Comparing Population Proportions

8.2 If both sample sizes are small, then the sampling distribution of \( \hat{p}_1 - \hat{p}_2 \) may not be normal.

8.4 The sample sizes are large enough if \( n_1 \hat{p}_1 \geq 15, \ n_1 \hat{q}_1 \geq 15 \) and \( n_2 \hat{p}_2 \geq 15, \ n_2 \hat{q}_2 \geq 15 \).

a. \( n_1 \hat{p}_1 = 10(,5) = 5 \geq 15, \ n_1 \hat{q}_1 = 10(,5) = 5 \geq 15 \)

\( n_2 \hat{p}_2 = 12(,5) = 6 \geq 15, \ n_2 \hat{q}_2 = 12(,5) = 6 \geq 15 \)

Since none of the products are greater than or equal to 15, the sample sizes are not large enough to assume normality.

b. \( n_1 \hat{p}_1 = 10(,1) = 1 \geq 15, \ n_1 \hat{q}_1 = 10(,9) = 9 \geq 15 \)

\( n_2 \hat{p}_2 = 12(,08) = 96 \geq 15, \ n_2 \hat{q}_2 = 12(,92) = 11.04 \geq 15 \)

Since none of the products are greater than or equal to 15, the sample sizes are not large enough to assume normality.

c. \( n_1 \hat{p}_1 = 30(,2) = 6 \geq 15, \ n_1 \hat{q}_1 = 30(,8) = 24 \geq 15 \)

\( n_2 \hat{p}_2 = 30(,3) = 9 \geq 15, \ n_2 \hat{q}_2 = 30(,7) = 21 \geq 15 \)

Since two of the products are not greater than or equal to 15, the sample sizes are not large enough to assume normality.

d. \( n_1 \hat{p}_1 = 100(,05) = 5 \geq 15, \ n_1 \hat{q}_1 = 100(,95) = 95 \geq 15 \)

\( n_2 \hat{p}_2 = 200(,09) = 18 \geq 15, \ n_2 \hat{q}_2 = 200(,91) = 182 \geq 15 \)

Since one of the products is not greater than or equal to 15, the sample sizes are not large enough to assume normality.

e. \( n_1 \hat{p}_1 = 100(,95) = 95 \geq 15, \ n_1 \hat{q}_1 = 100(,05) = 5 \geq 15 \)

\( n_2 \hat{p}_2 = 200(,91) = 182 \geq 15, \ n_2 \hat{q}_2 = 200(,09) = 18 \geq 15 \)

Since one of the products is not greater than or equal to 15, the sample sizes are not large enough to assume normality.

8.6 a. \( H_0: \ p_1 - p_2 = 0 \)
\( H_a: \ p_1 - p_2 \neq 0 \)

Copyright © 2013 Pearson Education, Inc.
Some preliminary calculations are:

\[
\hat{p}_1 = \frac{320}{800} = .40 \\
\hat{p}_2 = \frac{40}{800} = .50 \\
\hat{p} = \frac{320 + 400}{800 + 800} = .45
\]

The test statistic is

\[
z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(.40 - .50) - 0}{\sqrt{(.45)(.55)\left(\frac{1}{800} + \frac{1}{800}\right)}} = -4.02
\]

The rejection region requires \(\alpha / 2 = .05 / 2 = .025\) in each tail of the \(z\) distribution. From Table III, Appendix A, \(z_{.025} = 1.96\). The rejection region is \(z < -1.96\) or \(z > 1.96\).

Since the observed value of the test statistic falls in the rejection region \((z = -4.02 < -1.96)\), \(H_0\) is rejected. There is sufficient evidence to indicate that the proportions are unequal at \(\alpha = .05\).

b. The problem is identical to part a until the rejection region. The rejection region requires \(\alpha / 2 = .01 / 2 = .005\) in each tail of the \(z\) distribution. From Table III, Appendix A, \(z_{.005} = 2.58\). The rejection region is \(z < -2.58\) or \(z > 2.58\).

Since the observed value of the test statistic falls in the rejection region \((z = -4.02 < -2.58)\), \(H_0\) is rejected. There is sufficient evidence to indicate that the proportions are unequal at \(\alpha = .01\).

c. \(H_0: \ p_1 - p_2 = 0\) \\
\(H_a: \ p_1 - p_2 < 0\)

Test statistic as above: \(z = -4.02\)

The rejection region requires \(\alpha = .01\) in the lower tail of the \(z\) distribution. From Table III, Appendix A, \(z_{.01} = 2.33\). The rejection region is \(z < -2.33\).

Since the observed value of the test statistic falls in the rejection region \((z = -4.02 < -2.33)\), \(H_0\) is rejected. There is sufficient evidence to indicate that \(p_1 < p_2\) at \(\alpha = .01\).

d. For confidence coefficient .90, \(\alpha = .10\) and \(\alpha / 2 = .10 / 2 = .05\). From Table III, appendix A, \(z_{.05} = 1.645\). The 90% confidence interval is:

\[
(\hat{p}_1 - \hat{p}_2) \pm z_{.05} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = (.4 -.5) \pm 1.645 \sqrt{\frac{(4)(.6)}{800} + \frac{(5)(.5)}{800}}
\]

\[
\Rightarrow -.10 \pm .04 \Rightarrow (-.14, -.06)
\]
8.8 The sampling distribution \( \hat{p}_1 - \hat{p}_2 \) is approximately normal with:

\[
\mu_{(\hat{p}_1 - \hat{p}_2)} = p_1 - p_2 = .1 - .5 = -.4
\]

\[
\sigma_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} = \sqrt{\frac{.1(.9)}{100} + \frac{(.5)(.5)}{200}} = .046
\]

A sketch of the distribution is:

---

8.10 a. Let \( p_1 = \) proportion of all Democrats who prefer steak as their favorite barbeque food.

The point estimate of \( p_1 \) is \( \hat{p}_1 = \frac{662}{1250} = .530 \).

b. Let \( p_2 = \) proportion of all Republicans who prefer steak as their favorite barbeque food.

The point estimate of \( p_2 \) is \( \hat{p}_2 = \frac{586}{930} = .630 \).

c. The point estimate of the difference between the proportions of all Democrats and all Republicans who prefer steak as their favorite barbeque food is \( \hat{p}_1 - \hat{p}_2 = .530 - .630 = -.100 \).

d. For confidence coefficient .95, \( \alpha = .05 \) and \( \alpha/2 = .05/2 = .025 \). From Table III, Appendix A, \( z_{.025} = 1.96 \). The 95\% confidence interval is:

\[
(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \Rightarrow (.530 - .630) \pm 1.96 \sqrt{\frac{.53(.47)}{1250} + \frac{.63(.37)}{930}}
\]

\[
\Rightarrow -.100 \pm .042 \Rightarrow (-.142, -.058)
\]

e. We are 95\% confident that the difference between the proportions of all Democrats and all Republicans who prefer steak as their favorite barbeque food is between -.142 and -.058. Thus, there is evidence that Republicans prefer steak as their favorite barbeque food more than Democrats.
Comparing Population Proportions

f. 95% confidence means that if we were to take repeated samples of size 1,250 Democrats and 930 Republicans and form 95% confident intervals for the difference in the population parameters, 95% of the confidence intervals formed will contain the true difference and 5% will not.

8.12 a. Let \( p_1 = \) proportion of men who prefer keeping track of appointments in their heads and \( p_2 = \) proportion of women who prefer keeping track of appointments in their heads. To determine whether the percentage of men who prefer keeping track of appointments in their heads is larger than the corresponding percentage of women, we test:

\[
H_0: \ p_1 - p_2 = 0 \\
H_a: \ p_1 - p_2 > 0
\]

b. Some preliminary calculations:

\[
\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} = \frac{500(0.56) + 500(0.46)}{500 + 500} = 0.51
\]

The test statistic is

\[
z = \left( \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right) = \left( \frac{0.56 - 0.46}{0.51(0.49)\left(\frac{1}{500} + \frac{1}{500}\right)} \right) = 3.16
\]

c. The rejection region requires \( \alpha = 0.01 \) in the upper tail of the \( z \) distribution. From Table III, Appendix A, \( z_{0.01} = 2.33 \). The rejection region is \( z > 2.33 \).

d. The \( p \)-value is \( p = P(z \geq 3.16) \approx 0.5 - 0.5 = 0 \).

e. Since the observed value of the test statistic falls in the rejection region (\( z = 3.16 > 2.33 \), \( H_0 \) is rejected. There is sufficient evidence to indicate the percentage of men who prefer keeping track of appointments in their heads is larger than the corresponding percentage of women at \( \alpha = 0.01 \).

OR

Since the \( p \)-value is less than \( \alpha \) (\( p = 0 < \alpha = 0.01 \), \( H_0 \) is rejected. There is sufficient evidence to indicate the percentage of men who prefer keeping track of appointments in their heads is larger than the corresponding percentage of women at \( \alpha = 0.01 \).

8.14 Let \( p_1 = \) proportion of patients in the angioplasty group who had a heart attack and \( p_2 = \) proportion of patients in the medication only group who had a heart attack. The parameter of interest is \( p_1 - p_2 \), or the difference in the proportions of patients who had a heart attack between the 2 groups.

Some preliminary calculations are:

\[
\hat{p}_1 = \frac{x_1}{n_1} = \frac{211}{1,145} = 0.184 \\
\hat{p}_2 = \frac{x_2}{n_2} = \frac{202}{1,142} = 0.177
\]

Copyright © 2013 Pearson Education, Inc.
For confidence coefficient .95, \( \alpha = .05 \) and \( \alpha / 2 = .05 / 2 = .025 \). From Table III, Appendix A, \( z_{.025} = 1.96 \). The 95% confidence interval is:

\[
(.184 - .177) \pm 1.96 \sqrt{\frac{.184(.816)}{1,145} + \frac{.177(.823)}{1,142} \to .007 \pm .032 \to (-.025, .039)}
\]

We are 95% confident that the difference in the proportions of patients who had a heart attack between the 2 groups is between \(-.025\) and \(.039\). Since 0 is contained in the interval, there is no evidence that there is a difference in the proportion of heart attacks between those who received angioplasty and those who received medication only at \( \alpha = .05 \).

8.16 Let \( p_1 = \) proportion of world-class athletes not competing in the 1999 World Championships and \( p_2 = \) proportion of world-class athletes not competing in the 2000 Olympic Games. If the new EPO test is effective, then the proportion of nonparticipating athletes will increase.

Some preliminary calculations are:

\[
\hat{p}_1 = \frac{x_1}{n_1} = \frac{159}{830} = .1916 \quad \hat{p}_2 = \frac{x_2}{n_2} = \frac{133}{825} = .1612 \quad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{159 + 133}{830 + 825} = .1764
\]

To determine if the new test for EPO was effective in deterring an athlete’s participation in the 2000 Olympics, we test:

\[
H_0: \ p_1 - p_2 = 0 \\
H_a: \ p_1 - p_2 < 0
\]

The test statistic is

\[
z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.1916 - .1612}{\sqrt{.1764(.8236)\left(\frac{1}{830} + \frac{1}{825}\right)}} = 1.62
\]

The rejection region requires \( \alpha = .10 \) in the lower tail of the \( z \) distribution. From Table III, Appendix A, \( z_{.10} = 1.28 \). The rejection region is \( z < -1.28 \).

Since the observed value of the test statistic does not fall in the rejection region \( (z = 1.62 \not< -1.28) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate the new test was effective in deterring an athlete’s participation in the 2000 Olympics at \( \alpha = .10 \).

8.18 a. The estimate of the incumbency rate in the surrounding milk market is:

\[
\hat{p}_1 = \frac{x_1}{n_1} = \frac{91}{134} = .679
\]

The estimate of the incumbency rate in the Tri-county milk market is:

\[
\hat{p}_2 = \frac{x_2}{n_2} = \frac{50}{51} = .980
\]
b. To determine if the incumbency rate for the Tri-county milk market is greater than that of the surrounding milk market, we test:

\[ H_0: \ p_1 - p_2 = 0 \]
\[ H_a: \ p_1 - p_2 < 0 \]

From the printout, the test statistic is \( z = -4.30 \).

The \( p \)-value for the test is \( p \)-value = 0.000. Since the \( p \)-value is so small, \( H_0 \) is rejected. There is sufficient evidence to indicate the incumbency rate for the Tri-county milk market is greater than that of the surrounding milk market for any reasonable value of \( \alpha \). This is evidence that bid collusion was present.

8.20 Let \( p_1 \) = proportion of all commercials in 1998 that contain religious symbolism and \( p_2 \) = proportion of all commercials in 2008 that contain religious symbolism. Some preliminary calculations are:

\[
\hat{p}_1 = \frac{x_1}{n_1} = \frac{16}{797} = .020 \\
\hat{p}_2 = \frac{x_2}{n_2} = \frac{51}{1,499} = .034 \\
\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{16 + 51}{797 + 1,499} = .029
\]

To determine if the percentage of commercials that use religious symbolism has changed since the 1998 study, we test:

\[ H_0: \ p_1 - p_2 = 0 \]
\[ H_a: \ p_1 - p_2 \neq 0 \]

The test statistic is

\[
z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.020 - .034}{\sqrt{.029(.971)\left(\frac{1}{797} + \frac{1}{1,499}\right)}} = -1.90
\]

The \( p \)-value of the test is \( P(z \leq -1.90) + P(z \geq 1.90) = 2(.5 - .4713) = 2(.0287) = .0574 \).

Since no \( \alpha \) was given, we will use \( \alpha = .05 \). Since the \( p \)-value is larger than \( \alpha \) \( (p = .0574 > .05) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate the percentage of commercials that use religious symbolism has changed since the 1998 study at \( \alpha = .05 \).

8.22 We can estimate the values of \( p_1 \) and \( p_2 \) based on prior samples, by using an educated guess of the values, or most conservatively, estimating \( p_1 \) and \( p_2 \) with .5.

8.24 a. For confidence coefficient .99, \( \alpha = 1 - .99 = .01 \) and \( \alpha/2 = .01/2 = .005 \). From Table III, Appendix A, \( z_{.005} = 2.58 \).

\[
n_1 = n_2 = \frac{(z_{\alpha/2})^2 \left( p_1 q_1 + p_2 q_2 \right)}{(SE)^2} = \frac{2.58^2 \left(.4(1-.4) + .7(1-.7) \right)}{.01^2} = 2.99538
\]

\[
= 29,953.8 = 29,954
\]
b. For confidence coefficient .90, \( \alpha = 1 - .90 = .10 \) and \( \alpha/2 = .10/2 = .05 \). From Table III, Appendix A, \( z_{.05} = 1.645 \). Since we have no prior information about the proportions, we use \( p_1 = p_2 = .5 \) to get a conservative estimate. For a width of .05, the standard error is .5.

\[
\begin{align*}
  n_1 = n_2 &= \frac{(z_{\alpha/2})^2 (p_1q_1 + p_2q_2)}{(SE)^2} \\
  &= \frac{1.645^2 (.5(1-.5) + .5(1-.5))}{.02^2} = 2,164.82 = 2,165
\end{align*}
\]

c. From part b, \( z_{.05} = 1.645 \).

\[
\begin{align*}
  n_1 = n_2 &= \frac{(z_{\alpha/2})^2 (p_1q_1 + p_2q_2)}{(SE)^2} \\
  &= \frac{1.645^2 (.5(1-.5) + .5(1-.5))}{.03^2} = 1,00123 \\
  &= 1,112.48 = 1,113
\end{align*}
\]

8.26 For confidence coefficient .90, \( \alpha = 1 - .90 = .10 \) and \( \alpha/2 = .10/2 = .05 \). From Table III, Appendix A, \( z_{.05} = 1.645 \). For width = .04, the standard error is \( SE = .04 / 2 = .02 \). Since no information is known about the true proportions, we will use \( p_1 = p_2 = .5 \) to get a conservative estimate of these values.

\[
\begin{align*}
  n_1 = n_2 &= \frac{(z_{\alpha/2})^2 (p_1q_1 + p_2q_2)}{(SE)^2} \\
  &= \frac{1.645^2 (.5(.5) + .5(.5))}{.02^2} = 3,382.5 = 3,383
\end{align*}
\]

8.28 a. For confidence coefficient .80, \( \alpha = 1 - .80 = .20 \) and \( \alpha/2 = .20/2 = .10 \). From Table III, Appendix A, \( z_{.10} = 1.28 \). Since we have no prior information about the proportions, we use \( p_1 = p_2 = .5 \) to get a conservative estimate. For a width of .06, the standard error is .03.

\[
\begin{align*}
  n_1 = n_2 &= \frac{(z_{\alpha/2})^2 (p_1q_1 + p_2q_2)}{(SE)^2} \\
  &= \frac{(1.28)^2 (.5(1-.5) + .5(1-.5))}{.03^2} = 910.22 = 911
\end{align*}
\]

b. For confidence coefficient .90, \( \alpha = 1 - .90 = .10 \) and \( \alpha/2 = .10/2 = .05 \). From Table III, Appendix A, \( z_{.05} = 1.645 \). Using the formula for the sample size needed to estimate a proportion from Chapter 7,\n
\[
\begin{align*}
  n &= \frac{(z_{\alpha/2})^2 pq}{(SE)^2} = \frac{1.645^2 (.5(1-.5))}{.02^2} = \frac{.6765}{.0004} = 1,691.27 = 1,692
\end{align*}
\]

No, the sample size from part a is not large enough.

8.30 For confidence coefficient .95, \( \alpha = 1 - .95 = .05 \) and \( \alpha/2 = .05/2 = .025 \). From Table III, Appendix A, \( z_{.025} = 1.96 \). Since no prior information was given as to the values of the \( p \)'s, we will use \( p = .5 \) for both populations.

\[
\begin{align*}
  n_1 = n_2 &= \frac{(z_{\alpha/2})^2 (p_1q_1 + p_2q_2)}{(SE)^2} \\
  &= \frac{1.96^2 (.5(.5) + .5(.5))}{.02^2} = 4,802
\end{align*}
\]

In order to estimate the difference in proportions to within .02 using a 90% confidence interval, we would need to sample 4,802 observations from each population.

Copyright © 2013 Pearson Education, Inc.
The sample size $n$ will be large enough so that for every cell the expected cell count $E_i$ will be equal to 5 or more.

Using Table V, Appendix A:

a. $P(\chi^2 \leq 1.063623) = 1 - .90 = .10$

b. $P(\chi^2 > 30.5779) = .01$

c. $P(\chi^2 \geq 82.3581) = .90$

d. $P(\chi^2 < 18.4926) = 1 - .95 = .05$

The hypotheses of interest are:

$H_0: p_1 = .25, p_2 = .25, p_3 = .50$

$H_a$: At least one of the probabilities differs from the hypothesized value

$E_1 = np_{1,0} = 320(.25) = 80$

$E_2 = np_{2,0} = 320(.25) = 80$

$E_3 = np_{3,0} = 320(.50) = 160$

The test statistic is

$$\chi^2 = \sum \frac{(n_i - E_i)^2}{E_i} = \frac{(78 - 80)^2}{80} + \frac{(60 - 80)^2}{80} + \frac{(182 - 160)^2}{160} = 8.075$$

The rejection region requires $\alpha = .05$ in the upper tail of the $\chi^2$ distribution with $df = k - 1 = 3 - 1 = 2$. From Table V, Appendix A, $\chi^2_{0.05} = 5.99147$. The rejection region is $\chi^2 > 5.99147$.

Since the observed value of the test statistic falls in the rejection region $(\chi^2 = 8.075 > 5.99147)$, $H_0$ is rejected. There is sufficient evidence to indicate that at least one of the probabilities differs from its hypothesized value at $\alpha = .05$.

Some preliminary calculations are:

$E_1 = np_{1,0} = 400(.2) = 80$

$E_2 = np_{2,0} = 400(.4) = 160$

$E_3 = np_{3,0} = 400(.1) = 40$

$E_4 = np_{4,0} = 400(.3) = 120$

To determine if probabilities differ from the hypothesized values, we test:

$H_0$: $p_1 = .2, p_2 = .4, p_3 = .1, and p_4 = .3$

$H_a$: At least one of the probabilities differs from its hypothesized value

The test statistic is

$$\chi^2 = \sum \frac{(n_i - E_i)^2}{E_i} = \frac{(70 - 80)^2}{80} + \frac{(196 - 160)^2}{160} + \frac{(46 - 40)^2}{40} + \frac{(88 - 120)^2}{120} = 18.78$$
Chapter 8

The rejection region requires \( \alpha = .05 \) in the upper tail of the \( \chi^2 \) distribution with \( df = k - 1 = 4 - 1 = 3 \). From Table V, Appendix A, \( \chi^2_{0.05} = 7.81473 \). The rejection region is \( \chi^2 > 7.81473 \).

Since the observed value of the test statistic falls in the rejection region \( (\chi^2 = 18.78 > 7.81473) \), \( H_0 \) is rejected. There is sufficient evidence to indicate the probabilities differ from their hypothesized values at \( \alpha = .05 \).

8.40 a. The categorical variable is slime mold species. It has 6 levels: LE, TM, AC, AD, HC, and HS.

b. To determine if the relative frequency of occurrence of beetles differs for the six slime mold species, we test:

\[ H_0: p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = 1/6 \]

\[ H_a: \text{At least one } p_i \neq 1/6 \]

c. Some preliminary calculations are:

\[ E_i = np_{i,0} = 19(1/6) = 3.1667 \text{ for } i = 1, 2, \ldots, 6 \]

The test statistic is

\[
\chi^2 = \sum \frac{(n_i - E_i)^2}{E_i} = \frac{[3 - 3.1667]^2}{3.1667} + \frac{[2 - 3.1667]^2}{3.1667} + \frac{[7 - 3.1667]^2}{3.1667}
\]

\[
+ \frac{[3 - 3.1667]^2}{3.1667} + \frac{[1 - 3.1667]^2}{3.1667} + \frac{[3 - 3.1667]^2}{3.1667} = 6.5789
\]

d. The rejection region requires \( \alpha = .05 \) in the upper tail of the \( \chi^2 \) distribution with \( df = k - 1 = 6 - 1 = 5 \). From Table V, Appendix A, \( \chi^2_{0.05} = 11.0705 \). The rejection region is \( \chi^2 > 11.0705 \).

Since the observed value of the test statistic does not fall in the rejection region \( (\chi^2 = 6.5789 > 11.0705) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate that the relative frequency of occurrence of beetles differs for the six slime mold species at \( \alpha = .05 \).

e. From part c, the expected cell count for each cell is 3.1667. In order for the test statistic to have a \( \chi^2 \) distribution, the expected cell counts for all the cells must be 5 or more. Thus, the validity of the test is in question.

8.42 a. To determine if one performance measure is used more often than any other, we test:

\[ H_0: p_1 = p_2 = p_3 = p_4 = p_5 = .2 \]

\[ H_a: \text{At least one } p_i \neq .2 \]
Comparing Population Proportions

From the printout, the test statistic is \( \chi^2 = 1.66667 \) and the \( p \)-value is \( p = .797 \). Since the \( p \)-value is not less than \( \alpha = .10 \), \( H_0 \) is not rejected. There is insufficient evidence to indicate one performance measure is used more often than any other at \( \alpha = .10 \).

b. \( \hat{p}_1 = \frac{8}{30} = .267 \)

For confidence coefficient .90, \( \alpha = .10 \) and \( \alpha / 2 = .10 / 2 = .05 \). From Table III, Appendix A, \( z_{.05} = 1.645 \). The 90% confidence interval is:

\[
\hat{p}_1 \pm z_{.05} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n}} = .267 \pm 1.645 \sqrt{\frac{.267(.733)}{30}} = .267 \pm .133 \Rightarrow (.134, .400)
\]

We are 90% confident that the true proportion of museums world-wide that use total visitors as their performance measure is between .134 and .400.

8.44 Some preliminary calculations are:

\[
E_1 = np_{1,0} = 7,506(.70) = 5,254.2 \\
E_2 = np_{2,0} = 7,506 (.15) = 1,125.9 \\
E_3 = np_{3,0} = 7,506 (.10) = 750.6 \\
E_4 = np_{4,0} = 7,506 (.05) = 375.3
\]

To determine if the data refute the claim, we test:

\[
H_0: \ p_1 = .70, \ p_2 = .15, \ p_3 = .10, \text{ and } \ p_4 = .05 \\
H_a: \ \text{At least one } \ p_i \text{ differs from it hypothesized value}
\]

The test statistic is

\[
\chi^2 = \sum \frac{(n_i - E_i)^2}{E_i}
\]

\[
= \frac{(5,079 - 5,254.2)^2}{5,254.2} + \frac{(1,042 - 1,125.9)^2}{1,125.9} + \frac{(848 - 750.6)^2}{750.6} + \frac{(537 - 375.3)^2}{375.3}
\]

\[
= 94.40
\]

Since no \( \alpha \) level was given, we will use \( \alpha = .05 \). The rejection region requires \( \alpha = .05 \) in the upper tail of the \( \chi^2 \) distribution with \( df = k - 1 = 4 - 1 = 3 \). From Table V, Appendix A, \( \chi^2_{.05} = 7.81473 \). The rejection region is \( \chi^2 > 7.81473 \).

Since the observed value of the test statistic falls in the rejection region \( (\chi^2 = 94.40 > 7.81473) \), \( H_0 \) is rejected. There is sufficient evidence to indicate the data refute the claim about the proportions claimed by the sociologist at \( \alpha = .05 \).

8.46 a. \( \hat{p}_1 = \frac{88}{504} = .175 \)

For confidence coefficient .90, \( \alpha = .10 \) and \( \alpha / 2 = .10 / 2 = .05 \). From Table III, Appendix A, \( z_{.05} = 1.645 \). The 90% confidence interval is:
\[ \hat{p}_1 \pm z_{.05} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n}} = .175 \pm 1.645 \sqrt{\frac{.175(.825)}{504}} = .175 \pm .028 \Rightarrow (.147, .203) \]

We are 90% confident that the true proportion of melt ponds in the Canadian Arctic that have first-year ice is between .147 and .203.

b. Some preliminary calculations:

\[ E_1 = np_{1,0} = 504(.15) = 75.6 \quad E_2 = np_{2,0} = 504(.40) = 201.6 \]
\[ E_3 = np_{1,0} = 504(.45) = 226.8 \]

To determine if the true proportions follow the proportions given, we test:

\[ H_0: \ p_1 = .15, \ p_2 = .40, \ p_3 = .45 \]
\[ H_a: \ At \ least \ one \ p_i \ differs \ from \ its \ hypothesized \ value \]

The test statistic is

\[ \chi^2 = \sum \left( \frac{n_i - E_i}{E_i} \right)^2 = \frac{[88 - 75.6]^2}{75.6} + \frac{[196 - 201.6]^2}{201.6} + \frac{[220 - 226.8]^2}{226.8} = 2.39 \]

The rejection region requires \( \alpha = .01 \) in the upper tail of the \( \chi^2 \) distribution with df \( = k - 1 = 3 - 1 = 2 \). From Table V, Appendix A, \( \chi^2_{.01} = 9.21034 \). The rejection region is \( \chi^2 > 9.21034 \).

Since the observed value of the test statistic does not fall in the rejection region \( (\chi^2 = 2.39 < 9.21034) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate at least one proportion differs from its hypothesized value at \( \alpha = .01 \).

8.48 Some preliminary calculations are:

\[ E_1 = np_{1,0} = 435(.28) = 121.8 \quad E_2 = np_{2,0} = 435(.04) = 17.4 \]
\[ E_3 = np_{3,0} = 435(.02) = 8.7 \quad E_4 = np_{4,0} = 435(.66) = 287.1 \]

To determine if the House of Representatives is representative of the religious affiliation of their constituents in the U.S., we test:

\[ H_0: \ p_1 = .28, \ p_2 = .04, \ p_3 = .02, \ p_4 = .66 \]
\[ H_a: \ At \ least \ one \ p_i \ differs \ from \ its \ hypothesized \ value \]

The test statistic is

\[ \chi^2 = \sum \left( \frac{n_i - E_i}{E_i} \right)^2 = \frac{[117 - 121.8]^2}{121.8} + \frac{[61 - 17.4]^2}{17.4} + \frac{[30 - 8.7]^2}{8.7} + \frac{[227 - 287.1]^2}{287.1} = 174.17 \]
Since no $\alpha$ value was given, we will use $\alpha = .05$. The rejection region requires $\alpha = .05$ in the upper tail of the $\chi^2$ distribution with df = $k - 1 = 4 - 1 = 3$. From Table V, Appendix A, $\chi^2_{.05} = 7.81473$. The rejection region is $\chi^2 > 7.81473$.

Since the observed value of the test statistic falls in the rejection region ($\chi^2 = 174.14 > 7.81473$), $H_0$ is rejected. There is sufficient evidence to indicate at least one proportion differs from its hypothesized value at $\alpha = .05$.

8.50 A two-way contingency table presents multinomial count data classified on two scales of classification.

8.52 The statement “One goal of a contingency table analysis is to determine if the two classifications are dependent” is true.

8.54 a. $df = (r - 1)(c - 1) = (5 - 1)(5 - 1) = 16$. From Table V, Appendix A, $\chi^2_{16} = 26.2962$. The rejection region is $\chi^2 > 26.2962$.

b. $df = (r - 1)(c - 1) = (3 - 1)(6 - 1) = 10$. From Table V, Appendix A, $\chi^2_{10} = 15.9871$. The rejection region is $\chi^2 > 15.9871$.

c. $df = (r - 1)(c - 1) = (2 - 1)(3 - 1) = 2$. From Table V, Appendix A, $\chi^2_{2} = 9.21034$. The rejection region is $\chi^2 > 9.21034$.

8.56 a. To convert the frequencies to percentages, divide the numbers in each column by the column total and multiply by 100. Also, divide the row totals by the overall total and multiply by 100. The column totals are 25, 64, and 78, while the row totals are 96 and 71. The overall sample size is 167. The table of percentages are:

<table>
<thead>
<tr>
<th></th>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Row 1</strong></td>
<td>9/25 = 36%</td>
<td>34/64 = 53.1%</td>
<td>53/78 = 67.9%</td>
</tr>
<tr>
<td></td>
<td>96/167 = 57.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Row 2</strong></td>
<td>16/25 = 64%</td>
<td>30/64 = 46.9%</td>
<td>25/78 = 32.1%</td>
</tr>
<tr>
<td></td>
<td>71/167 = 42.5%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
b. 

![Chart of Percent vs Column]

<table>
<thead>
<tr>
<th>Column Number</th>
<th>Percent in Row A1</th>
<th>Percent in Row A2</th>
<th>Percent in Row A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>70</td>
<td>70</td>
<td>47.0%</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>53</td>
<td>32.5%</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>70</td>
<td>49.3%</td>
</tr>
</tbody>
</table>

Chart of Percent vs Column

8.58 To convert the frequencies to percentages, divide the numbers in each column by the column total and multiply by 100. Also, divide the row totals by the overall total and multiply by 100.

<table>
<thead>
<tr>
<th>Row</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>( \frac{40}{134} \cdot 100% = 29.9% )</td>
<td>( \frac{72}{163} \cdot 100% = 44.2% )</td>
<td>( \frac{42}{142} \cdot 100% = 29.6% )</td>
<td>( \frac{154}{439} \cdot 100% = 35.1% )</td>
</tr>
<tr>
<td>A2</td>
<td>( \frac{63}{134} \cdot 100% = 47.0% )</td>
<td>( \frac{53}{163} \cdot 100% = 32.5% )</td>
<td>( \frac{70}{142} \cdot 100% = 49.3% )</td>
<td>( \frac{186}{439} \cdot 100% = 42.4% )</td>
</tr>
<tr>
<td>A3</td>
<td>( \frac{31}{134} \cdot 100% = 23.1% )</td>
<td>( \frac{38}{163} \cdot 100% = 23.3% )</td>
<td>( \frac{30}{142} \cdot 100% = 21.1% )</td>
<td>( \frac{99}{439} \cdot 100% = 22.6% )</td>
</tr>
</tbody>
</table>

![Chart of Percent vs Column]

Since the columns are not the same height, there is evidence to indicate that the row and column classifications are dependent.
b. The graph for row A2 is:

![Chart of Percent vs Column]

Since the columns are not the same height, there is evidence to indicate that the row and column classifications are dependent.

c. The graph for row A3 is:

![Chart of Percent vs Column]

Since the columns are almost the same height, there is evidence to indicate that the row and column classifications are independent.

8.60 a. 3-photo group: \( \hat{p}_1 = \frac{19}{32} = .594 \); 6-photo group: \( \hat{p}_2 = \frac{19}{32} = .594 \);

12-photo group: \( \hat{p}_3 = \frac{15}{32} = .469 \)
The students in the 12-photos per page yielded the lowest proportion.

b. The contingency table is:

<table>
<thead>
<tr>
<th>Photo Group</th>
<th>Targeted Mugshot</th>
<th>Non-Targeted Mugshot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Selected</td>
<td>Not Selected</td>
</tr>
<tr>
<td>3-photo</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>6-photo</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>12-photo</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>Totals</td>
<td>53</td>
<td>43</td>
</tr>
</tbody>
</table>

c. Some preliminary calculations are:

\[
\hat{E}_{11} = \frac{R_1 C_1}{n} = \frac{32(53)}{96} = 17.667 \\
\hat{E}_{12} = \frac{R_1 C_2}{n} = \frac{32(43)}{96} = 14.333 \\
\hat{E}_{21} = \frac{R_2 C_1}{n} = \frac{32(53)}{96} = 17.667 \\
\hat{E}_{22} = \frac{R_2 C_2}{n} = \frac{32(43)}{96} = 14.333 \\
\hat{E}_{31} = \frac{R_3 C_1}{n} = \frac{32(53)}{96} = 17.667 \\
\hat{E}_{32} = \frac{R_3 C_2}{n} = \frac{32(43)}{96} = 14.333
\]

To determine if differences in the proportions who selected the target mugshot among the three photo groups exist, we test:

\[H_0: \text{Target mugshot and photo group are independent} \]
\[H_a: \text{Target mugshot and photo group are dependent} \]

The test statistic is

\[
\chi^2 = \sum \frac{[n_{ij} - \hat{E}_{ij}]^2}{\hat{E}_{ij}}\]

\[
= \frac{[19 - 17.667]^2}{17.667} + \frac{[13 - 14.333]^2}{14.333} + \frac{[19 - 17.667]^2}{17.667} \\
+ \frac{[13 - 14.333]^2}{14.333} + \frac{[15 - 17.667]^2}{17.667} + \frac{[17 - 14.333]^2}{14.333} = 1.348
\]

The rejection region requires \( \alpha = .10 \) in the upper tail of the \( \chi^2 \) distribution with \( df = (r - 1)(c - 1) = (3 - 1)(2 - 1) = 2 \). From Table V, Appendix A, \( \chi^2_{.10} = 4.60517 \). The rejection region is \( \chi^2 > 4.60517 \).

Since the observed value of the test statistics does not fall in the rejection region \( \chi^2 = 1.348 < 4.60517 \), \( H_0 \) is not rejected. There is insufficient evidence to indicate differences in the proportions who selected the target mugshot among the three photo groups exist at \( \alpha = .10 \).
8.62  a. One of the variables is Therapy which has 4 levels: Prayer, MIT, Prayer and MIT, and Standard. The second variable is Major adverse cardiovascular event which has 2 levels: Yes and No.

b. To determine if major adverse cardiovascular event depends on type of therapy, we test:

\[ H_0: \text{Major adverse cardiovascular event does not depend on type of therapy} \]
\[ H_a: \text{Major adverse cardiovascular event depends on type of therapy} \]

c. From the printout, the test statistic is \( \chi^2 = 1.828 \) and the \( p \)-value is \( p = .609. \)

Since the \( p \)-value is not less than \( \alpha = .10 \) \( (p = .609 \nsim .10), H_0 \) is not rejected. There is insufficient evidence to indicate that major adverse cardiovascular event depends on type of therapy at \( \alpha = .10 \). The researchers can conclude that there is no evidence to support the thought that MIT and prayer help to prevent major adverse cardiovascular events.

8.64  Some preliminary calculations are:

\[
\hat{E}_{11} = \frac{R_1C_1}{n} = \frac{234(40)}{437} = 21.419 \\
\hat{E}_{12} = \frac{R_1C_2}{n} = \frac{234(397)}{437} = 212.581 \\
\hat{E}_{21} = \frac{R_2C_1}{n} = \frac{203(40)}{437} = 18.581 \\
\hat{E}_{22} = \frac{R_2C_2}{n} = \frac{203(397)}{437} = 184.419
\]

To determine if the response rate of air traffic controllers to mid-air collision alarms differs for true and false alerts, we test:

\[ H_0: \text{Response rates and alerts are independent} \]
\[ H_a: \text{Response rates and alerts are dependent} \]

The test statistic is

\[
\chi^2 = \sum \frac{[n_{ij} - \hat{E}_{ij}]^2}{\hat{E}_{ij}} = \frac{[3 - 21.419]^2}{21.419} + \frac{[37 - 212.581]^2}{212.581} + \frac{[166 - 184.419]^2}{184.419} = 37.533
\]

The rejection region requires \( \alpha = .05 \) in the upper tail of the \( \chi^2 \) distribution with \( df = (r - 1)(c - 1) = (2 - 1)(2 - 1) = 1 \). From Table V, Appendix A, \( \chi^2_{0.05} = 3.84146 \). The rejection region is \( \chi^2 > 3.84146 \).
Since the observed value of the test statistics falls in the rejection region
\( (\chi^2 = 37.533 > 3.84146) \), \( H_0 \) is rejected. There is sufficient evidence to indicate the response rate of air traffic controllers to mid-air collision alarms differs for true and false alerts at \( \alpha = 0.05 \).

There is a difference in the response rates between true and false alerts. There is a higher response rate to true alerts than to false alerts. This indicates that air traffic controllers do not tend to ignore true alerts in the future.

8.66 Some preliminary calculations are:

\[
\hat{E}_{11} = \frac{R_{11}}{n} = \frac{32(32)}{96} = 10.67 \\
\hat{E}_{12} = \frac{R_{12}}{n} = \frac{32(64)}{96} = 21.33 \\
\hat{E}_{21} = \frac{R_{21}}{n} = \frac{32(32)}{96} = 10.67 \\
\hat{E}_{22} = \frac{R_{22}}{n} = \frac{32(64)}{96} = 21.33
\]

To determine whether the proportion of subjects who select menus consistent with the theory depends on goal condition, we test:

\[ H_0: \text{Consistency to theory and Goal condition are independent} \]
\[ H_a: \text{Consistency to theory and Goal condition are dependent} \]

The test statistic is

\[
\chi^2 = \sum \sum \frac{[n_{ij} - \hat{E}_{ij}]^2}{\hat{E}_{ij}} = \frac{[15 - 10.67]^2}{10.67} + \frac{[17 - 21.33]^2}{21.33} \\
+ \frac{[14 - 10.67]^2}{10.67} + \frac{[18 - 21.33]^2}{21.33} + \frac{[3 - 10.67]^2}{10.67} + \frac{[29 - 21.33]^2}{21.33} = 12.467
\]

The rejection region requires \( \alpha = 0.01 \) in the upper tail of the \( \chi^2 \) distribution with \( df = (r - 1)(c - 1) = (3 - 1)(2 - 1) = 2 \). From Table V, Appendix A, \( \chi^2_{0.01} = 9.21034 \). The rejection region is \( \chi^2 > 9.21034 \).

Since the observed value of the test statistic falls in the rejection region
\( (\chi^2 = 12.467 > 9.21034) \), \( H_0 \) is rejected. There is sufficient evidence that the proportion of subjects who select menus consistent with the theory depends on goal condition at \( \alpha = 0.01 \).

8.68 Some preliminary calculations are:

\[
\hat{E}_{11} = \frac{R_{11}}{n} = \frac{194(237)}{363} = 126.66 \\
\hat{E}_{12} = \frac{R_{12}}{n} = \frac{194(91)}{363} = 48.63
\]
Comparing Population Proportions

\[ \hat{E}_{13} = \frac{R_C}{n} = \frac{194(35)}{363} = 18.71 \quad \hat{E}_{21} = \frac{R_C}{n} = \frac{169(237)}{363} = 110.34 \]
\[ E_{22} = \frac{R_C}{n} = \frac{169(91)}{363} = 42.37 \quad \hat{E}_{23} = \frac{R_C}{n} = \frac{169(35)}{363} = 16.29 \]

To determine if the percentages of moths caught by the three traps depends on day of the week, we test:

- \( H_0: \) Traps and Day of Week are independent
- \( H_a: \) Traps and Day of Week are dependent

The test statistic is

\[ \chi^2 = \sum \sum \frac{(n_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} = \frac{[136 - 126.66]^2}{126.66} + \frac{[41 - 48.63]^2}{48.63} + \frac{[17 - 18.71]^2}{18.71} + \frac{[101 - 110.34]^2}{110.34} + \frac{[50 - 42.37]^2}{42.37} + \frac{[18 - 16.29]^2}{16.29} = 4.386 \]

Since no \( \alpha \) is given, we will use \( \alpha = .05 \). The rejection region requires \( \alpha = .05 \) in the upper tail of the \( \chi^2 \) distribution with \( df = (r - 1)(c - 1) = (2 - 1)(3 - 1) = 2 \). From Table V, Appendix A, \( \chi^2_{.05} = 5.99147 \). The rejection region is \( \chi^2 > 5.99147 \).

Since the observed value of the test statistic does not fall in the rejection region \( (\chi^2 = 4.386 > 5.99147) \), \( H_0 \) is not rejected. There is insufficient evidence that the percentages of moths caught by the three traps depend on day of the week at \( \alpha = .05 \).

8.70 a. \( \hat{E}_{11} = \frac{R_C}{n} = \frac{5(13)}{73} = .89 \quad \hat{E}_{12} = \frac{R_C}{n} = \frac{5(17)}{73} = 1.16 \)
\( \hat{E}_{13} = \frac{R_C}{n} = \frac{5(43)}{73} = 2.95 \quad \hat{E}_{21} = \frac{R_C}{n} = \frac{5(13)}{73} = 9.08 \)
\( \hat{E}_{22} = \frac{R_C}{n} = \frac{51(17)}{73} = 11.88 \quad \hat{E}_{23} = \frac{R_C}{n} = \frac{51(43)}{73} = 30.04 \)
\( \hat{E}_{31} = \frac{R_C}{n} = \frac{17(13)}{73} = 3.03 \quad \hat{E}_{32} = \frac{R_C}{n} = \frac{17(17)}{73} = 3.96 \)
\( \hat{E}_{33} = \frac{R_C}{n} = \frac{17(43)}{73} = 10.01 \)

No, the assumptions are met. Several cells have expected counts less than 5. In order for the test to be valid, all expected values need to be 5 or more.
b. The new table would be:

<table>
<thead>
<tr>
<th>Abundance Seedlings</th>
<th>Dwarf Shrub</th>
<th>Grasses</th>
<th>Herbs</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS &amp; SR</td>
<td>8</td>
<td>3</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>SA</td>
<td>5</td>
<td>14</td>
<td>32</td>
<td>51</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>17</td>
<td>43</td>
<td>73</td>
</tr>
</tbody>
</table>

No, the assumptions are still not met. There is one cell with an expected count less than 5.

c. The new table would be:

<table>
<thead>
<tr>
<th>Abundance Seedlings</th>
<th>Dwarf Shrub/Grasses</th>
<th>Type of Plant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS &amp; SR</td>
<td>11</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>SA</td>
<td>19</td>
<td>32</td>
<td>51</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>43</td>
<td>73</td>
</tr>
</tbody>
</table>

Now, all of the assumptions appear to be met.

d. To determine if the plant type and seedling abundance are related, we test:

- $H_0$: Plant type and seedling abundance are independent
- $H_a$: Plant type and seedling abundance are dependent
The test statistic is

\[ \chi^2 = \sum \sum \frac{(n_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} \]

\[ = \frac{(11 - 9.04)^2}{9.04} + \frac{(11 - 12.96)^2}{12.96} + \frac{(19 - 20.96)^2}{20.96} + \frac{(32 - 30.04)^2}{30.04} = 1.03 \]

The rejection region requires \( \alpha = .10 \) in the upper tail of the \( \chi^2 \) distribution with \( df = (r - 1)(c - 1) = (2 - 1)(2 - 1) = 1 \). From Table V, Appendix A, \( \chi^2_{10} = 2.70554 \). The rejection region is \( \chi^2 > 2.70554 \).

Since the observed value of the test statistic does not fall in the rejection region \( (\chi^2 = 1.03 > 2.70554) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate plant type and seedling abundance are related at \( \alpha = .10 \).

### 8.72 Some preliminary calculations are:

\[ \hat{E}_{i1} = \frac{R_iC_1}{n} = \frac{57(60)}{171} = 20.000 \]

\[ \hat{E}_{i2} = \frac{R_iC_2}{n} = \frac{57(111)}{171} = 37.000 \]

\[ \hat{E}_{21} = \frac{R_2C_1}{n} = \frac{58(60)}{171} = 20.351 \]

\[ \hat{E}_{22} = \frac{R_2C_2}{n} = \frac{58(111)}{171} = 37.649 \]

\[ \hat{E}_{31} = \frac{R_3C_1}{n} = \frac{56(60)}{171} = 19.649 \]

\[ \hat{E}_{32} = \frac{R_3C_2}{n} = \frac{56(111)}{171} = 36.351 \]

To determine if the option choice depends on emotional state, we test:

- \( H_0: \) Option choice and emotional state are independent
- \( H_a: \) Option choice and emotional state are dependent

The test statistic is

\[ \chi^2 = \sum \frac{(n_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} = \frac{(45 - 20)^2}{220} + \frac{(12 - 37)^2}{37} + \frac{(8 - 20.351)^2}{20.351} + \frac{(50 - 37.649)^2}{37.649} + \frac{(7 - 19.649)^2}{19.649} + \frac{(49 - 36.351)^2}{36.351} = 72.234 \]

The rejection region requires \( \alpha = .10 \) in the upper tail of the \( \chi^2 \) distribution with \( df = (r - 1)(c - 1) = (3 - 1)(2 - 1) = 2 \). From Table V, Appendix A, \( \chi^2_{10} = 4.60517 \). The rejection region is \( \chi^2 > 4.60517 \).
Since the observed value of the test statistics falls in the rejection region 
\( \chi^2 = 72.234 > 4.60517 \), \( H_0 \) is rejected. There is sufficient evidence to indicate the option 
choice depends on emotional state at \( \alpha = .10 \).

8.74  a. For First Trial (#1), the output from using MINITAB is:

**Tabulated statistics: Condition, Switch**

Using frequencies in Number

<table>
<thead>
<tr>
<th>Rows: Condition</th>
<th>Columns: Switch</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>All</td>
<td></td>
</tr>
<tr>
<td>Empty</td>
<td>17</td>
<td>10</td>
<td>27</td>
<td>20.50 6.50</td>
</tr>
<tr>
<td>Steroids</td>
<td>22</td>
<td>5</td>
<td>27</td>
<td>20.50 6.50</td>
</tr>
<tr>
<td>Steroids2</td>
<td>19</td>
<td>8</td>
<td>27</td>
<td>20.50 6.50</td>
</tr>
<tr>
<td>Vanish</td>
<td>24</td>
<td>3</td>
<td>27</td>
<td>20.50 6.50</td>
</tr>
<tr>
<td>All</td>
<td>82</td>
<td>26</td>
<td>108</td>
<td>82.00 26.00</td>
</tr>
</tbody>
</table>

Cell Contents: Count

Pearson Chi-Square = 5.876, DF = 3, P-Value = 0.118
Likelihood Ratio Chi-Square = 6.096, DF = 3, P-Value = 0.107

To determine if switching boxes depends on condition, we test:

\[ H_0: \text{ Switching boxes and condition are independent} \]
\[ H_a: \text{ Switching boxes and condition are dependent} \]

From the printout, the test statistic is \( \chi^2 = 5.876 \) and the \( p \)-value is \( p = .118 \).

Since no \( \alpha \) was given, we will use \( \alpha = .05 \). Since the \( p \)-value is not less than \( \alpha = .05 \) 
\( (p = .118 \not< .05) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate switching 
boxes depends on condition for First Trial at \( \alpha = .05 \).
For **Last Trial (#23)**, the output from using MINITAB is:

**Tabulated statistics: Condition, Switch**

Using frequencies in Number2

<table>
<thead>
<tr>
<th></th>
<th>Condition</th>
<th>Switch</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>All</td>
</tr>
<tr>
<td>Empty</td>
<td>4</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>8.25</td>
<td>18.75</td>
<td>27.00</td>
</tr>
<tr>
<td>Steroids</td>
<td>6</td>
<td>21</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>8.25</td>
<td>18.75</td>
<td>27.00</td>
</tr>
<tr>
<td>Steroids2</td>
<td>8</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>8.25</td>
<td>18.75</td>
<td>27.00</td>
</tr>
<tr>
<td>Vanish</td>
<td>15</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>8.25</td>
<td>18.75</td>
<td>27.00</td>
</tr>
<tr>
<td>All</td>
<td>33</td>
<td>75</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>33.00</td>
<td>75.00</td>
<td>108.00</td>
</tr>
</tbody>
</table>

Cell Contents: Count


Pearson Chi-Square = 12.000, DF = 3, P-Value = 0.007
Likelihood Ratio Chi-Square = 11.780, DF = 3, P-Value = 0.008

To determine if switching boxes depends on condition, we test:

\[ H_0: \text{Switching boxes and condition are independent} \]
\[ H_a: \text{Switching boxes and condition are dependent} \]

From the printout, the test statistic is \( \chi^2 = 12.000 \) and the \( p \)-value is \( p = 0.007 \).

Since no \( \alpha \) was given, we will use \( \alpha = 0.05 \). Since the \( p \)-value is less than \( \alpha = 0.05 \) \( (p = 0.007 < 0.05) \), \( H_0 \) is rejected. There is sufficient evidence to indicate switching boxes depends on condition for Last Trial at \( \alpha = 0.05 \).
b. For **Condition Empty**, the output from using MINITAB is:

**Tabulated statistics: Trial, Switch**

Using frequencies in Number1

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Yes</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>10.50</td>
<td>16.50</td>
<td>27.00</td>
</tr>
<tr>
<td>23</td>
<td>4</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>10.50</td>
<td>16.50</td>
<td>27.00</td>
</tr>
<tr>
<td>All</td>
<td>21</td>
<td>33</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>21.00</td>
<td>33.00</td>
<td>54.00</td>
</tr>
</tbody>
</table>

Cell Contents: Count

<table>
<thead>
<tr>
<th></th>
<th>Expected count</th>
</tr>
</thead>
</table>

Pearson Chi-Square = 13.169, DF = 1, P-Value = 0.000  
Likelihood Ratio Chi-Square = 13.924, DF = 1, P-Value = 0.000

To determine if switching boxes depends on trial number, we test:

\[ H_0: \text{Switching boxes and trial number are independent} \]
\[ H_a: \text{Switching boxes and trial number are dependent} \]

From the printout, the test statistic is \( \chi^2 = 13.169 \) and the \( p \)-value is \( p = .000 \).

Since no \( \alpha \) was given, we will use \( \alpha = .05 \). Since the \( p \)-value is less than \( \alpha = .05 \) \( (p = .000 < .05) \), \( H_0 \) is rejected. There is sufficient evidence to indicate switching boxes depends on Trial number for Empty condition at \( \alpha = .05 \).
For **Condition Vanish**, the output from using MINITAB is:

**Tabulated statistics: Trial, Switch**

Using frequencies in Number2  
Rows: Trial  Columns: Switch  

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Yes</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>19.50</td>
<td>7.50</td>
<td>27.00</td>
</tr>
<tr>
<td>23</td>
<td>15</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>19.50</td>
<td>7.50</td>
<td>27.00</td>
</tr>
<tr>
<td>All</td>
<td>39</td>
<td>15</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>39.00</td>
<td>15.00</td>
<td>54.00</td>
</tr>
</tbody>
</table>

Cell Contents:  
Count  
Expected count

Pearson Chi-Square = 7.477, DF = 1, P-Value = 0.006  
Likelihood Ratio Chi-Square = 7.878, DF = 1, P-Value = 0.005

To determine if switching boxes depends on trial number, we test:

- $H_0$: Switching boxes and trial number are independent  
- $H_a$: Switching boxes and trial number are dependent

From the printout, the test statistic is $\chi^2 = 7.477$ and the $p$-value is $p = .006$.

Since no $\alpha$ was given, we will use $\alpha = .05$. Since the $p$-value is less than $\alpha = .05$ ($p = .006 < .05$), $H_0$ is rejected. There is sufficient evidence to indicate switching boxes depends on Trial number for Vanish condition at $\alpha = .05$. 

Copyright © 2013 Pearson Education, Inc.
For **Condition Steroids**, the output from using MINITAB is:

**Tabulated statistics: Trial, Switch**

Using frequencies in Number3

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Yes</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22</td>
<td>5</td>
<td>27</td>
</tr>
<tr>
<td>14</td>
<td>13</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>6</td>
<td>21</td>
<td>27</td>
</tr>
<tr>
<td>14</td>
<td>13</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>28</td>
<td>26</td>
<td>54</td>
</tr>
<tr>
<td>28</td>
<td>26</td>
<td>54</td>
<td></td>
</tr>
</tbody>
</table>

Cell Contents: Count

Pearson Chi-Square = 18.989, DF = 1, P-Value = 0.000
Likelihood Ratio Chi-Square = 20.307, DF = 1, P-Value = 0.000

To determine if switching boxes depends on trial number, we test:

- $H_0$: Switching boxes and trial number are independent
- $H_a$: Switching boxes and trial number are dependent

From the printout, the test statistic is $\chi^2 = 18.989$ and the $p$-value is $p = .000$.

Since no $\alpha$ was given, we will use $\alpha = .05$. Since the $p$-value is less than $\alpha = .05$ ($p = .000 < .05$), $H_0$ is rejected. There is sufficient evidence to indicate switching boxes depends on Trial number for Steroids condition at $\alpha = .05$. 

Copyright © 2013 Pearson Education, Inc.
For **Condition Steroids2**, the output from using MINITAB is:

**Tabulated statistics: Trial, Switch**

Using frequencies in Number

**Rows:** Trial  **Columns:** Switch

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Yes</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>13.50</td>
<td>13.50</td>
<td>27.00</td>
</tr>
<tr>
<td>23</td>
<td>8</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>13.50</td>
<td>13.50</td>
<td>27.00</td>
</tr>
<tr>
<td>All</td>
<td>27</td>
<td>27</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>27.00</td>
<td>27.00</td>
<td>54.00</td>
</tr>
</tbody>
</table>

**Cell Contents:**

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>Expected count</th>
</tr>
</thead>
</table>

Pearson Chi-Square $= 8.963$, DF = 1, P-Value $= 0.003$
Likelihood Ratio Chi-Square $= 9.229$, DF = 1, P-Value $= 0.002$

To determine if switching boxes depends on trial number, we test:

**$H_0$:** Switching boxes and trial number are independent  
**$H_a$:** Switching boxes and trial number are dependent

From the printout, the test statistic is $\chi^2 = 8.963$ and the $p$-value is $p = .003$.

Since no $\alpha$ was given, we will use $\alpha = .05$. Since the $p$-value is less than $\alpha = .05$ ($p = .003 < .05$), $H_0$ is rejected. There is sufficient evidence to indicate switching boxes depends on Trial number for Steroids2 condition at $\alpha = .05$.

c. For all conditions, switching depends on trial number. For First trial, switching does not depend on condition. For Last trial, switching does depend on condition. Thus, both condition and trial number influence a subject to switch choices.

8.76 In a one-way chi-square analysis, one must hypothesize the proportions in each of the cells. In a two-way chi-square analysis, one does not need to specify the proportions in each of the cells. In a two-way chi-square analysis, one tests to see if the two categorical variables are independent or not.

8.78 a. Some preliminary calculations are:

If all the categories are equally likely,

\[ p_{1,0} = p_{2,0} = p_{3,0} = p_{4,0} = p_{5,0} = .2 \]

\[ E_1 = E_2 = E_3 = E_4 = E_5 = np_{1,0} = 150(.2) = 30 \]
To determine if the categories are not equally likely, we test:

\[ H_0: \ p_1 = p_2 = p_3 = p_4 = p_5 = .2 \]
\[ H_a: \ At \ least \ one \ probability \ is \ different \ from \ .2 \]

The test statistic is

\[ \chi^2 = \sum \frac{(n_i - E_i)^2}{E_i} = \frac{(28 - 30)^2}{30} + \frac{(35 - 30)^2}{30} + \frac{(33 - 30)^2}{30} + \frac{(25 - 30)^2}{30} + \frac{(29 - 30)^2}{30} = 2.133 \]

The rejection region requires \( \alpha = .10 \) in the upper tail of the \( \chi^2 \) distribution with \( df = k - 1 = 5 - 1 = 4 \). From Table V, Appendix A, \( \chi^2_{10} = 7.77944 \). The rejection region is \( \chi^2 > 7.77944 \).

Since the observed value of the test statistic does not fall in the rejection region \( (\chi^2 = 2.133 \not> 7.77944) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate the categories are not equally likely at \( \alpha = .10 \).

b. \( \hat{p}_2 = \frac{35}{150} = .233 \)

For confidence coefficient .90, \( \alpha = .10 \) and \( \alpha / 2 = .10 / 2 = .05 \). From Table III, Appendix A, \( z_{.05} = 1.645 \). The confidence interval is:

\[ \hat{p}_2 \pm z_{.05} \sqrt{\frac{\hat{p}_2 \hat{q}_2}{n_2}} = .233 \pm 1.645 \sqrt{\frac{.233(.767)}{150}} \Rightarrow .233 \pm .057 \Rightarrow (.176, .290) \]

8.80 From the printout, the 95% confidence interval for the difference in mean seabird densities of oiled and unoiled transects is \((-2.93, 2.49)\). We are 95% confident that the true difference in mean seabird densities of oiled and unoiled transects is between \(-2.93 \) and \(2.49\). Since 0 is contained in the interval, there is no evidence to indicate that the mean seabird densities are different for the oiled and unoiled transects.

8.82 a. To determine if the opinions of Internet users are evenly divided among the four categories, we test:

\[ H_0: \ p_1 = p_2 = p_3 = p_4 = .25 \]
\[ H_a: \ At \ least \ one \ of \ the \ probabilities \ differs \ from \ .25 \]

where
\[ p_1 = \text{proportion who agree strongly} \]
\[ p_2 = \text{proportion who agree somewhat} \]
\[ p_3 = \text{proportion who disagree somewhat} \]
\[ p_4 = \text{proportion who disagree strongly} \]
b. \[ E_1 = np_{1.0} = 328(.25) = 82 = E_2 = E_3 = E_4 = 82 \]

\[
\chi^2 = \sum \frac{(n_i - E_i)^2}{E_i} = \frac{(59 - 82)^2}{82} + \frac{(108 - 82)^2}{82} + \frac{(82 - 82)^2}{82} + \frac{(79 - 82)^2}{82} = 14.80
\]

The rejection region requires \( \alpha = .05 \) in the upper tail of the \( \chi^2 \) distribution with \( df = k - 1 = 4 - 1 = 3 \). From Table V, Appendix A, \( \chi^2_{0.05} = 7.81473 = 7.81473 \). The rejection region is \( \chi^2 > 7.81473 \).

Since the observed value of the test statistic falls in the rejection \( \chi^2 = 14.80 > 7.81473 \), \( H_0 \) is rejected. There is sufficient evidence to indicate the opinions of Internet users are not evenly divided at \( \alpha = .05 \).

C. The Type I error is concluding the opinions of Internet users are not evenly divided among the four categories when, in fact, they are.

The Type II error is concluding the opinions of Internet users are evenly divided among the four categories when, in fact, they are not.

d. 1. A multinomial experiment has been conducted. This is generally satisfied by taking a random sample from the population of interest.
   2. We must assume the sample size \( n \) is large enough so that the expected cell count, \( E(n_i) \), will be equal to 5 or more for every cell.

8.84 a. The two qualitative variables were grade and reading ability. Grade had two levels: 4th grade and 5th grade. Reading ability had two levels: normal and reading disability.

b. The contingency table would be:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Reading Ability</th>
<th>Normal</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th</td>
<td>32</td>
<td>55</td>
<td>87</td>
</tr>
<tr>
<td>5th</td>
<td>34</td>
<td>40</td>
<td>74</td>
</tr>
<tr>
<td>Totals</td>
<td>66</td>
<td>95</td>
<td>161</td>
</tr>
</tbody>
</table>

c. The expected cell counts for the cells would be:

\[
\hat{E}_{11} = \frac{RC_1}{n} = \frac{87(66)}{161} = 35.66 \quad \hat{E}_{12} = \frac{RC_2}{n} = \frac{87(95)}{161} = 51.34
\]

\[
\hat{E}_{21} = \frac{RC_1}{n} = \frac{74(66)}{161} = 30.34 \quad \hat{E}_{22} = \frac{RC_2}{n} = \frac{74(95)}{161} = 43.66
\]
d. The test statistic is

\[ \chi^2 = \sum \sum \frac{(n_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} = \frac{(32 - 35.66)^2}{35.66} + \frac{(55 - 51.34)^2}{51.34} + \frac{(34 - 30.34)^2}{30.34} + \frac{(40 - 43.66)^2}{43.66} \]

\[ = 1.38 \]

e. The rejection region requires \( \alpha = .10 \) in the upper tail of the \( \chi^2 \) distribution with
df \( = (r - 1)(c - 1) = (2 - 1)(2 - 1) = 1 \). From Table V, Appendix A, \( \chi^2_{10} = 2.70554 \). The rejection region is \( \chi^2 > 2.70554 \).

f. To determine if there is a link between reading disability and grade level, we test:

\[ H_0: \text{Reading Disability and Grade Level are independent} \]
\[ H_a: \text{Reading Disability and Grade Level are dependent} \]

The test statistic is \( \chi^2 = 1.38 \).

The rejection region is \( \chi^2 > 2.70554 \).

Since the observed value of the test statistic does not fall in the rejection region \( (\chi^2 = 1.38 \neq 2.70554) \), \( H_0 \) is not rejected. There is insufficient evidence of a link between reading disability and grade level at \( \alpha = .10 \).

8.86 a. The qualitative variable of interest is the response groups to “Made in the USA”. There are 4 levels of this variable: “Made in USA” means “100%” of labor and materials are produced in the US, “Made in USA” means “75 to 99%” of labor and materials are produced in the US, “Made in USA” means “50 to 74%” of labor and materials are produced in the US, and “Made in USA” means “less than 50%” of labor and materials are produced in the US.

b. From the problem, \( p_1 = .50, p_2 = .25, p_3 = .20, \) and \( p_4 = .05 \).

c. To determine whether the proportions of consumers in each of the categories differ from those claimed by the consumer advocate group, we test:

\[ H_0: p_1 = .50, p_2 = .25, p_3 = .20, p_4 = .05 \]
\[ H_a: \text{At least one } p_i \text{ differs from its hypothesized value} \]

d. Some preliminary calculations are:

\[ E_1 = np_{1,0} = 106(.5) = 53 \]
\[ E_2 = np_{2,0} = 106(.25) = 26.5 \]
\[ E_3 = np_{3,0} = 106(.20) = 21.2 \]
\[ E_4 = np_{4,0} = 106(.05) = 5.3 \]

\[ \chi^2 = \sum \frac{(n_i - E_i)^2}{E_i} = \frac{[64 - 53]^2}{53} + \frac{[20 - 26.5]^2}{26.5} + \frac{[18 - 21.2]^2}{21.2} + \frac{[4 - 5.3]^2}{5.3} = 4.679 \]
Comparing Population Proportions

- The rejection region requires \( \alpha = .10 \) in the upper tail of the \( \chi^2 \) distribution with \( df = k - 1 = 4 - 1 = 3 \). From Table V, Appendix A, \( \chi^2_{.10} = 6.25139 \). The rejection region is \( \chi^2 > 6.25139 \).

- Since the observed value of the test statistic does not fall in the rejection region \( (\chi^2 = 4.679 \neq 6.25139) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate the proportions of consumers in each of the categories differ from those claimed by the consumer advocate group at \( \alpha = .10 \).

- \( \hat{p}_1 = \frac{64}{106} = .604 \)

For confidence coefficient .90, \( \alpha = .10 \) and \( \alpha / 2 = .10 / 2 = .05 \). From Table III, Appendix A, \( z_{.05} = 1.645 \). The 90% confidence interval is:

\[
\hat{p}_1 \pm z_{.05} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1}} \Rightarrow .604 \pm 1.645 \sqrt{\frac{.604(.396)}{106}} \Rightarrow .604 \pm .078 \Rightarrow (.526, .682)
\]

We are 90% confident that the true proportion of consumers who believe “Made in USA” means “100%” of labor and materials are produced in the US is between .526 and .682.

8.88 a. Let \( p_1 \) = proportion of female students who switched due to loss of interest in SME and \( p_2 \) = proportion of male students who switched due to lack of interest in SME.

Some preliminary calculations are:

\[
\hat{p}_1 = \frac{x_1}{n_1} = \frac{74}{172} = .430; \quad \hat{p}_2 = \frac{x_2}{n_2} = \frac{72}{163} = .442; \quad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{74 + 72}{172 + 163} = .436
\]

To determine if the proportion of female students who switch due to lack of interest in SME differs from the proportion of males who switch due to a lack of interest, we test:

\[ H_0: p_1 - p_2 = 0 \]
\[ H_a: p_1 - p_2 \neq 0 \]

The test statistic is

\[
z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(.430 - .442) - 0}{\sqrt{.436(.564)\left(\frac{1}{172} + \frac{1}{163}\right)}} = -0.22
\]

The rejection region requires \( \alpha / 2 = .10 / 2 = .05 \) in each tail of the \( z \) distribution. From Table III, Appendix A, \( z_{.05} = 1.645 \). The rejection region is \( z < -1.645 \) or \( z > 1.645 \).

Since the observed value of the test statistic does not fall in the rejection region \( (z = -0.22 < -1.645) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate the proportion of female students who switch due to lack of interest in SME differs from the proportion of males who switch due to a lack of interest in SME at \( \alpha = .10 \).
b. Let \( p_1 \) = proportion of female students who switched due to low grades in SME and \( p_2 \) = proportion of male students who switched due to low grades in SME.

Some preliminary calculations are:

\[
\hat{p}_1 = \frac{x_1}{n_1} = \frac{33}{172} = .192; \quad \hat{p}_2 = \frac{x_2}{n_2} = \frac{44}{163} = .270
\]

For confidence coefficient .90, \( \alpha = .10 \) and \( \alpha / 2 = .05 \). From Table III, Appendix A, \( z_{.05} = 1.645 \). The confidence interval is:

\[
(\hat{p}_1 - \hat{p}_2) \pm z_{.05} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \Rightarrow (.192 - .270) \pm 1.645 \sqrt{\frac{.192(.808)}{172} + \frac{.270(.730)}{163}}
\]

\[
\Rightarrow -.078 \pm .076 \Rightarrow (-.154, -.002)
\]

We are 90% confident that the difference between the proportions of female and male switchers who lost confidence due to low grades in SME is between \(-.154\) and \(-.002\). Since the interval does not include 0, there is evidence to indicate the proportion of female switchers due to low grades is less than the proportion of male switchers due to low grades.

8.90 Some preliminary calculations are:

\[
\hat{E}_{11} = \frac{RC_1}{n} = \frac{95(118)}{262} = 42.79 \quad \hat{E}_{12} = \frac{RC_2}{n} = \frac{95(144)}{262} = 52.21
\]

\[
\hat{E}_{21} = \frac{RC_1}{n} = \frac{69(118)}{262} = 31.08 \quad \hat{E}_{22} = \frac{RC_2}{n} = \frac{69(144)}{262} = 37.92
\]

\[
\hat{E}_{31} = \frac{RC_1}{n} = \frac{42(118)}{262} = 18.92 \quad \hat{E}_{32} = \frac{RC_2}{n} = \frac{42(144)}{262} = 23.08
\]

\[
\hat{E}_{41} = \frac{RC_1}{n} = \frac{56(118)}{262} = 25.22 \quad \hat{E}_{42} = \frac{RC_2}{n} = \frac{56(144)}{262} = 30.78
\]

To determine if a pig farmer’s education has an impact on the size of the pig farm, we test:

\[
H_0: \text{ Education and size of farm are independent} \\
H_a: \text{ Education and size of farm are dependent}
\]

The test statistic is

\[
\chi^2 = \sum \sum \frac{[n_{ij} - \hat{E}_{ij}]^2}{\hat{E}_{ij}} = \frac{(42 - 42.79)^2}{42.79} + \frac{(53 - 52.21)^2}{52.21} + \frac{(27 - 31.08)^2}{31.08} + \frac{(42 - 37.92)^2}{37.92}
\]

\[
+ \frac{(22 - 18.92)^2}{18.92} + \frac{(20 - 23.08)^2}{23.08} + \frac{(27 - 25.22)^2}{25.22} + \frac{(29 - 30.78)^2}{30.78} = 2.142
\]

Copyright © 2013 Pearson Education, Inc.
The rejection region requires \( \alpha = .05 \) in the upper tail of the \( \chi^2 \) distribution with \( df = (r - 1)(c - 1) = (4 - 1)(2 - 1) = 3 \). From Table V, Appendix A, \( \chi^2_{0.05} = 7.81473 \). The rejection region is \( \chi^2 > 7.81473 \).

Since the observed value of the test statistic does not fall in the rejection region \( (\chi^2 = 2.142 < 7.81473) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate that the a pig farmer’s education has an impact on the size of the pig farm at \( \alpha = .05 \).

To convert the frequencies to percentages, divide the numbers in each row by the row total and multiply by 100. Also, divide the column totals by the overall total and multiply by 100.

<table>
<thead>
<tr>
<th>Education Level</th>
<th>No College</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1,000</td>
<td>( \frac{42}{95} \cdot 100% = 44.2% )</td>
<td>( \frac{53}{95} \cdot 100% = 55.8% )</td>
</tr>
<tr>
<td>1,000-2,000</td>
<td>( \frac{27}{69} \cdot 100% = 39.1% )</td>
<td>( \frac{42}{69} \cdot 100% = 60.9% )</td>
</tr>
<tr>
<td>2,001-5,000</td>
<td>( \frac{22}{42} \cdot 100% = 52.4% )</td>
<td>( \frac{20}{42} \cdot 100% = 47.6% )</td>
</tr>
<tr>
<td>&gt;5,000</td>
<td>( \frac{27}{56} \cdot 100% = 48.2% )</td>
<td>( \frac{29}{56} \cdot 100% = 51.8% )</td>
</tr>
<tr>
<td>Totals</td>
<td>( \frac{118}{262} \cdot 100% = 45.0% )</td>
<td>( \frac{144}{262} \cdot 100% = 55.0% )</td>
</tr>
</tbody>
</table>

Using MINITAB, a bar chart of those who had no college is:

The graph supports the result of the test of hypothesis. Within each level of Farm Size, the percent of observations with no education are very similar.
8.92 a. The contingency table is:

<table>
<thead>
<tr>
<th>Altitude</th>
<th>Flight Response</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>&lt; 300</td>
<td>85</td>
<td>105</td>
</tr>
<tr>
<td>300-600</td>
<td>77</td>
<td>121</td>
</tr>
<tr>
<td>≥ 600</td>
<td>17</td>
<td>59</td>
</tr>
<tr>
<td>Totals</td>
<td>179</td>
<td>285</td>
</tr>
</tbody>
</table>

b. Some preliminary calculations are:

\[
\hat{E}_{11} = \frac{RC_1}{n} = \frac{190(179)}{464} = 73.297 \\
\hat{E}_{12} = \frac{RC_2}{n} = \frac{190(285)}{464} = 116.703 \\
\hat{E}_{21} = \frac{R_1C}{n} = \frac{198(179)}{464} = 76.384 \\
\hat{E}_{22} = \frac{R_1C}{n} = \frac{198(285)}{464} = 121.616 \\
\hat{E}_{31} = \frac{RC_1}{n} = \frac{76(179)}{464} = 29.319 \\
\hat{E}_{32} = \frac{RC_2}{n} = \frac{76(285)}{464} = 46.681
\]

To determine if flight response of the geese depends on the altitude of the helicopter, we test:

\[H_0: \text{Flight response and Altitude of helicopter are independent}\]
\[H_a: \text{Flight response and Altitude of helicopter are dependent}\]

The test statistic is

\[
\chi^2 = \sum \sum \frac{(n_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} = \frac{(85 - 73.297)^2}{73.297} + \frac{(105 - 116.703)^2}{116.703} + \frac{(77 - 76.384)^2}{76.384}
\]

\[
+ \frac{(121 - 121.616)^2}{121.616} + \frac{(17 - 29.319)^2}{29.319} + \frac{(59 - 46.681)^2}{46.681} = 11.477
\]

The rejection region requires \(\alpha = .01\) in the upper tail of the \(\chi^2\) distribution with df = \((r - 1)(c - 1) = (3 - 1)(2 - 1) = 2\). From Table V, Appendix A, \(\chi^2_{0.01} = 9.21034\). The rejection region is \(\chi^2 > 9.21034\).

Since the observed value of the test statistic falls in the rejection region \((\chi^2 = 11.477 > 9.21034)\), \(H_0\) is rejected. There is sufficient evidence to indicate that the flight response of the geese depends on the altitude of the helicopter at \(\alpha = .01\).
Comparing Population Proportions

The contingency table is:

<table>
<thead>
<tr>
<th>Lateral Distance</th>
<th>Flight Response</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>&lt; 1000</td>
<td>37</td>
<td>243</td>
</tr>
<tr>
<td>1000-2000</td>
<td>68</td>
<td>37</td>
</tr>
<tr>
<td>2000-3000</td>
<td>44</td>
<td>4</td>
</tr>
<tr>
<td>≥ 3000</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>Totals</td>
<td>179</td>
<td>285</td>
</tr>
</tbody>
</table>

d. Some preliminary calculations are:

\[
\hat{E}_{11} = \frac{R_1C_1}{n} = \frac{280(179)}{464} = 108.017 \\
\hat{E}_{12} = \frac{R_1C_2}{n} = \frac{280(285)}{464} = 171.983 \\
\hat{E}_{21} = \frac{R_2C_1}{n} = \frac{105(179)}{464} = 40.506 \\
\hat{E}_{22} = \frac{R_2C_2}{n} = \frac{105(285)}{464} = 64.494 \\
\hat{E}_{31} = \frac{R_3C_1}{n} = \frac{48(179)}{464} = 18.517 \\
\hat{E}_{32} = \frac{R_3C_2}{n} = \frac{48(285)}{464} = 29.483 \\
\hat{E}_{41} = \frac{R_4C_1}{n} = \frac{31(179)}{464} = 11.959 \\
\hat{E}_{42} = \frac{R_4C_2}{n} = \frac{31(285)}{464} = 19.041
\]

To determine if flight response of the geese depends on the lateral distance of the helicopter, we test:

\[H_0: \text{Flight response and Lateral distance of the helicopter are independent}\]

\[H_a: \text{Flight response and Lateral distance of the helicopter are dependent}\]

The test statistic is:

\[\chi^2 = \sum \sum \left[ \frac{n_{ij} - \hat{E}_{ij}}{\hat{E}_{ij}} \right]^2 \]

\[= \frac{(37 - 108.017)^2}{108.017} + \frac{(243 - 171.983)^2}{171.983} + \frac{(68 - 40.506)^2}{40.506} + \frac{(37 - 64.494)^2}{64.494} + \frac{(44 - 18.517)^2}{18.517} + \frac{(4 - 29.494)^2}{29.494} + \frac{(30 - 11.959)^2}{11.959} + \frac{(1 - 19.041)^2}{19.041} \]

\[= 207.814\]

The rejection region requires \(\alpha = .01\) in the upper tail of the \(\chi^2\) distribution with \(df = (r - 1)(c - 1) = (4 - 1)(2 - 1) = 3\). From Table V, Appendix B, \(\chi^2_{0.01} = 11.3449\). The rejection region is \(\chi^2 > 11.3449\).
Since the observed value of the test statistic falls in the rejection region \( (\chi^2 = 207.814 > 11.3449) \), \( H_0 \) is rejected. There is sufficient evidence to indicate that the flight response of the geese depends on the lateral distance of the helicopter at \( \alpha = .01 \).

e. Using SAS, the contingency table for altitude by response with the row percents is:

```
Table of ALTGRP by RESPONSE

<table>
<thead>
<tr>
<th>ALTGRP</th>
<th>RESPONSE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LOW</td>
<td>HIGH</td>
<td>Total</td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Row Pct</td>
<td>LOW</td>
<td>HIGH</td>
<td>Total</td>
</tr>
<tr>
<td>&lt;300</td>
<td>85</td>
<td>105</td>
<td>190</td>
</tr>
<tr>
<td></td>
<td>18.32</td>
<td>22.63</td>
<td>40.95</td>
</tr>
<tr>
<td></td>
<td>44.74</td>
<td>55.26</td>
<td></td>
</tr>
<tr>
<td>300-600</td>
<td>77</td>
<td>121</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td>16.59</td>
<td>26.08</td>
<td>42.67</td>
</tr>
<tr>
<td></td>
<td>38.89</td>
<td>61.11</td>
<td></td>
</tr>
<tr>
<td>600+</td>
<td>17</td>
<td>59</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>3.66</td>
<td>12.72</td>
<td>16.38</td>
</tr>
<tr>
<td></td>
<td>22.37</td>
<td>77.63</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>179</td>
<td>285</td>
<td>464</td>
</tr>
<tr>
<td></td>
<td>38.58</td>
<td>61.42</td>
<td>100.00</td>
</tr>
</tbody>
</table>
```

Statistics for Table of ALTGRP by RESPONSE

<table>
<thead>
<tr>
<th>Statistic</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td>2</td>
<td>11.4770</td>
<td>0.0032</td>
</tr>
<tr>
<td>Likelihood Ratio Chi-Square</td>
<td>2</td>
<td>12.1040</td>
<td>0.0024</td>
</tr>
<tr>
<td>Mantel-Haenszel Chi-Square</td>
<td>1</td>
<td>10.2104</td>
<td>0.0014</td>
</tr>
<tr>
<td>Phi Coefficient</td>
<td></td>
<td>0.1573</td>
<td></td>
</tr>
<tr>
<td>Contingency Coefficient</td>
<td></td>
<td>0.1554</td>
<td></td>
</tr>
<tr>
<td>Cramer's V</td>
<td></td>
<td>0.1573</td>
<td></td>
</tr>
</tbody>
</table>

Sample Size = 464

From the row percents, it appears that the lower the plane, the lower the response. For altitude < 300m, 55.26% of the geese had a high response. For altitude 300-600m, 61.11% of the geese had a high response. For altitude 600+m, 77.63% of the geese had a high response. Thus, instead of setting a minimum altitude for the planes, we need to set a maximum altitude. For this data, the lowest response is at an altitude of < 300 meters.
Comparing Population Proportions

Using SAS, the contingency table for lateral distance by response with the row percents is:

The FREQ Procedure
Table of LATGRP by RESPONSE

LATGRP     RESPONSE
----------+--------+--------+
         | LON    | HIGH   | Total
----------+--------+--------+
<1000    |     37 |    242 |    279
| 7.99   | 52.27  | 60.26 |
| 13.26  | 86.74  |
----------+--------+--------+
1000-2000|     68 |     37 |    105
| 14.69  | 7.99   | 22.68 |
| 64.76  | 35.24  |
----------+--------+--------+
2000-3000|     44 |      4 |     48
| 9.50   | 0.86   | 10.37 |
| 91.67  | 8.33   |
----------+--------+--------+
3000+    |     30 |      1 |     31
| 6.48   | 0.22   | 6.70  |
| 96.77  | 3.23   |
----------+--------+--------+
Total    |    179 |    284 |    463
| 38.66  | 61.34  | 100.00|

Statistics for Table of LATGRP by RESPONSE

<table>
<thead>
<tr>
<th>Statistic</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td>3</td>
<td>207.0800</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Likelihood Ratio Chi-Square</td>
<td>3</td>
<td>226.8291</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mantel-Haenszel Chi-Square</td>
<td>1</td>
<td>189.2843</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Phi Coefficient</td>
<td></td>
<td>0.6688</td>
<td></td>
</tr>
<tr>
<td>Contingency Coefficient</td>
<td></td>
<td>0.5559</td>
<td></td>
</tr>
<tr>
<td>Cramer’s V</td>
<td></td>
<td>0.6688</td>
<td></td>
</tr>
</tbody>
</table>

Effective Sample Size = 463
Frequency Missing = 1

From the row percents, it appears that the greater the lateral distance, the lower the response. For a lateral distance of 3000+m only 3.23% of the geese had a high response. Thus, the further away the plane is laterally, the lower the response. For this data, the lowest response is when the plane is further than 3000 meters.

Thus, the recommendation would be a maximum height of 300 m and a minimum lateral distance of 3000 m.

8.94 Some preliminary calculations are:

\[
\hat{E}_{11} = \frac{RC_1}{n} = 538.87 \\
\hat{E}_{12} = \frac{RC_2}{n} = 2267.13 \\
\hat{E}_{21} = \frac{RC_1}{n} = 128.67 \\
\hat{E}_{22} = \frac{RC_2}{n} = 541.33
\]
\[ \hat{E}_{31} = \frac{R_{c_1}}{n} = 151.14 \quad \hat{E}_{32} = \frac{R_{c_2}}{n} = 635.86 \]
\[ \hat{E}_{41} = \frac{R_{c_1}}{n} = 236.21 \quad \hat{E}_{42} = \frac{R_{c_2}}{n} = 993.79 \]
\[ \hat{E}_{51} = \frac{R_{c_1}}{n} = 132.12 \quad \hat{E}_{52} = \frac{R_{c_2}}{n} = 558.88 \]
\[ \hat{E}_{61} = \frac{R_{c_1}}{n} = 251.00 \quad \hat{E}_{62} = \frac{R_{c_2}}{n} = 1056.00 \]

To determine if carrying a homemade weapon in custody depends on gang classification score we test:

- **H\(_0\):** Handmade weapon and gang score are independent
- **H\(_a\):** Handmade weapon and gang score are dependent

The test statistic is

\[ \chi^2 = \sum \sum \frac{(n_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} = \frac{(255 - 538.87)^2}{538.87} + \frac{(2551 - 2267.13)^2}{2267.13} + \cdots + \frac{(831 - 1056.00)^2}{1056.00} \]
\[ = 461.636 \]

The rejection region requires \( \alpha = .01 \) in the upper tail of the \( \chi^2 \) distribution with df = \((r - 1)(c - 1) = (6 - 1)(2 - 1) = 5 \). From Table V, Appendix A, \( \chi^2_{0.01} = 15.0863 \). The rejection region is \( \chi^2 > 15.0863 \).

Since the observed value of the test statistic falls in the rejection region \(( \chi^2 = 461.363 > 15.0863 \)\), \( H_0 \) is rejected. There is sufficient evidence to indicate that carrying a homemade weapon in custody depends on gang classification score at \( \alpha = .01 \).
To convert the frequencies to percentages, divide the numbers in each row by the row total and multiply by 100. Also, divide the column totals by the overall total and multiply by 100.

<table>
<thead>
<tr>
<th>Row</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(\frac{255}{2806} \cdot 100% = 9.1%)</td>
<td>(\frac{2551}{2806} \cdot 100% = 90.9%)</td>
</tr>
<tr>
<td>1</td>
<td>(\frac{110}{670} \cdot 100% = 16.4%)</td>
<td>(\frac{560}{670} \cdot 100% = 83.6%)</td>
</tr>
<tr>
<td>2</td>
<td>(\frac{151}{787} \cdot 100% = 19.2%)</td>
<td>(\frac{636}{787} \cdot 100% = 80.8%)</td>
</tr>
<tr>
<td>3</td>
<td>(\frac{271}{1230} \cdot 100% = 22.0%)</td>
<td>(\frac{959}{1230} \cdot 100% = 88.0%)</td>
</tr>
<tr>
<td>4</td>
<td>(\frac{175}{688} \cdot 100% = 25.4%)</td>
<td>(\frac{513}{688} \cdot 100% = 74.6%)</td>
</tr>
<tr>
<td>5</td>
<td>(\frac{476}{1307} \cdot 100% = 36.4%)</td>
<td>(\frac{831}{1307} \cdot 100% = 63.6%)</td>
</tr>
<tr>
<td>Totals</td>
<td>(\frac{1438}{7488} \cdot 100% = 19.2%)</td>
<td>(\frac{6050}{7488} \cdot 100% = 80.8%)</td>
</tr>
</tbody>
</table>

Using MINITAB, a bar chart of those who had no college is:

The graph supports this conclusion. The higher the gang classification score, the more likely it is that a homemade weapon was carried.
8.96 Let \( p_1 \) = proportion of patients receiving Zyban who were not smoking one year later and \( p_2 \) = proportion of patients not receiving Zyban who were not smoking one year later.

Some preliminary calculations are:

\[
\hat{p}_1 = \frac{x_1}{n_1} = \frac{71}{309} = .230 \\
\hat{p}_2 = \frac{x_2}{n_2} = \frac{37}{306} = .121
\]

\[
\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{71 + 37}{309 + 306} = .176 \\
\hat{q} = 1 - \hat{p} = 1 - .176 = .824
\]

To determine if the antidepressant drug Zyban helped cigarette smokers kick their habit, we test:

\[ H_0: p_1 - p_2 = 0 \]
\[ H_a: p_1 - p_2 > 0 \]

The test statistic is

\[
z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(.230 - .121) - 0}{\sqrt{.176(.824)\left(\frac{1}{309} + \frac{1}{306}\right)}} = 3.55
\]

The rejection region requires \( \alpha = .05 \) in the upper tail of the \( z \) distribution. From Table III, Appendix A, \( z_{.05} = 1.645 \). The rejection region is \( z > 1.645 \).

Since the observed value of the test statistic falls in the rejection region (\( z = 3.55 > 1.645 \)), \( H_0 \) is rejected. There is sufficient evidence to indicate that the antidepressant drug Zyban helped cigarette smokers kick their habit at \( \alpha = .05 \).

8.98 Using MINITAB, the results of the analyses are:

**Tabulated statistics: Stops, Kills**

Using frequencies in Fr

<table>
<thead>
<tr>
<th>Rows: Stops</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32</td>
<td>33</td>
<td>19</td>
<td>5</td>
<td>2</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>28.31</td>
<td>34.88</td>
<td>18.71</td>
<td>6.57</td>
<td>2.53</td>
<td>91.00</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>36</td>
<td>18</td>
<td>8</td>
<td>3</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>27.69</td>
<td>34.12</td>
<td>18.29</td>
<td>6.43</td>
<td>2.47</td>
<td>89.00</td>
</tr>
<tr>
<td>All</td>
<td>56</td>
<td>69</td>
<td>37</td>
<td>13</td>
<td>5</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>56.00</td>
<td>69.00</td>
<td>37.00</td>
<td>13.00</td>
<td>5.00</td>
<td>180.00</td>
</tr>
</tbody>
</table>

Cell Contents: Count

Pearson Chi-Square = 2.171, DF = 4, P-Value = 0.704
Likelihood Ratio Chi-Square = 2.182, DF = 4, P-Value = 0.702

* NOTE * 2 cells with expected counts less than 5
First, we check to see if the assumption about the expected cells is met. From the table, there are two expected cell counts that are less than 5. Thus, the results of the test are suspect.

To determine if the number of kills is related to whether the trial was stopped or not, we test:

\[ H_0: \text{Number of kills and whether the trial was stopped or not are independent} \]
\[ H_a: \text{Number of kills and whether the trial was stopped or not are dependent} \]

The test statistic is \( \chi^2 = 2.171 \) and the \( p \)-value is \( p = .704 \) (from the printout).

Since this \( p \)-value is so large, \( H_0 \) is not rejected. There is insufficient evidence to indicate that the number of kills is related to whether the trial was stopped or not at \( \alpha \leq .10 \).

8.100 If \( x \) is normal,

\[
\begin{align*}
P(x < -2) &= .5 - .4472 = .0228 \\
P(-2 \leq x < -1) &= .4772 - .3413 = .1359 \\
P(-1 \leq x < 0) &= .3413 \\
P(0 \leq x < 1) &= .3413 \\
P(1 \leq x < 2) &= .4772 - .3413 = .1359 \\
P(x \geq 2) &= .5 - .4772 = .0228
\end{align*}
\]

(Using Table III, Appendix A)

\[
\begin{align*}
E_1 &= np_{1,0} = 200(.0228) = 4.56 = E_6 \\
E_2 &= np_{2,0} = 200(.1359) = 27.18 = E_5 \\
E_3 &= np_{3,0} = 200(.3413) = 68.26 = E_4
\end{align*}
\]

To determine if \( x \) is normally distributed, we test:

\[ H_0: p_1 = p_5 = .0228, p_2 = p_5 = .1359, p_3 = p_4 = .3413 \]
\[ H_a: \text{At least one of the proportions differs from its hypothesized value} \]

The test statistic is

\[
\chi^2 = \sum \frac{(n_i - E_i)^2}{E_i} = \frac{(7 - 4.56)^2}{4.56} + \frac{(20 - 27.18)^2}{27.18} + \frac{(61 - 68.26)^2}{68.26} + \frac{(77 - 68.26)^2}{68.26} \\
+ \frac{(26 - 27.18)^2}{27.18} + \frac{(9 - 4.56)^2}{4.56} = 9.47
\]

The rejection region requires \( \alpha = .05 \) in the upper tail of the \( \chi^2 \) distribution with df = \( k - 1 = 6 - 1 = 5 \). From Table V, Appendix A, \( \chi^2_{0.05} = 11.0705 \). The rejection region is \( \chi^2 > 11.0705 \).

Since the observed value of the test statistic does not fall in the rejection region \( (\chi^2 = 9.47 \not\geq 11.0705) \), \( H_0 \) is not rejected. There is insufficient evidence to indicate that \( x \) is not normally distributed at \( \alpha = .05 \).