# Answers

# Chapter 1 Collecting Data in Reasonable Ways

## **SECTION 1.1**

#### **Exercise Set 1**

**1.1** observational study; the person conducting the study merely recorded whether or not the boomers sleep with their phones within arm's length and whether or not people used their phones to take photos.

**1.2** observational study; children were not assigned to different experimental groups.

**1.3** experiment; the researchers assigned different toddlers to experimental conditions (adult played with/ talked to the robot or the adult ignored the robot).

**1.4** observational study; the researchers did not assign people to experimental conditions.

**1.5** experiment; because the researchers assigned study participants to one of three experimental groups (meditation, distraction task, or relaxation technique).

#### **Additional Exercises**

**1.11** observational study; there was no assignment of subjects to experimental groups.

**1.13** experiment; the study participants were assigned to one of the two experimental groups (how much would you pay for the mug or how much would you sell the mug for).

## **SECTION 1.2**

#### **Exercise Set 1**

**1.15** (a) census(b) population characteristic

**1.16** The sample is the 2,121 children between the ages of 1 and 4, and the population is all children between the ages of 1 and 4.

**1.17** No. The 6,000 people who sent hair samples were volunteers and were not chosen at random.

**1.18** There are several reasonable approaches. One possibility is to use a list of all the students at the school and to write all of the names on otherwise identical slips of paper. Thoroughly mix the slips of paper, and select 150 slips. Include the individuals whose names are on the slips of paper in the sample.

#### 1.19 (a) all U.S. women

(**b**) only women from Maryland, Minnesota, Oregon, and Pennsylvania were included in the sample.

(c) Given that only women from four states were included in the sample, the sample is not likely to be representative of the population of interest.

(d) Selection bias is present because the selection method excluded women from all states other than Maryland, Minnesota, Oregon, and Pennsylvania.

#### **Additional Exercises**

**1.25** The population is all 7,000 property owners. The sample is the 500 property owners selected for the survey.

**1.27** The population is the 5,000 bricks in the lot. The sample is the 100 bricks chosen for inspection.

**1.29** Bias introduced through the two different sampling methods may have contributed to the different results. The online sample could suffer from voluntary response bias in that perhaps only those who feel very strongly would take the time to go to the website and register their vote. In addition, younger people might be more technologically savvy, and therefore the website might over-represent the views of younger people (particularly students) who support the parade. The telephone survey responses might over-represent the view of permanent residents (as students might only use cell phones and not have a local phone number).

## **SECTION 1.3**

#### Exercise Set 1

**1.30** Random assignment allows the researcher to create groups that are equivalent, so that the subjects in each experimental group are as much alike as possible. This ensures that the experiment does not favor one experimental condition (playing Unreal Tournament 2004 or Tetris) over another.

**1.31** (a) Allowing participants to choose which group they want to be in could introduce systematic differences between the two experimental conditions (tai chi group or control group), resulting in potential confounding. Those who would choose to do tai chi might, in some way, be different from those who would choose the control group. (b) Because the purpose of this experiment is to determine whether the tai chi treatment has an effect on immunity to a virus, a control group is needed to provide a baseline against which the treatment group can be compared.

**1.32** (a) The attending nurse was responsible for administering medication after judging the degree of pain and nausea, so the researchers did not want the nurse's personal beliefs about the different surgical procedures to influence measurements.

(b) Because the children who had the surgery could easily determine whether the surgical procedure was laparoscopic repair or open repair based on the type of incision.

**1.33** There are several possible approaches. One possibility is to write the subjects names on otherwise identical slips of paper. Mix the slips of paper thoroughly and draw out slips one at a time. The names on the first 15 slips are assigned to the experimental condition of listening to a Mozart piano sonata for 24 minutes. The names on the next 15 slips are assigned to the experimental condition of listening to popular music for the same length of time. The remaining 15 names are assigned to the relaxation with no music experimental condition.

**1.34** (1) Do ethnic group and gender influence the type of care that a heart patient receives? (2) The experimental conditions are gender and race. (3) The response variable is the type of care the heart patient received. (4) The experimental units are the 720 primary care doctors. It is not clear how the physicians were chosen. (5) Yes, the design incorporates random assignment of doctors to view one of the four different videos through rolling a four-sided die. (6) No control group was used. There is no need for a control group in this study. (7) There is no indication that the study includes blinding, but there is no need for blinding in this experiment.

#### **Additional Exercises**

**1.41** (a) Some surgical procedures are more complex and require a greater degree of concentration; music with a vocal component might be more distracting when the surgical procedure is more complex.

(b) The temperature of the room might affect the comfort of the surgeon; if the surgeon is too hot or too cold, she or he might be uncomfortable, and therefore more easily distracted by the vocal component.

(c) If the music is too loud, the surgeon might be distracted and unable to focus, regardless of the presence or absence of the vocal component. If the music is too soft, the surgeon might try to concentrate on listening to the vocal component and therefore pay more attention to the music rather than to the surgical procedure.

**1.43** This experiment could not have been double-blind because the surgeon would know whether or not there was a vocal component to the music.

1.45 (a) Probably not, because the judges might not believe that Denny's food is as good as that of other restaurants.(b) Experiments are often blinded in this way to eliminate preconceptions about particular experimental treatments.

## **SECTION 1.4**

#### **Exercise Set 1**

**1.46** This was an observational study, so cause-and-effect conclusions cannot be drawn.

**1.47** (a) It is not reasonable to conclude that watching *Oprah* causes a decrease in cravings for fattening foods. This was an observational study, so cause-and-effect conclusions cannot be drawn.

(b) It is not reasonable to generalize the results of this survey to all women in the United States because not all women watch daytime talk shows. It is not reasonable to generalize these results to all women who watch daytime talk shows because not all women who watch daytime shows access DietSmart.com.

**1.48** The researcher would have had to assign the nine cyclists at random to one of the three experimental conditions (chocolate milk, Gatorade, or Endurox).

#### 1.49

**Study 1:** This is an observational study; random selection was used; this was not an experiment so there were no experimental groups; no, because this was not an experiment, cause-and-effect cannot be concluded; it is reasonable to generalize to the population of students at this particular large college.

**Study 2:** This study was an experiment; random selection was not used; there was no random assignment to experimental conditions (the grouping was based on gender); the conclusion is not appropriate because of confounding of gender and treatment (women ate pecans, and men did not eat pecans); it is not reasonable to generalize to a larger population.

**Study 3:** This is an observational study; no random selection; no random assignment to experimental groups; the conclusion is not appropriate because this was an observational study, and therefore cause-and-effect conclusions cannot be drawn; cannot generalize to any larger population.

**Study 4:** This is an experiment; no random selection; there was random assignment to experimental groups; yes, because this was an experiment with random assignment of subjects to experimental groups, we can draw cause-and-effect conclusions; cannot generalize to a larger population.

**Study 5:** This is an experiment; there was random selection from students enrolled at a large college; random assignment of subjects to experimental groups was used; because this was a simple comparative experiment with random assignment of subjects to experimental groups, we can draw cause-and-effect conclusions; there was random selection of students, so we can generalize conclusions from this study to the population of all students enrolled at the large college.

#### **Additional Exercises**

**1.55** Would need to know if dieters were randomly assigned to the experimental conditions (large fork or small fork) and if the study participants were randomly selected from the population of dieters.

**1.57** There was no random selection from some population.

**1.59** Yes, because this was an experiment and there was random assignment of subjects to experimental groups.

## ARE YOU READY TO MOVE ON? CHAPTER 1 REVIEW EXERCISES

**1.61** (a) experiment, there is random assignment of subjects to experimental conditions.

(b) observational study, there was no assignment of subjects to experimental conditions

(c) observational study, there was no assignment of subjects to experimental conditions.

(d) experiment, there was random assignment of study participants to experimental conditions.

1.63 (a) population characteristic
(b) statistic
(c) population characteristic
(d) statistic
(e) statistic

**1.65** The council president could assign a unique identifying number to each of the names on the petition, numbered from 1 to 500. On identical slips of paper, write the numbers 1 to 500, with each number on a single slip of paper. Thoroughly mix the slips of paper and select 30 numbers. The 30 numbers correspond to the unique numbers assigned to names on the petition. These 30 are the names that would be in the sample.

**1.67** Without random assignment of the study participants to experimental conditions, confounding could impact the conclusions of the study. For example, people who would choose an attractive avatar might be more outgoing and willing to engage than someone who would choose an unattractive avatar.

1.69 (a) The alternate assignment to the experimental groups (large serving bowls, small serving bowls) would probably produce groups that are similar.(b) Blinding ensures that individuals do not let personal beliefs

influence their measurements. The research assistant who weighed the plates and estimated the calorie content of the food might (intentionally or not) have let personal beliefs influence the estimate of the calorie content of the food on the plate.

**1.71** (a) (1) Does using hand gestures help children learn math? (2) Using hand gestures and not using hand gestures. (3) Number correct on the six-problem test. (4) The 128 children in the study; they were selected because they were the children who answered all six questions on the pretest incorrectly. (5) Yes, the children were assigned randomly to one of the two experimental groups. (6) Yes, the control group is the experimental condition of not using any hand gestures. (7) There was no blinding. It would not be possible to include blinding of subjects in this experiment (the children would know

whether or not they were using hand gestures), and there is no need to blind the person recording the response because the test was graded with each answer correct or incorrect, so there is no subjectivity in recording the responses. (b) The conclusions are reasonable because the subjects were assigned to the treatment groups at random.

**1.73** (a) No, the 60 games selected were the 20 most popular (by sales) for each of three different gaming systems. The study excluded the games that were not in the top 20 most popular (by sales).

(**b**) It is not reasonable to generalize to all video games because of the exclusion of those games not in the top 20 (by sales).

#### 1.75

**Study 1:** observational study; no random selection; no random assignment to experimental groups; not reasonable to conclude that taking calcium supplements is the cause of the increased heart attack risk; not reasonable to generalize conclusions from this study to a larger population.

**Study 2:** observational study; there was random selection from the population of people living in Minneapolis who receive Social Security; no random assignment of subjects to experimental groups; not reasonable to conclude that taking calcium supplements is the cause of the increased heart attack risk; it is reasonable to generalize the results of this study to the population of people living in Minneapolis who receive Social Security.

Study 3: experiment; there was random selection from the population of people living in Minneapolis who receive Social Security; no random assignment of subjects to experimental groups; not reasonable to conclude that taking calcium supplements is the cause of the increased risk of heart attack because the participants in this study who did not have a previous history of heart problems were given the calcium supplement, and those with a history of heart problems were not given the supplement. It is not possible to determine the role of the calcium supplement because only those study participants who did not have a history of heart problems were given the supplement; it is possible to generalize the results from this study to the population of all people living in Minneapolis who receive Social Security. However, it is unclear (due to the confounding described in Question 4) what the conclusion would be.

**Study 4:** experiment; no random selection from some larger population; there was random assignment of study participants to experimental groups; it is reasonable to conclude that taking calcium supplements is the cause of the increased risk of heart attack; it is not reasonable to generalize conclusions from this study to some larger population.

# Chapter 2 Graphical Methods for Describing Data Distributions

## **SECTION 2.1**

#### **Exercise Set 1**

**2.1** (a) numerical, discrete (b) categorical (c) numerical, continuous (d) numerical, continuous (e) categorical

2.2 (a) discrete (b) continuous (c) discrete (d) discrete

#### 2.3

**Data Set 1:** one variable; categorical; summarize the data distribution; bar chart

**Data Set 2:** one variable; numerical; compare groups; comparative dotplot or comparative stem-and-leaf display.

**Data Set 3:** two variables; numerical; investigate relationship; scatterplot.

**Data Set 4:** one variable; categorical, compare groups, comparative bar chart.

**Data Set 5:** one variable; numerical; summarize the data distribution; dotplot, stem-and-leaf display or histogram.

#### **Additional Exercises**

2.7 (a) numerical (b) numerical (c) categorical(d) numerical (e) categorical

**2.9** (a) categorical (b) numerical (c) numerical(d) categorical

2.11 (a) numerical (b) numerical (c) numerical(d) categorical (e) categorical (f) numerical (g) categorical

**2.13** one variable; numerical; compare groups, comparative dotplot or comparative stem-and-leaf display.

**2.15** one variable; numerical, summarize the data distribution; dotplot, stem-and-leaf plot or histogram.

## **SECTION 2.2**

#### **Exercise Set 1**



(**b**) One example: "Senior Satisfaction! Over 80% say they would enroll again."

#### 2.17

Additional Exercises





Credit card fraud is the most commonly occurring identity theft type. Although phone/utility, bank, and employment fraud each constitute a relatively large portion of overall type of identity theft, the collective "other fraud" category is greater than any one of these other three.

2.23 The relative frequency distribution is:

Nonfamilies 0.20
Nomannines 0.29
Married with Children 0.27
Married without Children 0.29
Single Parent 0.15



## **SECTION 2.3**

#### Exercise Set 1



(b) Because the costs of meat and poultry products (represented by the squares on the dotplot) are generally smaller than the costs for other sources of protein, they do appear to be a good value.

#### 2.25 (a)



(b) Both distributions are skewed toward larger values. The 2007 ticket sales are centered at about \$210 million dollars, which is higher than the center of the 2008 ticket sales, which are centered around \$150 million dollars. The lowest ticket sales for both 2007 and 2008 are approximately \$127 million dollars. Ticket sales for 2008 have a maximum value of approximately \$533 million dollars, which is much higher than the highest ticket sales for 2007. Without this extreme value, the spreads between the lowest and highest values are approximately equal.

28	8
29	
30	
31	
32	8
33	0
34	178
35	00145678899
36	238999
37	0034566777
38	01124558
39	00259
40	045
41	2
42	2
	Legend: 34   1 = 34.1 years
	28 29 30 31 32 33 34 35 36 37 38 39 40 41 42

The distribution of median ages is centered at approximately 37 years old, with values ranging from 28.8 to 42.2 years. The distribution is approximately symmetric, with one unusual value of 28.8 years.

#### 2.27 (a)

Very Large Urban Area Large Urban Area 1 023478 2 369 3 0033589 8 99 4 0366 8711 5 012355 6 9730 7 2 8 9 3

Legend: 4|6 = 46 extra hours per year

(b) The statement "The larger the urban areas, the greater the extra travel time during peak period travel" is generally consistent with the data. Although there is overlap between the times for the very large and large urban areas, the extra travel times for the very large urban areas are generally greater than those for the large urban areas.





(c) The histograms are similar in shape. A notable difference is that the Credit Bureau data show that 7% of students have credit card balances of at least \$7,000, but no survey respondent indicated a balance of at least \$7,000.
(d) Yes, because students with credit card balances of \$7,000 or more might be too embarrassed to admit that they have such a high balance.

**2.29** (a) If the exam is quite easy, the scores would be clustered at the high end of the scale, with a few low scores for the students who did not study. The histogram would be negatively skewed.

(**b**) If the exam is difficult, the scores would be clustered around a much lower value, with only a few high scores. The histogram would be positively skewed.

(c) In this case, the histogram would be bimodal, with a cluster of high scores and a cluster of low scores.

#### **Additional Exercises**

**2.37** The distribution of wind speed is positively skewed and bimodal. There are peaks in the 35–40 m/s and 60–65 m/s intervals.

Frequency





In general, the Disney movies have longer tobacco exposure times than the non-Disney movies. Disney movies have a typical value of approximately 80 seconds, which is larger than a typical value for the non-Disney movies of 50 seconds. In addition, there is more variability in the Disney tobacco exposure times than the others. Disney movies vary between 6 and 548 seconds, which is greater than the observed spread for the non-Disney movies, which vary between 1 and 205 seconds. Finally, there appears to be one outlier in the Disney movies (548 seconds) and no outliers in the non-Disney movies.

## SECTION 2.4

#### **Exercise Set 1**



The scatterplot shows the expected positive relationship between grams of fat and calories. The relationship is weak.



As was observed in the calories versus fat scatterplot, there is also a weak, positive relationship between calories and sodium. The relationship between calories and sodium appears to be a little stronger than the calories versus fat relationship.



There is no apparent relationship between sodium and fat.



The lower-left region corresponds to healthier fast-food choices. This region corresponds to food items with fewer than 3 grams of fat and fewer than 900 milligrams of sodium.



There has been a steady downward trend in the percent of people who smoke among people who did not graduate from high school, from a high of 44% to 29% in 2005.



(c) There has been a steady downward trend in the percent of people age 25 or older who smoke, regardless of education level. In 1960, regardless of education level, the percentage was approximately the same (about 44–48%). Over time, however, the differences have become more pronounced. In 2005, those people with bachelor's degrees or higher had the lowest smoking rate (10%), followed by those with some college (21%). The highest rates of smoking were found among those who either did not graduate from high school (29%) or graduated high school but did not attend college (27%).

#### **Additional Exercises**



(**b**) The percentage of households with computers has increased over time, from a low of approximately 8% in 1985 to over 50% in 2000. The rate of increase has also increased over time.

**2.49** There is a relatively weak positive relationship between poverty rate and dropout rate. There are two states that have poverty rates between 10% and 15% but have much higher dropout rates (over 15%) compared to other states with comparable poverty rates.

## **SECTION 2.5**

#### **Exercise Set 1**

2.51 (a) categorical

(b) A bar chart was used because the response is a categorical variable, and dotplots are used for numerical responses.

(c) This is not a correct representation of the response data because the percent values add up to over 100% (they add to 107%).

**2.52** (a) Overall score is numerical. Grade is categorical. (b) The figure is equivalent to a segmented bar graph because the bar is divided into segments, with different shaded regions representing the different grades ("Top of the Class," "Passing," "Barely Passing," and "Failing"), and the height of each segment is equal to the frequency for that category (for example, there are five school districts in the Top of the Class category, three in the Passing category, and so on), making the area of each shaded region proportional to the relative frequencies for each grade.



One alternate assignment of grades is to require that "Top of the Class" schools earn grades of 72 or higher, "Passing" schools earn between 66 and 71, "Barely Passing" schools earn between 61 and 65, and "Failing" schools earn 60 or below. This alternative is suggested because there appear to be clusters of dots on the dotplot that correspond to the suggested ranges.

**2.53** Answers will vary. For example: For teens ages 12 to 17 years, the percentage of cell phone owners increases with age in each of the years 2004, 2006, and 2008. In addition, within each age group, the percentage of teens owning cell phones increased between 2004 and 2008. The largest increase in percentage of teens owning cell phones was among 12 year olds, and the smallest percentage increase was among 13 year olds.

#### **Additional Exercises**

**2.57** (a) The areas in the display are not proportional to the values they represent. The "no" category seems to represent more than 68%.



2.59 (a)





(c) The time-series plot best shows the trend over time.

## ARE YOU READY TO MOVE ON? CHAPTER 2 REVIEW EXERCISES

#### 2.61

Data Set 1: one variable; numerical; summarize the data distribution; a dotplot, stem-and-leaf display or histogram would be an appropriate graphical display.

Data Set 2: two variables; numerical; investigate the relationship; a scatterplot would be an appropriate graphical display.

Data Set 3: two variables; categorical; compare groups; comparative bar chart.

Data Set 4: one variable; categorical; summarize the data distribution; a bar chart would be an appropriate graphical display.

Data Set 5: One variable; numerical; compare groups; a comparative dotplot or comparative stem-and-leaf display would be an appropriate graphical display.



(**b**) The "In Default" bar in the "For-Profit Colleges" category is taller than either of the other "In Default" bars.

<b>2.65</b> (a)	0	8
	1	4467788999
	2	000011222333334445557778
	3	01222334457
	4	02458
		Legend: $1 4 = 14$ cents per gallon

(b) The center is approximately 24 cents per gallon, and most states have a tax that is near the center value, with tax values ranging from 8 cents per gallon to 48 cents per gallon. The distribution is approximately symmetric. (c) The only value that might be considered unusual is the 8 cents per gallon tax in Alaska.

2.67 (a)



#### (b) About 114.

(c) There is a lot of variability, with quality rating ranging between a low score of 84 defects per 100 vehicles and a high score of 170 defects.

(d) Two brands (Land Rover and Mitsubishi) seem to stand out as having a much greater value for number of defects. Four brands (Acura, Lexus, Mercedes-Benz, and Porsche) have smaller values for number of defects.



The histogram is centered at approximately 790, with values that range between approximately 720 and 880. The distribution is bimodal and positively skewed.



There is a weak negative relationship between customer satisfaction (as measured by the APEAL rating) and number of defects. Brands with a higher number of defects per 100 vehicles tend to have lower satisfaction ratings.



(b) The segmented bar graph in part (a) is more informative because it is easier to get a sense of the percentages of each ethnicity enrolled. Specifically, in the original graphical display, with the "Nonwhite" category further subdivided, it is difficult to compare the "Nonwhite" breakout categories with the other categories represented in the pie chart.
(c) The pie chart combined with the segmented bar graph could have been chosen because some of the pie slices might be very thin and hard to see, and too many pieces could be difficult to visually process.

**2.71** The display is misleading because the area principle is violated. The areas of the cocaine mounds are not proportional to the relative frequencies being represented.

# Chapter 3 Numerical Methods for Describing Data Distributions

## **SECTION 3.1**

#### **Exercise Set 1**

**3.1** The distribution is approximately symmetric with no outliers, so the mean and standard deviation should be used to describe the center and spread, respectively.

**3.2** The distribution is skewed with an outlier, so the median and interquartile range should be used.

**3.3** The distribution is skewed with a possible outlier, so the median and interquartile range should be used.

**3.4** The average may not be the best measure of a typical value for this data set because the distribution is clearly skewed.

#### **Additional Exercises**

**3.9** The distribution is skewed, so median and interquartile range should be used.

**3.11** The distribution is roughly symmetric with no obvious outliers, so the mean and standard deviation should be used.

## **SECTION 3.2**

#### Exercise Set 1

**3.12**  $\bar{x} = 51.33$  ounces; this is a typical value for the amount of alcohol poured. s = 15.22 ounces; this represents how much, on average, the values in the data set spread out, or deviate, from the mean.

**3.13** (a)  $\bar{x} = 59.23$  ounces; this is a typical value for the amount of alcohol poured. s = 16.71 ounces; this represents how much, on average, the values in the data set spread out, or deviate, from the mean.

(**b**) Individuals pouring alcohol into short, wide glasses pour, on average, more alcohol than when pouring into tall, slender glasses.

#### **3.14** (a) $\bar{x} = 59.85$ hours, s = 14.78 hours

(**b**)  $\bar{x} = 56.67$  hours, s = 9.75 hours. When Los Angeles was excluded from the data set, the mean and standard deviation both decreased quite a bit. This suggests that using the mean and standard deviation as measures of center and spread for data sets with outliers can be risky, because outliers seem to have a significant impact on those measures.

**3.15** Answers will vary. One possible answer: The mean is \$444, but it is likely that some parents spend only a small amount. There is probably a lot of variability in amount spent, so we would expect a large value for the standard deviation.

#### **Additional Exercises**

**3.21** (a)  $\bar{x} = 287.714$ ; the deviations from the mean are 209.286, -94.714, 40.286, -132.714, 38.286, -42.714, -17.714.

(c)  $s^2 = 12,601.905, s = 112.258.$ 

**3.23** The deviations are exactly the same as the corresponding deviations for the original data set. Since the deviations are the same, the new variance and standard deviation are also the same as the old variance and standard deviation. Subtracting the same number from or adding the same number to every value in a data set does not change the value of the variance or standard deviation.

## **SECTION 3.3**

#### **Exercise Set 1**

**3.25** (a) median = 433,246.5. Half of the newspapers had average weekly circulations of less than 433,246.5, and the other half had average weekly circulations of more than 433,246.5.

(**b**) The median is preferable because the distribution is skewed and contains outliers.

(c) It is not reasonable to generalize from this sample to the population of daily U.S. newspapers because these newspapers were not randomly selected. They are the top 20 American newspapers in average weekday circulation.

**3.26** Lower quartile = 10,478; 25% of the catsups have sodium contents lower than 10,478. Upper quartile = 11,778; 75% of the catsups have sodium contents lower than 11,778. The interquartile range is iqr = 11,778 - 10,478 = 1,300; the range of the middle 50% of the catsup sodium contents is 1,300.

**3.27** median = 142; half of the values of number of minutes used in cell phone calls in one month are less than or equal to 142 minutes, and half of the data values of number of minutes used in cell phone calls are greater than or equal to 142 minutes. iqr = 195; the middle 50% of the data values have a range of 195 minutes.

**3.28** median = 21%; half of the tips were below 21%, and the remaining half were above 21%. *iqr* = 24.85%; the middle 50% of tips had a range of 24.85%.

#### **Additional Exercises**

**3.33** The large difference between the mean and median indicates that there were some parents who spent large amounts of money on school supplies.

**3.35** (a) mean will be greater than the median. (b)  $\bar{x} = 370.69$  seconds, *median* = 369.5 seconds. (c) The largest time could be increased by any amount and not affect the sample median because the position of the middle value will not change if the largest value is increased. The largest time could be decreased to 370 seconds without changing the value of the median.

## **SECTION 3.4**

#### **Exercise Set 1**

**3.36** Minimum = 0, lower quartile = 14, median = 33.5, upper quartile = 63, maximum = 151.



The boxplot shows that there is one outlier (170 defects), and the value of the largest non-outlier is 146 defects. The middle 50% of the data values range between about 106 and 126 defects. The distribution is positively skewed.

**3.38** (a) No, they are not outliers. For this data set, values are outliers if they are greater than 32.3 + 1.5(32.3 - 20) = 50.75 cents, or less than 20 - 1.5(32.3 - 20) = 1.55 cents



The distribution is positively skewed.

**3.39** (a) lower quartile = 16.05 inches, upper quartile = 21.93 inches. iqr = 21.93 - 16.05 = 5.88 inches. (b) For this data set, values are outliers if they are greater than 21.93 + 1.5(5.88) = 30.75 inches or less than 16.05 - 1.5(5.88) = 7.23 inches. The value 31.57 inches is an outlier.



The modified boxplot shows one outlier at the high end of the scale. The distribution of inches of rainfall is slightly positively skewed.

#### 3.40



Both distributions (short, wide and tall, slender) are skewed, although the direction of skew is different for the two distributions. The amount of alcohol poured into short, wide glasses tends to be more than the amount poured into tall, slender glasses.

#### **Additional Exercises**



(b) The most noticeable difference between the wireless percent for the three geographical regions is that the Middle States region is negatively skewed and has a smaller interquartile range than the East and West regions. The Eastern region has the smallest median (11.4%), and the Middle States and Western regions have medians that are about the same (16.9% and 16.3%).

**3.49** The fact that the mean is so much higher than the median indicates that the distribution is positively skewed.

## **SECTION 3.5**

#### **Exercise Set 1**

**3.50** First national aptitude test: z = 1.5. Second national aptitude test: z = 1.875. The student performed better on the second national aptitude test relative to the other test takers because the *z*-score for the second test is higher than for the first test.

**3.51** (a) 40 minutes is 1 standard deviation above the mean; 30 minutes is 1 standard deviation below the mean. The values that are 2 standard deviations away from the mean are 25 and 45 minutes.

(**b**) Approximately 95% of times are between 25 and 45 minutes; approximately 0.3% of times are less than 20 minutes or greater than 50 minutes; approximately 0.15% of times are less than 20 minutes.

**3.52** The 10th percentile of \$0 indicates that 10% of students have \$0 or less of student debt. The 25th percentile (which is the lower quartile) indicates that 25% of students have \$0 or less of student debt. The 50th percentile (the median) indicates that 50% of students have \$11,000 or less of student debt. The 75th percentile (the upper quartile) indicates that 75% of students have \$24,600 or less of student debt. The 90th percentile indicates that 90% of students have \$39,300 or less of student debt.





(**b**) Note: percentiles were estimated using midpoints of the histogram bar intervals; (i) 86th percentile is approximately 20.5 minutes; (ii) 15th percentile is approximately

17.5 minutes; (iii) 90th percentile is approximately 21.5 minutes; (iv) 95th percentile is approximately 25.5 minutes; (v) 10th percentile is approximately 17.5 minutes

#### **Additional Exercises**

**3.59** (a) 1,100 gallons; (b) 1,400 gallons; (c) 1,700 gallons

**3.61** (a) 120 (b) 20 (c) -0.5 (d) 97.5 (e) Since a sco

(e) Since a score of 40 is 3 standard deviations below the mean, that corresponds to a percentile of 0.15%. Therefore, there were relatively few scores below 40.

## ARE YOU READY TO MOVE ON? CHAPTER 3 REVIEW EXERCISES

**3.63**  $\bar{x} = 792.03$  which is a typical or representative value for the APEAL rating. s = 36.70 which represents how much, on average, the values in the data set spread out, or deviate, from the mean APEAL rating.

#### **3.65** (a) $\bar{x} = 27.31\%$ , s = 23.83%

(b) After removing the 105% tip, the new mean and standard deviation are  $\bar{x}_{new} = 23.23\%$  and  $s_{new} = 15.70\%$ . These values are much smaller than the mean and standard deviation computed with 105 included. This suggests that the mean and standard deviation can change dramatically when outliers are present (or removed) from the data set and, therefore, are probably not the best measures of center and spread to use in this situation.

**3.67** (a) median = 140 seconds; half the values are less than 140 seconds and half the values greater than 140 seconds. iqr = 100 seconds; the middle 50% of the data values have a range of 100 seconds. (b) There is an outlier in the data set.

**3.69** (a) Median = 8 grams/serving; lower quartile = 7 grams/serving; upper quartile = 12 grams/serving; interquartile range = 12 - 7 = 5 grams/serving (b) Median = 10 grams/serving; lower quartile = 6 grams/serving; upper quartile = 13 grams/serving; interquartile range = 13 - 6 = 7 grams/serving (c) There are no outliers in the sugar content data. (d) The minimum value and lower quartile are the same because the smallest five values in the data set are all equal to 7.



The sugar content in grams/serving is much more variable than the fiber content in grams/serving. The boxplot of fiber content shows that the minimum and lower quartiles are equal to each other, which is not observed in the sugar content. The distribution of sugar content values is approximately symmetric, which is different from the skewed fiber content distribution.

**3.71** (a) The 25th percentile indicates that 25% of fulltime female workers age 25 or older with an associate degree earn \$26,800 or less. The 50th percentile indicates that 50% of full-time female workers age 25 or older with an associate degree earn \$36,800 or less. The 75th percentile indicates that 75% of full-time female workers age 25 or older with an associate degree earn \$51,100 or less.

(**b**) The 25th, 50th, and 75th percentile values for men are all greater than the corresponding percentiles for female workers, indicating that full-time employed men age 25 or older with an associate degree, in general, earn more than full-time employed women age 25 or older with an associate degree.

# **Chapter 4 Describing Bivariate Numerical Data**

## **SECTION 4.1**

#### Exercise Set 1

4.1 Scatterplot 1: (i) Yes (ii) Yes (iii) Negative Scatterplot 2: (i) Yes (ii) No (iii) – (i) Yes (ii) Yes (ii) Positive Scatterplot 4: (i) Yes (ii) Yes (iii) Positive

**4.2** (a) Negative correlation, because as interest rates rise, the number of loan applications might decrease.

(**b**) Close to zero, because there is no reason to believe that height and IQ should be related.

(c) Positive correlation, because taller people tend to have larger feet.

(d) Positive correlation, because as the minimum daily temperature increases, the cooling cost would also increase.

**4.3** (a) There is a moderately strong positive linear relationship between school achievement test score and midlife IQ.

(b) r = 0.6, because the article says that r = 0.64 indicated a very strong relationship, higher than the correlation between height and weight in adults. Therefore, a correlation that is moderately strong (r = 0.6) with a positive association (taller people tend to weigh more) is consistent with the statement. **4.4** (a) r = 0.335; there is a weak, positive linear relationship.

(**b**) The conclusion that "heavier logging led to large forest fires" cannot be justified because correlation does not imply causation.



There is a weak, negative linear relationship between satisfaction rating and quality rating (number of defects). **(b)** r = -0.239; there is a weak, negative linear relationship between satisfaction rating and quality rating (number of defects).

**4.6** No, the statement is not correct. A correlation of 0 indicates that there is not a linear relationship between two variables. There could be a strong nonlinear relationship (for example, a quadratic relationship) between the two variables.

**4.7** No, it is not reasonable to conclude that increasing alcohol consumption will increase income. Correlation measures the strength of association, but association does not imply causation.

4.8 The correlation between college GPA and academic self-worth (r = 0.48) indicates that there is a weak or moderate positive linear relationship between those variables. This tells us that athletes with higher a GPA tend to feel better about themselves academically than those with lower grades. The correlation between college GPA and high school GPA (r = 0.46) indicates that there is a weak or moderate positive relationship between those variables as well. This tells us that those athletes with higher high school GPA tend to also have a higher college GPA. Finally, the correlation between college GPA and a measure of tendency to procrastinate (r = -0.36) indicates that there is a weak negative linear relationship between those variables. Athletes with a lower college GPA tend to procrastinate more than athletes with a higher college GPA.

#### **Additional Exercises**

**4.15** (a) (b) *r* = 0.001

(c) The correlation coefficient of r = 0.001 indicates that there is essentially no linear relationship between mare weight and foal weight. The scatterplot shows that there is no obvious relationship (linear or otherwise) between mare weight and foal weight.

**4.17** r = 0.987; this value is consistent with the previous answer, because the correlation coefficient is large (close to 1) and positive, which indicates a strong positive association between household debt and corporate debt.

**4.19** The sample correlation coefficient would be closest to -0.9. Cars traveling at a faster rate of speed will travel the length of the highway segment more quickly than those who are traveling more slowly, and the correlation would be strong.

## **SECTION 4.2**

#### **Exercise Set 1**

**4.21** It makes sense to use the least squares regression line to summarize the relationship between x and y for Scatterplot 1, but not for Scatterplot 2. Scatterplot 1 shows a linear relationship between x and y, but Scatterplot 2 shows a curved relationship between x and y.

**4.22** It would be larger because the least squares regression line is *the* line with the minimum value for the sum of the squared vertical deviations from the line. All other lines would have larger values for the sum of the squared vertical deviations.

**4.23** (a)  $\hat{y} = -5.0 + 0.017x$ 

**(b)** 30.7 therms.

(c) 0.017 therms.

(d) No, because the regression line was determined based on house sizes between 1000 and 3000 square feet. There is no guarantee that the linear relationship will continue outside this range of house sizes.

**4.24** (a) The response variable (y) is birth weight, and the predictor variable (x) is mother's age.

(b) It is reasonable to use a line to summarize the relationship because the scatterplot shows a clear linear relationship between birth weight and mother's age. (c)  $\hat{y} = -1163.4 + 245.15x$ 

(d) The slope of 245.15 is the amount, on average, by which the birth weight increases when the mother's age increases by one year.

(e) It is not appropriate to interpret the intercept of the least squares regression line. The intercept is the birth weight for a mother who is zero years old, which is impossible. In addition, the intercept is negative, indicating a negative birth weight, which is also impossible.

(**f**) 3,249.3 grams

(g) 2,513.85 grams

(h) No, because 23 years is well outside the range of ages used in determining the least squares regression line.

4.25 (a) The response variable is the cost of medical care, and the predictor variable is the measure of pollution.
(b) There is a moderate, negative linear relationship.
(c) ŷ = 1082.2 - 4.691x

(d) The slope is negative, and it is consistent with the observed negative association in the scatterplot.

(e) No, the association between medical cost and pollution level is negative, which indicates that people over the age of 65 in more polluted areas tend to have lower medical costs. (f) \$918.02

(g) No, because the value 60 is far outside the range of data values in the sample.

**4.26** (a)  $\hat{y} = 11.48 + 0.970x$ (b) 496.48. (c) 302.48

#### **Additional Exercises**

**4.33**  $\hat{y} = 13.5 - 0.195x$ 

**4.35** Age is a better predictor of number of cell phone calls. The linear relationship between age and number of cell phone calls is stronger than the relationship between age and number of text messages sent.

**4.37** The slope is the change in predicted price for each additional mile from the Bay, so the slope would be -4,000.

## **SECTION 4.3**

#### **Exercise Set 1**

**4.39** (a)  $\hat{y} = 1.33878 - 0.007661x$ 

(b)  $r^2 = 0.099$ ; Approximately 9.9% of the variability in telomere length can be explained by the linear relationship between telomere length and perceived stress. (c)  $s_e = 0.159$ ; a typical amount by which telomere length will deviate from the least squares regression line is 0.159.

(d) negative; because the slope of the least squares regression line is negative. weak; because r = -0.315.

**4.40** A small value of  $s_e$  indicates that residuals tend to be small. Because residuals represent the difference between an observed y value and a predicted y value, the value of  $s_e$  tells us how much accuracy we can expect when using the least squares regression line to make predictions.

**4.41** It is important to consider both  $r^2$  and  $s_e$  when evaluating the usefulness of the least squares regression line because a large  $r^2$  (which indicates the proportion of variability in *y* that can be explained by the linear relationship between *x* and *y*) tells us that knowing the value of *x* is helpful in predicting *y*, and a small  $s_e$  indicates that residuals tend to be small.

**4.42** (a) Yes, the scatterplot looks reasonably linear.



**(b)**  $\hat{y} = 492.80 + 14.763x$ 

(c) The residuals are -7.55, -12.14, 26.87, 9.00, and -16.17. There is a curved pattern in the residual plot. The curvature indicates that the relationship between median distance walked and representative age is not linear.



**4.43** (a) The pattern for girls differs from boys in that the girls' scatterplot shows more apparent nonlinearity in the pattern.

**(b)**  $\hat{y} = 480 + 12.525x$ 

(c) The residuals are -37.70, 10.63, 50.56, 8.52, -32.01. The curvature of the residual plot indicates that a curve is more appropriate than a line for describing the relationship between median distance walked and representative age.



**4.44** (b)  $\hat{y} = -0.03443 + 0.5803x$ . The predicted nitrogen retention for a flying squirrel whose nitrogen intake is 0.06 grams is 0.000388 grams. The residual associated with the observation (0.06, 0.01) is 0.009612. (c) The observation (0.25, 0.11) is potentially influential because that point has an *x* value that is far away from the rest of the data set.

(d)  $\hat{y} = -0.037 + 0.627(0.06) = 0.00062$ ; This prediction is larger than the prediction made in Part (b).

**4.45** (a) There appears to be a linear relationship. (b)  $\hat{y} = 18.483 + 0.0028655x$ 

(c) The observation (3928, 46.8) is not influential, because the *x*-value for that observation is not far from the rest of the data. In addition, removal of the potentially influential point produces a least squares regression line with a *y*-intercept and slope similar to the original line.

(d) Those points are not considered influential even though they are far from the rest of the data because they follow the trend of the remaining data points. Removal of those points would produce a least squares regression line similar to the line found using the full data set.

(e)  $s_e = 9.16217$ ; A typical deviation from the least squares regression line is 9.16217 percentage points.

(f)  $r^2 = 0.832$ ; Approximately 83.2% of the variability in percentage transported can be explained by the linear relationship between percentage transported and number of salmon.

## ARE YOU READY TO MOVE ON? CHAPTER 4 REVIEW EXERCISES

4.57	Scatterplot 1:	(i) Yes (ii) Yes (iii) Negative
	Scatterplot 2:	(i) Yes (ii) No (iii) –
	Scatterplot 3:	(i) Yes (ii) Yes (iii) Positive
	Scatterplot 4:	(i) Yes (ii) Yes (iii) Positive

**4.59** (a) r = -0.10; there is a weak, negative linear relationship. Because the relationship is negative, larger arch heights tend to be paired with smaller, average hopping heights.

(b) The correlations coefficients support the conclusion since they are all fairly close to 0.

**4.61** (a) r = 0.944; there is a strong positive linear relationship between sugar consumption and depression rate. (b) No, because you can't conclude that a cause-and-effect relationship exists just based on a strong correlation. (c) These countries were not a random sample of all countries, and it is unlikely that they are representative.

**4.63** (a) negative, because it is likely that the work is more stressful and less enjoyable for nurses with high patient-to-nurse ratios.

(**b**) negative, because it is likely that the patients at hospitals where the patient-to-nurse ratio is high will not get as much individual attention and will be less satisfied with their care.

(c) negative, because it is likely that the quality of patient care suffers when patient-to-nurse ratio is high.

**4.65** There is no evidence that the form of the relationship between x and y remains the same outside the range of the data.

**4.67** (a) size, because the relationship between price and size (r = 0.700) is stronger than the relationship between price and land-to-building ratio (r = -0.332). (b)  $\hat{y} = 1.33 + 0.00525x$ 

**4.69** (a)  $\hat{y} = 14.2 + 0.790x$  where  $\hat{y}$  is the predicted number of transplants and *x* is the year (using 1, 2, ..., 10 to represent the years). The number of transplants has increased steadily over time, with a predicted increase of about 790 per year.

(**b**) The residual plot shows a curved pattern, suggesting that the relationship between year and number of transplants is nonlinear.

**4.71** (a) No, because the value of outpatient cost-to-charge ratio (54) is not far from the other outpatient values in the data set.

(**b**) Yes, it is an outlier, because it has a large residual—it is far away from the least squares regression line.

(c) No, because even though it has an outpatient value that is larger than the others in the data set, this data point is consistent with the linear pattern formed by the other data values.

(d) No, because it falls quite close to the least squares regression line. This data point has a small residual.

# **Chapter 5 Probability**

## **SECTION 5.1**

#### **Exercise Set 1**

5.1 (a) In the long run, about 1% of people who suffer cardiac arrest in New York City will survive.(b) Because 1% of 2329 is 23.29, approximately 23 people

who suffered cardiac arrest in New York City survived.

**5.2** In the long run, 86% of the time this particular flight that flies between Phoenix and Atlanta will arrive on time.

**5.3** (a) 0.83 (b) 0.83 (c) 0.17 (d) 0.48 (e) 0.52

#### **Additional Exercises**

**5.7** Approximately 167; *P*(observing six when a fair die is rolled) =  $\frac{1}{6}$ 

**5.9** (a) 0.17 (b) 0.85 (c) 0.15 (d) 0.75

## **SECTION 5.2**

#### Exercise Set 1

**5.11** (a) 0.225 (b) 0.775 (c) 0.06 (d) 0.715 (e) 0.234 (f) The airline should not be particularly worried. The probability of selecting one person who was delayed overnight is approximately  $\frac{75}{8000} = 0.0094$ , and the probability of selecting additional people who were delayed overnight decreases from 0.0094.

**5.12** (a)  $S = \{AB, AC, AD, AE, BC, BD, BE, CD, CE, DE\}$ (b) Yes. (c) 0.3

(d) No, the probability does not change. (e) The probability increases to 0.6.

**5.13** 0.11

#### **Additional Exercises**

**5.17** (a) 0.62 (b) 0.38

**5.19** (a) 0.0119 (b) 0.000002 (c) 0.012

**5.21** (a)  $S = \{AB, AC, AD, AE, AF, BC, BD, BE, BF, CD, CE, CF, DE, DF, EF\}$  (b) Yes. (c) 0.0667 (d) 0.4 (e) 0.533

## **SECTION 5.3**

#### **Exercise Set 1**

**5.22** (a) Not mutually exclusive, because there are seniors who are computer science majors.

(**b**) Not mutually exclusive, because there are female students who are computer science majors.

(c) Mutually exclusive, because a student cannot have a college residence that is more than 10 miles from campus and live in a college dormitory that is on campus.
(d) Mutually exclusive, because college football teams are male teams and females would not be on the team.

**5.23** The events are dependent because the probability that a male experiences pain daily is not equal to the probability that a female experiences pain daily.

**5.24** (a) 0.70 (b) 0.189 (c) 0.5 (d) 0.811 (e) 0.799 (f) 0.101

#### 5.25 (a)

	<i>E</i> (stop at first light)	<i>Not E</i> (does not stop at first light)	Total
F (stop at second light)	150	150	300
at second light) Total	250 <b>400</b>	450 <b>600</b>	700 1000

(b) (i) 0.55 (ii) 0.45 (iii) 0.40 (iv) 0.25

#### 5.26 (a)

	Cable TV	Not Cable TV	Total
Internet Service	250	170	420
Not Internet Service	550	30	580
Total	800	200	1000

#### **(b) (i)** 0.25 **(ii)** 0.72

#### **5.27** (a) (i) 0 (ii) 0.64

(b) No, events *A* and *B* are not mutually exclusive. For mutually exclusive events,  $P(A \cup B) = P(A) + P(B)$ . If *A* and *B* were mutually exclusive, then  $P(A \cup B) = P(A) + P(B) = 0.26 + 0.34 = 0.60$ , which is different from the value of  $P(A \cup B)$  that is given.

#### **5.28** 0.000625

#### 5.29 (a)

	Cash (M)	Credit Card (not M)	Total
Extended Warranty (E)	75	85	160
No Extended Warranty (not E)	395	445	840
Total	470	530	1000

(**b**) 0.555. In the long run, approximately 55.5% of customers will either pay with cash or purchase an extended warranty, or both.

#### **Additional Exercises**

**5.39** If the events *T* and *P* were mutually exclusive, you could add the probabilities to obtain  $P(T \cup P)$ . Since P(T) + P(P) = 0.83 + 0.56 = 1.39 is greater than 1 (and probabilities cannot be greater than 1), you know that *T* and *P* cannot be mutually exclusive.

**5.41** (a) No, the events *O* and *N* are not independent. If they were independent,  $P(O \cap N) = P(O) \cdot P(N) = (0.7)(0.07) = 0.049$ , which is not equal to the value of  $P(O \cap N)$  given in the exercise.

**(b)** 

	Ν	Not N	Total
0	40	660	700
Not O	30	270	300
Total	70	930	1000

(c) 0.730. In the long run, 73% of the time airline ticket purchasers will buy their ticket online or not show up for a flight, or both.

## **SECTION 5.4**

#### Exercise Set 1

**5.43** (a) The conditional probability is the 80%, because you are told that 80% *of those receiving such a citation* attended traffic school. The key phrase is "of those receiving such a citation," which is what indicates that the 80% is a conditional probability. (b) P(F) = 0.20 and P(E|F) = 0.80

**5.44** (a) (i) 0.62 (ii) 0.36 (iii) 0.40 (iv) 0.29 (v) 0.31 (b) (i) The probability that a randomly selected Honda Civic buyer is male. (ii) The probability that a randomly selected Honda Civic buyer purchased a hybrid. (iii) The probability that, of the males, a randomly selected buyer purchased a hybrid. (iv) The probability that, of the females, a randomly selected buyer purchased a hybrid. (v) The probability that, of the hybrid purchasers, a randomly selected buyer is female.

(c) These probabilities are not equal.

**5.45** (a) P(L) = 0.493: The probability that the goalkeeper jumps to the left.

P(C) = 0.063: The probability that the goalkeeper stays in the center.

P(R) = 0.444: The probability that the goalkeeper jumps to the right.

P(B|L) = 0.142: The probability that, given that the goalkeeper jumps to the left, the goal was blocked.

P(B|C) = 0.333: The probability that, given that the goalkeeper stays in the center, the goal was blocked.

P(B|R) = 0.126: The probability that, given that the goalkeeper jumps to the right, the goal was blocked. **(b)** 

	Blocked (B)	Not Blocked (not B)	Total
Goalkeeper Jumped Left (L)	70	423	493
Goalkeeper Stayed Center (C)	21	42	63
Goalkeeper Jumped Right (R)	56	388	444
Total	147	853	1000

(c) 0.147.

(d) It seems as if the best strategy is for the goalkeeper to stay in the center when defending a kick.

5.46 (a)

	Uses Alternative Therapies	Does Not Use Alternative Therapies	Total
HS/Less	315	7,005	7,320
College: 1–4 yrs	393	4,400	4,793
College: ≥5 yrs	120	975	1,095
Total	828	12,380	13,208

(b) (i) 0.083 (ii) 0.063 (iii) 0.110 (iv) 0.043 (v) 0.380 (vi) 0.087 (c) For *H* and *A* to be independent,  $P(H \cap A) = P(H) \cdot P(A)$ .  $P(H \cap A) = \frac{315}{13,208} = 0.024$  and  $P(H) \cdot P(A) = (0.554)(0.063) = 0.0349$ , which is not equal to 0.024. The events *H* and *A* are not independent.

#### **Additional Exercises**

**5.51** P(A|B) would be larger than P(B|A) because it is more likely that a professional basketball player is over 6 feet tall than for an individual who is over 6 feet tall to be a professional basketball player.

**5.53** (a) 0.490 (b) 0.3125 (c) 0.240 (d) 0.100

#### **5.55** (a) 0.46

(b) It more likely that he or she is in the first priority group.
(c) It is not particularly likely that a student in the third priority group would get more than nine units during the first attempt to register because, of those in the third priority group, only 36% of students get more than nine units.

## **SECTION 5.5**

#### **Exercise Set 1**

**5.57** Use of the following conditional probabilities can provide justification for the given conclusion. P(very harmful|current smoker) = 0.625, P(very harmful|former smoker) = 0.788, and P(very harmful|never smoked) = 0.869.

## **SECTION 5.6**

#### **Exercise Set 1**

**5.59** (a) 0.85 (b) 0.19 (c) 0.6889 (d) 0.87

#### **5.60** (a) Answers will vary.

(b) This is not a fair way of distributing licenses because those companies/individuals who are requesting multiple licenses are given approximately the same chance to get two or three licenses as an individual has to get a single license. Perhaps companies who request multiple licenses should be required to submit one application per license.

#### **Additional Exercises**

**5.63** Results from the simulation will vary. The exact answer is 0.8468 (which can be verified using a tree diagram).

**5.65** (a) Results from the simulation will vary. Based on one simulation, P(more than five pairs) = 0.43. (b) Results from the simulation will vary. Based on one simulation, P(researchers will incorrectly conclude that Treatment 2 is better) = 0.17.

## ARE YOU READY TO MOVE ON? CHAPTER 5 REVIEW EXERCISES

**5.67** This probability interpretation is misleading because it demonstrates a misunderstanding of the law of averages. A better interpretation is that on average, a passenger will be safe on 11,000,000 flights. Some passengers will be safe on more flights, and others on fewer flights.

#### **5.69** (a) 0.07 (b) 0.30. (c) 0.57

**5.71** (i) P(win game 1) = 0.02 and P(win game 2) = 0.022, therefore it is advantageous to play game 2 because the probability of winning is higher. (ii) In playing either of these games 100 times, you would expect to lose money overall.

**5.73** The outcomes *selected smoker who is trying to quit uses a nicotine aid* and *selected smoker who has attempted to quit begins smoking again within two weeks* are dependent events. An individual smoker who is trying to quit smoking is somewhat more likely to begin smoking again within two weeks if she or he does not use a nicotine aid (62%) than if she or he does use a nicotine aid (60%).

**5.75** (a) P(S) = 0.987, P(A) = 0.354, and  $P(S \cap A) = 0.347$ 

(b)

	Specialist Opinion (S)	No Specialis Opinion (not S)	t Total
Pathology Results (A)	347	7	354
No Pathology Results (not A)	640	6	646
Total	987	13	1000

(c) (i) 0.006 (ii) 0.994

#### **5.77** (a) (i) 0 (ii) 0.81

(**b**) *A* and *B* are not mutually exclusive events. If they were mutually exclusive,  $P(A \cap B) = 0$ ,

(c) *A* and *B* are not mutually exclusive events. If they were mutually exclusive,  $P(A \cup B) = P(A) + P(B) = 0.65 + 0.57 = 1.22$ , but probabilities can't be greater than 1.

#### 5.79 (a)

	Expedited Shipping (E)	No Expedited Shipping (not E)	Total
Gift Wrap (G)	31	89	120
No Gift Wrap (not G)	229	651	880
Total	260	740	1000

(**b**) 0.349. This is the probability that, in the long run, a customer will request expedited shipping or gift wrap, or both.

**5.81** (a) There are two conditional probabilities. One is the 39.9% who say they sometimes use Twitter to communicate with students, and the other is the 27.5% who say they sometimes use Twitter as a learning tool in the classroom. The probabilities are conditional because the description of them begins with the phrase "*Of those who use Twitter*..." (b) (i) 0.307 (ii) 0.693 (iii) 0.399 (iv) 0.275 (c)

	Uses Twitter (T)	Does Not Use Twitter ( <i>not T</i> )	Total
Communicates with Students (C)	122	0	122
with Students ( <i>not C</i> )	185 <b>307</b>	693 <b>693</b>	878 1000

$$P(C) = \frac{122}{1000} = 0.122.$$

(**d**)

	Uses Twitter (T)	Does Not Use Twitter (not T)	Total
Learning Tool in Classroom (L)	84	0	84
Does Not Use as Learning Tool in Classroom ( <i>not L</i> ) Total	223 <b>307</b>	693 <b>693</b>	916 1000

$$P(L) = \frac{84}{1000} = 0.084$$

5.83 (a)

	Confidence High ( <i>H</i> )	Confidence Low ( <i>not H</i> )	Total
Diagnosis Correct (C)	98	163	261
Diagnosis Incorrect (I)	54	685	739
Total	152	848	1000

**(b)** 
$$P(C|H) = \frac{98}{152} = 0.645$$

(c)

	Confidence High ( <i>H</i> )	Confidence Low ( <i>not H</i> )	Total
Diagnosis Correct (C)	266	229	495
Diagnosis Incorrect (I)	127	378	505
Total	393	607	1000

$$P(C|H) = \frac{266}{393} = 0.677$$

This probability is higher than the probability computed in Part (b). This indicates that a faculty member with high confidence in a correct diagnosis is more likely to be correct than a medical student who is also highly confident in her/his diagnosis.

## Chapter 6 Random Variables and Probability Distributions

NOTE: Your numerical answers may sometimes differ slightly from those in the answer section, depending on whether a table, a graphing calculator, or statistical software is used to compute probabilities.

## **SECTION 6.1**

#### **Exercise Set 1**

6.1 (a) discrete (b) continuous (c) discrete (d) discrete (e) continuous

**6.2** The possible values for x are x = 1, 2, 3, ... (the positive integers).

Answers will vary. One possible answer is

Outcomes	x
S	1
LS	2
RLS	3
RRS	3
LRLRS	5

**6.3** (a) 3, 4, 5, 6, 7 (b) -3, -2, -1, 1, 2, 3 (c) 0, 1, 2 (d) 0, 1

#### **Additional Exercises**

- 6.7 (a) discrete (b) continuous (c) continuous (d) discrete
- **6.9** (a) 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

**(b)** -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5 **(c)** 1, 2, 3, 4, 5, 6

## **SECTION 6.2**

#### **Exercise Set 1**

**6.10** (a) p(4) = 0.01

(b) It is the probability of randomly selecting a carton of one dozen eggs and finding exactly 1 broken egg.
(c) P(y ≤ 2) = 0.95; the probability that a randomly selected carton of eggs contains 0, 1, or 2 broken eggs is 0.95.
(d) P(y < 2) = 0.85; it is smaller because it does not include the possibility of 2 broken eggs.</li>
(e) Exactly 10 unbroken eggs is equivalent to exactly 2 broken eggs; p(2) = 0.10.

(f) At least 10 unbroken eggs is equivalent to 10, 11, or 12 unbroken eggs, or 2, 1, or 0 broken eggs;  $P(y \le 2) = 0.95$ .

6.11 (a) For 1,000 graduates, you would expect to see approximately 450 graduates who contributed nothing, 300 graduates to contribute \$10, 200 graduates to contribute \$25, and 50 graduates to contribute \$50.
(b) \$0 (c) 0.25 (d) 0.55

**6.12** (a) (1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4) (b) Because the bottles are randomly selected, the outcomes are equally likely and each has probability 1/6.



(b) The most likely outcomes are 0 and 1. (c)  $P(x \ge 2) = 0.1808$ 

#### **Additional Exercises**

**6.19** (a)  $k = \frac{1}{15}$  (b)  $P(y \le 3) = 0.4$ (c)  $P(2 \le y \le 4) = 0.6$ 

## **SECTION 6.3**

#### **Exercise Set 1**







(b) The probability that a randomly selected individual waits between 4 and 7 minutes for service at a bank is 0.26.

**6.23** (a) h = 2 (b) P(x > 0.5) = 0.25(c)  $P(x \le 0.25) = 0.4375$ **6.24** (a)  $\frac{1}{2}(0.40)(5) = 1$ (b) P(x < 0.20) = 0.5; P(x < 0.1) = 0.125; P(x > 0.3) = 0.125(c) P(0.10 < x < 0.20) = 0.375

#### **Additional Exercises**

**6.29** The probability P(x < 1) is the smallest. P(x > 3) and P(2 < x < 3) are equal, and larger than the other two probabilities.

**6.31**  $P(2 < x < 3) = P(2 \le x \le 3) < P(x < 2) < P(x > 7)$ . The smallest two probabilities are both equal to 1/10, the third probability is equal to 2/10, and the fourth probability is equal to 3/10.

## **SECTION 6.4**

#### Exercise Set 1

**6.32** (a)  $\mu_y = 0.56$ ; this is the mean value of the number of broken eggs in the population of egg cartons. (b) P(y < 0.56) = P(y = 0) = 0.65. This is not particularly surprising because, in the long run, 65% of egg cartons contain no broken eggs.

(c) This computation of the mean is incorrect because it assumes that the numbers of broken eggs (0, 1, 2, 3, or 4) are all equally likely.

#### **6.33** (a) $\mu_{\rm r} = 16.38, \sigma_{\rm r} = 1.9984$

(b) The mean,  $\mu_x = 16.38$  cubic feet represents the longrun average storage space of freezers sold by this particular appliance dealer. The standard deviation,  $\sigma_x = 1.9984$  cubic feet, represents a typical amount by which the storage space in freezers purchased deviates from the mean.

**6.34** Answers may vary. Two possible probability distributions are shown below.

Probability Distribution 1 (  $\mu_x = 3$  and  $\sigma_x = 1.265$ ):

x	1	2	3	4	5
<b>р(</b> х)	0.15	0.20	0.30	0.20	0.15

Probability Distribution 2 ( $\mu_{y} = 3$  and  $\sigma_{y} = 1.643$ ):

x	1	2	3	4	5
р(х)	0.30	0.15	0.10	0.15	0.30

#### Additional Exercises

**6.39** (a)  $\mu_x = 2.3$  represents the long-run average number of lots ordered per customer.

(**b**)  $\sigma_x^2 = 0.81$ ,  $\sigma_x = 0.9$  lots. A typical deviation from the mean is about 0.9 lots.

### **SECTION 6.5**

#### **Exercise Set 1**

6.41 (a) 0.9599
(b) 0.2483
(c) 0.1151
(d) 0.9976
(e) 0.6887
(f) 0.6826
(g) approximately 1

6.42 (a) 0.9909
(b) 0.9909
(c) 0.1093
(d) 0.1267
(e) 0.0706
(f) 0.0228
(g) 0.9996
(h) approximately 1

6.43 (a) 0.5
(b) 0.9772
(c) 0.9772
(d) 0.8185
(e) 0.9938
(f) approximately 1

**6.44** (a) At most 60 wpm: 0.5; Less than 60 wpm: 0.5 (b) 0.8185

(c) 0.0013; it would be surprising, because the probability of finding such a typist is very small.

(d) The probability that a randomly selected typist has a typing speed that exceeds 75 wpm is 0.1587. The probability that both typists have typing speeds that exceed 75 wpm is (0.1587)(0.1587) = 0.0252. (e) typing speeds of 47.376 wpm or less.

**6.45** 0.3173

**6.46** The proportion of corks produced by this machine that are defective is approximately 0. The second machine produces fewer defective corks.

#### **Additional Exercises**

**6.53** It isn't reasonable to think that traffic flow is approximately normal because traffic flow can't be negative, and 0 is less than 2 standard deviations below the mean.

**6.55** Since these values are times, they must all be positive. In the normal distribution with mean 9.9 and standard deviation 6.2, approximately 5.5% of processing times would be less than or equal to zero.

6.57 (a) 0.1359
(b) 0.228
(c) 0.5955

**6.59** (a) 0.9332 (b) 72.82 minutes

**6.61**  $P(x < 4.9) = 0.0228; P(x \ge 5.2) \approx 0$ 

**6.63** To get an A, your score must be greater than 85.252, so you received an A.

6.65 The bulbs should be replaced after 657.9 hours.

## **SECTION 6.6**

#### Exercise Set 1

**6.67** (a) The plot does not look linear. This supports the author's statement.



(**b**) r = 0.921; *critical* r (from Table 6.2) is 0.911; it is reasonable to think that the population distribution is normal.



(c) Because both normal probability plots are approximately linear, it seems reasonable that both risk behavior scores and PANAS scores are approximately normally distributed.

#### **Additional Exercises**

**6.71** (a) Yes, the normal probability plot appears linear. (**b**) r = 0.994; *critical* r for n = 6 lies between 0.832 (*critical r* for n = 5) and 0.880 (*critical r* for n = 10). The computed correlation coefficient is larger than the *critical r* for n = 6, so it is reasonable to think that the fuel efficiency distribution is approximately normal.

6.73 Based on the linearity of the normal probability plot and the fact that the correlation coefficient for the (normal score, disk diameter) pairs (r = 0.987) exceeds the *critical* r of 0.941 (from Table 6.2), it is reasonable to think that disk diameter is normally distributed.

## **SECTION 6.7**

#### **Exercise Set 1**

6.74 (b)	( <b>a</b> ) 0, 1, 1	2, 3, 4, a	and 5.			
x	0	1	2	3	4	5
<b>p(</b> x)	0.2373	0.3955	0.2637	0.0879	0.0146	0.0010

**6.75** (a) P(x = 2) = 0.0486; the probability that exactly two of the four randomly selected households have cable TV is 0.0486.

**(b)** 
$$P(x = 4) = 0.6561$$
  
**(c)**  $P(x \le 3) = 0.3439$ 

**6.76** (a) P(x = 2) = 0.2637**(b)**  $P(x \le 1) = 0.6328$ (c)  $P(2 \le x) = 0.3672$ (d)  $P(x \neq 2) = 0.7363$ 6.77 (a) 0.7359 (b) 0.3918 (c) 0.0691

6.78 (a) There is not a fixed number of trials, which is required for the binomial distribution. This setting is geometric. **(b)** (i) p(4) = 0.0623

(ii)  $P(x \le 4) = 0.2836$ (iii) P(x > 4) = 0.7164(iv)  $P(x \ge 4) = 0.7787$ 

(c) The differences between the four probabilities are shown in bold font.

(i) The probability that it takes exactly four songs until the first song by the particular artist is played is 0.0623. (ii) The probability that it takes at most four songs until the first song by the particular artist is played is 0.2836. (iii) The probability that it takes more than four songs until the first song by the particular artist is played is 0.7164.

(iv) The probability that it takes at least four songs until the first song by the particular artist is played is 0.7787.

6.79 (a) geometric

**(b)** P(x = 3) = 0.1084(c) P(x < 4) = 0.3859

(d)  $P(x > 3) = 1 - P(x \le 3) = 0.6141$ 

#### Additional Exercises

**6.87** The binomial probability distribution with n = 5 and p = 0.5:

x	0	1	2	3	4	5	
<b>p(x)</b>	0.03125	0.15625	0.31250	0.31250	0.15625	0.03125	

**6.89** This scenario is sampling without replacement. Since more than 5% of the population is being sampled (the percentage of the population being sampled is 2,000/10,000 = 20%), the binomial distribution will not give a good approximation of the probability distribution of the number of invalid signatures.

**6.91** (a)  $P(x \le 7) + P(x \ge 18) = 0.0216 + 0.0216 =$ 0.0432.

**(b)** 0.8542, 0.8542

(c) 0.1548; 0.1548; The probabilities are large compared to the probabilities in Part (b) because p (the probability of a head) in Part (c) is closer to 0.5 than in Part (b).

(d) Changing the rule for fair to  $7 \le x \le 18$  increases the likelihood that the coin will be judged fair and decreases the probability that the coin is judged to not be fair. Although the new rule is more likely to judge a fair coin fair, it is also more likely to judge a biased coin as fair.

## **SECTION 6.8**

#### **Exercise Set 1**

6.93 (a) 0.026
(b) 0.7580
(c) 0.7366
(d) 0.9109

**6.94** (a) 0.8849 (b) 0.8739 (c) 0.0110 (d) 0.5434

6.95 (a) 0.3248
(b) 0.2763
(c) 0.0012

#### **Additional Exercises**

**6.99** (a) Both np = (60)(0.7) = 42 and n(1 - p) = 60(1 - 0.7) = 18 are at least 10. (b) (i) 0.1121

(**ii**) 0.4440

(iii) 0.5560

(c) The probability in (i) represents the probability of exactly 42 correct; the probability in (ii) is the probability of less than 42 correct, the probability in (iii) represents the probability of 42 or fewer correct.

(d) The normal approximation is not appropriate here because n(1 - p) = 60(1 - 0.96) = 2.4, which is less than 10. (e) The small probability in Part (d) indicates that it is extremely unlikely for someone who is not faking the test to correctly answer 42 or fewer questions, compared with the probability that a person who is faking the test correctly answers 42 or fewer questions (0.556).

## ARE YOU READY TO MOVE ON? CHAPTER 6 REVIEW EXERCISES

**6.101** The depth, *x*, can take on any value between 0 and 100 ( $0 \le x \le 100$ ), inclusive. *x* is continuous.

**6.103** (a) The plane can accommodate 100 passengers, so if 100 or fewer passengers show up for the flight, everyone can be accommodated;  $P(x \le 100) = 0.82$ .

**(b)** P(x > 100) = 0.18.

(c) The first person on the standby list:  $P(x \le 99) = 0.65$ . The third person on the standby list:  $P(x \le 97) = 0.27$ .

**6.105** (a) P(x < 10) = 0.5; P(x > 15) = 0.25(b) P(7 < x < 12) = 0.25(c) c = 18 minutes

**6.107** (a) Supplier 1 (b) Supplier 2

(c) I would recommend Supplier 1 because the bulbs will last, on average, longer than those from Supplier 2. Additionally, the bulbs from Supplier 1 have less variability in their lifetimes, so there is more consistency in the bulb lifetimes when compared with Supplier 2.
(d) Approximately 1000 hours.
(e) Approximately 100 hours.

**6.109** (a) 0.284 (b) 0.091 (c) 0.435 (d) 29.93 mm

**6.111** (a) 0.159 (b) 51.4 minutes (c) 41.6 minutes

**6.113** Yes, because pattern in the normal probability plot is curved rather than linear.

# Chapter 7 An Overview of Statistical Inference— Learning from Data

## **SECTION 7.1**

#### **Exercise Set 1**

7.1 The inferences made involve estimation.

7.2 (a) American teenagers between the ages of 12 and 17.
(b) The percentage of teens who own a cell phone, the percentage of teens who use a cell phone to send and receive text messages, and the percentage of teens ages 16–17 who have used a cell phone to text while driving.
(c) No, the actual percentage of teens owning a cell phone is probably not exactly 75%. The value of a sample statistic won't necessarily be equal to the population value.
(d) I would expect the estimate of the percentage of teens who own a cell phone to be more accurate. The sample contained teens aged 12–17; however, only those teens aged 16 and 17 were asked about texting while driving. The number of teens aged 16 and 17 is a subset of the overall sample, and so the estimate is based on a smaller sample.

**7.3** The inference made is one that involves hypothesis testing.

**7.4** (a) People driving along stretches of highway that have digital billboards.

(b) The time required to respond to road signs is greater when digital billboards are present.

(c) Answers will vary. One possible answer is: In addition to the information provided, I would like to know if the subjects were randomly selected or not, and whether or not the subjects were randomly assigned to any experimental groups. I would also like to know if there was a control group in which the subjects did not see digital billboards or a group in which subjects did not see any change in the display on the digital billboard.

(d) lower

**7.5** (a) Answers will vary. One possible answer is "What proportion of students approve of a recent decision made by the university to increase athletic fees in order to upgrade facilities?"

(**b**) Answers will vary. One possible answer is "Do more than half of students approve of the recent decision made by the university to increase athletic fees in order to upgrade facilities?"

#### **Additional Exercises**

**7.11** The inference made is one that involves hypothesis testing.

## **SECTION 7.2**

#### **Exercise Set 1**

**7.13** Estimate a population mean. The data are numerical, rather than categorical.

**7.14** The type of data determines what graphical, numerical, and inferential methods are appropriate.

7.15 The other two questions are (1) Q: Question Type (estimation problem or hypothesis testing problem), and (2)S: Study Type (sample data or experiment data).

- 7.16 Q: Estimation
  - S: Sample data
  - T: One variable, categorical data
  - N: One sample
- 7.17 Q: Hypothesis testing
  - S: Sample data
  - T: One variable, numerical data
  - N: Two samples
- 7.18 Q: Hypothesis testing
  - S: Sample data
  - T: Two variables, categorical and numerical
  - N: One sample

#### **Additional Exercises**

- 7.25 Q: Estimation
  - S: Sample data
  - T: One variable, categorical
  - N: One sample

#### 7.27 Q: Hypothesis testing

- S: Experiment
- T: One variable, categorical
- N: Two treatments
- 7.29 Q: Hypothesis testing
  - S: Experiment data
  - T: One variable, numerical
  - N: Two treatments

## ARE YOU READY TO MOVE ON? CHAPTER 7 REVIEW EXERCISES

**7.31** Estimation, because the researchers were interested in the mean rating given to the wine under each of the five musical conditions. The researchers did not indicate that they had a claim that they wanted to test.

**7.33** (a) Answers will vary. One possible answer is: "What proportion of people who purchased season tickets for home games of the New York Yankees purchased alcoholic beverages during the game?"

(b) Answers will vary. One possible answer is: "Do fewer than 50% of people who purchased season tickets for home games of the New York Yankees drive to the game?"

7.35 (a) The proportion of people who take a garlic supplement who get a cold is lower than the proportion of those who do not take a garlic supplement and who get a cold.
(b) Yes, it is possible that the conclusion is incorrect. The observed difference in treatment effects may be due to chance variability in the response variable and the random assignment to treatments, and not due to the treatment.
(c) greater

**7.37** The type of data collected determines not only the type of numerical and graphical methods that can be applied but also the particular inferential method or methods that can be used. Numerical data require different inferential methods than categorical data, and univariate data require different inferential methods than bivariate data.

- **7.39** Q: Hypothesis testing
  - S: Experiment data
    - T: One variable, categorical
    - N: Two treatments

**7.41 (a)** Estimate, Method, Check, Calculate, Communicate results

(b) The difference is that the "estimate" step is replaced with "hypotheses," where you determine the hypotheses you wish to test, rather than defining the population characteristic or treatment effect you wish to estimate.

# Chapter 8 Sampling Variability and Sampling Distributions

## **SECTION 8.1**

#### Exercise Set 1

**8.1** No, because the value of  $\hat{p}$  will vary from sample to sample.

**8.2** (a) The histogram on the left; it has values that tend to deviate more from the center, and it is more spread out than the histogram on the right.

(b) n = 75, because the histogram on the right seems to have less sample-to-sample variability, is centered at 0.55, and has tall bars close to 0.55.

8.3 (a) a population proportion.
(b) p = 0.22.

**8.4** (a) a sample proportion. (b)*p*<sup>•</sup> = 0.38.

**8.5** *p* is the proportion of successes in the entire population, and  $\hat{p}$  is the proportion of successes in the sample.

#### **Additional Exercises**

**8.11** Sample statistics are computed from a sample. Since one sample from the population is likely to differ from other possible samples taken from the same population, the sample statistics computed from different samples are likely to be different.

**8.13** (a) a sample proportion. (b)  $\hat{p} = 0.45$ .

## **SECTION 8.2**

#### **Exercise Set 1**

8.15 (a)  $\mu_{\hat{\rho}} = 0.65; \sigma_{\hat{\rho}} = 0.151$ (b)  $\mu_{\hat{\rho}} = 0.65; \sigma_{\hat{\rho}} = 0.107$ (c)  $\mu_{\hat{\rho}} = 0.65; \sigma_{\hat{\rho}} = 0.087$ (d)  $\mu_{\hat{\rho}} = 0.65; \sigma_{\hat{\rho}} = 0.067$ (e)  $\mu_{\hat{\rho}} = 0.65; \sigma_{\hat{\rho}} = 0.048$ (f)  $\mu_{\hat{\rho}} = 0.65; \sigma_{\hat{\rho}} = 0.034$ 

**8.16** For *p* = 0.65: 30, 50, 100, and 200. For *p* = 0.2: 50 100, and 200.

**8.17** (a)  $\mu_{\hat{p}} = 0.07$ ;  $\sigma_{\hat{p}} = 0.0255$ (b) No, because np = (100)(0.07) = 7 is less than 10. (c) The mean does not change. The standard deviation will decrease to 0.0180. The mean does not change because the sampling distribution is always centered at the population value regardless of the sample size. The standard deviation of the sampling distribution will decrease as the sample size increases because the sample size (*n*) is in the denominator of the formula for standard deviation.

(d) Yes, because np = (200)(0.07) = 14 and n(1 - p) = (200)(1 - 0.07) = 186 are both greater than 10.

**8.18** (a)  $\mu_{\hat{p}} = 0.15$ ;  $\sigma_{\hat{p}} = 0.0357$ (b) Yes, because np = (100)(0.15) = 15 and n(1-p) = (100)(1-0.15) = 85 are both greater than 10.

(c) The mean does not change. The standard deviation will decrease to 0.0252. The mean does not change because the sampling distribution is always centered at the population value regardless of the sample size. The standard deviation of the sampling distribution will decrease as the sample size

increases because the sample size (n) is in the denominator of the formula for standard deviation

(d) Yes, because np = (200)(0.15) = 30 and n(1 - p) = (200)(1 - 0.15) = 170 are both greater than 10.

#### **Additional Exercises**

**8.23** n = 100 and p = 0.5.

**8.25** For *p* = 0.2: 50 and 100. For *p* = 0.8: 50 and 100. For *p* = 0.6: 25, 50, and 100.

**8.27** For samples of size n = 40: p = 0.45 and p = 0.70. For samples of size n = 75: p = 0.20, p = 0.45, and p = 0.70.

## **SECTION 8.3**

#### **Exercise Set 1**

8.29 (a)  $\sigma_{\hat{p}} = \sqrt{\frac{0.48(1-0.48)}{500}} = 0.0223$ (b) More sample-to-sample variability in the sample proportions because  $\sigma_{\hat{p}}$  is now greater than  $\sigma_{\hat{p}}$  when n = 500. (c) The sample size is smaller because, in order for  $\sigma_{\hat{p}}$  to be greater, the denominator of  $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$  must be smaller than when n = 500.

#### 8.30

What You Know	How You Know It
The sampling distribution of $\hat{p}$ is centered at the actual (but unknown) value of the population proportion.	Rule 1 states that $\mu_{\hat{p}} = p$ . This is true for random samples, and the description of the study says that the sample was selected at random.
An estimate of the standard deviation of $\hat{p}$ , which describes how much the $\hat{p}$ values spread out around the population proportion $p$ , is 0.0153.	Rule 2 states that $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ . In this exercise, $n = 1000$ . The value of $p$ is not known. However, $\hat{p}$ provides an estimate of $p$ that can be used to estimate the standard deviation of the sampling distribution. Specifically, $\hat{p} =$ 0.37, so $\sigma_{\hat{p}} \approx \sqrt{\frac{0.37(1-0.37)}{1000}}$ = 0.0153. This standard deviation provides information about how tightly the $\hat{p}$ values from different random samples will cluster around the value of $p$ .
The sampling distribution of $\hat{p}$ is approximately normal.	Rule 3 states that the sampling distribution of $\hat{p}$ is approximately normal if <i>n</i> is large and <i>p</i> is not too close to 0 or 1. Here the sample size is 1000. The sample includes 370 successes and 630 failures, which are both much greater than 10. So we conclude that the sampling distribution of $\hat{p}$ is approximately normal.

#### **8.31** (a) $\hat{p} = 0.346$

(b) No. From Rule 1, we know that the sampling distribution is centered at p. From Rule 2, the standard deviation of the sampling distribution of  $\hat{p}$  is  $\sigma_{\hat{p}} = 0.066$ . By Rule 3, the sampling distribution of  $\hat{p}$  is approximately normal because the sample includes 18 successes and 34 failures, which are both greater than 10. Since the sampling distribution of  $\hat{p}$  is approximately normal and is centered at the actual population proportion p, about 95% of all possible random samples of size n = 52 will produce a sample proportion that is within 2(0.066) =0.132 of the actual value of the population proportion. This margin of error of 0.132 is over twice the value of 0.05, so it is not reasonable to think that this estimate is within 0.05 of the actual value of the population proportion.

**8.32** From Rule 1, we know that the sampling distribution of  $\hat{p}$  is centered at the population value p = 0.61. Rule 2 tells us that  $\sigma_{\hat{r}} = 0.0126$ . Finally, the sampling distribution of  $\hat{p}$  is approximately normal because np = 1,500(0.61) = $915 \ge 10$  and  $n(1 - p) = 1,500(1 - 0.61) = 585 \ge 10$ . The probability of observing a sample proportion of 0.55 or smaller just by chance (due to sampling variability) is

 $P(\hat{p} < 0.55) = P\left(z < \frac{0.55 - 0.61}{0.0126}\right) = P(z < -4.76) \approx 0.$ 

Therefore, it seems likely that the proportion of California high school graduates who attend college the year after graduation is different from the national figure.

#### **Additional Exercises**

8.37 No, it is not likely that this estimate is within 0.05 of the actual value of the population proportion. From Rule 1, we know that the sampling distribution is centered at p. From Rule 2, the standard deviation of the sampling distribution of  $\hat{p}$  is  $\sigma_{\hat{p}} = 0.049$ . By Rule 3, the sampling distribution of  $\hat{p}$  is approximately normal because  $np = 100(0.38) = 38 \ge 10$  and  $n(1 - p) = 100(1 - 0.38) = 62 \ge 10$ . Since the sampling distribution of  $\hat{p}$  is approximately normal and is centered at the actual population proportion p, we now know that about 95% of all possible random samples of size n = 100 will produce a sample proportion that is within 2(0.049) = 0.098 of the actual value of the population proportion. This margin of error is nearly twice the value of 0.05, so it is unlikely that this estimate is within 0.05 of the population proportion.

**8.39** By Rule 1, we know that the sampling distribution of  $\hat{p}$  is centered at the unknown value of p, the true proportion of social network users who believe that it is not OK to "friend" your boss. Rule 2 says that the standard deviation of the sampling distribution of  $\hat{p}$  is  $\sigma_{\hat{p}} =$  $\sqrt{\frac{p(1-p)}{n}}$ . The value of p is not known. However,  $\hat{p}$  provides an estimate of p that can be used to estimate the

standard deviation of the sampling distribution. In this

case,  $\hat{p} = 0.56$ , so the estimate of the standard deviation of the sampling distribution is  $\sigma_{\hat{\alpha}} = 0.014$ . Finally, Rule 3 says that the sampling distribution of  $\hat{p}$  is approximately normal if n is large and p is not too close to 0 or 1. Here the sample size is 1,200. The sample includes 672 successes (56% of 1,200) and 528 failures (44% of 1,200), which are both much greater than 10. Therefore, we can conclude that the sampling distribution of  $\hat{p}$  is approximately normal. The probability of observing a sample proportion at least as large as what we actually observed (  $\hat{p} = 0.56$ ) if the

true value were 0.5 is  $P(\hat{p} \ge 0.56) = P\left(z \ge \frac{0.56 - 0.5}{0.014}\right) =$  $P(z \ge 4.29) \approx 0.000$ . It is highly unlikely that a sample proportion as large as 0.56 would be observed if the true value were 0.5 (or less). Therefore, it seems plausible that the proportion of social network users who believe that it is not OK to "friend" your boss is greater than 0.5.

## ARE YOU READY TO MOVE ON? CHAPTER 8 REVIEW EXERCISES

**8.41** No, because the value of  $\hat{p}$  varies from sample to sample.

**8.43** Different samples will likely yield different values of  $\hat{p}$ , which is the concept of sampling variability (or sample-to-sample variability). However, there is only one true value for the population proportion.

8.45 (a) Population proportion. **(b)** p = 0.21.

**8.47** For both p = 0.70 and p = 0.30: n = 50, n = 100, and n = 200.

**8.49** (a)  $\mu_{\hat{p}} = p = 0.25; \sigma_{\hat{p}} = 0.031$ (b) Yes, because  $np = 200(0.25) = 50 \ge 10$  and  $n(1-p) = 200(1-0.25) = 150 \ge 10.$ (c) The change in sample size does not affect the mean but does affect the standard deviation of the sampling distribution of  $\hat{p}$ . The new standard deviation is  $\sigma_{\hat{p}} = 0.061$ . (d) Yes, because  $np = 50(0.25) = 12.5 \ge 10$  and  $n(1-p) = 50(1-0.25) = 37.5 \ge 10.$ 

**8.53** By Rule 1, we know that the sampling distribution of  $\hat{p}$  is centered at the unknown value of p. Rule 2 says that the standard deviation of the sampling distribution of  $\hat{p}$  is  $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ . The value of p is not known. However,  $\hat{p}$ provides an estimate of p that can be used to estimate the standard deviation of the sampling distribution. In this case,  $\hat{p} = 0.637$ , and the estimate of the standard deviation of the sampling distribution is  $\sigma_{\hat{n}} = 0.011$ . Finally, Rule 3 says that the sampling distribution of  $\hat{p}$  is approximately normal if *n* is large and *p* is not too close to 0 or 1. Here the sample size is 2,013. The sample includes 1,283 successes and 730 failures, which are both much greater than 10.

Therefore, we can conclude that the sampling distribution of  $\hat{p}$  is approximately normal. The probability of observing a sample proportion at least as large as what we observed if the true value were 0.5 is equal to  $P(\hat{p} \ge 0.637) =$ 

 $P\left(z \ge \frac{0.637 - 0.5}{0.011}\right) = P(z \ge 12.45) \approx 0.000$ . It is highly

unlikely that a sample proportion as large as 0.637 would be observed if the true value were 0.5 (or less). Therefore, it seems plausible that the proportion of adult Americans who believe rudeness is a worsening problem is greater than 0.5.

# Chapter 9 Estimating a Population Proportion

NOTE: Answers may vary slightly if you are using statistical software, a graphing calculator or depending on how the values of sample statistics are rounded when performing hand calculations. Don't worry if you don't match these numerical answers exactly (but your answers should be relatively close).

## **SECTION 9.1**

#### **Exercise Set 1**

**9.1** An unbiased statistic with a smaller standard error is preferred because it is likely to result in an estimate that is closer to the actual value of the population characteristic than an unbiased statistic that has a larger standard error.

9.2 Statistics II and III

**9.3** Statistic I, because it has a smaller bias than Statistics II and III.

**9.4** *n* = 200.

**9.5** (a) The formula for the standard error of  $\hat{p}$  is  $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ . The quantity p(1-p) reaches a maximum value when p = 0.5.

(b) The standard error of  $\hat{p}$  is the same when p = 0.2 as when p = 0.8 because when p = 0.2, (1 - p) = (1 - 0.2) =0.8. Similarly, when p = 0.8, (1 - p) = (1 - 0.8) = 0.2. So, the quantity p(1 - p) is the same in both cases.

**9.6** n = 400 and p = 0.8

#### **Additional Exercises**

**9.13** A biased statistic might be chosen over an unbiased statistic if the bias is not too large, and the standard error of the biased statistic is much smaller than the standard error of the unbiased statistic. In this case, the observed value of the biased statistic might be closer to the actual value than the value of an unbiased statistic.

**9.15** *n* = 200.

## **SECTION 9.2**

#### **Exercise Set 1**

**9.17** Statement 1: Incorrect, because the value 0.0157 is the standard error of  $\hat{p}$ , and therefore approximately 32% of all possible values of  $\hat{p}$  would differ from the value of the actual population proportion by more than 0.0157 (using properties of the normal distribution).

#### Statement 2: Correct

Statement 3: Incorrect, because the phrase "will never differ from the value of the actual population proportion" is wrong. The value 0.0307 is the margin of error and indicates that in about 95% of all possible random samples, the estimation error will be less than the margin of error. In about 5% of the random samples, the estimation error will be greater than the margin of error.

**9.18** (a) 0.049  
(b) 
$$n = 100$$
  
(c)  $1/\sqrt{2} \approx 0.707$ 

**9.19** (a) 
$$\hat{p} = 0.260$$

**(b)** 
$$\sigma_{a} = 0.019$$

(c) margin of error = 0.037. The estimate of the proportion of all businesses that have fired workers for misuse of the Internet is unlikely to differ from the actual population proportion by more than 0.037.

9.20 (a) yes
(b) no
(c) no
(d) no

**9.21** (a)  $\hat{p} = 0.404$ 

(**b**) The sample was selected in such a way that makes it representative of the population of U.S. college students. Additionally, there are 2,998 successes and 4,423 failures in the sample, which are both at least 10.

(c) margin of error = 0.011

(d) It is unlikely that the estimated proportion of U.S. college students who use the Internet more than 3 hours per day ( $\hat{p} = 0.404$ ) will differ from the actual population proportion by more than 0.011 (or 1.1%).

#### **9.22** (a) $\hat{p} = 0.277$

(**b**) The sample is a random sample from the population of American children. Additionally, there are 1,720 successes and 4,492 failures in the sample, which are both greater than 10.

(c) margin of error = 0.011

(d) It is unlikely that the estimated proportion of American children who indicated that they eat fast food on a typical day ( $\hat{p} = 0.277$ ) will differ from the actual population proportion by more than 0.011.

#### **Additional Exercises**

**9.29** marign of error = 0.024. It is unlikely that the estimated proportion of adults who believe that the shows are mostly or totally made up ( $\hat{p} = 0.82$ ) will differ from the actual population proportion by more than 0.024.

**9.31** In this case, margin of error = 0.027, which rounds to 3%.

**9.33** In this case, margin of error = 0.032, which rounds to 3%. The margin of error tells you that it is unlikely that the estimate will differ from the actual population proportion by more than 0.03.

## **SECTION 9.3**

#### **Exercise Set 1**

**9.34** (a) The confidence intervals are centered at  $\hat{p}$ . In this case, the intervals are not centered in the same place because two different samples were taken, each yielding a different value of  $\hat{p}$ .

(**b**) Interval 2 conveys more precise information about the value of the population proportion because Interval 2 is narrower than Interval 1.

(c) A smaller sample size produces a larger margin of error. In this case, Interval 1 (being wider than Interval 2) was based on the smaller sample size.

(d) Interval 1 would have the higher confidence level, because the *z* critical value for higher confidence is larger, resulting in a wider confidence interval.

**9.35** (a) 95% (b) n = 100

**9.36** The method used to construct this interval estimate is successful in capturing the actual value of the population proportion about 95% of the time.

9.37 (a) yes
(b) no
(c) no
(d) no
9.38 (a) 1.645
(b) 2.58
(c) 1.28

**9.39** Question type (Q): Estimation; Study type (S): Sample data; Type of data (T): One categorical variable; Number of samples or treatments (N): One sample.

**9.40** Estimate (E): The proportion of hiring managers and human resources professionals who use social networking sites to research job applicants, *p*, will be estimated.

Method (M): Because the answers to the four key questions are estimation, sample data, one categorical variable, and one sample (see Exercise 9.39), consider a 95% confidence interval for the proportion of hiring managers and human

resource professionals who use social networking sites to research job applicants.

Check (C): The sample is representative of hiring managers and human resource professionals. In addition, the sample includes 1,200 successes and 1,467 failures, which are both greater than 10. The two required conditions are satisfied.

Calculations (C): (0.4311,0.4689)

Communicate Results (C):

Interpret confidence interval: You can be 95% confident that the actual proportion of hiring managers and human resources professionals is somewhere between 0.4311 and 0.4689.

Interpret confidence level: The method used to construct this interval estimate is successful in capturing the actual value of the population proportion about 95% of the time.

**9.41** (a) (0.3449, 0.3951). You can be 90% confident that the actual proportion of college freshmen who carry a credit card balance is somewhere between 0.3449 and 0.3951.

(**b**) (0.454, 0.506). You can be 90% confident that the actual proportion of college seniors who carry a credit card balance is somewhere between 0.454 and 0.506.

(c) The two confidence intervals from (a) and (b) do not have the same width because the standard errors, and hence the margins of error, are based on two different values for  $\hat{p}$ .

**9.42** (a) (0.2039, 0.2561). You can be 95% confident that the actual proportion of U.S. adults for whom math was their most favorite subject is somewhere between 0.2039 and 0.2561.

(**b**) (0.3401, 0.3999). You can be 95% confident that the actual proportion of U.S. adults for whom math was their least favorite subject is somewhere between 0.3401 and 0.3999.

**9.43** (a) (0.1206, 0.2254). You can be 95% confident that the actual proportion of all adult Americans who planned to purchase a Valentine's Day gift for their pet is somewhere between 0.1206 and 0.2254.

(b) The 95% confidence interval computed for the actual sample size would have been narrower than the confidence interval computed in Part (a) because the standard error, and hence the margin of error, would have been smaller.

#### **Additional Exercises**

9.55 (a) (0.494, 0.546). You can be 90% confident that the actual proportion of adult Americans who would say that lying is never justified is somewhere between 0.494 and 0.546.
(b) (0.6252, 0.6748). You can be 90% confident that the actual proportion of adult Americans who would say that it is often or sometimes OK to lie to avoid hurting someone's feelings is somewhere between 0.6252 and 0.6748.

(c) The confidence interval in Part (a) indicates that it is plausible that at least 50% of adult Americans would say that lying is never justified, and the confidence interval in part (b) indicates that it is also plausible that well over 50% of adult Americans would say that it is often or sometimes OK to lie to avoid hurting someone's feelings. These are contradictory responses.

**9.57** (0.597, 0.643). You can be 95% confident that the actual proportion of all Australian children who would say that they watch TV before school is between 0.597 and 0.643. For the method used to construct the interval to be valid, the sample must have either been randomly selected from the population of interest or the sample must have been selected in such a way that it should result in a sample that is representative of the population.

**9.59** (0.1446, 0.1894). You can be 90% confident that the actual proportion of all full-time workers so angered in the past year that they wanted to hit a coworker is between 0.1446 and 0.1894.

## **SECTION 9.4**

#### Exercise Set 1

**9.61** Assuming a 95% confidence level, and using a conservative estimate of p = 0.5, n = 2,401.

**9.62** Assuming a 95% confidence level, and using the preliminary estimate of p = 0.27, n = 303. Using the conservative estimate of p = 0.5, n = 385.

**9.63** Assuming a 95% confidence level, and using a conservative estimate of p = 0.5, n = 97.

#### **Additional Exercises**

**9.67** Assuming a 95% confidence level, and using the preliminary estimate of p = 0.32, n = 335. Using the conservative estimate of p = 0.5, n = 385.

**9.69** (a) (0.1186, 0.2854). You can be 95% confident that the actual proportion of all such patients under 50 years old who experience a failure within the first 2 years is between 0.1186 and 0.2854.

(**b**) (0.0107, 0.0611). You can be 99% confident that the actual proportion of all such patients age 50 or older who experience a failure within the first 2 years is between 0.0107 and 0.0611.

(c) For a 95% confidence level, and using a preliminary estimate of p = 0.202, n = 689.

## ARE YOU READY TO MOVE ON? CHAPTER 9 REVIEW EXERCISES

#### 9.71 (a) Statistic II is unbiased.

(**b**) Statistic I will tend to be closer to the population value than Statistic II.

**9.73**  $\hat{p}$  from a random sample of size 400 tends to be closer to the actual value when p = 0.7.

#### 9.75 Statement 1: Correct

Statement 2: Incorrect. About 95% of all possible sample proportions computed from random samples of size 500 will be within 0.0429 of the actual population proportion (this is the margin of error for a 95% confidence level). About 5% of the sample proportions would differ from the actual population proportion by more than 0.0429.

Statement 3: Incorrect. About 68% of all possible sample proportions computed from random samples of size 500 will be within 0.0219 of the actual population proportion (this is the standard deviation of the sampling distribution). About 32% of the sample proportions would differ from the actual population proportion by more than 0.0219.

#### **9.77** (a) $\hat{p} = 0.18$

(b) The two conditions are: (1) Large sample size: The number of successes and failures in the sample are at least 10. In this case, there are 722 successes, and 3,291 failures, which are both much greater than 10. (2) Random selection of the sample, or the sample is representative of the population: We are told that the sample is representative of the population. (c) margin of error = 0.0119

(d) It is unlikely that the estimate  $\hat{p} = 0.18$  differs from the value of the actual proportion of adults Americans who say they have seen a ghost by more than 0.0119.

(e) (0.17, 0.19). You can be 90% confident that the actual proportion of all adult Americans who have seen a ghost is somewhere between 0.17 and 0.19.

(f) A 99% confidence interval would be wider because the *z* critical value for 99% confidence (z = 2.58) is larger than the *z* critical value for 90% confidence (z = 1.645).

**9.79** (a) As the confidence level increases, the width of the confidence interval also increases.

(**b**) As the sample size increases, the width of the confidence interval decreases.

(c) When  $\hat{p} = 0.5$ , the margin of error (for a fixed sample size) is maximum and decreases symmetrically as  $\hat{p}$  decreases toward 0 or increases toward 1. A larger margin of error results in a wider confidence interval.

**9.81** Question type (Q): Estimation; Study type (S): Sample data; Type of data (T): One categorical variable; Number of samples or treatments (N): One sample.

**9.83** (a) (0.6426, 0.6774). You can be 90% confident that the actual proportion of all Americans ages 8 to 18 who own a cell phone is somewhere between 0.6426 and 0.6774. (b) (0.7443, 0.7757). You can be 90% confident that the actual proportion of all Americans ages 8 to 18 who own an MP3 music player is somewhere between 0.7443 and 0.7757. (c) The interval in Part (b) is narrower because  $\hat{p}$  of 0.76 is farther from 0.5 than  $\hat{p}$  of 0.66. For a fixed sample size and confidence level, the confidence interval is widest

when  $\hat{p} = 0.5$  and decreases symmetrically as  $\hat{p}$  moves closer to 0 or 1.

# Chapter 10 Asking and Answering Questions About a Population Proportion

NOTE: Answers may vary slightly if you are using statistical software, a graphing calculator or depending on how the values of sample statistics are rounded when performing hand calculations. Don't worry if you don't match these numerical answers exactly (but your answers should be relatively close).

## **SECTION 10.1**

#### **Exercise Set 1**

**10.1**  $\hat{p}$  is a sample statistic. Hypotheses are about population characteristics.

**10.2** 
$$H_0: p = \frac{1}{3}$$
 and  $H_a: p > \frac{1}{3}$   
**10.3**  $H_0: p = 0.67$  and  $H_a: p \neq 0.67$ 

**10.4** (a) There is convincing evidence that the proportion of American adults who favor drafting women is less than 0.5.

(b) Yes

(c) Yes

**10.5** (a) The conclusion is consistent with testing  $H_0$ : concealed weapons laws do not reduce crime versus  $H_a$ : concealed weapons laws reduce crime.

(**b**) The null hypothesis was not rejected, since no evidence was found that the laws were reducing crime.

**10.6** The sample data provide convincing evidence against the null hypothesis. If the null hypothesis were true, the sample data would be very unlikely.

#### **Additional Exercises**

10.13 (a) legitimate

(**b**) not legitimate

(c) not legitimate

(**d**) legitimate

(e) not legitimate

**10.15**  $H_0$ : p = 0.6 versus  $H_a$ : p > 0.6. In order to make the change, the university requires evidence that *more than* 60% of the faculty are in favor of the change.

## **SECTION 10.2**

#### Exercise Set 1

**10.17** Not rejecting the null hypothesis when it is not true.

**10.18** A small significance level, because  $\alpha$  is the probability of a Type I error.

**10.19** (a) Before filing charges of false advertising against the company, the consumer advocacy group would require convincing evidence that more than 10% of the flares are defective.

(b) A Type I error is thinking that more than 10% of the flares are defective when in fact 10% (or fewer) of the flares are defective. This would result in the expensive and time-consuming process of filing charges of false advertising against the company when the company advertising is not false. A Type II error is not thinking that more than 10% of the flares are defective when in fact more than 10% of the flares are defective. This would result in the consumer advocacy group not filing charges when the company advertising was false.

**10.20** (a) A Type I error would be thinking that less than 90% of the TV sets need no repair when in fact (at least) 90% need no repair. The consumer agency might take action against the manufacturer when the manufacturer is not at fault. A Type II error would be *not* thinking that less than 90% of the TV sets need no repair when in fact less than 90% need no repair. The consumer agency would not take action against the manufacturer when the manufacturer is making untrue claims about the reliability of the TV sets. (b) Taking action against the manufacturer when the manufacturer is not at fault could involve large and unnecessary legal costs to the consumer agency.  $\alpha = 0.01$  should be recommended.

10.21 (a) Type I error, 0.091

(**b**) Thinking that a woman does have cancer in the other breast when she really does not, 0.097.

#### **Additional Exercises**

**10.27** A Type I error is rejecting a true null hypothesis and a Type II error is not rejecting a false null hypothesis.

10.29 Answers will vary.

**10.31** The "38%" value given in the article is the proportion of *all* felons; in other words, it is a *population* proportion. Therefore, you know that the population proportion is less than 0.4, and there is no need for a hypothesis test.

## **SECTION 10.3**

#### **Exercise Set 1**

**10.32** (a) The sampling distribution of  $\hat{p}$  will be approximately normal with a mean of 0.25 and a standard deviation of 0.0125.

(**b**) A sample proportion of  $\hat{p} = 0.24$  would not be surprising if p = 0.25, because it is less than one standard deviation below 0.25. This is not unusual for a normal distribution.

(c) A sample proportion of  $\hat{p} = 0.20$  would be surprising if p = 0.25, because it is 4 standard deviations below 0.25. This would be unusual for a normal distribution. (d) If p = 0.25,  $P(\hat{p} \le 0.22) = P(z \le -2.4) = 0.0082$ . Because this probability is small, there is convincing evidence that the goal is not being met.

**10.33** (a) The sampling distribution of  $\hat{p}$  will be approximately normal with a mean of 0.75 and a standard deviation of 0.0256.

(**b**) A sample proportion of  $\hat{p} = 0.83$  would be surprising if  $\hat{p} = 0.75$ , because it is more than 3 standard deviations above 0.75. This would be unusual for a normal distribution. (**c**) A sample proportion of  $\hat{p} = 0.79$  would not be surprising if p = 0.75, because it is about 1.5 standard

deviations above 0.75. This is not unusual for a normal distribution. (d) If  $r_{1} = 0.75$ ,  $P(2 \ge 0.80) = P(2 \ge 1.05) = 0.0256$ .

(d) If p = 0.75,  $P(\hat{p} \ge 0.80) = P(z \ge 1.95) = 0.0256$ . Because this probability is small, there is convincing evidence that the null hypothesis is not true.

#### **Additional Exercises**

**10.37** (a) The sampling distribution of  $\hat{p}$  will be approximately normal with a mean of 0.5 and a standard deviation of 0.1118.

(**b**) Answers will vary. One reasonable answer would be sample proportions greater than 0.7236 (these are sample proportions that are more than two standard deviations above 0.5).

## **SECTION 10.4**

#### Exercise Set 1

**10.38** (a) A *P*-value of 0.0003 means that it is very unlikely (probability = 0.0003), assuming that  $H_0$  is true, that you would get a sample result at least as inconsistent with  $H_0$  as the one obtained in the study.  $H_0$  would be rejected. (b) A *P*-value of 0.350 means that it is not particularly unlikely (probability = 0.350), assuming that  $H_0$  is true, that you would get a sample result at least as inconsistent with  $H_0$  as the one obtained in the study. There is no reason to reject  $H_0$ .

**10.39** (a)  $H_0$ :  $p = \frac{2}{3}$  and  $H_a$ :  $p > \frac{2}{3}$ (b) The null hypothesis would be rejected because the

*P*-value (0.013) is less than  $\alpha = 0.05$ .

**10.40** This step involves using the answers to the four key questions (QSTN) to identify an appropriate method.

#### **Additional Exercises**

**10.45** A *P*-value of 0.0002 means that it is very unlikely (probability = 0.0002), assuming that  $H_0$  is true, that you

would get a sample result at least as inconsistent with  $H_0$  as the one obtained in the study. This is strong evidence against the null hypothesis.

## **SECTION 10.5**

#### **Exercise Set 1**

**10.47** Estimation, sample data, one numerical variable, one sample. A hypothesis test for a population proportion would not be appropriate.

**10.48** Hypothesis testing, sample data, one categorical variable, one sample. A hypothesis test for a population proportion would be appropriate.

**10.49** Estimation, sample data, one categorical variable, one sample. A hypothesis test for a population proportion would not be appropriate.

**10.50** (a) Large-sample *z* test is not appropriate.

- (**b**) Large-sample *z* test is appropriate.
- (c) Large-sample *z* test is appropriate.
- (d) Large-sample *z* test is not appropriate.

**10.51** (a) 0.2912 (b) 0.1788 (c) 0.0233 (d) 0.0125 (e) 0.9192

**10.52** (a)  $H_0$ : p = 0.4,  $H_a$ : p < 0.4, z = -2.969, *P*-value = 0.001, reject  $H_0$ . There is convincing evidence that the proportion of all adult Americans who would answer the question correctly is less than 0.4.

(**b**)  $H_0: p = \frac{1}{3}, H_a: p > \frac{1}{3}, z = 2.996, P$ -value = 0.001, reject  $H_0$ . There is convincing evidence that more than one-third of adult Americans would select a wrong answer.

**10.53**  $H_0$ : p = 0.25,  $H_a$ : p > 0.25, z = 0.45, *P*-value = 0.328, fail to reject  $H_0$ . There is no convincing evidence that more than 25% of Americans ages 16 to 17 have sent a text message while driving.

**10.54**  $H_0$ : p = 0.5,  $H_a$ : p > 0.5, z = 14.546, *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that a majority of adult Americans prefer to watch movies at home.

**10.55** (a) z = -1.897, *P*-value = 0.0289, reject  $H_0$ (b) z = -0.6, *P*-value = 0.274, fail to reject  $H_0$ (c) Both results *suggest* that fewer than half of adult Americans believe that movie quality is getting worse. However, getting 470 out of 1,000 people responding this way (as opposed to 47 out of 100) provides much *stronger* evidence.

**10.56**  $H_0: p = 0.5, H_a: p < 0.5, z = -2.536$ , *P*-value = 0.006, reject  $H_0$ . There is convincing evidence that the proportion of all adult Americans who want car Web access is less than 0.5. The marketing manager is not correct in his claim.

#### **Additional Exercises**

**10.67** (a) z = 2.530, *P*-value = 0.0057, reject  $H_0$ (b) No. The survey only included women ages 22 to 35.

**10.69**  $H_0: p = 0.75, H_a: p > 0.75, z = 4.13, P$ -value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that more than three-quarters of American adults believe that lack of respect and courtesy is a serious problem.

**10.71** z = 1.069, *P*-value = 0.143, fail to reject  $H_0$ 

**10.73**  $H_0: p = 0.25, H_a: p > 0.25, z = 2.202, P$ -value = 0.014, fail to reject  $H_0$ . There is not convincing evidence that more than one-fourth of American adults see a lottery or sweepstakes win as their best chance of accumulating \$500,000.

**10.75**  $H_0$ : p = 0.68,  $H_a$ :  $p \neq 0.68$ , z = 7.761, *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the proportion of religion surfers who belong to a religious community is different from 0.68.

## ARE YOU READY TO MOVE ON? CHAPTER 10 REVIEW EXERCISES

**10.77**  $H_0: p = 0.37, H_a: p \neq 0.37$ 

**10.79** Failing to reject the null hypothesis means that you are not convinced that the null hypothesis is false. This is not the same as being convinced that it is true.

10.81 (a) The researchers failed to reject H<sub>0</sub>.
(b) Type II error
(c) Yes

**10.83** (a) The sampling distribution of  $\hat{p}$  will be approximately normal with a mean of 0.25 and a standard deviation of 0.0160.

(b) A sample proportion of  $\hat{p} = 0.27$  would not be surprising if p = 0.25, because it is only 1.25 standard deviations above 0.25. This is not unusual for a normal distribution.

(c) A sample proportion of  $\hat{p} = 0.31$  would be surprising if p = 0.25, because it is 3.75 standard deviations above 0.25. This would be unusual for a normal distribution.

(d) If p = 0.25,  $P(\hat{p} \ge 0.33) = P(z \ge 5) \approx 0$ . Because this probability is approximately 0, there is convincing evidence that more than 25% of law enforcement agencies review social media activity.

**10.85** Hypothesis testing, sample data, one numerical variable, one sample. A hypothesis test for a population proportion would not be appropriate.

**10.87** (a)  $H_0$ : p = 0.25,  $H_a$ : p > 0.25(b) Yes (c) z = 3.33, *P*-value  $\approx 0$ (d) reject  $H_0$  **10.89**  $H_0: p = 0.40, H_a: p > 0.40, z = 4.186,$ *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the response rate is greater than 40%.

# Chapter 11 Asking and Answering Questions About the Difference Between Two Population Proportions

NOTE: Answers may vary slightly if you are using statistical software, a graphing calculator or depending on how the values of sample statistics are rounded when performing hand calculations. Don't worry if you don't match these numerical answers exactly (but your answers should be relatively close).

## **SECTION 11.1**

#### **Exercise Set 1**

**11.1** (a) Estimation, sample data, one categorical variable, two samples. A large-sample confidence interval for a difference in proportions should be considered. (b) (-0.050, -0.002). You can be 90% confident that the actual difference between the proportion of male U.S. residents living in poverty and this proportion for females is between -0.050 and -0.002. Because both endpoints of this interval are negative, you would estimate that the proportion of men living in poverty is smaller than the proportion of women living in poverty by somewhere between 0.002 and 0.050.

**11.2** (a) Yes.  $\hat{p}_1 = 0.20$ , and  $\hat{p}_2 = 0.15$ . Then  $n_1\hat{p}_1 = 222.4$ ,  $n_1(1 - \hat{p}_1) = 889.6$ ,  $n_2\hat{p}_2 = 166.8$  and

 $\begin{array}{l} \text{Then } n_1p_1 = 222.4, n_1(1 - p_1) = 639.6, n_2p_2 = 100.6 \text{ and} \\ n_2(1 - \hat{p}_2) = 945.2 \text{ are all greater than 10.} \\ \textbf{(b)} \ (0.018, 0.082) \end{array}$ 

(c) Zero is not included in the confidence interval. This means that you can be confident that the proportion of Americans ages 12 and older who owned an MP3 player was greater in 2006 than in 2005.

(d) You can be 95% confident that the proportion of Americans ages 12 and older who owned an MP3 player was greater in 2006 than in 2005 by somewhere between 0.018 and 0.082.

**11.3** (a) (0.536, 0.744). You can be 95% confident that the proportion of avid mountain bikers who have low sperm count is higher than the proportion for nonbikers by somewhere between 0.536 and 0.744.

(**b**) No, because this was an observational study, and it is not a good idea to draw cause-and-effect conclusions from an observational study.

#### **Additional Exercises**

**11.7** (0.047, 0.113). You can be 99% confident that the proportion of high school students who believed marijuana use is very distracting in 2009 was greater than this proportion in 2011 by somewhere between 0.047 and 0.114.

**11.9** (a) (-0.089, 0.009). Because 0 is included in this interval, there may be no difference in the proportion of high school graduates who were unemployed in 2008 and the proportion who were unemployed in 2009.

(**b**) Wider, because the confidence level is greater and the sample sizes are smaller.

## **SECTION 11.2**

#### **Exercise Set 1**

**11.11** (a)  $H_0: p_1 - p_2 = 0, H_a: p_1 - p_2 > 0$ 

(b) Yes, there are more than 10 successes (those who get less than 7 hours of sleep) and 10 failures (those who do not get less than 7 hours of sleep) in each sample.

(c) z = 3.63, *P*-value = 0.000, reject  $H_0$ .

(d) There is convincing evidence that the proportion of workers who usually get less than 7 hours of sleep a night is higher for those who work more than 40 hours per week than for those who work between 35 and 40 hours per week.

**11.12**  $H_0: p_T - p_0 = 0, H_a: p_T - p_0 < 0, z = -1.667,$ *P*-value = 0.048, reject  $H_0$ . There is convincing evidence that the proportion who are satisfied is higher for those who reserve a room online.

**11.13**  $H_0: p_C - p_H = 0$   $H_a: p_C - p_H > 0$  z = 3.800, *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the proportion experiencing suburn is higher for college graduates than for those without a high school degree.

#### **Additional Exercises**

**11.17**  $H_0: p_1 - p_2 = 0, H_a: p_1 - p_2 \neq 0, z = -0.574,$ *P*-value = 0.566. Fail to reject  $H_0$ . There is not convincing evidence of a difference between the proportion of young adults who think that their parents would provide financial support for marriage and the proportion of parents who say they would provide financial support for marriage.

**11.19** Since the values given are population characteristics, an inference procedure is not applicable. It is *known* that the rate of Lou Gehrig's disease among soldiers sent to the war is higher than for those not sent to the war.

## ARE YOU READY TO MOVE ON? CHAPTER 11 REVIEW EXERCISES

**11.21**  $H_0: p_1 - p_2 = 0, H_a: p_1 - p_2 < 0, z = -8.543,$ *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the proportion of teens who approve of the proposed laws is less than the proportion of parents of teens who approve.

**11.23** (a)  $H_0: p_G - p_B = 0, H_a: p_G - p_B \neq 0, z = -0.298,$ *P*-value = 0.766, fail to reject  $H_0$ . There is not convincing evidence that the proportion who think that newspapers are boring is different for teenage girls and boys.

(**b**)  $H_0: p_G - p_B = 0, H_a: p_G - p_B \neq 0, z = -2.022,$ *P*-value = 0.043, reject  $H_0$ . There is convincing evidence that the proportion who think that newspapers are boring is different for teenage girls and boys.

(c) Assuming that the population proportions are equal, you are much less likely to get a difference in sample proportions as large as the one given when the samples are very large than when the samples are relatively small.

## Chapter 12 Asking and Answering Questions About a Population Mean

NOTE: Answers may vary slightly if you are using statistical software or a graphing calculator, or depending on how the values of sample statistics are rounded when performing hand calculations. Don't worry if you don't match these numerical answers exactly (but your answers should be relatively close).

## **SECTION 12.1**

#### **Exercise Set 1**

**12.1** (a) 100, 3.333 (b) 100, 2.582 (c) 100, 1.667 (d) 100, 1.414 (e) 100, 1.000 (f) 100, 0.500

**12.2** The sampling distribution of  $\bar{x}$  will be approximately normal for the sample sizes in Parts (c)–(f), since those sample sizes are all greater than or equal to 30.

**12.3** (1) The mean of the sampling distribution of  $\bar{x}$  is equal to the population mean  $\mu$ . (2)  $\sigma_{\bar{x}} \approx \frac{s}{\sqrt{n}} = \frac{6.62}{\sqrt{212}} = 0.455$ . (3) The sampling distribution of  $\bar{x}$  is approximately normal because the sample size (n = 212) is greater than 30.

**12.4** The quantity  $\mu$  is the population mean, while  $\mu_{\overline{x}}$  is the mean of the  $\overline{x}$  distribution. It is the mean value of  $\overline{x}$  for all possible random samples of size *n*.

**12.5** (a)  $\mu_{\bar{x}} = \mu = 0.5$ ,  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.289}{\sqrt{16}} = 0.07225$ . (b) When n = 50,  $\mu_{\bar{x}} = \mu = 0.5$  and  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.289}{\sqrt{50}} = 0.041$ . Since  $n \ge 30$  the distribution of  $\bar{x}$  is approximately normal.

#### **Additional Exercises**

**12.11** (a) 200, 4.330 (b) 200, 3.354 (c) 200, 3.000 (d) 200, 2.371 (e) 200, 1.581 (f) 200, 0.866

**12.13**  $\bar{x}$  is the mean of a single sample, while  $\mu_{\bar{x}}$  is the mean of the  $\bar{x}$  distribution. It is the mean value of  $\bar{x}$  for all possible random samples of size *n*.

**12.15** (a) 0.8185, 0.0013 (b) 0.9772, 0.0000

**12.17**  $P(0.49 < \bar{x} < 0.51) = 0.9974$ ; the probability that the manufacturing line will be shut down unnecessarily is 1 - 0.9974 = 0.0026.

## **SECTION 12.2**

#### **Exercise Set 1**

12.19 (a) 90%
(b) 95%
(c) 1%
(d) 5%

**12.20 (a)** 2.12 (b) 2.81 (c) 1.78

#### **12.21** (a) 115.0

(**b**) The 99% confidence interval is wider than the 90% confidence interval. The 90% interval is (114.4, 115.6) and the 99% interval is (114.1, 115.9).

**12.22** (a) The fact that the mean is much greater than the median suggests that the distribution of times spent volunteering in the sample was positively skewed. (b) With the sample mean much greater than the sample median, and with the sample regarded as representative of the population, it seems very likely that the population is strongly positively skewed and, therefore, not normally distributed. (c) Since  $n = 1086 \ge 30$ , the sample size is large enough for the one-sample *t* confidence interval to be appropriate, even though the population distribution is not approximately normal.

(d) (5.232, 5.968). You can be 98% confident that the mean time spent volunteering for the population of parents of school-age children is between 5.232 and 5.968 hours.

**12.23** (7.411, 8.069). You can be 95% confident that the mean time spent studying for the final exam is between 7.411 and 8.069 hours.

**12.24** (41.439, 44.921). You can be 90% confident that the mean percentage of study time that occurs in the 24 hours prior to the final exam is between 41.439% and 44.921%.

**12.25** Using (sample range)/4 = 162.5 as an estimate of the population standard deviation, a sample size of 1,015 is needed.

**12.26** (a) This could happen if the sample distribution was very positively skewed.

(b) No. Since the sample distribution is very skewed to the right, it is very unlikely that the variable anticipated Halloween expense is approximately normally distributed.(c) Yes, since the sample is large.

(d) (39.821, 53.479). You can be 99% confident that the mean anticipated Halloween expense for the population of Canadian residents is between 39.821 and 53.479 dollars.

#### **Additional Exercises**

**12.35** (a) 2.15 (b) 2.86 (c) 1.71

**12.37** (5.776, 28.224). You can be 95% confident that the mean number of months elapsed since the last visit for the population of students participating in the program is between 5.776 and 28.224.

**12.39** (3.301, 5.899). You can be 95% confident that the mean number of partners on a mating flight is between 3.301 and 5.899.

## **SECTION 12.3**

#### **Exercise Set 1**

**12.41** (a) 0.040 (b) 0.019 (c) 0.000 (d) 0.130

**12.42** (a)  $H_0: \mu = 98, H_a: \mu \neq 98, t = 0.748, P$ -value = 0.468, fail to reject  $H_0$ . There is not convincing evidence that the mean heart rate after 15 minutes of Wii Bowling is different from the mean after 6 minutes of walking on a treadmill.

**12.43** (a)  $H_0: \mu = 66, H_a: \mu > 66, t = 8.731, P-value \approx 0,$  reject  $H_0$ . There is convincing evidence that the mean heart rate after 15 minutes of Wii Bowling is higher than the mean resting heart rate.

**12.44**  $H_0$ :  $\mu = 3.57 H_a$ :  $\mu > 3.57$ , t = 2.042, *P*-value = 0.040, reject  $H_0$ . There is convincing evidence that the mean price of a Big Mac in Europe is higher than \$3.57.

**12.45** (a)  $H_0: \mu = 12.5, H_a: \mu > 12.5, t = 1.26, P-value = 0.103, fail to reject <math>H_0$ . There is not convincing evidence that the mean time spent online is greater than 12.5 hours. (b)  $H_0: \mu = 12.5, H_a: \mu > 12.5, t = 3.16, P-value = 0.0008$ , reject  $H_0$ . There is convincing evidence that the mean time spent online is greater than 12.5 hours. (c) The standard deviation was much larger in Part (a), making it more difficult to detect a difference.

**12.46** By saying that listening to music reduces pain levels, the authors are telling you that the study resulted in convincing evidence that pain levels are reduced when music is being listened to. (In other words, the results of the study were statistically significant.) By saying, however, that the magnitude of the positive effects was small, the authors are telling you that the effect was not practically significant.

#### **Additional Exercises**

**12.53**  $H_0: \mu = 20, H_a: \mu > 20, t = 0.85, P$ -value = 0.205, fail to reject  $H_0$ . There is not convincing evidence that the mean support ime is greater than 2 minutes.

**12.55**  $H_0: \mu = 5, H_a: \mu < 5, t = -5.051, P$ -value = 0.000, reject  $H_0$ . There is convincing evidence that the mean attention span for teenage boys is less than 5 minutes.

**12.57**  $H_0: \mu = 30, H_a: \mu < 30, t = -1.160, P$ -value = 0.149, fail to reject  $H_0$ . There is not convincing evidence that the mean fuel efficiency under these circumstances is less than 30 miles per gallon.

**12.59**  $H_0: \mu = 15, H_a: \mu > 15, t = 5.682, P$ -value = 0.000, reject  $H_0$ . There is convincing evidence that the mean time to 100°F is greater than 15 minutes.

## ARE YOU READY TO MOVE ON? CHAPTER 12 REVIEW EXERCISES

**12.61** (a)  $\mu_{\rm r} = 2, \sigma_{\rm r} = 0.267$ 

(**b**) In each case  $\mu_{\overline{x}} = 2$ . When n = 20,  $\sigma_{\overline{x}} = 0.179$ , and when n = 100,  $\sigma_{\overline{x}} = 0.08$ . All three centers are the same, and the larger the sample size, the smaller the standard deviation of  $\overline{x}$ . Since the distribution of  $\overline{x}$  when n = 100 is the one with the smallest standard deviation of the three, this sample size is most likely to result in a value of  $\overline{x}$  close to  $\mu$ .

#### 12.63 (a) Narrower.

(b) The statement is not correct. The population mean, μ, is a constant, and it is not appropriate to talk about the probability that it falls within a certain interval.
(c) The statement is not correct. You can say that *in the long run* about 95 out of every 100 samples will result in confidence intervals that will contain μ, but we cannot say that in 100 such samples, *exactly* 95 will result in confidence intervals that contain μ.

**12.65** (54.172, 55.928). You can be 98% confident that the mean amount of money spent per graduation gift in 2007 was between \$54.172 and \$55.928.

#### **12.67** 753

**12.69**  $H_0: \mu = 20, H_a: \mu > 20, t = 14.836, P-value = 0.000, reject <math>H_0$ . There is convincing evidence that the mean wrist extension for all people using the new mouse design is greater than 20 degrees. To generalize the result to the population of Cornell students, you need to assume that the 24 students used in the study are representative of all students at the university. To generalize the result to the population of all university students, you need to assume that the 24 students used in the study are representative of all university students, you need to assume that the 24 students used in the study are representative of all university students.

# **Chapter 13**

NOTE: Answers may vary slightly if you are using statistical software or a graphing calculator, or depending on how the values of sample statistics are rounded when performing hand calculations. Don't worry if you don't match these numerical answers exactly (but your answers should be relatively close).

## SECTION 13.1

#### **Exercise Set 1**

**13.1** Studies 1 and 4 have samples that are independently selected.

**13.2** Scenario 1: The appropriate test is not about a difference in population means. There is only one sample. Scenario 2: The appropriate test is about a difference in population means. Scenario 3: The appropriate test is not about a difference in population means. There are two samples, but the variable is categorical rather than numerical.

**13.3**  $H_0: \mu_M - \mu_F = 0, H_a: \mu_M - \mu_F > 0, t = 0.49,$ *P*-value = 0.314, fail to reject  $H_0$ . There is not convincing evidence that the mean time is greater for males than for females.

**13.4**  $H_0: \mu_F - \mu_{NF} = 0, H_a: \mu_F - \mu_{NF} < 0, t = -9.29,$ *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that the mean time spent studying is less for Facebook users than for those who do not use Facebook.

**13.5**  $H_0: \mu_M - \mu_F = 0, H_a: \mu_M - \mu_F \neq 0, t = -3.77,$ *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that the mean wait time differs for males and females.

**13.6**  $H_0: \mu_C - \mu_A = 0, H_a: \mu_C - \mu_A \neq 0, t = -1.43,$ *P*-value = 0.156, fail to reject  $H_0$ . There is not convincing evidence that the mean reaction time differs for those talking on a cell phone and those who have a blood alcohol level of 0.08%. **13.7**  $H_0: \mu_{GA} - \mu_{NGA} = 0, H_a: \mu_{GA} - \mu_{NGA} > 0, t = 3.34,$ *P*-value = 0.001, reject  $H_0$ . There is convincing evidence that the mean percentage of the time spent with the previous partner is greater for genetically altered voles than for voles that were not genetically altered.

#### **Additional Exercises**

**13.15**  $H_0: \mu_{PI} - \mu_{PP} = 0, H_a: \mu_{PI} - \mu_{PP} < 0, t = -2.63,$ *P*-value = 0.005, for a = 0.05 or 0.01, reject  $H_0$ . There is convincing evidence that the mean range of motion is less for pitchers than for position players.

**13.17**  $H_0: \mu_w - \mu_{WIO} = 0, H_a: \mu_w - \mu_{WIO} < 0, t = -3.36,$ *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that the mean brain volume is smaller for children with ADHD than for children without ADHD.

**13.19** (a) Since boxplots are roughly symmetrical and since there are no outliers in either sample, the assumption of normality is plausible, and it is reasonable to carry out a two-sample t test.

(**b**)  $H_0: \mu_{2009} - \mu_{1999} = 0$ ,  $H_a: \mu_{2009} - \mu_{1999} > 0$ , t = 3.332, *P*-value = 0.001, reject  $H_0$ . There is convincing evidence that the mean time spent using electronic media was greater in 2009 than in 1999.

**13.21** (a)  $\mu_1$  = mean payment for claims not involving errors;  $\mu_2$  = mean payment for claims involving errors;  $H_0: \mu_1 - \mu_2 = 0; H_a: \mu_1 - \mu_2 < 0.$ 

(b) Answer: (ii) 2.65. Since the samples are large, a *t* distribution with a large number of degrees of freedom is used, which can be approximated with the standard normal distribution. P(z > 2.65) = 0.004, which is the *P*-value given. None of the other possible values of *t* gives the correct *P*-value.

**13.23** (a)  $H_0: \mu_M - \mu_F = 0, H_a: \mu_M - \mu_F > 0, t = 10.359, P-value \approx 0, reject <math>H_0$ . There is convincing evidence that the mean percentage of the time spent playing with the police car is greater for males than females. (b)  $H_0: \mu_M - \mu_F = 0, H_a: \mu_M - \mu_F < 0, t = -16.316, P-value \approx 0$ , reject  $H_0$ . There is convincing evidence that

*P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the mean percentage of the time spent playing with the doll is greater for females than for males.

(c)  $H_0: \mu_M - \mu_F = 0, H_a: \mu_M - \mu_F \neq 0, t = 4.690,$ *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the mean percentage of the time spent playing with the furry dog is different for males and females.

(d) The results do seem to provide evidence of a gender basis in the monkeys' choices of how much time to spend playing with each toy, with the male monkeys spending significantly more time with the "masculine toy" than the female monkeys, and with the female monkeys spending significantly more time with the "feminine toy" than the male monkeys. However, the data also provide evidence of a difference between male and female monkeys in the time they choose to spend playing with a "neutral toy." It is possible that it was some attribute other than masculinity/femininity in the toys that was attracting the different genders of monkey in different ways.

(e) The given mean time playing with the police car and mean time playing with the doll for female monkeys are sample means for the same sample of female monkeys. The two-sample *t* test can only be performed when there are two independent random samples.

**13.25** (a)  $H_0: \mu_s - \mu_F = 0, H_a: \mu_s - \mu_F \neq 0, t = -6.565,$ *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the mean appropriateness score assigned to wearing a hat in class is different for students and faculty.

**(b)**  $H_0: \mu_s - \mu_F = 0, H_a: \mu_s - \mu_F > 0, t = 6.249,$ *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the mean appropriateness score assigned to addressing an instructor by his or her first name is higher for students than for faculty.

**13.27**  $H_0: \mu_G - \mu_{WIO} = 0, H_a: \mu_G - \mu_{WIO} > 0, t = 2.429,$ *P*-value = 0.0104, with a = 0.01, fail to reject  $H_0$ . There is not convincing evidence that the mean number of goals scored was higher for games in which Gretzky played than for games in which he did not play.

## **SECTION 13.2**

#### **Exercise Set 1**

**13.28**  $H_0: \mu_d = 0, H_a: \mu_d \neq 0, t = 0.23, P-value = 0.825,$  fail to reject  $H_0$ . There is not convincing evidence that the mean reading differs for the Mini-Wright meter and the Wright meter.

**13.29**  $H_0: \mu_d = 0, H_a: \mu_d > 0, t = 4.458, P-value = 0.001, reject <math>H_0$ . There is convincing evidence that the mean time to exhaustion is greater after chocolate milk than after carbohydrate replacement drink.

**13.30**  $H_0: \mu_d = 0, H_a: \mu_d < 0, t = -3.11, P$ -value = 0.006, reject  $H_0$ . There is convincing evidence that the mean MPF at brain location 1 is higher after diesel exposure.

**13.31**  $H_0: \mu_d = 0, H_a: \mu_d < 0, t = -4.11, P$ -value = 0.001, reject  $H_0$ . There is convincing evidence that the mean MPF at brain location 2 is higher after diesel exposure.

**13.32** (a) Eosinophils. Virtually every child showed a reduction in eosinophils percentage, with many of the reductions by large amounts.

(b)  $FE_{NO}$ . While the majority seems to have shown a reduction, there are several who showed increases, and some of those by non-negligible amounts.

#### **Additional Exercises**

**13.37**  $H_0: \mu_d = 0, H_a: \mu_d > 0, t = 3.95, P$ -value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that the mean blood pressure is higher in a dental setting than in a medical setting.

**13.39**  $H_0: \mu_d = 0, H_a: \mu_d < 0, t = -4.56$ , *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that the mean memory score is higher after completion of the program than before the program.

**13.41** (a)  $H_0: \mu_d = 0, H_a: \mu_d \neq 0, t = -1.151, P$ -value = 0.294, fail to reject  $H_0$ . There is not convincing evidence that the mean percentage of exams earning college credit at central coast high schools in 2002 differed from the percentage in 1997.

(**b**) No. The sample consisted of schools from the central coast area only, and so the conclusion cannot be generalized to all California high schools.

(c) The boxplot shows that the distribution of differences contains an outlier, and so it is not reasonable to use the paired t test on this data set.

**13.43**  $H_0: \mu_d = 0, H_a: \mu_d > 0, t = 8.134, P-value \approx 0.000,$ reject  $H_0$ . There is convincing evidence that the mean costto-charge ratio for Oregon hospitals is lower for outpatient care than for inpatient care.

**13.45**  $H_0: \mu_d = 0, H_a: \mu_d \neq 0, t = -1.34, P$ -value = 0.193, fail to reject  $H_0$ . There is not convincing evidence that the mean number of seeds detected differs for the two methods.

## **SECTION 13.3**

#### Exercise Set 1

**13.46** (0.423, 2.910). You can be 98% confident that the difference between the mean number of hours per day spent using electronic media in 2009 and 1999 is between 0.423 and 2.910.

**13.47** (-538.606, -219.394). You can be 95% confident that the mean daily calorie intake for teens who do not eat fast food on a typical day minus the mean daily calorie intake for teens who do eat fast food on a typical day is between -538.606 and -219.394.

**13.48** (-3.069, -0.791). You can be 90% confident that the mean difference in MPF at brain location 1 is between -3.069 and -0.791.

**13.49** (-2.228, -0.852). You can be 90% confident that the mean difference in MPF at brain location 1 is between -2.228 and -0.852.

**13.50** (a) There is a single sample of girls, with two data values for each girl.

(b) (-1.029, -0.631). You can be 95% confident that the mean difference between the number of science classes a girl intends to take and the number she thinks boys should take is between -1.029 and -0.631.

#### **Additional Exercises**

**13.57** A 95% confidence interval for the difference in means is (-0.230, 0.010). Because 0 is included in the interval, it is possible that there is no difference in the mean GPA for working and nonworking students.

**13.59** (2.241, 5.009). You can be 90% confident that the mean number of words remembered 1 hour later is greater than the mean number recalled 24 hours later by somewhere between 2.241 and 5.009 words.

## ARE YOU READY TO MOVE ON? CHAPTER 13 REVIEW EXERCISES

**13.61** (a)  $H_0: \mu_1 - \mu_2 = 10$  versus  $H_a: \mu_1 - \mu_2 > 10$ (b)  $H_0: \mu_1 - \mu_2 = -10$  versus  $H_a: \mu_1 - \mu_2 < -10$ 

13.63 Scenario 1: no. Scenario 2: no. Scenario 3: yes.

**13.65**  $H_0: \mu_B - \mu_M = 0, H_a: \mu_B - \mu_M \neq 0, t = 6.22,$ *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that the mean sodium content is not the same for meal purchases at Burger King and McDonalds.

**13.67**  $H_0: \mu_d = 0, H_a: \mu_d > 0, t = 2.68, P-value = 0.010,$ reject  $H_0$ . There is convincing evidence that the mean energy expenditure is greater for the conventional shovel than for the perforated shovel.

**13.69** (a)  $H_0: \mu_d = 0, H_a: \mu_d > 0, t = 4.451, P$ -value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that, on average, male online daters overstate their height.

(b) (-0.210, 0.270). You can be 95% confident that for female online daters, the mean difference in reported height and actual height is between -0.210 and 0.270 inches.

(c)  $H_0: \mu_M - \mu_F = 0, H_a: \mu_M - \mu_F > 0, t = 3.094,$ *P*-value = 0.001, reject  $H_0$ . There is convincing evidence that the mean height difference is greater for males than for females.

(d) In Part (a), the male profile heights and the male actual heights are paired (according to which individual has the actual height and the height stated in the profile), and with paired samples, you use the paired t test. In Part (c), you were dealing with two independent samples (the sample of males and the sample of females), and therefore, the two-sample t test was appropriate.

**13.71** (a) (-0.186, 0.111). You can be 90% confident that the mean difference in pH for surface and subsoil is between -0.186 and 0.111. Because 0 is included in the interval, it is possible that there is no difference in mean pH.

(**b**) You must assume that the distribution of differences across all locations is normal.

# Chapter 14

NOTE: Answers may vary slightly if you are using statistical software or a graphing calculator, or depending on how the values of sample statistics are rounded when performing hand calculations. Don't worry if you don't match these numerical answers exactly (but your answers should be relatively close).

## **SECTION 14.1**

#### **Exercise Set 1**

**14.1** The mean quiz score for those in the texting group is enough lower than the mean for the no-texting group that a difference this large is not likely to have occurred by chance (due just to the random assignment) when there is no real difference between the treatment means.

**14.2** The difference in mean calorie intake for the 4-hour and 8-hour sleep groups could have occurred by chance just due to the random assignment.

#### **Additional Exercises**

14.5 (a) The two treatments are small fork and large fork.(b) The mean amount of food for those in the small fork group is enough larger than the mean for the small fork group that a difference this large is not likely to have occurred by chance (due just to the random assignment) when there is no real difference between the treatment means.

**14.7** The mean value assigned to the mug for those in the group given the mug is enough higher than the mean for the other group that a difference this large is not likely to have occurred by chance (due just to the random assignment) when there is no real difference between the treatment means.

## **SECTION 14.2**

#### **Exercise Set 1**

**14.9**  $H_0: \mu_N - \mu_T = 0, H_a: \mu_N - \mu_T > 0, t = 21.78,$ *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that the mean test score is higher for the new teaching method.

**14.10**  $H_0: \mu_G - \mu_{NG} = 0, H_a: \mu_G - \mu_{NG} > 0, t = 11.753,$ *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that learning the gesturing approach to solving problems of this type results in a higher mean number of correct responses.

**14.11**  $H_0: p_w - p_B = 0, H_a: p_w - p_B > 0, z = 4.245,$ *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that the proportion receiving a response is higher for white-sounding first names.

#### **Additional Exercises**

**14.15**  $H_0: \mu_T - \mu_{NT} = 0, H_a: \mu_T - \mu_{NT} < 0, t = -6.14,$ *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that the mean quiz score is lower for students who text. **14.17** No. It is not appropriate to use the two-sample *z* test because the groups are not large enough. You are not told the sizes of the groups, but you know that each is, at most, 81. The sample proportion for the fish oil group is 0.05, and 81(0.05) = 4.05, which is less than 10. So, the conditions for the two-sample *z* test are not satisfied.

**14.19** (a)  $H_0: \mu_s - \mu_D = 0, H_a: \mu_s - \mu_D \neq 0, t = -11.952$ , *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that mean elongations for the square knot and the Duncan loop for Maxon thread are not the same. (b)  $H_0: \mu_s - \mu_D = 0, H_a: \mu_s - \mu_D \neq 0, t = -68.803$ , *P*-value  $\approx 0.000$ , reject  $H_0$ . There is convincing evidence that mean elongations for the square knot and the Duncan loop for Ticron thread are not the same.

## SECTION 14.3

#### **Exercise Set 1**

**14.21** (-814.629, 524.096). You can be 95% confident that the difference in mean food intake for the two sleep conditions is between -814.629 and 524.096. Because 0 is included in the confidence interval, it is possible that there is no difference in the treatment means.

**14.22** (a) If the vertebroplasty group had been compared to a group of patients who did not receive any treatment, and if, for example, the people in the vertebroplasty group experienced a greater pain reduction on average than the people in the "no treatment" group, then it would be impossible to tell whether the observed pain reduction in the vertebroplasty group was caused by the treatment or merely by the subjects' knowledge that some treatment was being applied. By using a placebo group, it is ensured that the subjects in both groups have the knowledge of some "treatment," so that any differences between the pain reduction in the two groups can be attributed to the nature of the vertebroplasty treatment.

(b) (-0.687, 1.287). You can be 95% confident that the difference in mean pain intensity three days after treatment for the vertebroplasty treatment and the fake treatment is between -0.687 and 1.287.

**14.23** (a) At 14 days: (-1.186, 0.786). You can be 95% confident that the difference in mean pain intensity 14 days after treatment for the vertebroplasty treatment and the fake treatment is between -1.186 and 0.786. At one month: (-1.722, 0.322). You can be 95% confident that the difference in mean pain intensity one month after treatment for the vertebroplasty treatment and the fake treatment is between -1.722 and 0.322.

(b) The fact that all of the intervals contain zero tells you that you do not have convincing evidence of a difference in the mean pain intensity for the vertebroplasty treatment and the fake treatment at any of the three times.

14.24 (a) If people believe that one of the treatments is more effective than the other (for example that the injection is more effective than the spray), it might affect how they interpret and report symptoms.
(b) (0.033, 0.061). You can be 99% confident that the proportion of children who get sick with the flu after being vaccinated with an injection minus the proportion of children who get sick with the flu after being vaccinated with the nasal spray is between 0.033 and 0.061. Since zero is not included in the interval, you can say that the proportions of children who get the flu are different for the two vaccination methods.

#### **Additional Exercises**

**14.27** (a) (-0.135, 0.575). You can be 95% confident that the difference in mean overall course evaluation is between -0.135 and 0.575.

(**b**) The interval does not support the statement made in the title. Zero is in the confidence interval, indicating that there may be no real difference between the two treatment means.

**14.29** (-20.176, -11.544). You can be 90% confident that the difference in mean quiz score for the two treatments is between -20.176 and -11.544.

**14.31** Statement 3 is correct. The claim is that  $\mu_1 > \mu_2$ . This means that both endpoints for a confidence interval for  $\mu_1 - \mu_2$  would be positive.

## ARE YOU READY TO MOVE ON? CHAPTER 14 REVIEW EXERCISES

**14.33** The difference in mean sleep duration for the caffeine and no-caffeine groups could have occurred by chance just due to the random assignment. This does not mean you should be convinced the treatment means are equal—only that there is not convincing evidence that they are different.

**14.35**  $H_0: p_E - p_S = 0, H_a: p_E - p_S > 0, z = 1.209,$ *P*-value = 0.113, fail to reject  $H_0$ . There is not convincing evidence that the proportion of patients who improve is greater for the experimental treatment than for the standard treatment.

**14.37** (a) The two treatments are white-sounding names and black-sounding names.

(**b**) For a confidence level of 95%, (0.018, 0.048). You can be 95% confident that the difference in the proportions of responses for white-sounding names and black-sounding names is between 0.018 and 0.048.

# **Chapter 15**

NOTE: Answers may vary slightly if you are using statistical software or a graphing calculator, or depending on how the values of sample statistics are rounded when performing hand calculations. Don't worry if you don't match these numerical answers exactly (but your answers should be relatively close).

## **SECTION 15.1**

#### **Exercise Set 1**

15.1 (a) fail to reject H<sub>0</sub>
(b) reject H<sub>0</sub>
(c) reject H<sub>0</sub>
15.2 (a) 0.085
(b) 0.075

(c) 0.025

**15.3**  $H_0: p_1 = p_2 = p_3 = p_4 = p_5 = 0.2, X^2 = 20.686,$ *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the Twitter category proportions are not all equal.

**15.4** (a)  $H_0$ :  $p_1 = p_2 = p_3 = p_4 = p_5 = 0.2$ ,  $X^2 = 25.241$ , *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the proportions of home runs hit are not the same for all five directions.

(b) For home runs going to left center and center field, the observed counts are significantly lower than the numbers that would have been expected if the proportion of home runs hit was the same for all five directions; while for right field, the observed count is much high than the number that would have been expected.

**15.5**  $H_0: p_1 = 0.35, p_2 = 0.51, p_3 = 0.14, X^2 = 25.486,$ *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that one or more of the age groups buys a disproportionate share of lottery tickets.

#### Additional Exercises

**15.11** (**a**) 0.001 (**b**) 0.001 (**c**) *P*-value > 0.100

**15.13** Reject  $H_0$ . There is convincing evidence that the response proportions are not each 0.5.

**15.15**  $p_1 = \frac{9}{16}, p_2 = \frac{3}{16}, p_3 = \frac{3}{16}, p_4 = \frac{1}{16}, X^2 = 1.469,$ *P*-value = 0.690, fail to reject  $H_0$ . The data are consistent with Mendel's laws.

**15.17**  $H_0: p_1 = p_2 = p_3 = p_4 = 0.25, X^2 = 11.242,$ *P*-value = 0.0105, fail to reject  $H_0$ . There is not convincing evidence that there is a color preference.

## SECTION 15.2

#### **Exercise Set 1**

**15.19**  $H_0$ : The proportions falling into the three credit card response categories are the same for all three years,  $H_a$ : The proportions falling into the three credit card response categories are not all the same for all three years,  $X^2 = 5.204$ , *P*-value = 0.267, fail to reject  $H_0$ . There is not convincing evidence that the proportions falling into the three credit card response categories are different for the three years.

**15.20** (a)  $H_0$ : donation proportions are the same for all three gift types,  $H_a$ : donation proportions are not the same for all three gift types,  $X^2 = 96.506$ , *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the donation proportions are not the same for all three gift types.

(b) The result of Part (a) tells you that the level of the gift seems to make a difference. Looking at the data given, 12% of those receiving no gift made a donation, 14% of those receiving a small gift made a donation, and 21% of those receiving a large gift made a donation. So, it seems that the most effective strategy is to include a large gift, with the small gift making very little difference compared to no gift at all.

**15.21** (a)  $H_0$ : field of study and smoking status are independent,  $H_a$ : field of study and smoking status are not independent,  $X^2 = 90.853$ , *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that field of study and smoking status are not independent.

(b) The particularly high contributions to the chi-square statistic (in order of importance) come from the field of communication, languages, in which there was a disproportionately high number of smokers; from the field of mathematics, engineering, and sciences, in which there was a disproportionately low number of smokers; and from the field of social science and human services, in which there was a disproportionately high number of smokers.

**15.22**  $H_0$ : Locus of control and compulsive buyer behavior are independent,  $H_a$ : Locus of control and compulsive buyer behavior are not independent,  $X^2 = 5.402$ , *P*-value = 0.020, fail to reject  $H_0$ . There is not convincing evidence that there is an association between locus of control and compulsive buyer behavior.

#### **Additional Exercises**

15.27 Answers will vary.

**15.29** In a chi-square goodness-of-fit test, one population is compared to fixed category proportions. In a chi-square test of homogeneity, two or more populations are compared.

**15.31**  $H_0$ : There is no association between use of ART and whether baby is premature,  $H_a$ : There is an association between use of ART and whether baby is premature,  $X^2 =$  326.111, *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that there is an association between use of ART and whether the baby is premature.

**15.33**  $H_0$ : The proportions who believe that the story described a rape are the same for the three photo treatments,  $H_a$ : The proportions who believe that the story described a rape are not the same for the three photo treatments,  $X^2 = 28.810$ , *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the proportions who believe that the story described a rape are not the same for the three photo treatments.

**15.35**  $H_0$ : There is no association between city and vehicle type,  $H_a$ : There is an association between city and vehicle type,  $X^2 = 49.813$ , *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that there is an association between city and vehicle type.

**15.37**  $H_0$ : Age of children and parental response are independent,  $H_a$ : Age of children and parental response are not independent,  $X^2 = 10.091$ , *P*-value = 0.006, reject  $H_0$ . There is convincing evidence that there is an association between age of children and parental response.

**15.39** (a) Since the study was conducted using separate random samples of male and female inmates, this is a test for homogeneity.

(**b**)  $H_0$ : The proportions falling into the crime categories are the same for male and female inmates,  $H_a$ : The proportions falling into the crime categories for male and female inmates are not all the same,  $X^2 = 28.511$ , *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that the proportions falling into the crime categories for male and female inmates are not all the same.

## ARE YOU READY TO MOVE ON? CHAPTER 15 REVIEW EXERCISES

**15.41**  $H_0: p_1 = p_2 = \cdots = p_8 = \frac{1}{8}, X^2 = 166.958,$ *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that accidents are not equally likely to occur in each of the eight

time periods. **15.43** (a)  $H_0: p_1 = p_2 = \dots = p_{12} = \frac{1}{12}, X^2 = 82.163,$ *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that fatal bicycle accidents are not equally likely to occur in which of the events.

each of the months. (b)  $H_0$ :  $p_1 = \frac{31}{366}$ ,  $p_2 = \frac{29}{366}$ ,  $p_3 = \frac{31}{366}$ ,  $p_4 = \frac{30}{366}$ ,  $p_5 = \frac{31}{366}$ ,  $p_6 = \frac{30}{366}$ ,  $p_7 = \frac{31}{366}$ ,  $p_8 = \frac{31}{366}$ ,  $p_9 = \frac{30}{366}$ ,  $p_{10} = \frac{31}{366}$ ,  $p_{11} = \frac{30}{366}$ ,  $p_{12} = \frac{31}{366}$ 

(c)  $X^2 = 78.511$ , *P*-value  $\approx 0$ , reject  $H_0$ . There is convincing evidence that fatal bicycle accidents do not occur in the 12 months in proportion to the lengths of the months.

**15.45**  $H_0$ : The spirituality category proportions are the same for natural scientists and social scientists,  $H_a$ : The

spirituality category proportions are not all the same for natural scientists and social scientists,  $X^2 = 3.091$ , *P*-value = 0.378, fail to reject  $H_0$ . There is not convincing evidence that the spirituality category proportions are different for natural scientists and social scientists.

**15.47**  $H_0$ : There is no association between whether or not a child is overweight and the number of sweet drinks consumed,  $H_a$ : There is an association between whether or not a child is overweight and the number of sweet drinks consumed,  $X^2 = 3.030$ , *P*-value = 0.387, fail to reject  $H_0$ . There is not convincing evidence that there is an association between whether or not a child is overweight and the number of sweet drinks consumed.

**15.49** (a)  $H_0$ : The proportions in the water consumption categories are the same for males and females,  $H_a$ : The proportions in the water consumption categories are not the same for males and females, df = 4.

(**b**) The *P*-value for the test was 0.086, which is greater than the new significance level of 0.05. So, for a significance

level of 0.05, you do not have convincing evidence of a difference between males and females with regard to water consumption.

#### 15.51 (a)

	Napped	Did Not Nap	Row Total
Men	283	461	744
Women	231	513	744

(**b**)  $H_0$ : There is no association between gender and napping,  $H_a$ : There is an association between gender and napping,  $X^2 = 8.034$ , *P*-value = 0.005, if a = 0.05, reject  $H_0$ . There is convincing evidence that there is an association between gender and napping.

(c) Yes. There is evidence of an association between gender and napping in the population. The number who nap was greater than expected for men and less than expected for women.