SUPPLEMENTARY EXERCISES

for

INTRODUCTION TO THE PRACTICE OF STATISTICS

Fourth Edition

David S. Moore

and

George P. McCabe

These exercises appeared in the third edition of *Introduction to the Practice of Statistics*. They were replaced in the fourth edition in the interest of freshness, but they remain high-quality exercises that supplement those in the text.

Chapter 1

Section 1

S1.1 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?

S1.2 Here is part of a data set that describes motor vehicles:

Vehicle	Type	Where made	City MPG	Highway MPG
Acura 2.5TL	Compact	Foreign	20	25
Buick Skylark	Compact	Domestic	22	32
Audi A8 Quattro	Midsize	Foreign	17	25
Chrysler Concorde	Large	Domestic	19	27

Identify the individuals. Then list the variables recorded for each individual and classify each variable as categorical or quantitative.

S1.3 You want to compare the "size" of several statistics textbooks. Describe at least three possible numerical variables that describe the "size" of a book. In what *unit* would you measure each variable? What *measuring instrument* does each require? Describe a variable that is appropriate for estimating how long it would take you to read the book. Describe a variable that helps decide whether the book will fit easily into your book bag.

S1.4 You are studying the relationship between political attitudes and length of hair among male students. You will measure political attitudes with a standard questionnaire. How will you measure length of hair? Give precise instructions that an assistant could follow. Include a statement of the *unit* and the *measuring instrument* that your assistant is to use.

S1.5 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.

S1.6 Bicycle riding has become more popular. Is it also getting safer? During 1988, 24,800,000 people rode a bicycle at least six times and 949 people were killed in bicycle accidents. In 1992, there were 54,632,000 riders and 903 bicycle fatalities.
(a) Compare the death rates for the two years. What do you conclude?
(b) Although the data come from the same government publication, we suspect that some change in measurement took place between 1988 and 1992 that makes the data for the two years not directly comparable. Why should we be suspicious?

S1.7 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?

S1.8 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.

S1.9 In 1993 there were 90,523 deaths from accidents in the United States. Among these were 41,893 deaths from motor vehicle accidents, 13,141 from falls, 3807 from drowning, 3900 from fires, and 7382 from poisoning. (Data from the 1996 *Statistical Abstract of the United States*.)
(a) Find the percent of accidental deaths from each of these causes, rounded to the nearest percent. What percent of accidental deaths were due to other causes?
(b) Make a well-labeled bar graph of the distribution of causes of accidental deaths. Be sure to include an "other causes" bar.
(c) Would it also be correct to use a pie chart to display these data? Explain your answer.

S1.10 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:

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? (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.)

S1.11 Table S1 contains the percent of residents 65 years of age and over in each of the 50 states.
(a) Make a stemplot of these data using percent as stems and tenths of a percent as leaves. Then make a second stemplot splitting each stem in two (for leaves 0 to 4 and leaves 5 to 9). Which display do you prefer?
(b) Describe the shape, center, and spread of the distribution. Is it roughly symmetric or distinctly skewed? Which states are clear outliers? Use what you know about the states to explain these outliers.

Table S1 Percent of residents aged 65 and over in the states (1998)

State	Percent	State	Percent	State	Percent
Alabama	13.1	Louisiana	11.5	Ohio	13.4
Alaska	5.5	Maine	14.1	Oklahoma	13.4
Arizona	13.2	Maryland	11.5	Oregon	13.2
Arkansas	14.3	Massachusetts	14.0	Pennsylvania	15.9
California	11.1	Michigan	12.5	Rhode Island	15.6
Colorado	10.1	Minnesota	12.3	South Carolina	12.2
Connecticut	14.3	Mississippi	12.2	South Dakota	14.3
Delaware	13.0	Missouri	13.7	Tennessee	12.5
Florida	18.3	Montana	13.3	Texas	10.1
Georgia	9.9	Nebraska	13.8	Utah	8.8
Hawaii	13.3	Nevada	11.5	Vermont	12.3
Idaho	11.3	New Hampshire	12.0	Virginia	11.3
Illinois	12.4	New Jersey	13.6	Washington	11.5
Indiana	12.5	New Mexico	11.4	West Virginia	15.2
Iowa	15.1	New York	13.3	Wisconsin	13.2
Kansas	13.5	North Carolina	12.5	Wyoming	11.5
Kentucky	12.5	North Dakota	14.4		

S1.12 A manufacturing company is reviewing the salaries of its full-time employees below the executive level at a large plant. The clerical staff is almost entirely female, while a majority of the production workers and technical staff are male. As a result, the distributions of salaries for male and female employees may be quite different. The table below gives the frequencies and relative frequencies for women and men. Make histograms from these data, choosing the type that is most appropriate for comparing the two distributions. Then describe the overall shape of the two salary distributions and the chief differences between them.

Salary	Women		Men	
($1000)	Number	%	Number	%
10–15	89	11.8	26	1.1
15–20	192	25.4	221	9.0
20–25	236	31.2	677	27.9
25–30	111	14.7	823	33.6
30–35	86	11.4	365	14.9
35–40	25	3.3	182	7.4
40–45	11	1.5	91	3.7
45–50	3	0.4	33	1.4
50–55	2	0.3	19	0.8
55–60	0	0.0	11	0.4
60–65	0	0.0	0	0.0
65–70	1	0.1	3	0.1
Total	756	100.1	2451	100.0

S1.13 The years around 1970 brought unrest to many U.S. cities. Here are data on the number of civil disturbances in each 3-month period during the years 1968 to 1972:

Period		Count	Period		Count
1968	Jan.–Mar.	6	1970	July–Sept.	20
	Apr.–June	46		Oct.–Dec.	6
	July–Sept	25	1971	Jan.–Mar.	12
	Oct.–Dec.	3		Apr.–June	21
1969	Jan.–Mar.	5		July–Sept.	5
	Apr.–June	27		Oct.–Dec.	1
	July–Sept.	19	1972	Jan.–Mar.	3
	Oct.–Dec.	6		Apr.–June	8
1970	Jan.–Mar.	26		July–Sept.	5
	Apr–Jun	24		Oct–Dec	5

(a) Make a time plot of these counts. Connect the points in your plot by straight line segments to make the pattern clearer.
(b) Describe the trend and seasonal variation in this time series. Can you suggest an explanation for the seasonal variation in civil disorders?

S1.14 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. Here are the number of home runs Ruth hit in each year:

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

Earlier we examined the distribution of Ruth's home run totals during his Yankee years. Now make a time plot and describe its main features.

S1.15 There are seasonal patterns in the growth of children. We can examine the pattern for some children using data from the Egyptian village of Kalama. The following tables give the average weights and heights of Egyptian toddlers who reach the age of 24 months in each month of the year.[1] (Data courtesy of Linda D. McCabe and the Agency for International Development.)

Month	Weight (kg)	Month	Weight (kg)	Month	Weight (kg)
Jan.	11.26	May	10.87	Sept.	10.87
Feb.	11.39	June	10.80	Oct.	10.74
Mar.	11.37	July	10.77	Nov.	10.89
Apr.	11.10	Aug.	10.79	Dec.	11.25

Month	Height (cm)	Month	Height (cm)	Month	Height (cm)
Jan.	79.80	May	80.10	Sept.	79.94
Feb.	79.66	June	80.23	Oct.	79.80
Mar.	79.91	July	80.07	Nov.	79.39
Apr.	80.00	Aug.	79.79	Dec.	79.51

(a) Make a time plot of the weight data. Describe any seasonal variation that appears in the plot. In particular, which months correspond to the highest weights? Which correspond to the lowest?
(b) Repeat the analysis of (a) for the height data.
(c) Describe the similarities and differences between the weight plot and the height plot.

S1.16 In villages such as Kalama, the prevalence of illness may slow the growth of children. In particular, diarrhea is a serious problem and a major cause of death for small children. The following table gives the average percent of days ill with diarrhea per month for the children whose weight and height information appears in the previous exercise:

Month	Percent	Month	Percent	Month	Percent
Jan.	2.54	May	8.42	Sept.	1.20
Feb.	1.82	June	6.87	Oct.	2.51
Mar.	2.28	July	4.51	Nov.	2.34
Apr.	6.86	Aug.	3.42	Dec.	1.27

(a) Make a time plot. Describe any seasonal variation that appears in the plot. Which months have the highest average percent of days ill with diarrhea? Which months have the lowest?
(b) Egyptian physicians say there are two peaks in the frequency of diarrhea, a primary peak and a secondary, lower peak. Is there evidence in these data to support this claim?
(c) Compare your plot with those of the previous exercise. Can you formulate any theories concerning the relationship between diarrhea and the growth of these children? Note that illness in one month may not have an effect on growth until subsequent months. On the other hand, a child who is not growing well may be more susceptible to illness.

Section 2

S1.17 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.

S1.18 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?

S1.19 Exercise S1.10 presented data on the growth of chicks fed normal corn (the control group) and a new variety with better protein quality (the experimental group). Here are the weight gains in grams:

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) 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?

S1.20 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.

S1.21 The business magazine *Forbes* estimates (November 6, 1995) that the "average" household wealth of its readers is either about \$800,000 or about \$2.2 million, depending on which "average" it reports. Which of these numbers is the mean wealth and which is the median wealth? Explain your answer.

S1.22 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.

S1.23 The following are the golf scores of 12 members of a women's golf team in tournament play:

89 90 87 95 86 81 102 105 83 88 91 79

(a) Present the distribution by a stemplot and describe its main features.
(b) Compute the mean, variance, and standard deviation of these golf scores.
(c) Then compute the median, the quartiles, and the IQR. Are there any suspected outliers by the $1.5 \times IQR$ criterion?
(d) Based on the shape of the distribution, would you report the standard deviation or the quartiles as a measure of spread?

S1.24 Table S1 gives the percent of each state's residents who are at least 65 years old. Give a brief graphical and numerical description of the distribution. Explain your choice of numerical measures.

S1.25 In 1995, the SATs were "recentered" to locate the mean scores for individuals close to the center of the 200 to 800 range of possible scores. To see the effect of recentering, we enter the state results for the mathematics and verbal SATs for both 1990 and 1996 into a statistical software package. Here are the results. (This output is from the Minitab statistical software; other software produces similar results.)

```
SATV-1990
     N     MEAN   MEDIAN   STDEV       MIN       MAX        Q1        Q3
    51   448.16   443.00   30.82    397.00    511.00    422.00    476.00

SATM-1990
     N     MEAN   MEDIAN   STDEV       MIN       MAX        Q1        Q3
    51   497.39   490.00   34.57    437.00    577.00    470.00    523.00

SATV-1996
     N     MEAN   MEDIAN   STDEV       MIN       MAX        Q1        Q3
    51   531.90   525.00   33.76    480.00    596.00    501.00    565.00

SATM-1996
     N     MEAN   MEDIAN   STDEV       MIN       MAX        Q1        Q3
    51   529.30   521.00   34.83    473.00    600.00    500.00    557.00
```

Use the output to make side-by-side boxplots of the four distributions. What was the effect of recentering on the distribution of verbal SAT scores in the states? How did the distributions of state verbal and mathematics scores compare before recentering? How do they compare after recentering

S1.26 Give an approximate five-number summary of the salary data in Exercise S1.12 by pretending that all salaries fall at the midpoint of their class. You see that you can find approximate numerical descriptions from grouped data when the actual data are not available.

Section 3

S1.27 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.

S1.28 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.)

S1.29 Give an interval that contains the middle 95% of NOX levels in the exhaust of trucks using the model described in Exercise S1.27.

S1.30 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 S1.28.

S1.31 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 relative frequency of applicants whose score X satisfies each of the following conditions:
(a) $X > 700$
(b) $X < 500$
(c) $500 < X < 800$

S1.32 A patient is said to be hypokalemic (low potassium in the blood) if the measured level of potassium is 3.5 or less. (The units for this measure are meq/l, or milliequivalents per liter.) An individual's potassium level is not a constant, however, but varies from day to day. In addition, the measurement procedure itself has some variation. Suppose that the overall variation follows a normal distribution. Judy has a mean potassium level of 3.8 with a standard deviation of 0.2. If she is measured on many days, on what proportion of days will the measurement suggest that Judy is hypokalemic?

S1.33 The Acculturation Rating Scale for Mexican Americans (ARSMA) is a psychological test that evaluates the degree to which Mexican Americans are adapted to Mexican/Spanish versus Anglo/English culture. The range of possible scores is 1.0 to 5.0, with higher scores showing more Anglo/English acculturation. The distribution of ARSMA scores in a population used to develop the test is approximately normal with mean 3.0 and standard deviation 0.8. A researcher believes that Mexicans will have an average score near 1.7, and that first-generation Mexican Americans will average about 2.1 on the ARSMA scale. What proportion of the population used to develop the test has scores below 1.7? Between 1.7 and 2.1?

S1.34 How high a score on the ARSMA test of the previous exercise must a Mexican American obtain to be among the 30% of the population used to develop the test who are most Anglo/English in cultural orientation? What scores make up the 30% who are most Mexican/Spanish in their acculturation?

S1.35 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 Table 1.4, IPS page 31, and report your conclusions.

S1.36 Exercise S1.10 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?

S1.37 Explain why the normal quantile plot of the DRP reading scores (Figure 1.38, IPS page 90), suggests that the mean and standard deviation are satisfactory measures of center and spread for these data. Then calculate the mean and standard deviation from the scores given in IPS Exercise 1.26, page 30.

Chapter 2

Section 1

S2.1 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) Make a scatterplot of the data, taking the first-round score as the explanatory variable.
(b) Is there an association between the two scores? If so, is it positive or negative? Explain why you would expect scores in two rounds of a tournament to have an association like that you observed.
(c) The plot shows one outlier. Circle it. The outlier may occur because a good golfer had an unusually bad round or because a weaker golfer had an unusually good round. Can you tell from the data given whether the outlier is from a good player or from a poor player? Explain your answer.

S2.2 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. (Data from 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?

S2.3 In 1974, the Franklin National Bank failed. Franklin was one of the 20 largest banks in the nation, and the largest ever to fail. Could Franklin's weakened condition have been detected in advance by simple data analysis? The table below gives the total assets (in billions of dollars) and net income (in millions of dollars) for the 20 largest banks in 1973, the year before Franklin failed. Franklin is bank number 19. (D. E. Booth, *Regression Methods and Problem Banks*, COMAP, Lexington, Mass., 1986.)

Bank	1	2	3	4	5	6	7	8	9	10
Assets	49.0	42.3	36.3	16.4	14.9	14.2	13.5	13.4	13.2	11.8
Income	218.8	265.6	170.9	85.9	88.1	63.6	96.9	60.9	144.2	53.6
Bank	11	12	13	14	15	16	17	18	19	20
Assets	11.6	9.5	9.4	7.5	7.2	6.7	6.0	4.6	3.8	3.4
Income	42.9	32.4	68.3	48.6	32.2	42.7	28.9	40.7	13.8	22.2

(a) We expect banks with more assets to earn higher income. Make a scatterplot of these data that displays the relation between assets and income. Mark Franklin (Bank 19) with a separate symbol.
(b) Describe the overall pattern of your plot. Are there any banks with unusually high or low income relative to their assets? Does Franklin stand out from other banks in your plot?

S2.4 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?

S2.5 When animals of the same species live together, they often establish a clear pecking order. Lower-ranking individuals defer to higher-ranking animals, usually avoiding open conflict. A researcher on animal behavior wants to study the relationship between pecking order and physical characteristics such as weight. He confines four chickens in each of seven pens and observes the pecking order that emerges in each pen. Here is a table of the weights (in grams) of the chickens, arranged by pecking order. That is, the first row gives the weights of the dominant chickens in the seven pens, the second row the weights of the number 2 chicken in each pen, and so on. (Data collected by D. L. Cunningham, Cornell University.)

	Weight (g)						
Pecking order	Pen 1	Pen 2	Pen 3	Pen 4	Pen 5	Pen 6	Pen 7
1	1880	1300	1600	1380	1800	1000	1680
2	1920	1700	1830	1520	1780	1740	1460
3	1600	1500	1520	1520	1360	1520	1760
4	1830	1880	1820	1380	2000	2000	1800

(a) Make a plot of these data that is appropriate to study the effect of weight on pecking order. Include in your plot any means that might be helpful.
(b) We might expect that heavier chickens would tend to stand higher in the pecking order. Do these data give clear evidence for or against this expectation?

Section 2

S2.6 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

If you did not make a scatterplot in Exercise 2.6, do so now. Find the correlation between the Round 1 and Round 2 scores. Remove Player 7's scores and find the correlation for the remaining 10 players. Explain carefully why removing this single case substantially increases the correlation.

S2.7 The British government conducts regular surveys of household spending. Table S2 shows the average weekly household spending on tobacco products and alcoholic beverages for each of the 11 regions of Britain. (Data from British official statistics, *Family Expenditure Survey*, Department of Employment, 1981.)
(a) Make a scatterplot of spending on tobacco y against spending on alcohol x.
(b) Describe the pattern of the plot and any important deviations.
(c) Find the correlation. Then compute the correlation for the 10 regions omitting Northern Ireland. Explain why this r differs so greatly from the r for all 11 cases.

S2.8 IPS Figure 2.1 (page 107) is a scatterplot of each state's mean SAT mathematics score against the percent of the state's high school graduates who take the exam. The correlation between these variables is $r = -0.864$. Explain why r is negative and quite strong. What important features of the relationship does r fail to describe?

S2.9 Financial experts use statistical measures to describe the performance of investments such as mutual funds. In the past, fund managers feared that investors would not understand statistical descriptions, but mounting pressure to give better information is moving standard deviations and correlations into the public eye.
(a) The T. Rowe Price mutual fund group reports the standard deviation of yearly percent returns for its funds. Recently, Equity Income Fund had standard deviation 9.94%, and Science &

Table S2 Spending by region on alcohol and tobacco

Region	Alcohol	Tobacco
North	£6.47	£4.03
Yorkshire	6.13	3.76
Northeast	6.19	3.77
East Midlands	4.89	3.34
West Midlands	5.63	3.47
East Anglia	4.52	2.92
Southeast	5.89	3.20
Southwest	4.79	2.71
Wales	5.27	3.53
Scotland	6.08	4.51
Northern Ireland	4.02	4.56

Technology Fund had standard deviation 23.77%. Explain to someone who knows no statistics how these standard deviations help investors compare the two funds.
(b) Some mutual funds act much like the stock market as a whole, as measured by a market index such as the Standard & Poor's 500 stock index (S&P 500). Others are very different from the overall market. We can use correlation to describe the association. Monthly returns from Fidelity Magellan Fund, the largest mutual fund, have correlation $r = 0.85$ with the S&P 500. Fidelity Small Cap Stock Fund has correlation $r = 0.55$ with the S&P 500. Explain to someone who knows no statistics how these correlations help investors compare the two funds.

Section 3

S2.10 Concrete road pavement gains strength over time as it cures. Highway builders use regression lines to predict the strength after 28 days (when curing is complete) from measurements made after 7 days. Let x be strength after 7 days (in pounds per square inch) and y the strength after 28 days. One set of data gives this least-squares regression line:

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

(a) Draw a graph of this line, with x running from 3000 to 4000 pounds per square inch.
(b) Explain what the slope $b = 0.96$ in this equation says about how concrete gains strength as it cures.
(c) A test of some new pavement after 7 days shows that its strength is 3300 pounds per square inch. Use the equation of the regression line to predict the strength of this pavement after 28 days. Also draw the "up and over" lines from $x = 3300$ on your graph, as in IPS Figure 2.12.

S2.11 Manatees are large, gentle sea creatures that live along the Florida coast. Many manatees are killed or injured by powerboats. Table S3 gives data on powerboat registrations (in thousands) and the number of manatees killed by boats in Florida in the years 1977 to 1990.
(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

Table S3 Florida powerboats and manatee deaths, 1977–1990

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

by the number of boats registered? Does it appear that the number of manatees killed can be predicted accurately from power boat registrations?
(c) Find the least-squares regression line. Predict the number of manatees that will be killed by boats in a year when 716,000 powerboats are registered.
(d) Suppose that in some far future year 2 million powerboats are registered in Florida. Use the regression line to predict manatees killed. Explain why this prediction is very unreliable.
(e) Here are four more years of manatee data, in the same form as in Table S3:

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 power boat registrations. In fact, powerboat registrations remained at 716,000 for the next three years. Compare the mean manatee deaths in these years with your prediction from part (c). How accurate was your prediction?

S2.12 Joan is concerned about the amount of energy she uses to heat her home in the Midwest. She keeps a record of the natural gas she consumes each month over one year's heating season. Because the months are not all the same length, she divides each month's consumption by the number of days in the month to get the average number of cubic feet of gas used per day. Demand for heating is strongly influenced by the outside temperature. From local weather records, Joan obtains the average number of heating degree-days per day for each month. (One heating degree-day is accumulated for each degree a day's average temperature falls below 65° F. An average temperature of 20° F, for example, corresponds to 45 degree-days.) Here are Joan's data (provided by Robert Dale, Purdue University):

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 consumed	520	610	870	850	880	490	450	250	110

(a) Make a scatterplot of these data. There is a strongly linear pattern with no outliers.
(b) Find the equation of the least-squares regression line for predicting gas use from degree-days. Draw this line on your graph. Explain in simple language what the slope of the regression line

tells us about how gas use responds to outdoor temperature.
(c) Joan adds insulation in her attic during the summer, hoping to reduce her gas consumption. The next February, there are an average of 40 degree-days per day and her gas consumption is 870 cubic feet per day. Predict from the regression equation how much gas the house would have used at 40 degree-days per day last winter before the extra insulation. Did the insulation reduce gas consumption?

S2.13 Find the mean and standard deviation of the degree-day and gas consumption data in the previous exercise. Find the correlation between the two variables. Use these five numbers to find the equation of the regression line for predicting gas use from degree-days. Verify that your work agrees with your results in the previous exercise. Use the same five numbers to find the equation of the regression line for predicting degree-days from gas use. What units does each of these slopes have?

Section 4

S2.14 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?

S2.15 Table S4 gives data on the amount of beef consumed (pounds per person) and average retail price of beef (dollars per pound) in the United States for the years 1970 to 1993. Because all prices were generally rising during this period, the prices given are "real prices" in 1993 dollars. These are dollars with the buying power that a dollar had in 1993.
(a) Economists expect consumption of an item to fall when its real price rises. Make a scatterplot of beef consumption y against beef price x. Do you see a relationship of the type expected?
(b) Find the equation of the least-squares line and draw the line on your plot. What proportion of the variation in beef consumption is explained by regression on beef price?

Table S4 Price and consumption of beef, 1970–1993

Year	Price per pound (1993 dollars)	Consumption (lbs. per capita)	Year	Price per pound (1993 dollars)	Consumption (lbs. per capita)
1970	3.721	84.62	1982	3.570	77.03
1971	3.789	83.93	1983	3.396	78.64
1972	4.031	85.27	1984	3.274	78.41
1973	4.543	80.51	1985	3.069	79.19
1974	4.212	85.58	1986	2.989	78.83
1975	4.106	88.15	1987	3.032	73.84
1976	3.698	94.36	1988	3.057	72.65
1977	3.477	91.76	1989	3.096	69.34
1978	3.960	87.29	1990	3.107	67.78
1979	4.423	78.10	1991	3.059	66.79
1980	4.098	76.56	1992	2.931	66.48
1981	3.731	77.26	1993	2.934	65.06

(c) Although it appears that price helps explain consumption, the scatterplot seems to show some nonlinear patterns. Find the residuals from your regression in (b) and plot them against time. Connect the successive points by line segments to help see the pattern. Are there systematic effects of time remaining after we regress consumption on price? (A partial explanation is that beef production responds to price changes only after some time lag.)

S2.16 The price of seafood varies with species and time. The following table gives the prices in cents per pound received in 1970 and 1980 by fishermen and vessel owners for several species.

Species	1970 price	1980 price
Cod	13.1	27.3
Flounder	15.3	42.4
Haddock	25.8	38.7
Menhaden	1.8	4.5
Ocean perch	4.9	23.0
Salmon, chinook	55.4	166.3
Salmon, coho	39.3	109.7
Tuna, albacore	26.7	80.1
Clams, soft	47.5	150.7
Clams, blue, hard	6.6	20.3
Lobsters, American	94.7	189.7
Oysters, eastern	61.1	131.3
Sea scallops	135.6	404.2
Shrimp	47.6	149.0

(a) Plot the data with the 1970 price on the x axis and the 1980 price on the y axis.
(b) Describe the overall pattern. Are there any outliers? If so, circle them on your graph. Do these unusual points have large residuals from a fitted line? Are they influential in the sense

that removing them would change the fitted line?
(c) Compute the correlation for the entire set of data. What percent of the variation in 1980 prices is explained by the 1970 prices?
(d) Recompute the correlation discarding the cases that you circled in (b). Do these observations have a strong effect on the correlation? Explain why or why not.
(e) Does the correlation provide a good measure of the relationship between the 1970 and 1980 prices for this set of data? Explain your answer.

Section 5

S2.17 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. Present your suggestion in a diagram like one of those in IPS Figure 2.29, page 180.

S2.18 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? Draw a diagram like one of those in IPS Figure 2.29, page 180, to illustrate the relationships among the variables.

S2.19 A group of college students believes that herb tea has remarkable powers. To test this belief, they make weekly visits to a local nursing home, visiting with the residents and serving them herb tea. The nursing home staff reports that after several months many of the residents are more cheerful and healthy. A skeptical sociologist commends the students for their good deeds but scoffs at the idea that herb tea helped the residents. It's all confounding, says the sociologist. Identify the explanatory and response variables in this informal study. Then explain what other variables are confounded with the explanatory variable.

S2.20 It has been suggested that electromagnetic fields of the kind present near power lines can cause leukemia in children. Experiments with children and power lines are not ethical. Careful studies have found no association between exposure to electromagnetic fields and childhood leukemia. (See Gary Taubes, "Magnetic field-cancer link: will it rest in peace?" *Science*, 277 (1997), p. 29.) Describe the kind of information you would seek to investigate the claim that electromagnetic fields are associated with cancer.

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. Table S5, based on information from the Statistical Abstract of the United States, *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).*

S2.21 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

Table S5 Mean teachers' pay and per-pupil educational spending by state, 1995 (thousands of dollars)

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				

the same states outliers in both distributions?

S2.22 (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?

S2.23 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?

Section 6

S2.24 (Exact exponential growth) A clever courtier, offered a reward by an ancient king of Persia, asked for a grain of rice on the first square of a chess board, 2 grains on the second square, then 4, 8, 16, and so on.
(a) Make a table of the number of grains on each of the first 10 squares of the board.
(b) Plot the number of grains on each square against the number of the square for squares 1 to 10, and connect the points with a smooth curve. This is an exponential curve.
(c) How many grains of rice should the king deliver for the 64th (and final) square?
(d) Take the logarithm of each of your numbers of grains from (a). Plot these logarithms against the number of squares from 1 to 10. You should get a straight line.
(e) From your graph in (d) find the approximate values of the slope b and the intercept a for the line. Use the equation $y = a + bx$ to predict the logarithm of the amount for the 64th square. Check your result by taking the logarithm of the amount you found in (c).

S2.25 (Exact exponential growth) The rate of return on U.S. stocks between 1970 and 1995 was 11.34% per year, compounded annually.[2]
(a) If you invested \$1000 in an index fund that represents the entire stock market in 1970, how much would your investment be worth at the end of 1995?
(b) The result of (a) does not take into account inflation, the declining buying power of the dollar. The dollar amount needed to have the same buying power that \$1000 had in 1970 increased at the rate of 5.62% per year, compounded annually, between 1970 and 1995. How much money in 1995 matched the value of \$1000 in 1970?

S2.26 Biological populations can grow exponentially if not restrained by predators or lack of food. The gypsy moth outbreaks that occasionally devastate the forests of the Northeast illustrate approximate exponential growth. It is easier to count the number of acres defoliated by the moths than to count the moths themselves. Here are data on an outbreak in Massachusetts. (Data provided by Chuck Schwalbe, U.S. Department of Agriculture.)

Year	Acres
1978	63,042
1979	226,260
1980	907,075
1981	2,826,095

(a) Plot the number y of acres defoliated against the year x. The pattern of growth appears exponential.
(b) Verify that y is being multiplied by about 4 each year by calculating the ratio of acres defoliated each year to the previous year. (Start with 1979 to 1978, when the ratio is 226,260/63,042 = 3.6.)
(c) Take the logarithm of each number y and plot the logarithms against the year x. The linear pattern confirms that the growth is exponential.
(d) Find the least-squares line fitted to the four points on your graph from (c). Use this line to predict the number of acres defoliated in 1982. (Predict log y by substituting $x = 1982$ in the equation. Then use the fact that $y = 10^{\log y}$ to predict y.) The actual number for 1982 was 1,383,265, far less than the prediction. The exponential growth of the gypsy moth population was cut off by a viral disease, and the population quickly collapsed back to a low level.

S2.27 We read much about the threat of exploding federal spending on social insurance programs (chiefly social security and Medicare). Here are the amounts spent, in millions of dollars:

Year	1960	1965	1970	1975	1980	1985	1990
Spending	14,307	21,807	45,246	99,715	191,162	313,108	422,257

(a) Examine these data graphically. Does the pattern appear closer to linear growth or to exponential growth?
(b) Plot the logarithms of social insurance spending. Fit the least-squares line and add it to your graph. During which periods was growth faster than the long-term trend? Slower than the trend?
(c) The government has taken some actions to slow growth in spending on Medicare. The 1990 amount lies below the long-term trend line. The latest available figure for federal social insurance

spending is 495,710 for 1992. Use your line from (b) to predict 1992 spending. Was the actual amount spent less than the prediction?

S2.28 The following table shows the growth of the population of Europe (millions of persons) between 400 B.C. and 1950.

Date	Pop.	Date	Pop.	Date	Pop.
400 B.C.	23	1200	61	1600	90
1 A.D.	37	1250	69	1650	103
200	67	1300	73	1700	115
700	27	1350	51	1750	125
1000	42	1400	45	1800	187
1050	46	1450	60	1850	274
1100	48	1500	69	1900	423
1150	50	1550	78	1950	594

(a) Plot population against time.
(b) The graph shows that the population of Europe dropped at the collapse of the Roman Empire (A. D. 200 to 500) and at the time of the Black Death (A. D. 1348). Growth has been uninterrupted since 1400. Plot the logarithm of population against time, beginning in 1400.
(c) Was growth exponential between 1400 and 1950? If not, what overall pattern do you see?

S2.29 The number of motor vehicles (cars, trucks, and buses) registered in the United States has grown as follows (vehicle counts in millions):

Year	1940	1945	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995
Vehicles	32.4	31.0	49.2	62.7	73.9	90.4	108.8	132.9	155.8	171.7	188.8	203.1

(a) Plot the number of vehicles against time. Also plot the logarithm of the number of vehicles against time.
(b) Look at the years 1950 to 1980. Was the growth in motor vehicle registrations more nearly linear or more nearly exponential during this period? (Use a ruler to compare the two graphs.)
(c) Examine the logarithm graph. The year 1940 fits the pattern well, but 1945 and the years after 1980 do not. All years after 1980 have negative residuals. Does the pattern after 1980 suggest exponential growth at a lower rate? Why?
(d) The year 1945 is a single unusual point. Can you suggest an explanation for the negative residual in 1945?

S2.30 The least-squares line fitted to the 1950 to 1980 data from the previous exercise is approximately

$$\log y = -30.62 + (0.01657 \times \text{year})$$

Use this line to predict motor vehicle registrations in 1945 and 1995. As your work in the previous exercise shows, these predictions are too high.

S2.31 The productivity of American agriculture has grown rapidly due to improved technology (crop varieties, fertilizers, mechanization). Here are data on the output per hour of labor on

American farms. The variable is an "index number" that gives productivity as a percent of the 1967 level.

Year	1940	1945	1950	1955	1960	1965	1970	1975	1980	1985	1990
Productivity	21	27	35	47	67	91	113	137	166	217	230

Plot these data and also plot the logarithms of the productivity values against year. Then briefly describe the pattern of growth that your plots suggest.

Chapter 3

Section 1

S3.1 A letter to the editor of *Organic Gardening* magazine (August 1980) said, "Today I noticed about eight stinkbugs on the sunflower stalks. Immediately I checked my okra, for I was sure that they'd be under attack. There wasn't one stinkbug on them. I'd never read that stinkbugs are attracted to sunflowers, but I'll surely interplant them with my okra from now on."

Explain briefly why this anecdote does not provide good evidence that sunflowers attract stinkbugs away from okra. In your explanation, suggest some factors that might account for all of the bugs being on the sunflowers.

S3.2 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?

S3.3 What is the preferred treatment for breast cancer that is detected in its early stages? The most common treatment was once mastectomy (removal of the breast). It is now usual to remove the tumor and nearby lymph nodes, followed by radiation. To study whether these treatments differ in their effectiveness, a medical team examines the records of 25 large hospitals and compares the survival times after surgery of all women who have had either treatment.
(a) What are the explanatory and response variables?
(b) Explain carefully why this study is not an experiment.
(c) Do you think this study will show whether a mastectomy causes longer average survival time? Explain your answer carefully.

S3.4 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?

S3.5 A study of the effect of abortions on the health of subsequent children was conducted as follows. The researchers obtained names of women who had had abortions from medical records in New York City hospitals. They then searched birth records to locate all women in this group who bore a child within five years of the abortion. Then hospital records were examined again for information about the health of the newborn child. Was this an observational study or an experiment? Why?

S3.6 In a study of the relationship between physical fitness and personality, middle-aged college faculty who have volunteered for an exercise program are divided into low-fitness and high-fitness groups on the basis of a physical examination. All subjects then take the Cattell Sixteen Personality Factor Questionnaire, and the results for the two groups are compared. Is this study an observational study or an experiment? Explain your answer.

Section 2

For each of the experimental situations described in Exercises S3.7 to S3.9, identify the experimental units or subjects, the factors, the treatments, and the response variables.

S3.7 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.

S3.8 What are the effects of repeated exposure to an advertising message? The answer may depend both on the length of the ad and on how often it is repeated. An experiment investigated this question using undergraduate students as subjects. All subjects viewed a 40-minute television program that included ads for a digital camera. Some subjects saw a 30-second commercial; others, a 90-second version. The same commercial was repeated either 1, 3, or 5 times during the program. After viewing, all of the subjects answered questions about their recall of the ad, their attitude toward the camera, and their intention to purchase it.

S3.9 New varieties of corn with altered amino acid patterns may have higher nutritive value than standard corn, which is low in the amino acid lysine. An experiment compares two new varieties, called opaque-2 and floury-2, with normal corn. Corn-soybean meal diets using each type of corn are prepared at three different protein levels: 12%, 16%, and 20%. There are thus nine diets in all. Researchers assign 10 one-day-old male chicks to each diet and record their weight gains after 21 days. The weight gain of the chicks is a measure of the nutritive value of their diet.

S3.10 Exercise S3.7 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.)

S3.11 Will providing child care for employees make a company more attractive to women, even those who are unmarried? You are designing an experiment to answer this question. You prepare recruiting material for two fictitious companies, both in similar businesses in the same location. Company A's brochure does not mention child care. There are two versions of Company B's material, identical except that one describes the company's on-site child-care facility. Your subjects are 40 unmarried women who are college seniors seeking employment. Each subject will read recruiting material for both companies and choose the one she would prefer to work for. You will give each version of Company B's brochure to half the women. You suspect that a higher percentage of those who read the description that includes child care will choose Company B.
(a) Outline the design of the experiment. Be sure to identify the response variable.
(b) The names of the subjects appear below. Do the randomization required by your design and list the subjects who will read the version that mentions child care. (If you use Table B, begin at line 121.)

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

S3.12 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.

S3.13 The following situations were not experiments. Can an experiment be done to answer the questions raised? If so, briefly outline its design. If not, explain why an experiment is not feasible.
(a) The "gender gap" issue of Exercise S3.2.
(b) The comparison of two surgical procedures for breast cancer in Exercise S3.3.

S3.14 You have 100 students from a management course available to serve as subjects in the experiment of Exercise S3.8 on effectiveness of television advertising. Outline a completely randomized design for this experiment in detail. (You need not do any randomization.)

S3.15 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.

S3.16 Some investment advisors believe that charts of past trends in the prices of securities can help predict future prices. Most economists disagree. In an experiment to examine the effects of using charts, business students trade (hypothetically) a foreign currency at computer screens. There are 20 student subjects available, named for convenience A, B, C, . . . , T. Their goal is to make as much money as possible, and the best performances are rewarded with small prizes. The student traders have the price history of the foreign currency in dollars in their computers. They may or may not also be given software that charts past trends. Describe *two* designs for this experiment, a completely randomized design and a matched pairs design in which each student serves as his or her own control. In both cases, carry out the randomization required by your design.

S3.17 Men and women often react differently to advertising. You can use this fact to improve on the completely randomized design for the advertising experiment of Exercises S3.8 and S3.14. Describe the design you recommend. (You need not do any randomization.)

S3.18 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?

S3.19 Is the number of days a letter takes to reach another city affected by the time of day it is mailed and whether or not the zip code is used? Describe briefly the design of a two-factor experiment to investigate this question. Be sure to specify the treatments exactly and to tell how you will handle lurking variables such as the day of the week on which the letter is mailed.

S3.20 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.

S3.21 A medical study of heart surgery investigates the effect of drugs called beta-blockers on

the pulse rate of the patient during surgery. The pulse rate will be measured at a specific point during the operation. The investigators decide to use as subjects 30 patients facing heart surgery who have consented to take part in the study. You have a list of these patients, numbered 1 to 30 in alphabetical order.
(a) Outline in graphical form a completely randomized experimental design for this study.
(b) Carry out the randomization required by your design and report the result. (If you use Table B, enter at line 125.)

S3.22 You are participating in the design of a medical experiment to investigate whether a calcium supplement in the diet will reduce the blood pressure of middle-aged men. Preliminary work suggests that calcium may be effective and that the effect may be greater for black men than for white men.
(a) Outline in graphic form the design of an appropriate experiment.
(b) Choosing the sizes of the treatment groups requires more statistical expertise. We will learn more about this aspect of design in later chapters. Explain in plain language the advantage of using larger groups of subjects.

S3.23 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.

S3.24 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.

Section 3

S3.25 A sociologist wants to know the opinions of employed adult women about government funding for day care. She obtains a list of the 520 members of a local business and professional women's club and mails a questionnaire to 100 of these women selected at random. Only 48 questionnaires are returned. What is the population in this study? What is the sample from whom information is actually obtained? What is the rate (percentage) of nonresponse?

S3.26 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) Each week, the Gallup Poll questions a sample of about 1500 adult U.S. residents in order to discover the opinions of Americans on a wide variety of issues.
(b) Every 10 years, the Census Bureau tries to gather basic information from every household in the United States. But a "long form" requesting much additional information is sent to a sample of about 17% of U.S. households.
(c) A machinery manufacturer purchases voltage regulators from a supplier. There are reports that variation in the output voltage of the regulators is affecting the performance of the finished products. To assess the quality of the supplier's production, the manufacturer chooses a sample of 5 regulators from the last shipment for careful laboratory analysis.

S3.27 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.)

Agarwal	Dewald	Hixson	Puri
Anderson	Fernandez	Klassen	Rodriguez
Baxter	Frank	Liang	Rubin
Bowman	Fuhrmann	Moser	Santiago
Bruvold	Goel	Naber	Shen
Casella	Gupta	Petrucelli	Shyr
Choi	Hicks	Pliego	Sundheim

S3.28 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.

S3.29 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.

S3.30 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.)

S3.31 A university's financial aid office wants to know how much it can expect students to earn from summer employment. This information will be used to set the level of financial aid. The population contains 3478 students who have completed at least one year of study but have not yet graduated. A questionnaire will be sent to an SRS of 100 of these students, drawn from an alphabetized list.
(a) Describe how you will label the students in order to select the sample.
(b) Use Table B, beginning at line 105, to select the *first five* students in the sample.

S3.32 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.

S3.33 You are on the staff of a member of Congress who is considering a controversial bill that would provide for government-sponsored insurance to cover care in nursing homes. You report that 1128 letters dealing with the issue have been received, of which 871 oppose the legislation. "I'm surprised that most of my constituents oppose the bill. I thought it would be quite popular," says the congresswoman. Are you convinced that a majority of the voters oppose the bill? State briefly how you would explain the statistical issue to the congresswoman.

Section 4

S3.34 The Bureau of Labor Statistics announces that last month it interviewed all members of the labor force in a sample of 60,000 households; **6.2%** of the people interviewed were unemployed. Is the bold number a parameter or a statistic?

S3.35 A researcher investigating the effects of a toxic compound in food conducts a randomized comparative experiment with young male white rats. A control group is fed a normal diet, while the experimental group is fed a diet with 2500 parts per million of the toxic material. After 8 weeks, the mean weight gain is **335** grams for the control group and **289** grams for the experimental group. Are the bold numbers parameters or statistics?

S3.36 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?

S3.37 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.

S3.38 An opinion poll asks, "Are you afraid to go outside at night within a mile of your home because of crime?" Suppose that the proportion of all adult U. S. residents who would say "Yes" to this question is $p = 0.4$.
(a) Use Table B to simulate the result of an SRS of 20 adults. Be sure to explain clearly which digits you used to represent each of "Yes" and "No." What proportion of your 20 responses were "Yes"?
(b) Repeat (a) using different lines in Table B until you have simulated the results of 10 SRSs of size 20 from the same population. Compute the proportion of "Yes" responses in each sample. Find the mean of these 10 proportions. Is it close to p?

S3.39 We will illustrate sampling variability and sampling distributions by a small example. Table 1.10 (IPS page 97) gives the positions and weights of the 95 players on a college football team. Call the quarterbacks, running backs, and receivers "offensive backs." There are 22 offensive backs on the team. The 90 players are our population.
(a) The proportion of offensive backs in the population is a parameter p. What is the value of p?
(b) Choose an SRS of 15 players from the team. The proportion who are offensive backs is a statistic $\hat{p}$. What is the value of $\hat{p}$?
(c) Now use software to choose 20 SRSs of size 15. For each sample, find the proportion $\hat{p}$ of offensive backs. Make a histogram of the 20 values of $\hat{p}$. This is a rough approximation to the sampling distribution of this statistic.
(d) Find the mean of the 20 $\hat{p}$ values. Is it close to the parameter p? Mark the value of p on the axis of your histogram. How is the lack of bias of $\hat{p}$ apparent from your graph?

Chapter 4

Section 1

S4.1 When we toss a penny, experience shows that the probability (long-term proportion) of a head is close to 1/2. Suppose now that we toss the penny repeatedly until we get a head. What is the probability that the first head comes up in an odd number of tosses (1, 3, 5, and so on)? To find out, repeat this experiment 50 times, and keep a record of the number of tosses needed to get a head on each of your 50 trials.
(a) From your experiment, estimate the probability of a head on the first toss. What value should we expect this probability to have?
(b) Use your results to estimate the probability that the first head appears on an odd-numbered toss.

S4.2 Obtain 10 identical thumbtacks and toss all 10 at once. Do this 50 times and record the number that land point up on each trial.
(a) What is the approximate probability that at least one lands point up?
(b) What is the approximate probability that more than one lands point up?

Section 2

S4.3 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.

S4.4 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.

S4.5 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.

S4.6 Buy a hot dog and record how many calories it has. Give a reasonable sample space S for your possible results. (Table 1.8 on page xx contains some typical values to guide you.)

S4.7 If you draw an M&M candy at random from a bag of the candies, the candy you draw will have one of six colors. The probability of drawing each color depends on the proportion of each color among all candies made.
(a) The table below gives the probability of each color for a randomly chosen plain M&M:

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?

S4.8 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.

S4.9 Las Vegas Zeke, when asked to predict the Atlantic Coast Conference basketball champion, follows the modern practice of giving probabilistic predictions. He says, "North Carolina's probability of winning is twice Duke's. North Carolina State and Virginia each have probability 0.1 of winning, but Duke's probability is three times that. Nobody else has a chance." Has Zeke given a legitimate assignment of probabilities to the eight teams in the conference? Explain your answer.

S4.10 A sociologist studying social mobility in Denmark finds that the probability that the son of a lower-class father remains in the lower class is 0.46. What is the probability that the son moves to one of the higher classes?

S4.11 Government data assign a single cause for each death that occurs in the United States. The data show that the probability is 0.45 that a randomly chosen death was due to cardiovascular (mainly heart) disease, and 0.22 that it was due to cancer. What is the probability that a death was due either to cardiovascular disease or to cancer? What is the probability that the death was due to some other cause?

S4.12 Select a first-year college student at random and ask what his or her academic rank was in high school. Here are the probabilities, based on proportions from a large sample survey of first-year students:

Rank	Top 20%	Second 20%	Third 20%	Fourth 20%	Lowest 20%
Probability	0.41	0.23	0.29	0.06	0.01

(a) What is the sum of these probabilities? Why do you expect the sum to have this value?
(b) What is the probability that a randomly chosen first-year college student was not in the top 20% of his or her high school class?
(c) What is the probability that a first-year student was in the top 40% in high school?
(d) Now choose two first-year college students at random. What is the probability that both were in the top 20% of their high school classes?

S4.13 Choose an American farm at random and measure its size in acres. Here are the probabilities that the farm chosen falls in several acreage categories:

Acres	< 10	10–49	50–99	100–179	180–499	500–999	1000–1999	≥ 2000
Probability	0.09	0.20	0.15	0.16	0.22	0.09	0.05	0.04

Let A be the event that the farm is less than 50 acres in size, and let B be the event that it is 500 acres or more.
(a) Find $P(A)$ and $P(B)$.
(b) Describe A^c in words and find $P(A^c)$ by the complement rule.
(c) Describe $\{A \text{ or } B\}$ in words and find its probability by the addition rule.

S4.14 Choose an American worker at random and classify his or her occupation into one of the following classes. These classes are used in government employment data.

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

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

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?

S4.15 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?

S4.16 Choose at random a U.S. resident at least 25 years of age. We are interested in the events

$A = \{$The person chosen completed 4 years of college$\}$
$B = \{$The person chosen is between 55 years old or older$\}$

Government data allow us to assign probabilities to these events.
(a) Explain why $P(A) = 0.230$.
(b) Find $P(B)$.
(c) Find the probability that the person chosen is at least 55 years old *and* has 4 years of college education, $P(A \text{ and } B)$. Are the events A and B independent?

S4.17 A six-sided die has four green and two red faces and is balanced so that each face is equally likely to come up. The die will be rolled several times. You must choose one of the following three sequences of colors; you will win $25 if the first rolls of the die give the sequence that you have chosen.

RGRRR
RGRRRG
GRRRRR

Which sequence do you choose? Explain your choice. (In a psychological experiment, 63% of 59 students who had not studied probability chose the second sequence. This is evidence that our intuitive understanding of probability is not very accurate. This and similar experiments are reported by A. Tversky and D. Kahneman, "Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment," *Psychological Review*, 90 (1983), pp. 293–315.)

S4.18 The type of medical care a patient receives may vary with the age of the patient. A large study of women who had a breast lump investigated whether or not each woman received a mammogram and a biopsy when the lump was discovered. Here are some probabilities estimated by the study. The entries in the table are the probabilities that *both* of two events occur; for example, 0.321 is the probability that a patient is under 65 years of age *and* the tests were done. The four probabilities in the table have sum 1 because the table lists all possible outcomes.

	Tests done?	
	Yes	No
Age under 65	0.321	0.124
Age 65 or over	0.365	0.190

(a) What is the probability that a patient in this study is under 65? That a patient is 65 or over?
(b) What is the probability that the tests were done for a patient? That they were not done?
(c) Are the events $A = \{\text{the patient was 65 or older}\}$ and $B = \{\text{the tests were done}\}$ independent? Were the tests omitted on older patients more or less frequently than would be the case if testing were independent of age?

S4.19 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

Section 3

S4.20 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?

S4.21 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?

Section 4

S4.22 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?

S4.23 The distribution of grades (A = 4, B = 3, and so on) in a large course is as follows:

Grade	0	1	2	3	4
Probability	0.10	0.15	0.30	0.30	0.15

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

S4.24 The Tri-State Pick 3 lottery game offers a choice of several bets. You choose a three-digit number. The lottery commission announces the winning three-digit number, chosen at random, at the end of each day. The "box" pays \$83.33 if the number you choose has the same digits as the winning number, in any order. Find the expected payoff for a \$1 bet on the box. (Assume that you chose a number having three distinct digits.)

S4.25 The baseball player Tony Gwynn got a hit about 35% of the time over his entire career. After he has failed to hit safely in six straight at-bats, a TV commentator said, "Tony is due for a hit by the law of averages." Is that right? Why?

S4.26 Find the standard deviation σ_X of the distribution of grades in Exercise E4.2.

S4.27 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.

S4.28 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.

S4.29 The psychologist Amos Tversky did many studies of our perception of chance behavior. In its obituary of Tversky (June 6, 1996), the *New York Times* cited the following example.
(a) Tversky asked subjects to choose between two public health programs that affect 600 people. One has probability 1/2 of saving all 600 and probability 1/2 that all 600 will die. The other is guaranteed to save exactly 400 of the 600 people. Find the mean number of people saved by the first program.
(b) Tversky then offered a different choice. One program has probability 1/2 of saving all 600 and probability 1/2 of losing all 600, while the other will definitely lose exactly 200 lives. What is the difference between this choice and that in (a)?
(c) Given option (a), most subjects choose the second program. Given option (b), most subjects choose the first program. Do the subjects appear to use means in making their decisions? Why do you think their choices differed in the two cases?

S4.30 The original simple form of the Connecticut state lottery (ignoring a few gimmicks) awarded the following prizes for each 100,000 tickets sold. The winners were chosen by drawing tickets at random.

1	\$5000 prize
18	\$200 prizes
120	\$25 prizes
270	\$20 prizes

If you hold one ticket in this lottery, what is your probability of winning anything? What is the mean amount of your winnings?

S4.31 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.

S4.32 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.

S4.33 The real return on an investment is its rate of increase corrected for the effects of inflation. You believe that the annual real return X on a portfolio of stocks will vary in the future with mean $\mu_X = 0.11$ and standard deviation $\sigma_X = 0.28$. (That is, you expect stocks to give an average return of 11% in the future, but with large variation from year to year.) You further think that the annual real return Y on Treasury bills will vary with mean $\mu_Y = 0.02$ and standard deviation $\sigma_Y = 0.05$. Even though this is not realistic, assume that returns on stocks and Treasury bills vary independently.
(a) If you put half of your assets into stocks and half into Treasury bills, your overall return will be $Z = 0.5X + 0.5Y$. Calculate μ_Z and σ_Z.
(b) You decide that you are willing to take more risk (greater variation in the return) in exchange for a higher mean return. Choose an allocation of your assets between stocks and Treasury bills that will accomplish this and calculate the mean and standard deviation of the overall return. (If you put a proportion α of your assets into stocks, the total return is $Z = \alpha X + (1 - \alpha)Y$.)

Section 5

S4.34 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.

S4.35 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.

S4.36 In the language of government statistics, the "labor force" includes all civilians over 16 years of age who are working or looking for work. Select a member of the U.S. labor force at random. Let A be the event that the person selected is white, and B the event that he or she is employed. In 1995, 84.6% of the labor force was white. Of the whites in the labor force, 95.1% were employed. Among nonwhite members of the labor force, 91.9% were employed.
(a) Express each of the percents given as a probability involving the events A and B; for example, $P(A) = 0.846$.
(b) Draw a tree diagram for the outcomes of recording first the race (white or nonwhite) of a randomly chosen member of the labor force and then whether or not the person is employed.
(c) Find the probability that the person chosen is an employed white. Also find the probability that an employed nonwhite is chosen. What is the probability $P(B)$ that the person chosen is employed?

S4.37 Use your results from the previous Exercise and the definition of conditional probability to find the probability $P(A \mid B)$ that a randomly selected member of the labor force is white, given that he or she is employed. (Alternatively, you can use Bayes's rule.)

S4.38 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.)

S4.39 In the setting of the previous Exercise, 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.)

S4.40 Psychologists use probability models to describe learning in animals. In one experiment, a rat is placed in a dark compartment and a door leading to a light compartment is opened. The rat will not move to the light compartment without reason. A bell is rung and if after 5 seconds the rat has not moved, it gets a shock through the floor of the dark compartment. To avoid the shock, rats soon learn to move when the bell is rung. A simple model for learning says that the rat can only be in one of two states:

> **State A:** The rat will not move from the dark compartment until it receives a shock.
>
> **State B:** The rat has learned to respond to the bell and moves immediately when the bell rings.

A rat starts in State A and eventually changes to State B. The change from State A to State B is the result of learning. A rat in State A gets a shock every time the bell rings; after it changes to State B it never gets a shock. Suppose that a rat has probability 0.2 of learning each time it is shocked. Let the random variable X be the number of shocks that this rat receives.[3]
(a) If a rat receives exactly 4 shocks, which state was the rat in at the end of each of Trials 1, 2, 3, and 4?
(b) Use the result of (a) and the multiplication rule to find $P(X = 4)$, the probability that the rat receives exactly 4 shocks.

(c) Based on your work in (a) and (b), give the probability distribution of X. That is, for any positive whole number x, what is $P(X\overline{x})$?

S4.41 Here is the distribution of the adjusted gross income X (in thousands of dollars) reported on individual federal income tax returns in 1993:

Income	< 10	10–24	25–49	50–99	≥ 100
Probability	0.29	0.27	0.25	0.14	0.05

(a) What is the probability that a randomly chosen return shows an adjusted gross income of \$50,000 or more?
(b) Given that a return shows an income of at least \$50,000, what is the conditional probability that the income is at least \$100,000?

S4.42 The distribution of blood types among white Americans is approximately as follows: 37% type A, 13% type B, 44% type O, and 6% type AB. Suppose that the blood types of married couples are independent and that both the husband and wife follow this distribution.
(a) An individual with type B blood can safely receive transfusions only from persons with type B or type O blood. What is the probability that the husband of a woman with type B blood is an acceptable blood donor for her?
(b) What is the probability that in a randomly chosen couple the wife has type B blood and the husband has type A?
(c) What is the probability that one of a randomly chosen couple has type A blood and the other has type B?
(d) What is the probability that at least one of a randomly chosen couple has type O blood?

S4.43 Exercise E4.14 gives the probability distribution of the sex and occupation of a randomly chosen American worker. Use this distribution to answer the following questions:
(a) Given that the worker chosen holds a managerial (Class A) job, what is the conditional probability that the worker is female?
(b) Classes D and E include most mechanical and factory jobs. What is the conditional probability that a worker is female, given that he or she holds a job in one of these classes?

S4.44 Toss a balanced coin 10 times. What is the probability of a run of 3 or more consecutive heads? What is the distribution of the length of the longest run of heads? What is the mean length of the longest run of heads? These are quite difficult questions if we must rely on mathematical calculations of probability. Computer simulation of the probability model can provide approximate answers.

Simulate 50 repetitions of tossing a balanced coin 10 times. (The key word in most software is "Bernoulli.") Now examine your 50 repetitions. Record the length of the longest run of heads in each trial. Combine your results with those of other students so that you have the results of several hundred repetitions.
(a) Make a table of the (approximate) probability distribution of the length X of the longest run of heads in 10 coin tosses. (The proportion of each outcome in many repetitions is approximately equal to its probability.) Draw a probability histogram of this distribution.
(b) What is your estimate of the probability of a run of 3 or more heads?
(c) Find the mean μ_X from your table of probabilities in (a). Then find the average of the 50

values of X in your 50 repetitions. How close to the mean μ_X was the average obtained in 50 repetitions?

Chapter 5

Section 1

S5.1 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?

S5.2 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.

S5.3 The Gallup Poll once found that about 15% of adults jog. Suppose that in fact the proportion of the adult population who jog is $p = 0.15$.
(a) What is the probability that the sample proportion $\hat{p}$ of joggers in an SRS of size $n = 200$ lies between 13% and 17%? (Use software or the normal approximation.)
(b) You just found the probability that an SRS of size 200 gives a sample proportion $\hat{p}$ within ± 2 percentage points of the population proportion $p = 0.15$. Find this probability for SRSs of sizes 800, 1600, and 3200. What general conclusion can you draw from your calculations?

S5.4 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 $\pm 3\%$. What is the probability that an SRS of size 300 has a sample proportion $\hat{p}$ that is within $\pm 3\%$ 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?

S5.5 According to government data, 25% of employed women have never been married.
(a) If 10 employed women are selected at random, what is the probability that exactly 2 have never been married? (Use software, Table C, or the binomial probability formula.)
(b) What is the probability that 2 or fewer have never been married?
(c) What is the probability that at least 8 have been married?

S5.6 According to government data, 21% of American children under the age of six live in households with incomes less than the official poverty level. A study of learning in early childhood chooses an SRS of 300 children.
(a) What is the mean number of children in the sample who come from poverty-level households? What is the standard deviation of this number?
(b) Use the normal approximation to calculate the probability that at least 80 of the children in the sample live in poverty. Be sure to check that you can safely use the approximation.

S5.7 Ray is a basketball player who has made about 70% of his free throws over several years. In a tournament game he makes only 2 of 6 free throws. Ray's coach says this was just bad luck. Suppose that Ray's free throws are independent trials with probability 0.7 of a success on each trial. What is the probability that he makes 2 or fewer in 6 attempts? Do you think that his tournament performance is just chance variation?

Section 2

S5.8 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).

S5.9 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?

S5.10 A bottling company uses a filling machine to fill plastic bottles with a popular cola. The bottles are supposed to contain 300 milliliters (ml). In fact, the contents vary according to a normal distribution with mean $\mu = 298$ ml and standard deviation $\sigma = 3$ ml.
(a) What is the probability that an individual bottle contains less than 295 ml?
(b) What is the probability that the mean contents of the bottles in a six-pack is less than 295 ml?

S5.11 Judy's doctor is concerned that she may suffer from hypokalemia (low potassium in the blood). There is variation both in the actual potassium level and in the blood test that measures the level. Judy's measured potassium level varies according to the normal distribution with $\mu = 3.8$ and $\sigma = 0.2$. A patient is classified as hypokalemic if the potassium level is below 3.5.
(a) If a single potassium measurement is made, what is the probability that Judy is diagnosed as hypokalemic?
(b) If measurements are made instead on 4 separate days and the mean result is compared with the criterion 3.5, what is the probability that Judy is diagnosed as hypokalemic?

S5.12 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?

S5.13 The study habits portion of the Survey of Study Habits and Attitudes (SSHA) psychological test consists of two sets of questions. One set measures "delay avoidance" and the other measures "work methods." A subject's study habits score is the sum $X + Y$ of the delay avoidance score X and the work methods score Y. The distribution of X in a broad population of first-year college students is close to $N(25, 10)$, while the distribution of Y in the same population is close to $N(25, 9)$.
(a) If a subject's X and Y scores were independent, what would be the distribution of the study habits score $X + Y$?
(b) Using the distribution you found in (a), what percent of the population have a study habits score of 60 or higher?
(c) In fact, the X and Y scores are strongly correlated. In this case, does the mean of $X + Y$ still have the value you found in (a)? Does the standard deviation still have the value you found in (a)?

Chapter 6

Section 1

S6.1 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.

S6.2 The 1990 census "long form" asked the total 1989 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 (alas, it is too simple to be true) 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 $\overline{x} = \$23,453$, 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 1989 mean income of Middletown householders who reported income.

S6.3 Refer to the previous problem. Give a 99% confidence interval for the total 1989 income of the households that reported income in Middletown.

S6.4 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%.

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?

S6.5 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.

S6.6 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.

S6.7 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.

Section 2

S6.8 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.

S6.9 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.

S6.10 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.

S6.11 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.

Section 4

S6.12 Example 6.12 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 $\mu = 460$.

S6.13 Example 6.16 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.

S6.14 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 $\mu = 1$.

S6.15 Use the result of Exercise S6.15 to give the probabilities of Type I and Type II errors for the test discussed there. Take the alternative hypothesis to be $\mu = 295$.

S6.16 Use the result of Exercise S6.14 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 $\mu = 460$.

Chapter

S6.17 A study compares two groups of mothers with young children who were on welfare two years ago. One group attended a voluntary training program offered free of charge at a local vocational school and advertised in the local news media. The other group did not choose to attend the training program. The study finds a significant difference ($P < 0.01$) between the proportions of the mothers in the two groups who are still on welfare. The difference is not only significant but quite large. The report says that with 95% confidence the percent of the nonattending group still on welfare is $21\% \pm 4\%$ higher than that of the group who attended the program. You are on the staff of a member of Congress who is interested in the plight of welfare mothers and who asks you about the report.
(a) Explain briefly and in nontechnical language what "a significant difference ($P < 0.01$)" means.
(b) Explain clearly and briefly what "95% confidence" means.
(c) Is this study good evidence that requiring job training of all welfare mothers would greatly reduce the percent who remain on welfare for several years?

S6.18 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.

S6.19 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.

S6.20 Use the same procedure for generating data as in the previous exercise. Now test the null hypothesis that $\mu = 22.5$. Explain your results.

S6.21 Figure 6.2 demonstrates the behavior of a confidence interval in repeated sampling by showing the results of 25 samples from the same population. Now you will do a similar demonstration. Suppose that (unknown to the researcher) the mean SAT-M score of all California high school seniors is $\mu = 460$, and that the standard deviation is known to be $\sigma = 100$. The scores vary normally.
(a) Simulate the drawing of 25 SRSs of size $n = 100$ from this population.
(b) The 95% confidence interval for the population mean 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?

S6.22 In the previous exercise you simulated the SAT-M scores of 25 SRSs of 100 California seniors. Now use these samples to demonstrate the behavior of a significance test. We know that the population of all SAT-M scores is normal with standard deviation $\sigma = 100$.
(a) Use your software to carry out a test of

$$H_0\colon \mu = 460$$

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

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?

S6.23 Suppose that in fact the mean SAT-M score of California high school seniors is $\mu = 480$. Would the test in the previous exercise usually detect a mean this far from the hypothesized value? This is a question about the power of the test.
(a) Simulate the drawing of 25 SRSs from a normal population with mean $\mu = 480$ and $\sigma = 100$. These represent the results of sampling when in fact the alternative $\mu = 480$ is true.
(b) Repeat on these new data the test of

$$H_0\colon \mu = 460$$

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

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?

Chapter 7

Section 1

S7.1 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.

S7.2 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.

S7.3 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.

S7.4 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.

S7.5 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.

S7.6 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.

S7.7 The bank in the previous exercise tested a new idea on a sample of 250 customers. Suppose that the bank wanted to be quite certain of detecting a mean increase of $\mu = \$100$ in the amount charged, at the $\alpha = 0.01$ significance level. Perhaps a sample of only $n = 50$ customers would accomplish this. Find the approximate power of the test with $n = 50$ against the alternative $\mu = \$100$ as follows:
(a) What is the t critical value for the one-sided test with $\alpha = 0.01$ and $n = 50$?
(b) Write the criterion for rejecting H_0: $\mu = 0$ in terms of the t statistic. Then take $s = 108$ (an estimate based on the data in Exercise 7.21) and state the rejection criterion in terms of $\overline{x}$.
(c) Assume that $\mu = 100$ (the given alternative) and that $\sigma = 108$ (an estimate from the data inthe 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?

S7.8 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.)

S7.9 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.

S7.10 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?

S7.11 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?

S7.12 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.

S7.13 The paper reporting the results on ARSMA used in Exercise S7.11 does not give the raw data or any discussion of normality. You would like to replace the t procedure used in Exercise 7.36 by a sign test. Can you do this from the available information? Carry out the sign test and state your conclusion, or explain why you are unable to carry out the test.

S7.14 Exercise S7.12 reports a small study comparing ARSMA and BI, two tests of the acculturation of Mexican Americans. Would this study usually detect a difference in mean scores of 0.2? To answer this question, calculate the approximate power of the test (with $n = 22$ subjects and $\alpha = 0.05$) of

$$H_0\colon \mu = 0$$

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

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 7.36, 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 7.36) by a normal probability calculation. This is the approximate power.

S7.15 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.)

Section 2

S7.16 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.

S7.17 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.

S7.18 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.

S7.19 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.

S7.20 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.

S7.21 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.

S7.22 The data on weights of skilled and novice rowers in Exercise S7.19 can be analyzed by the pooled t procedures, which assume equal population variances. Report the value of the t statistic, its degrees of freedom, and its P-value, and then state your conclusion. (The pooled procedures should not be used for the comparison of knee velocities in Exercise S7.18, because the sample standard deviations in the two groups are different enough to cast doubt on the assumption of a common population standard deviation.)

Section 3

S7.23 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?

S7.24 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.

S7.25 Exercise S7.18 records the results of comparing a measure of rowing style for skilled and novice female competitive rowers. Is there significant evidence of inequality between the standard deviations of the two populations?
(a) State H_0 and H_a.
(b) Calculate the F statistic. Between which two levels does the P-value lie?

S7.26 Answer the same questions for the weights of the two groups, recorded in Exercise S7.19.

S7.27 The observed inequality between the sample standard deviations of male and female SAT mathematics scores in Exercise S7.20 is clearly significant. You can say this without doing any calculations. Find F and look in Table E. Then explain why the significance of F could be seen without arithmetic.

Chapter

S7.28 Data on the numbers of manatees killed by boats each year are given in Exercise S1.11. After a long period of increasing numbers of deaths, the pattern flattens somewhat. In fact, the total for 1990 is 47, less than the total of 50 for 1989. Perhaps the trend has now reversed.

We would like to do a significance test to compare these two counts. Theoretical considerations suggest that the standard errors ($\sigma/\sqrt{n}$) for these types of counts can be approximated by the square root of the count. So, for example, the 1990 count, 47, has a standard error that is approximately $\sqrt{47}$. Use this approximation to perform an approximate two-sample z test for the difference between the 1989 and 1990 deaths. Find an approximate 95% confidence interval for the difference. What do you conclude?

Chapter 8

Section 1

S8.1 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: $p = 0.3$.

S8.2 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.)

S8.3 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?

Section 2

S8.4 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?

S8.5 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?

S8.6 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?

S8.7 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.

S8.8 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?

S8.9 Refer to the study of undergraduate student summer employment described in the previous exercise. Similar results from a smaller number of students may not have the same statistical significance. Specifically, suppose that 72 of 80 men surveyed were employed and 59 of 73 women surveyed were employed. The sample proportions are essentially the same as in the earlier exercise.
(a) Compute the z statistic for these data and report the P-value. What do you conclude?
(b) Compare the results of this significance test with your results in Exercise 8.42. What do you observe about the effect of the sample size on the results of these significance tests?

S8.10 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?

Chapter

S8.11 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.

S8.12 In the Section 6.1 we studied the effect of the sample size on the margin of error of the confidence interval for a single proportion. In this exercise we perform some calculations to observe this effect for the two-sample problem. As in the exercise above, suppose that $\hat{p}_1 = 0.6$,

$\hat{p}_2 = 0.4$, and n represents the common value of n_1 and n_2. Compute the 95% confidence intervals for the difference in the two proportions for $n = 15, 25, 50, 75, 100$, and 500. For each interval calculate the margin of error. Summarize and explain your results.

S8.13 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.

S8.14 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.

Chapter 9

Chapter

S9.1 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.

S9.2 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.

S9.3 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.

S9.4 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.

S9.5 Aluminum is suspected as a factor in the development of Alzheimer's disease. In one study, researchers compared a group of Alzheimer's patients with a carefully selected control group of people who did not have Alzheimer's but were similar in other ways. (Selection of a matching control group is a difficult task. In epidemiological studies such as this, however, experiments are not possible.) The focus of the study was on the use of antacids that contain aluminum. Each subject was classified according to the use of these antacids. The two-way table below gives the data. (Data from Amy Borenstein Graves et al., "The association between aluminum-containing products and Alzheimer's disease," *Journal of Clinical Epidemiology*, 43 (1990), pp. 35–44.)

	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?

S9.6 Are there gender differences in the progress of students in doctoral programs? A major university classified all students entering Ph.D. programs in a given year by their status 6 years later. The categories used were as follows: completed the degree, still enrolled, and dropped out. Here are the data:

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?

S9.7 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?

S9.8 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.

S9.9 Refer to Exercise S9.5 (page xxx), 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.

Chapter 10

Chapter

S10.1 Manatees are large sea creatures that live in the shallow water along the coast of Florida. Many manatees are injured or killed each year by powerboats. Exercise S2.11 gives data on manatees killed and powerboat registrations (in thousands of boats) in Florida for the period 1977 to 1990.
(a) Make a scatterplot of boats registered and manatees killed. Is there a strong straight-line pattern?
(b) Find the equation of the least-squares regression line. Draw this line on your scatterplot.
(c) Is there strong evidence that the mean number of manatees killed increases as the number of powerboats increases? State this question as null and alternative hypotheses about the slope of the population regression line, obtain the t statistic, and give your conclusion.
(d) Predict the number of manatees that will be killed if there are 716,000 powerboats registered. In 1991, 1992, and 1993, the number of powerboats remained at 716,000. The numbers of manatees killed were 53, 38, and 35. Compare your prediction with these data. Does the comparison suggest that measures taken to protect the manatees in these years were effective?

S10.2 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.

S10.3 Exercise 2.56 (page xxx) gives the following data from a study of two methods for measuring the blood flow in the stomachs of dogs:

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.

S10.4 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.

S10.5 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.

S10.6 Return to the data on current versus voltage given in the Ohm's law experiment of Exercise S10.4.
(a) Compute all values for the ANOVA table.
(b) State the null hypothesis tested by the ANOVA F statistic, and explain in plain language what this hypothesis says.

(c) What is the distribution of this F statistic when H_0 is true? Find an approximate P-value for the test of H_0.

S10.7 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).

S10.8 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.

Chapter 11

Chapter

S11.1 One model for subpopulation means for the computer science study is described in Example 11.1 as

$$\mu_{\text{GPA}} = \beta_0 + \beta_1\text{HSM} + \beta_2\text{HSS} + \beta_3\text{HSE}$$

(a) Give the model for the subpopulation mean GPA for students having high school grade scores HSM = 9 (A−), HSS = 8 (B+), and HSE = 7 (B).
(b) Using the parameter estimates given in Figure 11.4, calculate the estimate of this subpopulation mean. Then briefly explain in words what your numerical answer means.

S11.2 Use the model given in the previous exercise to do the following:
(a) For students having high school grade scores HSM = 6 (B−), HSS = 7 (B), and HSE = 8 (B+), express the subpopulation mean in terms of the parameters β_j.
(b) Calculate the estimate of this subpopulation mean using the b_j given in Figure 11.4. Briefly explain the meaning of the number you obtain.

Chapter 12

Chapter

S12.1 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.

S12.2 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.

S12.3 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.

S12.4 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.

S12.5 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.

S12.6 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 '***'.
```

YEAR Comparison	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence 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?

S12.7 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.

S12.8 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.

S12.9 For each part of Exercise S12.1, outline the ANOVA table, giving the sources of variation and the degrees of freedom. (Do not compute the numerical values for the sums of squares and mean squares.)

S12.10 For each part of Exercise S12.2, outline the ANOVA table, giving the sources of variation and the degrees of freedom. (Do not compute the numerical values for the sums of squares and mean squares.)

S12.11 Return to the change-of-majors study described in Example 12.3.
(a) State H_0 and H_a for ANOVA.
(b) Outline the ANOVA table, giving the sources of variation and the degrees of freedom.
(c) What is the distribution of the F statistic under the assumption that H_0 is true?
(d) Using Table E, find the critical value for an $\alpha = 0.05$ test.

S12.12 Return to the survey of college students described in Example 12.4.
(a) State H_0 and H_a for ANOVA.
(b) Outline the ANOVA table, giving the sources of variation and the degrees of freedom.
(c) What is the distribution of the F statistic under the assumption that H_0 is true?
(d) Using Table E, find the critical value for an $\alpha = 0.05$ test.

S12.13 Return to the nematode experiment described in Exercise S12.3. Suppose that when entering the data into the computer, you accidentally entered the first observation as 108 rather than 10.8.
(a) Run the ANOVA with the incorrect observation. Summarize the results.
(b) Compare this run with the results obtained with the correct data set. What does this illustrate about the effect of outliers in an ANOVA?
(c) Compute a table of means and standard deviations for each of the four treatments using the incorrect data. How would this table have helped you to detect the incorrect observation?

S12.14 With small numbers of observations in each group, it can be difficult to detect deviations from normality and violations of the equal standard deviations assumption for ANOVA. Return to the nematode experiment described in Exercise S12.3. The log transformation is often used for variables such as the growth of plants. In many cases this will tend to make the standard deviations more similar across groups and to make the data within each group look more normal. Rerun the ANOVA using the logarithms of the recorded values. Answer the questions given in Exercise S12.3. Compare these results with those obtained by analyzing the raw data.

S12.15 You are planning a study of the SAT mathematics scores of four groups of students. From Example 12.3, we know that the standard deviations of the three groups considered in that study were 86, 67, and 83. In Example 12.5, we found the pooled standard deviation to be 82.5. Since the power of the F test decreases as the standard deviation increases, use $\sigma = 90$ for the calculations in this exercise. This choice will lead to sample sizes that are perhaps a little larger than we need but will prevent us from choosing sample sizes that are too small to detect

the effects of interest. You would like to conclude that the population means are different when $\mu_1 = 620$, $\mu_2 = 600$, $\mu_3 = 580$, and $\mu_4 = 560$.
(a) Pick several values for n (the number of students that you will select from each 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.

S12.16 Refer to the previous exercise. Repeat all parts for the alternative $\mu_1 = 610$, $\mu_2 = 600$, $\mu_3 = 590$, and $\mu_4 = 580$.

Chapter 13

Chapter

S13.1 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.

S13.2 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.

S13.3 For each part of Exercise S13.1, 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.)

S13.4 For each part of Exercise S13.2, 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.)

S13.5 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?

S13.6 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?