

APPENDIX B

Solutions to Odd-Numbered End-of-Chapter Exercises

Chapter 1

- 1.01 summarize
1.03 cases
1.05 (a) X causes Y ; (b) Y causes X ; (c) some other variable (Z or a confounding variable) causes both X and Y .
1.07 random assignment
1.09 dependent variable (or outcome variable)
1.11 groups
1.13 independent = cause; effect = dependent
1.15 grouping
1.17 NOIR
1.19 ordinal
1.21 arbitrary; absolute
1.23 population
1.25 statistic; parameter
1.27 descriptive; inferential
1.29 two
1.31 (a) Do metal and plastic handcuffs differ in how much abrasion they cause?
(b) Type of handcuff is the explanatory variable; amount of abrasion is the outcome variable.
(c) Experimental
(d) N/A
1.33 (a) Are “conscientious” students smarter than “nonconscientious” students?
(b) Type of student (conscientious vs. nonconscientious) is the explanatory variable; IQ is the outcome variable.
(c) Quasi-experimental
(d) Socioeconomic status may be a confounding variable. Richer students may be able to afford to be conscientious (to buy all the books, not to have to work so they have more time for studying) and richer students may have gone to better schools, received a better education, and ended up with a higher IQ.
1.35 (a) Are a country’s wealth and its greenhouse gas production related?
(b) GDP is the explanatory variable; tons of CO₂ produced is the outcome variable.
(c) Correlational
(d) Population could be a confounding variable. Over the last 50 years, as there were more people in the country, there would be more people to produce and consume goods, increasing both GDP and greenhouse gases.
1.37 (a) Does type of information received affect voting behavior?
(b) Type of information (none vs. positive vs. negative) is the explanatory variable; voting for or against is the outcome variable.
(c) Experimental
(d) N/A
1.39 Nominal
1.41 Ordinal
1.43 Interval
1.45 Ordinal
1.47 Nominal
1.49 (a) Sample
(b) Parameters
(c) Inferential
1.51 Parameter
1.53 $N = 6$
1.55 $\sum X^2 = 8^2 + 9^2 + 5^2 + 4^2 + 7^2 + 8^2 = 299.00$
1.57 $\sum X = 13 + 18 + 11 = 42.00$
1.59
$$\frac{\sum X}{N} = \frac{(13 + 18 + 11)}{3} = \frac{42.0000}{3} = 14.00$$

1.61 12.68
1.63 121.01
1.65 22.47
1.67 2.53
1.69 c
1.71 (a) Do higher levels of physical distress in cities lead to higher levels of social distress?
(b) Amount of graffiti (high, moderate, or low) is the explanatory variable, and teenage pregnancy rate is the outcome variable.
(c) quasi-experimental
(d) The proportion of the city population that has a high school degree could be a confounding variable. Less-educated teenage girls may be more likely to get pregnant. City residents who have not completed high school may be less likely to be employed and more likely to deface property.

B-2 Appendix B

Chapter 2

- 2.01 count
2.03 values; frequencies
2.05 f_c
2.07 nominal
2.09 grouped
2.11 details
2.13 midpoint
2.15 how many
2.17 nominal; ordinal
2.19 in-between (fractional)
2.21 interval; ratio
2.23 range
2.25 the interval width
2.27 bar graph
2.29 wider; tall
2.31 Y
2.33 touch
2.35 nominal
2.37 modality; skewness; kurtosis
2.39 bimodal
2.41 negative
2.43 ungrouped
2.45 Don't forget a title!

Frequency Distribution for Number of Psychiatric Diagnoses in 45 Residents of a Psychiatric Hospital

Number of Diagnoses	Frequency	Cumulative Frequency	Percentage	Cumulative Percentage
5	1	45	2.22	100.00
4	2	44	4.44	97.78
3	9	42	20.00	93.33
2	19	33	42.22	73.33
1	12	14	26.67	31.11
0	2	2	4.44	4.44

2.47

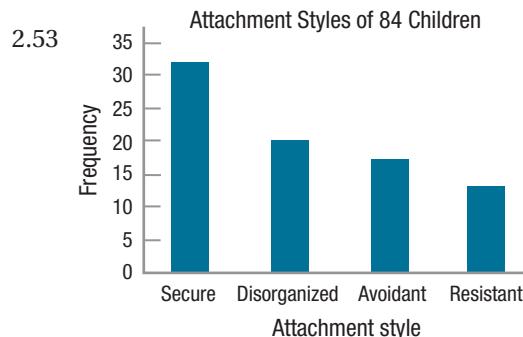
Grouped Frequency Distribution for Number of Chairs in 41 College Classrooms (Interval Width = 20)

Number of Chairs	Interval Midpoint	Frequency	Cumulative Frequency	Percentage	Cumulative Percentage
90–109	99.50	1	41	2.44	100.00
70–89	79.50	1	40	2.44	97.56
50–69	59.50	11	39	26.83	95.12
30–49	39.50	17	28	41.46	68.29
10–29	19.50	11	11	26.83	26.83

- 2.49 (a) Discrete
(b) Continuous
(c) Discrete
(d) Discrete
(e) Continuous

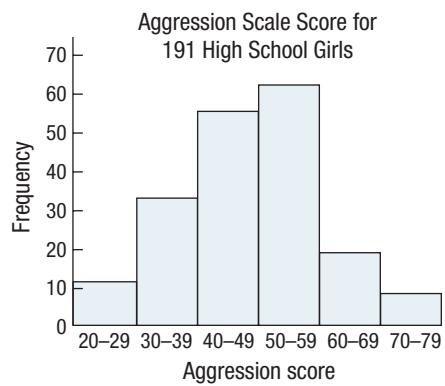
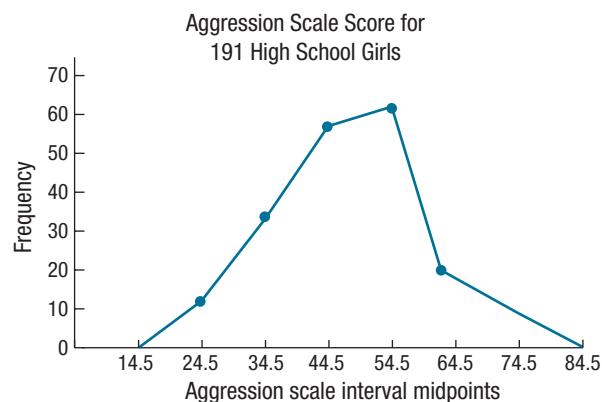
2.51

	Real Lower Limit	Real Upper Limit	Interval Width	Midpoint
a.	30.50	34.50	4.00	32.50
b.	25.00	45.00	20.00	35.00
c.	2.495	2.995	.500	2.745
d.	10.50	11.50	1.00	11.00



Since the data are nominal, the order is arbitrary and discussing the shape of the distribution does not make sense.

- 2.55 Either a histogram or frequency polygon can be made for these data.



The distribution is unimodal, doesn't seem too peaked or too flat, and is not skewed.

2.57	1 12
	2 13779
	3 003459
	4 357
	5 33
	6 7
	7 5
	8 4
	9 6

The distribution is unimodal, doesn't seem too peaked or too flat, and is positively skewed.

- 2.59 b
2.61 e
2.63 Shortest time: 4 minutes: Leave at 10:16:59, return at 10:21:00. 241 seconds, or 4 minutes and one second.
Longest time: 6 minutes: Leave at 10:16:00, return at 10:21:59. 359 seconds, or 5 minutes and 59 seconds.

Chapter 3

- 3.01 central tendency; variability
3.03 variability
3.05 0
3.07 less
3.09 mean
3.11 mode
3.13 interval; ratio
3.15 is
3.17 more
3.19 population; sample
3.21 variance
3.23 cases
3.25 tightly
3.27 square
3.29 σ
3.31 corrected; population
3.33 deviation score
3.35 The data are ratio-level, but there is an outlier, so the median is the most appropriate measure of central tendency. $Mdn = 4.00$.
3.37 They are ratio-level data, and there isn't anything unusual in terms of skewness or modality, so the mean is the most appropriate measure of central tendency.

$$M = \frac{\Sigma X}{N} = 23.17$$

$$3.39 M = \frac{80 + 88 + 76 + 65 + 59 + 77}{6} = 74.17$$

- 3.41 $N = 47$, so median is the score associated with case number $\frac{47 + 1}{2} = 24$. Case number 24

has a value of 102, so $Mdn = 102.00$. The value of 102 is also the most commonly occurring score, so mode = 102.00.

- 3.43 (a) Range = $201 - 37 = \$164$ million. (b) The interquartile range is a better measure of variability to report because it trims off the extreme scores. And the value of \$201 million seems to be an outlier, more than \$50 million higher than the next value.

- 3.45 (a) $N = 5$
(b) $\Sigma(X - M) = (47 - 50.00) + (53 - 50.00) + (67 - 50.00) + (45 - 50.00) + (38 - 50.00)$
 $= 0.00$
(c) $\Sigma(X - M)^2 = (47 - 50.00)^2 + (53 - 50.00)^2 + (67 - 50.00)^2 + (45 - 50.00)^2 + (38 - 50.00)^2$
 $= 476.00$

B-4 Appendix B

$$3.47 s^2 = \frac{\Sigma(X - M)^2}{N - 1}$$
$$= \frac{(57 - 63)^2 + (60 - 63)^2 + (67 - 63)^2 + (68 - 63)^2}{4 - 1}$$
$$= 28.67$$

$$3.49 M = \frac{\Sigma X}{N} = \frac{21 + 18 + 12 + 9}{4} = 15.00$$
$$s^2 = \frac{\Sigma(X - M)^2}{N - 1}$$
$$= \frac{(21 - 15.00)^2 + (18 - 15.00)^2 + (12 - 15.00)^2 + (9 - 15.00)^2}{4 - 1}$$
$$= 30.00$$

$$3.51 s = \sqrt{s^2} = \sqrt{6.55} = 2.56$$

$$3.53 s = \sqrt{\frac{\Sigma(X - M)^2}{N - 1}}$$
$$= \sqrt{\frac{(24.00 - 37.20)^2 + (36.00 - 37.20)^2 + (42.00 - 37.20)^2 + (50.00 - 37.20)^2 + (34.00 - 37.20)^2}{5 - 1}}$$
$$= 9.65$$

3.55 d

3.57 b

3.59 (a) The larger sample, $N = 100$, will have a smaller correction for s^2 to approximate the population variance. (b) As sample size gets larger, the sample probably captures more of the variability that exists in the population, so the formula “corrects” the standard deviation less.

Chapter 4

- 4.01 raw
- 4.03 mean; standard deviation
- 4.05 zero
- 4.07 mean; standard deviation
- 4.09 mean; median
- 4.11 decreases
- 4.13 rare
- 4.15 random
- 4.17 0.01
- 4.19 the same
- 4.21 less than
- 4.23 at or below

- 4.25 probability
- 4.27 the total number of possible outcomes
- 4.29 fraction; proportion; percentage
- 4.31 1.00; .00 (or, 100%; 0%)
- 4.33 .05; likely
- 4.35 3
- 4.37 $z = \frac{X - M}{s} = \frac{9.5 - 7}{2} = 1.25$
- 4.39 $X = M + (z \times s) = 12.75 + (1.45 \times 3.33)$
$$= 17.58$$
- 4.41 $z = \frac{X - M}{s} = \frac{73 - 100}{15} = -1.80$
- 4.43 $X = M + (z \times s) = 100 + (0 \times 15) = 100.00$

(One could also complete this problem without doing any math as a z score of zero is right at the mean, and the mean for IQ scores is 100.)
- 4.45 $z_{\text{SAT}} = \frac{X - M}{s} = \frac{620 - 500}{100} = 1.20$
$$z_{\text{Spelling}} = \frac{X - M}{s} = \frac{72 - 60}{15} = 0.80$$

Hillary did better on the math subtest of the SAT than she did on the spelling test.
- 4.47 9.01%
- 4.49 80.23%
- 4.51 47.50%
- 4.53 $z = \frac{X - M}{s} = \frac{90 - 77}{11} = 1.18$

11.90% of cases in a normal distribution have a z score ≥ 1.18 . Thus, 11.90% of Americans should have high diastolic blood pressure (90 or higher).
- 4.55 $PR = 1.70$
- 4.57 $z = \frac{X - M}{s} = \frac{113 - 100}{15} = 0.87$
$$PR = 80.78$$
$$PR = 42.07$$
- 4.59 PR of 88.5 = z score of 1.20
 $SAT = X = M + (z \times s) = 500 + (1.20 \times 100)$
$$= 620.00$$
- 4.61 $p(\text{boy}) = \frac{18}{28} = .64$
- 4.63 This is 5% on each side of the midpoint. Use column B: ± 0.13
- 4.65 First find z scores for middle 84% (42% on each side of the midpoint). Use column B: ± 1.41 . Then calculate IQ scores: $100 + (-1.41 \times 15) = 78.85$ and $100 + (1.41 \times 15) = 121.15$.

- The middle 84% of IQ scores fall from an IQ score of 78.85 to 121.15. (That could also be written as 100 ± 21.15 .)
- 4.67 That's 2% in each tail. Use column C: ± 2.05
- 4.69 That's 7.5% in each tail. Using column C, z scores = ± 1.44 . Then calculate SAT subtest scores: $SAT = 500 + (-1.44 \times 100) = 356.00$ and $SAT = 500 + (1.44 \times 100) = 644.00$. The SAT scores are 356.00 and 644.00.
- 4.71 0.135%, using column C
- 4.73 Using column B, 14.80% of scores fall from the mean to a z score of 0.38, so $2 \times 14.80 = 29.60\%$ of scores fall within 0.38 standard deviations of the mean.
- 4.75 Using column B, 19.15% of scores fall from the mean to a z score of 0.5, so $2 \times 48.21 = 38.30\%$ of scores fall within a half standard deviation of the mean. Thus, $p(\text{Not with a half standard deviation}) = 1 - .3830 = .6170$, or 61.7%.
- 4.77 That's 5% in each tail. Using column C, the z score should be 1.645.
- 4.79 That's 5% in the upper tail. Using column C, the z score should be 1.645.
- 4.81 1.00
- 4.83 c
- 4.85 $z = -5.00$. (z scores are deviation scores and the sum of a set of deviation scores is always zero. The first four cases sum to 5, so the one missing case must have a value of -5 in order to bring the sum of all five z scores to zero.)
- 4.87 There are more genius dogs than genius cats. Cats, with a smaller standard deviation, have scores that cluster closer to the mean. Thus, they have fewer scores that fall far away from the mean, which means fewer cats will score in the genius range. (This also means that there are more stupid dogs than stupid cats.)
- 4.89 Since the normal distribution is symmetric, the mean is $\frac{15 + 45}{2} = 30$. Assuming a very, very rare event has a z score of 3, the standard deviation is $\frac{45 - 30}{3} = 5$.
- ## Chapter 5
- 5.01 population; subset
- 5.03 attributes; proportion
- 5.05 convenience
- 5.07 random number table ("computer" is also an acceptable answer)
- 5.09 population
- 5.11 differ
- 5.13 70
- 5.15 random
- 5.17 small; large
- 5.19 standard error of the mean
- 5.21 30
- 5.23 mean
- 5.25 s
- 5.27 range; sample; population
- 5.29 1.96
- 5.31 $M = 10.5$, 95%CI [7.32, 13.68]
- 5.33 07, 18, 32, 20, 39, 02, 31, 45, 14, and 24
- 5.35 Consent rate is less than 70%, so the director should not pay attention to the results (or should at least be skeptical).
- 5.37 $\sigma_M = \frac{\sigma}{\sqrt{N}} = \frac{4}{\sqrt{88}} = 0.43$
- 5.39 $s_M = \frac{s}{\sqrt{N}} = \frac{7.50}{\sqrt{72}} = 0.88$
- 5.41 $95\%CI_{\mu} = M \pm (1.96 \times \sigma_M)$
 $= 17.00 \pm (1.96 \times 4)$
 $= 17.00 \pm 7.84$
 $M = 17.00$, 95%CI [9.16, 24.84]
- 5.43 $95\%CI_{\mu} = M \pm (1.96 \times s_M)$
 $= M \pm \left(1.96 \times \frac{s}{\sqrt{n}}\right)$
 $= -12.00 \pm \left(1.96 \times \frac{17}{\sqrt{72}}\right)$
 $= -12.00 \pm 3.93$
 $M = -12.00$, 95%CI [-15.93, -8.07]
- 5.45 c
- 5.47 As the population standard deviation increases, so does the standard error of the mean as long as the sample size is held constant.

■ B-6 Appendix B

$$5.49 \text{ 90\%CI}_{\mu} = M \pm (1.645 \times \sigma_M)$$

$$= M \pm \left(1.645 \times \frac{\sigma}{\sqrt{n}} \right)$$

$$= 100.00 \pm \left(1.645 \times \frac{12}{\sqrt{81}} \right)$$

$$= 100.00 \pm 2.19$$

$$M = 100.00, \text{ 90\%CI [97.81, 102.19]}$$

5.51 Increase the sample size, which would reduce the standard error of the mean.

Chapter 6

6.01 explanation; facts

6.03 hypothesis

6.05 H_0 ; H_1

6.07 specific; negative

6.09 alternative

6.11 disprove

6.13 true

6.15 sampling error

6.17 (1) Pick a test, (2) check the assumptions, (3) state the hypotheses, (4) set the decision rule, (5) calculate the value of the test statistic, (6) interpret the results.

6.19 assumptions

6.21 nondirectional

6.23 one; two

6.25 critical value

6.27 interpretation

6.29 population

6.31 influence

6.33 population mean

6.35 zero

6.37 rare zone (or extreme sections)

6.39 rare

6.41 -1.96 ; $+1.96$

6.43 z ; $+1.96$

6.45 α

6.47 sample mean; population mean

6.49 has rejected

6.51 statistically different

6.53 $p < .05$

6.55 does not

6.57 be rejected

6.59 can't

6.61 5%

6.63 larger; easier

6.65 low; reject

6.67 5% (or .05)

6.69 Type II; Type I

6.71 it should be rejected

6.73 Type I; Type II

6.75 a single-sample z test

6.77 The random sample assumption is not violated. The independence assumption is not violated. School readiness is a psychological variable and probably is normally distributed. Since no assumptions are violated, it is OK to proceed with the single-sample z test.

6.79 $H_0: \mu = 60$

$H_1: \mu \neq 60$

6.81 If $z \leq -1.96$ or if $z \geq 1.96$, reject H_0 .

If $-1.96 < z < 1.96$, fail to reject H_0 .

$$6.83 \sigma_M = \frac{\sigma}{\sqrt{N}} = \frac{10}{\sqrt{58}} = 1.31$$

$$6.85 z = \frac{M - \mu}{\sigma_M} = \frac{100 - 120}{17.5} = -1.14$$

$$6.87 \text{ (a)} \sigma_M = \frac{\sigma}{\sqrt{N}} = \frac{5}{\sqrt{28}} = 0.94$$

$$\text{ (b)} z = \frac{M - \mu}{\sigma_M} = \frac{12 - 10}{0.94} = 2.13$$

6.89 (a) $2.37 \geq 1.96$, so H_0 rejected;

(b) $z (N=23) = 2.37, p < .05$

6.91 (a) There was a statistically significant difference; (b) the sample mean is higher than the population mean.

6.93 A researcher compared the mean weight of American women who want to lose weight (178 pounds) to the average weight of American women (164 pounds). The difference was statistically significant, $z (N = 123) = 3.68, p < .05$. Women in America who do want to lose weight are heavier than American women in general. In future studies, it would be important to make sure that the sample is representative of women who want to lose weight.

6.95 (a) He made a Type I error. (b) He should have concluded that there is not enough evidence to say a difference exists.

6.97 power = $1 - \beta = 1 - .75 = .25$

6.99 **Step 1** Pick a Test

Comparing a sample mean to a population mean where the population standard deviation is known calls for a single-sample z test.

Step 2 Check the Assumptions

The random samples assumption is not violated. There is no evidence of violation of the independence of observations assumption. The normality assumption is not violated, as it seems reasonable to assume sodium consumption is normally distributed. No assumptions are violated, so it is OK to proceed with the single-sample z test.

Step 3 List the Hypotheses

The hypotheses are nondirectional.

$$H_0: \mu_{\text{Dieters}} = 3,400$$

$$H_1: \mu_{\text{Dieters}} \neq 3,400$$

Step 4 Set the Decision Rule

It is a two-tailed test and the dietitian is willing to make a Type I error 5% of the time, so $\alpha = .05$:

Reject H_0 if $z \leq -1.96$ or if $z \geq 1.96$.

Fail to reject H_0 if $-1.96 < z < 1.96$.

Step 5 Calculate the Test Statistic

$$\sigma_M = \frac{\sigma}{\sqrt{N}} = \frac{270}{\sqrt{172}} = 20.59$$

$$z = \frac{M - \mu}{\sigma_M} = \frac{2,900 - 3,400}{20.59} = -24.28$$

Step 6 Interpret the Results

A dietitian compared the sodium intake of a random sample of dieting Americans ($M = 2,900$ mg sodium) to the mean daily sodium intake of Americans ($\mu = 3,400$ mg sodium). The difference was a statistically significant one ($z(N = 172) = -24.28$, $p < .05$). Dieters' sodium intake is lower than for Americans in general. From this study, it is not clear whether this is because they were dieting or if their sodium intake were lower to begin with. Future research should investigate this.

- 6.101 b (If $\beta = .60$, then power = .40.)
- 6.103 b (The size of the effect, the difference between the two means, is larger.)
- 6.105 correct
- 6.107 incorrect; Type II error
- 6.109 There is a way to calculate beta, but we haven't covered it. I hope you'll take a more advanced statistics class some time and learn

how to calculate it. So, why ask this question?

Because I don't want you to assume that $\beta = .05$, like α . And, I don't want you to assume that $\beta = .20$.

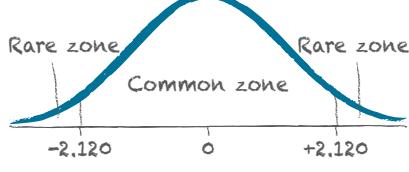
- 6.111 (a) This is a one-tailed test because he has made a prediction about the direction of the results before collecting any data.
 $(b) H_0: \mu_{\text{Athletes}} \geq 78$
 $H_1: \mu_{\text{Athletes}} < 78$
- 6.113 The statistician is comparing the mean of a sample to the mean of a population, and the population standard deviation is known. This sounds appropriate for a single-sample z test. However, before using the test, the assumptions must be checked. The random samples assumption is not violated, and neither is the independence of observations assumption. However, the normality assumption is violated. With random numbers, all numbers are equally likely. So, a 50 is as likely to occur as is a 0 or a 99. This assumption is robust to violations, but this distribution is not even close to normally distributed (it's flat, not normal), so the single-sample z test should not be done.

Chapter 7

- 7.01 does not
- 7.03 sampling error
- 7.05 is
- 7.07 independent
- 7.09 robust
- 7.11 two
- 7.13 does
- 7.15 \neq
- 7.17 increases
- 7.19 tails; I; sample size (or degrees of freedom)
- 7.21 one
- 7.23 2.5
- 7.25 N
- 7.27 s_M
- 7.29 facts
- 7.31 effect size; confidence interval
- 7.33 alternative hypothesis
- 7.35 sample mean; population mean
- 7.37 $p < .05$
- 7.39 not enough evidence
- 7.41 independent variable; dependent variable



B-8 Appendix B

- 7.43 no
 7.45 decreases
 7.47 II
 7.49 outcome variable; explanatory variable
 7.51 95; population
 7.53 effect size
 7.55 small
 7.57 Single-sample t test
 7.59 (a) The random sample assumption is violated. The independence assumption is not violated. The normality assumption is not violated. It is OK to proceed with the hypothesis test since the only assumption violated is a robust one.
 (b) With a convenience sample, it is doubtful that the sample is representative of the population at the college. Results can't be generalized to all students at the college.
- 7.61 $H_0: \mu_{\text{Nuns}} = 25$
 $H_1: \mu_{\text{Nuns}} \neq 25$
- 7.63 $H_0: \mu_{\text{NewTreatment}} \geq 6.30$
 $H_1: \mu_{\text{NewTreatment}} < 6.30$
- 7.65 $t_{cv} = 2.120$
- 
- 7.67 If $t \leq -2.000$ or if $t \geq 2.000$, reject H_0 . If $-2.000 < t < 2.000$, fail to reject H_0 .
- 7.69 $s_M = \frac{s}{\sqrt{N}} = \frac{12}{\sqrt{23}} = 2.50$
- 7.71 $t = \frac{M - \mu}{s_M} = \frac{10 - 12}{1.25} = -1.60$
- 7.73 $s_M = \frac{s}{\sqrt{N}} = \frac{1}{\sqrt{18}} = 0.24$
- $t = \frac{M - \mu}{s_M} = \frac{12 - 10}{0.24} = 8.33$
- 7.75 Rejected
 7.77 Not rejected
 7.79 $t(14) = 2.15, p < .05$
 7.81 $t(68) = 1.99, p > .05$
- 7.83 $d = \frac{M - \mu}{s} = \frac{90 - 100}{15} = -0.67$
- 7.85 $r^2 = \frac{t^2}{t^2 + df} \times 100 = \frac{3.45^2}{3.45^2 + 16} \times 100 = 42.66\%$

- 7.87 $95\% \text{ CI}_{\mu_{\text{Diff}}} = (M - \mu) \pm (t_{cv} \times s_M)$
 $= (45 - 50) \pm (2.093 \times 2.24)$
 $= [-9.69, -0.31]$
- 7.89 A nurse practitioner found that people who are heavy salt users have a statistically significantly higher blood pressure ($M = 138$) than do people in the general population [$\mu = 120$; $t(23) = 5.50, p < .05$]. The 18-point increase in blood pressure is a large increase. This study suggests that heavy salt consumption puts one at risk for hypertension. In the larger population from which this sample came, the mean difference might be as small as 11 points or as large as 25 points. Even an 11-point increase in blood pressure has meaningful negative consequences for health. To increase confidence that salt consumption is associated with an increase in blood pressure, it would be a good idea to replicate this study. When doing so, a larger sample would allow the researcher to get a better estimate of the size of the effect in the population.

7.91 Step 1 Pick a Test

Comparing a sample mean to a population mean when the population standard deviation is not known calls for a single-sample t test.

Step 2 Check the Assumptions

(1) The random sample assumption is violated, but the single-sample t test is robust to violations. (2) There is no evidence of violation of the independence of observations. (3) For the normality assumption, it seems plausible that hours spent in online social media is normally distributed.

Step 3 List the Hypotheses

The psychologist believes that valedictorians spend less time online in social media, so she has directional hypotheses.

$$H_0: \mu_{\text{Valedictorians}} \geq 18.68$$

$$H_1: \mu_{\text{Valedictorians}} < 18.68$$

Step 4 Set the Decision Rule

Doing a one-tailed test; using the standard alpha level, $.05; df = 31 - 1 = 30$. $t_{cv} = -1.697$

- Reject the null hypothesis if $t \leq -1.697$.
- Fail to reject the null hypothesis if $t > -1.697$.

Step 5 Calculate the Test Statistic

$$s_M = \frac{s}{\sqrt{N}} = \frac{6.80}{\sqrt{31}} = 1.22$$

$$t = \frac{M - \mu}{s_M} = \frac{16.24 - 18.68}{1.22} = -2.00$$

Step 6 Interpret the Results(1) *Was the null hypothesis rejected?*

As $-2.00 \leq -1.697$, the null hypothesis is rejected, and the difference is statistically significant. Valedictorians spend significantly less time in online social media than the average high school student. $t(30) = -2.00$, $p < .05$ (one-tailed).

(2) *How big is the effect?* The d value of -0.36 , calculated below, is a weak-to-medium effect:

$$d = \frac{M - \mu}{s} = \frac{16.24 - 18.68}{6.80} = -0.36$$

$$r^2 = \frac{t^2}{t^2 + df} \times 100 = \frac{2.00^2}{2.00^2 + 30} \times 100 = 12\%$$

(3) *How wide is the confidence interval?*

$$\begin{aligned} 95\% \text{ CI}_{\mu_{\text{Diff}}} &= (M - \mu) \pm (t_{cv} \times s_M) \\ &= (16.24 - 18.68) \pm (2.042 \times 1.22) \\ &= [-4.93, 0.05] \end{aligned}$$

The confidence interval, calculated above, is roughly from -5 to 0 . Thus, it is not clear how strong the effect is. If it is five hours a week less, that seems like a lot of time. Zero, obviously, is no difference.

Note that this confidence interval captures zero, even though the null hypothesis was rejected. How can that be? It is possible because the test was a one-tailed test, but confidence intervals are always two-tailed. When a researcher does a one-tailed test, he or she can change the degree of confidence to fit the test. With this one one-tailed test, all 5% of the rare zone was in one end. If the researcher put 5% in both ends, for a total of 10%, 90% would be left in the middle for a 90% confidence interval. To calculate the 90% confidence interval instead of the 95%, substitute the critical value of t found in Step 4, 1.697, into the confidence interval equation. This gives a confidence interval that doesn't capture 0,

as it ranges from -4.51 to -0.37 . Here are the calculations:

$$\begin{aligned} 90\% \text{ CI}_{\mu_{\text{Diff}}} &= (M - \mu) \pm (t_{cv} \times s_M) \\ &= (16.24 - 18.68) \pm (1.697 \times 1.22) \\ &= [-4.51, -0.37] \end{aligned}$$

Putting it all together: “A psychologist investigating time management in high school students compared the average amount of time spent by valedictorians per week in online social media ($M = 16.24$) to the known average for the population of high school students in general ($\mu = 18.68$). The difference was statistically significant [$t(30) = -2.00$, $p < .05$ (one-tailed)], indicating that students who are doing very well in school, like valedictorians, spend less time with social media and, presumably, more time on school work.

“Unfortunately, the sample size, 31, in this study was relatively small, making the confidence interval relatively wide. To resolve this, I recommend replicating the study with a larger sample of valedictorians, chosen in such a way that they represent the United States.”

- 7.93 The less variability there is, (a) the smaller the standard error of the mean, (b) the larger the t value, (c) the bigger the effect size, (d) and the narrower the confidence interval. Thus, when there is less variability it is easier to reject the null hypothesis, find a large effect, and have a more precise estimate of the population value.
- 7.95 The distance from the sample mean to the population mean has no impact on the standard error of the mean or the width of the confidence interval, but it does have an impact on the t value and the effect size. Having a greater distance between sample and population mean means there is a larger effect and that it is easier to reject the null hypothesis.
- 7.97 1.959 is the largest value of t that guarantees failure to reject the null hypothesis. (1.960 is the critical value of t for $\alpha = .05$, two-tailed, $df = \infty$.)
- 7.99 For the 90% confidence interval, use Equation 7.4, but instead of using t_{cv} , two-tailed, $\alpha = .05$, substitute the critical value of t , two-tailed, $\alpha = .10$. For the 99% confidence interval, substitute the critical value of t , two-tailed,

B-10 Appendix B

$\alpha = .05$. Here are the two equations for $N = 21$ (i.e., $df = 20$) and $s_M = 1$:

$$90\% \text{ CI}_{\mu_{\text{Diff}}} = (M - \mu) \pm (t_{cv} \times s_M) \\ = (M - \mu) \pm (1.725 \times 1) = (M - \mu) \pm 1.725$$

$$99\% \text{ CI}_{\mu_{\text{Diff}}} = (M - \mu) \pm (t_{cv} \times s_M) \\ = (M - \mu) \pm (2.845 \times 1) = (M - \mu) \pm 2.845$$

Chapter 8

8.01 mean

8.03 sample; population

8.05 independent; paired

8.07 paired

8.09 independent; mean

8.11 random samples

8.13 normally distributed

8.15 non

8.17 $\mu_1 \neq \mu_2$

8.19 rare; common; t

8.21 2

8.23 \geq

8.25 The pooled variance

8.27 means

8.29 rejected

8.31 5

8.33 $p < .05$; $p > .05$

8.35 0

8.37 outcome; explanatory

8.39 population means

8.41 small

8.43 t -statistic

8.45 a paired-samples t test (the student and parent are from the same family)

8.47 an independent-samples t test (no subject belongs to both groups)

8.49 The random samples assumption is violated. The independence assumption is not violated. The normality of distribution assumption is not violated. The homogeneity of variance assumption is violated, so much so that it is not OK to continue with the test.

8.51 (a) $H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

(b) The null hypothesis says that the two population means are the same, that is, there is no difference in effectiveness between the two treatments. Any difference between sample means is due to sampling error. The alternative hypothesis says that the two population means

are different because the two populations are differentially effective. The difference between sample means is too large to be due to sampling error.

- 8.53 (a) $H_0: \mu_{\text{Fluoride}} \geq \mu_{\text{Nonfluoride}}$
 $H_1: \mu_{\text{Fluoride}} < \mu_{\text{Nonfluoride}}$

(b) The null hypothesis says that the population mean for the number of cavities for the fluoride group will be greater than or the same as that for the nonfluoride group because fluoride has no effect, or a negative effect, on cavities. The alternative hypothesis states that the population mean for the number of cavities for the fluoride group will be less than that for the nonfluoride group because fluoride reduces cavities.

- 8.55 $df = n_1 + n_2 - 2 = 223 + 252 - 2 = 473$. Use $df = 450$. $t_{cv} = 1.965$.

- 8.57 $df = n_1 + n_2 - 2 = 46 + 46 - 2 = 90$. $t_{cv} = 2.632$.

- 8.59 (a) If $t \leq -2.086$ or if $t \geq 2.086$, then reject H_0 .
(b) If $-2.086 < t < 2.086$, then fail to reject H_0 .

- 8.61 $df = n_1 + n_2 - 2 = 12 + 13 - 2 = 23$

$$s_{\text{pooled}}^2 = \frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{df} \\ = \frac{7.4^2(12 - 1) + 8.2^2(13 - 1)}{23} \\ = 61.27$$

- 8.63 $N = n_1 + n_2 = 45 + 58 = 103$

$$s_{M_1-M_2} = \sqrt{s_{\text{pooled}}^2 \left[\frac{N}{n_1 \times n_2} \right]} \\ = \sqrt{5.63 \left[\frac{103}{45 \times 58} \right]} = 0.47$$

- 8.65 $N = n_1 + n_2 = 45 + 58 = 103$

$$df = N - 2 = 101$$

$$s_{M_1-M_2} = \sqrt{\left[\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{df} \right] \left[\frac{N}{n_1 \times n_2} \right]} \\ = \sqrt{\frac{5.98^2(45 - 1) + 7.83^2(58 - 1)}{101} \left[\frac{103}{45 \times 58} \right]} \\ = 1.41$$

$$8.67 t = \frac{M_1 - M_2}{s_{M_1-M_2}} = \frac{68 - 57}{2.34} = 4.70$$

8.69 $N = n_1 + n_2 = 72 + 60 = 132$

$df = N - 2 = 130$

$$\begin{aligned}s_{M_1-M_2} &= \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{df}} \left[\frac{N}{n_1 \times n_2} \right] \\&= \sqrt{\frac{4.6^2(72 - 1) + 3.3^2(60 - 1)}{130}} \left[\frac{132}{72 \times 60} \right] \\&= 0.71\end{aligned}$$

- 8.71 (a) H_0 rejected
 (b) Statistically significant
 (c) μ_2 is probably bigger than μ_1 .

8.73 $t(21) = 2.07, p > .05$

8.75 $t(8) = 2.31, p < .05$

8.77 (a) $d = \frac{M_1 - M_2}{\sqrt{s_{\text{pooled}}^2}} = \frac{17 - 12}{\sqrt{4.00}} = 2.50$.

(b) Very large effect

8.79 $df = N - 2 = 73 - 2 = 71$

$$\begin{aligned}r^2 &= \frac{t^2}{t^2 + df} \times 100 \\&= \frac{9.87^2}{9.87^2 + 71} \times 100 = 57.8\%\end{aligned}$$

8.81 (a) $95\% \text{ CI}_{\mu_{\text{Diff}}} = (M_1 - M_2) \pm (t_{cv} \times s_{M_1-M_2})$
 $= (31 - 24) \pm (2.045 \times 2.88)$
 $= [1.11, 12.89]$

(b) Reject H_0 .

- 8.83 (Note: One does not need to use everything that is calculated in an interpretation. Interpretation has a subjective element. What one person chooses to emphasize may differ from what another does.)

An education researcher investigated whether the color used to grade a test has an impact on first graders' self-esteem. The first graders took a third-grade math test and had 25% of their answers marked wrong, half of the students with a red pen and half of the students with pencil. They then took a self-esteem test. The "pencil" students had higher self-esteem ($M = 29.00$) than the "red ink" students ($M = 23.00$). This difference was statistically significant [$t(25) = 3.02, p < .05$]. The effect is a strong one and suggests that the color used to grade a test has an impact on self-esteem, with an effect of between 1.90 and 10.1 points

with 95% confidence. Unfortunately, because there was no control group, it is unclear if grading in pencil raises self-esteem or using red ink lowers it. Future research should use larger sample sizes, include such a control group, and utilize random assignment of the students to the treatment conditions. Until that research is done, it seems advisable for teachers to use pencils, not red ink, to grade tests and papers.

8.85 **Step 1** Pick a Test.

Comparing the means of two independent samples calls for an independent-samples t test.

Step 2 Check the Assumptions.

- Random samples: violated, but robust to violations. Plus, sample is representative of Americans, which is the objective.
- Independence of observations: not violated, each participant is in the study only once.
- Normality: Cholesterol is a physical characteristic; reasonable to assume it is normally distributed. Plus, robust assumption, especially when $N > 50$.
- Homogeneity of variance: the 2 standard deviations are about equal—not violated.

Step 3 List the Hypotheses.

$H_0: \mu_{\text{American}} = \mu_{\text{Mediterranean}}$

$H_1: \mu_{\text{American}} \neq \mu_{\text{Mediterranean}}$

Step 4 Set the Decision Rule.

Set $\alpha = .05$, two-tailed, $df = 70: t_{cv} = 1.994$.

- If $t \leq -1.994$ or if $t \geq 1.994$, reject H_0 .
- If $-1.994 < t < 1.994$, fail to reject H_0 .

Step 5 Calculate the Test Statistic.

$N = n_1 + n_2 = 36 + 36 = 72$

$df = N - 2 = 70$

$$\begin{aligned}s_{M_1-M_2} &= \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{df}} \left[\frac{N}{n_1 \times n_2} \right] \\&= \sqrt{\frac{24^2(36 - 1) + 26^2(36 - 1)}{70}} \left[\frac{72}{36 \times 36} \right] \\&= 5.90 \\t &= \frac{M_1 - M_2}{s_{M_1-M_2}} = \frac{230 - 190}{5.90} = 6.78\end{aligned}$$

B-12 Appendix B

Step 6 Interpret the Results.

Was the null hypothesis rejected?

$6.78 \geq 1.994$, so H_0 is rejected. Results are statistically significant. Cholesterol levels of those who follow a Mediterranean diet are lower than for those who follow an American diet.

How big is the effect?

$$d = \frac{M_1 - M_2}{\sqrt{s_{\text{Pooled}}^2}} = \frac{190 - 230}{\sqrt{626.0000}} = -1.60$$

How wide is the confidence interval?

$$\begin{aligned} 95\% \text{CI}_{\mu_{\text{Diff}}} &= (M_1 - M_2) \pm (t_{cv} \times s_{M_1 - M_2}) \\ &= (190 - 230) \pm (1.994 \times 5.90) \\ &= [-51.76, -28.24] \end{aligned}$$

Write a four-point interpretation:

This study used a representative sample of Americans to compare the cholesterol levels of those who ate a Mediterranean diet to those who ate an American diet. The mean cholesterol level of those who followed the Mediterranean diet was 40 points lower with a mean of 190 compared to a mean cholesterol level of 230 for those following an American diet. This difference was statistically significant [$t(70) = 6.78, p < .05$] and a 40-point difference is a very meaningful difference. The Mediterranean diet is associated with lower cholesterol, a level sufficiently lower that it means the difference between high cholesterol and normal cholesterol.

This study, because it did not use random assignment, cannot give conclusions about cause and effect. The report suggests that the Mediterranean diet is associated with reduced cholesterol, but doesn't prove such. Based on this study, it would be reasonable to proceed to a study in which participants are randomly assigned to different diets.

- 8.87 d. There is probably a difference between the two population means.
- 8.89 (a) The first test is an independent-samples t test and the second test is a paired-samples t test. (b) The first test answers the question, "Does one store carry more expensive merchandise than the other?" The second test answers the question, "For the same items,

is there a difference in price between the two stores?"

- 8.91 If the confidence interval is a 99% confidence interval but the alpha level for the test is set at .05, which is equivalent to a 95% confidence interval, the confidence interval might capture zero though the null hypothesis was rejected. Another way this could occur would be with a one-tailed test as confidence intervals are two-tailed. A one-tailed test with $\alpha = .05$ has a rare zone that encompasses the extreme 5%, while a 95% confidence interval reaches out into the tail an additional 2.5%.
- 8.93 (a) Yes
(b) There was evidence of an effect.
(c) Small.
(d) Because the confidence interval is so narrow, we are fairly certain that effect was small.
(e) Given how large the sample size was, it seems unlikely that a nonrepresentative sample was obtained. Not very worried about an erroneous conclusion being drawn.

Chapter 9

- 9.01 independent; paired-samples t test
9.03 confounding
9.05 individual differences
9.07 higher
9.09 sampling error
9.11 within; between
9.13 populations; samples
9.15 negative; small
9.17 one
9.19 s_{M_D}
9.21 does not
9.23 $>$
9.25 sample
9.27 individual differences; dependent variable
9.29 confidence interval for the difference between population means
9.31 small
9.33 paired-samples t test (each subject rated both types of coffee)
9.35 single-sample t test (the mean is being compared to the known USDA standard)
9.37 It is not OK to proceed with the paired-samples t test.

- The random samples assumption is violated, but this assumption is robust to violation.
- The nonrobust independence of observations assumption is violated, so it is not OK to proceed with the test. Why was it violated? All students are in the same classes, so presumably they were assigned the same homework, which would influence the online activity of all students.
- For the normality assumption, we can assume that the difference scores are normally distributed.

9.39 $H_0: \mu_{\text{School}} = \mu_{\text{Social}}$

$H_1: \mu_{\text{School}} \neq \mu_{\text{Social}}$

9.41 -5.00; 4.00; -2.00; 1.00; -1.00 or: 5.00; -4.00; 2.00; -1.00; 1.00

9.43 9

9.45 2.101

9.47 1.984

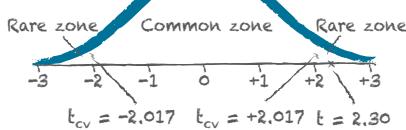
$$9.49 s_{M_D} = \frac{s_D}{\sqrt{N}} = \frac{2.57}{\sqrt{17}} = 0.62$$

$$9.51 t = \frac{M_1 - M_2}{s_{M_D}} = \frac{18 - 14}{5.34} = 0.75$$

$$9.53 s_{M_D} = \frac{s_D}{\sqrt{N}} = \frac{7.00}{\sqrt{5}} = 3.13$$

$$t = \frac{M_1 - M_2}{s_{M_D}} = \frac{28 - 25}{3.13} = 0.96$$

9.55 (a)



(b) The null hypothesis is rejected.

9.57 $t(4) = 3.21, p < .05$

9.59 $t(68) = 1.99, p > .05$

9.61 μ_2 is probably greater than μ_1 .

9.63 There is not enough evidence to say a difference exists in the means of the two populations.

$$9.65 95\% \text{CI} \mu_{\text{Diff}} = (M_1 - M_2) \pm (t_{cv} \times s_{M_D})$$

$$= (108 - 100) \pm (2.052 \times 5.00)$$

$$= [-2.26, 18.26]$$

9.67 Rejected

9.69 A sleep therapist studied whether an herbal tea that was advertised as a sleep aid had

any impact on sleep onset for people with insomnia. Using 23 pairs of insomniacs matched on a number of potential confounding variables, he found that those who used the tea went to sleep in 19.70 minutes and those who didn't took 21.20 minutes. This 1.5-minute difference was not statistically significant [$t(22) = 1.32, p > .05$], meaning that the study did not provide sufficient evidence that the tea had any impact on sleep onset.

Nonetheless, the study was suggestive that the tea might have some small effect, so it would be reasonable to replicate the study with a larger sample size. It would also be advisable to make the study double blind so that neither experimenter nor participant would know who was receiving the supposed sleep aid.

9.71 Step 1 Pick a Test

There are two samples, feet receiving standard treatment and feet receiving experimental treatment, the samples are dependent, and the dependent variable is measured at the ratio level. This calls for a paired-samples t test.

Step 2 Check the Assumptions

- Random samples: Violated, but robust to violations.
- Independence of observations: Not violated.
- Normality: Willing to assume not violated.

Step 3 List the Hypotheses

- $H_0: \mu_E = \mu_C$
- $H_1: \mu_E \neq \mu_C$

Step 4 Set the Decision Rule

$$df = N - 1 = 30 - 1 = 29$$

$\alpha = .05$, two-tailed

$$t_{cv} = \pm 2.045$$

Decision rule:

- If $t \leq -2.045$ or if $t \geq 2.045$, reject H_0 .
- If $-2.045 < t < 2.045$, fail to reject H_0 .

Step 5 Calculate the Test Statistic

$$s_{M_D} = \frac{s_D}{\sqrt{N}} = \frac{8.65}{\sqrt{30}} = 1.58$$

$$t = \frac{M_1 - M_2}{s_{M_D}} = \frac{88 - 72}{1.58} = 10.13$$

Step 6 Interpret the Results

Was the null hypothesis rejected?

B-14 Appendix B

As $10.13 \geq 2.045$, the null hypothesis is rejected. The difference is statistically significant and is in favor of the new treatment working better. In APA format, the results are $t(29) = 10.13, p < .05$.

How wide is the confidence interval? How big is the effect?

$$95\% \text{CI}_{\mu_{\text{Diff}}} = (M_1 - M_2) \pm (t_{cv} \times s_{M_D}) \\ = (88 - 72) \pm (2.045 \times 1.58) = [12.77, 19.23]$$

Write a four-point interpretation:

A dermatologist evaluated a new treatment for athlete's foot by taking 30 people with athlete's foot on both feet and randomly assigning which foot received standard treatment and which received the experimental treatment. He measured outcome as the percentage of reduction of symptoms and found a statistically significant improvement in outcome with the new treatment [$t(29) = 10.13, p < .05$]. The new treatment led to an 88% reduction in symptoms vs. a 72% reduction in symptoms with the old, standard treatment. The new treatment seems to work substantially better than the standard treatment. But, to increase confidence in robustness of this finding, the study should be replicated.

- 9.73 e (the width does not depend on the differences in the mean)
9.75 d (all else being equal, a larger standard error of the mean gives a wider confidence interval)
9.77 $M_1 - M_2 = -2.00$ (that's the midpoint of the confidence interval)

$$t = \frac{M_1 - M_2}{s_{M_D}} = \frac{-2.00}{4.00} = -.50$$

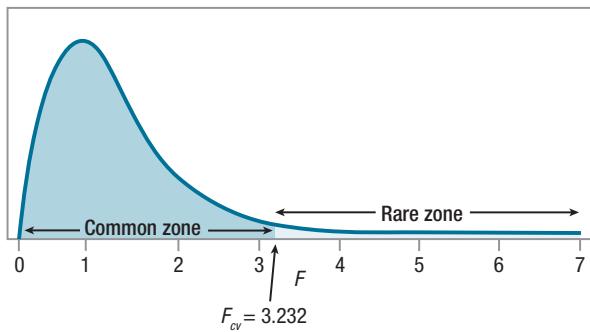
Chapter 10

- 10.01 two
10.03 ANOVA
10.05 way; factor
10.07 Type I
10.09 post-hoc
10.11 variability
10.13 individual differences
10.15 Between-groups
10.17 individual differences
10.19 large
10.21 Fratio
10.23 one
10.25 t
10.27 an independent-samples t test
10.29 cannot

- 10.31 homogeneity of variance
10.33 non
10.35 rejected
10.37 between
10.39 $df_{\text{Between}}; df_{\text{Within}}$
10.41 sum of squares
10.43 SS_{Total}
10.45 treatment effect
10.47 ANOVA summary table
10.49 mean square
10.51 effect
10.53 the alternative hypothesis
10.55 $>$
10.57 between-groups
10.59 100; 0
10.61 statistically significant
10.63 Type II; Type I (order does matter)
10.65 statistically significant
10.67 smallest
10.69 (e) This is a between-subjects, one-way ANOVA because the means of three or more independent samples are being compared.
10.71 (f) This is none of the tests learned about so far. There are three levels of one independent variable being compared on an interval-level dependent variable, but the samples are dependent. This type of test, a repeated-measures ANOVA, is covered in the next chapter.
10.73 The random samples assumption is violated. The independence of observations assumption is not violated. There is no reason to think the homogeneity of variance assumption is violated, but it would be nice to know the standard deviation for each group. Because the only violated assumption is robust, we can proceed with the between-subjects, one-way ANOVA.
10.75 $H_0: \mu_1 = \mu_2 = \mu_3$.
 $H_1:$ At least one population mean differs from at least one other population mean.
10.77 (a) $df_{\text{Total}} = N - 1 = 48 - 1 = 47$
(b) $df_{\text{Between}} = k - 1 = 4 - 1 = 3$
(c) $df_{\text{Within}} = N - k = 48 - 4 = 44$
10.79 (a) $df_{\text{Total}} = N - 1 = 60 - 1 = 59$
(b) $df_{\text{Between}} = k - 1 = 4 - 1 = 3$
(c) $df_{\text{Within}} = N - k = 60 - 4 = 56$
10.81 $F_{cv} = 5.123$
10.83 $df_{\text{Between}} = 4 - 1 = 3; df_{\text{Within}} = 50 - 4 = 46$
 $F_{cv} = 2.807$

10.85 $F_{cv} = 3.232$

If $F \geq 3.232$, reject H_0 . If $F < 3.232$, fail to reject H_0 .



10.87

	Group 1		Group 2		Group 3		
	X	X ²	X	X ²	X	X ²	
108	11,664	100	10,000	99	9,801		
102	10,404	105	11,025	95	9,025		Grand
				91	8,281	X	X ²
Sum	210	22,068	205	21,025	285	27,1067	700
n	2		2		3		7
							70,200

10.89 $SS_{\text{Total}} = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$

$$= 75953 - \frac{821^2}{9} = 1059.556$$

$$SS_{\text{Between}} = \sum \left(\frac{(\Sigma X_{\text{Group}})^2}{n_{\text{Group}}} \right) - \frac{(\Sigma X)^2}{N}$$

$$= \left(\frac{216^2}{2} + \frac{276^2}{3} + \frac{329^2}{4} \right) - \frac{821^2}{9} = 886.81$$

$$SS_{\text{Within}} = \sum \left(\Sigma X_{\text{Group}}^2 - \frac{(\Sigma X_{\text{Group}})^2}{n_{\text{Group}}} \right)$$

$$= \left(23360 - \frac{216^2}{2} \right) + \left(25448 - \frac{276^2}{3} \right)$$

$$= \left(27145 - \frac{329^2}{4} \right) = 172.75$$

10.91 $SS_{\text{Within}} = SS_{\text{Total}} - SS_{\text{Between}} = 98.75 - 40.33 = 58.42$

10.93 $MS_{\text{Between}} = \frac{SS_{\text{Between}}}{df_{\text{Between}}} = \frac{2,378.99}{3} = 793.00$

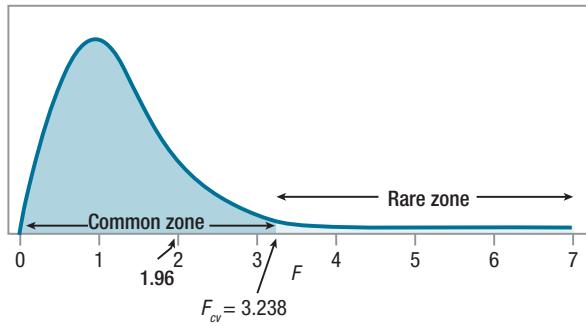
10.95 $MS_{\text{Within}} = \frac{SS_{\text{Within}}}{df_{\text{Within}}} = \frac{78.95}{32} = 2.47$

10.97 $F = \frac{MS_{\text{Between}}}{MS_{\text{Within}}} = \frac{38.88}{17.44} = 2.23$

10.99

Source of Variability	Sum of Squares	Degrees of Freedom	Mean Square	F ratio
Between groups	172.80	2	86.40	0.16
Within groups	6,410.80	12	534.23	
Total	6,583.60	14		

10.101 $F = 1.96$ falls in the common zone, so fail to reject the null hypothesis.



10.103 $7.64 \geq 3.467$, so reject H_0 .

10.105 $F(3, 17) = 5.34, p < .05$

10.107 $F(3, 44) = 3.28, p < .05$

10.109 The results are statistically significant, so conclude that it is probable that at least one population has a mean different from at least one other population mean.

10.111 $r^2 = \frac{SS_{\text{Between}}}{SS_{\text{Total}}} \times 100 = \frac{128.86}{413.67} \times 100 = 31.15\%$

10.113 $r^2 = \frac{SS_{\text{Between}}}{SS_{\text{Total}}} \times 100 = \frac{59.98}{328.58} \times 100 = 18.25\%$

10.115 Yes, complete a post-hoc test because the results were statistically significant.

10.117 $q = 5.19$

10.119 $HSD = q \sqrt{\frac{MS_{\text{Within}}}{n}} = 3.55 \sqrt{\frac{10.44}{8}} = 4.06$

10.121 $13.09 - 8.89 = 4.20$. $4.20 < 4.37$, so the difference is not statistically significant. There's not enough evidence to conclude that these two populations have a mean difference.

10.123 • M_1 vs. M_2 : $67.86 - 53.56 = 14.30$, which is greater than 8.30, so it is a statistically significant difference. The mean of population 1 is probably greater than the mean of population 2.

B-16 Appendix B

- M_1 vs. M_3 : $67.86 - 61.55 = 6.31$, which is less than 8.30, so it is not a statistically significant difference. There is not enough evidence to conclude that these two population means are different.
 - M_2 vs. M_3 : $61.55 - 53.56 = 7.99$, which is less than 8.30, so it is not a statistically significant difference. There is not enough evidence to conclude that the two population means are different.
- 10.125 "An addictions researcher obtained samples of alcoholics, smokers, and heroin addicts who were in treatment for the second time. From each participant she found out how long, in months, it took before relapse occurred after the first treatment. The alcoholics relapsed, on average, in 4.63 months, the smokers in 4.91 months, and the heroin addicts in 5.17 months. Using a between-subjects, one-way ANOVA, she found no statistically significant effect for the type of addiction on the mean time to relapse [$F(2, 22) = 0.07, p > .05$]. These data suggest, for those who fail in treatment, that there is no difference in the addictiveness of these three substances. Future research should use a larger number of participants and follow them prospectively after treatment ends."

Step 1 Pick a Test

The test compares means of three independent samples, so use a between-subjects, one-way ANOVA.

Step 2 Check the Assumptions

- The random samples assumption is violated, but this assumption is robust. Need to be careful about population to which the results are generalized.
- Independence of observations is not violated.
- Normality is not violated. It's reasonable to assume that satisfaction scores are normally distributed.
- Homogeneity of variance is not violated. All standard deviations are equal.

Only the random samples assumption is violated, but it is robust, so it is OK to

proceed with a between-subjects, one-way ANOVA.

Step 3 List the Hypotheses

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_1 : At least one population mean is different from at least one other.

Step 4 Set the Decision Rule

df_{Between} , numerator $df = k - 1 = 3 - 1 = 2$

df_{Within} , denominator $df = N - k = 12 - 3 = 9$

$df_{\text{Total}} = N - 1 = 12 - 1 = 11$

Set alpha at .05.

$$F_c = 4.256$$

Decision rule: If $F \geq 4.256$, reject H_0 ; if $F < 4.256$, fail to reject H_0 .

Step 5 Calculate the Test Statistic

$$SS_{\text{Total}} = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$$

$$= 78350 - \frac{960^2}{12} = 1550$$

$$SS_{\text{Between}} = \Sigma \left(\frac{(\Sigma X_{\text{Group}})^2}{n_{\text{Group}}} \right) - \frac{(\Sigma X)^2}{N}$$
$$= \left(\frac{260^2}{4} + \frac{360^2}{4} + \frac{340^2}{4} \right) - \frac{960^2}{12} = 1400$$

$$SS_{\text{Within}} = \Sigma \left(\Sigma X_{\text{Group}}^2 - \frac{(\Sigma X_{\text{Group}})^2}{n_{\text{Group}}} \right)$$
$$= \left(16950 - \frac{260^2}{4} \right) + \left(32450 - \frac{360^2}{4} \right)$$
$$= \left(28950 - \frac{340^2}{4} \right) = 150$$

Source of Variability	Sum of Squares	Degrees of Freedom	Mean Square	F ratio
Between groups	1,400.00	2	700.00	41.99
Within groups	150.00	9	16.67	
Total	1,550.00	11		

Step 6 Interpret the Results

- Was the null hypothesis rejected?

- Yes, $41.99 \geq 4.256$

- APA format: $F(2, 9) = 41.99, p > .05$

- Conclude that at least one population mean differs from at least one other mean.

- How big is the effect?

- Big!

$$r^2 = \frac{SS_{\text{Between}}}{SS_{\text{Total}}} \times 100 = \frac{1,400.00}{1,550.00} \times 100 = 90.32\%$$

- Where is the effect?

- $q = 3.95$

$$HSD = q \sqrt{\frac{MS_{\text{Within}}}{n}} = 3.95 \sqrt{\frac{16.67}{4}} = 8.06$$

- The premium brand in the premium bottle is statistically significantly more satisfying than the store brand in the store bottle.
- The store brand in the premium bottle is statistically significantly more satisfying than the store brand in the store bottle.
- There is not enough evidence to conclude that there is a difference in satisfaction between the premium brand and the store brand when both are in the premium bottle.
- Putting it all together: A consumer researcher investigated the effect of labeling on satisfaction with a product. Comparing a store brand shampoo in a store brand bottle, a premium brand shampoo in a premium brand bottle, and a store brand shampoo in a premium bottle, the mean satisfaction scores were, respectively, 65.00, 90.00, and 85.00. Using a between-subjects, one-way ANOVA, she found a strong and statistically significant effect of labeling [$F(2, 9) = 41.99$, $p < .05$]. There was no statistically significant difference between the store brand and the premium brand when both were in the premium bottle, but both were found to be statistically significantly more satisfying than the store shampoo in a store brand bottle. In rating shampoos, consumers seem to be more influenced by their perception of the shampoo from the bottle's label than they are by the shampoo in the bottle. It would be interesting, in future research, to use the same participants in all three conditions to see if the effect exists when they are directly comparing all three samples.

10.129 c. SS_{Total} must be largest; means squares are always no larger than corresponding sum of squares.

10.131 Could be true

10.133 False

10.135 False

10.137 Could be true

Chapter 11

11.01 independent-samples *t* test

11.03 Repeated-measures

11.05 rows; columns

11.07 powerful; greater

11.09 denominator

11.11 a repeated-measures ANOVA; a paired-samples *t* test

11.13 robust

11.15 normally distributed; population

11.17 is not

11.19 greater than or equal to

11.21 was

11.23 was

11.25 treatment; residual

11.27 n ; k ; N

11.29 df_{Total}

11.31 subjects

11.33 rare

11.35 variability

11.37 sum of squares; degrees of freedom

11.39 $MS_{\text{Treatment}}$; MS_{Residual}

11.41 decision rule

11.43 was

11.45 should

11.47 was not

11.49 eta squared

11.51 r squared

11.53 medium

11.55 rejected

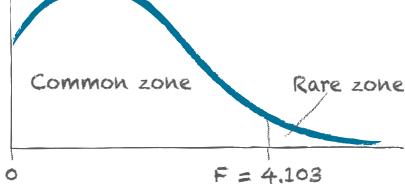
11.57 not statistically

11.59 Between-subjects, one-way ANOVA (the means of three or more independent samples are being compared)

11.61 Independent-samples *t* test (the means of two independent samples are being compared)

11.63 (a) The random samples assumption is not violated. The independence of observations assumption is violated as participants met together

B-18 Appendix B

- in groups. The normality assumption is not violated.
- (b) Should not proceed with the one-way, repeated-measures ANOVA because the independence of observations assumption is violated.
- 11.65 H_0 : All population means are equal.
 H_1 : At least one population mean is different from at least one other.
- 11.67 $df_{\text{Subjects}} = (n - 1) = 8 - 1 = 7$
 $df_{\text{Treatment}} = (k - 1) = 3 - 1 = 2$
 $df_{\text{Residual}} = (n - 1)(k - 1) = (8 - 1)(3 - 1) = 14$
 $df_{\text{Total}} = (N - 1) = (24 - 1) = 23$
- 11.69 Remember, $N = n \times k$. So, $N = 15 \times 4 = 60$
 $df_{\text{Subjects}} = (n - 1) = 15 - 1 = 14$
 $df_{\text{Treatment}} = (k - 1) = 4 - 1 = 3$
 $df_{\text{Residual}} = (n - 1)(k - 1) = (15 - 1)(4 - 1) = 42$
 $df_{\text{Total}} = (N - 1) = (60 - 1) = 59$
- 11.71 (a) $F_{cv} = 4.103$
(b)

- 11.73 If $F \geq 3.287$, reject H_0 .
If $F < 3.287$, fail to reject H_0 .
- 11.75 $SS_{\text{Total}} = 301.48$
- $$MS_{\text{Treatment}} = \frac{SS_{\text{Treatment}}}{df_{\text{Treatment}}} = \frac{48.48}{2} = 24.24$$
- $$MS_{\text{Residual}} = \frac{SS_{\text{Residual}}}{df_{\text{Residual}}} = \frac{115.50}{22} = 5.25$$
- $$F = \frac{MS_{\text{Treatment}}}{MS_{\text{Residual}}} = \frac{24.24}{5.25} = 4.62$$
- | Source of Variability | Sum of Squares | Degrees of Freedom | Mean Square | F ratio |
|-----------------------|----------------|--------------------|-------------|---------|
| Subjects | 137.50 | 11 | | |
| Treatment | 48.48 | 2 | 24.24 | 4.62 |
| Residual | 115.50 | 22 | 5.25 | |
| Total | 301.48 | 35 | | |
- 11.77 H_0 was rejected.
- 11.79 Based on df in summary table, $F_{cv} = 3.232$. H_0 is rejected.
- 11.81 (a) $F(5, 45) = 7.84, p < .05$
(b) Statistically significant
- 11.83 Based on df in summary table, $F_{cv} = 3.555$.
(a) $F(2, 18) = 11.24, p < .05$
(b) Statistically significant
- 11.85 The results of the ANOVA are not statistically significant. There is not enough evidence to conclude that the population means are different.
- 11.87 (a) $\eta^2 = \frac{SS_{\text{Treatment}}}{SS_{\text{Total}}} \times 100 = \frac{9.89}{86.98} \times 100 = 11.37\%$
(b) Medium
- 11.89 (a) $\eta^2 = \frac{SS_{\text{Treatment}}}{SS_{\text{Total}}} \times 100 = \frac{46.00}{336.00} \times 100 = 13.69\%$
(b) Medium
- 11.91 (a) The effect size is small, so there is just the normal level of worry about a Type II error.
(b) Do not recommend replication with a larger sample size.
- 11.93 $q = 4.79$
- 11.95 $HSD = q \sqrt{\frac{MS_{\text{Residual}}}{n}} = 2.90 \sqrt{\frac{12.27}{8}} = 3.59$
- 11.97 (a)
 - M_1 vs. $M_2 = 14.98 - 12.83 = 2.15$. $2.15 < 3.78$: Not statistically significant.
 - M_1 vs. $M_3 = 12.83 - 8.22 = 4.61$. $4.61 \geq 3.78$: Statistically significant.
 - M_2 vs. $M_3 = 14.98 - 8.22 = 6.76$. $6.76 \geq 3.78$: Statistically significant.
(b)
 - Not enough evidence to conclude there is a difference in means between population 1 and population 2.
 - Population 1 probably has a bigger mean than population 3.
 - Population 2 probably has a bigger mean than population 3.
- 11.99 "In this study, rats were reared in three different environments—a standard laboratory setting, an enriched laboratory setting, and a mimicked wild setting—to see if environment had an impact on the rats' intelligence. There was a statistically significant and strong effect of environment on mean intelligence [$F(2, 14) = 9.61, p < .05$]. Rats reared in the enriched environment earned

statistically higher scores ($M = 108.00$) on a variety of behavioral tasks than either the rats reared in a standard lab setting ($M = 95.00$) or the rats reared in an environment meant to mimic the wild ($M = 94.00$). There was no difference in the mean intelligence of these latter two types of rats. An enriched environment has a positive impact on rat intelligence and a standard laboratory environment does not have a negative impact on rat intelligence, as compared to a mimicked wild environment. This research should be replicated to make sure the effect is stable. Future research should attempt to determine what aspects of the enriched environment lead to improved performance."

11.101 Step 1 Pick a Test

Comparing the means of three dependent samples calls for a one-way, repeated-measures ANOVA.

Step 2 Check the Assumptions

- The random samples assumption is violated, but it is robust, so it is OK to proceed.
- The independence of observations assumption is not violated.
- The normality assumption is not violated. It is reasonable to assume a normal distribution for urges to drink alcohol.

Step 3 List the Hypotheses

$$H_0: \mu_1 = \mu_2 = \mu_3.$$

H_1 : At least one population mean is different from at least one other.

Step 4 Set the Decision Rule

$$df_{\text{Subjects}} = n - 1 = 10 - 1 = 9$$

$$df_{\text{Treatment}} = k - 1 = 3 - 1 = 2$$

$$df_{\text{Residual}} = (n - 1)(k - 1) = (10 - 1)(3 - 1) = 18$$

$$N = n \times k = 10 \times 3 = 30$$

$$df_{\text{Total}} = N - 1 = 30 - 1 = 29$$

Set $\alpha = .05$.

$$F_{cv} = 3.555$$

If $F \geq 3.555$, reject H_0 .

If $F < 3.555$, fail to reject H_0 .

Step 5 Calculate the Test Statistic

$$SS_{\text{Total}} = 646.67$$

$$MS_{\text{Treatment}} = \frac{SS_{\text{Treatment}}}{df_{\text{Treatment}}} = \frac{26.67}{2} = 13.34$$

$$MS_{\text{Residual}} = \frac{SS_{\text{Residual}}}{df_{\text{Residual}}} = \frac{88.00}{18} = 4.89$$

$$F = \frac{MS_{\text{Treatment}}}{MS_{\text{Residual}}} = \frac{13.34}{4.89} = 2.73$$

Source of Variability	Sum of Squares	Degrees of Freedom	Mean Square	F ratio
Subjects	532.00	9		
Treatment	26.67	2	13.34	2.73
Residual	88.00	18	4.89	
Total	646.67	29		

Step 6 Interpret the Results

$2.73 < 3.555$, so fail to reject the null hypothesis.

In APA format: $F(2, 18) = 2.73, p > .05$.

$$\eta^2 = \frac{SS_{\text{Treatment}}}{SS_{\text{Total}}} \times 100 = \frac{26.67}{646.67} \times 100 = 4\%$$

This is more than a small effect, so some concern about Type II error does exist. A study was conducted to see if different types of treatments for alcoholism were more or less effective in terms of an impact on urges to drink. Thirty alcoholics were matched on the severity of their addiction and then assigned to Alcoholics Anonymous, individual psychotherapy, or anti-urge medication. At the end of treatment the mean number of urges per day for the three treatments, respectively, was 8.00, 6.00, and 6.00. The results were not statistically significant, meaning that there was not enough evidence to show that one treatment was more effective than another in controlling urges [$F(2, 18) = 2.73, p > .05$.] However, the sample size in this study was small and might not have been sufficient to pick up the effect if one treatment was just slightly better or worse than the others. Thus, it is recommended that the study be replicated with a larger sample size.

- 11.103 If $df_{\text{Treatment}} = 3$, then $k = 4$. Given $df_{\text{Residual}} = 45$ and $df_{\text{Treatment}} = 3$, then $\frac{45}{3} = n - 1$. If $15 = n - 1$, then $n = 16$. $N = k \times n = 4 \times 16 = 64$.

B-20 Appendix B

- 11.105 There is no need to be concerned about Type II error. The null hypothesis was rejected, so there should be concern about the possibility of Type I error.
- 11.107 The difference between a *HSD* value calculated at the $\alpha = .01$ level and the $\alpha = .05$ level is the *q* value. *q* is bigger at .01 than .05, so *HSD* is bigger at .01 than .05. As a result of this, two means have to be farther apart to be considered a statistically significant difference at the .01 level. And, this makes sense—farther apart means are more likely to represent a real difference between populations, so there is less chance of a Type I error. And the probability of a Type I error is what alpha represents.

Chapter 12

- 12.01 Two
- 12.03 independent; within
- 12.05 two; one
- 12.07 main; interaction
- 12.09 not parallel
- 12.11 main
- 12.13 independent
- 12.15 independence of observations
- 12.17 three
- 12.19 at least; at least
- 12.21 interaction effect
- 12.23 the number of rows
- 12.25 degrees of freedom; degrees of freedom
- 12.27 between groups; within groups
- 12.29 squares
- 12.31 $F_{\text{Interaction}}$
- 12.33 statistically significant
- 12.35 .05 (or 5%)
- 12.37 eta squared (η^2)
- 12.39 9
- 12.41 statistically significant
- 12.43 statistically significant
- 12.45 between-subjects, one-way ANOVA
- 12.47 between-subjects, two-way ANOVA
- 12.49 (a) $\text{Row}_1 = \frac{\text{Add up all the cell means in a row}}{\text{The number of cells in a row}}$
 $= \frac{25 + 35 + 45}{3} = 45.00$

$$\text{Row}_2 = \frac{\text{Add up all the cell means in a row}}{\text{The number of cells in a row}}$$

$$= \frac{25 + 15 + 5}{3} = 15.00$$

(b)

$$\text{Column}_1 = \frac{\text{Add up all the cell means in a column}}{\text{The number of cells in a column}}$$

$$= \frac{25 + 25}{2} = 25.00$$

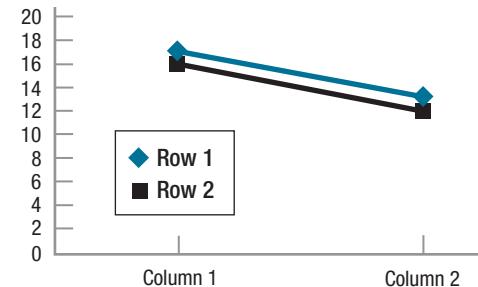
$$\text{Column}_2 = \frac{\text{Add up all the cell means in a column}}{\text{The number of cells in a column}}$$

$$= \frac{35 + 15}{2} = 25.00$$

$$\text{Column}_3 = \frac{\text{Add up all the cell means in a column}}{\text{The number of cells in a column}}$$

$$= \frac{45 + 5}{2} = 25.00$$

- 12.51 (a) There is no evidence of an interaction.



- (b) There appears to be a statistically significant columns main effect.

- (c) The main effect of columns should be interpreted.

- 12.53 (a) The random samples assumption was violated. The independence of observations assumption was not violated. It is assumed that the normality assumption was not violated. The homogeneity of variance assumption was violated.
- (b) It is OK to proceed with the test. The random samples assumption for between-subjects, two-way ANOVA is robust to violations. And, the homogeneity of variance assumption is robust to violation as long as the total sample size is large.

- 12.55 (a) H_0 _{Rows}: All row population means are the same.
 H_1 _{Rows}: At least one row population mean is different from at least one other row population mean.
- (b) H_0 _{Columns}: All column population means are the same.
 H_1 _{Columns}: At least one column population mean is different from at least one other column population mean.
- (c) H_0 _{Interaction}: There is no interactive effect of the two explanatory variables on the dependent variable in the population.
 H_1 _{Interaction}: The two explanatory variables, in the population, interact to affect the dependent variable in at least one cell.
- 12.57 (a) $df_{\text{Rows}} = R - 1 = 3 - 1 = 2$
(b) $df_{\text{Columns}} = C - 1 = 4 - 1 = 3$
(c) $df_{\text{Interaction}} = df_{\text{Rows}} \times df_{\text{Columns}} = 2 \times 3 = 6$
(d) $N = n \times R \times C = 8 \times 3 \times 4 = 96$
 $df_{\text{Within}} = N - (R \times C) = 96 - (3 \times 4) = 84$
(e) $df_{\text{Between}} = df_{\text{Rows}} + df_{\text{Columns}} + df_{\text{Interaction}}$
 $= 2 + 3 + 6 = 11$
(f) $df_{\text{Total}} = N - 1 = 96 - 1 = 95$
- 12.59 (a) $F_{cv\text{ Rows}} = 3.259$
(b) $F_{cv\text{ Columns}} = 3.259$
(c) $F_{cv\text{ Interaction}} = 2.634$
- 12.61 If $F_{\text{Interaction}} \geq 2.668$, reject H_0 .
If $F_{\text{Interaction}} < 2.668$, fail to reject H_0 .
- 12.63 $MS_{\text{Within}} = \frac{SS_{\text{Within}}}{df_{\text{Within}}} = \frac{168.48}{40} = 4.21$
- 12.65 $F_{\text{Rows}} = \frac{SS_{\text{Rows}}}{df_{\text{Rows}}} = \frac{17.30}{3.33} = 5.20$
- 12.67 $df_{\text{Rows}} = R - 1 = 2 - 1 = 1$
 $df_{\text{Columns}} = C - 1 = 2 - 1 = 1$
 $df_{\text{Interaction}} = df_{\text{Rows}} \times df_{\text{Columns}} = 1 \times 1 = 1$
 $N = n \times R \times C = 5 \times 2 \times 2 = 20$
 $df_{\text{Within}} = N - (R \times C) = 20 - (2 \times 2) = 16$
 $df_{\text{Between}} = df_{\text{Rows}} + df_{\text{Columns}} + df_{\text{Interaction}}$
 $= 1 + 1 + 1 = 3$
 $df_{\text{Total}} = N - 1 = 20 - 1 = 19$
 $MS_{\text{Rows}} = \frac{SS_{\text{Rows}}}{df_{\text{Rows}}} = \frac{15.00}{1} = 15.00$
 $MS_{\text{Cols}} = \frac{SS_{\text{Cols}}}{df_{\text{Cols}}} = \frac{12.00}{1} = 12.00$

$$MS_{\text{Interaction}} = \frac{SS_{\text{Interaction}}}{df_{\text{Interaction}}} = \frac{34.00}{1} = 34.00$$

$$MS_{\text{Within}} = \frac{SS_{\text{Within}}}{df_{\text{Within}}} = \frac{120.00}{16} = 7.50$$

$$F_{\text{Rows}} = \frac{MS_{\text{Rows}}}{MS_{\text{Within}}} = \frac{15.00}{7.50} = 2.00$$

$$F_{\text{Cols}} = \frac{MS_{\text{Cols}}}{MS_{\text{Within}}} = \frac{12.00}{7.50} = 1.60$$

$$F_{\text{Interaction}} = \frac{MS_{\text{Interaction}}}{MS_{\text{Within}}} = \frac{34.00}{7.50} = 4.53$$

Source of Variability	Sum of Squares	Degrees of Freedom	Mean Square	F ratio
Between groups	61.00	3		
Rows	15.00	1	15.00	2.00
Columns	12.00	1	12.00	1.60
Interaction	34.00	1	34.00	4.53
Within groups	120.00	16	7.50	
Total	181.00	19		

12.69 The null hypothesis was rejected.

12.71 $F(2, 66) = 3.70, p < .05$

12.73 $F(3, 168) = 2.45, p > .05$

12.75 (a) $F(2, 30) = 3.21, p > .05$

(b) $F(1, 30) = 1.08, p > .05$

(c) $F(2, 30) = 3.93, p < .05$

12.77 Interpret the interaction effect.

12.79 Interpret the interaction effect.

$$\begin{aligned} 12.81 \quad (a) \eta^2_{\text{Rows}} &= \frac{SS_{\text{Rows}}}{SS_{\text{Total}}} \times 100 \\ &= \frac{3.78}{47.83} \times 100 = 7.90\% \end{aligned}$$

(b) This is a medium effect.

$$\begin{aligned} 12.83 \quad (a) \eta^2_{\text{Rows}} &= \frac{SS_{\text{Rows}}}{SS_{\text{Total}}} \times 100 \\ &= \frac{33.00}{344.00} \times 100 = 9.59\% \end{aligned}$$

$$\begin{aligned} (b) \eta^2_{\text{Columns}} &= \frac{SS_{\text{Columns}}}{SS_{\text{Total}}} \times 100 \\ &= \frac{54.00}{344.00} \times 100 = 15.70\% \end{aligned}$$

B-22 Appendix B

- (c) $\eta^2_{\text{Interaction}} = \frac{SS_{\text{Interaction}}}{SS_{\text{Total}}} \times 100$
 $= \frac{45.00}{344.00} \times 100 = 13.08\%$
- 12.85 (a) $q_{\text{Rows}} = 3.44$
(b) $q_{\text{Columns}} = 2.85$
(c) $q_{\text{Interaction}} = 4.22$
- 12.87 $HSD_{\text{Rows}} = q_{\text{Rows}} \sqrt{\frac{MS_{\text{Within}}}{n_{\text{Rows}}}} = 3.49 \sqrt{\frac{4.56}{12}} = 2.15$
- 12.89 HSD value should only be calculated for the interaction effect.
- 12.91 (a)
 - Row 1 vs. Row 2 ($117.66 - 113.63 = 4.03$) is not a statistically significant difference.
 - Row 1 vs. Row 3 ($128.91 - 117.66 = 11.25$) is a statistically significant difference.
 - Row 2 vs. Row 3 ($128.91 - 113.63 = 15.28$) is a statistically significant difference.(b) The mean of Row 3 is higher than that of both Row 1 and Row 2.
- 12.93 To aid in interpretation, make a graph.



Men and women who went shopping for groceries with and without children were surveyed regarding how much they enjoyed the shopping experience. How much shoppers enjoyed the experience depended on the interaction of one's sex with whether there were children on the shopping trip ($F(1, 16) = 41.15, p < .05$). Men shopping with children and women shopping by themselves rated their shopping trips, on average as somewhat enjoyable. These ratings were statistically significantly higher than men shopping without children or

women shopping with children, who rated, on average, their shopping trips as somewhat less than enjoyable. It is possible that men and women feel differently about shopping with children, but another explanation is that novel situations are more enjoyable. For men, shopping with children tends to be novel and for women shopping without children is more likely to be novel. Future research should determine the novelty of the shopping experience as well as the presence of children.

12.95 Step 1 Pick a Test

Comparing the means of four independent samples formed by two crossed independent variables calls for a between-subjects, two-way ANOVA.

Step 2 Check the Assumptions

OK to proceed with the test.

- The random samples assumption is violated unless the population is meant to be the local high school.
- The independence of observations assumption is not violated.
- It is reasonable to assume that this psychological variable is normally distributed, so the normality assumption is not violated.
- The homogeneity of variance assumption is not violated.

Step 3 List the Hypotheses

H_0^{Rows} : All row population means are the same.

H_1^{Rows} : At least one row population mean is different from at least one other row population mean.

H_0^{Columns} : All column population means are the same.

H_1^{Columns} : At least one column population mean is different from at least one other column population mean.

$H_0^{\text{Interaction}}$: There is no interactive effect of the two explanatory variables on the dependent variable in the population.

$H_1^{\text{Interaction}}$: The two explanatory variables, in the population, interact to affect the dependent variable in at least one cell.

Step 4 Set the Decision Rule

$$df_{\text{Rows}} = 2 - 1 = 1$$

$$df_{\text{Columns}} = 2 - 1 = 1$$

$$df_{\text{Interaction}} = 1 \times 1 = 1$$

$$df_{\text{Within}} = 24 - (2 \times 2) = 20$$

All three F ratios have $df_{\text{numerator}} = 1$ and $df_{\text{denominator}} = 20$. $F_{cv} = 4.351$.

The decision rule for each of the three effects is:

- If $F \geq 4.351$, reject H_0 .
- If $F < 4.351$, fail to reject H_0 .

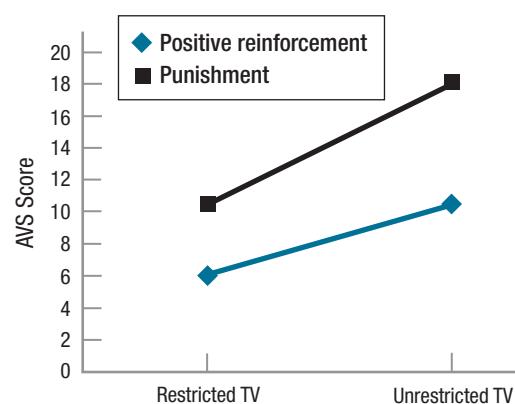
Step 5 Calculate the Value of the Test Statistic

Source of Variability	Sum of Squares	Degrees of Freedom	Mean Square	F ratio
Between groups	447.12	3		
Rows	210.04	1	210.04	20.51
Columns	222.04	1	222.04	21.68
Interaction	15.04	1	15.04	1.47
Within groups	204.83	20	10.24	
Total	651.95	23		

Step 6 Interpret the Results

- Calculate the row means and the column means:
 - Row 1, positive reinforcement = $\frac{6.00 + 10.50}{2} = 8.25$
 - Row 2, punishment = $\frac{10.33 + 18.00}{2} = 14.17$
 - Column 1, restricted TV = $\frac{6.00 + 10.33}{2} = 8.17$
 - Column 2, unrestricted TV = $\frac{10.50 + 18.00}{2} = 14.25$
- Determine statistical significance:
 - The rows effect (use of reinforcement vs. punishment) is statistically significant:
 - $F(1, 20) = 20.51, p < .05$
 - The mean Approval of Violence Scale (AVS) score for teens who were raised with positive reinforcement is 8.25 and is statistically lower than the mean of 14.17 for teens who were raised with punishment.

- Being raised with punishment leads to a statistically significant higher approval of the use of violence to solve disagreements.
- The columns effect (Restrict TV or not) is statistically significant:
 - $F(1, 20) = 21.68, p < .05$
 - The mean AVS score for teens who were restricted in what they could watch on TV is 8.17 and is statistically lower than the mean of 14.25 for teens who were unrestricted in their TV watching.
- Being unrestricted in what one watches on TV leads to a statistically significant higher approval of the use of violence to solve disagreements.
- The interaction effect is not statistically significant:
 - $F(1, 20) = 1.47, p > .05$
 - There is not enough evidence to conclude that type of discipline and restricting TV watching interact to affect the acceptance of violence as a way to solve disagreements.
- Inspection of the graph to aid in interpretation
 - The lines are mostly parallel, so there is no interaction effect.
 - The "Punishment" line is higher than the "Positive reinforcement" line, suggesting a main effect for type of discipline.
 - The "Unrestricted" dots are higher than the "Restricted" dots, suggesting a main effect for restricted TV.



B-24 Appendix B

- Calculating effect sizes:

- $$\eta^2_{\text{Rows}} = \frac{SS_{\text{Rows}}}{SS_{\text{Total}}} \times 100 = \frac{210.04}{651.95} \times 100 = 32.22\%$$

- This is a large effect.

- $$\eta^2_{\text{Columns}} = \frac{SS_{\text{Columns}}}{SS_{\text{Total}}} \times 100 = \frac{222.04}{651.95} \times 100 = 34.06\%$$

- This is a large effect.

- $$\eta^2_{\text{Interaction}} = \frac{SS_{\text{Interaction}}}{SS_{\text{Total}}} \times 100 = \frac{15.04}{651.95} \times 100 = 2.31\%$$

- This is a small effect.
- There is little reason to worry about Type II error for the interaction effect.
- Determine whether post-hoc tests are needed:
 - There are only two rows, so there is no need to calculate a *HSD* value for the statistically significant row effect: the positive reinforcement row has a statistically lower mean than the punishment row.
 - There are only two columns, so there is no need to calculate a *HSD* value for the statistically significant column effect: the restricted column has a statistically lower mean than the unrestricted column.
 - The interaction effect was not statistically significant, so there is no need to calculate the *HSD* value of the cells for the interaction effect.

The interpretation:

A developmental psychologist investigated the impact of (a) type of parental discipline, positive reinforcement vs. punishment, and (b) restrictions on television viewing on teenagers' acceptance of violence as a way to solve disagreements. It was found that each variable by itself had a large and independent effect on acceptance of violence. Teens whose parents

used punishment to discipline were statistically significantly more accepting of violence than were teens whose parents used positive reinforcement $F(1, 20) = 20.51, p < .05$. Teens whose television viewing was unrestricted were more accepting of the use of violence than were teens whose television viewing had been restricted $F(1, 20) = 21.68, p < .05$. This study suggests using positive reinforcement and restricting television viewing could lead to teenagers who are less likely to approve of the use of violence. But, nothing was manipulated in this study, so no conclusions regarding cause and effect can be drawn. Future research should attempt to manipulate either television viewing or type of discipline to see the impact on violence acceptance.

12.97 (a)

- A is given as 4.00.
- If $\frac{A + B}{2} = 8.00$ and $A = 4.00$, then

$$B = 12.00.$$

- If $A = 4.00$ and $\frac{A + C}{2} = 6.00$, then
 $C = 8.00$.

- If $B = 12.00$ and $\frac{B + D}{2} = 9.00$, then
 $D = 6.00$.

(b) $M_{\text{Grand}} = \frac{A + B + C + D}{4}$

$$= \frac{4.00 + 12.00 + 8.00 + 6.00}{4}$$
$$= 7.50$$

12.99 (a)

- Given $A = 1.00$ and $\frac{A + D}{2} = 2.50$, then
 $D = 4.00$.

- The other cells can't be calculated.

- (b) The grand mean, which is 4.00, can be calculated by finding the average of the row means or the column means.

12.101 e.

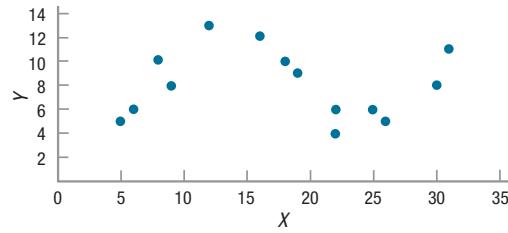
12.103 (a) $SS_{\text{Between}} = 17.50 + 13.75 + 22.25 = 53.50$

(b) $SS_{\text{Total}} = 53.50 + 44.75 = 98.25$

Chapter 13

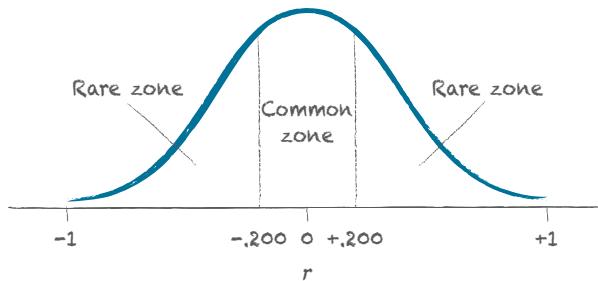
- 13.01 two; a
- 13.03 association
- 13.05 systematically
- 13.07 cause and effect
- 13.09 Scatterplots

- 13.11 $X; Y$
 13.13 circular shape
 13.15 undetermined
 13.17 .75
 13.19 quantifies; two variables
 13.21 weaker
 13.23 positive (direct)
 13.25 outlier
 13.27 unrestricted
 13.29 can
 13.31 ρ (rho)
 13.33 no
 13.35 alternate
 13.37 $-1; 1; 0$
 13.39 I
 13.41 reject
 13.43 scatterplot
 13.45 decision rule
 13.47 10
 13.49 negative; inverse
 13.51 insufficient; relationship
 13.53 r^2 (coefficient of determination)
 13.55 100
 13.57 ρ
 13.59 was not
 13.61 should
 13.63 .54
 13.65 partial correlation
 13.67



- 13.69 independent-samples t test
 13.71 Pearson r

- 13.73 (a) Random samples is not violated; independence of observations is not violated; normality is probably violated for age in college students; linearity is not violated.
 (b) The normality assumption is robust if the sample size is large, so it is OK to proceed with Pearson r .
 13.75 $H_0: \rho = 0; H_1: \rho \neq 0$
 13.77 $df = N - 2 = 63 - 2 = 61$
 13.79 $r_{cv} = \pm .666$
 13.81 $r_{cv} = .291$
 13.83 If $r \leq -.438$ or $r \geq .438$, reject H_0 ; If $-.438 < r < .438$, fail to reject H_0 .
 13.85



		X Deviation Scores	Y Deviation Scores
X	Y		
12	6	3.00	1.00
8	3	-1.00	-2.00
7	6	-2.00	1.00

		X Deviation Scores	Y Deviation Scores	Product of Deviation Scores
X	Y			
10	65	-2.33	11.67	-27.19
18	55	5.67	1.67	9.47
9	40	-3.33	-13.33	44.39
$M = 12.33$				$\Sigma = 26.67$

X	Y	X Deviation Scores	Y Deviation Scores	Product of Deviation Scores	Squared X Deviation Scores	Squared Y Deviation Scores
21	110	-4.33	10.67	-46.20	18.75	113.85
38	90	12.67	-9.33	-118.21	160.53	87.05
17	98	-8.33	-1.33	11.08	69.39	1.77
$M = 25.33$				$\Sigma = -153.33$	$\Sigma = 248.67$	$\Sigma = 202.67$

■ **B-26** Appendix B

13.93

X	Y	X Deviation Scores	Y Deviation Scores	Product of Deviation Scores	Squared X Deviation Scores	Squared Y Deviation Scores
88	23	1.67	-5.00	-8.35	2.79	25.00
93	28	6.67	0.00	0.00	44.49	0.00
78	33	-8.33	5.00	-41.65	69.39	25.00
M = 86.33	M = 28.00			$\Sigma = -50.00$	$\Sigma = 116.67$	$\Sigma = 50.00$
$SS_x SS_y = 5,833.50$						
$\sqrt{SS_x SS_y} = 76.38$						

$$13.95 \quad r = \frac{\Sigma[(X - M_x)(Y - M_y)]}{\sqrt{SS_x SS_y}} = \frac{255.70}{337.29} = .76$$

$$13.97 \quad \sqrt{SS_x SS_y} = \sqrt{12.44 \times 28.81} \\ = \sqrt{358.3964} = 18.93$$

$$r = \frac{\Sigma[(X - M_x)(Y - M_y)]}{\sqrt{SS_x SS_y}} = \frac{-12.03}{18.93} = -.64$$

- 13.99 (a) Failed to reject null hypothesis;
(b) Results not statistically significant.

$$13.101 \quad r(20) = -.67, p < .05$$

- 13.103 (a) Null hypothesis rejected; (b) Results statistically significant.

- 13.105 (a) There is not enough evidence to conclude a relationship exists between X and Y in the population; (b) There is no need to comment on the direction of the relationship because not enough evidence exists to conclude there is a relationship.

- 13.107 (a) $r^2 = r^2 \times 100 = .32^2 \times 100 = 10.24\%$;
(b) Medium effect.

$$13.109 \quad z_r = -0.50$$

$$13.111 \quad s_r = \frac{1}{\sqrt{N-3}} = \frac{1}{\sqrt{81-3}} = .11$$

$$13.113 \quad 95\%CI_{z_r} = z_r \pm 1.96s_r = -0.34 \pm 1.96(.24) \\ = [-0.81, 0.13]$$

$$13.115 \quad z_r = 0.60$$

$$s_r = \frac{1}{\sqrt{N-3}} = \frac{1}{\sqrt{45-3}} = .15$$

$$95\%CI_{z_r} = z_r \pm 1.96s_r = 0.60 \pm 1.96(.15) \\ = [0.31, 0.89]$$

$$13.117 \quad [-.69, -.39]$$

- 13.119 (a) Null hypothesis was not rejected; (b) Results were not statistically significant;
(c) Relationship in the population could

be weak; (d) Unsure of population value of correlation.

13.121 Power > .99; $\beta < .01$

13.123 120 cases

13.125 "In this study, the mental health level of a large, random sample of American men was measured, and the men were surveyed regarding how much aerobic exercise they engaged in each week. A moderately strong, direct, and statistically significant relationship was found between the two variables [$r(1,000) = .36, p < .05$], indicating that men who had higher levels of mental health also had higher levels of exercise. Because this is a correlational study, it is not clear if being mentally healthy leads to exercise, exercising leads to mental health, or a third variable, like physical health, leads to both. Future research should try to tease apart the direction of this relationship."

$$13.127 \quad r_{xy-z} = \frac{r_{xy} - (r_{xz} \times r_{yz})}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}} \\ = \frac{.28 - (.20 \times .22)}{\sqrt{(1 - .20^2)(1 - .22^2)}} = .25$$

13.129 After adjusting for socioeconomic status, there is a statistically significant, but small, relationship between the IQ of the student and the number of absences.

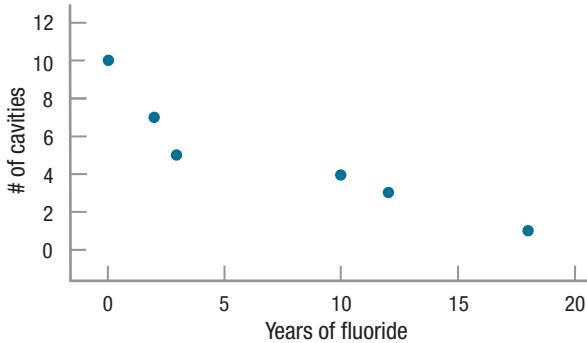
13.131 **Step 1** Pick a Test

Examining the relationship between two interval and/or ratio variables: Pearson r .

Step 2 Check the Assumptions

- Random samples: Violated.
- Independence of observations: Not violated.

- Normality: Willing to assume normality for both variables.
- Linearity: Not violated—see the scatterplot below. It is OK to proceed with Pearson r .



Step 3 List the Hypotheses

- $H_0: \rho = 0$.
- $H_1: \rho \neq 0$.

Step 4 Set the Decision Rule

A nondirectional (two-tailed) test, willing to have a 5% chance of a Type I error, so $\alpha = .05$, $df = N - 2 = 6 - 2 = 4$, so $r_{\alpha} = \pm .811$.

- If $r \leq -.811$ or $r \geq .811$, reject H_0 .
- If $-.811 < r < .811$, fail to reject H_0 .

Step 5 Calculate the Test Statistic

First, calculate means for X and Y so that deviation scores can be calculated.

$$r = \frac{\Sigma[(X - M_X)(Y - M_Y)]}{\sqrt{SS_X SS_Y}} = \frac{-102.00}{110.34} = -.9244 = -.92$$

Step 6 Interpret the Results

Was the null hypothesis rejected?

- $-.92 \leq -.811$, so reject H_0 .

How big is the effect?

$$\begin{aligned} r^2 &= r^2 \times 100 = -.92^2 \times 100 = .8464 \times 100 \\ &= 84.64\%. \end{aligned}$$

- This is a very large effect.

How wide is the confidence interval?

$$z_r = -1.59.$$

$$\begin{aligned} s_r &= \frac{1}{\sqrt{N - 3}} = \frac{1}{\sqrt{6 - 3}} = \frac{1}{\sqrt{3.00}} = \frac{1}{1.7321} \\ &= .5773 = .58 \end{aligned}$$

$$\begin{aligned} 95\% \text{ CI}_{z_r} &= z_r \pm 1.96 s_r = -1.59 \pm 1.96(.58) \\ &= -1.59 \pm 1.1368 = \text{from } -2.7268 \text{ to } \\ &\quad -0.4532 = \text{from } -2.73 \text{ to } -0.45. \end{aligned}$$

- In r units: from $-.99$ to $-.42$.

A four-point interpretation:

This study examined, in a small sample of adults from one dental practice, the relationship between the number of years of fluoride use in childhood and the number of cavities in permanent teeth. There was a very strong, statistically significant, inverse relationship

X	Y	X Deviation Scores	Y Deviation Scores	Product of Deviation Scores	Squared X Deviation Scores	Squared Y Deviation Scores
0	10	-7.50	5.00	-37.50	56.25	25.00
18	1	10.50	-4.00	-42.00	110.25	16.00
2	7	-5.50	2.00	-11.00	30.25	4.00
12	3	4.50	-2.00	-9.00	20.25	4.00
3	5	-4.50	0.00	0.00	20.25	0.00
10	4	2.50	-1.00	-2.50	6.25	1.00
$M = 7.50$	$M = 5.00$			$\Sigma = -102.00$	$\Sigma = 243.50$	$\Sigma = 50.00$
$SS_X SS_Y = 243.50 \times 50.00$						
$= 12,175.00$						
$\sqrt{SS_X SS_Y} = 110.34$						

B-28 Appendix B

- [$r(4) = -.92, p < .05$]. One plausible interpretation is that the more years of fluoride one has as a child, the fewer cavities one develops. However, as this is a correlational study, other explanations are possible. Perhaps parents who make sure their children receive fluoride also ensure that their children brush their teeth. Though the effect found in this sample was large, the sample size was small and the data derived from only one dental practice. Future research should replicate this study with a much larger and more representative sample.
- 13.133 There is insufficient evidence to conclude that X and Y are related in the population; thus, she should not draw any conclusion regarding the direction of the relationship.
- 13.135 g. r ranges from -1 to 1 . -1.26 is an impossible value.
- 13.137 A small effect requires at least 800 cases.
A medium effect requires at least 85 cases.
A large effect requires at least 30 cases.

Chapter 14

- 14.01 association; predicting; variable; variable
14.03 statistically significant
14.05 mean
14.07 minimizes; squared errors
14.09 error
14.11 predicted score; predictor
14.13 positive
14.15 Y-axis
14.17 point
14.19 the same
14.21 error
14.23 width
14.25 multiple regression; simple regression
14.27 ≈ 45
14.29 $b = r\left(\frac{s_Y}{s_X}\right) = -.47\left(\frac{3.65}{9.66}\right) = -.47 \times 0.3778$
 $= -0.1776 = -0.18$
14.31 Every penny increase in the price of gas was associated with families driving, on average, 35 fewer miles on their summer vacation.
14.33 $a = M_Y - bM_X = 17.50 - (4.33 \times 5.00)$
 $= 17.50 - 21.6500 = -4.1500 = -4.15$

- 14.35 $Y' = bX + a = 12.98 X + (-5.00)$
 $= 12.98X - 5.00$
14.37 First, calculate b and a :
 $b = r\left(\frac{s_Y}{s_X}\right) = -.24\left(\frac{3.98}{11.00}\right)$
 $= -.24 \times 0.3618 = -0.0868 = -0.09$
 $a = M_Y - bM_X = 25.00 - (-0.09 \times 55.00)$
 $= 25.00 - (-4.9500) = 25.00 + 4.9500$
 $= 29.9500 = 29.95$
Then complete the regression equation:
 $Y' = bX + a = -0.09X + 29.95$
14.39 $Y' = bX + a = 0.37X + 15 = (0.37 \times 25) + 15$
 $= 9.2500 + 15 = 24.2500 = 24.25$
14.41 (a)

(b) Y' values can be calculated for X values from 10 to 80.

14.43 First calculate the Y' values for the two endpoints:
 $Y' = 2.50X - 12.50 = (2.50 \times 30) - 12.50$
 $= 75.0000 - 12.50 = 62.5000 = 62.50$
 $Y' = 2.50X - 12.50 = (2.50 \times 70) - 12.50$
 $= 175.0000 - 12.50 = 162.5000 = 162.50$
The two endpoints are (30, 62.50) and (70, 162.50)
Then draw the regression line:

14.45 Residual = $Y - Y' = 12.96 - 13.43 = -0.4700$
 $= -0.47$

- 14.47 $s_{Y-Y'} = s_Y \sqrt{1 - r^2} = 10.55\sqrt{1 - .28^2}$
 $= 10.55\sqrt{1 - .0784} = 10.55\sqrt{.9216}$
 $= 10.55 \times .9600 = 10.1280 = 10.13$
- 14.49 The range of scores, from 400 to 1600, covers 1,200 points. If the average error is 40 points, that's off by 3% of the range. This seems to be a relatively small amount of error.
- 14.51 $Y' = (3.42 \times \text{variable 1}) + (-0.76 \times \text{variable 2}) + 5.74 = (3.42 \times 56.66) + (-0.76 \times 88.99) + 5.74 = 193.7772 + (-67.6324) + 5.74 = 193.7772 - 67.6324 + 5.74 = 131.8848 = 131.88$
- 14.53 b
- 14.55 c
- 14.57 The college applauds Shemekia's knowledge of statistics. But, the error could just as easily go the other way and Shemekia's first year GPA would be 2.52, well below their admission cutoff score.
- 14.59 Yes. Start with the value on the y -axis. Then move to the right horizontally until you intersect the regression line. Then move vertically downward until you hit the x -axis. This seems to be your predicted value of X for a given value of Y .
- 15.39 Examining the relationship between two nominal-level variables: chi-square test of independence. (Or, comparing two independent samples on a nominal-level dependent variable: chi-square test of independence.)
- 15.41 Looking at the relationship between two ordinal-level variables: Spearman r
- 15.43 (a) Random samples assumption is not violated; (b) Independence of observations assumption is not violated; (c) All cells have expected frequencies of at least 5 is violated, so can't proceed with the planned test.
- 15.45 H_0 : In the population of small towns, 2% experienced a murder in the past 10 years.
 H_1 : In the population of small towns, the percentage experiencing a murder in the past 10 years is not equal to 2%.
- 15.47 $df = k - 1 = 4 - 1 = 3$
- 15.49 $df = (R - 1) \times (C - 1) = (3 - 1) \times (3 - 1) = 2 \times 2 = 4$
- 15.51 (a) $\chi^2_{cv} = 5.991$ (b) If $\chi^2 \geq 5.991$, reject H_0 . If $\chi^2 < 5.991$, fail to reject H_0 .
- 15.53 $f_{\text{ExpectedA}} = \frac{\%_{\text{ExpectedA}} \times N}{100} = \frac{63 \times 97}{100} = 61.11$
 $f_{\text{ExpectedB}} = \frac{\%_{\text{ExpectedB}} \times N}{100} = \frac{37 \times 97}{100} = 35.89$

$$\begin{aligned} 15.55 \quad \chi^2 &= \sum \frac{(f_{\text{Observed}} - f_{\text{Expected}})^2}{f_{\text{Expected}}} \\ &= \frac{(9 - 6.97)^2}{6.97} + \frac{(14 - 15.58)^2}{15.58} + \frac{(18 - 18.45)^2}{18.45} \\ &= 0.76 \end{aligned}$$

$$\begin{aligned} 15.57 \quad f_{\text{ExpectedA}} &= \frac{\%_{\text{ExpectedA}} \times N}{100} \\ &= \frac{35 \times 90}{100} = 31.50 \\ f_{\text{ExpectedB}} &= \frac{\%_{\text{ExpectedB}} \times N}{100} \\ &= \frac{65 \times 90}{100} = 58.50 \end{aligned}$$

B-30 Appendix B

$$\begin{aligned}\chi^2 &= \sum \frac{(f_{\text{Observed}} - f_{\text{Expected}})^2}{f_{\text{Expected}}} \\ &= \frac{(40 - 31.50)^2}{31.50} + \frac{(50 - 58.50)^2}{58.50} \\ &= 3.53\end{aligned}$$

- 15.59 $\chi^2(1, N = 43) = 4.81, p < .05$
15.61 $\chi^2(4, N = 55) = 9.49, p < .05$
15.63 There are more cases than expected in Category I and fewer than expected in Category II.
15.65 "This study investigated if fortunetellers had the extrasensory ability to tell whether a randomly drawn card was red or black. Though the fortunetellers judged the color accurately 58% of the time, there was insufficient evidence to show that they performed better than chance [$\chi^2(1, N = 124) = 3.23, p > .05$]. The possibility exists that fortune tellers have some slight extrasensory abilities, too slight to be found in this sample. The study should be replicated with a larger sample size in order to have a better chance of finding the effect if it exists."
15.67

Step 1 Pick a Test

Seeing if the selection of one's favorite cola by the sample matches what would occur if the selection were made randomly. This calls for a chi-square goodness-of-fit test.

Step 2 Check the Assumptions

1. Random samples is not violated.
2. Independence of observations is not violated.
3. All expected frequencies more than 5 is not violated.

Step 3 List the Hypotheses

- H_0 : The distribution of cola choices in the population is random.
- H_1 : The distribution of cola choices in the population differs from random.

Step 4 Set the Decision Rule

$$df = k - 1 = 2 - 1 = 1$$

$$\chi^2_{cv} = 3.841 \text{ with } \alpha = .05$$

If $\chi^2 \geq 3.841$, then reject H_0 . If $\chi^2 < 3.841$, then fail to reject H_0 .

Step 5 Calculate the Test Statistic

$$\begin{aligned}f_{\text{ExpectedCorrect}} &= \frac{\%_{\text{ExpectedCorrect}} \times N}{100} \\ &= \frac{25 \times 880}{100} = 220.00\end{aligned}$$

$$\begin{aligned}f_{\text{ExpectedIncorrect}} &= \frac{\%_{\text{ExpectedIncorrect}} \times N}{100} \\ &= \frac{75 \times 880}{100} = 660.00\end{aligned}$$

$$\begin{aligned}\chi^2 &= \sum \frac{(f_{\text{Observed}} - f_{\text{Expected}})^2}{f_{\text{Expected}}} \\ &= \frac{(279 - 220.00)^2}{220.00} + \frac{(601 - 660.00)^2}{660.00} \\ &= 21.10\end{aligned}$$

Step 6 Interpret the Results

- $21.10 \geq 3.841$, so reject H_0 .
- $\chi^2(1, N = 880) = 21.10, p < .05$

Comparing observed frequency to expected frequency shows that more people than expected correctly identified their favorite cola:

The ability of consumers to differentiate their favorite cola from three competing brands was tested in this study. If consumers could not differentiate and were making decisions randomly, then they would have selected their preferred cola 25% of the time. They showed a statistically significant improvement over this [$\chi^2(1, N = 880) = 21.10, p < .05$], correctly choosing their cola 32% of the time. Though an effect exists, it doesn't seem very powerful as the majority of consumers could not correctly identify their preferred cola. It would be interesting to retest this sample and find out whether the people who identified the cola correctly can do so consistently.

	Support Gay Marriage: Yes	Support Gay Marriage: No
	20s	
40s		
60s		

$$15.71 N_{\text{Row1}} = 34 + 86 = 120$$

$$N_{\text{Row2}} = 67 + 113 = 180$$

$$N_{\text{Column1}} = 34 + 67 = 101$$

$$N_{\text{Column2}} = 86 + 113 = 199$$

$$15.73 f_{\text{ExpectedA}} = \frac{N_{\text{Row1}} \times N_{\text{Column1}}}{N} = \frac{34 \times 30}{72}$$

$$= 14.1667 = 14.17$$

$$f_{\text{ExpectedB}} = \frac{N_{\text{Row1}} \times N_{\text{Column2}}}{N} = \frac{34 \times 42}{72}$$

$$= 19.8333 = 19.83$$

$$f_{\text{ExpectedC}} = \frac{N_{\text{Row2}} \times N_{\text{Column1}}}{N} = \frac{38 \times 30}{72}$$

$$= 15.8333 = 15.83$$

$$f_{\text{ExpectedD}} = \frac{N_{\text{Row2}} \times N_{\text{Column2}}}{N} = \frac{38 \times 42}{72}$$

$$= 22.1667 = 21.17$$

$$15.75 \chi^2 = \sum \frac{(f_{\text{Observed}} - f_{\text{Expected}})^2}{f_{\text{Expected}}}$$

$$= \frac{(28 - 25.00)^2}{25.00} + \frac{(12 - 15.00)^2}{15.00}$$

$$+ \frac{(12 - 15.00)^2}{15.00} + \frac{(12 - 9.00)^2}{9.00}$$

$$= 2.56$$

$$15.77 f_{\text{Expected}} = \frac{N_{\text{Row1}} \times N_{\text{Column1}}}{N}$$

$$= \frac{35 \times 32}{71} = 15.7746 = 15.77$$

$$f_{\text{Expected}} = \frac{N_{\text{Row1}} \times N_{\text{Column2}}}{N}$$

$$= \frac{35 \times 39}{71} = 19.2253 = 19.22$$

$$f_{\text{Expected}} = \frac{N_{\text{Row2}} \times N_{\text{Column1}}}{N}$$

$$= \frac{36 \times 32}{71} = 16.2253 = 16.23$$

$$f_{\text{Expected}} = \frac{N_{\text{Row2}} \times N_{\text{Column2}}}{N}$$

$$= \frac{36 \times 39}{71} = 19.7746 = 19.77$$

$$\chi^2 = \sum \frac{(f_{\text{Observed}} - f_{\text{Expected}})^2}{f_{\text{Expected}}}$$

$$= \frac{(14 - 15.77)^2}{15.77} + \frac{(21 - 19.22)^2}{19.22}$$

$$+ \frac{(18 - 16.23)^2}{16.23} + \frac{(18 - 19.77)^2}{19.77}$$

$$= 0.72$$

15.79 Type I

15.81 Outcome A occurs less frequently than expected for Group I and more frequently than expected for Group II, while Outcome B occurs more frequently than expected for Group I and less frequently than expected for Group II.

$$15.83 \text{ (a)} V = \sqrt{\frac{\chi^2}{N \times df_{\text{RC}}}} = \sqrt{\frac{8.92}{78 \times 1}} = .34$$

(b) Medium effect

(c) No need to worry about Type II error. (As the null hypothesis was rejected, worry about Type I error.)

15.85 "A tennis journalist gathered a random sample of points played at major tournaments in order to investigate whether the server's advantage diminished as the length of the rally increased. She classified rallies as short (the ball went back and forth no more than 2 times) or long (the ball went back and forth more than 2 times). She found that when the rally was short, the server won the point about 62% of the time. In contrast, the server won only 48% of the time on long rallies. This difference is statistically significant [$\chi^2(1, N = 209) = 3.93, p < .05$], and it is a meaningful difference. Future research might examine if the same effect is found in non-professional players."

15.87 Step 1 Pick a Test

Comparing a nominal-level dependent variable between two independent samples calls for a chi-square test of independence.

Step 2 Check the Assumptions

- Random samples is violated. Can proceed, but limitations on generalizability.
- Independence of observations is not violated.

B-32 Appendix B

- Expected frequencies ≥ 5 is not violated.

Step 3 List the Hypotheses

H_0 : There is no difference in satisfaction with the scar left by sutures vs. staples.

H_1 : There is some difference in satisfaction with the scar left by sutures vs. staples.

Step 4 Set the Decision Rule

$$df = (R - 1) \times (C - 1) = (2 - 1) \times (2 - 1) = 1 \times 1 = 1$$

$$\chi^2_{cv} = 3.841 \text{ with } \alpha = .05$$

If $\chi^2 \geq 3.841$, then reject H_0 . If $\chi^2 < 3.841$, then fail to reject H_0 .

Step 5 Calculate the Test Statistic

$$f_{\text{Expected}} = \frac{N_{\text{Row1}} \times N_{\text{Column1}}}{N} = \frac{17 \times 21}{42} = 8.50$$

$$f_{\text{Expected}} = \frac{N_{\text{Row1}} \times N_{\text{Column2}}}{N} = \frac{21 \times 25}{42} = 12.50$$

$$f_{\text{Expected}} = \frac{N_{\text{Row2}} \times N_{\text{Column1}}}{N} = \frac{21 \times 17}{42} = 8.50$$

$$f_{\text{Expected}} = \frac{N_{\text{Row2}} \times N_{\text{Column2}}}{N} = \frac{21 \times 25}{42} = 12.50$$

$$\chi^2 = \sum \frac{(f_{\text{Observed}} - f_{\text{Expected}})^2}{f_{\text{Expected}}}$$

$$= \frac{(8 - 8.50)^2}{8.50} + \frac{(13 - 12.50)^2}{12.50}$$

$$+ \frac{(9 - 8.50)^2}{8.50} + \frac{(12 - 12.50)^2}{12.50}$$

$$= 0.0988 = 0.10$$

Step 6 Interpret the Results

- $0.10 < 3.841$; fail to reject H_0 .
- $\chi^2(1, N = 42) = 0.10, p > .05$

$$V = \sqrt{\frac{\chi^2}{N \times df_{RC}}} = \sqrt{\frac{0.10}{42 \times 1}} = .05$$

In this study, patient satisfaction with scars left by sutures vs. staples was measured.

There was no statistical difference in level of satisfaction [$\chi^2(1, N = 42) = 0.10, p > .05$], suggesting that either could be used to close incisions, at the preference of the surgeon and/or the patient. A positive of this study is that patients were randomly assigned to receive sutures or staples, but only one

surgeon provided both. Thus, it is possible that he is better at one technique over the other. Therefore, this study should be replicated using multiple surgeons and with the quality of the suturing and the stapling assessed.

15.89	12	8	20
	14	16	30
	26	24	

15.91 This table can't be completed. If the marginal frequencies are known, a table can be completed if there are as many cells filled in as degrees of freedom (in this case, 2).

15.93 If observed frequencies and expected frequencies are all the same, then $\chi^2 = 0$ and the null hypothesis is not rejected.

Chapter 16

- 16.01 describe
16.03 two variables
16.05 explanatory variable (or independent variable)
16.07 context
16.09 shape; level of measurement
16.11 is
16.13 Descriptive
16.15 Correlational
16.17 Experimental
16.19 Descriptive
16.21 (a) Ratio; (b) Continuous
16.23 (a) Nominal; (b) Discrete
16.25 (a) Interval; (b) Continuous
16.27 Standard deviation is the best answer. Also acceptable are interquartile range, variance, and range.
16.29 Bar graph
16.31 chi-square goodness-of-fit test
16.33 between-subjects, one-way ANOVA
16.35 independent-samples *t* test
16.37 single-sample *t* test
16.39 Pearson *r*
16.41 independent-samples *t* test
16.43 independent-samples *t* test
16.45 No such test appears in the book.
16.47 Pearson *r*
16.49 chi-square test of independence