12

Inference for Proportions

12.1 (a) Population: the 175 residents of Tonya's dorm; p is the proportion who like the food. (b) $\hat{p} = 0.28$.

12.2 (a) The population is the 2400 students at Glen's college, and p is the proportion who believe tuition is too high. (b) $\hat{p} = 0.76$.

12.3 (a) The population is the 15,000 alumni, and p is the proportion who support the president's decision. (b) $\hat{p} = 0.38$.

12.4 (a) No—the population is not large enough relative to the sample. (b) Yes—we have an SRS, the population is 48 times as large as the sample, and the success count (38) and failure count (12) are both greater than 10. (c) No—there were only 5 or 6 "successes" in the sample.

12.5 (a) No—np_0 and $n(1 - p_0)$ are less than 10 (they both equal 5). (b) No—the expected number of failures is less than 10 ($n(1 - p_0) = 2$). (c) Yes—we have an SRS, the population is more than 10 times as large as the sample, and $np_0 = n(1 - p_0) = 10$.

12.6 (a) $\text{SE}_{\hat{p}} = \sqrt{(0.54)(0.46)/1019} \doteq 0.01561$, so the 95% confidence interval is $0.54 \pm (1.96)(0.01561) = 0.51$ to 0.57. The margin of error is about 3%, as stated.

(b) We weren't given sample sizes for each gender. (However, students who know enough algebra can get a good estimate of those numbers by solving the system $x + y = 1019$ and $0.65x + 0.43y = 550$: approximately 508 men and 511 women.)

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

12.7 (a) The methods can be used here, since we assume we have an SRS from a large population, and all relevant counts are more than 10. For TVs in rooms: $\hat{p}_1 \doteq 0.66$ and $\text{SE}_{\hat{p}} = \sqrt{(0.66)(0.34)/1048} \doteq 0.01463$, so the 95% confidence interval is $0.66 \pm (1.96)(0.01463) \doteq 0.631$ to 0.689. For preferring Fox: $\hat{p}_2 \doteq 0.18$ and $\text{SE}_{\hat{p}} = \sqrt{(0.18)(0.82)/1048} \doteq 0.01187$, so the 95% confidence interval is $0.18 \pm (1.96)(0.01187) \doteq 0.157$ to 0.203.

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

(c) We test $H_0: p = 0.5$ versus $H_a: p > 0.5$. The test statistic is $z = (0.66-0.50)/\sqrt{\frac{(0.5)(0.5)}{1048}} \doteq 10.36$, which gives very strong evidence against H_0 ($P < 0.0002$); we conclude that more than half of teenagers have TVs in their rooms. (Additionally, the interval from (a) does not include 0.50 or less.) With the TI-83, $z = 10.379$ and $P = 1.577 \times 10^{-25}$.

12.8 (a) $\hat{p} = .66$, and since $n\hat{p} = 132$ and $n(1 - \hat{p}) = 68$ are both greater than 10, the confidence interval based on z can be used. The 95% confidence interval for p is $.66 \pm (1.96)\sqrt{((.66)(.34)/200)} = .66 \pm 0.06565$, or 0.59435 to 0.72565.

(b) Yes; the 95% confidence interval 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).

12.9 (a) $\hat{p} = \frac{15}{84} \doteq 0.1786$, and $SE_{\hat{p}} = \sqrt{\hat{p}(1 - \hat{p})/84} \doteq 0.0418$. (b) Checking conditions, $n\hat{p} = 15$ and $n(1 - \hat{p}) = 69$ are both at least 10. Provided that there are at least $(84)(p) = 840$ applicants in the population of interest, we are safe constructing the confidence inerval. $\hat{p} \pm 1.645\ SE_{\hat{p}} =$ 0.1098 to 0.2473.

12.10 $n = (\frac{1.645}{0.04})^2(0.7)(0.3) \doteq 355.2$—use $n = 356$. With $\hat{p} = 0.5$, $SE_{\hat{p}} \doteq 0.0265$, so the true margin of error is $(1.645)(0.0265) = 0.0436$.

12.11 (a) 1051.7—round up to 1052. (b) 1067.1—round up to 1068; 16 additional people.

12.12 450.2—round up to 451.

12.13 (a) We do not know that the examined records came from an SRS, so we must be cautious in drawing emphatic conclusions. Both $n\hat{p}$, $n(1 - \hat{p})$ are at least 10. $\hat{p} = \frac{542}{1711} \doteq 0.3168$; $SE_{\hat{p}} = \sqrt{\hat{p}(1 - \hat{p})/1711} \doteq 0.01125$; the interval is $\hat{p} \pm 1.960\ SE_{\hat{p}} = 0.2947$ to 0.3388. (b) No: We do not know; for example, what percentage of cyclists who were *not* involved in fatal accidents had alcohol in their systems.

12.14 $\hat{p} = 0.05227$, $z = -3.337$, and $P < 0.0005$—very strong evidence against H_0, and in favor of H_a: $p < \frac{1}{10}$.

12.15 (a) Checking conditions: we are treating the Gallup sample as an SRS; the population of adults is much larger than 10(1785); $n\hat{p}$ and $n(1 - \hat{p})$ are both at least 10. The interval is 39.0% to 45.0%. (b) Since 50% falls above the 99% confidence interval, this is strong evidence against H_0: $p = 0.5$ in favor of H_a: $p < 0.5$. (In fact, $z = -6.75$ and P is tiny.) (c) 16589.4—round up to 16590. The use of $p^* = 0.5$ is reasonable because our confidence interval shows that the actual p is in the range 0.3 to 0.7.

12.16 (a) The distribution is approximately normal with mean $\mu = p = 0.14$ and standard deviation $\sigma = \sqrt{(0.14)(0.86)/9224} \doteq 0.003613$.

(b) For testing H_0: $p = 0.14$ versus H_a: $p > 0.14$, we have $\hat{p} \doteq 0.27$, and the test statistic is $z = (0.27 - 0.14)/0.003613 \doteq 36$. We have very strong (in fact, overwhelming) evidence that Harleys are more likely to be stolen.

12.17 (a) H_0: $p = 0.5$ vs. H_a: $p > 0.5$, $z = 1.697$, $P = 0.0448$—reject H_0 at the 5% level. (b) 0.5071 to 0.7329. (c) The coffee should be presented in random order—some should get the instant coffee first, and others the fresh-brewed first.

12.18 (a) $n = (\frac{2.576}{0.015})^2(0.2)(0.8) \doteq 4718.8$—use $n = 4719$. (b) $2.576\sqrt{\frac{(0.1)(0.9)}{4719}} \doteq 0.01125$.

12.19 (a) Let p = Shaq's free-throw percentage during the season following his off-season training. We wish to test H_0: $p = .533$ vs. H_a: $p > .533$. $np_0 = (39)(.533) = 27.087$ and $n(1 - p_0) = (39)(.467) = 18.213$ are both greater than 10, so the one-sample z test may be used. The test statistic is $z = (.667 - .533)/\sqrt{((.533)(.467)/39)} = 1.673$, and the P-value $= 0.04715$. There is some evidence that Shaq has, in fact, improved his free-throw percentage.

(b) Type I error: concluding that Shaq has improved his free-throwing when in fact he has not. Type II error: concluding that Shaq has not improved his free-throwing when in fact he has.

(c) We seek the power of the test when $p = 0.6$. With $\alpha = 0.05$, we reject H_0 in favor of H_a when $z > 1.645$, that is, when $\hat{p} > 0.533 + (1.645)(\sqrt{((.533)(.467)/39)} = 0.6644$. Then, $P(\text{reject } H_0 \text{ when } p = 0.6) = P(\hat{p} > 0.6644 \text{ when } p = 0.6) = P(Z > 0.821) = 0.2058$.

(d) $P(\text{Type I error}) = \alpha = 0.05$; $P(\text{Type II error when } p = 0.6) = 1 - 0.2058 = 0.7942$.

12.20 This exercise is a follow-up to Activity 12.

12.21

X	n	$\hat{p}$	z	P-value
14	50	.28	−.752	.2261
98	350	.28	−1.998	.0233
140	500	.28	−2.378	.0088

Although Tonya, Frank, and Sarah all recorded the same sample proportion, $\hat{p} = .28$, the P-values were all quite different. Conclude: For a given sample proportion, the larger the sample size, the smaller the P-value.

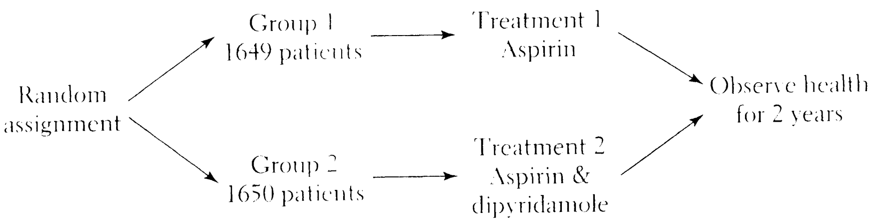
12.22 (a) $\hat{p}_1 = \frac{6}{53} \doteq 0.1132$ and $\hat{p}_2 = \frac{45}{108} \doteq 0.4167$. (b) $\mathrm{SE} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{53} + \frac{\hat{p}_2(1-\hat{p}_2)}{108}} \doteq 0.06438$, so the 95% confidence interval is $(0.1132 - 0.4167) \pm (1.96)(0.06438) = -0.4297$ to -0.1773. 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) $45/206 \doteq 21.8\%$ did not respond. If those who did not respond were different from those who did, then not having them represented in our sample makes our conclusions suspect.

12.23 $\hat{p}_1 = \frac{54}{72} = 0.75$ and $\hat{p}_2 = \frac{10}{17} \doteq 0.5882$. Then $\mathrm{SE} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{72} + \frac{\hat{p}_2(1-\hat{p}_2)}{17}} \doteq 0.1298$, so the 90% confidence interval is $(0.75 - 0.5882) \pm (1.645)(0.1298) = -0.0518$ to 0.3753. These methods can be used because the populations of mice are certainly more than 10 times as large as the samples, and the counts of successes and failures are more than 5 in both samples.

12.24 The population-to-sample ratios are large enough, and all relevant counts are at least 10. $\hat{p}_1 = 0.3895$, $\hat{p}_2 = 0.3261$, and the interval is -0.0208 to 0.14764.

12.25 To test H_0: $p_1 = p_2$ versus H_a: $p_1 > p_2$, we find $\hat{p}_1 = \frac{223}{33809} \doteq 0.006596$, $\hat{p}_2 = \frac{7}{1541} \doteq 0.004543$ and $\hat{p} = \frac{223+7}{33809+1541} \doteq 0.006506$. Then $\mathrm{SE} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{33809} + \frac{1}{1541}\right)} \doteq 0.002094$, so $z = (\hat{p}_1 - \hat{p}_2)/\mathrm{SE} \doteq 0.98$. This gives $P = 0.1635$, so we don't have enough evidence to conclude that the death rates are different.

12.26 (a) Diagram below. (b) To test H_0: $p_1 = p_2$ versus H_a: $p_1 \neq p_2$, we find $\hat{p}_1 = \frac{206}{1649} \doteq 0.1249$, $\hat{p}_2 = \frac{157}{1650} \doteq 0.09515$, and $\hat{p} = \frac{206+157}{1649+1650} \doteq 0.11003$. Then $\mathrm{SE} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{1649} + \frac{1}{1650}\right)} \doteq 0.010897$, so $z = (\hat{p}_1 - \hat{p}_2)/\mathrm{SE} \doteq 2.73$. This gives $P = 0.0064$, which is very strong evidence that the proportions are different. (c) With the same hypotheses (for the proportions of deaths), this time we find $\hat{p}_1 = \frac{182}{1649} \doteq 0.1104$, $\hat{p}_2 = \frac{185}{1650} \doteq 0.1121$, and $\hat{p} = \frac{182+185}{1649+1650} \doteq 0.11125$. Then $\mathrm{SE} \doteq 0.010949$, so $z \doteq -0.16$; this gives $P = 0.8728$—virtually no evidence of a difference.



12.27 For home computer access, we test H_0: $p_1 = p_2$ versus H_a: $p_1 \neq p_2$ and find $\hat{p}_1 = \frac{86}{131} \doteq 0.6565$, $\hat{p}_2 = \frac{1173}{1916} \doteq 0.6122$, and $\hat{p} = \frac{86+1173}{131+1916} \doteq 0.6150$. Then $\mathrm{SE} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{131} + \frac{1}{1916}\right)} \doteq 0.04394$, so $z = (\hat{p}_1 - \hat{p}_2)/\mathrm{SE} \doteq 1.01$. This gives $P = 0.3124$—little evidence that the proportions are different. With the same hypotheses for the proportions with PC access at work, we find

$\hat{p}_1 = \frac{100}{131} \doteq 0.7634$, $\hat{p}_2 = \frac{1132}{1916} \doteq 0.5908$, and $\hat{p} = \frac{100+1132}{131+1916} \doteq 0.6019$. Then SE $\doteq 0.044207$, so $z \doteq 3.90$. This gives $P < 0.0004$, so we have very strong evidence of a difference (specifically, that higher-income blacks have more access at work).

12.28 (a) For eventual contact, we test H_0: $p_1 = p_2$ versus H_a: $p_1 < p_2$ and find $\hat{p}_1 = \frac{58}{100} = 0.58$, $\hat{p}_2 = \frac{200}{291} \doteq 0.6873$, and $\hat{p} = \frac{58+200}{100+291} \doteq 0.6598$. Then SE $= \sqrt{\hat{p}(1-\hat{p})(\frac{1}{100} + \frac{1}{291})} \doteq 0.05492$, so $z = (\hat{p}_1 - \hat{p}_2)/\text{SE} \doteq -1.95$. This gives $P = 0.0254$—fairly strong evidence that the proportions and different.

(b) With the same hypotheses for the proportions completing the survey, we find $\hat{p}_1 = \frac{33}{100} = 0.33$, $\hat{p}_2 = \frac{134}{291} \doteq 0.4605$, and $\hat{p} = \frac{33+134}{100+291} \doteq 0.4271$. Then SE $\doteq 0.05734$, so $z \doteq -2.28$. This gives $P = 0.0113$, so we have strong evidence of a difference.

(c) For survey completion, the standard error for a confidence interval is SE $= \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{100} + \frac{\hat{p}_2(1-\hat{p}_2)}{291}} \doteq 0.05536$, so, for example, a 95% confidence interval for the difference $p_1 - p_2$ is $(0.33 - 0.4605) \pm (1.96)(0.05536) = -0.2390$ to -0.0220. Although the difference for "eventual contact" was not significant, we might examine the difference anyway: The confidence interval standard error is SE $\doteq 0.05634$, so a 95% confidence interval for the difference $p_1 - p_2$ is $(0.58 - 0.6873) \pm (1.96)(0.05634) = -0.2177$ to 0.0031. This indicates that the differences could be very small, but they might be substantially large, and so have the potential to reduce nonresponse rates in surveys. At the very least, further study is warranted.

12.29 (a) H_0: $p_1 = p_2$ vs. H_a: $p_1 > p_2$; the populations are much larger than the samples, and 17 (the smallest count) is greater than 5.

(b) $\hat{p} = 0.0632$, $z = 2.926$, and $P = 0.0017$—the difference is statistically significant.

(c) 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.

12.30 (a) H_0: $p_1 = p_2$ vs. H_a: $p_1 \neq p_2$; $P = 0.0028$—the difference is statistically significant. (b) 0.1172 to 0.3919.

12.31 H_0: $p_1 = p_2$ vs. H_a: $p_1 \neq p_2$; $P = 0.9805$—insufficient evidence to reject H_0.

12.32 (a) H_0: $p_1 = p_2$ vs. H_a: $p_1 > p_2$; $P = 0.0335$—reject H_0 (at the 5% level). (b) -0.0053 to 0.2336.

12.33 The population-to-sample ratio is large enough, and the smallest count is 10—twice as big as it needs to be to allow the z procedures. Fatal heart attacks: $z = -2.67$, $P = 0.0076$. Non-fatal heart attacks: $z = -4.58$, $P < 0.000005$. Strokes: $z = 1.43$, $P = 0.1525$. The proportions for both kinds of heart attacks were significantly different; the stroke proportions were not.

12.34 (a) -9.91% to -4.09%. Since 0 is not in this interval, we would reject H_0: $p_1 = p_2$ at the 1% level. (b) -0.7944 to -0.4056. Since 0 is not in this interval, we would reject H_0: $\mu_1 = \mu_2$ at the 1% level.

12.35 (a) $\hat{p} = \frac{5690}{12{,}931} \doteq 0.4400$ and $\text{SE}_{\hat{p}} = \sqrt{(0.4400)(0.5600)/12{,}931} \doteq 0.004365$, so the 95% confidence interval is $0.4400 \pm (1.96)(0.004365) \doteq 0.4315$ to 0.4486.

(b) $\hat{p} \approx 0.4400$, $\hat{p}_2 \approx \frac{1051}{3285} = 0.3200$, and SE ≈ 0.0092; the 95% confidence interval is $(0.4400 - 0.3200) \pm (1.96)(0.0092)$, or 0.102 to 0.138.

(c) 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.

12.36 (a) We must make sure that we draw the samples from a representative cross-section of all Illinois high-school freshmen and seniors. A multistage/cluster sampling procedure could be used in order to ensure that all geographical areas of the state are equally likely to contain sampled individuals.

(b) $\hat{p} = .02025$, and since $n\hat{p} = 34$ and $n(1 - \hat{p}) = 1645$ are both greater than 10, the confidence interval based on z can be used. The 95% confidence interval for p is $.02025 \pm (1.96)\sqrt{((.02025)(.97975)/1679)} = .02025 \pm 0.00674$, or 0.01351 to 0.02699.

(c) Letting p_1 = the proportion of freshmen who have used steroids and p_2 = the proportion of seniors who have used steroids, we test H_0: $p_1 = p_2$ vs. H_a: $p_1 \neq p_2$. From the data, $\hat{p}_1 = .02025$, $\hat{p}_2 = .01057$, and $\hat{p} = .01905$. $n_1\hat{p}$, $n_1(1 - \hat{p})$, $n_2\hat{p}$, and $n_2(1 - \hat{p})$ are all greater than 5, so a normal approximation can be used. Test statistic $z = (.02025 - .01757)/\sqrt{((.01905)(.98095)(1/1679 + 1/1366)} = 0.5382$, and the P-value = 0.59. There is no reason to reject H_0; the difference between p_1, p_2 is not significant.

12.37 (a) $\hat{p} \doteq 0.1486$ and $SE_{\hat{p}} = \sqrt{(0.1486)(0.8514)/148} \doteq 0.02923$, so the 95% confidence interval is $0.1486 \pm (1.96)(0.02923) \doteq 0.0913$ to 0.2059.

(b) $n = (\frac{1.96}{0.04})^2(0.1486)(0.8514) \doteq 303.7$—use $n = 304$. (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) 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.

12.38 H_0: $p_1 = p_2$ vs. H_a: $p_1 \neq p_2$; $P = 0.6981$—insufficient evidence to reject H_0.

12.39 (a) $\hat{p}_m = 0.1415$, $\hat{p}_w = 0.1667$; $P = 0.6981$. (b) $z = 2.12$, $P = 0.0336$. (c) From (a): -0.1056 to 0.1559. From (b): 0.001278 to 0.049036. The larger samples make the margin of error (and thus the length of the confidence interval) smaller.

12.40 For testing H_0: $p = 1/3$ versus H_a: $p > 1/3$, we have $\hat{p} \doteq 0.3786$, and the test statistic is $z = (0.3786 - 1/3)/\sqrt{\frac{(1/3)(2/3)}{803}} \doteq 2.72$. This gives $P = 0.0033$—very strong evidence that more than one-third of this group never use condoms.

12.41 (a) 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) No: $t = -0.8658$, which gives a P-value close to 0.4.

12.42 No—the data is not based on an SRS, and thus the z procedures are not reliable in this case. In particular, a voluntary response sample is typically biased.

12.43 $\hat{p} = \frac{427}{3160} \doteq 0.1351$, and $SE_{\hat{p}} = \sqrt{\hat{p}(1 - \hat{p})/3160} \doteq 0.006081$, so the 99% confidence interval is $0.1351 \pm (2.576)(0.006081) = 0.1194$ to 0.1508.

12.44 To test H_0: $p_1 = p_2$ vs. H_a: $p_1 < p_2$, we find $\hat{p}_1 = \frac{40}{244} \doteq 0.1639$, $\hat{p}_2 = \frac{87}{245} \doteq 0.3551$, and the pooled value $\hat{p} = \frac{40 + 87}{244 + 245} \doteq 0.2597$. Then $SE = \sqrt{\hat{p}(1 - \hat{p})(\frac{1}{244} + \frac{1}{245})} \doteq 0.03966$, so $z = (\hat{p}_1 - \hat{p}_2)/SE \doteq -4.82$. This gives a tiny P-value (7.2×10^{-7}), so we conclude that bupropion increases the success rate.

12.45 (a) $p_0 = \frac{143{,}611}{181{,}535} \doteq 0.7911$. (b) $\hat{p} = \frac{339}{870} \doteq 0.3897$, $\sigma_{\hat{p}} \doteq 0.0138$, and $z = (\hat{p} - p_0)/\sigma_{\hat{p}} \doteq -29.1$, so $P \doteq 0$ (regardless of whether H_a is $p < p_0$ or $p \neq p_0$). This is very strong evidence against H_0; we conclude that Mexican Americans are underrepresented on juries. (c) $\hat{p}_1 = \frac{339}{870} \doteq 0.3897$, while $\hat{p}_2 = \frac{143{,}611 - 339}{181{,}535 - 870} \doteq 0.7930$. Then $\hat{p} \doteq 0.7911$ (the value of p_0 from part (a)), $s_p = 0.0138$, and $z \doteq -29.2$—and again, we have a tiny P-value and reject H_0.

13

Inference for Tables

13.1 (a) (i) $0.20 < P < 0.25$. (ii) $P = 0.235$. (b) (i) $0.02 < P < 0.025$. (ii) $P = 0.0204$. (c) (i) $P > 0.25$. (ii) $P = 0.3172$.

13.2 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. Expected counts: 140.5, 281.5, 32, 46. $X^2 = 161.77$, df = 3. P-value $= 7.6 \times 10^{-35} \approx 0.0000$. Reject H_0. The two distributions are different.

13.3 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. Expected counts: 21 GG, 42 Gg, 21 gg. $X^2 = 5.43$, df = 2. P-value $= P(X_2^2 > 5.43) = 0.0662$. There is no compelling reason to reject H_0 (though the P-value **is** a little on the low side).

13.4 H_0: The ethnicity distribution of the Ph.D. degree in 1994 is the same as it was in 1981. H_a: The ethnicity distribution of the Ph.D. degree in 1994 is different from the distribution in 1981. Expected counts = 300 × (1981 percents) = 237, 12, 4, 8, 1, 38. $X^2 = 61.98$, df = 5. P-value $= P(X_5^2 > 61.98) = 4.734 \times 10^{-12} \approx 0.0000$. We reject H_0 and conclude that the ethnicity distribution of the Ph.D. degree has changed from 1981 to 1994. (b) 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.

13.5 Use a χ^2 goodness of fit test. (b) Use a one-proportion z test. (c) You can construct the interval; however, your ability to generalize may be limited by the fact that your sample of bags is not an SRS. M&M's may be packaged by weight rather than count.

13.6 H_0: The age-group distribution in 1996 is the same as the 1980 distribution. H_a: The age-group distribution in 1996 is different from the 1980 distribution. One simulation produced observed counts: 37, 35, 15, 13. The expected counts: 41.39, 27.68, 19.64, and 11.28 are stored in list L_4, and the difference terms $(O - E)^2/E$ are assigned to L_5.