Chapter 10

Statistical Inference About Means and Proportions with Two Populations

Learning Objectives

1. Be able to develop interval estimates and conduct hypothesis tests about the difference between two population means when σ_1 and σ_2 are known.

2. Know the properties of the sampling distribution of $\bar{x}_1 - \bar{x}_2$.

3. Be able to use the *t* distribution to conduct statistical inferences about the difference between two population means when σ_1 and σ_2 are unknown.

4. Learn how to analyze the difference between two population means when the samples are independent and when the samples are matched.

5. Be able to develop interval estimates and conduct hypothesis tests about the difference between two population proportions.

6. Know the properties of the sampling distribution of $\bar{p}_1 - \bar{p}_2$.

Solutions

1. a. $\bar{x}_1 - \bar{x}_2$ = 13.6 – 11.6 = 2

b. $z_{\alpha/2} = z_{.05} = 1.645$

$$\bar{x}_1 - \bar{x}_2 \pm 1.645\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$2 \pm 1.645\sqrt{\frac{(2.2)^2}{50} + \frac{(3)^2}{35}}$$

2 ± .98 (1.02 to 2.98)

c. $z_{\alpha/2} = z_{.025} = 1.96$

$$2 \pm 1.96\sqrt{\frac{(2.2)^2}{50} + \frac{(3)^2}{35}}$$

2 ± 1.17 (.83 to 3.17)

2. a. $$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(25.2 - 22.8) - 0}{\sqrt{\frac{(5.2)^2}{40} + \frac{6^2}{50}}} = 2.03$$

b. p-value = 1.0000 – .9788 = .0212

c. p-value $\leq$ .05, reject H_0.

3. a. $$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{(8.4)^2}{80} + \frac{(7.6)^2}{70}}} = -1.53$$

b. p-value = 2(.0630) = .1260

c. p-value > .05, do not reject H_0.

4. a. μ_1 = population mean for smaller cruise ships

μ_2 = population mean for larger cruise ships

$\bar{x}_1 - \bar{x}_2$ = 85.36 – 81.40 = 3.96

b. $$z_{.025}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$1.96\sqrt{\frac{(4.55)^2}{37} + \frac{(3.97)^2}{44}} = 1.88$$

c. 3.96 ± 1.88 (2.08 to 5.84)

5. a. $\bar{x}_1 - \bar{x}_2$ = 135.67 – 68.64 = 67.03

b. $$z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = 2.576\sqrt{\frac{(35)^2}{40} + \frac{(20)^2}{30}} = 17.08$$

c. 67.03 ± 17.08 (49.95 to 84.11)

We estimate that men spend \$67.03 more than women on Valentine's Day with a margin of error of \$17.08.

6. μ_1 = mean hotel price in Atlanta

μ_2 = mean hotel price in Houston

H_0: $\mu_1 - \mu_2 \geq 0$

H_a: $\mu_1 - \mu_2 < 0$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(91.71 - 101.13) - 0}{\sqrt{\frac{20^2}{35} + \frac{25^2}{40}}} = -1.81$$

p-value = .0351

p-value $\leq$ .05; reject H_0. The mean price of a hotel room in Atlanta is lower than the mean price of a hotel room in Houston.

7. a. μ_1 = population mean satisfaction score for Publix customers

μ_2 = population mean satisfaction score for Trader Joe's customers

H_0: $\mu_1 - \mu_2 = 0$

H_a: $\mu_1 - \mu_2 \neq 0$

b. $\bar{x}_1 - \bar{x}_2$ = 86 – 85 = 1

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(86 - 85) - 0}{\sqrt{\frac{12^2}{250} + \frac{12^2}{300}}} = 0.97$$

For this two-tailed test, p-value is two times the upper-tail area at $z = 0.97$.

p-value = 2(1.0000 – .8340) = .3320

p-value > .05; do not reject H_0. There is not sufficient evidence to conclude that the population mean satisfaction scores differ for the two retailers.

c. $\bar{x}_1 - \bar{x}_2 \pm z_{.025}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

$$(86-85) \pm 1.96\sqrt{\frac{12^2}{250}+\frac{12^2}{300}}$$

$$1 \pm 2.01 \qquad -1.01 \text{ to } 3.01)$$

Because the 95% confidence interval contains zero, the difference between the mean customer satisfaction score of Publix and Trader Joe's is not statistically significant at the 95% confidence level.

8. a. This is an upper-tail hypothesis test.

H_0: $\mu_1 - \mu_2 \leq 0$

H_a: $\mu_1 - \mu_2 > 0$

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} = \frac{(76-73)}{\sqrt{\frac{6^2}{60}+\frac{6^2}{60}}} = 2.74$$

p-value = area in upper tail at $z = 2.74$

p-value = 1.0000 - .9969 = .0031

Because .0031 $\leq \alpha = .05$, we reject the null hypothesis. The difference is significant. We can conclude that customer service has improved for Rite Aid.

b. This is another upper-tail test, but it only involves one population.

H_0: $\mu \leq 75.7$

H_a: $\mu > 75.7$

$$z=\frac{(\bar{x}_1-75.7)}{\sqrt{\frac{\sigma^2}{n_1}}}=\frac{(76-75.7)}{\sqrt{\frac{6^2}{60}}}=.39$$

p-value = area in upper tail at z = .39

p-value = 1.0000 – .6517 = .3483

Because .3483 > α = .05, we cannot reject the null hypothesis. The difference is not statistically significant.

c. This is an upper-tail test similar to the one in part a.

H_0: $\mu_1-\mu_2\le 0$

H_a: $\mu_1-\mu_2>0$

$$z=\frac{(\bar{x}_1-\bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}=\frac{(77-75)}{\sqrt{\frac{6^2}{60}+\frac{6^2}{60}}}=1.83$$

p-value = area in upper tail at z = 1.83

p-value = 1.0000 – .9664 = .0336

Because .0336 ≤ α = .05, we reject the null hypothesis. The difference is significant. We can conclude that customer service has improved for Expedia.

d. We will reject the null hypothesis of "no increase" if the p-value ≤ .05. For an upper-tail hypothesis test, the p-value is the area in the upper tail at the value of the test statistic. A value of z = 1.645 provides an upper-tail area of .05. So, we must solve the following equation for $\bar{x}_1-\bar{x}_2$.

$$z=\frac{(\bar{x}_1-\bar{x}_2)}{\sqrt{\frac{6^2}{60}+\frac{6^2}{60}}}=1.645$$

$$\bar{x}_1 - \bar{x}_2 = 1.645\sqrt{\frac{6^2}{60} + \frac{6^2}{60}} = 1.80$$

This tells us that as long as the year 2 score for a company exceeds the year 1 score by 1.80 or more the difference will be statistically significant.

e. The increase from year 1 to year 2 for J.C. Penney is not statistically significant because it is less than 1.80. We cannot conclude that customer service has improved for J.C. Penney.

9. a. $\bar{x}_1 - \bar{x}_2 = 22.5 - 20.1 = 2.4$

b. $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{2.5^2}{20} + \frac{4.8^2}{30}\right)^2}{\frac{1}{19}\left(\frac{2.5^2}{20}\right)^2 + \frac{1}{29}\left(\frac{4.8^2}{30}\right)^2} = 45.8$$

Use $df = 45$.

c. $t_{.025} = 2.014$

$$t_{.025}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2.014\sqrt{\frac{2.5^2}{20} + \frac{4.8^2}{30}} = 2.1$$

d. 2.4 ± 2.1 (.3 to 4.5)

10. a. $$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(13.6 - 10.1) - 0}{\sqrt{\frac{5.2^2}{35} + \frac{8.5^2}{40}}} = 2.18$$

b. $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{5.2^2}{35} + \frac{8.5^2}{40}\right)^2}{\frac{1}{34}\left(\frac{5.2^2}{35}\right)^2 + \frac{1}{39}\left(\frac{8.5^2}{40}\right)^2} = 65.7$$

Use $df = 65$.

c. Using t table, area in tail is between .01 and .025 $\therefore$ two-tail p-value is between .02

and .05.

Exact p-value corresponding to $t = 2.18$ is .0329

d. p-value $\leq$ .05, reject H_0.

11. a. $\bar{x}_1 = \frac{54}{6} = 9 \quad \bar{x}_2 = \frac{42}{6} = 7$

b. $s_1 = \sqrt{\frac{\Sigma(x_i - \bar{x}_1)^2}{n_1 - 1}} = 2.28$

$s_2 = \sqrt{\frac{\Sigma(x_i - \bar{x}_2)^2}{n_2 - 1}} = 1.79$

c. $\bar{x}_1 - \bar{x}_2 = 9 - 7 = 2$

d. $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{2.28^2}{6} + \frac{1.79^2}{6}\right)^2}{\frac{1}{5}\left(\frac{2.28^2}{6}\right)^2 + \frac{1}{5}\left(\frac{1.79^2}{6}\right)^2} = 9.5$$

Use $df = 9$, $t_{.05} = 1.833$

$\bar{x}_1 - \bar{x}_2 \pm 1.833\sqrt{\frac{2.28^2}{6} + \frac{1.79^2}{6}}$

2 ± 2.17 (–.17 to 4.17)

12. a. $\bar{x}_1 - \bar{x}_2 = 36.2 - 29.9 = 6.3$

b. $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{13.5^2}{50} + \frac{11.9^2}{40}\right)^2}{\frac{1}{49}\left(\frac{13.5^2}{50}\right)^2 + \frac{1}{39}\left(\frac{11.9^2}{40}\right)^2} = 87.1$$

Use $df = 87$, $t_{.025} = 1.988$

6

$6.3 \pm 1.988\sqrt{\frac{13.5^2}{50} + \frac{11.9^2}{40}}$

6.3 ± 5.3 (1 to 11.6)

13. a. $\bar{x}_1 = \frac{\Sigma x_i}{n_1} = \frac{425}{10} = 42.5$

$s_1 = \sqrt{\frac{\Sigma(x_i - \bar{x}_1)^2}{n_1 - 1}} = \sqrt{\frac{438.56}{10-1}} = 6.98$

$\bar{x}_2 = \frac{\Sigma x_i}{n_2} = \frac{267.6}{12} = 22.3$

$s_2 = \sqrt{\frac{\Sigma(x_i - \bar{x}_2)^2}{n_2}} = \sqrt{\frac{225.96}{12-1}} = 4.53$

b. $\bar{x}_1 - \bar{x}_2 = 42.5 - 22.3 = 20.2$ or \$20,200

The mean annual cost to attend private colleges is \$20,200 more than the mean annual cost to attend public colleges.

c.
$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{6.98^2}{10} + \frac{4.53^2}{12}\right)^2}{\frac{1}{9}\left(\frac{6.98^2}{10}\right)^2 + \frac{1}{11}\left(\frac{4.53^2}{12}\right)^2} = 14.9$$

Use $df = 14$, $t_{.025} = 2.145$

$$(\bar{x}_1 - \bar{x}_2) \pm t_{.025}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$20.2 \pm 2.145\sqrt{\frac{6.98^2}{10} + \frac{4.53^2}{12}}$$

20.2 ± 5.5 (14.7 to 25.7)

95% confidence interval, private colleges have a population mean annual cost \$14,700 to \$25,700 more expensive than public colleges.

14. a. H_0: $\mu_1 - \mu_2 \geq 0$

H_a: $\mu_1 - \mu_2 < 0$

b.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(48{,}537 - 55{,}317) - 0}{\sqrt{\frac{18{,}000^2}{110} + \frac{10{,}000^2}{30}}} = -2.71$$

c.

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{18{,}000^2}{110} + \frac{10{,}000^2}{30}\right)^2}{\frac{1}{110 - 1}\left(\frac{18{,}000^2}{110}\right)^2 + \frac{1}{30 - 1}\left(\frac{10{,}000^2}{30}\right)^2}$$

$$= 85.20$$

Rounding down, we will use a t distribution with 85 degrees of freedom. From the t table we see that $t = -2.71$ corresponds to a p-value between .005 and 0.

Exact p-value corresponding to $t = -2.71$ is .004.

d. p-value $\leq$ .05, reject H_0. We conclude that the salaries of finance majors are lower than the salaries of business analytics majors.

15. a. μ_1 = population mean change in hotels rates in Europe

μ_2 = population mean change in hotels rates in the United States

H_0: $\mu_1 - \mu_2 = 0$

H_a: $\mu_1 - \mu_2 \neq 0$ Research hypothesis

b. $\bar{x}_1 - \bar{x}_2 = 0.039 - 0.047 = -0.008$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(0.039 - 0.047) - 0}{\sqrt{\frac{0.0301^2}{47} + \frac{0.0364^2}{53}}} = -1.134$$

$$df = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{0.0301^2}{47}+\frac{0.0364^2}{53}\right)^2}{\frac{1}{47-1}\left(\frac{0.0301^2}{47}\right)^2+\frac{1}{53-1}\left(\frac{0.0364^2}{53}\right)^2} = 97.5$$

Rounding down, the degrees of freedom = 97. Because this is a two-tailed test, the *p*-value is two times the lower-tail area. Using *t* table, area in lower tail is between .10 and .20; therefore, *p*-value is between .20 and .40. Using software, *p*-value =.26.

p-value > .01, do not reject H_0. Conclusion: The mean change in hotels rates in Europe and the United States are not different.

16. a. μ_1 = population mean verbal score parents college grads

μ_2 = population mean verbal score parents high school grads

H_0: $\mu_1 - \mu_2 \le 0$

H_a: $\mu_1 - \mu_2 > 0$

b. $\bar{x}_1 = \frac{\Sigma x_i}{n} = \frac{8400}{16} = 525$

$\bar{x}_2 = \frac{\Sigma x_i}{n} = \frac{5844}{12} = 487$

$\bar{x}_1 - \bar{x}_2$ = 525 – 487 = 38 points higher if parents are college grads

c. $s_1 = \sqrt{\frac{\Sigma(x_i-\bar{x}_1)^2}{n_1-1}} = \sqrt{\frac{52962}{16-1}} = \sqrt{3530.8} = 59.42$

$s_2 = \sqrt{\frac{\Sigma(x_i-\bar{x}_2)^2}{n_2-1}} = \sqrt{\frac{29456}{12-1}} = \sqrt{2677.82} = 51.75$

$$t = \frac{(\bar{x}_1-\bar{x}_2)-D_0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} = \frac{(525-487)-0}{\sqrt{\frac{59.42^2}{16}+\frac{51.75^2}{12}}} = 1.80$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{59.42^2}{16} + \frac{51.75^2}{12}\right)^2}{\frac{1}{15}\left(\frac{59.42^2}{16}\right)^2 + \frac{1}{11}\left(\frac{51.75^2}{12}\right)^2} = 25.3$$

Use $df = 25$.

Using t table, p-value is between .025 and .05.

Exact p-value corresponding to $t = 1.80$ is .0420/

d. p-value $\leq$ .05, reject H_0. Conclude higher population mean verbal scores for students whose parents are college grads.

17. a. H_0: $\mu_1 - \mu_2 \leq 0$

H_a: $\mu_1 - \mu_2 > 0$

b.
$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(6.82 - 6.25) - 0}{\sqrt{\frac{.64^2}{16} + \frac{.75^2}{10}}} = 1.99$$

c.
$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{.64^2}{16} + \frac{.75^2}{10}\right)^2}{\frac{1}{15}\left(\frac{.64^2}{16}\right)^2 + \frac{1}{9}\left(\frac{.75^2}{10}\right)^2} = 16.9$$

Use $df = 16$

Using t table, p-value is between .025 and .05.

Exact p-value corresponding to $t = 1.99$ is .0320.

d. p-value $\leq$ .05, reject H_0. The consultant with more experience has a higher population mean rating.

18. a. Let μ_1 = population mean minutes late for delayed Delta flights

μ_2 = population mean minutes late for delayed Southwest flights

H_0: $\mu_1 - \mu_2 = 0$

H_a: $\mu_1 - \mu_{22} \neq 0$

b. $\bar{x}_1 = \frac{\sum_{i=1}^{n} x_i}{n_1} = \frac{1265}{25} = 50.6$ minutes

$\bar{x}_2 = \frac{\sum_{i=1}^{n} x_i}{n_2} = \frac{1056}{20} = 52.8$ minutes

The difference between sample mean delay times is 50.6 – 52.8 = –2.2 minutes, which indicates the sample mean delay time is 2.2 minutes less for Delta.

c. Sample standard deviations: $s_1 = 26.57$ and $s_2 = 20.11$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(50.6 - 52.8) - 0}{\sqrt{\frac{26.57^2}{25} + \frac{20.11^2}{20}}} = -.32$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{26.57^2}{25} + \frac{20.11^2}{20}\right)^2}{\frac{1}{24}\left(\frac{26.57^2}{25}\right)^2 + \frac{1}{19}\left(\frac{20.11^2}{20}\right)^2} = 42.9$$

Use $df = 42$.

p-value for this two-tailed test is two times the lower-tail area for $t = -.32$.

Using t table, p-value is greater than 2(.20) = .40.

Exact p-value corresponding to $t = -.32$ with 42 df is .7506.

p-value > .05, do not reject H_0. We cannot reject the assumption that the population mean delay times are the same at Delta and Southwest Airlines. There is no statistical evidence that one airline does better than the other in terms of its population mean delay time.

19. a. 1, 2, 0, 0, 2

b. $\bar{d} = \sum d_i / n = 5/5 = 1$

c. $s_d = \sqrt{\dfrac{\sum(d_i - \bar{d})^2}{n-1}} = \sqrt{\dfrac{4}{5-1}} = 1$

d. $t = \dfrac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \dfrac{1-0}{1/\sqrt{5}} = 2.24$

$df = n - 1 = 4$

Using t table, p-value is between .025 and .05.

Exact p-value corresponding to $t = 2.24$ is .0443.

Reject H_0; conclude $\mu_d > 0$.

20. a. 3, –1, 3, 5, 3, 0, 1

b. $\bar{d} = \sum d_i / n = 14/7 = 2$

c. $s_d = \sqrt{\dfrac{\sum(d_i - \bar{d})^2}{n-1}} = \sqrt{\dfrac{26}{7-1}} = 2.08$

d. $\bar{d} = 2$

e. With 6 degrees of freedom $t_{.025} = 2.447$

$2 \pm 2.447\left(2.082/\sqrt{7}\right)$

2 ± 1.93 (.07 to 3.93)

21. Difference = rating after – rating before

H_0: $\mu_d \leq 0$

H_a: $\mu_d > 0$

$\bar{d} = .625$ and $s_d = 1.30$

$t = \dfrac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \dfrac{.625 - 0}{1.30/\sqrt{8}} = 1.36$

$df = n - 1 = 7$

Using t table, p-value is between .10 and .20.

Exact p-value corresponding to $t = 1.36$ is .1080.

Do not reject H_0; we cannot conclude that seeing the commercial improves the mean potential to purchase.

$$\bar{d} = \frac{\sum_i d_i}{n} = \frac{-114.06}{25} = -4.56$$

22. a. Let

$$d_i = \frac{end\ of\ first\ quarter\ price - beginning\ of\ first\ quarter\ price}{beginning\ of\ first\ quarter\ price}$$

$$\bar{d} = \frac{\sum_i d_i}{n} = \frac{-1.74}{25} = -0.07$$

b.

$$s_d = \sqrt{\frac{\sum_i (d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{0.2666}{25-1}} = 0.105$$

With $df = 24$, $t_{.025} = 2.064$

$$\bar{d} \pm t_{.025} \frac{s_d}{\sqrt{n}} = -0.07 \pm 2.064\left(\frac{0.105}{\sqrt{25}}\right) = -0.07 \pm 0.04$$

Confidence interval: (–0.11 to –0.03)

The 95% confidence interval shows that the population mean percentage change in the price per share of stock is a decrease of 3% to 11%. This may be a harbinger of a further stock market swoon.

23. a. μ_1 = population mean grocery expenditures

μ_2 = population mean dining-out expenditures

H_0: $\mu_d = 0$

H_a: $\mu_d \neq 0$

b. $t = \dfrac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \dfrac{850 - 0}{1123/\sqrt{42}} = 4.91$

$df = n - 1 = 41$

p-value ≈ 0

Conclude that there is a difference between the annual population mean expenditures for groceries and for dining out.

c. Groceries has the higher mean annual expenditure by an estimated $850.

$$\bar{d} \pm t_{.025} \frac{s_d}{\sqrt{n}}$$

$$850 \pm 2.020 \frac{1123}{\sqrt{42}}$$

850 ± 350 (500 to 1200)

24. a. Difference = Current year airfare – Previous year airfare

H_0: $\mu_d \leq 0$

H_a: $\mu_d > 0$

Differences 30, 63, –42, 10, 10, –27, 50, 60, 60, –30, 62, 30

$$\bar{d} = \frac{\sum d_i}{n} = \frac{276}{12} = 23$$

$$s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{16{,}558}{12-1}} = 38.80$$

$$t = \frac{\bar{d} - 0}{s_d / \sqrt{n}} = \frac{23 - 0}{38.80 / \sqrt{12}} = 2.05$$

$df = n - 1 = 11$

Using t table, p-value is between .05 and .025.

Exact p-value corresponding to $t = 2.05$ is .0325.

Because p-value < .05, reject H_0. We can conclude that there has been a significance increase in business travel airfares over the one-year period.

b. Current year: $\bar{x} = \sum x_i / n = 5844 / 12 = \$487$

Previous year: $\bar{x} = \sum x_i / n = 5568 / 12 = \$464$

c. One-year increase = \$487 – \$464 = \$23

\$23/\$464 = .05, or a 5% increase in business travel airfares for the one-year period.

25. a. Difference = math score – writing score

H_0: $\mu_d = 0$

H_a: $\mu_d \neq 0$

Use difference data: 66, 52, 65, –38, 28, –24, 50, 40, –5, 31, 55, –20

$$\bar{d} = \frac{\sum d_i}{n} = \frac{300}{12} = 25$$

$$s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n - 1}} = \sqrt{\frac{15{,}100}{12 - 1}} = 37.05$$

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{25 - 0}{37.05 / \sqrt{12}} = -2.34$$

$df = n - 1 = 11$

Using t table, lower-tail area is between .025 and .01.

Thus, the two-tailed test p-value is between .05 and .02.

Exact p-value corresponding to $t = -2.34$ is .0392.

p-value $\leq$ .05, reject H_0. Conclude that there is a significant difference between the population mean scores for the SAT math test and the SAT writing test.

b. $\bar{d} = 25$

$\bar{x}_M = \frac{\sum x_i}{n} = \frac{6168}{12} = 514$ for the math test

$\bar{x}_W = \frac{\sum x_i}{n} = \frac{5868}{12} = 489$ for the writing test

The SAT math test has a higher mean score than the SAT writing test.

26. a. H_0: $\mu_d = 0$

H_a: $\mu_d \neq 0$

Differences: –2, –1, –5, 1, 1, 0, 4, –7, –6, 1, 0, 2, –3, –7, –2, 3, 1, 2, 1, –4

$$\bar{d} = \sum d_i / n = -21/20 = -1.05$$

$$s_d = \sqrt{\frac{\sum(d_i - \bar{d})^2}{n-1}} = 3.3162$$

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-1.05 - 0}{3.3162/\sqrt{20}} = -1.42$$

$df = n - 1 = 19$

Using t table, area in tail is between .05 and .10.

Two-tail p-value must be between .10 and .20.

Exact p-value corresponding to $t = -1.42$ is .1718.

Cannot reject H_0. There is no significant difference between the mean scores for the first and fourth rounds.

b. $\bar{d} = -1.05$; first round scores were lower than fourth round scores.

c. $\alpha = .05$

$df = 19 \qquad t = 1.729$

Margin of error $= t_{.025}\dfrac{s_d}{\sqrt{n}} = 1.729\dfrac{3.3162}{\sqrt{20}} = 1.28$

Yes, just check to see if the 90% confidence interval includes a difference of zero. If it does, the difference is not statistically significant.

90% Confidence interval: –1.05 ± 1.28 (–2.33, .23)

The interval does include 0, so the difference is not statistically significant.

27. a. Difference = Price deluxe – Price standard

H_0: $\mu_d = 10$

H_a: $\mu_d \neq 10$

$\bar{d} = 8.86$ and $s_d = 2.61$

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{8.86 - 10}{2.61 / \sqrt{7}} = -1.16$$

$df = n - 1 = 6$

Using t table, area is between .10 and .20.

Two-tail p-value is between .20 and .40.

Exact p-value corresponding to $t = -1.16$ is .2901.

Do not reject H_0; we cannot reject the hypothesis that a \$10 price differential exists.

b. 95% confidence interval

$\bar{d} \pm t_{.025} = s_d / \sqrt{n}$

$8.86 \pm 2.447(2.61)/\sqrt{7}$

8.86 ± 2.41 or (6.45 to 11.27)

28. a. $\bar{p}_1 - \bar{p}_2 = .48 - .36 = .12$

b. $$\bar{p}_1 - \bar{p}_2 \pm z_{.05}\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$$

$$.12 \pm 1.645\sqrt{\frac{.48(1-.48)}{400} + \frac{.36(1-.36)}{300}}$$

$.12 \pm .0614$ (.0586 to .1814)

c. $$.12 \pm 1.96\sqrt{\frac{.48(1-.48)}{400} + \frac{.36(1-.36)}{300}}$$

$.12 \pm .0731$ (.0469 to .1931)

29. a. $$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{200(.22) + 300(.16)}{200 + 300} = .1840$$

$$z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.22 - .16}{\sqrt{.1840(1-.1840)\left(\frac{1}{200} + \frac{1}{300}\right)}} = 1.70$$

p - value = 1.0000 – .9554 = .0446

b. p-value $\leq$ .05; reject H_0.

30. $\bar{p}_1 = 220/400 = .55$ $\bar{p}_2 = 192/400 = .48$

$$\bar{p}_1 - \bar{p}_2 \pm z_{.025}\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$$

$$.55 - .48 \pm 1.96\sqrt{\frac{.55(1-.55)}{400} + \frac{.48(1-.48)}{400}}$$

$.07 \pm .0691$ (.0009 to .1391)

Seven percent more executives are predicting an increase in full-time jobs. The confidence interval shows the difference may be from 0.09% to 13.91%. Because 0% does not lie strictly within this confidence interval (i.e., the lower limit of the confidence interval is not negative), we can state with 95% confidence that the percentage of executives optimistic about hiring is larger in the current year than in the previous year.

31. a. The point estimate of the proportion of women who trust recommendations made on Pinterest is $\bar{p}_1$ = 117/150 = .78.

b. The point estimate of the proportion of men who trust recommendations made on Pinterest is $\bar{p}_2$ = 102/170 = .60.

c. $\bar{p}_1 - \bar{p}_2$ = .78 – .60 = .18

$$.18 \pm z_{.025}\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1}+\frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$$

$$.18 \pm 1.96\sqrt{\frac{.78(.22)}{150}+\frac{.60(.40)}{170}}$$

.18 ± .0991 (.0809 to .2791)

The 95% confidence interval estimate of the difference between the proportion of women and men who trust recommendations made on Pinterest is .18 ± .0991 or (.0809 to .2791).

32. Let p_1 = the population proportion of tuna that is mislabeled

p_2 = the population proportion of mahi mahi that is mislabeled

a. The point estimate of the proportion of tuna that is mislabeled is $\bar{p}_1$ = 99/220 = .45.

b. The point estimate of the proportion of mahi mahi that is mislabeled is $\bar{p}_2$ = 56/160 = .35.

c. $\bar{p}_1 - \bar{p}_2$ = .45 – .35 = .10

$$.10 \pm z_{.025}\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1}+\frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$$

$$.10 \pm 1.96\sqrt{\frac{.45(.55)}{220}+\frac{.35(.65)}{160}}$$

.10 ± .0989 (.0011 to .1989)

The 95% confidence interval estimate of the difference between the proportion of tuna and mahi mahi that is mislabeled is .10 ± .0989 or (.0011 to .1989).

33. Let p_1 = the population proportion of voters in rural Minnesota voted in the 2016 election

p_2 = the population proportion of voters in urban Minnesota voted in the 2016 election

a. H_0: $p_1 \le p_2$

H_a: $p_1 > p_2$

b. $\bar{p}_1$ = 663/884 = .75 75% of voters in rural Minnesota voted in the 2016 election.

c. $\bar{p}_2$ = 414/575 = .72 72% of voters in urban Minnesota voted in the 2016 election.

d. $\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1+n_2} = \frac{663+414}{884+575} = .7382$

$$z = \frac{\bar{p}_1-\bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{.75-.72}{\sqrt{.7382(1-.7382)\left(\frac{1}{884}+\frac{1}{575}\right)}} = 1.27$$

Upper-tail p-value is the area to the right of the test statistic.

Using normal table with z = 1.27: p-value = 1 – .8980 = .1020.

p-value > α ; do not reject H_0.

We cannot conclude that voters from rural Minnesota voted more frequently than voters from urban Minnesota in the 2016 presidential election.

34. Let p_1 = the population proportion of wells drilled in 2012 that were dry

p_2 = the population proportion of wells drilled in 2018 that were dry

a. H_0: $p_1 - p_2 \le 0$

H_a: $p_1 - p_2 > 0$

b. $\bar{p}_1$ = 24/119 = .2017

c. $\bar{p}_2$ = 21/162 = .1111

d. $\bar{p} = \dfrac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \dfrac{24+18}{119+162} = .1495$

$$z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{.2017-.1111}{\sqrt{.1495(1-.1495)\left(\frac{1}{119}+\frac{1}{162}\right)}} = 2.10$$

Upper-tail p-value is the area to the right of the test statistic.

Using normal table with z = 2.10: p-value = 1 – .9821 = .0179.

p-value < .05, so reject H_0 and conclude that wells drilled in 2012 were dry more frequently than wells drilled in 2018—that is, the frequency of dry wells has decreased over the eight years from 2012 to 2018.

35. a. Let p_1= population proportion of rooms occupied for current year

p_2= population proportion of rooms occupied for previous year

H_0: $p_1 - p_2 \le 0$

H_a: $p_1 - p_2 > 0$

b. $\bar{p}_1$= 1470/1750 = .84 (current year)

$\bar{p}_2$= 1458/1800 = .81 (previous year)

c. $\bar{p} = \dfrac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \dfrac{1750(.84)+1800(.81)}{1750+1800} = .8248$

$$z=\frac{\bar{p}_1-\bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}=\frac{.84-.81}{\sqrt{.8248(1-.8248)\left(\frac{1}{1750}+\frac{1}{1800}\right)}}=2.35$$

p-value is are in the upper tail at $z = 2.35$.

p-value = 1.0000 – .9906 = .0094

p-value $\leq .05$, reject H_0. There has been an increase in the hotel occupancy rate.

d. $$\bar{p}_1-\bar{p}_2 \pm z_{.025}\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1}+\frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$$

$$.84-.81\pm 1.96\sqrt{\frac{.84(1-.84)}{1750}+\frac{.81(1-.81)}{1800}}$$

$.03 \pm .025$ (.005 to .055)

Officials would likely be pleased with the occupancy statistics. The trend for the current year is an increase in hotel occupancy rates compared to last year. The point estimate is a 3% increase with a 95% confidence interval from .5% to 5.5%.

36. a. Let p_1 = population proportion of men expecting to get a raise or promotion this year

p_2 = population proportion of women expecting to get a raise or promotion this year

H_0: $p_1 - p_2 \leq 0$

H_a: $p_1 - p_2 > 0$

b. $\bar{p}_1$ = 104/200 = .52 (52%)

$\bar{p}_2$ = 74/200 = .37 (37%)

c. $$\bar{p}=\frac{n_1\bar{p}_1+n_2\bar{p}_2}{n_1+n_2}=\frac{104+74}{200+200}=.445$$

$$z=\frac{\bar{p}_1-\bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}=\frac{.52-.37}{\sqrt{.445(1-.445)\left(\frac{1}{200}+\frac{1}{200}\right)}}=3.02$$

p-value = 1.0000 – .9987 = .0013

Reject H_0. There is a significant difference between the population proportions with a great proportion of men expecting to get a raise or a promotion this year.

37. a. Let p_1 = population proportion of Carl's loans that default

p_2 = population proportion of Norm's loans that default

H_0: $p_1 - p_2 = 0$

H_a: $p_1 - p_2 \neq 0$

b. $\bar{p}_1$ = 9/60 = 0.15 (15%)

$\bar{p}_2$ = 7/80 = 0.0875 (7.5%)

c. $\bar{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2} = \frac{9+7}{60+80} = .114$

$$z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.15 - .085}{\sqrt{.114(1-.114)\left(\frac{1}{60} + \frac{1}{80}\right)}} = 1.15$$

p-value = 2(1.0000 – .875) = 0.25

Do no reject H_0. We cannot conclude there is a significant difference between the population default proportions in the loans approved by Carl and the loans approved by Norm.

38. H_0: $\mu_1 - \mu_2 = 0$

H_a: $\mu_1 - \mu_2 \neq 0$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(4.1 - 3.4) - 0}{\sqrt{\frac{(2.2)^2}{120} + \frac{(1.5)^2}{100}}} = 2.79$$

p-value = 2(1.0000 – .9974) = .0052

p-value $\leq$ .05, reject H_0. A difference exists with system B having the lower mean checkout time.

39. a. $\bar{x}_1 = \frac{\Sigma x_i}{n_1} = \frac{21{,}548}{33} = 652.97$ Mean monthly lease payment for an SUV in 2015

$\bar{x}_2 = \frac{\Sigma x_i}{n_2} = \frac{24{,}841}{46} = 540.02$ Mean monthly lease payment for an SUV in 2016

Difference = 652.97 – 540.02 = 112.95

Using sample mean prices, the 2015 mean monthly lease payment for an SUV is $112.95 more than in 2016.

b. $s_1 = \sqrt{\frac{\Sigma(x_i - \bar{x}_1)^2}{n_1 - 1}} = 101.02$

$s_2 = \sqrt{\frac{\Sigma(x_i - \bar{x}_2)^2}{n_2}} = 93.48$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{(101.02)^2}{33} + \frac{(93.48)^2}{46}} = 22.343$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{101.02^2}{33} + \frac{93.48^2}{46}\right)^2}{\frac{1}{32}\left(\frac{101.02^2}{33}\right)^2 + \frac{1}{45}\left(\frac{93.48^2}{46}\right)^2} = 65.75$$

Use $df = 65$, $t_{.005} = 2.654$.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{.005}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 112.95 \pm 59.30 \text{ (53.65 to 172.25)}$$

We are 99% confident that monthly lease payments for an SUV were between $53.65 and $172.25 higher in 2015 than in 2016.

c. Using the confidence interval of part b, we can with 99% confidence state that 2015 is higher than 2016 by 53.65 to 172.25 and therefore did decline from2015 t0 2016.

To answer this question, we can also conduct a one-tailed hypothesis test. No value for the level of significance (α) has been given, but most people would agree that a p-value $\leq$ .01 would justify concluding that monthly lease payments for an SUV have decreased from 2015 to 2016.

μ_1 = population mean monthly lease payment for an SUV in 2015

μ_2 = population mean monthly lease payment for an SUV in 2016.

$H_0: \mu_1 - \mu_2 \leq 0$

$H_a: \mu_1 - \mu_2 > 0$ Research hypothesis

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{112.95}{\sqrt{\frac{101.02^2}{33} + \frac{93.48^2}{46}}} = 5.06$$

Degrees of freedom = 65 (calculated in part b).

p-value is upper-tail area.

Using t table: p-value is less than .005. Using software, p-value≈ 0. p-value $\leq$ .01, Reject H_0. Thus, we are justified in concluding that mean monthly lease payments for SUVs are higher in 2015 than 2016 and therefore, did decline from 2015 to 2016.

40. a. $H_0: \mu_1 - \mu_2 \leq 0$

$H_a: \mu_1 - \mu_2 > 0$

b. $n_1 = 30$ $n_2 = 30$

$\bar{x}_1 = 16.23$ $\bar{x}_2 = 15.70$

$s_1 = 3.52$ $s_2 = 3.31$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(16.23 - 15.70) - 0}{\sqrt{\frac{(3.52)^2}{30} + \frac{(3.31)^2}{30}}} = .60$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{3.52^2}{30} + \frac{3.31^2}{30}\right)^2}{\frac{1}{29}\left(\frac{3.52^2}{30}\right)^2 + \frac{1}{29}\left(\frac{3.31^2}{30}\right)^2} = 57.8$$

Use $df = 57$.

Using t table, p-value is greater than .20.

Exact p-value corresponding to $t = .60$ is .2754.

p-value > .05, do not reject H_0. Cannot conclude that the mutual funds with a load have a greater mean rate of return.

41. a. $n_1 = 10$ $n_2 = 8$

$\bar{x}_1 = 21.2$ $\bar{x}_2 = 22.8$

$s_1 = 2.70$ $s_2 = 3.55$

$\bar{x}_1 - \bar{x}_2 = 21.2 - 22.8 = -1.6$

Kitchens are less expensive by \$1600.

b.

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{2.70^2}{10} + \frac{3.55^2}{8}\right)^2}{\frac{1}{9}\left(\frac{2.70^2}{10}\right)^2 + \frac{1}{7}\left(\frac{3.55^2}{8}\right)^2} = 12.9$$

Use $df = 12$, $t_{.05} = 1.782$

$$-1.6 \pm 1.782\sqrt{\frac{2.70^2}{10} + \frac{3.55^2}{8}}$$

-1.6 ± 2.7 $(-4.3$ to $1.1)$

42. a. $\bar{d} = \frac{\Sigma d_i}{n} = \frac{280}{20} = 14$

b. $s_d = \sqrt{\frac{\Sigma(d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{54{,}880}{19}} = 53.744$

$df = n - 1 = 19$, $t_{.05} = 1.729$

$$\bar{d} \pm t_{.05} \frac{s_d}{\sqrt{n}} = 14 \pm 1.729 \frac{53.744}{\sqrt{20}}$$

14 ± 20.78 (–6.78 to 34.78)

c. H_0: $\mu_d = 0$

H_a: $\mu_d \neq 0$

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{14 - 0}{53.744 / \sqrt{20}} = 1.165$$

Using t table with $df = 19$, the area in upper tail is between .20 and .10. Thus, for the two-tailed test, the p-value is between .20 and .40.

Using software, the exact p-value for $t = 1.165$ is .258.

Cannot reject H_0; cannot concluded that there is a difference between the mean scores for the no sibling and with sibling groups.

43. a. Let p_1 = population proportion saying financial security more than fair in recent year

p_2 = population proportion saying financial security more than fair in previous year

$$H_0: p_1 - p_2 = 0$$
$$H_a: p_1 - p_2 \neq 0$$

b.

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{1000(.41) + 900(.35)}{1000 + 900} = .3816$$

$$z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(.41 - .35)}{\sqrt{.3816(1-.3816)\left(\frac{1}{1000} + \frac{1}{900}\right)}} = 2.69$$

p-value for this two-tailed test is two times the area in the upper tail at $z = 2.69$.

p-value = 2(1.0000 – .9964) = .0072.

p-value $\leq$ .05, reject H_0. Conclude the population proportions are not equal.

There has been a change in the population proportion saying that their financial security is more than fair.

c. $\bar{p}_1 - \bar{p}_2 \pm z_{.025}\sqrt{\dfrac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \dfrac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$

$(.41 - .35) \pm 1.96\sqrt{\dfrac{.41(1-.41)}{1000} + \dfrac{.35(1-.35)}{900}}$

$.06 \pm .0436$

95% Confidence Interval (.0164 to .1036)

Based on the results, the population proportion of adults saying that their financial security is more than fair has increased between 1.64% and 10.36%.

44. a. $\bar{p}_1 = 76/400 = .19$

$\bar{p}_2 = 90/900 = .10$

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{76+90}{400+900} = .1277$$

$$z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \frac{.19 - .10}{\sqrt{.1277(1-.1277)\left(\dfrac{1}{400} + \dfrac{1}{900}\right)}} = 4.49$$

p-value ≈ 0

Reject H_0; there is a difference between claim rates.

b. $\bar{p}_1 - \bar{p}_2 \pm z_{.025}\sqrt{\dfrac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \dfrac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$

$.19 - .10 \pm 1.96\sqrt{\dfrac{.19(1-.19)}{400} + \dfrac{.10(1-.10)}{900}}$

$.09 \pm .0432$ (.0468 to .1332)

Claim rates are higher for single males.

45. $\bar{p}_1 = 9/142 = .0634$

$\overline{p}_2 = 5/268 = .0187$

$$\overline{p} = \frac{n_1\overline{p}_1 + n_2\overline{p}_2}{n_1 + n_2} = \frac{9+5}{142+268} = .0341$$

$$z = \frac{\overline{p}_1 - \overline{p}_2}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{.0634 - .0187}{\sqrt{.0341(1-.0341)\left(\frac{1}{142}+\frac{1}{268}\right)}} = 2.37$$

p-value = 2(1.0000 – .9911) = .0178

p-value $\leq$ .02, reject H_0. There is a significant difference in drug resistance between the two states. Alabama has the higher drug-resistance rate.

46. Let p_1 = the population proportion of American adults under 30 years old

p_2 = the population proportion of Americans over 30 years old

a. From the file *Computer News*, there are 109 Yes responses for each age group. The total number of respondents of the under 30 years group is n_1 = 150, whereas the 30 and over group had is n_2 = 200 total respondents.

American adults under 30 years old: $\overline{p}_1$ = 109/150 = .727

Americans over 30 years old: $\overline{p}_2$ = 109/200 = .545

b. $\overline{p}_1 - \overline{p}_2 = .727 - .545 = .182$

$$s_{\overline{p}_1 - \overline{p}_2} = \sqrt{\frac{.727(1-.727)}{150} + \frac{.545(1-.545)}{200}} = .0506$$

Confidence interval: .182 ± 1.96(.0506) or.182 ± .0992 (.0824 to .2809)

c. Because the confidence interval in part b does not include 0 and both values are negative, conclude that the proportion of American adults under 30 years old who use a computer to gain access to news is greater than the proportion of Americans over 30 years old that use a computer to gain access to news.