Transmission type	Number of cylinders	City MPG	Highway MPG
⋮					
Acura NSX	Two-seater	Automatic	6	17	24
BMW 330I	Compact	Manual	6	20	30
Cadillac Seville	Midsize	Automatic	8	18	26
Ford F150 2WD	Standard pickup truck	Automatic	6	16	19
⋮					

(a) What are the individuals in this data set?
(b) For each individual, what variables are given? Which of these variables are categorical and which are quantitative?

1.30 Here are the first lines of a professor's data set at the end of a statistics course:

```
NAME            MAJOR   POINTS  GRADE

ADVANI, SURA    COMM    397     B
BARTON, DAVID   HIST    323     C
BROWN, ANNETTE  BIOL    446     A
CHIU, SUN       PSYC    405     B
CORTEZ, MARIA   PSYC    461     A
```

What are the individuals and the variables in these data? Which variables are categorical and which are quantitative?

1.31 How can we help wood surfaces resist weathering, especially when restoring historic wooden buildings? A study of this question prepared wooden panels and then exposed them to the weather. Here are some of the variables recorded. Which of these variables are categorical, and which are quantitative?
(a) Type of wood (yellow poplar, pine, cedar)
(b) Water repellent (solvent-based, water-based)
(c) Paint thickness (millimeters)
(d) Paint color (white, gray, light blue)
(e) Weathering time (months)

1.32 You are preparing to study the television-viewing habits of college students. Describe two categorical variables and two quantitative variables that you might measure for each student. Give the units of measurement for the quantitative variables.

1.33 You want to measure the "physical fitness" of college students. Describe several variables you might use to measure fitness. What instrument or instruments does each measurement require?

1.34 Popular magazines rank colleges and universities on their "academic quality" in serving undergraduate students. Describe five variables that you would like to see measured for each college if you were choosing where to study. Give reasons for each of your choices.

1.35 Is driving becoming more dangerous? Traffic deaths declined for years, bottoming out at 39,250 killed in 1992, then began to increase again. In 2002, 42,815 people died in traffic accidents. But more vehicles drove more miles in 2002 than in 1992. In fact, the government says that motor vehicles traveled 2247 billion miles in 1992 and 2830 billion miles in 2002. Reports on transportation use deaths per 100 million miles as a measure of risk. Compare the rates for 1992 and 2002. What do you conclude?

1.36 What color is your car? Here are the most popular colors for vehicles made in North America during the 2003 model year.

Color	Percent of vehicles
Silver	20.1
White	18.4
Black	11.6
Medium/dark gray	11.5
Light brown	8.8
Medium/dark blue	8.5
Medium red	6.9

What percent of vehicles had other colors? Display these data in a bar graph. Is it also correct to use a pie chart if a category for "Other" is included? If you are using software, make a pie chart if it is correct to do so.

1.37 Favorite vehicle colors differ among types of vehicle. Here are data on the most popular colors in 2003 for luxury cars and for SUVs, trucks, and vans. The entry "–" means "less than 1%." Be creative: make one bar graph that compares the two vehicle types as well as comparing colors. Arrange your graph so that it is easy to compare the two types of vehicle.

Color	Luxury car percent	SUV/truck/van percent
Black	10.9	11.6
Light brown	–	6.3
Medium/dark blue	3.8	9.3
Medium/dark gray	23.3	8.8
Medium/dark green	–	7.0
Medium red	3.9	6.2
White	30.4	22.3
Silver	18.8	17.0

1.38 The Department of Education estimates the "average unmet need" for undergraduate students—the cost of school minus estimated family contributions and financial aid. Here are the averages for full-time students at four types of institution in the 1999–2000 academic year:

Public 2-year	Public 4-year	Private nonprofit 4-year	Private for-profit
$2747	$2369	$4931	$6548

Make a bar graph of these data. Write a one-sentence conclusion about the unmet needs of students. Explain clearly why it is incorrect to make a pie chart.

1.39 The J.D. Power Initial Quality Survey polls more than 50,000 buyers of new motor vehicles 90 days after their purchase. A two-page questionnaire asks about "things gone wrong." Here are data on problems per 100 vehicles for vehicles made by Toyota and by General Motors in recent years. Toyota has been the industry leader in quality. Make two time plots in the same graph to compare Toyota and GM. What are the most important conclusions you can draw from your graph?

Year	1998	1999	2000	2001	2002	2003	2004
GM	187	179	164	147	130	134	120
Toyota	156	134	116	115	107	115	101

1.40 Study habits of students. You are planning a survey to collect information about the study habits of college students. Describe two categorical variables and two quantitative variables that you might measure for each student. Give the units of measurement for the quantitative variables.

1.41 Physical fitness of students. You want to measure the "physical fitness" of college students. Describe several variables you might use to measure fitness. What instrument or instruments does each measurement require?

1.42 An aging population. The population of the United States is aging, though less rapidly than in other developed countries. Here is a stemplot of the percents of residents aged 65 and over in the 50 states, according to the 2000 census. The stems are whole percents and the leaves are tenths of a percent.

5	7
6	
7	
8	5
9	679
10	6
11	02233677
12	0011113445789
13	00012233345568
14	034579
15	36
16	
17	6

(a) There are two outliers: Alaska has the lowest percent of older residents, and Florida has the highest. What are the percents for these two states?
(b) Ignoring Alaska and Florida, describe the shape, center, and spread of this distribution.

1.43 Split the stems. Make another stemplot of the percent of residents aged 65 and over in the states other than Alaska and Florida by splitting stems 8 to 15 in the plot from the previous exercise. Which plot do you prefer? Why?

1.44 Tornado damage. The states differ greatly in the kinds of severe weather that afflict them. The following table shows the average property damage caused

by tornadoes per year over the period from 1950 to 1999 in each of the 50 states and Puerto Rico. (National Climatic Data Center, storm events data base. See `sciencepolicy.colorado.edu/sourcebook/tornadoes.html`.) (To adjust for the changing buying power of the dollar over time, all damages were restated in 1999 dollars.)
(a) What are the top five states for tornado damage? The bottom five?
(b) Make a histogram of the data, by hand or using software, with classes "$0 \leq$ damage < 10," "$10 \leq$ damage < 20," and so on. Describe the shape, center, and spread of the distribution. Which states may be outliers? (To understand the outliers, note that most tornadoes in largely rural states such as Kansas cause little property damage. Damage to crops is not counted as property damage.)
(c) If you are using software, also display the "default" histogram that your software makes when you give it no instructions. How does this compare with your graph in (b)?

Average property damage per year due to tornadoes

State	Damage ($millions)	State	Damage ($millions)	State	Damage ($millions)
Alabama	51.88	Louisiana	27.75	Ohio	44.36
Alaska	0.00	Maine	0.53	Oklahoma	81.94
Arizona	3.47	Maryland	2.33	Oregon	5.52
Arkansas	40.96	Massachusetts	4.42	Pennsylvania	17.11
California	3.68	Michigan	29.88	Puerto Rico	0.05
Colorado	4.62	Minnesota	84.84	Rhode Island	0.09
Connecticut	2.26	Mississippi	43.62	South Carolina	17.19
Delaware	0.27	Missouri	68.93	South Dakota	10.64
Florida	37.32	Montana	2.27	Tennessee	23.47
Georgia	51.68	Nebraska	30.26	Texas	88.60
Hawaii	0.34	Nevada	0.10	Utah	3.57
Idaho	0.26	New Hampshire	0.66	Vermont	0.24
Illinois	62.94	New Jersey	2.94	Virginia	7.42
Indiana	53.13	New Mexico	1.49	Washington	2.37
Iowa	49.51	New York	15.73	West Virginia	2.14
Kansas	49.28	North Carolina	14.90	Wisconsin	31.33
Kentucky	24.84	North Dakota	14.69	Wyoming	1.78

1.45 Use an applet for the tormado damage data. The *One-Variable Statistical Calculator* applet on the text CD and Web site will make stemplots and histograms. It is intended mainly as a learning tool rather than as a replacement for statistical software. The histogram function is particularly useful because you can change the number of classes by dragging with the mouse. The tornado damage data from Exercise 1.44 are available in the applet. Choose this data set and go to the "Histogram" tab.
(a) Sketch the default histogram that the applet first presents. If the default graph does not have nine classes, drag it to make a histogram with nine classes and sketch the result. This should agree with your histogram in part (b) of the previous exercise.
(b) Make a histogram with one class and also a histogram with the greatest number

of classes that the applet allows. Sketch the results.
(c) Drag the graph until you find the histogram that you think best pictures the data. How many classes did you choose? Sketch your final histogram.

1.46 Caliornia temperatures. The following table contains data on the mean annual temperatures (degrees Fahrenheit) for the years 1951 to 2000 at two locations in California: Pasadena and Redding. (Data from the U.S. Historical Climatology Network, archived at `www.co2science.org`. Despite claims made on this site, temperatures at most U.S. locations show a gradual increase over the past century.) Make time plots of both time series and compare their main features. You can see why discussions of climate change often bring disagreement.

Mean annual temperatures (°F) in two California cities

	Mean Temperature			Mean Temperature	
Year	Pasadena	Redding	Year	Pasadena	Redding
1951	62.27	62.02	1976	64.23	63.51
1952	61.59	62.27	1977	64.47	63.89
1953	62.64	62.06	1978	64.21	64.05
1954	62.88	61.65	1979	63.76	60.38
1955	61.75	62.48	1980	65.02	60.04
1956	62.93	63.17	1981	65.80	61.95
1957	63.72	62.42	1982	63.50	59.14
1958	65.02	64.42	1983	64.19	60.66
1959	65.69	65.04	1984	66.06	61.72
1960	64.48	63.07	1985	64.44	60.50
1961	64.12	63.50	1986	65.31	61.76
1962	62.82	63.97	1987	64.58	62.94
1963	63.71	62.42	1988	65.22	63.70
1964	62.76	63.29	1989	64.53	61.50
1965	63.03	63.32	1990	64.96	62.22
1966	64.25	64.51	1991	65.60	62.73
1967	64.36	64.21	1992	66.07	63.59
1968	64.15	63.40	1993	65.16	61.55
1969	63.51	63.77	1994	64.63	61.63
1970	64.08	64.30	1995	65.43	62.62
1971	63.59	62.23	1996	65.76	62.93
1972	64.53	63.06	1997	66.72	62.48
1973	63.46	63.75	1998	64.12	60.23
1974	63.93	63.80	1999	64.85	61.88
1975	62.36	62.66	2000	66.25	61.58

1.47 What do you miss in the histogram? Make a histogram of the mean annual temperatures at Pasadena for the years 1951 to 2000 (data appear in Exercise 1.46). Describe the distribution of temperatures. Then explain why this histogram misses very important facts about temperatures in Pasadena.

1.48 Change the scale of the axis. 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 for Pasadena in Exercise 1.46, one that makes mean temperature appear to increase very rapidly and one that shows only a slow increase. The moral of this exercise is: *pay close attention to the scales when you look at a time plot.*

1.49 Fish in the Bering Sea. "Recruitment," the addition of new members to a fish population, is an important measure of the health of ocean ecosystems. Here are data on the recruitment of rock sole in the Bering Sea between 1973 and 2000 (National Oceanic and Atmospheric Administration, `www.noaa.gov`):

Year	Recruitment (millions)	Year	Recruitment (millions)	Year	Recruitment (millions)	Year	Recruitment (millions)
1973	173	1980	1411	1987	4700	1994	505
1974	234	1981	1431	1988	1702	1995	304
1975	616	1982	1250	1989	1119	1996	425
1976	344	1983	2246	1990	2407	1997	214
1977	515	1984	1793	1991	1049	1998	385
1978	576	1985	1793	1992	505	1999	445
1979	727	1986	2809	1993	998	2000	676

(a) Make a graph to display the distribution of rock sole recruitment, then describe the pattern and any striking deviations that you see.
(b) Make a time plot of recruitment and describe its pattern. As is often the case with time series data, a time plot is needed to understand what is happening.

1.50 Oil wells. How much oil the wells in a given field will ultimately produce is key information in deciding whether to drill more wells. Here are the estimated total amounts of oil recovered from 64 wells in the Devonian Richmond Dolomite area of the Michigan basin, in thousands of barrels (J. Marcus Jobe and Hutch Jobe, "A statistical approach for additional infill development," *Energy Exploration and Exploitation,* 18 (2000), pp. 89–103):

21.7	53.2	46.4	42.7	50.4	97.7	103.1	51.9	43.4	69.5
156.5	34.6	37.9	12.9	2.5	31.4	79.5	26.9	18.5	14.7
32.9	196.0	24.9	118.2	82.2	35.1	47.6	54.2	63.1	69.8
57.4	65.6	56.4	49.4	44.9	34.6	92.2	37.0	58.8	21.3
36.6	64.9	14.8	17.6	29.1	61.4	38.6	32.5	12.0	28.3
204.9	44.5	10.3	37.7	33.7	81.1	12.1	20.1	30.5	7.1
10.1	18.0	3.0	2.0						

Graph the distribution and describe its main features.

1.51 Guinea pigs. The following table gives the survival times in days of 72 guinea pigs after they were injected with tubercle bacilli in a medical experiment. (T. Bjerkedal, "Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli," *American Journal of Hygiene,* 72 (1960), pp. 130–148.) Make a suitable graph and describe the shape, center, and spread of the distribution of survival times. Are there any outliers?

Survival times (days) of guinea pigs

43	45	53	56	56	57	58	66	67	73
74	79	80	80	81	81	81	82	83	83
84	88	89	91	91	92	92	97	99	99
100	100	101	102	102	102	103	104	107	108
109	113	114	118	121	123	126	128	137	138
139	144	145	147	156	162	174	178	179	184
191	198	211	214	243	249	329	380	403	511
522	598								

Section 1.2

1.52 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.

1.53 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.

1.54 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

446 552 527 504 450 583 501 545 549 506

(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 $\overline{x}$ and s to summarize these data. Find the standard deviation s. Do you agree that this is a suitable summary?

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

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

and for 20 first-year college men:

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 these data, or use your result from Exercise 1.15.

(b) Find the mean $\overline{x}$ and the median M for both sets of SSHA scores. What feature of each distribution explains the fact that $\overline{x} > M$?
(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 $1.5 \times IQR$ 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?

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

1.57 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 $\overline{x}$ 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.

1.58 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.)

1.59 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%?

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

Year	1948	1952	1956	1960	1964	1968	1972
Percent	49.6	55.1	57.4	49.7	61.1	43.4	60.7
Year	1976	1980	1984	1988	1992	1996	2000
Percent	50.1	50.7	58.8	53.9	43.2	49.2	47.9

(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?

1.61 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:

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 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 following table 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 1.62 to 1.64 analyze these data (*Consumer Reports*, June 1986, pp. 366–367).*

Beef hot dogs		Meat hot dogs		Poultry hot dogs	
Calories	Sodium	Calories	Sodium	Calories	Sodium
186	495	173	458	129	430
181	477	191	506	132	375
176	425	182	473	102	396
149	322	190	545	106	383
184	482	172	496	94	387
190	587	147	360	102	542
158	370	146	387	87	359
139	322	139	386	99	357
175	479	175	507	170	528
148	375	136	393	113	513
152	330	179	405	135	426
111	300	153	372	142	513
141	386	107	144	86	358
153	401	195	511	143	581
190	645	135	405	152	588
157	440	140	428	146	522
131	317	138	339	144	545
149	319				
135	298				
132	253				

1.62 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.

1.63 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.

1.64 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?

1.65 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 $\overline{x}$ 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 $\overline{x}$ and s? Do these distributions seem to fit the criteria?

1.66 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.

1.67 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_{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_{new}. (One horsepower is 746 watts.) What is the power in watts of a 140-horsepower engine?

1.68 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_{new} between x and the "normal" body temperature of 98.6°.
(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.

1.69 This is a standard deviation contest. You must choose four numbers from the whole numbers 0 to 10, with repeats allowed.
(a) Choose four numbers that have the smallest possible standard deviation.
(b) Choose four numbers that have the largest possible standard deviation.
(c) Is more than one choice possible in either (a) or (b)? Explain.

1.70 This exercise requires a calculator with a standard deviation button or statistical software on a computer. The observations

$$10,001 \quad 10,002 \quad 10,003$$

have mean $\overline{x} = 10,002$ and standard deviation $s = 1$. Adding a 0 in the center of each number, the next set becomes

$$100,001 \quad 100,002 \quad 100,003$$

The standard deviation remains $s = 1$ as more 0s are added. Use your calculator or computer to calculate the standard deviation of these numbers, adding extra 0s until you get an incorrect answer. How soon did you go wrong? This demonstrates that calculators and computers cannot handle an arbitrary number of digits correctly.

1.71 College tuition and fees. A study examined the tuition and fees charged by the 56 four-year colleges in the state of Massachusetts. Here are those charges (in dollars), arranged in increasing order:

4,123	4,186	4,324	4,342	4,557	4,884	5,397	6,129
6,963	6,972	8,232	13,584	13,612	15,500	15,934	16,230
16,696	16,700	17,044	17,500	18,550	18,750	19,145	19,300
19,410	19,700	19,700	19,910	20,234	20,400	20,640	20,875
21,165	21,302	22,663	23,550	24,324	25,840	26,965	27,522
27,544	27,904	28,011	28,090	28,420	28,420	28,900	28,906
28,950	29,060	29,338	29,392	29,600	29,624	29,630	29,875

Find the five-number summary and make a boxplot. What distinctive feature of the histogram do these summaries miss? Remember that numerical summaries are not a substitute for looking at the data.

1.72 Outliers in percent of older residents. The stemplot in Exercise 1.42 displays the distribution of the percents of residents aged 65 and over in the 50 states. Stemplots help you find the five-number summary because they arrange the observations in increasing order.
(a) Give the five-number summary of this distribution.
(b) Does the $1.5 \times IQR$ rule identify Alaska and Florida as suspected outliers? Does it also flag any other states?

1.73 Tormados and property damage. Exercise 1.44 gives the average property damage caused by tornadoes over a 50-year period in each of the states. The distribution is strongly skewed to the right.
(a) Give the five-number summary. Explain why you can see from these five numbers that the distribution is right-skewed.
(b) A histogram or stemplot suggests that a few states are outliers. Show that there are *no* suspected outliers according to the $1.5 \times IQR$ rule. You see once again that a rule is not a substitute for plotting your data.
(c) Find the mean property damage. Explain why the mean and median differ so greatly for this distribution.

1.74 Mean versus median for oil wells. Exercise 1.50 gives data on the total oil recovered from 64 wells. Your graph in that exercise shows that the distribution is clearly right-skewed.
(a) Find the mean and median of the amounts recovered. Explain how the relationship between the mean and the median reflects the shape of the distribution.
(b) Give the five-number summary and explain briefly how it reflects the shape of the distribution.

1.75 Effects of logging in Borneo. "Conservationists have despaired over destruction of tropical rainforest by logging, clearing, and burning." These words begin a report on a statistical study of the effects of logging in Borneo. Researchers compared forest plots that had never been logged (Group 1) with similar plots nearby that had been logged 1 year earlier (Group 2) and 8 years earlier (Group 3). All plots were 0.1 hectare in area. Here are the counts of trees for plots in each group (we thank Ethan J. Temeles of Amherst College for providing the data. His work is described in Ethan J. Temeles and W. John Kress, "Adaptation in a plant-hummingbird association," *Science,* 300 (2003), pp. 630–633):

Group 1:	27	22	29	21	19	33	16	20	24	27	28	19
Group 2:	12	12	15	9	20	18	17	14	14	2	17	19
Group 3:	18	4	22	15	18	19	22	12	12			

Give a complete comparison of the three distributions, using both graphs and numerical summaries. To what extent has logging affected the count of trees? The researchers used an analysis based on $\bar{x}$ and s. Explain why this is reasonably well justified.

1.76 Running and heart rate. How does regular running affect heart rate? The RUNNERS data set contains heart rates for four groups of people:

Sedentary females
Sedentary males
Female runners (at least 15 miles per week)
Male runners (at least 15 miles per week)

The heart rates were measured after 6 minutes of exercise on a treadmill. There are 200 subjects in each group. Give a complete comparison of the four distributions, using both graphs and numerical summaries. How would you describe the effect of running on heart rate? Is the effect different for men and women?

1.77 Guinea pigs. Exercise 1.51 gives the survival times of 72 guinea pigs in a medical study. Survival times—whether of cancer patients after treatment or of car batteries in everyday use—are almost always right-skewed. Make a graph to verify that this is true of these survival times. Then give a numerical summary that is appropriate for such data. Explain in simple language, to someone who knows no statistics, what your summary tells us about the guinea pigs.

1.78 Weight gain. A study of diet and weight gain deliberately overfed 16 volunteers for eight weeks. The mean increase in fat was $\bar{x} = 2.39$ kilograms and the standard deviation was $s = 1.14$ kilograms. What are $\bar{x}$ and s in pounds? (A kilogram is 2.2 pounds.)

1.79 Guinea pigs. Find the **quintiles** (the 20th, 40th, 60th, and 80th percentiles) of the guinea pig survival times in Exercise 1.51. For quite large sets of data, the quintiles or the **deciles** (10th, 20th, 30th, etc. , percentiles) give a more detailed summary than the quartiles.

1.80 Changing units from inches to centimeters. Changing the unit of length from inches to centimeters multiplies each length by 2.54 because there are 2.54 centimeters in an inch. This change of units multiplies our usual measures of spread by 2.54. This is true of IQR and the standard deviation. What happens to the variance when we change units in this way?

Section 1.3

1.81 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.

1.82 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.)

1.83 Give an interval that contains the middle 95% of NOX levels in the exhaust of trucks using the model described in Exercise 1.81.

1.84 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 1.82.

1.85 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?

1.86 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):

Decade	Mean	Std. dev.
1910s	.266	.0371
1940s	.267	.0326
1970s	.261	.0317

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

1.87 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)?

1.88 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?)

1.89 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) $500 < X < 800$

1.90 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) $-1.66 < Z < 2.85$

1.91 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) $-2.25 < Z < 1.77$

1.92 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.

1.93 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.

1.94 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?

1.95 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%?

1.96 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?

1.97 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?

1.98 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?

1.99 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%?

1.100 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.)

1.101 The distribution of Internet access costs in Exercise 1.61 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.

1.102 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.

1.103 Exercise 1.16 presents data on the weight gains of chicks fed two types of corn. The researchers use $\overline{x}$ and s to summarize each of the two distributions. Make a Normal quantile plot for each group and report your findings. Is use of $\overline{x}$ and s justified?

1.104 Sketch density curves that might describe distributions with the following shapes:
(a) Symmetric, but with two peaks (that is, two strong clusters of observations).
(b) Single peak and skewed to the left.

1.105 Remember that it is areas under a density curve, not the height of the curve, that give proportions in a distribution. To illustrate this, sketch a density curve that has its mode (peak point) at 0 on the horizontal axis but has greater area within 0.25 on either side of 1 than within 0.25 on either side of 0.

1.106 If you ask a computer to generate "random numbers" between 0 and 1, you will get observations from a uniform distribution, whose density curve is 1 between 0 and 1. Use areas under this density curve to answer the following questions.
(a) Why is the total area under this curve equal to 1?
(b) What proportion of the observations lie above 0.75?
(c) What proportion of the observations lie between 0.25 and 0.75?

1.107 Many random number generators allow users to specify the range of the random numbers to be produced. Suppose that you specify that the outcomes are to be distributed uniformly between 0 and 2. Then the density curve of the outcomes has constant height between 0 and 2, and height 0 elsewhere.
(a) What is the height of the density curve between 0 and 2? Draw a graph of the density curve.
(b) Use your graph from (a) and the fact that areas under the curve are proportions of outcomes to find the proportion of outcomes that are less than 1.
(c) Find the proportion of outcomes that lie between 0.5 and 1.3.

1.108 One reason that Normal distributions are important is that they describe how the results of an opinion poll would vary if the poll were repeated many times. About 40% of adult Americans say they are afraid to go out at night because of crime. Take many randomly chosen samples of 1050 people. The proportions of people in these samples who stay home for fear of crime will follow the Normal distribution with mean 0.4 and standard deviation 0.015. Use this fact and the 68–95–99.7 rule to answer these questions.
(a) In many samples, what percent of samples give results above 0.4? Above 0.43?
(b) In a large number of samples, what range contains the central 95% of proportions of people who stay home because of crime?

1.109 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) $Z < 1.85$ (this is a cumulative proportion)
(b) $Z > 1.85$
(c) $Z > -0.66$
(d) $-0.66 < Z < 1.85$

1.110 Using either Table A or your calculator or software, find the proportion of observations from a standard Normal distribution for each of the following events. In each case, sketch a standard Normal curve and shade the area representing the proportion.
(a) $Z \leq -2$ (this is a cumulative proportion)
(b) $Z \geq -2$
(c) $Z > 1.67$
(d) $-2 < Z < 1.67$

1.111 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) 20% of the observations fall below z.
(b) 30% of the observations fall above z.

1.112 The variable Z has a standard Normal distribution.
(a) Find the number z that has cumulative proportion 0.8.
(b) Find the number z such that the event $Z > z$ has proportion 0.45.

There are two major tests of readiness for college, the ACT and the SAT. ACT scores are reported on a scale from 1 to 36. The distribution of ACT scores for more than 1 million students in a recent high school graduating class was roughly Normal with mean $\mu = 20.8$ *and standard deviation* $\sigma = 4.8$*. SAT scores are reported on a scale from 400 to 1600. The SAT scores for 1.4 million students in the same graduating class were roughly Normal with mean* $\mu = 1026$ *and standard deviation* $\sigma = 209$*. Exercises 1.113 to 1.122 are based on this information.*

1.113 Tonya scores 1318 on the SAT. Jermaine scores 27 on the ACT. Assuming that both tests measure the same thing, who has the higher score?

1.114 Jacob scores 16 on the ACT. Emily scores 670 on the SAT. Assuming that both tests measure the same thing, who has the higher score?

1.115 Jose scores 1287 on the SAT. Assuming that both tests measure the same thing, what score on the ACT is equivalent to Jose's SAT score?

1.116 Maria scores 28 on the ACT. Assuming that both tests measure the same thing, what score on the SAT is equivalent to Maria's ACT score?

1.117 Reports on a student's ACT or SAT usually give the percentile as well as the actual score. The percentile is just the cumulative proportion stated as a percent: the percent of all scores that were lower than this one. Tonya scores 1318 on the SAT. What is her percentile?

1.118 Reports on a student's ACT or SAT usually give the percentile as well as the actual score. The percentile is just the cumulative proportion stated as a percent: the percent of all scores that were lower than this one. Jacob scores 16 on the ACT. What is his percentile?

1.119 It is possible to score higher than 1600 on the SAT, but scores 1600 and above are reported as 1600. What proportion of SAT scores are reported as 1600?

1.120 It is possible to score higher than 36 on the ACT, but scores 36 and above are reported as 36. What proportion of ACT scores are reported as 36?

1.121 What SAT scores make up the top 10% of all scores?

1.122 How well must Abigail do on the ACT in order to place in the top 20% of all students?

1.123 Changing the mean of a Normal distribution by a moderate amount can greatly change the percent of observations in the tails. Suppose that a college is looking for applicants with SAT Math scores 750 and above.
(a) In 2003, the scores of men on the math SAT followed a Normal distribution with mean 537 and standard deviation 116. What percent of men scored 750 or better?
(b) Women's scores that year had a Normal distribution with mean 503 and standard deviation 110. What percent of women scored 750 or better? You see that the percent of men above 750 is more than two and one-half times the percent of women with such high scores.

1.124 The yearly rate of return on stock indexes (which combine many individual stocks) is approximately Normal. Between 1900 and 2002, U.S. common stocks had a mean yearly return of 8.3%, with a standard deviation of about 20.3%. Take this Normal distribution to be the distribution of yearly returns over a long period.
(a) In what range do the middle 95% of all yearly returns lie?
(b) The market is down for the year if the return is less than zero. In what percent

of years is the market down?
(c) In what percent of years does the index gain 25% or more?

1.125 Binge drinking survey. One reason that Normal distributions are important is that they describe how the results of an opinion poll would vary if the poll were repeated many times. About 20% of college students say they are frequent binge drinkers. Think about taking many randomly chosen samples of 1600 students. The proportions of college students in these samples who say they are frequent binge drinkers will follow the Normal distribution with mean 0.20 and standard deviation 0.01. Use this fact and the 68–95–99.7 rule to answer these questions.
(a) In many samples, what percent of samples give results above 0.2? Above 0.22?
(b) In a large number of samples, what range contains the central 95% of proportions of students who say they are frequent binge drinkers?

1.126 Heights of women. The heights of women aged 20 to 29 are approximately Normal with mean 64 inches and standard deviation 2.7 inches. Men the same age have mean height 69.3 inches with standard deviation 2.8 inches. What are the z-scores for a woman 6 feet tall and a man 6 feet tall? What information do the z-scores give that the actual heights do not?

1.127 Proportion of women with high cholesterol. Too much cholesterol in the blood increases the risk of heart disease. Young women are generally less afflicted with high cholesterol than other groups. The cholesterol levels for women aged 20 to 34 follow an approximately Normal distribution with mean 185 milligrams per deciliter (mg/dl) and standard deviation 39 mg/dl. (Results for 1988 to 1991 from a large sample survey, reported in National Center for Health Statistics, *Health, United States, 1995,* 1996.) (a) Cholesterol levels above 240 mg/dl demand medical attention. What percent of young women have levels above 240 mg/dl?
(b) Levels above 200 mg/dl are considered borderline high. What percent of young women have blood cholesterol between 200 and 240 mg/dl?

1.128 Proportion of men with high cholesterol. Middle-aged men are more susceptible to high cholesterol than the young women of the previous exercise. The blood cholesterol levels of men aged 55 to 64 are approximately Normal with mean 222 mg/dl and standard deviation 37 mg/dl. What percent of these men have high cholesterol (levels above 240 mg/dl)? What percent have borderline high cholesterol (between 200 and 240 mg/dl)?

1.129 Logging in Borneo. The study of the effects of logging on tree counts in the Borneo rain forest (Exercise 1.75) used statistical methods that are based on Normal distributions. Make Normal quantile plots for each of the three groups of forest plots. Are the three distributions roughly Normal? What are the most prominent deviations from Normality that you see?

Chapter 1 Review Exercises

1.130 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	1344677889
13	0012455566789999
14	11222344445789
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.

1.131 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?

1.132 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.

1960	1920	1991	1919	1985	1985	1975	1980	1975	1985
1930	1993	1985	1975	1970	1970	1975	1980	1940	1940
1980	1919	1980	1950	1940	1950	1993	1985	1975	1960
1919	1950	1960	1975	1950	1919	1920	1985	1970	1975
1930	1975	1960	1920	1940	1950	1985	1990	1950	1970
1985	1920	1950	1980	1975	1950	1950	1919	1919	1985
1985	1991	1980	1960	1940	1960	1930	1998	1994	1960
1919	1975	1919	1950	1975	1930	1919	1970	1920	1930
1950	1975	1970	1985	1919	1960	1930	1980	1960	1950
1996	1940	1950	1998	1930	1919	1930	1950	1950	1920

(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?

1.133 The following table 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 $\overline{x}$ and s as a quick description of this distribution?

Player	Salary	Player	Salary	Player	Salary
Jeter	$12,600,000	Posada	$4,050,000	Spencer	$320,000
B. Williams	12,357,143	Stanton	2,450,000	T. Williams	320,000
Clemens	10,300,000	Hernandez	2,050,000	Almanzar	270,000
Mussina	10,000,000	Watson	1,700,000	Bellinger	230,000
Rivera	9,150,000	Mendoza	1,600,000	Einertson	206,000
Justice	7,000,000	Oliver	1,100,000	Choate	204,750
Pettitte	7,000,000	Rodriguez	850,000	Coleman	204,000
O'Neill	6,500,000	Soriano	630,000	Jimenez	200,000
Knoblauch	6,000,000	Sojo	500,000	Parker	200,000
Martinez	6,000,000	Boehringer	350,000	Seabol	200,000
Brosius	5,250,000				

1.134 At the time the salaries in the previous exercise were announced, one U.S. 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?

1.135 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?

1.136 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?

1.137 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 following table 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?

County	Population	County	Population	County	Population
Alameda	1,443,741	Marin	247,289	San Mateo	707,161
Alpine	1,208	Mariposa	17,130	Santa Barbara	399,347
Amador	35,100	Mendocino	86,265	Santa Clara	1,682,585
Butte	203,171	Merced	210,554	Santa Cruz	255,602
Calaveras	40,554	Modoc	9,449	Shasta	163,256
Colusa	18,804	Mono	12,853	Sierra	3,555
Contra Costa	948,816	Monterey	401,762	Siskiyou	44,301
Del Norte	27,507	Napa	124,279	Solano	394,542
El Dorado	156,299	Nevada	92,033	Sonoma	458,614
Fresno	799,407	Orange	2,846,289	Stanislaus	446,997
Glenn	26,453	Placer	248,399	Sutter	78,930
Humboldt	126,518	Plumas	20,824	Tehama	56,039
Imperial	142,361	Riverside	1,545,387	Trinity	13,022
Inyo	17,945	Sacramento	1,223,499	Tulare	368,021
Kern	661,645	San Benito	53,234	Tuolumne	54,501
Kings	129,461	San Bernardino	1,709,434	Ventura	753,197
Lake	58,309	San Diego	2,813,833	Yolo	168,660
Lassen	33,828	San Francisco	776,733	Yuba	60,219
Los Angeles	9,519,338	San Joaquin	563,598		
Madera	123,109	San Luis Obispo	246,681		

1.138 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?

1.139 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_{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?

1.140 The Florida State University Seminoles have been among the more successful teams in college football. The following table 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?

QB	235	QB	220	QB	215	QB	228	K	175	K	205
K	220	K	185	RB	200	RB	188	RB	190	RB	215
RB	190	RB	225	RB	225	RB	240	RB	237	R	205
R	185	R	185	R	190	R	195	R	201	R	190
R	195	R	180	OL	291	OL	280	OL	300	OL	320
OL	325	OL	285	OL	305	OL	305	OL	290	OL	310
OL	315	OL	285	OL	290	OL	325	OL	310	OL	256
OL	305	OL	300	DB	170	DB	207	DB	185	DB	175
DB	180	DB	190	DB	210	DB	200	DB	180	DB	195
DB	185	DB	170	DB	180	DB	190	DB	190	LB	220
LB	212	LB	233	LB	190	LB	195	LB	215	LB	220
LB	240	LB	230	LB	220	LB	225	LB	200	DL	260
DL	245	DL	264	DL	254	DL	215	DL	250	DL	295
DL	240	DL	275	DL	255	DL	285	DL	245	DL	270
DL	250	DL	250	TE	255	TE	245	TE	260	TE	245

1.141 Distance-learning courses. The 222 students enrolled in distance-learning courses offered by a college ranged from 18 to 64 years of age. The mode of their ages was 19. The median age was 31. (Julie Reinhart and Paul Schneider, "Student satisfaction, self-efficacy, and the perception of the two-way audio/visual distance learning environment," *Quarterly Review of Distance Education,* 2 (2001), pp. 357–365.) Explain how this can happen.

1.142 By-products from DDT. By-products from the pesticide DDT were major threats to the survival of birds of prey until use of DDT was banned at the end of 1972. Can time plots show the effect of the ban? Here are two sets of data for bald eagles nesting in the forests of northwestern Ontario. (James W. Grier, "Ban of DDT and subsequent recovery of reproduction of bald eagles," *Science,* 218 (1982), pp. 1232–1235.) The following data set gives the mean number of young per breeding area:

Year	1966	1967	1968	1969	1970	1971	1972	1973
Young	1.26	0.73	0.89	0.84	0.54	0.60	0.54	0.78
Year	1974	1975	1976	1977	1978	1979	1980	1981
Young	0.46	0.77	0.86	0.96	0.82	0.98	0.93	1.12

The following data are measurements of the chemical DDE (the by-product of DDT that most threatens birds of prey) from bald eagle eggs in the same area of Canada. These are in parts per million (ppm). There are often several measurements per year.

Year	1967	1967	1968	1971	1971	1972	1976	1976	1976	1976
DDE	44	95	121	125	95	87	13.3	16.4	50.4	59.8
Year	1976	1977	1977	1980	1980	1980	1981	1981	1981	
DDE	56.4	0.6	23.8	16.6	14.5	24.0	7.8	48.2	53.4	

Make time plots of eagle young and of mean DDE concentration in eggs. How does the effect of banning DDT appear in your plots?

1.143 Damage caused by tornadoes. The average damage caused by tornadoes in the states (see Exercise 1.44) and the estimated amount of oil recovered from different oil wells (Exercise 1.50) both have right-skewed distributions. Choose one of these data sets. Make a Normal quantile plot. How is the skewness of the distribution visible in the plot? Based on the plot, which observations (if any) would you call outliers?

1.144 Proportions older than 65. We know that the distribution of the percents of state residents over 65 years of age has a low outlier (Alaska) and a high outlier (Florida). The stemplot in Exercise 1.42 looks unimodal and roughly symmetric.
(a) Sketch what a Normal quantile plot would look like for a distribution that is Normal except for two outliers, one in each direction.
(b) If your software includes Normal quantile plots, make a plot of the percent-over-65 data and discuss what you see.

1.145 SAT Mathematics scores and grade point averages. The CSDATA data set contains information on 234 computer science students. We are interested in comparing the SAT Mathematics scores and grade point averages of female students with those of male students. Make two sets of side-by-side boxplots to carry out these comparisons. Write a brief discussion of the male-female comparisons. Then make Normal quantile plots of grade point averages and SAT Math scores separately for men and women. Which students are clear outliers? Which of the four distributions are approximately Normal if we ignore outliers?

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 low scores are better.)

Player	1	2	3	4	5	6	7	8	9	10	11	12
Round 1	89	90	87	95	86	81	102	105	83	88	91	79
Round 2	94	85	89	89	81	76	107	89	87	91	88	80

(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 which 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.)

Country	Alcohol from wine	Heart disease deaths	Country	Alcohol from wine	Heart disease deaths
Australia	2.5	211	Netherlands	1.8	167
Austria	3.9	167	New Zealand	1.9	266
Belgium	2.9	131	Norway	0.8	227
Canada	2.4	191	Spain	6.5	86
Denmark	2.9	220	Sweden	1.6	207
Finland	0.8	297	Switzerland	5.8	115
France	9.1	71	United Kingdom	1.3	285
Iceland	0.8	211	United States	1.2	199
Ireland	0.7	300	West Germany	2.7	172
Italy	7.9	107			

(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 following table 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.

State	Mean NAEP score	Percent proficient	Percent poverty	State	Mean NAEP score	Percent proficient	Percent poverty
Alabama	218	14	21.8	Missouri	229	24	14.4
Arizona	219	17	23.6	Montana	230	25	21.2
Arkansas	217	14	13.1	Nebraska	226	24	14.8
California	214	15	22.3	Nevada	220	16	12.8
Connecticut	234	32	13.4	New Mexico	214	12	23.5
Georgia	220	18	24.7	New York	227	22	28.9
Hawaii	216	14	14.5	North Carolina	232	28	21.3
Idaho	227	21	17.4	North Dakota	231	25	17.2
Illinois	225	22	12.2	Ohio	231	26	16.0
Indiana	234	31	12.6	Oklahoma	225	17	19.9
Iowa	233	28	14.2	Oregon	227	24	19.4
Kansas	232	30	13.3	Rhode Island	225	23	20.5
Kentucky	221	17	16.7	South Carolina	220	18	17.6
Louisiana	218	14	29.8	Tennessee	220	13	14.5
Maine	231	24	12.0	Texas	233	27	20.1
Maryland	222	22	8.1	Utah	227	24	11.8
Massachusetts	235	33	15.0	Vermont	232	30	12.2
Michigan	231	29	14.8	Virginia	230	25	7.9
Minnesota	235	34	12.6	West Virginia	225	18	25.7
Mississippi	211	9	19.3	Wyoming	229	25	13.0

(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.)

Flow rate	0.31	0.85	1.26	2.47	3.75
Eroded soil	0.82	1.95	2.18	3.01	6.07

(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.)

Person	1	2	3	4	5	6	7	8	9
Age	29	39	44	37	42	17	38	43	51
Incubation	13	46	43	34	20	20	18	72	19
Outcome	D	S	S	D	D	S	D	S	D
Person	10	11	12	13	14	15	16	17	18
Age	30	32	59	33	31	32	32	36	50
Incubation	36	48	44	21	32	86	48	28	16
Outcome	D	D	S	D	D	S	D	S	D

(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 following table 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.)

Board color	Insects trapped					
Lemon yellow	45	59	48	46	38	47
White	21	12	14	17	13	17
Green	37	32	15	25	39	41
Blue	16	11	20	21	14	7

(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?

2.8 In each of the following situations, is it more reasonable to simply explore the relationship between the two variables or to view one of the variables as an explanatory variable and the other as a response variable? In the latter case, which is the explanatory variable and which is the response variable?
(a) The amount of time spent studying for a statistics exam and the grade on the exam.
(b) The weight in kilograms and height in centimeters of a person.
(c) Inches of rain in the growing season and the yield of corn in bushels per acre.
(d) A student's scores on the SAT math exam and the SAT verbal exam.
(e) A family's income and the years of education their eldest child completes.

2.9 City and highway gas mileage. The following table gives the city and highway gas mileages for minicompact and two-seater cars. We expect a positive association between the city and highway mileages of a group of vehicles. We have already seen that the Honda Insight is a different type of car, so omit it as you work with these data.
(a) Make a scatterplot that shows the relationship between city and highway mileage, using city mileage as the explanatory variable. Use different plotting symbols for the two types of cars.
(b) Interpret the plot. Is there a positive association? Is the form of the plot roughly linear? Is the form of the relationship similar for the two types of cars? What is the most important difference between the two types?

Fuel economy (miles per gallon) for 2004 model vehicles

Two-Seater Cars			Minicompact Cars		
Model	City	Highway	Model	City	Highway
Acura NSX	17	24	Aston Martin Vanquish	12	19
Audi TT Roadster	20	28	Audi TT Coupe	21	29
BMW Z4 Roadster	20	28	BMW 325CI	19	27
Cadillac XLR	17	25	BMW 330CI	19	28
Chevrolet Corvette	18	25	BMW M3	16	23
Dodge Viper	12	20	Jaguar XK8	18	26
Ferrari 360 Modena	11	16	Jaguar XKR	16	23
Ferrari Maranello	10	16	Lexus SC 430	18	23
Ford Thunderbird	17	23	Mini Cooper	25	32
Honda Insight	60	66	Mitsubishi Eclipse	23	31
Lamborghini Gallardo	9	15	Mitsubishi Spyder	20	29
Lamborghini Murcielago	9	13	Porsche Cabriolet	18	26
Lotus Esprit	15	22	Porsche Turbo 911	14	22
Maserati Spyder	12	17			
Mazda Miata	22	28			
Mercedes-Benz SL500	16	23			
Mercedes-Benz SL600	13	19			
Nissan 350Z	20	26			
Porsche Boxster	20	29			
Porsche Carrera 911	15	23			
Toyota MR2	26	32			

2.10 Biological clocks. Many plants and animals have "biological clocks" that coordinate activities with the time of day. When researchers looked at the length of the biological cycles in the plant *Arabidopsis* by measuring leaf movements, they found that the length of the cycle is not always 24 hours. The researchers suspected that the plants adapt their clocks to their north-south position. Plants don't know geography, but they do respond to light, so the researchers looked at the relationship between the plants' cycle lengths and the length of the day on June 21st at their locations. The data file includes data on cycle length and day length, both in hours, for 146 plants. (We thank C. Robertson McClung of Dartmouth College for supplying the data. The study is reported in Todd P. Michael et al., "Enhanced fitness conferred by naturally occurring variation in the circadian clock," *Science,* 302 (2003), pp. 1049–1053.) Plot cycle length as the response variable against day length as the explanatory variable. Does there appear to be a positive association? Is it a strong association? Explain your answers.

2.11 Two problems with feet. Metatarsus adductus (call it MA) is a turning in of the front part of the foot that is common in adolescents and usually corrects itself. Hallux abducto valgus (call it HAV) is a deformation of the big toe that is not common in youth and often requires surgery. Perhaps the severity of MA can help predict the severity of HAV. The following table gives data on 38 consecutive patients who came to a medical center for HAV surgery. (Alan S. Banks et al., "Juvenile hallux abducto valgus association with metatarsus adductus," *Journal of*

the American Podiatric Medical Association, 84 (1994), pp. 219–224.) Using X-rays, doctors measured the angle of deformity for both MA and HAV. They speculated that there is a positive association—more serious MA is associated with more serious HAV.
(a) Make a scatterplot of the data in the following table. (Which is the explanatory variable?)
(b) Describe the form, direction, and strength of the relationship between MA angle and HAV angle. Are there any clear outliers in your graph?
(c) Do you think the data confirm the doctors' speculation?

Two measurements of foot deformities

HAV angle	MA angle	HAV angle	MA angle	HAV angle	MA angle
28	18	21	15	16	10
32	16	17	16	30	12
25	22	16	10	30	10
34	17	21	7	20	10
38	33	23	11	50	12
26	10	14	15	25	25
25	18	32	12	26	30
18	13	25	16	28	22
30	19	21	16	31	24
26	10	22	18	38	20
28	17	20	10	32	37
13	14	18	15	21	23
20	20	26	16		

2.12 Fuel consumption and speed. 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.)

Speed (km/h)	Fuel used (liters/100 km)	Speed (km/h)	Fuel used (liter/100 km)
10	21.00	90	7.57
20	13.00	100	8.27
30	10.00	110	9.03
40	8.00	120	9.87
50	7.00	130	10.79
60	5.90	140	11.77
70	6.30	150	12.83
80	6.95		

(a) Make a scatterplot. (Which variable should go on the x axis?)
(b) Describe the form of the relationship. In what way is it not linear? Explain why

the form of the relationship makes sense.
(c) It does not make sense to describe the variables as either positively associated or negatively associated. Why not?
(d) Is the relationship reasonably strong or quite weak? Explain your answer.

2.13 Worms and plant growth. To demonstrate the effect of nematodes (microscopic worms) on plant growth, a botanist introduces different numbers of nematodes into 16 planting pots. He then transplants a tomato seedling into each pot. Here are data on the increase in height of the seedlings (in centimeters) 14 days after planting: (Data provided by Matthew Moore.)

Nematodes	Seedling growth			
0	10.8	9.1	13.5	9.2
1,000	11.1	11.1	8.2	11.3
5,000	5.4	4.6	7.4	5.0
10,000	5.8	5.3	3.2	7.5

(a) Make a scatterplot of the response variable (growth) against the explanatory variable (nematode count). Then compute the mean growth for each group of seedlings, plot the means against the nematode counts, and connect these four points with line segments.
(b) Briefly describe the conclusions about the effects of nematodes on plant growth that these data suggest.

2.14 Mutual funds. Fidelity Investments, like other large mutual funds companies, offers many "sector funds" that concentrate their investments in narrow segments of the stock market. These funds often rise or fall by much more than the market as a whole. We can group them by broader market sector to compare returns. Here are percent total returns for 23 Fidelity "Select Portfolios" funds for the year 2003, a year in which stocks rose sharply: (Compiled from Fidelity data in the *Fidelity Insight* newsletter, 20 (2004), No. 1.)

Market sector	Fund returns (percent)						
Consumer	23.9	14.1	41.8	43.9	31.1		
Financial services	32.3	36.5	30.6	36.9	27.5		
Technology	26.1	62.7	68.1	71.9	57.0	35.0	59.4
Natural resources	22.9	7.6	32.1	28.7	29.5	19.1	

(a) Make a plot of total return against market sector (space the four market sectors equally on the horizontal axis). Compute the mean return for each sector, add the means to your plot, and connect the means with line segments.
(b) Based on the data, which of these market sectors were the best places to invest in 2003? Hindsight is wonderful.
(c) Does it make sense to speak of a positive or negative association between market sector and total return?

2.15 Mutual funds in another year. The data for 2003 in the previous exercise make sector funds look attractive. Stocks rose sharply in 2003, after falling sharply in 2002 (and also in 2001 and 2000). Let's look at the percent returns for both 2003 and 2002 for these same 23 funds. Here they are:

2002 return	2003 return	2002 return	2003 return	2002 return	2003 return
−17.1	23.9	−0.7	36.9	−37.8	59.4
−6.7	14.1	−5.6	27.5	−11.5	22.9
−21.1	41.8	−26.9	26.1	−0.7	36.9
−12.8	43.9	−42.0	62.7	64.3	32.1
−18.9	31.1	−47.8	68.1	−9.6	28.7
−7.7	32.3	−50.5	71.9	−11.7	29.5
−17.2	36.5	−49.5	57.0	−2.3	19.1
−11.4	30.6	−23.4	35.0		

Do a careful graphical analysis of these data: side-by-side comparison of the distributions of returns in 2002 and 2003 and also a description of the relationship between the returns of the same funds in these two years. What are your most important findings? (The outlier is Fidelity Gold Fund.)

Section 2.2

2.16 *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.)

Femur	38	56	59	64	74
Humerus	41	63	70	72	84

(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).

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

Femur	0.38	0.56	0.59	0.64	0.74
Humerus	410	630	700	720	840

(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.

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

Player	1	2	3	4	5	6	7	8	9	10	11
Round 1	89	90	87	95	86	81	105	83	88	91	79
Round 2	94	85	89	89	81	76	89	87	91	88	80

Make a scatterplot of the data. 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.

2.19 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?

2.20 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`.)

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

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?

2.21 You are going to use the *Correlation and Regression* applet to make different scatterplots with 10 points that have correlation close to 0.7. *Many patterns can*

have the same correlation. Always plot your data before you trust a correlation.
(a) Stop after adding the first two points. What is the value of the correlation? Why does it have this value no matter where the two points are located?
(b) Make a lower-left to upper-right pattern of 10 points with correlation about $r = 0.7$. (You can drag points up or down to adjust r after you have 10 points.) Make a rough sketch of your scatterplot.
(c) Make another scatterplot with nine points in a vertical stack at the left of the plot. Add one point far to the right and move it until the correlation is close to 0.7. Make a rough sketch of your scatterplot.
(d) Make yet another scatterplot with 10 points in a curved pattern that starts at the lower left, rises to the right, then falls again at the far right. Adjust the points up or down until you have a quite smooth curve with correlation close to 0.7. Make a rough sketch of this scatterplot also.

2.22 Go to the *Correlation and Regression* applet. Click on the scatterplot to create a group of 10 points in the lower-left corner of the scatterplot with a strong straight-line pattern (correlation about 0.9).
(a) Add one point at the upper right that is in line with the first 10. How does the correlation change?
(b) Drag this last point down until it is opposite the group of 10 points. How small can you make the correlation? Can you make the correlation negative? *A single outlier can greatly strengthen or weaken a correlation. Always plot your data to check for outlying points.*

2.23 Coffee prices and deforestation. Coffee is a leading export from several developing countries. When coffee prices are high, farmers often clear forest to plant more coffee trees. Here are data for five years on prices paid to coffee growers in Indonesia and the rate of deforestation in a national park that lies in a coffee-producing region: (Data from a plot in James A. Levine, Norman L. Eberhardt, and Michael D. Jensen, "Role of nonexercise activity thermogenesis in resistance to fat gain in humans," *Science,* 283 (1999), pp. 212–214.)

Price (cents per pound)	Deforestation (percent)
29	0.49
40	1.59
54	1.69
55	1.82
72	3.10

(a) Make a scatterplot. Which is the explanatory variable? What kind of pattern does your plot show?
(b) Find the correlation r step-by-step. That is, find the mean and standard deviation of the two variables. Then find the five standardized values for each variable and use the formula for r. Explain how your value for r matches your graph in (a).
(c) 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 (b).

2.24 Coffee prices in dollars or euros. Coffee is currently priced in dollars. If it were priced in euros, and the dollar prices in the previous exercise were translated into the equivalent prices in euros, would the correlation between coffee price and percent deforestation change? Explain your answer.

2.25 Mutual funds. Mutual fund reports often give correlations to describe how the prices of different investments are related. You look at the correlations between three Fidelity funds and the Standard & Poor's 500 stock index, which describes stocks of large U.S. companies. The three funds are Dividend Growth (stocks of large U.S. companies), Small Cap Stock (stocks of small U.S. companies), and Emerging Markets (stocks in developing countries). For 2003, the three correlations are $r = 0.35$, $r = 0.81$, and $r = 0.98$. (Compiled from Fidelity data in the *Fidelity Insight* newsletter, 20 (2004), No. 1.)
(a) Which correlation goes with each fund? Explain your answer.
(b) The correlations of the three funds with the index are all positive. Does this tell you that stocks went up in 2003? Explain your answer.

2.26 Mutual funds. Exercise 2.15 gives data on the returns from 23 Fidelity "sector funds" in 2002 (a down-year for stocks) and 2003 (an up-year).
(a) Make a scatterplot if you did not do so in the previous exercise. Fidelity Gold Fund, the only fund with a positive return in both years, is an extreme outlier.
(b) To demonstrate that correlation is not resistant, find r for all 23 funds and then find r for the 22 funds other than Gold. Explain from Gold's position in your plot why omitting this point makes r more negative.

2.27 Gas mileage and speed. The table below gives data on gas mileage against speed for a small car. Make a scatterplot and find the correlation r. Explain why r is close to zero despite a strong relationship between speed and gas used.

Speed (km/h)	Fuel used (liters/100 km)	Speed (km/h)	Fuel used (liter/100 km)
10	21.00	90	7.57
20	13.00	100	8.27
30	10.00	110	9.03
40	8.00	120	9.87
50	7.00	130	10.79
60	5.90	140	11.77
70	6.30	150	12.83
80	6.95		

2.28 Effect of a change in units. Consider again the correlation r between the speed of a car and its gas consumption from the data in the previous exercise.
(a) Transform the data so that speed is measured in miles per hour and fuel consumption in gallons per mile. (There are 1.609 kilometers in a mile and 3.785 liters in a gallon.) Make a scatterplot and find the correlation for both the original and the transformed data. How did the change of units affect your results?
(b) Now express fuel consumption in miles per gallon. (So each value is $1/x$ if x is gallons per mile.) Again make a scatterplot and find the correlation. How did this

change of units affect your results?
(*Lesson:* The effects of a linear transformation of the form $x_{\text{new}} = a + bx$ are simple. The effects of a nonlinear transformation are more complex.)

2.29 City and highway gas mileage. The table below gives the city and highway gas mileages for 21 two-seater cars, including the Honda Insight gas-electric hybrid car.

Fuel economy (miles per gallon) for 2004 model vehicles

Two-Seater Cars			Minicompact Cars		
Model	City	Highway	Model	City	Highway
Acura NSX	17	24	Aston Martin Vanquish	12	19
Audi TT Roadster	20	28	Audi TT Coupe	21	29
BMW Z4 Roadster	20	28	BMW 325CI	19	27
Cadillac XLR	17	25	BMW 330CI	19	28
Chevrolet Corvette	18	25	BMW M3	16	23
Dodge Viper	12	20	Jaguar XK8	18	26
Ferrari 360 Modena	11	16	Jaguar XKR	16	23
Ferrari Maranello	10	16	Lexus SC 430	18	23
Ford Thunderbird	17	23	Mini Cooper	25	32
Honda Insight	60	66	Mitsubishi Eclipse	23	31
Lamborghini Gallardo	9	15	Mitsubishi Spyder	20	29
Lamborghini Murcielago	9	13	Porsche Cabriolet	18	26
Lotus Esprit	15	22	Porsche Turbo 911	14	22
Maserati Spyder	12	17			
Mazda Miata	22	28			
Mercedes-Benz SL500	16	23			
Mercedes-Benz SL600	13	19			
Nissan 350Z	20	26			
Porsche Boxster	20	29			
Porsche Carrera 911	15	23			
Toyota MR2	26	32			

(a) Make a scatterplot of highway mileage y against city mileage x for all 21 cars. There is a strong positive linear association. The Insight lies far from the other points. Does the Insight extend the linear pattern of the other cars, or is it far from the line they form?
(b) Find the correlation between city and highway mileages both without and with the Insight. Based on your answer to (a), explain why r changes in this direction when you add the Insight.

Section 2.3

2.30 (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?

2.31 (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.

2.32 (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.)

2.33 (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?

2.34 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.

2.35 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

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

fit the data well. (William M. Lewis and Michael C. Grant, "Acid precipitation in the western United States," *Science*, 207 (1980), pp. 176–177.)
(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.

2.36 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.

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

(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 powerboat 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 previous exercise:

1991	716	53	1993	716	35
1992	716	38	1994	735	49

Add these points to your scatterplot. Florida took stronger measures to protect manatees during these years. Do you see any evidence that these measures succeeded?
(f) In part (c) you predicted manatee deaths in a year with 716,000 powerboat 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?

2.37 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:

Year	1935	1940	1945	1950	1955	1960	1965	1970	1975	1980
Population	32.1	30.5	24.4	23.0	19.1	15.6	12.4	9.7	8.9	7.2

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

2.38 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.)

Nitrates	50	50	100	200	400	800	1200	1600	2000	2000
Absorbance	7.0	7.5	12.8	24.0	47.0	93.0	138.0	183.0	230.0	226.0

(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?

2.39 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 $\overline{x}$ for the existing points and y greater than the mean $\overline{y}$ 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 diets 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.*

The following three exercises 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.)

2.40 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.

2.41 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.

2.42 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.

2.43 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:

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

(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).

2.44 First test and final exam. Here are data for eight students from anelementary statistics course:

First test score	153	144	162	149	127	118	158	153
Final exam score	145	140	145	170	145	175	170	160

(a) Plot the data with the first test scores on the x axis and the final exam scores on the y axis.
(b) Find the least-squares regression line for predicting the final exam score using the first test score.
(c) Graph the least-squares regression line on your plot.

2.45 Second test and final exam. Refer to the previous exercise. Here are the data for the second test and the final exam for the same students:

Second test score	158	162	144	162	136	158	175	153
Final exam score	145	140	145	170	145	175	170	160

(a) Plot the data with the second test scores on the x axis and the final exam scores on the y axis.
(b) Find the least-squares regression line for predicting the final exam score using the second test score.
(c) Graph the least-squares regression line on your plot.

2.46 The effect of an outlier. Refer to the previous two exercises. Add a ninth student whose scores on the second test and final exam would lead you to classify the additional data point as an outlier. Recalculate the least-squares regression line with this additional case and summarize the effect it has on the least-squares regression line.

2.47 The effect of a different point. Examine the data in Exercise 2.45 and add a diferent ninth student who has low scores on the second test and the final exam, and fits the overall pattern of the other scores in the data set. Recalculate the least-squares regression line with this additional case and summarize the effect it has on the least-squares regression line.

2.48 Problems with feet. Metatarsus adductus (call it MA) is a turning in of the front part of the foot that is common in adolescents and usually corrects itself. Hallux abducto valgus (call it HAV) is a deformation of the big toe that is not common in youth and often requires surgery. Perhaps the severity of MA can help predict the severity of HAV. The following table gives data on 38 consecutive patients who came to a medical center for HAV surgery. (Alan S. Banks et al., "Juvenile hallux abducto valgus association with metatarsus adductus," *Journal of the American Podiatric Medical Association,* 84 (1994), pp. 219–224.) Using

X-rays, doctors measured the angle of deformity for both MA and HAV. They speculated that there is a positive association—more serious MA is associated with more serious HAV. A scatterplot the severity of the mild foot deformity called MA can help predict the severity of the more serious deformity called HAV.

Two measurements of foot deformities

HAV angle	MA angle	HAV angle	MA angle	HAV angle	MA angle
28	18	21	15	16	10
32	16	17	16	30	12
25	22	16	10	30	10
34	17	21	7	20	10
38	33	23	11	50	12
26	10	14	15	25	25
25	18	32	12	26	30
18	13	25	16	28	22
30	19	21	16	31	24
26	10	22	18	38	20
28	17	20	10	32	37
13	14	18	15	21	23
20	20	26	16		

(a) Find the equation of the least-squares regression line for predicting HAV angle from MA angle. Make a scatterplot of the data with the lease-squares regression line.
(b) A new patient has MA angle 25 degrees. What do you predict this patient's HAV angle to be?
(c) Does knowing MA angle allow doctors to predict HAV angle accurately? Explain your answer from the scatterplot, then calculate a numerical measure to support your finding.

2.49 Predict final exam scores. In Professor Friedman's economics course the correlation between the students' total scores before the final examination and their final examination scores is $r = 0.55$. The pre-exam totals for all students in the course have mean 270 and standard deviation 30. The final exam scores have mean 70 and standard deviation 9. Professor Friedman has lost Julie's final exam but knows that her total before the exam was 310. He decides to predict her final exam score from her pre-exam total.
(a) What is the slope of the least-squares regression line of final exam scores on pre-exam total scores in this course? What is the intercept?
(b) Use the regression line to predict Julie's final exam score.
(c) Julie doesn't think this method accurately predicts how well she did on the final exam. Calculate r^2 and use the value you get to argue that her actual score could have been much higher or much lower than the predicted value.

2.50 Pesticide decay. Fenthion is a pesticide used to control the olive fruit fly. There are government limits on the amount of pesticide residue that can be present in olive products. Because the pesticide decays over time, producers of olive oil might simply store the oil until the fenthion has decayed. The simple exponential decay model says that the concentration C of pesticide remaining after time t is

$$C = C_0 e^{-kt}$$

where C_0 is the initial concentration and k is a constant that determines the rate of decay. This model is a straight line if we take the logarithm of the concentration:

$$\log C = \log C_0 - kt$$

(The logarithm here is the natural logarithm, not the common logarithm with base 10.) Here are data on the concentration (milligrams of fenthion per kilogram of oil) in specimens of Greek olive oil: (Data from plots in Chaido Lentza-Rizos, Elizabeth J. Avramides, and Rosemary A. Roberts, "Persistence of fenthion residues in olive oil," *Pesticide Science,* 40 (1994), pp. 63–69.)

Days stored	Concentration				
28	0.99	0.99	0.96	0.95	0.93
84	0.96	0.94	0.91	0.91	0.90
183	0.89	0.87	0.86	0.85	0.85
273	0.87	0.86	0.84	0.83	0.83
365	0.83	0.82	0.80	0.80	0.79

(a) Plot the natural logarithm of concentration against days stored. Notice that there are several pairs of identical data points. Does the pattern suggest that the model of simple exponential decay describes the data reasonably well, at least over this interval of time? Explain your answer.
(b) Regress the logarithm of concentration on time. Use your result to estimate the value of the constant k.

2.51 The decay product is toxic. Unfortunately, the main product of the decay of the pesticide fenthion is fenthion sulfoxide, which is also toxic. Here are data on the total concentration of fenthion and fenthion sulfoxide in the same specimens of olive oil described in the previous exercise:

Days stored	Concentration				
28	1.03	1.03	1.01	0.99	0.99
84	1.05	1.04	1.00	0.99	0.99
183	1.03	1.02	1.01	0.98	0.98
273	1.07	1.06	1.03	1.03	1.02
365	1.06	1.02	1.01	1.01	0.99

(a) Plot concentration against days stored. Your software may fill the available space in the plot, which in this case hides the pattern. Try a plot with vertical scale from 0.8 to 1.2. Be sure your plot takes note of the pairs of identical data points.
(b) What is the slope of the least-squares line for predicting concentration of fenthion and fenthion sulfoxide from days stored? Explain why this value agrees with the graph.
(c) What do the data say about the idea of reducing fenthion in olive oil by storing the oil before selling it?

Section 2.4

2.52 The following table 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.

Beef hot dogs		Meat hot dogs		Poultry hot dogs	
Calories	Sodium	Calories	Sodium	Calories	Sodium
186	495	173	458	129	430
181	477	191	506	132	375
176	425	182	473	102	396
149	322	190	545	106	383
184	482	172	496	?94	387
190	587	147	360	102	542
158	370	146	387	?87	359
139	322	139	386	?99	357
175	479	175	507	170	528
148	375	136	393	113	513
152	330	179	405	135	426
111	300	153	372	142	513
141	386	107	144	?86	358
153	401	195	511	143	581
190	645	135	405	152	588
157	440	140	428	146	522
131	317	138	339	144	545
149	319				
135	298				
132	253				

(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?

2.53 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.)

Spheres	4.0	4.7	6.3	8.2	12.0	15.9	17.4	18.1	20.2	23.9
Vein	3.3	8.3	4.5	9.3	10.7	16.4	15.4	17.6	21.0	21.7

(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?

2.54 The following table 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 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.

State	Mean NAEP score	Percent proficient	Percent poverty	State	Mean NAEP score	Percent proficient	Percent poverty
Alabama	218	14	21.8	Missouri	229	24	14.4
Arizona	219	17	23.6	Montana	230	25	21.2
Arkansas	217	14	13.1	Nebraska	226	24	14.8
California	214	15	22.3	Nevada	220	16	12.8
Connecticut	234	32	13.4	New Mexico	214	12	23.5
Georgia	220	18	24.7	New York	227	22	28.9
Hawaii	216	14	14.5	North Carolina	232	28	21.3
Idaho	227	21	17.4	North Dakota	231	25	17.2
Illinois	225	22	12.2	Ohio	231	26	16.0
Indiana	234	31	12.6	Oklahoma	225	17	19.9
Iowa	233	28	14.2	Oregon	227	24	19.4
Kansas	232	30	13.3	Rhode Island	225	23	20.5
Kentucky	221	17	16.7	South Carolina	220	18	17.6
Louisiana	218	14	29.8	Tennessee	220	13	14.5
Maine	231	24	12.0	Texas	233	27	20.1
Maryland	222	22	8.1	Utah	227	24	11.8
Massachusetts	235	33	15.0	Vermont	232	30	12.2
Michigan	231	29	14.8	Virginia	230	25	7.9
Minnesota	235	34	12.6	West Virginia	225	18	25.7
Mississippi	211	9	19.3	Wyoming	229	25	13.0

(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?

2.55 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.

State	Poverty percent	M.D.'s per 100,000	State	Poverty percent	M.D.'s per 100,000
Alabama	15.1	198	Montana	15.9	190
Alaska	8.6	167	Nebraska	11.0	218
Arizona	15.2	202	Nevada	11.0	173
Arkansas	16.4	190	New Hampshire	8.9	237
California	15.3	247	New Jersey	8.5	295
Colorado	8.6	238	New Mexico	20.8	212
Connecticut	8.4	354	New York	15.7	387
Delaware	10.1	234	North Carolina	13.0	232
D.C.	19.7	737	North Dakota	13.9	222
Florida	13.3	238	Ohio	11.4	235
Georgia	13.7	211	Oklahoma	13.5	169
Hawaii	11.9	265	Oregon	13.1	225
Idaho	13.9	154	Pennsylvania	10.6	291
Illinois	10.4	260	Rhode Island	11.4	338
Indiana	8.3	195	South Carolina	12.8	207
Iowa	8.7	173	South Dakota	11.7	184
Kansas	10.5	203	Tennessee	13.2	246
Kentucky	13.8	209	Texas	15.6	203
Louisiana	18.2	246	Utah	7.9	200
Maine	10.4	223	Vermont	9.6	305
Maryland	7.6	374	Virginia	9.8	241
Massachusetts	10.9	412	Washington	9.2	235
Michigan	10.3	224	West Virginia	16.7	215
Minnesota	9.1	249	Wisconsin	8.5	227
Mississippi	16.8	163	Wyoming	11.9	171
Missouri	11.1	230			

(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?

2.56 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?

2.57 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.)

Day	1	2	3	4	5	6	7	8	9	10	11	12
Rural	NA	67	42	33	46	NA	43	54	NA	NA	NA	NA
City	39	68	42	34	48	82	45	NA	NA	60	57	NA
Day	13	14	15	16	17	18	19	20	21	22	23	24
Rural	38	88	108	57	70	42	43	39	NA	52	48	56
City	39	NA	123	59	71	41	42	38	NA	57	50	58
Day	25	26	27	28	29	30	31	32	33	34	35	36
Rural	44	51	21	74	48	84	51	43	45	41	47	35
City	45	69	23	72	49	86	51	42	46	NA	44	42

(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 14th 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?

2.58 Go to the *Correlation and Regression* applet. Click on the scatterplot to create a group of 10 points in the lower-left corner of the scatterplot with a strong straight-line pattern (correlation about 0.9). Now click the "Show least-squares line" box to display the regression line.
(a) Add one point at the upper right that is far from the other 10 points but exactly on the regression line. Why does this outlier have no effect on the line even though it changes the correlation?
(b) Now drag this last point down until it is opposite the group of 10 points. You see that one end of the least-squares line chases this single point, while the other

end remains near the middle of the original group of 10. What makes the last point so influential?

2.59 A multimedia statistics learning system includes a test of skill in using the computer's mouse. The software displays a circle at a random location on the computer screen. The subject tries to click in the circle with the mouse as quickly as possible. A new circle appears as soon as the subject clicks the old one. The following table gives data for one subject's trials, 20 with each hand. Distance is the distance from the cursor location to the center of the new circle, in units whose actual size depends on the size of the screen. Time is the time required to click in the new circle, in milliseconds.
(a) We suspect that time depends on distance. Make a scatterplot of time against distance, using separate symbols for each hand.
(b) Describe the pattern. How can you tell that the subject is right-handed?
(c) Find the regression line of time on distance separately for each hand. Draw these lines on your plot. Which regression does a better job of predicting time from distance? Give numerical measures that describe the success of the two regressions.
(d) It is possible that the subject got better in later trials due to learning. It is also possible that he got worse due to fatigue. Plot the residuals from each regression against the time order of the trials (down the columns in the following table). Is either of these systematic effects of time visible in the data?

Time	Distance	Hand	Time	Distance	Hand
115	190.70	right	240	190.70	left
96	138.52	right	190	138.52	left
110	165.08	right	170	165.08	left
100	126.19	right	125	126.19	left
111	163.19	right	315	163.19	left
101	305.66	right	240	305.66	left
111	176.15	right	141	176.15	left
106	162.78	right	210	162.78	left
96	147.87	right	200	147.87	left
96	271.46	right	401	271.46	left
95	40.25	right	320	40.25	left
96	24.76	right	113	24.76	left
96	104.80	right	176	104.80	left
106	136.80	right	211	136.80	left
100	308.60	right	238	308.60	left
113	279.80	right	316	279.80	left
123	125.51	right	176	125.51	left
111	329.80	right	173	329.80	left
95	51.66	right	210	51.66	left
108	201.95	right	170	201.95	left

2.60 Fuel consumption and speed. The table following this exercise gives data on the fuel consumption y of a car at various speeds x. The relationship is strongly curved: Fuel used decreases with increasing speed at low speeds, then increases again as higher speeds are reached. The equation of the least-squares regression line for these data is

$$\hat{y} = 11.058 - 0.01466x$$

The residuals, in the same order as the observations, are

10.09	2.24	−0.62	−2.47	−3.33	−4.28	−3.73	−2.94
−2.17	−1.32	−0.42	0.57	1.64	2.76	3.97	

(a) Make a scatterplot of the observations and draw the regression line on your plot. The line is a poor description of the curved relationship.
(b) Check that the residuals have sum zero (up to roundoff error).
(c) Make a plot of the residuals against the values of x. Draw a horizontal line at height zero on your plot. The residuals show the same pattern about this line as the data points show about the regression line in the scatterplot in (a).

Speed (km/h)	Fuel used (liters/100 km)	Speed (km/h)	Fuel used (liter/100 km)
10	21.00	90	7.57
20	13.00	100	8.27
30	10.00	110	9.03
40	8.00	120	9.87
50	7.00	130	10.79
60	5.90	140	11.77
70	6.30	150	12.83
80	6.95		

2.61 Pesticide in olive oil. The table following this exercise gives data on the concentration of the pesticide fenthion in Greek olive oil that has been stored for various lengths of time. The exponential decay model used to describe how concentration decreases over time proposes a curved relationship between storage time and concentration. Do the residuals from fitting a regression line show a curved pattern? The least-squares line for predicting concentration is

$$\hat{y} = 0.965 - 0.00045x$$

(a) The first batch of olive oil was stored for 28 days and had fenthion concentration 0.99 mg/kg. What is the predicted concentration for this batch? What is the residual?
(b) The residuals, arranged as in the data table, are:

Days stored	Residual				
28	0.0378	0.0378	0.0078	−0.0022	−0.0222
84	0.0329	0.0129	−0.0171	−0.0171	−0.0271
183	0.0072	−0.0128	−0.0228	−0.0328	−0.0328
273	0.0275	0.0175	−0.0025	−0.0125	−0.0125
365	0.0286	0.0186	−0.0014	−0.0014	−0.0114

Check that your residual from (a) agrees (up to roundoff error) with the value 0.0378 given here. Verify that the residuals sum to zero (again up to roundoff error).
(c) Make a residual plot. Is a curved pattern visible? Is the curve very strong? (Software often makes the pattern hard to see because it fills the entire plot area. Try a plot with vertical scale from −0.1 to 0.1.)

Days stored	Concentration				
28	0.99	0.99	0.96	0.95	0.93
84	0.96	0.94	0.91	0.91	0.90
183	0.89	0.87	0.86	0.85	0.85
273	0.87	0.86	0.84	0.83	0.83
365	0.83	0.82	0.80	0.80	0.79

2.62 City and highway gas mileage. The following table gives the city and highway gas mileages for 21 two-seater cars, including the Honda Insight gas-electric hybrid car.

Fuel economy (miles per gallon) for 2004 model vehicles

Two-Seater Cars			Minicompact Cars		
Model	City	Highway	Model	City	Highway
Acura NSX	17	24	Aston Martin Vanquish	12	19
Audi TT Roadster	20	28	Audi TT Coupe	21	29
BMW Z4 Roadster	20	28	BMW 325CI	19	27
Cadillac XLR	17	25	BMW 330CI	19	28
Chevrolet Corvette	18	25	BMW M3	16	23
Dodge Viper	12	20	Jaguar XK8	18	26
Ferrari 360 Modena	11	16	Jaguar XKR	16	23
Ferrari Maranello	10	16	Lexus SC 430	18	23
Ford Thunderbird	17	23	Mini Cooper	25	32
Honda Insight	60	66	Mitsubishi Eclipse	23	31
Lamborghini Gallardo	9	15	Mitsubishi Spyder	20	29
Lamborghini Murcielago	9	13	Porsche Cabriolet	18	26
Lotus Esprit	15	22	Porsche Turbo 911	14	22
Maserati Spyder	12	17			
Mazda Miata	22	28			
Mercedes-Benz SL500	16	23			
Mercedes-Benz SL600	13	19			
Nissan 350Z	20	26			
Porsche Boxster	20	29			
Porsche Carrera 911	15	23			
Toyota MR2	26	32			

(a) Make a scatterplot of highway mileage (response) against city mileage (explanatory) for all 21 cars.
(b) Use software or a graphing calculator to find the regression line for predicting highway mileage from city mileage and also the 21 residuals for this regression. Make a residual plot with a horizontal line at zero. (The "stacks" in the plot are due to the fact that mileage is measured only to the nearest mile per gallon.)
(c) Which car has the largest positive residual? The largest negative residual?
(d) The Honda Insight, an extreme outlier, does not have the largest residual in either direction. Why is this not surprising?

2.63 City and highway gas mileage. Continue your work in the previous exercise. Find the regression line for predicting highway mileage from city mileage for the 20 two-seater cars other than the Honda Insight. Draw both regression lines on your scatterplot. Is the Insight very influential for the least-squares line? (Look at the position of the lines for city mileages between 10 and 30 mpg, values that cover most cars.) What explains your result?

2.64 Stride rate of runners. Runners are concerned about their form when racing. One measure of form is the stride rate, the number of steps taken per second. As running speed increases, the stride rate should also increase. In a study of 21 of the best American female runners, researchers measured the stride rate for different speeds. The following table gives the speeds (in feet per second) and the mean stride rates for these runners: (R. C. Nelson, C. M. Brooks, and N. L. Pike, "Biomechanical comparison of male and female distance runners," in P. Milvy (ed.), *The Marathon: Physiological, Medical, Epidemiological, and Psychological Studies,* New York Academy of Sciences, 1977, pp. 793–807.)

Speed	15.86	16.88	17.50	18.62	19.97	21.06	22.11
Stride rate	3.05	3.12	3.17	3.25	3.36	3.46	3.55

(a) Plot the data with speed on the x axis and stride rate on the y axis. Does a straight line adequately describe these data?
(b) Find the equation of the regression line of stride rate on speed. Draw this line on your plot.
(c) For each of the speeds given, obtain the predicted value of the stride rate and the residual. Verify that the residuals add to zero.
(d) Plot the residuals against speed. Describe the pattern. What does the plot indicate about the adequacy of the linear fit? Are there any potentially influential observations?

2.65 Stride rate and running speed. The previous exercise gives data on the mean stride rate of a group of 21 elite female runners at various running speeds. Find the correlation between speed and stride rate. Would you expect this correlation to increase or decrease if we had data on the individual stride rates of all 21 runners at each speed? Why?

Section 2.5

For exercises on data analysis for two-way tables, use the exercises for Chapter 9 and ignore the parts that ask questions about statistical inference.

Section 2.6

2.66 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.

2.67 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?

2.68 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.

2.69 CEO compensation and layoffs. "Based on an examination of twenty-two companies that announced large layoffs during 1994, Downs found a strong (.31) correlation between the size of the layoffs and the compensation of the CEOs." (Kevin Phillips, *Wealth and Democracy,* Broadway Books, 2002, p. 151.) This correlation is probably explained by common response to a lurking variable, the size of the company as measured by its number of employees. Explain how common response could create the observed correlation. Use a diagram to illustrate your explanation.

2.70 Health and income. An article entitled "The Health and Wealth of Nations" says: "The positive correlation between health and income per capita is one of the best-known relations in international development. This correlation is commonly thought to reflect a causal link running from income to health. . . . Recently, however, another intriguing possibility has emerged: that the health-income correlation is partly explained by a causal link running the other way—from health to income." (David E. Bloom and David Canning, "The health and wealth of nations," *Science,* 287 (2000), pp. 1207–1208.)

2.71 Self-esteem and work performance. People who do well tend to feel good about themselves. Perhaps helping people feel good about themselves will help them do better in their jobs and in life. Raising self-esteem became for a time a goal in many schools and companies. Can you think of explanations for the association between high self-esteem and good performance other than "self-esteem causes better work"?

Chapter 2 Review Exercises

The following three exercises 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.)

Year	Overseas % return	U.S. % return	Year	Overseas % return	U.S. % return
1971	29.6	14.6	1986	69.4	18.6
1972	36.3	18.9	1987	24.6	5.1
1973	−14.9	−14.8	1988	28.5	16.8
1974	−23.2	−26.4	1989	10.6	31.5
1975	35.4	37.2	1990	−23.0	−3.1
1976	2.5	23.6	1991	12.8	30.4
1977	18.1	−7.4	1992	−12.1	7.6
1978	32.6	6.4	1993	32.9	10.1
1979	4.8	18.2	1994	6.2	1.3
1980	22.6	32.3	1995	11.2	37.6
1981	−2.3	−5.0	1996	6.4	23.0
1982	−1.9	21.5	1997	2.1	33.4
1983	23.7	22.4	1998	20.3	28.6
1984	7.4	6.1	1999	27.2	21.0
1985	56.2	31.6	2000	−14.0	−9.1

2.72 (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?

2.73 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?

2.74 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). The following three exercises are based on these data.

State	Region	Pay	Spend	State	Region	Pay	Spend
Me.	NE	32.0	6.41	N.H.	NE	29.0	6.13
Vt.	NE	35.4	7.37	Mass.	NE	42.2	6.17
R.I.	NE	40.7	7.36	Conn.	NE	50.0	8.50
N.Y.	MA	47.6	9.45	N.J.	MA	46.1	9.86
Pa.	MA	44.5	7.20	Ohio	ENC	36.8	5.62
Ind.	ENC	36.8	6.00	Ill.	ENC	39.4	5.26
Mich.	ENC	47.4	6.93	Wis.	ENC	37.7	7.00
Minn.	WNC	35.9	5.11	Iowa	WNC	31.5	5.56
Mo.	WNC	31.2	4.97	N.Dak.	WNC	26.3	4.60
S.Dak.	WNC	26.0	4.84	Nebr.	WNC	30.9	5.38
Kans.	WNC	34.7	5.76	Del.	SA	39.1	7.17
Md.	SA	40.7	6.72	D.C.	SA	43.7	8.21
Va.	SA	34.0	5.66	W.Va.	SA	31.9	6.52
N.C.	SA	30.8	4.95	S.C.	SA	30.3	4.93
Ga.	SA	32.6	5.40	Fla.	SA	32.6	5.72
Ky.	ESC	32.3	5.61	Tenn.	ESC	32.5	4.54
Ala.	ESC	31.1	4.46	Miss.	ESC	26.8	4.12
Ark.	WSC	28.9	4.26	La.	WSC	26.5	4.71
Okla.	WSC	28.2	4.38	Tex.	WSC	31.2	5.42
Mont.	MN	28.8	5.83	Idaho	MN	29.8	6.03
Wyo.	MN	31.3	6.07	Colo.	MN	34.6	5.50
N.Mex.	MN	28.5	5.42	Ariz.	MN	32.2	4.25
Utah	MN	29.1	3.67	Nev.	MN	34.8	5.13
Wash.	PA	36.2	5.81	Oreg.	PA	38.6	6.25
Calif.	PA	41.1	4.73	Alaska	PA	48.0	9.93
Hawaii	PA	38.5	6.16				

2.75 Make a stemplot or histogram for teachers' pay. Is the distribution roughly symmetric or clearly skewed? Find the five-number summary. Are there any suspected 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?

2.76 (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?

2.77 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?

2.78 Lamb's-quarter is a common weed that interferes with the growth of corn. An agriculture researcher planted corn at the same rate in 16 small plots of ground, then weeded the plots by hand to allow a fixed number of lamb's-quarter plants to grow in each meter of corn row. No other weeds were allowed to grow. Here are the yields of corn (bushels per acre) in each of the plots:

Weeds per meter	Corn yield	Weeds per meter	Corn yield	Weeds per meter	Corn yield	Weeds per meter	Corn yield
0	166.7	1	166.2	3	158.6	9	162.8
0	172.2	1	157.3	3	176.4	9	142.4
0	165.0	1	166.7	3	153.1	9	162.8
0	176.9	1	161.1	3	156.0	9	162.4

(a) What are the explanatory and response variables in this experiment?
(b) Make side-by-side stemplots of the yields, after rounding to the nearest bushel. What do you conclude about the effect of this weed on corn yield?
(c) Make a scatterplot of corn yield against weeds per meter. Find the least-squares regression line and add it to your plot. The advantage of regression over the side-by-side comparison in (b) is that we can use the fitted model to draw conclusions for counts of weeds other than the ones the researcher actually used. What does the slope of the fitted line tell us about the effect of lamb's-quarter on corn yield?
(d) Predict the yield for corn grown under these conditions with six lamb's-quarter plants per meter of row.

2.79 Stock prices and earnings. In the long run, the price of a company's stock ought to parallel changes in the company's earnings. The following table gives data on the annual growth rates in earnings and in stock prices (both in percent) for major industry groups as set by Standard & Poor's. (H. Bradley Perry, "Analyzing growth stocks: what's a good growth rate?" *AAII Journal,* 13 (1991), pp. 7–10.)
(a) Make a graph showing how earnings growth explains growth in stock price. Does it appear to be true that (on the average in the long run) stock price growth parallels earnings growth?
(b) What percent of the variation in stock price growth among industry groups can be explained by the linear relationship with earnings growth?
(c) If stock prices exactly followed earnings, the slope of the least-squares line for predicting price growth from earnings growth would be 1. Explain why. What is the slope of the least-squares line for these data?
(d) What is the correlation between earnings growth and price growth? If we had data on all of the individual companies in these 20 industries, would the correlation be higher or lower? Why?

Percent growth in stock price and earnings for industry groups

Industry	Earnings growth (%)	Price growth (%)
Auto	3.3	2.9
Banks	8.6	6.5
Chemicals	6.6	3.1
Computers	10.2	5.3
Drugs	11.3	10.0
Electrical equipment	8.5	8.2
Food	7.6	6.5
Household products	9.7	10.1
Machinery	5.1	4.7
Oil: domestic	7.4	7.3
Oil: international	7.7	7.7
Oil equipment/services	10.1	10.8
Railroad	6.6	6.6
Retail: department stores	10.1	9.5
Retail: food	6.9	6.9
Soft drinks	12.7	12.0
Steel	−1.0	−1.6
Tobacco	12.3	1.7
Utilities: electric	2.8	1.4
Utilities: gas	5.2	6.2

2.80 Running speed and stride rate. The following table gives data on the relationship between running speed (feet per second) and stride rate (steps taken per second) for elite female runners:

Speed	15.86	16.88	17.50	18.62	19.97	21.06	22.11
Stride rate	3.05	3.12	3.17	3.25	3.36	3.46	3.55

Here are the corresponding data from the same source for male runners:

Speed	15.86	16.88	17.50	18.62	19.97	21.06	22.11
Stride rate	2.92	2.98	3.03	3.11	3.22	3.31	3.41

(a) Plot the data for both groups on one graph using different symbols to distinguish between the points for females and those for males.
(b) Suppose now that the data came to you without identification as to gender. Compute the least-squares line from all of the data and plot it on your graph.
(c) Compute the residuals from this line for each observation. Make a plot of the residuals against speed. How does the fact that the data come from two distinct groups show up in the residual plot?

2.81 Wood flakes as a building material. Wood scientists are interested in replacing solid wood building material with less expensive products made from wood flakes. They collected the following data to examine the relationship between the length (in inches) and the strength (in pounds per square inch) of beams made from wood flakes: (Data provided by Jim Bateman and Michael Hunt, Purdue University.)

Length	5	6	7	8	9	10	11	12	13	14
Strength	446	371	334	296	249	254	244	246	239	234

(a) Make a scatterplot that shows how the length of a beam affects its strength.
(b) Describe the overall pattern of the plot. Are there any outliers?
(c) Fit a least-squares line to the entire set of data. Graph the line on your scatterplot. Does a straight line adequately describe these data?
(d) The scatterplot suggests that the relation between length and strength can be described by *two* straight lines, one for lengths less than 9 inches and another for lengths 9 inches or greater. Fit least-squares lines to these two subsets of the data, and draw the lines on your plot. Do they describe the data adequately? What question would you now ask the wood experts?

2.82 Global investing. One reason to invest abroad is that markets in different countries don't move in step. When American stocks go down, foreign stocks may go up. So an investor who holds both bears less risk. That's the theory. Now we read: "The correlation between changes in American and European share prices has risen from 0.4 in the mid-1990s to 0.8 in 2000." ("Dancing in step," *Economist,* March 22, 2001.) Explain to an investor who knows no statistics why this fact reduces the protection provided by buying European stocks.

2.83 Stock prices in Europe and the United States. The same article that claims that the correlation between changes in stock prices in Europe and the United States was 0.8 in 2000 goes on to say: "Crudely, that means that movements on Wall Street can explain 80% of price movements in Europe." Is this true? What is the correct percent explained if $r = 0.8$?

2.84 SAT scores and grade point averages. Can we predict college grade point average from SAT scores and high school grades? The CSDATA data set contains information on this issue for a large group of computer science students. We will look only at SAT Mathematics scores as a predictor of later college GPA, using the variables SATM and GPA from CSDATA. Make a scatterplot, obtain r and r^2, and draw on your plot the least-squares regression line for predicting GPA from SATM. Then write a brief discussion of the ability of SATM alone to predict GPA. (In Chapter 11 we will see how combining several explanatory variables improves our ability to predict.)

2.85 Sexual imagery in magazine ads. In what ways do advertisers in magazines use sexual imagery to appeal to youth? One study classified each of 1509 full-page or larger ads as "not sexual" or "sexual," according to the amount and style of the clothing of the male or female model in the ad. The ads were also classified according to the target readership of the magazine. (Tom Reichert, "The prevalence of sexual imagery in ads targeted to young adults," *Journal of Consumer Affairs,* 37 (2003), pp. 403–412.) Here is the two-way table of counts:

	Magazine readership			
Model clothing	Women	Men	General interest	Total
Not sexual	351	514	248	1113
Sexual	225	105	66	396
Total	576	619	314	1509

(a) Summarize the data numerically and graphically.
(b) All of the ads were taken from the March, July, and November issues of six magazines in one year. Discuss how this fact influences your interpretation of the results.

2.86 Age of the intended readership. The ads in the study described in the previous exercise were also classified according to the age group of the intended readership. Here is a summary of the data:

	Magazine readership age group	
Model clothing	Young adult	Mature adult
Not sexual (percent)	72.3%	76.1%
Sexual (percent)	27.2%	23.9%
Number of ads	1006	503

Using parts (a) and (b) of the previous exercise as a guide, analyze these data and write a report summarizing your work.

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.)

Anesthetic	A	B	C	D
Death rate	1.7%	1.7%	3.4%	1.9%

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?

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.

3.11 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.)

3.12 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.)

3.13 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.)

Abbott	Decker	Herrera	Lucero	Richter
Abdalla	Devlin	Hersch	Masters	Riley
Alawi	Engel	Hurwitz	Morgan	Samuels
Broden	Fuentes	Irwin	Nelson	Smith
Chai	Garrett	Jiang	Nho	Suarez
Chuang	Gill	Kelley	Ortiz	Upasani
Cordoba	Glover	Kim	Ramdas	Wilson
Custer	Hammond	Landers	Reed	Xiang

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 childcare 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.)

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

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.

3.27 Jamie is a hard-core computer programmer. He and all his friends prefer Jolt Cola (caffeine equivalent to two cups of coffee) to either Coke or Pepsi (caffeine equivalent to less than one cup of coffee). Explain why Jamie's experience is not good evidence that most young people prefer Jolt to Coke or Pepsi.

3.28 When the discussion turns to the pros and cons of wearing automobile seat belts, Herman always brings up the case of a friend who survived an accident because he was not wearing a seat belt. The friend was thrown out of the car and landed on a grassy bank, suffering only minor injuries, while the car burst into flames and was destroyed. Explain briefly why this anecdote does not provide good evidence that it is safer not to wear seat belts.

3.29 Several large observational studies suggested that women who take hormones such as estrogen after menopause have lower risk of a heart attack than women who do not take hormones. Hormone replacement became popular. But in 2002, several careful experiments showed that hormone replacement does not reduce heart attacks. The National Institutes of Health, after reviewing the evidence, concluded that the observational studies were wrong. Taking hormones after menopause quickly fell out of favor.
(a) Explain the difference between an observational study and an experiment to compare women who do and don't take hormones after menopause.
(b) Suggest some characteristics of women who choose to take hormones that might affect the rate of heart attacks. In an observational study, these characteristics are confounded with taking hormones.

3.30 Moderate use of alcohol is associated with better health. Some studies suggest that drinking wine rather than beer or spirits confers added health benefits.
(a) Explain the difference between an observational study and an experiment to compare people who drink wine with people who drink beer.
(b) Suggest some characteristics of wine drinkers that might benefit their health. In an observational study, these characteristics are confounded with drinking wine.

3.31 Try to find information on this question: what percent of college undergraduates work part-time or full-time while they are taking classes? Start with the Web site of the National Center for Education Statistics, `nces.ed.gov`. Keep a log of your search for this information.

3.32 Doctors identify "chronic tension-type headaches" as headaches that occur almost daily for at least six months. Can antidepressant medications or stress management training reduce the number and severity of these headaches? Are both together more effective than either alone? Investigators compared four treatments: antidepressant alone, placebo alone, antidepressant plus stress management, and placebo plus stress management. Outline the design of the experiment. The headache sufferers named below have agreed to participate in the study. Use software or Table B at line 130 to randomly assign the subjects to the treatments.

Acosta	Duncan	Han	Liang	Padilla	Valasco
Asihiro	Durr	Howard	Maldonado	Plochman	Vaughn
Bennett	Edwards	Hruska	Marsden	Rosen	Wei
Bikalis	Farouk	Imrani	Montoya	Solomon	Wilder
Chen	Fleming	James	O'Brian	Trujillo	Willis
Clemente	George	Kaplan	Ogle	Tullock	Zhang

3.33 How does smoking marijuana affect willingness to work? Canadian researchers persuaded young adult men who used marijuana to live for 98 days in a "planned environment." The men earned money by weaving belts. They used their earnings

to pay for meals and other consumption and could keep any money left over. One group smoked two potent marijuana cigarettes every evening. The other group smoked two weak marijuana cigarettes. All subjects could buy more cigarettes but were given strong or weak cigarettes, depending on their group. Did the weak and strong groups differ in work output and earnings?
(a) Outline the design of this experiment.
(b) Here are the names of the 20 subjects. Use software or Table B at line 131 to carry out the randomization your design requires.

Abbott	Decker	Gutierrez	Lucero	Rosen
Afifi	Engel	Hwang	McNeill	Thompson
Brown	Fluharty	Iselin	Morse	Travers
Chen	Gerson	Kaplan	Quinones	Ullmann

3.34 Eye cataracts are responsible for over 40% of blindness around the world. Can drinking tea regularly slow the growth of cataracts? We can't experiment on people, so we use rats as subjects. Researchers injected 18 young rats with a substance that causes cataracts. One group of the rats also received black tea extract; a second group received green tea extract; and a third got a placebo, a substance with no effect on the body. The response variable was the growth of cataracts over the next six weeks. Yes, both tea extracts did slow cataract growth.
(a) Outline the design of this experiment.
(b) Use software or Table B, starting at line 142, to assign rats to treatments.

3.35 Workers who survive a lay-off of other employees at their location may suffer from "survivor guilt." A study of survivor guilt and its effects used as subjects 120 students who were offered an opportunity to earn extra course credit by doing proofreading. Each subject worked in the same cubicle as another student, who was an accomplice of the experimenters. At a break midway through the work, one of three things happened:

> Treatment 1: The accomplice was told to leave; it was explained that this was because she performed poorly.
>
> Treatment 2: It was explained that unforeseen circumstances meant there was only enough work for one person. By "chance," the accomplice was chosen to be laid off.
>
> Treatment 3: Both students continued to work after the break.

The subjects' work performance after the break was compared with performance before the break.
(a) Outline the design of this completely randomized experiment.
(b) If you are using software, choose the subjects for Treatment 1. If not, use Table B at line 123 to choose the first four subjects for Treatment 1.

3.36 Twenty-four public middle schools agree to participate in the experiment described in Exercise 3.35. Use a diagram to outline a completely randomized design for this experiment. Then do the randomization required to assign schools to treatments. If you use Table B, start at line 105.

3.37 Stores advertise price reductions to attract customers. What type of price cut is most attractive? Market researchers prepared ads for athletic shoes announcing different levels of discounts (20%, 40%, 60%, or 80%). The student subjects who read the ads were also given "inside information" about the fraction of shoes on sale (25%, 50%, 75%, or 100%). Each subject then rated the attractiveness of the sale on a scale of 1 to 7.
(a) There are two factors. Make a sketch that displays the treatments formed by all combinations of levels of the factors.
(b) Outline a completely randomized design using 80 student subjects. Use software or Table B at line 133 to choose the subjects for the first treatment.

3.38 We often see players on the sidelines of a football game inhaling oxygen. Their coaches think this will speed their recovery. We might measure recovery from intense exercise as follows: Have a football player run 100 yards three times in quick succession. Then allow three minutes to rest before running 100 yards again. Time the final run. Because players vary greatly in speed, you plan a matched pairs experiment using 20 football players as subjects. Describe the design of such an experiment to investigate the effect of inhaling oxygen during the rest period. Why should each player's two trials be on different days? Use Table B at line 170 to decide which players will get oxygen on their first trial.

3.39 Calcium is important to the development of young girls. To study how the bodies of young girls process calcium, investigators used the setting of a summer camp. Calcium was given in Hawaiian Punch at either a high or a low level. The camp diet was otherwise the same for all girls. Suppose that there are 60 campers.
(a) Outline a completely randomized design for this experiment.
(b) Describe a matched pairs design in which each girl receives both levels of calcium (with a "washout period" between). What is the advantage of the matched pairs design over the completely randomized design?
(c) The same randomization can be used in different ways for both designs. Label the subjects 01 to 60. You must choose 30 of the 60. Use Table B at line 160 to choose just the first 5 of the 30. How are the 30 subjects chosen treated in the completely randomized design? How are they treated in the matched pairs design?

3.40 Twenty overweight females have agreed to participate in a study of the effectiveness of four reducing regimens, A, B, C, and D. The researcher first calculates how overweight each subject is by comparing the subject's actual weight with her "ideal" weight. The subjects and their excess weights in pounds are

Birnbaum	35	Hernandez	25	Moses	25	Smith	29
Brown	34	Jackson	33	Nevesky	39	Stall	33
Brunk	30	Kendall	28	Obrach	30	Tran	35
Dixon	34	Loren	32	Rodriguez	30	Wilansky	42
Festinger	24	Mann	28	Santiago	27	Williams	22

The response variable is the weight lost after eight weeks of treatment. Because the initial amount overweight will influence the response variable, a block design is appropriate.

(a) Arrange the subjects in order of increasing excess weight. Form five blocks by grouping the four least overweight, then the next four, and so on.
(b) Use Table B to do the required random assignment of subjects to the four reducing regimens separately within each block. Be sure to explain exactly how you used the table.

3.41 Fractures of the spine are common and serious among women with advanced osteoporosis (low mineral density in the bones). Can taking strontium renelate help? A large medical trial assigned 1649 women to take either strontium renelate or a placebo each day. All of the subjects had osteoporosis and had had at least one fracture. All were taking calcium supplements and receiving standard medical care. The response variables were measurements of bone density and counts of new fractures over three years. The subjects were treated at 10 medical centers in 10 different countries. Outline an appropriate design for this experiment. Explain why this is the proper design.

3.42 Final Fu. Your friends are big fans of "Final Fu," MTV2's martial arts competition. To what extent do you think you can generalize your preferences for this show to all students at your college?

3.43 Compost tea. Compost tea is rich in microorganisms that help plants grow. It is made by soaking compost in water. (Based on a study conducted by Brent Ladd, a Water Quality Specialist with the Purdue University Department of Agricultural and Biological Engineering.) Design a comparative experiment that will provide evidence about whether or not compost tea works for a particular type of plant that interests you. Be sure to provide all details regarding your experiment, including the response variable or variables that you will measure.

3.44 Measuring water quality in streams and lakes. Water quality of streams and lakes is an issue of concern to the public. Although trained professionals typically are used to take reliable measurements, many volunteer groups are gathering and distributing information based on data that they collect. (Based on a study conducted by Sandra Simonis under the direction of Professor Jon Harbor from the Purdue University Earth and Atmospheric Sciences Department.) You are part of a team to train volunteers to collect accurate water quality data. Design an experiment to evaluate the effectiveness of the training. Write a summary of your proposed design to present to your team. Be sure to include all of the details that they will need to evaluate your proposal.

3.45 Eye cataracts. Eye cataracts are responsible for over 40% of blindness around the world. Can drinking tea regularly slow the growth of cataracts? We can't experiment on people, so we use rats as subjects. Researchers injected 21 young rats with a substance that causes cataracts. One group of the rats also received black tea extract; a second group received green tea extract; and a third got a placebo, a substance with no effect on the body. The response variable was the growth of cataracts over the next six weeks. Yes, both tea extracts did slow cataract growth. (Geetha Thiagarajan et al., "Antioxidant properties of green and black tea, and their potential ability to retard the progression of eye lens cataract," *Experimental*

Eye Research, 73 (2001), pp. 393–401.)
(a) Outline the design of this experiment.
(b) Use software or Table B, starting at line 120, to assign rats to treatments.

3.46 Treatment of clothing fabric. A maker of fabric for clothing is setting up a new line to "finish" the raw fabric. The line will use either metal rollers or natural-bristle rollers to raise the surface of the fabric; a dyeing cycle time of either 30 minutes or 40 minutes; and a temperature of either 150° or 175°C. An experiment will compare all combinations of these choices. Four specimens of fabric will be subjected to each treatment and scored for quality.
(a) What are the factors and the treatments? How many individuals (fabric specimens) does the experiment require?
(b) Outline a completely randomized design for this experiment. (You need not actually do the randomization.)

3.47 Treatment of pain for cancer patients. Health care providers are giving more attention to relieving the pain of cancer patients. An article in the journal *Cancer* surveyed a number of studies and concluded that controlled-release morphine tablets, which release the painkiller gradually over time, are more effective than giving standard morphine when the patient needs it. (Carol A. Warfield, "Controlled-release morphine tablets in patients with chronic cancer pain," *Cancer,* 82 (1998), pp. 2299–2306.) The "methods" section of the article begins: "Only those published studies that were controlled (i.e., randomized, double blind, and comparative), repeated-dose studies with CR morphine tablets in cancer pain patients were considered for this review." Explain the terms in parentheses to someone who knows nothing about medical trials.

3.48 Saint-John's-wort and depression. Does the herb Saint-John's-wort relieve major depression? Here are some excerpts from the report of a study of this issue. (R. C. Shelton et al., "Effectiveness of St. John's wort in major depression," *Journal of the American Medical Association,* 285 (2001), pp. 1978–1986.) The study concluded that the herb is no more effective than a placebo.
(a) "Design: Randomized, double-blind, placebo-controlled clinical trial. . . ." Explain the meaning of each of the terms in this description.
(b) "Participants . . . were randomly assigned to receive either Saint-John's-wort extract ($n = 98$) or placebo ($n = 102$). . . . The primary outcome measure was the rate of change in the Hamilton Rating Scale for Depression over the treatment period." Based on this information, use a diagram to outline the design of this clinical trial.

3.49 The Monday effect on stock prices. Puzzling but true: stocks tend to go down on Mondays. There is no convincing explanation for this fact. A recent study looked at this "Monday effect" in more detail, using data on the daily returns of stocks on several U.S. exchanges over a 30-year period. Here are some of the findings:

> *To summarize, our results indicate that the well-known Monday effect is caused largely by the Mondays of the last two weeks of the month. The mean Monday return of the first three weeks of the month is, in general, not significantly different from zero and is generally significantly higher*

> *than the mean Monday return of the last two weeks. Our finding seems to make it more difficult to explain the Monday effect.*

(K. Wang, Y. Li, and J. Erickson, "A new look at the Monday effect," *Journal of Finance,* 52 (1997), pp. 2171–2186.) A friend thinks that "significantly" in this article has its plain English meaning, roughly "I think this is important." Explain in simple language what "significantly higher" and "not significantly different from zero" actually tell us here.

3.50 Five-digit zip codes and delivery time of mail. Does adding the five-digit postal zip code to an address really speed up delivery of letters? Does adding the four more digits that make up "zip + 4" speed delivery yet more? What about mailing a letter on Monday, Thursday, or Saturday? Describe the design of an experiment on the speed of first-class mail delivery. For simplicity, suppose that all letters go from you to a friend, so that the sending and receiving locations are fixed.

Section 3.2

3.51 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?

3.52 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?

3.53 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?

3.54 A newspaper advertisement for *USA Today: The Television Show* said:

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

Explain why this opinion poll is almost certainly biased.

3.55 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.)

Abate	Dubois	Hixson	Putnam
Anderson	Fernandez	Klassen	Rodriguez
Baxter	Frank	Liang	Rubin
Bowen	Fuhrmann	Moser	Santiago
Bruvold	Goel	Naber	Shen
Casella	Gupta	Petrucelli	Shyr
Choi	Hicks	Pliego	Sundheim

3.56 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.)

3.57 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.

Acosta	Dewald	Huang	Puri
Ali	Fleming	Kim	Richards
Baxter	Fonseca	Lujan	Rodriguez
Bowman	Gates	Mourning	Santiago
Brams	Goel	Nunez	Shen
Cortez	Gomez	Peters	Vega
Cross	Hernandez	Pliego	Watanabe

3.58 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.)

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

3.59 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.

3.60 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.

3.61 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.)

3.62 A committee on community relations in a college town plans to survey local businesses about the importance of students as customers. From telephone book listings, the committee chooses 150 businesses at random. Of these, 73 return the questionnaire mailed by the committee. What is the population for this sample survey? What is the sample? What is the rate (percent) of nonresponse?

3.63 A Gallup Poll asked, "Do you think the U.S. should take the leading role in world affairs, take a major role but not the leading role, take a minor role, or take no role at all in world affairs?" Gallup's report said, " Results are based on telephone interviews with 1,002 national adults, aged 18 and older, conducted Feb. 9-12, 2004."
(a) What is the population for this sample survey? What was the sample size?
(b) Gallup notes that the order of the four possible responses was rotated when the question was read over the phone. Why was this done?

3.64 For each of the following sampling situations, identify the population as exactly as possible. That is, say what kind of individuals the population consists of and say exactly which individuals fall in the population. If the information given is not complete, complete the description of the population in a reasonable way.

(a) An opinion poll contacts 1161 adults and asks them, "Which political party do you think has better ideas for leading the country in the twenty-first century?"
(b) A furniture maker buys hardwood in large lots. The supplier is supposed to dry the wood before shipping—wood that is not dry won't hold its size and shape. The furniture maker chooses 5 pieces of wood from each lot and tests their moisture content. If any piece exceeds 12% moisture content, the entire lot is sent back.
(c) The American Community Survey (ACS) will replace the census "long form" starting with the 2010 census. The main part of the ACS contacts 250,000 addresses by mail each month, with follow-up by phone and in person if there is no response. Each household answers questions about their housing, economic, and social status.

3.65 An opinion poll calls 1800 randomly chosen residential telephone numbers, then asks to speak with an adult member of the household. The interviewer asks, "How many movies have you watched in a movie theater in the past 12 months?"
(a) What population do you think the poll has in mind?
(b) In all, 1131 people respond. What is the rate (percent) of nonresponse?
(c) What source of response error is likely for the question asked?

3.66 You want to ask a sample of college students the question "How much do you trust information about health that you find on the Internet—a great deal, somewhat, not much, or not at all?" You try out this and other questions on a pilot group of 10 students chosen from your class. The class members are

Anderson	Eckstein	Johnson	Puri
Arroyo	Fernandez	Kim	Richards
Batista	Fullmer	Molina	Rodriguez
Bell	Gandhi	Morgan	Samuels
Burke	Garcia	Nguyen	Shen
Calloway	Glauser	Palmiero	Velasco
Delluci	Helling	Percival	Washburn
Drasin	Husain	Prince	Zhao

Choose an SRS. If you use Table B, start at line 139.

3.67 Popularity of news personalities. A Gallup Poll conducted telephone interviews with 1001 U.S. adults aged 18 and over on July 24-27, 2006. One of the questions asked whether the respondents had a favorable or an unfavorable opinion of 17 news personalities. Diane Sawyer received the highest rating, with 80% of the respondents giving her a favorable rating. (From `poll.gallup.com` on August 8, 2006.) (a) What is the population for this sample survey? What was the sample size?
(b) The report on the survey states that 8% of the respondents either never heard of Sawyer or had no opinion about her. When they included only those who provided an opinion, Sawyer's approval percent rose to 88% and she was still at the top of the list. Charles Gibson, on the other hand, was ranked eighth on the original list, with a 55% favorable rating. When only those providing an opinion were counted, his rank rose to second, with 87% approving. Discuss the advantages and disadvantages of the two different ways of reporting the approval percent. State which one you prefer and why.

3.68 The Excite Poll. The Excite Poll can be found online at `poll.excite.com`. The question appears on the screen, and you simply click buttons to vote “Yes,” “No,” “Not sure,” or “Don’t care.” On July 22, 2006, the question was “Do you agree or disagree with proposed legislation that would discontinue the U.S. penny coin?” In all, 631 said “Yes,” another 564 said “No,” and the remaining 65 indicated that they were not sure.
(a) What is the sample size for this poll?
(b) Compute the percent of responses in each of the possible response categories.
(c) Discuss the poll in terms of the population and sample framework that we have studied in this chapter.

Section 3.3

In the following four exercises, state whether each boldface number is a parameter or a statistic.

3.69 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.

3.70 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.

3.71 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.

3.72 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.

3.73 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?

3.74 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?

3.75 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?
In the following four exercises, state whether each boldface number is a parameter or a statistic.

3.76 An opinion poll uses random digit dialing equipment to dial 2000 randomly chosen residential telephone numbers. Of these, **621** are unlisted numbers. This isn't surprising, because **35%** of all residential numbers are unlisted.

3.77 A study of voting chose 663 registered voters at random shortly after an election. Of these, **72%** said they had voted in the election. Election records show that only **56%** of registered voters voted in the election.

3.78 The Tennessee STAR experiment randomly assigned children to regular or small classes during their first four years of school. When these children reached high school, **40.2%** of blacks from small classes took the ACT or SAT college entrance exams. Only **31.7%** of blacks from regular classes took one of these exams.

3.79 How does caffeine affect our bodies? In a matched pairs experiment, subjects pushed a button as quickly as they could after taking a caffeine pill and again after taking a placebo pill. The mean pushes per minute were $\boldsymbol{\overline{x} = 283}$ for the placebo and $\boldsymbol{\overline{x} = 311}$ for caffeine.

3.80 Opinions of Hispanics. A New York Times News Service article on a poll concerned with the opinions of Hispanics includes this paragraph:

> *The poll was conducted by telephone from July 13 to 27, with 3,092 adults nationwide, 1,074 of whom described themselves as Hispanic. It has a margin of sampling error of plus or minus three percentage points for the entire poll and plus or minus four percentage points for Hispanics. Sample sizes for most Hispanic nationalities, like Cubans or Dominicans, were too small to break out the results separately.*

(Adam Nagourney and Janet Elder, "New York Times CBS Poll: What Hispanics Believe," found online at `Hispanic.cc`.)
(a) Why is the "margin of sampling error" larger for Hispanics than for all 3092 respondents?

(b) Why would a very small sample size prevent a responsible news organization from breaking out results for Cubans?

3.81 Real estate ownership. An agency of the federal government plans to take an SRS of residents in each state to estimate the proportion of owners of real estate in each state's population. The populations of the states range from less than 500,000 people in Wyoming to about 35 million in California.
(a) Will the variability of the sample proportion vary from state to state if an SRS of size 2000 is taken in each state? Explain your answer.
(b) Will the variability of the sample proportion change from state to state if an SRS of 1/10 of 1% (0.001) of the state's population is taken in each state? Explain your answer.

Section 3.4

3.82 Serving as an experimental subject for extra credit. Students taking Psychology 001 are required to serve as experimental subjects. Students in Psychology 002 are not required to serve, but they are given extra credit if they do so. Students in Psychology 003 are required either to sign up as subjects or to write a term paper. Serving as an experimental subject may be educational, but current ethical standards frown on using "dependent subjects" such as prisoners or charity medical patients. Students are certainly somewhat dependent on their teachers. Do you object to any of these course policies? If so, which ones, and why?

3.83 The 2000 census. The 2000 census long form asked 53 detailed questions, for example:

> *Do you have COMPLETE plumbing facilities in this house, apartment, or mobile home; that is, 1) hot and and cold piped water, 2) a flush toilet, and 3) a bathtub or shower?*

The form also asked your income in dollars, broken down by source, and whether any "physical, mental, or emotional condition" causes you difficulty in "learning, remembering, or concentrating." Some members of Congress objected to these questions, even though Congress had approved them.

Give brief arguments on both sides of the debate over the long form: the government has legitimate uses for such information, but the questions seem to invade people's privacy.

Chapter 3 Review Exercises

3.84 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.

3.85 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.

3.86 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.

3.87 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. 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.

3.88 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.

3.89 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?

3.90 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.

3.91 As the millennium approached, *Time* magazine launched a Person of the Century Poll on the Internet. Web users could choose the person “who most influenced the course of history” during the twentieth century. The top choice was Elvis Presley, with 625,045 votes. Further down the list, Linus Torvalds (originator of the Linux operating system for PCs) beat out Nelson Mandela and Princess Diana. *Time* ignored an online campaign to get out the vote for Jesus, leaving him off the list. Use this example in a brief discussion of the weaknesses of voluntary response samples.

3.92 Cash bonues for the unemployed. Will cash bonuses speed the return to work of unemployed people? The Illinois Department of Employment Security designed an experiment to find out. The subjects were 10,065 people aged 20 to 54 who were filing claims for unemployment insurance. Some were offered $500 if they found a job within 11 weeks and held it for at least 4 months. Others could tell potential employers that the state would pay the employer $500 for hiring them.

A control group got neither kind of bonus. (Based on Stephen A. Woodbury and Robert G. Spiegelman, "Bonuses to workers and employers to reduce unemployment: randomized trials in Illinois," *American Economic Review,* 77 (1987), pp. 513–530.)
(a) Suggest a few response variables of interest to the state and outline the design of the experiment.
(b) How will you label the subjects for random assignment? Use Table B at line 167 to choose the first 3 subjects for the first treatment.

3.93 Prostate treatment study using Canada's national health records. A large observational study used records from Canada's national health care system to compare the effectiveness of two ways to treat prostate disease. The two treatments are traditional surgery and a new method that does not require surgery. The records described many patients whose doctors had chosen one or the other method. The study found that patients treated by the new method were significantly more likely to die within 8 years. (Based on Christopher Anderson, "Measuring what works in health care," *Science,* 263 (1994), pp. 1080–1082.)
(a) Further study of the data showed that this conclusion was wrong. The extra deaths among patients who received the new treatment could be explained by lurking variables. What lurking variables might be confounded with a doctor's choice of surgical or nonsurgical treatment?
(b) You have 300 prostate patients who are willing to serve as subjects in an experiment to compare the two methods. Use a diagram to outline the design of a randomized comparative experiment.

3.94 What type of study? What is the best way to answer each of the questions below: an experiment, a sample survey, or an observational study that is not a sample survey? Explain your choices.
(a) Are people generally satisfied with how things are going in the country right now?
(b) Do college students learn basic accounting better in a classroom or using an online course?
(c) How long do your teachers wait on the average after they ask their class a question?

3.95 Discolored french fries. Few people want to eat discolored french fries. Potatoes are kept refrigerated before being cut for french fries to prevent spoiling and to preserve flavor. But immediate processing of cold potatoes causes discoloring due to complex chemical reactions. The potatoes must therefore be brought to room temperature before processing. Design an experiment in which tasters will rate the color and flavor of french fries prepared from several groups of potatoes. The potatoes will be fresh picked, stored for a month at room temperature, or stored for a month refrigerated. They will then be sliced and cooked either immediately or after an hour at room temperature.
(a) What are the factors and their levels, the treatments, and the response variables?
(b) Describe and outline the design of this experiment.
(c) It is efficient to have each taster rate fries from all treatments. How will you use randomization in presenting fries to the tasters?

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.)

4.6 Hold a penny upright on its edge under your forefinger on a hard surface, then snap it with your other forefinger so that it spins for some time before falling. Based on 50 spins, what is the probability of heads?

4.7 You may feel that it is obvious that the probability of a head in tossing a coin is about 1/2 because the coin has two faces. Such opinions are not always correct. The previous exercise asked you to spin a penny rather than toss it—that changes the probability of a head. Now try another variation. Stand a nickel on edge on a hard, flat surface. Pound the surface with your hand so that the nickel falls over. What is the probability that it falls with heads upward? Make at least 50 trials to estimate the probability of a head.

4.8 Many Internet sites give the probabilities of being dealt various five-card poker hands. For example, the probability of being dealt two pairs is approximately 1/21. Explain is simple language what "probability 1/21" means. Also explain why it does *not* mean that in 21 deals you will get exactly one two-pair hand.

4.9 About 30% of adult Internet users are between 18 and 29 years of age. Suppose the probability that a randomly chosen Internet user is in this age group is exactly 0.3. Use the *Probability* applet to make a study of short-term variability and long-term regularity as follows.
(a) Set the probability of heads to 0.3. Each head stands for an Internet user who is between 18 and 29 and each tail is a user who is not. Set the number of tosses to 20. Click "Toss" to make 20 tosses. What was the proportion of heads? Do this 25 times, keep a record of the 25 proportions of heads, and make a stemplot of these numbers. *Lesson:* In the short run (20 repetitions) proportions are quite variable and are often not close to the probability.
(b) With the probability of heads still set to 0.3, make 200 tosses. (Set tosses to 40 and click "Toss" five times without a reset.) What was the proportion of heads? Do this 25 times and make a stemplot of the 25 proportions of heads. *Lesson:* More repetitions make proportions less variable and generally closer to the probability. Of course, 200 repetitions is not "the long run," but you see the idea.

4.10 If you use statistical software, you can 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 Yes/No trials, each with probability p of a "Yes," and records just the count of "Yes" outcomes.
(a) Simulate 100 draws of 20 Internet users from the population in Exercise 4.8.

(That is, ask the software to generate 100 binomial observations, each with $n = 20$ trials and probability $p = 0.3$ of a "Yes.") Record the count in the 18 to 29 age group on each draw. Convert the counts into percents of the 20 Internet users in each trial who are 18 to 29. Make a histogram of these 100 percents. Describe the shape, center, and spread of this distribution.
(b) Now simulate drawing 320 Internet users. (That is, set $n = 320$ and $p = 0.3$.) Do this 100 times and record the percent in the 18 to 29 age group for each of the 100 draws. Make a histogram of the percents and describe the shape, center, and spread of the distribution.
(c) In what ways are the distributions in parts (a) and (b) alike? In what ways do they differ? (Because regularity emerges in the long run, we expect the results of drawing 320 subjects to be less variable than the results of drawing 20 subjects.)

4.11 The color of candy. It is reasonable to think that packages of M&M's Milk Chocolate Candies are filled at the factory with candies chosen at random from the very large number produced. So a package of M&M's contains a number of repetitions of a random phenomenon: choosing a candy at random and noting its color. What is the probability that an M&M's Milk Chocolate Candy is green? To find out, buy one or more packs. How many candies did you examine? How many were green? What is your estimate of the probability that a randomly chosen candy is green?

Section 4.2

4.12 The percent return on U.S. common stocks in the next year is random. The following table 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.

Year	% return	Year	% return
1971	14.6	1986	18.6
1972	18.9	1987	5.1
1973	−14.8	1988	16.8
1974	−26.4	1989	31.5
1975	37.2	1990	−3.1
1976	23.6	1991	30.4
1977	−7.4	1992	7.6
1978	6.4	1993	10.1
1979	18.2	1994	1.3
1980	32.3	1995	37.6
1981	−5.0	1996	23.0
1982	21.5	1997	33.4
1983	22.4	1998	28.6
1984	6.1	1999	21.0
1985	31.6	2000	−9.1

4.13 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 corresponding to extreme discomfort.

4.14 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.

4.15 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.

4.16 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, $P(\mathrm{H}) = 0.55$ and $P(\mathrm{T}) = 0.45$.
(b) When two coins are tossed, $P(\mathrm{HH}) = 0.4$, $P(\mathrm{HT}) = 0.4$, $P(\mathrm{TH}) = 0.4$, and $P(\mathrm{TT}) = 0.4$.
(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 $P(6) = 1/6$.

4.17 Here are several assignments of probabilities to the six faces of a die:

Outcome	1	2	3	4	5	6
Probabilities 1	1/3	0	1/6	0	1/6	1/3
Probabilities 2	1/6	1/6	1/6	1/6	1/6	1/6
Probabilities 3	1/7	1/7	1/7	1/7	1/7	1/7
Probabilities 4	1/3	1/3	−1/6	−1/6	1/3	1/3

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.

4.18 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:

Language	Spanish	French	German	All others	None
Probability	0.26	0.09	0.03	0.03	0.59

(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?

4.19 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 following table gives the probability of each color for a randomly chosen plain M&M:

Color	Brown	Red	Yellow	Green	Orange	Blue
Probability	0.3	0.2	0.2	0.1	0.1	?

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:

Color	Brown	Red	Yellow	Green	Orange	Blue
Probability	0.2	0.1	0.2	0.1	0.1	?

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?

4.20 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?

4.21 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?

4.22 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/`.)

Color	Silver	White	Black	Gray	Blue	Medium red
Probability	0.201	0.184	0.116	0.088	0.085	0.069

(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?

4.23 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?

4.24 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 following table gives the probabilities that a randomly chosen worker falls into each of 12 sex-by-occupation classes.

Class	A	B	C	D	E	F
Male	0.14	0.11	0.06	0.11	0.12	0.03
Female	0.09	0.20	0.08	0.01	0.04	0.01

(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?

4.25 The Pick 4 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*?

4.26 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?

4.27 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?

4.28 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?

4.29 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?

4.30 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.)

Side	broad convex	broad concave	narrow flat	narrow hollow
Probability	0.4	0.4	0.1	0.1

4.31 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.

4.32 A randomly chosen subject arrives for a study of exercise and fitness. Describe a sample space for each of the following. (In some cases, you may have some freedom in your choice of S.)
(a) The subject is either female or male.
(b) After 10 minutes on an exercise bicycle, you ask the subject to rate his or her effort on the Rate of Perceived Exertion (RPE) scale. RPE ranges in whole-number steps from 6 (no exertion at all) to 20 (maximal exertion).
(c) You measure VO2, the maximum volume of oxygen consumed per minute during exercise. VO2 is generally between 2.5 liters per minute and 6 liters per minute.
(d) You measure the maximum heart rate (beats per minute).

4.33 Choose a student at random from a large statistics class. Give a reasonable sample space S for answers to each of these questions. (In some cases you may have some freedom in specifying S.)
(a) Are you right-handed or left-handed?
(b) What is your height in centimeters? (One inch is 2.54 centimeters.)
(c) How much money in coins (not bills) are you carrying?
(d) How many minutes did you study last night?

4.34 Role-playing games like Dungeons & Dragons use many different types of dice. Suppose that a four-sided die has faces marked 1, 2, 3, 4. The intelligence of a character is determined by rolling this die twice and adding 1 to the sum of the spots.
(a) What is the sample space for rolling the die twice (spots on first and second rolls)?
(b) What is the sample space for the character's intelligence?

4.35 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) Roll a die and record the count of spots on the up-face: $P(1) = 0$, $P(2) = 1/6$, $P(3) = 1/3$, $P(4) = 1/3$, $P(5) = 1/6$, $P(6) = 0$.
(b) Choose a college student at random and record sex and enrollment status: $P(\text{female full-time}) = 0.56$, $P(\text{female part-time}) = 0.24$, $P(\text{male full-time}) = 0.44$, $P(\text{male part-time}) = 0.17$.
(c) Deal a card from a shuffled deck: $P(\text{clubs}) = 12/52$, $P(\text{diamonds}) = 12/52$, $P(\text{hearts}) = 12/52$, $P(\text{spades}) = 16/52$.

4.36 Role-playing games like Dungeons & Dragons use many different types of dice. Suppose that a four-sided die has faces marked 1, 2, 3, 4. The intelligence of a character is determined by rolling this die twice and adding 1 to the sum of the spots. The faces are equally likely and the two rolls are independent. In the previous exercise you gave the sample space S for the character's intelligence. Now give the assignment of

4.37 Dugout Lou thinks that the probabilities for the American League baseball champion are as follows. The Yankees have probability 0.6 of winning. The Red Sox and Angels have equal probabilities of winning. The Athletics and White Sox also have equal probabilities, but their probabilities are one-third that of the Red Sox and Angels. No other team has a chance. What is Lou's assignment of probabilities to teams?

4.38 When do you study? A student is asked on which day of the week he or she spends the most time studying. What is the sample space?

4.39 Sample space for heights. You record the height in inches of a randomly selected student. What is the sample space?

4.40 Phone-related accidents on Monday or Friday. Some states are considering laws that will ban the use of cell phones while driving because they believe that the ban will reduce phone-related car accidents. One study classified these types of accidents by the day of the week when they occurred. (From D. A. Redelmeier and R. J. Tibshirani, "Association between cellular-telephone calls and motor vehicle collisions," *New England Journal of Medicine,* 336 (1997) pp. 453–458.) For this example, we use the values from this study as our probability model. Here are the probabilities:

Day	Sun.	Mon.	Tues.	Wed.	Thur.	Fri.	Sat.
Probability	0.03	0.19	0.18	0.23	0.19	0.16	0.02

Find the probability that a phone-related accident occurred on a Monday or a Friday.

4.41 Not on Wednesday. Refer to the previous exercise. Find the probability that a phone-related accident occurred on a day other than a Wednesday.

4.42 Spam topics. A majority of email messages are now "spam." Choose a spam email message at random. Here is the distribution of topics (Robyn Greenspan, "The deadly duo: spam and viruses, October 2003," found online at `cyberatlas.internet.com`):

Topic	Adult	Financial	Health	Leisure	Products	Scams
Probability	0.145	0.162	0.073	0.078	0.210	0.142

(a) What is the probability that a spam email does not concern one of these topics?
(b) Corinne is particularly annoyed by spam offering "adult" content (that is, pornography) and scams. What is the probability that a randomly chosen spam email falls into one or the other of these categories?

4.43 Race in the census. The 2000 census allowed each person to choose from a long list of races. That is, in the eyes of the Census Bureau, you belong to whatever race you say you belong to. "Hispanic/Latino" is a separate category; Hispanics may be of any race. If we choose a resident of the United States at random, the 2000 census gives these probabilities:

	Hispanic	Not Hispanic
Asian	0.000	0.036
Black	0.003	0.121
White	0.060	0.691
Other	0.062	0.027

Let A be the event that a randomly chosen American is Hispanic, and let B be the event that the person chosen is white.
(a) Verify that the table gives a legitimate assignment of probabilities.
(b) What is $P(A)$?
(c) Describe B^c in words and find $P(B^c)$ by the complement rule.
(d) Express "the person chosen is a non-Hispanic white" in terms of events A and B. What is the probability of this event?

4.44 Are the events independent? The previous exercise assigns probabilities for the ethnic background of a randomly chosen resident of the United States. Let A be the event that the person chosen is Hispanic, and let B be the event that he or she is white. Are events A and B independent? How do you know?

Section 4.3

4.45 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?

4.46 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?

4.47 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:

Number of cars X	0	1	2	3	4	5
Probability	0.09	0.36	0.35	0.13	0.05	0.02

(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?

4.48 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?

4.49 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:

Years	4	5	6	7	8	9	10	11	12
Probability	0.010	0.007	0.007	0.013	0.032	0.068	0.070	0.041	0.752

(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?

4.50 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.

4.51 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) $P(0.14 \leq \hat{p} \leq 0.16)$

4.52 Choose an American household at random and let the random variable Y be the number of persons living in the household. Here is the distribution of Y:

Number of persons	1	2	3	4	5	6	7
Household probability	0.27	0.33	0.16	0.14	0.06	0.03	0.01
Family probability	0	0.44	0.22	0.20	0.09	0.03	0.02

(a) Express "more than one person lives in this household" in terms of Y. What is the probability of this event?
(b) What is $P(2 < Y \leq 4)$?
(c) What is $P(Y \neq 2)$?

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

Rooms	1	2	3	4	5	6	7	8	9	10
Owned	0.003	0.002	0.023	0.104	0.210	0.224	0.197	0.149	0.053	0.035
Rented	0.008	0.027	0.287	0.363	0.164	0.093	0.039	0.013	0.003	0.003

(a) Express "the unit has 5 or more rooms" in terms of X. What is the probability of this event?
(b) Express the event $\{X > 5\}$ in words. What is its probability?
(c) What important fact about discrete random variables does comparing your answers to (a) and (b) illustrate?

4.54 Let X be a random number between 0 and 1 produced by a uniform random number generator. Find the following probabilities:
(a) $P(X < 0.5)$
(b) $P(X \leq 0.5)$
(c) What important fact about continuous random variables does comparing your answers to (a) and (b) illustrate?

4.55 Let the random variable X be a uniform random number between 0 and 1. Find the following probabilities:
(a) $P(X \geq 0.27)$
(b) $P(X = 0.27)$
(c) $P(0.27 < X < 1.27)$
(d) $P(0.1 \leq X \leq 0.2$ or $0.8 \leq X \leq 0.9)$
(e) The probability that X is not in the interval 0.3 to 0.8.

4.56 Many random number generators allow users to specify the range of the random numbers to be produced. Suppose that you specify that the range is to be all numbers between 0 and 2. Call the random number generated Y. Then the density curve of the random variable Y has constant height between 0 and 2, and height 0 elsewhere.
(a) What is the height of the density curve between 0 and 2? Draw a graph of the density curve.
(b) Use your graph from (a) and the fact that probability is area under the curve to find $P(Y \leq 1)$.

(c) Find $P(0.5 < Y < 1.3)$.
(d) Find $P(Y \geq 0.8)$.

4.57 Owner-occupied and rented housing units. How do rented housing units differ from units occupied by their owners? Here are the distributions of the number of rooms for owner-occupied units and renter-occupied units in San Jose, California: (From the Census Bureau's 1998 American Housing Survey.)

Rooms	1	2	3	4	5	6	7	8	9	10
Owned	0.003	0.002	0.023	0.104	0.210	0.224	0.197	0.149	0.053	0.035
Rented	0.008	0.027	0.287	0.363	0.164	0.093	0.039	0.013	0.003	0.003

Make probability histograms of these two distributions, using the same scales. What are the most important differences between the distributions for owner-occupied and rented housing units?

4.58 Find the probabilities. Let the random variable X be the number of rooms in a randomly chosen owner-occupied housing unit in San Jose, California. The previous exercise gives the distribution of X.
(a) Express "the unit has 6 or more rooms" in terms of X. What is the probability of this event?
(b) Express the event $\{X > 6\}$ in words. What is its probability?
(c) What important fact about discrete random variables does comparing your answers to (a) and (b) illustrate?

4.59 Households and families in government data. In government data, a household consists of all occupants of a dwelling unit, while a family consists of two or more persons who live together and are related by blood or marriage. So all families form households, but some households are not families. Here are the distributions of household size and of family size in the United States:

Number of persons	1	2	3	4	5	6	7
Household probability	0.27	0.33	0.16	0.14	0.06	0.03	0.01
Family probability	0	0.44	0.22	0.20	0.09	0.03	0.02

Make probability histograms for these two discrete distributions, using the same scales. What are the most important differences between the sizes of households and families?

4.60 Select the members of a student advisory board. Weary of the low turnout in student elections, a college administration decides to choose an SRS of three students to form an advisory board that represents student opinion. Suppose that 40% of all students oppose the use of student fees to fund student interest groups, and that the opinions of the three students on the board are independent. Then the probability is 0.4 that each opposes the funding of interest groups.
(a) Call the three students A, B, and C. What is the probability that A and B support funding and C opposes it?
(b) List all possible combinations of opinions that can be held by students A, B, and C. (*Hint:* There are eight possibilities.) Then give the probability of each of

these outcomes. Note that they are not equally likely.
(c) Let the random variable X be the number of student representatives who oppose the funding of interest groups. Give the probability distribution of X.
(d) Express the event "a majority of the advisory board opposes funding" in terms of X and find its probability.

Section 4.4

4.61 Exercise 4.47 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?

4.62 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 the payoff X on a single play?
(b) What is the mean payoff μ_X?
(c) In the long run, how much does the casino keep from each dollar bet?

4.63 (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?

4.64 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:

Number of toys x_i	0	1	2	3	4	5
Probability p_i	0.03	0.16	0.30	0.23	0.17	0.11

Calculate the mean μ_X and the standard deviation σ_X.

4.65 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 distribution you found in (b) and by applying the addition rule to your results in (a). You should of course get the same answer.

4.66 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?

4.67 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.

4.68 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.

4.69 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.

4.70 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.

4.71 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:

Temperature	540°	545°	550°	555°	560°
Probability	0.1	0.25	0.3	0.25	0.1

(a) Find the mean temperature μ_X and the standard deviation σ_X.
(b) The target temperature is 550°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 μ_Y and standard deviation σ_Y of the temperature of the flame in the Fahrenheit scale?

4.72 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?

4.73 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 μ_Z and σ_Z? Is the average Z more or less variable than the reading Y of the less variable scale?

4.74 Here is the distribution of grades (A = 4, B = 3, and so on) in Statistics 101 at North Carolina State University:

Value of X	0	1	2	3	4
Probability	0.01	0.05	0.30	0.43	0.21

Find the average (that is, the mean) grade in this course.

4.75 Typographical and spelling errors can be either "nonword errors" or "word errors." A nonword error is not a real word, as when "the" is typed as "teh." A word error is a real word, but not the right word, as when "lose" is typed as "loose." When undergraduates are asked to write a 250-word essay (without spell-checking), the number of nonword errors has the following distribution:

Errors	0	1	2	3	4
Probability	0.1	0.2	0.3	0.3	0.1

The number of word errors has this distribution:

Errors	0	1	2	3
Probability	0.4	0.3	0.2	0.1

What are the mean numbers of nonword errors and word errors in an essay?

4.76 Find the mean and standard deviation of the total number of errors (nonword errors plus word errors) in an essay if the error counts have the distributions given in the previous exercise.
(a) The counts of nonword and word errors are independent.
(b) Students who make many nonword errors also tend to make many word errors, so that the correlation between the two error counts is 0.5.

4.77 You buy a hot stock for \$1000. The stock either gains 30% or loses 25% each day, each with probability 0.5. Its returns on consecutive days are independent of each other. You plan to sell the stock after two days.
(a) What are the possible values of the stock after two days, and what is the probability for each value? What is the probability that the stock is worth more after two days than the \$1000 you paid for it?
(b) What is the mean value of the stock after two days? You see that these two criteria give different answers to the question, "Should I invest?"

4.78 For each of the following situations, would you expect the random variables X and Y to be independent? Explain your answers.
(a) X is the rainfall (in inches) on November 6 of this year, and Y is the rainfall at the same location on November 6 of next year.
(b) X is the amount of rainfall today, and Y is the rainfall at the same location tomorrow.
(c) X is today's rainfall at the Orlando, Florida, airport, and Y is today's rainfall at Disney World just outside Orlando.

4.79 A time and motion study measures the time required for an assembly-line worker to perform a repetitive task. The data show that the time required to bring a part from a bin to its position on an automobile chassis varies from car to car with mean 11 seconds and standard deviation 2 seconds. The time required to attach the part to the chassis varies with mean 20 seconds and standard deviation 8 seconds.
(a) What is the mean time required for the entire operation of positioning and attaching the part?
(b) Industry quality programs strive to reduce variation in processes. A training program reduces the standard deviation for attaching the part from 8 seconds to 4 seconds. Will this change your result in (a) if the mean times don't change? Why is reducing the variation nonetheless worthwhile to the automaker?

(c) The study finds that the times required for the two steps are independent. A part that takes a long time to position, for example, does not take more or less time to attach than other parts. How would your answer in (a) change if the two variables were dependent with correlation 0.3?

4.80 Find the standard deviation of the time required for the two-step assembly operation studied in the previous exercise, assuming that the study shows the two times to be independent. Redo the calculation assuming that the two times are dependent, with correlation 0.3. Explain in nontechnical language why positive correlation increases the variability of the total time.

4.81 You have two instruments with which to measure the height of a tower. If the true height is 100 meters, measurements with the first instrument vary with mean 100 meters and standard deviation 1.2 meters. Measurements with the second instrument vary with mean 100 meters and standard deviation 0.85 meter. You make one measurement with each instrument. Your results are X_1 for the first and X_2 for the second, and are independent.
(a) To combine the two measurements, you might average them, $Y = (X_1 + X_2)/2$. What are the mean and standard deviation of Y?
(b) It makes sense to give more weight to the less variable measurement because it is more likely to be close to the truth. Statistical theory says that to make the standard deviation as small as possible you should weight the two measurements inversely proportional to their variances. The variance of X_2 is very close to half the variance of X_1, so X_2 should get twice the weight of X_1. That is, use

$$W = \frac{1}{3}X_1 + \frac{2}{3}X_2$$

What are the mean and standard deviation of W?

4.82 Means of the numbers of rooms in housing units. How do rented housing units differ from units occupied by their owners? Here are the distributions of the number of rooms for owner-occupied units and renter-occupied units in San Jose, California (from the Census Bureau's 1998 American Housing Survey):

Rooms	1	2	3	4	5	6	7	8	9	10
Owned	0.003	0.002	0.023	0.104	0.210	0.224	0.197	0.149	0.053	0.035
Rented	0.008	0.027	0.287	0.363	0.164	0.093	0.039	0.013	0.003	0.003

Find the mean number of rooms for both types of housing units. Make probability histograms of these two distributions, using the same scales. How do the means reflect the differences you see in the histograms?

4.83 Standard deviations of numbers of rooms in housing units. Refer to the previous exercise. Which of the two distributions of room counts appears more spread out in the probability histograms? Why? Find the standard deviation for both distributions. The standard deviation provides a numerical measure of spread.

Portfolio analysis. Here are the means, standard deviations, and correlations for the annual returns from three Fidelity mutual funds for the 10 years ending in Febru-

ary 2004. (Means and standard deviations from Fidelity Investments, `fidelity.com`*. Correlations from the* Fidelity Insight *newsletter,* `www.fidelityinsight.com`*. The correlations concern an unspecified period, and other online sources give different correlations, so these should be regarded as approximate at best.) Because there are three random variables, there are three correlations. We use subscripts to show which pair of random variables a correlation refers to.*

W = annual return on 500 Index Fund	$\mu_W = 11.12\%$	$\sigma_W = 17.46\%$
X = annual return on Investment Grade Bond Fund	$\mu_X = 6.46\%$	$\sigma_X = 4.18\%$
Y = annual return on Diversified International Fund	$\mu_Y = 11.10\%$	$\sigma_Y = 15.62\%$

Correlations

$\rho_{WX} = -0.22$ $\rho_{WY} = 0.56$ $\rho_{XY} = -0.12$

The following three exercises make use of these historical data.

4.84 Investing in a mix of U.S. stocks and foreign stocks. Many advisers recommend using roughly 20% foreign stocks to diversify portfolios of U.S. stocks. You see that the 500 Index (U.S. stocks) and Diversified International (foreign stocks) Funds had almost the same mean returns. A portfolio of 80% 500 Index and 20% Diversified International will deliver this mean return with less risk. Verify this by finding the mean and standard deviation of returns on this portfolio.

4.85 The effect of correlation. Diversification works better when the investments in a portfolio have small correlations. To demonstrate this, suppose that returns on 500 Index Fund and Diversified International Fund had the means and standard deviations we have given but were uncorrelated ($\rho_{WY} = 0$). Show that the standard deviation of a portfolio that combines 80% 500 Index with 20% Diversified International is then smaller than your result from the previous exercise. What happens to the mean return if the correlation is 0?

4.86 A portfolio with three investments. Portfolios often contain more than two investments. The rules for means and variances continue to apply, though the arithmetic gets messier. A portfolio containing proportions a of 500 Index Fund, b of Investment Grade Bond Fund, and c of Diversified International Fund has return $R = aW + bX + cY$. Because a, b, and c are the proportions invested in the three funds, $a + b + c = 1$. The mean and variance of the portfolio return R are

$$\begin{aligned} \mu_R &= a\mu_W + b\mu_X + c\mu_Y \\ \sigma_R^2 &= a^2\sigma_W^2 + b^2\sigma_X^2 + c^2\sigma_Y^2 + 2ab\rho_{WX}\sigma_W\sigma_X + 2ac\rho_{WY}\sigma_W\sigma_Y + 2bc\rho_{XY}\sigma_X\sigma_Y \end{aligned}$$

A basic well-diversified portfolio has 60% in 500 Index, 20% in Investment Grade Bond, and 20% in Diversified International. What are the (historical) mean and standard deviation of the annual returns for this portfolio? What does an investor gain by choosing this diversified portfolio over 100% U.S. stocks? What does the investor lose (at least in this time period)?

Section 4.5

4.87 Here is a two-way table of all suicides committed in a recent year by sex of the victim and method used.

	Male	Female
Firearms	15,802	2,367
Poison	3,262	2,233
Hanging	3,822	856
Other	1,571	571
Total	24,457	6,027

(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.

4.88 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 $\{A \text{ and } B\}$) is 0.3. What is the probability of the event $\{A \text{ or } B\}$ that Consolidated will win at least one of the jobs?

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

4.90 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.

4.91 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 $P(B \mid A) = 0.32$. Find the probability that a randomly chosen employed person is a woman holding a managerial or professional position.

4.92 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?

4.93 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 4.94 and 4.95 use this information.

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

4.94 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.

4.95 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 $P(\text{age 18 to 29} \mid \text{married})$. 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.

4.96 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?

4.97 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.)

4.98 In the setting of Exercise 4.96, what percent of sales are made to women? (Write this as a conditional probability.)

4.99 In the setting of Exercise 4.97, 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.)

4.100 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

$$\begin{aligned} P(A) &= 0.9 \\ P(B \mid A) &= 0.36 \\ P(B \mid A^c) &= 0 \end{aligned}$$

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.)

4.101 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?

4.102 In 2000, the Internal Revenue Service received 129,075,000 individual tax returns. Of these, 10,855,000 reported an adjusted gross income of at least \$100,000, and 240,000 reported at least \$1 million. If you know that a randomly chosen return shows an income of \$100,000 or more, what is the conditional probability that the income is at least \$1 million?

4.103 Julie is graduating from college. She has studied biology, chemistry, and computing and hopes to work as a forensic scientist applying her science background to crime investigation. Late one night she thinks about some jobs she has applied for. Let A, B, and C be the events that Julie is offered a job by

A = the Connecticut Office of the Chief Medical Examiner
B = the New Jersey Division of Criminal Justice
C = the federal Disaster Mortuary Operations Response Team

Julie writes down her personal probabilities for being offered these jobs:

$P(A) = 0.6$	$P(B) = 0.4$	$P(C) = 0.2$
$P(A \text{ and } B) = 0.1$	$P(A \text{ and } C) = 0.05$	$P(B \text{ and } C) = 0.05$
$P(A \text{ and } B \text{ and } C) = 0$		

Make a Venn diagram of the events A, B, and C. Mark the probabilities of every intersection involving these events and their complements. Use this diagram for Exercises 4.104 to 4.106.

4.104 What is the probability that Julie is offered at least one of the three jobs?

4.105 What is the probability that Julie is offered both the Connecticut and New Jersey jobs, but not the federal job?

4.106 If Julie is offered the federal job, what is the conditional probability that she is also offered the New Jersey job? If Julie is offered the New Jersey job, what is the conditional probability that she is also offered the federal job?

4.107 Here are the counts (in thousands) of earned degrees in the United States in the 2005–2006 academic year, classified by level and by the sex of the degree recipient:

	Bachelor's	Master's	Professional	Doctorate	Total
Female	?784	276	39	20	1119
Male	?559	197	44	25	?825
Total	1343	473	83	45	1944

(a) If you choose a degree recipient at random, what is the probability that the person you choose is a woman?
(b) What is the conditional probability that you choose a woman, given that the person chosen received a professional degree?
(c) Are the events "choose a woman" and "choose a professional degree recipient" independent? How do you know?

4.108 The previous exercise gives the counts (in thousands) of earned degrees in the United States in the 2005–2006 academic year. Use these data to answer the following questions.
(a) What is the probability that a randomly chosen degree recipient is a man?
(b) What is the conditional probability that the person chosen received a bachelor's degree, given that he is a man?
(c) Use the multiplication rule to find the probability of choosing a male bachelor's degree recipient. Check your result by finding this probability directly from the table of counts.

4.109 The probability that a randomly chosen student at the University of New Harmony is a woman is 0.6. The probability that the student is studying education is 0.15. The conditional probability that the student is a woman, given that the student is studying education, is 0.8. What is the conditional probability that the student is studying education, given that she is a woman?

4.110 The voters in a large city are 40% white, 40% black, and 20% Hispanic. (Hispanics may be of any race in official statistics, but in this case we are speaking of political blocks.) A black mayoral candidate anticipates attracting 30% of the white vote, 90% of the black vote, and 50% of the Hispanic vote. Draw a tree diagram with probabilities for the race (white, black, or Hispanic) and vote (for or against the candidate) of a randomly chosen voter. What percent of the overall vote does the candidate expect to get?

4.111 In the election described in the previous exercise, what percent of the candidate's votes come from black voters? (Write this as a conditional probability.)

4.112 At a self-service gas station, 40% of the customers pump regular gas, 35% pump midgrade, and 25% pump premium gas. Of those who pump regular, 30% pay at least $20. Of those who pump midgrade, 50% pay at least $20. And of those who pump premium, 60% pay at least \$20. What is the probability that the next customer pays at least \$20?

4.113 In the setting of the previous exercise, what percent of customers who pay at least \$20 pump premium? (Write this as a conditional probability.)

4.114 Gender and majors. The probability that a randomly chosen student at the University of New Harmony is a woman is 0.62. The probability that the student is studying education is 0.17. The conditional probability that the student is a woman, given that the student is studying education, is 0.8. What is the conditional probability that the student is studying education, given that she is a woman?

4.115 Cystic fibrosis. Cystic fibrosis is a lung disorder that often results in death. It is inherited but can be inherited only if both parents are carriers of an abnormal gene. In 1989, the CF gene that is abnormal in carriers of cystic fibrosis was identified. The probability that a randomly chosen person of European ancestry carries an abnormal CF gene is 1/25. (The probability is less in other ethnic groups.) The CF20m test detects most but not all harmful mutations of the CF gene. The test is positive for 90% of people who are carriers. It is (ignoring human error) never positive for people who are not carriers. Jason tests positive. What is the probability that he is a carrier?

4.116 Use Bayes's rule. Refer to the previous exercise. Jason knows that he is a carrier of cystic fibrosis. His wife, Julianne, has a brother with cystic fibrosis, which means the probability is 2/3 that she is a carrier. If Julianne is a carrier, each child she has with Jason has probability 1/4 of having cystic fibrosis. If she is not a carrier, her children cannot have the disease. Jason and Julianne have one child, who does not have cystic fibrosis. This information reduces the probability that Julianne is a carrier. Use Bayes's rule to find the conditional probability that Julianne is a carrier, given that she and Jason have one child who does not have cystic fibrosis.

Chapter 4 Review Exercises

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

Hand	Worthless	One pair	Two pairs	Better hands
Probability	0.50	0.42	0.05	?

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?

4.118 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?

4.119 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 μ_X and the standard deviation σ_X without finding the distribution of X. (Start with the distribution of each of the X_i.)

4.120 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`):

	Test result	
	+	−
Antibodies present	0.9985	0.0015
Antibodies absent	0.0060	0.9940

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.)

4.121 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?

4.122 Two wine tasters rate each wine they taste on a scale of 1 to 5. From data on their ratings of a large number of wines, we obtain the following probabilities for both tasters' ratings of a randomly chosen wine:

	Taster 2				
Taster 1	1	2	3	4	5
1	0.03	0.02	0.01	0.00	0.00
2	0.02	0.08	0.05	0.02	0.01
3	0.01	0.05	0.25	0.05	0.01
4	0.00	0.02	0.05	0.20	0.02
5	0.00	0.01	0.01	0.02	0.06

(a) Why is this a legitimate assignment of probabilities to outcomes?
(b) What is the probability that the tasters agree when rating a wine?
(c) What is the probability that Taster 1 rates a wine higher than 3? What is the probability that Taster 2 rates a wine higher than 3?

4.123 In the setting of the previous exercise, Taster 1's rating for a wine is 3. What is the conditional probability that Taster 2's rating is higher than 3?

4.124 Rotter Partners is planning a major investment. The amount of profit X is uncertain but a probabilistic estimate gives the following distribution (in millions of dollars):

Profit	1	1.5	2	4	10
Probability	0.1	0.2	0.4	0.2	0.1

(a) Find the mean profit μ_X and the standard deviation σ_X of the profit.
(b) Rotter Partners owes its source of capital a fee of \$200,000 plus 10% of the profits X. So the firm actually retains

$$Y = 0.9X - 0.2$$

from the investment. Find the mean and standard deviation of Y.

4.125 A grocery store gives its customers cards that may win them a prize when matched with other cards. The back of the card announces the following probabilities of winning various amounts if a customer visits the store 10 times:

Amount	\$1000	\$200	\$50	\$10
Probability	1/10,000	1/1000	1/100	1/20

(a) What is the probability of winning nothing?
(b) What is the mean amount won?
(c) What is the standard deviation of the amount won?

4.126 Profits from an investment. Rotter Partners is planning a major investment. The amount of profit X is uncertain but a probabilistic estimate gives the following distribution (in millions of dollars):

Profit	1	1.5	2	4	10
Probability	0.4	0.2	0.2	0.1	0.1

(a) Find the mean profit μ_X and the standard deviation σ_X of the profit.
(b) Rotter Partners owes its source of capital a fee of \$200,000 plus 10% of the profits X. So the firm actually retains

$$Y = 0.9X - 0.2$$

from the investment. Find the mean and standard deviation of Y.

4.127 Prizes for grocery store customers. A grocery store gives its customers cards that may win them a prize when matched with other cards. The back of the card announces the following probabilities of winning various amounts if a customer visits the store 10 times:

Amount	\$1000	\$250	\$100	\$10
Probability	1/10,000	1/1000	1/100	1/20

(a) What is the probability of winning nothing?
(b) What is the mean amount won?
(c) What is the standard deviation of the amount won?

CHAPTER 5

Section 5.1

5.1 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 $\overline{x} = 51,800$ 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.

5.2 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).

5.3 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 $\overline{x}$ 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 $\overline{x}$'s?

5.4 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?

5.5 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?

5.6 The scores of twelfth-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 twelfth-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 twelfth-graders. What is the probability that their mean score is higher than 300? Higher than 335?

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

5.8 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.)

5.9 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 $\overline{y} - \overline{x}$ between the mean weight gain $\overline{y}$ of the 20 chicks in the high-lysine group and the mean weight gain $\overline{x}$ 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 $\overline{y} - \overline{x}$?
(b) The weight gains are Normally distributed in both populations. What is the distribution of $\overline{x}$? Of $\overline{y}$? What is the distribution of $\overline{y} - \overline{x}$?
(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?

5.10 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 $\overline{x}$ 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 $\overline{y}$ for the 24 students in the B group?
(c) Use the results of (a) and (b), keeping in mind that $\overline{x}$ and $\overline{y}$ are independent, to find the distribution of the difference $\overline{y} - \overline{x}$ 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?

5.11 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?

5.12 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?

5.13 ACT, Inc., the producer of the ACT test of readiness for college work, also produces tests for eighth- and ninth-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 eighth-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 eighth-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)?

5.14 Antonio measures the alcohol content of whiskey for his Chemistry 101 lab. He actually measures the mass of 5 milliliters of whiskey—a chemical calculation then finds the percent alcohol from the mass. The standard deviation of students' measurements of mass is $\sigma = 10$ milligrams (mg). Antonio repeats the measurement 3 times and records the mean $\overline{x}$ of his 3 measurements.
(a) What is the standard deviation $\sigma_{\overline{x}}$ of Antonio's mean result?
(b) How many times must Antonio repeat the measurement to reduce the standard deviation of $\overline{x}$ to 5 mg? Explain to someone who knows no statistics the advantage of reporting the average of several measurements rather than the result of a single measurement.

5.15 An automatic grinding machine in an auto parts plant prepares axles with a target diameter $\mu = 40.125$ millimeters (mm). The machine has some variability, so the standard deviation of the diameters is $\sigma = 0.002$ mm. A sample of 4 axles is inspected each hour for process control purposes, and records are kept of the sample mean diameter. If the process mean is exactly equal to the target value, what will be the mean and standard deviation of the numbers recorded?

5.16 Averages are less variable than individual observations. Suppose that the axle diameters in the previous exercise vary according to a Normal distribution. In that case, the mean $\overline{x}$ of an SRS of axles also has a Normal distribution.
(a) Make a sketch of the Normal curve for a single axle. Add the Normal curve for the mean of an SRS of 4 axles on the same sketch.
(b) What is the probability that the diameter of a single randomly chosen axle differs from the target value by 0.004 mm or more?

(c) What is the probability that the mean diameter of an SRS of 4 axles differs from the target value by 0.004 mm or more?

5.17 Averages of several measurements are less variable than individual measurements. The true mass of the whiskey sample in Exercise 5.14 is 4.6 grams, or 4600 milligrams (mg). Antonio's measurements have the Normal distribution with mean 4600 mg and standard deviation 10 mg. In this case, the mean of his 3 measurements also has a Normal distribution.
(a) Sketch on the same graph the two Normal curves, for individual measurements and for means of 3 measurements.
(b) What is the probability that Antonio misses the true mass by more than 10 mg in either direction if he makes one measurement?
(c) What is the probability that the mean of three independent measurements misses the true mass by more than 10 mg?

5.18 North Carolina State University posts the grade distributions for its courses online. You can find that the distribution of grades in Statistics 101 in the fall 2003 semester was

Grade	A	B	C	D	F
Probability	0.21	0.43	0.30	0.05	0.01

(a) Using the common scale A = 4, B = 3, C = 2, D = 1, F = 0, take X to be the grade of a randomly chosen Statistics 101 student. Use the definitions of the mean and standard deviation for discrete random variables to find the mean μ and the standard deviation σ of grades in this course.
(b) Statistics 101 is a large course. We can take the grades of an SRS of 50 students to be independent of each other. If $\overline{x}$ is the average of these 50 grades, what are the mean and standard deviation of $\overline{x}$?
(c) What is the probability $P(X \geq 3)$ that a randomly chosen Statistics 101 student gets a B or better? What is the approximate probability $P(\overline{x} \geq 3)$ that the grade point average for 50 randomly chosen Statistics 101 students is B or better?

5.19 A $1 bet in a state lottery's Pick 3 game pays $500 if the three-digit number you choose exactly matches the winning number, which is drawn at random. Here is the distribution of the payoff X:

Payoff X	$0	$500
Probability	0.999	0.001

Each day's drawing is independent of other drawings.
(a) What are the mean and standard deviation of X?
(b) Joe buys a Pick 3 ticket every day. What does the law of large numbers say about the average payoff Joe receives from his bets?
(c) What does the central limit theorem say about the distribution of Joe's average payoff after 365 bets in a year?
(d) Joe comes out ahead for the year if his average payoff is greater than $1 (the amount he spent each day on a ticket). What is the probability that Joe ends the year ahead?

5.20 In response to the increasing weight of airline passengers, the Federal Aviation Administration in 2003 told airlines to assume that passengers average 190 pounds in the summer, including clothing and carry-on baggage. But passengers vary: the FAA gave a mean but not a standard deviation. A reasonable standard deviation is 35 pounds. Weights are not Normally distributed, especially when the population includes both men and women, but they are not very non-Normal. A commuter plane carries 19 passengers. What is the approximate probability that the total weight of the passengers exceeds 4000 pounds? (*Hint:* To apply the central limit theorem, restate the problem in terms of the mean weight.)

5.21 A hole in an engine block is 2.5 centimeters (cm) in diameter. Shafts manufactured to go through this hole must have 0.025 cm clearance for unforced fit. That is, shaft diameter cannot exceed 2.475 cm. The shafts vary in diameter according to the Normal distribution with mean 2.45 cm and standard deviation 0.01 cm.
(a) What percent of shafts will fit into the hole?
(b) Redo the problem assuming that the hole diameter also varies, independently of the shaft diameter, following the Normal distribution with mean 2.5 cm and standard deviation 0.01 cm. You must find the probability that the hole diameter exceeds the shaft diameter by at least 0.025 cm.

5.22 "Durable press" cotton fabrics are treated to improve their recovery from wrinkles after washing. Unfortunately, the treatment also reduces the strength of the fabric. The breaking strength of untreated fabric is Normally distributed with mean 58 pounds and standard deviation 2.3 pounds. The same type of fabric after treatment has Normally distributed breaking strength with mean 30 pounds and standard deviation 1.6 pounds. A clothing manufacturer tests 5 specimens of each fabric. All 10 strength measurements are independent.
(a) What is the probability that the mean breaking strength of the 5 untreated specimens exceeds 50 pounds?
(b) What is the probability that the mean breaking strength of the 5 untreated specimens is at least 25 pounds greater than the mean strength of the 5 treated specimens?

5.23 Many companies place advertisements to improve the image of their brand rather than to promote specific products. In a randomized comparative experiment, business students read ads that cited either the *Wall Street Journal* or the *National Enquirer* for important facts about a fictitious company. The students then rated the trustworthiness of the source on a 7-point scale. Suppose that in the population of all students scores for the *Journal* have mean 4.8 and standard deviation 1.5, while scores for the *Enquirer* have mean 2.4 and standard deviation 1.6.
(a) There are 30 students in each group. Although individual scores are discrete, the mean score for a group of 30 will be close to Normal. Why?
(b) What are the means and standard deviations of the sample mean scores $\overline{y}$ for the *Journal* group and $\overline{x}$ for the *Enquirer* group?
(c) We can take all 60 scores to be independent because students are not told each other's scores. What is the distribution of the difference $\overline{y} - \overline{x}$ between the mean scores in the two groups?
(d) Find $P(\overline{y} - \overline{x} \geq 1)$.

5.24 Linda invests her money in a portfolio that consists of 80% Fidelity 500 Index Fund and 20% Fidelity Diversified International Fund. Suppose that in the long run the annual real return X on the 500 Index Fund has mean 9% and standard deviation 20%, the annual real return Y on the Diversified International Fund has mean 10% and standard deviation 17%, and the correlation between X and Y is 0.6.
(a) The return on Linda's portfolio is $R = 0.8X + 0.2Y$. What are the mean and standard deviation of R?
(b) The distribution of returns is typically roughly symmetric but with more extreme high and low observations than a Normal distribution. The average return over a number of years, however, is close to Normal. If Linda holds her portfolio for 20 years, what is the approximate probability that her average return is less than 5%?
(c) The calculation you just made is not overly helpful, because Linda isn't really concerned about the mean return $\overline{R}$. To see why, suppose that her portfolio returns 12% this year and 6% next year. The mean return for the two years is 9%. If Linda starts with $1000, how much does she have at the end of the first year? At the end of the second year? How does this amount compare with what she would have if both years had the mean return, 9%? Over 20 years, there may be a large difference between the ordinary mean $\overline{R}$ and the *geometric mean,* which reflects the fact that returns in successive years multiply rather than add.

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

5.26 A particle moves along the line in a random walk. That is, the particle starts at the origin (position 0) and moves either right or left in independent steps of length 1. If the particle moves to the right with probability 3/4, its movement at the ith step is a random variable X_i with distribution

$$\begin{aligned} P(X_i = 1) &= 0.75 \\ P(X_i = -1) &= 0.25 \end{aligned}$$

The position of the particle after k steps is the sum of these random movements,

$$Y = X_1 + X_2 + \cdots + X_k$$

Use the central limit theorem to find the approximate probability that the position of the particle after 500 steps is at least 200 to the right.

5.27 An automatic grinding machine in an auto parts plant prepares axles with a target diameter $\mu = 40.135$ millimeters (mm). The machine has some variability, so the standard deviation of the diameters is $\sigma = 0.003$ mm. A sample of 4 axles is inspected each hour for process control purposes, and records are kept of the sample mean diameter. If the process mean is exactly equal to the target value, what will be the mean and standard deviation of the numbers recorded?

5.28 Averages are less variable than individual observations. Suppose that the axle diameters in the previous exercise vary according to a Normal distribution. In that case, the mean $\bar{x}$ of an SRS of axles also has a Normal distribution.
(a) Make a sketch of the Normal curve for a single axle. Add the Normal curve for the mean of an SRS of 4 axles on the same sketch.
(b) What is the probability that the diameter of a single randomly chosen axle differs from the target value by 0.006 mm or more?
(c) What is the probability that the mean diameter of an SRS of 4 axles differs from the target value by 0.006 mm or more?

5.29 The number of lightning strikes on a square kilometer of open ground in a year has mean 6 and standard deviation 2.4. (These values are typical of much of the United States.) Counts of strikes on separate areas are independent. The National Lightning Detection Network uses automatic sensors to watch for lightning in an area of 10 square kilometers.
(a) What are the mean and standard deviation of the total number of lightning strikes observed?
(b) What are the mean and standard deviation of the mean number of strikes per square kilometer?

5.30 The distribution of annual returns on common stocks is roughly symmetric, but extreme observations are more frequent than in a Normal distribution. Because the distribution is not strongly non-Normal, the mean return over even a moderate number of years is close to Normal. Annual real returns on the Standard & Poor's 500 stock Index over the period 1871 to 2004 have varied with mean 9.2% and standard deviation 20.6%. Andrew plans to retire in 45 years and is considering investing in stocks. What is the probability (assuming that the past pattern of variation continues) that the mean annual return on common stocks over the next 45 years will exceed 15%? What is the probability that the mean return will be less than 5%?

5.31 A hole in an engine block is 2.5 centimeters (cm) in diameter. Shafts manufactured to go through this hole must have 0.024 cm clearance for unforced fit. That is, shaft diameter cannot exceed 2.474 cm. The shafts vary in diameter according to the Normal distribution with mean 2.45 cm and standard deviation 0.01 cm.
(a) What percent of shafts will fit into the hole?
(b) Redo the problem assuming that the hole diameter also varies, independently of the shaft diameter, following the Normal distribution with mean 2.5 cm and standard deviation 0.01 cm. You must find the probability that the hole diameter exceeds the shaft diameter by at least 0.025 cm.

Section 5.2

5.32 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.

5.33 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.

5.34 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?

5.35 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.

5.36 (a) What is the mean number of Hispanics on randomly chosen committees of 15 workers in Exercise 5.34?
(b) What is the standard deviation σ of the count X of Hispanic members?
(c) Suppose that 10% of the factory workers were Hispanic. Then $p = 0.1$. What is σ in this case? What is σ 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?

5.37 (a) Find the mean number of graduates out of 20 players in the setting of Exercise 5.35 if the university's claim is true.
(b) Find the standard deviation σ 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 σ of the count of graduates? If $p = 0.99$, what is σ? 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?

5.38 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?

5.39 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.

5.40 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.

5.41 "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?

5.42 "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?

5.43 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 adviser 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.

5.44 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?

5.45 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?

5.46 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?

5.47 In each situation below, is it reasonable to use a binomial distribution for the random variable X? Give reasons for your answer in each case. If a binomial distribution applies, give the values of n and p.
(a) Most calls made at random by sample surveys don't succeed in talking with a live person. Of calls to New York City, only 1/12 succeed. A survey calls 500 randomly selected numbers in New York City. X is the number that reach a live person.
(b) At peak periods, 25% of attempted log-ins to an Internet service provider fail. Log-in attempts are independent and each has the same probability of failing. Darci logs in repeatedly until she succeeds. X is the number of the login attempt that finally succeeds.
(c) On a bright October day, Canada geese arrive to foul the pond at an apartment complex at the average rate of 12 geese per hour; X is the number of geese that arrive in the next three hours.

5.48 In each situation below, is it reasonable to use a binomial distribution for the random variable X? Give reasons for your answer in each case.
(a) An auto manufacturer chooses one car from each hour's production for a detailed quality inspection. One variable recorded is the count X of finish defects (dimples, ripples, etc.) in the car's paint.
(b) The pool of potential jurors for a murder case contains 100 persons chosen at random from the adult residents of a large city. Each person in the pool is asked whether he or she opposes the death penalty; X is the number who say "Yes."
(c) Joe buys a ticket in his state's Pick 3 lottery game every week; X is the number of times in a year that he wins a prize.

5.49 Typographic errors in a text are either nonword errors (as when "the" is typed as "teh") or word errors that result in a real but incorrect word. Spell-checking software will catch nonword errors but not word errors. Human proofreaders catch 70% of word errors. You ask a fellow student to proofread an essay in which you have deliberately made 20 word errors.
(a) If the student matches the usual 70% rate, what is the distribution of the number of errors caught? What is the distribution of the number of errors missed?
(b) Missing 9 or more out of 20 errors seems a poor performance. What is the probability that a proofreader who catches 70% of word errors misses 9 or more out of 20?

5.50 What kinds of Web sites do males aged 18 to 34 visit most often? Pornographic sites take first place, but about 50% of male Internet users in this age group visit an auction site such as eBay at least once a month. Interview a random sample of 12 male Internet users aged 18 to 34.
(a) What is the distribution of the number who have visited an online auction site in the past month?
(b) What is the probability that at least 8 of the 12 have visited an auction site in the past month?

5.51 Return to the proofreading setting of Exercise 5.49.
(a) What is the mean number of errors caught? What is the mean number of errors missed? You see that these two means must add to 20, the total number of errors.
(b) What is the standard deviation σ of the number of errors caught?
(c) Suppose that a proofreader catches 90% of word errors, so that $p = 0.9$. What is σ in this case? What is σ if $p = 0.99$? What happens to the standard deviation of a binomial distribution as the probability of a success gets close to 1?

5.52 Suppose that 50% of male Internet users aged 18 to 34 have visited an auction site at least once in the past month.
(a) If you interview 12 at random, what is the mean of the count X who have visited an auction site? What is the mean of the proportion $\hat{p}$ in your sample who have visited an auction site?
(b) Repeat the calculations in (a) for samples of size 120 and 1200. What happens to the mean count of successes as the sample size increases? What happens to the mean proportion of successes?

5.53 In the proofreading setting of Exercise 5.49, what is the smallest number of misses m with $P(X \geq m)$ no larger than 0.05? You might consider m or more misses as evidence that a proofreader actually catches fewer than 70% of word errors.

5.54 Some of the methods in this section are approximations rather than exact probability results. We have given rules of thumb for safe use of these approximations.
(a) You are interested in attitudes toward drinking among the 75 members of a fraternity. You choose 25 members at random to interview. One question is "Have you had five or more drinks at one time during the last week?" Suppose that in fact 20% of the 75 members would say "Yes." Explain why you *cannot* safely use the

$B(25, 0.2)$ distribution for the count X in your sample who say "Yes."
(b) The National AIDS Behavioral Surveys found that 0.2% (that's 0.002 as a decimal fraction) of adult heterosexuals had both received a blood transfusion and had a sexual partner from a group at high risk of AIDS. Suppose that this national proportion holds for your region. Explain why you *cannot* safely use the Normal approximation for the sample proportion who fall in this group when you interview an SRS of 500 adults.

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

5.56 The *Probability* applet on the text CD and Web site simulates tosses of a coin. You can choose the number of tosses n and the probability p of a head. You can therefore use the applet to simulate binomial random variables.

In an audit of sales records, the count of misclassified sales records has the binomial distribution with $n = 15$ and $p = 0.08$. Set these values for the number of tosses and probability of heads in the applet. Table C shows that the probability of getting a sample with exactly 1 misclassified record is 0.3734. This is the long-run proportion of samples with exactly 1 bad record. Click "Toss" and "Reset" repeatedly to simulate 20 samples. Record the number of bad records (the count of heads) in each of the 20 samples. What proportion of the 20 samples had exactly 1 bad record? Remember that probability tells us only what happens in the long run.

5.57 In 1998, Mark McGwire of the St. Louis Cardinals hit 70 home runs, a new major league record. Was this feat as surprising as most of us thought? In the three seasons before 1998, McGwire hit a home run in 11.6% of his times at bat. He went to bat 509 times in 1998. If he continues his past performance, McGwire's home run count in 509 times at bat has approximately the binomial distribution with $n = 509$ and $p = 0.116$.
(a) What is the mean number of home runs McGwire will hit in 509 times at bat?
(b) What is the probability that he hits 70 or more home runs?
(c) In 2001, Barry Bonds of the San Francisco Giants hit 73 home runs, breaking McGwire's record. This was surprising. In the three previous seasons, Bonds hit a home run in 8.65% of his times at bat. He batted 476 times in 2001. Considering his home run count as a binomial random variable with $n = 476$ and $p = 0.0865$, what is the probability of 73 or more home runs?

5.58 A believer in the "random walk" theory of the behavior of stock prices thinks that an index of stock prices has probability 0.65 of increasing in any year. Moreover, the change in the index in any given year is not influenced by whether it rose or fell in earlier years. Let X be the number of years among the next 6 years in which the index rises.
(a) What are n and p in the binomial distribution of X?
(b) Give the possible values that X can take and the probability of each value. Draw a probability histogram for the distribution of X.

(c) Find the mean of the number X of years in which the stock price index rises and mark the mean on your probability histogram.
(d) Find the standard deviation of X. What is the probability that X takes a value within one standard deviation of its mean?

5.59 "What do you think is the ideal number of children for a family to have?" A Gallup Poll asked this question of 1006 randomly chosen adults. Almost half (49%) thought two children was ideal. Suppose that $p = 0.49$ is exactly true for the population of all adults. Gallup announced a margin of error of ± 3 percentage points for this poll. What is the probability that the sample proportion $\hat{p}$ for an SRS of size $n = 1006$ falls between 0.46 and 0.52? You see that it is likely, but not certain, that polls like this give results that are correct within their margin of error.

5.60 Return to the Gallup Poll setting of the previous exercise. We are supposing that the proportion of all adults who think that two children is ideal is $p = 0.49$. What is the probability that a sample proportion $\hat{p}$ falls between 0.46 and 0.52 (that is, within ± 3 percentage points of the true p) if the sample is an SRS of size $n = 250$? Of size $n = 4000$? Combine these results with your work in Exercise 5.59 to make a general statement about the effect of larger samples in a sample survey.

5.61 The changing probabilities you found in the previous two exercises are due to the fact that the standard deviation of the sample proportion $\hat{p}$ gets smaller as the sample size n increases. If the population proportion is $p = 0.49$, how large a sample is needed to reduce the standard deviation of $\hat{p}$ to $\sigma_{\hat{p}} = 0.005$? (The 68–95–99.7 rule then says that about 95% of all samples will have $\hat{p}$ within 0.01 of the true p.)

5.62 A Gallup Poll finds that 30% of adults visited a casino in the past 12 months, and that 6% bet on college sports. These results come from a random sample of 1011 adults. For an SRS of size $n = 1011$:
(a) What is the probability that the sample proportion $\hat{p}$ is between 0.29 and 0.31 if the population proportion is $p = 0.30$?
(b) What is the probability that the sample proportion $\hat{p}$ is between 0.05 and 0.07 if the population proportion is $p = 0.06$?
(c) How does the probability that $\hat{p}$ falls within ± 0.01 of the true p change as p gets closer to 0?

5.63 The Harvard College Alcohol Study finds that 67% of college students support efforts to "crack down on underage drinking." The study took a sample of almost 15,000 students, so the population proportion who support a crackdown is very close to $p = 0.67$. The administration of your college surveys an SRS of 100 students and finds that 62 support a crackdown on underage drinking.
(a) What is the sample proportion who support a crackdown on underage drinking?
(b) If in fact the proportion of all students on your campus who support a crackdown is the same as the national 67%, what is the probability that the proportion in an SRS of 100 students is as small or smaller than the result of the administration's sample?
(c) A writer in the student paper says that support for a crackdown is lower on

your campus than nationally. Write a short letter to the editor explaining why the survey does not support this conclusion.

5.64 A selective college would like to have an entering class of 1200 students. Because not all students who are offered admission accept, the college admits more than 1200 students. Past experience shows that about 70% of the students admitted will accept. The college decides to admit 1500 students. Assuming that students make their decisions independently, the number who accept has the $B(1500, 0.7)$ distribution. If this number is less than 1200, the college will admit students from its waiting list.
(a) What are the mean and the standard deviation of the number X of students who accept?
(b) Use the Normal approximation to find the probability that at least 1000 students accept.
(c) The college does not want more than 1200 students. What is the probability that more than 1200 will accept?
(d) If the college decides to increase the number of admission offers to 1700, what is the probability that more than 1200 will accept?

5.65 The mailing list of an agency that markets scuba-diving trips to the Florida Keys contains 70% males and 30% females. The agency calls 30 people chosen at random from its list.
(a) What is the probability that 20 of the 30 are men? (Use the binomial probability formula.)
(b) What is the probability that the first woman is reached on the fourth call? (That is, the first 4 calls give MMMF.)

5.66 One way of checking the effect of undercoverage, nonresponse, and other sources of error in a sample survey is to compare the sample with known demographic facts about the population. The 2000 census found that 23,772,494 of the 209,128,094 adults (age 18 and over) in the United States called themselves "Black or African American."
(a) What is the population proportion p of blacks among American adults?
(b) An opinion poll chooses 1500 adults at random. What is the mean number of blacks in such samples? (Explain the reasoning behind your calculation.)
(c) Use a Normal approximation to find the probability that such a sample will contain 170 or fewer blacks. Be sure to check that you can safely use the approximation.

5.67 Here is a simple probability model for multiple-choice tests. Suppose that each student has probability p of correctly answering a question chosen at random from a universe of possible questions. (A strong student has a higher p than a weak student.) The correctness of an answer to a question is independent of the correctness of answers to other questions. Jodi is a good student for whom $p = 0.75$.
(a) Use the Normal approximation to find the probability that Jodi scores 70% or lower on a 100-question test.
(b) If the test contains 250 questions, what is the probability that Jodi will score 70% or lower?
(c) How many questions must the test contain in order to reduce the standard

deviation of Jodi's proportion of correct answers to half its value for a 100-item test?
(d) Laura is a weaker student for whom $p = 0.6$. Does the answer you gave in (c) for the standard deviation of Jodi's score apply to Laura's standard deviation also?

5.68 What kinds of Web sites do males aged 18 to 34 visit most often? Pornographic sites take first place, but about 50% of male Internet users in this age group visit an auction site such as eBay at least once a month (John Schwartz, "Leisure pursuits of today's young men," *New York Times,* March 29, 2004). Interview a random sample of 15 male Internet users aged 18 to 34.
(a) What is the distribution of the number who have visited an online auction site in the past month?
(b) What is the probability that at least 8 of the 15 have visited an auction site in the past month?

5.69 Suppose that 50% of male Internet users aged 18 to 34 have visited an auction site at least once in the past month.
(a) If you interview 15 at random, what is the mean of the count X who have visited an auction site? What is the mean of the proportion $\hat{p}$ in your sample who have visited an auction site?
(b) Repeat the calculations in (a) for samples of size 150 and 1500. What happens to the mean count of successes as the sample size increases? What happens to the mean proportion of successes?

Chapter 5 Review Exercises

5.70 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.)

5.71 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.

5.72 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?

5.73 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.

5.74 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 $\overline{x}$ is greater than L is only 0.01?

5.75 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?

5.76 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 $\overline{x}$ 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 non-Normal?

5.77 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)?

5.78 The probability that a randomly chosen driver will be involved in an accident in the next year is about 0.2. This is based on the proportion of millions of drivers who have accidents. "Accident" includes things like crumpling a fender in your own driveway, not just highway accidents. Carlos, David, Jermaine, Ramon, Scott, and Sean are college students who live together in an off-campus apartment. Last year, 3 of the 6 had accidents. What is the probability that 3 or more of 6 randomly chosen drivers have an accident in the same year? Why does your calculation not apply to drivers like the 6 students?

5.79 Consider that the total SAT scores of high school seniors in a recent year had mean $\mu = 1026$ and standard deviation $\sigma = 209$. The distribution of SAT scores is roughly Normal.
(a) Ramon scored 1100. If scores have a Normal distribution, what percentile of the distribution is this?
(b) Now consider the mean $\overline{x}$ of the scores of 70 randomly chosen students. If $\overline{x} = 1100$, what percentile of the sampling distribution of $\overline{x}$ is this?
(c) Which of your calculations, (a) or (b), is less accurate because SAT scores do not have an exactly Normal distribution? Explain your answer.

5.80 Although cities encourage carpooling to reduce traffic congestion, most vehicles carry only one person. For example, 70% of vehicles on the roads in the Minneapolis-St. Paul metropolitan area are occupied by just the driver.
(a) If you choose 10 vehicles at random, what is the probability that more than half (that is, 6 or more) carry just one person?
(b) If you choose 100 vehicles at random, what is the probability that more than half (that is, 51 or more) carry just one person?

5.81 The Census Bureau says that the 10 most common last names in the United States are (in order) Smith, Johnson, Williams, Jones, Brown, Davis, Miller, Wilson, Moore, and Taylor. These names account for 5.6% of all U.S. residents. Out of curiosity, you look at the authors of the textbooks for your current courses. There are 9 authors in all. Would you be surprised if none of the names of these authors were among the 10 most common? Give a probability to support your answer and explain the reasoning behind your calculation.

5.82 It is a striking fact that the first digits of numbers in legitimate records often follow a distribution known as Benford's law. Here it is:

First digit	1	2	3	4	5	6	7	8	9
Proportion	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

Fake records usually have fewer first digits 1, 2, and 3. What is the approximate probability, if Benford's law holds, that among 1200 randomly chosen invoices there are no more than 680 in amounts with first digit 1, 2, or 3?

5.83 According to genetic theory, the blossom color in the second generation of a certain cross of sweet peas should be red or white in a 3:1 ratio. That is, each plant has probability 3/4 of having red blossoms, and the blossom colors of separate plants are independent.
(a) What is the probability that exactly 6 out of 8 of these plants have red blossoms?
(b) What is the mean number of red-blossomed plants when 80 plants of this type are grown from seeds?
(c) What is the probability of obtaining at least 50 red-blossomed plants when 80 plants are grown from seeds?

5.84 The weight of the eggs produced by a certain breed of hen is Normally distributed with mean 65 grams (g) and standard deviation 5 g. If cartons of such eggs can be considered to be SRSs of size 12 from the population of all eggs, what is the probability that the weight of a carton falls between 750 g and 825 g?

5.85 Suppose (as is roughly true) that 20% of all Internet users have posted photos online. A sample survey interviews an SRS of 1555 Internet users.
(a) What is the actual distribution of the number X in the sample who have posted photos online?
(b) What is the probability that 300 or fewer of the people in the sample have posted photos online? (Use either software or a suitable approximation.)

5.86 A study of rush-hour traffic in San Francisco records the number of people in each car entering a freeway at a suburban interchange. Suppose that this number X has mean 1.5 and standard deviation 0.75 in the population of all cars that enter at this interchange during rush hours.
(a) Does the count X have a binomial distribution? Why or why not?
(b) Could the exact distribution of X be Normal? Why or why not?
(c) Traffic engineers estimate that the capacity of the interchange is 700 cars per hour. According to the central limit theorem, what is the approximate distribution of the mean number of persons $\overline{x}$ in 700 randomly selected cars at this interchange?
(d) The count of people in 700 cars is $700\overline{x}$. Use your result from (c) to give an approximate distribution for the count. What is the probability that 700 cars will carry more than 1075 people?

5.87 A sample survey interviews an SRS of 267 college women. Suppose (as is roughly true) that 70% of all college women have been on a diet within the last 12 months. What is the probability that 75% or more of the women in the sample have been on a diet?

5.88 A machine fastens plastic screw-on caps onto containers of motor oil. If the machine applies more torque than the cap can withstand, the cap will break. Both

the torque applied and the strength of the caps vary. The capping machine torque has the Normal distribution with mean 7 inch-pounds and standard deviation 0.9 inch-pounds. The cap strength (the torque that would break the cap) has the Normal distribution with mean 10 inch-pounds and standard deviation 1.2 inch-pounds.
(a) Explain why it is reasonable to assume that the cap strength and the torque applied by the machine are independent.
(b) What is the probability that a cap will break while being fastened by the capping machine?

5.89 The unique colors of the cashmere sweaters your firm makes result from heating undyed yarn in a kettle with a dye liquor. The pH (acidity) of the liquor is critical for regulating dye uptake and hence the final color. There are 5 kettles, all of which receive dye liquor from a common source. Past data show that pH varies according to a Normal distribution with $\mu = 4.22$ and $\sigma = 0.127$. You use statistical process control to check the stability of the process. Twice each day, the pH of the liquor in each kettle is measured, each time giving a sample of size 5. The mean pH $\overline{x}$ is compared with "control limits" given by the 99.7 part of the 68–95–99.7 rule for Normal distributions, namely $\mu_{\overline{x}} \pm 3\sigma_{\overline{x}}$. What are the numerical values of these control limits for $\overline{x}$?

5.90 Suppose (as is roughly true) that 88% of college men and 82% of college women were employed last summer. A sample survey interviews SRSs of 500 college men and 500 college women. The two samples are of course independent.
(a) What is the approximate distribution of the proportion $\hat{p}_F$ of women who worked last summer? What is the approximate distribution of the proportion $\hat{p}_M$ of men who worked?
(b) The survey wants to compare men and women. What is the approximate distribution of the difference in the proportions who worked, $\hat{p}_M - \hat{p}_F$? Explain the reasoning behind your answer.
(c) What is the probability that in the sample a higher proportion of women than men worked last summer?

5.91 The probability that a randomly chosen driver will be involved in an accident in the next year is about 0.2. This is based on the proportion of millions of drivers who have accidents. "Accident" includes things like crumpling a fender in your own driveway, not just highway accidents. Carlos, David, Jermaine, Ramon, and Scott are college students who live together in an off-campus apartment. Last year, 3 of the 5 had accidents. What is the probability that 3 or more of 5 randomly chosen drivers have an accident in the same year? Why does your calculation not apply to drivers like the 5 students?

5.92 Suppose that the total SAT scores of high school seniors in a recent year had mean $\mu = 1026$ and standard deviation $\sigma = 209$. The distribution of SAT scores is roughly Normal.
(a) Julie scored 1110. If scores have a Normal distribution, what percentile of the distribution is this?
(b) Now consider the mean $\bar{x}$ of the scores of 80 randomly chosen students. If $\bar{x} = 1110$, what percentile of the sampling distribution of $\bar{x}$ is this?

(c) Which of your calculations, (a) or (b), is less accurate because SAT scores do not have an exactly Normal distribution? Explain your answer.

5.93 Serving in a bomber crew in World War II was dangerous. The British estimated that the probability of an aircraft loss due to enemy action was 1/20 for each mission. A tour of duty for British airmen in Bomber Command was 30 missions. What is the probability that an airman would complete a tour of duty without being on an aircraft lost from enemy action?

5.94 A sample survey interviews an SRS of 280 college women. Suppose (as is roughly true) that 70% of all college women have been on a diet within the last 12 months. What is the probability that 75% or more of the women in the sample have been on a diet?

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 $\overline{x} = 2.36$. 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 σ. 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 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.

6.9 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. Quarterman, the margin of error is 2.8 percent, with a confidence level of 95 percent.*

(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.

6.10 The mean amount μ 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 $\overline{x}$?
(b) The 68–95–99.7 rule says that the probability is about 0.95 that $\overline{x}$ is within ______ of the population mean μ. Fill in the blank.
(c) About 95% of all samples will capture the true mean of all of the invoices in the interval $\overline{x}$ plus or minus ______. Fill in the blank.

6.11 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:

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

Suppose that the standard deviation of the population is known to be $\sigma = 4.5$ kg.
(a) What is $\sigma_{\overline{x}}$, the standard deviation of $\overline{x}$?
(b) Give a 95% confidence interval for μ, 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?

6.12 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.

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

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	

Assume that $\sigma = 10$ bushels per acre.
(a) Find the 90% confidence interval for the mean yield μ 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?

6.14 Suppose that the crop researchers in the previous exercise obtained the same value of $\overline{x}$ from a sample of 50 plots rather than 15.
(a) Compute the 95% confidence interval for the mean yield μ.
(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.)

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

6.16 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?

6.17 How large a sample of Julie's potassium levels in the previous exercise would be needed to estimate her mean μ within ± 0.06 with 95% confidence?

6.18 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.)

6.19 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?

6.20 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%±3% means.
(c) This confidence interval has the same form we have met earlier:

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

(Actually σ is estimated from the data, but we ignore this for now.) What is the standard deviation σ_{estimate} of the estimated percent?
(d) Does the announced margin of error include errors due to practical problems such as undercoverage and nonresponse?

6.21 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 μ 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.

6.22 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?

6.23 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.

6.24 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.

6.25 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?

6.26 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 $\overline{x} = \$9740$ 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 σ. The station calculates the 95% confidence interval for the mean pay μ that all citizens would propose for council members to be $9669 to $9811. Is this result trustworthy? Explain your answer.

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

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

6.29 Suppose that the sample mean is 50 and the standard deviation is assumed to be 5. Make a diagram that illustrates the effect of sample size on the width of a 95% interval. Use the following sample sizes: 10, 20, 40, and 100. Summarize what the diagram shows.

6.30 A study with 25 observations gave a mean of 70. Assume that the standard deviation is 15. Make a diagram that illustrates the effect of the confidence level on the width of the interval. Use 80%, 90%, 95%, and 99%. Summarize what the diagram shows.

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

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

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

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

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

15.8	13.6	15.6	19.1	22.4	15.6	22.5	17.2	19.4	22.6
19.4	18.0	14.6	18.7	21.0	14.8	22.6	21.5	14.3	20.9

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

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

21.0	19.0	18.7	39.2	45.8	19.8	48.4	21.0	29.1	35.7
31.6	49.0	16.0	34.6	36.3	19.0	43.3	37.5	16.5	34.5

Assume that the standard deviation is 10.3 mph. Estimate the mean speed at which this vehicle was driven with a margin of error for 95% confidence.

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

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

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

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

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

6.40 How large a sample of one-bedroom apartments in Exercise 6.32 would be needed to estimate the mean μ within $\pm$\$20 with 90% confidence?

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

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

6.43 A survey of users of the Internet found that males outnumbered females by nearly 2 to 1 (P. H. Lewis, "Technology" column, *New York Times*, May 29, 1995). 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. Quarterman, the margin of error is 2.8 percent, with a confidence level of 95 percent.*

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

6.44 The mean number of calories consumed by women in the United States who are 19 to 30 years of age is $\mu = 1791$ calories per day. The standard deviation is 31 calories (*Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids (Macronutrients),* Food and Nutrition Board, Institute of Medicine, 2002). You will study a sample of 200 women in this age range, and one of the variables you will collect is calories consumed per day.
(a) What is the standard deviation of the sample mean $\bar{x}$?
(b) The 68–95–99.7 rule says that the probability is about 0.95 that $\overline{x}$ is within ______ calories of the population mean μ. Fill in the blank.
(c) About 95% of all samples will capture the true mean of calories consumed per day in the interval $\overline{x}$ plus or minus ______ calories. Fill in the blank.

6.45 A Gallup Poll asked working adults about their job satisfaction (Chris Chambers, "Americans skeptical about mergers of big companies," Gallup News Service, November 1, 2000). One question was "All in all, which best describes how you feel about your job?" The possible answers were "love job," "like job," "dislike job," and "hate job." Fifty-nine percent of the sample responded that they liked their job. Material provided with the results of the poll noted:

> *Results are based on telephone interviews with 1,001 national adults, aged 18 and older, conducted Aug. 8–11, 2005. For results based on the total sample of national adults, one can say with 95% confidence that the maximum margin of sampling error is 3 percentage points.*

The Gallup Poll uses a complex multistage sample design, but the sample percent has approximately a Normal sampling distribution.
(a) The announced poll result was 59% ± 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 59%±3% means.
(c) This confidence interval has the same form we have met earlier:

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

What is the standard deviation σ_{estimate} of the estimated percent?
(d) Does the announced margin of error include errors due to practical problems such as undercoverage and nonresponse?

Section 6.2

6.46 Each of the following situations requires a significance test about a population mean μ. State the appropriate null hypothesis H_0 and alternative hypothesis H_a in each case.
(a) The mean area of the several thousand apartments in a new development is advertised to be 1250 square feet. A tenant group thinks that the apartments are smaller than advertised. They hire an engineer to measure a sample of apartments to test their suspicion.
(b) Larry's car averages 32 miles per gallon on the highway. He now switches to a new motor oil that is advertised as increasing gas mileage. After driving 3000 highway miles with the new oil, he wants to determine if his gas mileage actually has increased.
(c) The diameter of a spindle in a small motor is supposed to be 5 millimeters. If the spindle is either too small or too large, the motor will not perform properly. The manufacturer measures the diameter in a sample of motors to determine whether the mean diameter has moved away from the target.

6.47 In each of the following situations, a significance test for a population mean μ is called for. State the null hypothesis H_0 and the alternative hypothesis H_a in each case.
(a) Experiments on learning in animals sometimes measure how long it takes a mouse to find its way through a maze. The mean time is 18 seconds for one particular maze. A researcher thinks that a loud noise will cause the mice to complete the maze faster. She measures how long each of 10 mice takes with a noise as stimulus.

(b) The examinations in a large accounting class are scaled after grading so that the mean score is 50. A self-confident teaching assistant thinks that his students have a higher mean score than the class as a whole. His students this semester can be considered a sample from the population of all students he might teach, so he compares their mean score with 50.
(c) A university gives credit in French language courses to students who pass a placement test. The language department wants to know if students who get credit in this way differ in their understanding of spoken French from students who actually take the French courses. Some faculty think the students who test out of the courses are better, but others argue that they are weaker in oral comprehension. Experience has shown that the mean score of students in the courses on a standard listening test is 24. The language department gives the same listening test to a sample of 40 students who passed the credit examination to see if their performance is different.

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

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

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

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

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

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

6.52 Each of the following situations requires a significance test about a population mean μ. 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 two high, your product will not look good. You measure the wattage of a random sample of 20 bulbs.

6.53 In each of the following situations, a significance test for a population mean μ is called for. State the null hypothesis H_0 and the alternative hypothesis H_a in each case.
(a) Experiments on learning in animals sometimes measure how long it takes a mouse to find its way through a maze. The mean time is 18 seconds for one particular maze. A researcher thinks that a loud noise will cause the mice to complete the maze faster. She measures how long each of 10 mice takes with a noise as stimulus.
(b) The examinations in a large history class are scaled after grading so that the mean score is 50. A self-confident teaching assistant thinks that his students have a higher mean score than the class as a whole. His students this semester can be considered a sample from the population of all students he might teach, so he compares their mean score with 50.
(c) The Census Bureau reports that households spend an average of 31% of their total spending on housing. A homebuilders association in Cleveland wonders if the national finding applies in their area. They interview a sample of 40 households in the Cleveland metropolitan area to learn what percent of their spending goes toward housing.

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

(b) An education researcher randomly divides sixth-grade students into two groups for physical education class. He teaches both groups basketball skills, using the same methods of instruction in both classes. He encourages Group A with compliments and other positive behavior but acts cool and neutral toward Group B. He hopes to show that positive teacher attitudes result in a higher mean score on a test of basketball skills than do neutral attitudes.
(c) An economist believes that among employed young adults there is a positive correlation between income and the percent of disposable income that is saved. To test this, she gathers income and savings data from a sample of employed persons in her city aged 25 to 34.

6.55 A test of the null hypothesis H_0: $\mu = \mu_0$ gives test statistic $z = 1.8$.
(a) What is the P-value if the alternative is H_a: $\mu > \mu_0$?
(b) What is the P-value if the alternative is H_a: $\mu < \mu_0$?
(c) What is the P-value if the alternative is H_a: $\mu \neq \mu_0$?

6.56 The P-value for a two-sided test of the null hypothesis H_0: $\mu = 10$ is 0.06.
(a) Does the 95% confidence interval include the value 10? Why?
(b) Does the 90% confidence interval include the value 10? Why?

6.57 A 95% confidence interval for a population mean is $(28, 35)$.
(a) Can you reject the null hypothesis that $\mu = 34$ at the 5% significance level? Why?
(b) Can you reject the null hypothesis that $\mu = 36$ at the 5% significance level? Why?

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

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

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

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

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

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

6.62 A study examined the effect of exercise on how students perform on their final exam in statistics. The P-value was given as 0.87.
(a) State null and alternative hypotheses that could have been used for this study. (Note that there is more than one correct answer.)
(b) Do you reject the null hypothesis?
(c) What is your conclusion?
(d) What other facts about the study would you like to know for a proper interpretation of the results?

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

6.64 The mean yield of corn in the United States is about 120 bushels per acre. A survey of 40 farmers this year gives a sample mean yield of $\overline{x} = 123.8$ 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

$$H_0\colon \mu = 120$$

$$H_a\colon \mu \neq 120$$

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 non-Normal? Why?

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

(a) Assume that ACT scores vary Normally with standard deviation 6. Is the outcome $\overline{x} = 22.1$ good evidence that the population mean score is greater than 20? State H_0 and H_a, compute the test statistic and the P-value, and answer the question by interpreting your result.
(b) The results are in any case inconclusive because of the design of the study. The effects of the new course are confounded with any change from past years, such as other new courses or higher standards. Briefly outline the design of a better study of the effect of the new course on ACT scores.

6.66 There are other z statistics that we have not yet studied. We can use Table D to assess the significance of any z statistic. A study compares the habits of students who are on academic probation with students whose grades are satisfactory. One variable measured is the hours spent watching television last week. The null hypothesis is "no difference" between the means for the two populations. The alternative hypothesis is two-sided. The value of the test statistic is $z = -1.37$.
(a) Is the result significant at the 5% level?
(b) Is the result significant at the 1% level?

6.67 You measure the weights of 24 male runners. These runners are not a random sample from a population, but you are willing to assume that their weights represent the weights of similar runners. Here are their weights in kilograms:

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

Exercise 6.11 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

$$H_0\colon \mu = 61.3 \text{ kg}$$

$$H_a\colon \mu \neq 61.3 \text{ kg}$$

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

6.68 An old farmer claims to be able to detect the presence of water with a forked stick. In a test of this claim, he is presented with 5 identical barrels, some containing water and some not. He is right in 4 of the 5 cases.
(a) Suppose the farmer has probability p of being correct. If he is just guessing, $p = 0.5$. State an appropriate H_0 and H_a in terms of p for a test of whether he does better than guessing.
(b) If the farmer is simply guessing, what is the distribution of the number X of correct answers in 5 tries?

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

6.69 Here are several situations where there is an incorrect application of the ideas presented in this section. Write a short paragraph explaining what is wrong in each situation and why it is wrong.
(a) A climatologist wants to test the null hypothesis that it will rain tomorrow.
(b) A random sample of size 20 is taken from a population that is assumed to have a standard deviation of 15. The standard deviation of the sample mean is 15/20.
(c) A researcher tests the following null hypothesis: H_0: $\bar{x} = 10$.

6.70 Here are several situations where there is an incorrect application of the ideas presented in this section. Write a short paragraph explaining what is wrong in each situation and why it is wrong.
(a) A change is made that should improve student satisfaction with the way grades are processed at your college. The null hypothesis, that there is an improvement, is tested versus the alternative, that there is no change.
(b) A significance test rejected the null hypothesis that the sample mean is 25.
(c) A report on a study says that the results are statistically significant and the P-value is 0.95.

6.71 State the appropriate null hypothesis H_0 and alternative hypothesis H_a in each of the following cases.
(a) An experiment is designed to examine the effect of a diet high in soy products on the bone density of adult rats.
(b) The student newspaper at your college recently changed the format for their news stories. You take a random sample of students and select those who regularly read the newspaper. These are asked to indicate their opinions on the changes using a five-point scale: -2 if the new format is much worse than the old, -1 if the new format is somewhat worse than the old, 0 if the new format is the same as the old, $+1$ if the new format is somewhat better than the old, and $+2$ if the new format is much better than the old.
(c) The examinations in a large history class are scaled after grading so that the mean score is 75. A self-confident teaching assistant thinks that his students have a higher mean score than the class as a whole. His students this semester can be considered a sample from the population of all students he might teach, so he compares their mean score with 75.

6.72 State the null hypothesis H_0 and the alternative hypothesis H_a in each case.
(a) A national study reports that households spend an average of 30% of their food expenditures in restaurants. A restaurant association in your area wonders if the national finding applies locally. They interview a sample of 40 households and ask about their total food budget and the amount spent in restaurants.
(b) Experiments on learning in animals sometimes measure how long it takes a mouse to find its way through a maze. The mean time is 20 seconds for one particular maze. A researcher thinks that playing rap music will cause the mice to complete the maze faster. She measures how long each of 12 mice takes with the rap music as a stimulus.

(c) A dual X-ray absorptiometry (DXA) scanner is used to measure bone mineral density for people who may be at risk for osteoporosis. To be sure that the measurements are accurate, an object called a "phantom" that has known mineral density $\mu = 1.3$ grams per square centimeter is measured. This phantom is scanned 8 times.

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

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

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

6.76 A test statistic for a two-sided significance test for a population mean is $z = -1.4$. Sketch a standard Normal curve and mark this value of z on it. Find the P-value and shade the appropriate areas under the curve to illustrate your calculations.

6.77 The P-value for a significance test is 0.082.
(a) Do you reject the null hypothesis at level $\alpha = 0.05$?
(b) Do you reject the null hypothesis at level $\alpha = 0.01$?
(c) Explain your answers.

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

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

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

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

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

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

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

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

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

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

These students can be considered to be an SRS of the third-graders in a suburban school district. DRP scores are approximately Normal. Suppose that the standard deviation of scores in this school district is known to be $\sigma = 11$. The researcher believes that the mean score μ of all third-graders in this district is higher than the national mean, which is 32.
(a) State the appropriate H_0 and H_a to test this suspicion.
(b) Carry out the test. Give the P-value, and then interpret the result in plain language.

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

$$H_0\colon \mu = 1.4$$

$$H_a\colon \mu > 1.4$$

The calculated value of the test statistic is $z = 1.75$.
(a) Is the result significant at the 5% level?
(b) Is the result significant at the 1% level?

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

$$H_0\colon \mu = 0.5$$

$$H_a\colon \mu \neq 0.5$$

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

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

6.90 You have performed a two-sided test of significance and obtained a value of $z = -4.3$. Use Table D to find the approximate P-value for this test.

6.91 You have performed a one-sided test of significance and obtained a value of $z = 0.22$. Use Table D to find the approximate P-value for this test.

6.92 You will perform a significance test of

$$H_0\colon \mu = 0$$

versus

$$H_a\colon \mu > 0$$

(a) What values of z would lead you to reject H_0 at the 5% level?
(b) If the alternative hypothesis was

$$H_a\colon \mu \neq 0$$

what values of z would lead you to reject H_0 at the 5% level?
(c) Explain why your answers to parts (a) and (b) are different.

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

6.94 Refer to the previous exercise. Find the P-value for $z = -1.34$.

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

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	96.6	119.3	104.8	101.7

Assume (unrealistically) that you know that the standard deviation of readings for all detectors of this type is $\sigma = 9$.
(a) Give a 95% confidence interval for the mean reading μ for this type of detector.
(b) Is there significant evidence at the 5% level that the mean reading differs from the true value 105? State hypotheses and base a test on your confidence interval from (a).

6.96 A new supplier offers a good price on a catalyst used in your production process. You compare the purity of this catalyst with that of the catalyst offered by your current supplier. The P-value for a test of "no difference" is 0.27. 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.

6.97 The level of calcium in the blood in healthy young adults varies with mean about 9.5 milligrams per deciliter and standard deviation about $\sigma = 0.4$. A clinic in rural Guatemala measures the blood calcium level of 160 healthy pregnant women at their first visit for prenatal care. The mean is $\overline{x} = 9.57$. Is this an indication that the mean calcium level in the population from which these women come differs from 9.5?
(a) State H_0 and H_a.
(b) Carry out the test and give the P-value, assuming that $\sigma = 0.4$ in this population. Report your conclusion.
(c) Give a 95% confidence interval for the mean calcium level μ in this population. We are confident that μ lies quite close to 9.5. This illustrates the fact that a test

based on a large sample ($n = 160$ here) will often declare even a small deviation from H_0 to be statistically significant.

Section 6.3

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

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

$$H_0\colon \mu = 525$$

$$H_a\colon \mu > 525$$

in each of the following situations:
(a) The students' average score is $\overline{x} = 541.4$. Is this result significant at the 5% level?
(b) The average score is $\overline{x} = 541.5$. Is this result significant at the 5% level?
The difference between the two outcomes in (a) and (b) is of no importance. Beware attempts to treat $\alpha = 0.05$ as sacred.

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

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

6.102 A researcher looking for evidence of extrasensory perception (ESP) tests 500 subjects. Four of these subjects do significantly better ($P < 0.01$) than random guessing.
(a) Is it proper to conclude that these four people have ESP? Explain your answer.
(b) What should the researcher now do to test whether any of these four subjects have ESP?

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

6.104 You are the statistical expert on a team that is planning a study. After you have made a careful presentation of the mechanics of significance testing, one of the team members suggests using $\alpha = 0.50$ for the study because you would be more likely to obtain statistically significant results with this choice. Explain in simple terms why this would not be a good use of statistical methods.

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

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

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

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

$$H_0\colon \mu = 480$$

$$H_a\colon \mu > 480$$

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

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

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

$$H_0\colon \mu = 480$$

$$H_a\colon \mu > 480$$

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

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

6.112 The table in Exercise 1.44 gives average property damage per year due to tornadoes for each of the states. Is it appropriate to use the statistical methods we discussed in this chapter for these data? Explain why or why not.

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

6.114 A P-value of 0.90 is reported for a significance test for a population mean. Interpret this result.

Section 6.4

6.115 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

$$H_0\colon \mu = 450$$

$$H_a\colon \mu > 450$$

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

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

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 $\mu = 460$.

6.116 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 $\mu = 462$.

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

$$H_0\colon \mu = 300$$

$$H_a\colon \mu < 300$$

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

$$z = \frac{\overline{x} - 300}{3/\sqrt{6}}$$

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.

6.118 Use the result of the previous exercise to give the probabilities of Type I and Type II errors for the test discussed there. Take the alternative hypothesis to be $\mu = 295$.

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

$$H_0\colon \mu = 0$$

$$H_a\colon \mu > 0$$

You decide to reject H_0 if $\overline{x} > 0$ and to accept H_0 otherwise.
(a) Find the probability of a Type I error, that is, the probability that your test rejects H_0 when in fact $\mu = 0$.
(b) Find the probability of a Type II error when $\mu = 0.3$. This is the probability that your test accepts H_0 when in fact $\mu = 0.3$.
(c) Find the probability of a Type II error when $\mu = 1$.

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

6.121 You want to see if a redesign of the cover of a mail-order catalog will increase sales. A very large number of customers will receive the original catalog, and a random sample of customers will receive the one with the new cover. For planning purposes, you are willing to assume that the sales from the new catalog will be approximately Normal with $\sigma = 60$ dollars and that the mean for the original catalog will be $\mu = 40$ dollars. You decide to use a sample size of $n = 1000$. You wish to test

$$H_0\colon \mu = 40$$

$$H_a\colon \mu > 40$$

You decide to reject H_0 if $\overline{x} > 43.12$ and to accept H_0 otherwise.
(a) Find the probability of a Type I error, that is, the probability that your test rejects H_0 when in fact $\mu = 40$ dollars.
(b) Find the probability of a Type II error when $\mu = 45$ dollars. This is the probability that your test accepts H_0 when in fact $\mu = 45$.
(c) Find the probability of a Type II error when $\mu = 50$.
(d) The distribution of sales is not Normal, because many customers buy nothing. Why is it nonetheless reasonable in this circumstance to assume that the mean will be approximately Normal?

6.122 Consider a test of a hypothesis about the SAT scores of California high school students based on an SRS of 500 students. The hypotheses are

$$H_0\colon \mu = 450$$

$$H_a\colon \mu > 450$$

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

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

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

6.123 You want to see if a redesign of the cover of a mail-order catalog will increase sales. A very large number of customers will receive the original catalog, and a random sample of customers will receive the one with the new cover. For planning purposes, you are willing to assume that the sales from the new catalog will be approximately Normal with $\sigma = 50$ dollars and that the mean for the original catalog will be $\mu = 25$ dollars. You decide to use a sample size of $n = 900$. You wish to test

$$H_0\colon \mu = 25$$

$$H_a\colon \mu > 25$$

You decide to reject H_0 if $\overline{x} > 26$.
(a) Find the probability of a Type I error, that is, the probability that your test rejects H_0 when in fact $\mu = 25$ dollars.
(b) Find the probability of a Type II error when $\mu = 28$ dollars. This is the probability that your test accepts H_0 when in fact $\mu = 28$.

(c) Find the probability of a Type II error when $\mu = 30$.
(d) The distribution of sales is not Normal, because many customers buy nothing. Why is it nonetheless reasonable in this circumstance to assume that the mean will be approximately Normal?

6.124 You are in charge of marketing for a Web site that offers automated medical diagnoses. The program will scan the results of routine medical tests (pulse rate, blood pressure, urinalysis, etc.) and either clear the patient or refer the case to a doctor. You are marketing the program for use as part of a preventive-medicine system to screen many thousands of persons who do not have specific medical complaints. The program makes a decision about each patient.
(a) What are the two hypotheses and the two types of errors that the program can make? Describe the two types of errors in terms of "false positive" and "false negative" test results.
(b) The program can be adjusted to decrease one error probability at the cost of an increase in the other error probability. Which error probability would you choose to make smaller, and why? (This is a matter of judgment. There is no single correct answer.)

Chapter 6 Review Exercises

6.125 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 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?

6.126 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 μ. Repeat this process 100 times and then count the number of times that the confidence interval includes the value $\mu = 20$. Explain your results.

6.127 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.

6.128 Use the same procedure for generating data as in the previous exercise. Now test the null hypothesis that $\mu = 22.5$. Explain your results.

6.129 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 SATM 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 μ has the form $\overline{x} \pm m$. What is the margin of error m? (Remember that we know $\sigma = 100$.)
(c) Use your software to calculate the 95% confidence interval for μ 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 $\overline{x}$ 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 μ? In a very large number of samples, what percent of the confidence intervals would contain μ?

6.130 In the previous exercise you simulated the SATM 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 SATM scores is Normal with standard deviation $\sigma = 100$.
(a) Use your software to carry out a test of

$$H_0\colon \mu = 460$$

$$H_a\colon \mu \neq 460$$

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 $\overline{x}$ 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?

6.131 Suppose that in fact the mean SATM 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

$$H_0\colon \mu = 460$$

$$H_a\colon \mu \neq 460$$

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?

6.132 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 $\overline{x} = 12.9$ grams per 100 milliliters of blood and a standard deviation of 1.6. Taking the standard deviation to be the population value σ, 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?

6.133 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 α test when α, the alternative hypothesis, and all facts about the population remain unchanged.

6.134 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.)

6.135 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 μ has the form $\overline{x} \pm m$. What is the margin of error m? (Remember that we know $\sigma = 100$.)
(c) Use your software to calculate the 95% confidence interval for μ 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 $\overline{x}$ 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 μ? In a very large number of samples, what percent of the confidence intervals would contain μ?

6.136 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

$$H_0\colon \mu = 475$$

$$H_a\colon \mu \neq 475$$

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 $\overline{x}$ 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?

6.137 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

$$H_0\colon \mu = 475$$

$$H_a\colon \mu \neq 475$$

that you did in the previous exercise. How many of the 50 tests have P-values 0.05 or smaller? These tests reject the null hypothesis at the $\alpha = 0.05$ significance level, which is the correct conclusion.
(c) The power of the test against the alternative $\mu = 500$ is the probability that the test will reject H_0: $\mu = 475$ when in fact $\mu = 500$. Calculate this power. In a very large number of samples from a population with mean 500, what percent would reject H_0?

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

6.139 An agronomist examines the cellulose content of a variety of alfalfa hay. Suppose that the cellulose content in the population has standard deviation $\sigma = 8$ milligrams per gram (mg/g). A sample of 16 cuttings has mean cellulose content $\overline{x} = 140$ mg/g.

(a) Give a 95% confidence interval for the mean cellulose content in the population.
(b) A previous study claimed that the mean cellulose content was $\mu = 135$ mg/g, but the agronomist believes that the mean is higher than that figure. State H_0 and H_a and carry out a significance test to see if the new data support this belief.
(c) The statistical procedures used in (a) and (b) are valid when several assumptions are met. What are these assumptions?

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

32 33 40 35 24 36 31 30 20 25

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

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

$$H_0\colon \mu = 0 \quad \text{(no increase)}$$

$$H_a\colon \mu > 0 \quad \text{(an increase)}$$

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

6.142 When asked to explain the meaning of "statistically significant at the $\alpha =$ 0.05 level," a student says, "This means there is only probability 0.05 that the null hypothesis is true." Is this an essentially correct explanation of statistical significance? Explain your answer.

6.143 Use a computer to generate $n = 12$ observations from a Normal distribution with mean 20 and standard deviation 5: $N(20, 5)$. Find the 95% confidence interval for μ. Repeat this process 100 times and then count the number of times that the confidence interval includes the value $\mu = 20$. Explain your results.

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

6.145 Use the same procedure for generating data as in the previous exercise. Now test the null hypothesis that $\mu = 18$. Explain your results.

6.146 A study of late adolescents and early adults reported average months of full-time employment for individuals aged 18 to 26 (Sabrina Oesterle et al., "Volunteerism during the transition to adulthood: a life course perspective," *Social Forces,* 83 (2004), pp. 1123–1149). Here are the means:

Age	18	19	20	21	22	23	24	25	26
Months employed	2.9	4.2	5.0	5.3	6.4	7.4	8.5	8.9	9.3

Assume that the standard deviation for each of these means is 4.5 months and that each sample size is 750.
(a) Calculate the 95% confidence interval for each mean.
(b) Plot the means versus age. Draw a vertical line through the first mean extending up to the upper confidence limit and down to the lower limit. At the ends of the line, draw a short dash. Do the same for each of the other means.
(c) Write a summary of what the data show. Note that in circumstances such as this, it is common practice not to make any adjustments for the fact that several confidence intervals are being reported. Be sure to include comments about this in your summary.

6.147 Persons aged 55 and over represented 21.3% of the U.S. population in the year 2000. This group is expected to increase to 30.5% by 2025. In terms of actual numbers of people, the increase is from 58.6 million to 101.4 million. Restaurateurs have found this market to be important and would like to make their businesses attractive to older customers. One study used a questionnaire to collect data from people aged 50 and over (Barbara A. Almanza, Richard Ghiselli, and William Jaffe, "Foodservice design and aging baby boomers: importance and perception of physical amenities in restaurants," *Foodservice Research International,* 12 (2000), pp. 25–40). For one part of the analysis, individuals were classified into two age groups: 50 to 64 and 65 to 79. There were 267 people in the first group and 263 in the second. One set of items concerned ambiance, menu design, and service. A series of statements were rated on a 1 to 5 scale with 1 representing "strongly disagree" and 5 representing "strongly agree." In some cases the wording has been shortened in the following table. Here are the means:

Statement	50–64	65–79
Ambiance:		
Most restaurants are too dark	2.75	2.93
Most restaurants are too noisy	3.33	3.43
Background music is often too loud	3.27	3.55
Restaurants are too smoky	3.17	3.12
Tables are too small	3.00	3.19
Tables are too close together	3.79	3.81
Menu design:		
Print size is not large enough	3.68	3.77
Glare makes menus difficult to read	2.81	3.01
Colors of menus make them difficult to read	2.53	2.72
Service:		
It is difficult to hear the service staff	2.65	3.00
I would rather be served than serve myself	4.23	4.14
I would rather pay the server than a cashier	3.88	3.48
Service is too slow	3.13	3.10

First, examine the means of the people who are 50 to 64. Order the statements according to the means and describe the results. Then do the same for the older group. For each statement compute the z statistic and the associated P-value for the comparison between the two groups. For these calculations you can assume that the standard deviation of the difference is 0.08, so z is simply the difference in the means divided by 0.08. Note that you are performing 13 significance tests in this exercise. Keep this in mind when you interpret your results. Write a report summarizing your work.

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

$$H_0\colon \mu = 0$$

$$H_a\colon \mu > 0$$

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

$$H_0\colon \mu = 64$$

$$H_a\colon \mu \neq 64$$

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

$$H_0\colon \mu = 20$$

$$H_a\colon \mu < 20$$

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 $H_0\colon \mu = 0$ in terms of the t statistic. Then take $s = 108$ (an estimate based on the data in the previous exercise) and state the rejection criterion in terms of $\overline{x}$.
(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 68.24 52.73 21.50 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 nerves in DDT-poisoning syndrome in albino rats," *Toxicology and Applied Pharmacology*, 6 (1964), pp. 97–213.)

1.6 1.7 1.8 1.9

(a) Find the mean refractory period $\overline{x}$ 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 $\overline{x} = 1.67$ 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 $\overline{x} = 0.2519$ 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 7.11 does not give the raw data or any discussion of Normality. You would like to replace the t procedure used in Exercise 7.12 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 7.12 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

$$H_0\colon \mu = 0$$

$$H_a\colon \mu \neq 0$$

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 7.12, and restate the rejection criterion in terms of $\overline{x}$.
(c) Find the probability of this event when $\mu = 0.2$ (the alternative given) and $\sigma = 0.3$ (estimated from the data in Exercise 7.12) 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

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

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

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

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

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

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

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

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

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

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

(a) Make a stemplot of the data. The distribution is skewed to the right and has several high outliers. The *bootstrap* 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:

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

(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 7.30?

7.34 The cost of health care is the subject of many studies that use statistical methods. One such study estimated that the average length of service for home health care among people over the age of 65 who use this type of service is 96.0 days with a standard error of 5.1 days. Assuming that the degrees of freedom are large, calculate a 90% confidence interval for the true mean length of service. (A. N. Dey, "Characteristics of elderly home health care users," National Center for Health Statistics, 1996.)

7.35 The embryos of brine shrimp can enter a dormant phase in which metabolic activity drops to a low level. Researchers studying this dormant phase measured the level of several compounds important to normal metabolism. The results were reported in a table, with the note, "Values are means ± SEM for three independent samples." The table entry for the compound ATP was 0.84 ± 0.01. Biologists reading the article are presumed to be able to decipher this. (S. C. Hand and E. Gnaiger, "Anaerobic dormancy quantified in *Artemia* embryos," *Science*, 239 (1988), pp. 1425–1427.)
(a) What does the abbreviation "SEM" stand for?
(b) The researchers made three measurements of ATP, which had $\overline{x} = 0.84$. 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 following table gives the times required (in seconds) to move the indicator a fixed distance (data provided by Timothy Sturm, Purdue University):

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

(a) Each of the 25 students used both instruments. Discuss briefly how the experiment should be arranged and how randomization should be used.
(b) The project hoped to show that right-handed people find right-hand threads easier to use. State the appropriate H_0 and H_a about the mean time required to complete the task.
(c) Carry out a test of your hypotheses. Give the P-value and report your conclusions.

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

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

7.39 The tomato experts who carried out the field trial described in the previous exercise suspect that the relative lack of significance there is due to low power. They would like to be able to detect a mean difference in yields of 0.6 pound per plant at the 0.05 significance level. Based on the previous study, use 0.92 as an estimate of both the population σ and the value of s in future samples.
(a) What is the power of the test from Exercise 7.38 with $n = 12$ against the alternative $\mu = 0.6$?
(b) If the sample size is increased to $n = 30$ plots of land, what will be the power against the same alternative?

7.40 The following situations all require inference about a mean or means. Identify each as (1) a single sample, (2) matched pairs, or (3) two independent samples. The procedures of this section apply to cases (1) and (2). We will learn procedures for (3) in the next section.
(a) An education researcher wants to learn whether inserting questions before or after introducing a new concept in an elementary school mathematics text is more effective. He prepares two text segments that teach the concept, one with motivating questions before and the other with review questions after. Each text segment is used to teach a different group of children, and their scores on a test over the material are compared.
(b) Another researcher approaches the same problem differently. She prepares text segments on two unrelated topics. Each segment comes in two versions, one with questions before and the other with questions after. Each of a group of children is taught both topics, one topic (chosen at random) with questions before and the other with questions after. Each child's test scores on the two topics are compared to see which topic he or she learned better.
(c) To evaluate a new analytical method, a chemist obtains a reference specimen of known concentration from the National Institute of Standards and Technology. She then makes 20 measurements of the concentration of this specimen with the new method and checks for bias by comparing the mean result with the known concentration.
(d) Another chemist is evaluating the same new method. He has no reference specimen, but a familiar analytic method is available. He wants to know if the new and

old methods agree. He takes a specimen of unknown concentration and measures the concentration 10 times with the new method and 10 times with the old method.

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

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

15.8	13.6	15.6	19.1	22.4	15.6	22.5	17.2	19.4	22.6
19.4	18.0	14.6	18.7	21.0	14.8	22.6	21.5	14.3	20.9

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

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

21.0	19.0	18.7	39.2	45.8	19.8	48.4	21.0	29.1	35.7
31.6	49.0	16.0	34.6	36.3	19.0	43.3	37.5	16.5	34.5

Answer the questions given in the previous exercise.

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

7.45 For a sample of size 5, a test of a null hypothesis versus a two-sided alternative gives $t = 2.45$.
(a) Is the test result significant at the 5% level? Draw a sketch of the appropriate t distribution and illustrate your calculation with this sketch.
(b) Now assume that the same statistic was obtained for a sample size of $n = 10$. Assess the statistical significance of the result and illustrate the calculation with a

sketch. How did the statistical significance change with the sample size? Explain your answer.

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

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

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

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

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

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

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

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

t distribution.
(f) If you have software available, find the exact P-value.

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

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

725	764	853	559	1225	456	555	1168	682	676
808	882	1062	970	909	802	574	576	784	997
651	716	438	1220	1475	948	1050	976	572	483
1256	685	1144	848	1213	3223	1270	1108	350	421

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

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

7.57 Many organizations are doing surveys to determine the satisfaction of their customers. Attitudes toward various aspects of campus life were the subject of one such study conducted at Purdue University. Each item was rated on a 1 to 5 scale, with 5 being the highest rating. The average response of 1406 first-year students to

"Feeling welcomed at Purdue" was 3.9 with a standard deviation of 0.98. Assuming that the respondents are an SRS, give a 99% confidence interval for the mean of all first-year students.

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

2	5	7	–2	2	7	4	1	0	7	3	4	3	4	9	4	5
2	9	6	0	3	6	–1	3	4	6	7	–2	7	–3	3	4	4

(a) Display the data and summarize the distribution.
(b) Find the mean, the standard deviation, and the standard error of the mean.
(c) Give a 95% confidence interval for the mean improvement in reasoning scores.

7.59 Refer to the previous exercise. Test the null hypothesis that there is no improvement versus the alternative suggested by the neurobiological arguments. State the hypotheses, and give the test statistic with degrees of freedom and the P-value. What do you conclude? From your answer to part (c) of the previous exercise what can be concluded from this significance test?

7.60 The researchers studying vitamin C in CSB were also interested in a similar commodity called wheat-soy blend (WSB). Both of these commodities are mixed with other ingredients and cooked. Loss of vitamin C as a result of this process was another concern of the researchers. One preparation used in Haiti called gruel (or "bouillie" in Creole) can be made from WSB, salt, sugar, milk, banana, and other optional items to improve the taste. Samples of gruel prepared in Haitian households were collected. The vitamin C content (in milligrams per 100 grams of blend, dry basis) was measured before and after cooking. Here are the results:

Sample	1	2	3	4	5
Before	73	79	86	88	78
After	20	27	29	36	17

Set up appropriate hypotheses and carry out a significance test for these data. (It is not possible for cooking to increase the amount of vitamin C.)

7.61 Refer to the previous exercise. The fact that vitamin C is destroyed by cooking is neither new nor surprising, so the significance test you performed in the previous exercise simply confirms that this fact is evident in the small sample of data collected. The real question here concerns how much of the vitamin C is lost.
(a) Give a 95% confidence interval for the amount of vitamin C lost by preparing and cooking gruel in Haiti.
(b) The units for your interval are mg/100 g of blend. To meaningfully interpret the amount lost, it is necessary to know something about the amount that we started with. For example, a loss of 50 mg/100 g would probably not be of much concern if we started with 5000 mg/100 g. The specifications call for the blend to

contain 98 mg/100 g (dry basis). The difference between this specification and the "before" values above is due to sample variation in the manufacturing process and the handling of the product from the time it was manufactured until it was used to prepare gruel in these Haitian homes. Express the "after" data as percent of specification and give a 95% confidence interval for the mean percent.

7.62 A bank wonders whether omitting the annual credit card fee for customers who charge at least \$3000 in a year would increase the amount charged on its credit card. The bank makes this offer to an SRS of 500 of its existing credit card customers. It then compares how much these customers charge this year with the amount that they charged last year. The mean increase is \$565, and the standard deviation is \$267.
(a) Is there significant evidence at the 1% level that the mean amount charged increases under the no-fee offer? State H_0 and H_a and carry out a t test.
(b) Give a 95% confidence interval for the mean amount of the increase.
(c) The distribution of the amount charged is skewed to the right, but outliers are prevented by the credit limit that the bank enforces on each card. Use of the t procedures is justified in this case even though the population distribution is not Normal. Explain why.
(d) A critic points out that the customers would probably have charged more this year than last even without the new offer because the economy is more prosperous and interest rates are lower. Briefly describe the design of an experiment to study the effect of the no-fee offer that would avoid this criticism.

7.63 In a randomized comparative experiment on the effect of dietary calcium on blood pressure, 54 healthy white males were divided at random into two groups. One group received calcium; the other, a placebo. At the beginning of the study, the researchers measured many variables on the subjects. The paper reporting the study gives $\overline{x} = 114.9$ and $s = 9.3$ for the seated systolic blood pressure of the 27 members of the placebo group.
(a) Give a 95% confidence interval for the mean blood pressure of the population from which the subjects were recruited.
(b) What assumptions about the population and the study design are required by the procedure you used in (a)? Which of these assumptions are important for the validity of the procedure in this case?

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

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	96.6	119.3	104.8	101.7

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

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

Factory	Haiti	Factory	Haiti	Factory	Haiti
44	40	45	38	39	43
50	37	32	40	52	38
48	39	47	35	45	38
44	35	40	38	37	38
42	35	38	34	38	41
47	41	41	35	44	40
49	37	43	37	43	35
50	37	40	34	39	38
39	34	37	40	44	36

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

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

Teacher	Pretest	Posttest	Teacher	Pretest	Posttest
1	30	29	11	30	32
2	28	30	12	29	28
3	31	32	13	31	34
4	26	30	14	29	32
5	20	16	15	34	32
6	30	25	16	20	27
7	34	31	17	26	28
8	15	18	18	25	29
9	28	33	19	31	32
10	20	25	20	29	32

Summarize the data graphically and numerically. Then analyze the data using a significance test and a confidence interval. Write a short report summarizing your results.

7.67 Exercise 7.60 gives data on the amount of vitamin C in gruel made from wheat-soy blend in 5 Haitian households before and after cooking. Is there evidence that the median amount of vitamin C is less after cooking? State hypotheses, carry out a sign test, and report your conclusion.

7.68 Apply the sign test to the data in Exercise 7.68 to assess the effects of piano lessons on spatial-temporal reasoning.
(a) State the hypotheses two ways: in terms of a population median and in terms of the probability of an improvement in the test score.
(b) Carry out the sign test. Find the approximate P-value using the Normal approximation to the binomial distributions, and report your conclusion.

7.69 Use the sign test to assess whether the summer institute in Exercise 7.66 improves Spanish listening skills. State the hypotheses, give the P-value using the binomial table (Table C), and report your conclusion.

7.70 C-reactive protein (CRP) is a substance that can be measured in the blood. Values increase substantially within 6 hours of an infection and reach a peak within 24 to 48 hours after. In adults, chronically high values have been linked to an increased risk of cardiovascular disease. In a study of apparently healthy children aged 6 to 60 months in Papua New Guinea, CRP was measured in 90 children (data provided by Francisco Rosales of the Department of Nutritional Sciences, Penn State University). The units are milligrams per liter (mg/l). Here are the data from a random sample of 40 of these children.

0.00	3.90	5.64	8.22	0.00	5.62	3.92	6.81	30.61	0.00
73.20	0.00	46.70	0.00	0.00	26.41	22.82	0.00	0.00	3.49
0.00	0.00	4.81	9.57	5.36	0.00	5.66	0.00	59.76	12.38
15.74	0.00	0.00	0.00	0.00	9.37	20.78	7.10	7.89	5.53

(a) Look carefully at the data above. Do you think that there are outliers or is this a skewed distribution? Now use a histogram or stemplot to examine the distribution. Write a short summary describing the distribution.
(b) Do you think that the mean is a good characterization of the center of this distribution? Explain why or why not.
(c) Find a 95% confidence interval for the mean CRP. Discuss the appropriateness of using this methodology for these data.

7.71 Refer to the previous exercise. With strongly skewed distributions such as this, we frequently reduce the skewness by taking a log transformation. We have a bit of a problem here, however, because some of the data are recorded as 0.00 and the logarithm of zero is not defined. For this variable, the value 0.00 is recorded whenever the amount of CRP in the blood is below the level that the measuring instrument is capable of detecting. The usual procedure in this circumstance is to add a small number to each observation before taking the logs. Transform these data by adding 1 to each observation and then taking the logarithm. Use the questions in the previous exercise as a guide to your analysis, and prepare a summary contrasting this analysis with the one that you performed in the previous exercise.

7.72 In the Papua New Guinea study that provided the data for the previous two exercises, the researchers also measured serum retinol. A low value of this variable can be an indicator of vitamin A deficiency. Below are the data on the same sample of 40 children from this study. The units are micromoles per liter (μmol/l).

1.15	1.36	0.38	0.34	0.35	0.37	1.17	0.97	0.97	0.67
0.31	0.99	0.52	0.70	0.88	0.36	0.24	1.00	1.13	0.31
1.44	0.35	0.34	1.90	1.19	0.94	0.34	0.35	0.33	0.69
0.69	1.04	0.83	1.11	1.02	0.56	0.82	1.20	0.87	0.41

Analyze these data. Use the questions in the previous two exercises as a guide.

7.73 The level of various substances in the blood of kidney dialysis patients is of concern because kidney failure and dialysis can lead to nutritional problems. A researcher performed blood tests on several dialysis patients on 6 consecutive clinic visits (data from Joan M. Susic, "Dietary phosphorus intakes, urinary and peritoneal phosphate excretion and clearance in continuous ambulatory peritoneal dialysis patients," MS thesis, Purdue University, 1985). One variable measured was the level of phosphate in the blood. Phosphate levels for an individual tend to vary Normally over time. The data on one patient, in milligrams of phosphate per deciliter (mg/dl) of blood, are given below:

5.6 5.1 4.6 4.8 5.7 6.4

(a) Calculate the sample mean $\overline{x}$ and its standard error.
(b) Use the t procedures to give a 90% confidence interval for this patient's mean phosphate level.

7.74 The normal range of values for blood phosphate levels is 2.6 to 4.8 mg/dl. The sample mean for the patient in the previous exercise falls above this range. Is this good evidence that the patient's mean level in fact falls above 4.8? State H_0 and H_a and use the data in the previous exercise to carry out a t test. Between which levels from Table D does the P-value lie? Are you convinced that the patient's phosphate level is higher than normal?

Section 7.2

7.75 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.

7.76 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:

Group	n	$\overline{x}$	s
A	150	\$1987	\$392
B	150	\$2056	\$413

(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.

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

```
                              TTEST PROCEDURE

Variable: KNEE

GROUP       N                Mean              Std Dev           Std Error
----------------------------------------------------------------------
SKILLED    10          4.18283335           0.47905935          0.15149187
NOVICE      8          3.01000000           0.95894830          0.33903942

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

(a) The researchers believed that the knee velocity would be higher for skilled rowers. State H_0 and H_a.
(b) Give the value of the two-sample t statistic and its P-value (note that SAS provides two-sided P-values). What do you conclude?

(c) Give a 90% confidence interval for the mean difference between the knee velocities of skilled and novice female rowers.

7.78 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

GROUP       N            Mean            Std Dev          Std Error
----------------------------------------------------------------------
SKILLED    10      70.3700000         6.10034898         1.92909973
NOVICE      8      68.4500000         9.03999930         3.19612240

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

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

7.79 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):

Group	$\overline{x}$	s
Males	416	87
Females	386	74

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.

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

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

(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.

7.81 The data on weights of skilled and novice rowers in Exercise 7.78 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 7.77, because the sample standard deviations in the two groups are different enough to cast doubt on the assumption of a common population standard deviation.)

7.82 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.

7.83 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?

7.84 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:

Low fitness			High fitness		
4.99	5.53	3.12	6.68	5.93	5.71
4.24	4.12	3.77	6.42	7.08	6.20
4.74	5.10	5.09	7.32	6.37	6.04
4.93	4.47	5.40	6.38	6.53	6.51
4.16	5.30		6.16	6.68	

(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.

7.85 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:

Month	n	$\bar{x}$	$s/\sqrt{n}$
July	90	\$2.95	\$0.023
September	45	\$3.61	\$0.029

(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.

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

7.87 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.

7.88 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.

7.89 The following table 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.)

Beef hot dogs		Meat hot dogs		Poultry hot dogs	
Calories	Sodium	Calories	Sodium	Calories	Sodium
186	495	173	458	129	430
181	477	191	506	132	375
176	425	182	473	102	396
149	322	190	545	106	383
184	482	172	496	94	387
190	587	147	360	102	542
158	370	146	387	87	359
139	322	139	386	99	357
175	479	175	507	170	528
148	375	136	393	113	513
152	330	179	405	135	426
111	300	153	372	142	513
141	386	107	144	86	358
153	401	195	511	143	581
190	645	135	405	152	588
157	440	140	428	146	522
131	317	138	339	144	545
149	319				
135	298				
132	253				

(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?

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

Calcium group			Placebo group		
Begin	End	Decrease	Begin	End	Decrease
107	100	7	123	124	–1
110	114	–4	109	97	12
123	105	18	112	113	–1
129	112	17	102	105	–3
112	115	–3	98	95	3
111	116	–5	114	119	–5
107	106	1	119	114	5
112	102	10	114	112	2
136	125	11	110	121	–11
102	104	–2	117	118	–1
			130	133	–3

One group took a calcium supplement, and the other group received a placebo.
(a) Perform the significance test using a two-sample t test that does not require equal population standard deviations. Compare your P-value with the result $P = 0.059$ for the pooled t test.
(b) Give a 90% confidence interval for the difference in means, again using a procedure that does not require equal standard deviations. How does the margin of error of your interval compare with 5.6 mm, the margin of error for the pooled t test?

7.91 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.)

Group	n	$\overline{x}$	s
Children	10	-3.67	33.89
Adults	20	-23.17	50.74

(a) What is the standard error of the sample mean VOT for the 20 adult subjects? What is the standard error of the difference $\overline{x}_1 - \overline{x}_2$ 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).

7.92 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.)

7.93 Repeat the comparison of mean VOTs for children and adults in Exercise 7.91 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 7.91 from the two-sample t procedures?

7.94 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:

Group	n	$\overline{x}$	s
Males	675	\$3297.91	\$2394.65
Females	621	\$2380.68	\$1815.55

(a) Use the two-sample t procedures to give a 90% confidence interval for the difference between the mean summer earnings of male and female students.
(b) The distribution of earnings is strongly skewed to the right. Nevertheless, use of t procedures is justified. Why?
(c) Once the sample size was decided, the sample was chosen by taking every kth name from an alphabetical list of undergraduates. Is it reasonable to consider the sample as two SRSs chosen from the male and female undergraduate populations?
(d) What other information about the study would you request before accepting the results as describing all undergraduates?

7.95 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

12.207 16.869 25.050 22.429 8.456 20.589

The control group data were

11.074 9.686 12.064 9.351 8.182 6.642

Normal quantile plots 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

$$H_0\colon \mu_1 = \mu_2$$

$$H_a\colon \mu_1 \neq \mu_2$$

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

```
                         TTEST PROCEDURE

Variable: SPIKE

GROUP       N             Mean           Std Dev         Std Error
-----------------------------------------------------------------
DDT         6      17.60000000        6.34014839        2.58835474
CONTROL     6       9.49983333        1.95005932        0.79610839

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 $\overline{x}_i$ and s_i, verify the values given for the test statistic $t = 2.99$ and the degrees of freedom df $= 5.9$.

7.96 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:

Group	Sex	n	$\overline{x}$	s
1	Male	133	25.34	5.05
2	Female	162	24.94	5.44

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

7.97 In each of the following situations explain what is wrong and why.
(a) A researcher wants to test $H_0\colon \bar{x}_1 = \bar{x}_2$ versus the two-sided alternative $H_1\colon \bar{x}_1 \neq \bar{x}_2$.
(b) A study recorded the scores of 20 children who were similar in age. The scores of the 10 boys in the study were compared with the scores of all 20 children using

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

7.98 For each of the following, answer the question and give a short explanation of your reasoning.
(a) A 95% confidence interval for the difference between two means is reported as (1.6, 2.3). What can you conclude about the results of a significance test of the null hypothesis that the population means are equal versus the two-sided alternative?
(b) Will larger samples generally give a larger or smaller margin of error for the difference between two sample means?

7.99 For each of the following, answer the question and give a short explanation of your reasoning.
(a) A significance test for comparing two means gave $t = -3.11$ with 23 degrees of freedom. Can you reject the null hypothesis that the μ's are equal versus the two-sided alternative at the 5% significance level?
(b) Answer part (a) for the one-sided alternative that the difference in means is positive.

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

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

7.102 Do piano lessons improve the spatial-temporal reasoning of preschool children? We examined this question in Exercises 7.58 and 7.59 by analyzing the change in spatial-temporal reasoning of 34 preschool children after six months of piano lessons. Here we examine the same question by comparing the changes of those students with the changes of 44 children in a control group. Here are the data for the children who took piano lessons:

2	5	7	–2	2	7	4	1	0	7	3	4	3	4	9	4	5
2	9	6	0	3	6	–1	3	4	6	7	–2	7	–3	3	4	4

The control group scores are

1	–1	0	1	–4	0	0	1	0	–1	0	1	1	–3	–2
4	–1	2	4	2	2	2	–3	–3	0	2	0	–1	3	–1
5	–1	7	0	4	0	2	1	–6	0	2	–1	0	–2	

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

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

7.104 Refer to Exercises 7.58 and 7.59 and the previous two exercises. We have used four ways to address the question of interest. Discuss the relative merits of each approach.

7.105 In what ways are companies that fail different from those that continue to do business? A study compared various characteristics of 68 healthy and 33 failed firms. One of the variables was the ratio of current assets to current liabilities. Roughly speaking, this is the amount that the firm is worth divided by what it owes. The data are given in the following table.

Ratio of current assets to current liabilities

Healthy firms						Failed firms		
1.50	0.10	1.76	1.14	1.84	2.21	0.82	0.89	1.31
2.08	1.43	0.68	3.15	1.24	2.03	0.05	0.83	0.90
2.23	2.50	2.02	1.44	1.39	1.64	1.68	0.99	0.62
0.89	0.23	1.20	2.16	1.80	1.87	0.91	0.52	1.45
1.91	1.67	1.87	1.21	2.05	1.06	1.16	1.32	1.17
0.93	2.17	2.61	3.05	1.52	1.93	0.42	0.48	0.93
1.95	2.61	1.11	0.95	0.96	2.25	0.88	1.10	0.23
2.73	1.56	2.73	0.90	2.12	1.42	1.11	0.19	0.13
1.62	1.76	2.22	2.80	1.85	0.96	2.03	0.51	1.12
1.71	1.02	2.50	1.55	1.69	1.64	0.92	0.26	1.15
1.03	1.80	0.67	2.44	2.30	2.21	0.13	0.88	0.09
1.96	1.81							

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

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

Group	n	$\overline{x}$	s
Positive test	134	2733	599
Other	5974	3118	672

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

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

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

Here are the scores of the men:

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

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

7.108 Does bread lose its vitamins when stored? Small loaves of bread were prepared with flour that was fortified with a fixed amount of vitamins. After baking, the vitamin C content of two loaves was measured. Another two loaves were baked at the same time, stored for three days, and then the vitamin C content was measured. The units are milligrams per hundred grams of flour (mg/100 g) (Helen Park et al., "Fortifying bread with each of three antioxidants," *Cereal Chemistry,* 74 (1997), pp. 202–206). Here are the data:

Immediately after baking:	47.62,	49.79
Three days after baking:	21.25,	22.34

(a) When bread is stored, does it lose vitamin C? To answer this question, perform a two-sample t test for these data. Be sure to state your hypotheses, the test statistic with degrees of freedom, and the P-value.
(b) Give a 90% confidence interval for the amount of vitamin C lost.

7.109 Suppose that the researchers in the previous exercise could have measured the same two loaves of bread immediately after baking and again after three days. Assume that the data given had come from this study design. (Assume that the values given in the previous exercise are for first loaf and second loaf from left to right.)
(a) Explain carefully why your analysis in the previous exercise is *not correct* now,

even though the data are the same.
(b) Redo the analysis for the design based on measuring the same loaves twice.

7.110 Refer to the previous two exercises. The amount of vitamin E (in mg/100 g of flour) in the same loaves was also measured. Here are the data:

Immediately after baking:	94.6,	96.0
Three days after baking:	97.4,	94.3

(a) When bread is stored, does it lose vitamin E? To answer this question, perform a two-sample t test for these data. Be sure to state your hypotheses, the test statistic with degrees of freedom, and the P-value.
(b) Give a 90% confidence interval for the amount of vitamin E lost.

7.111 Refer to the previous three exercises. Some people claim that significance tests with very small samples never lead to rejection of the null hypothesis. Discuss this claim using the results of these two exercises.

7.112 Refer to the previous four exercises. The analysis of the loss of vitamin C when bread is stored is a rather unusual case involving very small sample sizes. There are only two observations per condition (immediately after baking and three days later). When the samples are so small, we have very little information to make a judgment about whether the population standard deviations are equal. The potential gain from pooling is large when the sample sizes are very small. Assume that we will perform a two-sided test using the 5% significance level.
(a) Find the critical value for the unpooled t test statistic that does not assume equal variances. Use the minimum of $n_1 - 1$ and $n_2 - 1$ for the degrees of freedom.
(b) Find the critical value for the pooled t test statistic.
(c) How does comparing these critical values show an advantage of the pooled test?

Section 7.3

7.113 The F statistic $F = s_1^2/s_2^2$ is calculated from samples of size $n_1 = 10$ and $n_2 = 21$. (Remember that n_1 is the numerator sample size.)
(a) What is the upper 5% critical value for this F?
(b) In a test of equality of standard deviations against the two-sided alternative, this statistic has the value $F = 2.45$. Is this value significant at the 10% level? Is it significant at the 5% level?

7.114 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.

7.115 Exercise 7.77 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?

7.116 Answer the same questions for the weights of the two groups, recorded in Exercise 7.78.

7.117 The observed inequality between the sample standard deviations of male and female SAT Mathematics scores in Exercise 7.79 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.

7.118 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?

7.119 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?

7.120 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.

7.121 Exercise 7.82 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.

7.122 A USDA survey used to estimate wheat prices in July and September is described in Exercise 7.85. Using the standard deviations you calculated there, perform the test for equality of standard deviations and summarize your conclusion.

7.123 The data for VOTs of children and adults in Exercise 7.91 show quite different sample standard deviations. How statistically significant is the observed inequality?

7.124 Suppose that you wanted to compare intramural basketball players and intramural soccer players on the "ego strength" personality factor described in Exercise 7.84. 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 7.84, 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.

7.125 An F statistic will be used to compare two variances. How large does the ratio of the largest to the smallest variance need to be for the significance test to reject the null hypothesis that the population variances are the same in the following settings? Use the 5% level of significance.
(a) The two sample sizes are 5.
(b) The two sample sizes are 10.
(c) The two sample sizes are 26.
(d) What do you conclude?

7.126 Return to the SSHA data in Exercise 7.107. SSHA scores are generally less variable among women than among men. We want to know whether this is true for this college.
(a) State H_0 and H_a. Note that H_a is one-sided in this case.
(b) Because Table E contains only upper critical values for F, a one-sided test requires that in calculating F the numerator s^2 belongs to the group that H_a claims to have the larger σ. Calculate this F.
(c) Compare F to the entries in Table E (no doubling of p) to obtain the P-value. Be sure the degrees of freedom are in the proper order. What do you conclude about the variation in SSHA scores?

7.127 In Exercise 7.106 data on cocaine use and birth weight are summarized. The study has been criticized because of several design problems. Suppose that you are designing a new study. Based on the results in Exercise 7.106, you think that the true difference in mean birth weights may be about 350 grams (g); a difference this large is clinically important. For planning purposes assume that you will have 75 women in each group and that the common standard deviation is 650 g, a guess that is between the two standard deviations in Exercise 7.106. If you use a pooled two-sample t test with a Type I error of 0.05, what is the power of the test for this design?

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

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

7.130 We studied the loss of vitamin C when bread is stored in Exercise 7.108. Recall that two loaves were measured immediately after baking and another two loaves were measured after three days of storage. These are very small sample sizes.
(a) Use Table E to find the value that the ratio of variances would have to exceed for us to reject the null hypothesis (at the 5% level) that the standard deviations are equal. What does this suggest about the power of the test?
(b) Perform the test and state your conclusion.

Chapter 7 Review Exercises

7.131 Data on the numbers of manatees killed by boats each year are given in Exercise 2.36. 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?

7.132 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?

7.133 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:

Group	n	$\overline{x}$	s
Nitrite	30	7880	1115
Control	30	8112	1250

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

7.134 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.)

Group	n	$\overline{x}$	s
Students	750	35.12	4.31
Workers	412	37.32	3.83

(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.

7.135 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. Landis, "Gas bubbles in fossil amber as possible indicators of the major gas composition of ancient air," *Science*, 239 (1988), pp. 1406–1409.)
(a) Graph the data, and comment on skewness and outliers.
(b) The t procedures will be only approximate in this case. Give a 90% t confidence interval for the mean percent of nitrogen in ancient air.

7.136 The table in Exercise 1.19 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.

7.137 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.

7.138 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:

Group	n	$\overline{x}$	s
Pets	26	193	68
Clinic	23	174	44

(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?

7.139 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)."

	Runners	Others
Abdomen	7.1 (1.0)	20.6 (9.0)
Thigh	6.1 (1.8)	17.4 (6.6)

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

7.140 The following table gives the levels of three pollutants in the exhaust of 46 randomly selected vehicles of the same type. You will investigate emissions of nitrogen oxides (NOX).

Amounts of three pollutants emitted by light-duty egines (grams per mile)

EN	HC	CO	NOX	EN	HC	CO	NOX
1	0.50	5.01	1.28	2	0.65	14.67	0.72
3	0.46	8.60	1.17	4	0.41	4.42	1.31
5	0.41	4.95	1.16	6	0.39	7.24	1.45
7	0.44	7.51	1.08	8	0.55	12.30	1.22
9	0.72	14.59	0.60	10	0.64	7.98	1.32
11	0.83	11.53	1.32	12	0.38	4.10	1.47
13	0.38	5.21	1.24	14	0.50	12.10	1.44
15	0.60	9.62	0.71	16	0.73	14.97	0.51
17	0.83	15.13	0.49	18	0.57	5.04	1.49
19	0.34	3.95	1.38	20	0.41	3.38	1.33
21	0.37	4.12	1.20	22	1.02	23.53	0.86
23	0.87	19.00	0.78	24	1.10	22.92	0.57
25	0.65	11.20	0.95	26	0.43	3.81	1.79
27	0.48	3.45	2.20	28	0.41	1.85	2.27
29	0.51	4.10	1.78	30	0.41	2.26	1.87
31	0.47	4.74	1.83	32	0.52	4.29	2.94
33	0.56	5.36	1.26	34	0.70	14.83	1.16
35	0.51	5.69	1.73	36	0.52	6.35	1.45
37	0.57	6.02	1.31	38	0.51	5.79	1.51
39	0.36	2.03	1.80	40	0.48	4.62	1.47
41	0.52	6.78	1.15	42	0.61	8.43	1.06
43	0.58	6.02	0.97	44	0.46	3.99	2.01
45	0.47	5.22	1.12	46	0.55	7.47	1.39

(a) Make a stemplot and, if your software allows, a Normal quantile plot of the NOX levels. Do the plots suggest that the distribution of NOX emissions is approximately Normal? Can you safely employ t procedures to analyze these data?
(b) Give a 95% confidence interval for the mean NOX level in vehicles of this type.
(c) Your supervisor hopes the average NOX level is less than 1 gram per mile. You will have to tell him that it's not so. Carry out a significance test to assess the strength of the evidence that the mean NOX level is greater than 1, and then write a short report to your supervisor based on your work in (b) and (c). (Your supervisor has never heard of P-values, so you must use plain language.)

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

7.142 In Exercises 7.127 and 7.128 you found the power for a study designed to compare birth weights of children born to cocaine users with those born to controls. Fix the sample size at 50 in each group and assume the standard deviation is 650 grams and the significance level is 0.05. Pick a set of alternatives that will give values of power ranging from fairly low values to fairly high values. Plot your results versus that alternative and give a short summary of what you have found.

7.143 Healthy bones are continually being renewed by two processes. Through bone formation, new bone is built; through bone resorption, old bone is removed. If one or both of these processes are disturbed, by disease, aging, or space travel, for example, bone loss can be the result. Osteocalcin (OC) is a biochemical marker for bone formation: higher levels of bone formation are associated with higher levels of OC. A blood sample is used to measure OC, and it is much less expensive to obtain than direct measures of bone formation. The units are milligrams of OC per milliliter of blood (mg/ml). One study examined various biomarkers of bone turnover (C. M. Weaver et al., "Quantification of biochemical markers of bone turnover by kinetic measures of bone formation and resorption in young healthy females," *Journal of Bone and Mineral Research,* 12 (1997), pp. 1714–1720). Here are the OC measurements on 31 healthy females aged 11 to 32 years who participated in this study:

68.9	56.3	54.6	31.2	36.4	31.4	52.8	38.4
35.7	76.5	44.4	40.2	77.9	54.6	9.9	20.6
20.0	17.2	24.2	20.9	17.9	19.7	15.9	20.8
8.1	19.3	16.9	10.1	47.7	30.2	17.2	

(a) Display the data with a stemplot or histogram and a boxplot. Describe the distribution.
(b) Find a 95% confidence interval for the mean OC. Comment on the suitability of using this procedure for these data.

7.144 Refer to the previous exercise. Tartrate resistant acid phosphatase (TRAP) is a biochemical marker for bone resorption that is also measured in blood. Here are the TRAP measurements, in units per liter (U/l), for the same 31 females:

19.4	25.5	19.0	9.0	19.1	14.6	25.2	14.6
28.8	14.9	10.7	5.9	23.7	19.0	6.9	8.1
9.5	6.3	10.1	10.5	9.0	8.8	8.2	10.3
3.3	10.1	9.5	8.1	18.6	14.4	9.6	

(a) Display the data with a stemplot or histogram and a boxplot. Describe the distribution.
(b) Find a 95% confidence interval for the mean TRAP. Comment on the suitability of using this procedure for these data.

7.145 Refer to Exercise 7.143 and the OC data for 31 females. Variables that measure concentrations such as this often have distributions that are skewed to the right. For this reason it is common to work with the logarithms of the measured values. Here are the OC values transformed with the (natural) log:

4.23	4.03	4.00	3.44	3.59	3.45	3.97	3.65
3.58	4.34	3.79	3.69	4.36	4.00	2.29	3.03
3.00	2.84	3.19	3.04	2.88	2.98	2.77	3.03
2.09	2.96	2.83	2.31	3.86	3.41	2.84	

(a) Display the data with a stemplot and a boxplot. Describe the distribution.
(b) Find a 95% confidence interval for the mean OC. Comment on the suitability of using this procedure for these data.
(c) Transform the mean and the endpoints of the confidence interval back to the original scale, mg/ml. Compare this interval with the one you computed in Exercise 7.144.

7.146 Refer to Exercise 7.144 and the TRAP data for 31 females. Variables that measure concentrations such as this often have distributions that are skewed to the right. For this reason it is common to work with the logarithms of the measured values. Here are the TRAP values transformed with the (natural) log:

2.97	3.24	2.94	2.20	2.95	2.68	3.23	2.68
3.36	2.70	2.37	1.77	3.17	2.94	1.93	2.09
2.25	1.84	2.31	2.35	2.20	2.17	2.10	2.33
1.19	2.31	2.25	2.09	2.92	2.67	2.26	

(a) Display the data with a stemplot and a boxplot. Describe the distribution.
(b) Find a 95% confidence interval for the mean TRAP. Comment on the suitability of using this procedure for these data.
(c) Transform the mean and the endpoints of the confidence interval back to the original scale, U/l. Compare this interval with the one you computed in Exercise 7.145.

7.147 Polychlorinated biphenyls (PCBs) are a collection of compounds that are no longer produced in the United States but are still found in the environment. Evidence suggests that they can cause harmful health effects when consumed. Because PCBs can accumulate in fish, efforts have been made to identify areas where fish contain excessive amounts so that recommendations concerning consumption limits can be made. There are over 200 types of PCBs. Data from the Environmental Protection Agency National Study of Residues in Lake Fish are given in the data set PCB. More details about this data set can be found in the Data Appendix. Various lakes in the United States were sampled and the amounts of PCBs in fish were measured. The variable PCB is the sum of the amounts of all PCBs found in the fish. The units are parts per billion (ppb).
(a) Use graphical and numerical summaries to describe the distribution of this variable. Include a histogram with the location of the mean and the median clearly marked.
(b) Do you think it is appropriate to use methods based on Normal distributions for these data? Explain why or why not.
(c) Find a 95% confidence interval for the mean. Will this interval contain approximately 95% of the observations in the data set? Explain your answer.
(d) Transform the PCB variable with a logarithm. Analyze the transformed data and summarize your results. Do you prefer to work with the raw data or with logs for this variable? Give reasons for your answer.
(e) Visit the Web site `http://epa.gov/waterscience/fishstudy/` to find details about how the data were collected. Write a summary describing these details and discuss how the results from this study can be generalized to other settings.

7.148 Refer to the previous exercise. Not all types of PCBs are equally harmful. A scale has been developed to convert the raw amount of each type of PCB to a toxic equivalent score (TEQ). The PCB data set contains a variable TEQPCB that is the total TEQ from all PCBs found in each sample. Using the questions in the previous exercise, analyze these data and summarize the results.

7.149 In Exercise 7.70 you analyzed the C-reactive protein (CRP) scores for a random sample of 40 children who participated in a study in Papua New Guinea. Serum retinol for the same children was analyzed in Exercise 7.72. Data for all 90 children who participated in the study are given in the data set PNG, described in the Data Appendix. Researchers who analyzed these data along with data from several other countries were interested in whether or not infections (as indicated by high CRP values) were associated with lower levels of serum retinol. A child with a value of CRP greater than 5.0 mg/l is classified as recently infected. Those whose CRP is less than or equal to 5.0 mg/l are not. Compare the serum retinol levels of the infected and noninfected children. Include graphical and numerical summaries, comments on all assumptions, and details of your analyses. Write a short report summarizing your results.

7.150 Refer to the previous exercise. The researchers in this study also measured α1-acid glycoprotein (AGP). This protein is similar to CRP in that it is an indicator of infection. However, it rises more slowly than CRP and reaches a maximum 2 to 3 days after an infection. The units for AGP are grams per liter (g/l), and any value greater than 1.0 g/l is an indication of infection. Analyze the data on AGP in the data set PNG and write a report summarizing your results.

CHAPTER 8

Section 8.1

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

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

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

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

8.5 In the previous exercise we found that 11 customers from a random sample of 40 would be willing to buy a software upgrade that costs \$100. If the upgrade is to

be profitable, you will need to sell it to more than 20% of your customers. Do the sample data give good evidence that more than 20% are willing to buy?
(a) Formulate this problem as a hypothesis test. Give the null and alternative hypotheses. Will you use a one-sided or a two-sided alternative? Why?
(b) Carry out the significance test. Report the test statistic and the P-value.
(c) Should you proceed with plans to produce and market the upgrade?

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

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

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

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

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

8.11 Refer to the previous exercise. There we arbitrarily chose to state the hypotheses in terms of the proportion of households that have children. We could as easily have used the proportion of households that do not have children.
(a) Write hypotheses in terms of the proportion of households that do not have children to examine how well the sample represents the state in regard to having children in the household or not.

(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-year students at the university who would identify being well-off as an important personal goal.

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

8.16 Shereka, a starting player for a major college basketball team, made only 36.2% of her free throws last season. During the summer she worked on developing a softer shot in the hope of improving her free-throw accuracy. In the first eight games of this season Shereka made 22 free throws in 42 attempts. Let p be her probability of making each free throw she shoots this season.
(a) State the null hypothesis H_0 that Shereka's free-throw probability has remained the same as last year and the alternative H_a that her work in the summer resulted in a higher probability of success.
(b) Calculate the z statistic for testing H_0 versus H_a.
(c) Do you accept or reject H_0 for $\alpha = 0.05$? Find the P-value.
(d) Give a 90% confidence interval for Shereka's free-throw success probability for the new season. Are you convinced that she is now a better free-throw shooter than last season?
(e) What assumptions are needed for the validity of the test and confidence interval calculations that you performed?

8.17 Land's Beginning is a company that sells its merchandise through the mail. It is considering buying a list of addresses from a magazine. The magazine claims that at least 25% of its subscribers have high incomes (they define this to be household income in excess of \$100,000). Land's Beginning would like to estimate the proportion of high-income people on the list. Checking income is very difficult and expensive but another company offers this service. Land's Beginning will pay to find incomes for an SRS of people on the magazine's list. They would like the margin of error of the 95% confidence interval for the proportion to be 0.05 or less. Use the guessed value $p^* = 0.25$ to find the required sample size.

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

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

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

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

8.22 Explain what is wrong with each of the following:
(a) An approximate 95% confidence interval for an unknown proportion p is $\hat{p}$ plus or minus its standard error.
(b) You can use a significance test to evaluate the hypothesis H_0: $\hat{p} = 0.3$ versus the

two-sided alternative.
(c) The large-sample significance test for a population proportion is based on a t statistic.

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

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

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

8.26 A survey of 1280 student loan borrowers found that 448 had loans totaling more than $20,000 for their undergraduate education. Give a 95% confidence interval for the proportion of all student loan borrowers who have loans of $20,000 or more for their undergraduate education.

8.27 In the survey described in the previous exercise, there were 1050 borrowers whose total debt was $10,000 or more. Of these, 192 left school without completing a degree. Consider the population to be borrowers whose total debt was $10,000 or more. Find a 95% confidence interval for the proportion of borrowers who left school without completing a degree in this population.

8.28 Refer to Exercise 8.13. Would a 99% confidence interval be wider or narrower than the one that you found in that exercise? Verify your results by computing the interval.

8.29 Refer to Exercise 8.14. Would a 90% confidence interval be wider or narrower than the one that you found in that exercise? Verify your results by computing the interval.

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

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

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

8.33 An automobile manufacturer would like to know what proportion of its customers are dissatisfied with the service received from their local dealer. The customer relations department will survey a random sample of customers and compute a 95% confidence interval for the proportion that are dissatisfied. From past studies, they believe that this proportion will be about 0.25. Find the sample size needed if the margin of error of the confidence interval is to be about 0.02. Suppose 15% of the sample say that they are dissatisfied. What is the margin of error of the 95% confidence interval?

8.34 You have been asked to survey students at a large college to determine the proportion who favor an increase in student fees to support an expansion of the student newspaper. Each student will be asked whether he or she is in favor of the proposed increase. Using records provided by the registrar you can select a random sample of students from the college. After careful consideration of your resources, you decide that it is reasonable to conduct a study with a sample of 10 students. For this sample size, construct a table of the margins of error for 95% confidence intervals when $\hat{p}$ takes the values 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9.

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

8.36 Owning a cell phone. In a 2004 survey of 1200 undergraduate students throughout the United States, 89% of the respondents said they owned a cell phone. (2005 press release from *The Student Monitor*. It is available online at `studentmonitor.com/press/02.pdf`.) For 90% confidence, what is the margin of error?

8.37 Importance of cell phone "features and functions." Refer to the previous exercise, In that same survey, one question asked what aspect was most important when buying a cell phone. "Features and functions" was the choice for 336 students. Give a 95% confidence interval for the proportion of U.S. students who find "features and functions" the most important aspect when buying a phone.

8.38 Owning a cell phone, continued. Refer to the previous two exercises. It was reported that cell phone ownership by undergraduate students in 2003 was 83%. Do the sample data in 2004 give good evidence that this percent has increased?
(a) Give the null and alternative hypotheses.
(b) Carry out the significance test. Report the test statistic and the P-value.
(c) State your conclusion using $\alpha = 0.05$.

8.39 Working while enrolled in school. A 1993 nationwide survey by the National Center for Education Statistics reports that 72% of all undergraduates work while enrolled in school. (From the U.S. Department of Education. National Center for Education Statistics. Postsecondary Financing Strategies: How Undergraduates Combine Work, Borrowing, and Attendance, NCES 98088, by Stephanie Cuccaro-Alamin and Susan P. Choy. Project Officer, C. Dennis Carroll. Washington, DC: 1998.) You decide to test whether this percent is different at your university. In your random sample of 100 students, 77 said they were currently working.
(a) Give the null and alternative hypotheses for this study.
(b) Carry out the significance test. Report the test statistic and P-value.
(c) Does it appear that the percent of students working at your university is different at the $\alpha = 0.05$ level?

8.40 Can we use the large-sample confidence interval? In each of the following circumstances state whether you would use the large-sample confidence interval.
(a) $n = 50$, $X = 30$
(b) $n = 90$, $X = 15$
(c) $n = 10$, $X = 2$
(d) $n = 60$, $X = 50$
(e) $n = 25$, $X = 15$

8.41 More on whether to use the large-sample confidence interval. In each of the following circumstances state whether you would use the large-sample confidence interval.
(a) $n = 8$, $X = 4$
(b) $n = 500$, $X = 13$
(c) $n = 40$, $X = 18$
(d) $n = 15$, $X = 15$
(e) $n = 50$, $X = 22$

8.42 Gambling and college athletics. Gambling is an issue of great concern to those involved in intercollegiate athletics. Because of this, the National Collegiate Athletic Association (NCAA) surveyed student-athletes concerning their gambling-related behaviors. (Based on information in "NCAA 2003 national study of collegiate sports wagering and associated health risks," which can be found at the NCAA Web site, `ncaa.org.`) There were 5594 Division I male athletes in the survey. Of these, 3547 reported participation in some gambling behavior. This included playing cards, betting on games of skill, buying lottery tickets, and betting on sports.
(a) Find the sample proportion and the large-sample margin of error for 95% confidence. Explain in simple terms the meaning of the 95%.
(b) Because of the way that the study was designed to protect the anonymity of the student-athletes who responded, it was not possible to calculate the number of students who were asked to respond but did not. Does this fact affect the way that you interpret the results? Write a short paragraph explaining your answer.

8.43 Gambling and female athletes. In the study described in the previous exercise, 1447 out of a total of 3469 female student-athletes reported participation in some gambling activity.
(a) Use the large-sample methods to find an estimate of the true proportion with a 95% confidence interval.
(b) The margin of error for this sample is not the same as the margin of error calculated for the previous exercise. Explain why.

Section 8.2

8.44 In the 1996 regular baseball season, the World Series Champion New York Yankees played 80 games at home and 82 games away. They won 49 of their home games and 43 of the games played away. We can consider these games as samples from potentially large populations of games played at home and away. How much advantage does the Yankee home field provide?

(a) Find the proportion of wins for the home games. Do the same for the away games.
(b) Find the standard error needed to compute a confidence interval for the difference in the proportions.
(c) Compute a 90% confidence interval for the difference between the probability that the Yankees win at home and the probability that they win when on the road. Are you convinced that the 1996 Yankees were more likely to win at home?

8.45 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?

8.46 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?

8.47 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.

8.48 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:

	Men	Women
Employed	718	593
Not employed	79	139
Total	797	732

(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?

8.49 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 8.48. What do you observe about the effect of the sample size on the results of these significance tests?

8.50 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?

8.51 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:

Period	n	X(lied)
1	84	15
2	106	21

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

8.52 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.

8.53 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:

Population	n	X(natural)
1 (rural)	160	64
2 (urban)	261	89

(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.

8.54 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?

8.55 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?

8.56 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.

8.57 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?

8.58 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?

8.59 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.

8.60 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:

	Men	Women
Employed	728	603
Not employed	89	149
Total	817	752

(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?

8.61 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 8.60. What do you observe about the effect of the sample size on the results of these significance tests?

8.62 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 received the placebo tablet. (William S. Fields et al., "Controlled trial of aspirin in cerebral ischemia," *Stroke*, 8 (1977), pp. 301–315.) After six months of treatment, the attending physicians evaluated each patient's progress as either favorable or unfavorable. Of the 78 patients in the aspirin group, 63 had favorable outcomes; 43 of the 77 control patients had favorable outcomes.
(a) Compute the sample proportions of patients having favorable outcomes in the two groups.
(b) Give a 90% confidence interval for the difference between the favorable proportions in the treatment and control groups.
(c) The physicians conducting the study had concluded from previous research that aspirin was likely to increase the chance of a favorable outcome. Carry out a significance test to confirm this conclusion. State hypotheses, find the P-value, and write a summary of your results.

8.63 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.

8.64 Suppose that the experiment in 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 8.63. What do you observe about the effect of the sample size on the results of these significance tests?

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

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

8.66 In each of the following circumstances state whether you would use the large-sample confidence interval, the plus four method, or neither for a 95% confidence interval.
(a) $n_1 = 4$, $n_2 = 100$, $X_1 = 1$, and $X_2 = 65$
(b) $n_1 = 500$, $n_2 = 300$, $X_1 = 175$, and $X_2 = 208$
(c) $n_1 = 6$, $n_2 = 10$, $X_1 = 4$, and $X_2 = 2$
(d) $n_1 = 60$, $n_2 = 55$, $X_1 = 24$, and $X_2 = 37$
(e) $n_1 = 200$, $n_2 = 100$, $X_1 = 128$, and $X_2 = 94$

8.67 Suppose there are two binomial populations. For the first, the true proportion of successes is 0.4; for the second, it is 0.5. Consider taking independent samples from these populations, 50 from the first and 60 from the second.
(a) Find the mean and the standard deviation of the distribution of $\hat{p}_1 - \hat{p}_2$.
(b) This distribution is approximately Normal. Sketch this Normal distribution and mark the location of the mean.
(c) Find a value d for which the probability is 0.95 that the difference in sample proportions is within $\pm d$. Mark these values on your sketch.

8.68 In a study of pet ownership, the 595 pet owners and 1939 non–pet owners were classified according to gender. For the pet owners, there were 285 women, while for the non–pet owners, there were 1024 women. Find the proportion of pet owners who were women. Do the same for the non–pet owners. Give a 95% confidence interval for the difference in the two proportions.

8.69 Refer to the previous exercise. Redo it in terms of the proportions of men in each classification. Explain how you could have obtained these results from the calculations you did in Exercise 8.68.

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

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

8.72 Can we use the large-sample confidence interval? In each of the following circumstances state whether you would use the large-sample confidence interval.
(a) $n_1 = 30$, $n_2 = 30$, $X_1 = 10$, and $X_2 = 15$
(b) $n_1 = 15$, $n_2 = 10$, $X_1 = 10$, and $X_2 = 5$
(c) $n_1 = 25$, $n_2 = 20$, $X_1 = 11$, and $X_2 = 8$
(d) $n_1 = 40$, $n_2 = 40$, $X_1 = 20$, and $X_2 = 12$
(e) $n_1 = 50$, $n_2 = 50$, $X_1 = 40$, and $X_2 = 45$

8.73 More on whether to use the large-sample confidence interval. In each of the following circumstances state whether you would use the large-sample confidence interval.
(a) $n_1 = 25$, $n_2 = 25$, $X_1 = 12$, and $X_2 = 8$
(b) $n_1 = 25$, $n_2 = 25$, $X_1 = 17$, and $X_2 = 12$
(c) $n_1 = 60$, $n_2 = 30$, $X_1 = 30$, and $X_2 = 15$
(d) $n_1 = 60$, $n_2 = 55$, $X_1 = 45$, and $X_2 = 37$
(e) $n_1 = 200$, $n_2 = 100$, $X_1 = 128$, and $X_2 = 94$

8.74 Comparing cell phone ownership in 2003 and 2004. In Exercise 8.38, you were asked to compare the 2004 proportion of cell phone owners (89%) with the 2003 estimate (83%). It would be more appropriate to compare these two proportions using the methods of this section. Given that the sample size of each SRS is 1200 students, compare these two years with a significance test, and give an estimate of the difference in proportions of undergraduate cell phone owners with a 95% margin of error. Write a short summary of your results.

8.75 Gender and gambling behaviors among student-athletes. Gambling behaviors of Division I intercollegiate student-athletes were analyzed in Exercises 8.42 and 8.43. Use the methods of this section to compare the males and females with a significance test, and give an estimate of the difference in proportions of student-athletes who participate in any gambling activity with a 95% margin of error. In Exercise 8.42 it is noted that we do not have any information available to assess nonresponse. Consider the possibility that the response rates differ by gender and by whether or not the person participates in any gambling activity. Write a short summary of how these differences might affect inference.

8.76 Pet ownership and marital status. In a study of pet ownership, the 595 pet owners and 1939 non–pet owners were classified according to whether or not they were married. For the pet owners, 53.3% were married, while for the non–pet owners, 57.7% were married. Find a 95% confidence interval for the difference. Write a short summary of your work.

8.77 Pet ownership and marital status: the significance test. In the previous exercise we compared the proportion of pet owners who were married with the proportion of non–pet owners who were married in the Health ABC Study. Use a significance test to make the comparison and summarize the results of your analysis.

8.78 A comparison of the proportion of frequent binge drinkers. In the published report on binge drinking, survey results from both 1993 and 1999 are presented. Using the table below, test whether the proportions of frequent binge drinkers are different at the 5% level. Also construct a 95% confidence interval for the difference. Write a short summary of your results.

Year	n	X
1993	14,995	2,973
1999	13,819	3,140

8.79 A comparison of the proportion of frequent binge drinkers, revisited. Refer to the previous exercise. Redo the exercise in terms of the proportion of nonfrequent binge drinkers in each classification. Explain how you could have obtained these results from the calculations you did in the previous exercise.

Chapter 8 Review Exercises

8.80 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.

8.81 In 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. Suppose that $\hat{p}_1 = 0.6$, $\hat{p}_2 = 0.4$, and n represents the common value of n_1 and n_2. Compute the 95% confidence intervals for the difference in the two proportions for $n = 15$, 25, 50, 75, 100, and 500. For each interval calculate the margin of error. Summarize and explain your results.

8.82 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.

8.83 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.

8.84 "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`.) 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.

8.85 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.

8.86 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.

8.87 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.

8.88 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?

8.89 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.

8.90 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 $p - p_0$.

8.91 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:

	Black	White
Vegetarian	42	135
Not vegetarian	47	88

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?

8.92 A study examined the association between high blood pressure and increased risk of death from cardiovascular disease. There were 2676 men with low blood pressure and 3338 men with high blood pressure. In the low-blood-pressure group, 21 men died from cardiovascular disease; in the high-blood-pressure group, 55 died.
(a) Compute the 95% confidence interval for the difference in proportions.
(b) Do the study data confirm that death rates are higher among men with high blood pressure? State hypotheses, carry out a significance test, and give your conclusions.

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

8.94 Different kinds of companies compensate their key employees in different ways. Established companies may pay higher salaries, while new companies may offer stock options that will be valuable if the company succeeds. Do high-tech companies tend to offer stock options more often than other companies? One study looked at a random sample of 200 companies. Of these, 91 were listed in the *Directory of Public High Technology Corporations* and 109 were not listed. Treat these two groups as SRSs of high-tech and non-high-tech companies. Seventy-three of the high-tech companies and 75 of the non-high-tech companies offered incentive stock options to key employees. Give a 95% confidence interval for the difference in the proportions of the two types of companies that offer stock options. Then compare the two types of companies using a significance test. Be sure to state your hypotheses, the test statistic and the P-value. Write a short summary of your conclusions.

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

8.96 Many new products introduced into the market are targeted toward children. The choice behavior of children with regard to new products is of particular interest to companies that design marketing strategies for these products. As part of one study, children in different age groups were compared on their ability to sort new products into the correct product category (milk or juice). Here are some of the data:

Age group	n	Number who sorted correctly
4- to 5-year-olds	50	10
6- to 7-year-olds	53	28

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

8.97 The Gallup Poll in Exercise 8.95 reported that in a similar poll in 1973, 43% of Americans said that their ideal family would include three or more children. Give the margin of error for this poll, assuming that the sample size was the same. Then compare the proportions with a significance test and give a 95% confidence interval for the difference. Write a summary of your results.

8.98 In a random sample of 875 students from a large public university, it was found that 411 of the students changed majors during their college years.
(a) Give a 95% confidence interval for the proportion of students at this university

who change majors.
(b) Express your results from (a) in terms of the *percent* of students who change majors.
(c) University officials concerned with counseling students are interested in the number of students who change majors rather than the proportion. The university has 37,000 undergraduate students. Convert the confidence interval you found in (a) to a confidence interval for the *number* of students who change majors during their college years.

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

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

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

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

8.103 As the previous problem noted, using the guessed value 0.5 for both $\hat{p}_1$ and $\hat{p}_2$ gives a conservative margin of error in confidence intervals for the difference between

two population proportions. You are planning a survey and will calculate a 95% confidence interval for the difference in two proportions when the data are collected. You would like the margin of error of the interval to be less than or equal to 0.10. You will use the same sample size n for both populations.
(a) How large a value of n is needed?
(b) Give a general formula for n in terms of the desired margin of error m and the critical value z^*.

8.104 What's wrong? For each of the following, explain what is wrong and why.
(a) A 90% confidence interval for the difference in two proportions includes errors due to nonresponse.
(b) A z statistic is used to test the null hypothesis that H_0: $\hat{p}_1 = \hat{p}_2$.
(c) If two sample proportions are equal, then the sample counts must be equal.

8.105 Proportion of male heavy lottery players. A study of state lotteries included a random digit dialing (RDD) survey conducted by the National Opinion Research Center (NORC). The survey asked 2406 adults about their lottery spending. (From Charles T. Clotfelter et al., "State lotteries at the turn of the century: report to the National Gambling Impact Study Commission," 1999.) A total of 248 individuals were classified as "heavy" players. Of these, 152 were male. The study notes that 48.5% of U.S. adults are male. For this analysis, assume that the 248 heavy lottery players are a random sample of all heavy lottery players and that the margin of error for the 48.5% estimate of the percent of males in the U.S. adult population is so small that it can be neglected. Use a significance test to compare the proportion of males among heavy lottery players with the proportion of males in the U.S. adult population. Construct a 95% confidence interval for the proportion. Write a summary of what you have found. Be sure to comment on the possibility that some people may be reluctant to provide information about their lottery spending and how this might affect the results.

8.106 Cell phone ownership: 2000 versus 2004. Refer to Exercise 8.74. The estimated proportion of undergraduates owning a cell phone in 2000 was 43%. We want to test whether the proportion of undergraduate cell phone owners has more than doubled in the last 4 years.
(a) Compute the quantity $\hat{p}_1 - 2\hat{p}_2$ where $\hat{p}_1$ is the 2004 estimate and $\hat{p}_2$ is the 2000 estimate
(b) Using the rules for variances, compute the standard error of this estimate.
(c) Compute the z statistic and P-value. What is your conclusion at the 5% level?

8.107 "No Sweat" garment labels. Following complaints about the working conditions in some apparel factories both in the United States and abroad, a joint government and industry commission recommended in 1998 that companies that monitor and enforce proper standards be allowed to display a "No Sweat" label on their products. Does the presence of these labels influence consumer behavior? A survey of U.S. residents aged 18 or older asked a series of questions about how likely they would be to purchase a garment under various conditions. For some conditions, it was stated that the garment had a "No Sweat" label; for others, there was no mention of such a label. On the basis of the responses, each person was classified as

a "label user" or a "label nonuser." (Marsha A. Dickson, "Utility of no sweat labels for apparel customers: profiling label users and predicting their purchases," *Journal of Consumer Affairs*, 35 (2001), pp. 96–119.) There were 296 women surveyed. Of these, 63 were label users. On the other hand, 27 of 251 men were classified as users.
(a) Give a 95% confidence interval for the difference in the proportions.
(b) You would like to compare the women with the men. Set up appropriate hypotheses, and find the test statistic and the P-value. What do you conclude?

8.108 Education of the customers. To devise effective marketing strategies it is helpful to know the characteristics of your customers. A study compared demographic characteristics of people who use the Internet for travel arrangements and of people who do not. (From Karin Weber and Weley S. Roehl, "Profiling people searching for and purchasing travel products on the World Wide Web," *Journal of Travel Research*, 37 (1999), pp. 291–298.) Of 1132 Internet users, 643 had completed college. Among the 852 nonusers, 349 had completed college.
(a) Do users and nonusers differ significantly in the proportion of college graduates?
(b) Give a 95% confidence interval for the difference in the proportions.

8.109 Income of the customers. The study mentioned in the previous exercise also asked about income. Among Internet users, 493 reported income of less than \$50,000 and 378 reported income of \$50,000 or more. (Not everyone answered the income question.) The corresponding numbers for nonusers were 477 and 200. Perform a significance test to compare the incomes of users with nonusers and also give an estimate of the difference in proportions with a 95% margin of error.

8.110 Nonresponse for the income question. Refer to the previous two exercises. Give the total number of users and the total number of nonusers for the analysis of education. Do the same for the analysis of income. The difference is due to respondents who chose "Rather not say" for the income question. Give the proportions of "Rather not say" individuals for users and nonusers. Perform a significance test to compare these and give a 95% confidence interval for the difference. People are often reluctant to provide information about their income. Do you think that this amount of nonresponse for the income question is a serious limitation for this study?

CHAPTER 9

Chapter 9 Exercises

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:

Rest	January	
of year	Up	Down
Up	35	13
Down	13	14

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):

	County type		
	Urban	Intermediate	Rural
Tetracycline	65	90	172
No tetracycline	149	136	158

(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):

	Nicotine (milligrams/day)		
Alcohol (ounces/day)	None	1–15	16 or more
None	105	7	11
0.01–0.10	58	5	13
0.11–0.99	84	37	42
1.00 or more	57	16	17

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):

	Nutritional status			
Illness	Normal	I	II	III and IV
URI	95	143	144	70
Diarrhea	53	94	101	48
URI and diarrhea	27	60	76	27
None	113	48	44	22
Total	288	345	365	167

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 following table 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):

	Aluminum-containing antacid use			
	None	Low	Medium	High
Alzheimer's patients	112	3	5	8
Control group	114	9	3	2

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 PhD 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:

Status	Men	Women
Completed	423	98
Still enrolled	134	33
Dropped out	238	98

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.

	Survival	
Age group	Dead	Survived
Under 1 year	17	3806
1–4 years	37	7113
5–9 years	3	2208
10–14 years	3	1888
15–19 years	8	2715
20–24 years	6	2209
25–29 years	9	1492
30 years and over	14	1636

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.

Years of school completed, by age (thousands of persons)

	Age group			
Education	25 to 34	35 to 54	55 and over	Total
Did not complete high school	5,325	9,152	16,035	30,512
Completed high school	14,061	24,070	18,320	56,451
College, 1 to 3 years	11,659	19,926	9,662	41,247
College, 4 or more years	10,342	19,878	8,005	38,225
Total	41,388	73,028	52,022	166,438

(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:

Undergraduate college enrollment (thousands of students)				
Age	2-year full-time	2-year part-time	4-year full-time	4-year part-time
Under 18	36	98	75	37
18 to 21	1126	711	3270	223
22 to 34	634	1575	2267	1380
35 and over	221	1092	390	915
Total	2017	3476	6002	2555

(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 2×2 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 4×3 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.

	Marital status				
Age (years)	Never married	Married	Widowed	Divorced	Total
18–24	9,289	3,046	19	260	12,613
25–39	6,948	21,437	206	3,408	32,000
40–64	2,307	26,679	2,219	5,508	36,713
≥ 65	768	7,767	8,636	1,091	18,264
Total	19,312	58,931	11,080	10,266	99,588

(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.

	Gender	
Method	Male	Female
Firearms	16,381	2,559
Poison	3,569	2,110
Hanging	3,824	803
Other	1,641	623

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.

	Alaska Airlines		America West	
	On time	Delayed	On time	Delayed
Los Angeles	497	62	694	117
Phoenix	221	12	4840	415
San Diego	212	20	383	65
San Francisco	503	102	320	129
Seattle	1841	305	201	61

(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:

	Pet ownership	
Patient status	No	Yes
Alive	28	50
Dead	11	3

(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 2×2 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. Malkeil, "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:

	Next year	
This year	Winner	Loser
Winner	85	35
Loser	37	83

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:

	Next year	
This year	Winner	Loser
Winner	96	148
Loser	145	99

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):

	Country			
Taster	Ireland	Portugal	Norway	Italy
Yes	558	345	185	402
No	225	109	81	134

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.

	Ethnic group			
Blood type	Hawaiian	Hawaiian-white	Hawaiian-Chinese	White
O	1,903	4,469	2,206	53,759
A	2,490	4,671	2,368	50,008
B	178	606	568	16,252
AB	99	236	243	5,001

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 milliequivalents per liter) and hyperkalemic if the level is too high (above 5.5 meq/l). 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:

	Potassium group		
	Hypokalemic	Normal	Hyperkalemic
n	1094	4689	27
Hypertension	88.3%	78.1%	40.7%
Heart failure	16.5%	24.7%	55.6%
Diabetes	20.6%	25.5%	29.6%
Gender (% female)	72.5%	68.0%	48.1%

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 2×3 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:

Year	1970	1972	1974	1976	1978	1980	1982	1984	1986
Women	164	195	226	283	302	342	369	385	412

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):

Year	1987	1988	1989	1990	1991	1992	1993
Women	60.0%	60.6%	61.6%	62.4%	63.0%	63.4%	63.2%

Year	1994	1995	1996	1997	1998	1999	2000
Women	63.3%	63.4%	63.8%	64.2%	64.4%	64.9%	65.9%

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.

9.33 The Census Bureau provides estimates of numbers of people in the United States classified in various ways. Let's look at college students. The following table gives us data to examine the relation between age and full-time or part-time status. The numbers in the table are expressed as thousands of U.S. college students.

U.S. college students by age and status		
	Status	
Age	Full-time	Part-time
15–19	3388	389
20–24	5238	1164
25–34	1703	1699
35 and over	762	2045

(a) What is the U.S. Census Bureau estimate of the number of full-time college students aged 15 to 19?
(b) Give the joint distribution of age and status for this table.
(c) What is the marginal distribution of age? Display the results graphically.
(d) What is the marginal distribution of status? Display the results graphically.

9.34 Refer to the previous exercise. Find the conditional distribution of status for each of the four age categories. Display the distributions graphically and summarize their differences and similarities.

9.35 Refer to the previous two exercises. Compute the conditional distribution of age for each of the two status categories. Summarize the distributions graphically and write a short paragraph describing the distributions and how they differ.

9.36 The following table gives some census data concerning the enrollment status of recent high school graduates aged 16 to 24 years. The table entries are in thousands of students.

Enrollment and gender		
Status	Men	Women
Two-year college, full-time	890	969
Two-year college, part-time	340	403
Four-year college, full-time	2897	3321
Four-year college, part-time	249	383
Graduate school	306	366
Vocational school	160	137

(a) How many male recent high school graduates aged 16 to 24 years were enrolled full-time in two-year colleges?
(b) Give the marginal distribution of gender for these students. Display the results graphically.
(c) What is the marginal distribution of status for these students? Display the results graphically.

9.37 Refer to the previous exercise. Find the conditional distribution of gender for each status. Describe the distributions graphically and write a short summary comparing the major features of these distributions.

9.38 Refer to the previous two exercises. Find the conditional distribution of status for each gender. Describe the distributions graphically and write a short summary comparing the major features of these distributions.

9.39 Here are the row and column totals for a two-way table with two rows and two columns:

a	b	200
c	d	200
200	200	400

Find *two different* sets of counts a, b, c, and d for the body of the table that give these same totals. This shows that the relationship between two variables cannot be obtained from the two individual distributions of the variables.

9.40 Construct a 3×3 table of counts where there is no apparent association between the row and column variables.

9.41 A marketing research firm conducted a survey of companies in its state. They mailed a questionnaire to 300 small companies, 300 medium-sized companies, and 300 large companies. The rate of nonresponse is important in deciding how reliable survey results are. Here are the data on response to this survey:

Size of company	Response		Total
	Yes	No	
Small	175	125	300
Medium	145	155	300
Large	120	180	300

(a) What was the overall percent of nonresponse?
(b) Describe how nonresponse is related to the size of the business. (Use percents to make your statements precise.)
(c) Draw a bar graph to compare the nonresponse percents for the three size categories.
(d) Using the total number of responses as a base, compute the percent of responses that come from each of small, medium, and large businesses.
(e) The sampling plan was designed to obtain equal numbers of responses from small, medium, and large companies. In preparing an analysis of the survey results, do you think it would be reasonable to proceed as if the responses represented companies of each size equally?

9.42 A study of the career plans of young women and men sent questionnaires to all 722 members of the senior class in the College of Business Administration at the University of Illinois. One question asked which major within the business program the student had chosen. Here are the data from the students who responded:

Major	Gender	
	Female	Male
Accounting	68	56
Administration	91	40
Economics	5	6
Finance	61	59

(a) Describe the differences between the distributions of majors for women and men with percents, with a graph, and in words.
(b) What percent of the students did not respond to the questionnaire? The nonresponse weakens conclusions drawn from these data.

9.43 Asia has become a major competitor of the United States and Western Europe in education as well as economics. Here are counts of first university degrees in science and engineering in the three regions:

Field	Region		
	United States	Western Europe	Asia
Engineering	61,941	158,931	280,772
Natural science	111,158	140,126	242,879
Social science	182,166	116,353	236,018

Direct comparison of counts of degrees would require us to take into account Asia's much larger population. We can, however, compare the distribution of degrees by field of study in the three regions. Do this using calculations and graphs, and write a brief summary of your findings.

9.44 Mountain View University has professional schools in business and law. Here is a three-way table of applicants to these professional schools, categorized by gender, school, and admission decision.

Business			Law		
	Admit			Admit	
Gender	Yes	No	Gender	Yes	No
Male	400	200	Male	90	110
Female	200	100	Female	200	200

(a) Make a two-way table of gender by admission decision for the combined professional schools by summing entries in the three-way table.
(b) From your two-way table, compute separately the percents of male and female applicants admitted. Male applicants are admitted to Mountain View's professional schools at a higher rate than female applicants.
(c) Now compute separately the percents of male and female applicants admitted by the business school and by the law school.
(d) Explain carefully, as if speaking to a skeptical reporter, how it can happen that Mountain View appears to favor males when this is not true within each of the professional schools.

9.45 Refer to the previous exercise. Make up a similar table for a hypothetical university having four different schools that illustrates the same point. Carefully summarize your table with the appropriate percents.

9.46 A university classifies its classes as either "small" (fewer than 40 students) or "large." A dean sees that 62% of Department A's classes are small, while Department B has only 40% small classes. She wonders if she should cut Department A's budget and insist on larger classes. Department A responds to the dean by pointing out that classes for third- and fourth-year students tend to be smaller than classes for first- and second-year students. The three-way following table gives the counts of classes by department, size, and student audience. Write a short report for the dean that summarizes these data. Start by computing the percents of small classes in the two departments, and include other numerical and graphical comparisons as needed. (Do not perform any statistical significance tests.) Here are the numbers of classes to be analyzed:

	Department A			Department B		
Year	Large	Small	Total	Large	Small	Total
First	2	0	2	18	2	20
Second	9	1	10	40	10	50
Third	5	15	20	4	16	20
Fourth	4	16	20	2	14	16

9.47 For each of the following situations give the degrees of freedom and an appropriate bound on the P-value (give the exact value if you have software available) for the X^2 statistic for testing the null hypothesis of no association between the row and column variables.

(a) a 4 by 5 table with $X^2 = 26.23$
(b) a 4 by 3 table with $X^2 = 26.23$
(c) a 5 by 4 table with $X^2 = 26.23$
(d) a 7 by 6 table with $X^2 = 26.23$

9.48 For each of the following situations give the degrees of freedom and an appropriate bound on the P-value (give the exact value if you have software available) for the X^2 statistic for testing the null hypothesis of no association between the row and column variables.
(a) a 2 by 2 table with $X^2 = 1.31$
(b) a 4 by 4 table with $X^2 = 17.54$
(c) a 2 by 8 table with $X^2 = 22.10$
(d) a 5 by 3 table with $X^2 = 12.61$

9.49 To be competitive in global markets, many U.S. corporations are undertaking major reorganizations. Often these involve "downsizing," sometimes called a "reduction in force" (RIF), where substantial numbers of employees are terminated. Federal and various state laws require that employees be treated equally regardless of their age. In particular, employees over the age of 40 years are in a "protected" class, and many allegations of discrimination focus on comparing employees over 40 with their younger coworkers. Here are the data for a recent RIF:

	Over 40	
Terminated	No	Yes
Yes	16	82
No	585	771

(a) Make a table that includes the following information for each group (over 40 or not): total number of employees, the proportion of employees who were terminated, and the standard error for the proportion.
(b) Perform the chi-square test for this two-way table. Give the test statistic, the degrees of freedom, the P-value, and your conclusion.

9.50 A major issue that arises in the kind of case described in the previous exercise concerns the extent to which the employees are similar. This often involves examining other variables. For example, suppose that the employees over 40 did not do as good a job as the younger workers. Let's examine the last performance appraisal to explore this idea. The possible values are as follows: "partially meets expectations," "fully meets expectations," "usually exceeds expectations," and "continually exceeds expectations." Because there were very few employees who partially exceeded expectations, we combine the first two categories. Here are the data:

	Over 40	
Performance appraisal	No	Yes
Partially or fully meets expectations	82	237
Usually exceeds expectations	357	492
Continually exceeds expectations	63	32

Analyze the data. Do the older employees appear to have lower performance evaluations?

9.51 Can you increase the response rate for a mail survey by contacting the respondents before they receive the survey? A study designed to address this question compared three groups of subjects. The first group received a preliminary letter about the survey, the second group was phoned, and the third received no preliminary contact. A positive response was defined as returning the survey within two weeks. Here are the counts:

	Intervention		
Response	Letter	Phone call	None
Yes	171	146	118
No	220	68	455
Total	391	214	573

Another study also attempted to evaluate the effect of a letter sent before the survey on the response rate. In this study subjects who received a prenotification letter were compared with subjects who received no letter. Here are the data:

Response	Letter	No letter
Yes	2570	2645
No	2448	2384
Total	5018	5029

(a) Summarize the results of each study graphically and numerically.
(b) Perform the appropriate significance tests, and report the results and the conclusions.
(c) Based on these two studies, write a report about methods to improve responses for surveys. The subjects in the first study were students from three Houston universities. The subjects in the second study were physicians in the United States. Be sure to include comments on the extent to which you think the results can be generalized to other populations of people from other places.

9.52 Gastric freezing was once a recommended treatment for ulcers in the upper intestine. A randomized comparative experiment found that 28 of the 82 patients who were subjected to gastric freezing improved, while 30 of the 78 patients in the control group improved. The hypothesis of "no difference" between the two groups can be tested in two ways: using a z statistic or using the X^2 statistic.
(a) State the appropriate hypothesis and a two-sided alternative, and carry out a z test. What is the P-value?
(b) Present the data in a 2×2 table. State the appropriate hypotheses and carry out the chi-square test. What is the P-value? Verify that the X^2 statistic is the square of the z statistic.
(c) What do you conclude about the effectiveness of gastric freezing as a treatment for ulcers?

9.53 At your college 29% of the students are in their first year, 27% in their second, 25% in their third, and 19% in their fourth year. You have taken a survey of students, and when you classify them by year of study, you have 54, 66, 56, and 30 students in the first, second, third, and fourth years, respectively. Use a goodness of fit test to examine how well your sample reflects the population of your college.

9.54 Computer software generated 500 random numbers that should look like they are from the standard Normal distribution. They are categorized into five groups: (1) less than or equal to -0.8, (2) greater than -0.8 and less than or equal to -0.2, (3) greater than -0.2 and less than or equal to 0.2, (4) greater than 0.2 and less than or equal to 0.8, and (5) greater than 0.8. The counts in the five groups are 98, 112, 79, 111, and 100, respectively. Find the probabilities for these five intervals using Table A. Then compute the expected number for each interval for a sample of 500. Finally, perform the goodness of fit test and summarize your results.

9.55 Computer software generated 500 random numbers that should look like they are from the uniform distribution on the interval 0 to 1. They are categorized into five groups: (1) less than or equal to 0.2, (2) greater than 0.2 and less than or equal to 0.4, (3) greater than 0.4 and less than or equal to 0.6, (4) greater than 0.6 and less than or equal to 0.8, and (5) greater than 0.8. The counts in the five groups are 113, 95, 108, 99, and 85, respectively. The probabilities for these five intervals are all the same. What is this probability? Compute the expected number for each interval for a sample of 500. Finally, perform the goodness of fit test and summarize your results.

9.56 Does using Rodham matter? In April 2006, the Opinion Research Corporation conducted a telephone poll for CNN of 1012 adult Americans. (See `i.a.cnn.net/cnn/2006/images/04/27/rel11e.pdf` for a summary of the poll.) Half those polled were asked their opinion of Hillary Rodham Clinton. The other half were asked their opinion of Hillary Clinton. The table below summarizes the results. A chi-square test was used to determine if opinions differed based on the name.

	Opinion			
Name	Favorable	Unfavorable	Never heard of	No opinion
Hillary Rodham Clinton	50%	42%	2%	6%
Hillary Clinton	46%	43%	2%	9%

(a) Computer software gives $X^2 = 4.23$. Can we comfortably use the chi-square distribution to compute the P-value? Explain.
(b) What are the degrees of freedom for X^2?
(c) Give an appropriate bound for the P-value using Table F and state your conclusions.

9.57 New treatment for cocaine addiction. Cocaine addiction is difficult to overcome. Addicts have been reported to have a significant depletion of stimulating neurotransmitters and thus continue to take cocaine to avoid feelings of depression and anxiety. A 3-year study with 72 chronic cocaine users compared an antidepressant drug called desipramine with lithium and a placebo. (Lithium is a standard drug to treat cocaine addiction. A placebo is a substance containing no medication, used so that the effect of being in the study but not taking any drug can be seen.) One-third of the subjects, chosen at random, received each treatment. (Data from D. M. Barnes, "Breaking the cycle of addiction," *Science,* 241 (1988), pp. 1029–1030.) Here are the results:

	Cocaine relapse?	
Treatment	Yes	No
Desipramine	10	14
Lithium	18	6
Placebo	20	4

(a) Compare the effectiveness of the three treatments in preventing relapse using percents and a bar graph. Write a brief summary.
(b) Can we comfortably use the chi-square test to test the null hypothesis that there is no difference between treatments? Explain.
(c) Perform the significance test and summarize the results.

9.58 Drinking status and class attendance. As part of the 1999 College Alcohol Study, students who drank alcohol in the last year were asked if drinking ever resulted in missing a class. (Results of this survey are reported in Henry Wechsler et al., "College Binge Drinking in the 1990s: A Continuing Problem," *Journal of American College Health*, 48 (2000), pp. 199–210.) The data are given in the following table:

	Drinking status		
Missed a class	Nonbinger	Occasional binger	Frequent binger
No	4617	2047	1176
Yes	446	915	1959

(a) Summarize the results of this table graphically and numerically.
(b) What is the marginal distribution of drinking status? Display the results graphically.
(c) Compute the relative risk of missing a class for occasional bingers versus nonbingers and for frequent bingers versus nonbingers. Summarize these results.
(d) Perform the chi-square test for this two-way table. Give the test statistic, degrees of freedom, the P-value, and your conclusion.

9.59 Air pollution from a steel mill. One possible effect of air pollution is genetic damage. A study designed to examine this problem exposed one group of mice to air near a steel mill and another group to air in a rural area and compared the numbers of mutations in each group. (Christopher Somers et al., "Air pollution induces heritable DNA mutations," *Proceedings of the National Academy of Sciences*, 99 (2002), pp. 15,904–15,907.) Here are the data for a mutation at the *Hm-2* gene locus:

	Location	
Mutation	Steel mill air	Rural air
Yes	30	23
No		
Total	96	150

(a) Fill in the missing entries in the table.
(b) Summarize the data numerically and graphically.

(c) Is there evidence to conclude that the location is related to the occurrence of mutations? Perform the significance test and summarize the results.

9.60 Secondhand stores. Shopping at secondhand stores is becoming more popular and has even attracted the attention of business schools. A study of customers' attitudes toward secondhand stores interviewed samples of shoppers at two secondhand stores of the same chain in two cities. The breakdown of the respondents by gender is as follows (William D. Darley, "Store-choice behavior for pre-owned merchandise," *Journal of Business Research,* 27 (1993), pp. 17–31):

Gender	City 1	City 2
Men	38	68
Women	203	150
Total	241	218

Is there a significant difference between the proportions of women customers in the two cities?
(a) State the null hypothesis, find the sample proportions of women in both cities, do a two-sided z test, and give a P-value using Table A.
(b) Calculate the X^2 statistic and show that it is the square of the z statistic. Show that the P-value from Table F agrees (up to the accuracy of the table) with your result from (a).
(c) Give a 95% confidence interval for the difference between the proportions of women customers in the two cities.

9.61 More on secondhand stores. The study of shoppers in secondhand stores cited in the previous exercise also compared the income distributions of shoppers in the two stores. Here is the two-way table of counts:

Income	City 1	City 2
Under \$10,000	70	62
\$10,000 to \$19,999	52	63
\$20,000 to \$24,999	69	50
\$25,000 to \$34,999	22	19
\$35,000 or more	28	24

Verify that the X^2 statistic for this table is $X^2 = 3.955$. Give the degrees of freedom and the P-value. Is there good evidence that customers at the two stores have different income distributions?

9.62 Why are animals brought to animal shelters? Euthanasia of healthy but unwanted pets by animal shelters is believed to be the leading cause of death for cats and dogs. A study designed to find factors associated with bringing a cat to an animal shelter compared data on cats that were brought to an animal shelter with data on cats from the same county that were not brought in. (Gary J. Patronek et al., "Risk factors for relinquishment of cats to an animal shelter," *Journal of the American Veterinary Medical Association,* 209 (1996), pp. 582–588.) One of the factors examined was the source of the cat: the categories were private owner or breeder, pet store, and other (includes born in home, stray, and obtained from a shelter). This kind of study is called a case-control study by epidemiologists. Here are the data:

	Source		
Group	Private	Pet store	Other
Cases	124	16	76
Controls	219	24	203

The same researchers did a similar study for dogs. (Gary J. Patronek et al., "Risk factors for relinquishment of dogs to an animal shelter," *Journal of the American Veterinary Medical Association,* 209 (1996), pp. 572–581.) The data are given in the following table:

	Source		
Group	Private	Pet store	Other
Cases	188	7	90
Controls	518	68	142

(a) Analyze the data for the dogs and the cats separately. Be sure to include graphical and numerical summaries. Is there evidence to conclude that the source of the animal is related to whether or not the pet is brought to an animal shelter?
(b) Write a discussion comparing the results for the cats with those for the dogs.
(c) These data were collected using a telephone interview with pet owners in Mishawaka, Indiana. The animal shelter was run by the Humane Society of Saint Joseph County. The control group data were obtained by a random digit dialing telephone survey. Discuss how these facts relate to your interpretation of the results.

9.63 More on why animals are brought to animal shelters. Refer to the previous exercise concerning the case-control study of factors associated with bringing a cat to an animal shelter and the similar study for dogs. The last category for the source of the pet was given as "Other" and includes born in home, stray, and obtained from a shelter. The following two-way table lists these categories separately for cats:

	Source				
Group	Private	Pet store	Home	Stray	Shelter
Cases	124	16	20	38	18
Controls	219	24	38	116	49

Here is the same breakdown for dogs:

	Source				
Group	Private	Pet store	Home	Stray	Shelter
Cases	188	7	11	23	56
Controls	518	68	20	55	67

Analyze these 2×5 tables and compare the results with those that you obtained for the 2×3 tables in Exercise 9.62. With a large number of cells, the chi-square test sometimes does not have very much power.

9.64 Student loans. A study of 865 college students found that 42.5% had student loans. (Data provided by Susan Prohofsky, from her PhD dissertation, "Selection of undergraduate major: the influence of expected costs and expected benefits," Purdue University, 1991.) The students were randomly selected from the approximately 30,000 undergraduates enrolled in a large public university. The overall purpose of the study was to examine the effects of student loan burdens on the choice of a career. A student with a large debt may be more likely to choose a field where starting salaries are high so that the loan can more easily be repaid. The following table classifies the students by field of study and whether or not they have a loan:

	Student loan	
Field of study	Yes	No
Agriculture	32	35
Child development and family studies	37	50
Engineering	98	137
Liberal arts and education	89	124
Management	24	51
Science	31	29
Technology	57	71

Carry out a complete analysis of the association between having a loan and field of study, including a description of the association and an assessment of its statistical significance.

9.65 Altruism and field of study. In one part of the study described in the previous exercise, students were asked to respond to some questions regarding their interests and attitudes. Some of these questions form a scale called PEOPLE that measures altruism, or an interest in the welfare of others. Each student was classified as low, medium, or high on this scale. Is there an association between PEOPLE score and field of study? Here are the data:

	PEOPLE score		
Field of study	Low	Medium	High
Agriculture	5	27	35
Child development and family studies	1	32	54
Engineering	12	129	94
Liberal arts and education	7	77	129
Management	3	44	28
Science	7	29	24
Technology	2	62	64

Analyze the data and summarize your results. Are there some fields of study that have very large or very small proportions of students in the high-PEOPLE category?

9.66 "No Sweat" label. Following complaints about the working conditions in some apparel factories both in the United States and abroad, a joint government and industry commission recommended in 1998 that companies that monitor and enforce proper standards be allowed to display a "No Sweat" label on their products. Does the presence of these labels influence consumer behavior? A survey of U.S.

residents aged 18 or older asked a series of questions about how likely they would be to purchase a garment under various conditions. For some conditions, it was stated that the garment had a "No Sweat" label; for others, there was no mention of such a label. On the basis of the responses, each person was classified as a "label user" or a "label nonuser." (Marsha A. Dickson, "Utility of no sweat labels for apparel customers: profiling label users and predicting their purchases," *Journal of Consumer Affairs,* 35 (2001), pp. 96–119.) There were 296 women surveyed. Of these, 63 were label users. On the other hand, 27 of 251 men were classified as users.
(a) Construct the 2×2 table of counts for this problem. Include the marginal totals for your table.
(b) Use a X^2 statistic to examine the question of whether or not there is a relationship between gender and use of No Sweat labels. Give the test statistic, degrees of freedom, the P-value, and your conclusion.
(c) You examined this question using the methods of the previous chapter in Exercise 8.107. Verify that if you square the z statistic you calculated for that exercise, you obtain the X^2 statistic that you calculated for this exercise.

CHAPTER 10

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 2.36 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 2.53 gives the following data from a study of two methods for measuring the blood flow in the stomachs of dogs:

Spheres	4.0	4.7	6.3	8.2	12.0	15.9	17.4	18.1	20.2	23.9
Vein	3.3	8.3	4.5	9.3	10.7	16.4	15.4	17.6	21.0	21.7

"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 σ is $s = 1.757$.
(a) Find $\overline{x}$ and $\sum(x_i - \overline{x})^2$ 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):

V	0.50	1.00	1.50	1.80	2.00
I	0.52	1.19	1.62	2.00	2.40

(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 β_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 10.4.
(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.)

Player	1	2	3	4	5	6	7	8	9	10	11	12
Round 1	89	90	87	95	86	81	102	105	83	88	91	79
Round 2	94	85	89	89	81	76	107	89	87	91	88	80

(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 β_0 in this line? What does this number say about overseas returns when the U.S. market is flat (0% return)?
(b) What is β_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`):

Student	1	2	3	4	5	6	7	8
Beers	5	2	9	8	3	7	3	5
BAC	0.10	0.03	0.19	0.12	0.04	0.095	0.07	0.06

Student	9	10	11	12	13	14	15	16
Beers	3	5	4	6	5	7	1	4
BAC	0.02	0.05	0.07	0.10	0.085	0.09	0.01	0.05

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):

Month	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June
Degree-days	15.6	26.8	37.8	36.4	35.5	18.6	15.3	7.9	0.0
Gas consumption	5.2	6.1	8.7	8.5	8.8	4.9	4.5	2.5	1.1

(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 β_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 $\mathrm{SE}_{b_1} = 0.00815$.
(a) Explain briefly what the slope β_1 of the population regression line represents. Then give a 90% confidence interval for β_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 $\mathrm{SE}_{b_0} = 0.20351$.
(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:

HR	94	96	95	95	94	95	94	104	104	106
VO2	0.473	0.753	0.929	0.939	0.832	0.983	1.049	1.178	1.176	1.292

HR	108	110	113	113	118	115	121	127	131
VO2	1.403	1.499	1.529	1.599	1.749	1.746	1.897	2.040	2.231

(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.

Year	Yield	Year	Yield	Year	Yield	Year	Yield
1957	48.3	1968	79.5	1979	109.5	1990	118.5
1958	52.8	1969	85.9	1980	91.0	1991	108.6
1959	53.1	1970	72.4	1981	108.9	1992	131.5
1960	54.7	1971	88.1	1982	113.2	1993	100.7
1961	62.4	1972	97.0	1983	81.1	1994	138.6
1962	64.7	1973	91.3	1984	106.7	1995	113.5
1963	67.9	1974	71.9	1985	118.0	1996	127.1
1964	62.9	1975	86.4	1986	119.4	1997	126.7
1965	74.1	1976	88.0	1987	119.8	1998	134.4
1966	73.1	1977	90.8	1988	84.6	1999	133.8
1967	80.1	1978	101.0	1989	116.3	2000	136.9

(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).

Year	Yield	Year	Yield	Year	Yield	Year	Yield
1957	23.2	1968	26.7	1979	32.1	1990	34.1
1958	24.2	1969	27.4	1980	26.5	1991	34.2
1959	23.5	1970	26.7	1981	30.1	1992	37.6
1960	23.5	1971	27.5	1982	31.5	1993	32.6
1961	25.1	1972	27.8	1983	26.2	1994	41.4
1962	24.2	1973	27.8	1984	28.1	1995	35.3
1963	24.4	1974	23.7	1985	34.1	1996	37.6
1964	22.8	1975	28.9	1986	33.3	1997	38.9
1965	24.5	1976	26.1	1987	33.9	1998	38.9
1966	25.4	1977	30.6	1988	27.0	1999	36.6
1967	24.5	1978	29.4	1989	32.3	2000	38.1

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 10.22 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 σ 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 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 10.16 is equal to the F statistic in your ANOVA table. (Any difference found is due to roundoff error.)
(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:

Month	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June
Degree-days	15.6	26.8	37.8	36.4	35.5	18.6	15.3	7.9	0.0
Gas consumption	5.2	6.1	8.7	8.5	8.8	4.9	4.5	2.5	1.1

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):

Femur	38	56	59	64	74
Humerus	41	63	70	72	84

(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.)

Speed (km/h)	Fuel used (liters/100 km)	Speed (km/h)	Fuel used (liters/100 km)
10	21.00	90	7.57
20	13.00	100	8.27
30	10.00	110	9.03
40	8.00	120	9.87
50	7.00	130	10.79
60	5.90	140	11.77
70	6.30	150	12.83
80	6.95		

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.

10.32 Find a 95% confidence interval for the slope in each of the following settings:
(a) $n = 25$, $\hat{y} = 12.3 + 16.10x$, and $\text{SE}_{b_1} = 8.05$
(b) $n = 25$, $\hat{y} = 1.3 + 6.10x$, and $\text{SE}_{b_1} = 8.05$
(c) $n = 125$, $\hat{y} = 12.3 + 16.10x$, and $\text{SE}_{b_1} = 8.05$

10.33 In Exercise 7.70 we examined the distribution of C-reactive protein (CRP) in a sample of 40 children from Papua New Guinea. Serum retinol values for the same children were studied in Exercise 7.72. One important question that can be addressed with these data is whether or not infections, as indicated by CRP, cause a decrease in the measured values of retinol, low values of which indicate a vitamin A deficiency. The data are given below.
(a) Examine the distributions of CRP and serum retinol. Use graphical and numerical methods.
(b) Forty percent of the CRP values are zero. Does this violate any assumption that we need to do a regression analysis using CRP to predict serum retinol? Explain your answer.
(c) Run the regression, summarize the results, and write a short paragraph explaining you conclusions.
(d) Explain the assumptions needed for your results to be valid. Examine the data with respect to these assumptions and report your results.

C-reactive protein and serum retinol

CRP	Retinol	CRP	Retinol	CRP	Retinol	CRP	Retinol	CRP	Retinol
0.00	1.15	30.61	0.97	22.82	0.24	5.36	1.19	0.00	0.83
3.90	1.36	0.00	0.67	0.00	1.00	0.00	0.94	0.00	1.11
5.64	0.38	73.20	0.31	0.00	1.13	5.66	0.34	0.00	1.02
8.22	0.34	0.00	0.99	3.49	0.31	0.00	0.35	9.37	0.56
0.00	0.35	46.70	0.52	0.00	1.44	59.76	0.33	20.78	0.82
5.62	0.37	0.00	0.70	0.00	0.35	12.38	0.69	7.10	1.20
3.92	1.17	0.00	0.88	4.81	0.34	15.74	0.69	7.89	0.87
6.81	0.97	26.41	0.36	9.57	1.90	0.00	1.04	5.53	0.41

10.34 In Exercise 7.143 we looked at the distribution of osteocalcin (OC), a biomarker for bone formation, in a sample of 31 healthy females aged 11 to 32 years. This biomarker is relatively inexpensive to measure, requiring only a single sample of blood. Measuring bone formation (VO+), on the other hand, is very expensive. Oral and intravenous administration of stable isotopes of calcium are needed, 25 blood samples over a period of two weeks are drawn, and the collection of all urine and fecal samples for two weeks is required. If a biomarker can reliably predict bone formation, then we could avoid the cost of the expensive VO+ measures. Studies designed to assess the effects of interventions intended to increase bone formation could include many more subjects if only the biomarker measurement is needed. The measured values of VO+ and OC for the 31 females in this study are given below.
(a) Use numerical and graphical summaries to describe the distributions of VO+ and OC.
(b) Plot the data. Give a reason for your choice of variables for the x and y axes. Describe the pattern and note any unusual observations. Do the assumptions needed for regression analysis appear to be approximately satisfied?
(c) Run the regression using OC to predict VO+. Summarize the results.

VO+ and osteocalcin

VO+	OC	VO+	OC	VO+	OC	VO+	OC
476	8.1	1032	40.2	624	17.2	285	9.9
694	10.1	445	20.6	479	15.9	403	19.7
753	17.9	896	31.2	572	16.9	391	20.0
687	17.2	968	19.3	512	24.2	513	20.8
628	20.9	985	44.4	838	30.2	878	31.4
1100	38.4	1251	76.5	870	47.7	2221	54.6
1303	54.6	2545	36.4	1606	68.9	1126	77.9
1682	52.8	2240	56.3	1557	35.7		

10.35 In Exercise 7.144 we looked at the distribution of tartrate resistant acid phosphatase (TRAP), a biomarker for bone resorption. The following table gives values for this biomarker and a measure of bone resorption VO–. Analyze these data using the questions in the previous exercise as a guide.

VO– and TRAP

VO–	TRAP	VO–	TRAP	VO–	TRAP	VO–	TRAP
407	3.3	874	5.9	445	6.3	351	6.9
980	8.1	493	8.1	572	8.2	634	8.8
1028	9.0	1116	9.0	857	9.5	536	9.5
701	9.6	934	10.1	477	10.1	254	10.3
766	10.5	496	10.7	924	14.4	954	14.6
918	14.6	1065	14.9	722	18.6	1486	19.0
1018	19.0	2236	19.1	903	19.4	960	23.7
1251	25.2	1761	25.5	1446	28.8		

10.36 Refer to the OC and VO+ data in Exercise 10.34. For variables such as these, it is common to work with the logarithms of the measured values. Reanalyze these data using the logs of both OC and VO+. Summarize your results and compare them with those you obtained in Exercise 10.34.

10.37 Refer to the TRAP and VO– data in Exercise 10.35. Reanalyze these data using the logs of both TRAP and VO–. Summarize your results and compare them with those you obtained in Exercise 10.35.

10.38 Refer to the data in Exercise 2.11. Metatarsus adductus (call it MA) is a turning in of the front part of the foot that is common in adolescents and usually corrects itself. Hallux abducto valgus (call it HAV) is a deformation of the big toe that is not common in youth and often requires surgery. Perhaps the severity of MA can help predict the severity of HAV. Using X-rays, doctors measured the angle of deformity for both MA and HAV. They speculated that there is a positive association—more serious MA is associated with more serious HAV.
(a) Make a scatterplot of the data. (Which is the explanatory variable?)
(b) Describe the form, direction, and strength of the relationship between MA angle and HAV angle. Are there any clear outliers in your graph?
(c) Give a statistical model that provides a framework for asking the question of interest for this problem.
(d) Translate the question of interest into null and alternative hypotheses.
(e) Test these hypotheses and write a short description of the results. Be sure to include the value of the test statistic, the degrees of freedom, the P-value, and a clear statement of what you conclude.

10.39 Refer to the previous exercise. Give a 95% confidence interval for the slope. Explain how this interval can tell you what to conclude from a significance test for this parameter.

10.40 A study reported correlations between several personality traits and scores on the Graduate Record Examination (GRE) for a sample of 342 test takers (D. E. Powers and J. K. Kaufman, “Do standardized multiple choice tests penalize deep-thinking or creative students?” Educational Testing Service Research Report RR–02–15 (2002)). Here is a table of the correlations:

	GRE score		
Personality trait	Analytical	Quantitative	Verbal
Conscientiousness	−0.17	−0.14	−0.12
Rationality	−0.06	−0.03	−0.08
Ingenuity	−0.06	−0.08	−0.02
Quickness	0.21	0.15	0.26
Creativity	0.24	0.26	0.29
Depth	0.06	0.08	0.15

For each correlation, test the null hypothesis that the corresponding true correlation is zero. Reproduce the table and mark the correlations that have $P < 0.001$ with ***, those that have $P < 0.01$ with **, and those that have $P < 0.05$ with *. Some critics of standardized tests have suggested that the tests penalize students who are "deep thinkers" and those who are very creative. Others have suggested that students who work quickly do better on these tests. Write a summary of the results of your significance tests, taking into account these comments.

CHAPTER 11

Chapter 11 Exercises

11.1 One model for subpopulation means for a computer science study is described as

$$\mu_{\text{GPA}} = \beta_0 + \beta_1\text{HSM} + \beta_2\text{HSS} + \beta_3\text{HSE}$$

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).

11.2 Use the model given in the previous exercise to express the subpopulation mean in terms of the parameters β_j for students having high school grade scores HSM = 6 (B−), HSS = 7 (B), and HSE = 8 (B+),

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 was 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 σ?
(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

$$\widehat{GPA} = b_0 + b_1\text{IQ} + b_2\text{C3}$$

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

$$\widehat{GPA} = b_0 + b_1\text{IQ} + b_2\text{C3} + b_3\text{SC}$$

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 Exercises 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 $= (\text{year} - 1978.5)^2$. (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

can happen.
(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?

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.)

Wages	LOS	Size	Wages	LOS	Size	Wages	LOS	Size
48.3355	94	Large	64.1026	24	Large	41.2088	97	Small
49.0279	48	Small	54.9451	222	Small	67.9096	228	Small
40.8817	102	Small	43.8095	58	Large	43.0942	27	Large
36.5854	20	Small	43.3455	41	Small	40.7000	48	Small
46.7596	60	Large	61.9893	153	Large	40.5748	7	Large
59.5238	78	Small	40.0183	16	Small	39.6825	74	Small
39.1304	45	Large	50.7143	43	Small	50.1742	204	Large
39.2465	39	Large	48.8400	96	Large	54.9451	24	Large
40.2037	20	Large	34.3407	98	Large	32.3822	13	Small
38.1563	65	Small	80.5861	150	Large	51.7130	30	Large
50.0905	76	Large	33.7163	124	Small	55.8379	95	Large
46.9043	48	Small	60.3792	60	Large	54.9451	104	Large
43.1894	61	Small	48.8400	7	Large	70.2786	34	Large
60.5637	30	Large	38.5579	22	Small	57.2344	184	Small
97.6801	70	Large	39.2760	57	Large	54.1126	156	Small
48.5795	108	Large	47.6564	78	Large	39.8687	25	Large
67.1551	61	Large	44.6864	36	Large	27.4725	43	Small
38.7847	10	Small	45.7875	83	Small	67.9584	36	Large
51.8926	68	Large	65.6288	66	Large	44.9317	60	Small
51.8326	54	Large	33.5775	47	Small	51.5612	102	Large

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.

11.28 In each of the following settings, give a 95% confidence interval for the coefficients of x_1 and x_2.
(a) $n = 25$, $\hat{y} = 10.6 + 12.1x_1 + 17.3x_2$, $\text{SE}_{b_1} = 7.2$, and $\text{SE}_{b_2} = 4.1$
(b) $n = 103$, $\hat{y} = 15.6 + 12.1x_1 + 7.3x_2$, $\text{SE}_{b_1} = 7.2$, and $\text{SE}_{b_2} = 4.1$

11.29 In each of the following situations, explain what is wrong and why.
(a) The multiple correlation gives the proportion of the variation in the response variable that is explained by the explanatory variables.
(b) In a multiple regression with a sample size of 50 and four explanatory variables, the test statistic for the null hypothesis H_0: $b_2 = 0$ is a t statistic that follows the $t(45)$ distribution when the null hypothesis is true.
(c) One of the assumptions for multiple regression is that the distribution of each explanatory variable should be Normal.

11.30 The fitted regression equation for a multiple regression is

$$\hat{y} = -2.6 + 4.1x_1 - 3.2x_2$$

(a) If $x_1 = 5$ and $x_2 = 3$, what is the predicted value of y?
(b) For the answer to part (a) to be valid, is it necessary that the values $x_1 = 5$ and $x_2 = 3$ correspond to a case in the data set? Explain why or why not.
(c) If you hold x_1 at a fixed value, what is the effect of an increase of one unit in x_2 on the predicted value of y?

11.31 Six explanatory variables are used to predict a response variable using a multiple regression. There are 200 observations.
(a) Write the statistical model that is the foundation for this analysis. Be sure to include a description of all assumptions.
(b) Outline the analysis of variance table giving the sources of variation and numerical values for the degrees of freedom.

11.32 A multiple regression analysis of 54 cases was performed with 3 explanatory variables. Suppose that SSM = 18 and SSE = 100.
(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) What are the degrees of freedom for this statistic?
(c) Find bounds on the P-value using Table E. Give the two entries from the table that you use to determine the bounds.

11.33 Refer to the previous exercise. What proportion of the variation in the response variable is explained by the explanatory variables?

11.34 Using a new software package, you ran a multiple regression. The output reported an F statistic with $P < 0.001$, but none of the t tests for the individual coefficients were significant. Does this mean that there is something wrong with the software? Explain your answer.

11.35 Banks charge different interest rates for different loans. A random sample of 2229 loans made for the purchase of new automobiles was studied to identify variables that explain the interest rate charged. All of these loans were made directly by the bank. A multiple regression was run with interest rate as the response variable and 13 explanatory variables.
(a) The F statistic reported is 71.34. State the null and alternative hypotheses for this statistic. Give the degrees of freedom and the P-value for this test. What do

you conclude?
(b) The value of R^2 is 0.297. What percent of the variation in interest rates is explained by the 13 explanatory variables?
(c) The researchers report a t statistic for each of the regression coefficients. State the null and alternative hypotheses tested by each of these statistics. What are the degrees of freedom for these t statistics? What values of t will lead to rejection of the null hypothesis at the 5% level?
(d) The following table gives the explanatory variables and the t statistics (these are given without the sign, assuming that all tests are two-sided) for the regression coefficients. Which of the explanatory variables are significantly different from zero in this model?

Variable	b	t
Intercept	15.47	
Loan size (in dollars)	−0.0015	10.30
Length of loan (in months)	−0.906	4.20
Percent down payment	−0.522	8.35
Cosigner (0 = no, 1 = yes)	−0.009	3.02
Unsecured loan (0 = no, 1 = yes)	0.034	2.19
Total payments (borrower's monthly installment debt)	0.100	1.37
Total income (borrower's total monthly income)	−0.170	2.37
Bad credit report (0 = no, 1 = yes)	0.012	1.99
Young borrower (0 = older than 25, 1 = 25 or younger)	0.027	2.85
Male borrower (0 = female, 1 = male)	−0.001	0.89
Married (0 = no, 1 = yes)	−0.023	1.91
Own home (0 = no, 1 = yes)	−0.011	2.73
Years at current address	−0.124	4.21

(e) The signs of many of these coefficients are what we might expect before looking at the data. For example, the negative coefficient of loan size means that larger loans get a smaller interest rate. This is very reasonable. Examine the signs of each of the statistically significant coefficients and give a short explanation of what they mean.

11.36 Refer to the previous exercise. The researchers also looked at loans made indirectly, that is, through an auto dealer. They studied 5664 indirect loans. Multiple regression was used to predict the interest rate using the same set of explanatory variables.
(a) The F statistic reported is 27.97. State the null and alternative hypotheses for this statistic. Give the degrees of freedom and the P-value for this test. What do you conclude?
(b) The value of R^2 is 0.141. What percent of the variation in interest rates is explained by the 13 explanatory variables? Compare this value with the percent explained for direct loans in the previous exercise.
(c) The researchers report a t statistic for each of the regression coefficients. State the null and alternative hypotheses tested by each of these statistics. What are the degrees of freedom for these t statistics? What values of t will lead to rejection of the null hypothesis at the 5% level?
(d) The following table gives the explanatory variables and the t statistics (these

are given without the sign, assuming that all tests are two-sided) for the regression coefficients. Which of the explanatory variables are significantly different from zero in this model?

Variable	b	t
Intercept	15.89	
Loan size (in dollars)	−0.0029	17.40
Length of loan (in months)	−1.098	5.63
Percent down payment	−0.308	4.92
Cosigner (0 = no, 1 = yes)	−0.001	1.41
Unsecured loan (0 = no, 1 = yes)	0.028	2.83
Total payments (borrower's monthly installment debt)	−0.513	1.37
Total income (borrower's total monthly income)	0.078	0.75
Bad credit report (0 = no, 1 = yes)	0.039	1.76
Young borrower (0 = older than 25, 1 = 25 or younger)	−0.036	1.33
Male borrower (0 = female, 1 = male)	−0.179	1.03
Married (0 = no, 1 = yes)	−0.043	1.61
Own home (0 = no, 1 = yes)	−0.047	1.59
Years at current address	−0.086	1.73

(e) The signs of many of these coefficients are what we might expect before looking at the data. For example, the negative coefficient of loan size means that larger loans get a smaller interest rate. This is very reasonable. Examine the signs of each of the statistically significant coefficients and give a short explanation of what they mean.

11.37 Refer to the previous two exercises. The authors conclude that banks take higher risks with indirect loans because they do not take into account borrower characteristics when setting the loan rate. Explain how the results of the multiple regressions lead to this conclusion.

11.38 Multiple regressions are sometimes used in litigation. In the case of *Cargill, Inc. v. Hardin,* the prosecution charged that the cash price of wheat was manipulated in violation of the Commodity Exchange Act. In a statistical study conducted for this case, a multiple regression model was constructed to predict the price of wheat using three supply and demand explanatory variables. Data for 14 years were used to construct the regression equation, and a prediction for the suspect period was computed from this equation. The value of R^2 was 0.989. The predicted value was reported as \$2.136 with a standard error of \$0.013. Express the prediction with a 95% interval. (The degrees of freedom were large for this analysis, so use 100 as the df to determine t^*.) The actual price for the period in question was \$2.13. The judge in this case decided that the analysis provided evidence that the price was not artificially depressed, and the opinion was sustained by the court of appeals. Write a short summary of the results of the analysis that relate to the decision and explain why you agree or disagree with it.

11.39 The prevalence of childhood obesity in industrialized nations is constantly rising. Since between 30% and 60% of obese children maintain their obesity into adulthood, there is great interest in better understanding the reasons for this rising trend. In one study, researchers looked at the relationship between a child's percent fat mass and several explanatory variables (C. Maffeis et al., "Distribution of food intake as a risk factor for childhood obesity," *International Journal of Obesity,* 24 (2000), pp. 75–80). These were the percent of energy intake at dinner, each parent's body mass index (BMI), an index of energy intake validity (EI/BMR), and gender. The following table summarizes the results of the multiple regression analysis:

	b	$s(b)$
Intercept	5.13	3.03
Sex (M=0, F=1)	4.69	0.51
Dinner (%)	0.08	0.02
EI/predicted BMR	-1.90	0.65
Mother's BMI (kg/m^2)	0.23	0.07
Father's BMI (kg/m^2)	0.27	0.09

In addition, it is reported that $R = 0.44$ and $F(5, 524) = 25.16$.
(a) How many children were used in this study?
(b) What percent of the variation in percent fat mass is explained by these explanatory variables?
(c) Interpret the sign of each of the regression coefficients given in the table. For EI/predicted BMR, data values ranged between 1.4 and 2.8 with a low value associated with underreporting of energy intake.
(d) Construct a 95% confidence interval for the difference in predicted percent fat mass when energy intake at dinner differs by 5% (assume all other variables are the same).

11.40 What factors predict substance abuse among high school students? One study designed to answer this question collected data from 89 high school seniors in a suburban Florida high school (Miguel A. Diego et al., "Academic performance, popularity, and depression predict adolescent substance abuse," *Adolescence,* 38 (2003), pp. 35–42). One of the response variables was marijuana use, which was rated on a four-point scale. A multiple regression analysis used grade point average (GPA), popularity, and a depression score to predict marijuana use. The results were reported in a table similar to this:

	b	t	P
GPA	−0.597	4.55	< 0.001
Popularity	0.340	2.69	< 0.01
Depression	0.030	2.69	< 0.01

A footnote to the table gives $R^2 = 0.34$, $F(3, 85) = 14.83$, and $P < 0.001$.
(a) State the null and alternative hypotheses that are tested by each of the t statistics. Give the results of these significance tests.
(b) Interpret the sign of each of the regression coefficients given in the table.
(c) In the expression $F(3, 85)$, what do the numbers 3 and 85 represent?
(d) State the null and alternative hypotheses that are tested by the F statistic. What is the conclusion?
(e) Each of the variables in this analysis was measured by having the students complete a questionnaire. Discuss how this might affect the results.

(f) How well do you think that these results can be applied to other populations of high school students?

11.41 Refer to the previous exercise. The researchers also studied cigarette use, alcohol use, and cocaine use. Here is a summary of the results for the individual regression coefficients:

		b	t	P
Cigarette	GPA	-0.340	2.16	< 0.05
	Popularity	0.338	2.24	< 0.05
	Depression	0.034	2.60	< 0.05
Alcohol	GPA	-0.321	3.83	< 0.001
	Popularity	0.185	2.29	< 0.05
	Depression	0.015	2.19	< 0.05
Cocaine	GPA	-0.583	5.99	< 0.001
	Popularity	-1.90	2.25	< 0.05
	Depression	0.002	0.27	Not sig.

And here are other relevant results:

Response variable	R^2	F	P
Cigarette	0.18	6.38	< 0.001
Alcohol	0.27	10.37	< 0.001
Cocaine	0.38	12.21	< 0.001

Using the questions given in the previous exercise, summarize the results for each of these response variables. Then write a short essay comparing the results for the four different response variables.

11.42 A study designed to determine how willing consumers are to pay a premium for non-biotech breakfast cereals (cereals that do not include gene-altered ingredients) included both U.S. and U.K. subjects (Wanki Moon and Siva K. Balasubramanian, "Willingness to pay for non-biotech foods in the U.S. and the U.K.," *Journal of Consumer Affairs,* 37 (2003), pp. 317–339). The response variable was a measure of how much extra they would be willing to pay, and the explanatory variables included items related to perceived risks and benefits, demographic variables, and country. Country was coded as 0 for the U.K. subjects and 1 for the U.S. subjects. The parameter estimate for country was reported as -0.2304 with $t = -4.196$. The total number of subjects was 1810.
(a) Interpret the regression coefficient. Are subjects in the U.S. more willing to pay extra for non-biotech breakfast cereals than U.K. subjects, or are they less willing?
(b) Use the t statistic to find a bound on the P-value. Explain the hypothesis tested by this statistic and summarize the result of the significance test.
(c) The U.S. data were collected using questionnaires that were sent to a nationally representative sample of 5200 households enrolled in the National Dairy Panel (NDP) Group. The response rate was 58%. The same questionnaire was used for an online survey of the 9000 U.K. customers enrolled in another NDP Group. The response rate was 28.5%. Several of the items used in the analysis included "Don't know" as a possible response. Respondents choosing this option were excluded from the analysis. Discuss the implications of these considerations on the results.

11.43 Although the benefits of physical exercise are well known, most people do not exercise and many who start exercise programs drop out after a short time. A study designed to determine factors associated with exercise enjoyment collected data from 282 female volunteers who were participants in not-for-credit aerobic dance classes at two university centers (Steven R. Wininger and David Pargman, "Assessment of factors associated with exercise enjoyment," *Journal of Music Therapy,* 40 (2003), pp. 57–73). Exercise enjoyment was the response variable, with a possible range of 18 to 136. Three explanatory variables were analyzed: satisfaction with the music used (range 4 to 28), satisfaction with the instructor (range 6 to 42), and identity, a variable that measured the extent to which the subject viewed herself as an exerciser. A table of correlations among the four variables was given, and the text noted that all were significant with $P < 0.01$. The coefficients for music (1.02), instructor (0.96), and identity (0.30) were given in another table, where it was noted that $R^2 = 0.33$.
(a) Can you give the fitted regression equation? If your answer is Yes, write the equation; if No, explain what additional information you would need.
(b) Does the fact that all of the correlations between the four variables are significant at $P < 0.01$ tell us that the regression coefficients for each of the three explanatory variables will be statistically significant? Explain your answer.
(c) The statistic for testing the null hypothesis that the population regression coefficients for the three explanatory variables are all zero is $F = 45.64$. Give the degrees of freedom for this statistic, and carry out the significance test. What do you conclude?
(d) What proportion of the variation in exercise enjoyment is explained by music, instructor, and identity?
(e) The authors of the study note that males were not included because there were too few of them in these classes. Do you think that these results would apply to males? Explain why or why not.

11.44 Labels providing nutrition facts give consumers information about the nutritional value of food products that they buy. A study of these labels collected data from 152 consumers who were sent information about a frozen chicken dinner. Each subject was asked to give an overall product nutrition score and also evaluated each of 10 nutrients on a 9-point scale, with higher values indicating that the product has a healthy value for the given nutrient. Composite scores for favorable nutrients (such as protein and fiber) and unfavorable nutrients (such as fat and sodium) were used in a multiple regression to predict the overall product nutrition score (Scot Burton et al., "Implications of accurate usage of nutrition facts panel information for food product evaluations and purchase intentions," *Journal of the Academy of Marketing Science,* 27 (1999), pp. 470–480). The following was reported in a table:

Explanatory variables	b	se	t	Model F	R^2
				33.7**	0.31
Unfavorable nutrients	0.82	0.12	6.8**		
Favorable nutrients	0.57	0.10	5.5**		
Constant	3.33	0.13	26.1**		

**$p < 0.01$

(a) What is the equation of the least-squares line?
(b) Give the null and alternative hypotheses associated with the entry labeled "Model F" and interpret this result.
(c) The column labeled "t" contains three entries. Explain what each of these means.
(d) What are the degrees of freedom associated with the t statistics that you explained in part (c)?

11.45 The product used in the previous exercise was described by the researchers as a poor-nutrition product. The label information for this product had high values for unfavorable nutrients such as fat and low values for favorable nutrients such as fiber. The researchers who conducted this study collected a parallel set of data from subjects who were provided label information for a good-nutrition product. This label had low values for the unfavorable nutrients and high values for the favorable ones. The same type of multiple regression model was run for the 162 consumers who participated in this part of the study. Here are the regression results:

Explanatory variables	b	se	t	Model F	R^2
				44.0**	0.36
Unfavorable nutrients	0.86	0.12	6.9**		
Favorable nutrients	0.66	0.10	6.9**		
Constant	3.96	0.12	32.8**		

**$p < 0.01$

For this analysis, answer the questions in parts (a) to (d) of the previous exercise.

11.46 Refer to the previous two exercises. When the researchers planned these studies, they expected both unfavorable nutrients and favorable nutrients to be positively associated with the overall product nutrition score. They also expected the unfavorable nutrients to have a stronger effect. Examine the regression coefficients and the associated t statistics for the two regression models. Then, use this information to discuss how well the researchers' expectations were fulfilled.

The following five exercises use the data given in the next exercise.

11.47 Online stock trading has increased dramatically during the past several years. An article discussing this new method of investing provided data on the major Internet stock brokerages who provide this service (Alan Levinsohn, "Online brokerage, the new core account?" *ABA Banking Journal,* September 1999, pp. 34–42). Following are some data for the top 10 Internet brokerages. The variables are Mshare, the market share of the firm; Accts, the number of Internet accounts in thousands; and Assets, the total assets in billions of dollars. These firms are not a random sample from any population but we will use multiple regression methods to develop statistical models that relate assets to the other two variables.

ID	Broker	Mshare	Accts	Assets
1	Charles Schwab	27.5	2500	219.0
2	E*Trade	12.9	909	21.1
3	TD Waterhouse	11.6	615	38.8
4	Datek	10.0	205	5.5
5	Fidelity	9.3	2300	160.0
6	Ameritrade	8.4	428	19.5
7	DLJ Direct	3.6	590	11.2
8	Discover	2.8	134	5.9
9	Suretrade	2.2	130	1.3
10	National Discount Brokers	1.3	125	6.8

(a) Plot assets versus accounts and describe the relationship.
(b) Perform a simple linear regression to predict assets from the number of accounts. Give the least-squares line and the results of the hypothesis test for the slope.
(c) Obtain the residuals from part (b) and plot them versus accounts. Describe the plot. What do you conclude?
(d) Construct a new variable that is the square of the number of accounts. Rerun the regression analysis with accounts and the square as explanatory variables. Summarize the results.

11.48 In the multiple regression you performed in the previous exercise, the P-value for the number of accounts was 0.8531, while the P-value for the square was 0.0070. Unless we have a strong theoretical reason for considering a model with a quadratic term and no linear term, we prefer not to do this. One problem with these two explanatory variables is that they are highly correlated. Here is a way to construct a version of the quadratic term that is less correlated with the linear term. We first find the mean for accounts, and then we subtract this value from accounts before squaring. The mean is 793.6, so the new quadratic explanatory variable will be $(\text{Accts} - 793.6)^2$. Run the multiple regression to predict assets using accounts and the new quadratic term. Compare these results with what you found in the previous exercise.

11.49 To one person, the plot of assets versus the number of accounts indicates that the relationship is curved. Another person might see this as a linear relationship with two outliers. Identify the two outliers and rerun the linear regression and the multiple regression with the linear and quadratic terms. Summarize your results.

11.50 Sometimes we attempt to model curved relationships by transforming variables. Take the logarithm of assets and the logarithm of the number of accounts. Does the relationship between the logs appear to be approximately linear? Analyze the data and provide a summary of your results. Be sure to include plots along with the results of your statistical inference.

11.51 Recall that the relationship between an explanatory variable and a response variable can depend on what other explanatory variables are included in the model.
(a) Use a simple linear regression to predict assets using the number of accounts. Give the regression equation and the results of the significance test for the regression coefficient.

(b) Do the same using market share to predict assets.
(c) Run a multiple regression using both the number of accounts and market share to predict assets. Give the multiple regression equation and the results of the significance tests for the two regression coefficients.
(d) Compare the results in parts (a), (b), and (c). If you had to choose one of these three models, which one do you prefer? Give an explanation for your answer.

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):

Nematodes	Seedling growth			
0	10.8	9.1	13.5	9.2
1,000	11.1	11.1	8.2	11.3
5,000	5.4	4.6	7.4	5.0
10,000	5.8	5.3	3.2	7.5

(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):

Form	n	$\overline{x}$	s	Min.	Q_1	Median	Q_3	Max.
1	79	31.78	4.45	18	29	32	35	42
2	81	32.88	4.40	20	30	33	36	42
3	81	34.47	4.29	24	32	35	38	46

Here is the SAS output for a one-way ANOVA run on the exam scores:

```
                                     Sum of          Mean
Source                      DF      Squares        Square  F Value    Pr > F
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

1     - 3          -4.3543     -2.6843     -1.0144   ***
1     - 2          -2.7617     -1.0917      0.5782
```

(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):

Year	n	$\overline{x}$	s	Min.	Q_1	Median	Q_3	Max.
76	59	6.80	.58	5.74	6.33	6.73	7.32	8.05
77	58	6.75	.68	3.95	6.39	6.80	7.23	8.10
78	58	6.76	.50	5.01	6.50	6.78	7.10	7.66
82	68	6.50	.55	5.15	6.11	6.53	6.83	7.82
87	70	6.40	.68	4.38	6.09	6.46	6.85	8.15

Here is the SAS output for a one-way ANOVA run on the logs of the lead concentrations:

```
                                   Sum of          Mean
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 '***'.

                         Simultaneous                Simultaneous
                            Lower     Difference        Upper
          YEAR            Confidence    Between      Confidence
       Comparison           Limit        Means          Limit

      76    - 78          -0.2745       0.0424         0.3592
      76    - 77          -0.2687       0.0482         0.3650
      76    - 82          -0.0046       0.3003         0.6052
      76    - 87           0.0995       0.4024         0.7052   ***

      78    - 76          -0.3592      -0.0424         0.2745
      78    - 77          -0.3124       0.0058         0.3240
      78    - 82          -0.0484       0.2579         0.5642
      78    - 87           0.0557       0.3600         0.6643   ***

      77    - 76          -0.3650      -0.0482         0.2687
      77    - 78          -0.3240      -0.0058         0.3124
      77    - 82          -0.0542       0.2521         0.5584
      77    - 87           0.0499       0.3542         0.6585   ***

      82    - 76          -0.6052      -0.3003         0.0046
      82    - 78          -0.5642      -0.2579         0.0484
      82    - 77          -0.5584      -0.2521         0.0542
      82    - 87          -0.1897       0.1021         0.3938

      87    - 76          -0.7052      -0.4024        -0.0995   ***
      87    - 78          -0.6643      -0.3600        -0.0557   ***
      87    - 77          -0.6585      -0.3542        -0.0499   ***
      87    - 82          -0.3938      -0.1021         0.1897
```

(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 12.1, 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 12.2, 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 nematode experiment described in Exercise 12.3. 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?

12.12 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 12.3. 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 12.3. Compare these results with those obtained by analyzing the raw data.

12.13 You are planning a study of the SAT mathematics scores of four groups of students. From a previous study, 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 group) 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.

12.14 Refer to the previous exercise. Repeat all parts for the alternative $\mu_1 = 610$, $\mu_2 = 600$, $\mu_3 = 590$, and $\mu_4 = 580$.

12.15 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 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.

12.16 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.

12.17 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.

12.18 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 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?

12.20 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.

12.21 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.

12.22 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.

12.23 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 following table gives data on the number of cereal leaf beetles trapped:

Color	Insects trapped					
Lemon yellow	45	59	48	46	38	47
White	21	12	14	17	13	17
Green	37	32	15	25	39	41
Blue	16	11	20	21	14	7

(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?

12.24 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?

12.25 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 PhD dissertation, "A multivariate study of exercise training effects on beta-endorphin and emotionality in psychologically normal, medically healthy men," Purdue University, 1983.) 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:

Source	Degrees of freedom	Sum of squares	Mean square	F
Groups	3	104,855.87		
Error	32	70,500.59		
Total				

(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?

12.26 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:

Source	Degrees of freedom	Sum of squares	Mean square	F
Groups	3		158.96	
Error	32		62.81	
Total				

(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?

12.27 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:

Country	n	$\overline{x}$	s
Egypt	46	3.7	2.5
Kenya	111	3.1	1.8
Mexico	52	2.9	1.8

(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.

12.28 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.

12.29 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:

Country	n	$\overline{x}$	s
Egypt	88	1217	327
Kenya	91	844	184
Mexico	54	1119	285

(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.

12.30 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?

12.31 In the exercise program study described in Exercise 12.25, the summary statistics for physical fitness scores are as follows:

Group	n	$\overline{x}$	s
Treatment (T)	10	291.91	38.17
Control (C)	5	308.97	32.07
Joggers (J)	11	366.87	41.19
Sedentary (S)	10	226.07	63.53

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.

12.32 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.

12.33 Exercise 12.26 gives the ANOVA table for depression scores from the exercise program study described in Exercise 12.25. Here are the summary statistics for the depression scores:

Group	n	$\overline{x}$	s
Treatment (T)	10	51.90	6.42
Control (C)	5	57.40	10.46
Joggers (J)	11	49.73	6.27
Sedentary (S)	10	58.20	9.49

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.

12.34 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.

12.35 You are planning a study of the weight gains of pregnant women during the third trimester of pregnancy similar to that described in Exercise 12.27. 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 alternative $\mu_1 = 2.6$, $\mu_2 = 3.0$, 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, and 200. What sample size would you choose for your study?

12.36 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?

12.37 Refer to the color attractiveness experiment described in Exercise 12.23. 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?

12.38 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 12.23. Compare these results with those obtained by analyzing the raw data.

12.39 For each of the following, explain what is wrong and why.
(a) Use one-way ANOVA when the response variable has only two possible values.

(b) You cannot use one-way ANOVA when there are more than three means to be compared.
(c) The pooled estimate s_p is a parameter of the ANOVA model.

12.40 For each of the following, explain what is wrong and why.
(a) The ANOVA F statistic tests the null hypothesis that the three sample means are equal.
(b) The mean squares in an ANOVA table will add, that is, $\text{MST} = \text{MSG} + \text{MSE}$.
(c) Within-group variation is the variation in the data due to the differences in the sample means.

12.41 A study compared 4 groups with 11 observations per group. An F statistic of 3.52 was reported.
(a) Give the degrees of freedom for this statistic and the entries from Table E that correspond to this distribution.
(b) Sketch a picture of this F distribution with the information from the table included.
(c) Based on the table information, how would you report the P-value?
(d) Can you conclude that all pairs of means are different? Explain your answer.

12.42 For each of the following situations, state how large the F statistic needs to be for rejection of the null hypothesis at the 0.05 level.
(a) Compare 6 groups with 2 observations per group.
(b) Compare 6 groups with 4 observations per group.
(c) Compare 6 groups with 11 observations per group.
(d) Summarize what you have learned about F distributions from this exercise.

12.43 For each of the following situations, find the F statistic and the degrees of freedom. Then draw a sketch of the distribution under the null hypothesis and shade in the portion corresponding to the P-value. State how you would report the P-value.
(a) Compare 5 groups with 8 observations per group, $\text{MSE} = 50$, and $\text{MSG} = 57$.
(b) Compare 3 groups with 7 observations per group, $\text{SSG} = 40$, and $\text{SSE} = 90$.

12.44 For each of the following situations, draw a picture of the ANOVA model. Use numerical values for the μ_i. To sketch the Normal curves, you may want to review the 68–95–99.7 rule.
(a) $\mu_1 = 10$, $\mu_2 = 25$, $\mu_3 = 24$, and $\sigma = 5$.
(b) $\mu_1 = 10$, $\mu_2 = 20$, $\mu_3 = 30$, $\mu_4 = 30.1$, and $\sigma = 5$.
(c) $\mu_1 = 10$, $\mu_2 = 25$, $\mu_3 = 24$, and $\sigma = 2$.

12.45 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) Last semester, an alcohol awareness program was conducted for three groups of students at an eastern university. Follow-up questionnaires were sent to the participants two months after each presentation. There were 220 responses from students in an elementary statistics course, 145 from a health and safety course, and 76 from a cooperative housing unit. One of the questions was "Did you discuss the presentation with any of your friends?" The answers were rated on a five-point scale with

1 corresponding to "not at all" and 5 corresponding to "a great deal."
(b) A study examined the effects of consuming three different varieties of onions on blood cholesterol levels. Five cross-bred (Large White by Landrace) pigs were randomly assigned to each treatment.
(c) A researcher for a video game developer researcher wanted to evaluate a prototype of a new game. Free copies were given to 25 eighth-graders, 25 third-year high school students, and 25 second-year college students who were regular video game players. Two weeks later they were asked to tell how much they liked the game on a 1 to 10 scale.

12.46 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) The effects of six different treatments designed to make fabrics stronger were compared. A batch of 120 samples of cloth was available for the experiment, and an equal number were randomly assigned to each of the treatments. The breaking strength of each cloth sample was measured.
(b) A waiter designed a study to see the effects of his behaviors on the amount of tips that he received. For some customers, he would tell a joke; for others, he would describe two of the food items as being particularly good that night; and for others he would behave normally. Using a table of random numbers, he assigned equal numbers of his next 30 customers to his different behaviors.
(c) A supermarket wants to compare the effects of providing free samples of cheddar cheese on sales. An experiment will be conducted from 5:00 P.M. to 6:00 P.M. for the next 20 weekdays. On each day, customers will be offered one of the following: a small cube of cheese pierced by a toothpick, a small slice of cheese on a cracker, a cracker with no cheese, or nothing.

12.47 Refer to Exercise 12.45. 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.

12.48 Refer to Exercise 12.46. 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.

12.49 Refer to Exercise 12.45. For each situation, discuss the method of obtaining the data and how this would affect the extent to which the results could be generalized.

12.50 Refer to Exercise 12.46. For each situation, discuss the method of obtaining the data and how this would affect the extent to which the results could be generalized.

12.51 An experiment was run to compare three groups. The sample sizes were 20, 18, and 15, and the corresponding estimated standard deviations were 220, 190, and 200.
(a) Is it reasonable to use the assumption of equal standard deviations when we

analyze these data? Give a reason for your answer.
(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?

12.52 Does bread lose its vitamins when stored? Small loaves of bread were prepared with flour that was fortified with a fixed amount of vitamins. After baking, the vitamin C content of two loaves was measured. Another two loaves were baked at the same time, stored for one day, and then the vitamin C content was measured. In a similar manner, two loaves were stored for three, five, and seven days before measurements were taken. The units are milligrams of vitamin C per hundred grams of flour (mg/100 g) (H. Park et al., "Fortifying bread with each of three antioxidants," *Cereal Chemistry,* 74 (1997), pp. 202–206). Here are the data:

Condition	Vitamin C (mg/100 g)	
Immediately after baking	47.62	49.79
One day after baking	40.45	43.46
Three days after baking	21.25	22.34
Five days after baking	13.18	11.65
Seven days after baking	8.51	8.13

(a) Give a table with sample size, mean, standard deviation, and standard error for each condition.
(b) Perform a one-way ANOVA for these data. Be sure to state your hypotheses, the test statistic with degrees of freedom, and the P-value.
(c) Summarize the data and the means with a plot. Use the plot and the ANOVA results to write a short summary of your conclusions.

12.53 Refer to the previous exercise. Use the Bonferroni or another multiple-comparisons procedure to compare the group means. Summarize the results.

12.54 Refer to Exercise 12.52. Measurements of the amounts of vitamin A (beta-carotene) and vitamin E in each loaf are given below. Use the analysis of variance method to study the data for each of these vitamins.

Condition	Vitamin A (mg/100 g)		Vitamin E (mg/100 g)	
Immediately after baking	3.36	3.34	94.6	96.0
One day after baking	3.28	3.20	95.7	93.2
Three days after baking	3.26	3.16	97.4	94.3
Five days after baking	3.25	3.36	95.0	97.7
Seven days after baking	3.01	2.92	92.3	95.1

12.55 Refer to the previous exercise.
(a) Explain why it is inappropriate to perform a multiple-comparisons analysis for the vitamin E data.
(b) Perform the Bonferroni or another multiple-comparisons procedure for the vitamin A data and summarize the results.

12.56 Refer to Exercises 12.52 to 12.55. Write a report summarizing what happens to vitamins A, C, and E after bread is baked. Include appropriate statistical inference results and graphs.

12.57 The air in poultry-processing plants often contains fungus spores. If the ventilation is inadequate, this can affect the health of the workers. To measure the presence of spores, air samples are pumped to an agar plate, and "colony-forming units (CFUs)" are counted after an incubation period. Here are data from the "kill room" of a plant that slaughters 37,000 turkeys per day, taken at four seasons of the year. The units are CFUs per cubic meter of air (Michael Wayne Peugh, "Field investigation of ventilation and air quality in duck and turkey slaughter plants," MS thesis, Purdue University, 1996).

Fall	Winter	Spring	Summer
1231	384	2105	3175
1254	104	701	2526
1088	97	842	1090

(a) Examine the data using graphs and descriptive measures. How do airborne fungus spores vary with the seasons?
(b) Is the effect of season statistically significant?

12.58 Refer to the previous exercise. There is not sufficient information to examine the distributions in detail, but it is not unreasonable to expect count data such as these to be skewed. Reanalyze the data after taking logs of the CFU counts. Summarize your work and compare the results you have found here with what you obtained in the previous exercise.

12.59 If a supermarket product is offered at a reduced price frequently, do customers expect the price of the product to be lower in the future? This question was examined by researchers in a study conducted on students enrolled in an introductory management course at a large midwestern university. For 10 weeks 160 subjects received information about the products. The treatment conditions corresponded to the number of promotions (1, 3, 5, or 7) that were described during this 10-week period. Students were randomly assigned to four groups (based on M. U. Kalwani and C. K. Yim, "Consumer price and promotion expectations: an experimental study," *Journal of Marketing Research,* 29 (1992), pp. 90–100). The table below gives the data.

Price promotion data

Number of promotions	Expected price (dollars)									
1	3.78	3.82	4.18	4.46	4.31	4.56	4.36	4.54	3.89	4.13
	3.97	4.38	3.98	3.91	4.34	4.24	4.22	4.32	3.96	4.73
	3.62	4.27	4.79	4.58	4.46	4.18	4.40	4.36	4.37	4.23
	4.06	3.86	4.26	4.33	4.10	3.94	3.97	4.60	4.50	4.00
3	4.12	3.91	3.96	4.22	3.88	4.14	4.17	4.07	4.16	4.12
	3.84	4.01	4.42	4.01	3.84	3.95	4.26	3.95	4.30	4.33
	4.17	3.97	4.32	3.87	3.91	4.21	3.86	4.14	3.93	4.08
	4.07	4.08	3.95	3.92	4.36	4.05	3.96	4.29	3.60	4.11
5	3.32	3.86	4.15	3.65	3.71	3.78	3.93	3.73	3.71	4.10
	3.69	3.83	3.58	4.08	3.99	3.72	4.41	4.12	3.73	3.56
	3.25	3.76	3.56	3.48	3.47	3.58	3.76	3.57	3.87	3.92
	3.39	3.54	3.86	3.77	4.37	3.77	3.81	3.71	3.58	3.69
7	3.45	3.64	3.37	3.27	3.58	4.01	3.67	3.74	3.50	3.60
	3.97	3.57	3.50	3.81	3.55	3.08	3.78	3.86	3.29	3.77
	3.25	3.07	3.21	3.55	3.23	2.97	3.86	3.14	3.43	3.84
	3.65	3.45	3.73	3.12	3.82	3.70	3.46	3.73	3.79	3.94

(a) Make a Normal quantile plot for the data in each of the four treatment groups. Summarize the information in the plots and draw a conclusion regarding the Normality of these data.
(b) Summarize the data with a table containing the sample size, mean, standard deviation, and standard error for each group.
(c) Is the assumption of equal standard deviations reasonable here? Explain why or why not.
(d) Run the one-way ANOVA. Give the hypotheses tested, the test statistic with degrees of freedom, and the P-value. Summarize your conclusion.

12.60 Refer to the previous exercise. Use the Bonferroni or another multiple-comparisons procedure to compare the group means. Summarize the results and support your conclusions with a graph of the means.

12.61 Recommendations regarding how long infants in developing countries should be breast-fed are controversial. If the nutritional quality of the breast milk is inadequate because the mothers are malnourished, then there is risk of inadequate nutrition for the infant. On the other hand, the introduction of other foods carries the risk of infection from contamination. Further complicating the situation is the fact that companies that produce infant formulas and other foods benefit when these foods are consumed by large numbers of customers. One question related to this controversy concerns the amount of energy intake for infants who have other foods introduced into the diet at different ages. Part of one study compared the energy intakes, measured in kilocalories per day (kcal/d), for infants who were breast-fed exclusively for 4, 5, or 6 months (based on a study by J. E. Stuff and B. L. Nichols reported in Chelsea Lutter, "Recommended length of exclusive breast-feeding, age of introduction of complementary foods and the weaning dilemma," World Health Organization, 1992). Here are the data:

Breast-fed for:	Energy intake (kcal/d)									
4 months	499	620	469	485	660	588	675	517	649	209
	404	738	628	609	617	704	558	653	548	
5 months	490	395	402	177	475	617	616	587	528	518
	370	431	518	639	368	538	519	506		
6 months	585	647	477	445	485	703	528	465		

(a) Make a table giving the sample size, mean, and standard deviation for each group of infants. Is it reasonable to pool the variances?
(b) Run the analysis of variance. Report the F statistic with its degrees of freedom and P-value. What do you conclude?

12.62 Refer to the previous exercise.
(a) Examine the residuals. Is the Normality assumption reasonable for these data?
(b) Explain why you do not need to use a multiple-comparisons procedure for these data.

12.63 Refer to Exercise 12.52, where we studied the effects of storage on the vitamin C content of bread. In this experiment 64 mg of vitamin C per 100 g of flour was added to the flour that was used to make each loaf.
(a) Convert the vitamin C amounts (mg/100 g) to percents of the amounts originally in the loaves by dividing the amounts in Exercise 12.52 by 64 and multiplying by 100. Calculate the transformed means, standard deviations, and standard errors and summarize them with the sample sizes in a table.
(b) Explain how you could have calculated the table entries directly from the table you gave in part (a) of Exercise 12.52.
(c) Analyze the percents using analysis of variance. Compare the test statistic, degrees of freedom, P-value, and conclusion you obtain here with the corresponding values that you found in Exercise 12.52.

12.64 Refer to the previous exercise and Exercise 12.54. The flour used to make the loaves contained 5 mg of vitamin A per 100 g of flour and 100 mg of vitamin E per 100 g of flour. Summarize the effects of transforming the data to percents for all three vitamins.

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 S1 = 0.015 inches by 2 inches and

S2 = 0.025 inches by 2 inches. (Data provided by Michael Hunt and Bob Lattanzi of the Purdue University Forestry Department.)

	Size of flakes	
Species	S1	S2
Aspen	308	278
	428	398
	426	331
Birch	214	534
	433	512
	231	320
Maple	272	158
	376	503
	322	220

(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 S1 = 0.015 inches by 2 inches and S2 = 0.025 inches by 2 inches.

	Size of flakes	
Species	S1	S2
Aspen	1296	1472
	1997	1441
	1686	1051
Birch	903	1422
	1246	1376
	1355	1238
Maple	1211	1440
	1827	1238
	1541	748

(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.
(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:

	35–39	40–44	45–49	50–54	55–59
White	131.0	132.3	135.2	139.4	142.0
Nonwhite	132.3	134.2	137.2	141.3	144.1

(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:

	35–39	40–44	45–49	50–54	55–59
White	89.4	90.2	90.9	91.6	91.4
Nonwhite	91.2	93.1	93.3	94.5	93.5

(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 non-white 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 following table. 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.

	Group	
Gender	PG	LU
Males	27.56	25.34
Females	29.25	24.94

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:

Source	Degrees of freedom	Sum of squares	Mean square	F
A (Gender)		62.40		
B (Group)		1,599.03		
AB				
Error		13,633.29		
Total		15,458.52		

(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 13.15, summarize the results of this study.

13.17 For each of the following, explain what is wrong and why.
(a) The FIT part of the model in a two-way ANOVA represents the variation that

is sometimes called error or residual.
(b) You should reject the null hypothesis that there is no interaction in a two-way ANOVA when the test statistic is small.
(c) Sums of squares are equal to mean squares divided by degrees of freedom.

13.18 For each of the following, explain what is wrong and why.
(a) The significance tests for the main effects in a two-way ANOVA have a chi-square distribution when the null hypothesis is true.
(b) You can perform a two-way ANOVA only when the sample sizes are the same in all cells.
(c) The cell means $\bar{x}_{ij}$ are parameters of the two-way ANOVA model.

13.19 A 2×4 ANOVA was run with 5 observations per cell.
(a) Give the degrees of freedom for the F statistic that is used to test for interaction in this analysis and the entries from Table E that correspond to this distribution.
(b) Sketch a picture of this distribution with the information from the table included.
(c) The calculated value of the F statistic is 2.59. How would you report the P-value?
(d) Would you expect a plot of the means to look parallel? Explain your answer.

13.20 For each of the following situations, state how large the F statistic needs to be for rejection of the null hypothesis at the 5% level. Sketch each distribution and indicate the region where you would reject.
(a) the main effect for the first factor in a 3×3 ANOVA with 4 observations per cell
(b) the interaction in a 3×3 ANOVA with 4 observations per cell
(c) the interaction in a 2×2 ANOVA with 251 observations per cell

13.21 Analysis of data for a 3×2 ANOVA with 5 observations per cell gave the F statistics in the following table:

Effect	F
A	1.21
B	3.63
AB	2.04

What can you conclude from the information given?

13.22 Does repetition of an advertising message increase its effectiveness? One theory suggests that there are two phases in the process. In the first phase, called "wearin," negative or unfamiliar views are transformed into positive views. In the second phase, called "wearout," the effectiveness of the ad is decreased because of boredom or other factors. One study designed to investigate this theory examined two factors. The first was familiarity of the ad, with two levels, familiar and unfamiliar; the second was repetition, with three levels, 1, 2, and 3 (Margaret C. Campbell and Kevin Lane Keller, "Brand familiarity and advertising repetition effects," *Journal of Consumer Research,* 30 (2003), pp. 292–304). One of the response variables collected was attitude toward the ad. This variable was the average of four items, each measured on a seven-point scale, anchored by bad–good, low quality–high quality, unappealing–appealing, and unpleasant–pleasant. Here are the means for attitude:

	Repetition		
Familiarity	1	2	3
Familiar	4.56	4.73	5.24
Unfamiliar	4.14	5.26	4.41

(a) Make a plot of the means and describe the patterns that you see.
(b) Does the plot suggest that there is an interaction between familiarity and repetition? If your answer is Yes, describe the interaction.

13.23 Refer to the previous exercise. In settings such as this, researchers collect data for several response variables. For this study, they also constructed variables that were called attitude toward the brand, total thoughts, support arguments, and counterarguments. Here are the means:

	Attitude to brand			Total			Support			Counter		
	Repetition			Repetition			Repetition			Repetition		
Familiarity	1	2	3	1	2	3	1	2	3	1	2	3
Familiar	4.67	4.65	5.06	1.33	1.93	2.55	0.63	0.67	0.98	0.54	0.70	0.49
Unfamiliar	3.94	4.79	4.26	1.52	3.06	3.17	0.76	1.40	0.64	0.52	0.75	1.14

For each of the four response variables, give a graphical summary of the means. Use this summary to discuss any interactions that are evident. Write a short report summarizing the effect of repetition on the response variables measured, using the data in this exercise and the previous one.

13.24 Refer to the previous exercise. Here are the standard deviations for attitude toward brand:

	Repetition		
Familiarity	1	2	3
Familiar	1.16	1.46	1.16
Unfamiliar	1.39	1.22	1.42

Find the pooled estimate of the standard deviation for these data. Use the rule for examining standard deviations in ANOVA from Chapter 12 to determine if it is reasonable to use a pooled standard deviation for the analysis of these data.

13.25 Refer to Exercise 13.23. Here are the standard deviations for total thoughts:

	Repetition		
Familiarity	1	2	3
Familiar	1.63	1.42	1.52
Unfamiliar	1.64	2.16	1.59

Find the pooled estimate of the standard deviation for these data. Use the rule for examining standard deviations in ANOVA from Chapter 12 to determine if it is reasonable to use a pooled standard deviation for the analysis of these data.

13.26 Refer to Exercises 13.22 and 13.23. The subjects were 94 adult staff members at a West Coast university. They watched a half-hour local news show from a different state that included the ads. The selected ads were judged to be “good” by some experts and had been shown in regions other than where the study was conducted. The real names of the products were replaced by either familiar or unfamiliar brand names by a professional video editor. The ads were pretested and no one in the pretest sample suggested that the ads were not real. Discuss each of these facts in terms of how you would interpret the results of this study.

13.27 Refer to Exercises 13.22 and 13.23. The ratings for this study were each measured on a seven-point scale, anchored by bad–good, low quality–high quality, unappealing–appealing, and unpleasant–pleasant. The results presented were averaged over three ads for different products: a bank, women’s clothing, and a health care plan. Write a short report summarizing the Normality assumption for two-way ANOVA and the extent to which it is reasonable for the analysis of these data.

13.28 One way to repair serious wounds is to insert some material as a scaffold for the body’s repair cells to use as a template for new tissue. Scaffolds made from extracellular material (ECM) are particularly promising for this purpose. Because they are made from biological material, they serve as an effective scaffold and are then resorbed. Unlike biological material that includes cells, however, they do not trigger tissue rejection reactions in the body. One study compared 6 types of scaffold material (S. Badylak et al., “Marrow-derived cells populate scaffolds composed of xenogeneic extracellular matrix,” *Experimental Hematology,* 29 (2001), pp. 1310–1318). Three of these were ECMs and the other three were made of inert materials. There were three mice used per scaffold type. The response measure was the percent of glucose phosphated isomerase (Gpi) cells in the region of the wound. A large value is good, indicating that there are many bone marrow cells sent by the body to repair the tissue. The experiment included additional groups of rats who received the same types of scaffold but were measured at different times. Here are the data for 4 weeks and 8 weeks after the repair:

	Gpi (%)					
Material	4 weeks			8 weeks		
ECM1	55	70	70	60	65	65
ECM2	60	65	65	60	70	60
ECM3	75	70	75	70	80	70
MAT1	20	25	25	15	25	25
MAT2	5	10	5	10	5	5
MAT3	10	15	10	5	15	10

(a) Make a table giving the sample size, mean, and standard deviation for each of the material-by-time combinations. Is it reasonable to pool the variances? Because the sample sizes in this experiment are very small, we expect a large amount of variability in the sample standard deviations. Although they vary more than we would prefer, we will proceed with the ANOVA.
(b) Make a plot of the means. Describe the main features of the plot.
(c) Run the analysis of variance. Report the F statistics with degrees of freedom and P-values for each of the main effects and the interaction. What do you conclude? Write a short paragraph summarizing the results of your analysis.

13.29 Refer to the previous exercise. Here are the data that were collected at 2 weeks, 4 weeks, and 8 weeks:

	Gpi (%)								
Material	2 weeks			4 weeks			8 weeks		
ECM1	70	75	65	55	70	70	60	65	65
ECM2	60	65	70	60	65	65	60	70	60
ECM3	80	60	75	75	70	75	70	80	70
MAT1	50	45	50	20	25	25	15	25	25
MAT2	5	10	15	5	10	5	10	5	5
MAT3	30	25	25	10	15	10	5	15	10

Rerun the analyses that you performed in the previous exercise. How does the addition of the data for 2 weeks change the conclusions? Write a summary comparing these results with those in the previous exercise.

13.30 Refer to the previous exercise. Analyze the data for each time period separately using a one-way ANOVA. Use a multiple-comparisons procedure where needed. Summarize the results.

13.31 How does the frequency that a supermarket product is promoted at a discount affect the price that customers expect to pay for the product? Does the percent reduction also affect this expectation? These questions were examined by researchers in a study conducted on students enrolled in an introductory management course at a large midwestern university. For 10 weeks 160 subjects received information about the products. The treatment conditions corresponded to the number of promotions (1, 3, 5, or 7) that were described during this 10-week period and the percent that the product was discounted (10%, 20%, 30%, and 40%). Ten students were randomly assigned to each of the $4 \times 4 = 16$ treatments (M. U. Kalwani and C. K. Yim, "Consumer price and promotion expectations: an experimental study," *Journal of Marketing Research,* 29 (1992), pp. 90–100). The following table gives the data.

Expected price data

Number of promotions	Percent discount	Expected price ($)									
1	40	4.10	4.50	4.47	4.42	4.56	4.69	4.42	4.17	4.31	4.59
1	30	3.57	3.77	3.90	4.49	4.00	4.66	4.48	4.64	4.31	4.43
1	20	4.94	4.59	4.58	4.48	4.55	4.53	4.59	4.66	4.73	5.24
1	10	5.19	4.88	4.78	4.89	4.69	4.96	5.00	4.93	5.10	4.78
3	40	4.07	4.13	4.25	4.23	4.57	4.33	4.17	4.47	4.60	4.02
3	30	4.20	3.94	4.20	3.88	4.35	3.99	4.01	4.22	3.70	4.48
3	20	4.88	4.80	4.46	4.73	3.96	4.42	4.30	4.68	4.45	4.56
3	10	4.90	5.15	4.68	4.98	4.66	4.46	4.70	4.37	4.69	4.97
5	40	3.89	4.18	3.82	4.09	3.94	4.41	4.14	4.15	4.06	3.90
5	30	3.90	3.77	3.86	4.10	4.10	3.81	3.97	3.67	4.05	3.67
5	20	4.11	4.35	4.17	4.11	4.02	4.41	4.48	3.76	4.66	4.44
5	10	4.31	4.36	4.75	4.62	3.74	4.34	4.52	4.37	4.40	4.52
7	40	3.56	3.91	4.05	3.91	4.11	3.61	3.72	3.69	3.79	3.45
7	30	3.45	4.06	3.35	3.67	3.74	3.80	3.90	4.08	3.52	4.03
7	20	3.89	4.45	3.80	4.15	4.41	3.75	3.98	4.07	4.21	4.23
7	10	4.04	4.22	4.39	3.89	4.26	4.41	4.39	4.52	3.87	4.70

(a) Summarize the means and standard deviations in a table and plot the means. Summarize the main features of the plot.
(b) Analyze the data with a two-way ANOVA. Report the results of this analysis.
(c) Using your plot and the ANOVA results, prepare a short report explaining how the expected price depends on the number of promotions and the percent of the discount.

13.32 Refer to the previous exercise. Rerun the analysis as a one-way ANOVA with $4 \times 4 = 16$ treatments. Summarize the results of this analysis. Use a multiple-comparisons procedure to describe combinations of number of promotions and percent discounts that are similar or different.