

## 9.2.7: Exercises

### Inference for a single proportion

**6.1 Vegetarian college students.** Suppose that 8% of college students are vegetarians. Determine if the following statements are true or false, and explain your reasoning.

- (a) The distribution of the sample proportions of vegetarians in random samples of size 60 is approximately normal since  $n \geq 30$ .
- (b) The distribution of the sample proportions of vegetarian college students in random samples of size 50 is right skewed.
- (c) A random sample of 125 college students where 12% are vegetarians would be considered unusual.
- (d) A random sample of 250 college students where 12% are vegetarians would be considered unusual.
- (e) The standard error would be reduced by one-half if we increased the sample size from 125 to 250.

**6.2 Young Americans, Part I.** About 77% of young adults think they can achieve the American dream. Determine if the following statements are true or false, and explain your reasoning.<sup>36</sup>

- (a) The distribution of sample proportions of young Americans who think they can achieve the American dream in samples of size 20 is left skewed.
- (b) The distribution of sample proportions of young Americans who think they can achieve the American dream in random samples of size 40 is approximately normal since  $n \geq 30$ .
- (c) A random sample of 60 young Americans where 85% think they can achieve the American dream would be considered unusual.
- (d) A random sample of 120 young Americans where 85% think they can achieve the American dream would be considered unusual.

**6.3 Orange tabbies.** Suppose that 90% of orange tabby cats are male. Determine if the following statements are true or false, and explain your reasoning.

- (a) The distribution of sample proportions of random samples of size 30 is left skewed.
- (b) Using a sample size that is 4 times as large will reduce the standard error of the sample proportion by one-half.
- (c) The distribution of sample proportions of random samples of size 140 is approximately normal.
- (d) The distribution of sample proportions of random samples of size 280 is approximately normal.

**6.4 Young Americans, Part II.** About 25% of young Americans have delayed starting a family due to the continued economic slump. Determine if the following statements are true or false, and explain your reasoning.<sup>37</sup>

- (a) The distribution of sample proportions of young Americans who have delayed starting a family due to the continued economic slump in random samples of size 12 is right skewed.
- (b) In order for the the distribution of sample proportions of young Americans who have delayed starting a family due to the continued economic slump to be approximately normal, we need random samples where the sample size is at least 40.
- (c) A random sample of 50 young Americans where 20% have delayed starting a family due to the continued economic slump would be considered unusual.
- (d) A random sample of 150 young Americans where 20% have delayed starting a family due to the continued economic slump would be considered unusual.
- (e) Tripling the sample size will reduce the standard error of the sample proportion by one-third.

<sup>36</sup>A. Vaughn. "Poll finds young adults optimistic, but not about money". In: *Los Angeles Times* (2011).

<sup>37</sup>Demos.org. "The State of Young America: The Poll". In: (2011).

**6.5 Prop 19 in California.** In a 2010 Survey USA poll, 70% of the 119 respondents between the ages of 18 and 34 said they would vote in the 2010 general election for Prop 19, which would change California law to legalize marijuana and allow it to be regulated and taxed. At a 95% confidence level, this sample has an 8% margin of error. Based on this information, determine if the following statements are true or false, and explain your reasoning.<sup>38</sup>

- (a) We are 95% confident that between 62% and 78% of the California voters in this sample support Prop 19.
- (b) We are 95% confident that between 62% and 78% of all California voters between the ages of 18 and 34 support Prop 19.
- (c) If we considered many random samples of 119 California voters between the ages of 18 and 34, and we calculated 95% confidence intervals for each, 95% of them will include the true population proportion of Californians who support Prop 19.
- (d) In order to decrease the margin of error to 4%, we would need to quadruple (multiply by 4) the sample size.
- (e) Based on this confidence interval, there is sufficient evidence to conclude that a majority of California voters between the ages of 18 and 34 support Prop 19.

**6.6 2010 Healthcare Law.** On June 28, 2012 the U.S. Supreme Court upheld the much debated 2010 healthcare law, declaring it constitutional. A Gallup poll released the day after this decision indicates that 46% of 1,012 Americans agree with this decision. At a 95% confidence level, this sample has a 3% margin of error. Based on this information, determine if the following statements are true or false, and explain your reasoning.<sup>39</sup>

- (a) We are 95% confident that between 43% and 49% of Americans in this sample support the decision of the U.S. Supreme Court on the 2010 healthcare law.
- (b) We are 95% confident that between 43% and 49% of Americans support the decision of the U.S. Supreme Court on the 2010 healthcare law.
- (c) If we considered many random samples of 1,012 Americans, and we calculated the sample proportions of those who support the decision of the U.S. Supreme Court, 95% of those sample proportions will be between 43% and 49%.
- (d) The margin of error at a 90% confidence level would be higher than 3%.

**6.7 Fireworks on July 4th.** In late June 2012, Survey USA published results of a survey stating that 56% of the 600 randomly sampled Kansas residents planned to set off fireworks on July 4th. Determine the margin of error for the 56% point estimate using a 95% confidence level.<sup>40</sup>

**6.8 Elderly drivers. In January 2011,** The Marist Poll published a report stating that 66% of adults nationally think licensed drivers should be required to retake their road test once they reach 65 years of age. It was also reported that interviews were conducted on 1,018 American adults, and that the margin of error was 3% using a 95% confidence level.<sup>41</sup>

- (a) Verify the margin of error reported by The Marist Poll.
- (b) Based on a 95% confidence interval, does the poll provide convincing evidence that more than 70% of the population think that licensed drivers should be required to retake their road test once they turn 65?

<sup>38</sup>Survey USA, Election Poll #16804, data collected July 8-11, 2010.

<sup>39</sup>Gallup, Americans Issue Split Decision on Healthcare Ruling, data collected June 28, 2012.

<sup>40</sup>Survey USA, News Poll #19333, data collected on June 27, 2012.

<sup>41</sup>Marist Poll, Road Rules: Re-Testing Drivers at Age 65?, March 4, 2011.

**6.9 Life after college.** We are interested in estimating the proportion of graduates at a mid-sized university who found a job within one year of completing their undergraduate degree. Suppose we conduct a survey and find out that 348 of the 400 randomly sampled graduates found jobs. The graduating class under consideration included over 4500 students.

- (a) Describe the population parameter of interest. What is the value of the point estimate of this parameter?
- (b) Check if the conditions for constructing a confidence interval based on these data are met.
- (c) Calculate a 95% confidence interval for the proportion of graduates who found a job within one year of completing their undergraduate degree at this university, and interpret it in the context of the data.
- (d) What does "95% confidence" mean?
- (e) Now calculate a 99% confidence interval for the same parameter and interpret it in the context of the data.
- (f) Compare the widths of the 95% and 99% confidence intervals. Which one is wider? Explain.

**6.10 Life rating in Greece.** Greece has faced a severe economic crisis since the end of 2009. A Gallup poll surveyed 1,000 randomly sampled Greeks in 2011 and found that 25% of them said they would rate their lives poorly enough to be considered

"suffering".<sup>42</sup>

- (a) Describe the population parameter of interest. What is the value of the point estimate of this parameter?
- (b) Check if the conditions required for constructing a confidence interval based on these data are met.
- (c) Construct a 95% confidence interval for the proportion of Greeks who are "suffering".
- (d) Without doing any calculations, describe what would happen to the confidence interval if we decided to use a higher confidence level.
- (e) Without doing any calculations, describe what would happen to the confidence interval if we used a larger sample.

**6.11 Study abroad.** A survey on 1,509 high school seniors who took the SAT and who completed an optional web survey between April 25 and April 30, 2007 shows that 55% of high school seniors are fairly certain that they will participate in a study abroad program in college.<sup>43</sup>

- (a) Is this sample a representative sample from the population of all high school seniors in the US? Explain your reasoning.
- (b) Let's suppose the conditions for inference are met. Even if your answer to part (a) indicated that this approach would not be reliable, this analysis may still be interesting to carry out (though not report). Construct a 90% confidence interval for the proportion of high school seniors (of those who took the SAT) who are fairly certain they will participate in a study abroad program in college, and interpret this interval in context.
- (c) What does "90% confidence" mean?
- (d) Based on this interval, would it be appropriate to claim that the majority of high school seniors are fairly certain that they will participate in a study abroad program in college?

<sup>42</sup>Gallup World, More Than One in 10 "Suffering" Worldwide, data collected throughout 2011.

<sup>43</sup>studentPOLL, College-Bound Students' Interests in Study Abroad and Other International Learning Activities, January 2008.

**6.12 Legalization of marijuana, Part I.** The 2010 General Social Survey asked 1,259 US residents: "Do you think the use of marijuana should be made legal, or not?" 48% of the respondents said it should be made legal.<sup>44</sup>

- (a) Is 48% a sample statistic or a population parameter? Explain.
- (b) Construct a 95% confidence interval for the proportion of US residents who think marijuana should be made legal, and interpret it in the context of the data.
- (c) A critic points out that this 95% confidence interval is only accurate if the statistic follows a normal distribution, or if the normal model is a good approximation. Is this true for these data? Explain.
- (d) A news piece on this survey's findings states, "Majority of Americans think marijuana should be legalized." Based on your confidence interval, is this news piece's statement justified?

**6.13 Public option, Part I.** A Washington Post article from 2009 reported that "support for a government-run health-care plan to compete with private insurers has rebounded from its summertime lows and wins clear majority support from the public." More specifically, the article says "seven in 10 Democrats back the plan, while almost nine in 10 Republicans oppose it. Independents divide 52 percent against, 42 percent in favor of the legislation." There were 819 Democrats, 566 Republicans and 783 Independents surveyed.<sup>45</sup>

- (a) A political pundit on TV claims that a majority of Independents oppose the health care public option plan. Do these data provide strong evidence to support this statement?
- (b) Would you expect a confidence interval for the proportion of Independents who oppose the public option plan to include 0.5? Explain.

**6.14 The Civil War.** A national survey conducted in 2011 among a simple random sample of 1,507 adults shows that 56% of Americans think the Civil War is still relevant to American politics and political life.<sup>46</sup>

- (a) Conduct a hypothesis test to determine if these data provide strong evidence that the majority of the Americans think the Civil War is still relevant.
- (b) Interpret the p-value in this context.

(c) Calculate a 90% confidence interval for the proportion of Americans who think the Civil War is still relevant. Interpret the interval in this context, and comment on whether or not the confidence interval agrees with the conclusion of the hypothesis test.

**6.15 Browsing on the mobile device.** A 2012 survey of 2,254 American adults indicates that 17% of cell phone owners do their browsing on their phone rather than a computer or other device.<sup>47</sup>

(a) According to an online article, a report from a mobile research company indicates that 38 percent of Chinese mobile web users only access the internet through their cell phones.<sup>48</sup> Conduct a hypothesis test to determine if these data provide strong evidence that the proportion of Americans who only use their cell phones to access the internet is different than the Chinese proportion of 38%.

(b) Interpret the p-value in this context.

(c) Calculate a 95% confidence interval for the proportion of Americans who access the internet on their cell phones, and interpret the interval in this context.

<sup>44</sup>National Opinion Research Center, *General Social Survey, 2010*.

<sup>45</sup>D. Balz and J. Cohen. "Most support public option for health insurance, poll finds". In: *The Washington Post* (2009).

<sup>46</sup>Pew Research Center Publications, *Civil War at 150: Still Relevant, Still Divisive, data collected between March 30 - April 3, 2011*.

<sup>47</sup>Pew Internet, *Cell Internet Use 2012, data collected between March 15 - April 13, 2012*.

<sup>48</sup>S. Chang. "The Chinese Love to Use Feature Phone to Access the Internet". In: *M.I.C Gadget* (2012).

**6.16 Is college worth it? Part I.** Among a simple random sample of 331 American adults who do not have a four-year college degree and are not currently enrolled in school, 48% said they decided not to go to college because they could not afford school.<sup>49</sup>

(a) A newspaper article states that only a minority of the Americans who decide not to go to college do so because they cannot afford it and uses the point estimate from this survey as evidence. Conduct a hypothesis test to determine if these data provide strong evidence supporting this statement.

(b) Would you expect a confidence interval for the proportion of American adults who decide not to go to college because they cannot afford it to include 0.5? Explain.

**6.17 Taste test.** Some people claim that they can tell the difference between a diet soda and a regular soda in the first sip. A researcher wanting to test this claim randomly sampled 80 such people. He then filled 80 plain white cups with soda, half diet and half regular through random assignment, and asked each person to take one sip from their cup and identify the soda as diet or regular. 53 participants correctly identified the soda.

(a) Do these data provide strong evidence that these people are able to detect the difference between diet and regular soda, in other words, are the results significantly better than just random guessing?

(b) Interpret the p-value in this context.

**6.18 Is college worth it? Part II.** Exercise 6.16 presents the results of a poll where 48% of 331 Americans who decide to not go to college do so because they cannot afford it.

(a) Calculate a 90% confidence interval for the proportion of Americans who decide to not go to college because they cannot afford it, and interpret the interval in context.

(b) Suppose we wanted the margin of error for the 90% confidence level to be about 1.5%. How large of a survey would you recommend?

**6.19 College smokers.** We are interested in estimating the proportion of students at a university who smoke. Out of a random sample of 200 students from this university, 40 students smoke.

(a) Calculate a 95% confidence interval for the proportion of students at this university who smoke, and interpret this interval in context. (Reminder: check conditions)

(b) If we wanted the margin of error to be no larger than 2% at a 95% confidence level for the proportion of students who smoke, how big of a sample would we need?

**6.20 Legalize Marijuana, Part II.** As discussed in Exercise 6.12, the 2010 General Social Survey reported a sample where about 48% of US residents thought marijuana should be made legal. If we wanted to limit the margin of error of a 95% confidence interval to 2%, about how many Americans would we need to survey ?

**6.21 Public option, Part II.** Exercise 6.13 presents the results of a poll evaluating support for the health care public option in 2009, reporting that 52% of Independents in the sample opposed the public option. If we wanted to estimate this number to within 1% with 90% confidence, what would be an appropriate sample size?

**6.22 Acetaminophen and liver damage.** It is believed that large doses of acetaminophen (the active ingredient in over the counter pain relievers like Tylenol) may cause damage to the liver. A researcher wants to conduct a study to estimate the proportion of acetaminophen users who have liver damage. For participating in this study, he will pay each subject \$20 and provide a free medical consultation if the patient has liver damage.

(a) If he wants to limit the margin of error of his 98% confidence interval to 2%, what is the minimum amount of money he needs to set aside to pay his subjects?

(b) The amount you calculated in part (a) is substantially over his budget so he decides to use fewer subjects. How will this affect the width of his confidence interval?

<sup>49</sup>Pew Research Center Publications, *Is College Worth It?*, data collected between March 15-29, 2011.

## Difference of two proportions

**6.23 Social experiment, Part I.** A "social experiment" conducted by a TV program questioned what people do when they see a very obviously bruised woman getting picked on by her boyfriend. On two different occasions at the same restaurant, the same couple was depicted. In one scenario the woman was dressed "provocatively" and in the other scenario the woman was dressed "conservatively". The table below shows how many restaurant diners were present under each scenario, and whether or not they intervened.

	Scenario		
	Provocative	Conservative	Total
Yes	5	15	20
No	15	10	25
Total	20	25	45

Explain why the sampling distribution of the difference between the proportions of interventions under provocative and conservative scenarios does not follow an approximately normal distribution.

**6.24 Heart transplant success.** The Stanford University Heart Transplant Study was conducted to determine whether an experimental heart transplant program increased lifespan. Each patient entering the program was officially designated a heart transplant candidate, meaning that he was gravely ill and might benefit from a new heart. Patients were randomly assigned into treatment and control groups. Patients in the treatment group received a transplant, and those in the control group did not. The table below displays how many patients survived and died in each group.<sup>50</sup>

	Control	treatment
alive	4	24
dead	30	45

A hypothesis test would reject the conclusion that the survival rate is the same in each group, and so we might like to calculate a confidence interval. Explain why we cannot construct such an interval using the normal approximation. What might go wrong if we constructed the confidence interval despite this problem?

**6.25 Gender and color preference.** A 2001 study asked 1,924 male and 3,666 female undergraduate college students their favorite color. A 95% confidence interval for the difference between the proportions of males and females whose favorite color is black ( $p_{\text{male}} - p_{\text{female}}$ ) was calculated to be (0.02, 0.06). Based on this information, determine if the following statements are true or false, and explain your reasoning for each statement you identify as false.<sup>51</sup>

- (a) We are 95% confident that the true proportion of males whose favorite color is black is 2% lower to 6% higher than the true proportion of females whose favorite color is black.
- (b) We are 95% confident that the true proportion of males whose favorite color is black is 2% to 6% higher than the true proportion of females whose favorite color is black.
- (c) 95% of random samples will produce 95% confidence intervals that include the true difference between the population proportions of males and females whose favorite color is black.
- (d) We can conclude that there is a significant difference between the proportions of males and females whose favorite color is black and that the difference between the two sample proportions is too large to plausibly be due to chance.
- (e) The 95% confidence interval for  $(p_{female} - p_{male})$  cannot be calculated with only the information given in this exercise.

<sup>50</sup>B. Turnbull et al. "Survivorship of Heart Transplant Data". In: *Journal of the American Statistical Association* 69 (1974), pp. 74 - 80.

<sup>51</sup>L. Ellis and C. Fiske. "Color preferences according to gender and sexual orientation". In: *Personality and Individual Differences* 31.8 (2001), pp. 1375-1379.

**6.26 The Daily Show.** A 2010 Pew Research foundation poll indicates that among 1,099 college graduates, 33% watch The Daily Show. Meanwhile, 22% of the 1,110 people with a high school degree but no college degree in the poll watch The Daily Show. A 95% confidence interval for  $(p_{collegegrad} - p_{HSorless})$ , where  $p$  is the proportion of those who watch The Daily Show, is (0.07, 0.15). Based on this information, determine if the following statements are true or false, and explain your reasoning if you identify the statement as false.<sup>52</sup>

- (a) At the 5% significance level, the data provide convincing evidence of a difference between the proportions of college graduates and those with a high school degree or less who watch The Daily Show.
- (b) We are 95% confident that 7% less to 15% more college graduates watch The Daily Show than those with a high school degree or less.
- (c) 95% of random samples of 1,099 college graduates and 1,110 people with a high school degree or less will yield differences in sample proportions between 7% and 15%.
- (d) A 90% confidence interval for  $(p_{collegegrad} - p_{HSorless})$  would be wider.
- (e) A 95% confidence interval for  $(p_{HSorless} - p_{collegegrad})$  is (-0.15, -0.07).

**6.27 Public Option, Part III.** Exercise 6.13 presents the results of a poll evaluating support for the health care public option plan in 2009. 70% of 819 Democrats and 42% of 783 Independents support the public option.

- (a) Calculate a 95% confidence interval for the difference between  $(p_D - p_I)$  and interpret it in this context. We have already checked conditions for you.
- (b) True or false: If we had picked a random Democrat and a random Independent at the time of this poll, it is more likely that the Democrat would support the public option than the Independent.

**6.28 Sleep deprivation, CA vs. OR, Part I.** According to a report on sleep deprivation by the Centers for Disease Control and Prevention, the proportion of California residents who reported insufficient rest or sleep during each of the preceding 30 days is 8.0%, while this proportion is 8.8% for Oregon residents. These data are based on simple random samples of 11,545 California and 4,691 Oregon residents. Calculate a 95% confidence interval for the difference between the proportions of Californians and Oregonians who are sleep deprived and interpret it in context of the data.<sup>53</sup>

**6.29 Offshore drilling, Part I.** A 2010 survey asked 827 randomly sampled registered voters in California "Do you support? Or do you oppose? Drilling for oil and natural gas off the Coast of California? Or do you not know enough to say?" Below is the distribution of responses, separated based on whether or not the respondent graduated from college.<sup>54</sup>

	College Grad	
	Yes	No

Support	154	132
Oppose	180	126
Do not know	104	131
Total	438	389

(a) What percent of college graduates and what percent of the non-college graduates in this sample do not know enough to have an opinion on drilling for oil and natural gas off the Coast of California?

(b) Conduct a hypothesis test to determine if the data provide strong evidence that the proportion of college graduates who do not have an opinion on this issue is different than that of non-college graduates.

<sup>52</sup>The Pew Research Center, *Americans Spending More Time Following the News*, data collected June 8-28, 2010.

<sup>53</sup>CDC, *Perceived Insufficient Rest or Sleep Among Adults - United States, 2008*.

<sup>54</sup>Survey USA, *Election Poll #16804*, data collected July 8-11, 2010.

**6.30 Sleep deprivation, CA vs. OR, Part II.** Exercise 6.28 provides data on sleep deprivation rates of Californians and Oregonians. The proportion of California residents who reported insufficient rest or sleep during each of the preceding 30 days is 8.0%, while this proportion is 8.8% for Oregon residents. These data are based on simple random samples of 11,545 California and 4,691 Oregon residents.

(a) Conduct a hypothesis test to determine if these data provide strong evidence the rate of sleep deprivation is different for the two states. (Reminder: check conditions)

(b) It is possible the conclusion of the test in part (a) is incorrect. If this is the case, what type of error was made?

**6.31 Offshore drilling, Part II.** Results of a poll evaluating support for drilling for oil and natural gas off the coast of California were introduced in Exercise 6.29.

	College Grad	
	Yes	No
Support	154	132
Oppose	180	126
Do not know	104	131
Total	438	389

(a) What percent of college graduates and what percent of the non-college graduates in this sample support drilling for oil and natural gas off the Coast of California?

(b) Conduct a hypothesis test to determine if the data provide strong evidence that the proportion of college graduates who support offshore drilling in California is different than that of noncollege graduates.

**6.32 Full body scan, Part I.** A news article reports that "Americans have differing views on two potentially inconvenient and invasive practices that airports could implement to uncover potential terrorist attacks." This news piece was based on a survey conducted among a random sample of 1,137 adults nationwide, interviewed by telephone November 7-10, 2010, where one of the questions on the survey was "Some airports are now using 'full-body' digital x-ray machines to electronically screen passengers in airport security lines. Do you think these new x-ray machines should or should not be used at airports?" Below is a summary of responses based on party affiliation.<sup>55</sup>

	Party Affiliation		
	Republican	Democrat	Independent
Should	264	299	351
Should not	38	55	77
Don't know/No answer	16	15	22

Total	318	369	450
-------	-----	-----	-----

- (a) Conduct an appropriate hypothesis test evaluating whether there is a difference in the proportion of Republicans and Democrats who think the full-body scans should be applied in airports. Assume that all relevant conditions are met.
- (b) The conclusion of the test in part (a) may be incorrect, meaning a testing error was made. If an error was made, was it a Type I or a Type II error? Explain.

<sup>55</sup>S. Condon. "Poll: 4 in 5 Support Full-Body Airport Scanners". In: CBS News (2010).

**6.33 Sleep deprived transportation workers.** The National Sleep Foundation conducted a survey on the sleep habits of randomly sampled transportation workers and a control sample of non-transportation workers. The results of the survey are shown below.<sup>56</sup>

		Transportation	Professionals		
	Control	Pilots	Truck Drivers	Train Operators	Bux/Taxi/Limo Drivers
Less than 6 hours of sleep	35	19	35	29	21
6 to 8 hours of sleep	193	132	117	119	131
More than 8 hours	64	51	51	32	58
Tota	292	202	203	180	210

Conduct a hypothesis test to evaluate if these data provide evidence of a difference between the proportions of truck drivers and non-transportation workers (the control group) who get less than 6 hours of sleep per day, i.e. are considered sleep deprived.

**6.34 Prenatal vitamins and Autism.** Researchers studying the link between prenatal vitamin use and autism surveyed the mothers of a random sample of children aged 24 - 60 months with autism and conducted another separate random sample for children with typical development. The table below shows the number of mothers in each group who did and did not use prenatal vitamins during the three months before pregnancy (periconceptional period).<sup>57</sup>

	Autism		
	Autism	Typical development	Total
No vitamin	111	70	181
Vitamin	143	159	302
Total	254	229	483

- (a) State appropriate hypotheses to test for independence of use of prenatal vitamins during the three months before pregnancy and autism.
- (b) Complete the hypothesis test and state an appropriate conclusion. (Reminder: verify any necessary conditions for the test.)
- (c) A New York Times article reporting on this study was titled "Prenatal Vitamins May Ward Off Autism". Do you nd the title of this article to be appropriate? Explain your answer. Additionally, propose an alternative title.<sup>58</sup>

**6.35 HIV in sub-Saharan Africa.** In July 2008 the US National Institutes of Health announced that it was stopping a clinical study early because of unexpected results. The study population consisted of HIV-infected women in sub-Saharan Africa who had been given single dose Nevirapine (a treatment for HIV) while giving birth, to prevent transmission of HIV to the infant. The study was a randomized comparison of continued treatment of a woman (after successful childbirth) with Nevirapine vs. Lopinavir, a second drug used to treat HIV. 240 women participated in the study; 120 were randomized to each of the two treatments. Twenty-four weeks after starting the study treatment, each woman was tested to determine if the HIV infection was becoming worse (an outcome called virologic failure). Twenty-six of the 120 women treated with Nevirapine experienced virologic failure, while 10 of the 120 women treated with the other drug experienced virologic failure.<sup>59</sup>

- (a) Create a two-way table presenting the results of this study.

- (b) State appropriate hypotheses to test for independence of treatment and virologic failure.
- (c) Complete the hypothesis test and state an appropriate conclusion. (Reminder: verify any necessary conditions for the test.)

<sup>56</sup>National Sleep Foundation, 2012 Sleep in America Poll: Transportation Workers Sleep, 2012.

<sup>57</sup>R.J. Schmidt et al. "Prenatal vitamins, one-carbon metabolism gene variants, and risk for autism". In: *Epidemiology* 22.4 (2011), p. 476.

<sup>58</sup>R.C. Rabin. "Patterns: Prenatal Vitamins May Ward Off Autism". In: *New York Times* (2011).

<sup>59</sup>S. Lockman et al. "Response to antiretroviral therapy after a single, peripartum dose of nevirapine". In: *Obstetrical & gynecological survey* 62.6 (2007), p. 361.

**6.36 Diabetes and unemployment.** A 2012 Gallup poll surveyed Americans about their employment status and whether or not they have diabetes. The survey results indicate that 1.5% of the 47,774 employed (full or part time) and 2.5% of the 5,855 unemployed 18-29 year olds have diabetes.<sup>60</sup>

- (a) Create a two-way table presenting the results of this study.
- (b) State appropriate hypotheses to test for independence of incidence of diabetes and employment status.
- (c) The sample difference is about 1%. If we completed the hypothesis test, we would find that the p-value is very small (about 0), meaning the difference is statistically significant. Use this result to explain the difference between statistically significant and practically significant findings.

### Testing for goodness of fit using chi-square

**6.37 True or false, Part I.** Determine if the statements below are true or false. For each false statement, suggest an alternative wording to make it a true statement.

- (a) The chi-square distribution, just like the normal distribution, has two parameters, mean and standard deviation.
- (b) The chi-square distribution is always right skewed, regardless of the value of the degrees of freedom parameter.
- (c) The chi-square statistic is always positive.
- (d) As the degrees of freedom increases, the shape of the chi-square distribution becomes more skewed.

**6.38 True or false, Part II.** Determine if the statements below are true or false. For each false statement, suggest an alternative wording to make it a true statement.

- (a) As the degrees of freedom increases, the mean of the chi-square distribution increases.
- (b) If you found  $\chi^2 = 10$  with  $df = 5$  you would fail to reject  $H_0$  at the 5% significance level.
- (c) When finding the p-value of a chi-square test, we always shade the tail areas in both tails.
- (d) As the degrees of freedom increases, the variability of the chi-square distribution decreases.

**6.39 Open source textbook.** A professor using an open source introductory statistics book predicts that 60% of the students will purchase a hard copy of the book, 25% will print it out from the web, and 15% will read it online. At the end of the semester he asks his students to complete a survey where they indicate what format of the book they used. Of the 126 students, 71 said they bought a hard copy of the book, 30 said they printed it out from the web, and 25 said they read it online.

- (a) State the hypotheses for testing if the professor's predictions were inaccurate.
- (b) How many students did the professor expect to buy the book, print the book, and read the book exclusively online?
- (c) This is an appropriate setting for a chi-square test. List the conditions required for a test and verify they are satisfied.
- (d) Calculate the chi-squared statistic, the degrees of freedom associated with it, and the p-value.
- (e) Based on the p-value calculated in part (d), what is the conclusion of the hypothesis test? Interpret your conclusion in this context.

<sup>60</sup>Gallup Wellbeing, *Employed Americans in Better Health Than the Unemployed*, data collected Jan. 2, 2011 - May 21, 2012.

**6.40 Evolution vs. creationism.** A Gallup Poll released in December 2010 asked 1019 adults living in the Continental U.S. about their belief in the origin of humans. These results, along with results from a more comprehensive poll from 2001 (that we will assume to be exactly accurate), are summarized in the table below:<sup>61</sup>

Response	Year	
	2010	2001
Humans evolved, with God guiding (1)	38%	37%
Humans evolved, but God had no part in process (2)	16%	12%
God created humans in present form (3)	40%	45%
Other / No opinion (4)	6%	6%

- Calculate the actual number of respondents in 2010 that fall in each response category.
- State hypotheses for the following research question: have beliefs on the origin of human life changed since 2001?
- Calculate the expected number of respondents in each category under the condition that the null hypothesis from part (b) is true.
- Conduct a chi-square test and state your conclusion. (Reminder: verify conditions.)

### Testing for independence in two-way tables

**6.41 Quitters.** Does being part of a support group affect the ability of people to quit smoking? A county health department enrolled 300 smokers in a randomized experiment. 150 participants were assigned to a group that used a nicotine patch and met weekly with a support group; the other 150 received the patch and did not meet with a support group. At the end of the study, 40 of the participants in the patch plus support group had quit smoking while only 30 smokers had quit in the other group.

- Create a two-way table presenting the results of this study.
- Answer each of the following questions under the null hypothesis that being part of a support group does not affect the ability of people to quit smoking, and indicate whether the expected values are higher or lower than the observed values.
  - How many subjects in the "patch + support" group would you expect to quit?
  - How many subjects in the "only patch" group would you expect to not quit?

**6.42 Full body scan, Part II.** The table below summarizes a data set we first encountered in Exercise 6.32 regarding views on full-body scans and political affiliation. The differences in each political group may be due to chance. Complete the following computations under the null hypothesis of independence between an individual's party affiliation and his support of full-body scans. It may be useful to first add on an extra column for row totals before proceeding with the computations.

	Party Affiliation		
	Republican	Democrat	Independent
Should	264	299	351
Should not	38	55	77
Don't know/No answer	16	15	22
Total	318	369	450

- How many Republicans would you expect to not support the use of full-body scans?
- How many Democrats would you expect to support the use of full-body scans?
- How many Independents would you expect to not know or not answer?

<sup>61</sup>Four in 10 Americans Believe in Strict Creationism, December 17, 2010, <http://www.gallup.com/poll/145286/Four-Americans-Believe-Strict-Creationism.aspx>.

**6.43 Offshore drilling, Part III.** The table below summarizes a data set we first encountered in Exercise 6.29 that examines the responses of a random sample of college graduates and nongraduates on the topic of oil drilling. Complete a chi-square test for

these data to check whether there is a statistically significant difference in responses from college graduates and non-graduates.

	College Grad	
	Yes	No
Support	154	132
Oppose	180	126
Do not know	104	131
Total	438	389

**6.44 Coffee and Depression.** Researchers conducted a study investigating the relationship between caffeinated coffee consumption and risk of depression in women. They collected data on 50,739 women free of depression symptoms at the start of the study in the year 1996, and these women were followed through 2006. The researchers used questionnaires to collect data on caffeinated coffee consumption, asked each individual about physician-diagnosed depression, and also asked about the use of antidepressants. The table below shows the distribution of incidences of depression by amount of caffeinated coffee consumption.<sup>62</sup>

	Caffeinated coffee consumption					
	≤ 1cup/week	2-6 cups/week	1 cup/day	2-3 cups/day	≥ 4 cups/day	Total
Yes	670	373	905	564	95	2,607
No	11,545	6,244	16,329	11,726	2,288	48,132
Total	12,215	6,617	17,234	12,290	2,383	50,739

- What type of test is appropriate for evaluating if there is an association between coffee intake and depression?
- Write the hypotheses for the test you identified in part (a).
- Calculate the overall proportion of women who do and do not suffer from depression.
- Identify the expected count for the highlighted cell, and calculate the contribution of this cell to the test statistic, i.e.  $\frac{(Observed - Expected)^2}{Expected}$ .
- The test statistic is  $X^2 = 20.93$ . What is the p-value?
- What is the conclusion of the hypothesis test?
- One of the authors of this study was quoted on the NYTimes as saying it was "too early to recommend that women load up on extra coffee" based on just this study.<sup>63</sup> Do you agree with this statement? Explain your reasoning.

<sup>62</sup>M. Lucas et al. "Coffee, caffeine, and risk of depression among women". In: *Archives of internal medicine* 171.17 (2011), p. 1571.

<sup>63</sup>A. O'Connor. "Coffee Drinking Linked to Less Depression in Women". In: *New York Times* (2011).

**6.45 Privacy on Facebook.** A 2011 survey asked 806 randomly sampled adult Facebook users about their Facebook privacy settings. One of the questions on the survey was, "Do you know how to adjust your Facebook privacy settings to control what people can and cannot see?" The responses are cross-tabulated based on gender.<sup>64</sup>

	Gender		Total
	Male	Female	
Yes	288	378	666
No	61	62	123
Not sure	10	7	17
Total	359	447	806

- (a) State appropriate hypotheses to test for independence of gender and whether or not Facebook users know how to adjust their privacy settings.
- (b) Verify any necessary conditions for the test and determine whether or not a chi-square test can be completed.

**6.46 Shipping holiday gifts.** A December 2010 survey asked 500 randomly sampled Los Angeles residents which shipping carrier they prefer to use for shipping holiday gifts. The table below shows the distribution of responses by age group as well as the expected counts for each cell (shown in parentheses).

	Age			
	18-34	35-54	55+	Total
USPS	72 (81)	97 (102)	76 (62)	245
UPS	52 (53)	76 (68)	34 (41)	162
FedEx	31 (21)	24 (27)	9 (16)	64
Something else	7 (5)	6 (7)	3 (4)	16
Not sure	3 (5)	6 (5)	4 (3)	13
Total	165	209	126	500

- (a) State the null and alternative hypotheses for testing for independence of age and preferred shipping method for holiday gifts among Los Angeles residents.
- (b) Are the conditions for inference using a chi-square test satisfied?

### Small sample hypothesis testing for a proportion

**6.47 Bullying in schools.** A 2012 Survey USA poll asked Florida residents how big of a problem they thought bullying was in local schools. 9 out of 191 18-34 year olds responded that bullying is no problem at all. Using these data, is it appropriate to construct a confidence interval using the formula  $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$  for the true proportion of 18-34 year old Floridians who think bullying is no problem at all? If it is appropriate, construct the confidence interval. If it is not, explain why.

<sup>64</sup>Survey USA, News Poll #17960, data collected February 16-17, 2011.

**6.48 Choose a test.** We would like to test the following hypotheses:

$$H_0 : p = 0.1$$

$$H_A : p \neq 0.1$$

The sample size is 120 and the sample proportion is 8.5%. Determine which of the below test(s) is/are appropriate for this situation and explain your reasoning.

- I. Z test for a proportion,  
i.e. proportion test using normal model
- II. Z test for comparing two proportions
- III.  $\chi^2$  test of independence
- IV. Simulation test for a proportion
- V. t test for a mean
- VI. ANOVA

**6.49 The Egyptian Revolution.** A popular uprising that started on January 25, 2011 in Egypt led to the 2011 Egyptian Revolution. Polls show that about 69% of American adults followed the news about the political crisis and demonstrations in Egypt closely during the first couple weeks following the start of the uprising. Among a random sample of 30 high school students, it was found that only 17 of them followed the news about Egypt closely during this time.<sup>65</sup>

- (a) Write the hypotheses for testing if the proportion of high school students who followed the news about Egypt is different than the proportion of American adults who did.

- (b) Calculate the proportion of high schoolers in this sample who followed the news about Egypt closely during this time.
- (c) Based on large sample theory, we modeled  $\hat{p}$  using the normal distribution. Why should we be cautious about this approach for these data?
- (d) The normal approximation will not be as reliable as a simulation, especially for a sample of this size. Describe how to perform such a simulation and, once you had results, how to estimate the p-value.
- (e) Below is a histogram showing the distribution of  $\hat{p}_{sim}$  in 10,000 simulations under the null hypothesis. Estimate the p-value using the plot and determine the conclusion of the hypothesis test.

<sup>65</sup>Gallup Politics, *Americans' Views of Egypt Sharply More Negative, data collected February 2-5, 2011.*

**6.50 Assisted Reproduction.** Assisted Reproductive Technology (ART) is a collection of techniques that help facilitate pregnancy (e.g. in vitro fertilization). A 2008 report by the Centers for Disease Control and Prevention estimated that ART has been successful in leading to a live birth in 31% of cases<sup>66</sup>. A new fertility clinic claims that their success rate is higher than average. A random sample of 30 of their patients yielded a success rate of 40%. A consumer watchdog group would like to determine if this provides strong evidence to support the company's claim.

- (a) Write the hypotheses to test if the success rate for ART at this clinic is significantly higher than the success rate reported by the CDC.
- (b) Based on large sample theory, we modeled  $\hat{p}$  using the normal distribution. Why is this not appropriate here?
- (c) The normal approximation would be less reliable here, so we should use a simulation strategy. Describe a setup for a simulation that would be appropriate in this situation and how the p-value can be calculated using the simulation results.
- (d) Below is a histogram showing the distribution of  $\hat{p}_{sim}$  in 10,000 simulations under the null hypothesis. Estimate the p-value using the plot and use it to evaluate the hypotheses.
- (e) After performing this analysis, the consumer group releases the following news headline: "Infertility clinic falsely advertises better success rates". Comment on the appropriateness of this statement.

<sup>66</sup>CDC. *2008 Assisted Reproductive Technology Report.*

## Hypothesis testing for two proportions

**6.51 Social experiment, Part II.** Exercise 6.23 introduces a "social experiment" conducted by a TV program that questioned what people do when they see a very obviously bruised woman getting picked on by her boyfriend. On two different occasions at the same restaurant, the same couple was depicted. In one scenario the woman was dressed "provocatively" and in the other scenario the woman was dressed "conservatively". The table below shows how many restaurant diners were present under each scenario, and whether or not they intervened.

	Scenario		
	Provocative	Conservative	Total
Yes	5	15	20
No	15	10	25
Total	20	25	45

A simulation was conducted to test if people react differently under the two scenarios. 10,000 simulated differences were generated to construct the null distribution shown. The value  $\hat{p}_{pr,sim}$  represents the proportion of diners who intervened in the simulation for the provocatively dressed woman, and  $\hat{p}_{con,sim}$  is the proportion for the conservatively dressed woman.

- (a) What are the hypotheses? For the purposes of this exercise, you may assume that each observed person at the restaurant behaved independently, though we would want to evaluate this assumption more rigorously if we were reporting these results.
- (b) Calculate the observed difference between the rates of intervention under the provocative and conservative scenarios:  $\hat{p}_{pr} - \hat{p}_{con}$ .
- (c) Estimate the p-value using the figure above and determine the conclusion of the hypothesis test.

**6.52 Is yawning contagious?** An experiment conducted by the MythBusters, a science entertainment TV program on the Discovery Channel, tested if a person can be subconsciously influenced into yawning if another person near them yawns. 50 people were randomly assigned to two groups: 34 to a group where a person near them yawned (treatment) and 16 to a group where there wasn't a person yawning near them (control). The following table shows the results of this experiment.<sup>67</sup>

	Group		
	Treatment	Control	Total
Yawn	10	4	14
Not Yawn	24	12	36
Total	34	16	50

A simulation was conducted to understand the distribution of the test statistic under the assumption of independence: having someone yawn near another person has no influence on if the other person will yawn. In order to conduct the simulation, a researcher wrote yawn on 14 index cards and not yawn on 36 index cards to indicate whether or not a person yawned. Then he shuffled the cards and dealt them into two groups of size 34 and 16 for treatment and control, respectively. He counted how many participants in each simulated group yawned in an apparent response to a nearby yawning person, and calculated the difference between the simulated proportions of yawning as  $\hat{p}_{trtmt;sim} - \hat{p}_{ctrl;sim}$ . This simulation was repeated 10,000 times using software to obtain 10,000 differences that are due to chance alone. The histogram shows the distribution of the simulated differences.

- What are the hypotheses?
- Calculate the observed difference between the yawning rates under the two scenarios.
- Estimate the p-value using the figure above and determine the conclusion of the hypothesis test.

<sup>67</sup>MythBusters, Season 3, Episode 28.

## Contributors

David M Diez (Google/YouTube), Christopher D Barr (Harvard School of Public Health), Mine Çetinkaya-Rundel (Duke University)

This page titled 9.2.7: Exercises is shared under a CC BY-SA 3.0 license and was authored, remixed, and/or curated by David Diez, Christopher Barr, & Mine Çetinkaya-Rundel via source content that was edited to the style and standards of the LibreTexts platform.

- 6.E: Inference for Categorical Data (Exercises)** by David Diez, Christopher Barr, & Mine Çetinkaya-Rundel has no license indicated.  
Original source: <https://www.openintro.org/book/os>.