Solutions

Chapter 1

Introduction

Answers to Check Your Understanding

page 4: **1.** The cars in the student parking lot. **2.** He measured the car's model (categorical), year (quantitative), color (categorical), number of cylinders (quantitative), gas mileage (quantitative), weight (quantitative), and whether it has a navigation system (categorical).

Answers to Odd-Numbered Introduction Exercises

1.1 Type of wood, type of water repellent, and paint color are categorical. Paint thickness and weathering time are quantitative.

1.3 (a) AP® Statistics students who completed a questionnaire on the first day of class. (b) Categorical: gender, handedness, and favorite type of music. Quantitative: height, homework time, and the total value of coins in a student's pocket. (c) The individual is a female who is right-handed. She is 58 inches tall, spends 60 minutes on homework, prefers Alternative music, and has 76 cents in her pocket.

1.5 Student answers will vary. For example, quantitative variables could be graduation rate and student-faculty ratio, and categorical variables could be region of the country and type of institution (2-year college, 4-year college, university).

1.7 b

Section 1.1

Answers to Check Your Understanding

page 14: **1.** Fly: 99/415 = 23.9%, Freeze time: 96/415 = 23.1%, Invisibility: 67/415 = 16.1%, Superstrength: 43/415 = 10.4%, Telepathy: 110/415 = 26.5%. **2.** A bar graph is shown below. It appears that telepathy, ability to fly, and ability to freeze time were the most popular choices, with about 25% of students choosing each one. Invisibility was the 4th most popular and superstrength was the least popular.


Superpower preference
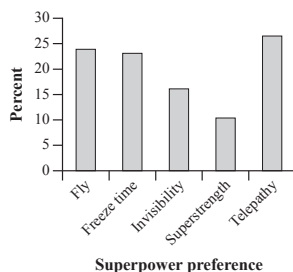
page 18: **1.** For the U.K. students: 54/200 = 27% said fly, 52/200 = 26% said freeze time, 30/200 = 15% said invisibility, 20/200 = 10% said superstrength, and 44/200 = 22% said telepathy. For the U.S. students: 45/215 = 20.9% said fly, 44/215 = 20.5% said freeze time, 37/215 = 17.2% said invisibility, 23/215 = 10.7% said superstrength, and 66/215 = 30.7% said telepathy. **2.** A bar graph is shown in the next column. 3. There is an association between country of origin and superpower preference. Students in the U.K. are more likely to choose flying and freezing time, while students in the U.S. are more likely to choose invisibility or telepathy. Superstrength is about equally unpopular in both countries.


Superpower preference
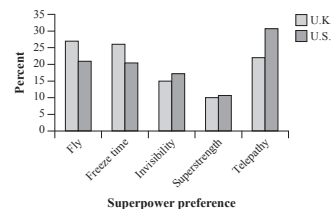
Answers to Odd-Numbered Section 1.1 Exercises

1.9 (a) 1% (b) A bar graph is given below. (c) Yes, because the numbers in the table refer to parts of a single whole.


Color
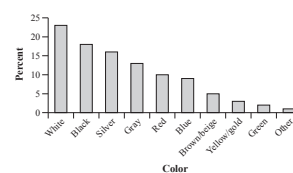
1.11 (a) A bar graph is given below. A pie chart would also be appropriate because the numbers in the table refer to parts of a single whole. (b) Perhaps induced or C-section births are scheduled for weekdays so doctors don't have to work as much on the weekend.

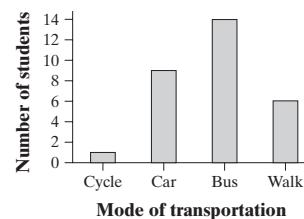
Day
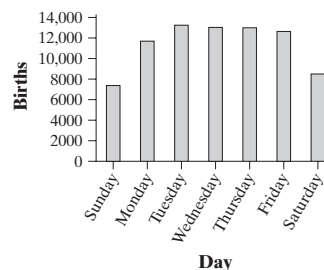
1.13 About 63% are Mexican and 9% are Puerto Rican.

1.15 (a) The given percents represent fractions of different age groups, rather than parts of a single whole. (b) A bar graph is given below.

Age group
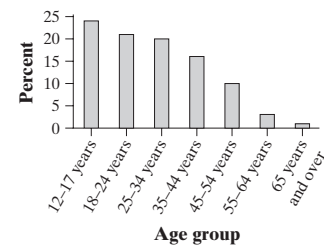
1.17 (a) The areas of the pictures should be proportional to the numbers of students they represent. (b) A bar graph is given below.

**Mode of transportation**

1.19 (a) 133 people; 36 buyers of coffee filters made of recycled paper. **(b)** 36.8% said "higher," 24.1% said "the same," and 39.1% said "lower." Overall, 60.9% of the members of the sample think the quality is the same or higher.

1.21 For buyers, 55.6% said higher, 19.4% said the same, and 25% said lower. For the nonbuyers, 29.9% said higher, 25.8% said the same, and 44.3% said lower. We see that buyers are much more likely to consider recycled filters higher in quality and much less likely to consider them lower in quality than nonbuyers.

1.23 Americans are much more likely to choose white/pearl and red, while Europeans are much more likely to choose silver, black, or gray. Preferences for blue, beige/brown, green, and yellow/gold are about the same for both groups.

1.25 A table and a side-by-side bar graph comparing the distributions of snowmobile use for environmental club members and nonmembers are shown below. There appears to be an association between environmental club membership and snowmobile use. The visitors who are members of an environmental club are much more likely to have never used a snowmobile and less likely to have rented or owned a snowmobile than visitors who are not in an environmental club.

	Not a member	**Member**
Never used	445/1221 = 36.4%	212/305 = 69.5%
Snowmobile renter	497/1221 = 40.7%	77/305 = 25.2%
Snowmobile owner	279/1221 = 22.9%	16/305 = 5.2%
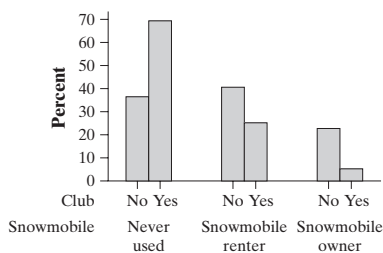
1.27 d
1.29 d
1.31 b
1.33 d
1.35 Answers will vary. Two possible tables are given below.

| 10	40
50	0

30	20
30	20

Section 1.2
Answers to Check Your Understanding
1. This distribution is skewed to the right and unimodal. **2.** The midpoint of the 28 values is between 1 and 2. **3.** The number of siblings varies from 0 to 6. **4.** There are two potential outliers at 5 and 6 siblings.

1. Both males and females have distributions that are skewed to the right, though the distribution for the males is more heavily skewed. The midpoint for the males (9 pairs) is less than the midpoint for the females (26 pairs). The number of shoes owned by females varies more (from 13 to 57) than for males (from 4 to 38). The male distribution has three likely outliers at 22, 35, and 38. The females do not have any likely outliers. **2.** b **3.** e **4.** c

1. One possible histogram is shown below. **2.** The distribution is roughly symmetric and bell-shaped. The typical IQ appears to be between 110 and 120 and the IQs vary from 80 to 150. There do not appear to be any outliers.
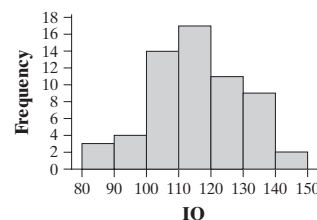
1. This is a bar graph because field of study is a categorical variable. **2.** No, because the variable is categorical and the categories could be listed in any order on the horizontal axis.

Answers to Odd-Numbered Section 1.2 Exercises
1.37 (a) The graph is shown below. **(b)** The distribution is roughly symmetric with a midpoint of 6 hours. The hours of sleep vary from 3 to 11. There do not appear to be any outliers.
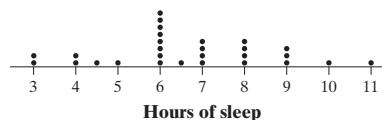
1.39 (a) This dot represents a game where the opposing team won by 1 goal. **(b)** All but 4 of the 25 values are positive, which indicates that the U.S. women's soccer team had a very good season. They won 21/25 = 84% of their games.

1.41 As coins get older, they are taken out of circulation and new coins are introduced, meaning that most coins will be from recent years with a few from previous years.

1.43 Both distributions are roughly symmetric and have about the same amount of variability. The center of the internal distribution is greater than the center of the external distribution, indicating that external rewards do not promote creativity. Neither distribution appears to have outliers.

1.45 (a) Otherwise, most of the data would appear on just a few stems, making it hard to identify the shape of the distribution. **(b)** Key: 12 | 1 means that 12.1% of that state's residents are aged 25 to 34. **(c)** The distribution of percent of residents aged 25−34 is roughly symmetric with a possible outlier at 16.0%. The center is around 13%. Other than the outlier at 16.0%, the values vary from 11.4% to 15.1%.

1.47 (a) The stemplots are given in the next column. The stemplot with split stems makes it easier to see the shape of the distribution. **(b)** The distribution is slightly skewed to the right with a center near 780 mm, and values that vary from around 600 mm to 960 mm. There do not appear to be any outliers. **(c)** In El Niño years, there is typically less rain than in other years (18 of 23 years).

Without splitting stems

```
6 | 0 3 5 5 7
7 | 0 1 2 4 4 8 8 9 9 9
8 | 1 1 3 6 6 7
9 | 0 6
```
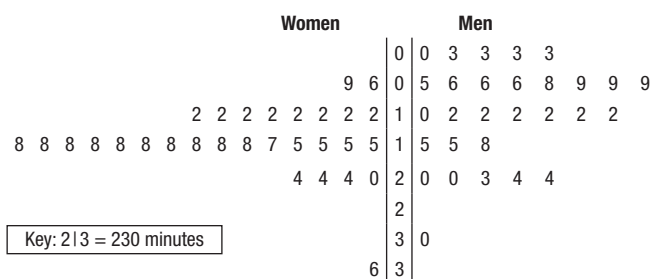
Key: 6 | 3 = 630 mm of rain

With splitting stems

```
6 | 0 3
6 | 5 5 7
7 | 0 1 2 4 4
7 | 8 8 9 9 9
8 | 1 1 3
8 | 6 6 7
9 | 0
9 | 6
```

1.49 (a) Most people will round their answers to the nearest 10 minutes (or 30 or 60). The students who claimed 300 and 360 minutes of studying on a typical weeknight may have been exaggerating. (b) The stemplots suggest that women (claim to) study more than men. The center for women (about 175 minutes) is greater than the center for men (about 120 minutes).

```
      Women              Men
                    0 | 0  3 3 3 3
              9 6 | 0 | 5  6 6 6 8 9 9 9
  2 2 2 2 2 2 2 2 | 1 | 0  2 2 2 2 2 2
8 8 8 8 8 8 8 8 8 8 7 5 5 5 5 | 1 | 5  5 8
            4 4 4 0 | 2 | 0  0 3 4 4
                    2 |
                    3 | 0
                6 | 3 |
```

Key: 2 | 3 = 230 minutes

1.51 (a) The distribution is slightly skewed to the left and unimodal. (b) The center is between 0% and 2.5%. (c) The highest return was between 10% and 12.5%. Ignoring the low outliers, the lowest return was between −12.5% and −10%. (d) About 37% of these months (102 out of 273) had negative returns.

1.53 (a) The histogram is given below. (b) The distribution of travel times is roughly symmetric. The center is near 23 minutes and the values vary from 15.5 to 30.9 minutes. There do not appear to be any outliers.

Travel time (minutes)
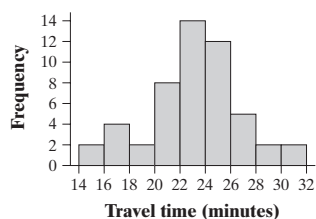
1.55 The histogram is given below. The distribution of DRP scores is roughly symmetric with the center around 35. The DRP scores vary from 14 to 54. There do not appear to be any outliers.


DRP scores
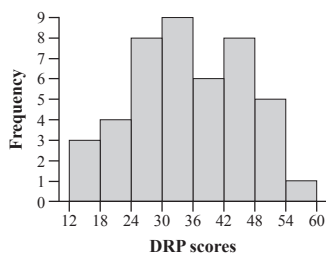
1.57 (a) The histogram is given below. The distribution of word lengths is skewed to the right and single-peaked. The center is around 4 letters, with words that vary from 1 to 15 letters. There do not appear to be any outliers. (b) There are more short words in Shakespeare's plays and more very long words in *Popular Science* articles.


Length of words (number of letters)
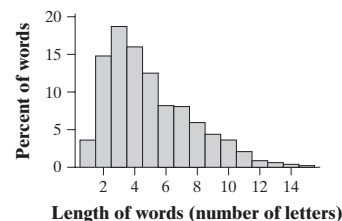
1.59 The scale on the horizontal axis is very different from one graph to the other.

1.61 A bar graph should be used because birth month is a categorical variable. A possible bar graph is given below.

Month
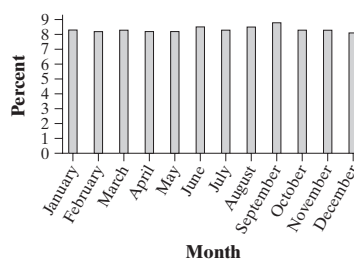
1.63 (a) The percents for women sum to 100.1% due to rounding errors. (b) Relative frequency histograms are shown below because there are considerably more men than women. (c) Both histograms are skewed to the right. The center of the women's distribution of salaries is less than the men's. The distributions of salaries are about equally variable, and the table shows that there are some outliers in each distribution who make between $65,000 and $70,000.

Salary ($ thousands)

Salary ($ thousands)
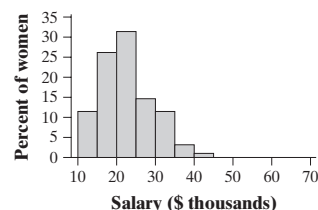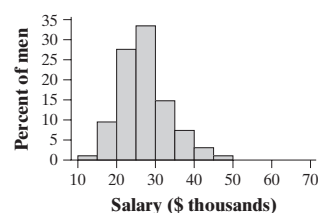
1.65 The distribution of age is skewed to the right for both males and females, meaning that younger people outnumber older people. Among the younger Vietnamese, there are more males than females. After age 35, however, females seem to outnumber the males, making the center of the female distribution a little greater than the male distribution. Both distributions have about the same amount of variability and no outliers.

1.67 **(a)** Amount of studying. We would expect some students to study very little, but most students to study a moderate amount. Any outliers would likely be high outliers, leading to a right-skewed distribution. **(b)** Right- versus left-handed. About 90% of the population is right-handed (represented by the bar at 0). **(c)** Gender. We would expect a more similar percentage of males and females than for the right-handed and left-handed students. **(d)** Heights. We expect many heights near the average and a few very short or very tall people.

1.69 a

1.71 c

1.73 d

1.75 **(a)** Major League Baseball players who were on the roster on opening day of the 2012 season. **(b)** 6. Two variables are categorical (team, position) and the other 4 are quantitative (age, height, weight, and salary).

1.77 **(a)** 71/858 = 8.3% were elite soccer players and 43/858 = 5.0% of the people had arthritis. **(b)** 10/71 = 14.1% of elite soccer players had arthritis and 10/43 = 23.3% of those with arthritis were elite soccer players.

Section 1.3

Answers to Check Your Understanding

page 53: **1.** Because the distribution is skewed to the right, we would expect the mean to be larger than the median. **2.** Yes. The mean is 31.25 minutes, which is greater than the median of 22.5 minutes. **3.** Because the distribution is skewed, the median would be a better measure of the center of the distribution.

page 59: **1.** The data in order are: 290, 301, 305, 307, 307, 310, 324, 345. The 5-number summary is 290, 303, 307, 317, 345. **2.** The IQR is 14 pounds. The range of the middle half of the data is 14 pounds. **3.** Any outliers occur below $303 - 1.5(14) = 282$ or above $317 + 1.5(14) = 338$, so 345 pounds is an outlier. **4.** The boxplot is given below.
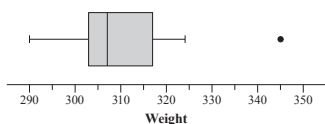
page 63: **1.** The mean is 75. **2.** The table is given below.

Observation	Deviation	Squared deviation
67	$67 - 75 = -8$	$(-8)^2 = 64$
72	$72 - 75 = -3$	$(-3)^2 = 9$
76	$76 - 75 = 1$	$1^2 = 1$
76	$76 - 75 = 1$	$1^2 = 1$
84	$84 - 75 = 9$	$9^2 = 81$
Total	**0**	**156**

3. The variance is $s_x^2 = \dfrac{156}{5-1} = 39$ inches squared and the standard deviation is $s_x = \sqrt{39} = 6.24$ inches.

4. The players' heights typically vary by about 6.24 inches from the mean height of 75 inches.

Answers to Odd-Numbered Section 1.3 Exercises

1.79 $\bar{x} = 85$

1.81 **(a)** median $= 85$ **(b)** $\bar{x} = 79.33$ and median $= 84$. The median did not change much but the mean did, showing that the median is more resistant to outliers than the mean.

1.83 The mean is $60,954 and the median is $48,097. The distribution of salaries is likely to be quite right skewed because of a few people who have a very large income, making the mean larger than the median.

1.85 The team's annual payroll is $1.2(25) = 30$ or $30 million. No, because the median only describes the middle value in the distribution. It doesn't provide specific information about any of the other values.

1.87 **(a)** Estimating the frequencies of the bars (from left to right) as 10, 40, 42, 58, 105, 60, 58, 38, 27, 18, 20, 10, 5, 5, 1, and 3, the mean is $\bar{x} = \dfrac{3504}{500} = 7.01$. The median is the average of the 250th and 251st values, which is 6. **(b)** Because the median is less than the mean, we would use the median to argue that shorter domain names are more popular.

1.89 **(a)** $IQR = 91 - 78 = 13$. The middle 50% of the data have a range of 13 points. **(b)** Any outliers are below $78 - 1.5(13) = 58.5$ or above $91 + 1.5(13) = 110.5$. There are no outliers.

1.91 **(a)** Outliers are anything below $3 - 1.5(40) = -57$ or above $43 + 1.5(40) = 103$, so 118 is an outlier. The boxplot is shown below. **(b)** The article claims that teens send 1742 texts a month, which is about 58 texts a day. Nearly all of the members of the class (21 of 25) sent fewer than 58 texts per day, which seems to contradict the claim in the article.
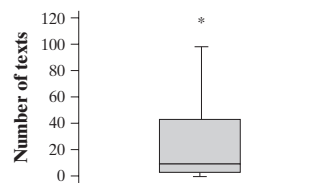
1.93 **(a)** Positive numbers indicate students who had more text messages than calls. Because the 1st quartile is about 0, roughly 75% of the students had more texts than calls, which supports the article's conclusion. **(b)** No. Students in statistics classes tend to be upperclassmen and their responses might differ from those of underclassmen.

1.95 **(a)** About 3% and -3.5%. **(b)** About 0.1%. **(c)** The stock fund is much more variable. It has higher positive returns, but also higher negative returns.

1.97 **(a)** $s_x = \sqrt{\dfrac{2.06}{6-1}} = 0.6419$ mg/dl. **(b)** The phosphate level typically varies from the mean by about 0.6419 mg/dl.

1.99 **(a)** Skewed to the right, because the mean is much larger than the median and Q_3 is much further from the median than Q_1. **(b)** The amount of money spent typically varies from the mean by about $21.70. **(c)** Any points below $19.06 - 1.5(26.66) = -20.93$ or above $45.72 + 1.5(26.66) = 85.71$ are outliers. Because the maximum of 93.34 is greater than 85.71, there is at least one outlier.

1.101 Yes. For example, in data set 1, 2, 3, 4, 5, 6, 7, 8 the IQR is 4. If 8 is changed to 88, the IQR will still be 4.

1.103 **(a)** One possible answer is 1, 1, 1, 1. **(b)** 0, 0, 10, 10. **(c)** For part (a), any set of four identical numbers will have $s_x = 0$. For part (b), however, there is only one possible answer. We want the values to be as far from the mean as possible, so our best choice is two values at each extreme.

1.105 *State:* Do the data indicate that men and women differ in their study habits and attitudes toward learning? *Plan:* We will draw side-by-side boxplots of the data about men and women; compute summary statistics; and compare the shape, center, and spread of both distributions. *Do:* The boxplots are given below, as is a table of summary statistics.
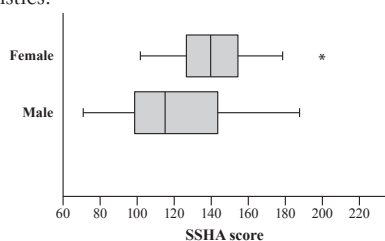


```
Variable   N    Mean    StDev   Minimum    Q₁     Median    Q₃    Maximum
Women     18  141.06   26.44    101.00  126.00  138.50  154.00  200.00
Men       20  121.25   32.85     70.00   98.00  114.50  143.00  187.00
```

Both distributions are slightly skewed to the right. Both the mean and median are higher for women than for men. The scores for men are more variable than the scores for women. There are no outliers in the male distribution and a single outlier at 200 in the female distribution. *Conclude*: Men and women differ in their study habits and attitudes toward learning. The typical score for females is about 24 greater than the typical score for males. Female scores are also more consistent than male scores.

1.107 d

1.109 e

1.111 A histogram is given below. This distribution is roughly symmetric with a center around 170 cm and values that vary from 145.5 cm to 191 cm. There do not appear to be any outliers.
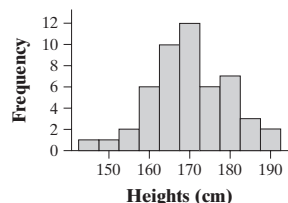

1.113 Women appear to be more likely to engage in behaviors that are indicative of good "habits of mind." They are especially more likely to revise papers to improve their writing. The difference is a little smaller for seeking feedback on their work, although the percentage is still higher for females.

Answers to Chapter 1 Review Exercises

R1.1 (a) Movies. (b) *Quantitative:* Year, time, box office sales. Categorical: Rating, genre. *Note*: Year might be considered categorical if we want to know how many of these movies were made each year rather than the average year. (c) This movie is *Avatar,* released in 2009. It was rated PG-13, runs 162 minutes, is an action film, and had box office sales of $2,781,505,847.

R1.2 A bar chart is given below.
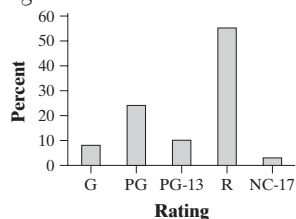
R1.3 (a) The "bars" are different widths. For example, the bar for "send/receive text messages" should be roughly twice the size of the bar for "camera" when it is actually about 4 times as large. (b) No, because they do not describe parts of a whole. Students were free to answer in more than one category. (c) A bar graph is given below.
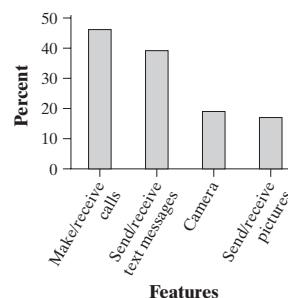
R1.4 (a) 148/219 = 67.6%. Marginal distribution, because it is part of the distribution of one variable for all categories of the other variable. (b) 78/82 = 95.1% of the younger students were Facebook users. 78/148 = 52.7% of the Facebook users were younger.

R1.5 There does appear to be an association between age and Facebook status. From both the table and the graph given below, we can see that as age increases, the percent of Facebook users decreases. For younger students, about 95% are members. That drops to 70% for middle students and drops even further to 31.3% for older students.

Age	Facebook user?	
	Yes	No
Younger (18–22)	95.1%	4.9%
Middle (23–27)	70.0%	30.0%
Older (28 and up)	31.3%	68.7%
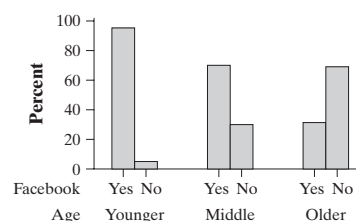
R1.6 (a) A stemplot is given below. (b) The distribution is roughly symmetric with one possible outlier at 4.88. The center of the distribution is between 5.4 and 5.5. The densities vary from 4.88 to 5.85. (c) Because the distribution is roughly symmetric, we can use the mean to estimate the Earth's density to be about 5.45 times the density of water.

```
48 | 8
49 |
50 | 7
51 | 0
52 | 6799
53 | 04469        Key: 48 | 8 = 4.88
54 | 2467
55 | 03578
56 | 12358
57 | 59
58 | 5
```

R1.7 (**a**) A histogram is given below. The survival times are right-skewed, as expected. The median survival time is 102.5 days and the range of survival times is $598 - 43 = 555$ days. There are several high outliers with survival times above 500.
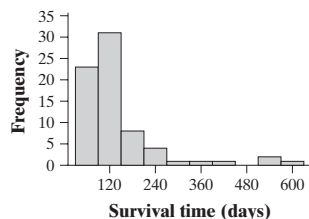
(**b**) The boxplot is given below.
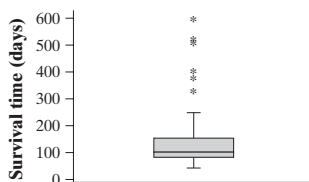
(**c**) Use the median and *IQR* to summarize the distribution because the outliers will have a big effect on the mean and standard deviation.

R1.8 (**a**) About 20% of low-income and 33% of high-income households. (**b**) The shapes of both distributions are skewed to the right; however, the skewness is much stronger in the distribution for low-income households. On average, household size is larger for high-income households. One-person households might have less income because they would include many young single people who have no job or retired single people with a fixed income.

R1.9 (**a**) The amount of mercury per can of tuna will typically vary from the mean by about 0.3 ppm. (**b**) Any point below $0.071 - 1.5(0.309) = -0.393$ or above $0.38 + 1.5(0.309) = 0.8435$ would be considered an outlier. There are no low outliers, but there are several high outliers. (**c**) The distribution of the amount of mercury in cans of tuna is highly skewed to the right. The median is 0.18 ppm and the *IQR* is 0.309 ppm.

R1.10 The distribution for light tuna is skewed to the right with several high outliers, while the distribution for albacore tuna is more symmetric with just a couple of high outliers. Because it has a greater center, the albacore tuna generally has more mercury. However, the light tuna has a much bigger spread of values, with some cans having as much as twice the amount of mercury as the largest amount in the albacore tuna.

Answers to Chapter 1 AP® Statistics Practice Test

T1.1 d
T1.2 e
T1.3 b
T1.4 b
T1.5 c
T1.6 c
T1.7 b
T1.8 c
T1.9 e
T1.10 b
T1.11 d
T1.12 (**a**) A histogram is given below. (**b**) Any point below $30 - 1.5(47) = -40.5$ or above $77 + 1.5(47) = 147.5$ is an outlier. So 151 minutes is an outlier. (**c**) Median and *IQR*, because the distribution is skewed and has a high outlier.
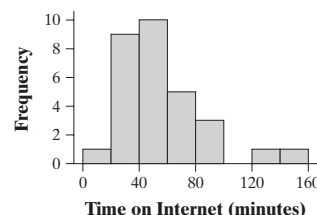
T1.13 (**a**) Row totals are 1154, 53, and 1207. Column totals are 785, 375, 47, and 1207. (**b**) Nondiabetic: 96.1% none and 3.9% one or more. Prediabetic: 96.5% none and 3.5% one or more. Diabetic: 80.9% none and 19.1% one or more. (**c**) The graph is given below.
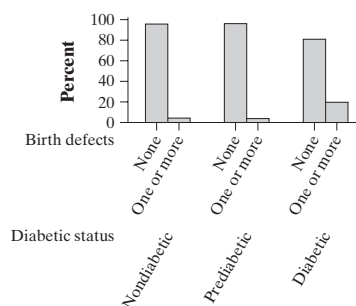
(**d**) Yes. Nondiabetics and prediabetics appear to have babies with birth defects at about the same rate. However, those with diabetes have a much higher rate of babies with birth defects.

T1.14 (**a**) Between 550 and 559 hours. (**b**) Because it has a higher minimum lifetime or because its lifetimes are more consistent (less variable). (**c**) Because it has a higher median lifetime.

T1.15 Side-by-side boxplots and descriptive statistics for both leagues are given below. Both distributions are roughly symmetric, although there are two low outliers in the NL. The data suggest that the number of home runs is somewhat less in the NL. All 5 numbers in the 5-number summary are less for the NL teams than for the AL teams. However, there is more variability among the AL teams.
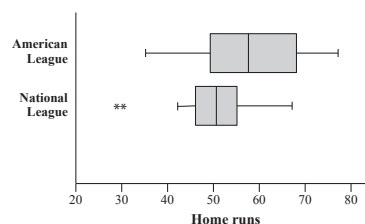
Variable	N	Mean	StDev	Minimum	Q_1	Median	Q_3	Maximum
American League	14	56.93	12.69	35.00	49.00	57.50	68.00	77.00
National League	14	50.14	11.13	29.00	46.00	50.50	55.00	67.00

Chapter 2

Section 2.1

Answers to Check Your Understanding

page 89: **1.** c **2.** Her daughter weighs more than 87% of girls her age and she is taller than 67% of girls her age. **3.** About 65% of calls lasted less than 30 minutes, which means that about 35% of calls lasted 30 minutes or longer. **4.** $Q_1 = 13$ minutes, $Q_3 = 32$ minutes, and $IQR = 19$ minutes.

page 91: **1.** $z = -0.466$. Lynette's height is 0.466 standard deviations below the mean height of the class. **2.** $z = 1.63$. Brent's height is 1.63 standard deviations above the mean height of the class. **3.** $-0.85 = \dfrac{74 - 76}{\sigma}$, so $\sigma = 2.35$ inches.

page 97: **1.** Shape will not change. However, it will multiply the center (mean, median) and spread (range, IQR, standard deviation) by 2.54. **2.** Shape and spread will not change. It will, however, add 6 inches to the center (mean, median). **3.** Shape will not change. However, it will change the mean to 0 and the standard deviation to 1.

### Answers to Odd-Numbered Section 2.1 Exercises

2.1 **(a)** She is at the 25th percentile, meaning that 25% of the girls had fewer pairs of shoes than she did. **(b)** He is at the 85th percentile, meaning that 85% of the boys had fewer pairs of shoes than he did. **(c)** The boy is more unusual because only 15% of the boys have as many or more than he has. The girl has a value that is closer to the center of the distribution.

2.3 A percentile only describes the relative location of a value in a distribution. Scoring at the 60th percentile means that Josh's score is better than 60% of the students taking this test. His correct percentage could be greater than 60% or less than 60%, depending on the difficulty of the test.

2.5 The girl weighs more than 48% of girls her age, but is taller than 78% of the girls her age.

2.7 **(a)** The student sent about 205 text messages in the 2-day period and sent more texts than about 78% of the students in the sample. **(b)** Locate 50% on the *y*-axis, read over to the points, and then go down to the *x*-axis. The median is approximately 115 text messages.

2.9 **(a)** $IQR \approx \$46 - \$19 = \$27$ **(b)** About the 26th percentile. **(c)** The histogram is below.
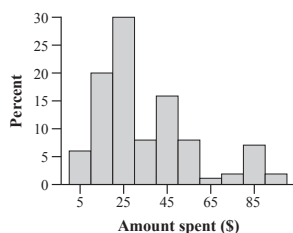
2.11 Eleanor. Her standardized score ($z = 1.8$) is higher than Gerald's ($z = 1.5$).

2.13 **(a)** Your bone density is far below average—about 1.5 times farther below average than a typical below-average density. **(b)** Solving $-1.45 = \dfrac{948 - 956}{\sigma}$ gives $\sigma = 5.52$ g/cm^2.

2.15 **(a)** He is at the 76th percentile, meaning his salary is higher than 76% of his teammates. **(b)** $z = 0.79$. Lidge's salary was 0.79 standard deviations above the mean salary.

2.17 Multiply each score by 4 and add 27.

2.19 **(a)** mean = 87.188 inches and median = 87.5 inches. **(b)** The standard deviation (3.20 inches) and IQR (3.25 inches) do not change because adding a constant to each value in a distribution does not change the spread.

2.21 **(a)** mean = 5.77 feet and median = 5.79 feet. **(b)** Standard deviation = 0.267 feet and IQR = 0.271 feet.

2.23 Mean = $\dfrac{9}{5}(25) + 32 = 77°F$ and standard deviation = $\dfrac{9}{5}(2) = 3.6°F$.

2.25 c

2.27 c

2.29 c

2.31 The distribution is skewed to the right with a center around 20 minutes and the range close to 90 minutes. The two largest values appear to be outliers.

Section 2.2

Answers to Check Your Understanding

page 107: **1.** It is legitimate because it is positive everywhere and it has total area under the curve = 1. **2.** 12% **3.** Point A in the graph below is the approximate median. About half of the area is to the left of A and half of the area is to the right of A. **4.** Point B in the graph below is the approximate mean (balance point). The mean is less than the median in this case because the distribution is skewed to the left.
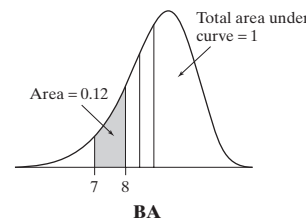
page 112: **1.** The graph is given below. **2.** Approximately $\dfrac{100\% - 68\%}{2} = 16\%$. **3.** Approximately $\dfrac{100\% - 68\%}{2} = 16\%$ have heights below 62 inches and approximately $\dfrac{100\% - 99.7\%}{2} = 0.15\%$ of young women have heights above 72 inches, so the remaining 83.85% have heights between 62 and 72 inches.
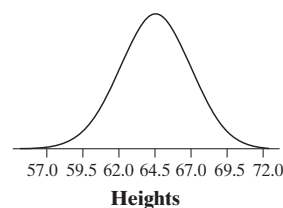
page 116: (All graphs are shown on the following page.) **1.** The proportion is 0.9177. **2.** The proportion is 0.9842. **3.** The proportion is $0.9649 - 0.2877 = 0.6772$. **4.** The *z*-score for the 20th percentile is $z = -0.84$. **5.** 45% of the observations are greater than $z = 0.13$.

Answers to Odd-Numbered Section 2.2 Exercises

2.33 Sketches will vary, but here is one example:

2.35 (a) It is on or above the horizontal axis everywhere, and the area beneath the curve is $\frac{1}{3} \times 3 = 1$. (b) $\frac{1}{3} \times 1 = \frac{1}{3}$. (c) Because $1.1 - 0.8 = 0.3$, the proportion is $\frac{1}{3} \times 0.3 = 0.1$.

2.37 Both are 1.5.

2.39 (a) Mean is C, median is B. (b) Mean is B, median is B.

2.41 The graph is shown below.

Men's height (inches)
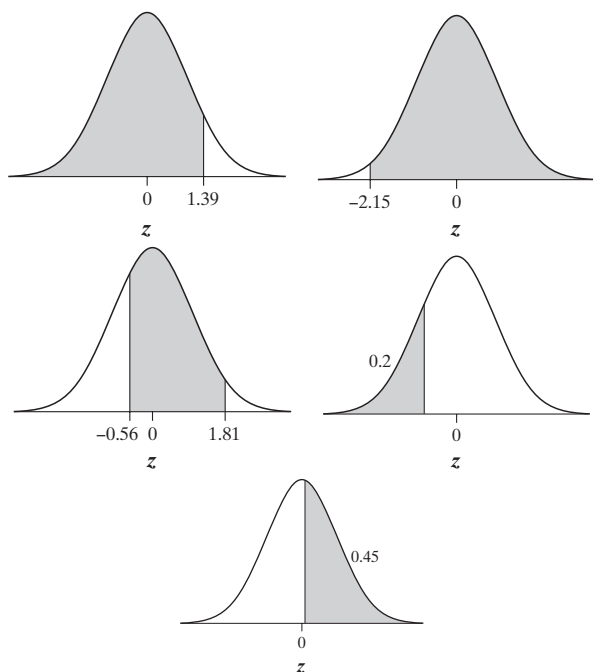
page 121: **1.** For 14-year-old boys, the amount of cholesterol follows a $N(170, 30)$ distribution and we want to find the percent of boys with cholesterol of more than 240 (see graph below). $z = \frac{240 - 170}{30} = 2.33$. From Table A, the proportion of z-scores above 2.33 is $1 - 0.9901 = 0.0099$. *Using technology:* `normalcdf (lower:240,upper:1000,μ:170,σ:30)` = 0.0098. About 1% of 14-year-old boys have cholesterol above 240 mg/dl. **2.** For 14-year-old boys, the amount of cholesterol follows a $N(170, 30)$ distribution and we want to find the percent of boys with cholesterol between 200 and 240 (see graph below). $z = \frac{200 - 170}{30} = 1$ and $z = \frac{240 - 170}{30} = 2.33$. From Table A, the proportion of z-scores between 1 and 2.33 is $0.9901 - 0.8413 = 0.1488$. *Using technology:* `normalcdf(lower:200,upper: 240,μ:170,σ:30)` = 0.1488. About 15% of 14-year-old boys have cholesterol between 200 and 240 mg/dl. **3.** For Tiger Woods, the distance his drives travel follows an $N(304, 8)$ distribution and the 80th percentile is the boundary value x with 80% of the distribution to its left (see graph below). A z-score of 0.84 gives the area closest to 0.80 (0.7995). Solving $0.84 = \frac{x - 304}{8}$ gives $x = 310.7$. *Using technology:* `invNorm(area:0.8,μ:304,σ:8)` = 310.7. The 80th percentile of Tiger Woods's drive lengths is about 310.7 yards.

Cholesterol levels

Cholesterol levels

2.43 (a) Between $69 - 2(2.5) = 64$ and $69 + 2(2.5) = 74$ inches. (b) About $\frac{100\% - 95\%}{2} = 2.5\%$. (c) About $\frac{100\% - 68\%}{2} = 16\%$ of men are shorter than 66.5 inches and $\frac{100\% - 95\%}{2} = 2.5\%$ are shorter than 64 inches, so approximately $16\% - 2.5\% = 13.5\%$ of men have heights between 64 inches and 66.5 inches. (d) Because $\frac{100\% - 68\%}{2} = 16\%$ of the area is to the right of 71.5, 71.5 is at the 84th percentile.

2.45 Taller curve: standard deviation ≈ 0.2. Shorter curve: standard deviation ≈ 0.5.

2.47 (a) 0.9978. (b) $1 - 0.9978 = 0.0022$ (c) $1 - 0.0485 = 0.9515$ (d) $0.9978 - 0.0485 = 0.9493$

2.49 (a) $0.9505 - 0.0918 = 0.8587$ (b) $0.9633 - 0.6915 = 0.2718$

2.51 (a) $z = -1.28$ (b) $z = 0.41$

2.53 (a) The length of pregnancies follows a $N(266, 16)$ distribution and we want the proportion of pregnancies that last less than 240 days (see graph below). $z = \frac{240 - 266}{16} = -1.63$. From Table A, the proportion of z-scores less than -1.63 is 0.0516. *Using technology:* `normalcdf(lower:-1000,upper:240,μ:266, σ:16)` = 0.0521. About 5% of pregnancies last less than 240 days, so 240 days is at the 5th percentile of pregnancy lengths.

Length of pregnancy (days)
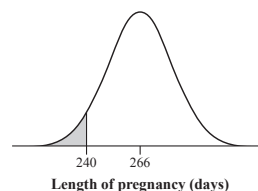
(b) The length of pregnancies follows a $N(266, 16)$ distribution and we want the proportion of pregnancies that last between 240 and 270 days (see the following graph). $z = \frac{240 - 266}{16} = -1.63$ and $z = \frac{270 - 266}{16} = 0.25$. From Table A, the proportion of z-scores between -1.63 and 0.25 is $0.5987 - 0.0516 = 0.5471$. *Using technology:* `normalcdf(lower:240,upper:270,`

μ:266,σ:16) = 0.5466. About 55% of pregnancies last between 240 and 270 days.

Length of pregnancy (days)
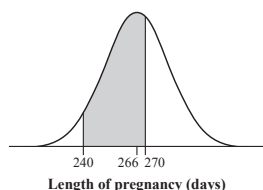
(c) The length of pregnancies follows a N(266, 16) distribution and we are looking for the boundary value x that has an area of 0.20 to the right and 0.80 to the left (see graph below). A z-score of 0.84 gives the area closest to 0.80 (0.7995). Solving $0.84 = \dfrac{x - 266}{16}$ gives $x = 279.44$. *Using technology:* `invNorm(area:0.8,\mu:266,\sigma:16)` = 279.47. The longest 20% of pregnancies last longer than 279.47 days.

Length of pregnancy (days)
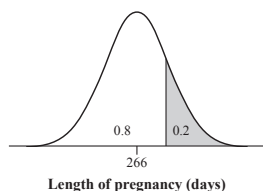
2.55 (a) For large lids, the diameter follows a N(3.98, 0.02) distribution and we want to find the percent of lids that have diameters less than 3.95 (see graph below). $z = \dfrac{3.95 - 3.98}{0.02} = -1.5$. From Table A, the proportion of z-scores below -1.5 is 0.0668. *Using technology:* `normalcdf(lower:-1000,upper:3.95,` μ:3.98,σ:0.02) = 0.0668. About 7% of the large lids are too small to fit.

Lid width (inches)

(b) For large lids, the diameter follows a N(3.98, 0.02) distribution and we want to find the percent of lids that have diameters greater than 4.05 (see graph below). $z = \dfrac{4.05 - 3.98}{0.02} = 3.5$. From Table A, the proportion of z-scor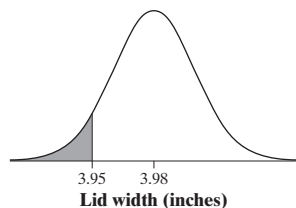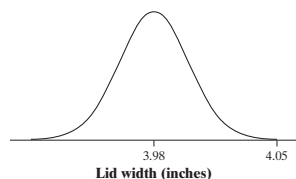es above 3.50 is approximately 0. *Using technology:* `normalcdf(lower:4.05,upper:1000,\mu:3.98,` σ:0.02) = 0.0002. Approximately 0% of the large lids are too big to fit.

Lid width (inches)

(c) Make a larger proportion of lids too small. If lids are too small, customers will just try another lid. But if lids are too large, the customer may not notice and then spill the drink.

2.57 (a) For large lids, the diameter follows a N(μ, 0.02) distribution and we want to find the value of μ that will result in only 1% of lids that are too small to fit (see graph below). A z-score of -2.33 gives the value closest to 0.01 (0.0099). Solving $-2.33 = \dfrac{3.95 - \mu}{0.02}$ gives $\mu = 4.00$. *Using technology:* `invNorm(area:0.01,\mu:0,\sigma:1)` gives $z = -2.326$. Solving $-2.326 = \dfrac{3.95 - \mu}{0.02}$ gives $\mu = 4.00$. The manufacturer should set the mean diameter to approximately $\mu = 4.00$ to ensure that only 1% of lids are too small. **(b)** For large lids, the diameter follows a N(3.98, σ) distribution and we want to find the value of σ that will result in only 1% of lids that are too small to fit (see graph below). A z-score of -2.33 gives the value closest to 0.01 (0.0099). Solving $-2.33 = \dfrac{3.95 - 3.98}{\sigma}$ gives $\sigma = 0.013$. *Using technology:* `invNorm(area:0.01,\mu:0,\sigma:1)` gives $z = -2.326$. Solving $-2.326 = \dfrac{3.95 - 3.98}{\sigma}$ gives $\sigma = 0.013$. A standard deviation of at most 0.013 will result in only 1% of lids that are too small to fit.

Diameter (inches)
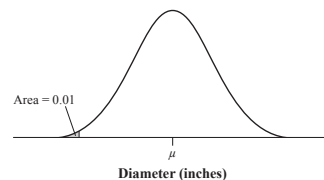
(c) Reduce the standard deviation. This will reduce the number of lids that are too small and the number of lids that are too big. If we make the mean a little larger as in part (a), we will reduce the number of lids that are too small, but we will increase the number of lids that are too big.

Diameter (inches)
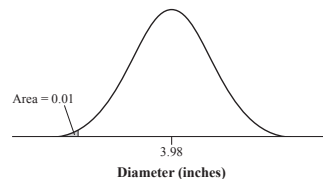
2.59 (a) $z = -1.28$ and $z = 1.28$ **(b)** Solving $-1.28 = \dfrac{x - 64.5}{2.5}$ gives $x = 61.3$ inches and solving $1.28 = \dfrac{x - 64.5}{2.5}$ gives $x = 67.7$ inches.

2.61 Solving $1.04 = \dfrac{60 - \mu}{\sigma}$ and $1.88 = \dfrac{75 - \mu}{\sigma}$ gives $\mu = 41.43$ minutes and $\sigma = 17.86$ minutes.

2.63 (a) A histogram is given below. The distribution of shark lengths is roughly symmetric and somewhat bell-shaped, with a mean of 15.586 feet and a standard deviation of 2.55 feet. **(b)** 30/44 = 68.2% , 42/44 = 95.5%, and 44/44 = 100%. These are very close to the 68−95−99.7 rule.
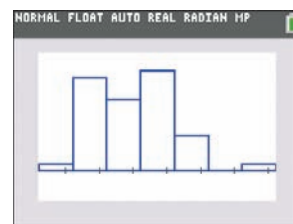
(c) A Normal probability plot is given below. Except for one small shark and one large shark, the plot is fairly linear, indicating that the distribution of shark lengths is approximately Normal.

Length of shark (feet)
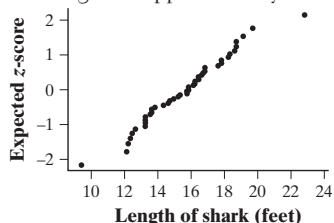
(d) All indicate that shark lengths are approximately Normal.

2.65 The distribution is close to Normal because the plot is nearly linear. There is a small "wiggle" between 120 and 130, with several values a little larger than would be expected in a Normal distribution. Also, the smallest value and the two largest values are a little farther from the mean than would be expected in a Normal distribution.

2.67 No. If it was Normal, then the minimum value should be around 2 or 3 standard deviations below the mean. However, the actual minimum has a z-score of just $z = -1.09$. Also, if the distribution was Normal, the minimum and maximum should be about the same distance from the mean. However, the maximum is much farther from the mean (20,209) than the minimum (8741).

2.69 b

2.71 b

2.73 a

2.75 For both kinds of cars, we see that the highway mileage is greater than the city mileage. The two-seater cars have a more variable distribution, both on the highway and in the city. Also the mileage values are slightly lower for the two-seater cars than for the minicompact cars, both on the highway and in the city, with a greater difference on the highway. All four distributions are roughly symmetric.

Answers to Chapter 2 Review Exercises

R2.1 **(a)** $z = 1.20$. Paul's height is 1.20 standard deviations above the average male height for his age. **(b)** 85% of boys Paul's age are shorter than Paul.

R2.2 **(a)** 58th percentile **(b)** $IQR = 11 - 2.5 = 8.5$ hours per week.

R2.3 **(a)** The shape of the distribution would not change.

Mean $= \dfrac{43.7}{3.28} = 13.32$ meters, median $= \dfrac{42}{3.28} = 12.80$ meters,

standard deviation $= \dfrac{12.5}{3.28} = 3.81$ meters,

$IQR = \dfrac{12.5}{3.28} = 3.81$ meters. **(b)** Mean $= 43.7 - 42.6 = 1.1$ feet; standard deviation $= 12.5$ feet, because subtracting a constant from each observation does not change the spread.

R2.4 **(a)** The median (line A in the graph below) should be slightly to the right of the main peak, with half of the area to the left and half to the right. **(b)** The mean (line B in the graph below) should be slightly to the right of the line for the median at the balancing point.

A B
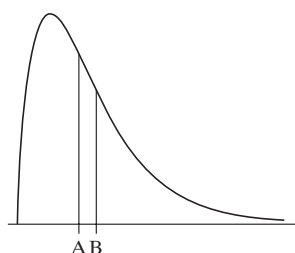
R2.5 **(a)** Between $336 - 3(3) = 327$ days and $336 + 3(3) = 345$ days. **(b)** About $\dfrac{100\% - 68\%}{2} = 16\%$.

R2.6 **(a)** $0.9616 - 0.0122 = 0.9494$ **(b)** If 35% of all values are greater than a particular z-value, then 65% are lower. A z-score of 0.39 gives the value closest to 0.65 (0.6517). *Using technology:* `invNorm(area:0.65,μ:0,σ:1)` gives $z = 0.385$.
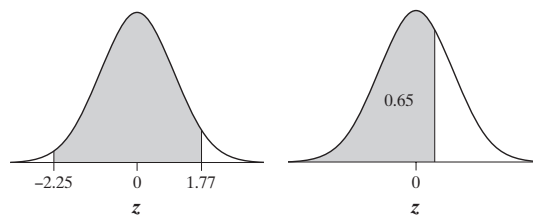
R2.7 **(a)** Birth weights follow a $N(3668, 511)$ distribution and we want to find the percent of babies with weights less than 2500 grams (see graph below). $z = \dfrac{2500 - 3668}{511} = -2.29$. From Table A, the proportion of z-scores below -2.29 is 0.0110. *Using technology:* `normalcdf(lower:-1000,upper:2500,μ:3668,σ:511) = 0.0111`. About 1% of babies will be identified as low birth weight.

Birth weights (grams)
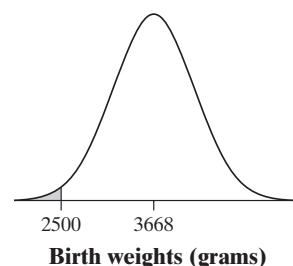
(b) Birth weights follow a $N(3668, 511)$ distribution. The 1st quartile is the boundary value with 25% of the area to its left. The 3rd quartile is the boundary value with 75% of the area to its left (see graph below). A z-score of -0.67 gives the value closest to 0.25 (0.2514). Solving $-0.67 = \dfrac{x - 3668}{511}$ gives $Q_1 = 3325.63$. A z-score of 0.67 gives the value closest to 0.75 (0.7486). Solving $0.67 = \dfrac{x - 3668}{511}$ gives $Q_3 = 4010.37$. *Using technology:* `invNorm(area:0.25,μ:3668,σ:511)` gives $Q_1 = 3323.34$ and `invNorm(area:0.75,μ:3668,σ:511)` gives $Q_3 = 4012.66$. The quartiles are $Q_1 = 3323.34$ grams and $Q_3 = 4012.66$ grams.

Birth weight (grams)
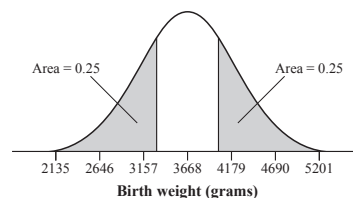
R2.8 **(a)** The amount of ketchup dispensed follows a $N(1.05, 0.08)$ distribution and we want to find the percent of times that the amount of ketchup dispensed will be between 1 and 1.2 ounces (see

graph below). $z = \dfrac{1.2 - 1.05}{0.08} = 1.88$ and $z = \dfrac{1 - 1.05}{0.08} = -0.63$.

From Table A, the proportion of z-scores between -0.63 and 1.88 is $0.9699 - 0.2643 = 0.7056$. *Using technology:* `normalcdf (lower:1,upper:1.2,μ:1.05,σ:0.08)` $= 0.7036$. About 70% of the time the dispenser will put between 1 and 1.2 ounces of ketchup on a burger.
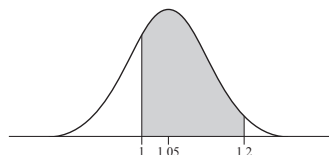
(b) The amount of ketchup dispensed follows a N(1.1, σ) distribution and we want to find the value of σ that will result in at least 99% of burgers getting between 1 and 1.2 ounces of ketchup (see graph below). Because the mean of 1.1 is in the middle of the interval from 1 to 1.2, we are looking for the middle 99% of the distribution. This leaves 0.5% in each tail. A z-score of -2.58 gives the value closest to 0.005 (0.0049). Solving $-2.58 = \dfrac{1 - 1.1}{\sigma}$ gives $\sigma = 0.039$. *Using technology:* `invNorm(area:0.005,μ:0,σ:1)` gives $z = -2.576$. Solving $-2.576 = \dfrac{1 - 1.1}{\sigma}$ gives $\sigma = 0.039$. A standard deviation of at most 0.039 ounces will result in at least 99% of burgers getting between 1 and 1.2 ounces of ketchup.
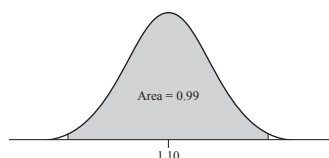
R2.9 If the distribution is Normal, the 10th and 90th percentiles must be equal distances above and below the mean. Thus, the mean is $\dfrac{25 + 475}{2} = 250$ points. The 10th percentile in a standard Normal distribution is $z = -1.28$. Solving $-1.28 = \dfrac{25 - 250}{\sigma}$, we get $\sigma = 175.8$. *Using technology:* `invNorm(area:0.10,μ:0,σ:1)` gives $z = -1.282$, so $-1.282 = \dfrac{25 - 250}{\sigma}$ and $\sigma = 175.5$.

R2.10 A histogram and Normal probability plot are given below. The histogram is roughly symmetric but not very bell-shaped. The Normal probability plot, however, is roughly linear. For these data, $\bar{x} = 0.8004$ and $s_x = 0.0782$. Although the percentage within 1 standard deviation of the mean (55.1%) is less than expected (68%), the percentage within 2 (93.9%) and 3 standard deviations (100%) match the 68−95−99.7 rule quite well. It is reasonable to say that these data are approximately Normally distributed.
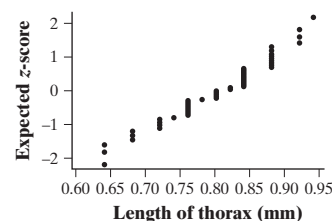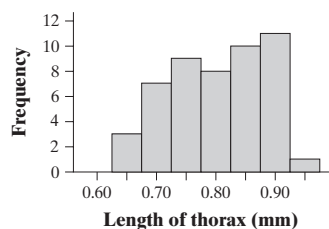
R2.11 The steep, nearly vertical portion at the bottom and the clear bend to the right indicate that the distribution of the data is right-skewed with several outliers and not approximately Normally distributed.

Answers to Chapter 2 AP® Statistics Practice Test

T2.1 e
T2.2 d
T2.3 b
T2.4 b
T2.5 a
T2.6 e
T2.7 c
T2.8 e
T2.9 e
T2.10 c
T2.11 **(a)** Jane's performance was better. Because her performance (40) exceeded the standard for the Presidential award (39), she performed above the 85th percentile. Matt's performance (40) met the standard for the National award (40), meaning he performed at the 50th percentile. **(b)** Because Jane's score has a higher percentile than Matt's score, she is farther to the right in her distribution than Matt is in his. Therefore, Jane's standardized score will likely be greater than Matt's.

T2.12 **(a)** For male soldiers, head circumference follows a N(22.8, 1.1) distribution and we want to find the percent of soldiers with head circumference less than 23.9 inches (see graph below). $z = \dfrac{23.9 - 22.8}{1.1} = 1$. From Table A, the proportion of z-scores below 1 is 0.8413. *Using technology:* `normalcdf(lower: −1000,upper:23.9,μ:22.8,σ:1.1)` $= 0.8413$. About 84% of soldiers have head circumferences less than 23.9 inches. Thus, 23.9 inches is at the 84th percentile.
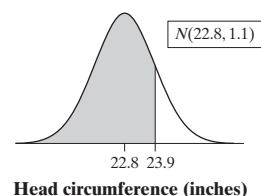
(b) For male soldiers, head circumference follows a N(22.8, 1.1) distribution and we want to find the percent of soldiers with head circumferences less than 20 inches or greater than 26 inches (see graph below). $z = \dfrac{20 - 22.8}{1.1} = -2.55$ and $z = \dfrac{26 - 22.8}{1.1} = 2.91$.

From Table A, the proportion of z-scores below $z = -2.55$ is 0.0054 and the proportion of z-scores above 2.91 is $1 - 0.9982 = 0.0018$, for a total of $0.0054 + 0.0018 = 0.0072$. *Using technology:* $1 - $ `normalcdf(lower:20,upper:26,μ:22.8,σ:1.1)` $= 1 - 0.9927 = 0.0073$. A little less than 1% of soldiers have head

circumferences less than 20 inches or greater than 26 inches and require custom helmets.

Head circumference (inches)
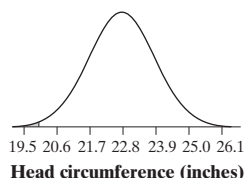
(c) For male soldiers, head circumference follows a $N(22.8, 1.1)$ distribution. The 1st quartile is the boundary value with 25% of the area to its left. The 3rd quartile is the boundary value with 75% of the area to its left (see graph below). A z-score of -0.67 gives the value closest to 0.25 (0.2514). Solving $-0.67 = \dfrac{x - 22.8}{1.1}$ gives $Q_1 = 22.063$. A z-score of 0.67 gives the value closest to 0.75 (0.7486). Solving $0.67 = \dfrac{x - 22.8}{1.1}$ gives $Q_3 = 23.537$. *Using technology:* `invNorm(area:0.25,μ:22.8,σ:1.1)` gives $Q_1 = 22.058$ and `invNorm(area:0.75,μ:22.8,σ:1.1)` gives $Q_3 = 23.542$. Thus, $IQR = 23.542 - 22.058 = 1.484$ inches.

Area = 0.25   Area = 0.25

Head circumference (inches)
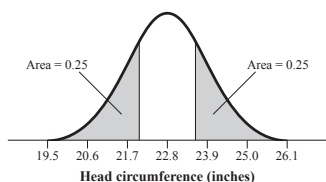
T2.13 No. First, there is a large difference between the mean and the median. In a Normal distribution, the mean and median are the same, but in this distribution the mean is 48.25 and the median is 37.80. Second, the distance between the minimum and the median is 35.80 but the distance between the median and the maximum is 167.10. In a Normal distribution, these distances should be about the same. Both of these facts suggest that the distribution is skewed to the right.

Chapter 3

Section 3.1

Answers to Check Your Understanding

page 144: **1.** Explanatory: number of cans of beer. Response: blood alcohol level. **2.** Explanatory: amount of debt and income. Response: stress caused by college debt.

page 149: **1.** Positive. The longer the duration of the eruption, the longer we should expect to wait between eruptions because long eruptions use more energy and it will take longer to build up the energy needed to erupt again. **2.** Roughly linear with two clusters. The clusters indicate that, in general, there are two types of eruptions—shorter eruptions that last around 2 minutes and longer eruptions that last around 4.5 minutes. **3.** Fairly strong. The points don't deviate much from the linear form. **4.** There are a few possible outliers around the clusters. However, there aren't many and potential outliers are not very distant from the main clusters of points.   **5.** How long the previous eruption was.

page 153: **(a)** $r \approx 0.9$. This indicates that there is a strong, positive linear relationship between the number of boats registered in Florida and the number of manatees killed. **(b)** $r \approx 0.5$. This indicates that there is a moderate, positive linear relationship between the number of named storms predicted and the actual number of named storms. **(c)** $r \approx 0.3$. This indicates that there is a weak, positive linear relationship between the healing rate of the two front limbs of the newts. **(d)** $r \approx -0.1$. This indicates that there is a weak, negative linear relationship between last year's percent return and this year's percent return in the stock market.

Answers to Odd-Numbered Section 3.1 Exercises

3.1 Explanatory: water temperature (quantitative). Response: weight change (quantitative).

3.3 **(a)** Positive. Students with higher IQs tend to have higher GPAs and vice versa because both IQ and GPA are related to mental ability. **(b)** Roughly linear, because a line through the scatterplot of points would provide a good summary. Moderately strong, because most of the points would be close to the line. **(c)** IQ ≈ 103 and GPA ≈ 0.4.

3.5 A scatterplot is shown below.
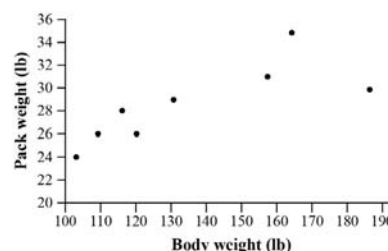
3.7 **(a)** There is a positive association between backpack weight and body weight. For students under 140 pounds, there seems to be a linear pattern in the graph. However, for students above 140 pounds, the association begins to curve. Because the points vary somewhat from the linear pattern, the relationship is only moderately strong. **(b)** The hiker with body weight 187 pounds and pack weight 30 pounds. This hiker makes the form appear to be nonlinear for weights above 140 pounds. Without this hiker, the association would look very linear for all body weights.

3.9 **(a)** A scatterplot is shown below. **(b)** The relationship is curved. Large amounts of fuel were used for low and high values of speed and smaller amounts of fuel were used for moderate speeds. This makes sense because the best fuel efficiency is obtained by driving at moderate speeds. **(c)** Both directions are present in the scatterplot. The association is negative for lower speeds and positive for higher speeds. **(d)** The relationship is very strong, with little deviation from a curve that can be drawn through the points.
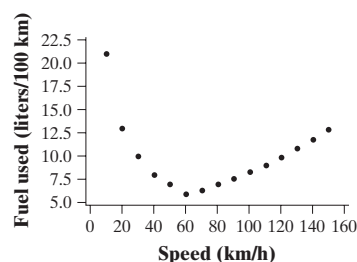
3.11 **(a)** Most of the southern states fall in the same pattern as the rest of the states. However, southern states typically have lower mean SAT math scores than other states with a similar percent of students taking the SAT. **(b)** West Virginia has a much lower mean

SAT Math score than the other states that have a similar percent of students taking the exam.

3.13 A scatterplot is shown below. There is a negative, linear, moderately strong relationship between the percent returning and the number of breeding pairs.
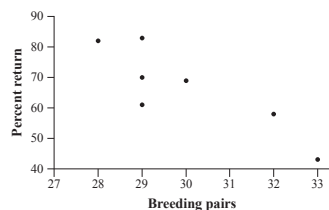
3.15 (a) $r = 0.9$ (b) $r = 0$ (c) $r = 0.7$ (d) $r = -0.3$ (e) $r = -0.9$

3.17 (a) Gender is a categorical variable and correlation r is for two quantitative variables. (b) The largest possible value of the correlation is $r = 1$. (c) The correlation r has no units.

3.19 (a) The scatterplot below shows a strong, positive linear relationship between the two measurements. It appears that all five specimens come from the same species. (b) The femur measurements have $\bar{x} = 58.2$ and $s_x = 13.2$. The humerus measurements have $\bar{y} = 66$ and $s_y = 15.89$. The sum of the z-score products is 3.97620, so the correlation coefficient is $r = (1/4)(3.97620) = 0.9941$. The very high value of the correlation confirms the strong, positive linear association between femur length and humerus length in the scatterplot from part (a).
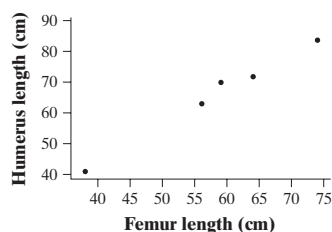
3.21 (a) There is a strong, positive linear association between sodium and calories. (b) It increases the correlation. It falls in the linear pattern of the rest of the data and observations with unusually small or unusually large values of x have a big influence on the correlation.

3.23 (a) The correlation would not change, because correlation is not affected by a change of units for either variable. (b) The correlation would not change, because it does not distinguish between explanatory and response variables.

3.25 (a) A scatterplot is shown below. (b) $r = 0$ (c) The correlation measures the strength of a *linear* association, but this plot shows a nonlinear relationship between speed and mileage.
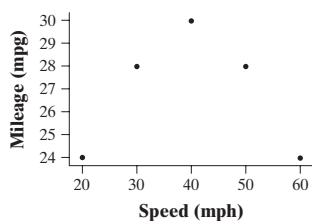
3.27 a

3.29 d

3.31 b

3.33 A histogram is shown below. The distribution is right-skewed, with several possible high outliers. Because of the skewness and outliers, we should use the median (5.4 mg) and *IQR* (5.5 mg) to describe the center and spread.
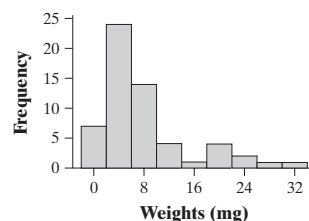
Section 3.2

Answers to Check Your Understanding

page 168: **1.** 40. For each additional week, we predict that a rat will gain 40 grams of weight. **2.** 100. The predicted weight for a newborn rat is 100 grams. **3.** $\hat{y} = 100 + 40(16) = 740$ grams **4.** 2 years = 104 weeks, so $\hat{y} = 100 + 40(104) = 4260$ grams. This is equivalent to 9.4 pounds (about the weight of a large newborn human). This is unreasonable and is the result of extrapolation.

page 172: The answer is given in the text.

page 174: **1.** $y - \hat{y} = 31{,}891 - 36{,}895 = -\$5004$ **2.** The actual price of this truck is \$5004 less than predicted based on the number of miles it has been driven. **3.** The truck with 44,447 miles and a price of \$22,896. This truck has a residual of $-\$8120$, which means that the line overpredicted the price by \$8120. No other truck had a residual that was farther below 0 than this one.

page 176: **1.** The backpack for this hiker was almost 4 pounds heavier than expected based on the weight of the hiker. **2.** Because there appears to be a negative-positive-negative pattern in the residual plot, a linear model is not appropriate for these data.

Answers to Odd-Numbered Section 3.2 Exercises

3.35 predicted weight $= 80 - 6$ (days)

**3.37** (a) 1.109. For each 1-mpg increase in city mileage, the predicted highway mileage will increase by 1.109 mpg. (b) 4.62 mpg. This would represent the highway mileage for a car that gets 0 mpg in the city, which is impossible. (c) 22.36 mpg

3.39 (a) -0.0053. For each additional week in the study, the predicted pH decreased by 0.0053 units. (b) 5.43. The predicted pH level at the beginning of the study (weeks $= 0$) is 5.43. (c) 4.635

3.41 No. 1000 months is well outside the observed time period and we can't be sure that the linear relationship continues after 150 weeks.

3.43 The line $\hat{y} = 1 - x$ is a much better fit. The sum of squared residuals for this line is only 3, while the sum of squared residuals for $\hat{y} = 3 - 2x$ is 18.

3.45 residual $= 5.08 - 5.165 = -0.085$. The actual pH value for that week was 0.085 less than predicted.

3.47 (a) The scatterplot (with regression line) is shown below. (b) $\hat{y} = 31.9 - 0.304x$. (c) For each increase of 1 in the percent of returning birds, the predicted number of new adult birds will decrease by 0.304. (d) residual $= 11 - 16.092 = -5.092$. In this colony, there were 5.092 fewer new adults than expected based on the percent of returning birds.
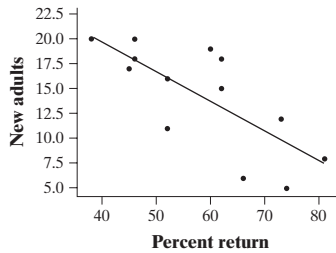
3.49 **(a)** Because there is no obvious leftover pattern in the residual plot shown below, a line is an appropriate model to use for these data. **(b)** The point with the largest residual (66% returning) has a residual of about −6. This means that the colony with 66% returning birds has about 6 fewer new adults than predicted based on the percent returning.
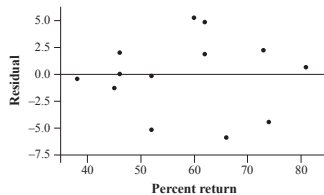
3.51 No. Because there is an obvious negative-positive-negative pattern in the residual plot, a linear model is not appropriate for these data.

3.53 **(a)** There is a positive, linear association between the two variables. There is more variation in the field measurements for larger laboratory measurements. **(b)** No. The points for the larger depths fall systematically below the line $y = x$, showing that the field measurements are too small compared to the laboratory measurements. **(c)** The slope would be closer to 0 and the y intercept would be larger.

3.55 **(a)** residual = 150.06 − 146.295 = 3.765. Yu-Na Kim's free skate score was 3.765 points higher than predicted based on her short program score. **(b)** Because there is no leftover pattern in the residual plot, a linear model is appropriate for these data. **(c)** When using the least-squares regression line with x = short program score to predict y = free skate score, we will typically be off by about 10.2 points. **(d)** About 73.6% of the variation in free skate scores is accounted for by the linear model relating free skate scores to short program scores.

3.57 r^2: About 56% of the variation in the number of new adults is accounted for by the linear model relating number of new adults to the percent returning. s: When using the least-squares regression line with x = percent returning to predict y = number of new adults, we will typically be off by 3.67 adults.

3.59 **(a)** $\hat{y} = 266.07 − 6.650x$, where y = percent of males that return the next year and x = number of breeding pairs. When $x = 30, \hat{y} = 66.57$. **(b)** R-Sq = 74.6% **(c)** $r = -\sqrt{0.746} = -0.864$. The sign is negative because the slope is negative. **(d)** When using the least-squares regression line with x = number of breeding pairs to predict y = percent returning, we will typically be off by 7.76%.

3.61 **(a)** $\hat{y} = 33.67 + 0.54x$. **(b)** If the value of x is 1 standard deviation below $\bar{x}$, the predicted value of y will be r standard deviations of y below $\bar{y}$. So the predicted value for the husband is $68.5 − 0.5(2.7) = 67.15$ inches.

3.63 **(a)** $r^2 = 0.25$. About 25% of the variation in husbands' heights is accounted for by the linear model relating husband's height to

wife's height. **(b)** When using the least-squares regression line with x = wife's height to predict y = husband's height, we will typically be off by 1.2 inches.

3.65 **(a)** $\hat{y} = x$ where y = final and x = midterm **(b)** If $x = 50$, $\hat{y} = 67.1$. If $x = 100, \hat{y} = 87.6$. **(c)** The student who did poorly on the midterm (50) is predicted to do better on the final (closer to the mean), while the student who did very well on the midterm (100) is predicted to do worse on the final (closer to the mean).

3.67 *State*: Is a linear model appropriate for these data? If so, how well does the least-squares regression line fit the data? *Plan*: We will look at the scatterplot and residual plot to see if the association is linear or nonlinear. Then, if a linear model is appropriate, we will use s and r^2 to measure how well the line fits the data. *Do*: The scatterplot below shows a moderately strong, positive linear association between the number of stumps and the number of clusters of beetle larvae. The residual plot doesn't show any obvious leftover pattern, confirming that a linear model is appropriate.
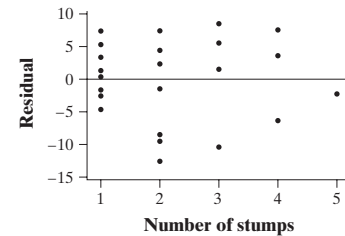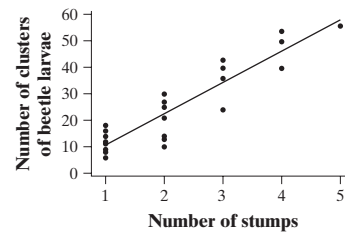
$\hat{y} = −1.29 + 11.89x$, where y = number of clusters of beetle larvae and x = number of stumps. $s = 6.42$, meaning that our predictions will typically be off by about 6.42 clusters when we use the line to predict the number of clusters of beetle larvae from the number of stumps. Finally, $r^2 = 0.839$, meaning 83.9% of the variation in the number of clusters of beetle larvae is accounted for by the linear model relating number of clusters of beetle larvae to the number of stumps. *Conclude*: The linear model relating number of clusters of beetle larvae to the number of stumps is appropriate and fits the data well, accounting for more than 80% of the variation in number of clusters of beetle larvae.

3.69 **(a)** A scatterplot is shown below. There is a moderate, positive linear association between HbA and FBG. There are possible outliers to the far right (subject 18) and near the top of the plot (subject 15).
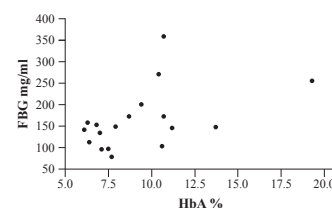
(b) Because the point is in the positive, linear pattern formed by most of the data values, it makes r closer to 1. Also, because the point is likely to be below the least-squares regression line, it will "pull down" the line on the right side, making the slope closer to 0. Without the outlier, r decreases from 0.4819 to 0.3837 as expected. Likewise, the equation changes from $\hat{y} = 66.4 + 10.4x$ to $\hat{y} = 52.3 + 12.1x$. **(c)** The point makes r closer to 0 because it is out of the linear pattern formed by most of the data values. Because this point's x coordinate is very close to $\bar{x}$ but the y coordinate is far above $\bar{y}$, it won't influence the slope very much but will increase the y intercept. Without the outlier, r increases from 0.4819 to 0.5684, as expected. Likewise, the equation changes from $\hat{y} = 66.4 + 10.4x$ to $\hat{y} = 69.5 + 8.92x$.

3.71 a

3.73 c

3.75 d

3.77 b

3.79 For these vehicles, the combined mileage follows a $N(18.7, 4.3)$ distribution and we want to find the percent of cars with lower mileage than 25 (see graph below). $z = \dfrac{25 - 18.7}{4.3} = 1.47$. From Table A, the proportion of z-scores below 1.47 is 0.9292. *Using technology:* `normalcdf(lower:-1000,upper:25,μ:18.7,σ:4.3)` $= 0.9286$. About 93% percent of vehicles get worse combined mileage than the Chevrolet Malibu.

N(18.7, 4.3)

18.7   25.0
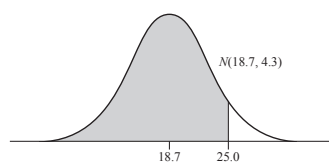
3.81 **(a)** A bar graph is given below. The people who use marijuana more are more likely to have caused accidents. **(b)** Association does not imply causation. For example, it could be that drivers who use marijuana more often are more willing to take risks than other drivers and that the willingness to take risks is what is causing the higher accident rate.
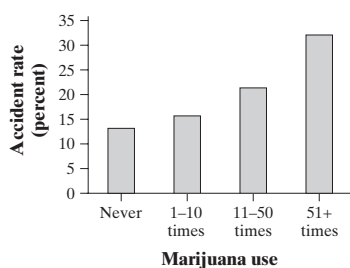
Answers to Chapter 3 Review Exercises

R3.1 **(a)** There is a moderate, positive linear association between gestation and life span. Without the outliers at the top and in the upper right, the association appears moderately strong, positive, and curved. **(b)** It makes r closer to 0 because it decreases the strength of what would otherwise be a moderately strong positive association. Because this point is close to $\bar{x}$ but far above $\bar{y}$, it won't affect the slope much but will increase the y intercept. Because it has such a large residual, it increases s. **(c)** Because it is in the positive, linear pattern formed by most of the data values, it will make r closer to 1. Also, because the point is likely to be above the least-squares regression line, it will "pull up" the line on the right side,

making the slope larger and the intercept smaller. Because this point is likely to have a small residual, it decreases s.

R3.2 **(a)** 0.0138. For each increase of 1 meter in dive depth, the predicted duration increases by 0.0138 minutes. **(b)** The y intercept suggests that a dive of 0 depth would last an average of 2.69 minutes; this obviously does not make any sense. **(c)** 5.45 minutes **(d)** If the variables are reversed, the correlation will remain the same. However, the slope and y intercept will be different.

R3.3 **(a)** $\hat{y} = 3704 + 12{,}188x$, where y represents the mileage of the cars and x represents the age. **(b)** residual $= 65{,}000 - 76{,}832 = -11{,}832$. This teacher has driven 11,832 fewer miles than predicted based on the age of the car. **(c)** $r = +\sqrt{0.837} = 0.915$. This shows that there is a strong, positive linear association between the age of cars and their mileage. **(d)** Yes, because there is no leftover pattern in the residual plot. **(e)** $s = 20{,}870.5$: When using the least-squares regression line with $x = $ car's age to predict $y = $ number of miles it has been driven, we will typically be off by about 20,870.5 miles. $r^2 = 83.7\%$: About 83.7% of the variability in mileage is accounted for by the linear model relating mileage to age.

R3.4 **(a)** The scatterplot is shown below. Average March temperature, because changes in March temperature probably have an effect on the date of first bloom.
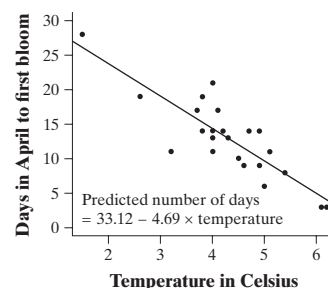
(b) $r = -0.85$ and $\hat{y} = 33.12 - 4.69x$, where y represents the number of days and x represents the temperature. r: There is a strong, negative linear association between the average March temperature and the days in April until first bloom. *Slope*: For every 1° increase in average March temperature, the predicted number of days in April until first bloom decreases by 4.69. *y intercept*: If the average March temperature was 0°C, the predicted number of days in April to first bloom is 33.12 (May 3). **(c)** No, $x = 8.2$ is well beyond the values of x we have in the data set. **(d)** residual $= 10 - 12.015 = -2.015$. In this year, the actual date of first bloom occurred about 2 days earlier than predicted based on the average March temperature. **(e)** There is no leftover pattern in the residual plot shown below, indicating that a linear model is appropriate.
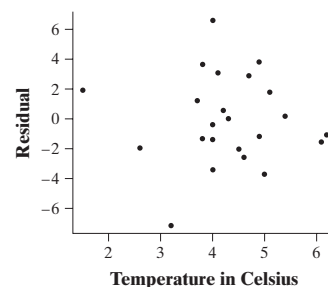
R3.5 (a) $\hat{y} = 30.2 + 0.16x$, where y = final exam score and x = total score before the final examination. (b) 78.2 (c) Of all the lines that the professor could use to summarize the relationship between final exam score and total points before the final exam, the least-squares regression line is the one that has the smallest sum of squared residuals. (d) Because $r^2 = 0.36$, only 36% of the variability in the final exam scores is accounted for by the linear model relating final exam scores to total score before the final exam. More than half (64%) of the variation in final exam scores is *not* accounted for, so Julie has reason to question this estimate.

R3.6 Even though there is a high correlation between number of calculators and math achievement, we shouldn't conclude that increasing the number of calculators will *cause* an increase in math achievement. It is possible that students who are more serious about school have better math achievement and also have more calculators.

Answers to Chapter 3 AP® Statistics Practice Test

T3.1 d
T3.2 e
T3.3 c
T3.4 a
T3.5 a
T3.6 c
T3.7 b
T3.8 e
T3.9 b
T3.10 c
T3.11 (a) A scatterplot with regression line is shown below. (b) $\hat{y} = 71.95 + 0.3833x$, where y = height and x = age. (c) 255.934 cm, or 100.76 inches (d) This was an extrapolation. Our data were based only on the first 5 years of life and the linear trend will not continue forever.
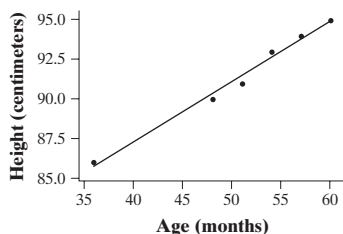
T3.12 (a) The point in the upper-right-hand corner has a very high silicon value for its isotope value. (b) (i) r would get closer to -1 because it does not follow the linear pattern of the other points. (ii) Because this point is "pulling up" the line on the right side of the plot, removing it will make the slope steeper (more negative) and the y intercept smaller (note that the y axis is to the *right* of the points in the scatterplot). (iii) Because this point has a large residual, removing it will make s a little smaller.

T3.13 (a) $\hat{y} = 92.29 - 0.05762x$, where y is the percent of the grass burned and x is the number of wildebeest. (b) For every increase of 1000 wildebeest, the predicted percent of grassy area burned decreases by about 0.058. (c) $r = -\sqrt{0.646} = -0.804$. There is a strong, negative linear association between the percent of grass burned and the number of wildebeest. (d) Yes, because there is no obvious leftover pattern in the residual plot.

Chapter 4
Section 4.1
Answers to Check Your Understanding

page 213: **1.** Convenience sampling. This could lead the inspector to overestimate the quality of the oranges if the farmer puts the best oranges on top. **2.** Voluntary response sampling. In this case, those who are happy that the UN has its headquarters in the U.S. already have what they want and so are less likely to respond. The proportion who answered "No" in the sample is likely to be higher than the true proportion in the U.S. who would answer "No."

page 223: **1.** You would have to identify 200 different seats, go to those seats in the arena, and find the people who are sitting there, which would take a lot of time. **2.** It is best to create strata where the people within a stratum are very similar to each other but different than the people in other strata. In this case, it would be better to take the lettered rows as the strata because each lettered row is the same distance from the court and so would contain only seats with the same (or nearly the same) ticket price. **3.** It is best if the people in each cluster reflect the variability found in the population. In this case, it would be better to take the numbered sections as the clusters because they include all different seat prices.

page 228: **1.** (a) Undercoverage (b) Nonresponse (c) Undercoverage **2.** By making it sound like they are not a problem in the landfill, this question will result in fewer people suggesting that we should ban disposable diapers. The proportion who would say "Yes" to this survey question is likely to be smaller than the proportion who would say "Yes" to a more fairly worded question.

Answers to Odd-Numbered Section 4.1 Exercises

4.1 Population: all local businesses. Sample: the 73 businesses that return the questionnaire.

4.3 Population: the 1000 envelopes stuffed during a given hour. Sample: the 40 randomly selected envelopes.

4.5 This is a voluntary response sample. In this case, it appears that people who strongly support gun control volunteered more often, causing the proportion in the sample to be greater than the proportion in the population.

4.7 This is a voluntary response sample and overrepresents the opinions of those who feel most strongly about the issue being surveyed.

4.9 (a) A convenience sample (b) The first 100 students to arrive at school likely had to wake up earlier than other students, so 7.2 hours is probably less than the true average.

4.11 (a) Number the 40 students from 01 to 40. Pick a starting point on the random number table. Record two-digit numbers, skipping numbers that aren't between 01 and 40 and any repeated numbers, until you have 5 unique numbers between 01 and 40. Use the 5 students corresponding to these numbers. (b) Using line 107, skip the numbers not in bold: 82 73 95 78 90 **20** 80 74 75 **11** 81 67 65 53 00 94 **38** **31** 48 93 60 94 **07**. Select Johnson (20), Drasin (11), Washburn (38), Rider (31), and Calloway (07).

4.13 (a) *Using calculator*: Number the plots from 1 to 1410. Use the command `randInt(1,1410)` to select 141 different integers from 1 to 1410 and use the corresponding 141 plots. (b) Answers will vary.

4.15 (a) False—although, on average, there will be four 0s in every set of 40 digits, the number of 0s can be less than 4 or greater than 4 by chance. (b) True—there are 100 pairs of digits 00 through 99,

and all are equally likely. (c) False—0000 is just as likely as any other string of four digits.

4.17 (a) It might be difficult to locate the 20 phones from among the 1000 produced that day. (b) The quality of the phones produced may change during the day, so that the last phones manufactured are not representative of the day's production. (c) Because each sample of 20 phones does not have the same probability of being selected. In an SRS, it is possible for 2 consecutive phones to be selected in a sample, but this is not possible with a systematic random sample.

4.19 Assign numbers 01 to 30 to the students. Pick a starting point on the random digit table. Record two-digit numbers, skipping any that aren't between 01 and 30 and any repeated numbers, until you have 4 unique numbers between 01 and 30. Use the corresponding four students. Then assign numbers 0 to 9 to the faculty members. Continuing on the table, record one-digit numbers, skipping any repeated numbers, until you have 2 unique numbers between 0 and 9. Use the corresponding faculty members. Starting on line 123 gives 08-Ghosh, 15-Jones, 07-Fisher, and 27-Shaw for the students and 1-Besicovitch and 0-Andrews for the faculty.

4.21 (a) Use the three types of seats as the strata because people who can afford more expensive tickets probably have different opinions about the concessions than people who can afford only the cheaper tickets. (b) A stratified random sample will include seats from all over the stadium, which would make it very time-consuming to obtain. A cluster sample of numbered sections would be easier to obtain, because the people selected for the sample would be sitting close together.

4.23 No. In an SRS, each possible sample of 250 engineers is equally likely to be selected, including samples that aren't exactly 200 males and 50 females.

4.25 (a) Cluster sampling. (b) To save time and money. In an SRS, the company would have to visit individual homes all over the rural subdivision instead of only 5 locations.

4.27 (a) It is unlikely, because different random samples will include different students and produce different estimates of the proportion of students who use Twitter. (b) An SRS of 100 students. Larger random samples give us better information about the population than smaller random samples.

4.29 Because you are sampling only from the lower-priced ticket holders, this will likely produce an estimate that is too small, as fans in the club seats and box seats probably spend more money at the game than fans in cheaper seats.

4.31 (a) 89.1% (b) Because the people who have long commutes are less likely to be at home and be included in the sample, this will likely produce an estimate that is too small.

4.33 We would not expect very many people to claim they have run red lights when they haven't, but some people will deny running red lights when they have. Thus, we expect that the sample proportion underestimates the true proportion of drivers who have run a red light.

4.35 (a) The wording is clear, but the question is slanted in favor of warning labels because of the first sentence stating that some cell phone users have developed brain cancer. (b) The question is clear, but it is slanted in favor of national health insurance by asserting it would reduce administrative costs and not providing any counterarguments. (c) The wording is too technical for many people to understand. For those who do understand the question, it is slanted because it suggests reasons why one should support recycling.

4.37 c

4.39 d

4.41 d

4.43 (a) For each additional day, the predicted sleep debt increases by about 3.17 hours. (b) The predicted sleep debt for a 5-day school week is $2.23 + 3.17(5) = 18.08$ hours. This is about 3 hours more than the researcher claimed for a 5-day week, so the students have reason to be skeptical of the research study's reported results.

Section 4.2

Answers to Check Your Understanding

page 237: **1.** Experiment, because a treatment (brightness of screen) was imposed on the laptops. **2.** Observational study, because students were not assigned to eat a particular number of meals with their family per week. **3.** Explanatory: number of meals per week eaten with their family. Response: GPA. **4.** There are probably other variables that are influencing the response variable. For example, students who have part-time jobs may not be able to eat many meals with their families and may not have much time to study, leading to lower grades.

page 247: **1.** Randomly assign the 29 students to two treatments: evaluating the performance in small groups or evaluating the performance alone. The response variable will be the accuracy of their final performance evaluations. To implement this design, use 29 equally sized slips of paper. Label 15 of them "small group" and 14 of them "alone." Then shuffle the papers and hand them out at random to the 29 students, assigning them to a treatment. **2.** The purpose of the control group is to provide a baseline for comparison. Without a group to compare to, it is impossible to determine if the small group treatment is more effective.

page 249: **1.** No. Perhaps seeing the image of their unborn child encouraged the mothers who had an ultrasound to eat a better diet, resulting in healthier babies. **2.** No. While the people weighing the babies at birth may not have known whether that particular mother had an ultrasound or not, the mothers knew. This might have affected the outcome because the mothers knew whether they had received the treatment or not. **3.** Treat all mothers as if they had an ultrasound, but for some mothers the ultrasound machine wouldn't be turned on. To avoid having mothers know the machine was turned off, the ultrasound screen would have to be turned away from all the mothers.

Answers to Odd-Numbered Section 4.2 Exercises

4.45 Experiment, because students were randomly assigned to the different teaching methods.

4.47 (a) Observational study, because mothers weren't assigned to eat different amounts of chocolate. (b) Explanatory: the mother's chocolate consumption. Response: the baby's temperament. (c) No, this study is an observational study so we cannot draw a cause-and-effect conclusion. It is possible that women who eat chocolate daily have less stressful lives and the lack of stress helps their babies to have better temperaments.

4.49 Type of school. For example, private schools tend to have smaller class sizes and students that come from families with higher socioeconomic status. If these students do better in the future, we wouldn't know if the better performance was due to smaller class sizes or higher socioeconomic status.

4.51 Experimental units: pine seedlings. Explanatory variable: light intensity. Response variable: dry weight at the end of the study. Treatments: full light, 25% light, and 5% light.

4.53 Experimental units: the individuals who were called. Explanatory variables: (1) information provided by interviewer; (2) whether caller offered survey results. Response variable: whether or not the call was completed. Treatments: (1) name/no offer; (2) university/no offer; (3) name and university/no offer; (4) name/offer; (5) university/offer; (6) name and university/offer.

4.55 Experimental units: 24 fabric specimens. Explanatory variables: (1) roller type; (2) dyeing cycle time; (3) temperature. Response variable: a quality score. Treatments: (1) metal, 30 min, 150°; (2) natural, 30 min, 150°; (3) metal, 40 min, 150°; (4) natural, 40 min, 150°; (5) metal, 30 min, 175°; (6) natural, 30 min, 175°; (7) metal, 40 min, 175°; (8) natural, 40 min, 175°.

4.57 There was no control group. We don't know if the improvement was due to the placebo effect or if the flavonols actually affected the blood flow.

4.59 (a) Write all names on slips of paper, put them in a container, and mix thoroughly. Pull out 40 slips of paper and assign these subjects to Treatment 1. Then pull out 40 more slips of paper and assign these subjects to Treatment 2. The remaining 40 subjects are assigned to Treatment 3. (b) Assign the students numbers from 1 to 120. Using the command RandInt (1,120) on the calculator, assign the students corresponding to the first 40 unique numbers chosen to Treatment 1, the students corresponding to the next 40 unique numbers chosen to Treatment 2, and the remaining 40 students to Treatment 3. (c) Assign the students numbers from 001 to 120. Pick a spot on Table D and read off the first 40 unique numbers between 001 and 120. The students corresponding to these numbers are assigned to Treatment 1. The students corresponding to the next 40 unique numbers between 001 and 120 are assigned to Treatment 2. The remaining 40 students are assigned to Treatment 3.

4.61 Random assignment. If players are allowed to choose which treatment they get, perhaps the more motivated players will choose the new method. If they improve more by the end of the study, the coach can't be sure if it was the exercise program or player motivation that caused the improvement.

4.63 *Comparison:* Researchers used a design that compared a low-carbohydrate diet with a low-fat diet. *Random assignment:* Subjects were randomly assigned to one of the two diets. *Control:* The experiment used subjects who were all obese at the beginning of the study and who all lived in the same area. *Replication:* There were 66 subjects in each treatment group.

4.65 Write the names of the patients on 36 identical slips of paper, put them in a hat, and mix them well. Draw out 9 slips. The corresponding patients will receive the antidepressant. Draw out 9 more slips. Those patients will receive the antidepressant plus stress management. The patients corresponding to the next 9 slips drawn will receive the placebo, and the remaining 9 patients will receive the placebo plus stress management. At the end of the experiment, record the number and severity of chronic tension-type headaches for each of the 36 subjects and compare the results for the 4 groups.

4.67 (a) Other variables include expense and condition of the patient. For example, if a patient is in very poor health, a doctor might choose not to recommend surgery because of the added complications. Then we won't know if a higher death rate is due to the treatment or the initial health of the subjects. (b) Write the names of all 300 patients on identical slips of paper, put them in a hat, and mix them well. Draw out 150 slips and assign the corresponding subjects to receive surgery. The remaining 150 subjects receive the new method. At the end of the study, count how many patients survived in each group.

4.69 The subjects developed rashes on the arm exposed to the placebo (a harmless leaf) simply because they thought they were being exposed to a poison ivy leaf. Likewise, most of the subjects didn't develop rashes on the arm that was exposed to poison ivy because they didn't think they were being exposed to the real thing.

4.71 Because the experimenter knew which subjects had learned the meditation techniques, he is not blind. If the experimenter believed that meditation was beneficial, he may subconsciously rate subjects in the meditation group as being less anxious.

4.73 (a) To make sure that the two groups were as similar as possible before the treatments were administered. (b) The difference in weight loss was larger than would be expected due to the chance variation created by the random assignment to treatments. (c) Even though the low-carb dieters lost 2 kg more over the year than the low-fat group, a difference of 2 kg could be due just to chance variation created by the random assignment.

4.75 (a) The different diagnoses, because the treatments were randomly assigned to patients within each diagnosis. (b) Using a randomized block design allows us to account for the variability in response due to differences in diagnosis by initially comparing the results within each block. In a completely randomized design, this variability will be unaccounted for, making it harder to determine if there is a difference in health and satisfaction due to the difference between doctors and nurse-practitioners.

4.77 (a) A randomized block design would help us account for the variability in yield that is due to the differences in fertility in the field, making it easier to determine if one variety is better than the others. (b) The rows. There should be a stronger association between row number and yield than column number and yield. (c) Let the digits 1 to 5 correspond to the five corn varieties A to E. Begin with line 111 on the random digit table, and assign the letters to the top row from left to right, ignoring numbers 0 and 6−9 and repeated numbers. Use a different line (111, 112, 113, 114, and 115) for each row. Top row (left to right): ADECB, second row: ECDAB, third row: BEDCA, fourth row: DEACB, bottom row: ADCBE.

4.79 (a) If all rats from litter 1 were fed Diet A and if these rats gained more weight, we would not know if this was because of the diet or because of genetics and initial health. (b) Use a randomized block design with the litters as blocks. For each of the litters, randomly assign half of the rats to receive Diet A and the other half to receive Diet B. This will allow researchers to account for the differences in weight gain caused by the differences in genetics and initial health.

4.81 (a) Matched pairs design. (b) In a completely randomized design, the differences between the students will add variability to the response, making it harder to detect if there is a difference caused by the treatments. In a matched pairs design, each student is compared with himself (or herself), so the differences between students are accounted for. (c) If all the students used the hands-free phone during the first session and performed worse, we wouldn't know if the better performance during the second session is due to the lack of phone or to learning from their mistakes the first time. By randomizing the order, some students will use the hands-free phone during the first session and others during the second session. (d) The simulator, route, driving conditions, and traffic flow were all kept the same for both sessions, preventing these variables from adding variability to the response variable.

4.83 (a) Randomly assign the 20 subjects into two groups of 10. Write the name of each subject on a note card, shuffle the cards, and select 10 to be assigned to the 70° environment. The remaining 10 subjects will be assigned to the 90° environment. Then the number of correct insertions will be recorded for each subject and the two groups compared. (b) All subjects will perform the task twice, once in each temperature condition. Randomly choose the order by flipping a coin. Heads: 70°, then 90°. Tails: 90°, then 70°. For each subject, compare the number of correct insertions in each environment.

4.85 (a) If the students find a difference between the two groups, they will not know if the difference is due to gender or the deodorant. (b) Each student should have one armpit randomly assigned to receive Deodorant A and the other Deodorant B. Because each gender uses both deodorants, there is no longer any confounding between gender and deodorant.

4.87 c

4.89 b

4.91 c

4.93 b

4.95 (a) For these seeds, the weights follow a N(525, 110) distribution and we want the proportion of seeds that weigh more than 500 mg (see graph below). $z = \dfrac{500 - 525}{110} = -0.23$. From Table A, the proportion of z-scores greater than -0.23 is $1 - 0.4090 = 0.5910$. *Using technology* normalcdf(lower:500,upper:10000, µ:525,σ:110) $= 0.5899$. About 59% of seeds will weigh more than 500 mg.

N(525,110)

500 525
Weight (mg)
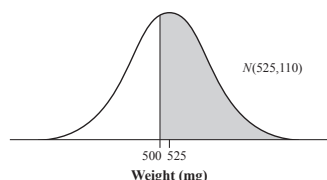
(b) For these seeds, the weights follow a N(525, 110) distribution and we are looking for the boundary value x that has an area of 0.10 to the left (see graph below). A z-score of -1.28 gives the closest value to 0.10 (0.1003). Solving $-1.28 = \dfrac{x - 525}{110}$ gives $x = 384.2$. *Using technology*: invNorm(area:0.10,µ:525,σ:110) = 384.0. The smallest weight among the remaining seeds should be about 384 mg.

Area = 0.10

N(525,110)

525
Weight (mg)
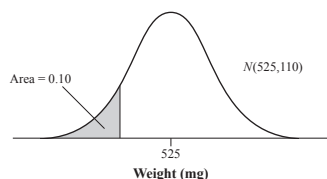
Section 4.3

Answers to Odd-Numbered Section 4.3 Exercises

4.97 If the study involves random sampling, we can make inferences about the population from which we sampled. If the study involves random assignment, we can make inferences about cause and effect.

4.99 Because this study involved random assignment to the treatments, we can infer that the difference between foster care or institutional care caused the difference in response.

4.101 Because this study did not involve random assignment to a treatment, we cannot infer cause and effect. Also, because the individuals were not randomly chosen, we cannot generalize to a larger population.

4.103 As daytime running lights become more common, they may be less effective at catching the attention of other drivers. Also, a driving simulator might not be very realistic.

4.105 Answers will vary.

4.107 Answers will vary.

4.109 Confidential. The person taking the survey knows who is answering the questions, but will not share the results of individuals with anyone else.

4.111 The subjects were not able to give informed consent. They did not know what was happening to them and they were not old enough to understand the ramifications.

4.113 The conditional distributions for males and females are displayed in the table and graph below. Men are more likely to view animal testing as justified if it might save human lives: over two-thirds of men agree or strongly agree with this statement, compared to slightly less than half of the women. The percentages who disagree or strongly disagree tell a similar story: 16% of men versus 30% of women.

Response	Male	Female
Strongly agree	14.7%	9.3%
Agree	52.3%	38.8%
Neither	16.9%	21.9%
Disagree	11.8%	19.3%
Strongly disagree	4.3%	10.7%
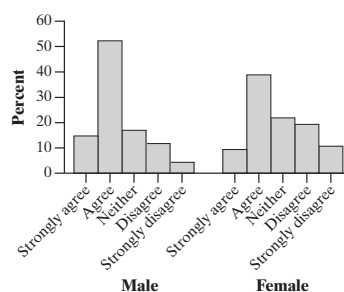
Answers to Chapter 4 Review Exercises

R4.1 (a) Population: all Ontario residents. Sample: the 61,239 people interviewed. (b) Because different samples will produce different estimates, it is unlikely that the percentages in the entire population would be exactly the same as the percentages in the sample. However, they should be fairly close.

R4.2 (a) Announce in a daily bulletin that there is a survey concerning student parking available in the main office for students who want to respond. Because those who feel strongly are more likely to respond, their opinions will be overrepresented. (b) Interview a group of students as they come in from the parking lot. People who already can park on campus might have different opinions about the parking situation than those who cannot.

R4.3 (a) Number the players from 01 to 25 in alphabetical order. Move from left to right, reading pairs of digits until you find three different pairs between 01 and 25, and select the corresponding players. (b) 17 (Musselman), 09 (Fuhrmann), and 23 (Smith).

R4.4 Stratified, because it is likely that the opinions of professors will vary based on which type of institution they are at. Then a stratified random sample will provide a more precise estimate than the other methods. Furthermore, the other methods might miss faculty from one particular type of institution.

R4.5 (a) People may not remember how many movies they watched in a movie theater in the past year. So shorten the amount of time that they ask about, perhaps 3 or 6 months. (b) This will underrepresent younger adults who use only cell phones. If younger adults go to movies more often than older adults, the estimated mean will be too small. (c) Because the frequent moviegoers will not be at home to respond, the estimated mean will be too small.

R4.6 (a) Different anesthetics were not randomly assigned to the subjects. (b) Type of surgery. If Anesthesia C is used more often with a type of surgery that has a higher death rate, we wouldn't know if the death rate was higher because of the anesthesia or the type of surgery.

R4.7 (a) Units: potatoes. Explanatory: storage method and time from slicing until cooking. Response: ratings of color and flavor. Treatments: (1) fresh/immediately, (2) fresh/after an hour, (3) room temperature/immediately, (4) room temperature/after an hour, (5) refrigerator/immediately, (6) refrigerator/after an hour. (b) Using 300 identical slips of paper, write "1" on 50 of them, "2" on 50 of them, and so on. Put the papers in a hat and mix well. Then select a potato and randomly select a slip from the hat to determine which treatment that potato will receive. Repeat this process for the remaining 299 potatoes, making sure not to replace the slips of paper into the hat. (c) Use a randomized block design with regular potatoes in one block and sweet potatoes in the other block. Randomly assign the 6 treatments within each block as in part (b).

R4.8 (a) No. The 1000 students were not randomly selected from any larger population. (b) Yes. The students were randomly assigned to the three treatments.

R4.9 (a) By giving some patients a treatment that should have no effect at all, but appears like the Saint-John's-wort, the researchers can account for the expectations of patients (the placebo effect) by comparing the results for the two groups. (b) To create two groups of subjects that are roughly equivalent at the beginning of the experiment. (c) The subjects should not know which treatment they are getting so that the researchers can account for the placebo effect. The researchers should be unaware of which subjects received which treatment so that they cannot influence how the results are measured. (d) The difference in improvement between the two groups wasn't large enough to rule out the chance variation caused by the random assignment to treatments.

R4.10 (a) Randomly assign 15 students to easy mazes and the other 15 to hard mazes. Use 30 identical slips of paper and write the name of each subject on a slip. Mix the slips in a hat, select 15 of them at random, and assign these subjects to hard mazes. The remaining 15 will be assigned to easy mazes. After the experiment, compare the time estimates of the two groups. (b) Each student does the activity twice, once with each type of maze. Randomly determine which set of mazes is used first by flipping a coin for each subject. Heads: easy, then hard. Tails: hard, then easy. After the experiment, compare each student's easy maze and hard maze time estimate. (c) The matched pairs design would be more likely to

detect a difference because it accounts for the variability between subjects.

R4.11 (a) This does not meet the requirements of informed consent because the subjects did not know the nature of the experiment before they agreed to participate. (b) All individual data should be kept confidential and the experiment should go before an institutional review board before being implemented.

Answers to Chapter 4 AP® Statistics Practice Test

T4.1 c
T4.2 e
T4.3 d
T4.4 c
T4.5 b
T4.6 b
T4.7 d
T4.8 d
T4.9 d
T4.10 b
T4.11 d
T4.12 (a) Experimental units: acacia trees. Treatments: placing either active beehives, empty beehives, or nothing in the trees. Response: damage to the trees caused by elephants. (b) Assign the trees numbers from 01 to 72 and use a random number table to pick 24 different two-digit numbers in this range. Those trees will get the active beehives. The trees corresponding to the next 24 different two-digit numbers from 01 to 72 will get the empty beehives, and the remaining 24 trees will remain empty. Compare the damage caused by elephants to the three groups of trees.

T4.13 (a) Not all possible samples of size 1067 were possible. For example, using their method, they could not have had all respondents from the east coast. (b) If the household members who typically answer the phone have a different opinion than those who don't typically answer the phone, their opinions will be overrepresented. (c) If people without phones or with cell phones only have different opinions than the group of people with residential lines, these opinions will be underrepresented.

T4.14 (a) Each of the 11 individuals will be a block in this matched pairs design, with the order of treatments randomly assigned. This was to help account for the variability in tapping speed caused by the differences in subjects. (b) If all the subjects got caffeine the second time, the researchers wouldn't know if the increase was due to the caffeine or due to practice with the task. (c) Yes. Neither the subjects nor the people who come in contact with them during the experiment (including those who record the number of taps) need to know the order in which the caffeine or placebo was administered.

Answers to Cumulative AP® Practice Test 1

AP1.1 d
AP1.2 e
AP1.3 b
AP1.4 c
AP1.5 a
AP1.6 c
AP1.7 e
AP1.8 e
AP1.9 d
AP1.10 d
AP1.11 d

AP1.12 b

AP1.13 b

AP1.14 a

AP1.15 (a) The distribution of gains for subjects using Machine A is roughly symmetric, while the distribution of gains for subjects using Machine B is skewed to the left. The center is greater for Machine B than for Machine A. The distribution for Machine B is more variable than the distribution for Machine A. (b) B. The typical gain using Machine B is greater than the typical gain using Machine A. (c) A. The spread for Machine A is less than the spread for Machine B. (d) Volunteers from one fitness center were used and these volunteers may be different in some way from the general population of those who are interested in cardiovascular fitness. To broaden their scope of inference, they should randomly select people from the population they would like to draw an inference about.

AP1.16 (a) Number the 60 retail sales districts with a two-digit number from 01 to 60. Using a table of random digits, read two-digit numbers until 30 unique numbers from 01 to 60 have been selected. The corresponding 30 districts are assigned to the monetary incentives group and the remaining 30 to the tangible incentives group. After a specified period of time, record the change in sales for each district and compare the two groups. (b) The districts labeled 07, 51, and 18 are the first three to be assigned to the monetary incentives group. (c) Pair the two districts with the largest sales, the next two largest, down to the two smallest districts. For each pair, pick one of the districts and flip a coin. If the flip is "heads," this district is assigned to the monetary incentives group. If it is "tails," this district is assigned to the tangible incentives group. The other district in the pair is assigned to the other group. After a specified period of time, record the change in sales for each district and compare within each pair.

AP1.17 (a) There is a very strong, positive, linear association between sales and shelf length. (b) $\hat{y} = 317.94 + 152.68x$, where y = weekly sales (in dollars) and x = shelf length (in feet). (c) $1081 (d) When using the least-squares regression line with x = shelf space to predict y = sales, we will typically be off by about $s = $23. (e) $$$ About 98.2% of the variation in weekly sales revenue can be accounted for by the linear model relating sales to shelf length. (f) It would be inappropriate to interpret the intercept, because the data represent sales based on shelf lengths of 3 to 6 feet and 0 feet falls substantially outside that domain.

Chapter 5

Section 5.1

Answers to Check Your Understanding

page 292: **1.** (a) If you asked a large sample of U.S. adults whether they usually eat breakfast, about 61% of them will answer yes. (b) The exact number of breakfast eaters will vary from sample to sample. **2.** (a) 0. If an outcome can never occur, then it will occur in 0% of the trials. (b) 1. If an outcome will occur on every trial, then it will occur in 100% of the trials. (c) 0.01. An outcome that occurs in 1% of the trials is very unlikely, but will occur every once in a while. (d) 0.6. An outcome that occurs in 60% of the trials will happen more than half of the time.

page 299: **1.** Assign the members of the AP® Statistics class the numbers 01−28 and the rest of the students numbers 29−95. Ignore the numbers 96−99 and 00. In Table D, read off 4 two-digit numbers, making sure that the second number is different than the first and that the fourth number is different than the third. Record whether all four numbers are between 01 and 28 or not. **2.** Assign the numbers 1−10 to Jeff Gordon, 11−40 to Dale Earnhardt, Jr., 41−60 to Tony Stewart, 61−85 to Danica Patrick, and 86−100 to Jimmie Johnson. Then proceed as in the example.

Answers to Odd-Numbered Section 5.1 Exercises

5.1 (a) If we use a polygraph machine on many, many people who are all telling the truth, the machine will say about 8% of the people are lying. (b) Answers will vary. A false positive would mean that a person telling the truth would be found to be lying. A false negative would mean that a person lying would be found to be telling the truth.

5.3 (a) If we look at many families like this, approximately 25% of them will have a first-born child that develops cystic fibrosis. (b) No. The number of children with cystic fibrosis could be smaller or larger than 4 by random chance.

5.5 (a) Answers will vary. (b) Spin the coin many more times.

5.7 In the short run, there was quite a bit of variability in the percentage of made free throws. However, this percentage became less variable and approached 0.30 as the number of shots increased.

5.9 No, he is incorrectly applying the law of large numbers to a small number of at-bats.

5.11 (a) There are 10,000 four-digit numbers (0000, 0001, . . . , 2873, . . . , 9999), and each is equally likely. (b) 2873. To many, 2873 "looks" more random than 9999—we don't "expect" to get the same number four times in a row. It would be best to choose a number that others would avoid so you don't have to split the pot with many other people.

5.13 (a) Let diamonds, spades, and clubs represent making a free throw and hearts represent missing. Deal one card from the deck. (b) Let 00−74 represent making the free throw and 75−99 represent missing. Read a two-digit number from Table D. (c) Let 1−3 represent the player making the free throw and 4 represent a miss. Generate a random integer from 1−4.

5.15 (a) There are 19 (not 18) numbers from 00 to 18, 19 (not 18) numbers from 19 to 37, and 3 (not 2) numbers from 38 to 40. (b) Repeats should not be skipped. For example, if the first number selected was 08, then the probability of selecting a left-hander on the next selection would be 9% (instead of 10%).

5.17 (a) Valid. The chance of rolling a 1, 2, or 3 is 75% on a 4-sided die and the 100 rolls represent the 100 randomly selected U.S. adults. (b) Not valid. The probability of heads is 50% rather than 60%. This method will underestimate the number of times she hits the center of the target.

5.19 (a) What is the probability that, in a random selection of 10 passengers, none from first class are chosen? (b) Number the first-class passengers 01−12 and the other passengers 13−76. Look up two-digit numbers in Table D until you have 10 unique numbers from 01 to 76. Count the numbers between 01 and 12. (c) 71 48 70 99 84 29 07 71 48 63 61 68 34 70 52. There is one person selected who is in first class. (d) It seems plausible that the actual selection was random, because 15/100 is not very small.

5.21 (a) Use a random integer generator to select 30 numbers from 1 to 365. Record whether or not there were any repeats in the sample. (b) Answers will vary. (c) Answers will vary.

5.23 (a) Obtaining a sample percentage of 55% or higher is not particularly unusual (probability ≈ 43/200) when 50% of all students recycle. (b) Obtaining a sample percentage of at least 63% is very unlikely (probability ≈ 1/200) when 50% of all students recycle.

5.25 *State*: What is the probability that, in a sample of 4 randomly selected U.S. adult males, at least one of them is red-green colorblind? *Plan*: Let 00−06 denote a colorblind man and 07−99 denote a non-colorblind man. Read 4 two-digit numbers from Table D for each sample and record whether or not the sample had at least one red-green colorblind man in it. *Do*: In our 50 samples, 15 had at least one colorblind man in them. *Conclude*: The probability that a sample of 4 men would have at least one colorblind man is approximately 15/50 = 0.30.

5.27 *State*: What is the probability that it takes 20 or more selections in order to find one man who is red-green colorblind? *Plan*: Let 0−6 denote a colorblind man and 7−99 denote a non-colorblind man. Use technology to pick integers from 0 to 99 until we get a number between 0 and 6. Count how many numbers there are in the sample. *Do*: In 16 of our 50 samples, it took 20 or more selections to get one colorblind man. *Conclude*: Not surprised. The probability of needing 20 or more selections to get one colorblind man is fairly large (approximately 16/50 = 0.32).

5.29 *State*: What is the probability that the random assignment will result in at least 6 men in the same group? *Plan*: Number the men 1−8 and the women 9−20. Use technology to pick 10 unique integers between 1 and 20 for one group. Record if there are at least 6 numbers between 1 and 8 in either group. *Do*: In our 50 repetitions, 9 had one group with 6 or more men in it. *Conclude*: Not surprised. The probability of getting 6 or more men in one group is fairly large (approximately 9/50 = 0.18).

5.31 c

5.33 b

5.35 c

5.37 (a) Population: adult U.S. residents. Sample: the 353,564 adults who were interviewed. (b) The people who do not have a telephone were excluded. This would lead to an underestimate of the proportion in the population who experienced stress a lot of the day yesterday if the people without phones are poorer and consequently experience more stress.

Section 5.2
Answers to Check Your Understanding
page 309: **1.** A person cannot have a cholesterol level of both 240 or above and between 200 and 239 at the same time. **2.** A person has either a cholesterol level of 240 or above, or they have a cholesterol level between 200 and 239. $P(A \text{ or } B) = 0.16 + 0.29 = 0.45$. **3.** $P(C) = 1 - 0.45 = 0.55$.

page 311: **1.**

	Face card	Non-face card	Total
Heart	3	10	13
Non-heart	9	30	39
Total	12	40	52

2. $P(F \text{ and } H) = 3/52 = 0.058$. **3.** The face cards that are hearts will be double-counted because F and H are not mutually exclusive. $P(F \text{ or } H) = \dfrac{12}{52} + \dfrac{13}{52} - \dfrac{3}{52} = \dfrac{22}{52} = 0.423$.

Answers to Odd-Numbered Section 5.2 Exercises
5.39 (a) (1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4), (4,1), (4,2), (4,3), (4,4). (b) Each outcome has probability $\dfrac{1}{16}$.

5.41 $P(A) = \dfrac{1}{16} + \dfrac{1}{16} + \dfrac{1}{16} + \dfrac{1}{16} = \dfrac{4}{16} = 0.25$.

5.43 (a) Legitimate. (b) Not legitimate: the total is more than 1. (c) Legitimate.

5.45 (a) $P(\text{type AB}) = 1 - 0.96 = 0.04$ (b) $P(\text{not type AB}) = 1 - P(\text{type AB}) = 1 - 0.04 = 0.96$ (c) $P(\text{type O or B}) = 0.49 + 0.20 = 0.69$

5.47 (a) $1 - 0.13 - 0.29 - 0.30 = 0.28$ (b) Using the complement rule, $1 - 0.13 = 0.87$.

5.49 (a) $P(\text{Female}) = \dfrac{275}{595} = 0.462$ (b) $P(\text{Eats breakfast regularly})$ $= \dfrac{300}{595} = 0.504$. (c) $P(\text{Female and breakfast}) = \dfrac{110}{595} = 0.185$.
(d) $P(\text{Female or breakfast}) = \dfrac{275}{595} + \dfrac{300}{595} - \dfrac{110}{595} = \dfrac{465}{595} = 0.782$.

5.51 (a)

	B	Not B	Total
E	10	10	20
Not E	8	10	18
Total	18	20	38

(b) $P(B) = \dfrac{18}{38} = 0.474$; $P(E) = \dfrac{20}{38} = 0.526$. (c) The ball lands in a spot that is black and even. $P(B \text{ and } E) = \dfrac{10}{38} = 0.263$. (d) If we add the probabilities of B and E, the spots that are black and even will be double-counted. $P(B \text{ or } E) = \dfrac{18}{38} + \dfrac{20}{38} - \dfrac{10}{38} = \dfrac{28}{38} = 0.737$.

5.53 (a)

(b) $P(B \cup M) = \dfrac{430}{595} = 0.723$. There is a 0.723 probability that we select a person who is a breakfast eater, a male, or both. (c) $P(B^C \cap M^C) = \dfrac{165}{595} = 0.277$. There is a 0.277 probability that we select a female who is not a breakfast eater.

5.55 (a)

	FB	Not FB	Total
YT	0.66	0.07	0.73
Not YT	0.19	0.08	0.27
Total	0.85	0.15	1

(b)

(c) $FB \cup YT$ (d) $P(FB \cup YT) = 0.85 + 0.73 - 0.66 = 0.92$.

5.57 c

5.59 c

5.61 The scatterplot for the average crawling age and average temperature is given below.

In this scatterplot, there appears to be a moderately strong, negative linear relationship between average temperature and average crawling age. The equation for the least-squares regression line is $\widehat{age}$ = $35.7 - 0.077$ (temp). We predict that babies will walk 0.077 weeks earlier for every degree warmer it gets.

Section 5.3

Answers to Check Your Understanding

page 321: **1.** $P(L) = \dfrac{3656}{10,000} = 0.3656$. There is a 0.3656 probability of selecting a course grade that is lower than a B. **2.** $P(E \mid L) = \dfrac{800}{3656} = 0.219$. $P(L \mid E) = \dfrac{800}{1600} = 0.50$. $P(L \mid E)$ gives the probability of getting a lower grade given that the student is studying engineering or physical science. Because this probability (0.50) is greater than $P(L) = 0.3656$, we can conclude that grades are lower in engineering and physical sciences.

page 326: **1.**

2. $P(\text{laptop}) = 0.30 + 0.175 + 0.175 = 0.65$.

3. $P(\text{made in CA} \mid \text{laptop}) = \dfrac{0.30}{0.65} = 0.462$.

page 328: **1.** Independent. Because we are replacing the cards, knowing what the first card was will not help us predict what the second card will be. **2.** Not independent. Once we know the suit of the first card, then the probability of getting a heart on the second card will change depending on what the first card was. **3.** Independent. $P(\text{right-handed}) = 24/28 = 6/7$ is the same as $P(\text{right-handed} \mid \text{female}) = 18/21 = 6/7$.

page 331: **1.** $P(\text{returned safely}) = 0.95$. So $P(\text{safe return on all 20 missions}) = 0.95^{20} = 0.3585$. **2.** No. Being a college student and being 55 or older are not independent events.

Answers to Odd-Numbered Section 5.3 Exercises

5.63 (a) $P(\text{almost certain}|M) = \dfrac{597}{2459} = 0.2428$.

(b) $P(F \mid \text{Some chance}) = \dfrac{426}{712} = 0.5983$.

5.65 (a) $P(D \mid F) = \dfrac{13}{17} = 0.7647$. Given that a senator is female, there is a 0.7647 probability that she is a Democrat.

(b) $P(F \mid D) = \dfrac{13}{60} = 0.2167$. Given that a senator is a Democrat, there is a 0.2167 probability that she is a female.

5.67 (a) $P(\text{not English}) = 1 - 0.59 = 0.41$.

(b) $P(\text{Spanish} \mid \text{other than English}) = \dfrac{0.26}{0.41} = 0.6341$.

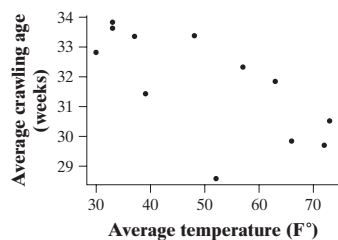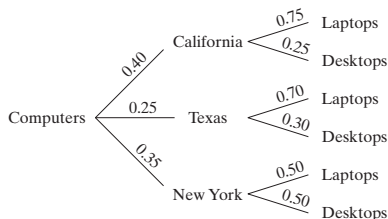**5.69** $P(B) < P(B \mid T) < P(T) < P(T \mid B)$. There are very few pro basketball players, so $P(B)$ should be smallest. If you are a pro basketball player, it is quite likely that you are tall, so $P(T \mid B)$ should be largest. Finally, it's much more likely to be over 6 feet tall than it is to be a pro basketball player if you're over 6 feet tall.

5.71 $P(\text{YT} \mid \text{FB}) = \dfrac{0.66}{0.85} = 0.7765$.

5.73 $P(\text{download music}) = 0.29$, $P(\text{don't care} \mid \text{download music}) = 0.67$.
$P(\text{download music} \cap \text{don't care}) = (0.29)(0.67) = 0.1943 = 19.43\%$.

5.75 (a) A tree diagram is below.

(b) $P(\text{one soft} \cap \text{one hard}) = \left(\frac{14}{20}\right)\left(\frac{6}{19}\right) + \left(\frac{6}{20}\right)\left(\frac{14}{19}\right) = \frac{168}{380} = 0.4421$

5.77 (a) A tree diagram is below.

(b) $P(\text{credit card}) = (0.88)(0.28) + (0.02)(0.34) + (0.10)(0.42) = 0.2952$ (c) $P(\text{premium gasoline}|\text{credit card}) = \dfrac{0.0420}{0.2952} = 0.142$.

5.79 (a) $P(\text{lactose intolerant}) = (0.82)(0.15) + (0.14)(0.70) + (0.04)(0.90) = 0.257$.

(b) $P(\text{Asian} \mid \text{lactose intolerant}) = \dfrac{0.036}{0.257} = 0.1401$.

5.81 $P(\text{antibody} \mid \text{positive}) =$
$$\dfrac{(0.01)(0.9985)}{(0.01)(0.9985) + (0.99)(0.006)} = 0.6270$$
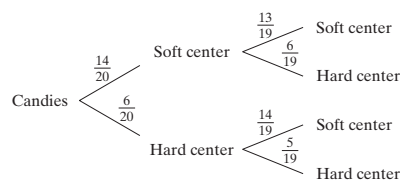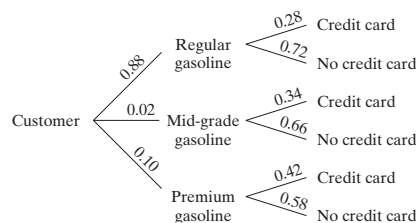
5.83 (a) $\dfrac{663}{2367} = 0.2801$ (b) $\dfrac{1421}{4826} = 0.2944$ (c) The events are not independent because the probabilities in parts (a) and (b) are not the same.

5.85 Not independent. From Exercise 5.65, we saw that $P(D \mid F) = 0.7647$, which is not the same as $P(D) = \dfrac{60}{100} = 0.60$.

5.87 Independent. $P(\text{sum of } 7 \mid \text{green is } 4) = 1/6 = 0.1667$, which equals $P(\text{sum of } 7) = 6/36 = 0.1667$.

5.89 $P(\text{all remain bright}) = (0.98)^{20} = 0.6676$

5.91 $P(\text{at least one universal donor}) = 1 - (0.928)^{10} = 0.5263$

5.93 No, because the events are not independent. If one show starts late, we can predict that the next show will start late as well.

5.95 (a) $P(\text{doubles}) = \dfrac{6}{36} = \dfrac{1}{6} = 0.167$

(b) $P(\text{no doubles first} \cap \text{doubles second}) = \dfrac{5}{6}\left(\dfrac{1}{6}\right) = \dfrac{5}{36} = 0.139$

(c) $P(\text{first doubles on third roll}) = \dfrac{5}{6}\left(\dfrac{5}{6}\right)\left(\dfrac{1}{6}\right) = \dfrac{25}{216} = 0.116$

(d) 4th: $\left(\dfrac{5}{6}\right)^3\left(\dfrac{1}{6}\right)$. 5th: $\left(\dfrac{5}{6}\right)^4\left(\dfrac{1}{6}\right)$. The probability that the first doubles are rolled on the kth roll is $\left(\dfrac{5}{6}\right)^{k-1}\left(\dfrac{1}{6}\right)$.

5.97 c

5.99 e

5.101 $P(\text{at least one is underweight}) = 1 - (1 - 0.131)^2 = 0.2448$

Answers to Chapter 5 Review Exercises

R5.1 When the weather conditions are like those seen today, it has rained on the following day about 30% of the time.

R5.2 (a) Let the numbers $00-14$ represent not wearing a seat belt and $15-99$ represent wearing a seat belt. Read 10 sets of two-digit numbers. For each set of 10 two-digit numbers, record whether there are two consecutive numbers between $00-14$ or not. (b) The first sample is 29 **07** 71 48 63 61 68 34 70 52 (not two consecutive). The second sample is 62 22 45 **10** 25 95 **05** 29 **09 08** (two consecutive). The third sample is 73 59 27 51 86 87 **13** 69 57 61 (not two consecutive).

R5.3 (a)

Difference	1	5	−3
Probability	$\dfrac{18}{36}$	$\dfrac{6}{36}$	$\dfrac{12}{36}$

(b) Die A. $P(A > B) = P(\text{positive difference}) = \dfrac{18}{36} + \dfrac{6}{36} = \dfrac{24}{36}$

R5.4 (a) Legitimate. All probabilities are between 0 and 1 and they add up to 1. (b) $P(\text{Hispanic}) = 0.001 + 0.006 + 0.139 + 0.003 = 0.149$ (c) $P(\text{not a non-Hispanic white}) = 1 - 0.674 = 0.326$ (d) People who are white and Hispanic will be double-counted. $P(\text{white or Hispanic}) = 0.813 + 0.149 - 0.139 = 0.823$.

R5.5 (a)
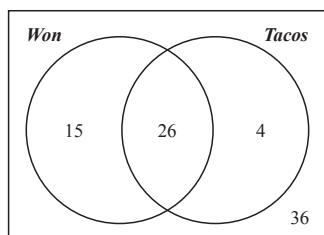
(b) $P(\text{lost and no tacos}) = 36/81 = 0.444$ (c) $P(\text{won or tacos}) = \dfrac{41}{81} + \dfrac{30}{81} - \dfrac{26}{81} = \dfrac{45}{81} = 0.556$.

R5.6 (a)
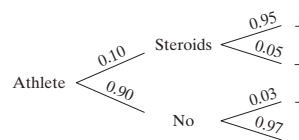
(b) $P(+) = (0.10)(0.95) + (0.90)(0.03) = 0.122$.

(c) $P(\text{steroids} \mid +) = \dfrac{0.095}{0.122} = 0.7787$.

R5.7 (a)

	Thick	Thin	Total
Mushrooms	2	2	4
No mushrooms	1	2	3
Total	3	4	7

(b) Not independent: $P(\text{mushrooms}) = \dfrac{4}{7} = 0.571$ does not equal $P(\text{mushrooms} \mid \text{thick crust}) = \dfrac{2}{3} = 0.667$.

(c) Independent: $P(\text{mushrooms}) = \dfrac{4}{8} = \dfrac{1}{2} = 0.50$ is equal to $P(\text{mushrooms} \mid \text{thick crust}) = \dfrac{2}{4} = \dfrac{1}{2} = 0.50$.

R5.8 (a) $P(\text{damage}) = \dfrac{209}{871} = 0.24$.

(b) $P(\text{damage} \mid \text{no cover}) = \dfrac{60}{211} = 0.2844$,

$P\left(\text{damage} \mid < \dfrac{1}{3}\right) = \dfrac{76}{234} = 0.3248$,

$P\left(\text{damage} \mid \dfrac{1}{3} \text{ to } \dfrac{2}{3}\right) = \dfrac{44}{221} = 0.1991$, and

$P\left(\text{damage} \mid > \dfrac{2}{3}\right) = \dfrac{29}{205} = 0.1415$. (c) Yes. It appears that deer do much more damage when there is no cover or less than 1/3 cover than when there is more cover.

R5.9 (a) $P(\text{up three consecutive years}) = (0.65)^3 = 0.274625$.

(b) $P(\text{same direction for 3 years}) = (0.65)^3 + (0.35)^3 = 0.3175$.

R5.10 (a) (A, A), (A, B), (B, A), (B, B)

Blood type	A	AB	B
Probability	0.25	0.5	0.25

(b) (A, A), (A, B), (O, A), (O, B)

Blood type	A	AB	B
Probability	0.5	0.25	0.25

$P(\text{at least 1 type B}) = 1 - P(\text{neither are type B}) = 1 - (0.75)^2 = 0.4375$.

Answers to Chapter 5 AP® Statistics Practice Test

T5.1 e
T5.2 d
T5.3 c
T5.4 b
T5.5 b
T5.6 c
T5.7 e
T5.8 e
T5.9 b
T5.10 c

T5.11 (a) Here is a completed table, with T indicating that the teacher wins and Y indicating that you win. $P(\text{teacher wins}) = \dfrac{27}{48} = 0.5625$.

	1	2	3	4	5	6	7	8
1	—	T	T	T	T	T	T	T
2	Y	—	T	T	T	T	T	T
3	Y	Y	—	T	T	T	T	T
4	Y	Y	Y	—	T	T	T	T
5	Y	Y	Y	Y	—	T	T	T
6	Y	Y	Y	Y	Y	—	T	T

(b) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \dfrac{27}{48} + \dfrac{8}{48} - \dfrac{5}{48} = \dfrac{30}{48}$.

(c) Not independent. $P(A) = \dfrac{27}{48} = 0.5625$ does not equal $P(A \mid B) = \dfrac{5}{8} = 0.625$.

T5.12 (a)

Part
- 0.60 → A: 0.10 Defective, 0.90 OK
- 0.30 → B: 0.30 Defective, 0.70 OK
- 0.10 → C: 0.40 Defective, 0.60 OK
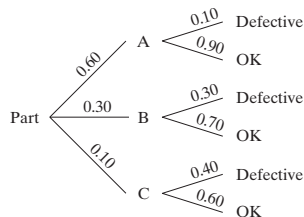
(b) $P(\text{defective}) = (0.60)(0.10) + (0.30)(0.30) + (0.10)(0.40) = 0.19$. (c) Machine B. $P(A \mid \text{defective}) = \dfrac{0.06}{0.19} = 0.3158$.

$P(B \mid \text{defective}) = \dfrac{0.09}{0.19} = 0.4737$. $P(C \mid \text{defective}) = \dfrac{0.04}{0.19} = 0.2105$.

T5.13 (a) Here is a two-way table that summarizes this information:

	Smokes	Does not smoke	Total
Cancer	0.08	0.04	0.12
No cancer	0.17	0.71	0.88
Total	0.25	0.75	1.00

$P(\text{gets cancer} \mid \text{smoker}) = \dfrac{0.08}{0.25} = 0.32$.

(b) $P(\text{smokes} \cup \text{gets cancer}) = 0.25 + 0.12 - 0.08 = 0.29$.

(c) $P(\text{cancer}) = 0.12$, so $P(\text{at least one gets cancer}) = 1 - P(\text{neither gets cancer}) = 1 - 0.88^2 = 0.2256$

T5.14 (a) Let $00-16$ represent out-of-state and $17-99$ represent in-state. Read two-digit numbers until you have found two numbers between 00 and 16. Record how many 2-digit numbers you had to read. (b) The first sample is 41 **05 09** (it took three cars). The second sample is 20 31 **06** 44 90 50 59 59 88 43 18 80 53 **11** (it took 14 cars). The third sample is 58 44 69 94 86 85 79 67 **05** 81 18 45 **14** (it took 13 cars).

Chapter 6
Section 6.1

Answers to Check Your Understanding
page 350: 1. $P(X \geq 3)$ is the probability that the student got either an A or a B. $P(X \geq 3) = 0.68$.
2. $P(X < 2) = 0.02 + 0.10 = 0.12$
3. The histogram below is skewed to the left. Higher grades are more likely, but there are a few lower grades.
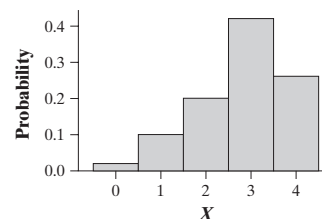
page 355: 1. $\mu_X = 1.1$. If many, many Fridays are randomly selected, the average number of cars sold will be about 1.1.
2. $\sigma_X = \sqrt{0.89} = 0.943$. The number of cars sold on a randomly selected Friday will typically vary from the mean (1.1) by about 0.943 cars.

Answers to Odd-Numbered Section 6.1 Exercises
6.1 (a)

Value	0	1	2	3	4
Probability	1/16	4/16	6/16	4/16	1/16

(b) The histogram below shows that this distribution is symmetric with a center at 2.
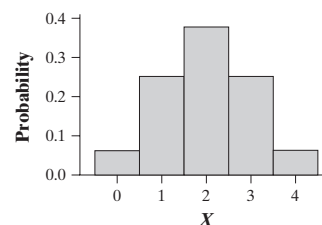
(c) $P(X \leq 3) = 15/16 = 0.9375$. There is a 0.9375 probability that you will get three or fewer heads in 4 tosses of a fair coin.
6.3 (a) $P(X \geq 1) = 0.9$. (b) The event $X \leq 2$ is "at most two non-word errors." $P(X \leq 2) = 0.6$. $P(X < 2) = 0.3$.
6.5 (a) All of the probabilities are between 0 and 1 and they sum to 1. (b) The histogram below is unimodal and skewed to the right.

Number of rooms in renter-occupied units
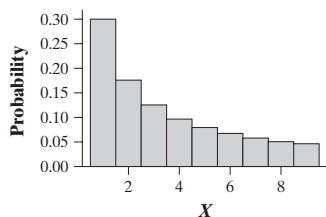
(c) The event $X \geq 6$ is the event that "the first digit in a randomly chosen record is a 6 or higher." $P(X \geq 6) = 0.222$. (d) $P(X \leq 5) = 0.778$.

6.7 (a) The outcomes that make up the event A are 7, 8, and 9. $P(A) = 0.155$. (b) The outcomes that make up the event B are 1, 3, 5, 7, and 9. $P(B) = 0.609$. (c) The outcomes that make up the event "A or B" are 1, 3, 5, 7, 8, and 9. $P(A \text{ or } B) = 0.660$. This is not the same as $P(A) + P(B)$ because the outcomes 7 and 9 are included in both events.

6.9 (a)

X	$-\$1$	$\$2$
Probability	0.75	0.25

(b) $E(X) = -\$0.25$. If the player makes many $1 bets, he will lose about $0.25 per $1 bet, on average.

6.11 $\mu_X = 2.1$. If many, many undergraduates performed this task, they would make about 2.1 nonword errors, on average.

6.13 (a) This distribution is symmetric and 5 is located at the center. (b) According to Benford's law, $E(X) = 3.441$. To detect a fake expense report, compute the sample mean of the first digits. A value closer to 5 suggests a fake report and a value near 3.441 is consistent with a truthful report. (c) $P(Y > 6) = 3/9 = 0.333$. Under Benford's law, $P(X > 6) = 0.155$. To detect a fake expense report, compute the proportion of first digits that begin with 7, 8, or 9. A value closer to 0.333 suggests a fake report and a value closer to 0.155 is consistent with a truthful report.

6.15 $\sigma_X = \sqrt{1.29} = 1.1358$. The number of nonword errors in a randomly selected essay will typically differ from the mean (2.1) by about 1.14 words.

6.17 (a) $\sigma_Y = \sqrt{6.667} = 2.58$. (b) $\sigma_X = \sqrt{6.0605} = 2.4618$. This would not be the best way to tell the difference between a fake and a real expense report because the standard deviations are similar.

6.19 (a) See the following histograms. The distribution of the number of rooms is roughly symmetric for owners and skewed to the right for renters. Renter-occupied units tend to have fewer rooms than owner-occupied units. There is more variability in the number of rooms for owner-occupied units.

Number of rooms in owner-occupied units
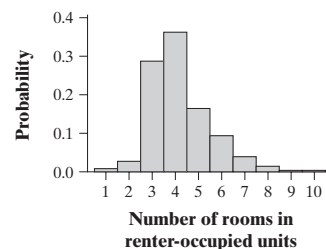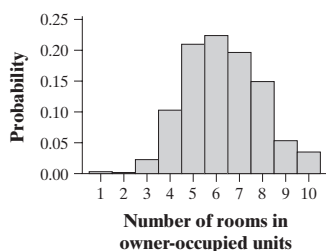
(b) Owner: $\mu_X = 6.284$ rooms. Renter: $\mu_Y = 4.187$ rooms. Single people and younger people are more likely to rent and need less space than people with families. (c) $\sigma_X = \sqrt{2.68934} = 1.6399$. The number of rooms in a randomly selected owner-occupied unit will typically differ from the mean (6.284) by about 1.6399 rooms. $\sigma_Y = \sqrt{1.71003} = 1.3077$. The number of rooms in a randomly selected renter-occupied unit will typically differ from the mean (4.187) by about 1.3077 rooms.

6.21 (a) $P(X > 0.49) = 0.51$. (b) $P(X \geq 0.49) = 0.51$. (c) $P(0.19 \leq X < 0.37 \text{ or } 0.84 < X \leq 1.27) = 0.18 + 0.16 = 0.34$

6.23 The time Y of a randomly chosen student has the $N(7.11, 0.74)$ distribution. We want to find $P(Y < 6)$. $z = \dfrac{6 - 7.11}{0.74} = -1.50$ and $P(Z > -1.50) = 0.0668$. *Using technology*: `normalcdf(lower: -1000,upper:6,μ:7.11,σ:0.74)` = 0.0668. There is about a 7% chance that this student will run the mile in under 6 minutes.

6.25 (a) The speed Y of a randomly chosen serve has the $N(115, 6)$ distribution. We want to find $P(Y > 120)$. $z = \dfrac{120 - 115}{6} = 0.83$ and $P(Z > 0.83) = 0.2033$. *Using technology*: `normalcdf (lower:120,upper:1000,μ:115,σ:6)` = 0.2023. There is a 0.2023 probability of selecting a serve that is greater than 120 mph. (b) The line above 120 has no area, so $P(Y \geq 120) = P(Y > 120) = 0.2023$. (c) We want to find c such that $P(Y \leq c) = 0.15$. Solving $-1.04 = \dfrac{c - 115}{6}$ gives $c = 108.76$. *Using technology*: `invNorm(area:0.15,μ:115,σ:6)` = 108.78. Fifteen percent of Nadal's serves will be less than or equal to 108.78 mph.

6.27 b

6.29 c

6.31 Yes. The mean difference (post $-$ pre) was 5.38 and the median difference was 3. This means that at least half of the students improved their reading scores.

6.33 predicted post-test = $17.897 + 0.78301(\text{pretest})$.

Section 6.2

Answers to Check Your Understanding

page 367:
1. $Y = 500X$. $\mu_Y = 500(1.1) = \$550$. $\sigma_Y = 500(0.943) = \$471.50$.
2. $T = Y - 75$. $\mu_T = 550 - 75 = \$475$. $\sigma_T = \$471.50$.

page 376: 1. $\mu_T = 1.1 + 0.7 = 1.8$. Over many Fridays, this dealership sells or leases about 1.8 cars in the first hour of business, on average.
2. $\sigma_T^2 = (0.943)^2 + (0.64)^2 = 1.2988$, so $\sigma_T = \sqrt{1.2988} = 1.14$.
3. $\mu_B = 500(1.1) + 300(0.7) = \$760$. $\sigma_B^2 = (500)^2(0.943)^2 + (300)^2(0.64)^2 = 259,176.25$, so $\sigma_B = \sqrt{259,176.25} = \$509.09$.

page 378: 1. $\mu_D = 1.1 - 0.7 = 0.4$. Over many Fridays, this dealership sells about 0.4 cars more than it leases during the first hour of business, on average.

2. $\sigma_D^2 = (0.943)^2 + (0.64)^2 = 1.2998$, so $\sigma_D = \sqrt{1.2998} = 1.14$.
3. $\mu_B = 500(1.1) - 300(0.7) = \$340. \ \sigma_B^2 = (500)^2(0.943)^2 + (300)^2(0.64)^2 = 259,176.25$, so $\sigma_B = \sqrt{259,176.25} = \$509.09$.

Answers to Odd-Numbered Section 6.2 Exercises

6.35 $\mu_Y = 2.54(1.2) = 3.048$ cm and $\sigma_Y = 2.54(0.25) = 0.635$ cm.
6.37 **(a)** The distribution shown below is skewed to the left. Most of the time, the ferry makes \$20 or \$25.
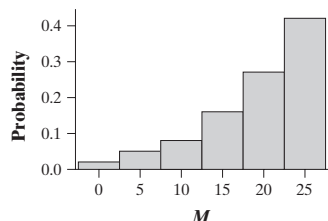
(b) $\mu_M = \$19.35$. If many ferry trips were selected at random, the ferry would collect about \$19.35 per trip, on average. **(c)** $\sigma_M = \$6.45$. The amounts collected on randomly selected ferry trips will typically vary by about \$6.45 from the mean (\$19.35).
6.39 **(a)** $\mu_G = 5(7.6) + 50 = 88.$ **(b)** $\sigma_G = 5(1.32) = 6.6.$ **(c)** $\sigma_G^2 = (5\sigma_X)^2 = 25\sigma_X^2$. The variance of G is 25 times the variance of X.
6.41 **(a)** $\mu_Y = -\$0.65$. If many ferry trips were selected at random, the ferry would lose about \$0.65 per trip, on average. **(b)** $\sigma_Y = \$6.45$. The amount of profit on randomly selected ferry trips will typically vary by about \$6.45 from the mean $(-\$0.65)$.
6.43 $\mu_Y = 6(3.87) - 20 = \$3.22. \ \sigma_Y = 6(1.29) = \$7.74$.
6.45 **(a)** $\mu_Y = 47.3°F. \ \sigma_Y = 4.05 °F$. **(b)** Y has the N(47.3, 4.05) distribution. We want to find $P(Y < 40)$. $z = \dfrac{40 - 47.3}{4.05} = -1.80$ and $P(Z < -1.80) = 0.0359$. *Using technology:* `normalcdf(lower:−1000,upper:40,µ:47.3,σ:4.05) = 0.0357`. There is a 0.0357 probability that the midnight temperature in the cabin is below 40°F.
6.47 **(a)** Yes. The mean of a sum is always equal to the sum of the means. **(b)** No, because it is not reasonable to assume that X and Y are independent.
6.49 $\mu_{Y_1+Y_2} = (-0.65) + (-0.65) = -\$1.30. \ \sigma^2_{Y_1+Y_2} = 6.45^2 + 6.45^2 = 83.205$, so $\sigma_{Y_1+Y_2} = \sqrt{83.205} = \$9.12$.
6.51 $\mu_{3X} = 3(2.1) = 6.3$ and $\sigma_{3X} = 3(1.136) = 3.408. \ \mu_{2Y} = 2(1.0) = 2.0$ and $\sigma_{2Y} = 2(1.0) = 2.0$. Thus, $\mu_{3X+2Y} = 6.3 + 2.0 = 8.3$ and $\sigma^2_{3X+2Y} = 3.408^2 + 2.0^2 = 15.6145$, so $\sigma_{3X+2Y} = \sqrt{15.6145} = 3.95$.
6.53 **(a)** $\mu_{Y-X} = 1.0 - 2.1 = -1.1$. If you were to select many essays, there would be about 1.1 fewer word errors than nonword errors, on average. $\sigma^2_{Y-X} = (1.0)^2 + (1.136)^2 = 2.2905$, so $\sigma_{Y-X} = \sqrt{2.2905} = 1.51$. The difference in the number errors will typically vary by about 1.51 from the mean (-1.1). **(b)** The outcomes that make up this event are $1 - 0 = 1, \ 2 - 0 = 2, 2 - 1 = 1, 3 - 0 = 3, 3 - 1 = 2, 3 - 2 = 1$. There is a 0.15 probability that a randomly chosen student will have more word errors than nonword errors.
6.55 The difference in score deductions for a randomly selected essay is $3X - 2Y. \ \mu_{3X} = 3(2.1) = 6.3$ and $\sigma_{3X} = 3(1.136) = 3.408. \ \mu_{2Y} = 2(1.0) = 2.0$ and $\sigma_{2Y} = 2(1.0) = 2.0$. Thus, $\mu_{3X-2Y} = 6.3 - 2.0 = 4.3$ and $\sigma^2_{3X-2Y} = 3.408^2 + 2.0^2 = 15.6145$, so $\sigma_{3X-2Y} = \sqrt{15.6145} = 3.95$.

6.57 $\mu_{X_1+X_2} = 303.35 + 303.35 = \$606.70$ and $\sigma^2_{X_1+X_2} = 9707.57^2 + 9707.57^2 = 188,473,830.6$, so $\sigma_{X_1+X_2} = \sqrt{188,473,830.6} = \$13,728.58. \ W = \dfrac{1}{2}(X_1 + X_2)$, so $\mu_W = \dfrac{1}{2}(606.70) = \$303.35$ and $\sigma_W = \dfrac{1}{2}(13,728.58) = \$6864.29$.
6.59 **(a)** Normal with mean $= 11 + 20 = 31$ seconds and standard deviation $= \sqrt{2^2 + 4^2} = 4.4721$ seconds. **(b)** We want to find the probability that the total time is less than 30 seconds. $z = \dfrac{30 - 31}{4.4721} = -0.22$ and $P(Z < -0.22) = 0.4129$. *Using technology:* `normalcdf(lower:−1000,upper:30,µ:31,σ: 4.4721) = 0.4115`. There is a 0.4115 probability of completing the process in less than 30 seconds for a randomly selected part.
6.61 Let $T = $ the total team swim time. $\mu_T = 55.2 + 58.0 + 56.3 + 54.7 = 224.2$ seconds and $\sigma_T^2 = (2.8)^2 + (3.0)^2 + (2.6)^2 + (2.7)^2 = 30.89$, so $\sigma_T = \sqrt{30.89} = 5.56$ seconds. Thus, T has the N(224.2, 5.56) distribution. We want to find $P(T < 220)$. $z = \dfrac{220 - 224.2}{5.56} = -0.76$ and $P(Z < -0.76) = 0.2236$. *Using technology:* `normalcdf(lower:−1000,upper:220,µ: 224.2,σ:5.56) = 0.2250`. There is a 0.2250 probability that the total team time is less than 220 seconds in a randomly selected race.
6.63 Let $D = X_1 - X_2 = $ the difference in NOX levels. $\mu_D = 1.4 - 1.4 = 0$ and $\sigma^2_{X_1-X_2} = \sigma^2_{X_1} + \sigma^2_{X_2} = 0.3^2 + 0.3^2 = 0.18$, so $\sigma_{X_1-X_2} = \sqrt{0.18} = 0.4243$. Thus, D has the N(0, 0.4243) distribution. We want to find $P(D > 0.8 \text{ or } D < -0.8) = P(D > 0.8) + P(D < -0.8)$. $z = \dfrac{0.8 - 0}{0.4243} = 1.89$ and $z = \dfrac{-0.8 - 0}{0.4243} = -1.89$ and $P(Z < -1.89 \text{ or } Z > 1.89) = 0.0588$. *Using technology:* `1 − normalcdf(lower:−0.8,upper:0.8,µ:0,σ: 0.4243) = 0.0594`. There is a 0.0594 probability that the difference is at least as large as the attendant observed.
6.65 c
6.67 **(a)** Fidelity Technology Fund, because its correlation is larger. **(b)** No, the correlation doesn't tell us anything about the values of the variables, only about the strength of the linear relationship between them.

Section 6.3
Answers to Check Your Understanding

page 389: **1.** Binomial. Binary? "Success" = get an ace. "Failure" = don't get an ace. Independent? Because you are replacing the card in the deck and shuffling each time, the result of one trial does not tell you anything about the outcome of any other trial. Number? $n = 10$. Success? The probability of success is $p = 4/52$ for each trial. **2.** Not binomial. Binary? "Success" = over 6 feet. "Failure" = not over 6 feet. Independent? Because we are selecting without replacement from a small number of students, the observations are not independent. Number? $n = 3$. Success? The probability of success will not change from trial to trial. **3.** Not binomial. Binary? "Success" = roll a 5. "Failure" = don't roll a 5. Independent? Because you are rolling a die, the outcome of any one trial does not tell you anything about the outcome of any other trial. Number? $n = 100$. Success? No. The probability of success changes when the corner of the die is chipped off.
page 397: **1.** Binary? "Success" = question answered correctly. "Failure" = question not answered correctly. Independent? The computer randomly assigned correct answers to the questions, so

knowing the result of one trial (question) should not tell you anything about the result on any other trial. Number? $n = 10$. Success? The probability of success is $p = 0.20$ for each trial.

2. $P(X = 3) = \binom{10}{3}(0.2)^3(0.8)^7 = 0.2013$. There is a 20% chance that Patti will answer exactly 3 questions correctly.

3. $P(X \geq 6) = 1 - P(X \leq 5) = 1 - 0.9936 = 0.0064$. There is only a 0.0064 probability that a student would get 6 or more correct, so we would be quite surprised if Patti was able to pass.

page 400: **1.** $\mu_X = 10(0.20) = 2$. If many students took the quiz, we would expect students to get about 2 answers correct, on average. **2.** $\sigma_X = \sqrt{10(0.20)(0.80)} = 1.265$. If many students took the quiz, we would expect individual students' scores to typically vary from the mean of 2 correct answers by about 1.265 correct answers. **3.** $P(X > 2 + 2(1.265)) = P(X > 4.53) = 1 - P(X \leq 4) = 1 - 0.9672 = 0.0328$.

page 408: **1.** Die rolls are independent, the probability of getting doubles is the same on each roll (1/6), and we are repeating the chance process until we get a success (doubles).

2. $P(T = 3) = \left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right) = 0.1157$. There is a 0.1157 probability that you will get the first set of doubles on the third roll of the dice. **3.** $P(T \leq 3) = \frac{1}{6} + \left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right) = 0.4213$.

Answers to Odd-Numbered Section 6.3 Exercises

6.69 Binomial. Binary? "Success" = seed germinates and "Failure" = seed does not germinate. Independent? Yes, because the seeds were randomly selected, knowing the outcome of one seed shouldn't tell us anything about the outcomes of other seeds. Number? $n = 20$ seeds. Success? $p = 0.85$.

6.71 Not binomial. Binary? "Success" = person is left-handed and "Failure" = person is right-handed. Independent? Because students are selected randomly, their handedness is independent. Number? There is not a fixed number of trials for this chance process because you continue until you find a left-handed student. Success? $p = 0.10$.

6.73 **(a)** Binomial. Binary? "Success" = reaching a live person and "Failure" = any other outcome. Independent? Knowing whether or not one call was completed tells us nothing about the outcome on any other call. Number? $n = 15$. Success? $p = 0.2$. **(b)** This is not a binomial setting because there are not a fixed number of attempts. The Binary, Independent, and Success conditions are satisfied, however, as in part (a).

6.75 $P(X = 4) = \binom{7}{4}(0.44)^4(0.56)^3 = 0.2304$. There is a 0.2304 probability that exactly 4 of the 7 elk survive to adulthood.

6.77 $P(X > 4) = \binom{7}{5}(0.44)^5(0.56)^2 + \cdots = 0.1402$. Because this probability isn't very small, it is not surprising for more than 4 elk to survive to adulthood.

6.79 **(a)** $P(X = 17) = \binom{20}{17}(0.85)^{17}(0.15)^3 = 0.2428$.

(b) $P(X \leq 12) = \binom{20}{0}(0.85)^0(0.15)^{20} + \cdots + \binom{20}{12}(0.85)^{12}(0.15)^8 = 0.0059$. Because this is such a low probability, Judy should be suspicious.

6.81 **(a)** $\mu_X = 15(0.20) = 3$. If we watched the machine make many sets of 15 calls, we would expect about 3 calls to reach a live person, on average. **(b)** $\sigma_X = \sqrt{15(0.20)(0.80)} = 1.55$. If we watched the machine make many sets of 15 calls, we would expect the number of calls that reach a live person to typically vary by about 1.55 from the mean (3).

6.83 **(a)** $\mu_Y = 15(0.80) = 12$. Notice that $\mu_X = 3$ and $12 + 3 = 15$ (the total number of calls). **(b)** $\sigma_Y = \sqrt{15(0.80)(0.20)} = 1.55$. This is the same value as σ_X, because $Y = 15 - X$ and adding a constant to a random variable doesn't change the spread.

6.85 **(a)** Binary? "Success" = win a prize and "Failure" = don't win a prize. Independent? Knowing whether one bottle wins or not should not tell us anything about the caps on other bottles. Number? $n = 7$. Success? $p = 1/16$. **(b)** $\mu_X = 1.167$. If we were to buy many sets of 7 bottles, we would get 1.167 winners per set, on average. $\sigma_X = 0.986$. If we were to buy many sets of 7 bottles, the number of winning bottles would typically differ from the mean (1.167) by 0.986. **(c)** $P(X \geq 3) = 1 - P(X \leq 2) = 0.0958$. Because 0.0958 isn't a very small probability, the clerk shouldn't be surprised. It is plausible to get 3 or more winners in a sample of 7 bottles by chance alone.

6.87 No. Because we are sampling without replacement and the sample size (10) is more than 10% of the population size (76), we should not treat the observations as independent.

6.89 If the sample is a small fraction of the population (less than 10%), the make-up of the population doesn't change enough to make the lack of independent trials an issue.

6.91 **(a)** Binary? "Success" = visit an auction site at least once a month and "Failure" = don't visit an auction site at least once a month. Independent? We are sampling without replacement, but the sample size (500) is far less than 10% of all males aged 18 to 34. Number? $n = 500$. Success? $p = 0.50$. **(b)** $np = 250$ and $n(1 - p) = 250$ are both at least 10. **(c)** $\mu_X = 250$ and $\sigma_X = 11.18$. Thus, X has approximately the $N(250, 11.18)$ distribution. We want to find $P(X \geq 235)$. $z = \dfrac{235 - 250}{11.18} = -1.34$ and $P(Z \geq -1.34) = 0.9099$ *Using technology:* `normalcdf(lower:235, upper:1000,μ:250,σ:11.18) = 0.9102`. There is a 0.9102 probability that at least 235 of the men in the sample visit an online auction site.

6.93 Let X be the number of 1s and 2s. Then X has a binomial distribution with $n = 90$ and $p = 0.477$ (in the absence of fraud). $P(X \leq 29) = 0.0021$. Because the probability of getting 29 or fewer invoices that begin with the digits 1 or 2 is quite small, we have reason to be suspicious that the invoice amounts are not genuine.

6.95 **(a)** Not geometric. We can't classify the possible outcomes on each trial (card) as "success" or "failure" and we are not selecting cards until we get a *single* success. **(b)** Games of 4-Spot Keno are independent, the probability of winning is the same in each game ($p = 0.259$), and Lola is repeating a chance process until she gets a success. $X =$ number of games needed to win once is a geometric random variable with $p = 0.259$.

6.97 **(a)** Let $X =$ the number of bottles Alan purchases to find one winner. $P(X = 5) = (5/6)^4(1/6) = 0.0804$.
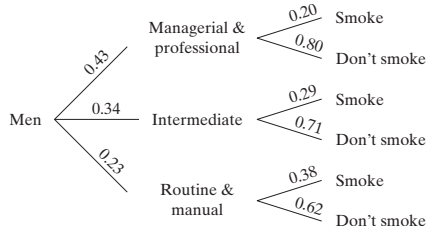**(b)** $P(X \leq 8) = (1/6) + \cdots + (5/6)^7(1/6) = 0.7674$.

6.99 **(a)** $\mu_X = \dfrac{1}{0.097} = 10.31$.
(b) $P(X \geq 40) = 1 - P(X \leq 39) = 0.0187$. Because the probability of not getting an 8 or 9 before the 40th invoice is small, we may begin to worry that the invoice amounts are fraudulent.

6.101 b

6.103 d
6.105 c
6.107 (a)

$P(\text{smoke}) = 0.43(0.20) + 0.34(0.29) + 0.23(0.38) = 0.272 = 27.2\%$

(b) $P(\text{routine and manual} \mid \text{smoke}) = \dfrac{(0.23)(0.38)}{0.272} = 0.321 = 32.1\%$

Answers to Chapter 6 Review Exercises

R6.1 (a) $P(X = 5) = 1 - 0.1 - 0.2 - 0.3 - 0.3 = 0.1$.
(b) Discrete, because it takes a fixed set of values with gaps in between. (c) $P(X \le 2) = 0.3$. $P(X < 2) = 0.1$. These are not the same because the outcome $X = 2$ is included in the first calculation but not the second. (d) $\mu_X = 1(0.1) + \cdots + 5(0.1) = 3.1$. $\sigma_X^2 = (1 - 3.1)^2(0.1) + \cdots + (5 - 3.1)^2(0.1) = 1.29$, so $\sigma_X = \sqrt{1.29} = 1.136$.
R6.2 (a) Temperature is a continuous random variable because it takes all values in an interval of numbers—there are no gaps between possible temperatures. (b) $P(X < 540) = P(X \le 540)$ because X is a continuous random variable. In this case, $P(X = 540) = 0$ because the line segment above $X = 540$ has no area. (c) Mean $= 550 - 550 = 0°C$. The standard deviation stays the same, $5.7°C$, because subtracting a constant does not change the variability. (d) In degrees Fahrenheit, the mean is $\mu_Y = \frac{9}{5}(550) + 32 = 1022°F$ and the standard deviation is $\sigma_Y = \left(\frac{9}{5}\right)(5.7) = 10.26°F$.
R6.3 (a) If you were to play many games of 4-Spot Keno, you would get a payout of about $0.70 per game, on average. If you were to play many games of 4-Spot Keno, the payout amounts would typically vary by about $6.58 from the mean ($0.70). (b) Let Y be the amount of Jerry's payout. $\mu_Y = 5(0.70) = \$3.50$ and $\sigma_Y = 5(6.58) = \$32.90$. (c) Let W be the amount of Marla's payout. $\mu_W = 0.70 + 0.70 + 0.70 + 0.70 + 0.70 = \$3.50$ and $\sigma_W^2 = 6.58^2 + 6.58^2 + 6.58^2 + 6.58^2 + 6.58^2 = 216.482$, so $\sigma_W = \sqrt{216.482} = \$14.71$.
(d) Even though their expected values are the same, the casino would probably prefer Marla since there is less variability in her strategy and her winnings are more predictable.
R6.4 (a) C follows a $N(10, 1.2)$ distribution and we want to find $P(C > 11)$. $z = \dfrac{11 - 10}{1.2} = 0.83$ and $P(Z > 0.83) = 0.2033$. *Using technology*: `normalcdf(lower:11,upper:1000, μ:10,σ:1.2) = 0.2023`. There is a 0.2023 probability that a randomly selected cap has a strength greater than 11 inch-pounds. (b) The machine that makes the caps and the machine that applies the torque are not the same. (c) $C - T$ is Normal with mean $10 - 7 = 3$ inch-pounds and standard deviation $\sqrt{0.9^2 + 1.2^2} = 1.5$ inch-pounds. (d) We want to find

$P(C - T < 0)$. $z = \dfrac{0 - 3}{1.5} = -2$ and $P(Z < -2) = 0.0228$. *Using technology*: `normalcdf(lower:−1000,upper:0, μ:3,σ:1.5) = 0.0228`. There is a 0.0228 probability that a randomly selected cap will break when being fastened by the machine.

R6.5 (a) Binary? "Success" = orange and "Failure" = not orange. Independent? The sample of size $n = 8$ is less than 10% of the large bag, so we can assume the outcomes of trials are independent. Number? $n = 8$. Success? $p = 0.20$. (b) $\mu_X = 8(0.2) = 1.6$. If we were to select many samples of size 8, we would expect to get about 1.6 orange M&M'S, on average. (c) $\sigma_X = \sqrt{8(0.2)(0.8)} = 1.13$. If we were to select many samples of size 8, the number of orange M&M'S would typically vary by about 1.13 from the mean (1.6).
R6.6 (a) $P(X = 0) = \binom{8}{0}(0.2)^0(0.8)^8 = 0.1678$. Because the probability is not that small, it would not be surprising to get no orange M&M'S in a sample of size 8. (b) $P(X \ge 5) = \binom{8}{5}(0.2)^5(0.80)^3 + \cdots = 0.0104$ Because the probability is small, it would be surprising to find 5 or more orange M&M'S in a sample of size 8.
R6.7 Let Y be the number of spins to get a "wasabi bomb." Y is a geometric random variable with $p = \dfrac{3}{12} = 0.25$. $P(Y \le 3) = (0.75)^2(0.25) + (0.75)(0.25) + 0.25 = 0.5781$.
R6.8 (a) Let X be the number of heads in 10,000 tosses. $\mu_X = 10{,}000(0.5) = 5{,}000$ and $\sigma_X = \sqrt{10{,}000(0.5)(0.5)} = 50$. (b) $np = 10{,}000(0.5) = 5{,}000$ and $n(1 - p) = 10{,}000(0.5) = 5000$ are both at least 10. (c) We want to find $P(X \le 4933 \text{ or } X \ge 5067)$. $z = \dfrac{4933 - 5000}{50} = -1.34$ and $z = \dfrac{5067 - 5000}{50} = 1.34$ and $P(Z \le -1.34) + P(Z \ge 1.34) = 0.1802$. *Using technology*: $1 -$ `normalcdf(lower:4933,upper:5067,μ:5000, σ:50) = 0.1802`. Because this probability isn't small, we don't have convincing evidence that Kerrich's coin was unbalanced—a difference this far from 5000 could be due to chance alone.

Answers to Chapter 6 AP® Statistics Practice Test

T6.1 b
T6.2 d
T6.3 d
T6.4 e
T6.5 d
T6.6 b
T6.7 c
T6.8 b
T6.9 b
T6.10 c
T6.11 (a) $P(Y \le 2) = 0.96$. (b) $\mu_Y = 0(0.78) + \cdots = 0.38$. If we were to randomly select many cartons of eggs, we would expect about 0.38 to be broken, on average. (c) $\sigma_Y^2 = (0 - 0.38)^2(0.78) + \ldots = 0.6756$. So $\sigma_Y = \sqrt{0.6756} = 0.8219$. If we were to randomly select many cartons of eggs, the number of broken eggs would typically vary by about 0.6756 from the mean (0.38). (d) Let X stand for the number of cartons inspected to find one carton with at least 2 broken eggs. X is a geometric random variable with $p = 0.11$. $P(X \le 3) = (0.11) + (0.89)(0.11) + (0.89)^2(0.11) = 0.2950$.

T6.12 (a) Binary? "Success" = dog first and "Failure" = not dog first. Independent? We are sampling without replacement, but 12 is less than 10% of all dog owners. Number? $n = 12$. Success? $p = 0.66$. (b) $P(X \le 4) = \binom{12}{0}(0.66)^0(0.34)^{12} + \cdots + \binom{12}{4}(0.66)^4(0.34)^8 = 0.0213$. Because this probability is small, it is unlikely to have only 4 or fewer owners greet their dogs first by chance alone. This gives convincing evidence that the claim by the *Ladies Home Journal* is incorrect.

T6.13 (a) $\mu_D = 50 - 25 = 25$ minutes, $\sigma_D^2 = 100 + 25 = 125$, and $\sigma_D = \sqrt{125} = 11.18$ minutes. (b) D follows a $N(25, 11.18)$ distribution and we want to find $P(D < 0)$. $z = \dfrac{0 - 25}{11.18} = -2.24$ and $P(Z < -2.24) = 0.0125$. *Using technology:* `normalcdf (lower:-1000,upper:0,µ:25,σ:11.18) = 0.0127`. There is a 0.0127 probability that Ed spent longer on his assignment than Adelaide did on hers.

T6.14 (a) Let X stand for the number of Hispanics in the sample. $\mu_X = 1200(0.13) = 156$ and $\sigma_X = \sqrt{1200(0.13)(0.87)} = 11.6499$. (b) 15% of 1200 is 180, so we want to find $P(X \ge 180) = \binom{1200}{180}(0.13)^{180}(0.87)^{1020} + \cdots = 0.0235$. Because this probability is small, it is unlikely to select 180 or more Hispanics in the sample just by chance. This gives us reason to be suspicious about the sampling process.

Chapter 7

Section 7.1

Answers to Check Your Understanding

page 425: **1.** Parameter: $\mu = 20$ ounces. Statistic: $\bar{x} = 19.6$ ounces. **2.** Parameter: $p = 0.10$, or 10% of passengers. Statistic: $\hat{p} = 0.08$, or 8% of the sample of passengers.

page 428: **1.** Individuals: M&M'S Milk Chocolate Candies; variable: color; and parameter of interest: proportion of orange M&M'S. The graph below shows the population distribution.
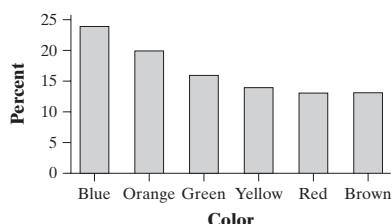
2. The graph below shows a possible distribution of sample data. For this sample there are 11 orange M&M'S, so $\hat{p} = \dfrac{11}{50} = 0.22$.
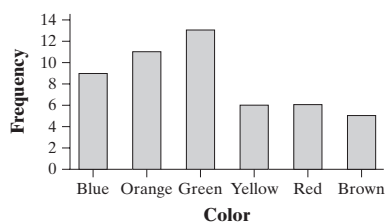
3. The middle graph is the approximate sampling distribution of $\hat{p}$ because the center of the distribution should be at approximately

0.20. The first graph shows the distribution of the colors for one sample and the third graph is centered at 0.40 rather than 0.20.

page 434: **1.** No. The mean of the approximate sampling distribution of the sample median (73.5) is not equal to the median of the population (75). **2.** Smaller. Larger samples provide more precise estimates because larger samples include more information about the population distribution. **3.** Skewed to the left and unimodal.

Answers to Odd-Numbered Section 7.1 Exercises

7.1 (a) *Population*: all people who signed a card saying that they intend to quit smoking. *Parameter*: the proportion of the population who actually quit smoking. *Sample*: a random sample of 1000 people who signed the cards. *Statistic*: the proportion of the sample who actually quit smoking; $\hat{p} = 0.21$. (b) *Population*: all the turkey meat. *Parameter*: minimum temperature in all of the turkey meat. *Sample*: four randomly chosen locations in the turkey. *Statistic*: minimum temperature in the sample of four locations; sample minimum $= 170°F$.

7.3 $\mu = 2.5003$ is a parameter and $\bar{x} = 2.5009$ is a statistic.

7.5 $\hat{p} = 0.48$ is a statistic and $p = 0.52$ is a parameter.

7.7 (a) 2 and 6 ($\bar{x} = 4$), 2 and 8 (5), 2 and 10 (6), 2 and 10 (6), 2 and 12 (7), 6 and 8 (7), 6 and 10 (8), 6 and 10 (8), 6 and 12 (9), 8 and 10 (9), 8 and 10 (9), 8 and 12 (10), 10 and 10 (10), 10 and 12 (11), 10 and 12 (11). (b) The sampling distribution of $\bar{x}$ is skewed to the left and unimodal. The mean of the sampling distribution is 8, which is equal to the mean of the population. The values of $\bar{x}$ vary from 4 to 11.
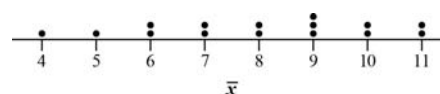
7.9 (a) In one simulated SRS of 100 students, there were 73 students who did all their assigned homework. (b) The distribution is reasonably symmetric and bell-shaped. It is centered at about 0.60. Values vary from about 0.47 to 0.74. There don't appear to be any outliers. (c) Yes, because there were no values of $\hat{p}$ less than or equal to 0.45 in the simulation. (d) Because it would be very surprising to get a sample proportion of 0.45 or less in an SRS of size 100 when $p = 0.60$, we should be skeptical of the newspaper's claim.

7.11 (a) A graph of the population distribution is shown below.
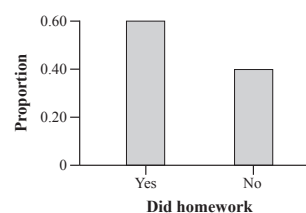
(b) Answers will vary. An example bar graph is given.
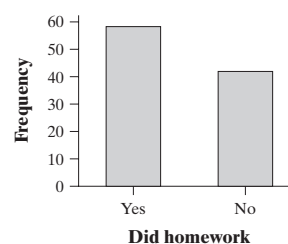
7.13 (a) Skewed to the right with a center at $9(°F)^2$. The values vary from about 2 to $27.5(°F)^2$. (b) A sample variance of $25(°F)^2$ provides convincing evidence that the manufacturer's claim is false and that the thermostat actually has more variability than claimed because a value this large was rare in the simulation.

7.15 If we chose many SRSs and calculated the sample mean $\bar{x}$ for each sample, we will not consistently underestimate μ or consistently overestimate μ.

7.17 A larger random sample will provide more information and, therefore, more precise results.

7.19 (a) Statistics ii and iii, because the means of their sampling distributions appear to be equal to the population parameter. (b) Statistic ii, because it is unbiased and has very little variability.

7.21 c

7.23 a

7.25 (a) We are looking for the percentage of values that are 2.5 standard deviations or farther below the mean in a Normal distribution. In other words, we are looking for $P(Z \leq -2.5)$. Using Table A, $P(Z \leq -2.5) = 0.0062$. Using technology: `normalcdf(lower: -1000,upper:-2.50,μ:0,σ:1) = 0.0062`. Less than 1% of healthy young adults have osteoporosis. (b) Let X be the BMD for women aged $70 - 79$ on the standard scale. Then X follows a $N(-2, 1)$ distribution and we want to find $P(X \leq -2.5)$.

$z = \dfrac{-2.5 - (-2)}{1} = -0.5$ and $P(Z \leq -0.5) = 0.3085$. Using technology: `normalcdf(lower:-1000,upper:-2.5, μ:-2,σ:1) = 0.3085`. About 31% of women aged $70 - 79$ have osteoporosis.

Section 7.2

Answers to Check Your Understanding

page 445: **1.** $\mu_{\hat{p}} = p = 0.75$. **2.** The standard deviation of the sampling distribution of $\hat{p}$ is $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}} = \sqrt{\dfrac{0.75(0.25)}{1000}} = 0.0137$. There are more than $10(1000) = 10{,}000$ young adult Internet users, so the 10% condition has been met. **3.** Yes. Both $np = 1000(0.75) = 750$ and $n(1-p) = 1000(0.25) = 250$ are at least 10. **4.** The sampling distribution would still be approximately Normal with mean 0.75. However, the standard deviation would be smaller by a factor of 3: $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}} = \sqrt{\dfrac{0.75(0.25)}{9000}} = 0.0046$.

Answers to Odd-Numbered Section 7.2 Exercises

7.27 (a) We would not be surprised to find 8 (32%) orange candies because values this small happened fairly often in the simulation. However, there were few samples in which there were 5 (20%) or fewer orange candies. So getting 5 orange candies would be surprising. (b) A sample of 50, because we expect to be closer to $p = 0.45$ in larger samples.

7.29 (a) $\mu_{\hat{p}} = p = 0.45$. (b) $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}} = \sqrt{\dfrac{0.45(0.55)}{25}} = 0.0995$. The 10% condition is met because there are more than $10(25) = 250$ candies in the large machine. (c) Yes, because $np = 25(0.45) = 11.25$ and $n(1-p) = 25(0.55) = 13.75$ are both at least 10. (d) The sampling distribution would still be approximately Normal with a mean of $\mu_{\hat{p}} = 0.45$. However, the standard deviation decreases to $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}} = \sqrt{\dfrac{0.45(0.55)}{100}} = 0.0497$.

7.31 (a) No, because more than 10% of the population $(10/76 = 13\%)$ was selected. (b) No, because the sample size was only $n = 10$. Neither np nor $n(1-p)$ will be at least 10.

7.33 The Large Counts condition is not met because $np = 15(0.3) = 4.5 < 10$.

7.35 (a) $\mu_{\hat{p}} = p = 0.70$. (b) $\sigma_{\hat{p}} = \sqrt{\dfrac{0.7(0.3)}{1012}} = 0.0144$. The 10% condition is met because the sample of size 1012 is less than 10% of the population of all U.S. adults. (c) Yes, because $np = 1012(0.70) = 708.4$ and $n(1-p) = 1012(0.30) = 303.6$ are both at least 10. (d) We want to find $P(\hat{p} \leq 0.67)$. $z = \dfrac{0.67 - 0.70}{0.0144} = -2.08$ and $P(Z \leq -2.08) = 0.0188$. Using technology: `normalcdf(lower:-1000,upper:0.67,μ:0.70,σ: 0.0144) = 0.0186`. There is a 0.0186 probability of obtaining a sample in which 67% or fewer say they drink the milk. Because this is a small probability, there is convincing evidence against the claim.

7.37 4048, because using $4n$ for the sample size halves the standard deviation $(\sqrt{4n} = 2\sqrt{n})$.

7.39 $\mu_{\hat{p}} = 0.70$. Because 267 is less than 10% of the population of college women, $\sigma_{\hat{p}} = \sqrt{\dfrac{0.7(0.3)}{267}} = 0.0280$. Because $np = 267(0.7) = 186.9$ and $n(1-p) = 267(0.3) = 80.1$ are both at least 10, the sampling distribution of $\hat{p}$ can be approximated by a Normal distribution. We want to find $P(\hat{p} \geq 0.75)$. $z = \dfrac{0.75 - 0.7}{0.0280} = 1.79$ and $P(Z \geq 1.79) = 0.0367$. Using technology: `normalcdf (lower:0.75,upper:1000,μ:0.7,σ:0.0280) = 0.0371`. There is a 0.0371 probability that 75% or more of the women in the sample have been on a diet within the last 12 months.

7.41 (a) $\mu_{\hat{p}} = 0.90$. Because 100 is less than 10% of the population of orders, $\sigma_{\hat{p}} = \sqrt{\dfrac{0.90(0.10)}{100}} = 0.03$. Because $np = 100(0.90) = 90$ and $n(1-p) = 100(0.10) = 10$ are both at least 10, the sampling distribution of $\hat{p}$ can be approximated by a Normal distribution. We want to find $P(\hat{p} \leq 0.86)$. $z = \dfrac{0.86 - 0.90}{0.03} = -1.33$ and $P(Z \leq -1.33) = 0.0918$. Using technology: `normalcdf (lower:-1000,upper:0.86,μ:0.90,σ:0.03) = 0.0912`. There is a 0.0912 probability that 86% or fewer of orders in an SRS of 100 were shipped within 3 working days. (b) Because the probability isn't very small, it is plausible that the 90% claim is correct and that the lower than expected percentage is due to chance alone.

7.43 a

7.45 b

7.47 The Venn diagram is shown below.
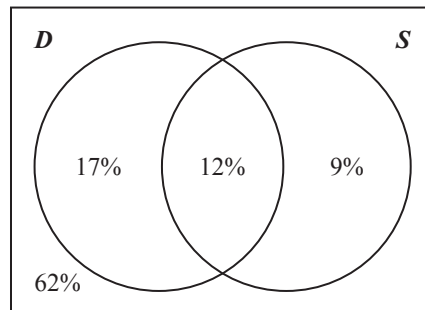
62% neither download nor share music files.

Section 7.3

Answers to Check Your Understanding

page 456: **1.** X = length of pregnancy follows a N(266, 16) distribution and we want to find $P(X > 270)$. $z = \dfrac{270 - 266}{16} = 0.25$ and $P(Z > 0.25) = 0.4013$. *Using technology*: `normalcdf(lower: 270,upper:1000,µ:266,σ:16)` = 0.4013. There is a 0.4013 probability of selecting a woman whose pregnancy lasts for more than 270 days. **2.** $\mu_{\bar{x}} = \mu = 266$ days **3.** The sample of size 6 is less than 10% of all pregnant women, so $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{16}{\sqrt{6}} = 6.532$ days. **4.** $\bar{x}$ follows a N(266, 6.532) distribution and we want to find $P(\bar{x} > 270)$. $z = \dfrac{270 - 266}{6.532} = 0.61$ and $P(Z > 0.61) = 0.2709$. *Using technology*: `normalcdf(lower:270,upper:1000, µ:266,σ:6.532)` = 0.2701. There is a 0.2701 probability of selecting a sample of 6 women whose mean pregnancy length exceeds 270 days.

Answers to Odd-Numbered Section 7.3 Exercises

7.49 $\mu_{\bar{x}} = \mu = 255$ seconds. Because the sample size (10) is less than 10% of the population of songs on David's iPod, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{60}{\sqrt{10}} = 18.974$ seconds.

7.51 $30 = \dfrac{60}{\sqrt{n}} \rightarrow \sqrt{n} = \dfrac{60}{30} = 2 \rightarrow n = 4$.

7.53 **(a)** Normal with $\mu_{\bar{x}} = \mu = 188$ mg/dl. Because the sample size (100) is less than 10% of all men aged 20 to 34, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{41}{\sqrt{100}} = 4.1$ mg/dl. **(b)** We want to find $P(185 \leq \bar{x} \leq 191)$. $z = \dfrac{185 - 188}{4.1} = -0.73$ and $z = \dfrac{191 - 188}{4.1} = 0.73$

$P(-0.73 \leq Z \leq 0.73) = 0.5346$. *Using technology*: `normalcdf(lower:185,upper:191,µ:188,σ:4.1)` = 0.5357. There is a 0.5357 probability that $\bar{x}$ estimates μ within ± 3 mg/dl.

(c) $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{41}{\sqrt{1000}} = 1.30$ mg/dl. So $\bar{x}$ follows a N(188, 1.30) distribution and we want to find $P(185 \leq \bar{x} \leq 191)$. $z = \dfrac{185 - 188}{1.30}$

$= -2.31$ and $z = \dfrac{191 - 188}{1.30} = 2.31$. $P(-2.31 \leq Z \leq 2.31)$

$= 0.9792$. *Using technology*: `normalcdf(lower:185, upper:191,µ:188,σ:1.30)` = 0.9790. There is a 0.9790 probability that $\bar{x}$ estimates μ within ± 3 mg/dl. The larger sample is better because it is more likely to produce a sample mean within 3 mg/dl of the population mean.

7.55 **(a)** Let X = amount of cola in a randomly selected bottle. X follows the N(298, 3) distribution and we want to find $P(X < 295)$. $z = \dfrac{295 - 298}{3} = -1$ and $P(Z < -1) = 0.1587$. *Using technology*: `normalcdf(lower:−1000,upper:295, µ:298,σ:3)` = 0.1587. There is a 0.1587 probability that a randomly selected bottle contains less than 295 ml. **(b)** $\mu_{\bar{x}} = \mu = 298$ ml. Because 6 is less than 10% of all bottles produced, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{3}{\sqrt{6}} = 1.2247$ ml. We want to find $P(\bar{x} < 295)$

using the N(298, 1.2247) distribution. $z = \dfrac{295 - 298}{1.2247} = -2.45$ and $P(Z < -2.45) = 0.0071$. *Using technology*: `normalcdf (lower:−1000,upper:295,µ:298,σ:1.2247)` = 0.0072. There is a 0.0072 probability that the mean contents of six randomly selected bottles are less than 295 ml.

7.57 No. The histogram of the sample values will look like the population distribution. The CLT says that the histogram of the sampling distribution of the *sample mean* will look more and more Normal as the sample size increases.

7.59 **(a)** Because the distribution of the play times of the population of songs is heavily skewed to the right and $n = 10 < 30$. **(b)** Because $n = 36 \geq 30$, the CLT applies. $\mu_{\bar{x}} = \mu = 225$ seconds. Because 36 is less than 10% of all songs on David's iPod, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{60}{\sqrt{36}} = 10$ seconds. We want to find $P(\bar{x} > 240)$ using the N(225, 10) distribution. $z = \dfrac{240 - 225}{10} = 1.50$ and $P(Z > 1.50) = 0.0668$. *Using technology*: `normalcdf(lower: 240,upper:1000,µ:225,σ:10)` = 0.0668. There is a 0.0668 probability that the mean play time is more than 240 seconds.

7.61 **(a)** We do not know the shape of the distribution of passenger weights. **(b)** We want to find $P(\bar{x} > 6000/30) = P(\bar{x} > 200)$. Because the sample size is large ($n = 30 \geq 30$), the distribution of $\bar{x}$ is approximately Normal with $\mu_{\bar{x}} = \mu = 190$ pounds. Because $n = 30$ is less than 10% of all possible passengers, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{35}{\sqrt{30}} = 6.3901$ pounds. $z = \dfrac{200 - 190}{6.3901} = 1.56$ and $P(Z > 1.56) = 0.0594$. *Using technology*: `normalcdf(lower: 200,upper:1000,µ:190,σ:6.3901)` = 0.0588. There is a 0.0588 probability that the mean weight exceeds 200 pounds.

7.63 Because the sample size is large ($n = 10,000 \geq 30$), the sampling distribution of $\bar{x}$ is approximately Normal. $\mu_{\bar{x}} = \mu = \$250$. Assuming 10,000 is less than 10% of all homeowners with fire insurance, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{1000}{\sqrt{10,000}} = \$10$. We want to find $P(\bar{x} \leq 275)$ using the N(250, 10) distribution. $z = \dfrac{275 - 250}{10} = 2.50$ and $P(Z \leq 2.50) = 0.9938$. *Using technology*: `normalcdf(lower: −1000,upper:275,µ:250,σ:10)` = 0.9938. There is a 0.9938 probability that the mean annual loss from a sample of 10,000 policies is no greater than \$275.

7.65 b

7.67 b

7.69 Didn't finish high school: $\dfrac{1062}{12,470} = 0.0852$; high school but no college: $\dfrac{1977}{37,834} = 0.0523$, less than a bachelor's degree: $\dfrac{1462}{34,439} = 0.0425$, college graduate: $\dfrac{1097}{40,390} = 0.0272$. The unemployment rate decreases with additional education.

7.71 $P(\text{in labor force} \mid \text{college graduate}) = \dfrac{40,390}{51,582} = 0.7830$.

Answers to Chapter 7 Review Exercises

R7.1 The population is the set of all eggs shipped in one day. The sample consists of the 200 eggs examined. The parameter is the proportion $p = 0.03$ of eggs shipped that day that had salmonella.

The statistic is the proportion $\hat{p} = \dfrac{9}{200} = 0.045$ of eggs in the sample that had salmonella.

R7.2 (a) A sketch of the population distribution is given below.

$N(3668, 511)$

2135 2646 3157 3668 4179 4690 5201
Birth weight (grams)
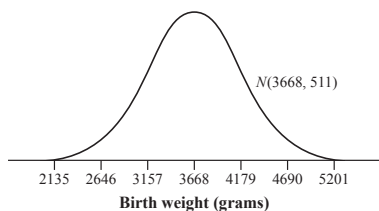
(b) Answers will vary. An example dotplot is given. (c) The dot at 2750 represents one SRS of size 5 from this population where the sample range was 2750 grams.

2600   3000   3400   3800   4200   4600
Birth weight (grams)
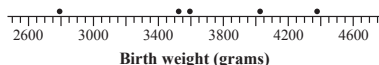
R7.3 (a) No, because sample range is always less than the actual range (3417). If it were unbiased, the distribution would be centered at 3417. (b) Take larger samples.

R7.4 (a) $\mu_{\hat{p}} = p = 0.15$. (b) Because the sample size of $n = 1540$ is less than 10% of the population of all adults,

$\sigma_{\hat{p}} = \sqrt{\dfrac{0.15(0.85)}{1540}} = 0.0091$. (c) Yes, because $np = 1540(0.15)$

$= 231$ and $n(1 - p) = 1540(0.85) = 1309$ are both at least 10.

(d) We want to find $P(0.13 \leq \hat{p} \leq 0.17)$. $z = \dfrac{0.13 - 0.15}{0.0091} = -2.20$

and $z = \dfrac{0.17 - 0.15}{0.0091} = 2.20$. The desired probability is

$P(-2.20 \leq Z \leq 2.20) = 0.9722$. *Using technology*: `normalcdf`
`(lower:0.13,upper:0.17,μ:0.15,σ:0.0091)`
$= 0.9720$. There is a 0.9720 probability of obtaining a sample in which between 13% and 17% are joggers.

R7.5 (a) $\mu_{\hat{p}} = p = 0.30$. Because 100 is less than 10% of the popu-

lation of travelers, $\sigma_{\hat{p}} = \sqrt{\dfrac{0.30(0.70)}{100}} = 0.0458$. Because

$np = 100(0.30) = 30$ and $n(1 - p) = 100(0.70) = 70$ are both at least 10, the sampling distribution of $\hat{p}$ can be approximated by a Normal distribution. We want to find $P(\hat{p} \leq 0.20)$.

$z = \dfrac{0.20 - 0.30}{0.0458} = -2.18$ and $P(Z \leq -2.18) = 0.0146$. *Using*

technology: `normalcdf(lower:−1000,upper:0.20,μ:`
`0.30,σ:0.0458) = 0.0145`. There is a 0.0145 probability that 20% or fewer of the travelers get a red light. (b) Because this is a small probability, there is convincing evidence against the agents' claim—it isn't plausible to get a sample proportion of travelers with a red light this small by chance alone.

R7.6 (a) X = WAIS score for a randomly selected individual follows a $N(100, 15)$ distribution and we want to find
$P(X \geq 105)$. $z = \dfrac{105 - 100}{15} = 0.33$ and $P(Z \geq 0.33) = 0.3707$.

Using technology: `normalcdf(lower:105,upper:1000,`
`μ:100,σ:15) = 0.3694`. There is a 0.3694 probability of selecting an individual with a WAIS score of at least 105.
(b) $\mu_{\bar{x}} = \mu = 100$. Because the sample of size 60 is less than

10% of all adults, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{15}{\sqrt{60}} = 1.9365$. (c) $\bar{x}$ follows a

$N(100, 1.9365)$ distribution and we want to find $P(\bar{x} \geq 105)$.
$z = \dfrac{105 - 100}{1.9365} = 2.58$ and $P(Z \geq 2.58) = 0.0049$. *Using technol-*
ogy: `normalcdf(lower:105,upper:1000,μ:100,σ:`
`1.9365) = 0.0049`. There is a 0.0049 probability of selecting a sample of 60 adults whose mean WAIS score is at least 105. (d) The answer to part (a) could be quite different depending on the shape of the population distribution. The answer to part (b) would be the same because the mean and standard deviation do not depend on the shape of the population distribution. Because of the large sample size ($60 \geq 30$), the answer for part (c) would still be fairly reliable due to the central limit theorem.

R7.7 (a) Because $n = 50 \geq 30$. (b) $\mu_{\bar{x}} = \mu = 0.5$. Because 50 is less than 10% of all traps, the standard deviation is

$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{0.7}{\sqrt{50}} = 0.0990$. Thus, $\bar{x}$ follows a $N(0.5, 0.0990)$

distribution and we want to find $P(\bar{x} \geq 0.6)$. $z = \dfrac{0.6 - 0.5}{0.0990} = 1.01$

and $P(Z \geq 1.01) = 0.1562$. *Using technology*: `normalcdf`
`(lower:0.6,upper:1000,μ:0.5,σ:0.0990) = 0.1562`.
There is a 0.1562 probability that the mean number of moths is greater than or equal to 0.6. (c) No. Because this probability is not small, it is plausible that the sample mean number of moths is this high by chance alone.

Answers to Chapter 7 AP® Statistics Practice Test

T7.1 c
T7.2 c
T7.3 c
T7.4 a
T7.5 b
T7.6 b
T7.7 b
T7.8 e
T7.9 c
T7.10 e
T7.11 A. Both A and B appear to be unbiased, and A has less variability than B.
T7.12 (a) We do not know the shape of the population distribution of monthly fees. (b) $\mu_{\bar{x}} = \mu = \$38$. Because the sample of size 500 is less than 10% of all households with Internet access,

$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{10}{\sqrt{500}} = 0.4472$. (c) Because the sample size is large

($n = 500 \geq 30$), the distribution of $\bar{x}$ will be approximately

Normal. (d) We want to find $P(\bar{x} > 39)$. $z = \dfrac{39 - 38}{0.4472} = 2.24$ and

$P(Z > 2.24) = 0.0125$. *Using technology*: `normalcdf(lower:`
`39,upper:1000,μ:38,σ:0.4472) = 0.0127`. There is a 0.0127 probability that the mean monthly fee exceeds $39.
T7.13 $\mu_{\hat{p}} = p = 0.22$. Because 300 is less than 10% of children

under the age of 6, $\sigma_{\hat{p}} = \sqrt{\dfrac{0.22(0.78)}{300}} = 0.0239$. Because

$np = 300(0.22) = 66$ and $n(1 - p) = 300(0.78) = 234$ are both at least 10, the sampling distribution of $\hat{p}$ can be approximated by a Normal distribution. We want to find $P(\hat{p} > 0.20)$. $z =$

$\dfrac{0.20 - 0.22}{0.0239} = -0.84$ and $P(Z > -0.84) = 0.7995$. *Using tech-*

nology: `normalcdf(lower:0.20,upper:1000,μ:0.22,σ:`
`0.0239) = 0.7987`. There is a 0.7987 probability that more than 20% of the sample are from poverty-level households.

Answers to Cumulative AP® Practice Test 2

AP2.1 a
AP2.2 d
AP2.3 e
AP2.4 b
AP2.5 c
AP2.6 e
AP2.7 c
AP2.8 a
AP2.9 d
AP2.10 c
AP2.11 b
AP2.12 c
AP2.13 d
AP2.14 c
AP2.15 d
AP2.16 c
AP2.17 e
AP2.18 a
AP2.19 c
AP2.20 b
AP2.21 a

AP2.22 (a) Observational study, because no treatments were imposed on the subjects. (b) Two variables are confounded when their effects on the cholesterol level cannot be distinguished from one another. For example, people who take omega-3 fish oil might also exercise more. Researchers would not know whether it was the omega-3 fish oil or the exercise that was the real explanation for lower cholesterol. (c) No. Even though the difference was statistically significant, this wasn't an experiment and taking fish oil is possibly confounded with exercise.

AP2.23 (a) P(type O or Hawaiian-Chinese) = 65,516/145,057 = 0.452. (b) P(type AB$|$Hawaiian) = 99/4670 = 0.021.
(c) P(Hawaiian) = 4670/145,057 = 0.032; P(Hawaiian$|$type B) = 178/17,604 = 0.010. Because these probabilities are not equal, the two events are not independent. (d) P(type A and white) = 50,008/145,057 = 0.345. P(at least one type A and white) = 1 − P(neither are type A and white) = 1 − (1 − 0.345)² = 0.571.

AP2.24 (a) The distribution of seed mass for the cicada plants is roughly symmetric, while the distribution for the control plants is skewed to the left. The median seed mass is the same for both groups. The cicada plants had a bigger range in seed mass, but the control plants had a bigger *IQR*. Neither group had any outliers. (b) The cicada plants. The distribution of seed mass for the cicada plants is roughly symmetric, which suggests that the mean should be about the same as the median. However, the distribution of seed mass for the control plants is skewed to the left, which will pull the mean of this distribution below its median toward the lower values. Because the medians of both distributions are equal, the mean for the cicada plants is greater than the mean for the control plants. (c) The purpose of the random assignment is to create two groups of plants that are roughly equivalent at the beginning of the experiment. (d) Benefit: controlling a source of variability. Different types of flowers will have different seed masses, making the response more variable if other types of plants were used. Drawback: we can't make inferences about the effect of cicadas on other types of plants, because other plants might respond differently to cicadas.

AP2.25 (a) We want to find $P(\bar{x} < 25,000/50) = P(\bar{x} < 500)$. Because the sample size is large ($n = 50 \geq 30$), the distribution of $\bar{x}$ is approximately Normal with $\mu_{\bar{x}} = \mu = 525$ pages. Because

$n = 50$ is less than 10% of all novels in the library,

$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{200}{\sqrt{50}} = 28.28$ pages. $z = \dfrac{500 - 525}{28.28} = -0.88$ and

$P(Z < -0.88) = 0.1894.$ *Using technology:* `normalcdf(lower:` `-1000,upper:500,μ:525,σ:28.28) = 0.1883`. There is a 0.1883 probability that the total number of pages in 50 novels is fewer than 25,000. (b) Let X be the number of novels that have fewer than 400 pages. X is a binomial random variable with $n = 50$ and $p = 0.30$. We want to find $P(X \geq 20)$. *Using technology:* $P(X \geq 20) = 1 - P(X \leq 19) = 1 -$ `binomcdf(trials:50,` `p:0.30, x value:19) = 0.0848`. There is a 0.0848 probability of selecting at least 20 novels that have fewer than 400 pages. *Note:* Using the Normal approximation, $P(X \geq 20) = 0.0614$.

Chapter 8

Section 8.1

Answers to Check Your Understanding
page 485: **1.** We are 95% confident that the interval from 2.84 to 7.55 g captures the population standard deviation of the fat content of Brand X hot dogs. **2.** If this sampling process were repeated many times, approximately 95% of the resulting confidence intervals would capture the population standard deviation of the fat content of Brand X hot dogs. **3.** False. Once the interval is calculated, it either contains σ or it does not contain σ.

Answers to Odd-Numbered Section 8.1 Exercises
8.1 Sample mean, $\bar{x} = 30.35$.

8.3 Sample proportion, $\hat{p} = \dfrac{36}{50} = 0.72$.

8.5 (a) Approximately Normal with mean $\mu_{\bar{x}} = 280$ and standard deviation $\sigma_{\bar{x}} = \dfrac{60}{\sqrt{840}} = 2.1$. (b) See graph below. (c) About 95% of the $\bar{x}$ values will be within 2 standard deviations of the mean. Therefore, $m = 2(2.1) = 4.2$. (d) About 95%.

273.7 275.8 277.9 280.0 282.1 284.2 286.3
$\bar{x}$
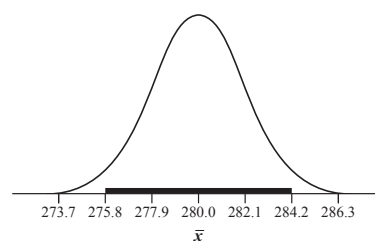
8.7 The sketch is given below. The interval with the value of $\bar{x}$ in the shaded region will contain the population mean (280), while the interval with the value of $\bar{x}$ outside the shaded region will not contain the population mean (280).

273.7 275.8 277.9 280.0 282.1 284.2 286.3
$\bar{x}$
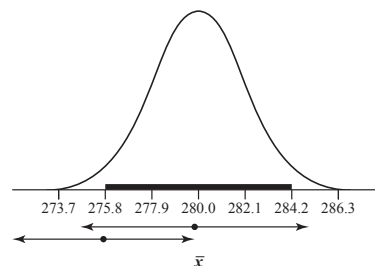
8.9 **(a)** We are 95% confident that the interval from 0.63 to 0.69 captures the true proportion of those who favor an amendment to the Constitution that would permit organized prayer in public schools. **(b)** Point estimate $= \hat{p} = \dfrac{0.63 + 0.69}{2} = 0.66$ and margin of error $= 0.69 - 0.66 = 0.03$. **(c)** Because the value $2/3 = 0.667$ (and values less than 2/3) are in the interval of plausible values, there is not convincing evidence that more than two-thirds of U.S. adults favor such an amendment.

8.11 Because only 84% of the intervals actually contained the true parameter, these were probably 80% or 90% confidence intervals.

8.13 Answers will vary. One practical difficulty is response bias: people might answer "yes" because they think they should, even if they don't really support the amendment.

8.15 *Interval*: We are 95% confident that the interval from 10.9 to 26.5 captures the true difference (girls − boys) in the mean number of pairs of shoes owned by girls and boys. *Level*: If this sampling process were repeated many times, approximately 95% of the resulting confidence intervals would capture the true difference (girls − boys) in the mean number of pairs of shoes owned by girls and boys.

8.17 Yes. Because the interval does not include 0 as a plausible value, there is convincing evidence of a difference in the mean number of shoes for boys and girls.

8.19 **(a)** Incorrect. The interval provides plausible values for the *mean* BMI of all women, not plausible values for individual BMI measurements. **(b)** Incorrect. We shouldn't use the results of one sample to predict the results for future samples. **(c)** Correct. A confidence interval provides an interval of plausible values for a parameter. **(d)** Incorrect. The population mean doesn't change and will either be a value between 26.2 and 27.4 100% of the time or 0% of the time. **(e)** Incorrect. We are 95% confident that the population mean is between 26.2 and 27.4, but that does not absolutely rule out any other possibility.

8.21 b

8.23 e

8.25 **(a)** Observational study, because there was no treatment imposed on the pregnant women or the children. **(b)** No. We cannot make any conclusions about cause and effect because this was not an experiment.

Section 8.2

Answers to Check Your Understanding

page 496: **1.** Random: not met because this was a convenience sample. 10%: met because the sample of 100 is less than 10% of the population at a large high school. Large Counts: met because 17 successes and 83 failures are both at least 10. **2.** Random: met because the inspector chose an SRS of bags. 10%: met because the sample of 25 is less than 10% of the thousands of bags filled in an hour. Large Counts: not met because there were only 3 successes, which is less than 10.

page 499: **1.** $p =$ the true proportion of all U.S. college students who are classified as frequent binge drinkers. **2.** Random: met because the statement says that the students were chosen randomly. 10%: met because the sample of 10,904 is less than 10% of all U.S. college students. Large Counts: met because 2486 successes and 8418 failures are both at least 10. **3.** $\dfrac{1 - 0.99}{2} = 0.005$ and the

closest area in Table A is 0.0051 (or 0.0049), corresponding to a critical value of $z^* = 2.57$ (or 2.58). *Using technology*: `invNorm(area:0.005, μ:0, σ:1) = −2.576`, so $z^* = 2.576$. $0.228 \pm 2.576\sqrt{\dfrac{0.228(1 - 0.228)}{10904}} = 0.228 \pm 0.010 = (0.218, 0.238)$. **4.** We are 99% confident that the interval from 0.218 and 0.238 captures the true proportion of all U.S. college students who are classified as frequent binge drinkers.

page 503: **1.** Solving $1.96\sqrt{\dfrac{0.80(0.20)}{n}} \leq 0.03$ for n gives $n \geq 682.95$. We should select a sample of at least 683 customers. **2.** The required sample size will be larger because the critical value is larger for 99% confidence (2.576) versus 95% confidence (1.96). The company would need to select at least 1180 customers.

Answers to Odd-Numbered Section 8.2 Exercises

8.27 Random: met because Latoya selected an SRS of students. 10%: not met because the sample size (50) is more than 10% of the population of seniors in the dormitory (175). Large Counts: met because $n\hat{p} = 14 \geq 10$ and $n(1 - \hat{p}) = 36 \geq 10$.

8.29 Random: may not be met because we do not know if the people who were contacted were a random sample. 10%: met because the sample size (2673) is less than 10% of the population of adult heterosexuals. Large Counts: not met because $n\hat{p} = 2673(0.002) \approx 5$ is not at least 10.

8.31 $\dfrac{1 - 0.98}{2} = 0.01$, and the closest area is 0.0099, corresponding to a critical value of $z^* = 2.33$. *Using technology*: `invNorm (area:0.01, μ:0, σ:1) = −2.326`, so $z^* = 2.326$.

8.33 **(a)** Population: seniors at Tonya's high school. Parameter: true proportion of all seniors who plan to attend the prom. **(b)** Random: the sample is a simple random sample. 10%: The sample size (50) is less than 10% of the population size (750). Large Counts: $n\hat{p} = 36 \geq 10$ and $n(1 - \hat{p}) = 14 \geq 10$.

(c) $0.72 \pm 1.645\sqrt{\dfrac{0.72(0.28)}{50}} = 0.72 \pm 0.10 = (0.62, 0.82)$.

(d) We are 90% confident that the interval from 0.62 to 0.82 captures the true proportion of all seniors at Tonya's high school who plan to attend the prom.

8.35 **(a)** S: $p =$ the true proportion of all full-time U.S. college students who are binge drinkers. P: One-sample z interval for p. Random: the students were selected randomly. 10%: the sample size (5914) is less than 10% of the population of all college students. Large Counts: $n\hat{p} = 2312 \geq 10$ and $n(1 - \hat{p}) = 3602 \geq 10$. D: (0.375, 0.407). C: We are 99% confident that the interval from 0.375 to 0.407 captures the true proportion of full-time U.S. college students who are binge drinkers. **(b)** Because the value 0.45 does not appear in our 99% confidence interval, it isn't plausible that 45% of full-time U.S. college students are binge drinkers.

8.37 Answers will vary. Response bias is one possibility.

8.39 S: $p =$ the true proportion of all students retaking the SAT who receive coaching. P: One-sample z interval for p. Random: the students were selected randomly. 10%: the sample size (3160) is less

than 10% of the population of all students taking the SAT twice. Large Counts: $n\hat{p} = 427 \geq 10$ and $n(1 - \hat{p}) = 2733 \geq 10$. D: (0.119, 0.151). C: We are 99% confident that the interval from 0.119 to 0.151 captures the true proportion of students retaking the SAT who receive coaching.

8.41 (a) We do not know the sample sizes for the men and for the women. (b) The margin of error for women alone would be greater than 0.03 because the sample size for women alone is smaller than 1019.

8.43 (a) Solving $1.645\sqrt{\dfrac{0.75(0.25)}{n}} \leq 0.04$ gives $n \geq 318$.

(b) Solving $1.645\sqrt{\dfrac{0.5(0.5)}{n}} \leq 0.04$ gives $n \geq 423$. In this case, the sample size needed is 105 people larger.

8.45 Solving $1.96\sqrt{\dfrac{0.5(0.5)}{n}} \leq 0.03$ gives $n \geq 1068$.

8.47 (a) Solving $0.03 = z^*\sqrt{\dfrac{0.64(0.36)}{1028}}$ gives $z^* = 2.00$. The confidence level is likely 95%, because 2.00 is very close to 1.96. (b) Teens are hard to reach and often unwilling to participate in surveys, so nonresponse is a major "practical difficulty" for this type of poll.
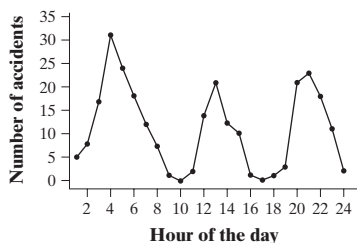
8.49 a

8.51 c

8.53 (a) A histogram of the number of accidents per hour is given below.

Number of accidents/hr
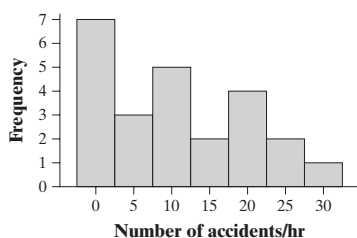
(b) A graph of the number of accidents is given below.

Hour of the day

(c) The histogram in part (a) shows that the number of accidents has a distribution that is skewed to the right. (d) The graph in part (b) shows that there is a cyclical nature to the number of accidents.

Section 8.3

Answers to Check Your Understanding

page 514: **1.** df = 21, $t^* = 2.189$. *Using technology:* `invT (area:0.02, df = 21) = −2.189`, so $t^* = 2.189$. **2.** df = 70, $t^* = 2.660$ (using df = 60). *Using technology:* `invT(area: 0.005, df = 70) = −2.648`, so $t^* = 2.648$.

page 522: S: We are trying to estimate μ = the true mean healing rate at a 95% confidence level. P: One-sample t interval for μ.

Random: The description says that the newts were randomly chosen. 10%: The sample size (18) is less than 10% of the population of newts. Normal/Large Sample: The histogram below shows no strong skewness or outliers, so this condition is met.

Healing rate (micrometers/hr)
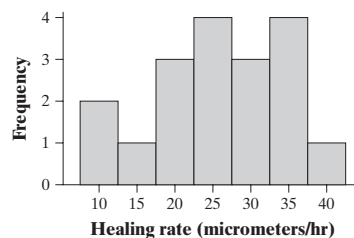
D: $25.67 \pm 2.110\left(\dfrac{8.32}{\sqrt{18}}\right) = 25.67 \pm 4.14 = (21.53, 29.81)$. C: We are 95% confident that the interval from 21.53 to 29.81 micrometers per hour captures the true mean healing time for newts.

page 524: Using $\sigma = 154$ and $z^* = 1.645$ for 90% confidence, $30 \geq 1.645\dfrac{154}{\sqrt{n}}$. Thus, $n \geq \left(\dfrac{1.645(154)}{30}\right)^2 = 71.3$, so take a sample of 72 students.

Answers to Odd-Numbered Section 8.3 Exercises

8.55 (a) $t^* = 2.262$. (b) $t^* = 2.861$. (c) $t^* = 1.671$ (*using technology:* $t^* = 1.665$).

8.57 Because the sample size is small ($n = 20 < 30$) and there are outliers in the data.

8.59 (a) No, because we are trying to estimate a population proportion, not a population mean. (b) No, because the 15 team members are not a random sample from the population. (c) No, because the sample size is small ($n = 25 < 30$) and there are outliers in the sample.

8.61 $SE_{\bar{x}} = \dfrac{9.3}{\sqrt{27}} = 1.7898$. If we take many samples of size 27, the sample mean blood pressure will typically vary by about 1.7898 from the population mean blood pressure.

8.63 (a) Because $19.03 = \dfrac{s_x}{\sqrt{23}}$, $s_x = 91.26$ cm. (b) They are using a critical value of $t^* = 1$. With df = 22, the area between $t = -1$ and $t = 1$ is approximately `tcdf(lower: −1, upper: 1, df: 22) = 0.67`. So, the confidence level is 67%.

8.65 (a) S:μ = the true mean percent change in BMC for breast-feeding mothers. P: One-sample t interval for μ. Random: the mothers were randomly selected. 10%: 47 is less than 10% of all breast-feeding mothers. Normal/Large Sample: $n = 47 \geq 30$. D: Using df = 40, (−4.575, −2.605). *Using technology:* (−4.569, −2.605) with df = 46. C: We are 99% confident that the interval from −4.569 to −2.605 captures the true mean percent change in BMC for breast-feeding mothers. (b) Because all of the plausible values in the interval are negative (indicating bone loss), the data give convincing evidence that breast-feeding mothers lose bone mineral, on average.

8.67 (a) S:μ = the true mean size of the muscle gap for the population of American and European young men. P: One-sample t interval for μ. Random: the young men were randomly selected. 10%: 200 is less than 10% of young men in America and Europe. Normal/Large Sample: $n = 200 \geq 30$. D: Using df = 100, (1.999, 2.701). *Using technology:* (2.001, 2.699) with df = 199. C: We are

95% confident that the interval from 2.001 to 2.699 captures the true mean size of the muscle gap for the population of American and European young men. **(b)** The large sample size ($n = 200 \geq 30$) allows us to use a t interval for μ.

8.69 S:μ = the true mean fuel efficiency for this vehicle. P: One-sample t interval for μ. Random: the records were selected at random. 10%: it is reasonable to assume that 20 is less than 10% of all records for this vehicle. Normal/Large Sample: the histogram does not show any strong skewness or outliers.
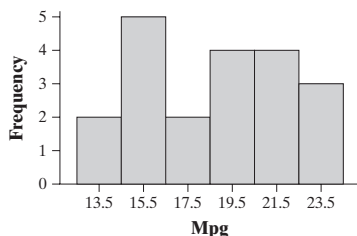
D: Using df = 19, (17.022, 19.938). C: We are 95% confident that the interval from 17.022 to 19.938 captures the true mean fuel efficiency for this vehicle.

8.71 **(a)** S: μ = the true mean difference in the estimates from these two methods in the population of tires. P: One-sample t interval for μ. Random: A random sample of tires was selected. 10%: the sample size (16) is less than 10% of all tires. Normal/Large Sample: The histogram of differences shows no strong skewness or outliers.
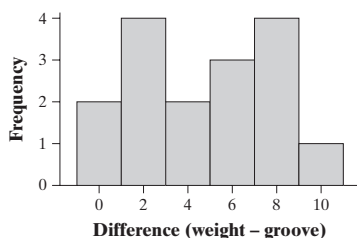
D: Using df = 15, (2.837, 6.275). C: We are 95% confident that the interval from 2.837 to 6.275 thousands of miles captures the true mean difference in the estimates from these two methods in the population of tires. **(b)** Because 0 is not included in the confidence interval, there is convincing evidence of a difference in the two methods of estimating tire wear.

8.73 Solving $2.576 \dfrac{7.5}{\sqrt{n}} \leq 1$ gives $n \geq 374$.

8.75 b

8.77 b

8.79 **(a)** Because the sum of the probabilities must be 1, $P(X = 7) = 0.57$. **(b)** $\mu_X = 5.44$. If we were to randomly select many young people, the average number of days they watched television in the past 7 days would be about 5.44. **(c)** Because the sample size is large ($n = 100 \geq 30$), we expect the mean number of days $\bar{x}$ for 100 randomly selected young people (aged 19 to 25) to be approximately Normally distributed with mean $\mu_{\bar{x}} = \mu = 5.44$. Because the sample size (100) is less than 10% of all young people aged 19 to 25, the standard deviation is $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{2.14}{\sqrt{100}} = 0.214$.

We want to find $P(\bar{x} \leq 4.96)$. $z = \dfrac{4.96 - 5.44}{0.214} = -2.24$ and

$P(Z \leq -2.24) = 0.0125.$ *Using technology*: `normalcdf(lower: -1000, upper:4.96, μ:5.44, σ:0.214) = 0.0124`. There is a 0.0124 probability of getting a sample mean of 4.96 or smaller. Because this probability is small, a sample mean of 4.96 or smaller would be surprising.

Answers to Chapter 8 Review Exercises

R8.1 **(a)** $\dfrac{1 - 0.94}{2} = 0.03$, and the closest area is 0.0301, corresponding to a critical value of $z^* = 1.88$. *Using technology*: `invNorm(area:0.03, μ:0, σ:1) = -1.881`, so $z^* = 1.881$. **(b)** Using Table B and 50 degrees of freedom, $t^* = 2.678$. *Using technology*: `invT(area:0.005, df:57) = -2.665`, so $t^* = 2.665$.

R8.2 **(a)** $\bar{x} = \dfrac{430 + 470}{2} = 450$ minutes. Margin of error $= 470 - 450 = 20$ minutes. Because $n = 30$, df = 29 and $t^* = 2.045$. Because $20 = 2.045 \dfrac{s_x}{\sqrt{30}}$, standard error $= \dfrac{s_x}{\sqrt{30}} = 9.780$ minutes and $s_x = 53.57$ minutes. **(b)** The confidence interval provided gives an interval estimate for the *mean* lifetime of batteries produced by this company, not individual lifetimes. **(c)** No. A confidence interval provides a statement about an unknown population mean, not another sample mean. **(d)** If we were to take many samples of 30 batteries and compute 95% confidence intervals for the mean lifetime, about 95% of these intervals will capture the true mean lifetime of the batteries.

R8.3 **(a)** p = the proportion of all adults aged 18 and older who would say that football is their favorite sport to watch on television. It may not equal 0.37 because the proportion who choose football will vary from sample to sample. **(b)** Random: The sample was random. 10%: The sample size (1000) is less than 10% of all adults. Large Counts: $n\hat{p} = 370 \geq 10$ and $n(1 - \hat{p}) = 630 \geq 10$.

(c) $0.37 \pm 1.96\sqrt{\dfrac{0.37(0.63)}{1000}} = (0.3401, 0.3999)$. **(d)** We are 95% confident that the interval from 0.3401 to 0.3999 captures the true proportion of all adults who would say that football is their favorite sport to watch on television.

R8.4 **(a)** μ = mean IQ score for the 1000 students in the school. **(b)** Random: the data are from an SRS. 10%: the sample size (60) is less than 10% of the 1000 students at the school. Normal/Large Sample: $n = 60 \geq 30$. **(c)** Using df = 50, $114.98 \pm 1.676\left(\dfrac{14.8}{\sqrt{60}}\right) = (111.778, 118.182)$. *Using technology*: (111.79, 118.17) with df = 59. **(d)** We are 90% confident that the interval from 111.79 to 118.17 captures the true mean IQ score for the 1000 students in the school.

R8.5 Solving $2.576\sqrt{\dfrac{0.5(0.5)}{n}} \leq 0.01$ gives $n \geq 16,590$.

R8.6 **(a)** S: p = the true proportion of all drivers who have run at least one red light in the last 10 intersections they have entered. P: One-sample z interval for p. Random: the drivers were selected at random. 10%: The sample size (880) is less than 10% of all drivers. Large Counts: $n\hat{p} = 171 \geq 10$ and $n(1 - \hat{p}) = 709 \geq 10$. D: (0.168, 0.220). C: We are 95% confident that the interval from 0.168 to 0.220 captures the true proportion of all drivers who have run at least one red light in the last 10 intersections they have

entered. **(b)** It is likely that more than 171 respondents have run red lights because some people may lie and say they haven't run a red light. The margin of error does not account for these sources of bias; it accounts only for sampling variability.

R8.7 (a) S: μ = the true mean measurement of the critical dimension for the engine crankshafts produced in one day. P: One-sample t interval for μ. Random: The data come from an SRS. 10%: the sample size (16) is less than 10% of all crankshafts produced in one day. Normal/Large Sample: the histogram shows no strong skewness or outliers.

Crankshaft measurement (mm)
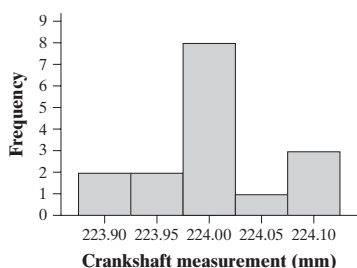
D: Using df $=$ 15, (223.969, 224.035). C: We are 95% confident that the interval from 223.969 to 224.035 mm captures the true mean measurement of the critical dimension for engine crankshafts produced on this day. **(b)** Because 224 is a plausible value in this interval, we don't have convincing evidence that the process mean has drifted.

R8.8 Solving $1.96\left(\dfrac{3000}{\sqrt{n}}\right) \leq 1000$ gives $n \geq 35$.

R8.9 (a) The margin of error must get larger to increase the capture rate of the intervals. **(b)** If we quadruple the sample size, the margin of error will decrease by a factor of 2.

R8.10 (a) When we use the sample standard deviation s_x to estimate the population standard deviation σ. **(b)** The t distributions are wider than the standard Normal distribution and they have a slightly different shape with more area in the tails. **(c)** As the degrees of freedom increase, the spread and shape of the t distributions become more like the standard Normal distribution.

Answers to Chapter 8 AP® Statistics Practice Test

T8.1 a
T8.2 d
T8.3 c
T8.4 d
T8.5 b
T8.6 a
T8.7 c
T8.8 d
T8.9 e
T8.10 d
T8.11 (a) S: p $=$ the true proportion of all visitors to Yellowstone who would say they favor the restrictions. P: One-sample z interval for p. Random: the visitors were selected randomly. 10%: the sample size (150) is less than 10% of all visitors to Yellowstone National Park. Large Counts: $n\hat{p} = 89 \geq 10$ and $n(1 - \hat{p}) = 61 \geq 10$. D: (0.490, 0.696). C: We are 99% confident that the interval from 0.490 to 0.696 captures the true proportion of all visitors who would say that they favor the restrictions. **(b)** Because there are values smaller than 0.50 in the confidence interval, the U.S.

Forest Service cannot conclude that more than half of visitors to Yellowstone National Park favor the proposal.
T8.12 (a) Because the sample size is large ($n = 48 \geq 30$), the Normal/Large Sample condition is met. **(b)** Maurice's interval uses a z critical value instead of a t critical value. Also, Maurice used the wrong value in the square root—it should be $n = 48$. *Correct:*

Using df $= 40, 6.208 \pm 2.021\left(\dfrac{2.576}{\sqrt{48}}\right) = (5.457, 6.959)$. *Using*

technology: (5.46, 6.956) with df $= 47$.
T8.13 S: μ = the true mean number of bacteria per milliliter of raw milk received at the factory. P: One-sample t interval for μ. Random: The data come from a random sample. 10%: the sample size (10) is less than 10% of all 1-ml specimens that arrive at the factory. Normal/Large Sample: the dotplot shows that there is no strong skewness or outliers.
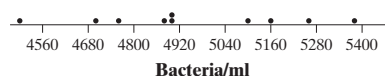
Bacteria/ml

D: Using df $= 9, (4794.37, 5105.63)$. C: We are 90% confident that the interval from 4794.37 to 5105.63 bacterial/ml captures the true mean number of bacteria in the milk received at this factory.

Chapter 9
Section 9.1
Answers to Check Your Understanding
page 541: **1. (a)** p = proportion of all students at Jannie's high school who get less than 8 hours of sleep at night. **(b)** $H_0{:}p = 0.85$ and $H_a{:}p \neq 0.85$. **2. (a)** μ = true mean amount of time that it takes to complete the census form. **(b)** $H_0{:}\mu = 10$ and $H_a{:}\mu > 10$.
page 549: **1.** Finding convincing evidence that the new batteries last longer than 30 hours on average, when in reality their true mean lifetime is 30 hours. **2.** Not finding convincing evidence that the new batteries last longer than 30 hours on average, when in reality their true mean lifetime >30 hours. **3.** Answers will vary. A consequence of a Type I error would be that the company spends the extra money to produce these new batteries when they aren't any better than the older, cheaper type. A consequence of a Type II error would be that the company would not produce the new batteries, even though they were better.

Answers to Odd-Numbered Section 9.1 Exercises
9.1 $H_0{:}\mu = 115$; $H_a{:}\mu > 115$, where μ is the true mean score on the SSHA for all students at least 30 years of age at the teacher's college.
9.3 $H_0{:}p = 0.12$; $H_a{:}p \neq 0.12$, where p is the true proportion of lefties at his large community college.
9.5 $H_0{:} \sigma = 3$; $H_a{:} \sigma > 3$, where σ is the true standard deviation of the temperature in the cabin.
9.7 The null hypothesis is always that there is "no difference" or "no change" and the alternative hypothesis is what we suspect is true. Correct: $H_0{:}p = 0.37$; $H_a{:}p > 0.37$.
9.9 Hypotheses are always about population parameters. Correct: $H_0{:}\mu = 1000$ grams; $H_a{:}\mu < 1000$ grams.
9.11 (a) The attitudes of older students do not differ from other students, on average. **(b)** Assuming the mean score on the SSHA for students at least 30 years of age at this school is really 115, there is a 0.0101 probability of getting a sample mean of at least 125.7 just by chance in an SRS of 45 older students.

9.13 $\alpha = 0.10$: Because the P-value of $0.2184 > \alpha = 0.10$, we fail to reject H_0. We do not have convincing evidence that the proportion of left-handed students at Simon's college is different from the national proportion. $\alpha = 0.05$: Because the P-value of $0.2184 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the proportion of left-handed students at Simon's college is different from the national proportion.

9.15 $\alpha = 0.05$: Because the P-value of $0.0101 < \alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean score on the SSHA for all students at least 30 years of age at the teacher's college >115. $\alpha = 0.01$: Because the P-value of $0.0101 > \alpha = 0.01$, we fail to reject H_0. We do not have convincing evidence that the true mean score on the SSHA for all students at least 30 years of age at the teacher's college >115.

9.17 Either H_0 is true or H_0 is false—it isn't true some of the time and not true at other times.

9.19 The P-value should be compared with a significance level (such as $\alpha = 0.05$), not the hypothesized value of p. Also, the data never "prove" that a hypothesis is true, no matter how large or small the P-value.

9.21 (a) $H_0\!:\!\mu = 6.7$; $H_a\!:\!\mu < 6.7$, where μ represents the mean response time for all accidents involving life-threatening injuries in the city. (b) I: Finding convincing evidence that the mean response time has decreased when it really hasn't. A consequence is that the city may not investigate other ways to reduce the mean response time and more people could die. II: Not finding convincing evidence that the mean response time has decreased when it really has. A consequence is that the city spends time and money investigating other methods to reduce the mean response time when they aren't necessary. (c) Type I, because people may end up dying as a result.

9.23 (a) $H_0\!:\!\mu = \$85,000$; $H_a\!:\!\mu > \$85,000$, where μ = the mean income of all residents near the restaurant. (b) I: Finding convincing evidence that the mean income of all residents near the restaurant exceeds $85,000 when in reality it does not. The consequence is that you will open your restaurant in a location where the residents will not be able to support it. II: Not finding convincing evidence that the mean income of all residents near the restaurant exceeds $85,000 when in reality it does. The consequence of this error is that you will not open your restaurant in a location where the residents would have been able to support it and you lose potential income.

9.25 d

9.27 c

9.29 (a) $P(\text{woman}) = 0.4168$, so $(24,611)(0.4168) = 10,258$ degrees were awarded to women. (b) No. $P(\text{woman}) = 0.4168$, which is not equal to $P(\text{woman} \mid \text{bachelors}) = 0.43$. (c) $P(\text{at least 1 of the 2 degrees earned by a woman})$
$= 1 - P(\text{neither degree is earned by a woman}) =$
$1 - \left(\dfrac{14,353}{24,611}\right)\left(\dfrac{14,352}{24,610}\right) = 0.6599$

Section 9.2

Answers to Check Your Understanding

page 560: S: $H_0\!:\!p = 0.20$ versus $H_a\!:\!p > 0.20$, where p is the true proportion of all teens at the school who would say they have electronically sent or posted sexually suggestive images of themselves. P: One-sample z test for p. Random: Random sample. 10%: The sample size $(250) < 10\%$ of the 2800 students. Large

Counts: $250(0.2) = 50 \geq 10$ and $250(0.8) = 200 \geq 10$. D: $z = \dfrac{0.252 - 0.20}{\sqrt{\dfrac{0.20(0.80)}{250}}} = 2.06$ and $P(Z \geq 2.06) = 0.0197$. C: Because the P-value of $0.0197 < \alpha = 0.05$, we reject H_0. We have convincing evidence that more than 20% of the teens in her school would say they have electronically sent or posted sexually suggestive images of themselves.

page 563: S: $H_0\!:\!p = 0.75$ versus $H_a\!:\!p \neq 0.75$, where p is the true proportion of all restaurant employees at this chain who would say that work stress has a negative impact on their personal lives. P: One-sample z test for p. Random: Random sample. 10%: The sample size $(100) < 10\%$ of all employees. Large Counts: $100(0.75) = 75 \geq 10$ and $100(0.25) = 25 \geq 10$.

D: $z = \dfrac{0.68 - 0.75}{\sqrt{\dfrac{0.75(0.25)}{100}}} = -1.62$ and $2P(Z \leq -1.62) = 0.1052$. C: Because the P-value of $0.1052 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true proportion of all restaurant employees at this large restaurant chain who would say that work stress has a negative impact on their personal lives is different from 0.75.

page 564: The confidence interval given in the output includes 0.75, which means that 0.75 is a plausible value for the population proportion that we are seeking. So both the significance test (which didn't rule out 0.75 as the proportion) and the confidence interval give the same conclusion. The confidence interval, however, gives a range of plausible values for the population proportion instead of only making a decision about a single value.

page 569: **1.** A Type II error. If a Type I error occurred, they would reject a good shipment of potatoes and have to wait to get a new delivery. However, if a Type II error occurred, they would accept a bad batch and make potato chips with blemishes. This might upset consumers and decrease sales. To minimize the probability of a Type II error, choose a large significance level such as $\alpha = 0.10$ **2.** (a) Increase. Increasing α to 0.10 makes it easier to reject the null hypothesis, which increases power. (b) Decrease. Decreasing the sample size means we don't have as much information to use when making the decision, which makes it less likely to correctly reject H_0. (c) Decrease. It is harder to detect a difference of $0.02\ (0.10 - 0.08)$ than a difference of $0.03\ (0.11 - 0.08)$.

Answers to Odd-Numbered Section 9.2 Exercises

9.31 Random: Random sample. 10%: The sample size $(60) < 10\%$ of all students. Large Counts: $60(0.80) = 48 \geq 10$ and $60(0.20) = 12 \geq 10$.

9.33 $np_0 = 10(0.5) = 5$ and $n(1 - p_0) = 10(0.5) = 5$ are both <10.

9.35 (a) $z = \dfrac{0.683 - 0.80}{\sqrt{\dfrac{0.80(0.20)}{60}}} = -2.27$ (b) $P(Z \leq -2.27) = 0.0116$.
Using technology: `normalcdf (lower:−1000, upper: −2.27, μ:0, σ:1) = 0.0116`. The graph is given below.

N(0, 1)

−2.27   0

z
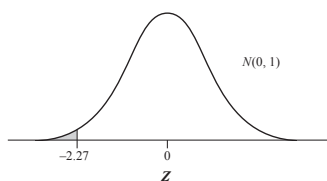
9.37 **(a)** P-value = 0.0143. 5%: Because the P-value of 0.0143 < α = 0.05, we reject H_0. There is convincing evidence that $p > 0.5$. 1%: Because the P-value of 0.0143 > α = 0.01, we fail to reject H_0. There is not convincing evidence that $p > 0.5$. **(b)** P-value = 0.0286. Because this P-value is still less than α = 0.05 and greater than α = 0.01, we would again reject H_0 at the 5% significance level and fail to reject H_0 at the 1% significance level.

9.39 S: H_0:p = 0.37 versus H_a:p > 0.37, where p = true proportion of all students who are satisfied with the parking situation after the change. P: One-sample z test for p. Random: Random sample. 10%: The sample size (200) < 10% of the population of size 2500. Large Counts: 200(0.37) = 74 ≥ 10 and 200(0.63) = 126 ≥ 10. D: z = 1.32, P-value = 0.0934. C: Because the P-value of 0.0934 > α = 0.05, we fail to reject H_0. We do not have convincing evidence that the true proportion of all students who are satisfied with the parking situation after the change > 0.37.

9.41 **(a)** S: H_0:p = 0.50 versus H_a:p > 0.50, where p is the true proportion of boys among first-born children. P: One-sample z test for p. Random: Random sample. 10%: The sample size (25,468) < 10% of all first-borns. Large Counts: 25,468(0.50) = 12,734 ≥ 10 and 25,468(0.50) = 12,734 ≥ 10. D: z = 5.49, P-value≈0. C: Because the P-value of approximately 0 < α = 0.05, we reject H_0. There is convincing evidence that first-born children are more likely to be boys. **(b)** First-born children, because that is the group that we sampled from.

9.43 Here are the corrections: H_a: $p > 0.75$; p = the true proportion of middle school students who engage in bullying behavior; 10%: the sample size (558) < 10% of the population of middle school students; np_0 = 558(0.75) = 418.5 ≥ 10 and $n(1 - p_0)$ = 558(0.25) = 139.5 ≥ 10; $z = \dfrac{0.7975 - 0.75}{\sqrt{\dfrac{(0.75)(0.25)}{558}}}$ = 2.59;

P-value = 0.0048. Because the P-value of 0.0048 < α = 0.05, we reject H_0. We have convincing evidence that more than three-quarters of middle school students engage in bullying behavior.

9.45 S: H_0:p = 0.60 versus H_a:p ≠ 0.60, where p is the true proportion of teens who pass their driving test on the first attempt. P: One-sample z test for p. Random: Random sample. 10%: The sample size (125) < 10% of all teens. Large Counts: 125(0.60) = 75 ≥ 10 and 125(0.40) = 50 ≥ 10. D: z = 2.01, P-value = 0.0444. C: Because our P-value of 0.0444 < α = 0.05, we reject H_0. There is convincing evidence that the true proportion of teens who pass the driving test on their first attempt is different from 0.60.

9.47 **(a)** D: (0.607,0.769). C: We are 95% confident that the interval from 0.607 to 0.769 captures the true proportion of teens who pass the driving test on the first attempt. **(b)** Because 0.60 is not in the interval, we have convincing evidence that the true proportion of teens who pass the driving test on their first attempt is different from 0.60.

9.49 No. Because the value 0.16 is included in the interval, we do not have convincing evidence that the true proportion of U.S. adults who would say they use Twitter differs from 0.16.

9.51 **(a)** p = the true proportion of U.S. teens aged 13 to 17 who think that young people should wait to have sex until marriage. **(b)** Random: Random sample. 10%: The sample size (439) < 10% of the population of all U.S. teens. Large Counts: 439(0.5) = 219.5 ≥ 10 and 439(0.5) = 219.5 ≥ 10. **(c)** Assuming that the true proportion of U.S. teens aged 13 to 17 who think that young people should wait to have sex until marriage is 0.50, there is a 0.011 probability of getting a sample proportion that is at least as different from 0.5 as the proportion in the sample. **(d)** Yes. Because the P-value of 0.011 < α = 0.05, we reject H_0. There is convincing evidence that the true proportion of U.S. teens aged 13 to 17 who think that young people should wait to have sex until marriage differs from 0.5.

9.53 **(a)** I: Finding convincing evidence that more than 37% of students were satisfied with the new parking arrangement, when in reality only 37% were satisfied. *Consequence:* The principal believes that students are satisfied and takes no further action. II: Failing to find convincing evidence that more than 37% are satisfied with the new parking arrangement, when in reality more than 37% are satisfied. *Consequence:* The principal takes further action on parking when none is needed. **(b)** If the true proportion of students that are satisfied with the new arrangement is really 0.45, there is a 0.75 probability that the survey provides convincing evidence that the true proportion > 0.37. **(c)** Increase the sample size or significance level.

9.55 P(Type I) = α = 0.05 and P(Type II) = 0.22.

9.57 **(a)** If the true proportion of Alzheimer's patients who would experience nausea is really 0.08, there is a 0.29 probability that the results of the study would provide convincing evidence that the true proportion < 0.10. **(b)** Increase the number of measurements taken (n) to get more information. **(c)** Decrease. If α is smaller, it becomes harder to reject the null hypothesis. This makes it harder to correctly reject H_0. **(d)** Increase. Because 0.07 is further from the null hypothesis value of 0.10, it will be easier to detect a difference between the null value and actual value.

9.59 c

9.61 b

9.63 **(a)** $X - Y$ has a Normal distribution with mean $\mu_{X-Y} = -0.2$ and standard deviation $\sigma_{X-Y} = \sqrt{(0.1)^2 + (0.05)^2} = 0.112$. To fit in a case, $X - Y$ must take on a negative number. **(b)** We want to find $P(X - Y < 0)$ using the $N(-0.2, 0.112)$ distribution. $z = \dfrac{0 - (-0.2)}{0.112} = 1.79$ and $P(Z < 1.79) = 0.9633$. *Using technology*: 0.9629. There is a 0.9629 probability that a randomly selected CD will fit in a randomly selected case. **(c)** P(all fit) = $(0.9629)^{100}$ = 0.0228. There is a 0.0228 probability that all 100 CDs will fit in their cases.

Section 9.3

Answers to Check Your Understanding

page 579: **1.** H_0:μ = 320 versus H_a:μ ≠ 320, where μ = the true mean amount of active ingredient (in milligrams) in Aspro tablets from this batch of production. **2.** Random: Random sample. 10%: The sample of size 36 < 10% of the population of all tablets in this batch. Normal/Large Sample: n = 36 ≥ 30. **3.** $t = \dfrac{319 - 320}{3/\sqrt{36}} = -2$ **4.** For this test, df = 35. Using Table B and df = 30, the tail area is between 0.025 and 0.05. Thus, the P-value for the two-sided test is between 0.05 and 0.10. *Using technology*: `2tcdf(lower:−1000,upper:−2,df:35)` =

$2(0.0267) = 0.0534$. Because the P-value of $0.0534 > \alpha = 0.05$, we fail to reject H_0. There is not convincing evidence that the true mean amount of the active ingredient in Aspro tablets from this batch of production differs from 320 mg.

page 583: **1.** S: $H_0: \mu = 8$ versus $H_a: \mu < 8$, where μ is the true mean amount of sleep that students at the professor's school get each night. P: One-sample t test for μ. Random: Random sample. 10%: The sample size $(28) < 10\%$ of the population of students. Normal/Large Sample: The histogram below indicates that there is not much skewness and no outliers.

Sleep (hours)
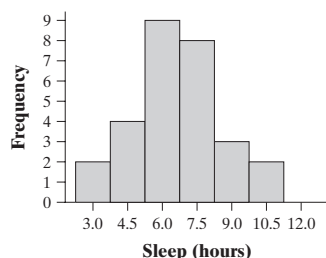
D: $\bar{x} = 6.643$ and $s_x = 1.981$. $t = -3.625$ and the P-value is between 0.0005 and 0.001. *Using technology*: P-value $= 0.0006$. C: Because our P-value of $0.0006 < \alpha = 0.05$, we reject H_0. There is convincing evidence that students at this university get less than 8 hours of sleep, on average.

page 586: **1.** S: $H_0: \mu = 128$ versus $H_a: \mu \neq 128$, where μ is the true mean systolic blood pressure for the company's middle-aged male employees. P: One-sample t test for μ. Random: Random sample. 10%: The sample size $(72) < 10\%$ of the population of middle-aged male employees. Normal/Large Sample: $n = 72 \geq 30$. D: $t = 1.10$ and P-value $= 0.275$. C: Because our P-value of $0.275 > \alpha = 0.05$, we fail to reject H_0. There is not convincing evidence that the mean systolic blood pressure for this company's middle-aged male employees differs from the national average of 128. **2.** We are 95% confident that the interval from 126.43 to 133.43 captures the true mean systolic blood pressure for the company's middle-aged male employees. The value of 128 is in this interval and therefore is a plausible mean systolic blood pressure for the males 35 to 44 years of age.

page 589: S: $H_0: \mu_d = 0$ versus $H_a: \mu_d > 0$, where μ_d is the true mean difference (air – nitrogen) in pressure lost. P: Paired t test for μ_d. Random: Treatments were assigned at random to each pair of tires. Normal/Large Sample: $n = 31 \geq 30$. D: $\bar{x} = 1.252$ and $s_x = 1.202$. $t = 5.80$ and P-value ≈ 0. C: Because the P-value of approximately $0 < \alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean difference in pressure (air – nitrogen) > 0. In other words, we have convincing evidence that tires lose less pressure when filled with nitrogen than when filled with air, on average.

Answers to Odd-Numbered Section 9.3 Exercises

9.65 Random: Random sample. 10%: The sample size $(45) < 10\%$ of the population size of 1000. Normal/Large Sample: $n = 45 \geq 30$.

9.67 The Random condition may not be met, because we don't know if this is a random sample of the atmosphere in the Cretaceous era. Also, the Normal/Large Sample condition is not met. The sample size < 30 and the histogram below shows that the data are strongly skewed to the left.

Nitrogen (percent)
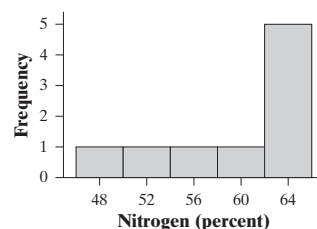
9.69 (a) $t = \dfrac{125.7 - 115}{29.8/\sqrt{45}} = 2.409$. (b) For this test, df $= 44$. Using Table B and df $= 40$, we have $0.01 < P$-value < 0.02. *Using technology*: $\texttt{tcdf(lower:2.409, upper:1000, df:44)} = 0.0101$.

9.71 (a) Using Table B and df $= 19$, we have $0.025 < P$-value < 0.05. *Using technology*: P-value $= 0.043$. 5%: Because the P-value of $0.043 < \alpha = 0.05$, we reject H_0. There is convincing evidence that $\mu < 5$. 1%: Because the P-value of $0.043 > \alpha = 0.01$, we fail to reject H_0. There is not convincing evidence that $\mu < 5$. (b) *Using technology*: P-value $= 0.086$. 5%: Because the P-value of $0.086 > \alpha = 0.05$, we fail to reject H_0. There is not convincing evidence that $\mu \neq 5$. 1%: same as part (a).

9.73 (a) S: $H_0: \mu = 25$ versus $H_a: \mu > 25$, where μ is the true mean speed of all drivers in a construction zone. P: One-sample t test for μ. Random: Random sample. 10%: The sample size $(10) < 10\%$ of all drivers. Normal/Large Sample: There is no strong skewness or outliers in the sample.
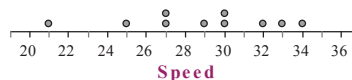
Speed

D: $\bar{x} = 28.8$ and $s_x = 3.94$. $t = 3.05$, df $= 9$, and the P-value is between 0.005 and 0.01 (0.0069). C: Because the P-value of $0.0069 < \alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean speed of all drivers in the construction zone > 25 mph. (b) Because we rejected H_0, it is possible we made a Type I error—finding convincing evidence that the true mean speed > 25 mph when it really isn't.

9.75 (a) S: $H_0: \mu = 1200$ versus $H_a: \mu < 1200$, where μ is the true mean daily calcium intake of women 18 to 24 years of age. P: One-sample t test for μ. Random: Random sample. 10%: The sample size $(36) < 10\%$ of all women aged 18 to 24. Normal/Large Sample: $n = 36 \geq 30$. D: $t = -6.73$ and P-value $= 0.000$. C: Because the P-value of approximately $0 < \alpha = 0.05$, we reject H_0. There is convincing evidence that women aged 18 to 24 are getting less than 1200 mg of calcium daily, on average. (b) Assuming that women aged 18 to 24 get 1200 mg of calcium per day, on average, there is about a 0 probability that we would observe a sample mean ≤ 856.2 mg by chance alone.

9.77 S: $H_0: \mu = 11.5$ versus $H_a: \mu \neq 11.5$, where μ is the true mean hardness of the tablets. P: One-sample t test for μ. Random: The tablets were selected randomly. 10%: The sample size $(20) < 10\%$ of all tablets in the batch. Normal/Large Sample: There is no strong skewness or outliers in the sample.

Hardness

D: $\bar{x} = 11.5164$ and $s_x = 0.0950$. $t = 0.77$, df $= 19$, and the *P*-value is between 0.40 and 0.50 (0.4494). *C:* Because our *P*-value of $0.4494 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true mean hardness of these tablets is different from 11.5.

9.79 *D:* With df $= 19$, (11.472,11.561). *C:* We are 95% confident that the interval from 11.472 to 11.561 captures the true mean hardness measurement for this type of pill. The confidence interval gives 11.5 as a plausible value for the true mean hardness μ, but it gives other plausible values as well.

9.81 *S:* $H_0{:}\mu = 200$ versus $H_a{:}\mu \neq 200$, where μ is the true mean response time of European servers. *P:* One-sample *t* interval to help us perform a two-sided test for μ. Random: The servers were selected randomly. 10%: The sample size (14) $< 10\%$ of all servers in Europe. Normal/Large Sample: The sample size is small, but a graph of the data reveals no strong skewness or outliers. *D:* (158.22, 189.64). *C:* Because our 95% confidence interval does not contain 200 milliseconds, we reject H_0 at the $\alpha = 0.05$ significance level. We have convincing evidence that the mean response time of European servers is different from 200 milliseconds.

9.83 (a) Yes. Because the *P*-value of $0.06 > \alpha = 0.05$, we fail to reject H_0: $\mu = 10$ at the 5% level of significance. Thus, the 95% confidence interval will include 10. (b) No. Because the *P*-value of $0.06 < \alpha = 0.10$, we reject $H_0{:}\mu = 10$ at the 10% level of significance. Thus, the 90% confidence interval would not include 10 as a plausible value.

9.85 (a) If all the subjects used the right thread first and they were tired when they used the left thread, then we wouldn't know if the difference in times was because of tiredness or because of the direction of the thread. (b) *S:* H_0: $\mu_d = 0$ versus H_a: $\mu_d > 0$, where μ_d is the true mean difference (left − right) in the time (in seconds) it takes to turn the knob with the left-hand thread and the right-hand thread. *P:* Paired *t* test for μ_d. Random: The order of treatments was determined at random. Normal/Large Sample: There is no strong skewness or outliers.

Difference in times
(left thread – right thread)
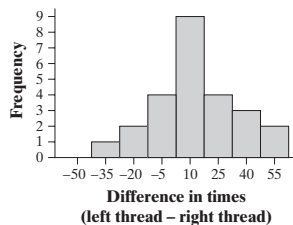
D: $\bar{x} = 13.32$ and $s_x = 22.94$. $t = 2.903$, df $= 24$, and the *P*-value is between 0.0025 and 0.005 (0.0039). *C:* Because the *P*-value of $0.0039 < \alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean difference (left − right) in time it takes to turn the knob > 0.

9.87 (a) $H_0{:}\mu_d = 0$ versus $H_a{:}\mu_d > 0$, where μ_d is the true mean difference in tomato yield (A − B). (b) df $= 9$. (c) *Interpretation*: Assuming that the average yield for both varieties is the same, there is a 0.1138 probability of getting a mean difference as large or larger than the one observed in this experiment. *Conclusion*: Because the *P*-value of $0.1138 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true mean difference in tomato yield (A − B) > 0. (d) I: Finding convincing evidence that Variety A tomato plants have a greater mean yield, when in reality there is no difference. II: Not finding convincing evidence that Variety A tomato plants have a higher mean yield, when in reality Variety A does have a greater mean yield. They might have made a Type II error.

9.89 Increase the significance level α or increase the sample size n.

9.91 When the sample size is very large, rejecting the null hypothesis is very likely, even if the actual parameter is only slightly different from the hypothesized value.

9.93 (a) No, in a sample of size $n = 500$, we expect to see about $(500)(0.01) = 5$ people who do better than random guessing, with a significance level of 0.01. (b) The researcher should repeat the procedure on these four to see if they again perform well.

9.95 b

9.97 d

9.99 c

9.101 a

9.103 (a) Not included. The margin of error does not account for undercoverage. (b) Not included. The margin of error does not account for nonresponse. (c) Included. The margin of error is calculated to account for sampling variability.

Answers to Chapter 9 Review Exercises

R9.1 (a) H_0: $\mu = 64.2$; H_a: $\mu \neq 64.2$, where $\mu =$ the true mean height of this year's female graduates from the local high school. (b) H_0: $p = 0.75$; H_a: $p < 0.75$, where $p =$ the true proportion of all students at Mr. Starnes's school who completed their math homework last night.

R9.2 Random: Random sample. 10%: The sample size (24) $< 10\%$ of the population of adults. Normal/Large Sample: The histogram below shows that the distribution is roughly symmetric with no outliers.

Time (seconds)
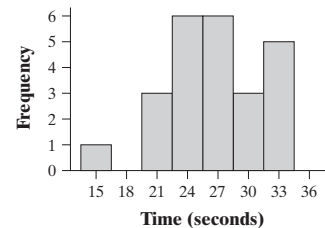
R9.3 (a) $H_0 : \mu = 300$ versus H_a: $\mu < 300$, where $\mu =$ the true mean breaking strength of these chairs. (b) I: Finding convincing evidence that the mean breaking strength <300 pounds, when in reality it is 300 pounds or higher. *Consequence*: falsely accusing the company of lying. II: Not finding convincing evidence that the mean breaking strength <300 pounds, when in reality it <300 pounds. *Consequence*: allowing the company to continue to sell chairs that don't work as well as advertised. (c) Because a Type II error is more serious, increase the probability of a Type I error by using $\alpha = 0.10$. (d) If the true mean breaking strength is 294 pounds, there is a 0.71 probability that we will find convincing

evidence that the true mean breaking strength < 300 pounds. **(e)** Increase the sample size or increase the significance level.
R9.4 (a) S: H_0: $p = 0.05$ versus H_a: $p < 0.05$, where p is the true proportion of adults who will get the flu after using the vaccine. P: One-sample z test for p. Random: Random sample. 10%: The sample size (1000) $< 10\%$ of the population of adults. Large Counts: $1000(0.05) = 50 \geq 10$ and $1000(0.95) = 950 \geq 10$. D: $z = -1.02$ and P-value $= 0.1539$. C: Because the P-value of $0.1539 > \a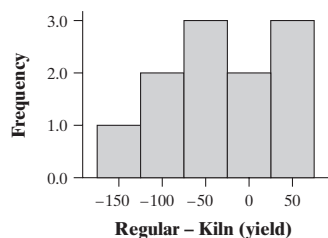lpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that fewer than 5% of adults who receive this vaccine will get the flu. **(b)** Because we failed to reject the null hypothesis, we could have made a Type II error—not finding convincing evidence that the true proportion of adults get the flu after using this vaccine <0.05, when in reality the true proportion <0.05. **(c)** Answers will vary.
R9.5 (a) Assuming that the roulette wheel is fair, there is a 0.0384 probability that we would get a sample proportion of reds at least this different from the expected proportion of reds (18/38) by chance alone. **(b)** Because the P-value of $0.0384 < \alpha = 0.05$, the results are statistically significant at the $\alpha = 0.05$ level. This means that we reject H_0 and have convincing evidence that the true proportion of reds is different than $p = 18/38$. **(c)** Because $18/38 = 0.474$ is one of the plausible values in the interval, this interval does not provide convincing evidence that the wheel is unfair. It does not, however, prove that the wheel is fair as there are many other plausible values in the interval that are not equal to 18/38. Also, the conclusion here is inconsistent with the conclusion in part (b) because the manager used a 99% confidence interval, which is equivalent to a test using $\alpha = 0.01$.
R9.6 (a) S: H_0:$\mu = 105$ versus H_a:$\mu \neq 105$, where μ is the true mean reading from radon detectors. P: One-sample t test for μ. Random: Random sample. 10%: The sample size (11) $< 10\%$ of all radon detectors. Normal/Large Sample: A graph of the data shows no strong skewness or outliers. D: $t = -0.06$, df $= 10$, and and P-value > 0.50 (0.9513). C: Because the P-value of $0.9513 > \alpha = 0.10$, we fail to reject H_0. We do not have convincing evidence that the true mean reading from the radon detectors is different than 105. **(b)** Yes. Because 105 is in the interval from 99.61 to 110.03, both the confidence interval and the significance test agree that 105 is a plausible value for the true mean reading from the radon detectors.
R9.7 (a) The random condition can be satisfied by randomly allocating which plot got the regular barley seeds and which one got the kiln-dried seeds within each pair of adjacent plots. **(b)** S: H_0: $\mu_d = 0$ versus H_a: $\mu_d < 0$, where μ_d is the true mean difference (regular $-$ kiln) in yield between regular barley seeds and kiln-dried barley seeds. P: Paired t test for μ_d. Random: Assumed. Normal/Large Sample: The histogram below shows no strong skewness or outliers.

Regular – Kiln (yield)

D: $\bar{x} = -33.7$ and $s_x = 66.2$. $t = -1.690$, df $= 10$, and the P-value is between 0.05 and 0.10 (0.0609). C: Because the P-value

of $0.0609 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true mean difference (regular $-$ kiln) in yield <0.

Answers to Chapter 9 AP® Statistics Practice Test

T9.1 b
T9.2 e
T9.3 c
T9.4 e
T9.5 b
T9.6 c
**T9.7** e
T9.8 d
T9.9 a
T9.10 c
T9.11 (a) S: H_0: $p = 0.20$ versus H_a: $p > 0.20$, where p is the true proportion of customers who would pay $100 for the upgrade. P: One-sample z test for p. Random: Random sample. 10%: The sample size (60) $< 10\%$ of this company's customers. Large Counts: $60(0.20) = 12 \geq 10$ and $60(0.8) = 48 \geq 10$. D: $z = 1.29$, P-value $= 0.0984$. C: Because the P-value of $0.0984 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that more than 20% of customers would pay $100 for the upgrade. **(b)** I: Finding convincing evidence that more than 20% of customers would pay for the upgrade, when in reality they would not. II: Not finding convincing evidence that more than 20% of customers would pay for the upgrade, when in reality more than 20% would. For the company, a Type I error is worse because they would go ahead with the upgrade and lose money. **(c)** Increase the sample size or increase the significance level.
T9.12 (a) Students may improve from Monday to Wednesday just because they have already done the task once. Then we wouldn't know if the experience with the test or the caffeine is the cause of the difference in scores. A better way to run the experiment would be to randomly assign half the students to get 1 cup of coffee on Monday and the other half to get no coffee on Monday. Then have each person do the opposite treatment on Wednesday. **(b)** S: H_0: $\mu_d = 0$ versus H_a: $\mu_d < 0$, where μ_d is the true mean difference (no coffee $-$ coffee) in the number of words recalled without coffee and with coffee. P: Paired t test for μ_d. Random: The treatments were assigned at random. Normal/Large Sample: The histogram below shows a symmetric distribution with no outliers.

No coffee – coffee
(words recalled)
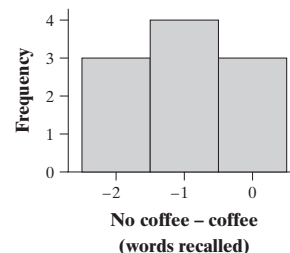
D: $\bar{x} = -1$ and $s_x = 0.816$. $t = -3.873$, df $= 9$, and the P-value is between 0.001 and 0.0025 (0.0019). C: Because the P-value of $0.0019 < \alpha = 0.05$, we reject H_0. We have convincing evidence that the mean difference (no coffee $-$ coffee) in word recall < 0.
T9.13 S: H_0: $\mu = \$158$ versus H_a: $\mu \neq \$158$, where μ is the true mean amount spent on food by households in this city. P: One-sample t test for μ. Random: Random sample. 10%: The sample size (50) $< 10\%$ of households in this small city. Normal/Large

Sample: $n = 50 \geq 30$. D: $t = 2.47$; using df $= 40$, the P-value is between 0.01 and 0.02 (using df $= 49$, 0.0168). C: Because the P-value of $0.0168 < \alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean amount spent on food per household in this city is different from the national average of \$158.

Chapter 10

Section 10.1

Answers to Check Your Understanding

page 619: S: $p_1 =$ true proportion of teens who go online every day and $p_2 =$ true proportion of adults who go online every day. P: Two-sample z interval for $p_1 - p_2$. Random: Independent random samples. 10%: $n_1 = 799 < 10\%$ of teens and $n_2 = 2253 < 10\%$ of adults. Large Counts: 503, 296, 1532, and 721 are all ≥ 10. D:

$$(0.63 - 0.68) \pm 1.645\sqrt{\frac{0.63(0.37)}{799} + \frac{0.68(0.32)}{2253}} =$$

$(-0.0824, -0.0176)$. C: We are 90% confident that the interval from -0.0824 to -0.0176 captures the true difference in the proportion of U.S. adults and teens who go online every day.

page 628: S: $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 > 0$, where p_1 is the true proportion of children like the ones in the study who do not attend preschool that use social services later and p_2 is the true proportion of children like the ones in the study who attend preschool that use social services later. P: Two-sample z test for $p_1 - p_2$. Random: Two groups in a randomized experiment. Large Counts: 49, 12, 38,

$$24 \text{ are all} \geq 10. \text{D}: z = \frac{(0.8033 - 0.6129) - 0}{\sqrt{\frac{0.7073(0.2927)}{61} + \frac{0.7073(0.2927)}{62}}} = 2.32$$

and P-value $= 0.0102$. C: Because the P-value of $0.0102 < \alpha = 0.05$, we reject H_0. There is convincing evidence that the true proportion of children like the ones in the study who do not attend preschool that use social services later is greater than the true proportion of children like the ones in the study who attend preschool that use social services later.

Answers to Odd-Numbered Section 10.1 Exercises

10.1 (a) Approximately Normal because $100(0.25) = 25$, $100(0.75) = 75$, $100(0.35) = 35$, and $100(0.65) = 65$ are all at least 10. **(b)** $\mu_{\hat{p}_1 - \hat{p}_2} = 0.25 - 0.35 = -0.10$. **(c)** Because $n_1 = 100 < 10\%$ of the first bag and $n_2 = 100 < 10\%$ of the second bag, $\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{0.25(0.75)}{100} + \frac{0.35(0.65)}{100}} = 0.0644$.

10.3 (a) Approximately Normal because $50(0.30) = 15$, $50(0.7) = 35$, $100(0.15) = 15$, and $100(0.85) = 85$ are all at least 10. **(b)** $\mu_{\hat{p}_C - \hat{p}_A} = 0.30 - 0.15 = 0.15$. **(c)** Because $n_C = 50 < 10\%$ of the jelly beans in the Child mix and $n_A = 100 < 10\%$ of the jelly beans in the Adult mix, $\sigma_{\hat{p}_C - \hat{p}_A} = \sqrt{\frac{0.3(0.7)}{50} + \frac{0.15(0.85)}{100}} = 0.0740$.

10.5 The data do not come from independent random samples or two groups in a randomized experiment. Also, there were less than 10 successes (3) in the group from the west side of Woburn.

10.7 There were less than 10 failures (0) in the treatment group, less than 10 successes (8) in the control group, and less than 10 failures in the control group (4).

10.9 (a) $\text{SE}_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{0.26(1 - 0.26)}{316} + \frac{0.14(1 - 0.14)}{532}} = 0.0289$.
If we were to take many random samples of 316 young adults and 532 older adults, the difference in the sample proportions of young adults and older adults who use Twitter will typically be 0.0289 from the true difference. **(b)** S: $p_1 =$ true proportion of young adults who use Twitter and $p_2 =$ true proportion of older adults who use Twitter. P: Two-sample z interval for $p_1 - p_2$. Random: Two independent random samples. 10%: $n_1 = 316 < 10\%$ of all young adults and $n_2 = 532 < 10\%$ of all older adults. Large Counts: 82, 234, 74, 458 are all at least 10. D: (0.072, 0.168). C: We are 90% confident that the interval from 0.072 to 0.168 captures the true difference in the proportions of young adults and older adults who use Twitter.

10.11 (a) S: $p_1 =$ true proportion of young men who live in their parents' home and $p_2 =$ true proportion of young women who live in their parents' home. P: Two-sample z interval for $p_1 - p_2$. Random: Reasonable to consider these independent random samples. 10%: $n_1 = 2253 < 10\%$ of the population of young men and $n_2 = 2629 < 10\%$ of the population of young women. Large Counts: 986, 1267, 923, 1706 are all at least 10. D: (0.051, 0.123). C: We are 99% confident that the interval from 0.051 to 0.123 captures the true difference in the proportions of young men and young women who live in their parents' home. **(b)** Because the interval does not contain 0, there is convincing evidence that the true proportion of young men who live in their parents' home is different from the true proportion of young women who live in their parents' home.

10.13 $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 \neq 0$, where p_1 is the true proportion of all teens who would say that they own an iPod or MP3 player and p_2 is the true proportion of all young adults who would say that they own an iPod or MP3 player.

10.15 P: Two-sample z test for $p_1 - p_2$. Random: Independent random samples. 10%: $n_1 = 800 < 10\%$ of all teens and $n_2 = 400 < 10\%$ of all young adults. Large Counts: 632, 168, 268, and 132 are all at least 10. D: $z = 4.53$ and P-value ≈ 0. C: Because the P-value of close to $0 < \alpha = 0.05$, we reject H_0. There is convincing evidence that the true proportion of teens who would say that they own an iPod or MP3 player is different from the true proportion of young adults who would say that they own an iPod or MP3 player.

10.17 D: (0.066, 0.174). C: We are 95% confident that the interval from 0.066 to 0.174 captures the true difference in proportions of teens and young adults who own iPods or MP3 players. Because 0 is not included in the interval, it is consistent with the results of Exercise 15.

10.19 S: $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 > 0$, where p_1 is the true proportion of 6- to 7-year-olds who would sort correctly and p_2 is the true proportion of 4- to 5-year-olds who would sort correctly. P: Two-sample z test for $p_1 - p_2$. Random: Independent random samples. 10%: $n_1 = 53 < 10\%$ of all 6- to 7-year olds and $n_2 = 50 < 10\%$ of all 4- to 5-year-olds. Large Counts: 28, 25, 10, 40 are all ≥ 10. D: $z = 3.45$ and P-value $= 0.0003$. C: Because the P-value of $0.0003 < \alpha = 0.05$, we reject H_0. We have convincing evidence that the true proportion of 6- to 7-year-olds who would sort correctly is greater than the true proportion of 4- to 5-year-olds who would sort correctly.

10.21 (a) S: $H_0: p_A - p_B = 0$ versus $H_a: p_A - p_B > 0$, where p_A is the true proportion of students like these who would pass the driver's license exam when taught by instructor A and p_B is the true

proportion of students like these who would pass the driver's license exam when taught by instructor B. *P*: Two-sample *z* test for $p_A - p_B$. Random: Two groups in a randomized experiment. Large Counts: 30, 20, 22, 28 are all $\geq$ 10. *D*: $z = 1.60$ and *P*-value = 0.0547. *C*: Because the *P*-value of 0.0547 $> \alpha = 0.05$, we fail to reject H_0. There is not convincing evidence that the true proportion of students like these who would pass using instructor A is greater than the true proportion who would pass using instructor B. **(b)** I: Finding convincing evidence that instructor A is more effective than instructor B, when in reality the instructors are equally effective. II: Not finding convincing evidence that instructor A is better, when in reality instructor A is more effective. It is possible we made a Type II error.

10.23 (a) Two-sample *z* test for $p_1 - p_2$. Random: Two groups in a randomized experiment. Large Counts: 44, 44, 21, 60 are all $\geq$ 10. **(b)** If no difference exists in the true pregnancy rates of women who are being prayed for and those who are not, there is a 0.0007 probability of getting a difference in pregnancy rates as large or larger than the one observed in the experiment by chance alone. **(c)** Because the *P*-value of 0.0007 $< \alpha = 0.05$, we reject H_0. There is convincing evidence that the pregnancy rates among women like these who are prayed for are higher than the pregnancy rates for those who are not prayed for. **(d)** Knowing they were being prayed for might have affected their behavior in some way that would have affected whether they became pregnant or not. Then we wouldn't know if it was the prayer or the other behaviors that caused the higher pregnancy rate.

10.25 a

10.27 c

10.29 (a) $\hat{y} = -13{,}832 + 14{,}954x$, where $\hat{y}$ = the predicted mileage and x = the age in years of the cars. **(b)** For each year older the car is, the predicted mileage will increase by 14,954 miles. **(c)** Residual = $-25{,}708$. The student's car had 25,708 fewer miles than expected, based on its age.

Section 10.2

Answers to Check Your Understanding

page 644: S:μ_1 = the true mean price of wheat in July and μ_2 = the true mean price of wheat in September. *P*: Two-sample *t* interval for $\mu_1 - \mu_2$. Random: Independent random samples. 10%: $n_1 = 90 < 10\%$ of all wheat producers in July and $n_2 = 45 < 10\%$ of all wheat producers in September. Normal/Large Sample: $n_1 = 90 \geq 30$ and $n_2 = 45 \geq 30$. *D*: Using df = 40, $(-0.759, -0.561)$. Using df = 100.45, $(-0.756, -0.564)$. *C*: We are 99% confident that the interval from -0.756 to -0.564 captures the true difference in mean wheat prices in July and September.

page 649: S: $H_0:\mu_1 - \mu_2 = 0$ versus $H_a:\mu_1 - \mu_2 > 0$, where μ_1 is the true mean breaking strength for polyester fabric buried for 2 weeks and μ_2 is the true mean breaking strength for polyester fabric buried for 16 weeks. *P*: Two-sample *t* test. Random: Two groups in a randomized experiment. Normal/Large Sample: The dotplots below show no strong skewness or outliers in either group.

2 weeks · · ·
16 weeks · · · · ·
102 108 114 120 126 132 138
Breaking strength
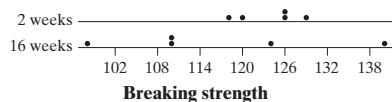
D: $\bar{x}_1 = 123.8$, $s_1 = 4.60$, $\bar{x}_2 = 116.4$, $s_2 = 16.09$. $t = 0.989$. Using df = 4, the *P*-value is between 0.15 and 0.20. Using df = 4.65,

P-value = 0.1857. *C*: Because the *P*-value of 0.1857 $> \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true mean breaking strength of polyester fabric that is buried for 2 weeks is greater than the true mean breaking strength for polyester fabric that is buried for 16 weeks.

Answers to Odd-Numbered Section 10.2 Exercises

10.31 (a) Because the distributions of *M* and *B* are Normal, the distribution of $\bar{x}_M - \bar{x}_B$ is also Normal. **(b)** $\mu_{\bar{x}_M - \bar{x}_B} = 188 - 170 = 18$ mg/dl. **(c)** Because $25 < 10\%$ of all 20- to 34-year-old males and $36 < 10\%$ of all 14-year-old boys, $\sigma_{\bar{x}_M - \bar{x}_B} = \sqrt{\dfrac{(41)^2}{25} + \dfrac{(30)^2}{36}} = 9.60$ mg/dl.

10.33 Random: Two independent random samples. 10%: $20 < 10\%$ of all males at the school and $20 < 10\%$ of all females at the school. Normal/Large Sample: not met because there are fewer than 30 observations in each group and the stemplot for Males shows several outliers.

10.35 Random: not met because these data are not from two *independent* random samples. Knowing the literacy percent for females in a country helps us predict the literacy percent for males in that country. 10%: not met because 24 is more than 10% of Islamic countries. Normal/Large Sample: not met because the samples sizes are both small and both distributions are skewed to the left and have an outlier (see boxplots below).


Female · |—[|]—|
Male · |—[|]—|
0   20   40   60   80   100
Percent literate
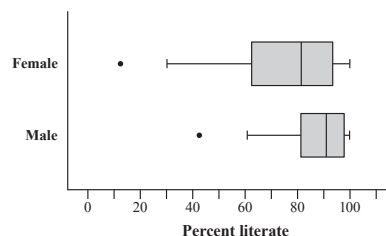
10.37 (a) The distributions of percent change are both slightly skewed to the left. People drinking red wine generally have more polyphenols in their blood, on average. The distribution of percent change for the white wine drinkers is a little bit more variable. **(b)** S: μ_1 = the true mean change in polyphenol level in the blood of people like those in the study who drink red wine and μ_2 = the true mean polyphenol level in the blood of people like those in the study who drink white wine. *P*: Two-sample *t* interval for $\mu_1 - \mu_2$. Random: Two groups in a randomized experiment. Normal/Large Sample: The dotplots given in the problem do not show strong skewness or outliers. *D*: $\bar{x}_1 = 5.5$, $s_1 = 2.517$, $\bar{x}_2 = 0.23$, $s_2 = 3.292$. Using df = 8, $(2.701, 7.839)$. Using df = 14.97, $(2.845, 7.689)$. *C*: We are 90% confident that the interval from 2.845 to 7.689 captures the true difference in mean change in polyphenol level for men like these who drink red wine and men like these who drink white wine. **(c)** Because all of the plausible values in the interval are positive, this interval supports the researcher's belief that red wine is more effective than white wine.

10.39 (a) Earnings amounts cannot be negative, yet the standard deviation is almost as large as the distance between the mean and 0. However, the sample sizes are both very large $(675 \geq 30 \text{ and } 621 \geq 30)$. **(b)** S: μ_1 = the true mean summer earnings of male students and μ_2 = the true mean summer earnings of female students. *P*: Two-sample *t* interval for $\mu_1 - \mu_2$. Random:

Reasonable to consider these independent random samples. 10%: $n_1 = 675 < 10\%$ of male students at a large university and $n_2 = 621 < 10\%$ of female students at a large university. Normal/Large Sample: $n_1 = 675 \geq 30$ and $n_2 = 621 \geq 30$. D: Using df = 100, (412.68, 635.58). Using df = 1249.21, (413.62, 634.64). C: We are 90% confident that the interval from \$413.62 to \$634.64 captures the true difference in mean summer earnings of male students and female students at this large university. (c) If we took many random samples of 675 males and 621 females from this university and each time constructed a 90% confidence interval in this same way, about 90% of the resulting intervals would capture the true difference in mean earnings for males and females.

10.41 (a) S: $H_0{:}\mu_1 - \mu_2 = 0$ versus $H_a{:}\mu_1 - \mu_2 < 0$, where μ_1 is the true mean time to breeding for the birds relying on natural food supply and μ_2 is the true mean time to breeding for birds with food supplementation. P: Two-sample t test. Random: Two groups in a randomized experiment. Normal/Large Sample: Neither distribution displays strong skewness or outliers.

Time to breeding
(days behind caterpillar peak)
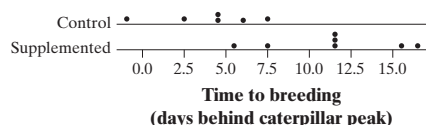
D: $\bar{x}_1 = 4.0$, $s_1 = 3.11$, $\bar{x}_2 = 11.3$, $s_2 = 3.93$. $t = -3.74$. Using df = 5, the P-value is between 0.005 and 0.01. Using df = 10.95, P-value = 0.0016. C: Because the P-value of 0.0016 < $\alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean time to breeding is less for birds relying on natural food supply than for birds with food supplements. (b) Assuming that the true mean time to breeding is the same for birds relying on natural food supply and birds with food supplements, there is a 0.0016 probability that we would observe a difference in sample means of -7.3 or smaller by chance alone.

10.43 S: H_0: $\mu_1 - \mu_2 = 0$ versus H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1 is the true mean number of words spoken per day by female students and μ_2 is the true mean number of words spoken per day by male students. P: Two-sample t test. Random: Independent random samples. 10%: $n_1 = 56 < 10\%$ of females at a large university and $n_2 = 56 < 10\%$ of males at a large university. Normal/Large Sample: $n_1 = 56 \geq 30$ and $n_2 = 56 \geq 30$. D: $t = -0.248$. Using df = 50, P-value > 0.50. Using df = 106.20, P-value = 0.8043. C: Because the P-value of 0.8043 > $\alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true mean number of words spoken per day by female students is different than the true mean number of words spoken per day by male students at this university.

10.45 (a) The distribution for the activities group is slightly skewed to the left, while the distribution for the control group is slightly skewed to the right. The center of the activities group is higher than the center of the control group. The scores in the activities group are less variable than the scores in the control group. (b) S: H_0: $\mu_1 - \mu_2 = 0$ versus H_a: $\mu_1 - \mu_2 > 0$, where μ_1 is the true mean DRP score for third-grade students like the ones in the experiment who do the activities and μ_2 is the true mean DRP score for third-grade students like the ones in the experiment who don't do the activities. P: Two-sample t test. Random: Two groups in a randomized experiment. Normal/Large Sample: No strong skewness or outliers in either boxplot. D: $t = 2.311$. Using df = 20, the P-value

is between 0.01 and 0.02. Using df = 37.86, P-value = 0.0132. C: Because the P-value of 0.0132 < $\alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean DRP score for third-grade students like the ones in the experiment who do the activities is greater than the true mean DRP score for third-grade students like the ones in the experiment who don't do the activities. (c) Because this was a randomized controlled experiment, we can conclude that the activities caused the increase in the mean DRP score.

10.47 D: Using df = 50, $(-3563, 2779)$. Using df = 106.2, $(-3521, 2737)$. C: We are 95% confident that the interval from -3521 to 2737 captures the true difference between mean number of words spoken per day by female students and the mean number of words spoken per day by male students. This interval allows us to determine if 0 is a plausible value for the difference in means and also provides other plausible values for the difference in mean words spoken per day.

10.49 (a) S: H_0: $\mu_1 - \mu_2 = 10$ versus H_a: $\mu_1 - \mu_2 > 10$, where μ_1 is the true mean cholesterol reduction for people like the ones in the study when using the new drug and μ_2 is the true mean cholesterol reduction for people like the ones in the study when using the current drug. P: Two-sample t test. Random: Two groups in a randomized experiment. Normal/Large Sample: No strong skewness or outliers. D: $t = 0.982$. Using df = 13, the P-value is between 0.15 and 0.20. Using df = 26.96, P-value = 0.1675. C: Because the P-value of 0.1675 > $\alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true mean cholesterol reduction is more than 10 mg/dl greater for the new drug than for the current drug. (b) Type II error. It is possible that the difference in mean cholesterol reduction is more than 10 mg/dl greater for the new drug than the current drug, but we didn't find convincing evidence that it was.

10.51 (a) The researchers randomly assigned the subjects to create two groups that were roughly equivalent at the beginning of the experiment. (b) Only about 5 out of the 1000 differences were ≥ 4.15, P-value ≈ 0.005. Because the P-value of 0.005 < $\alpha = 0.05$, we have convincing evidence that the true mean rating for students like these that are provided with internal reasons is higher than the true mean rating for students like these that are provided with external reasons. (c) Because we found convincing evidence that the mean is higher for students with internal reasons when it is possible that there is no difference in the means, we could have made a Type I error.

10.53 (a) Two-sample. Two distinct groups of cars in a randomized experiment. (b) Paired. Both treatments are applied to each subject. (c) Two-sample. Two distinct groups of women.

10.55 (a) Paired, because we have two scores for each student. (b) S: H_0: $\mu_d = 0$ versus H_a: $\mu_d > 0$, where μ_d is the true mean increase in SAT verbal scores of students who were coached. P: Paired t test for μ_d. Random: Random sample. 10%: $n_d = 427 < 10\%$ of students who are coached. Normal/Large Sample: $427 \geq 30$. D: $t = 10.16$. Using df = 426, P-value ≈ 0. C: Because the P-value of approximately 0 < $\alpha = 0.05$, we reject H_0. There is convincing evidence that students who are coached increase their scores on the SAT verbal test, on average.

10.57 a

10.59 b

10.61 (a) One-sample z interval for a proportion. (b) Paired t test for the mean difference. (c) Two-sample z interval for the difference in proportions. (d) Two-sample t test for a difference in means.

10.63 (a) P(at least one mean outside interval) $= 1 - P$(neither mean outside interval) $= 1-(0.95)^2 = 1-0.9025 = 0.0975$. **(b)** Let $X =$ the number of samples that must be taken to observe one falling above $\mu_{\bar{x}} + 2\sigma_{\bar{x}}$. Then X is a geometric random variable with $p = 0.025$. $P(X = 4) = (1 - 0.025)^3(0.025) = 0.0232$. **(c)** Let $X =$ the number of sample means out of 5 that fall outside this interval. X is a binomial random variable with $n = 5$ and $p = 0.32$. We want $P(X \geq 4) = 1 - P(X \leq 3) = 1 - $ binomcdf (trials:5, p:0.32,x value:3) $= 1 - 0.961 = 0.039$. This is a reasonable criterion because when the process is under control, we would only get a "false alarm" about 4% of the time.

10.65 (a) Perhaps the people who responded are prouder of their improvements and are more willing to share. This could lead to an overestimate of the true mean improvement. **(b)** This was an observational study, not an experiment. The students (or their parents) chose whether or not to be coached; students who choose coaching might have other motivating factors that help them do better the second time.

Answers to Chapter 10 Review Exercises

R10.1 (a) Paired t test for the mean difference. **(b)** Two-sample z interval for the difference in proportions. **(c)** One-sample t interval for the mean. **(d)** Two-sample t interval for the difference between two means.

R10.2 (a) $\text{SE}_{\hat{p}_1 - \hat{p}_2} = \sqrt{\dfrac{0.832(1 - 0.832)}{220} + \dfrac{0.581(1 - 0.581)}{117}}$

$= 0.0521$. If we were to take many random samples of 220 Hispanic female drivers in New York and 117 Hispanic female drivers in Boston, the difference in the sample proportions who wear seatbelts will typically be 0.0521 from the true difference in proportions of all Hispanic female drivers in New York and Boston who wear seat belts. **(b)** S: $p_1 =$ proportion of all Hispanic female drivers in New York who wear seat belts and $p_2 =$ proportion of all Hispanic female drivers in Boston who wear seat belts. P: Two-sample z interval for $p_1 - p_2$. Random: Independent random samples. 10%: $n_1 = 220 < 10\%$ of all Hispanic female drivers in New York and $n_2 = 117 < 10\%$ of all Hispanic female drivers in Boston. Large Counts: 183, 37, 68, 49 are all ≥ 10. D: $(0.149, 0.353)$. C: We are 95% confident that the interval from 0.149 to 0.353 captures the true difference in the proportions of Hispanic women drivers in New York and Boston who wear their seat belts.

R10.3 (a) The women in the study were randomly assigned to one of the two treatments. **(b)** Because both groups are large ($n_C = 45 \geq 30$ and $n_A = 45 \geq 30$), the sampling distribution of $\bar{x}_C - \bar{x}_A$ should be approximately Normal. **(c)** Assuming no difference exists in the true mean ratings of the product for women like these who read or don't read the news story, there is less than a 0.01 probability of observing a difference as large as or larger than 0.49 by chance alone.

R10.4 (a) S:$\mu_1 =$ the true mean NAEP quantitative skills test score for young men and $\mu_2 =$ the true mean NAEP quantitative skills test score for young women. P: Two-sample t interval for $\mu_1 - \mu_2$. Random: Reasonable to consider these independent random samples. 10%: $n_1 = 840 < 10\%$ of all young men and $n_2 = 1077 < 10\%$ of all young women. Normal/Large Sample: $n_1 = 840 \geq 30$ and $n_2 = 1077 \geq 30$. D: Using df $= 100$, $(-6.80, 2.14)$. Using df $= 1777.52$, $(-6.76, 2.10)$. C: We are 90% confident that the interval from -6.76 to 2.10 captures the true difference in the mean NAEP quantitative skills test score for young men and the

mean NAEP quantitative skills test score for young women. **(b)** Because 0 is in the interval, we do not have convincing evidence of a difference in mean score for male and female young adults.

R10.5 (a) S: $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 < 0$, where p_1 is the true proportion of patients like these who take AZT and develop AIDS and p_2 is the true proportion of patients like these who take placebo and develop AIDS. P: Two-sample z test for $p_1 - p_2$. Random: Two groups in a randomized experiment. Large Counts: $17, 418, 38, 397$ are all ≥ 10. D: $z = -2.91$, P-value $= 0.0018$. C: Because the P-value of $0.0018 < \alpha = 0.05$, we reject H_0. We have convincing evidence that taking AZT lowers the proportion of patients like these who develop AIDS compared to a placebo. **(b)** I: Finding convincing evidence that AZT lowers the risk of developing AIDS, when in reality it does not. Consequence: patients will pay for a drug that doesn't help. II: Not finding convincing evidence that AZT lowers the risk of developing AIDS, when in reality it does. Consequence: patients won't take the drug when it could actually delay the onset of AIDS. It is possible that we made a Type I error.

R10.6 (a) The Large Counts condition is not met because there are only 7 failures in the control area. **(b)** The Normal/Large Sample condition is not met because both sample sizes are small and there are outliers in the male distribution.

R10.7 (a) Even though each subject has two scores (before and after), the two groups of students are independent. **(b)** The distribution for the control group is slightly skewed to the right, while the distribution for the treatment group is roughly symmetric. The center for the treatment group is greater than the center for the control group. The differences in the control group are more variable than the differences in the treatment group. **(c)** S: $H_0: \mu_1 - \mu_2 = 0$ versus $H_a: \mu_1 - \mu_2 > 0$, where $\mu_1 =$ the true mean difference in test scores for students like these who get the treatment message and $\mu_2 =$ the true mean difference in test scores for students like these who get the neutral message. P: Two-sample t test for $\mu_1 - \mu_2$. Random: Two groups in a randomized experiment. Normal/Large Sample: Neither boxplot showed strong skewness or any outliers. D: Using the differences, $\bar{x}_1 = 11.4$, $s_1 = 3.169$, $\bar{x}_2 = 8.25$, $s_2 = 3.69$. $t = 1.91$. Using df $= 7$, the P-value is between 0.025 and 0.05. Using df $= 13.92$, P-value $= 0.0382$. C: Because the P-value of $0.0382 < \alpha = 0.05$, we reject H_0. There is convincing evidence that the true mean difference in test scores for students like these who get the treatment message is greater than the true mean difference in test scores for students like these who get the neutral message. **(d)** We cannot generalize to all students who failed the test because our sample was not a random sample of all students who failed the test.

Answers to Chapter 10 AP® Statistics Practice Test

T10.1 e
T10.2 b
T10.3 a
T10.4 a
T10.5 e
T10.6 e
**T10.7** c
T10.8 c
T10.9 b
T10.10 a
T10.11 (a) S: $\mu_1 =$ the true mean hospital stay for patients like these who get heating blankets during surgery and $\mu_2 =$ the true mean hospital stay for patients like these who have core temperatures reduced during surgery. P: Two-sample t interval for $\mu_1 - \mu_2$.

Random: Two groups in a randomized experiment. Normal/Large Sample: $n_1 = 104 \geq 30$ and $n_2 = 96 \geq 30$. D: Using df = 80, $(-4.17, -1.03)$. Using df = 165.12, $(-4.16, -1.04)$. C: We are 95% confident that the interval from -4.16 to -1.04 captures the true difference in mean length of hospital stay for patients like these who get heating blankets during surgery and those who have their core temperatures reduced during surgery. **(b)** Yes. Because 0 is not in the interval, we have convincing evidence that the true mean hospital stay for patients like these who get heating blankets during surgery is different than the true mean hospital stay for patients like these who have core temperatures reduced during surgery. **(c)** If we were to repeat this experiment many times and calculate 95% confidence intervals for the difference in means each time, about 95% of the intervals would capture the true difference in mean hospital stay for patients like these who get heating blankets during surgery and mean hospital stay for patients like these who have core temperatures reduced during surgery.

T10.12 **(a)** S: $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 > 0$, where p_1 is the true proportion of cars that have the brake defect in last year's model and p_2 is the true proportion of cars that have the brake defect in this year's model. P: Two-sample z test for $p_1 - p_2$. Random: Independent random samples. 10%: $n_1 = 100 < 10\%$ of last year's model and $n_2 = 350 < 10\%$ of this year's model. Large Counts: 20, 80, 50, 300 are all ≥ 10. D: $z = 1.39$, P-value $= 0.0822$. C: Because the P-value of $0.0822 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true proportion of brake defects is smaller in this year's model compared to last year's model. **(b)** I: Finding convincing evidence that there is a smaller proportion of brake defects in this year's car model, when in reality there is not. This might result in more accidents because people think that their brakes are safe. II: Not finding convincing evidence that there is a smaller proportion of brake defects in this year's model, when in reality there is a smaller proportion. This might result in reduced sales of this year's model.

T10.13 **(a)** $H_0: \mu_1 - \mu_2 = 0$ versus $H_a: \mu_1 - \mu_2 < 0$, where $\mu_1 =$ the true mean rent for one-bedroom apartments in the area of her college campus and $\mu_2 =$ the true mean rent for two-bedroom apartments in the area of her college campus. **(b)** Two-sample t test for $\mu_1 - \mu_2$. Random: Independent random samples. 10%: $n_1 = 10 < 10\%$ of all one-bedroom apartments in this area and $n_2 = 10 < 10\%$ of all two-bedroom apartments in this area. Normal/Large Sample: The dotplots below show no strong skewness or outliers in either distribution.
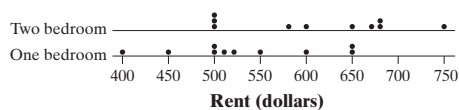

(c) Assuming the true mean rent of the two types of apartments is really the same, there is a 0.029 probability of getting an observed difference in mean rents as large as or larger than the one in this study. **(d)** Because the P-value of $0.029 < \alpha = 0.05$, Pat should reject H_0. She has convincing evidence that the true mean rent of two-bedroom apartments is greater than the true mean rent of one-bedroom apartments in the area of her college campus.

Answers to Cumulative AP® Practice Test 3

AP3.1 e
AP3.2 e
AP3.3 d

AP3.4 c
AP3.5 d
AP3.6 d
AP3.7 c
AP3.8 a
AP3.9 d
AP3.10 c
AP3.11 b
AP3.12 c
AP3.13 c
AP3.14 d
AP3.15 d
AP3.16 e
AP3.17 b
AP3.18 b
AP3.19 e
AP3.20 c
AP3.21 a
AP3.22 d
AP3.23 b
AP3.24 e
AP3.25 a
AP3.26 b
AP3.27 c
AP3.28 d
AP3.29 a
AP3.30 b
AP3.31 S: $H_0: \mu_d = 0$ versus $H_a: \mu_d < 0$, where μ_d is the true mean change in weight (after $-$ before) in pounds for people like these who follow a five-week crash diet. P: Paired t test for μ_d. Random: Random sample. 10%: $n_d = 15$ is less than 10% of all dieters. Normal/Large Sample: There is no strong skewness or outliers.
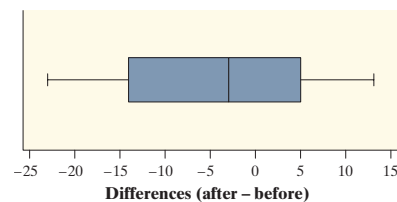
D: $\bar{x} = -3.6$ and $s_x = 11.53$. $t = -1.21$. Using df = 14, the P-value is between 0.10 and 0.15 (0.1232). C: Because the P-value of 0.1232 is greater than $\alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true mean change in weight (after $-$ before) for people like these who follow a five-week crash diet is less than 0.

AP3.32 **(a)** Observational study. No treatments were imposed on the individuals in the study. **(b)** $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 < 0$, where p_1 is the true proportion of VLBW babies who graduate from high school by age 20 and p_2 is the true proportion of non-VLBW babies who graduate from high school by age 20. P: Two-sample z test for $p_1 - p_2$. Random: Independent random samples. 10%: $n_1 = 242$ is less than 10% of all VLBW babies and $n_2 = 233$ is less than 10% of all non-VLBW babies. Large Counts: 179, 63, 193, 40 are all ≥ 10. Do: $z = -2.34$ and P-value $= 0.0095$. *Conclude:* Because the P-value of 0.0095 is less than $\alpha = 0.05$, we reject H_0. We have convincing evidence that the true proportion of VLBW babies who graduate from high school by

age 20 is less than the true proportion of non-VLBW babies who graduate from high school by age 20.

AP3.33 (a) $\hat{y} = -73.64 + 5.7188x$, where $\hat{y} =$ predicted distance and $x =$ temperature (degrees Celsius) (b) For each increase of 1°C in the water discharge temperature, the predicted distance from the nearest fish to the outflow pipe increases by about 5.7188 meters. (c) Yes. The residual plot shows no leftover pattern. (d) residual $= 78 - 92.21 = -14.21$ meters. The actual distance on this afternoon was 14.21 meters closer than expected, based on the temperature of the water.

AP3.34 (a) Define $W =$ the weight of a randomly selected gift box. Then $\mu_W = 8(2) + 2(4) + 3 = 27$ ounces and $\sigma_W = \sqrt{8(0.5^2) + 2(1^2) + 0.2^2} = 2.01$ ounces. (b) We want to find $P(W > 30)$ using the $N(27, 2.01)$ distribution. $z = \dfrac{30 - 27}{2.01} = 1.49$ and $P(W > 30) = 0.0681$. *Using technology:* 0.0678. There is a 0.0678 probability of randomly selecting a box that weighs more than 30 ounces. (c) P(at least one box is greater than 30 ounces) $= 1 - P$(none of the boxes is greater than 30 ounces) $= 1 - (1 - 0.0678)^5 = 1 - (0.9322)^5 = 0.2960$. (d) Because the distribution of W is Normal, the distribution of $\overline{W}$ will also be Normal, with mean $\mu_{\overline{W}} = 27$ ounces and standard deviation $\sigma_{\overline{W}} = \dfrac{2.01}{\sqrt{5}} = 0.899$. We want to find $P(\overline{W} > 30)$. $z = \dfrac{30 - 27}{0.899} = 3.34$ and $P(Z > 3.34) = 0.0004$. There is a 0.0004 probability of randomly selecting 5 boxes that have a mean weight of more than 30 ounces.

AP3.35 (a) S: $H_0: \mu_A - \mu_B = 0$ versus $H_a: \mu_A - \mu_B \neq 0$, where μ_A is the true mean annualized return for stock A and μ_B is the true mean annualized return for stock B. P: Two-sample t test. Random: Independent random samples. 10%: $n_A = 50$ is less than 10% of all days in the past 5 years and $n_B = 50$ is less than 10% of all days in the past 5 years. Normal/Large Sample: $n_A = 50 \geq 30$ and $n_B = 50 \geq 30$. D: $t = 2.07$. Using df $= 40$, the P-value is between 0.04 and 0.05. Using df $= 90.53$, P-value $= 0.0416$. C: Because the P-value of 0.0416 is less than $\alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean annualized return for stock A is different than the true mean annualized return for stock B. (b) $H_o: \sigma_A - \sigma_B = 0$ vs. $H_a: \sigma_A - \sigma_B > 0$, where σ_A is the true standard deviation of returns for stock A and σ_B is the true standard deviation of returns for stock B. (c) When the standard deviation of stock A is greater than the standard deviation of stock B, the variance of stock A will be bigger than the variance of stock B. Thus, values of F that are significantly greater than 1 would indicate that the price volatility for stock A is higher than that for stock B. (d)$F = \dfrac{(12.9)^2}{(9.6)^2} = 1.806$. (e) In the simulation, a test statistic of 1.806 or greater occurred in only 6 out of the 200 trials. Thus, the approximate P-value is $6/200 = 0.03$. Because the approximate P-value of 0.03 is less than $\alpha = 0.05$, we reject H_0. There is convincing evidence that the true standard deviation of returns for stock A is greater than the true standard deviation of returns for stock B.

Chapter 11
Section 11.1
Answers to Check Your Understanding

page 684: **1.** H_0: The company's claimed color distribution for its Peanut M&M'S is correct versus H_a: The company's claimed color distribution is not correct. **2.** The expected count of both blue and orange candies is $46(0.23) = 10.58$, for green and yellow is $46(0.15) = 6.9$, and for red and brown is $46(0.12) = 5.52$.

3. $\chi^2 = \dfrac{(12 - 10.58)^2}{10.58} + \dfrac{(7 - 10.58)^2}{10.58} + \dfrac{(13 - 6.9)^2}{6.9} + \dfrac{(4 - 6.9)^2}{6.9}$
$+ \dfrac{(8 - 5.52)^2}{5.52} + \dfrac{(2 - 5.52)^2}{5.52} = 11.3724$

page 687: **1.** The expected counts are all at least 5. df $= 6-1 = 5$.
2.
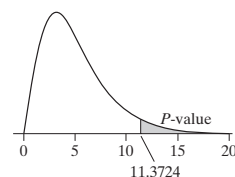
Chi-square distribution with 5 df

3. The P-value is between 0.025 and 0.05 (0.0445). **4.** Because the P-value of $0.0445 < \alpha = 0.05$, we reject H_0. There is convincing evidence that the color distribution of M&M'S® Peanut Chocolate Candies is different from what the company claims.
page 691: S: H_0: The distribution of eye color and wing shape is the same as what the biologists predict versus H_a: The distribution of eye color and wing shape is not what the biologists predict. P: Chi-square test for goodness of fit. Random: Random sample. 10%: $n = 200 < 10\%$ of all fruit flies. Large Counts: 112.5, 37.5, 37.5, 12.5 all ≥ 5. D: $\chi^2 = 6.1867$, df $= 3$, the P-value is between 0.10 and 0.15 (0.1029). C: Because the P-value of $0.1029 > \alpha = 0.01$, we fail to reject H_0. We do not have convincing evidence that the distribution of eye color and wing shape is different from what the biologists predict.

Answers to Odd-Numbered Section 11.1 Exercises

11.1 (a) H_0: The company's claimed distribution for its deluxe mixed nuts is correct versus H_a: The company's claimed distribution is not correct. (b) Cashews: $150(0.52) = 78$, almonds: $150(0.27) = 40.5$, macadamia nuts: $150(0.13) = 19.5$, brazil nuts: $150(0.08) = 12$.

11.3 $\chi^2 = \dfrac{(83 - 78)^2}{78} + \dfrac{(29 - 40.5)^2}{40.5} + \dfrac{(20 - 19.5)^2}{19.5} + \dfrac{(18 - 12)^2}{12}$
$= 6.599$

11.5 (a) Expected counts are all at least 5 and df $= 3$.
(b)
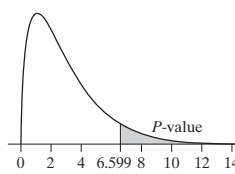
Chi-square distribution with 3 df

(c) The P-value is between 0.05 and 0.10 (0.0858). (d) Because the P-value of $0.0858 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the company's claimed distribution for its deluxe mixed nuts is incorrect.

11.7 S: H_0: Nuthatches do not prefer particular types of trees when searching for seeds and insects versus H_a: Nuthatches do prefer particular types of trees when searching for seeds and insects. P: Chi-square test for goodness of fit. Random: Random sample. 10%: $n = 156 < 10\%$ of all nuthatches. Large Counts: 84.24, 62.4, 9.36 all $\geq$ 5. D: $\chi^2 = 7.418$. With df = 2, the P-value is between 0.02 and 0.025 (0.0245). C: Because the P-value of $0.0245 < \alpha = 0.05$, we reject H_0. There is convincing evidence that nuthatches prefer particular types of trees when they are searching for seeds and insects.

11.9 Time spent doing homework is quantitative. Chi-square tests for goodness of fit should be used only for distributions of categorical data.

11.11 (a) S: H_0: The first digit of invoices from this company follow Benford's law versus H_a: The first digit of invoices from this company do not follow Benford's law. P: Chi-square test for goodness of fit. Random: Random sample. 10%: Assume $n = 250 < 10\%$ of all invoices from this company. Large Counts: 75.25, 44, 31.25, 24.25, 19.75, 16.75, 14.5, 12.75, 11.5 all $\geq$ 5. D: $\chi^2 = 21.563$. With df = 8, the P-value is between 0.005 and 0.01 (0.0058). C: Because the P-value of $0.0058 < \alpha = 0.05$, we reject H_0. There is convincing evidence that the first digit of invoices from this company do not follow Benford's law. *Follow-up analysis*: The largest contributors to the statistic are amounts with first digit 3, 4 and 7. There are more invoices that start with 3 or 4 than expected and fewer invoices that start with 7 than expected. (b) I: Finding convincing evidence that the company's invoices do not follow Benford's law (suggesting fraud), when in reality they are consistent with Benford's law. A consequence is falsely accusing this company of fraud. II: Not finding convincing evidence that the invoices do not follow Benford's law (suggesting fraud), when in reality they do not. A consequence is allowing this company to continue committing fraud. A Type I error would be more serious for the accountant.

11.13 (a) H_0: The true distribution of flavors for Skittles candies is the same as the company's claim versus H_a: The true distribution of flavors for Skittles candies is not the same as the company's claim. (b) Expected counts all = 12. (c) Using df = 4, χ^2 statistics greater than 9.49 would provide significant evidence at the $\alpha = 0.05$ level and χ^2 values greater than 13.28 would provide significant evidence at the $\alpha = 0.01$ level. (d) Answers will vary.

11.15 S: H_0: All 12 astrological signs are equally likely versus H_a: All 12 astrological signs are not equally likely. P: Chi-square test for goodness of fit. Random: Random sample. 10%: $n = 4344 < 10\%$ of all people in the United States. Large Counts: All expected counts = 362, which are $\geq$ 5. D: $\chi^2 = 19.76$. With df = 11, the P-value is between 0.025 and 0.05 (0.0487). C: Because the P-value of $0.0487 < \alpha = 0.05$, we reject H_0. There is convincing evidence that the 12 astrological signs are not equally likely. *Follow-up analysis*: The largest contributors to the statistic are Aries and Virgo. There are fewer Aries ($321 - 362 = -41$) and more Virgos ($402 - 362 = 40$) than we would expect.

11.17 S: H_0: Mendel's 3:1 genetic model is correct versus H_a: Mendel's 3:1 genetic model is not correct. P: Chi-square test for goodness of fit. Conditions are met. D: $\chi^2 = 0.3453$. With df = 1, the P-value > 0.25 (0.5568). C: Because the P-value of $0.5568 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that Mendel's 3:1 genetic model is wrong.

11.19 d

11.21 c

11.23 The distribution of English grades for the heavy readers is skewed to the left, while the distribution of English grades for the light readers is roughly symmetric. The center of the distribution of English grades is greater for the heavy readers than for the light readers. The English grades are more variable for the light readers. There is one low outlier in the heavy reading group but no outliers in the light reading group.

11.25 (a) For each additional book read, the predicted English GPA increases by about 0.024. The predicted English grade for a student who has read 0 books is about 3.42. (b) residual $= 2.85 - 3.828 = -0.978$. This student's English GPA is 0.978 less than predicted, based on the number of books this student has read. (c) Not very strong. On the scatterplot, the points are quite spread out from the line. Also, the value of r^2 is 0.083, which means that only 8.3% of the variation in English grades is accounted for by the linear model relating English GPA to number of books read.

Section 11.2

Answers to Check Your Understanding

page 699: **1.** Main: 0.060 several times a month or less, 0.236 at least once a week, 0.703 at least once a day. Commonwealth: 0.121 several times a month or less, 0.250 at least once a week, 0.628 at least once a day. **2.** Because there was such a big difference in the sample size from the two different types of campuses. **3.** Students on the main campus are more likely to be everyday users of Facebook. Also, those on the commonwealth campuses are more likely to use Facebook several times a month or less.
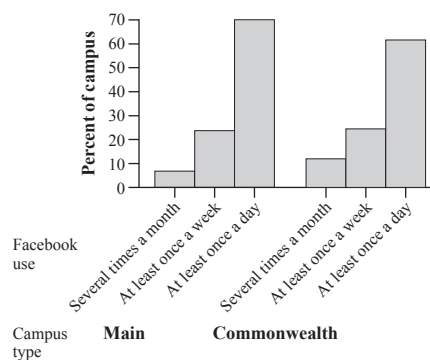
page 705: **1.** H_0: There is no difference in the distributions of Facebook use among students at the main campus and students at the commonwealth campuses versus H_a: There is a difference in the distributions of Facebook use among students at the main campus and students at the commonwealth campuses. **2.** $\dfrac{(131)(910)}{1537} = 77.56$, $\dfrac{(131)(627)}{1537} = 53.44$, $\dfrac{(372)(910)}{1537} = 220.25$, $\dfrac{(372)(627)}{1537} = 151.75$, $\dfrac{(1034)(910)}{1537} = 612.19$, $\dfrac{(1034)(627)}{1537} = 421.81$

3. $\chi^2 = \dfrac{(55 - 77.56)^2}{77.56} + \cdots + \dfrac{(394 - 421.81)^2}{421.81} = 19.49$

4. With df = 2, the P-value < 0.0005 (0.000059). **5.** Assuming that no difference exists in the distributions of Facebook use between students on Penn State's main campus and students at Penn State's commonwealth campuses, there is a 0.000059 probability of observing samples that show a difference in the distributions of Facebook

use among students at the main campus and the commonwealth campuses as large or larger than the one found in this study. **6.** Because the *P*-value of $0.000059 < \alpha = 0.05$, we reject H_0. There is convincing evidence that the distribution of Facebook use is different among students at Penn State's main campus and students at Penn State's commonwealth campuses.

page 711: **1.**
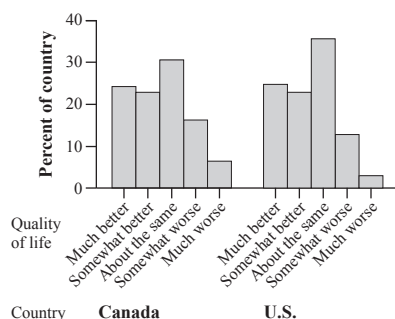
2. S: H_0: There is no difference in the distribution of quality of life for patients who have suffered a heart attack in Canada and the U.S. versus H_a: There is a difference. . . . P: Chi-square test for homogeneity. Random: Independent random samples. 10%: $n_1 = 311 < 10$ % of all Canadian heart attack patients and $n_2 = 2165 < 10\%$ of all U.S. heart attack patients. Large Counts: 77.37, 538.63, 71.47, 497.53, 109.91, 765.09, 41.70, 290.30, 10.55, 73.45 all ≥ 5. D: $\chi^2 = 11.725$. With df = 4, the *P*-value is between 0.01 and 0.02 (0.0195). C: Because the *P*-value of $0.0195 > \alpha = 0.01$, we fail to reject H_0. There is not convincing evidence that a difference exists in the distribution of quality of life for heart attack patients in Canada and the United States.

page 717: S: H_0: There is no association between an exclusive territory clause and business survival versus H_a: There is an association. . . . P: Chi-square test for independence. Random: Random sample. 10%: We assume that $n = 170 < 10\%$ of all new franchise firms. Large Counts: 102.74, 20.26, 39.26, 7.74 all ≥ 5. D: $\chi^2 = 5.911$. Using df = 1, the *P*-value is between 0.01 and 0.02 (0.0150). C: Because the *P*-value of $0.0150 > \alpha = 0.01$, we fail to reject H_0. There is not convincing evidence of an association between exclusive territory clause and business survival.

Answers to Odd-Numbered Section 11.2 Exercises

11.27 (a) Female: 0.209, 0.104, 0.313, 0.373. Male: 0.463, 0.269, 0.075, 0.194.
(b)
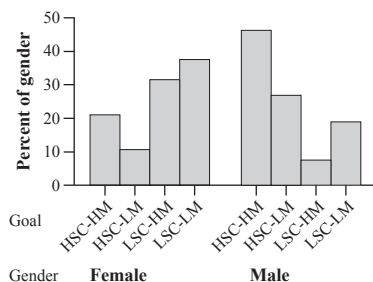
(c) In general, it appears that females were classified mostly as low social comparison, whereas males were classified mostly as high social comparison. However, about an equal percentage of males and females were classified as high mastery.

11.29 (a) H_0: There is no difference in the distribution of sports goals for male and female undergraduates at this university versus H_a: There is a difference. . . . (b) 22.5, 12.5, 13, 19, 22.5, 12.5, 13, 19. (c) $\chi^2 = 24.898$.

11.31 (a) Random: Independent random samples. 10%: $n_1 = 67 < 10\%$ of all males and $n_2 = 67 < 10\%$ of all females at the university. Large Counts: All expected counts ≥ 5. (b) With df = 3, the *P*-value < 0.0005 (0.000016). (c) Assuming that no difference exists in the distributions of goals for playing sports among males and females, there is a 0.000016 probability of observing independent random samples that show a difference in the distributions of goals for playing sports among males and females as large or larger than the one found in this study. (d) Because the *P*-value of $0.000016 < \alpha = 0.05$, we reject H_0. There is convincing evidence of a difference in the distribution of goals for playing sports among male and female undergraduates at this university.

11.33 (a) Cold: 0.593 hatched. Neutral: 0.679 hatched. Hot: 0.721 hatched. As the temperature warms up from cold to neutral to hot, the proportion of eggs that hatch appears to increase. (b) S: H_0: There is no difference in the true proportion of eggs that hatch in cold, neutral, or hot water versus H_a: There is a difference. . . . P: Chi-square test for homogeneity. Random: 3 groups in a randomized experiment. Large Counts: 18.63, 38.63, 71.74, 8.37, 17.37, 32.26 all ≥ 5. D: $\chi^2 = 1.703$. With df = 2, the *P*-value > 0.25 (0.4267). C: Because the *P*-value of $0.4267 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that there is a difference in the true proportions of eggs that hatch in cold, neutral, or hot water.

11.35 We do not have the actual counts of the travelers in each category. We also do not know if the sample was taken randomly or if the samples are independent.

11.37 (a) The data are given in the table below. The best success rate is for the patch plus the drug (0.355), followed by the drug alone (0.303). The patch alone (0.164) is just a little better than the placebo (0.156).

	Nicotine Patch	Drug	Patch plus drug	Placebo	Total
Success	40	74	87	25	**226**
Failure	204	170	158	135	**667**
Total	**244**	**244**	**245**	**160**	**893**

(b) Each of the four treatments has the same probability of success for smokers like these. (c) S: H_0: The true proportions of smokers like these who are able to quit for a year are the same for each of the four treatments versus H_a: The true proportions are not the same. . . . P: Chi-square test for homogeneity. Random: 4 groups in a randomized experiment. Large Counts: 61.75, 61.75, 62, 40.49, 182.25, 182.25, 183, 119.51 all ≥ 5. D: $\chi^2 = 34.937$. With df = 3, the *P*-value < 0.0005. C: Because the *P*-value of approximately $0 < \alpha = 0.05$, we reject H_0. There is convincing evidence that the true proportions of smokers like these who are able to quit for a year are not the same for each of the four treatments.

11.39 The largest component comes from those who had success using both the patch and the drug (25 more than expected). The next largest component comes from those who had success using just the patch (21.75 less than expected).

11.41 Buyers are much more likely to think the quality of recycled coffee filters is higher, while nonbuyers are more likely to think the quality is the same or lower.
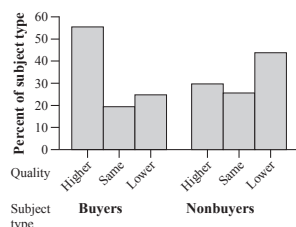
11.43 (a) H_0: There is no association between beliefs about the quality of recycled products and whether or not a person buys recycled products in the population of adults versus H_a: There is an association. . . . (b)13.26, 35.74, 8.66, 23.34, 14.08, 37.92 (c) $\chi^2 = 7.64$. With df = 2, the P-value is between 0.02 and 0.025 (0.022). (d) Because the P-value of $0.022 < \alpha = 0.05$, we reject H_0. There is convincing evidence of an association between beliefs about the quality of recycled products and whether or not a person buys recycled products in the population of adults.

11.45 S: H_0: There is no association between education level and opinion about a handgun ban in the adult population versus H_a: There is an association. . . . P: Chi-square test for independence. Random: Random sample. 10%: $n = 1201 < 10\%$ of all adults. Large Counts: 46.94, 86.19, 187.36, 94.29, 71.22, 69.06, 126.81, 275.64, 138.71, 104.78 all ≥ 5. D: $\chi^2 = 8.525$. With df = 4, the P-value is between 0.05 and 0.10 (0.0741). C: Because the P-value of $0.0741 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that there is an association between educational level and opinion about a handgun ban in the adult population.

11.47 (a) Independence, because the data come from a single random sample. (b) H_0: There is no association between gender and where people live in the population of young adults versus H_a: There is an association. . . . (c) Random: Random sample. 10% : $n = 4854 < 10\%$ of all young adults. Large Counts: The expected counts are all at least 5. (d) *P-value*: If no association exists between gender and where people live in the population of young adults, there is a 0.012 probability of getting a random sample of 4854 young adults with an association as strong or even stronger than the one found in this study. *Conclusion*: Because the P-value of $0.012 < \alpha = 0.05$, we reject H_0. There is convincing evidence that an association exists between gender and where people live in the population of young adults.

11.49 (a) *Hypotheses*: H_0: There is no difference in the improvement rates for patients like these who receive gastric freezing and those who receive the placebo versus H_a: There is a difference. . . . *P-value*: Assuming that no difference exists in the improvement rates between those receiving gastric freezing and those receiving the placebo, there is a 0.570 probability of observing a difference in improvement rates as large or larger than the difference observed in the study by chance alone. *Conclusion*: Because the P-value of 0.570 is larger than $\alpha = 0.05$, we fail to reject H_0. There is not convincing evidence that a difference exists in the improvement rates for patients like these who receive gastric freezing and those who receive the placebo. (b) The P-values are equal and $z^2 = (-0.57)^2 = 0.3249 \approx \chi^2 = 0.322$.

11.51 d

11.53 d

11.55 a

11.57 (a) One-sample t interval for a mean. (b) Two-sample z test for the difference between two proportions.

11.59 (a) Experiment, because a treatment (type of rating scale) was deliberately imposed on the students who took part in the study. (b) Several of the expected counts are less than 5.

Answers to Chapter 11 Review Exercises

R11.1 S: H_0: The proposed 1:2:1 genetic model is correct versus H_a: The proposed 1:2:1 genetic model is not correct. P: Chi-square test for goodness of fit. Random: Random sample. $10\% : n = 84 < 10\%$ of all yellow-green parent plants. Large Counts: 21, 42, 21 all ≥ 5. D: $\chi^2 = 6.476$. Using df = 2, the P-value is between 0.025 and 0.05 (0.0392). C: Because the P-value of $0.0392 > \alpha = 0.01$, we fail to reject H_0. We do not have convincing evidence that the proposed 1:2:1 genetic model is not correct.

R11.2 Several of the expected counts are less than 5.

R11.3 (a)

|	Stress management	Exercise	Usual care	Total
Suffered cardiac event	3	7	12	**22**
No cardiac event	30	27	28	**85**
Total	**33**	**34**	**40**	**107**

(b) The success rate was highest for stress management (0.909), followed by exercise (0.794) and usual care (0.70). (c) S: H_0: The true success rates for patients like these are the same for all three treatments versus H_a: The true success rates are not all the same. . . . P: Chi-square test for homogeneity. Random: 3 groups in a randomized experiment. Large Counts: 6.79, 6.99, 8.22, 26.21, 27.01, 31.78 all ≥ 5. D: $\chi^2 = 4.840$. With df = 2, the P-value is between 0.05 and 0.10 (0.0889). C: Because the P-value $> \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true success rates for patients like these are not the same for all three treatments.

R11.4 (a) The data could have been collected from 3 independent random samples—a random sample of ads from magazines aimed at young men, a random sample of ads from magazines aimed at young women, and a random sample of ads aimed at young adults in general. In each sample, the ads would be classified as sexual or not sexual. (b) The data could have been collected from a single random sample of ads from magazines aimed at young adults. Then each ad in the sample would be classified as sexual or not sexual, and the magazine that the ad was from would be classified as aimed at young men, young women, or young adults in general. (c) $\frac{351}{576} = 0.6094 = 60.94\%$. $\frac{(1113)(576)}{1509} = 424.8$. $\frac{(351-424.8)^2}{424.8} = 12.82$. (The difference is due to rounding error.) (d) The "sexual, Women" cell. There were 225 observed ads in this cell, which was 73.8 more than expected.

R11.5 (a)
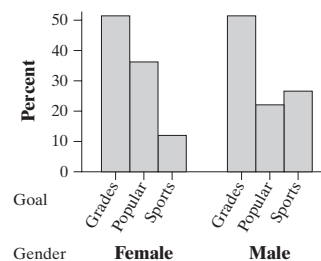
Both groups of children have the largest percentage reporting grades as the goal. But after that, boys were more likely to pick sports, whereas girls were more likely to pick being popular. **(b)** S: H_0: There is no association between gender and goals for 4th, 5th, and 6th grade students versus H_a: There is an association. . . . P: Chi-square test for independence. Random: Random sample. $10\% : n = 478 < 10\%$ of all 4th, 5th, and 6th grade students. Large Counts: 129.70, 117.30, 74.04, 66.96, 47.26, 42.74 all ≥ 5. D: $\chi^2 = 21.455$. With df $= 2$, the P-value < 0.0005 (0.00002). C: Because the P-value of $0.00002 < \alpha = 0.05$, we reject H_0. There is convincing evidence that an association exists between gender and goals for 4th, 5th, and 6th grade students.

Answers to Chapter 11 AP® Statistics Practice Test

T11.1 b
T11.2 c
T11.3 e
**T11.4** d
T11.5 c
T11.6 c
T11.7 b
T11.8 a
T11.9 d
T11.10 d
T11.11 S: H_0: The distribution of gas types is the same as the distributor's claim versus H_a: The distribution of gas types is not the same as the distributor's claim. P: Chi-square test for goodness of fit. Random: Random sample. $10\% : n = 400 < 10\%$ of all customers at this distributor's service stations. Large Counts: 240, 80, 80 all ≥ 5. D: $\chi^2 = 13.15$. With df $= 2$, the P-value is between 0.001 and 0.0025 (0.0014). C: Because the P-value of $0.0014 < \alpha = 0.05$, we reject H_0. There is convincing evidence that the distribution of gas type is not the same as the distributor claims.
T11.12 **(a)** Random assignment was used to create three roughly equivalent groups at the beginning of the study.
(b)
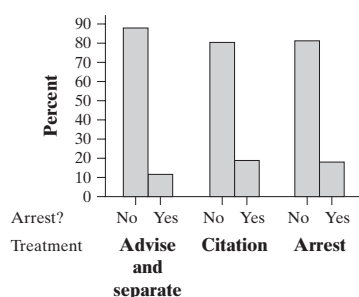
(c) H_0: The true proportion of spouse abusers like the ones in the study who will be arrested within 6 months is the same for all three police responses versus H_a: The true proportions are not all the same. **(d)** *P-value*: If the true proportion of spouse abusers like the ones in the study who will be arrested within 6 months is the same for all three police responses, there is a 0.0796 probability of getting differences between the three groups as large as or larger than the ones observed by chance alone. *Conclusion:* Because the P-value of 0.0796 is larger than $\alpha = 0.05$, we fail to reject H_0. There is not convincing evidence that true proportion of spouse abusers like the ones in the study who will be arrested within 6 months is not the same for all three police responses.

T11.13 **(a)** S: H_0: There is no association between smoking status and educational level among French men aged 20 to 60 years versus H_a: There is an association. . . . P: Chi-square test for independence. Random: Random sample. $10\% : n = 459 < 10\%$ of all French men aged 20 to 60 years. Large Counts: 59.48, 44.21, 42.31, 50.93, 37.85, 36.22, 42.37, 31.49, 30.14, 34.22, 25.44, 24.34 all ≥ 5. D: $\chi^2 = 13.305$. With df $= 6$, the P-value is between 0.025 and 0.05 (0.0384). C: Because the P-value of $0.0384 < \alpha = 0.05$, we reject H_0. There is convincing evidence of an association between smoking status and educational level among French men aged 20 to 60 years.

Chapter 12
Section 12.1
Answers to Check Your Understanding
page 752: S: $\beta =$ slope of the population regression line relating fat gain to change in NEA. P: t interval for the slope. Linear: There is no leftover pattern in the residual plot. Independent: The sample size ($n = 16$) is less than 10% of all healthy young adults. Normal: The histogram of the residuals shows no strong skewness or outliers. Equal SD: Other than one point with a large positive residual, the residual plot shows roughly equal scatter for all x values. Random: Random sample. D: With df $= 14$, $(-0.005032, -0.001852)$. C: We are 95% confident that the interval from -0.005032 to -0.001852 captures the slope of the population regression line relating fat gain to change in NEA.
page 757: S: H_0: $\beta = 0$ versus H_a: $\beta < 0$, where β is the slope of the true regression line relating fat gain to NEA change. P: t test for the slope β. D: $t = -4.64$. P-value $\approx 0.000/2 \approx 0$. C: Because the P-value of approximately 0 is less than $\alpha = 0.05$, we reject H_0. There is convincing evidence that the slope of the true regression line relating fat gain to NEA change is negative.

Answers to Odd-Numbered Section 12.1 Exercises
12.1 The Equal SD condition is not met because the SD of the residuals clearly increases as the laboratory measurement (x) increases.
12.3 Linear: There is no leftover pattern in the residual plot. Independent: Knowing the BAC for one subject should not help us predict the BAC for another subject. Normal: The histogram of the residuals shows no strong skewness or outliers. Equal SD: The residual plot shows roughly equal scatter for all x values. Random: These data come from a randomized experiment.
12.5 α is the true y intercept, which measures the true mean BAC level if no beers had been drunk ($a = -0.012701$). β is the true slope, which measures how much the true mean BAC changes with the drinking of one additional beer ($b = 0.018$). Finally, σ is the true standard deviation of the residuals, which measures how much the observed values of BAC typically vary from the population regression line ($s = 0.0204$).
12.7 **(a)** $SE_b = 0.0024$. If we repeated the experiment many times, the slope of the sample regression line would typically vary by about 0.0024 from the slope of the true regression line for predicting BAC from the number of beers consumed. **(b)** With df $= 14$, $0.018 \pm 2.977(0.0024) = (0.011, 0.025)$. **(c)** We are 99% confident that the interval from 0.011 to 0.025 captures the slope of the true regression line for predicting BAC from the number of beers consumed. **(d)** If we repeated the experiment many times

and computed a confidence interval for the slope each time, about 99% of the resulting intervals would contain the slope of the true regression line for predicting BAC from the number of beers consumed.

12.9 S: β = the slope of the population regression line relating number of clusters of beetle larvae to number of stumps. P: t interval for β. D: With df = 21, (8.678, 15.11). C: We are 99% confident that the interval from 8.678 to 15.11 captures the slope of the population regression line relating number of clusters of beetle larvae to number of stumps.

12.11 (a) $\hat{y}$ = −1.286 + 11.894(5) = 58.184 clusters. (b) s = 6.419, so we would expect our prediction to be off from the actual number of clusters by about 6.419 clusters.

12.13 (a) $\hat{y}$ = 166.483 − 1.0987x, where $\hat{y}$ is the predicted corn yield and x is the number of weeds per meter. Slope: for each additional weed per meter, the predicted corn yield will decrease by about 1.0987 bushels/acre. y intercept: if there are no weeds per meter, we would predict a corn yield of 166.483 bushels/acre. (b) When using weeds per meter to predict corn yield, the actual yield will typically vary from the predicted yield by about 7.98 bushels/acre. (c) S: H_0: β = 0 versus H_a: β < 0, where β is the slope of the true regression line relating corn yield to weeds per meter. P: t test for β. D: t = −1.92. P-value = 0.075/2 = 0.0375. C: Because the P-value of 0.0375 is less than α = 0.05, we reject H_0. There is convincing evidence that the slope of the true regression line relating corn yield to weeds per meter is negative.

12.15 S: H_0: β = 0 versus H_a: β < 0, where β is the slope of the population regression line relating heart disease death rate to wine consumption in the population of countries. P: t test for β. Linear: There is no leftover pattern in the residual plot. Independent: The sample size (n = 19) is less than 10% of all countries. Normal: The histogram of residuals shows no strong skewness or outliers. Equal SD: The residual plot shows that the standard deviation of the death rates might be a little smaller for large values of wine consumption, x, but it is hard to tell with so few data values. Random: Random sample.
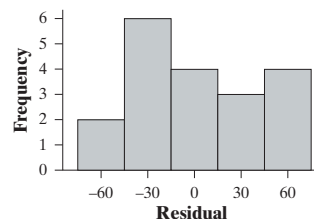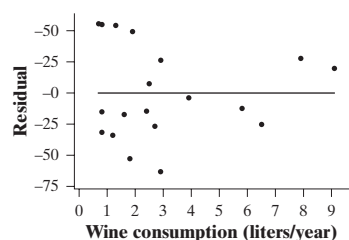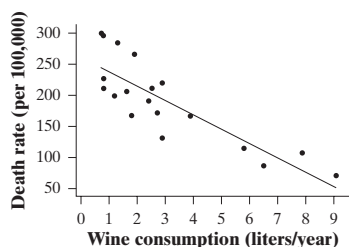
D: t = −6.46, df = 17, and P-value ≈ 0. C: Because the P-value of approximately 0 is less than α = 0.05, we reject H_0. There is convincing evidence of a negative linear relationship between wine consumption and heart disease death rate in the population of countries.

12.17 (a) With df = 19, 11,630.6 ± 2.093(1249) = (9016.4, 14,244.8). (b) Because the automotive group claims that people drive 15,000 miles per year, this says that for every increase of 1 year, the mileage would increase by 15,000 miles. (c) t = −2.70. With df = 19, the P-value is between 0.01 and 0.02 (0.0142). Because the P-value of 0.0142 is less than α = 0.05, we reject H_0. We have convincing evidence that the slope of the population regression line relating miles to years is not equal to 15,000. (d) Yes. Because the interval in part (a) does not include the value 15,000, the interval also provides convincing evidence that the slope of the population regression line relating miles to years is not equal to 15,000.

12.19 c

12.21 a

12.23 b

12.25 (a) The two treatments (say the color, read the word) were deliberately assigned to the students. (b) He used a randomized block design where each student was a block. He did this to help account for the different abilities of students to read the words or to say the color they were printed in. (c) To help average out the effects of the order in which people did the two treatments. If every subject said the color of the printed word first and were frustrated by this task, the times for the second treatment might be worse. Then we wouldn't know the reason the times were longer for the second treatment—because of frustration or because the second method actually takes longer.

12.27 There is a small number of differences (n_d = 16 < 30) and there is an outlier.

12.29 (a) (i) $\dfrac{295}{1526}$ = 0.1933. (ii) $\dfrac{295 + 77 + 212}{1526}$ = 0.3827.

(iii) $\dfrac{212}{305}$ = 0.6951. (b) No. The probability that a person is a snowmobile owner (295/1526 = 0.1933) is different from the probability that the person is a snowmobile owner given that he or she belongs to an environmental organization (16/305 = 0.0525).

(c) (i) P(both are owners) = $\left(\dfrac{295}{1526}\right)\left(\dfrac{294}{1525}\right)$ = 0.0373

(ii) P(at least one belongs to an environmental organization)

= 1 − P(neither belong) = 1 − $\left(\dfrac{1221}{1526}\right)\left(\dfrac{1220}{1525}\right)$ = 0.3599

Section 12.2

Answers to Check Your Understanding

page 782: **1.** Option 1: $\widehat{\text{premium}} = -343 + 8.63(58) = \$157.54$

Option 2: $\widehat{\ln(\text{premium})} = -12.98 + 4.416(\ln 58) = 4.9509$
$\rightarrow \hat{y} = e^{4.9509} = \$141.30$

Option 3: $\widehat{\ln(\text{premium})} = -0.063 + 0.0859(58) = 4.9192$
$\rightarrow \hat{y} = e^{4.9192} = \$136.89$

2. Exponential (Option 3), because the scatterplot showing ln(premium) versus age was the most linear and this model had the most randomly scattered residual plot.

Answers to Odd-Numbered Section 12.2 Exercises

12.31 **(a)** The scatterplot shows a fairly strong, positive, slightly curved association between length and period with one very unusual point (106.5, 2.115) in the top right corner.
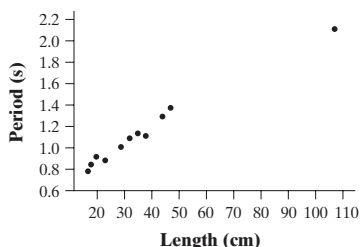
(b) The class used the square root of $x =$ length. **(c)** The class used the square of $y =$ period.

12.33 **(a)** 1: $\hat{y} = -0.08594 + 0.21\sqrt{x}$, where y is the period and x is the length. 2: $\widehat{y^2} = -0.15465 + 0.0428x$, where y is the period and x is the length. **(b)** 1: $\hat{y} = -0.08594 + 0.21\sqrt{80} = 1.792$ seconds. 2: $\widehat{y^2} = -0.15465 + 0.0428(80) = 3.269$, so $\hat{y} = \sqrt{3.269} = 1.808$ seconds.

12.35 **(a)** The scatterplot of log(period) versus log(length) is roughly linear and the residual plot shows no obvious leftover patterns. **(b)** $\widehat{\log y} = -0.73675 + 0.51701 \log(x)$, where y is the period and x is the length.

12.37 $\widehat{\log y} = -0.73675 + 0.51701 \log(80) = 0.24717$. Thus, $\hat{y} = 10^{0.24717} = 1.77$ seconds.

12.39 $\widehat{\log y} = 1.01 + 0.72 \log(127) = 2.525$. Thus, $\hat{y} = 10^{2.525} = 334.97$ grams.

12.41 **(a)** The relationship between bacteria count and time is strong, negative, and curved with a possible outlier in the top left-hand corner.
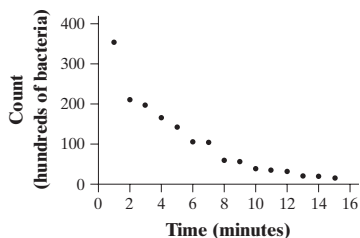

(b) Because the scatterplot of ln(count) versus time is fairly linear. **(c)** $\ln y = 5.97316 - 0.218425x$, where y is the count of

surviving bacteria and x is time in minutes. **(d)** $\widehat{\ln y} = 5.97316 - 0.218425(17) = 2.26$, so $\hat{y} = e^{2.26} = 9.58$ or 958 bacteria.

12.43 **(a)** Exponential, because the scatterplot of log(height) versus bounce number is more linear. **(b)** $\widehat{\log y} = 0.45374 - 0.11716x$, where $y =$ height in feet and $x =$ bounce number. **(c)** $\widehat{\log y} = 0.45374 - 0.11716(7) = -0.36638$, so $\hat{y} = 10^{-0.36638} = 0.43$ feet. **(d)** The trend in the residual plot suggests that the residual for $x = 7$ would be positive, meaning that the predicted height will be less than the actual height.

12.45 **(a)** There is a strong, positive curved relationship between heart weight and length of left ventricle for mammals.
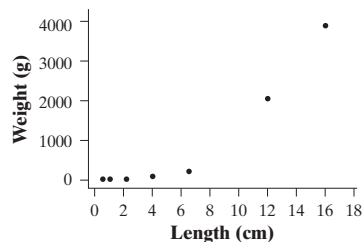
(b) Two scatterplots are given below. Because the relationship between ln(weight) and ln(length) is roughly linear, heart weight and length seem to follow a power model.
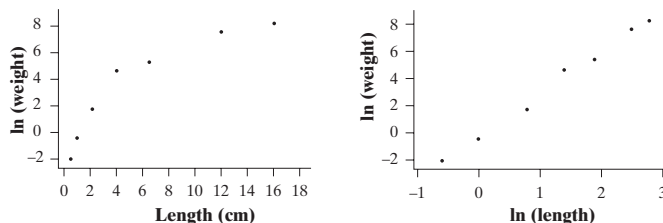
(c) $\widehat{\ln y} = -0.314 + 3.1387 \ln x$, where y is the weight of the heart and x is the length of the cavity of the left ventricle. **(d)** $\widehat{\ln y} = -0.314 + 3.1387 \ln (6.8) = 5.703$, so $\hat{y} = e^{5.703} = 299.77$ grams.

12.47 c

12.49 e

12.51 **(a)** For Marcela, $X =$ the length of her shower on a randomly selected day follows a Normal distribution with mean 4.5 minutes and standard deviation 0.9 minutes. We want to find $P(3 < X < 6)$. $z = \dfrac{3 - 4.5}{0.9} = -1.67$ and $z = \dfrac{6 - 4.5}{0.9} = 1.67$, so $P(3 < X < 6) = 0.9050$. *Using technology:* 0.9044. There is a 0.9044 probability that Marcela's shower lasts between 3 and 6 minutes. **(b)** Solving $-0.67 = \dfrac{Q_1 - 4.5}{0.9}$ gives $Q_1 = 3.897$ minutes. Solving $0.67 = \dfrac{Q_3 - 4.5}{0.9}$ gives $Q_3 = 5.103$ minutes. *Using technology:* $Q_1 = 3.893$ minutes and $Q_3 = 5.107$ minutes. Thus, an outlier is any value above $5.107 + 1.5(5.107 - 3.893) = 6.928$. Because $7 > 6.928$, a shower of 7 minutes would be considered an outlier for Marcela. **(c)** $P(X > 7) = 0.0027$. Let $Y =$ the number of days that Marcela's shower is 7 minutes or higher. Y is a binomial random variable with $n = 10$ and $p = 0.0027$. $P(Y \geq 2) = 1 - P(Y \leq 1) = 1 - \texttt{binomcdf(trials: 10, p: 0.0027, x value: 1)} = 0.0003$. **(d)** $\bar{x}$ follows a N(4.5, 0.285) distribution and we want to find $P(\bar{x} > 5)$. $z = \dfrac{5 - 4.5}{0.285} = 1.75$ and $P(Z > 1.75) =$. *Using technology:* 0.0397. There is a 0.0397 probability that the mean length of Marcela's showers on these 10 days exceeds 5 minutes.

12.53 (a) S: $p =$ true proportion of all AP® teachers attending this workshop who have tattoos. P: One-sample z interval for p. Random: Random sample. 10%: The sample size ($n = 98$) is less than 10% of the population of teachers at this workshop (1100). Large Counts: 23 and 75 are both ≥ 10. D: (0.151, 0.319). C: We are 95% confident that the interval from 0.151 to 0.319 captures the true proportion of AP® teachers at this workshop who have tattoos. (b) Yes. Because the value 0.14 is not included in the interval, we have convincing evidence that the true proportion of teachers at the workshop who have a tattoo is not 0.14. (c) If we had two more failures, the interval will shift to lower values and might include the value 0.14. However, the new interval is (0.148, 0.312), which does not include the value 0.14. So the answer would not change if we got responses from the 2 nonresponders.

Answers to Chapter 12 Review Exercises

R12.1 (a) There is a moderately strong, positive linear relationship between the thickness and the velocity. (b) $\hat{y} = 70.44 + 274.78x$, where y is the velocity and x is the thickness. (c) Residual = $104.8 - 180.352 = -75.552$, so the line overpredicts the velocity by 75.552 ft/sec. (d) The linear model is appropriate. The scatterplot shows a linear relationship and the residual plot has no leftover patterns. (e) Slope: For each increase of an inch in thickness, the predicted velocity increases by 274.78 feet/second. s: When using the least-squares regression line with $x =$ thickness to predict $y =$ velocity, we will typically be off by about 56.36 feet per second. r^2: About 49.3% of the variation in velocity is accounted for by the linear relationship relating velocity to thickness. SE_b: If we take many different random samples of 12 pistons and compute the least-squares regression line for each sample, the estimated slope will typically vary from the slope of the population regression line for predicting velocity from thickness by about 88.18.

R12.2 S: $H_0:\beta = 0$ versus $H_a:\beta \neq 0$, where β is the slope of the population regression line relating thickness to velocity. P: t test for β. Linear: The residual plot shows no leftover patterns. Independent: Knowing the velocity for one piston should not help us predict the velocity for another piston. Also, the sample size ($n = 12$) is less than 10% of the pistons in the population. Normal: We are told that the Normal probability plot of the residuals is roughly linear. Equal SD: The residual plot shows roughly equal scatter for all x values. Random: The data come from a random sample. D: $t = 3.116$. With df $= 10$, the P-value is between 0.01 and 0.02 (0.0109). C: Because the P-value of 0.0109 is less than $\alpha = 0.05$, we reject H_0. There is convincing evidence of a linear relationship between thickness and gate velocity in the population of pistons formed from this alloy of metal.

R12.3 D: With df $= 12 - 2 = 10$, (78.315, 471.245). C: We are 95% confident that the interval from 78.315 to 471.245 captures the slope of the population regression line for predicting velocity from thickness for the population of pistons formed from this alloy of metal. Because 0 is not in the interval, we reject 0 as a plausible value for the slope of the population regression line, as in R12.2.

R12.4 The Linear condition is violated because there is clear curvature to the scatterplot and an obvious curved pattern in the residual plot. The Random condition may not be met because we weren't told if the sample was selected at random.

R12.5 (a) Yes, because there is no leftover pattern in the residual plot. (b) $\hat{y} = -0.000595 + 0.3\left(\dfrac{1}{x^2}\right)$. Here, $y =$ intensity and $x =$ distance. (c) $\hat{y} = -0.000595 + 0.3\left(\dfrac{1}{(2.1)^2}\right) = 0.0674$ candelas.

R12.6 (a)
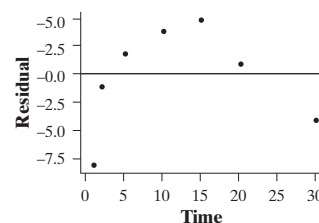
(b) There is a leftover pattern in the residual plot, so the relationship between practice time and percent of words recalled is not linear. (c) Power, because the scatterplot showing ln(recall) versus ln(time) is more linear than the scatterplot showing ln(recall) versus time. (d) Power: $\widehat{\ln y} = 3.48 + 0.293 \ln(25) = 4.423$ and $\hat{y} = e^{4.423} = 83.35$ percent of words recalled. Exponential: $\widehat{\ln y} = 3.69 + 0.0304(25) = 4.45$ and $\hat{y} = e^{4.45} = 85.63$ percent of words recalled. Based on my answer to part (c), I think the power model will give a better prediction.

Answers to Chapter 12 AP® Statistics Practice Test

T12.1 c
T12.2 b
T12.3 d
T12.4 a
T12.5 d
T12.6 d
T12.7 e
T12.8 d
T12.9 d
T12.10 c
T12.11 (a) $\hat{y} = 4.546 + 4.832x$, where y is the weight gain and x is the dose of growth hormone. (b) (i) For each 1-mg increase in growth hormone, the predicted weight gain increases by about 4.832 ounces. (ii) If a chicken is given no growth hormone ($x = 0$), the predicted weight gain is 4.546 ounces. (iii) When using the least-squares regression line with $x =$ dose of growth hormone to predict $y =$ weight gain, we will typically be off by about 3.135 ounces. (iv) If we repeated this experiment many times, the sample slope will typically vary by about 1.0164 from the true slope of the least-squares regression line with $y =$ weight gain and $x =$ dose of growth hormone. (v) About 38.4% of the variation in weight gain is accounted for by the linear model relating weight gain to the dose of growth hormone. (c) S: $H_0:\beta = 0$ versus $H_a:\beta \neq 0$, where β is the slope of the true regression line relating $y =$ weight gain to $x =$ dose of growth hormone. P: t test for β. D: $t = 4.75$, df $= 13$, and P-value $= 0.0004$. C: Because the P-value of 0.0004 is less than $\alpha = 0.05$, we reject H_0. There is convincing evidence of a linear relationship between the dose of growth hormone and weight gain for chickens like these. (d) D: With df $= 13$, (2.6373, 7.0273). C: We are 95% confident that the interval from 2.6373 to 7.0273 captures the slope of the true regression line relating $y =$ weight gain to $x =$ dose of growth hormone for chickens like these.

T12.12 (a) There is clear curvature evident in both the scatterplot and the residual plot. (b) 1: $\hat{y} = 2.078 + 0.0042597(30)^3 = 117.09$ board feet. 2: $\widehat{\ln y} = 1.2319 + 0.113417(30) = 4.63441$ and $\hat{y} = e^{4.63441} = 102.967$ board feet. (c) The residual plot for Option 1 is much more scattered, while the plot for Option 2 shows curvature, meaning that the model from Option 1 relating the amount of usable lumber to cube of the diameter is more appropriate.

Answers to Cumulative AP® Practice Test 4

AP4.1 e
AP4.2 c
AP4.3 e
AP4.4 a
AP4.5 b
AP4.6 e
AP4.7 d
AP4.8 a
AP4.9 e
AP4.10 a
AP4.11 b
AP4.12 d
AP4.13 e
AP4.14 d
AP4.15 b
AP4.16 a
AP4.17 d
AP4.18 c
AP4.19 a
AP4.20 b
AP4.21 e
AP4.22 e
AP4.23 b
AP4.24 c
AP4.25 d
AP4.26 e
AP4.27 c
AP4.28 c
AP4.29 b
AP4.30 b
AP4.31 d
AP4.32 a
AP4.33 e
AP4.34 b
AP4.35 a
AP4.36 b
AP4.37 d
AP4.38 a
AP4.39 d
AP4.40 d

AP4.41 S: $H_0:\mu_1 - \mu_2 = 0$ vs. $H_a:\mu_1 - \mu_2 \neq 0$, where $\mu_1 =$ true mean difference in electrical potential for diabetic mice and $\mu_2 =$ true mean difference in electrical potential for normal mice. P: Two-sample t test for $\mu_1 - \mu_2$. Random: Independent random samples. 10%: $n_1 = 24$ is less than 10% of all diabetic mice and $n_2 = 18$ is less than 10% of all normal mice. Normal/Large Sample Size: No outliers or strong skewness. D: $t = 2.55$. Using df $= 23$, the P-value is between 0.01 and 0.02. Using df $= 38.46$, P-value $= 0.0149$. C: Because the P-value of 0.0149 is less than $\alpha = 0.05$, we reject H_0. There is convincing evidence that the true mean difference in electric potential for diabetic mice is different than for normal mice.

AP4.42 (a) $H_0:p_1 - p_2 = 0$ vs. $H_a:p_1 - p_2 < 0$, where $p_1 =$ the true proportion of women like the ones in the study who were physically active as teens that would suffer a cognitive decline and $p_2 =$ the true proportion of women like the ones in the study who were not physically active as teens that would suffer a cognitive decline. (b) A two-sample z test for $p_1 - p_2$. (c) No. Because the participants were mostly white women from only four states, the findings may

not be generalizable to women in other racial and ethnic groups or who live in other states. (d) Two variables are confounded when their effects on the response variable cannot be distinguished from one another. For example, women who were physically active as teens might have also done other things differently as well, such as eating a healthier diet. We would be unable to determine if it was their physically active youth or their healthier diet that slowed their level of cognitive decline.

AP4.43 (a) Because the first question called it a "fat tax," people may have reacted negatively because they believe this is a tax on those who are overweight. The second question provides extra information that gets people thinking about the obesity problem in the U.S. and the increased health care that could be provided as a benefit with the tax money. Better: "Would you support or oppose a tax on non-diet sugared soda?" (b) This method samples only people at fast-food restaurants. They may go to these restaurants because they like the sugary drinks and wouldn't want to pay a tax on their favorite beverages. Thus, it is likely that the proportion of those who would oppose such a tax will be overestimated with this method. Better: take a random sample of all New York State residents. (c) Use a stratified random sampling method in which each state is a stratum.

AP4.44 (a) $P(S) = (0.1)(0.3) + (0.4)(0.2) + (0.5)(0.1) = 0.16$.
(b) $P(C|S) = \dfrac{(0.5)(0.1)}{(0.1)(0.3) + (0.4)(0.2) + (0.5)(0.1)} = 0.3125$.

AP4.45 (a) No. The scatterplot exhibits a strong curved pattern. (b) B, because the scatterplot shows a much more linear pattern and its residual plot shows no leftover patterns.

(c) $\widehat{\ln(\text{weight})} = 15.491 - 1.5222 \ln (3700) = 2.984$, thus $\widehat{\text{weight}} = e^{2.984} = 19.77$mg. (d) About 86.3% of the variation in ln(seed weight) is accounted for by the linear model relating ln(seed weight) to ln(seed count).

AP4.46 (a) Let $X =$ diameter of a randomly selected lid. Because X follows a Normal distribution, the sampling distribution of $\bar{x}$ also follows a Normal distribution. $\mu_{\bar{x}} = 4$ inches and $\sigma_{\bar{x}} = \dfrac{0.02}{\sqrt{25}} = 0.004$ 0.004 inches. (b) We want to find $P(\bar{x} < 3.99 \text{ or } \bar{x} > 4.01)$ using the N(4, 0.004) distribution. $z = \dfrac{3.99 - 4}{0.004} = -2.50$ and $z = \dfrac{4.01 - 4}{0.004} = 2.50$. $P(Z < -2.50 \text{ or } Z > 2.50) = 0.0124$. Assuming that the machine is working properly, there is a 0.0124 probability that the mean diameter of a sample of 25 lids is less than 3.99 inches or greater than 4.01 inches. (c) We want to find $P(4 < \bar{x} < 4.01)$ using the N(4, 0.004) distribution. $z = \dfrac{4 - 4}{0.004} = 0$ and $z = \dfrac{4.01 - 4}{0.004} = 2.50$. $P(0 < Z < 2.50) = 0.4938$. Assuming that the machine is working properly, there is a 0.4938 probability that the mean diameter of a sample of 25 lids is between 4.00 and 4.01 inches. (d) Let $Y =$ the number of samples (out of 5) in which the sample mean is between 4.00 and 4.01. The random variable Y has a binomial distribution with $n = 5$ and $p = 0.4938$. *Using technology*: $P(X \geq 4) = 1 - P(X \leq 3) = 1 - $ `binomcdf` `(trials:5, p:0.4938, x value:3) = 0.1798`. (e) Because the probability found in part (b) is less than the probability found in part (d), getting a sample mean below 3.99 or above 4.01 is more convincing evidence that the machine needs to be shut down. This event is much less likely to happen by chance when the machine is working correctly. (f) Answers will vary.