AP Statistics
Packet 12/13

<u>Inference for Proportions</u>
Inference for a Population Proportion
Comparing Two Proportions
<u>Inference for Tables: Chi-Square Procedures</u>
Test for Goodness of Fit
Inference for Two-Way Tables

HW #25 6 – 8, 10, 11, 13, 21

12.6 **EQUALITY FOR WOMEN?** Have efforts to promote equality for women gone far enough in the United States? A poll on this issue by the cable network MSNBC contacted 1019 adults. A newspaper article about the poll said, "Results have a margin of sampling error of plus or minus 3 percentage points."

(a) Overall, 54% of the sample (550 of 1019 people) answered "Yes." Find a 95% confidence interval for the proportion in the adult population who would say "Yes" if asked. Is the report's claim about the margin of error roughly right? (Assume that the sample is an SRS.)

- p = true proportion of adults who would say "Yes" if asked whether efforts to promote equality for women have gone far enough in the U.S.
- SRS
  1019 < 10% of adult population
  $n\hat{p} = 550 > 10$ and $n(1-\hat{p}) = 469 > 10$
- One-proportion z interval
- $$95\%\ CI = 0.54 \pm 1.96\sqrt{\frac{(0.54)(0.46)}{1019}}$$
  $$= (0.51, 0.57)$$
- We are 95% confident that the true proportion of adults who would say "Yes" if asked whether efforts to promote equality for women have gone far enough in the U.S. is between 51 and 57%.

The margin of error is about 3% as stated.

(b) The news article said that 65% of men, but only 43% of women, think that efforts to promote equality have gone far enough. Explain why we do not have enough information to give confidence intervals for men and women separately.

We weren't given sample sizes for each gender.

(c) Would a 95% confidence interval for women alone have a margin of error less than 0.03, about equal to 0.03, or greater than 0.03? Why? You see that the news article's statement about the margin of error for poll results is a bit misleading.

The margin of error for women alone would be greater than 0.03 since the sample size is smaller.

12.7 **TEENS AND THEIR TV SETS** *The New York Times* and CBS News conducted a nationwide poll of 1048 randomly selected 13- to 17-year-olds. Of these teenagers, 692 had a television in their room and 189 named Fox as their favorite television network. We will act as if the sample were an SRS.

(a) Use your calculator to give 95% confidence intervals for the proportion of all people in this age group who have a TV in their room and the proportion who would choose Fox as their favorite network. Check that we can use our methods. *(Check assumptions.)*

Check assumptions for TV in room:	Check assumptions for prefer FOX:
SRS 1048 < 10% of all 13- to 17-year olds $n\hat{p}$ = 692 > 10, $n(1-\hat{p})$ = 356 >10	SRS 1048 < 10% of all 13- to 17-year olds $n\hat{p}$ = 189 > 10, $n(1-\hat{p})$ = 859 >10
$95\%\ CI\ \ for\ TV\ in\ room = 0.66 \pm 1.96\sqrt{\frac{(0.66)(0.34)}{1048}}$ $= (0.631, 0.689)$	$95\%\ CI\ \ for\ \ preferring\ FOX = 0.18 \pm 1.96\sqrt{\frac{(0.18)(0.82)}{1048}}$ $= (0.157, 0.203)$

(b) The news article says, "In theory, in 19 cases out of 20, the poll results will differ by no more than three percentage points in either direction from what would have been obtained by seeking out all American teenagers." Explain how your results agree with this statement.

In both cases, the margin of error for a 95% CI ("19 cases out of 20") was (no more than) 3%.

(c) Is there good evidence that more than half of all teenagers have a TV in their room?

- p = proportion of teenagers with TV's in their rooms
- $H_0:\ p = 0.5\ \ v\ \ H_A: p > 0.5$
- Assumptions checked above
- One-proportion z test.
- $z = \dfrac{0.66 - 0.5}{\sqrt{\dfrac{(0.5)(0.5)}{1048}}} = 10.36$
- p-value = P(Z >10.36) < 0.0002
- since the p-value is so small, we reject H_0, and conclude that more than half of teenagers have TVs in their rooms.

(Additionally, the interval from (a) does not include 0.50 or less)

12.8 **WE WANT TO BE RICH** In a recent year, 73% of first-year college students responding to a national survey identified "being very well-off financially" as an important personal goal. A state university finds that 132 of an SRS of 200 of its first-year students say that this goal is important.

(a) Give a 95% confidence interval for the proportion of all first-year students at the university who would identify being well-off as an important goal.

- p = true proportion of all first-year students at this university who identify being well-off as an important goal
- SRS
  200 < 10% of first-year students at this university
  $n\hat{p} = 132 > 10$ and $n(1-\hat{p}) = 68 > 10$
- One-proportion z interval
- $$95\%\ CI = 0.66 \pm 1.96\sqrt{\frac{(0.66)(0.34)}{200}}$$
  $$= (0.594, 0.726)$$
- We are 95% confident that the true proportion of all first-year students at this university who identify being well-off as an important goal is between 59.4 and 72.6%.

(b) Is there good evidence that the proportion of all first-year students at this university who think being very well-off is important differs from the national value.

Yes; the 95% confidence interval for this university contains only values that are less than 0.73, so it is likely that for this particular population, p differs from 0.73 (specifically, is less than 0.73).

12.10 **STARTING A NIGHT CLUB** A college student organization wants to start a nightclub for students under the age of 21. To assess support for this proposal, they will select an SRS of students and ask each respondent if he or she would patronize this type of establishment. They expect that about 70% of the student body would respond favorably. What sample size is required to obtain a 90% confidence interval with an approximate margin of error of 0.04. Suppose that 50% of the sample responds favorably. Calculate the margin of error of the 90% confidence interval.

Solving for *n:* $1.645\sqrt{\frac{(0.7)(0.3)}{n}} = 0.04$ we find $n \geq 356$

If p = 0.5, then the true margin of error = $1.645\sqrt{\frac{(0.5)(0.5)}{356}} = 0.0436$

12.11 **SCHOOL VOUCHERS** A national opinion poll found that 44% of all American adults agree that parents should be given vouchers good for education at any public or private school of their choice. The result was based on a small sample. How large an SRS is required to obtain a margin of error of 0.03 (that is, ±3%) in a 95% confidence interval?

(a) Answer this question using the previous poll's result as the guessed value p^*.

Solving for *n:* $1.96\sqrt{\frac{(0.44)(0.56)}{n}} = 0.03$ we find $n \geq 1052$

(b) Do the problem again using the conservative guess $p^* = 0.05$. By how much do the two sample sizes differ.

Solving for *n:* $1.96\sqrt{\frac{(0.5)(0.5)}{n}} = 0.03$ we find $n \geq 1068$

An additional 16 people are needed.

12.13 **DRUNKEN CYCLISTS** In the United States approximately 900 people die in bicycle accidents each year. One study examined the records of 1711 bicyclists aged 15 or older who were fatally injured in bicycle accidents between 1987 and 1991 and were tested for alcohol. Of these, 542 tested positive for alcohol (blood alcohol concentration of 0.01% or higher).

(a) Find a 95% confidence interval for *p*.

- *p* = true proportion of all cyclists involved in fatal accidents who had alcohol in their systems
- SRS – we do NOT know that the examined records come from as SRS, so we proceed with caution.
  1711 is probably not < 10% of all bicyclists fatally injured between 1987 and 1991, so again proceed with caution
  $n\hat{p} = 542 > 10$ and $n(1-\hat{p}) = 1169 > 10$
- One-proportion *z* interval
- $95\%\ CI = \frac{542}{1711} \pm 1.96\sqrt{\frac{\frac{542}{1711}\cdot\frac{1169}{1711}}{1711}}$
  $= (0.295, 0.339)$
- If our assumptions had been satisfied we would be 95% confident that the true proportion of all cyclists involved in fatal accidents who had alcohol in their systems is between 29.5 and 33.9%.

(b) Can you conclude from your statistical analysis of this study that alcohol causes fatal bicycle accidents? Explain.

No: We do not know; for example, what percentage of cyclists who were not involved in fatal accidents had alcohol in their systems.

12.21 **COLLEGE FOOD** Tonya, Frank, and Sarah are investigating student attitudes toward college food for an assignment in their introductory statistics class. Based on comments overheard from other students, they believe that fewer than 1 in 3 students like college food. To test this hypothesis, each selects an SRS of students who regularly eat in the cafeteria, and asks them if they like college food. Fourteen in Tonya's SRS of 50 replied, "Yes," while 98 in Frank's sample of 350, and 140 in Sarah's sample of 500 said they like college food. Use your calculator to perform a test of significance on all three results and fill in a table like this:

X	n	$\hat{p}$	z	*P*-value
14	50	0.28	-0.752	0.2261
98	350	0.28	-1.998	0.0233
140	500	0.28	-2.378	0.0088

Describe your findings in a short narrative.

Although Tonya, Frank, and Sarah all recorded the same sample proportion, $\hat{p} = 0.28$, the *p*-values were all quite different. Conclude: For a given sample proportion, the larger the sample size, the smaller the *p*-value.

HW #26   22, 23, 25 – 27, 29

12.22 **IN-LINE SKATERS** A study of injuries to in-line skaters used data from the National Electronic Injury Surveillance System, which collects data from a random sample of hospital emergency rooms. In the six-month study period, 206 people came to the sample hospitals with injuries from in-line skating. We can think of these people as an SRS of all people injured while skating. Researchers were able to interview 161 of these people. Wrist injuries (mostly fractures) were the most common.

(a) The interviews found that 53 people were wearing wrist guards and 6 of these had wrist injuries. Of the 108 who did not wear wrist guards, 45 had wrist injuries. What are the two sample proportions of wrist injuries?

$$\hat{p}_1 = \frac{6}{53} = 0.1132 \quad and \ \hat{p}_2 = \frac{45}{108} = 0.4167$$

(b) Give a 95% confidence interval for the difference between the two population proportions of wrist injuries. State carefully what populations your inference compares. We would like to draw conclusions about all in-line skaters, but we have data only for injured skaters.

- p_1 = true proportion of all people who went to the ER with wrist injuries who were wearing wrist guards, p_2 = true proportion of all people who went to the ER with wrist injuries who were NOT wearing wrist guards

- SRSs, given
  independent samples
  53 < 10% of people wearing wrist guards who came to the ER with skating injuries, 108 < 10% of people NOT wearing wrist guards who came to the ER with skating injuries
  $n_1\hat{p}_1 = 6 > 5,\ n_1(1-\hat{p}_1) = 47 > 5,\ n_2\hat{p}_2 = 45 > 5,\ n_2(1-\hat{p}_2) = 63 > 5$

- Two-proportion *z* interval

- $$95\%\,CI = (0.1132 - 0.4167) \pm 1.96\sqrt{\frac{(0.1132)(0.8868)}{53} + \frac{(0.4167)(0.5833)}{108}}$$
  $$= (-0.430, -0.177)$$

- We are 95% confident that the true proportion of skaters with wrist injuries is lower by 43 to 18% for skaters who are wearing wrist guards.

  As the problem notes, this interval should at least apply to the two groups about whom we have information: skaters (with and without wrist guards) with severe enough injuries to go to the emergency room.

(c) What was the percent of nonresponse among the original sample of 206 injured skaters? Explain why nonresponse may bias your conclusions.

45/206 = 0.218 did not respond. If those who did not respond were different in some way from those who did, then not having them represented in our sample makes our conclusions suspect.

12.23 **LYME DISEASE** Lyme disease is spread in the northeastern United States by infected ticks. The ticks were infected mainly by feeding on mice, so more mice result in more infected ticks. The mouse population in turn rises and falls with the abundance of acorns, their favored food. Experimenters studied two similar forest areas in a year when the acorn crop failed. They added hundreds of thousands of acorns to one area to imitate an abundant acorn crop, while leaving the other area untouched. The next spring, 54 of the 72 mice trapped in the first area were in breeding condition, versus 10 of the 17 mice trapped in the second area. Give a 90% confidence interval for the difference between the proportion of mice ready to breed in good acorn years and bad acorn years.

- p_1 = true proportion of mice ready to breed in a good acorn year, p_2 = true proportion of mice ready to breed in a bad acorn year

- SRSs, assumed
  independent samples
  72 < 10% of mice in good year, 17 < 10% of mice in bad years
  $n_1\hat{p}_1 = 54 > 5,\ n_1(1-\hat{p}_1) = 18 > 5,\ n_2\hat{p}_2 = 10 > 5,\ n_2(1-\hat{p}_2) = 7 > 5$

- Two-proportion *z* interval

- $$90\%\ CI = (0.75 - 0.5882) \pm 1.645\sqrt{\frac{(0.75)(0.25)}{72} + \frac{(0.5882)(0.4118)}{17}}$$
  $$= (-0.0518, 0.3753)$$

- We are 90% confident that the true proportion of mice who are ready to breed in good acorn years is between -5% and 38% greater than the proportion of mice are ready to breed in bad acorn years.

12.25 **THE GOLD COAST** A historian examining British colonial records for the Gold Coast in Africa suspects that the death rate was higher among African miners than among European miners. In the year 1936, there were 223 deaths among 33,809 African miners and 7 deaths among 1541 European miners on the Gold Coast.

Consider this year as a sample from the pre-war era in Africa. Is there good evidence that the proportion of African miners who died was higher than the proportion of European miners who died?

- p_A = true proportion of African miners who died in 1936 on the Gold Coast, p_E = true proportion of European miners who died in 1936 on the Gold Coast

- $H_0 : p_A = p_E$
  $H_A : p_A > p_E$

- SRSs, assumed
  independent samples
  33,809 < 10% of African miners, 1541 < 10% of European miners
  $n_A\hat{p}_A = 223 > 5,\ n_A(1-\hat{p}_A) = 33586 > 5,\ n_E\hat{p}_E = 7 > 5,\ n_E(1-\hat{p}_E) = 1534 > 5$

- Two-proportion *z* test

- $$\hat{p} = \frac{223+7}{33809+1541} = 0.006506$$
  $$z = \frac{0.006596 - 0.004543}{\sqrt{(0.006506)(0.993494)\left(\frac{1}{33809}+\frac{1}{1541}\right)}} = 0.9805$$

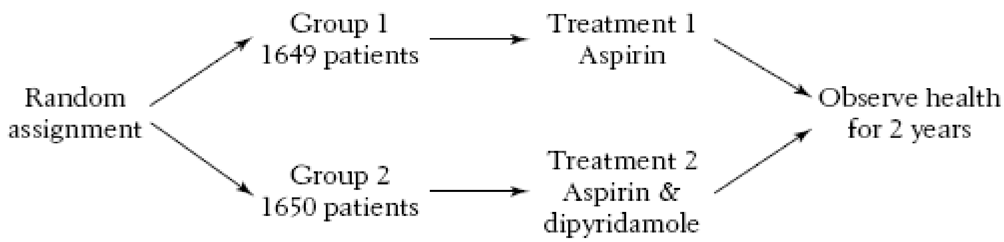
- *p*-value = P(Z > .9805) = 0.1635

- The *p*-value is large, so we fail to reject H_0, we could easily see a result like this by chance alone. We do not have evidence to conclude that the death rates for African and European miners are different.

12.26 **PREVENTING STROKES** Aspirin prevents blood from clotting and so helps prevent strokes. The Second European Stroke Prevention Study asked whether adding another anticlotting drug named diprydamole would be more effective for patients who had already had a stroke. Here are the data on strokes and deaths during the two years of the study:

	Number of patients	Number of strokes	Number of deaths
Aspirin alone	1649	206	182
Aspirin + dipyridamole	1650	157	185

(a) The study was a randomized comparative experiment. Outline the design of the study.



(b) Use your calculator to determine if there is a significant difference in the proportion of strokes in the two groups.

- p_1 = true proportion of strokes among patients who used aspirin alone,
  p_2 = true proportion of strokes among patients who used aspirin + dipyridamole
- $H_0: p_1 = p_2 \qquad H_A: p_1 \neq p_2$
- SRSs, assumed
  independent samples
  1649 < 10% of patients using aspirin alone, 1650 < 10% of patients using aspirin + dipyridamol
  pooling we have $\hat{p} = \dfrac{206+157}{1649+1650} = 0.11$
  $n_1\hat{p} = 1649(.11) = 181.4 > 5,\ n_1\hat{q} = 1649(.89) = 1467.6 > 5$
  $n_2\hat{p} = 1650(.11) = 181.5 > 5,\ n_2\hat{q} = 1650(.89) = 1467.5 > 5$
- Two-proportion *z* test
- $z = \dfrac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\hat{p}\hat{q}\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}} = \dfrac{0.125 - 0.095 - (0)}{\sqrt{(.11)(.89)\left(\dfrac{1}{1649}+\dfrac{1}{1650}\right)}} = 2.73$
- *p*-value = 2 P(z > 2.73) = 0.0064
- The *p*-value is small, so we reject H_0. We have strong evidence to conclude that the proportions for stroke rates are different.

(c) Is there a significant difference in death rates for the two groups?

- p_1 = true proportion of deaths among patients who used aspirin alone, p_2 = true proportion of deathhs among patients who used aspirin + dipyridamole
- $H_0: p_1 = p_2 \qquad H_A: p_1 \neq p_2$
- Assumptions are satisfied analogous to above
- z = -0.16
- *p*-value = 2P(z < -0.16) = 0.8728
- The *p*-value is very large, so we fail to reject H_0. We have no evidence that the proportions for death rates are different.

12.27 **ACCESS TO COMPUTERS** A sample survey by Nielsen Media Research looked at computer access and use of the Internet. Whites were significantly more likely than blacks to own a home computer, but the black-white difference in computer access at work was not significant. The study team then looked separately at the households with at least $40,000 income. The sample contained 1916 white and 131 black households in this class. Here are the sample counts for these households.

	Blacks	Whites
Own home computer	86	1173
Computer access at work	100	1132

Do higher-income blacks and whites differ significantly at the 5% level in the proportion who own home computers? Do they differ significantly in the proportion who have computer access at work?

- p_b = true proportion of blacks with $40,000 income with home computer,
  p_w = true proportion of whites with $40,000 income with home computer
- $H_0: p_b = p_w$
  $H_A: p_b \neq p_w$
- SRSs, assumed
  independent samples
  131 < 10% of blacks with $40,000 income,
  1916 < 10% of whites with $40,000 income

pooling

$$\hat{p} = \frac{86+1173}{131+1916} = 0.615$$

$n_b\hat{p} = 131(.615) = 80.5 > 5,\ n_b\hat{q} = 131(.385) = 50.5 >$

$n_w\hat{p} = 1916(.615) = 1178 > 5,\ n_w\hat{q} = 1916(.385) = 73$

- Two-proportion *z* test
- $$z = \frac{0.6565 - 0.6122}{\sqrt{(0.615)(0.385)\left(\frac{1}{131} + \frac{1}{1916}\right)}} = 1.01$$
- $p\text{-}value = 2P(z > 1.01) = 0.3124$
- The *p*-value is large, so we fail to reject H_0.
  We do not have evidence to conclude there is a difference in the proportions of higher-income blacks and whites who have home computers

- p_b = true proportion of blacks with $40,000 income with computer access at work,
  p_w = true proportion of whites with $40,000 income with computer access at work
- $H_0: p_b = p_w$
  $H_A: p_b \neq p_w$
- SRSs, assumed
  independent samples
  131 < 10% of blacks with $40,000 income,
  1916 < 10% of whites with $40,000 income

pooling

$$\hat{p} = \frac{100+1132}{131+1916} = 0.602$$

$n_b\hat{p} = 131(.602) = 79 > 5,\ n_b\hat{q} = 131(.398) = 52 > 5$

$n_w\hat{p} = 1916(.602) = 1153 > 5,\ n_w\hat{q} = 1916(.398) = 763$

- Two-proportion *z* test
- $$z = \frac{0.7634 - 0.5908}{\sqrt{(0.6019)(0.3981)\left(\frac{1}{131} + \frac{1}{1916}\right)}} = 3.90$$
- *p*-value < 0.0004
- The *p*-value is small, so we reject H_0. We hav
  strong evidence to conclude that there is a difference in the proportions of higher-income blacks and whites who have computer access a
  work. (Specifically, higher-income blacks hav
  more access at work.)

12.29 **TREATING AIDS** The drug AZT was the first drug that seemed effective in delaying the onset of AIDS. Evidence for AZT's effectiveness came from a large randomized comparative experiment. The subjects were 1300 volunteers who were infected with HIV, the virus that causes AIDS, but did not yet have AIDS. The study assigned 435 of the subjects at random to take 500 milligrams of AZT each day, and another 435 to take a placebo. (The others were assigned to a third treatment, a higher dose of AZT. We will compare only two groups.) At the end of the study, 38 of the placebo subjects and 17 of the AZT subjects had developed AIDS. We want to test the claim that taking AZT lowers the proportion of infected people who will develop AIDS in a given period of time.

(a) State hypotheses, and check that you can safely use the *z* procedure.

- p_1 = true proportion of HIV infected patients who develop AIDS while taking AZT, p_2 = true proportion of HIV infected patients who develop AIDS while taking a placebo
- $H_0: p_1 = p_2$
  $H_A: p_1 < p_2$
- SRSs, assumed
  independent samples
  435 < 10% of HIV infected patients taking AZT, 435 < 10% of of HIV infected patients taking a placebo
  pooling
  $$\hat{p} = \frac{38+17}{435+435} = 0.0632$$
  $n_1\hat{p} = 435(0.0632) = 27.5 > 5,$
  $n_1\hat{q} = 435(0.9368) = 407.5 > 5$
  same for $n_2\hat{p}$ *and* $n_2\hat{q}$

(b) How significant is the evidence that AZT is effective?

$\hat{p}$ = 0.0632,

$$z = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.087 - 0.039 - (0)}{\sqrt{(.0632)(.9368)\left(\frac{1}{435}+\frac{1}{435}\right)}} = 2.926$$

$p = P(z > 2.926) = 0.0017$

The difference is statistically significant. With such a small *p*-value we reject H_0 and conclude that the proportions are different.

(c) The experiment was double-blind. Explain what this means.

Neither the subjects nor the researchers who had contact with them knew which subjects were getting which drug—if anyone had known, they might have confounded the outcome by letting their expectations or biases affect the results.

HW #27   35, 37, 39, 41

12.35 **POLICE RADAR AND SPEEDING** Do drivers reduce excessive speed when they encounter police radar? Researchers studied the behavior of drivers on a rural interstate highway in Maryland where the speed limit was 55 miles per hour. They measured speed with an electronic device hidden in the pavement and, to eliminate large trucks, considered only vehicles less than 20 feet long. During some time periods, police radar was set up at the measurement location. Here are some of the data:

	Number of vehicles	Number over 65 mph
No radar	12,931	5,690
Radar	3,285	1,051

(a) Give a 95% confidence interval for the proportion of vehicles going faster than 65 miles per hour when no radar is present.

- p = true proportion of vehicles going faster than 65 miles per hour when no radar is present.
- SRS – assumed
  12931 < 10% of all motorists
  $n\hat{p} = 5690 > 10$ and $n(1-\hat{p}) = 7241 > 10$
- One-proportion z interval
- $95\%\, CI = 0.44 \pm 1.96\sqrt{\frac{(0.44)(0.56)}{12,931}}$
  $= (0.432, 0.449)$
- We are 95% confident that the true proportion of vehicles going faster than 65 miles per hour when no radar is present is between 43.2 and 44.9%.

(b) Give a 95% confidence interval for the effect of radar, as measured by the difference in proportions of vehicles going faster than 65 miles per hour with and without radar.

- p_1 = true proportion of vehicles going faster than 65 miles per hour when no radar is present,
  p_2 = true proportion of vehicles going faster than 65 miles per hour when radar is present.
- SRSs, assumed
  independent samples ??
  12,931 < 10% of all motorists in no radar zones
  3285 < 10% of all motorists in radar zones
  $n_1\hat{p}_1 = 5690 > 5$, $n_1(1-\hat{p}_1) = 7241 > 5$, $n_2\hat{p}_2 = 1051 > 5$, $n_2(1-\hat{p}_2) = 2234 > 5$
- Two-proportion z interval
- $95\%\, CI = (0.44 - 0.32) \pm 1.96\sqrt{\frac{(0.44)(0.56)}{12,931} + \frac{(0.32)(0.68)}{3285}}$
  $= (0.102, 0.138)$
- We are 95% confident that the 10.2 to 13.8% of drivers are more apt to speed when no radar is present, then when radar is present.

(c) The researchers chose a rural highway so that cars would be separated rather than in clusters where some cars might slow because they see other cars slowing. Explain why such clusters might make inference invalid.

In a cluster of cars, where one driver's behavior might affect the others, we do not have independence—one of the important properties of a random sample.

12.37 **SMALL-BUSINESS FAILURES, I** A study of the survival of small businesses chose an SRS from the telephone directory's Yellow Pages listings of food-and-drink businesses in 12 counties in central Indiana. For various reasons, they study got no response from 45% of the businesses chosen. Interviews were completed with 148 businesses. Three years later, 22 of these businesses had failed.

(a) Give a 95% confidence interval for the percent of all small businesses in this class that fail within three years.

- *p* = true proportion of small businesses in this class that fail within three years
- SRS – assumed
  148 < 10% of small food-and-drink businesses
  $n\hat{p} = 22 > 10$ and $n(1-\hat{p}) = 126 > 10$
- One-proportion *z* interval
- $95\%\ CI = 0.1486 \pm 1.96\sqrt{\dfrac{(0.1486)(0.8514)}{148}}$
  $= (0.0913, 0.2059)$
- We are 95% confident that the true proportion of small businesses in this class that fail within three years is between 9.1 and 20.6%.

(b) Based on the results of this study, how large a sample would you need to reduce the margin of error to 0.04?

Solving for *n:* $1.96\sqrt{\dfrac{(0.1486)(0.8514)}{n}} = 0.04$ we find $n \geq 304$

(We should not use $p^* = 0.5$ here since we have evidence that the true value of *p* is not in the range 0.3 to 0.7.)

(c) The authors hope that their findings describe the population of all small businesses. What about the study makes this unlikely? What population do you think the study findings describe?

Aside from the 45% nonresponse rate, the sample comes from a limited area in Indiana, focuses on only one kind of business, and leaves out any businesses not in the Yellow Pages (there might be a few of these; perhaps they are more likely to fail). It is more realistic to believe that this describes businesses that match the above profile; it might generalize to food-and-drink establishments elsewhere, but probably not to hardware stores and other types of business.

12.39 **SIGNIFICANT DOES NOT MEAN IMPORTANT** Never forget that even small effects can be statistically significant if the samples are large. To illustrate this fact, return to the study of 148 small businesses from the preceding exercise. Of these, 106 were headed by men and 42 were headed by women. During a three-year period, 15 of the men's businesses and 7 of the women's businesses failed.

(a) Find the proportions of failures for businesses headed by women and businesses headed by men. These sample proportions are quite close to each other. Test the hypothesis that the sample proportions of women's and men's businesses fail. (Use the two-sided alternative.) The test is very far from being significant.

$\hat{p}_m = 0.1415, \quad \hat{p}_w = 0.1667; \quad p = 0.6981$

(b) Now suppose that the same sample proportions came from a sample 30 times as large. That is, 210 out of 1260 businesses headed by women and 450 out of 3180 businesses headed by men fail. Verify that the proportions of failures are exactly the same as in (a). Repeat the *z* test for the new data, and show that it is now significant at the $\alpha = 0.05$ level.

$z = 2.12$, $p = 0.0336$

(c) It is wise to use a confidence interval to estimate the size of an effect, rather than just giving a *P*-value. Give 95% confidence intervals for the difference between the proportions of women's and men's businesses that fail for the settings of both (a) and (b). What is the effect of larger samples on the confidence interval?

From (a) (-0.1056, 0.1559)

From (b) (0.001278, 0.049036) The larger samples make the margin of error (and thus the length of the confidence interval) smaller

12.41 **MEN VERSUS WOMEN** The National Assessment of Educational Progress (NAEP) Young Adult Literacy Assessment Survey interviewed a random sample of 1917 people 21 to 25 years old. The sample contained 840 men, of whom 775 were fully employed. There were 1077 women, and 680 of them were fully employed.

(a) Use your calculator to find a 99% confidence interval to describe the difference between the proportions of young men and young women who are fully employed. Is the difference statistically significant at the 1% significance level?

99 % CI is 0.2465 to 0.3359—since 0 is not in this interval, we would reject H_0: $p_1 = p_2$ at the 1% level (in fact, P is practically 0).

(b) The mean and standard deviation of scores on the NAEP's test of quantitative skills were $\bar{x}_1 = 272.40$ *and* $s_1 = 59.2$ for the men in the sample. For the women, the results were $\bar{x}_2 = 274.73$ *and* $s_2 = 59.2$. Is the difference between the mean scores for men and women significant at the 1% level? No: $t = -0.8658$, which gives a *p*-value close to 0.4, not significant.

HW #28 1 – 4, 7, 8, 13

13.1 FINDING *P*-VALUES

(a) Find the *p*-value corresponding to $X^2 = 1.41$ for a chi-square distribution with 1 degree of freedom: (i) using Table E and (ii) with your graphing calculator.

(i) $0.20 < p < 0.25$ (ii) $p = 0.235$

(b) Find the area to the right of $X^2 = 19.62$ under the chi-square curve with 9 degrees of freedom: (i) using Table E and (ii) with your graphing calculator.

(i) $0.02 < p < 0.025$ (ii) $p = 0.0204$

(c) Find the *p*-value corresponding to $X^2 = 7.04$ for a chi-square distribution with 6 degrees of freedom: (i) using Table E and (ii) with your graphing calculator.

(i) $p > 0.25$ (ii) $p = 0.3172$

13.2 **ARE YOU MARRIED?** According to the March 2000 Current Population Survey, the marital status of the U.S. adult population is as follows:

Marital Status:	Never married	Married	Widowed	Divorced
Percent:	28.1	56.3	6.4	9.2

A random sample of 500 U.S. males, aged 25 to 29 years old, yielded the following frequency distribution:

Marital Status:	Never married	Married	Widowed	Divorced
Percent:	260	220	0	20

Perform a goodness of fit test to determine if the marital status distribution of U.S. males 25 to 29 years old differs from that of the U.S. adult population.

- H_0: The marital-status distribution of 25- to 29-year-old U.S. males is the same as that of the population as a whole.

  H_A: The marital-status distribution of 25- to 29-year-old U.S. males is different from that of the population as a whole.
- SRS, given

  Independent samples

  Expected counts are all ≥ 5 [140.5 281.5 32 46]
- Chi-square goodness of fit
- $$X^2 = \sum \frac{(\text{O}-\text{E})^2}{E} = \frac{(220-140.5)^2}{140.5} + \ldots + \frac{(20-46)^2}{46}$$

  $= 161.77, \quad with\ df = 3$
- *p*-value = $P(c^2 > 161.77) \approx 0$
- Since the *p*-value is so small, we reject H_0. We have evidence to show that the marital-status distribution of 25- to 29-year-old U.S. males is different from that of the population as a whole.

13.3 **GENETICS: CROSSING TOBACCO PLANTS** Researchers want to cross two yellow-green tobacco plants with genetic makeup (Gg). Here is a Punnet square for this genetic experiment:

	G	**g**
G	**GG**	**Gg**
g	**gG**	**gg**

This shows that the expected ratio of green (GG) to yellow-green (Gg) to albino (gg) tobacco plants is 1:2:1. When the researchers perform the experiment, the resulting offspring are 22 green, 50 yellow-green, and 12 albino seedlings. Use a chi-square goodness of fit test to assess the validity of the researchers' genetic model.

- H_0: The genetic model is valid (the different colors occur in the stated ratio of 1:2:1)
  H_A: The genetic model is not valid

- SRS, given
  Independent samples
  Expected counts are all ≥ 5 [21 42 21]

- Chi-square goodness of fit

- $X^2 = \sum \frac{(O-E)^2}{E} = \frac{(22-21)^2}{21} + \frac{(50-42)^2}{42} + \frac{(12-21)^2}{21}$
  $= 5.43, \quad \text{with } df = 2$

- *p*-value = $P(c^2 > 5.43) = 0.0662$

- Since the *p*-value is greater than 5%, we fail to reject H_0. We do not have evidence to show that the genetic model is not valid.

13.4 In recent years, a national effort has been made to enable more members of minority groups to have increased educational opportunities. You want to know if the policy of "affirmative action" and similar initiatives have had any effect in this regard. You obtain information on the ethnicity distribution of holders of the highest academic degree, the doctor of philosophy degree, for the year 1981:

Race/ethnicity	Percent
White, non-Hispanic	78.9
Black, non-Hispanic	3.9
Hispanic	1.4
Asian of Pacific Islander	2.7
American Indian/Alaskan Native	0.4
Nonresident alien	12.8

A random sample of 300 doctoral degree recipients in 1994 showed the following frequency distribution:

Race/ethnicity	Percent
White, non-Hispanic	189
Black, non-Hispanic	10
Hispanic	6
Asian of Pacific Islander	14
American Indian/Alaskan Native	1
Nonresident alien	80

(a) Perform a goodness of fit test to determine if the distribution of doctoral degrees in 1994 is significantly different from the distribution in 1981. Note: Although two expected counts are less than 5, the chi-square test still gives adequate results in this setting.

- H_0: The ethnicity distribution of the PhD degree in 1994 is the same as it was in 1981.
  H_A: The ethnicity distribution of the PhD degree in 1994 is not the same as it was in 1981.

- SRS, given
  Independent samples
  Expected counts: [236.7 11.7 4.2 8.1 1.2 38.4]

- Chi-square goodness of fit

- $$X^2 = \sum \frac{(O-E)^2}{E} = \frac{(189-236.7)^2}{236.7} + \ldots + \frac{(80-38.4)^2}{38.4}$$
  $$= 61.98, \quad with\ df = 5$$

- p-value = $P(c^2 > 61.98) \approx 0$

- Since the *p*-value is so small, we reject H_0. We have evidence to show that the ethnicity distribution of the PhD degree has changed from 1981 to 1994.

(b) In what categories have the greatest changes occurred and in what direction?

The greatest change is that many more nonresident aliens than expected received the Ph.D. degree in 1994 over the 1981 figures. To a lesser extent, a smaller proportion of white, non-Hispanics received the Ph.D. degree in 1994.

13.7 **IS YOUR RANDOM NUMBER GENERATOR WORKING?** In this exercise you will use your calculator to simulate sampling from the following uniform distribution:

X:	0	1	2	3	4	5	6	7	8	9
P(X):	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

You will then perform a goodness of fit test to see if a randomly generated sample distribution comes form a population that is different from this distribution.

(a) State the null and alternative hypotheses for this test.

$H_0: p_0 = p_1 = p_2 = \ldots = p_9 = 0.01$

H_A: At least one of the p's is not equal to 0.1.

(b) Use the randint function to randomly generate 200 digits from 0 to 9, and store these values in L4.

Using randInt (0, 9, 200) → L4, we obtained these counts for digits 0 to 9:

19, 17, 23, 22, 19, 20, 25, 12, 27, 16.

(c) Plot the data as a histogram with Window dimensions sets as follows: X[-0.5,9.5]$_1$ and Y[-5, 30]$_5$. (You may have to increase the vertical scale.) Then TRACE to see the frequencies of each digit. Record these frequencies (observed values) in L1.

(d) Determine the expected counts for a sample size of 200, and store them in L2.

They should all = 20

(e) Complete a goodness of fit test. Report your chi-square statistic, the *P*-value, and your conclusion with regard to the null and alternative hypotheses.

- SRS
  independent sample
  All expected counts ≥ 5
- $X^2 = \sum \frac{(O-E)^2}{E} = \frac{(19-20)^2}{20} + \ldots + \frac{(16-20)^2}{20}$
  $= 8.9, \quad with\ df = 9$
- *p*-value = 0.447
- *p*-value is large, so we fail to reject H_0. There is no evidence that the sample data were generated from a distribution that is different from the uniform distribution.

13.8 **ROLL THE DICE** Simulate rolling a fair, six-sided die 300 times on your calculator. Plot a histogram of the results, and then perform a goodness of fit test of the hypothesis that the die is fair.

Use randint(1, 6, 300) → L1
To simulate rolling a fair die 300 times.
In our simulation, we obtained the following frequency distribution:

side	1	2	3	4	5	6
freq	57	46	55	54	45	43

- H_0 : The die is fair, all the *p*'s = 1/6.
  H_A : The die is not fair
- SRS
  Independent samples
  All expected counts = 50 ≥ 5
- Chi-square goodness of fit
- $$X^2 = \sum \frac{(\text{O} - \text{E})^2}{E} = \frac{(57-50)^2}{50} + \ldots + \frac{(43-40)^2}{50}$$
  $= 3.6, \quad with\ df = 5$
- *p*-value = $P(c^2 > 3.6) = 0.608$
- Since the *p*-value is so large, we fail to reject H_0. We have no evidence to show that the die is not fair.

13.13 **CARNIVAL GAMES** A "wheel of fortune" at a carnival is divided into four equal parts:

Part I: Win a doll
Part II: Win a candy bar
Part III: Win a free ride
Part IV: Win nothing

You suspect that the wheel is unbalanced (i.e., not all parts of the wheel are equally likely to be landed upon when the wheel is spun.) The results of 500 spins of the wheel are as follows:

Part:	I	II	III	IV
Frequency:	95	105	135	165

Perform a goodness of fit test. Is there evidence that the wheel is not in balance?

- H_0 : The wheel is balanced. (All four outcomes are uniformly distributed)
  H_A : The wheel is not balanced.
- SRS
  Independent samples
  All expected counts = 125 ≥ 5
- Chi-square goodness of fit
- $$X^2 = \sum \frac{(\text{O} - \text{E})^2}{E} = \frac{(95-125)^2}{125} + \ldots + \frac{(165-125)^2}{125}$$
  $= 24, \quad with\ df = 3$
- *p*-value = $2.5 \times 10^{-5} \approx 0$
- Since the *p*-value is so small, we reject H_0. We have strong evidence to show that the wheel is not balanced. Since "Part IV: Win nothing" shows the greatest deviation from the expected result, there may be reason to suspect that the carnival game operator may have tampered with the wheel to make it harder to win.

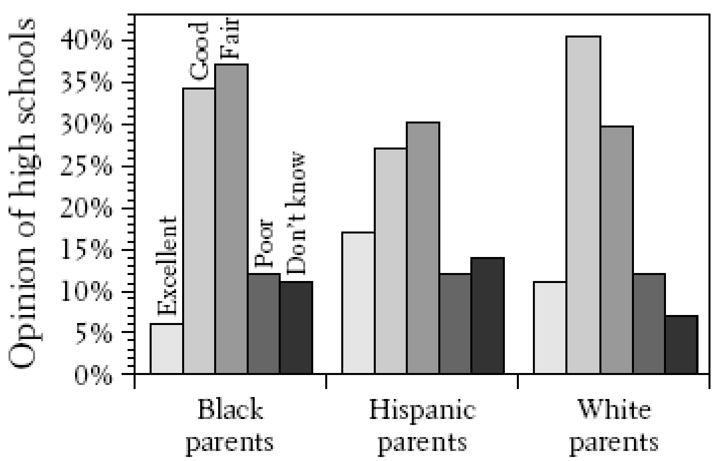
HW #29 18, 19, 20

13.18 **HOW ARE SCHOOLS DOING?** The nonprofit group Public Agenda conducted telephone interviews with 3 randomly selected groups of parents of high school children. There were 202 black parents, 202 Hispanic parents, and 201 white parents. One question asked was, "Are the high schools in your state doing as excellent, good, fair, or poor job, or don't you know enough to say?" Here are the survey results:

	Black parents	Hispanic parents	White parents
Excellent	12	34	22
Good	69	55	81
Fair	75	61	60
Poor	24	24	24
Don't know	22	28	14
Total	202	202	201

Write a brief analysis of these results. Include a graph or graphs, a test of significance, and your own discussion of the most important findings.

Various graphs can be made; one possibility is shown below.
For the null hypothesis "There is no relationship between race and opinions about schools," we find $X^2 = 22.426$ (df = 8) and $p = 0.004$ (Minitab output on the next page).
With such a small *p*-value, we reject the null hypothesis. We have evidence that there is a relationship between race and opinions about schools; specifically, blacks are less likely, and Hispanics more likely, to consider schools "excellent," while Hispanics and whites differ in percentage considering schools "good" (whites are higher) and percentage who "don't know" (Hispanics are higher). Also, a higher percentage of blacks rated schools as "fair."



Minitab output

```
         Black   Hispanic   White   Total
    1       12         34      22      68
         22.70      22.70   22.59

    2       69         55      81     205
         68.45      68.45   68.11

    3       75         61      60     196
         65.44      65.44   65.12

    4       24         24      24      72
         24.04      24.04   23.92

    5       22         28      14      64
         21.37      21.37   21.26

Total      202        202     201     605

ChiSq =  5.047 + 5.620 + 0.015 +
         0.004 + 2.642 + 2.441 +
         1.396 + 0.301 + 0.402 +
         0.000 + 0.000 + 0.000 +
         0.019 + 2.058 + 2.481 = 22.426

df = 8, p = 0.004
```

13.19 **EXTRACURRICULAR ACTIVITIES AND GRADES** North Carolina State University studied student performance in a course required by its chemical engineering major. One question of interest is the relationship between time spent in extracurricular activities and whether a student earned a C or better in the course. Here are the data for the 119 students who answered a question about extracurricular activities:

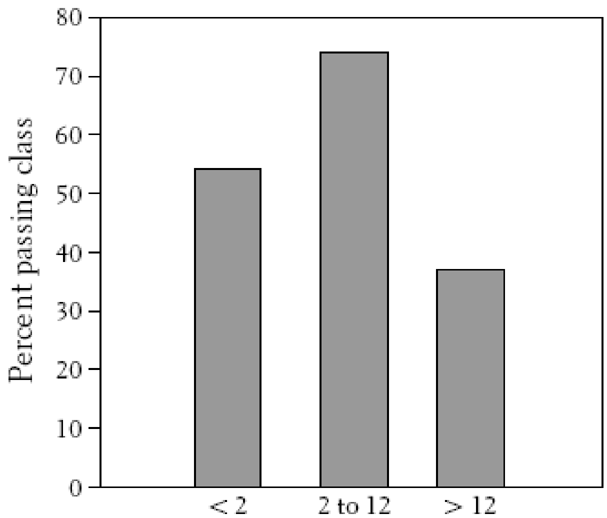
	Extracurricular activities (hours per week)		
	<2	2 to 12	>12
C or better	11	68	3
D or F	9	23	5

(a) This is an $r \times c$ table. What are the numbers r and c? $r = 2, c = 3$

(b) Find the proportion of successful students (C or better) in each of the three extracurricular activity groups. What kind of relationship between extracurricular activities and succeeding in the course do these proportions seem to show?

55.0%, 74.7%, and 37.5%. Some (but not too much) time spent in extracurricular activities seems to be beneficial.

(c) Make a bar graph to compare the three proportions of successes.



(d) What null hypothesis will a chi-square procedure test in this setting?

H_0: There is no association between amount of time spent on extracurricular activities and grades earned in the course.

Ha: There is an association

(e) Find the expected counts if this hypothesis is true, and display them in a two-way table.

	< 2	2 - 12	> 12
C or better	13.78	62.71	5.51
D or F	6.22	28.29	2.49

(f) Compare the observed counts with the expected counts. Are there large deviations between them? These deviations are another way of describing the relationship you described in (b).

The first and last columns have lower numbers than we expect in the "passing" row (and higher numbers in the "failing" row), while the middle column has this reversed—more passed than we would have expected if the proportions were all equal.

13.20 **SMOKING BY STUDENTS AND THEIR PARENTS** How are the smoking habits of students related to their parents' smoking? Here are data from a survey of students in eight Arizona high schools:

	Student smokes	Student does not smoke
Both parents smoke	400	1380
One parent smokes	416	1823
Neither parent smokes	188	1168

(a) This is an $r \times c$ table. What are the numbers r and c? $r = 3$, $c = 2$

(b) Calculate the proportion of students who smoke in each of the three parent groups. Then describe in words the association between parent smoking and student smoking.

22.5%, 18.6%, and 13.9%. A student's likelihood of smoking increases when one parent smokes, and increases even more when both smoke.

(c) Make a graph to display the association

Omitted

(d) Explain in words what the null hypothesis H_0 says about smoking.

The null hypothesis says that parents' smoking habits have no effect on their children.

(e) Find the expected counts in H_0 is true, and display them in a two-way table similar to the table of observed counts.

	Student smokes	Student does not smoke
Both parents smoke	332.49	1447.51
One parent smokes	418.22	1820.78
Neither parent smokes	253.29	1102.71

(f) Compare the tables of observed and expected counts. Explain how the comparison expresses the same association you see in (b) and (c).

In column 1, row 1, the expected count is much smaller than the actual count; meanwhile, the actual count is lower than expected in the lower left. This agrees with what we observed before: Children of non-smokers are less likely to smoke.

HW #30   24, 25, 27, 29, 30

13.24 **TREATING ULCERS** Gastric freezing was once a recommended treatment for ulcers in the upper intestine. Use of gastric freezing stopped after experiments showed it had no effect. One randomized comparative experiment found that 28 of the 82 gastric-freezing patients improved, while 30 of the 78 patients in the placebo group improved. We can test the hypothesis of "no difference" between the two groups in two ways: Using the two-sample *z* statistic or using the chi-square statistic.

(a) State the null hypothesis with a two-sided alternative and carry out the *z* test. What is the *P*-value from Table A?

H_0: $p_1 = p_2$ vs. Ha: $p_1 \neq p_2$. $z = -0.5675$ and $p = 0.5704$.

(b) Present the data in a 2×2 table. Use the chi-square statistic to test the hypothesis from (a). Verify that the X^2 statistic is the square of the *z* statistic. Use your calculator to verify that the chi-square *P*-value agrees with the *z* result.

$$X^2 = \sum \frac{(\mathrm{O}-\mathrm{E})^2}{E} = \frac{(28-29.73)^2}{29.73} + \ldots + \frac{(48-49.72)^2}{49.72}$$

$$= 0.323, \ \textit{which equals } z^2, \quad \textit{with df} = 1$$

p-value = 0.5704. (same as in (a))

With such a large *p*-value, we fail to reject H_0. There is no evidence that gastric freezing is beneficial.

	Improved	Did not improve
Gastric freezing	28 29.73	54 52.28
Placebo	30 28.27	48 49.72

(c) What do you conclude about the effectiveness of gastric freezing as a treatment for ulcers?

Gastric freezing is not significantly more (or less) effective than a placebo treatment.

13.25 **STRESS AND HEART ATTACKS** You read a newspaper article that describes a study of whether stress management can help reduce heart attacks. The 107 subjects all had reduced blood flow to the heart and so were at risk of a heart attack. They were assigned at random to three groups. The article goes on to say:

> One group took a four-month stress management program, another underwent a four-month exercise program, and the third received usual heart care from their personal physicians.
> In the next three years, only three of the 33 people in the stress management groups suffered "cardiac events," defined as a fatal or non-fatal attacks or a surgical procedure such as a bypass or angioplasty. In the same period, seven of the 34 people in the exercise groups and 12 out of the 40 patients in usual care suffered such events.

(a) Use the information in the news article to make a two-way table that describes the study results.

	Cardiac event?		
Group	**Yes**	**No**	**TOTAL**
Stress management	3	30	**33**
Exercise	7	27	**34**
Usual care	12	28	**40**
TOTAL	**22**	**85**	**107**

(b) What are the success rates of the three treatments in avoiding cardiac events.

Success rates (% of "No"s): 90.91%, 79.41%, 70%.

(c) Find the expected cell counts under the null hypothesis that there is no difference among the treatments. Verify that the expected counts meet our guideline for use of the chi-square test.

Group	**Expected Yes**	**Expected No**
Stress management	6.785	26.215
Exercise	6.991	27.009
Usual care	8.224	31.776

SRS, independent samples. All expected counts ≥ 5.

(d) Is there a significant difference among the success rates for the three treatments? Give appropriate statistical evidence to support your answer.

$X^2 = 4.84$ (df = 2), p-value = 0.0889.

Though the success rate for the stress management group is slightly higher than for the other two groups, there does not appear to be a significant difference among the success rates.

13.27 **DO YOU USE COCAINE?** Sample surveys on sensitive issues can give different results depending on how the question is asked. A University of Wisconsin study divided 2400 respondents into 3 groups at random. All were asked if they had ever used cocaine. One group of 800 was interviewed by phone; 21% said they had used cocaine. Another 800 people were asked the question in a one-on-one personal interview; 25% said "Yes." The remaining 800 were allowed to make anonymous written response; 28% said "Yes." Are there statistically significant differences among these proportions? Give appropriate statistical evidence to support your conclusion.

- H_0 : all proportions are equal

  H_A : some proportions are different

- SRS
  Independent samples
  All expected counts ≥ 5

	Observed			Expected	
	Yes	No		Yes	No
Phone	168	632		197.3	602.7
One-on-one	200	600		197.3	602.7
Anonymous	224	576		197.3	602.7

- c^2 - test
- $$X^2 = \sum \frac{(\text{O} - \text{E})^2}{E} = \frac{(168 - 197.3)^2}{197.3} + \ldots + \frac{(576 - 602.7)^2}{602.7}$$
  $= 10.62, \quad with\ df = 2$
- p-value $= 0.0049$
- With such a small p-value, we reject H_0. There is evidence that the contact method makes a difference in response.

13.29 **SECONDHAND STORES, I** Shopping at secondhand stores is becoming more popular and has even attracted the attention of business schools. A study of customers' attitudes toward secondhand stores interviewed samples of shoppers at two secondhand stores of the same chain in two cities. The breakdown of respondents by sex is as follows:

	City I	City II
Men	38	68
Women	203	150
Total	241	218

Is there a significant difference between the proportions of women customers in the two cities?

(a) State the null hypothesis, find the sample proportions of women in both cities. Assume all assumptions are met and use your calculator to do a two-sided *z* test, and find the *p*-value.

$H_0 : p_1 = p_2$, *where* p_1 *and* p_2 are the proportions of women customers in each city.

$\hat{p}_1 = 0.8423, \quad \hat{p}_2 = 0.6881$

$z = 3.9159$

p-value = 0.00009

With such a small *p*-value, we would reject the null hypothesis and conclude the proportions of women customers in each city are not the same.

(b) Use your calc to find the chi-square statistic X^2 and show that it is the square of the *z* statistic. Show that the *p*-value from this calculation agrees with your result from (a).

$X^2 = 15.334$, which equals z^2 from part (a); Again the *p*-value = 0.00009

(c) Use your calc to give a 95% confidence interval for the difference between the proportions of women customers in the two cities. We can see an advantage here to using the *z* methods. 95% CI = (0.0774, 0.2311)

13.30 **SECONDHAND STORES, II** The study of shoppers in secondhand stores cited in the previous exercise also compared the income distributions of shoppers in the two stores. Here is a two-way table of counts:

Income	City I	City II
Under $10,000	70	62
$10,000 to $19,999	52	63
$20,000 to $24,999	69	50
$25,000 to $34,999	22	19
$35,000 or more	28	24

A statistical calculator gives the chi-square statistic for this table as $X^2 = 3.955$. Is there good evidence that customers at the two stores have different income distributions?

With 4 degrees of freedom; $p = 0.4121$. There is not enough evidence to reject H_0 at any reasonable level of significance; the difference in the two income distributions is not statistically significant.