Chapter 9: Inferences from Two Samples

Section 9-1: Two Proportions

1. The samples are simple random samples that are independent. For each of the two groups, the number of successes is at least 5 and the number of failures is at least 5. (Depending on what we call a success, the four numbers are 33, 115, 201,229 and 200,745 and all of those numbers are at least 5.) The requirements are satisfied.

2. $n_1 = 201{,}229,\ \hat{p}_1 = \dfrac{33}{201{,}299} = 0.000163992,\ \hat{q}_1 = 1 - 0.000163992 = 0.999836;$

$n_2 = 200{,}745,\ \hat{p}_2 = \dfrac{115}{200{,}745} = 0.000572866,\ \hat{q}_2 = 1 - 0.000572866 = 0.999427;$

$\overline{p} = \dfrac{33+115}{201{,}299+200{,}745} = 0.000368183,\ \overline{q} = 1 - 0.000368183 = 0.999632$

3. a. $H_0\colon p_1 = p_2;\ H_1\colon p_1 < p_2$

   b. There is sufficient evidence to support the claim that the rate of polio is less for children given the Salk vaccine than it is for children given a placebo. The Salk vaccine appears to be effective.

4. a. hypothesis test

   b. The P-value method and the critical value method are equivalent in the sense that they will always lead to the same conclusion, but the confidence interval method is not equivalent to them.

   c. 0.90, or 90%

   d. Because the confidence interval limits do not contain 0, there appears to be a significant difference between the two proportions. Because the confidence interval consists of negative values only, it appears that the first proportion is less than the second proportion. There is sufficient evidence to support the claim that the rate of polio is less for children given the Salk vaccine than it is for children given a placebo.

5. $H_0\colon p_1 = p_2;\ H_1\colon p_1 > p_2;$ population$_1$ = vinyl gloves, population$_2$ = latex gloves;

   Test statistic: $z = 12.82;$ P-value $= 0.0000;$ Critical value: $z = 2.33;$ Reject H_0. There is sufficient evidence to support the claim that vinyl gloves have a greater virus leak rate than latex gloves.

6. $H_0\colon p_1 = p_2;\ H_1\colon p_1 > p_2;$ population$_1$ = surgery, population$_2$ = splints;

   Test statistic: $z = 3.12;$ P-value $= 0.0009;$ Critical value: $z = 2.33;$ Reject H_0. There is sufficient evidence to support the claim that the success rate is better with surgery.

For Exercises 7–22, assume that the data fit the requirements for the statistical methods for two proportions unless otherwise indicated.

7. a. $H_0\colon p_1 = p_2;\ H_1\colon p_1 < p_2;$ population$_1$ = ginkgo, population$_2$ = placebo;

   Test statistic: $z = -1.66;$ P-value $= 0.0484$ (Table: 0.0485); Critical value: $z = -2.33;$ Fail to reject H_0. There is not sufficient evidence to support the claim that the rate of dementia among those who use ginkgo is less than the rate of dementia among those who use a placebo. There is not sufficient evidence to support the claim that ginkgo is effective in preventing dementia.

$$\overline{p} = \frac{246+277}{1545+1524} = \frac{523}{3069};\ \overline{q} = 1 - \frac{523}{3069} = \frac{2546}{3069};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\overline{pq}}{n_1} + \dfrac{\overline{pq}}{n_2}}} = \frac{\left(\dfrac{246}{1545} - \dfrac{277}{1524}\right) - 0}{\sqrt{\dfrac{\left(\dfrac{523}{3069}\right)\left(\dfrac{2546}{3069}\right)}{1545} + \dfrac{\left(\dfrac{523}{3069}\right)\left(\dfrac{2546}{3069}\right)}{1524}}} = -1.66$$

7. (continued)

b. 98% CI: $-0.0541 < p_1 - p_2 < 0.000904$; (Table: $-0.0542 < p_1 - p_2 < 0.000909$) Because the confidence interval limits include 0, there does not appear to be a significant difference between dementia rates for those treated with ginkgo and those given a placebo. There is not sufficient evidence to support the claim that the rate of dementia among those who use ginkgo is less than the rate of dementia among those who use a placebo. There is not sufficient evidence to support the claim that ginkgo is effective in preventing dementia.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{246}{1545} - \frac{277}{1524}\right) \pm 2.326\sqrt{\frac{\left(\frac{246}{1545}\right)\left(\frac{1299}{1545}\right)}{1545} + \frac{\left(\frac{277}{1524}\right)\left(\frac{1297}{1524}\right)}{1524}}$$

c. The sample results suggest that ginkgo is not effective in preventing dementia.

8. a. $H_0\colon p_1 = p_2$; $H_1\colon p_1 \neq p_2$; population$_1$ = Lipitor, population$_2$ = Placebo;
Test statistic: $z = -0.73$; P-value $= 0.4368$ (Table: 0.4645); Critical values: $z = \pm 1.96$; Fail to reject H_0. There is not sufficient evidence to reject the claim that Lipitor and the placebo have the same rates of infection.

$$\bar{p} = \frac{7+27}{94+270} = \frac{17}{182};\ \bar{q} = 1 - \frac{17}{182} = \frac{165}{182};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{7}{94} - \frac{27}{270}\right) - 0}{\sqrt{\frac{\left(\frac{17}{182}\right)\left(\frac{165}{182}\right)}{90} + \frac{\left(\frac{17}{182}\right)\left(\frac{165}{182}\right)}{270}}} = -0.73$$

b. 95% CI: $-0.0895 < p_1 - p_2 < -0.0385$ (Table: $-0.0890 < p_1 - p_2 < -0.0379$); Because the confidence interval limits contain 0, there does not appear\ to be a significant difference between the rates of infection for Lipitor and the placebo.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \frac{7}{94} - \frac{27}{270} \pm 1.96\sqrt{\frac{\left(\frac{7}{94}\right)\left(\frac{87}{94}\right)}{94} + \frac{\left(\frac{27}{270}\right)\left(\frac{243}{270}\right)}{270}}$$

c. The sample results suggest that Lipitor and the placebo have the same infection rates.

9. a. $H_0\colon p_1 = p_2$; $H_1\colon p_1 > p_2$; population$_1$ = sustained care, population$_2$ = standard care;
Test statistic: $z = 2.64$; P-value $= 0.0041$; Critical value: $z = 2.33$; Reject H_0. There is sufficient evidence to support the claim that the rate of success for smoking cessation is greater with the sustained care program.

$$\bar{p} = \frac{51+30}{198+199} = \frac{81}{397};\ \bar{q} = 1 - \frac{81}{397} = \frac{316}{397};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{51}{198} - \frac{30}{199}\right) - 0}{\sqrt{\frac{\left(\frac{81}{397}\right)\left(\frac{316}{397}\right)}{198} + \frac{\left(\frac{81}{397}\right)\left(\frac{316}{397}\right)}{199}}} = 2.64$$

b. 98% CI: $0.0135 < p_1 - p_2 < 0.200$ (Table: $0.0134 < p_1 - p_2 < 0.200$); Because the confidence interval limits do not contain 0, there is a significant difference between the two proportions. Because the interval consists of positive numbers only, it appears that the success rate for the sustained care program is greater than the success rate for the standard care program.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{51}{198} - \frac{30}{199}\right) \pm 2.33\sqrt{\frac{\left(\frac{51}{198}\right)\left(\frac{147}{198}\right)}{198} + \frac{\left(\frac{30}{199}\right)\left(\frac{169}{199}\right)}{199}}$$

9. (continued)

c. Based on the samples, the success rates of the programs are 25.8% (sustained care) and 15.1% (standard care). That difference does appear to be substantial, so the difference between the programs does appear to have practical significance.

10. a. $H_0\colon p_1 = p_2$; $H_1\colon p_1 \neq p_2$; population$_1$ = women, population$_2$ = men;
Test statistic: $z = 11.35$; P-value = 0.0000 (Table: 0.0001); Critical values: $z = \pm 2.576$; Reject H_0. There is sufficient evidence to support the claim that men and women have different rates of developing thyroid disease.

$$\bar{p} = \frac{1397+436}{2739+1352} = \frac{1833}{4091};\ \bar{q} = 1 - \frac{1833}{4091} = \frac{2258}{4091};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{1397}{2739} - \frac{436}{1352}\right) - 0}{\sqrt{\frac{\left(\frac{1833}{4091}\right)\left(\frac{2258}{4091}\right)}{2739} + \frac{\left(\frac{1833}{4091}\right)\left(\frac{2258}{4091}\right)}{1352}}} = 11.35$$

b. 95% CI: $0.147 < p_1 - p_2 < 0.229$; Because the confidence interval limits do not contain 0, there is a significant difference between the two proportions. There is sufficient evidence to support the claim that men and women have different rates of developing thyroid disease and it appear the rate among women is higher.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{1397}{2739} - \frac{436}{1352}\right) \pm 2.576\sqrt{\frac{\left(\frac{1397}{2739}\right)\left(\frac{1342}{2739}\right)}{1273} + \frac{\left(\frac{436}{1273}\right)\left(\frac{837}{1273}\right)}{1273}}$$

11. a. $H_0\colon p_1 = p_2$; $H_1\colon p_1 > p_2$; population$_1$ = over age 55, population$_2$ = under age 25;
Test statistic: $z = 6.44$; P-value = 0.0000 (Table: 0.0001); Critical value: $z = 2.33$; Reject H_0. There is sufficient evidence to support the claim that the proportion of people over 55 who dream in black and white is greater than the proportion of those under 25.

$$\bar{p} = \frac{68+13}{306+298} = \frac{81}{604};\ \bar{q} = 1 - \frac{81}{604} = \frac{523}{604};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{68}{306} - \frac{13}{298}\right) - 0}{\sqrt{\frac{\left(\frac{81}{604}\right)\left(\frac{523}{604}\right)}{306} + \frac{\left(\frac{81}{604}\right)\left(\frac{523}{604}\right)}{298}}} = 6.44$$

b. 98% CI: Because the confidence interval limits do not include 0, it appears that the two proportions are not equal. Because the confidence interval limits include only positive values, it appears that the proportion of people over 55 who dream in black and white is greater than the proportion of those under 25.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{68}{306} - \frac{13}{298}\right) \pm 2.33\sqrt{\frac{\left(\frac{68}{306}\right)\left(\frac{238}{306}\right)}{306} + \frac{\left(\frac{13}{298}\right)\left(\frac{285}{298}\right)}{298}}$$

c. The results suggest that the proportion of people over 55 who dream in black and white is greater than the proportion of those under 25, but the results cannot be used to verify the cause of that difference.

12. a. $H_0\colon p_1 = p_2$; $H_1\colon p_1 \neq p_2$; population$_1$ = OxyContin, population$_2$ = placebo;
Test statistic: $z = 1.78$; P-value = 0.0757 (Table: 0.0750); Critical values: $z = \pm 1.96$; Fail to reject H_0. There is not sufficient evidence to support the claim that there is a difference between the rates of nausea for those treated with OxyContin and those given a placebo.

12. (continued)

$$\overline{p}=\frac{52+5}{227+45}=\frac{57}{272};\ \overline{q}=1-\frac{57}{272}=\frac{215}{272};$$

$$z=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{\overline{p}\overline{q}}{n_1}+\frac{\overline{p}\overline{q}}{n_2}}}=\frac{\left(\frac{52}{227}-\frac{5}{45}\right)-0}{\sqrt{\frac{\left(\frac{57}{272}\right)\left(\frac{215}{272}\right)}{227}+\frac{\left(\frac{57}{272}\right)\left(\frac{215}{272}\right)}{45}}}=1.78$$

b. 95% CI: $0.0111 < p_1 - p_2 < 0.225$; Because the confidence interval limits do not contain 0, there is a significant difference between the two proportions.

$$(\hat{p}_1-\hat{p}_2)\pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}=\left(\frac{52}{227}-\frac{5}{45}\right)\pm 1.96\sqrt{\frac{\left(\frac{52}{227}\right)\left(\frac{175}{227}\right)}{227}+\frac{\left(\frac{5}{45}\right)\left(\frac{40}{45}\right)}{45}}$$

c. The conclusions from parts (a) and (b) are different. The conclusion from part (a) results from a hypothesis test instead of an estimate of the difference between the two rates of nausea, so there does not appear to be sufficient evidence to conclude that there is a difference between the rates of nausea for those treated with OxyContin and those given a placebo.

13. a. $H_0\colon p_1 = p_2$; $H_1\colon p_1 > p_2$; population$_1$ = wearing seatbelt, population$_2$ = not wearing seatbelt; Test statistic: $z = 6.11$; P-value = 0.0000 (Table: 0.0001); Critical value: $z = 1.645$; Reject H_0. There is sufficient evidence to support the claim that the fatality rate is higher for those not wearing seat belts.

$$\overline{p}=\frac{31+16}{2823+7765}=\frac{47}{10{,}588};\ \overline{q}=1-\frac{47}{10{,}588}=\frac{10{,}541}{10{,}588};$$

$$z=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{\overline{p}\overline{q}}{n_1}+\frac{\overline{p}\overline{q}}{n_2}}}=\frac{\left(\frac{31}{2823}-\frac{16}{7765}\right)-0}{\sqrt{\frac{\left(\frac{47}{10{,}588}\right)\left(\frac{10{,}541}{10{,}588}\right)}{2823}+\frac{\left(\frac{47}{10{,}588}\right)\left(\frac{10{,}541}{10{,}588}\right)}{7765}}}=6.11$$

b. 90% CI: $0.00559 < p_1 - p_2 < 0.0123$; Because the confidence interval limits do not include 0, it appears that the two fatality rates are not equal. Because the confidence interval limits include only positive values, it appears that the fatality rate is higher for those not wearing seat belts.

$$(\hat{p}_1-\hat{p}_2)\pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}=\left(\frac{31}{2823}-\frac{16}{7765}\right)\pm 1.645\sqrt{\frac{\left(\frac{31}{2823}\right)\left(\frac{2792}{2823}\right)}{2823}+\frac{\left(\frac{16}{7765}\right)\left(\frac{7749}{7765}\right)}{7765}}$$

c. The results suggest that the use of seat belts is associated with fatality rates lower than those associated with not using seat belts.

14. a. H_0: $p_1 = p_2$; H_1: $p_1 \neq p_2$; population$_1$ = day, population$_2$ = night;
Test statistic: $z = 18.26$; P-value = 0.0000 (Table: 0.0002); Critical values: $z = \pm 2.576$ (Table: $z = \pm 2.575$); Reject H_0. There is sufficient evidence to warrant rejection of the claim that the survival rates are the same for day and night.

$$\bar{p} = \frac{11{,}604 + 4139}{58{,}593 + 28{,}155} = \frac{15{,}743}{86{,}748};\ \bar{q} = 1 - \frac{15{,}743}{86{,}748} = \frac{71{,}005}{86{,}748};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{11{,}604}{58{,}593} - \frac{4139}{28{,}155}\right) - 0}{\sqrt{\frac{\left(\frac{15{,}743}{86{,}748}\right)\left(\frac{71{,}005}{86{,}748}\right)}{58{,}593} + \frac{\left(\frac{15{,}743}{86{,}748}\right)\left(\frac{71{,}005}{86{,}748}\right)}{28{,}155}}} = 18.26$$

b. 99% CI: $0.0441 < p_1 - p_2 < 0.0579$; Because the confidence interval limits do not contain 0, there appears to be a significant difference between the two proportions. There is sufficient evidence to warrant rejection of the claim that the survival rates are the same for day and night.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} =$$

$$\left(\frac{11{,}604}{58{,}593} - \frac{4139}{28{,}155}\right) \pm 2.576\sqrt{\frac{\left(\frac{11{,}604}{58{,}593}\right)\left(\frac{46{,}989}{58{,}593}\right)}{58{,}593} + \frac{\left(\frac{4139}{28{,}155}\right)\left(\frac{24{,}016}{28{,}155}\right)}{28{,}155}}$$

c. The data suggest that for in-hospital patients who suffer cardiac arrest, the survival rate is not the same for day and night. It appears that the survival rate is higher for in-hospital patients who suffer cardiac arrest during the day.

15. a. H_0: $p_1 = p_2$; H_1: $p_1 \neq p_2$; population$_1$ = echinacea, population$_2$ = placebo;
Test statistic: $z = 0.57$; P-value = 0.5720 (Table: 0.5868); Critical values: $z = \pm 1.96$; Fail to reject H_0. There is not sufficient evidence to support the claim that Echinacea treatment has an effect.

$$\bar{p} = \frac{40 + 88}{45 + 103} = \frac{32}{37};\ \bar{q} = 1 - \frac{32}{37} = \frac{5}{37};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{40}{45} - \frac{88}{103}\right) - 0}{\sqrt{\frac{\left(\frac{32}{37}\right)\left(\frac{5}{37}\right)}{45} + \frac{\left(\frac{32}{37}\right)\left(\frac{5}{37}\right)}{103}}} = 0.57$$

b. 95% CI: $-0.0798 < p_1 - p_2 < 0.149$; Because the confidence interval limits do contain 0, there is not a significant difference between the two proportions. There is not sufficient evidence to support the claim that Echinacea treatment has an effect.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{40}{45} - \frac{88}{103}\right) \pm 1.96\sqrt{\frac{\left(\frac{40}{45}\right)\left(\frac{5}{45}\right)}{45} + \frac{\left(\frac{88}{103}\right)\left(\frac{15}{103}\right)}{103}}$$

c. Echinacea does not appear to have a significant effect on the infection rate. Because it does not appear to have an effect, it should not be recommended.

16. a. H_0: $p_1 = p_2$; H_1: $p_1 < p_2$; population$_1$ = used bednet, population$_2$ = did not use bednet;
Test statistic: $z = -2.44$; P-value = 0.0074; Critical value: Reject H_0. There is sufficient evidence to support the claim that the incidence of malaria is lower for infants who use the bednets.

$$\overline{p} = \frac{15+27}{343+294} = \frac{6}{91};\ \overline{q} = 1 - \frac{6}{91} = \frac{85}{91};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\overline{p}\overline{q}}{n_1} + \frac{\overline{p}\overline{q}}{n_2}}} = \frac{\left(\frac{15}{343} - \frac{27}{294}\right) - 0}{\sqrt{\frac{\left(\frac{6}{91}\right)\left(\frac{85}{91}\right)}{343} + \frac{\left(\frac{6}{91}\right)\left(\frac{85}{91}\right)}{294}}} = -2.44$$

b. 98% CI: $-0.0950 < p_1 - p_2 < -0.00125$; Because the confidence interval does not include 0 and it includes only negative values, it appears that the rate of malaria is lower for infants who use the bednets.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{15}{343} - \frac{27}{294}\right) \pm 2.33\sqrt{\frac{\left(\frac{15}{343}\right)\left(\frac{328}{343}\right)}{343} + \frac{\left(\frac{27}{294}\right)\left(\frac{267}{294}\right)}{294}}$$

c. The bednets appear to be effective.

17. a. H_0: $p_1 = p_2$; H_1: $p_1 < p_2$; population$_1$ = used left ear, population$_2$ = used right ear;
Test statistic: $z = -7.94$; P-value = 0.0000 (Table: 0.0001); Critical value: $z = -2.33$; Reject H_0. There is sufficient evidence to support the claim that the rate of right-handedness for those who prefer to use their left ear for cell phones is less than the rate of right-handedness for those who prefer to use their right ear for cell phones.

$$\overline{p} = \frac{166+436}{216+452} = \frac{301}{334};\ \overline{q} = 1 - \frac{301}{334} = \frac{33}{334};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\overline{p}\overline{q}}{n_1} + \frac{\overline{p}\overline{q}}{n_2}}} = \frac{\left(\frac{166}{216} - \frac{436}{452}\right) - 0}{\sqrt{\frac{\left(\frac{301}{334}\right)\left(\frac{33}{334}\right)}{216} + \frac{\left(\frac{301}{334}\right)\left(\frac{33}{334}\right)}{452}}} = -7.94$$

b. 98% CI: $-0.266 < p_1 - p_2 < -0.126$; Because the confidence interval limits do not contain 0, there is a significant difference between the two proportions. Because the interval consists of negative numbers only, it appears that the claim is supported. The difference between the populations does appear to have practical significance.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{166}{216} - \frac{436}{452}\right) \pm 2.33\sqrt{\frac{\left(\frac{166}{216}\right)\left(\frac{50}{216}\right)}{216} + \frac{\left(\frac{436}{452}\right)\left(\frac{16}{452}\right)}{452}}$$

18. a. $H_0: p_1 = p_2$; $H_1: p_1 < p_2$; population$_1$ = single bill, population$_2$ = multiple bills;
Test statistic: $z = -1.85$; P-value $= 0.0324$ (Table: 0.0322); Critical value: $z = -1.645$; Reject H_0. There is sufficient evidence to support the claim that when given a single large bill, a smaller proportion of women in China spend some or all of the money when compared to the proportion of women in China given the same amount in smaller bills.

$$\bar{p} = \frac{60+68}{75+75} = \frac{64}{75}; \bar{q} = 1 - \frac{64}{75} = \frac{11}{75};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{60}{75} - \frac{68}{75}\right) - 0}{\sqrt{\frac{\left(\frac{64}{75}\right)\left(\frac{11}{75}\right)}{75} + \frac{\left(\frac{64}{75}\right)\left(\frac{11}{75}\right)}{75}}} = -1.85$$

b. 90% CI: $-0.201 < p_1 - p_2 < -0.0127$; Because the confidence interval does not include 0 and it includes only negative values, it appears that the first proportion is less than the second proportion. There is sufficient evidence to support the claim that when given a single large bill, a smaller proportion of women in China spend some or all of the money when compared to the proportion of women in China given the same amount in smaller bills.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{60}{75} - \frac{68}{75}\right) \pm 1.645\sqrt{\frac{\left(\frac{60}{75}\right)\left(\frac{15}{75}\right)}{75} + \frac{\left(\frac{68}{75}\right)\left(\frac{7}{75}\right)}{75}}$$

c. Because the P-value $= 0.0324$ (Table: 0.0322), the difference is significant at the 0.05 significance level, but not at the 0.01 significance level. The conclusion does change.

19. a. $H_0: p_1 = p_2$; $H_1: p_1 > p_2$; population$_1$ = oxygen, population$_2$ = placebo;
Test statistic: $z = 9.97$; P-value $= 0.0000$ (Table: 0.0001); Critical value: $z = 2.33$; Reject H_0. There is sufficient evidence to support the claim that the cure rate with oxygen treatment is higher than the cure rate for those given a placebo. It appears that the oxygen treatment is effective.

$$\bar{p} = \frac{116+29}{150+148} = \frac{145}{298}; \bar{q} = 1 - \frac{145}{298} = \frac{153}{298};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{116}{150} - \frac{29}{148}\right) - 0}{\sqrt{\frac{\left(\frac{145}{298}\right)\left(\frac{153}{298}\right)}{150} + \frac{\left(\frac{145}{298}\right)\left(\frac{153}{298}\right)}{148}}} = 9.97$$

b. 98% CI: $0.467 < p_1 - p_2 < 0.687$; Because the confidence interval limits do not include 0, it appears that the two cure rates are not equal. Because the confidence interval limits include only positive values, it appears that the cure rate with oxygen treatment is higher than the cure rate for those given a placebo. It appears that the oxygen treatment is effective.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{116}{150} - \frac{29}{148}\right) \pm 2.33\sqrt{\frac{\left(\frac{116}{150}\right)\left(\frac{34}{150}\right)}{150} + \frac{\left(\frac{29}{148}\right)\left(\frac{119}{148}\right)}{148}}$$

c. The results suggest that the oxygen treatment is effective in curing cluster headaches.

20. a. $H_0: p_1 = p_2; H_1: p_1 \neq p_2$; population$_1$ = aspirin, population$_2$ = placebo;
Test statistic: $z = -5.19$; P-value $= 0.0000$ (Table: 0.0002); Critical values: $z = \pm 1.96$; Reject H_0. There is sufficient evidence to warrant rejection of the claim that aspirin has no effect on myocardial infarctions.

$$\bar{p} = \frac{139+239}{11,037+11,034} = \frac{18}{1051}; \bar{q} = 1 - \frac{18}{1051} = \frac{1033}{1051};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{139}{11,037} - \frac{239}{11,034}\right) - 0}{\sqrt{\frac{\left(\frac{18}{1051}\right)\left(\frac{1033}{1051}\right)}{11,037} + \frac{\left(\frac{18}{1051}\right)\left(\frac{1033}{1051}\right)}{11,034}}} = -5.19$$

b. 95% CI: $-0.0125 < p_1 - p_2 < -0.00564$; Because the confidence interval limits do not contain 0, there is a significant difference between the two proportions. There is sufficient evidence to warrant rejection of the claim that aspirin has no effect on myocardial infarctions.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{139}{11,037} - \frac{239}{11,034}\right) \pm 1.96\sqrt{\frac{\left(\frac{139}{11,037}\right)\left(\frac{10,898}{11,037}\right)}{11,037} + \frac{\left(\frac{239}{11,034}\right)\left(\frac{10,795}{11,034}\right)}{11,034}}$$

c. It appears that aspirin has an effect. Because the treatment group had a lower rate of myocardial infarctions, the aspirin treatment appears to be associated with a lower rate of myocardial infarctions. It should be noted that the treatment group and placebo group included only male physicians, so the results may or may not apply to the general population.

21. a. $H_0: p_1 = p_2; H_1: p_1 < p_2$; population$_1$ = male, population$_2$ = female;
Test statistic: $z = -1.17$; P-value $= 0.1214$ (Table: 0.1210); Critical value: $z = -2.33$; Fail to reject H_0. There is not sufficient evidence to support the claim that the rate of left-handedness among males is less than that among females.

$$\bar{p} = \frac{23+65}{240+520} = \frac{11}{95}; \bar{q} = 1 - \frac{11}{95} = \frac{84}{95};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{23}{240} - \frac{65}{520}\right) - 0}{\sqrt{\frac{\left(\frac{11}{95}\right)\left(\frac{84}{95}\right)}{240} + \frac{\left(\frac{11}{95}\right)\left(\frac{84}{95}\right)}{520}}} = -1.17$$

b. 98% CI: $-0.0848 < p_1 - p_2 < -0.0264$ (Table: $-0.0849 < p_1 - p_2 < -0.0265$); Because the confidence interval limits include 0, there does not appear to be a significant difference between the rate of left-handedness among males and the rate among females. There is not sufficient evidence to support the claim that the rate of left-handedness among males is less than that among females.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{23}{240} - \frac{65}{520}\right) \pm 2.33\sqrt{\frac{\left(\frac{23}{240}\right)\left(\frac{217}{240}\right)}{240} + \frac{\left(\frac{65}{520}\right)\left(\frac{455}{520}\right)}{520}}$$

c. The rate of left-handedness among males does not appear to be less than the rate of left-handedness among females.

22. a. H_0: $p_1 = p_2$; H_1: $p_1 > p_2$; population$_1$ = helicopter, population$_2$ = ground;
Test statistic: $z = 10.75$; P-value $= 0.0000$ (Table: 0.0001); Critical value: $z = 2.33$; Reject H_0. There is sufficient evidence to support the claim that the rate of fatalities is higher for patients transported by helicopter.

$$\bar{p} = \frac{7813+17,775}{61,909+161,566} = \frac{25,588}{223,475}; \bar{q} = 1-\frac{25,588}{223,475} = \frac{197,887}{223,475};$$

$$z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1}+\frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{7813}{61,909}-\frac{17,775}{161,566}\right)-0}{\sqrt{\frac{\left(\frac{25,588}{223,475}\right)\left(\frac{197,887}{223,475}\right)}{61,909}+\frac{\left(\frac{25,588}{223,475}\right)\left(\frac{197,887}{223,475}\right)}{161,566}}} = 10.75$$

b. 98% CI: $0.0126 < p_1 - p_2 < 0.0198$; Because the confidence interval limits do not contain 0, there is a significant difference between the two proportions. Because the entire range of values in the confidence interval consists of positive numbers, it appears that the rate of fatalities is higher for patients transported by helicopter.

$$(\hat{p}_1-\hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}} =$$

$$\left(\frac{7813}{61,909}-\frac{17,775}{161,566}\right) \pm 2.33\sqrt{\frac{\left(\frac{7813}{61,909}\right)\left(\frac{54,096}{61,909}\right)}{61,909}+\frac{\left(\frac{17,775}{161,566}\right)\left(\frac{143,791}{161,566}\right)}{161,566}}$$

c. The fatality rate for the helicopter sample is 0.126, or 12.6%, and the fatality rate for the ground services sample is 0.110, or 11.0%. The large sample sizes result in a significant difference, but it does not appear that the difference has very much practical significance. Also, it is possible that the most serious of the serious traumatic injuries led to helicopter transportation, and that could partly explain the higher rate of fatalities with helicopter transportation.

23. $n = \frac{z_{\alpha/2}^2}{2E^2} = \frac{1.96^2}{2 \cdot 0.03^2} = 2135$; The samples should include 2135 men and 2135 women.

24. a. The method of this section requires that both samples must have at least 5 successes and 5 failures, but the group not exposed to yawning includes a frequency of 4, which violates that requirement.

b. P-value $= 0.3729$ (Table: 0.3745); which is not close to the P-value of 0.5128 from Fisher's exact test.

H_0: $p_1 = p_2$; H_1: $p_1 > p_2$; population$_1$ = yawning, population$_2$ = not yawning;

$$\bar{p} = \frac{10+4}{34+16} = \frac{7}{25}; \bar{q} = 1-\frac{7}{25} = \frac{18}{25};$$

$$z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1}+\frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{10}{34}-\frac{4}{16}\right)-0}{\sqrt{\frac{\left(\frac{7}{25}\right)\left(\frac{18}{25}\right)}{34}+\frac{\left(\frac{7}{25}\right)\left(\frac{18}{25}\right)}{16}}} = 0.32$$

25. a. 95% CI: $0.0227 < p_1 - p_2 < 0.217$; population$_1$ = first sample, population$_2$ = second sample; Because the confidence interval limits do not contain 0, it appears that $p_1 = p_2$ can be rejected.

$$(\hat{p}_1-\hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{112}{200}-\frac{88}{200}\right) \pm 1.96\sqrt{\frac{\left(\frac{112}{200}\right)\left(\frac{88}{200}\right)}{200}+\frac{\left(\frac{88}{200}\right)\left(\frac{112}{200}\right)}{200}}$$

25. (continued)

b. First sample: 95% CI: $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = \frac{112}{200} \pm 1.96\sqrt{\frac{\left(\frac{112}{200}\right)\left(\frac{88}{200}\right)}{200}} \Rightarrow 0.491 < p_1 < 0.629$

Second sample: 95% CI: $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = \frac{88}{200} \pm 1.96\sqrt{\frac{\left(\frac{88}{200}\right)\left(\frac{112}{200}\right)}{200}} \Rightarrow 0.371 < p_2 < 0.509$

Because the confidence intervals do overlap, it appears that $p_1 = p_2$ cannot be rejected.

c. H_0: $p_1 = p_2$; H_1: $p_1 \neq p_2$; population$_1$ = first sample, population$_2$ = second sample;
Test statistic: $z = 2.40$; P-value $= 0.0164$; Critical values: $z = \pm 1.96$; Reject H_0. There is sufficient evidence to reject $p_1 = p_2$.

$$\bar{p} = \frac{112+88}{200+200} = \frac{1}{2};\ \bar{q} = 1 - \frac{1}{2} = \frac{1}{2};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{112}{200} - \frac{88}{200}\right) - 0}{\sqrt{\frac{\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}{200} + \frac{\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}{200}}} = 2.40$$

d. Reject $p_1 = p_2$. The least effective method is using the overlap between the individual confidence intervals.

26. Hypothesis test: H_0: $p_1 = p_2$; H_1: $p_1 \neq p_2$; population$_1$ = first sample, population$_2$ = second sample;
Test statistic: $z = -1.9615$; P-value $= 0.0498$ (Table: 0.05); Critical values: $z = \pm 1.96$; Reject H_0. There is sufficient evidence to reject $p_1 = p_2$.

$$\bar{p} = \frac{10+1404}{20+2000} = \frac{7}{10};\ \bar{q} = 1 - \frac{7}{10} = \frac{3}{10};$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{\left(\frac{10}{20} - \frac{1404}{2000}\right) - 0}{\sqrt{\frac{\left(\frac{7}{10}\right)\left(\frac{3}{10}\right)}{20} + \frac{\left(\frac{7}{10}\right)\left(\frac{3}{10}\right)}{2000}}} = -1.9615$$

95% CI: $-0.422 < p_1 - p_2 < 0.0180$; population$_1$ = first sample, population$_2$ = second sample; which suggests that we should not reject $p_1 = p_2$ (because 0 is included).

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \left(\frac{10}{20} - \frac{1404}{2000}\right) \pm 1.96\sqrt{\frac{\left(\frac{10}{20}\right)\left(\frac{10}{20}\right)}{20} + \frac{\left(\frac{1404}{2000}\right)\left(\frac{596}{2000}\right)}{2000}}$$

The hypothesis test and confidence interval lead to different conclusions about the equality of $p_1 = p_2$.

Section 9-2: Two Means: Independent Samples

1. Only part (c) describes independent samples.

2. a. Because the confidence interval does not include 0, it appears that there is a significant difference between the mean level of hemoglobin in women and the mean level of hemoglobin in men.

   b. We have 95% confidence that the interval from -1.76 g/dL to -1.62 g/dL actually contains the value of the difference between the two population means $(\mu_1 - \mu_2)$.

   c. $1.62 \text{ g/dL} < \mu_1 - \mu_2 < 1.76 \text{ g/dL}$

3. a. yes

   b. yes

   c. 98%

4. The critical value of $t = 1.685$ are more conservative than $t = 1.665$ in the sense that rejection of the null hypothesis requires a *greater difference* between the sample means. The sample evidence must be stronger with $t = 1.685$.

5. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = low humidity, population$_2$ = high humidity; Test statistic: $t = -0.452$; P-value $= 0.6521$ (Table: P-value > 0.20); Critical values: $t = \pm 1.987$ (Table: $t = \pm 2.014$); Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the two groups are from populations with the same mean. The result suggests that the increased humidity does not help in the treatment of croup.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(0.98 - 1.09) - 0}{\sqrt{\frac{1.22^2}{46} + \frac{1.11^2}{46}}} = -0.452 \text{ (df = 45)}$$

6. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = echinacea, population$_2$ = placebo; Test statistic: $t = -6.794$; P-value $= 0.0000$ (Table: P-value < 0.005); Critical values: $t = \pm 1.964$ (Table: $t = \pm 1.960$); Reject H_0. There is sufficient evidence to support the claim that echinacea has an effect on "days of fever."

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(0.81 - 1.50) - 0}{\sqrt{\frac{1.50^2}{337} + \frac{1.16^2}{370}}} = -6.794 \text{ (df = 336)}$$

7. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 < \mu_2$; population$_1$ = exposed to cocaine, population$_2$ = not exposed to cocaine; Test statistic: $t = -2.908$; P-value $= 0.0019$ (Table: P-value < 0.005); Critical value: $t = -1.649$ (Table: $t \approx -1.653$); Reject H_0. There is sufficient evidence to support the claim that the children exposed to cocaine have a lower mean score.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(7.3 - 8.2) - 0}{\sqrt{\frac{3.0^2}{190} + \frac{3.0^2}{186}}} = -2.908 \text{ (df = 185)}$$

b. 90% CI: $-1.4 < \mu_1 - \mu_2 < -0.4$; The confidence interval consists of negative numbers only and does not include 0. There is sufficient evidence to support the claim that the children exposed to cocaine have a lower mean score.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (7.3 - 8.2) \pm 1.653\sqrt{\frac{3.0^2}{190} + \frac{3.0^2}{186}} \text{ (df = 185)}$$

8. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 > \mu_2$; population$_1$ = magnet treatment, population$_2$ = sham treatment; Test statistic: $t = 0.132$; P-value $= 0.4480$ (Table: P-value > 0.10); Critical value: $t = 1.691$ (Table: $t = 1.729$); Fail to reject H_0. There is not sufficient evidence to support the claim that the magnets are effective in reducing pain.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(0.49 - 0.44) - 0}{\sqrt{\frac{0.96^2}{20} + \frac{1.4^2}{20}}} = 0.132 \text{ (df = 19)}$$

b. 90% CI: $-0.59 < \mu_1 - \mu_2 < 0.69$ (Table: $-0.61 < \mu_1 - \mu_2 < 0.71$);

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (0.49 - 0.44) \pm 1.729\sqrt{\frac{0.96^2}{20} + \frac{1.4^2}{20}} \text{ (df = 19)}$$

c. Magnets do not appear to be effective in treating back pain. It is valid to argue that the magnets might appear to be effective if the sample sizes were larger.

9. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 > \mu_2$; population$_1$ = king-size, population$_2$ = 100-mm;
Test statistic: $t = 8.075$; P-value = 0.0000 (Table: P-value < 0.005); Critical value: $t = 1.678$ (Table: $t = 1.711$); Reject H_0. There is sufficient evidence support the claim that the unfiltered king-size cigarettes have a mean tar content greater than that of 100-mm cigarettes. The result suggests that the filters are effective in reducing tar content, assuming that both types of cigarettes are about the same size.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(21.1 - 13.2) - 0}{\sqrt{\frac{3.2^2}{25} + \frac{3.7^2}{25}}} = 8.075 \text{ (df = 24)}$$

b. 90% CI: 6.3 mg < $\mu_1 - \mu_2$ < 9.5 mg (Table: 6.2 mg < $\mu_1 - \mu_2$ < 9.6 mg); The confidence interval includes positive numbers only, which suggests that the mean tar content of unfiltered king-size cigarettes is greater than the mean for 100-mm cigarettes.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (21.1 - 13.2) \pm 1.711\sqrt{\frac{3.2^2}{25} + \frac{3.7^2}{25}} \text{ (df = 24)}$$

10. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = paroxetine, population$_2$ = placebo;
Test statistic: $t = -1.334$; P-value = 0.1866 (Table: 0.10 < P-value < 0.20); Critical values: $t = \pm 2.159$ (Table: $t = \pm 2.023$); Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim the paroxetine and placebo groups have the same mean depression scale scores. Paroxetine does not appear to be an effective treatment for depression.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(20.38 - 21.57) - 0}{\sqrt{\frac{3.91^2}{33} + \frac{3.87^2}{43}}} = -1.334 \text{ (df = 42)}$$

11. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = treatment, population$_2$ = placebo;
Test statistic: $t = 5.045$; P-value = 0.0000 (Table: P-value < 0.01); Critical values: $t = \pm 2.058$ (Table: $t = \pm 2.080$); Reject H_0. There is sufficient evidence to warrant rejection of the claim that there is a significant different between the treatment and control groups. We cannot conclude that the *cause* is due to the treatment.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(4.20 - 1.71) - 0}{\sqrt{\frac{2.20^2}{22} + \frac{0.72^2}{22}}} = 5.045 \text{ (df = 21)}$$

b. 95% CI: $1.47 < \mu_1 - \mu_2 < 3.51$ (Table: $1.46 < \mu_1 - \mu_2 < 3.52$); Because the interval does not include 0, there appears to be a significant difference between the two population means. It does appear that there are significantly more errors made by those treated with alcohol.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (4.20 - 7.71) \pm 2.080\sqrt{\frac{2.20^2}{22} + \frac{0.72^2}{22}} \text{ (df = 21)}$$

12. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 < \mu_2$; population$_1$ = heavy users, population$_2$ = light users;
Test statistic: $t = -2.790$; P-value = 0.0031 (Table: P-value < 0.005); Critical value: $t = -2.357$ (Table: $t = -2.390$); Reject H_0. There is sufficient evidence to support the claim that the heavy users of marijuana have a lower mean of items sorted than light marijuana users.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(51.3 - 53.3) - 0}{\sqrt{\frac{4.5^2}{65} + \frac{3.6^2}{64}}} = -2.790 \text{ (df = 63)}$$

12. (continued)

b. 98% CI: $-3.96 < \mu_1 - \mu_2 < -0.310$ (Table: $-3.713 < \mu_1 - \mu_2 < -0.286$); Because the confidence interval includes only negative numbers, there is sufficient evidence to support the claim that the heavy users of marijuana have a lower mean of items sorted than light marijuana users.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (51.3 - 53.3) \pm 2.390\sqrt{\frac{4.5^2}{65} + \frac{3.6^2}{64}} \text{ (df = 63)}$$

13. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 > \mu_2$; population$_1$ = exposed, population$_2$ = not exposed; Test statistic: $t = 1.845$; P-value = 0.0351 (Table: P-value < 0.05); Critical value: $t = 1.673$ (Table: $t = 1.685$); Reject H_0. There is sufficient evidence to support the claim that nonsmokers exposed to tobacco smoke have a higher mean cotinine level than nonsmokers not exposed to tobacco smoke.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(60.58 - 16.35) - 0}{\sqrt{\frac{138.08^2}{40} + \frac{62.53^2}{40}}} = 1.845 \text{ (df = 39)}$$

b. 90% CI: 4.12 ng/mL $< \mu_1 - \mu_2 <$ 84.34 ng/mL (Table: 3.85 ng/mL $< \mu_1 - \mu_2 <$ 84.61 ng/mL);

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (60.58 - 16.35) \pm 1.685\sqrt{\frac{138.08^2}{40} + \frac{62.53^2}{40}} \text{ (df = 39)}$$

c. Exposure to second-hand smoke appears to have the effect of being associated with greater amounts of nicotine than for those not exposed to secondhand smoke.

14. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = female, population$_2$ = male; Test statistic: $t = 0.674$; P-value = 0.5015 (Table: P-value > 0.20); Critical values: $t = \pm 1.979$ (Table: $t = \pm 1.995$); Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that females and males have the same mean BMI.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(29.10 - 28.38) - 0}{\sqrt{\frac{7.39^2}{70} + \frac{5.37^2}{80}}} = 0.647 \text{ (df = 69)}$$

b. 95% CI: $-1.39 < \mu_1 - \mu_2 < 2.83$ (Table: $-1.41 < \mu_1 - \mu_2 < 2.85$); Because the confidence interval includes 0, there is not sufficient evidence to warrant rejection of the claim that the two samples are from populations with the same mean.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (29.10 - 28.38) \pm 1.995\sqrt{\frac{7.39^2}{70} + \frac{5.37^2}{80}} \text{ (df = 69)}$$

c. Based on the available sample data, it appears that males and females have the same mean BMI, but we can only conclude that there isn't sufficient evidence to say that they are different.

15. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 > \mu_2$; population$_1$ = low lead, population$_2$ = high lead; Test statistic: $t = 2.282$; P-value = 0.0132 (Table: P-value < 0.05); Critical value: $t = 1.673$ (Table: $t = 1.725$); Reject H_0. There is sufficient evidence to support the claim that the mean IQ score of people with low blood lead levels is higher than the mean IQ score of people with high blood lead levels.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(92.88462 - 86.90476) - 0}{\sqrt{\frac{15.34451^2}{78} + \frac{8.988352^2}{21}}} = 2.282 \text{ (df = 20)}$$

15. (continued)

b. 90% CI: $1.6 < \mu_1 - \mu_2 < 10.4$ (Table: $1.5 < \mu_1 - \mu_2 < 10.5$);

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \left(92.88462 - 86.90476\right) \pm 1.725\sqrt{\frac{15.34451^2}{78} + \frac{8.988352^2}{21}} \text{ (df} = 20)$$

c. Yes, it does appear that exposure to lead has an effect on IQ scores.

16. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 < \mu_2$; population$_1$ = wearing seatbelts, population$_2$ = not wearing seatbelts; Test statistic: $t = -2.330$ P-value = 0.0102 (Table: P-value < 0.025); Critical value: $t = -1.649$; (Table: $t = -1.660$); Reject H_0. There is sufficient evidence to support the claim that children wearing seat belts have a lower mean length of time in an ICU than do children not wearing seat belts.

$$t = \frac{\left(\bar{x}_1 - \bar{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{\left(0.83 - 1.39\right) - 0}{\sqrt{\frac{1.77^2}{123} + \frac{3.06^2}{290}}} = -2.330 \text{ (df} = 122)$$

b. 90% CI: $-0.96 \text{ day} < \mu_1 - \mu_2 < -0.16 \text{ day}$;

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \left(0.83 - 1.39\right) \pm 1.657\sqrt{\frac{1.77^2}{123} + \frac{3.06^2}{290}} \text{ (df} = 122)$$

c. It appears that after motor vehicle crashes, children wearing seat belts spend less time in intensive care units than children who don't wear seat belts. Children should wear seat belts (except for young children who should use properly installed car seats)!

17. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = Disney movies, population$_2$ = other movies; Test statistic: $t = 0.462$; P-value = 0.6465 (Table: P-value > 0.20); Critical values: $t = \pm 2.012$ (Table: $t = \pm 2.120$); Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that Disney animated children's movies and other animated children's movies have the same mean time showing tobacco use.

$$t = \frac{\left(\bar{x}_1 - \bar{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{\left(61.6 - 49.3\right) - 0}{\sqrt{\frac{118.8^2}{33} + \frac{69.3^2}{17}}} = 0.462 \text{ (df} = 16)$$

b. 95% CI: $-41.3 \text{ sec} < \mu_1 - \mu_2 < 65.9 \text{ sec}$ (Table: $-44.2 \text{ sec} < \mu_1 - \mu_2 < 68.8 \text{ sec}$)

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \left(61.6 - 49.3\right) \pm 2.120\sqrt{\frac{118.8^2}{33} + \frac{69.3^2}{17}} \text{ (df} = 16)$$

c. The times appear to be from a population with a distribution that is not normal (the sample is right-skewed), but the methods in this section are robust against departures from normality. (Results obtained by using other methods confirm that the results obtained here are quite good, even though the non-Disney times appear to violate the normality requirement.)

18. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 > \mu_2$; population$_1$ = Pennsylvania, population$_2$ = New York; Test statistic: $t = 3.265$; P-value = 0.0024 (Table: P-value > 0.005); Critical value: $t = 1.746$ (Table: $t = 1.796$); Reject H_0. There is sufficient evidence to support the claim that the mean amount of strontium-90 from Pennsylvania residents is greater than the mean amount from New York residents.

$$t = \frac{\left(\bar{x}_1 - \bar{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{\left(147.6 - 136.4\right) - 0}{\sqrt{\frac{10.6^2}{12} + \frac{5.2^2}{12}}} = 3.265 \text{ (df} = 11)$$

18. (continued)

b. 95% CI: $3.917\text{ mBq/g} < \mu_1 - \mu_2 < 18.417\text{ mBq/g}$

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (147.6 - 136.4) \pm 1.796\sqrt{\frac{10.6^2}{12} + \frac{5.2^2}{12}} \text{ (df} = 11)$$

19. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 < \mu_2$; population$_1$ = popes, population$_2$ = monarchs; Test statistic: $t = -1.810$; P-value $= 0.0442$ (Table: P-value > 0.025); Critical value: $t = -2.574$ (Table: $t = -2.650$); Fail to reject H_0. There is not sufficient evidence to support the claim that the mean longevity for popes is less than the mean for British monarchs after coronation

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(13.1 - 11.7) - 0}{\sqrt{\frac{8.96^2}{24} + \frac{18.60^2}{14}}} = -1.810 \text{ (df} = 13)$$

b. 98% CI: $-23.6\text{ years} < \mu_1 - \mu_2 < 4.0\text{ years}$ (Table: $-23.6\text{ years} < \mu_1 - \mu_2 < 4.4\text{ years}$)

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (13.1 - 11.7) \pm 1.796\sqrt{\frac{8.96^2}{24} + \frac{18.60^2}{14}} \text{ (df} = 13)$$

20. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = easy to difficult, population$_2$ = difficult to easy; Test statistic: $t = -2.657$; P-value $= 0.0114$ (Table: P-value < 0.02); Critical values ($\alpha = 0.05$): $t = \pm 2.023$ (Table: $t = \pm 2.131$); The conclusion depends on the choice of the significance level. There is a significant difference between the two population means at the 0.05 significance level, but not at the 0.01 significance level.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(27.115 - 31.728) - 0}{\sqrt{\frac{6.857^2}{25} + \frac{4.260^2}{16}}} = -2.657 \text{ (df} = 15)$$

21. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = female, population$_2$ = male; Test statistic: $t = -0.863$; P-value $= 0.3887$ (Table: P-value > 0.20); Critical values: $t = \pm 1.968$ (Table: $t = \pm 1.984$); Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that women and men have the same mean diastolic blood pressure.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(70.16 - 71.32) - 0}{\sqrt{\frac{11.22^2}{147} + \frac{11.99^2}{153}}} = -0.863 \text{ (df} = 146)$$

22. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 < \mu_2$; population$_1$ = girls, population$_2$ = boys; Test statistic: $t = -3.450$; P-value $= 0.0003$ (Table: P-value < 0.005); Critical value ($\alpha = 0.05$): $t = -1.649$; (Table: $t = -1.653$); Reject H_0. There is sufficient evidence to support the claim that at birth, girls have a lower mean weight than boys.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(3037.07 - 3272.82) - 0}{\sqrt{\frac{706.268^2}{205} + \frac{660.154^2}{195}}} = -3.450 \text{ (df} = 194)$$

23. a. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 > \mu_2$; population$_1$ = low lead, population$_2$ = high lead;
Test statistic: $t = 1.705$; P-value = 0.0457 (Table: P-value < 0.05); Critical value: $t = 1.661$ (Table: $t = 1.987$); df = $78 + 21 - 2 = 977$; Reject H_0. There is sufficient evidence to support the claim that the mean IQ score of people with low blood lead levels is higher than the mean IQ score of people with high blood lead levels.

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1)+(n_2-1)} = \frac{(78-1)15.34451^2 + (21-1)8.988352^2}{(78-1)+(21-1)} = 203.564;$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = \frac{(92.88462 - 86.90476) - 0}{\sqrt{\frac{203.564}{78} + \frac{203.564}{21}}} = 1.705 \text{ (df = 97)}$$

b. 90% CI: $0.15 < \mu_1 - \mu_2 < 11.81$;

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (92.88462 - 86.90476) \pm 1.661\sqrt{\frac{203.564}{78} + \frac{203.564}{21}} \text{ (df = 97)}$$

c. Yes, it does appear that exposure to lead has an effect on IQ scores. With pooling, df increases dramatically to 97, but the test statistic decreases from 2.282 to 1.705 (because the estimated standard deviation increases from 2.620268 to 3.507614), the P-value increases to 0.0457, and the 90% confidence interval becomes wider. With pooling, these results do not show greater significance.

24. df = 38.9884; Using "df = smaller of $n_1 - 1$ and $n_2 - 1$" is a more conservative estimate of the number of degrees of freedom (than the estimate obtained with Formula 9-1) in the sense that the confidence interval is wider, so the difference between the sample means needs to be more extreme to be considered a significant difference.

$$\text{df} = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1} + \frac{\left(s_2^2/n_2\right)^2}{n_2-1}} = \frac{\left(6.857^2/25 + 4.260^2/16\right)^2}{\frac{\left(6.857^2/25\right)^2}{25-1} + \frac{\left(4.260^2/16\right)^2}{16-1}} = 38.9884$$

25. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = treatment, population$_2$ = placebo;
Test statistic: $t = 15.322$; P-value = 0.0000 (Table: P-value < 0.01); Critical values: $t = \pm 2.080$; Reject H_0. There is sufficient evidence to warrant rejection of the claim that the two populations have the same mean.

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1)+(n_2-1)} = \frac{(22-1)0.015^2 + (22-1)0^2}{(22-1)+(22-1)} = 0.000125;$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = \frac{(0.049 - 0.000) - 0}{\sqrt{\frac{0.000125}{22} + \frac{0.000125}{22}}} = 15.322$$

Section 9-3: Two Dependent Samples (Matched Pairs)

1. Only parts (a) and (c) are true.

2. $\bar{d} = -0.28$°F and $s_d = 0.36$°F; μ_d represents the mean of the differences from the population of paired data.

Temperature (°F) at 8 am	97.8	99.0	97.4	97.4	97.5
Temperature (°F) at 12 am	98.6	99.5	97.5	97.3	97.6
Difference (°F)	–0.8	–0.5	–0.1	0.1	–0.1

3. The results will be the same.

4. df = $n - 1 = 5 - 1 = 4$; $t_{\alpha/2} = 4.604$

5. H_0: $\mu_d = 0$; H_1: $\mu_d \neq 0$; difference = right − left;
Test statistic: $t = -17.339$; P-value = 0.0001 (Table: P-value < 0.01); Critical values: $t = \pm 4.604$; Reject H_0. There is sufficient evidence to support the claim of a difference in measurements between the two arms. The right and left arms should yield the same measurements, but the given data show that this is not happening for this person.
$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} = \frac{-72.2 - 0}{9.31128/\sqrt{5}} = -2.609 \text{ (df = 4)}$$

6. a. H_0: $\mu_d = 0$ cm; H_1: $\mu_d > 0$ cm; difference = president − opponent;
Test statistic: $t = 1.304$; P-value = 0.1246 (Table: P-value > 0.10); Critical value: $t = 2.105$; Fail to reject H_0. There is not sufficient evidence to support the claim that for the population of heights of presidents and their main opponents, the differences have a mean greater than 0 cm (or that presidents tend to be taller than their opponents).
$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} = \frac{3.67 - 0}{6.890/\sqrt{6}} = 1.304 \text{ (df = 5)}$$
b. 90% CI: $-2.0 \text{ cm} < \mu_d < 9.3 \text{ cm}$; The confidence interval includes 0, so it is possible that $\mu_d = 0$.
$$\bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}} = 3.67 \pm 2.015\frac{6.890}{\sqrt{6}} \text{ (df = 5)}$$

7. a. H_0: $\mu_d = 0°\text{F}$; H_1: $\mu_d \neq 0°\text{F}$; difference = 8 AM − 12 AM;
Test statistic: $t = -7.499$; P-value = 0.0003 (Table: P-value < 0.01); Critical values: $t = \pm 2.447$; Reject H_0. There is sufficient evidence to warrant rejection of the claim that there is no difference between body temperatures measured at 8 AM and at 12 AM. There appears to be a difference.
$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} = \frac{-1.49 - 0}{0.524/\sqrt{7}} = -7.499 \text{ (df = 6)}$$
b. 95% CI: $-1.97°\text{F} < \mu_d < -1.00°\text{F}$; The confidence interval consists of negative numbers only and does not include 0.
$$\bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}} = -1.49 \pm 2.447\frac{0.524}{\sqrt{7}} \text{ (df = 6)}$$

8. a. H_0: $\mu_d = 0$ words; H_1: $\mu_d < 0$ words; difference = male − female;
Test statistic: $t = -0.472$; P-value = 0.3265 (Table: P-value > 0.10); Critical value: $t = -1.895$; Fail to reject H_0. There is not sufficient evidence to support the claim that among couples, males speak fewer words in a day than females.
$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} = \frac{-1677.75 - 0}{10{,}052.87/\sqrt{8}} = -0.472 \text{ (df = 7)}$$
b. 90% CI: $-8411.5 \text{ words} < \mu_d < 5056.0 \text{ words}$ (Table: $-8413.0 \text{ words} < \mu_d < 5057.5 \text{ words}$); The confidence interval includes 0 word, so it is possible that $\mu_d = 0$ word.
$$\bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}} = -1677.75 \pm 1.895\frac{10{,}052.87}{\sqrt{8}} \text{ (df = 7)}$$

9. H_0: $\mu_d = 0$ in.; H_1: $\mu_d \neq 0$ in.; difference = mother − daughter;
Test statistic: $t = -1.379$; P-value = 0.2013 (Table: P-value > 0.20); Critical values: $t = \pm 2.262$; Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that there is no difference in heights between mothers and their first daughters.

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-0.95 - 0}{2.179/\sqrt{10}} = -1.379 \text{ (df = 9)}$$

10. H_0: $\mu_d = 0$ in.; H_1: $\mu_d \neq 0$ in.; difference = father − son;
Test statistic: $t = 0.034$; P-value = 0.9737 (Table: P-value > 0.20); Critical values: $t = \pm 2.262$; Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that there is no difference in heights between fathers and their first sons.

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{0.02 - 0}{1.863/\sqrt{10}} = 0.034 \text{ (df = 9)}$$

11. 95% CI: -6.5 admissions $< \mu_d < -0.2$ admissions; Because the confidence interval does not include 0 admission, it appears that there is sufficient evidence to warrant rejection of the claim that when the 13th day of a month falls on a Friday, the numbers of hospital admissions from motor vehicle crashes are not affected. Hospital admissions do appear to be affected.

$$\bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}} = -3.333 \pm 2.571\frac{3.011}{\sqrt{6}} \text{ (difference = 6th − 13th, df = 5)}$$

12. H_0: $\mu_d = 0$; H_1: $\mu_d > 0$; difference = before − after;
Test statistic: $t = 6.371$; P-value = 0.0000 (Table: P-value < 0.005); Critical value: $t = 2.718$; Reject H_0. There is sufficient evidence to support the claim that captopril is effective in lowering systolic blood pressure.

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{18.58 - 0}{10.103/\sqrt{12}} = 6.371 \text{ (df = 11)}$$

13. 99% CI: $-66.7 \text{ cm}^3 < \mu_d < 49.7 \text{ cm}^3$ (Table: $-66.8 \text{ cm}^3 < \mu_d < 49.8 \text{ cm}^3$); Because the confidence interval includes the mean of the differences could be equal to 0 cm^3, so there does not appear to be a significant difference.

$$\bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}} = -8.5 \pm 3.249\frac{56.679}{\sqrt{10}} \text{ (difference = first born − second born, df = 9)}$$

14. 95% CI: $0.69 < \mu_d < 5.66$; Because the confidence interval limits do not contain 0 and they consist of positive values only, it appears that the “before” measurements are greater than the “after” measurements, so hypnotism does appear to be effective in reducing pain.

$$\bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}} = 3.125 \pm 2.365\frac{2.911}{\sqrt{8}} \text{ (difference = before − after, df = 7)}$$

15. 99% CI: -4.16 in. $< \mu_d < 2.16$ in.; Because the confidence interval limits contain 0, there is not sufficient evidence to support a claim that there is a difference between self-reported heights and measured heights. We might believe that males would tend to exaggerate their heights, but the given data do not provide enough evidence to support that belief.

$$\bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}} = -1.0 \pm 3.106\frac{3.520}{\sqrt{12}} \text{ (difference = reported − measured, df = 11)}$$

16. a. H_0: $\mu_d = 0$ cm; H_1: $\mu_d > 0$ cm; difference = regular − kiln dried;
Test statistic: $t = -1.532$; P-value = 0.1565 (Table: P-value > 0.10); Critical values: $t = \pm 2.228$; Fail to reject H_0. There is not sufficient evidence to reject the claim that the mean yields for the two types of seed are different.

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-1.091 - 0}{2.362/\sqrt{11}} = -1.532 \text{ (df} = 10)$$

b. 90% CI: $-2.678 \text{ cwt/acre} < \mu_d < 0.496 \text{ cwt/acre}$;

$$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}} = -1.091 \pm 2.228 \frac{2.362}{\sqrt{11}} \text{ (df} = 10)$$

c. It does not appear that either type of seed is better.

17. a. H_0: $\mu_d = 0°\text{F}$; H_1: $\mu_d \neq 0°\text{F}$; difference = 8 AM − 12 AM;
Test statistic: $t = -8.485$; P-value = 0.0000 (Table: P-value < 0.01); Critical values: $t = \pm 1.996$ (Table: $t = \pm 1.994$); Reject H_0. There is sufficient evidence to warrant rejection of the claim of no difference between body temperatures measured at 8 am and at 12 am. There appears to be a difference.

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-0.85 + 0}{0.833/\sqrt{69}} = -8.485 \text{ (df} = 68)$$

b. 95% CI: $-1.05°\text{F} < \mu_d < -0.65°\text{F}$; The confidence interval consists of negative numbers only and does not include 0.

$$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}} = -0.85 \pm 1.994 \frac{0.833}{\sqrt{69}} \text{ (df} = 68)$$

18. H_0: $\mu_d = 0$ in.; H_1: $\mu_d \neq 0$ in.; difference = mother − daughter;
Test statistic: $t = -4.090$; P-value = 0.0001 (Table: P-value < 0.01); Critical values: $t = \pm 1.978$ (Table: $t \approx \pm 1.974$); Reject H_0. There is sufficient evidence to warrant rejection of the claim of no difference in heights between mothers and their first daughters.

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-0.93 - 0}{2.636/\sqrt{134}} = -4.090 \text{ (df} = 133)$$

19. H_0: $\mu_d = 0$ in.; H_1: $\mu_d \neq 0$ in.; difference = father − son;
Test statistic: $t = -6.347$; P-value = 0.0000 (Table: P-value < 0.01); Critical values: $t = \pm 1.978$ (Table: $t \approx \pm 1.984$); Reject H_0. There is sufficient evidence to warrant rejection of the claim of no difference in heights between fathers and their first sons.

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-1.366 - 0}{2.491/\sqrt{134}} = -6.347 \text{ (df} = 133)$$

20. H_0: $\mu_d = 0$; H_1: $\mu_d > 0$; difference = tobacco − alcohol;
Test statistic: $t = 1.617$; P-value = 0.0561 (Table: $0.05 < P\text{-value} < 0.10$); Critical value: 1.677; Fail to reject H_0. There is not sufficient evidence to support the claim that more movie time is devoted to tobacco use than alcohol use.

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{24.98 - 0}{109.219/\sqrt{50}} = 1.617 \text{ (df} = 49)$$

21. For the temperatures in degrees Fahrenheit and the temperatures in degrees Celsius, the test statistic of $t = 0.124$ is the same, the P-value of 0.9023 is the same, the critical values of $t = \pm 2.028$ are the same, and the conclusions are the same, so the hypothesis test results are the same in both cases. The confidence intervals are $-0.25°\text{F} < \mu_d < 0.28°\text{F}$ and $-0.14°\text{C} < \mu_d < 0.16°\text{C}$. The confidence interval limits of $-0.14°\text{C}$ and $0.16°\text{C}$ have numerical values that are $5/9$ of the numerical values of $-0.25°\text{F}$ and $0.28°\text{F}$.

22. a. No, it is possible that values of x are not normally distributed and values of y are not normally distributed, but values of $x - y$ are normally distributed.

 b. Answers vary, but here is a typical result: 95% CI: $-0.24°\text{F} < \mu_d < 0.26°\text{F}$; which is reasonably close to the confidence interval found in Exercise 25.

Section 9-4: Two Variances or Standard Deviations

1. a. No, the numerator will always be larger than the denominator in the fraction.

 b. No, both variances are nonnegative, so their quotient cannot be negative.

 c. The two samples have standard deviations (or variances) that are very close in value.

 d. skewed right

2. a. $s_1^2 = 55.99469 \text{ cm}^2$ and $s_2^2 = 50.42392 \text{ cm}^2$

 b. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 \neq \sigma_2$; population$_1$ = women, population$_2$ = men

 c. $F = s_1^2 / s_2^2 = 7.48296^2 / 7.10098^2 = 1.1105$

 d. Fail to reject H_0. There is not sufficient evidence to support the claim that heights of men and heights of women have different variances.

3. No, unlike some other tests that have a requirement that samples must be from normally distributed populations or the samples must have more than 30 values, the F test has a requirement that the samples must be from normally distributed populations, regardless of how large the samples are.

4. The F test is very sensitive to departures from normality, which means that it works poorly by leading to wrong conclusions when either or both of the populations have a distribution that is not normal.

5. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 > \sigma_2$; population$_1$ = treatment, population$_2$ = placebo; Test statistic: $F = s_1^2 / s_2^2 = 2.20^2 / 0.72^2 = 9.3364$; P-value = 0.0000; Critical value: $F = 2.0842$ (Table: $2.0540 < F < 2.0960$); Reject H_0. There is sufficient evidence to support the claim that the treatment group has errors that vary more than the errors of the placebo group.

6. a. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 > \sigma_2$; population$_1$ = exposed, population$_2$ = not exposed; Test statistic: $F = s_1^2 / s_2^2 = 119.50^2 / 62.53^2 = 3.6552$; P-value = 0.0000; Critical value: $F = 1.7045$ (Table: $1.8752 < F < 2.0739$); Fail to reject H_0. There is sufficient evidence to support the claim that the variation of cotinine in smokers is greater than the variation of cotinine in nonsmokers not exposed to tobacco smoke.

 b. The sample is not from a normally distributed population as required, so the results in part (a) are highly questionable.

7. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 \neq \sigma_2$; population$_1$ = Atkins, population$_2$ = Zone; Test statistic: $F = s_1^2 / s_2^2 = 12^2 / 9^2 = 1.7778$; P-value = 0.0762; Upper critical value: $F = 1.8907$ (Table: $1.8752 < F < 2.0739$); Reject H_0. There is not sufficient evidence to warrant rejection of the claim that subjects from both treatment groups have ages with the same amount of variation. If treatment groups have different characteristics, comparisons of treatments become unclear, because differences might be due to the treatments or they might be due to the different group characteristics.

8. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 > \sigma_2$; population$_1$ = low lead, population$_2$ = high lead;
Test statistic: $F = s_1^2/s_2^2 = 15.34451^2/8.988352^2 = 2.9144$; P-value = 0.0045; Critical value: $F = 1.9246$ (Table: $1.8963 < F < 1.9464$); Reject H_0. There is sufficient evidence to support the claim that IQ scores of people with low lead levels vary more than IQ scores of people with high lead levels.

9. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 > \sigma_2$; population$_1$ = sham treatment, population$_2$ = magnet treatment;
Test statistic: $F = s_1^2/s_2^2 = 1.4^2/0.96^2 = 2.1267$; P-value = 0.0543; Critical value: $F = 2.1682$ (Table: $2.1555 < F < 2.2341$); Fail to reject H_0. There is not sufficient evidence to support the claim that those given a sham treatment have pain reductions that vary more than the pain reductions for those treated with magnets.

10. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 \neq \sigma_2$; population$_1$ = low humidity, population$_2$ = high humidity;
Test statistic: $F = s_1^2/s_2^2 = 1.22^2/1.11^2 = 1.2080$; P-value = 0.5288; Upper critical value: $F = 2.4300$ (Table: $1.6668 < F < 1.8752$); Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the low humidity and high humidity groups have the same variation.

11. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 > \sigma_2$; population$_1$ = filtered, population$_2$ = non-filtered;
Test statistic: $F = s_1^2/s_2^2 = 0.31^2/0.16^2 = 3.7539$; P-value = 0.0400; Critical value: $F = 3.4445$; Reject H_0. There is sufficient evidence to support the claim that king-size cigarettes with filters have amounts of nicotine that vary more than the amounts of nicotine in non-filtered king-size cigarettes.

12. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 \neq \sigma_2$; population$_1$ = placebo, population$_2$ = zinc;
Test statistic: $F = s_1^2/s_2^2 = 728^2/669^2 = 1.1842$; P-value = 0.3738; Upper critical value: $F = 2.8621$ (Table: $2.7559 < F < 2.8621$); Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the treatment group and placebo group have the same variation.

13. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 > \sigma_2$; population$_1$ = men, population$_2$ = women;
Test statistic: $F = s_1^2/s_2^2 = 0.89^2/0.66^2 = 1.8184$; P-value = 0.0774; Critical value: $F = 1.9983$ (Table: $1.9926 < F < 2.0772$); Fail to reject H_0. There is not sufficient evidence to support the claim that men have body temperatures that vary more than the body temperatures of women.

14. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 > \sigma_2$; population$_1$ = Pennsylvania, population$_2$ = New York;
Test statistic: $F = s_1^2/s_2^2 = 10.638^2/0.5.213^2 = 4.1643$ P-value = 0.0130; Critical value: $F = 2.8179$ (Table: $2.7876 < F < 2.8536$); Reject H_0. There is sufficient evidence to support the claim that the amount of strontium-90 for Pennsylvania residents has greater variation than the amount of strontium-90 for New York residents.

15. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 \neq \sigma_2$; population$_1$ = monarchs, population$_2$ = popes;
Test statistic: $F = s_1^2/s_2^2 = 18.603^2/8.960^2 = 4.3103$; P-value = 0.0023; Upper critical value: $F = 2.5308$ (Table: $2.4665 < F < 2.5699$); Reject H_0. There is sufficient evidence to warrant rejection of the claim that both populations of longevity times have the same variation.

16. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 \neq \sigma_2$; population$_1$ = easy to difficult, population$_2$ = difficult to easy;
Test statistic: $F = s_1^2/s_2^2 = 6.857^2/4.260^2 = 2.5908$; P-value = 0.0599; Upper critical value: $F = 2.7006$; Fail to reject H_0. There is not sufficient evidence to support a claim that the two populations of scores have different amounts of variation.

17. a. Calculations not shown.

    b. $c_1 = 3, c_2 = 0$

    c. Critical value $= \dfrac{\log(0.05/2)}{\log\left(\dfrac{25}{25+16}\right)} = 7.4569$

    d. $c_1 = 3 < 7.4569$; Fail to reject H_0. There is not sufficient evidence to support a claim that the two populations of scores have different amounts of variation.

18. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = easy to difficult, population$_2$ = difficult to easy; Test statistic: $t = 1.403$; P-value = 0.1686 (Table: P-value > 0.10); Critical values: $t = \pm 2.024$ (Table: $t = \pm 2.131$); Fail to reject H_0. There is not sufficient evidence to support a claim that the two populations of scores have different amounts of variation.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{(4.914 - 3.296) - 0}{\sqrt{\dfrac{22.715^2}{25} + \dfrac{6.759^2}{16}}} = 1.403 \text{ (df} = 15)$$

19. $F_L = \dfrac{1}{2.4374} = 0.4103, F_R = 2.7006$

Chapter Quick Quiz

1. H_0: $p_1 = p_2$; H_1: $p_1 \neq p_2$; population$_1$ = women, population$_2$ = men;

2. $x_1 = 258, x_2 = 282, \hat{p}_1 = \dfrac{258}{1121} = 0.230, \hat{p}_2 = \dfrac{282}{1084} = 0.260, \overline{p} = \dfrac{258+282}{1121+1084} = 0.245$

3. P-value = 0.1015 (Table: 0.1010)

4. a. 95% CI: $-0.0659 < p_1 - p_2 < 0.00591$

   b. The confidence interval includes the value of 0, so it is possible that the two proportions are equal. There is not a significant difference.

5. Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that for the people who were aware of the statement, the proportion of women is equal to the proportion of men.

6. True, since $n > 30$.

7. False, the requirements are $np \geq 5$ and $nq \geq 5$.

8. Because the data consist of matched pairs, they are dependent.

9. H_0: $\mu_d = 0$; H_1: $\mu_d \neq 0$

10. a. $t = \dfrac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$

    b. $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

    c. $z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\overline{p}\overline{q}}{n_1} + \dfrac{\overline{p}\overline{q}}{n_2}}}$

    d. $F = \dfrac{s_1^2}{s_2^2}$

Review Exercises

1. H_0: $p_1 = p_2$; H_1: $p_1 \neq p_2$; population$_1$ = blind, population$_2$ = not blind;
Test statistic: $z = -4.20$; P-value = 0.0000 (Table: 0.0002); Critical values: $z = \pm 2.576$; Reject H_0. There is sufficient evidence to warrant rejection of the claim that the acceptance rate is the same with or without blinding. Without blinding, reviewers know the names and institutions of the abstract authors, and they might be influenced by that knowledge.

$$\bar{p} = \frac{13,200(0.267)+13,433(0.290)}{13,200+13,433} = 0.637;\ \bar{q} = 1-0.637 = 0.363;$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2)-(p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1}+\frac{\bar{p}\bar{q}}{n_2}}} = \frac{(0.267-0.290)-0}{\sqrt{\frac{(0.637)(0.363)}{13,200}+\frac{(0.637)(0.363)}{13,433}}} = -4.20$$

2. 98% CI: $-0.0372 < p_1 - p_2 < -0.00892$; The confidence interval limits do not contain 0, so it appears that there is a significant difference between the two proportions.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}} = (0.267-0.290) \pm 2.576\sqrt{\frac{0.267(0.733)}{13,200}+\frac{(0.290)(0.710)}{13,433}}$$

3. 95% CI: $-25.33\text{ cm} < \mu_1 - \mu_2 < -7.51\text{ cm}$ (Table: $-25.70\text{ cm} < \mu_1 - \mu_2 < -7.14\text{ cm}$); With 95% confidence, we conclude that the mean height of women is less than the mean height of men by an amount that is between 7.51 cm and 25.33 cm (Table: 7.14 cm and 25.70 cm).

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} = (162.35-178.77) \pm 2.262\sqrt{\frac{11.847^2}{10}+\frac{5.302^2}{10}}\ \text{(df = 9)};$$

population$_1$ = women, population$_2$ = men

4. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 < \mu_2$; population$_1$ = women, population$_2$ = men;
Test statistic: $t = -4.001$; P-value = 0.0008 (Table: P-value < 0.005); Critical value: $t = -2.666$ (Table: $t = -2.821$); Reject H_0. There is sufficient evidence to support the claim that women have heights with a mean that is less than the mean height of men.

$$t = \frac{(\bar{x}_1 - \bar{x}_2)-(\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} = \frac{(162.35-178.77)-0}{\sqrt{\frac{11.847^2}{10}+\frac{5.302^2}{10}}} = -4.001\ \text{(df = 9)}$$

5. a. H_0: $\mu_d = 0$ kg; H_1: $\mu_d \neq 0$ kg; difference = pre − post;
Test statistic: $t = 2.301$; P-value = 0.0469; Critical values $(\alpha = 0.05)$: $t = \pm 2.262$; Reject H_0. There is sufficient evidence to conclude that there is a difference between pre-training and post-training weights.

$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} = \frac{2.0-0}{2.749/\sqrt{10}} = 2.301\ \text{(df = 9)}$$

b. 95% CI: $0.0\text{ kg} < \mu_d < 4.0\text{ kg}$; The confidence interval consists of negative numbers only and does not include 0.

$$\bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}} = 2.0 \pm 2.262\frac{2.749}{\sqrt{10}}\ \text{(df = 9)}$$

6. H_0: $\sigma_1 = \sigma_2$; H_1: $\sigma_1 \neq \sigma_2$; population$_1$ = women, population$_2$ = men;
Test statistic: $F = s_1^2/s_2^2 = 11.847^2/5.302^2 = 4.9933$; P-value = 0.0252; Upper critical value : $F = 4.0260$; Reject H_0. There is sufficient evidence to warrant rejection of the claim that women and men have heights with the same variation.

Cumulative Review Exercises

1. a. Because the sample data are matched with each column consisting of heights from the same family, the data are dependent.

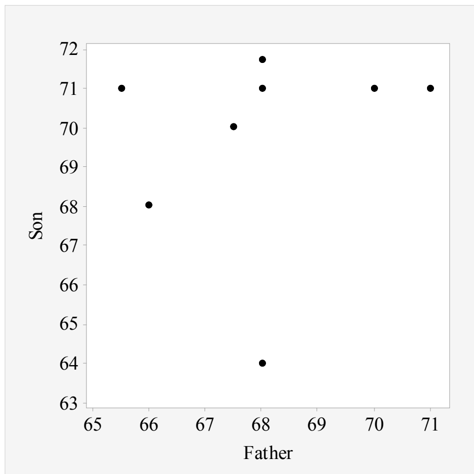
   b. $\overline{x} = \dfrac{64.0+68.0+70.0+71.0+71.0+71.0+71.0+71.7}{8} = 69.7 \text{ in.}; \; Q_2 = \dfrac{71.0+71.0}{2} = 71.0 \text{ in.};$

   $\text{range} = 71.7 - 64.0 = 7.7 \text{ in.}; \; s = \sqrt{\dfrac{(64.0-69.7)^2 + \cdots + (71.7-69.7)^2}{14-1}} = 2.6 \text{ in.}; \; s^2 = 6.6 \text{ in}^2$

   c. ratio

   d. continuous

2. There does not appear to be a correlation or association between the heights of fathers and the heights of their sons.



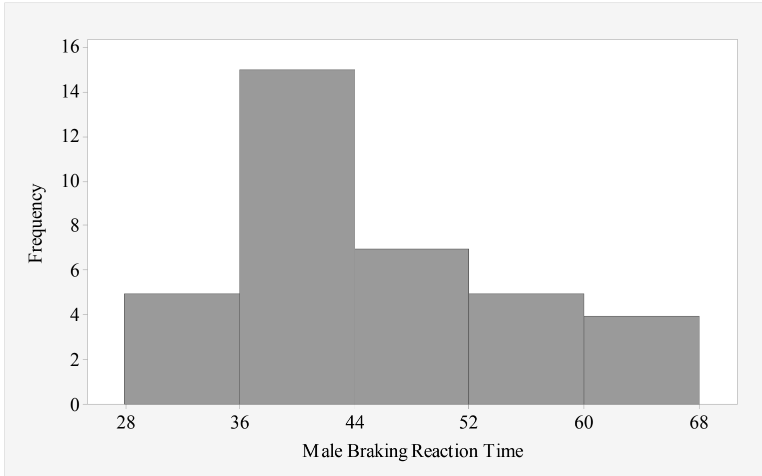
3. 90% CI: $\overline{x} \pm t_{\alpha/2}\dfrac{s}{\sqrt{n}} = 69.7 \pm 2.306 \cdot \dfrac{2.570}{\sqrt{8}} \Rightarrow 67.6 \text{ in.} < \mu < 71.9 \text{ in.};$ We have 95% confidence that the limits of 67.6 in. and 71.9 in. actually contain the true value of the mean height of all adult sons.

4. H_0: $\mu_d = 0$ in.; H_1: $\mu_d \neq 0$ in.; difference = father − son;
   Test statistic: $t = -1.712$; P-value = 0.1326 (Table: P-value > 0.10); Critical values: $t = \pm 2.365$; Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that differences between heights of fathers and their sons have a mean of 0. There does not appear to be a difference between heights of fathers and their sons.

   $$t = \frac{\overline{d} - \mu_d}{s_d/\sqrt{n}} = \frac{-1.7125 - 0}{2.847/\sqrt{8}} = -1.712 \text{ (df = 7)}$$

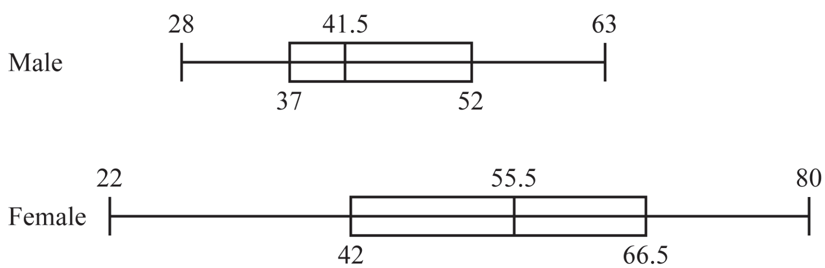
5. Because the points lie reasonably close to a straight-line pattern, and there is no other pattern that is not a straight-line pattern and there are no outliers, the sample data appear to be from a population with a normal distribution.

6. The shape of the histogram indicates that the sample data appear to be from a population with a distribution that is approximately normal.



7. Because the points are reasonably close to a straight-line pattern and there is no other pattern that is not a straight-line pattern, it appears that the braking reaction times of females are from a population with a normal distribution.

8. Because the boxplots overlap, there does not appear to be a significant difference between braking reaction times of males and females, but the braking reaction times for males appear to be generally lower than the braking reaction times of females.



9. H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$; population$_1$ = male, population$_2$ = female;
Test statistic: $t = -3.259$; P-value $= 0.0019$ (Table: P-value < 0.005); Critical values $t = \pm 2.664$ (Table: $t = \pm 2.724$); Reject H_0. There is sufficient evidence to warrant rejection of the claim that males and females have the same mean braking reaction time. Males appear to have lower reaction times.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(44.361 - 54.278) - 0}{\sqrt{\frac{9.472^2}{36} + \frac{15.611^2}{36}}} = -3.259 \text{ (df = 35)}$$

10. a. The sample sizes are greater than 30 and the data appear to be from a populations that meet the loose requirement of being normally distributed.

Males: 99%% CI: $\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} = 44.631 \pm 2.724 \cdot \frac{9.472}{\sqrt{36}} \Rightarrow 40.1 < \mu < 48.7$

Females: 99%% CI: $\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} = 54.278 \pm 2.724 \cdot \frac{15.611}{\sqrt{36}} \Rightarrow 47.2 < \mu < 61.4$

The confidence intervals overlap, so there does not appear to be significant difference between the mean braking reaction times of males and females.

10. (continued)

b. 99% CI: $-18.0 < \mu_1 - \mu_2 < -1.8$ (Table: $—18.2 < \mu_1 - \mu_2 < -1.6$); Because the confidence interval consists of negative numbers and does not include 0, there appears to be a significant difference between the mean braking reaction times of males and females.

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \left(44.361 - 54.278\right) \pm 2.724\sqrt{\frac{9.472^2}{36} + \frac{15.611^2}{36}} \text{ (df = 35);}$$

population$_1$ = men, population$_2$ = women

c. The results from part (b) are better.